

Paweł Stacewicz

Algorytmy genetyczne : inspiracje ewolucyjnych modeli umysłu

Filozofia Nauki 9/3, 145-157

2001

Artykuł został zdigitalizowany i opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

Paweł Stacewicz

Algorytmy genetyczne Inspiracje ewolucyjnych modeli umysłu

1. WSTĘP

Niniejsza praca należy do pogranicza filozofii umysłu i badań nad sztuczną inteligencją. Obydwie dziedziny łączą bardzo silne związki, czego najbardziej skrajnym wyrazem jest tzw. obliczeniowa teoria umysłu. W jej ramach usiłuje się sprecyzować, a czasami wręcz sformalizować, ogólne założenie o podobieństwie między umysłem a komputerem.

Przekonująca filozoficzna krytyka tej teorii stawia przed informatykami zadanie przybliżenia filozofom tych technik sztucznej inteligencji, które wychodzą poza zasadę algorytmiczności. Techniki takie istnieją. Jedną z nich są prezentowane w niniejszej pracy algorytmy genetyczne. Zdaniem autora mogą one stanowić istotny wkład informatyki w rozpoczęty już proces tworzenia ewolucyjnych modeli umysłu.

2. OBLICZENIOWE MODELE UMYŚLU

Pojęcie umysłu należy do najbardziej pojemnych i wieloznacznych. Przyczyną tego faktu jest z jednej strony historia tego terminu — reprezentuje on sobą to wszystko, co niegdyś próbowano określić mianem duszy — z drugiej strony zaś wielopłaszczyznowość współczesnych badań. Liczba nauk, które zajmują się czynnościami, strukturami i wytworami umysłu sięga dziesiątek, a ich rozpiętość tematyczna — od neurofizjologii po logikę i kulturoznawstwo — utrudnia nie tylko wymianę wyników, ale nawet uzgodnienie stosownej terminologii. Całą rzecz komplikuje wielość stanowisk filozoficznych wobec umysłu, które bądź stanowią syntezę wyników

pewnych nauk szczegółowych, bądź się od nich dystansują, usiłując stosować analizy czysto filozoficzne (np. fenomenologiczne).

Współczesną próbą odpowiedzi na zarysowane wyżej trudności są modele umysłu.¹ Tworzy się je w ramach takich nauk szczegółowych i interdyscyplinarnych, jak psychologia, informatyka, cybernetyka czy kognitywistyka. Proponowane modele mają różne stopnie ogólności. W ujęciach najwęższych usiłuje się opisać tylko wybrane zdolności umysłu (np. zdolność rozumienia języka), w ujęciach szerszych modeluje się sam intelekt, w propozycjach idących najdalej uwzględnia się wiele różnych, powiązanych ze sobą zjawisk psychicznych — np. procesy motywacyjne, konieczność zaspokajania różnorodnych potrzeb, mechanizm generowania problemów.

CHARAKTERYSTYKA MODELI OBLICZENIOWYCH. Ważną klasę modeli umysłu stanowią tzw. modele obliczeniowe nazywane również komputacyjnymi lub teorioinformacyjnymi. Zakłada się w nich, że myślenie — utożsamiane wstępnie z rozwiązywaniem problemów — polega na wykonywaniu algorytmów, a dokładniej na przekształcaniu symboli według zbioru ściśle określonych reguł. Przekształceń takich może dokonywać dowolny system zdolny pobierać dane z otoczenia, konsekwentnie stosować do nich zadane reguły i generować wyniki. Jego podłoże materialne — biologiczne, mechaniczne czy elektroniczne — nie odgrywa tak wielkiej roli, jak struktura zapewniająca konsekwentne stosowanie reguł. Same reguły mają postać hipotetyczną (jeśli A, to B), co sprawia, że system regułowy może operować na wielu różnych danych i pełnić różne funkcje. Przekształcane symbole reprezentują wszelkie informacje niezbędne do rozwiązania stawianego maszynie problemu, a zatem zbiory danych, ślady pamięciowe, rozwiązania cząstkowe i pomocnicze, twierdzenia logiczne, wzory matematyczne, itp. Z punktu widzenia systemu są one realizowane jako stany fizyczne jego elementów, np. rozkład napięć na elementach elektronicznych tworzących pamięć systemu.

Cechy omawianych modeli nawiązują oczywiście do funkcji komputerów cyfrowych. Przypomnijmy, że mimo niezwyklej złożoności komputerów i oszałamiającej wręcz różnorodności ich zastosowań istnieje stosunkowo prosta konstrukcja myślowa, która reprezentuje ich zasady działania. Jest to tzw. maszyna Turinga.² Jej idea

¹ *Za Małą Encyklopedią Logiki*, poprzez modelowanie rozumiemy takie upraszczanie zjawisk (procesów i przedmiotów), aby uzyskać o nich pewną dodatkową, użyteczną wiedzę, albo też skuteczniej rozwiązywać dotyczące ich problemy. Upraszczając zjawiska, wybiera się tylko niektóre ich cechy lub wzajemne relacje, resztę zaś pomija. Wśród modeli wyróżnia się teoretyczne i realne. Te pierwsze stanowią zbiór założeń upraszczających (mówiących o tym, które cechy są istotne i czemu służy ich wybór), te drugie zaś są zbiorami przedmiotów spełniających te założenia. Z tego powodu można je traktować jako realizacje modeli teoretycznych.

² Maszyna Turinga składa się z nieskończonej taśmy, na której są zapisywane symbole pewnego alfabetu, głowicy przesuwającej się po taśmie oraz pamięci stanów. Operacje głowicy są kontrolowane przez program, który składa się z instrukcji typu: „jeśli w bieżącej (wskazywanej przez głowicę) klatce taśmy stoi symbol x , a maszyna znajduje się w stanie y , to wpisz w klatkę symbol x' , zmień stan na y' i przesuń się o jedną klatkę w prawo lub w lewo. Maszyna rozpoczyna pracę od

wyraża najprościej to, co rozumie się przez mechaniczne przekształcanie symboli (obliczanie) zgodnie z zadaniem programem. Z tego względu obliczeniowe modele umysłu nazywa się również modelami redukującymi umysł do maszyny Turinga.

Drugi, równie abstrakcyjny rys, opisywanych modeli to ich logicyzm. Ponieważ formalnym wzorem rozumowań opartych na regułach jest logika, jądrem modeli obliczeniowych czyni się jej zasady (aksjomaty i reguły wyvodu). Zauważmy jednak, że zasady logiki rozumie się tu czysto syntaktycznie — jako zasady przekształcania wyrażań na podstawie ich struktury. Co do znaczenia wyrażań (symboli) zakłada się, że są one ściśle określone i niezmiennie, a jako takie albo stanowią obiektywny element świata, albo są dostarczane z zewnątrz przez osobę, która znaczenie ich określiła. W pierwszym wypadku opisane wyżej systemy byłyby w stanie rozwiązywać wszelkie problemy, w drugim stanowiłyby tylko narzędzia pomocnicze, których wyniki interpretowałby użytkownik.

WADY MODELI OBLICZENIOWYCH. Trzy ostatnie zdania opisują jedno z najbardziej istotnych ograniczeń modeli obliczeniowych.³ Ograniczoniom tym, z uwagi na częściowo negatywny charakter niniejszego artykułu przyjrzymy się dokładniej. Podstawowy zarzut krytyków brzmi następująco: umysł ujmowany obliczeniowo wykonuje po prostu zadany z góry program, a to wyklucza jego spontaniczną aktywność i interakcję ze światem zewnętrznym (która mogłaby ten program zmienić). Postulując taką naturę umysłu, ignoruje się klasyczne już wyniki psychologów. Przypomnijmy chociażby badania J. Piageta, w których niezwykle mocno akcentuje się aktywność poznawczą podmiotu i L. Wygotskiego, w których podkreśla się wagę społecznych uwarunkowań funkcji poznawczych. Filozofia z kolei uświadamia nam istnienie pewnych pierwotnych założeń modeli obliczeniowych. Mianowicie, jeśli uznamy odrębność umysłu od świata i zdolność do jego odzwierciedlenia, to wyłączając z modeli umysłu jakiegokolwiek twórcze interakcje z otoczeniem, musimy zgodzić się na dwa twierdzenia: 1) umysł zawiera pewne ściśle określone wzory rzeczywistości, do których jest w stanie dotrzeć; 2) świat, który umysł odzwierciedla, jest w swojej istocie niezmienny, przy czym — biorąc pod uwagę dzisiejszą wiedzę — niezmiennosc ta musi się wyrażać w prawach zmian.

Inna, często wymieniana, wada modeli obliczeniowych, to ich abiologiczny charakter. Chodzi tu głównie o ignorowanie struktury i zasad działania mózgu, a także praw ewolucji i rozwoju, które wpływają na ostateczny kształt mózgu. Wobec ogromu danych neurobiologii i genetyki zarzut ten brzmi niezwykle mocno. Podkreślmy tu, że pomysł maszyny Turinga, a więc podstawy modeli obliczeniowych,

wyróżnionego stanu początkowego i zbioru symboli na taśmie, które reprezentują dane rozwiązywanego problemu, zaś zatrzymuje się, gdy osiągnie pewien wyróżniony stan końcowy. Końcowy zbiór symboli na taśmie reprezentuje rozwiązanie problemu.

³ Przekonującą krytykę modeli obliczeniowych można znaleźć w książkach: Edelman, G. M., *Przenikliwe powietrze, jasny ogień. O materii umysłu*, PIW, Warszawa 1998, s. 304—308; Hetmański, M., *Umysł a maszyny. Krytyka obliczeniowej teorii umysłu*, Wydawnictwo UMCS, Lublin 2000, s. 136—164.

wywodzi się nie tyle z uogólnienia (czy uproszczenia) przebiegu rzeczywistych operacji mózgowych, co z analizy ich wyniku, tj. zasad rozumowania algorytmicznego. Wynik ten zaś nie musi być ani jedyny, ani najbardziej efektywny.⁴

Przedstawione argumenty krytyczne zostały dobrane tak, aby ukazać czytelnikowi pewien brakujący element obliczeniowych modeli umysłu. Elementem tym jest dynamiczna i celowa zmiana reguł przekształcania symboli czyli uczenie się.⁵ Nic nie stoi na przeszkodzie, aby jego zasady nawiązywały (przynajmniej częściowo) do prawideł ewolucji. Zaskakujące może wydać się to, że kwestię tę dostrzegał sam A. Turing, a więc w jakimś sensie pionier referowanych wyżej koncepcji. Niech jego spostrzeżenia posłużą za wstęp do dalszej części artykułu:

Każdą taką maszynę (tj. maszynę, którą chcemy uczynić inteligentną — P.S.) trzeba poddać próbie uczenia, ocenić jak dobre osiąga wyniki, i porównać to z wynikami innych maszyn. Przyjawszy następujące odpowiedniości: budowa mechanicznego dziecka — materiał dziecizny; zmiany w budowie maszyny — mutacje; dobór naturalny — opinia eksperymentatora; cały ten proces możemy porównać do ewolucji naturalnej.⁶

3. ALGORYTMY GENETYCZNE

Idea algorytmów genetycznych została wprowadzona do informatyki pod koniec lat 60-tych XX wieku przez J. Hollanda. Kierowany przez niego zespół już na początku badań postawił sobie dwa, dość ambitne, cele: 1) opisać i wyjaśnić istotę procesów adaptacyjnych w przyrodzie; 2) stworzyć użyteczne oprogramowanie działające na wzór systemów biologicznych. Jakkolwiek cele te są dalekie od pełnej realizacji, techniki ewolucyjne doczekały się wielu zastosowań, z których większość mieści się w formule sztucznej inteligencji. Poniżej przedstawimy podstawy algorytmów genetycznych, kładąc nacisk na fakt, że można je uznać za pewne narzędzie do rozwiązywania problemów.⁷ Na osobnej stronie umieścimy schemat algorytmu wraz z opisem najczęściej stosowanych operacji genetycznych.

⁴ Niezależnie od ograniczeń czy wad obliczeniowe modele umysłu mają wiele zalet. Najważniejszą z nich pozostaje ścisły związek z komputerami, a więc najbardziej zaawansowanymi systemami do sterowania, przetwarzania informacji i symulacji ludzkiego intelektu. W szczególności poprzez analogię z komputerowymi metodami reprezentacji wiedzy powstaje możliwość ścisłego opisu tzw. reprezentacji umysłowych i przeprowadzanych na nich operacji.

⁵ Zauważmy, że w wyniku uczenia się ulega zmianie sens symboli przekształcanych przez systemy algorytmiczne. Sens zmienia się, ponieważ używanie tych samych symboli prowadzi do innych wyników i reguł działania.

⁶ Turing, A., *Maszyna licząca a inteligencja*, [w:] B. Chwedeńczuk (red.), *Filozofia umysłu*, PWN, Warszawa 1999, s. 295.

⁷ Algorytmy genetyczne opracowano na podstawie książek: Goldberg, D. E., *Algorytmy genetyczne i ich zastosowania*, Wydawnictwa Naukowo-Techniczne, Warszawa 1998, oraz Michalewicz, Z., *Genetic Algorithms + Data Structures = Evolution Programs*, Springer Verlag, Berlin 1992.

PODSTAWY ALGORYTMÓW GENETYCZNYCH. Zasady działania algorytmów genetycznych nawiązują do znanych z przyrody mechanizmów reprodukcji, dziedziczenia cech i doboru naturalnego. W przeciwieństwie do algorytmu tradycyjnego, który zmierza do rozwiązania problemu krok po kroku, zgodnie z zadaną strategią, algorytm genetyczny generuje całe populacje rozwiązań i sprawdza ich dopasowanie do wymogów problemu. Generowanie kolejnych populacji odbywa się na zasadach przypominających proces ewolucji: rozwiązania mutują, wymieniają między sobą informacje, rozmnażają się, giną. Spośród nich selekcjonuje się wyniki coraz bardziej zbliżone do oczekiwanego — rolę środowiska, do którego w teorii ewolucji przystosowują się przedstawiciele danego gatunku, przejmuje specjalna funkcja oceny rozwiązań.

Prześledźmy sposób działania algorytmu dokładniej. Załóżmy, że dany jest problem P oraz pewna przestrzeń potencjalnych jego rozwiązań. Rozwiązania te nazywa się *chromosomami* i koduje najczęściej w postaci ciągów zerojedynkowych. Algorytm rozpoczyna działanie od pewnej grupy chromosomów (przeważnie wybranych losowo) określanej mianem *populacji początkowej*. W kolejnych krokach algorytmu rozpatrywane są pojedyncze populacje, w obrębie których chromosomy ewoluują. Oznacza to, że są poddawane działaniu operatorów genetycznych, takich jak *mutacja*, *krzyżowanie* i *inwersja*. W przypadku zerojedynkowej reprezentacji chromosomów mutacja polega na zamianie losowo wybranych zer na jedynki i odwrotnie, inwersja oznacza odwrócenie kolejności w wybranych sekwencjach zer i jedynek, zaś krzyżowanie odpowiada rozmnażaniu — dwa losowo wybrane chromosomy dzielą się na części, z których jest składane potomstwo. W wyniku takich lub podobnych operacji w populacji pojawiają się nowe osobniki, które bądź są do niej dopisywane, bądź zastępują chromosomy dotychczas w niej istniejące. Każdy z chromosomów nowopowstałej populacji zostaje oceniony pod kątem dopasowania do problemu P . Ocena ta staje się podstawą dla przeprowadzenia *selekcji*, która ogólnie rzecz biorąc polega na wyborze z populacji bieżącej chromosomów o największej wartości funkcji oceny. Dokładniej zaś o wyborze decyduje prawdopodobieństwo — chromosomy lepsze są wybierane z prawdopodobieństwem większym, chromosomy gorsze z mniejszym, a zatem im wyższa funkcja oceny chromosomu tym szerszą reprezentację zyskuje on wśród wyselekcjonowanych. Efektem selekcji jest nowa populacja, która zupełnie podobnie jak jej poprzedniczka zostanie przetworzona w kolejnym kroku algorytmu. Algorytm kończy pracę, kiedy zostanie spełnione pewne kryterium finalne, na przykład w populacji pojawi się osobnik o wystarczająco dużej funkcji oceny lub średnia ocena populacji osiągnie pewną zadawalającą wartość.

MODYFIKACJE OPERACJI GENETYCZNYCH. Przedstawiony wyżej schemat opisuje szeroki wachlarz algorytmów genetycznych. Podstawowe różnice między nimi dotyczą sposobu kodowania potencjalnych rozwiązań, operacji selekcji oraz operatorów odpowiadających za przetwarzanie chromosomów. Przyjrzyjmy się wybiórczo różnym odmianom tych składników algorytmu.

Stosunkowo najczęściej wykorzystywaną i najlepiej opisaną teoretycznie metodą kodowania chromosomów jest notacja binarna. Przyczyna jej popularności tkwi w istnieniu teorii (tzw. teorii schematów Hollanda), która uzasadnia zbieżność algorytmów operujących na ciągach zerojedynkowych, pozwala oszacować jej szybkość oraz wskazuje na możliwe przeszkody w poprawnym działaniu algorytmów. Mimo braku teorii opisującej inne metody kodowania w wielu zastosowaniach wygodniej jest operować na całych fragmentach chromosomów, które mogą być reprezentowane przez liczby w systemie dziesiętnym.

Niezależnie od tego prowadzi się badania nad wprowadzeniem do algorytmów genetycznych występującej w przyrodzie diploidalności. Chromosomy są przetwarzane i dziedziczone jako dwa komplementarne łańcuchy genów, wśród których występują geny dominujące i recesywne. Przy ocenie chromosomu bierze się jednak pod uwagę tylko wypadkową informacji obydwu łańcuchów, czyli osobnika, w którym jedne geny się ujawniły, a inne pozostały ukryte. Metoda taka zapobiega zapomnieniu przez chromosomy pewnych informacji, które okazały się użyteczne we wcześniejszych etapach przetwarzania.

Operacja selekcji zazwyczaj opiera się na tzw. *regule ruletki*. W pojedynczym kroku algorytmu poszczególnym chromosomom są przypisywane prawdopodobieństwa selekcji proporcjonalne do ich oceny względem populacji. Prawdopodobieństwa te można sobie wyobrazić jako mniejsze lub większe wycinki koła ruletki o polu równym jedności. Koło jest puszczone w ruch tyle razy, ile chromosomów liczy populacja. Za każdym razem, kiedy wskaźnik ruletki zatrzyma się na danym wycinku, do nowej populacji jest dołączany osobnik odpowiadający temu wycinkowi. Losowy charakter całej procedury powoduje, że niektóre osobniki mogą zostać wybrane kilkukrotnie, a inne w ogóle.

Opisana metoda doczekała się wielu modyfikacji, które polegają z grubsza na pewnych „udoskonaleniach” losowego charakteru selekcji. Najprostsza z nich ogranicza się do każdorazowego dołączania do populacji chromosomu o najwyższej funkcji oceny. W bardziej zaawansowanych modyfikacjach do populacji dołącza się oczekiwane liczby kopii poszczególnych osobników (wyznaczone na podstawie ich prawdopodobieństw selekcji), a dopiero pozostałe miejsca obsadza się według zasady ruletki.

Populacja, wśród której dokonuje się selekcji zawiera osobniki, które powstały poprzez wykonanie na innych osobnikach pewnych operacji rekonfigurujących. Do najbardziej typowych zalicza się opisane wcześniej krzyżowanie, mutację i inwersję. Częstotliwość ich wykonywania regulują odpowiednie prawdopodobieństwa. Na przykład jeśli prawdopodobieństwo krzyżowania wynosi 0,3, to wybór chromosomów do tej operacji polega na losowaniu dla każdego chromosomu liczby z przedziału $\langle 0,1 \rangle$ i sprawdzaniu, czy jest ona mniejsza od 0,3. Jeśli tak, chromosom będzie brał udział w krzyżowaniu. Prawdopodobieństwa operacji rekonfigurujących albo stanowią niezmienny element algorytmu (przypadek najczęstszy), albo są zmieniane

globalnie wraz z postęпами algorytmu, albo są dołączane do poszczególnych chromosomów i podlegają ewolucji.

Osobnym kierunkiem rozszerzeń tradycyjnego algorytmu genetycznego jest modelowanie gatunków. Począwszy od pewnego etapu ewolucji chromosomy dzieli się na grupy osobników podobnych, zaś ewolucja odbywa się tylko w ramach grup. Dzięki temu algorytm może poszukiwać jednocześnie wielu lokalnych ekstremów funkcji oceny.

Na koniec podkreślmy jeszcze raz, że cel konkretnego algorytmu wyznacza funkcja oceny. Z punktu widzenia analogii biologicznych funkcja ta modeluje środowisko, w którym podlega selekcji materiał genetyczny. Z punktu widzenia zastosowań definiuje ona najbardziej istotne cechy rozwiązywanych problemów.

Schemat algorytmu genetycznego

PARAMETRY ALGORYTMU: metoda kodowania rozwiązań, funkcja ich oceny, metoda selekcji, prawdopodobieństwa mutacji i krzyżowania, warunek końcowy

1. Stwórz populację początkową P_t ; $t = 0$
2. Dopóki nie jest spełniony warunek końcowy wykonuj 3 ... 6
 3. Przetwórz (P_t)
 4. Dokonaj_Oceny (P_t)
 5. Dokonaj_Selekcji (P_t)
 6. $t = t + 1$
7. Rozwiązaniem problemu jest najlepszy chromosom z P_t

MUTACJA

1. Dla $i = 1, \dots, n$ wykonuj kroki 2 i 3
2. Generuj losowo liczbę r z przedziału $[0,1]$
3. Jeśli ($r < P_M$), dokonaj mutacji w C_i

KRZYŻOWANIE

1. Dla $i = 1, \dots, n$ wykonuj kroki 2 i 3
2. Generuj losowo liczbę r z przedziału $[0,1]$
3. Jeśli ($r < P_K$), wybierz C_i do krzyżowania ($KR = KR \cup C_i$)
4. Skrzyżuj kolejne pary chromosomów z KR

SELEKCJA

1. Oblicz $O(C_i)$, $i = 1, \dots, n$. Oblicz $F = \sum_{i=1..n} O(C_i)$
2. Oblicz prawdopodobieństwa wyborów $p_i = O(C_i) / F$
3. Oblicz prawdopodobieństwa łączne $q_i = \sum_{j=1..i} p_j$
4. Dla $i = 1, \dots, n$ wykonuj kroki 5, 6, 7

5. Generuj losową liczbę r z przedziału $[0,1]$
6. Szukaj j spełniającego warunek: $q_j \leq r < q_{j+1}$
7. $SEL = SEL \cup C_j$

Oznaczenia

C_1, \dots, C_n — chromosomy, $P_t = \{C_1, \dots, C_n\}$ — populacja w kroku t

KR, SEL — zbiory chromosomów wybranych odpowiednio: do krzyżowania i selekcji

P_M, P_K — prawdopodobieństwa odpowiednio: mutacji i krzyżowania

$O(C_i)$ — ocena i -tego chromosomu

4. W KIERUNKU EWOLUCYJNYCH MODELI UMYSŁU

CECHY MODELI EWOLUCYJNYCH. Algorytmy genetyczne wchodzą w skład wielu zaawansowanych systemów informatycznych. W aktualnych badaniach szczególną wagę przywiązuje się do tzw. systemów hybrydowych, w ramach których usiłuje się łączyć algorytmy genetyczne z sieciami neuronowymi i narzędziami logicznymi. Upowszechnianie się tego typu rozwiązań będzie zapewne oddziaływało na filozofię, która zwróci się w stronę ewolucyjnych modeli umysłu. Uprzedzając nieco ten zwrot, wymienimy cechy, których nie wykazywały modele obliczeniowe, a które będą przysługiwać systemom opartym na algorytmach genetycznych.

Przed wszystkim ich zasady działania będą odwoływały się do obserwacji biologicznych. Algorytmiczne techniki komputerowe zostaną w nich wykorzystane na wyższym poziomie, w celu modelowania lub realizacji procesów biologicznych. Naczelne zasady dziedziczenia i doboru naturalnego zostaną zapewne uzupełnione o inne prawidłowości, dotyczące na przykład rozwoju struktur neuronowych czy metod ich uczenia.

Systemy ewolucyjne pozwolą opisać dynamikę rozwoju. Opisywane układy będą jednak doskonaliły się «same», a nie w wyniku zmiany ich parametrów przez eksperymentatora. Rozwój będzie polegał na selekcji, a nie wykonywaniu instrukcji. Dokładnie tak, jak w wypadku algorytmów genetycznych, system będzie testował wstępnie wiele własnych działań czy cech i wybierał spośród nich najlepsze. Innymi słowy, gotowy program działania zastąpi zbiór celów, które trzeba osiągnąć oraz zbiór ograniczeń, których nie można przekroczyć. Prawdopodobnie podstawowe strategie algorytmu genetycznego — wymiana informacji między chromosomami, dziedziczenie i selekcja — będą wykorzystywane na różnych poziomach systemu i w różnym stopniu. Na przykład inne mechanizmy będą odpowiadały za przyrost doświadczenia osobniczego, a inne za długoterminowe, rozciągnięte na wiele populacji, utrwalanie się cech gatunkowych.

Systemy ewolucyjne będą rozwijały się i działały w ciągłej interakcji ze środowiskiem. Dlatego też, analizując konkretny system „od wewnątrz”, czy to z punktu wi-

dzenia jego aktualnego stanu, czy to z punktu widzenia jego zasad działania, nie będzie można przewidzieć do końca jego zachowania. Za wiele zmian będzie odpowiadało bowiem zmienne i nieprzewidywalne środowisko. Podobna nieprzewidywalność i indywidualizacja będzie dotyczyć rozwoju systemu.

Trzy opisane cechy technik ewolucyjnych — biologizm, dynamika oparta na selekcji i interaktywność — stawiają zagadnienie komputerowego modelowania umysłu w zupełnie nowym świetle. Problem tkwi jednak w braku pełnej matematycznej teorii algorytmów genetycznych, która wskazywałaby między innymi na ich ograniczenia, a także opisywałaby najbardziej efektywne operacje genetyczne i metody kodowania chromosomów. Dopiero taka teoria i zrodzona w jej wyniku efektywna koncepcja algorytmów genetycznych mogłaby stanowić dla modeli ewolucyjnych to, czym dla modeli obliczeniowych jest maszyna Turinga.

EWOLUCJA UMYSŁU WEDŁUG D. DENNETTA. Brak wzmiankowanej teorii nie przeszkadza tworzyć dość rozbudowanych wizji ewolucji umysłu, które wychodzą znacznie poza uogólnienia technik informatycznych, a powstają na skrzyżowaniu wielu nauk, w tym filozofii. Jedną z nich zawdzięczamy D. Dennetowi. Filozof ten stawia tezę, iż świadome umysły ludzkie powstały w wyniku ewolucji (tj. komplikowania się i łączenia w większe całości) prostych, nieświadomych systemów intencjonalnych.⁸ Odtworzenie tego procesu jest możliwe dzięki założeniu, że systemy takie istnieją do dziś — są to na przykład układy sterujące zachowaniem prostszych organizmów niż człowiek (ameb, roślin, zwierząt) oraz składniki bardziej złożonych całości (komórki, systemy sensoryczne, systemy metaboliczne). D. Dennett, oprócz wielu innych ciekawych pomysłów, proponuje hierarchię systemów (istot) definiujących poszczególne etapy ewolucji umysłu.⁹ Systemy te wykazują coraz większą łatwość przystosowywania się do zmian środowiska oraz coraz szerszy i bardziej efektywny repertuar przekazywania własnych doświadczeń przyszłym pokoleniom.¹⁰

Dennetowską hierarchię bytów otwierają tzw. *istoty darwinowskie* — najprostsze systemy, które już w momencie narodzin dysponują w pełni ukształtowanymi wzorcami działania. W ich przypadku cały przekaz i doskonalenie umiejętności odbywa się na poziomie faktycznego działania. Te istoty, którym natura dała sposobność wykonywania funkcji lepiej dostosowanych do środowiska po prostu żyją dłużej i dają

⁸ Poprzez system intencjonalny D. Dennett rozumie każdy system regulowany informacyjnie (działający w zależności od odbieranych i przetwarzanych informacji) oraz nastawiony na cel (realizujący pewne cele). Jako przykłady wymienia zarówno systemy naturalne, właściwe istotom żywym (np. system rozprzodczania płynów u roślin czy układ wzrokowy zwierząt), jak i systemy sztuczne (np. system regulujący zachowanie termostatu czy program komputerowy do gry w szachy).

⁹ Dennett, D., *Natura umysłów*, Wydawnictwo CiS, Warszawa 1997, s. 52—70.

¹⁰ Z perspektywy możliwości człowieka (w pełni ukształtowanej osoby) istoty te można uznać za pewne aspekty ludzkich możliwości, badane przez różne działy psychologii, kulturoznawstwa itd. Z perspektywy historycznej są to, według Dennetta, kolejne ogniwa rzeczywistej ewolucji systemów intencjonalnych.

więcej potomstwa. Nieco wyżej od nich stoją tzw. *istoty skinnerowskie*, czyli systemy o wrodzonych dyspozycjach do działań, a nie ich gotowych wzorcach. Dyspozycje te można utożsamiać z pewnymi rozkładami wzmocnień tj. dziedziczonymi tendencjami do utrwalania jednych zachowań i osłabiania innych. O rzeczywistych, ujawniających się cechach (fenotypie) istoty skinnerowskiej decydują więc: środowisko życia oraz wrodzony rozkład celów działania i wzmocnień (genotyp). Istoty te uczą się metodą prób i błędów — w wyniku prostego warunkowania wzrasta prawdopodobieństwo działań prowadzących do pożądanego skutku i właściwie wzmocnianych. Na kolejnym etapie ewolucji powstają tzw. *istoty popperowskie*. W odróżnieniu od swoich poprzedniczek wykazują one zdolność tworzenia wewnętrznych reprezentacji obiektów środowiska. Dzięki niej są w stanie dokonywać bezpiecznej preselekcji zachowań, czyli wypróbowywać działania i obserwować ich skutki w zastępczym środowisku wewnętrznym. Wewnętrzne reprezentacje mają postać sieci informacyjnej, która częściowo jest dziedziczona, a częściowo kształtuje się w interakcji z otoczeniem. Operowanie nimi wymaga pewnych prostych mechanizmów percepcji i uwagi. Kolejny etap ewolucji umysłu należy do tzw. *istot gregoriańskich* — pierwszych organizmów zdolnych wytwarzać wokół siebie ochronny system kultury. Dennett wybija na pierwszy plan dwie ich cechy: zdolność używania narzędzi oraz nawyk utrwalania swoich doświadczeń w środowisku zewnętrznym. Istoty te po raz pierwszy w historii ewolucji przełamują barierę dziedziczenia bezpośredniego (genetycznego) — większość własnych osiągnięć utrwalają i przekazują za pośrednictwem kultury. Dzięki kulturze również dysponują bogatym repertuarem technik uczenia się. Należą do nich: wykorzystywanie narzędzi, wypróbowywanie rozwiązań w konstruowanym samodzielnie środowisku, uczenie się przez prostą komunikację i obserwację innych.

Dennettowską hierarchię bytów zamykają istoty świadome. Powstanie świadomości warunkuje uprzednia konieczność zaawansowanego komunikowania się i wykształcenie wystarczająco rozbudowanych struktur językowych. W referowanych poglądach przyjmuje się, iż do powstania świadomości wystarcza sama ewolucja. Nie istnieje zatem żadna pierwotna iskra świadomości, która warunkowałaby osiągnięcie przez pewne systemy szczebla umysłów ludzkich. Istoty świadome potrafią uczyć się nieporównanie szybciej i skuteczniej niż ich poprzedniczki. Rozwinięty język pozwala im sprawnie przekazywać wiedzę i utrwalać ją na przyszłość. Dzięki jego wieloznacznym strukturom umieją przenosić efekty uczenia się w nowe dziedziny i nowe sytuacje (uczenie się z transferem). Dzięki poczuciu ciągłości własnych działań mogą nie tylko wykorzystywać wiedzę przekazywaną dziedzicznie czy kulturowo, ale również wykorzystywać własne doświadczenia (uczenie przez samoobserwację).

EWOLUCJA MÓZGU WEDŁUG G. M. EDELMANA. Drugi przykład badań osadzonych w nurcie ewolucjonizmu pochodzi z biologii i wiąże się z osobą G. M. Edelmana. Biolog ten uważa, iż kluczem do tajemnic umysłu jest ewolucja i rozwój mózgu, a te ma wyjaśnić jego teoria selekcji grup neuronowych w mózgu.¹¹ Poniżej scharaktery-

¹¹ Edelman, G. M., *Przenikliwe powietrze, jasny ogień. O materii umysłu*, PIW, Warszawa 1998,

zujemy krótko jej trzy podstawowe zasady, które wyznaczają jednocześnie trzy główne stadia rozwojowe mózgu.

W etapie pierwszym tworzy się struktura neuroanatomiczna mózgu czyli zbiór odpowiednio rozmieszczonych względem siebie i połączonych ze sobą sieci neuronowych. Etap ten przebiega niezwykle dynamicznie, w jego trakcie neurony dzielą się, obumierają i rywalizują o miejsce w powstającej strukturze. Kod genetyczny narzuca jedynie zbiór ograniczeń na proces selekcji neuronów i połączeń, nie wyznacza dokładnego kształtu struktury, ale jej najbardziej ogólne, gatunkowe cechy. W drugim etapie, wskutek przyrostu doświadczenia, następuje różnicowanie siły połączeń między neuronami czyli selektywne wzmacnianie lub osłabianie synaps. W ten sposób zostają wyłonione obwody neuronów o ściśle określonych funkcjach. W trzecim etapie wykształcają się dwukierunkowe połączenia między mapami mózgu czyli grupami sąsiadujących ze sobą neuronów o cechach nabytych w etapach 1 i 2. Dzięki połączeniom między mapami w odpowiedzi na dany zbiór bodźców może się uaktywniać wiele map realizujących różne funkcje, co prowadzi do bardziej złożonych i bardziej kompleksowych decyzji (np. o ruchu organizmu), niż gdyby została uaktywniona tylko jedna mapa. Przekaz sygnałów między mapami zachodzi równolegle, a pozytywny wynik podjętych decyzji prowadzi do wzmacniania odpowiednich połączeń. To z kolei owocuje w przyszłości podobnym przebiegiem sygnałów i rozkładem pobudzeń. A zatem wszelki przepływ sygnałów (związany z działaniem, przypominaniem sobie, wyobrażaniem, itp.) nie da się oddzielić od procesu uczenia się.

G. M. Edelman niezwykle mocno podkreśla fakt, iż ludzki mózg jest systemem selekcyjnym (generującym i wybierającym rozwiązania), a nie algorytmicznym (realizującym programy). Jednocześnie odróżnia selekcję na poziomie kodu genetycznego określającego ogólne cechy ludzkiego mózgu (selekcja gatunkowa) od selekcji towarzyszącej kształtowaniu się struktur konkretnego mózgu (selekcja osobnicza).¹² Druga z nich określa przebieg tzw. „małego cyklu ewolucji”, który rozciąga się od poczęcia do śmierci organizmu, pierwsza natomiast wyznacza „duży cykl ewolucji”, w którym mogą powstawać nowe ośrodki mózgowie. Tylko wyniki selekcji gatunkowej utrwalają się w postaci zmian kodu genetycznego.

Podobnie jak D. Dennett, Edelman proponuje hierarchię systemów neuronalnych, które składają się obecnie na mózg i system nerwowy, a wyłaniały się i łączyły ze sobą w trakcie ewolucji. Aby dać jakiś jej obraz, wymienimy te systemy wraz z realizowanymi przez nie funkcjami:¹³ układ limbiczny i pień mózgu (regulacja funkcji życiowych organizmu), układ wzgórzowo-korowy (percepcja), mózdzek, zwoje pod-

s. 116—138.

¹² Określenia „selekcja gatunkowa” i „selekcja osobnicza” pochodzą od autora niniejszego artykułu. Ta pierwsza rozgrywa się w populacji jednostek o zróżnicowanych zdolnościach, ta druga odpowiada za wybór grup neuronów realizujących konkretne funkcje mózgu. Mechanizm i zakres obydwu procesów jest różny, ale zasada podobna: rywalizacja między zróżnicowanymi jednostkami (osobnikami lub grupami neuronów) w celu dostosowania organizmu do potrzeb środowiska.

¹³ Edelman, G.M., *op. cit.*, s. 163—172.

stawy i hipokamp (czasoprzestrzenna koordynacja wyników percepcji i działań organizmu), podsystemy kory mózgowej, np. kora czołowa (myślenie pojęciowe), połączenia zwrotne między systemami korowymi a wzgórzowo-korowymi (koordynacja bieżących wyników percepcji i danych pamięciowych), ośrodki Broca i Wernickiego (mowa). Wyłanianiu się wymienionych struktur towarzyszyło doskonalenie trzech kluczowych procesów mózgowych: percepcji, pamięci i uczenia się. Ich mechanizm daje się opisać, według Edelmana, wyłącznie w kategoriach przesyłania sygnałów między grupami neuronów, a ich rozwój jako tworzenie dodatkowych połączeń między wymienionymi strukturami w interakcji ze środowiskiem.

Edelman określa siebie jako przeciwnika wszelkich obliczeniowych teorii umysłu i nadużywania tezy o bezpośrednich podobieństwach między komputerami a ludźmi. Niemniej jednak, swoją teorię weryfikuje przy użyciu symulowanych komputerowo automatów, które są złożonymi, wielofunkcyjnymi sieciami sztucznych neuronów.

STATUS MODELI EWOLUCYJNYCH. Powyższe przykłady zostały dobrane tak, aby uświadomić czytelnikowi wielopłaszczyznowość ewolucyjnych ujęć umysłu. W szczególności koncepcja Dennetta pokazuje, w jak sugestywny i efektywny sposób można połączyć osiągnięcia biologów, psychologów, informatyków i filozofów nauki.¹⁴ Niezależnie od tego, jaki kierunek i zakres przyjmą zarysowane wyżej poszukiwania, wszelkie ujęcia ewolucyjne wydają się ważne, bo przywracają utraconą równowagę zainteresowań badawczych.

Oto w dotychczasowych badaniach nad umysłem dominowały ujęcia substancjalne i funkcjonalne. Zwolennicy pierwszego stanowiska upatrywali istoty myśli w jej nośniku — utożsamiając go na przykład z odpowiednio złożoną materią lub scharakteryzowaną negatywnie (w opozycji do własności materii) substancją duchową. Innymi słowy starali się odpowiedzieć na pytanie „dzięki czemu lub poprzez co myślimy?”. Funkcjonalisci z kolei wskazywali na zasady procesu myślenia i jego cele. Akcentowali zatem inne pytanie: „jak i po co myślimy?”. Dopiero w ujęciach ewolucyjnych wyeksponowano problem genezy zdolności do myślenia. Wydaje się zaś, że bez znajomości niezbędnych etapów ewolucji umysłu, trudno myśleć o konstrukcji sztucznych systemów inteligentnych.

LITERATURA

- Bolc, L., Zaremba, J., *Wprowadzenie do uczenia się maszyn*, Akademicka Oficyna Wydawnicza RM, Warszawa 1992.
- Bobryk, J., *Locus umysłu*, Zakład Narodowy Imienia Ossolińskich, Wrocław 1987.
- Bolter, J. D., *Człowiek Turinga. Kultura Zachodu w wieku komputera*, PWN, Warszawa 1990.
- Dennett, D. C., *Natura umysłów*, Wydawnictwo CiS, Warszawa 1997.
- Chwedeńczuk, B. (red.), *Filozofia umysłu*, PWN, Warszawa 1999.

¹⁴ Koncepcję Dennetta można traktować nawet jako program badawczy nowej, interdyscyplinarnej dziedziny ewolucyjnych badań nad umysłem, która dorównywałaby swym rozmachem cybernetyce.

- Edelman, G. M., *Przenikliwe powietrze, jasny ogień. O materii umysłu*, PIW, Warszawa 1998.
- Goldberg, D. E., *Algoritmy genetyczne i ich zastosowania*, Wydawnictwa Naukowo-Techniczne, Warszawa 1998.
- Hetmański, M., *Umysł a maszyny. Krytyka obliczeniowej teorii umysłu*, Wydawnictwo UMCS, Lublin 2000.
- Michalewicz, Z., *Genetic Algorithms + Data Structures = Evolution Programs*, Springer Verlag, Berlin 1992.
- Scott, A. S., *Schody do umysłu*, Wydawnictwa Naukowo-Techniczne, Warszawa 1999.
- Searle, J. R., *Umysł, język, społeczeństwo*, Wydawnictwo CiS, Warszawa 1999.
- Simon, H., *The Sciences of the Artificial*, The MIT Press, Cambridge (Mass.), London 1996.
- Tadeusiewicz, R., *Sieci neuronowe*, Akademicka Oficyna Wydawnicza RM, Warszawa 1993.