

**David Danks, 2014, *Unifying the Mind: Cognitive Representations as Graphical Models*, Cambridge, MA: The MIT Press, ss. xi + 287**

Początki prac nad graficzną reprezentacją zależności przyczynowych, która byłaby uogólnieniem stosowanych w powszechnie w naukach społecznych różnych odmian metody regresji, rozpoczęły się w połowie lat 1980. Prowadził je zespół kierowany przez C. Glymoura w Carnegie-Mellon University oraz, niezależnie, współpracownicy J. Pearla w University of California w Los Angeles. Efektem tych prac są m.in. algorytmy, zwł. tzw. PC, pozwalające na podstawie macierzy korelacji zmiennych ustalić możliwe rodzaje zależności przyczynowych oraz odpowiadającą im reprezentację graficzną w postaci grafów skierowanych, w których węzły reprezentują poszczególne zmienne („skutki”, „przyczyny”), a strzałki kierunek oddziaływań przyczynowych.

Zamierzenie autora *Unifying the Mind* jest niezwykle ambitne. Stara się on wykazać, że ta reprezentacja jest czymś znacznie więcej niż tylko pewną wyidealizowaną abstrakcją. Danks uważa, że modele graficzne zależności przyczynowych są podstawą reprezentacji mentalnej, na której opierają się

wszystkie procesy poznawcze, zachodzące w umyśle człowieka. Zasadniczy argument, jaki został zaprezentowany w książce, polega na wykazaniu słuszności tego stanowiska w odniesieniu do trzech klas kluczowych procesów poznawczych, jakimi są: myślenie przyczynowe, tworzenie pojęć oraz podejmowanie decyzji.

Struktura książki odzwierciedla tok argumentacji Danksa. Pierwsze trzy rozdziały wprowadzają w tematykę i dyskusje filozoficzne na temat reprezentacji mentalnej, obliczeniowych modeli poznawczych oraz modeli graficznych, a także unifikacji poznawczej dla różnych mechanizmów reprezentacji. Tę unifikację Danks traktuje jako jeden z rodzajów relacji, jaka zachodzi między różnymi teoriami. Spośród różnych możliwych relacji dla proponowanej unifikacji autor *Unifying the Mind* proponuje relację ograniczania jednej teorii przez drugą pod względem określonej własności, zwłaszcza prawdziwości, odrzuca natomiast relację redukcji jednej teorii do innej.

Rozdział czwarty „Causal Cognition” omawia pierwszy z trzech wy-

branych przez Danksa obszarów zastosowań modeli graficznych jako bazowej reprezentacji mentalnej. Dla uniknięcia trudności związanych z dyskusją filozoficzną nad naturą przyczynowości, autor proponuje ograniczenie się do samej reprezentacji zależności przyczynowych, bez rozstrzygnięcia ich odniesienia przedmiotowego oraz natury ontologicznej. Podstawowy podział tak rozumianej przyczynowości obejmuje: uczenie się przyczynowe oraz rozumowanie przyczynowe, chociaż, jak podkreśla autor, nie jest to w pełni rozłączne odróżnienie. Zwłaszcza, że oba rodzaje poznania przyczynowego traktuje on jako operacje na modelach graficznych: tworzenie modelu oraz wyprowadzanie z niego konsekwencji i predykcji. W obu przypadkach przywołuje wnioski z wcześniejszych dyskusji na temat acyklicznych grafów skierowanych jako właściwych modeli zależności przyczynowych, jak sam podkreśla: „poznanie przyczynowe było stosunkowo łatwym celem, gdyż jest to obszar nauk kognitywnych, w którym modele graficzne są już najczęściej wykorzystywane i przyjmowane” (s. 98).

Kolejny rozdział „Pojęcia, kategorie i wnioskowanie” dotyczy przede wszystkim pojęć jako reprezentacji opartej na modelach graficznych. Jak podkreśla Danks, wykazanie tego ma szczególne znaczenie dla całego argumentu, gdyż pojęcia stanowią składnik niemal wszystkich reprezentacji poznawczych. Stara się więc wykazać, że wszystkie ważniejsze teorie pojęć można ująć jako obliczeniowe teorie opierające się na modelach graficznych.

Wykazuje również, że wszystkie operacje na pojęciach można przedstawić jako operacje na modelach graficznych. Posługuje się przy tym przykładami pojęć, jak pies czy kot, wprowadzając kilka różnych odmian węzłów, które reprezentują m.in. zbiory zmiennych oraz zaktualizowane lub potencjalne tylko atrybuty tych zmiennych. Jedną z głównych relacji, jaką Danks stara się zdefiniować posługując się aparaturą pojęciową modeli graficznych, jest relacja podobieństwa. Mimo że w jej definicji posługuje się on prawdopodobieństwem warunkowym, to jednak trudno uznać za przekonującą argumentację, że jest to wystarczająca zbieżność, by utożsamić ją z modelowaną zależnością przyczynową.

Rozdział szósty pt. „Podejmowanie decyzji za pomocą modeli graficznych” dostarcza Danksowi kolejnego argumentu za uznaniem modeli graficznych za podstawowy rezerwuar reprezentacji wykorzystywanych w podstawowych procesach poznawczych. Podobnie jak było to w przypadku poznania przyczynowego, podejmowanie decyzji posiada bogatą literaturę, w której nie tylko wykorzystuje się modele graficzne do prezentowania drzew decyzyjnych, ale także do prezentacji różnych wartości prawdopodobieństwa dla poszczególnych scenariuszy. Zwłaszcza z punktu widzenia bayesowskiej teorii decyzji ten sposób prezentowania problemów teoriodecyzyjnych jest bardzo naturalny. Dlatego też wydaje się, że podobnie jak miało to miejsce w rozdziale czwartym, zadanie Danksa jest znacznie ułatwione istniejącym już

stanem literatury. Mimo tego jednak w swojej argumentacji wprowadza on reprezentację graficzną, która – oprócz standardowych dla acyklicznych grafów skierowanych węzłów i strzałek – posługuje się kilkoma różnymi rodzajami węzłów (trójkąty, kwadraty, elipsy, romby) dla oznaczenia różnego rodzaju czynników wpływających na podejmowanie decyzji bądź reprezentujących podjęcie decyzji.

Mimo więc pewnych naturalnie nasuwających się spostrzeżeń dotyczących wykorzystania reprezentacji graficznej w poznaniu przyczynowym, tworzeniu pojęć i operacjach na nich oraz podejmowaniu decyzji, nasuwają się istotne wątpliwości co do słuszności argumentacji Danksa. Przede wszystkim w jego argumentacji daje się zauważyć brak spójności w rozumieniu kluczowych dla całego argumentu „modeli graficznych”. W każdym z trzech rozważanych przypadków są one inaczej definiowane, z innymi rodzajami zmiennych oraz zależności między nimi. Podważa to ostateczny wniosek Danksa o istnieniu wspólnego dla różnych procesów poznawczych rezerwuaru reprezentacji w postaci modeli graficznych, które – jak się okazuje – są w podanych zastosowaniach odmiennymi sposobami reprezentacji.

W kolejnym rozdziale „Unifying Cognition” Danks stara się wykazać, że poszukiwana przez niego w poprzednich trzech rozdziałach zbieżność strukturalna reprezentacji w podstawowych obszarach poznania nie ma wyłącznie charakteru matematycznego, lecz „metafizyczny”. Jed-

nym z zasadniczych powodów słuszności prezentowanego poglądu ma być fakt, że rzeczywiście różne procesy poznawcze korzystają ze wspólnego rezerwuaru reprezentacji, a nie tylko posiadają wspólną czy izomorficzną strukturę, opartą na „silosie” reprezentacji swoistym dla danego rodzaju procesu poznawczego.

Rozdział ósmy jest krytyką innych podejść do reprezentacji mentalnej jako alternatyw dla modeli graficznych. Danks wyróżnia dwa rodzaje podejść alternatywnych. Pierwsza to unifikacje oparte na wspólnym dla różnych procesów poznawczych schemacie. Do najważniejszych reprezentantów tej grupy zalicza on koneksjonizm oraz bayesianizm. Druga grupa to unifikacje oparte na procesie, w których zakłada się istnienie kilku podstawowych procesów, do których można zredukować wszystkie złożone procesy poznawcze. W tej grupie znalazły się takie podejścia, jak: ACT-R, Soar oraz Sigma. Danks dochodzi ostatecznie do wniosku, że w przypadku niektórych z tych podejść mamy do czynienia z parcjalnymi unifikacjami, lecz jedynie proponowane w książce podejście może osiągnąć unifikację procesów i schematów.

W dwóch zamykających książkę rozdziałach Danks stara się zarysować podstawowe założenia teorii, która wyjaśniłaby neuronalne podłoże reprezentacji modeli graficznych. Charakteryzuje również najważniejsze pytania otwarte, które w przyszłości musi podjąć proponowana teoria unifikacji. Jednym z ważniejszych jest odniesienie do

zjawisk społecznych, co wymagałoby m.in. „badania możliwych ról opartych na modelach graficznych reprezentacji poznawczych w tworzeniu języka, jego rozumieniu oraz przetwarzaniu” (s. 226).

Korpus książki zamyka apendyks matematyczny, który podaje ścisłe definicje podstawowych pojęć dotyczących modeli graficznych oraz dowody najważniejszych twierdzeń matematycznych, do których w książce odwołuje się Danks. Na końcu książki dołączono bibliografię oraz indeks.

Książka *Unifying the Mind* jest poniekąd podsumowaniem istotnego etapu w rozwoju modeli graficznych. Do ich upowszechnienia w literaturze przyczyniła się znacząco praca J. Pearla *Causality* (pierwsze wydanie w 2000, drugie – 2009), która opisuje mate-

matyczną architekturę tych modeli. Danks jest jednym z najbliższych współpracowników C. Glymoura, co tłumaczy kontekst podjęcia ambitnej filozoficznie próby wykorzystującej tę aparaturę pojęciową, znacznie wykraczającą poza pracę Pearla. Mimo aplikacyjnej użyteczności modeli graficznych w reprezentacji wielu zagadnień związanych z badaniem zależności przyczynowych, podjęta przez Danksa próba – jako jedna z pierwszych w literaturze – jest daleka od stworzenia jednolitej koncepcji, która byłaby bez wątpienia bardziej zadowalającą koncepcją niż omówione w książce alternatywy.

*Paweł Kawalec*  
Wydział Filozofii  
Katolicki Uniwersytet Lubelski  
Jana Pawła II