

Marek Lechniak

Katolicki Uniwersytet Lubelski Jana Pawła II

Ku adekwatnej koncepcji zmiany przekonań

Wśród głównych zadań epistemologii współcześnie wymienia się m.in. adekwatne ujęcie dynamiki stanów epistemicznych i doksastycznych oraz zrozumienie racjonalności pozyskiwania wiedzy¹. W niniejszym artykule, sytuującym się na pograniczu epistemologii i logiki, podejmę problem dynamiki stanów epistemicznych. Celem moich rozważań będzie analiza tego, na czym polega dynamika stanów przekonaniowych i w jaki sposób jest ona analizowana w formalnych teoriach zmian przekonaniowych. Okaze się, że wiele podstawowych momentów zmiany stanów naszej wiedzy nie jest (i – być może – nie może być) adekwatnie ujęte przez dominujące formalne teorie zmian przekonaniowych. Wskażemy na koniec kierunki, w jakich, jak mniemamy, winna rozwijać się adekwatna, nie nazbyt idealizująca stany przekonaniowe, teoria zmian wiedzy.

Rozważmy następujące przykłady zmiany wiedzy²:

- (a) zmiana poglądów pewnych badaczy co do efektu cieplarnianego;
- (b) perswazja zmierzająca do zmian przekonań społecznych;
- (c) „rewolucja naukowa” na przykładzie Galileusza;
- (d) rewizja przekonań co do aksjomatyzowalności wiedzy w rezultacie twierdzeń limitacyjnych Gödla.

¹ Por. V. F. Hendricks, J. Symons, *Where's the bridge? Epistemology and epistemic logic*, „Philosophical Studies” 128 (2006), s. 137-167. Autorzy obok wymienionych celów epistemologii dodają jeszcze obronę wiedzy i odpowiedź na wyzwania sceptycyzmu. Wskazują również, że obrona przed sceptycyzmem dominuje u autorów należących do *mainstreamu* epistemologii, posługujących się eksperymentami myślowymi, tradycyjną analizą pojęciową oraz metodami opartymi na analizie intuicji, podczas gdy badanie dynamiki wiedzy uprawiają zwłaszcza filozofowie posługujący się metodami formalnymi; obie zaś orientacje w uprawianiu epistemologii w równym stopniu są zainteresowane rozważaniami nad racjonalnością (*ibidem*, s. 139-140).

² Chodzi tu oczywiście o szeroko pojętą wiedzę – wiedza spełniająca tzw. definicję klasyczną może się zmieniać jedynie kumulatywnie i w tym sensie jest mało interesująca dla badacza dynamiki stanów epistemicznych.

Wymienione przykłady dotyczą różnorodnych dziedzin szeroko pojętej wiedzy. Z jednej strony mamy zmiany w poglądach naukowych, czyli przekonaniach kompetentnych badaczy wyrażanych w teoriach naukowych [(c), (d), i w pewnym stopniu (a)], z drugiej zaś przykłady zmian w statystycznie badanej „opinii społecznej”, czyli jakichś przekonań potocznie żywionych przez ludzkie, niewyspecjalizowane i ograniczone poznawczo indywidua. Na pozór przykłady te nie powinny pojawić się razem; jednak przyjrzenie się bliższe, np. zmianom opinii badaczy w sprawie efektu cieplarnianego jako skutku działalności ludzkiej, pokazuje, że te różnice nie są tak wielkie, jak można by sądzić.

Zacznijmy od analizy tego przykładu³.

Filozoficzna teoria rewizji przekonań powinna stosować się do problemu globalnego ocieplenia przez wyjaśnienie tego, w jaki sposób większość naukowców doszła do akceptacji następujących wniosków: (1) Ziemia ociepla się; (2) Ocieplenie będzie miało dewastujący wpływ na społeczność ludzką; (3) Emisja gazów cieplarnianych jest głównym powodem ocieplenia; (4) Redukcja emisji jest najlepszym ze sposobów redukcji negatywnych jej wpływów na zmiany klimatyczne. Dodatkowo teoria ta powinna dostarczyć wglądu nie tylko w to, w jaki sposób naukowcy dochodzą do przyjęcia tych przekonań, ale także dlaczego niektórzy naukowcy i spora rzesza liderów biznesu i polityki nie przyjmuje ich.⁴

Ci ostatni dzielą się, zdaniem omawianych autorów, na tych, którzy twierdzą, że ocieplenie może być wyjaśnione fluktuacjami energii słonecznej, własnościami ruchu orbitalnego Ziemi, tych, którzy nie wierzą, że dwutlenek węgla emitowany przez ludzi może wpłynąć na efekt cieplarniany oraz tych, którzy akceptują większość twierdzeń naukowych na ten temat, a utrzymują, że nie ma wyraźnego kryzysu i nie ma potrzeby przeprowadzania kosztownych akcji przewycięzania go. Przeciwnicy tezy wskazują, że istnieje wiele znaków zapytania co do raportu IPCC (Międzyrządowego Zespołu do spraw Zmian Klimatu) i wskazują, że ludzie nie mogą mieć istotnego wpływu na ostatnie ocieplenie klimatu, rządy nie powinny wprowadzać zbyt rygorystycznych regulacji, jako że te ograniczają wolność gospodarczą i w efekcie osłabiają państwo. Z drugiej strony, aktywiści walczący z globalnym ociepleniem wierzą, że naukowe przewidywania co do jego skutków są wiarygodne i planeta zmierza do kryzysu, przy czym uważają za oczywiste, że to ludzie mają znaczący wpływ na wzrost temperatury na Ziemi.

Nieco inna sytuacja panuje w wypadku pewnych działań perswazyjnych zmierzających do zmiany przekonań (postaw) społecznych. Aronson i Pratkanis wskazują na cztery techniki wywierania wpływu:

³ Por. P. Thagard, S. Findlay, *Changing minds about climate change: Belief revision, coherence and emotion*, [w:] E. Olsson, S. Enquist (eds.), *Belief revision meets philosophy of science*, Springer, Dordrecht 2011, s. 329-345. Ściśle mówiąc, autorzy referowanego artykułu zajmują dość jednostronne stanowisko, przyznające rację zwolennikom tezy o związku między działalnością ludzką a globalnym ociepleniem sugerując przy tym, że przeciwnicy tej tezy są powiązani finansowo z biznesem paliwowym (nie wydaje się to rzetelne, bo tak samo można by pokazać „nieczystość” działań zwolenników tezy).

⁴ *Ibidem*, s. 330.

- (a) perswazję wstępną – przejęcie kontroli nad sytuacją i stworzenie atmosfery sprzyjającej komunikatowi;
 - (b) wskazanie na wiarygodność źródła – zadbanie o właściwy wizerunek nadawcy w oczach odbiorców (nadawca musi pokazać się jako dobrze poinformowany, godny zaufania);
 - (c) skonstruowanie i przekazanie komunikatu, który skupi uwagę i myśli odbiorców na treściach, na których zależy nadawcy;
 - (d) kontrolowanie uczuć odbiorców według zasady „wzbudź w odbiorcach określone emocje, a następnie podsuń im sposób reakcji na te emocje, który pokrywa się z działaniem, do którego próbujesz ich nakłonić”⁵.
- Jeszcze inaczej jest w wypadku tzw. rewolucji naukowych.

Historyk nauki, który bada dawne prace naukowe z punktu widzenia współczesnej historiografii może nabrać przekonania, że kiedy paradygmat ulega zmianie, wraz z nim zmienia się i świat. Kierując się nowym paradygmatem, uczeni stosują nowe przyrządy i widzą nowe obszary rzeczywistości. Co ważniejsze, w okresie rewolucji naukowej, posługując się dobrze znanymi przyrządami i badając obszary, które badali dawniej, dostrzegają coś zupełnie innego. Wygląda to tak, jak gdyby zawodowa społeczność uczonych przeniosła się na inną planetę, gdzie przedmioty dobrze znane ukazują się w innym świetle, wraz z innymi, wcześniej nie znanymi.⁶

Ta zmiana, obrazowo opisana przez Kuhna, polega na tym, że obraz świata przed rewolucją naukową jest niewspółmierny z obrazem świata po rewolucji; sytuację świetnie ilustruje przykład dotyczący ruchu wahadła.

Arystotelicy, którzy uważali, że ciężar dzięki swej naturze porusza się z góry na dół, aby tam osiągnąć stan naturalnego spoczynku, twierdzili, że takie huśtające się ciało ma po prostu trudności ze spadaniem. Uwięzione na łańcuchu, osiągnąć może stan spoczynku dopiero po dłuższym czasie ruchu wymuszonego. Natomiast Galileusz, patrząc na kołyszący się ciężar, widział wahadło – ciało, któremu niemal udaje się powtarzać ten sam ruch w nieskończoność. [...] Z własności wahadła wyprowadził na przykład swój dowód niezależności prędkości spadania od ciężaru oraz od stosunku między wysokością a prędkością końcową w ruchu po równi pochylej. [...] Wydaje się, że rola geniuszu [Galileusza – M.L.] polegała tu raczej na wykorzystaniu możliwości percepcyjnych, jakie stworzyła średniowieczna zmiana paradygmatu. Galileusz nie wyrósł całkowicie na gruncie arystotelizmu. Przeciwnie, uczono go analizy ruchu w kategoriach teorii impetu, późnośredniowiecznego paradygmatu, który głosił, że ciało ważkie porusza się nieprzerwanym ruchem dzięki sile wszczepionej mu przez ciało, które go wprowadziło w ruch.⁷

⁵ Por. A. Pratkanis, E. Aronson, *Wiek propagandy. Używanie i nadużywanie perswazji na co dzień*, przeł. J. Radzicki, M. Szuster Wydawnictwo Naukowe PWN, Warszawa 2003, s. 46-62; autorzy, wskazując, że owe cztery techniki odpowiadają czterem aspektom perswazji wyróżnionym przez Arystotelesa: *atechnoi* (okoliczności przygotowujące grunt dla perswazji), *ethos*, *logos* i *pathos*.

⁶ T. Kuhn, *Struktura rewolucji naukowych*, przeł. H. Ostromęcka, Fundacja Aletheia, Warszawa 2001, s. 197. Przedstawiona zmiana paradygmatu funkcjonuje tu jedynie jako przykład zmiany mającej charakter zmiany pojęciowej; abstrahujemy zaś od samej zasadności i trafności Kuhnowskiej koncepcji zmian wiedzy.

⁷ *Ibidem*, s. 210-211.

Jak wskazuje Kuhn, każdy z badaczy, Arystoteles i Galileusz, obserwował jakby co innego – pierwszy spadający, uwięziony kamień, drugi – wahadło. Zmiana wiedzy polega tu więc na całkowicie różnych wizjach (i w konsekwencji, niemożliwych do porównania różnych opisach) tej samej rzeczywistości (autor porównuje tę zmianę do *gestalt switch*) – nie ma więc dyskusji, czy dodać, czy usunąć jakieś sądy; mamy radykalnie odmienny sposób widzenia świata wyrażony za pomocą odmiennej „ramy” pojęciowej.

Z kolei czasem udowodnienie jakiegoś twierdzenia „pieczętuje dyskusję”, zastępując przekonania ich negacjami. Taką sytuację można ilustrować np. udowodnieniem w 1931 r. twierdzenia o niepełności przez Gödla⁸; w rezultacie tego twierdzenia musiano zrezygnować z poglądu, że każdy system wiedzy jest aksjomatyzowalny na rzecz poglądu z nim sprzecznego (upraszczając, Gödel wykazał, że nawet ubogie niesprzeczne systemy logiczne, jak np. arytmetyka liczb naturalnych nie mogą być w pełni aksjomatyzowalne, co było poważnym uderzeniem w założenia programu Hilberta).

Przywołane przykłady ilustrują złożoność problematyki zmiany poglądów (postaw). Właściwie w każdej z podanych sytuacji można zauważyć, że zmiana poglądów może mieć więcej niż jeden wymiar. Może to być zmiana zbioru uznawanych zdań, zmiana używanego aparatu pojęciowego, zmiana nastawień emocjonalnych prowadząca do zmiany postaw (a więc i akceptowanych sądów). Przy czym zmiany mogą mieć różną głębokość – od sytuacji, gdy w ramach jednego systemu pojęciowego zmienia się zbiór akceptowanych (dezakceptowanych) sądów do sytuacji, gdy pewien zbiór sądów zostaje w całości porzucony i zastąpiony innym zbiorem sądów opartym na odmiennej aparaturze pojęciowej, mogą im towarzyszyć zmiany nastawienia emocjonalnego odbiorców (perswazja) lub też mogą mieć charakter nieperswazyjny.

Zmiana wiedzy w ujęciu propozycjonalnym

Często w literaturze podkreślany jest fakt, że współczesna epistemologia i logika prezentuje propozycjonalne ujęcie wiedzy. W ujęciu tym wiedza czy przekonania traktowane są jako zbiór sądów, reprezentowanych przez zdania. Analizując powyższe przykłady, już na tzw. pierwszy rzut oka można wskazać, że zmiana typu rewolucyjnego nie może być reprezentowana w modelu wiedzy traktowanej jako system zdań; wyrażane jest to twierdzeniem o niewspółmierności teorii na-

⁸ Przypomnijmy, że twierdzenie to głosi, iż każdy niesprzeczny system aksjomatyczny zawierający arytmetykę liczb naturalnych jest niepełny, a więc, że istnieją zdania prawdziwe tego systemu, które nie dają się w nim udowodnić. „Twierdzenie Gödla można też sformułować jako twierdzenie stwierdzające, że zbiorem nieaksjomatyzowalnym jest zarówno zbiór zdań prawdziwych arytmetyki liczb naturalnych, jak również zbiór wyrażeń prawdziwych każdego systemu logiki zawierającego arytmetykę liczb naturalnych”. Por. L. Borkowski, *Logika formalna*, PWN, Warszawa 1970, s. 378-379.

ukowych. Po prostu znaczenia podstawowych terminów teorii przed rewolucją i po rewolucji są tak radykalnie różne, że, jak pokazywano na przykładzie z wahadłem, uczeni z różnych paradygmatów „widzą” w tym samym doświadczeniu inne obiekty. Jeśliby zatem użyć języka propozycjonalnego do opisu zmiany, dwa zbiory zdań, reprezentujące stan wiedzy przed rewolucją i po rewolucji, są całkowicie nieporównywalne. Podobna sytuacja zachodzi nierzadko w sytuacji perswazji: w definicjach perswazyjnych (najczęściej niejawnych) zmienia się już to znaczenie powtarzanych w argumentacjach terminów, już to ładunek emocjonalny związany z używanymi terminami. Choć na pozór zdania używane w argumentacjach nie zmieniają się (co do litery), to zmieniają się (często w sposób niezauważony dla odbiorców) sprzężone z nimi przekonania⁹.

Jeśli chodzi o modele zdaniowe przekonań, można wśród nich można odróżnić dwa typy:

- (a) modele fundacjonalistyczne;
- (b) modele koherencyjne.

W modelu fundacjonalistycznym¹⁰ bierzemy pod uwagę przekonanie wraz ze ścieżką jego uzasadnienia (uprawomocnienia). Główna teza fundacjonalizmu głosi, że podstawą uznania jakiegos przekonania jest odwołanie się do przekonania źródłowych, które stanowią podstawę dla uznania innych przekonań. Teoria fundacjonalistyczna odrzuca założenie o konserwatyzmie uprawomocnienia; nie można uznać zdania za uprawomocnione tylko dlatego, że po prostu zostało uznane. Uprawomocnienia są przy tym podważalne: podmiot może mieć uprawomocnione przekonanie *q*, po czym uznać inne przekonanie *p*, które podważy uprawomocnienie podmiotu do uznawania tego pierwszego przekonania. W takim wypadku zdaniem fundacjonalistów, winniśmy zaprzestać uznawania *q* (o ile nie będziemy mieć innego przekonania jako podstawy dla *q*). Uprawomocnienia przekonań nie powinny być koliste ani ciągnąć się w nieskończoność – łańcuch uprawomocnień winien kończyć się na przekonaniach, które nie wymagają uzasadnienia¹¹ albo mogą być uprawomocnione inaczej.

Z kolei ujęcie koherencyjne zakłada, iż jedyną racją dla uznania przekonania jest spójność; uderzająca różnica między jednym a drugim ujęciem jest ta, że teoria fundacjonalistyczna głosi, że jest się uprawomocnionym do trwania w uznawaniu czegoś, jedynie gdy ma się specjalne racje za trwaniem w akceptacji tego przekonania, podczas gdy teoria koherentystyczna głosi, iż jest się uprawomocnionym do trwania w uznawaniu czegoś tak długo, jak długo nie ma specjalnych powodów

⁹ Mechanizm perswazji językowej polega na tym, że „wykorzystuje się w niej istniejące skojarzenia emocjonalne, wiązane z określonymi wyrażeniami języka, przenosząc je na inne wyrażenia lub inne zakresy tych samych wyrażen; a także zmieniając nasilenie, rodzaj lub biegun ładunku emocjonalnego, związanego na gruncie mowy potocznej z danym wyrażeniem językowym”. Por. np. T. Pawłowski, *Tworzenie pojęć w naukach humanistycznych*, PWN, Warszawa 1986, s. 150.

¹⁰ Por. G. Harman, *Change in view*, Bradford Book The MIT Press, Cambridge MA 1986, s. 29-42.

¹¹ Przekonanie jest bazowe, gdy ma wewnętrzne uprawomocnienie nieodwołujące się do innych przekonań.

do tego, aby przestać w to wierzyć. Dopiero wtedy, gdy przekonania są niespójne (sprzeczne), powinno się dokonać minimalnej zmiany pośród przekonań po to, by wyeliminować tę niespójność. Już tu widać, że każdy z modeli jest w pewnym stopniu nieadekwatny, gdyż dla realnie istniejących, w jakimś stopniu racjonalnych podmiotów, zarówno treść (wartość logiczna?) informacji, jak i jej uzasadnienie powinny być ważne dla jej akceptacji czy dezakceptacji.

Jednym z najszerzej dyskutowanych w literaturze modeli koherencyjnych zmiany przekonaniowej jest tzw. model AGM¹². Opiera się on na następujących założeniach:

- Aktualne przekonania podmiotu reprezentowane są przez stan przekonaniowy, reprezentowany z kolei przez zbiór zdań¹³, którego podstawowym kryterium racjonalności jest niesprzeczność. Nie zakłada się żadnych relacji między zdaniami a stanami świata pozaprzekonaniowego (czyli abstrahuje się od prawdziwości akceptowanych zdań).
- Podmiot poznający traktowany jest jako suma mereologiczna kolejnych, następujących po sobie stanów przekonaniowych. Zmiana przekonaniowa rozpatrywana jest jako przejście od jednego stanu do stanu następnego.
- Każda zmiana przekonań jest wywoływana wyłącznie przez tzw. epistemiczną daną wejściową, która reprezentowana jest przez zdanie; każda zatem nowa informacja (wyrażona w zdaniu) jest traktowana jako funkcja odwzorowująca stary stan przekonaniowy na nowy stan przekonaniowy.
- Kryterium racjonalności zmiany jest jej minimalność, która jest potrzebna po to, aby stan przekonaniowy mógł przyjąć daną wejściową, która wywołała zmianę i powrócić do równowagi (*equilibrium*). Innymi słowy, nowa informacja wytrąca stan przekonaniowy z równowagi, a zasada minimalności zmiany jest kryterium umożliwiającym stanowi epistemicznemu powrót do tej równowagi.
- Wszelka zmiana jest determinowana przez rodzaj danych wejściowych – można wyróżnić trzy podstawowe rodzaje zmiany przekonaniowej: (a) dodanie przekonania do zbioru przekonań (ekspansja), (b) usunięcie przekonania ze zbioru przekonań (kontrakcja) i (c) zastąpienie przekonania przekonaniem z nim sprzecznym (rewizja).

¹² Por. P. Gärdenfors, *Knowledge in flux*, A Bradford Book The MIT Press Cambridge MA 1988; obszernie przedstawienie modelu AGM wraz z rozbudowaną analizą trudności tego modelu można znaleźć w M. Lechniak, *Przekonania i zmiana przekonań*, Wyd. KUL Lublin 2011, s. 289-383. Oczywiście, istnieją inne modele zmian przekonaniowych, np. modele Bayesiańskie, których tu nie analizuję; nawiasem mówiąc, jak pokazuje Gärdenfors, ich założenia wcale nie odbiegają daleko od założeń modelu AGM.

¹³ Już w pierwszych latach istnienia koncepcji AGM okazało się, że zamiast zdaniowej reprezentacji przekonań można przyjąć model sądowy (propozycjonalny), gdzie stan przekonaniowy reprezentowany jest przez system sfer; wykazał to Grove w 1986 r., a Gärdenfors w *Knowledge in flux* podaje już ten wynik; por. A. Grove, *Two Modelling for Theory Change*, Aucland Philosophy Papers 1986.

- Zbiór przekonaniowy jest reprezentowany przez niesprzeczny i logicznie domknięty (na klasyczną konsekwencję) zbiór zdań¹⁴.

Dzięki powyższym założeniom otrzymuje się elegancki formalnie, aksjomatycznie scharakteryzowany model zmian przekonaniowych¹⁵. Najkrócej mówiąc, ekspansja jest tu określona za pomocą definicji $K + A = Cn(K \cup \{A\})$, czyli stanowi zwykle teoriomnogościowe rozszerzenie zbioru zdań K . Z kolei rewizję (usunięcie przekonania i zastąpienie go przekonaniem z nim sprzecznym) można zdefiniować za pomocą tzw. definicji Leviego $K * A = (K - \neg A) + A$, zaś zbiór przekonaniowy po kontrakcji (usunięciu przekonania) może być opisany następującym wzorem $K - A = K \cap K * \neg A$ ¹⁶. Niestety, prostota i elegancja tego modelu nie idzie w parze z jego adekwatnością¹⁷. Za nieadekwatnością modelu przemawiają następujące fakty:

- (a) Sam model stanu przekonaniowego jako logicznie domkniętego zbioru zdań jest zbyt ubogi do reprezentowania zmian przekonaniowych. Modelowanie usunięcia przekonania wyrażonego zdaniem złożonym $A \wedge B$ wymaga, przy zachowaniu warunku minimalności zmiany, decyzji, czy usuwamy zdanie A , czy B ; tak więc konieczne jest uzupełnienie zbiorów zdań o relację preferencji określoną na zbiorach przekonaniowych, która porządkuje zdania w zbiorze według ich stopnia odporności na zmianę – mimo że przekonania są uznawane z tym samym stopniem asercji (jako pewne), mają one różną odporność na zmianę (zakorzenienie epistemiczne) wyznaczoną przez czynniki pozalogiczne.
- (b) W modelu AGM pomija się sprawę wartości logicznej zdań wejściowych i w ogóle oceny wartości informacji wejściowej; podmiot (a właściwie jego stan przekonaniowy) przyjmuje każdą nową daną wejściową. Trudno uznać taki pozbawiony „barier immunologicznych” system zdań za racjonalny. Choć faktycznie realne podmioty poznające przyjmują najczęściej usłyszane informacje, to jednak dokonują wstępnej oceny źródła tych informacji oraz stosunku informacji do już posiadanych zdań. W sytuacji uznania źródła informacji za niewiarygodne albo w sytuacji, gdy nowa informacja pozostaje w sprzeczności z tym, co już wiemy, zamiast ścieżki spontanicznej „włącza” się ścieżka refleksyjna, w której źródło informacji zostaje poddane analizie albo sama informacja

¹⁴ Ten ostatni warunek, czyli logiczne domknięcie, powoduje, że teoria AGM jest narażona na wszystkie zarzuty związane z logiczną wszechwiedzą. Z drugiej strony trudno sobie wyobrazić brak tego warunku – wówczas moglibyśmy postulować jedynie niesprzeczność aktualną stanu przekonaniowego.

¹⁵ Aksjomaty teorii AGM dla ekspansji, rewizji i kontrakcji można znaleźć w: M. Lechniak, *op. cit.*, s. 300-308.

¹⁶ Innymi słowy rezultatem operacji kontrakcji jest przecięcie (iloczyn) wyjściowego zbioru zdań z zbiorem zdań po rewizji zbioru wyjściowego ze względu na negację zdania A .

¹⁷ Przy czym nie musimy tu mówić o adekwatności względem realnych zdań niewyidealizowanych psychologicznych podmiotów ludzkich, ale adekwatności względem podmiotów poznających albo teorii naukowych, traktowanych jako zbiory zdań żywionych przez społeczność naukową.

- zostanie dokładnie zbadana (w celu niegenerowania sprzeczności czy też uniknięcia zniszczenia przez niezbadane, niesprawdzone informacje)¹⁸.
- (c) W modelu AGM racjonalność zmiany została zredukowana wyłącznie do niesprzeczności systemu przekonaniowego, a więc do kryterium wewnętrznego. Tymczasem na racjonalność zmian przekonaniowych wpływa nie tylko koherencja zbioru zdań uznawanych przez podmiot, ale i ich uzasadnienie (racja, na podstawie której zdanie zostało uznane). W teorii AGM kryterium uzasadnienia racjonalnie nabytych przekonań jest ignorowane. Dochodzenie do przekonań uzasadnionych (uprawnionych) bywa niekiedy procesem żmudnym i wcale nie musi mieć zagwarantowanego sukcesu w postaci nowego przekonania podmiotu.
- (d) Ponieważ wszystkie przekonania podmiotu poznającego tworzą jeden (logicznie domknięty) zbiór, zmiana przekonań dotyczących np. przedszkola mojego dziecka winna wpłynąć w jakiś sposób na moje przekonania dotyczące sytuacji politycznej na Bliskim Wschodzie. Podmiot przekonań winien być więc nieustannie poddawany zmianom przekonaniowym. Tymczasem bardziej realne podmioty mają zwykle przekonania dość starannie podzielone na niekoniecznie ze sobą powiązane podzbiory (jest to też sposób na obronę przed ciągle zagrażającą podmiotowi sprzecznością przekonań)¹⁹. Inną trudnością, którą generuje model, jest niemożność rozpatrywania w nim przekonań introspekcyjnych podmiotu czy też sekwencji zmian przekonaniowych (jest to tzw. *one-shot theory*).

Podsumowując te uwagi, trzeba podkreślić, że charakteryzowany przez model AGM podmiot przekonań ma dość osobliwy charakter. Choć jest podmiotem dynamicznym o nieograniczonych zdolnościach dedukcyjnych, pozostaje obojętny na wartość logiczną danych wejściowych. Jest to przy tym podmiot nieintrospekcyjny; dopuszczenie danych introspekcyjnych generuje w AGM nieprzewidywalne trudności związane z występowaniem zdań Moore'owskich²⁰, które to trudności (jak pokazał Segerberg) można oddalić, choć nie usunąć. Można by sądzić, że typem podmiotu opisywanym przez tę teorię jest jakiś podmiot sztuczny, zdolny do inferencji, gromadzący wszelkie dane jedynie według kryterium niesprzeczności.

Niektóre z trudności wewnętrznych modelu AGM próbowano usunąć. Dotyczy to np. domknięcia logicznego zbioru (zaproponowano jako remedium model tzw. baz przekonaniowych BBD²¹, gdzie oddziela się pewien zbiór przekonań

¹⁸ Por. P. Thagard, *Explanatory coherence*, „Behavioral and Brain Sciences” 12 (1989), s. 435-502.

¹⁹ Por. R. Stalnaker, *Inquiry*, A Bradford Book The MIT Press, Cambridge MA 1987, rozdz. V.

²⁰ Por. K. Segerberg, *Moore problems in full dynamic doxastic logic*, [w:] J. Malinowski, A. Pietruszczak (eds.), *Essays in logic and ontology*, Rodopi Amsterdam, New York 2006, s. 95-110.

²¹ Por. S. O. Hansson, *Taking belief bases seriously*, [w:] D. Prawitz, D. Westerstaal, *Logic and philosophy in Uppsala*, Kluwer Acad. Publ., Dordrecht 1994, s. 13-28.

bazowych i innych przekonania z nich wyprowadzonych), umieszczenia przekonania z różnych dziedzin w jednym zbiorze przekonaniowym (proponycja tzw. *splitting languages*²², czyli języków rozdzielonych – zakłada się tam, że przekonania dotyczące różnych dziedzin są wyrażone w różnych językach i dzięki temu nie wchodzi z sobą w interakcje), czy próba radzenia sobie ze zdaniem introspekcyjnymi w odpowiadających teorii AGM dynamicznych logikach doksastycznych. Jednakże podstawowa wada AGM, tzn. całkowita otwartość systemu przekonania na nowe informacje, abstrahowanie od wartości logicznej tych informacji czy ich uzasadnienia, pozostaje²³.

Dlatego warto w tym miejscu wygłosić kilka uwag o fundacjonalistycznych modelach przekonania i ich zmian. W podejściu tym dzieli się przekonania na dwa zbiory: tych, które są uprawomocnione przez inne przekonania oraz tych, które takiego uprawomocnienia nie mają. Zakłada się przy tym, że

[...] każde wyprowadzone przekonanie winno być wsparte pewnym niekolistym argumentem opierającym się na przekonaniach bazowych. Owa niekolistość pozwala również użyć argumentów, aby zdeterminować zmianę dla całego zbioru przekonania. Przekonanie powinno być pozostawione, nawet gdy usunie się jedno z jego uzasadnień, tak długo, jak długo choćby jedno z niezależnych uzasadnień pozostaje, a powinno być usunięte po usunięciu ostatniego z jego uzasadnień. Innymi słowy przekonanie wyprowadzone (niebazowe) powinno być utrzymywane wtedy, gdy ma ono co najmniej jeden niekolisty argument oparty na przekonaniu fundamentalnym.²⁴

Przykładem systemu opartego na założeniu fundacjonalistycznym jest Doyle'a RMS system (*reason maintenance system*), który w zamierzeniu autora miał pomagać w rewizji stanów bazy danych całego systemu poprzez użycie rejestru inferencji i obliczeń do tego, by wyznaczyć ścieżkę konsekwencji zmian stanu początkowego. Poprzez utrzymanie śladu tego, jak informacja została wyliczona z tego stanu, system rekonstruuje informację wyprowadzalną z danej informacji. Stany systemu składają się z dwóch typów elementów: węzłów (*N*) i racji (*R*). RMS używa węzłów do reprezentowania informacji (przekonania, reguł, procedur, danych z bazy danych), a racji do reprezentowania inferencji, a ściślej, do reprezentowania reguł inferencji. Ponieważ węzły nie muszą reprezentować przekonania, nie nakłada się żadnej logiki na węzły. Natomiast relacje między węzłami są *explicite* wskazywane przez racje (mogą one kodować również relacje logiczne między węzłami). Stan systemu jest uznawany za uprawomocniony o ile zbiór węzłów spełnia każdą

²² Por. R. Parikh, *Beliefs, belief revision and splitting languages*, „Proceedings Logic, Language and Computation”, CSLI 1999, s. 266-278.

²³ Oczywiście, można traktować AGM jako teorię „zabawkową”, powstałą dla względów sportowych i mówić, że postawione tu zarzuty do niej się nie odnoszą. Myślę, że nie byłoby to zgodne z zamierzeniami twórców AGM, którzy stworzyli poważną teorię zmiany przekonaniowej i zasługują też na „poważną” krytykę. Fakt, że wielu najpoważniejszych logików współczesnych włożyło wiele pracy, żeby tę teorię naprawić i pozbawić wskazanych tu trudności, zdaje się tu wystarczającym argumentem.

²⁴ Por. J. Doyle, *Reason Maintenance and Belief Revision. Foundations vs Coherence Theories*, [w:] P. Gärdenfors (ed.), *Belief Revision*, Cambridge Univ. Press, Cambridge 1992, s. 34.

z racji, a każdy węzeł N jest wsparty przez niekolisty argument z ważnych racji R. Ilekroć rozumujący podmiot dodaje lub usuwa racje, system uaktualnia zbiór węzłów do nowego stanu uprawomocnionego²⁵.

Inny przykład fundacjonalistycznego systemu zmiany przekonań (a nawet całej epistemologii opartej na założeniu fundacjonalizmu) można znaleźć w pracach J. Pollocka. Wskazuje on, że założenie koherencjonizmu, iż wszystkie przekonania mają jednakowy status mimo przyjmowanego ich uporządkowania ze względu na poziom zakorzenienia epistemicznego, jest słabo umocowane epistemologicznie. Według Pollocka²⁶ struktura dochodzenia do przekonania jest następująca: jedynym źródłem przekonań są percepcje, które zaopatrują podmiot w reprezentacje świata.

Posiadanie percepcji dostarcza nam podważalnego uzasadnienia dla przekonania, że świat jest taki, jaki jest reprezentowany przez percepcje. Jest to wyrażone przez powiedzenie, że percepcja dostarcza podważalnej racji dla przekonania, które wspiera. Zauważmy, że mamy tu „rację za” rozumianą jako relacja między stanami mentalnymi, a nie między treściami tych stanów. Jest to ważne, gdyż rozumiemy całkiem różnie w różnych stanach o tej samej treści, np. percepcja, że jest coś czerwonego przede mną, pragnienie, żeby było coś czerwonego przede mną, czy przekonanie, że coś czerwonego jest przede mną”, a przekonania uzyskane drogą świadectwa (trzecioosobowe) mogą być zredukowane do percepcji, dającej rację do myślenia, że osoba S składa raport, że *p*; informacja wyprowadzona z trzecioosobowego raportu wywodzi się z percepcyjnych danych wejściowych poprzez wieloetapową argumentację.²⁷

Tak więc ostatecznymi przesłankami argumentów są percepcje. Podstawowe dla tej koncepcji jest odróżnienie schematów podważalnych od schematów niepodważalnych. Rozumowanie w zgodzie ze schematami rozumowania podważalnego dostarcza racji dla wniosku i uprawomocnia konkluzję pod nieobecność sprzeciwiających się jej informacji. Ze schematami rozumowań podważalnych sprzężony jest zbiór tzw. anulatorów, które mogą służyć podważaniu uzasadnienia wniosku. Jeśli z danym rozumowaniem nie skojarzono zbioru anulatorów, jest ono niepodważalne (np. rozumowania oparte na prawach logiki). Są dwa rodzaje anulatorów: „odrzucający”, czyli argument za przeciwną konkluzją (racja za odrzuceniem konkluzji) i anulator „przecinający”, który atakuje związek między przesłankami a wnioskiem rozumowania podważalnego (analogicznie do dwóch strategii atakowania argumentacji w teorii dyskusji – głoszenia tezy opozycyjnej do tezy zwalczanej oraz atakowania (przez oponenta) uzasadnienia tezy bronionej).

²⁵ *Ibidem*, s. 34-35 oraz literatura na temat szczegółów technicznych systemu RMS.

²⁶ Por. np. J. Pollock, *Defeasible Reasoning*, [w:] J. Adler, L. Rips (eds.), *Reasoning: Studies of Human Inference and its Foundations*, Cambridge University Press, Cambridge 2008, s. 451-470 czy J. Pollock, A. Gillies, *Belief Revision and Epistemology*, „Synthese” 122 (2000), s. 69-92.

²⁷ Por. J. Pollock, A. Gillies, *op. cit.*, s. 74.

Status argumentu podważalnego determinowany jest zaś przez siłę tego argumentu i porównanie jej ze stopniem uzasadnienia anulatorów²⁸.

Trzeba tu wskazać, że sama teoria fundacjonalistyczna również poddana była ostrej krytyce. Stawiano jej m. in. następujące zarzuty²⁹:

- (1) ludzie najczęściej nie działają w taki sposób, jak gdyby brali cały czas pod uwagę ścieżkę uzasadnienia (i raczej rzadko pamiętają racje dla swoich przekonań). obrońcy fundacjonalizmu odpowiadają na to, iż choć być może, gdy idzie o realnych ludzi, tak faktycznie jest, to zarzut ten w odniesieniu do badanych w sztucznej inteligencji systemów rozumowania jest pozbawiony zasadności;
- (2) teoria fundacjonalistyczna nie gwarantuje zmiany konserwatywnej, czyli tego, że przekonania trwają i nie zmieniają się bez ważnego powodu; obrońcy fundacjonalizmu odpowiadają na ten zarzut, że konserwatywizm w słabszym sensie, wyrażający się w stwierdzeniu, że brak wcześniejszych wskazań co do fałszywości przekonania jest powodem dla podtrzymywania przekonania, jest respektowany przez ujęcie fundacjonalistyczne;
- (3) magazynowanie w pamięci wyprowadzeń przekonań i używanie ich do uaktualniania przekonań winno prowadzić do wielkiego obciążenia pamięci podmiotu;
- (4) mimo że w teoriach koherencyjnych nie podaje się racji dla przekonań, to jednak teoria koherencyjna może być rozwinięta, aby takich racji dla przekonań dostarczać; niestety, jak dotąd, trudno znaleźć takie rozwinięcie teorii koherentystycznej.

Omówione tu podejścia do zmiany poglądów sytuują się po stronie koncepcji propozycjonalnych. Mają one swoje mocne strony – główną zaletą jest możliwość oparcia ich na logice zdań. Szczególnie dotyczy to koncepcji koherentystycznych, które wprost stosują logikę zdaniową do reprezentowania przekonań, a relację między systemami zdaniowymi do reprezentowania zmian wiedzy. W modelach fundacjonalistycznych rola logiki jest mniejsza, jako że zwykle są to modele sentencjalne, a nie propozycjonalne – po prostu mówimy tu o relacjach między zdaniami, a nie między treściami zdań (sądami). Tymczasem, jak wskazywano wyżej, zmiana taka, jak w rewolucji naukowej, zdaje się trudna do sprowadzenia do któregoś z omawianych tu modeli. Dlatego warto tu zatrzymać się nad próbami logiczno-epistemologicznego opisu zmiany konceptualnej.

²⁸ Por. *ibidem*. W cytowanym artykule, podobnie jak w wielu książkach Pollocka można znaleźć formalne, bardziej techniczne przedstawienie jego systemu.

²⁹ Por. J. Doyle, *op. cit.*, s. 37-46.

Zmiana konceptualna³⁰

Jak zauważono na początku, rola zmiany pojęciowej w rozwoju nauki ujawniła się w sposób wyraźny po raz pierwszy w pracach T. Kuhna, który wskazał na szczególny rodzaj zmian teorii naukowych, zwany rewolucją. Rewolucja polega na całkowitej zmianie standardów i metod, tak że racjonalna ewolucja rywalizujących poglądów przy użyciu standardów zewnętrznych jawi się jako niemożliwa. Jak wskazuje Thagard, dopiero komputacyjne narzędzia nauk kognitywnych umożliwiają adekwatne ujęcie zmiany pojęciowej (logiczny empiryzm oferował redukcję symplifikującą zmianę pojęciową, podczas gdy historyczne ujęcie zmiany nie jest zbyt ściśle). Sama zmiana pojęciowa nie musi zachodzić jedynie przy zmianie teorii naukowej, ale dokonuje się też podczas rozwoju człowieka, co opisywane jest na gruncie psychologii rozwojowej³¹.

Trzeba tu zwrócić uwagę, że zmiana pojęciowa nie może być, zdaniem Thagarda, zredukowana do zmiany przekonań.

Zalóżmy, że chcemy dokonać bardziej realistycznego psychologicznie podejścia do rewizji przekonań, które mogłoby ująć, dlaczego pewne rewizje są trudniejsze do przeprowadzenia niż inne i dlaczego pewne rewizje mają bardziej globalny zasięg. Te aspekty rewizji przekonań mogą być zrozumiane jedynie przez zbadanie, jak przekonania są organizowane przez pojęcia. Może istnieć różnica pomiędzy rozstrzygnięciem, że wieloryby są ssakami a rozstrzygnięciem, że wieloryby mają płetwy, różnica, którą można zrozumieć w terminach całościowej struktury naszego systemu, odnosząc wieloryby do ssaków w sposób bardziej fundamentalny niż po prostu dochodząc do przekonania, że wieloryby są ssakami.³²

Pojęcia Thagard³³ rozumie jako reprezentacje umysłowe tego, co jest reprezentowane przez słowa (nazwy), podczas gdy sądy reprezentują znaczenie zdań. Wskazuje na liczne funkcje, spełniane przez pojęcia i analizuje, czy sprowadzając zmianę pojęciową do prostej, zdaniowo wyrażanej rewizji przekonań, funkcje te pozostałyby spełnione po takiej redukcji czy też nie; odpowiedzi mogą być argumentami za wagą przyjęcia struktury pojęciowej jako czynnika doniosłego dla zmian wiedzy.

³⁰ Przedstawione tu P. Thagarda ujęcie zmiany pojęciowej charakteryzuje się pewną całościowością i stanowi spójną próbę teorii zmiany poglądów, obejmującą zarówno zmianę na poziomie propozycjonalnym, jak i pojęciowym. Za tą koncepcją przemawia również to, że jej autor przez wiele lat podejmuje próby stworzenia działających aplikacji komputerowych do analizy zmiany poglądów (program ECHO, czyli *Explanatory Coherence by Harmony* – chodzi tu oczywiście o „Harmony”, którą Thagard przekreśla na cześć G. Harmana – *Optimization*).

³¹ Por. P. Thagard, *Conceptual Revolutions*, Princeton University Press, Princeton 1992, s. 4-5. Rola pojęć w poznaniu jest szeroko analizowana także w polskiej literaturze; wystarczy przypomnieć pracę Z. Chlewińskiego, *Umysł. Dynamiczna organizacja pojęć*, Wydawnictwo Naukowe PWN, Warszawa 1999 czy ostatnio wydaną obszerną pracę zbiorową J. Bremer, A. Chuderski (red.), *Pojęcia. Jak reprezentujemy i kategoryzujemy świat*, Universitas, Kraków 2011; niestety w cytowanych pracach kwestia zmiany pojęciowej nie jest podejmowana.

³² Por. P. Thagard, *Conceptual revolutions*, s. 21.

³³ Por. *ibidem*, s. 22-24.

Thagard wymienia dziesięć funkcji poznawczych, w których spełnianiu ważną rolę grają pojęcia. Są to: kategoryzacja, uczenie się, zapamiętywanie, inferencja dedukcyjna, wyjaśnianie, rozwiązywanie problemów, generalizacja, rozumowanie przez analogię, zrozumienie językowe i produktywność. Widać tu różne typy działalności poznawczej: działania na treściach nazw, rozumowania czy działania na znaczeniach nazw. Dwa ostatnie działania z tej listy, jako działania o charakterze czysto językowym, Thagard pomija, koncentrując się na pierwszych ośmiu funkcjach. Niektóre z nich zdają się wręcz niemożliwe bez odwoływania się do struktur pojęciowych człowieka. Tak więc kategoryzacja umożliwi nam rozpoznawanie obiektów jako rzeczy określonych kategorii. Jeśli mamy zaklasyfikować płamę w morzu jako wieloryba (a nie np. łódź podwodną), pojęcie przekonania i jego zmiany nie wydaje się bardzo przydatne (przyjęcie przekonania może być co najwyżej rezultatem podjęcia pewnej decyzji pojęciowej). Jeszcze „gorzej” sprawy mają się z procesem uczenia się. Przy koncepcji rewizji przekonaniowej nie pojawia się, zdaniem Thagarda, pytanie o źródło pojęć jako istotnych członów przekonania. Pojęcia są formowane na co najmniej dwa sposoby: jednym jest uczenie się z przykładów, drugim – kombinowanie nowych pojęć z pojęć już posiadanych. To ostatnie zadanie jest bardzo złożone – nowe pojęcie zyskuje nową strukturę, a nie jest tylko sumą starych pojęć³⁴; tworzenie pojęć w żadnej mierze nie jest zatem sprowadzalne do dodawania nowych przekonania³⁵. Podobnie, według Thagarda, pojęcia pełnią istotną rolę w zrozumieniu roli pamięci w zmianie przekonania. Zdolność do przypominania jest warunkiem koniecznym rewizji przekonania – organizacja pojęć jest dla pamięci istotna jako że jest ważna dla aktywacji pojęć umożliwiającej rewizję przekonania opartą na pamięci. Z kolei gdy idzie o sprawności dedukcyjne realnego podmiotu, w realnym świecie muszą istnieć ograniczenia (wymogi nałożone na to, co może zostać wydedukowane), po to, aby uniknąć eksplozji wykładniczo rosnącej liczby przekonania. Czasem, jak podaje Thagard, same inferencje nie muszą dokonywać się na poziomie zdań, ale za pomocą procesów na poziomie struktur pojęciowych; struktury pojęciowe umożliwiają inferencje niejako obok schematów zdaniowych; podobnie ma się rzecz z wyjaśnianiem i rozwiązywaniem problemów.

Wydaje się, że w niektórych sprawach Thagard ma tu rzeczywiście rację, choć rzeczywistość jest nieco bardziej skomplikowana. Mianowicie formalne systemy wykazują wysoki, zbyt wysoki stopień idealizacji ujęcia przekonania i jego zmiany, choć przekładalność też tych systemów na inne ujęcia zdaje się być argumentem za ich solidnością. Tak więc model zmiany przekonania AGM jest

³⁴ Por. J. Holland, K. Holyoak, R. Nisbett, P. Thagard, *Induction. Processes of inference, learning and discovery*, The MIT Press, Cambridge MA 1986, s. 138-143.

³⁵ Dyskusja tu przypomina nieco dyskusję nad problemem, czy słuszna jest allogeniczna czy idiogenetyczna koncepcja sądu. Argumenty Thagarda zdają się przemawiać na korzyść koncepcji allogenicznej; oczywiście idiogeniści odparliby te argumenty przez wykazywanie, że pojęcia mają naturę „zwinionych sądów”. Tu przytaczamy argumenty Thagarda jako przykład koncepcji spoza mainstreamu epistemologiczno-logicznego.

odtworzalny np. w logikach dynamicznych i ma solidną semantykę formalną. Z drugiej strony, jego podstawową wadą jest oparcie na rachunku zdań, co uniemożliwia jakiegokolwiek mówienie w nim o przedmiotach, ich nazwach i relacjach między nimi, co jak się zdaje, jest podstawowe dla analizy na poziomie konceptualnym. To wymagałoby wprowadzenia języka rachunku predykatów do modelu i przynajmniej częściowo np. mereologii (wiele relacji między pojęciami nie musi zakładać relacji teoriomnogościowych, ale mereologiczne). Tak więc przynajmniej trzy pierwsze operacje poznawcze wskazane przez Thagarda domagają się innego, bogatszego niż czysto propozycjonalne, ujęcia.

Thagard zauważa, że rewolucja naukowa może być opisana za pomocą następujących twierdzeń:

- (a) rewolucje naukowe dotyczą zasadniczych transformacji systemów pojęciowych i propozycjonalnych;
- (b) systemy pojęciowe są pierwotnie ustrukturyzowane za pomocą uporządkowania rodzajowego (*kind-hierarchy*) oraz uporządkowania mereologicznego część – całość (*part – hierarchy*)³⁶;
- (c) nowe pojęcia teoretyczne rodzą się za pomocą kombinacji konceptualnej, w której nowe pojęcia są wyprowadzane z części starych pojęć (np. pojęcie fali dźwiękowej, która nie jest obserwowalna, to rezultat połączenia pojęcia dźwięku z pojęciem fali, które są wyprowadzone z doświadczenia);
- (d) systemy propozycjonalne są zasadniczo organizowane przez relacje koherencji eksplanacyjnej³⁷, jako że, według Thagarda, przy zasadniczych zmianach teorii naukowych, najważniejsze są relacje między sądami co do sprzeczności i wyjaśniania³⁸;

³⁶ „Istotnym składnikiem rewolucji kopernikańskiej było nowe sklasyfikowanie Ziemi jako planety, Newton odróżnił masę od ciężaru i odmiennie ujął grawitację jako rodzaj siły dośrodkowej. Teoria względności przyjęła konceptualizację odmienną niż Newtonowska mechanika ujmując masę i energię jako przejawy masy-energii. Co więcej, znaczenie relacji część-całość zmieniło się w sposób dramatyczny przez zastąpienie zdroworozsądkowych pojęć czasu i przestrzeni przez zintegrowaną czasoprzestrzeń. Ostatecznie teoria kwantowa zatarła różnicę między falą a cząstką” (por. P. Thagard, *Conceptual revolutions*, s. 7).

³⁷ Por. P. Thagard, *Coherence in thought and action*, MIT Press, Cambridge MA 2000, s. 41-48. Prezentację koncepcji Thagarda w języku polskim można znaleźć np. w: M. Urbański, *Paula Thagarda koncepcja rozumowania abdukcyjnego*, „Studia z kognitywistyki i filozofii umysłu”, 6 (2012) nr 1, s. 97-120.

³⁸ Koherencja eksplanacyjna charakteryzowana jest przez Thagarda za pomocą następujących zasad:

- (E1) zasada symetrii: koherencja eksplanacyjna jest relacją symetryczną;
- (E2) zasada wyjaśniania: (a) hipoteza jest koherentna z wyjaśnianymi przez siebie obserwacjami lub innymi hipotezami; (b) hipotezy, które wspólnie wyjaśniają to samo *eksplanandum* są ze sobą koherentne; (c) im więcej hipotez potrzeba, by wyjaśnić jedno *eksplanandum*, tym niższy poziom ich wzajemnej koherencji;
- (E3) zasada podobieństwa: podobne hipotezy wyjaśniające podobne zjawiska są ze sobą koherentne;
- (E4) zasada priorytetu danych: poziom akceptacji elementów, charakteryzujących wyniki obserwacji, nie musi zależeć od związków koherencji;
- (E5) zasada sprzeczności: elementy sprzeczne nie są ze sobą koherentne;
- (E6) zasada konkurencji: jeśli elementy P i Q wyjaśniają to samo *eksplanandum*, a nie są ze sobą powiązane eksplanacyjnie, to nie są ze sobą koherentne;

- (e) nowe hipotezy teoretyczne są tworzone zwykle przez abdukcję³⁹;
- (f) przejście do nowych systemów konceptualnych i propozycjonalnych następuje z powodu tego, że nowe sądy i pojęcia mają większą koherencję wyjaśniającą. Ta ostatnia teza wskazuje, że tym, co decyduje o zwycięstwie nowych teorii, rozwijanych (w przypadku rewolucji) niezależnie od starych teorii, jest zwiększenie stopnia koherencji wyjaśniającej.

Thagard wskazuje wielu stopni zmiany pojęciowej, rozpiętej na skali od prostego dodania nowego przypadku do zmiany zasady organizującej drzewo pojęciowe. Oto one⁴⁰:

1. Dodanie nowego przypadku, np. że ta oto plama widoczna w pewnej odległości jest wielorybem;
2. Dodanie nowej słabej reguły, np. wieloryby mogą być znalezione w Oceanie Arktycznym;
3. Dodanie nowej mocnej reguły, która gra często rolę w rozwiązywaniu problemu i wyjaśnianiu, np. że wieloryby jedzą sardynki;
4. Dodanie nowej relacji część-całość, np. że wieloryby mają śledziona;
5. Dodanie nowej relacji rodzajowej, np. że delfin jest rodzajem wieloryba;
6. Dodanie nowego pojęcia, np. narwał;
7. Zapadnięcie się części hierarchii rodzajowej, zrezygnowanie z wcześniejszego rozróżnienia;
8. Reorganizacja hierarchii przez przeskok dziedzinowy, czyli przerwienie pojęcia z jednej gałęzi drzewa hierarchicznego do innej;
9. Przelączenie drzewa, czyli zmiana zasady organizującej drzewa hierarchicznego.

Jak zauważa Thagard, zmiany 1–5 są możliwe jedynie, gdy dysponujemy pojęciem np. wieloryba jako oddzielnej kategorii bytu. Pojęcia mogą być dodawane z różnych powodów, wliczając w to ich łączenie. W XIX w. naukowcy

(E7) zasada akceptacji: akceptacja elementu zależy od jego koherencji z reprezentacjami, do systemu których należy.

Warto tu zauważyć, że teoria koherencji eksplanacyjnej jest najbliższa poglądom Susan Haack, dla których Haack ukuła termin *fundherentism*. Tworząc syntezę koherentyzmu z fundacjonalizmem, Haack z koherentyzmu bierze wgląd, iż nie istnieją niepowątpiewalne prawdy, a przekonania są uprawomocnione w takim zakresie, w jakim łączą się one z innymi przekonaniem. Z empirycznego fundacjonalizmu bierze zaś pogląd, że nie wszystkie przekonania mają równy wkład w uprawomocnienie przekonań i że doświadczenie zmysłowe zajmuje specjalną, jeśli nie uprzywilejowaną rolę. Fundherentyzm wyrażony jest przez dwie zasady:

- (FH1) Doświadczenie podmiotu jest istotne dla uprawomocnienia jego przekonań empirycznych, ale nie istnieje potrzeba uprzywilejowanej klasy przekonań empirycznych uprawomocnionych wyłącznie przez wsparcie empiryczne, niezależnie od wsparcia innych przekonań;
- (FH2) Uprawomocnienie nie jest wyłącznie jednokierunkowe, ale dotyczy perswazyjnych relacji wzajemnego wsparcia; por. S. Haack, *Evidence and inquiry*, Basil Blackwell Press Oxford 1983, s. 19.

³⁹ Przypomnijmy, że rozumowanie abdukcyjne prowadzi do wniosków, za pomocą których próbujemy nadawać sens zadziwiającym zjawiskom, osiągnąć zrozumienie zaskakujących wydarzeń bądź tłumaczyć sobie budzące zaciekawienie informacje. Por. M. Urbański, *op. cit.*, s. 97.

⁴⁰ Por. P. Thagard, *Conceptual revolutions*, s. 35-37.

twierdzili, że elektryczność i magnetyzm są tym samym i stworzyli połączone pojęcie elektromagnetyzmu. Nowe pojęcia mogą także zostać wprowadzone z racji eksplanacyjnych, np. gdy pojęcie fali dźwiękowej zostało utworzone jako część wyjaśniania tego, dlaczego dźwięki zachowują się tak, jak się zachowują. Pojęcia takie, jak: tlen, elektron, kwark czy gen zostały wprowadzone jako część teorii wyjaśniających, co uzasadnia zmianę typu 6. Z kolei przykładem zmiany typu 7, czyli osłabienia lub zapadnięcia się pewnego rozróżnienia jest osłabienie przez Darwina różnicy: gatunek – odmiana czy Newtonowskie odrzucenie Arystotelesowskiej różnicy między ruchem naturalnym a wymuszonym. Dalej przeskok dziedzinowy ma miejsce w wypadku rewolucji Kopernikańskiej, gdzie „Ziemia” została przerwana do kategorii planet (ze swojej pozycji *sui generis*). W końcu zmiana typu 9, czyli zmiana zasady organizującej drzewo pojęciowe, jest zmianą najbardziej dramatyczną; np. Darwin nie przeklasyfikował ludzi do zwierząt, ale zmienił znaczenie klasyfikacji; przed Darwinem rodzaj był pojęciem wyrażającym pierwotne podobieństwo, po Darwinie rodzaj stał się pojęciem historycznym (posiadanie wspólnego przodka stało się co najmniej tak ważne jak podobieństwo wyglądu).

Trzeba tu dodać, że choć zmiany typu 1–3 mogą być rozpatrywane w kategoriach zmian przekonaniowych, to pozostałe rodzaje zmian tak traktowane być nie mogą, jako że dotyczą hierarchii pojęciowej; szczególnie dotyczy to zmiany zasady organizującej hierarchię:

Darwin nie zaatakował po prostu struktury kreacjonistycznej, ale wybudował alternatywny gmach, który wyparł kreacjonizm w całości. A więc przyjęcie nowego systemu konceptualnego jest zmianą bardziej całościową niż zmiana przekonań dokonywana po kawałku.⁴¹

Przytoczone tu rozważania Thagarda zdają się dostarczać mocnych racji dla uznania pojęcia zmiany pojęciowej za co najmniej tak ważne jak zmiana przekonaniowa. Można zapytać teraz o sposoby formalnego ujęcia zmiany pojęciowej. Thagard proponuje system oparty na swojej koncepcji koherencji eksplanacyjnej. Jest to model koneksjonistyczny, na którym oparł program ECHO (wykorzystujący sztuczną sieć neuronową, w której neuronom odpowiadają jednostki [*units*] połączone przez powiązania [*link*]). Rozwiązanie to, mimo co najmniej dwudziestu pięciu lat funkcjonowania, nie rozprzestrzeniło się poza krąg badaczy z nim związanych. Wskazano też wiele argumentów krytycznych wobec koncepcji koherencji eksplanacyjnej; niektóre z nich umieszczono jako dodatek do głównego artykułu Thagarda *Explanatory Coherence*⁴². Wśród wielu zarzutów na uwagę zasługują następujące:

⁴¹ *Ibidem*, s. 36.

⁴² „Behavioral and Brain Sciences” 12 (1989), s. 435-502. Dyskusja z koncepcją Thagarda znajduje się na stronach 467-490, a dalej są odpowiedzi autora.

- (a) teoria arbitralnie przyjmuje dążenie do maksymalizacji koherencji jako jedyne kryterium akceptowalności hipotez; tymczasem nie mniej ważnym celem naukowców jest dążenie do zwiększenia nowości predykcyjnej hipotezy (czyli jej walorów heurystycznych);
- (b) teoria nie uwzględnia operacji logicznych na hipotezach (inaczej mówiąc, umożliwia ona jedynie badanie pojedynczych hipotez i ich mocy wyjaśniającej. Jak wskazuje L. J. Cohen, wad tych nie mają ujęcia probabilistyczne (np. Bayesowskie) akceptacji;
- (c) teoria zakłada pojęcie „eksplanacja” jako niedefiniowalne;
- (d) system ECHO pracuje na już zwerbalizowanych wynikach zabiegów eksplanacyjnych badaczy: „tym, co Thagard w rzeczywistości modeluje nie jest rozumowanie badaczy, ale struktura ich argumentacji, prezentowana w tym, co można nazwać »kontekst perswazji«”;
- (e) podobnie ECHO nie modeluje procesu myślenia, ale raczej jego rezultat;
- (f) model koneksjonistyczny myślenia przyjęty przez Thagarda nie daje nic szczególnego koncepcji eksplanacjonistycznej, nie wnosi nic do pojęcia wyjaśniania i mógłby być zastąpiony przez modele probabilistyczne.

Podane tu trudności zasługują na uwagę, wskazując że choć sama analiza zmiany konceptualnej dokonana przez Thagarda zasługuje na uwagę, to niekoniecznie trzeba przyjąć proponowane przez niego rozwiązania. Jak się wydaje, dla logika przynajmniej, ważną przeszkodą jest „alogiczność” modelu koherencji eksplanacyjnej, odpornego na działania logiczne. Nie za bardzo widać też, na poziomie działania programu ECHO, w jaki sposób zachować specyfikę zmiany pojęciowej tak, żeby nie można było jej przełożyć na język propozycjonalny. Wydaje się, że lepsze efekty mogłoby dać podejście stosowane w formalnych ontologiach stosowanych, oparte na dogłębnej analizie ontologicznej badanego zjawiska i konceptualizacji jej wyników w języku I rzędu. Ale to zadanie na inny artykuł.

Marek Lechniak

Towards Adequate Theory of Belief Change

Abstract

In this article we deal with the concept of change of belief. Main theories of belief change represent it as a propositional change, i.e. the belief state is represented as a set of propositions and changes are represented as adding, contracting and contracting plus adding (revising) proposition to the set of propositions. Nevertheless, this model has many of defects and we point some of them. Admittedly, there are changes that can't be adequately represented in such a way,

for example conceptual changes; P. Thagard gave some attempts of analysis of that kind of belief changes. We present and criticize some theses of Thagard's theory and try to extract what conditions should be fulfilled the conceptual change be represented in more adequate way.

Keywords: AGM theory, belief change, coherentism, conceptual change, foundationalism.