

RZECZYWISTE WZORCE INTENCJONALNOŚCI

W naturalistycznych nurtach filozofii umysłu żywa jest kwestia pojmowania intencjonalności. Samą intencjonalność rozumie się tradycyjnie jako posiadanie treści lub bycie na jakiś temat (*aboutness*). Pod znakiem zapytania stawiane są jednak tradycyjne tezy na temat niematerialnego charakteru własności intencjonalnych, jej obiektów, związku intencjonalności ze świadomością. Trzy pytania wydają się szczególnie charakterystyczne w tym kontekście: (1) Czy intencjonalność jest rzeczywiście przysługującą własnością, czy też własnością przypisywaną tylko ze względów pragmatycznych? (2) Jakiego rodzaju jest to własność? W szczególności, czy jest to własność redukowalna do własności fizycznych? (3) Jakim obiektom należy przypisywać (przysługują) stany intencjonalne? Na jakich zasadach się to odbywa? (4) Czy własności intencjonalne są redukowalne do własności reprezentacji umysłowych i relacji między nimi?

Na powyższe pytania odpowiada między innymi koncepcja Daniela Dennetta, który jest uznawany za zwolennika tezy o czysto instrumentalistycznym statusie własności intencjonalnych. Początkowe prace tego autora rzeczywiście tego rodzaju wrażenie mogły sprawiać, lecz dziś autor ten bywa zaliczany do obozu realistów²⁸. Zamierzam niżej pokazać, iż Dennett w istocie pozostaje realistą co do stanów intencjonalnych, a jego koncepcja zawiera odpowiedzi na pytania, które postawiłem powyżej.

Zacznę od przedmiotu sporu o realność przekonań i pragnień. Owe przekonania i pragnienia mają być realnie istniejącymi wzorcami, dlatego też przeanalizuję przykład – grę w życie – na którym Dennett eksplikuje znaczenie tezy o realności istnienia wzorców. Na tym przykładzie pokażę też, iż wzorce intencjonalne są emergentne względem wzorców fizycznych i funkcjonalnych (do wzorców funkcjonalnych zaliczy Dennett reprezentacje umysłowe). Owe

²⁸ N. Nelkin, *Consciousness and the Origins of Thought*, Cambridge 1996, s. 196 i nn.; A. Clark, *Mindware: An Introduction to the Philosophy of Cognitive Science*, Oxford 2000, s. 50–60.

wzorze okażą się w koncepcji Dennetta algorytmami kompresji informacji o świecie – intencjonalność jest tedy realnym wzorcem, który jest emergentny (i jednocześnie mikroredukowalny) w stosunku do wzorców funkcjonalnych i fizycznych. Jego zdaniem intencjonalność przypisywana może być zasadnie tylko wówczas, gdy rzeczywiście jest emergentna, czyli gdy opis w kategoriach przekonaniach i pragnień daje nowe predykcje w stosunku do predykcji możliwych przy użyciu opisu w kategoriach fizycznych lub funkcjonalno-konstrukcyjnych.

Spór o realność przekonań i pragnień

Czy własności intencjonalne realnie istnieją? A jeśli istnieją, to jakiego rodzaju są własnościami? Wydawałoby się, że przekonania i pragnienia Dennett traktuje w sposób instrumentalistyczny. We wczesnych pracach podkreślał, że służą one predykcji zachowania pewnych systemów, co nie znaczy, iż mają jakieś konkretne odniesienie przedmiotowe w tych systemach. Otóż obserwator przyjmuje wobec nich tzw. nastawienie intencjonalne – to znaczy stara się je opisywać i przewidywać za pomocą przypisywanych im przekonań i pragnień. Niemniej korelaty tych przekonań i pragnień w postaci jednostkowych obiektów – zdaniem autora *The Intentional Stance* – istnieć nie muszą, aby takie predykcje były skuteczne. Antyrealistyczny wydzźwięk mają deklaracje Dennetta dotyczące jego teorii intencjonalności. Gdyby psychologia nastawienia intencjonalnego miała być redukcją zachowawczą psychologii potocznej, to nie jest to redukcja zachowująca wszystkie podziały kategoriałne zawarte w językach naturalnych. Autor *Brainstorms* dopuszcza, że część intencjonalnych predykatów potocznych okaże się nazwami pustymi. Taki pogląd można pociągnąć w dwóch (co najmniej) kierunkach:

- Nie ma przekonań i pragnień jako konkretnych obiektów fizycznych takich jak cząstki elementarne czy obiektów teoretycznych takiego rodzaju, jak fale radiowe; przekonania jako korelaty zdań w języku myśli przetwarzanym przez mózgi nie istnieją.
- Nie ma przekonań i pragnień jako stanów intencjonalnych aktywnych przyczynowo; przekonania i pragnienia traktuje się fikcjonalnie.

Pierwszy pogląd stanowiłby negację koncepcji języka myśli Jerry'ego Fodora; drugi zaś – zasługiwałby na miano eliminatywizmu, głoszonego chociażby przez Paula i Patricję Churchlandów²⁹ czy Richarda Rorty'ego w je-

²⁹ P. M. Churchland, *The ontological status of intentional states: Nailing folk psychology to its perch*, „The Behavioral and Brain Sciences” 1988, nr 11 (3), s. 507–508; *Problem umysłu i ciała*, tłum. T. Baszniak, w: B. Chwedeńczuk, *Fragmenty filozofii*

go wczesnych artykulach³⁰. Paul Churchland wypowiada się w sposób nie pozostawiający żadnych wątpliwości: „Przekonania i pragnienia dzielą los (*are of a piece with*) flogistonu, ciepłika i esencji chemicznych”³¹. Dennett jednak nie idzie tak daleko. Przyznaje przekonaniom i pragnieniom tyleż realności, co środkom ciężkości. Przyznaje, że przekonania nie są konkretami. Określa je w terminologii Reichenbacha jako *abstracta*, obiekty abstrakcyjne – w przeciwieństwie do illatów, postulowanych obiektów teoretycznych³². Przypomnijmy, co Reichenbach pisał na temat tych kategorii ontologicznych: „Dla takich bytów [jak np. elektryczność – *M.M.*] używam nazwy *illata*, co znaczy „rzeczy wywnioskowane”. Są one ściśle odróżnione z jednej strony od *konkretów*, które składają się na świat rzeczy obserwowalnych, z drugiej zaś od *abstraktów*, które są połączeniami konkretów i nie dają się obserwować bezpośrednio, ponieważ są całościami pojęciowymi [...]. Illata natomiast nie są połączeniem konkretów, lecz odrębnymi bytami, wywnioskowanymi z konkretów. Ich istnienie jest tylko uprawdopodobnione przez konkrety”³³.

Abstrakty, w przeciwieństwie do illatów, nie mogą mieć przypisanej lokalizacji przestrzennej³⁴. Illata znajdują się natomiast w przestrzeni, lecz nie są bezpośrednio obserwowalne. I tak na przykład abstraktami są liczby, kryzysy finansowe, illatami zaś – siła, elektryczność, fale radiowe, opór elektryczny. Przez „połączenia konkretów” rozumie Reichenbach z pewnością nie tyle fizyczne złączenia co najmniej dwóch konkretów (to byłyby wszak tylko konkrety), ile raczej coś w rodzaju kompleksów czy związków między konkretami. Oznacza to, że są one definiowane za pomocą relacji między konkretami.

analizycznej, tom II: *Filozofia umysłu*, Warszawa 1995; P. S. Churchland, *Dennett's instrumentalism: A frog at the bottom of the mug*, „The Behavioral and Brain Sciences” 1983, nr 3, s. 358–359.

³⁰ R. Rorty, *Idenryczność ciała i umysłu, prywatność a kategorie ontologiczne*. tłum. M. Szczubialka, w: B. Chwedeńczuk, *op. cit.*

³¹ P. M. Churchland, *The ontological...*, *op. cit.*, s. 508.

³² D. Dennett, *Brainchildren: Essays on Designing Minds*, Cambridge (Mass.) 1998, s. 96–97.

³³ H. Reichenbach, *Powstanie filozofii naukowej*, tłum. H. Krahelska, Warszawa 1960, s. 272.

³⁴ Jest to spory problem dla Dennetta, gdyż środki ciężkości i południki mają, jak się zdaje, przypisywaną lokalizację! Stąd być może powyższe uwagi Churchlanda, interpretującego środki ciężkości jako illata. Być może jednak Dennett nielokalizowalności abstraktów nie uznaje za ich cechę definicyjną, w przeciwieństwie do Reichenbacha. Może bowiem uznać, że illata z konieczności są ściśle lokalizowalne, zaś abstrakty stanowią obiekty wskazujące lokalizację illatów (lecz same nielokalizowalne). I tak południk wskazuje położenie obiektów geograficznych, lecz sam nie ma żadnej lokalizacji, gdyż trudno mówić np. o jego powierzchni.

Oponenci Dennetta zarzucali mu m.in. brak argumentów na rzecz regulującej definicji przekonań jako abstraktów. Alexander Rosenberg zwracał uwagę, że na podstawie przydatności predykcyjnej równie dobrze można by wnioskować o tym, że przekonania mają status illatów³⁵. Przekonania i inne stany intencjonalne, co gorsza, nie są tak dobrze zdefiniowane, jak abstrakty w fizyce: problem polega na tym, że wcale *dokładnie* nie wiadomo, jak są powiązane przekonania ze stanami układu nerwowego, illatami neuronauki; jeśli zaś nie wiadomo tego w sposób precyzyjny, to nazwanie przekonań abstraktami jest czczą deklaracją³⁶. Dennett w odpowiedzi dowodził, że w *zasadzie* wiadomo, jak abstrakty-przekonania są powiązane z illatami: są one mianowicie w takiej relacji, jak tak zwany poziom wiedzy do poziomu symboli (programu) w kognitywistycznej teorii umysłu Alana Newella.

Metafora wielopoziomowej analizy czy wyjaśniania zjawisk zasługuje na sporą uwagę. Rzeczywiste wzorce w sensie Dennetta należy analizować na wielu poziomach. Niektórzy filozofowie twierdzą, że typowe dla emergentyzmu metafory warstw czy poziomów są sprzeczne z ujęciem rzeczywistości jako zawierającej wzorce³⁷. Jak będę się starał pokazać, tego rodzaju ujęcie nie tylko jest zgodne z fizykalizmem, ale również z łagodnym emergentyzmem – nie ma mowy o żadnych sprzecznościach, nawet w metaforycznym sensie (niektórzy inni autorzy też optują za wielopoziomową, emergentystyczną wizją umysłu)³⁸.

Wzorce, o których mówi Dennett, to regularności występujące w ciągach danych. Są one czymś, co może być – z definicji – rozpoznane jako wzorzec³⁹. Niewykrywalne regularności czy wzorce są więc uznawane za coś, co nie może być wzorcem, albo nie jest w istocie wzorcem w sensie właściwym tego słowa⁴⁰. W jaki sposób rozpoznaje się regularności w danych? I czy jeden zbiór danych należy opisywać jako zawierający jeden wzorzec czy też wiele wzorców?

³⁵ A. Rosenberg, *Will the argument for abstracta please stand up?*, „The Behavioral and Brain Sciences” 1988, nr 11 (3), s. 526–527.

³⁶ *Ibidem*, s. 527.

³⁷ D. Stoljar, *Physicalism*, w: E. N. Zalta (red.), *The Stanford Encyclopedia of Philosophy (Spring 2001 Edition)*, URL = <<http://plato.stanford.edu/archives/spr2001/entries/physicalism/>>.

³⁸ Por. U. Żegleń, *Filozofia umysłu. Dyskusja z naturalistycznymi koncepcjami umysłu*, Toruń 2003.

³⁹ D. Dennett, *Brainchildren...*, *op. cit.*, s. 100.

⁴⁰ Haugeland (J. Haugeland, *Pattern and Being*, w: B. Dahlboom, *Dennett and His Critics*, Blackwell, 1993 s. 58) twierdzi, że u Dennetta są dwie definicje wyrażenia „wzorzec”: (a) to, co może zostać rozpoznane; (b) coś, co jest w porządku nielosowym; Haugeland dodaje, że obie definicje tworzą jedno pojęcie. Definicja operacyjna (a) ma być

Odpowiadając na te pytania, Dennett nawiązuje do definicji ciągu losowego⁴¹ Gregory'ego Chaitina⁴²:

(CL) Ciąg liczb jest losowy, jeśli najmniejszy algorytm określający go komputerowi ma taką samą długość w bitach, co sam ciąg.

Jak widać z definicji, niekompresowalne ciągi są losowe (natomiast problem, czy dany ciąg jest losowy, jest nierozstrzygalny – jak dowiódł także Chaitin). Przez kompresję danych rozumie się w informatyce zmniejszenie długości ciągu kompresowalnego – to jest zapisanie go w krótszej formie, którą można rozkodować przy użyciu odpowiedniego algorytmu dekompresującego. Wyróżnia się dwa rodzaje kompresji: bezstratną, gdzie po dekompresji dane wyjściowe są identyczne z wejściowymi, oraz stratną, gdzie pomija się część informacji, które są dla danego zastosowania nieistotne (np. niesłyszalne dla ludzkiego ucha elementy nagrania muzycznego). Pojęcie kompresji danych jest więc kluczowe dla definicji wzorca u Dennetta: rzeczywiste wzorce to dokładnie kompresowalne ciągi. Ciąg kompresowalny ma z definicji pewien algorytm, który go kompresuje i dekompresuje; jest rozpoznawalny przez ten algorytm kompresujący – kryterium rozpoznania jest tutaj skompresowanie ciągu. W ten sposób Dennett może bronić tezy, że w istocie w ogóle nie ma wzorców nierozpoznawalnych.

O dwóch rodzajach kompresji wspomniałem nie bez kozery. Kompresję stratną stosuje się wtedy, gdy część kompresowanych danych można uznać za szum, niepotrzebny dodatek do wartościowych informacji. Odrzucając szum, można zapisać tylko wartościowe informacje. Skoro tak, to istnieje duża dowolność interpretacji wzorców w ciągach danych: nie dość, iż można szukać dowolnej metody kompresji, to jeszcze można dowolne części danych traktować jako szum. Oczywiście, potraktowanie 90% danych jako zbędnego szumu może budzić wątpliwości, lecz gdyby była to cena za zrekonstruowanie starego nagrania – na przykład przemówienia Józefa Piłsudskiego – to przy-

niepotrzebna, skoro istnieje definicja teoretyczna (b). Sądzę jednak, iż (a) jest jedynie wstępnym opisem wzorca, zaś (b) – dopiero jego definicją właściwą. Dennett zauważa, że definicja losowości rozwiązuje pozorny problem Haugelanda (D. Dennett, *Bach from the drawing board*, w: B. Dahlboom, *op. cit.*, s. 213).

⁴¹ D. Dennett, *Brainchildren...*, *op. cit.*, s. 103.

⁴² Definicja Chaitina powstała niezależnie od analogicznych definicji Kolmogorowa i Solomonoffa. W matematyce rozwiniętymi już dziś dziedzinami są zarówno teoria złożoności Kolmogorowa (korzystająca z tej definicji), jak i algorytmiczna teoria informacji Chaitina. Definicja Chaitina w postaci ścisłej wymaga dokładnego podania metody kodowania liczb naturalnych (w których kodowalne są wszelkie inne komputerowe struktury danych); zdefiniowania języka programowania dla odpowiedniej maszyny, na której będzie realizowany algorytm. Dla moich celów jest to niepotrzebne: nie zajmuję się matematycznymi szczegółami tej koncepcji.

znamy, iż jest to operacja rzeczywiście polegająca na odrzucaniu szumu i pozostawianiu wartościowych informacji.

Wzorcy w świecie „gry w życie”

Należy zdać sobie sprawę z niezwykle szerokiego zakresu pojęcia „wzorec”, czyli kompresowalnego ciągu danych. Dennett przedstawia szereg przykładów wzorców. Nawiązuje m.in. do tak zwanej „gry w życie”. Pokróćce przedstawię więc zarys zasad „gry w życie”, a następnie dosyć dokładnie zajmę się Dennettowską argumentacją opartą na tym przykładzie – pozwoli mi to wrócić do sprawy relacji między illatami a abstraktami w odniesieniu do stanów intencjonalnych. Gra w życie to gra logiczna w rodzaju pasjansa. Nie chodzi w niej o wygraną. Istnieje plansza podzielona na kwadratowe komórki, które mogą być zaczernione lub białe: na planszy istnieje zawsze pewna konfiguracja początkowa. Skoro komórki są kwadratowe, to każda z nich ma osiem komórek sąsiednich. W każdym ruchu w grze realizuje się jedną zasadę: każda komórka, aby określić czynność wykonywaną w następnym ruchu, liczy, ile komórek jest zaczernionych w chwili obecnej. Jeśli odpowiedź wynosi dwa, komórka pozostaje w stanie obecnym (zaczerniona lub biała) w następnym ruchu. Jeśli odpowiedź wynosi trzy, komórka będzie zaczerniona w następnym ruchu, bez względu na jej stan obecny. W pozostałych okolicznościach komórka jest biała.

Skąd nazwa „gra w życie”? Otóż, zaczernianie i wybielanie komórek można rozumieć jako narodziny i śmierć; obserwowanie planszy zmieniającej się zgodnie z tymi zasadami daje wrażenie powstawania kolejnych generacji komórek i ich kolonii. Gra więc należy do kategorii gier symulacyjnych – gier przypominających rzeczywiste procesy życiowe⁴³. Conway, tworząc grę w życie, tak dobrał zasady, aby trudno było przewidzieć zachowanie populacji komórek, a jednocześnie żeby było ono interesujące. Początkowo w grę w życie grano za pomocą papierowych plansz i kwadratowych, dwukolorowych pionów, lecz obecnie najczęściej do tego celu wykorzystuje się komputerowe symulacje.

W zależności od wybranej konfiguracji początkowej następują różne zmiany układu na planszy. Niektóre układy są stabilne i w ogóle się nie zmieniają – Conway określa je jako „martwą naturę”. Inne, początkowo niesymetryczne, stają się układami symetrycznymi. Po wielu ruchach niektóre

⁴³ M. Gardner, *The fantastic combinations of John Conway's new solitaire game „life”*. „Scientific American” 1970, nr 223, s. 120–123.

całkowicie wymierają, inne oscylują między kilkoma konfiguracjami. Wykryto szereg ruchomych figur na planszy; figur „żarlocznych”, figur stabilnych i tak dalej. Starannie dobierając początkowe rozłożenie komórek zaczerpionych, można stworzyć frapujące układy. Wszystkie te układy, oczywiście, są konsekwencją przyjętej zasady gry w życie i konfiguracji początkowej.

Jednak przewidzenie zachowania komórek okazuje się bardzo trudne bez obserwacji symulowanych komputerowo układów. Przez obserwację odkryto stabilne, przemieszczające się figury – nazywane malowniczo „szybowcami”, „pompami”, „ciuchciami”. Stanowią one nowy poziom ontologiczny⁴⁴; zamiast mówić o jednej zasadzie zaczerpnienia i wybielania komórek na planszy, można się zastanawiać, jak ułożyć na przykład szybowiec, żeby doleciał do pompy. Na tym poziomie powstaje nowy język – zawierający swoiste nazwy, skróty drobiazgowych opisów konfiguracji. Na poziomie fizyki świata gry w życie nie mówi się o ruchu ani o szybowcach. Planszę opisuje się statycznie w postaci współrzędnych komórek zaczerpionych, traktowanych jak indywidua; na poziomie wyższym pojawiają się ruchome pompy i szybowce. Czy w świecie gry w życie ruch jest *rzeczywisty*? Czy tylko *pozorny*?

Mamy do czynienia z równie realnymi wzorcami – na obu poziomach – fizyki świata życia i jego – by tak rzec – makrofizyki. Przewidywania, w których indywiduami są szybowce i pompy, są mniej ściśle od przewidywań na poziomie współrzędnych⁴⁵, lecz znacznie wygodniejsze. Pojawiają się tutaj prawa przybliżone dotyczące makrofizyki świata życia, czyli algorytmy kompresji stratnej informacji o konfiguracji komórek na planszy. Przy owej kompresji tracimy część informacji – mianowicie przechodzimy do nastawienia funkcjonalnego względem obiektów na planszy. Pytamy, jak działają owe ruchowe konfiguracje.

Nie jest do końca jasne, jak Dennett chce uzasadnić tezę o mniejszej ściśłości praw dotyczących figur świata życia (mniejszej od ściśłości prawa fizyki tego świata). Mniej wprawne oko dłużej musi wylawiać pojawiające się na planszy obiekty, lecz gdy już je wyróżnimy, to działają one zawsze tak samo – zgodnie z podstawowym prawem fizyki świata życia. Niemniej jednak prawidłowości dotyczące szybowców i ciuchów zwykle nie są tak ściśle i precyzyjnie wyrażane, co opisy konfiguracji komórek na planszy, lecz zupełnie wystarczają projektantom ciekawych układów w grze w życie.

⁴⁴ D. Dennett, *Brainchildren...*, op. cit., s. 107.

⁴⁵ *Ibidem*, s. 108.

Pojawienie się nowego poziomu w grze w życie – czy w jej naukowej kontynuacji, sztucznym życiu – należy określić mianem „emergencji”⁴⁶. Poziom emergentny w świecie życia jest całkowicie uzależniony od poziomu podstawowego – w końcu nie trzeba mówić o zjawiskach emergentnych, aby określić konfigurację na planszy w kolejnym ruchu. Niemniej trudności związane kontekstowością, nieliniowością i lokalnością mikrooddziaływań są na tyle duże, że przewidywanie wyglądu planszy – bez obliczania jej układu, a jedynie na podstawie ogólnej znajomości interakcji między makrofigurami – jest praktycznie beznadziejnie trudne. Ryzyko epistemiczne tkwiące w prawach wyższego poziomu bierze się stąd, iż bez obserwacji symulowanych konfiguracji nie potrafimy przewidzieć zachowania pewnych obiektów *a priori*. Czy jednak dokładna empiryczna obserwacja przyniesie nam rzetelne prawa wysokiego poziomu – bez żadnych klauzul *ceteris paribus* (przy innych warunkach niezmiennych)? Otóż tego rodzaju prawa przestałyby być, automatycznie, prawami wysokiego poziomu. Zysk płynący z ujęcia komórek jako pewnego rodzaju obiektów złożonych wiąże się z ignorowaniem wielu możliwych sytuacji, w których te komórki przestaną zachowywać się jako „ciuchcie” czy „bomby”: wypowiedzi na temat „szybowców” z natury swej muszą być wypowiedziami *ceteris paribus*. Dlatego właśnie poziom figur w grze w życie jest poziomem realnym, naprawdę emergentnym, a nie tylko epifenomenem⁴⁷.

Odpowiednio, tworząc konfigurację świata życia, można uzyskać dosyć zaskakujące rezultaty. Możliwe jest mianowicie zbudowanie automatu, który powielalby się samoczynnie, a jednocześnie stanowiłby realizację uniwersalnej maszyny Turinga, czyli tego modelu komputera cyfrowego, który jest w stanie wykonać każdy algorytm dla tego rodzaju komputerów⁴⁸. Na planszy świata życia może powstać uniwersalny komputer, choć oczywiście plansza będzie musiała być bardzo dużych rozmiarów.

Von Neumann pokazał już, że w zasadzie dwuwymiarowa uniwersalna maszyna Turinga może być utworzona z automatów komórkowych, więc „tylko” problemem „inżynierskim” było pokazanie, jak w zasadzie utworzyć je z prostszych automatów komórkowych zdefiniowanych w świecie życia. Strumienie szybowców mogą na przykład stanowić taśmę, a czytelnikiem taśm może być ogromny moduł złożony z martwej natury, pożeraczy, szybowców

⁴⁶ M. Bedau, *The scientific and philosophical scope of artificial life*, „Leonardo” 2002, t. 35, nr 4, s. 395–400.

⁴⁷ Dokładniejsza definicja nazwy „emergencja” znajduje się niżej. W tym miejscu dla uproszczenia można powiedzieć, że w grze w życie mogą występować zjawiska nieliniowe w relacjach część (komórki planszy) – całość (oddziałujące figury), a zmiany ilościowo znikome prowadzą do ogromnych zmian jakościowych.

⁴⁸ S. Franklin, *Artificial Minds*, Cambridge Mass., 1995.

i innych elementów. Jak to wygląda? Poundstone wylicza, że cała konstrukcja, autoreprodukująca się maszyna ucieleśniająca uniwersalną maszynę Turinga, byłaby wielkości 10^{13} pikseli⁴⁹.

Uniwersalna maszyna Turinga może obliczyć każdą funkcję obliczalną (jeśli przyjmiemy, jak większość matematyków, tezę Churcha), tak więc może ona posłużyć – jak powiada Dennett – do realizacji programu gry w szachy. Konfiguracja komórek na planszy świata życia może grać z nami w królewską grę. Problemem jest tylko rozpoznanie przez nas, w jaki sposób owa maszyna reprezentowałaby swoje wejście i wyjście, tj. w jaki sposób przedstawiałaby sobie i nam układ figur szachowych na szachownicy. Ktoś, kto nie wie, że uniwersalna maszyna Turinga została zrealizowana przez gigantyczną konfigurację na planszy świata życia, nie byłby w stanie łatwo odgadnąć, że migoczące, mgliste obrazy złożone z miliardów pikseli to komputer grający w szachy.

Lecz z perspektywy osoby, która stawia hipotezę, że ta ogromna tablica czarnych punktów jest komputerem szachowym, dostępne są niezwykle skuteczne metody przewidywania przyszłości tej konfiguracji. Pierwszym krokiem może być przejście z ontologii szybowców i pozeraczy do ontologii symboli i stanów maszyny. Przyjmując to wyższe nastawienie funkcjonalne do konfiguracji, można przewidywać jej przyszłość jako maszyny Turinga. Drugim i jeszcze skuteczniejszym krokiem może być przejście do ontologii pozycji na szachownicy, możliwych ruchów szachowych i racji ich oceny; następnie, przyjmując nastawienie intencjonalne do konfiguracji, można przewidzieć jej przyszłość jako szachisty podejmującego intencjonalne działania – wykonującego ruchy w szachach i usiłującego uzyskać mata⁵⁰.

Dennett znowu przypomina o ryzyku tkwiącym w obu krokach – z jednej strony program szachowy, interpretowany intencjonalnie, może być niecałkowicie racjonalny; z drugiej zaś – konfiguracja komórek stanowiąca maszynę Turinga może zostać uszkodzona, jeśli na planszy znajduje się jakiś osamotniony „pozeracz” czy inna wędrująca figura, która zmieni układ komórek maszyny⁵¹. To ryzyko nie jest szczególnie dla gry w życie – wiąże się w ogóle z intencjonalnym interpretowaniem komputerowego programu grającego w szachy: program może być źle napisany, a komputer może się zepsuć.

Wzorce w grze w życie są rzeczywiste i występują na wielu poziomach. Ich wykrycie wymaga przyjęcia niekiedy śmiałych hipotez – jak w wypadku interpretacji ogromu migających komórek jako maszyny Turinga grającej w szachy. Dennett dodaje, że osoba stojąca przed ogromną planszą gry w życie jest

⁴⁹ D. Dennett, *Brainchildren...*, *op. cit.*, s. 108–109.

⁵⁰ *Ibidem*.

w położeniu Marsjanina, który potrafiłby przewidywać i interpretować działania ludzkie wyłącznie na poziomie mikrofizycznym: dopóki Marsjanin nie wpadnie na to, że istnieje intencjonalny poziom opisu, dopóty będzie stał przed niezwykle trudnym zadaniem obliczania toru każdej cząstki.

Wielopoziomowa ontologia świata życia: emergencja

Pojęcia emergencji i redukcji są stosunkowo nieostre; czasem uznaje się je za przeciwstawne i wykluczające, a z pewnością można mówić o wielu odmianach emergencji czy redukcji⁵². Tutaj przyjmę definicje regulujące za Williamem Wimsattem, który twierdzi, że emergencja i redukcja nie wykluczają się⁵³. Jego definicje tych pojęć są więc nietypowe, lecz unikają wady, która tkwi w koncepcjach standardowych: pojęcia te mają rzeczywiste egzemplifikacje w praktyce naukowej⁵⁴. Wimsattowi nie idzie o możliwość redukcji w zasadzie, lecz o taką racjonalną rekonstrukcję nauki, w której jest miejsce dla pojęć redukcji i emergencji.

Wimsatt w swojej analizie zauważa, że można wyróżnić dwa podstawowe znaczenia emergencji: zgodne i niezgodne z redukcją⁵⁵. I tak emergencja, układu *U* to taka jego własność, że nie można zidentyfikować lub wyjaśnić własności *U* wyższego poziomu za pomocą lub w kategoriach konfiguracji części niższego poziomu *U* oraz ich znanych własności monadycznych i relacyjnych. Jest to więc *nieredukowalność* własności wyższego poziomu do własności niższego poziomu. Natomiast emergencja, zakłada możliwość międzypoziomowej identyfikacji własności (a właściwie mikroredukcji)⁵⁶. Własności relacyj-

⁵² W. Strawński, *Jedność nauki, redukcja, emergencja. Z metodologicznych i ontologicznych problemów integracji wiedzy*. Warszawa 1997; R. Poczobut, *O odmianach emergencji*, „Roczniki Filozoficzne” 2002, t. 50, nr 1.

⁵³ W. Wimsatt, *Emergence as non-aggregativity and the biases of reductionism(s)*, „Foundations of Science” 2000, nr 5, s. 269–297. Odwołanie do tego autora jest również usprawiedliwione zasadami interpretacji tekstów: Dennett i Wimsatt często powołują się wzajemnie na swoje prace, choć nie zawsze dotyczą one problematyki redukcji i emergencji (W. Wimsatt, *Reductionism, levels of organization and the mind-body problem*, w: G. Globus, I. Savodnik i G. Maxwell (red.), *Consciousness and the Brain*, New York 1976, s. 199–267; D. Dennett, *Brainchildren...*, *op. cit.*, s. 104). Dennett nie wykluczył jako błędnej eksplikacji jego idei za pomocą koncepcji Wimsatta (Dennett, korespondencja prywatna 2004).

⁵⁴ W. Wimsatt, *Reductionism*, *op. cit.*; *idem*, *Reductive Explanation: A Functional Account*, w: *PSA-1974*, A. C. Michalos, C. A. Hooker, G. Pearce i R. S. Cohen (red.) [*Boston Studies in the Philosophy of Science*, 30], Reidel, Dordrecht 1976, s. 671–710.

⁵⁵ W. Wimsatt, *Reductionism...*, *op. cit.*, s. 208.

ne części w identyfikacji składników nie są wystarczająco symetryczne, proste i konserwatywne, aby uznać identyfikację składników różnych poziomów za „sumującą się” czy „agregacyjną”⁵⁷. Lista warunków pozwalających uznać, że układ jest złożony w sposób „sumujący się”, odpowiada przynajmniej częściowo różnym sposobom eksplikacji emergencji. Układ musi spełniać następujące warunki, aby uznać, że jego własności „sumują się”⁵⁸:

1. Podstawienie lub przestawienie takich samych części układu nie powoduje zmiany własności całego układu.
2. Występuje zachowanie jakościowego podobieństwa własności układu (tożsamość lub zmiana wartości w wypadku własności ilościowej) przy dodawaniu i usuwaniu części.
3. Występuje niezmienność własności układu przy operacjach polegających na rozkładaniu i ponownym sumowaniu części.
4. Występuje liniowość – brak oddziaływań między częściami systemu, które wpływałyby na tę własność układu.

Zawsze i tylko wtedy, gdy nie jest spełniony choć jeden z warunków 1... 4, mamy do czynienia z własnością emergentną (i emergencją). Wimsatt dodaje, że istnienie powszechnej i częstej agregacyjności jest złudne – jest to najczęściej wytwór niesprawdzonych i założonych z góry stałych, idealizacji i przeoczonych możliwości wariacji. Jak pokazują badania w teorii chaosu, układów dynamicznych jest niesłychanie wiele, lecz ich budowa wcale nie przeczy mechaniczności czy fizyczności⁵⁹. Można przypuszczać, że niededukowalność, o której mówili emergentyści, przy jednoczesnej determinowalności⁶⁰, daje się wyeksplikować poprawnie właśnie przy użyciu pojęć z teorii układów dynamicznych. Dopiero wówczas pojęcie emergencji wolne byłoby od groźby sprzeczności wewnętrznej, która pojawiałaby się, gdyby mikrodedukowalność całości do części pociągała za sobą dedukowalność cech całości; w układach nieliniowych takiej możliwości nie ma.

Bez definicji pojęcia poziomu tezy o międzypoziomowej możliwości czy niemożliwości identyfikacji części układu są słabo uchwytne. Jakie jednak kryterium można tutaj wskazać? Wimsatt dowodzi, że poziomy organizacji należy charakteryzować jako „lokalne maksima regularności i przewidywalności w przestrzeni fazowej różnych sposobów organizacji materii”⁶¹.

⁵⁶ Por. W. Strawiński, *op. cit.*, s. 137.

⁵⁷ „Emergencję₂” nazywam dla uproszczenia w pozostałych częściach pracy po prostu „emergencją”.

⁵⁸ W. Wimsatt, *Emergence...*, *op. cit.*, s. 269–297.

⁵⁹ M. Tempczyk, *Teoria chaosu a filozofia*, Warszawa 1998.

⁶⁰ W. Strawiński, *op. cit.*, s. 194.

⁶¹ W. Wimsatt, *Reductive Explanation...*, *op. cit.*, s. 680.

Sformułowanie to odwołuje się do teorii układów dynamicznych, gdzie przestrzeń fazowa (nazywana też przestrzenią stanów) to jedno z podstawowych pojęć⁶². Każdy układ dynamiczny można opisać za pomocą odpowiedniej liczby zmiennych (najczęściej w równaniach różniczkowych). Abstrakcyjna przestrzeń powstała z możliwych wartości zmiennych nosi właśnie nazwę „przestrzeni fazowej” czy „przestrzeni stanów” układu. Zachowanie układu można opisać jako zbiór trajektorii w przestrzeni fazowej; kiedy zacznie się ono w określonym stanie, jego ewolucja dynamiczna utworzy ciąg stanów będący trajekcją w przestrzeni fazowej. Trajektorie w systemie dynamicznym zwykle są zbieżne w co najmniej jednym obszarze w przestrzeni fazowej; te obszary nazywa się „atraktorami”, zaś obszary przestrzeni fazowej, w których zbiegają się wszystkie trajektorie, nazywa się „basenami atrakcji” danego atraktora⁶³. Za pomocą teorii układów dynamicznych wyjaśnia się więc pojęcia poziomów: Wimsatt wskazuje, że w przestrzeni stanów dowolnego układu dynamicznego mogą występować lokalne maksima regularności i przewidywalności – owe obszary będą poziomami organizacji układu. Dodaje, że siły doboru naturalnego (a na niższych poziomach rozmaite czynniki związane ze stabilnością układu) sprawiają, iż większość łatwo definiowalnych obiektów znajdzie się w sąsiedztwie (przestrzeni fazowej) poziomów organizacji, a najprostsze i najsilniejsze teorie będą dotyczyły obiektów na tych poziomach⁶⁴.

Regularność okazuje się tedy jednym z elementów definiujących poziomy ontologiczne w kategoriach układów dynamicznych. Jeśli zaś przez „regularność” będziemy rozumieć „nielosowość”, to zdefiniujemy ją też w kategoriach algorytmicznej teorii informacji. Regularne okażą się te dane, które będzie można efektywnie skompresować (nielosowe w sensie Chaitina). Przewidywalność zaś występuje tam, gdzie znajdziemy algorytm kompresji owych ciągów i będziemy w stanie go ekstrapolować na nowe próbki danych. Istnienie algorytmu kompresji dla danego ciągu danych jest równoznaczne z istnieniem mechanizmu wykrywającego regularność w ciągu danych, czyli pewnego detektora. „Detektor” rozumiem jako urządzenie dające binarną odpowiedź na zadane wejście. Takim detektorem może być nawet model matematyczny (abstrakcyjny), o ile zostanie zrealizowany przez odpowiedni układ

⁶² M. Tempczyk, *op. cit.*, s. 33–41.

⁶³ M. Silberstein, J. McGeever, *The search for ontological emergence*, „Philosophical Quarterly” 1999, t. 49, nr 195, s. 182–201.

⁶⁴ Ujęcie Wimsatta stanowi istotne uzupełnienie dla prostszego kryterium, opartego jedynie na rozmiarze obiektów występujących na określonych poziomach organizacji (W. Strawiański, *op. cit.*, s. 152–153).

fizyczny⁶⁵. Idąc dalej tym tropem, możemy opisywać poziomy organizacji układu jako powstające dzięki hierarchicznej strukturze detektorów, gdzie detektor poziomu n otrzymuje wejście tylko z detektora poziomu $n-1$ (w takiej organizacji jednak mamy do czynienia z liniową hierarchią detektorów: to założenie może być niekiedy fałszywe). Jak zauważają Bonabeau i Dessalles, pojęcie emergencji jest blisko sprzęgnięte z pojęciem złożoności algorytmicznej (złożoność algorytmiczną, czyli złożoność Kolmogorowa definiuje się przez długość najkrótszego algorytmu rekonstruującego dany ciąg liczb) i względne wobec algorytmu posiadanego w danym czasie przez obserwatora. Emergencję wykrywa się wówczas, gdy przewidywane zachowanie układu odbiega od zachowania faktycznego, a faktyczne zachowanie można wykryć, tj. opisać algorytmicznie, nowym algorytmem.

Opis planszy gry w życie w kategoriach figur, a nie tylko współrzędnych zaczernionych komórek, jest istotnie prostszy – mniej złożony algorytmicznie, gdyż dane można bardziej skompresować. Ta konsekwencja stanie się jaśniejsza, gdy zauważymy, że detektor (algorytm kompresujący) każdego poziomu wyróżnianego przez Dennetta jako wejście przyjmuje tylko dane z bezpośrednio niższego poziomu, czyli operuje na mniejszym zbiorze danych, na wyjściu dając skompresowany zapis (czyli jeszcze krótszy). Emergentne poziomy opisu świata życia okazują się więc produktem algorytmów opisu tego świata; wykryte w danych wzorce są realne, gdyż istnieją dla nich odpowiednie algorytmy. Nie da się jednak wykluczyć, że jeden sposób opisu (jeden algorytm) będzie równie dobry, jak inny. Mamy wszak dowolność uznawania czegoś za wartościowe dane, a czego innego – za szum⁶⁶. Dane, które kompresujemy, mogą być równie dobrze kompresowane na kilka sposobów – w taki sposób Dennett parafrazuje Quine'owską tezę o niezdeteminowaniu przekładu⁶⁷.

⁶⁵ E. Bonabeau, J. L. Dessalles, *Detection and emergence*, „Intellectica” 1997, t. 2, nr 25, s. 85–94.

⁶⁶ D. Dennett, *Brainchildren...*, *op. cit.*, s. 115–116.

⁶⁷ Mam jednak wątpliwości, czy Dennett rzeczywiście w tej parafrazie ujmuje intencje Quine'a, zwłaszcza że nie rozumie „niezdeteminowania przekładu” w żadnym z sensów proponowanych w – poprawnych moim zdaniem – analizach Zabłudowskiego (A. Zabłudowski, *On Quine's indeterminacy doctrine*, „Philosophical Review” 1989, nr 98); D. Dennett, *Brainchildren...*, *op. cit.*, s. 118.

Realność wzorców intencjonalnych jako algorytmów kompresji

Dennett wyróżnia trzy nastawienia w odniesieniu do pewnych obiektów fizycznych: fizyczne, funkcjonalne i intencjonalne. Jeśli obiekt jest intencjonalny, to można go analizować na poziomie funkcjonalnym i fizycznym. Jeśli jest funkcjonalny, to jest analizowalny na poziomie fizycznym. Przeciwne zależności nie zachodzą: nie każdy system fizyczny jest analizowalny na poziomie intencjonalnym. Choć w zasadzie każdy obiekt fizyczny można by tak analizować i opisywać, to jednak Dennett podaje dosyć przekonujące kryterium: opisy intencjonalne pewnych obiektów wcale nie są bardziej oszczędne od opisów funkcjonalnych ani nie stanowią podstaw dla nowych predykcji. Istnienie owych opisów – czy możliwości opisów – jest kwestią obiektywną, mierzalną w kategoriach informatycznych – przy użyciu złożoności algorytmicznej.

Nastawienie intencjonalne to strategia poznawcza, w której posługujemy się językiem mówiącym o abstraktach; nastawienie funkcjonalne – to strategia dotycząca illatów. Części funkcjonalne urządzeń są bowiem obiektami przestrzennymi (niekoniecznie jednak czasowymi). Dopiero w nastawieniu fizycznym mówimy o konkretach. Poziomy opis systemu intencjonalnego opisują zaś abstrakty, definiowane przez illata nastawienia funkcjonalnego.

Opisy intencjonalne mają być w takiej relacji do opisów funkcjonalnych (psychologii subosobowej), jak opisy figur na planszy świata życia do opisów współrzędnych zaczerpniętych komórek. W świecie życia illatami są komórki, zaś relacje między illatami pozwalają na zdefiniowanie abstrakcyjnego poziomu opisu – figur na planszy. Takimi samymi abstraktami mają być przekonania i pragnienia. Dennett odrzuca tezę, że pracą elementarnych struktur mózgu można efektywnie i bezpośrednio przewidywać stany intencjonalne. Przekonania są obiektami abstrakcyjnymi, takimi jak figury na planszy życia. Nie jest jednak tak, że należy je wyeliminować na rzecz opisu czysto fizycznego – algorytmy opisu intencjonalnego są przecież wydajniejsze dla pewnych danych wejściowych od algorytmów fizycznych.

Czym takie podejście różni się choćby od stanowiska Fodora? Dennett nie rozstrzyga, że przekonania jako stany intencjonalne są realizowane przez proste zdaniowe korelaty w języku myśli przetwarzanym przez mózg: stanom intencjonalnym niekoniecznie odpowiadają jednoznacznie zdaniowe reprezentacje mentalne. Fodor uznaje, że umysł jest obiektem, w którym nie występują na tyle skomplikowane relacje, żeby należało mówić o emergencji poziomu przekonań i pragnień względem poziomu psychologii subosobowej. Struktura algorytmiczna według Fodora w ścisły sposób odwzorowuje kompetencje ję-

zykowe⁶⁸. Przekonania muszą być kodowane w strukturach neurologicznych jako zdania języka myśli. Jeśli tak, to struktura układu, jakim jest umysł, musi być liniowa. Co więcej, organizmy nie posługujące się językiem, a mające stały intencjonalny (psy, krowy, żaby), muszą mieć wewnętrzne procesory języka myśli. Oba te założenia wydają się całkowicie nieprawdopodobne w świetle faktu, że mózgi są wytworami ewolucji, a język naturalny w ewolucji organizmów żywych pojawia się stosunkowo późno. Ewolucja poza tym nie wytwarza układów zbudowanych efektywnie i w prosty sposób, lecz raczej złożone emergentnie⁶⁹.

Dennett jest sceptyczny wobec realizmu Fodora: Fodor przypisuje realność przekonaniom także na poziomie subosobowym, czyli na poziomie operacji obliczeniowych umysłu. Zdaniem Dennetta jednak przekonania na poziomie subosobowym mogą być kodowane inaczej, niekoniecznie przez zdania w logice pierwszego rzędu. Funkcje przekonań na poziomie subosobowo-obliczeniowym mogą być realizowane przez inne struktury, na przykład przez wyrafinowane sieci neuronowe (wydaje się, że Fodor nie dopuszcza wielorakiej realizowalności kompetencji językowej przez różnego rodzaju algorytmy). Co więcej, człowiekowi można przypisywać nieskończenie wiele przekonań na mocy idealizacyjnego założenia o racjonalności, a w skończonym mózgu nie może mieścić się nieskończenie wiele illatów, które są wszak obiektami przestrzennymi⁷⁰. Nie można także ustalić, kiedy dokładnie człowiek nabywa pewne przekonania (np. przekonanie, że prawnicy zwykle noszą buty) – a należałoby móc to ustalać, gdyby przekonania były illatami w mózgu⁷¹.

Czy jednak stanowisko Dennetta, scharakteryzowane jako realizm wobec rzeczywistych wzorców wykrywanych przez algorytmy kompresji, rzeczywiście jest realizmem? Jednym z kryteriów realności jakiegos zjawiska, w przeciwieństwie do jego epifenomenalności, jest jego moc przyczynowa. Patricia Churchland przedstawia następujący argument, który ma świadczyć, że Dennett musi być instrumentalistą albo musi postulować istnienie języka myśli⁷². Obie możliwości są wedle niej równie nieatrakcyjne:

1. Nastawienie intencjonalne wobec x polega na postulowaniu reprezentacji zdaniowych u x .

⁶⁸ A. Cusins, *Dennett's realisation theory of the relation between folk and scientific psychology*. „The Behavioral and Brain Sciences” 1988, t. 11, nr 3, s. 508–509.

⁶⁹ W. Wimsatt, *Emergence...*, *op. cit.*

⁷⁰ D. Dennett, *Author's response. Science, philosophy, and interpretation*, „The Behavioral and Brain Sciences” 1988, t. 11, nr 3, s. 537.

⁷¹ *Ibidem*, s. 537.

⁷² P. S. Churchland, *Dennett's instrumentalism...* *op. cit.*, s. 537.

2. Jeśli zwierzętom przypisuje się przekonania i pragnienia, wówczas zgodnie z tą koncepcją reprezentacji mają one wrodzony i wewnętrzny język, a ich rozumowanie polega na wewnętrznej manipulacji symbolami językowymi.
3. Jeśli teoria przekonań i pragnień jest jedynie metodą organizowania danych, a nie postulowania przyczynowo aktywnych wewnętrznych stanów organizmu, to teoria ta jest instrumentalistyczna: przekonania i pragnień nie należy rozumieć jako realnych stanów mózgu, lecz jako stany wirtualne.
4. Teorie instrumentalistyczne traktują swoje obiekty teoretyczne jako użyteczne teoretyczne fikcje, czyli są fikcjonalistyczne.

Mamy więc do wyboru: albo Dennett uzna, że grzechotniki przetwarzają informację w języku myślenia (a wówczas uzna przekonania za realne stany mózgu), albo będzie instrumentalistą. W powyższym argumentcie „stany wirtualne” mózgu prawdopodobnie mają znaczyć to samo, co „abstrakcyjne stany” mózgu. W związku z tym zapewne Dennett, po pierwsze, nie uzna tego za zarzut, tylko za rekapitulację jego koncepcji; po drugie zaś, zaprzeczy, jakoby była ona instrumentalistyczna. Czy jednak może twierdzić, iż nie jest ona koncepcją epifenomenalistyczną ani fikcjonalistyczną?

Dennett odżegnuje się od fikcjonalizmu⁷³. W odpowiedzi Churchland pisał, że nie identyfikuje przekonań z żadnymi przyczynowo aktywnymi wewnętrznymi stanami organizmu. To byłoby świadectwem, iż przyznaje się do instrumentalizmu. Rzeczywiście, początkowo Dennett nazywał swoje stanowisko „instrumentalizmem”⁷⁴, lecz przedstawiona wyżej koncepcja rzeczywistych wzorców miała być tego stanowiska korektą – na rzecz „łagodnego” realizmu. Uzasadnienie tezy, iż przekonania nie są przyczynowo aktywnymi wewnętrznymi stanami organizmu, wskazuje, że nie jest ona epifenomenalistyczna: na tej samej zasadzie bowiem Dennett nie identyfikuje środka ciężkości lampy z żadnym wewnętrznym jej stanem ani cząsteczką. Dennett nie chce identyfikować bezpośrednio przekonań ze stanami mózgow, czyli uznawać przekonania za illata; przekonania są abstraktami, czyli nie są identyfikowalne jako nieprzestrzenne z żadnymi wewnętrznymi stanami mózgu (dodajmy: jako będące na innym poziomie emergentnym).

Jeśli tak, to czy nie popada Dennett w epifenomenalizm? Wydaje się na pierwszy rzut oka, iż abstrakty nie mają mocy przyczynowych⁷⁵. Liczba 4 nie jest przyczynowo aktywna. Dennett jednak uznaje, że abstrakty mają mo-

⁷³ D. Dennett, *Author's response. Taking the intentional stance seriously*, „The Behavioral and Brain Sciences” 1983, nr 3, s. 380.

⁷⁴ D. Dennett, *Brainstorms*, Cambridge Mass., 1978.

⁷⁵ Por. też A. Clark, *op. cit.*, s. 53–54.

ce przyczynowe⁷⁶: różnice w położeniu środka ciężkości mogą powodować przewidywalne zmiany zachowania obiektu. Jego zdaniem, jako faktyczna przyczyna liczy się to, co podaje się w wyjaśnieniach – a abstraktów używa się w poprawnych wyjaśnieniach fizycznych bardzo często. Jest to nietypowa i szeroka koncepcja przyczynowości, lecz w jej świetle Dennett nie jest epifenomenalista.

Dennett nie utożsamia posiadania przekonań z posiadaniem reprezentacji zdaniowych, ale nie popada w ten sposób w fikcjonalizm, instrumentalizm ani epifenomenalizm. Czemu jednak ob staje przy tym, że posiadanie przekonań nie sprowadza się do posiadania odpowiednich reprezentacji umysłowych? Negowanie posiadania reprezentacji może prowadzić go prostą drogą do fikcjonalizmu, do niemożności dostrzeżenia różnicy między wysuszonym trawnikiem a spragnionym człowiekiem.

Dennett podkreśla, że sukcesu predykcyjnego nastawienia intencjonalnego w danej dziedzinie nie należy mylić z potwierdzeniem hipotezy o istnieniu określonego systemu manipulacji reprezentacjami⁷⁷. W przeciwnym razie psychologia poziomu osobowego zostanie sprowadzona do poziomu subsobowego, a wówczas utracimy to, co ona nam daje: wyższą kompresję informacji na temat systemów intencjonalnych. Zamiast mówić o przekonaniach i pragnieniach w ogólny, intencjonalny sposób, będziemy musieli określać ich funkcjonalną konstrukcję, czyli będziemy musieli zajmować się bardziej złożonymi kwestiami. Nie jest to potrzebne do przewidywania zachowania, choć przydatne dla rozwoju samej psychologii subsobowej.

Myślę, że to jest argument, na mocy którego Dennett odrzuca utożsamienie intencjonalności z konkretnym systemem reprezentacji umysłowych, a także do zanegowania różnicy między pierwotną a pochodną intencjonalnością. Z jego perspektywy bowiem niczym nie różni się lista zakupów zapisana na papierze (tradycyjnie ujmowana jako obiekt pochodnie intencjonalny) od listy zapamiętanej i przechowywanej w umyśle (tradycyjnie ujmowanej jako intencjonalna pierwotnie). Na czym ma bowiem polegać różnica między oboma reprezentacjami? Jedynie na medium i sposobie kodowania, a to dla Dennetta za mało⁷⁸.

Oponenci Dennetta często wskazują, że nie definiuje on istoty intencjonalności, zajmując się jedynie zasadami przypisywania stanów intencjonalnych: Dennett pokazuje nam, jak systemy intencjonalne są rozpoznawane na podsta-

⁷⁶ D. Dennett, *Brainstorms*, op. cit., s. 358.

⁷⁷ D. Dennett, *Author's response. Taking... op. cit.*, s. 381.

⁷⁸ D. Dennett, J. Haugeland, *Intentionality*, w: R. L. Gregory (red.), *The Oxford Companion to the Mind*, Oxford 1987.

wie ich własności podstawialności i racjonalności. Nie podaje jednak nam istoty intencjonalności. System intencjonalny, w istocie, zawiera reprezentację celu (lub pragnienie), które w sposób instrumentalny steruje zachowaniem zwierzęcia⁷⁹.

Niemniej jednak Dennett uważa, że najbardziej obiecującym sposobem naturalizacji intencjonalności jest interpretacja stanów i procesów umysłowych jako reprezentacji, może nawet w języku myśli!⁸⁰ Jednocześnie dodaje, że dalszym krokiem takiej naturalizacji jest zaprzeczenie, iż reprezentacje umysłowe mają jakąkolwiek pierwotną lub wewnętrzną intencjonalność – w przeciwieństwie do pochodnej intencjonalności reprezentacji publicznych, takich jak zdania, obrazy i mapy. Dlaczego więc Dennett nie definiuje systemu intencjonalnego jako mającego reprezentacje umysłowe?

Odpowiedź Dennetta będzie zapewne brzmieć: gdyż reprezentacje umysłowe to poziom psychologii subosobowej. Przewidując zachowanie racjonalnych podmiotów, nie zajmujemy się rodzajem reprezentacji umysłowych. Względna separacja tych dwóch rodzajów psychologii prowadzi więc do komplikacji teoretycznych u Dennetta, komplikacji, które skłaniają niekiedy do interpretowania jego wypowiedzi w sposób instrumentalistyczny. Drugim uzasadnieniem odrzucenia definicji systemu intencjonalnego jako mającego reprezentacje jest możliwość przejawiania racjonalnego zachowania przez systemy o różnych systemach reprezentacji czy też systemy pozbawione reprezentacji, czyli – mówiąc żargonem filozofii umysłu – możliwość wielorakiej realizacji stanów intencjonalnych. Dennett nie chce *a priori* odrzucić różnych programów badawczych sztucznej inteligencji⁸¹; tymczasem w ramach sztucznej inteligencji istnieją podejścia radykalnie proceduralne. Proceduralizm Rodneya Brooksa polega na tym, że inteligencję modeluje się nie według logicznej reprezentacji, lecz stopniowo tworząc sztuczne stworzenia oddziałujące ze światem (np. sztuczne owady), szybko aktualizujące swoje wejścia sensoryczne, działające na podstawie równoległych układów analizy tych wejść. Stworzenia Brooksa nie mają *żadnej* zakodowanej reprezentacji, lecz często zachowują się w sposób, który chcielibyśmy nazwać racjonalnym w prostym sensie: zachowują się tak, jakby postępowały celowo. Jeśli ten program badań nad sztuczną inteligencją się powiedzie, to wówczas sfalsyfikuje to tezy Fodora, ale nie Dennetta, który oddzielał sprawę intencjonalnej interpretacji od spo-

⁷⁹ D. McFarland, *Intentions as goals*, „The Behavioral and Brain Sciences” 1983, nr 3, s. 369–370; por. też H. L. Roitblat, *Intentions and adaptations*, „The Behavioral and Brain Sciences” 1983, nr 3, s. 375.

⁸⁰ D. Dennett, J. Haugeland, *op. cit.*

⁸¹ D. Dennett, *Brainchildren...*, *op. cit.*, s. 215–234.

sobu, w jaki zachowanie intencjonalne jest realizowane. Hipotezy o istnieniu określonego sposobu reprezentacji mogą wydawać się pozaempiryczne, lecz są jedynie uprawdopodobnione pewnymi rozumowaniami. Nie wystarczy zdefiniować intencjonalność jako określony system reprezentacji; rzecz w tym, czy wszystkie obiekty zachowujące się w sposób stosunkowo racjonalny (celowy), mają reprezentacje. To jest jednak hipoteza empiryczna dla psychologii subosobowej. Łagodny realizm polega więc na przyjęciu tezy o możliwości wielorakiej realizacji stanów intencjonalnych i uznaniu stanów intencjonalnych za abstrakty definiowane przez funkcjonalne illata.

Czy jednak Dennett, oddzielając psychologię osobową od subosobowej, nie odcina się od realizmu? Wszak musi jakoś definiować przekonania jako abstrakty – a definiuje je przez illata psychologii subosobowej, czyli sam sobie przeczy... Ten zarzut można chyba łatwo odeprzeć, mówiąc, iż przekonania są na tyle abstrakcyjnymi cechami całościowo pojmowanych układów intencjonalnych, że mogą wiązać się z różnymi illatami: ich relację definiuje się dokładnie jako relację między jednym a drugim poziomem wyjaśniania, co wystarcza do ustalenia ich abstrakcyjnego charakteru.

W świetle powyższych analiz wydaje się, że Dennett rzeczywiście może zostać uznany za realistę w kwestii istnienia stanów intencjonalnych, nie zaś instrumentalistę. Owe stany okazują się u niego emergentne względem stanów funkcjonalnych, którymi mogą, lecz nie muszą, być reprezentacje umysłowe. Hasłowo rzecz ujmując, między intencjonalnością a reprezentacjami umysłowymi zachodzi nie tylko relacja emergencji, ale i wielorakiej realizacji. Konsekwencją tych poglądów są także pewne nieintuicyjne elementy koncepcji Dennetta, jak negacja zasady koneksji (koniecznego współwystępowania stanów świadomych i intencjonalnych) oraz różnic między intencjonalnością pochodną i pierwotną. Pełna ocena zasadności stanowiska autora *The Intentional Stance* wymagałaby odniesienia się właśnie do tych zagadnień, ale to z kolei wykracza poza ramy tego artykułu.