

## **DLACZEGO WARTO PRZECZYTAĆ?**

Książka ta, choć ukazuje się w profesjonalnej serii informatycznej, wiele ma do powiedzenia tym humanistom, których pasjonuje zagadka ludzkiego umysłu i jego fenomenalnych możliwości. Płodne i nowoczesne podejście do tej zagadki polega na konstruowaniu informatycznych modeli umysłu. Na tym polu badań inżynieria splata się z filozofią: pierwsza daje sprawdzalność i precyzję, druga kieruje ku fundamentalnym problemom rzeczywistości.

Czy komputery i roboty potrafią się uczyć, czyli rozwiązywać inteligentnie nowe, zaskakujące dla nich problemy? A jeśli tak, to do jakiego progu dociera ta zdolność? Czy istnieje jakaś nieprzekraczalna dla nich granica, którą potrafi jednak pokonać umysł ludzki?

Szukając odpowiedzi na te pytania, autor rozważa i przystępnie opisuje nowoczesne techniki modelowania umysłu, dające wgląd w sekrety jego uzdolnień. Mogą to być systemy eksperckie, które czerpią swoją siłę z logiki, a umieją zastąpić z powodzeniem lekarza, prawnika czy analityka giełdowego. Mogą to być sztuczne sieci neuronowe, dzięki którym zaczynamy rozumieć zdumiewającą plastyczność ludzkiego mózgu. Ale również algorytmy genetyczne, pozwalające wejrzeć "komputerowo" w nabytą ewolucyjnie moc obliczeniową żywych komórek. Autor pokazuje umiejętnie, jak w informatycznym lustrze odbijają się różne sfery świata. Warto więc zajrzeć z nim za drugą stronę lustra.

## **ZAWARTOŚĆ KSIĄŻKI** (bez podrozdziałów 2-go poziomu)

### **CZĘŚĆ I. Sztuczne systemy uczące się**

#### **1. Systemy informatyczne**

- 1.1. Podstawowe pojęcia informatyki
- 1.2. Specyfika systemów informatycznych
- 1.3. Narzędzia czy maszyny autonomiczne?
- 1.4. Podsumowanie

#### **2. W stronę sztucznych systemów inteligentnych**

- 2.1. Zagadnienie sztucznej inteligencji
- 2.2. Badania nad sztuczną inteligencją
- 2.3. Logicyzm czy naturalizm?
- 2.4. Podsumowanie

#### **3. Badania nad systemami uczącymi się**

- 3.1. Przedmiot i zakres badań
- 3.2. Strategie automatycznego uczenia się
- 3.3. Automatyczne uczenie się pojęć
- 3.4. Podsumowanie

## CZĘŚĆ II. Od systemów informatycznych do modeli umysłu

### 4. Umysł jako system poznawczy

- 4.1. Pojęcie umysłu
- 4.2. Interdyscyplinarne badania nad umysłem
- 4.3. Podsumowanie

### 5. Uczenie się jako przejaw otwartości umysłu

- 5.1. Pojęcie uczenia się
- 5.2. Wybrane badania nad uczeniem się
- 5.3. Podsumowanie

### 6. Modelowanie czynności umysłowych

- 6.1. Modelowanie
- 6.2. Teoriainformacyjna strategia modelowania
- 6.3. Podsumowanie

### 7. Wybrane teoriainformacyjne modele umysłu

- 7.1. Modele regułowe
- 7.2. Modele sieciowe
- 7.3. Modele ewolucyjne
- 7.4. Podsumowanie

### 8. Ku otwartym modelom umysłu

- 8.1. Modele otwarte a systemy uczące się
- 8.2. W stronę otwartych modeli ewolucyjnych
- 8.3. Podsumowanie

## WPROWADZENIE

Przed zapoznaniem się z treścią książki warto, by czytelnik zyskał ogólną orientację w jej tematyce i obranej przez autora linii rozumowania. Temu służy poniższe wprowadzenie.

1. Ogólną koncepcję monografii odzwierciedla jej tytuł: „*Umysł a modele maszyn uczących się*”. Tytuł ten ma charakter relacyjny, a jego brzmienie zwraca uwagę na fakt, że w pracy postawiono naprzeciw siebie: **(a) umysł** – rozpatrywany pod kątem zdolności do uczenia się, oraz **(b) informatyczne systemy uczące się** – rozpatrywane w postaci teoretycznej, czyli pod kątem właściwych im zasad przetwarzania danych.

2. Próbując określić bliżej pierwszy człon relacji tytułowej, tj. umysł, trzeba stwierdzić, że jest on rozumiany w duchu współczesnej kognitywistyki, a więc jako złożony **system**

**poznawczy.** System ten składa się z mózgu (podsystemu biologicznego) oraz pewnej złożonej zawartości informacyjnej – częściowo wrodzonej, a częściowo nabywanej w interakcji ze środowiskiem – która w określony sposób „konfiguruje” mózg, a w rezultacie warunkuje takie a nie inne zachowania podmiotu. [zob. rozdział 4]

Kluczową własność tak pojętego umysłu stanowi **zdolność do uczenia się**, która sprawia, iż obdarzony umysłem człowiek jest systemem **otwartym**; wykazuje zatem takie cechy, jak innowacyjność, wynalazczość i odkrywczność; a odwołując się do innego rodzaju terminologii, potrafi wchodzić w skuteczną interakcję z otoczeniem i adaptować się do jego wymagań. [zob. rozdział 5]

3. Jeśli chodzi o drugi człon relacji tytułowej, czyli systemy informatyczne, to w książce skupiono się na systemach trojakiemu rodzaju, wyodrębnianych m.in. w ramach badań nad **sztuczną inteligencją**. Są to: **(a)** systemy **regułowe** – przetwarzające dane (symboliczne) za pomocą przejrzystych reguł typu „jeśli przesłanka, to konkluzja”, a implementowane często w postaci systemów eksperckich; **(b)** systemy **koneksyjne** (sieciowe) – przetwarzające dane (subsymboliczne) w sposób rozproszony i równoległy, a realizowane praktycznie za pomocą sztucznych sieci neuronowych; oraz **(c)** systemy **ewolucyjne** (selekcyjne) – przetwarzające dane na zasadach przypominających naturalną ewolucję, a kontrolowane m.in. za pomocą algorytmów genetycznych<sup>1</sup>. [zob. rozdziały 1, 2 i 7]

4. Każdy z w/w systemów można nazwać **uczącym się** pod warunkiem, że wyposaży się go – na co pozwalają współczesne techniki informatyczne – w specjalne moduły do rozbudowy wiedzy systemu w interakcji ze środowiskiem. Zależnie od typu układu proces ten, tj. uczenie się, inaczej przebiega. I tak: w przypadku systemów regułowych uczenie się polega na zapamiętywaniu wyników różnego rodzaju **wnioskowań** (np. dedukcyjnych lub indukcyjnych); w przypadku układów koneksyjnych – na stopniowej **reorganizacji struktury** sieci (tj. struktury połączeń między elementami przetwarzającymi) ze względu na coraz większą skuteczność jej działań; w przypadku systemów ewolucyjnych – na generowaniu całych populacji próbnych rozwiązań, poddawaniu ich **symulowanej ewolucji**, a następnie selekcjonowaniu i zapamiętywaniu tych wyników, które zapewniają systemowi coraz większą skuteczność. [zob. rozdziały 3 i 7]

5. Z punktu widzenia badacza-interpretatora (np. filozofa) na szczególną uwagę zasługują **relacje metodologiczne** między wymienionymi wyżej sztucznymi systemami uczącymi się a umysłem – co sygnalizuje zwięźle zawarte w tytule książki słówko „a”. W książce położono nacisk na relację **modelowania**. Znaczy to, że pojęcia *stricte* informatyczne – takie jak algorytm, struktura danych czy aplikacja – traktuje się w tekście jako cenne poznawczo narzędzia opisu czynności umysłowych. Same systemy zaś (a także sterujące ich pracą oprogramowanie) stawia się w roli modeli, czyli uproszczonych obrazów aktywności poznawczej ludzi<sup>2</sup>. [zob. rozdziały 6, 7 i 8]

---

<sup>1</sup> Wybór tych trzech rodzajów systemów jest w pewnym sensie arbitralny. Zdaniem autora jednak oddaje on dość dobrze współczesne tendencje badawcze informatyków, a ponadto obrazuje pewną ważną dystynkcję między logicystycznym (odwołującym się do logiki) a naturalistycznym (odwołującym się do empirii) nurtem badań nad sztuczną inteligencją. Rozróżnienie to – reprezentowane, z jednej strony, przez systemy regułowe, a z drugiej strony, przez układy sieciowe i ewolucyjne – będzie przewijać się przez wiele partii monografii.

<sup>2</sup> Na marginesie trzeba dodać, że relacja modelowania nie jest ani jedyłą, ani szczególnie uprzywilejowaną relacją typu (umysł – system informatyczny). Oprócz niej rozpatruje się także relację sztucznej realizacji (z jej perspektywy systemy informatyczne pozwalają w sztuczny, inżynierski sposób zrealizować pewne funkcje poznawcze; zazwyczaj w inny sposób niż ma to miejsce w naturze) oraz relację wspomagania (z jej perspektywy systemy informatyczne są jedynie narzędziami wzmacniającymi przyrodzone zdolności człowieka).

6. Omówione w książce cząstkowe modele umysłu – cząstkowe, bo obrazujące tylko pewien wycinek aktywności poznawczej – wywodzą się ze współczesnej wersji **metafory** umysłu jako mechanizmu; współczesnej w tym sensie, że umysł zostaje przyrównany do **mechanizmu informatycznego**. Precyzując tę metaforę sięga się współcześnie do różnych teorii przetwarzania danych, traktując je jako podstawę formalizacji (tj. maksymalnie ścisłego opisu) budowanego modelu. Innymi słowy, konstruując model sięga się do pojęć informatycznych – zarówno ogólnych, jak szczegółowych (tj. dotyczących konkretnej techniki przetwarzania danych). [zob. rozdział 6]

Stosownie do trzech wyróżnionych typów systemów informatycznych (zob. pkt 3) zaprezentowano w książce trzy rodzaje modeli: **(a)** modele **regułowe** – wywodzące się z metafory „umysłu jako maszyny logicznej”, a konstruowane w postaci regułowych (symbolicznych i zalgebraizowanych) systemów eksperckich; **(b)** modele **sieciowe** – wywodzące się z metafory „umysłu jako maszyny mózgo-podobnej”, a konstruowane w postaci subsymbolicznych, rozproszonych i równoległych układów konekcyjnych; **(c)** modele **ewolucyjne** – nawiązujące do teorii ewolucji, a projektowane w postaci sztucznych systemów ewolucyjnych.

Każdy z modeli oceniono pod kątem adekwatnego opisu uczenia się jako czynności niezbędnej do zapewnienia otwartego charakteru umysłu. [zob. rozdział 7]

7. Metodologiczna analiza zagadnienia modelowania umysłu za pomocą systemów informatycznych ujawnia dwa **kluczowe problemy**, które mają charakter **infinitystyczny**, a które możemy nazwać roboczo *nieskończonym regresem wewnętrznym* i *nieskończonym regresem zewnętrznym*. Przewyciężenie tych problemów stanowi warunek konieczny opisu (a także sztucznej realizacji) umysłu jako otwartego systemu informatycznego.

8. Problem pierwszy, nazwany wyżej *regresem wewnętrznym*, ma swoje źródło w *stricte formalnych* właściwościach systemów uczących się. Otóż w przypadku systemów takich obok zwykłych algorytmów, odpowiedzialnych za działanie systemu, muszą być realizowane także algorytmy „wyższego poziomu”, opisujące zmianę kształtu pewnych elementów właściwego algorytmu (np. parametrów reguł symbolicznych).

Powstaje zatem sytuacja następująca. Za zmianę pewnych elementów algorytmu pierwszego poziomu, czyli algorytmu działania, odpowiada schemat drugiego poziomu (pierwszy z algorytmów uczenia się); za zmianę pewnych elementów algorytmu drugiego poziomu odpowiada schemat trzeciego poziomu (drugi z algorytmów uczenia się); za zmiany w obrębie tego ostatniego odpowiada schemat czwartego poziomu... i tak potencjalnie w **nieskończoność**.

Teoretycznie rzecz biorąc zatem, każdemu modelowi umysłu w postaci systemu uczącego się „grozi” **nieskończony regres** poziomów uczenia się – regres, który z formalnego punktu widzenia uniemożliwia skonstruowanie jakiegokolwiek systemu uczącego się samodzielnie, tj. bez ostatecznej kontroli z zewnątrz.

Aby problemu tego uniknąć, należy wskazać właściwy typ algorytmu, który byłby wykorzystany na ostatnim poziomie uczenia się. W książce argumentuje się na rzecz tezy, że powinien być to **schemat ewolucyjny** – schemat łączący w sobie lokalny indeterminizm (probabilizm) z globalnym ukierunkowaniem zmian. [zob. rozdział 8]

9. Problem drugi, nazwany wyżej *regresem zewnętrznym*, wynika z kolei z istnienia **sprzężenia zwrotnego** między umysłem i jego modelem.

Oto człowiek posługując się modelem – np. konstruując go, używając go, interpretując uzyskiwane za jego pomocą wyniki – tak naprawdę **doskonali swój umysł** (lub przynajmniej go zmienia). W ten sposób umysł, będący z założenia systemem otwartym, staje się

nieustannie „czymś więcej” niż jego tymczasowe modele. Innymi słowy, model będący obrazem dziedziny mentalnej w czasie  $t$  stopniowo przestaje być obrazem adekwatnym, ponieważ dziedzina ta zmienia się wskutek używania modelu. A dotyczy to zarówno umysłu konkretnego człowieka, jak i abstrakcyjnego pojęcia umysłu jako ogółu zdolności, czynności i struktur poznawczych ludzi na pewnym etapie rozwoju. Powstaje zatem groźny problem **nieskończonego regresu** – regresu polegającego na konieczności konstruowania kolejnych modeli, z których żaden nie może okazać się ostateczny.

Jako metodę rozwiązania tego problemu wskazano w książce ponownie na techniki ewolucyjne – techniki, które mogą stanowić podstawę, tzw. **populacyjnego schematu** modelowania. [zob. rozdział 8]