

Fred Dretske

Błędna reprezentacja*

Epistemologia zajmuje się wiedzą: w jaki sposób udaje się nam mieć rację? Jest jeszcze głębsze pytanie: jak udaje nam się pomylić? Jak systemy fizyczne mogą błędnie reprezentować środowisko?

Problem nie dotyczy tego, jak – na przykład – wykres *d* może błędnie reprezentować świat *w*. Rozważmy inny system *r*, mający już moce reprezentacyjne: tutaj wykres *d* może być wykorzystywany jako rozwinięcie systemu *r*, zwiększające jego siłę wyrazu, a więc biorące udział w sukcesach i niepowodzeniach reprezentacyjnych systemu *r*. Gdy do tego dojdzie, *d* może znaczyć, że *w* jest *F*, podczas gdy w rzeczywistości *w* nie jest *F*, ale znaczenie *d* jest ostatecznie wyprowadzalne z *r*. Wykres przedstawiający strukturę bezrobocia na przestrzeni ostatnich dziesięciu lat może błędnie reprezentować tę sytuację, ale zdolność wykresu do błędnego reprezentowania wynika z jego roli jako narzędzia wyrazu służącego podmiotom działającym, użytkownikom języka, którzy mają już tę zdolność.

Nie, problem dotyczy raczej zdolności reprezentacyjnych systemu, o ile zdolności te nie wywodzą się z operacji reprezentacyjnych o innym źródle. Dopóki nie mamy błędnego pojęcia, jak to jest możliwe, nie mamy pojęcia, jak naturalnie ewoluujące systemy biologiczne nabyły zdolność do wyrażania przekonań. Przekonanie jest – tu zakładam, że jest – zdolnością niewyprowadzaną reprezentacyjnie, której urzeczywistnienie może skutkować błędem reprezentacyjnym.

Zdolność do błędnej reprezentacji stanowi część, prawdopodobnie niewielką, ogólnego problemu znaczenia czy intencjonalności. Gdy mamy znaczenie, możemy, opisując i wyjaśniając zachowania ludzi, zwierząt, a może nawet funkcjonowanie maszyn, obdarzać nim do woli tego rodzaju systemy. Mając intencjonalność, możemy przyjąć (używając języka

* Pierwodruk: F.I. Dretske, *Misrepresentation*, w: *Belief: form, content, and function*, red. R. Bogdan, Clarendon Press, Oxford 1986, s. 17-37. Przekład publikujemy za uprzejmą zgodą autora.

Dennetta) nastawienie intencjonalne¹. Ale co (oprócz intencjonalności) daje nam (w przeciwieństwie np. do maszyn) zdolność do przyjęcia takiego nastawienia? Zdolność do przyjmowania takiego nastawienia jest wyrazem, a nie analizą intencjonalności. Znaczenie zapożyczony w systemach, wobec których przyjmujemy odpowiednie postawy, mówi nam nie więcej o pierwotnej zdolności do błędnego reprezentowania niż źle umiejscowiona pineska na mapie wojskowej. Poszukujemy więc takiego sposobu popełniania błędów przez naturę, gdzie błędna reprezentacja już nie odsyła dalej. Gdy to zrozumiemy, powinniśmy zrozumieć zarazem, jak szare komórki mogą błędnie reprezentować pogodę na jutrzejszy piknik.

1. Znaki naturalne

Znaki naturalnie coś znaczą bez naszej ingerencji². Woda nie płynie pod górę, dlatego płynąca na północ rzeka oznacza nachylenie w dół w kierunku północnym. Cień skierowany na wschód oznacza, że słońce jest na zachodzie. Nagłe przechylenie pasażerów w jednym kierunku oznacza przyspieszenie pociągu w kierunku przeciwnym. Waga tych zdarzeń czy też warunki decydujące o ich znaczeniu są niezależne od naszej interpretacji – a nawet od tego, czy jesteśmy w stanie je zinterpretować, czy nawet rozpoznać. Dentysta może użyć rentgena do diagnozy stanu prawego górnego zęba trzonowego, choć występowanie ciemnych plam oznacza zaawansowaną próchnicę, niezależnie od tego, czy ktokolwiek to zauważy. Rozszerzanie metalu wskazuje na rosnącą temperaturę (i w tym sensie oznacza wzrost temperatury), niezależnie od tego, czy ktokolwiek, obserwując rozszerzenie metalu, jest przekonany o wzroście temperatury. Coś znaczyło coś innego, jeszcze zanim inteligentne organizmy, zdolne do wykorzystania tego faktu (konstruujące termometry), zamieszkiwały Ziemię. Jeśli szukamy ostatecznego źródła znaczenia, a wraz z nim zrozumienia zdolności systemu do błędnego reprezentowania, to w tym miejscu z pewnością warto rozpocząć poszukiwania.

¹ D.C. Dennett, *Intentional systems*, „The Journal of Philosophy” 1971, t. LXVIII, nr 4, s. 87-106. Przedrukowane też w: D.C. Dennett, *Brainstorms. Philosophical Essays on Mind and Psychology*, MIT Press, Cambridge, Mass. 1978.

² Wymaga to pewnego zastrzeżenia, do którego wkrótce powrócę. To, co znak naturalny znaczy, często zależy od nas, od tego, co wiemy na temat odpowiednich alternatywnych możliwości, lub od tego, w jaki sposób używamy odpowiedniego przyrządu. Jeśli jednak niczego nie wiemy lub jeśli znak występuje w działaniu urządzenia, które nie ma normalnego użycia, znak nadal coś znaczy – ale nie to, co uznajemy za jego znaczenie w epistemicznie (lub funkcjonalnie) bogatszych warunkach. Powrócę do tej sprawy niżej w przypisie 8.

Znaki naturalne są mniej lub bardziej wiarygodnymi wskaźnikami, a ich znaczenie sprowadza się do tego, na co wskazują. Moc znaczenia naturalnego polega na oznaczaniu czegoś – np. odry Tomasza – co gwarantują odpowiednie obiektywne okoliczności, pewne relacje o charakterze praw, zachodzące między znakiem (lub własnością tego znaku) a warunkiem konstytuującym jego znaczenie (chorowanie Tomasza na odrę). W większości przypadków relacja ta ma charakter prawa lub związku przyczynowo-skutkowego, takiego, który pozwala utrzymywać w mocy kontrfaktyczne stwierdzenie w takim sensie, że gdyby nie zaszedł pierwszy warunek (gdyby Tomasz nie miał odry), drugi też by nie zaszedł (nie miałby charakterystycznych czerwonych plam na twarzy). Czasami zdarza się, że są to jedynie regularności, choć nie o charakterze praw, to mimo wszystko na tyle przekonujące, by wspierać utrzymanie związku znaku i znaczenia. Dzwonienie dzwonka u drzwi oznacza kogoś stojącego za nimi po części na mocy faktu, choć przypuszczalnie nie prawa przyrody, że zwierzęta (na przykład wiewiórki czy dziecięły) nie dzwonią regularnie do drzwi, gdy żerują. Gdyby wiewiórki zmieniły swoje przyzwyczajenia (gdyby dzwonki do drzwi były wykonane np. z orzechów), wówczas naciśnięcie dzwonka do drzwi nie oznaczałoby tego, co oznacza obecnie. Ale skoro tak się rzeczy obecnie mają, to możemy (przeważnie) powiedzieć, że dzwonek nie dzwoniłby, gdyby nikogo nie było za drzwiami, tzn. że dzwonek wskazuje czyjąś obecność za drzwiami i w związku z tym zachodzi właśnie to, co dzwonek oznacza. Opisana kontrfaktyczna zależność między dzwonieniem dzwonka i czyjąś obecnością za drzwiami jest odbiciem prawidłowości, która jednak nie ma ani charakteru konwencji, ani charakteru prawa. Mimo to dzwonek do drzwi zachowuje swoje znaczenie naturalne dopóty, dopóki utrzymuje się ta prawidłowość.

Poza tym nie mam nic więcej systematycznego do powiedzenia na temat tego, co stanowi naturalne znaczenie zdarzenia lub stanu³. Przejdę teraz do – jak mam nadzieję – stosunkowo dobrze znanego pojęcia. W razie konieczności odniosę się do konkretnych przykładów. Mój projekt ma na celu sprawdzenie, jak głęboko można zrozumieć błędną reprezentację, dyspozycję (stanu, zdarzenia, sytuacji) r do błędnego oznaczania (a więc wskazywania) w jako F (a zatem błędnego reprezentowania F) w kategoriach znaczenia naturalnego, że w jest F . Tylko wtedy, gdy (jeśli) mój projekt się powiedzie, a przynajmniej zarysują się perspektywy jego powodzenia, być może dokładniejszego przyjrzenia będą wymagały wszystkie założenia przemyczone na początku.

³ Pełniejsze ujęcie w: F.I. Dretske, *Knowledge and the Flow of Information*, wyd. II, MIT Press, Cambridge, Mass. 1982, rozdz. 1 i 2.

Choć znaczenie naturalne wydaje się obiecującym punktem wyjścia, trudno jest znaleźć do niego drogę. Znaki naturalne, choć coś oznaczają i (w tym sensie) mogą reprezentować *w* (przez wskazanie lub oznaczenie *w* jako *F*), są niezdolne do błędnego reprezentowania czegokolwiek. Znaki naturalne albo oznaczają, albo nie oznaczają w ogóle. Plamki na twarzy Tomasza z pewnością mogą oznaczać odrę Tomasza, ale oznaczają ją tylko wtedy, gdy Tomasz rzeczywiście ma odrę. Jeśli Tomasz nie ma odry, wówczas plamki nie oznaczają odry. Być może oznaczają tylko zbyt dużą ilość zjedzonych przez Tomasza słodyczy.

Grice wyraża tę kwestię, twierdząc, że zdarzenie (egzemplarz pewnego znaku naturalnego) znaczy, że *P* (w sensie naturalnego sensu „znaczenia”, które będę od tej pory określał skrótowo jako „znaczenie_n”), tylko jeśli *P* zachodzi⁴. Grice zestawia ten sens znaczenia ze znaczeniem nienaturalnym, gdzie znak może znaczyć, że *P*, mimo że *P* jest fałszywe. Jeśli zastrzeżemy słowo „znaczenie” (bez indeksów dolnych) dla znaczenia tego rodzaju, że coś może znaczyć, że *w* jest *F*, gdy *w* nie jest *F*, a więc dla znaczenia dopuszczającego możliwość błędnego reprezentowania, to wydaje się, że przy użyciu pojęcia znaczenia_n nieszczególnie można zrozumieć znaczenie.

Mówiąc o znakach i ich znaczeniu naturalnym, odnoszę się do konkretnych zdarzeń, stanów rzeczy lub warunków – takich jak ta konkretna ścieżka, te konkretne chmury i ten konkretny dym. Typ znaku (na przykład dym) może oznaczać w pewnym naturalnym sensie obecność ognia, nawet gdy każdy egzemplarz wybranego typu nie znaczy_n, że tam jest ogień (bo bywa, że tam ognia nie ma). Jednak to znaczenie, odniesione do konkretnego typu, bez względu na jego właściwą analizę, nie pomaga nam w zrozumieniu błędnej reprezentacji, o ile pojedyncze egzemplarze tego typu nie mają znaczenia odniesionego do danego typu: o ile poszczególne podmuchy dymu nie znaczą_n, że tam jest ogień, gdy w rzeczywistości go tam nie ma. Nam jednak nie o to chodzi. Rejestracja „pustego” wskaźnika poziomu benzyny (tego typu zdarzenia) może oznaczać pusty zbiornik, ale w sytuacji, gdy zbiornik nie jest pusty, żadna konkretna rejestracja „pustego” wskazana za sprawą wskaźnika nie znaczy_n, że zbiornik jest pusty. Toteż żadna konkretna rejestracja wskaźnika nie reprezentuje błędnie ilości benzyny w baku (przy znaczeniu_n, że jest on pusty, gdy w rzeczywistości taki nie jest).

Czasem niezdolność konkretnych znaków naturalnych do błędnego reprezentowania bywa przesłonięta przez sposób, w jaki je wykorzystujemy w urządzeniach technicznych. Nasze zainteresowanie tym, czy *w* staje się *F*, a jeśli tak, to kiedy tak się dzieje, sprawia, że konstruujemy przyrząd,

⁴ H.P. Grice, *Meaning*, „The Philosophical Review” 1957, t. LXVI, nr 3, s. 377-388.

którego różne stany są zaprojektowane w taki sposób, że funkcjonują jako znaki naturalne warunku w . Ze względu na sposób, w jaki używamy tego urządzenia, zwykle twierdzimy, że pewna konkretna rejestracja tego, że d jest G (zakładając, że jest to znak naturalny tego, że w jest F), znaczy, że w jest F , nawet gdy – na skutek dysfunkcji lub niewłaściwego użycia – system nie działa wystarczająco dobrze i w nie jest F . Nie jest to, rzecz jasna, znaczenie_n konkretnej pozycji wskaźnika. To jest to, co wskaźnik ma znaczyć_n; to, do znaczenia_n czego został on zaprojektowany; to, co (najprawdopodobniej) egzemplarze danego typu normalnie znaczą_n, ale nie to, co po prostu znaczą_n.

W razie zwarcia dzwonienie do drzwi (bez względu na to, co miało wskazywać wedle projektu, i bez względu na to, co normalnie wskazuje) nie wskazuje na naciśnięty dzwonek. Dzwonienie do drzwi znaczy_n (wskazuje), że prąd elektryczny płynie w obwodzie dzwonka (to zawsze znaczyło_n), lecz obecnie dzwonienie nie znaczy_n, że dzwonek został naciśnięty. Przepływ prądu obecnie znaczy_n – z pewnością tak byśmy to oceniali, gdybyśmy mogli zobaczyć dzwonek i zobaczyć, że nie został naciśnięty – że system jest dysfunkcyjny, czyli nastąpiło zwarcie gdzieś w instalacji elektrycznej. Zdanie „Ktoś jest za drzwiami” może znaczyć, że ktoś jest za drzwiami, nawet gdy nikogo tam nie ma, ale samo dzwonienie dzwonka nie może znaczyć, że nikogo tam nie ma. Przynajmniej o ile dotyczy to znaczenia_n. Jeśli dzwonek nie jest wciśnięty, wtedy musimy szukać czegoś innego, co dzwonienie dzwonka znaczyłoby_n. Często cofamy się do bliższego znaczenia_n, do pewnego warunku lub stanu rzeczy w normalnym łańcuchu przyczynowym wcześniejszych zdarzeń (na przykład do przepływu prądu lub do przyczyny przepływu prądu – jaką może być zwarcie) i określamy je jako znaczenie_n dzwoniącego dzwonka.

2. Znaczenie funkcjonalne

Na to można by odparować, że co prawda dzwonienie dzwonka nie może znaczyć_n, że ktoś jest za drzwiami, gdy rzeczywiście nikogo tam nie ma. Jednak w pewnym pokrewnym sensie znaczenia, dzwonienie dzwonka nadal to dokładnie znaczy, niezależnie od tego, czy ktoś tam jest, czy nie. Jeśli to nie jest znaczenie naturalne (znaczenie_n), to jest to znaczenie blisko z nim spokrewnione.

Niezależnie od tego, czy jest ono rzeczywiście spokrewnione, czy nie, z pewnością jest rodzajem znaczenia, dotyczącym systemów lub składników systemów, które mają możliwe do zidentyfikowania funkcje. Jeszcze raz rozważmy wskaźnik poziomu paliwa. Ma on funkcję: przekazywanie informacji o ilości paliwa w zbiorniku. Jeśli wszystko działa prawidłowo,

położenie wskazówki jest znakiem naturalnym stanu zawartości zbiornika. Odchylenie wskazówki w lewo znaczy_n, że zbiornik jest pusty, odchylenie jej zaś w prawo znaczy_n, że zbiornik jest pełny. Podobnie działa to dla pośrednich pozycji. Czasem jednak coś idzie nie tak: połączenia chodzą luzno, akumulator jest rozładowany, przewody przerywają. Wskaźnik zaczyna rejestrować „brak paliwa”, gdy zbiornik jest wciąż napełniony. Jeśli tak się zdarzy, zwykle powiemy, że wskaźnik błędnie reprezentuje stan zawartości zbiornika. Wskaźnik pokazuje, że zbiornik jest pusty, gdy w istocie tak nie jest. Znaczy to (oczywiście nie w sensie znaczenia_n, ale w pewnym sensie wciąż to znaczy), że zbiornik jest pusty.

Jeśli [własność] G przysługująca [przedmiotowi] d normalnie jest naturalnym znakiem, że w jest F , gdy właśnie to normalnie znaczy_n, wówczas istnieje sens, w którym znaczy właśnie to, niezależnie, czy w jest F , jeśli funkcją d jest wskazywanie warunku w . Nazwijmy ten rodzaj znaczenia znaczeniem_f – indeks wskazuje, że jest to znaczenie funkcjonalnie pochodne.

(Z_f) d jest G znaczy_f, że w jest F = funkcją d jest wskazywanie warunku w , a sposób, w jaki d wykonuje tę funkcję, sprowadza się po części do wskazywania, że w jest F przez to, że (d) jest G .

Pozycja wskazówki uszkodzonego wskaźnika poziomu paliwa znaczy_f, że zbiornik jest pusty, gdyż funkcją wskaźnika jest wskazywanie pozostałej ilości paliwa w zbiorniku, a sposób, w jaki [wskaźnik] wykonuje tę funkcję, sprowadza się po części do wskazywania pustego zbiornika, gdy wskaźnik rejestruje „brak paliwa”⁵. Z tego samego powodu i w tym samym sensie dzwonenie dzwonka informuje (tzn. znaczy_f), że ktoś jest za drzwiami, nawet wtedy, gdy nikogo tam nie ma.

To, czy Z_f stanowi jakikolwiek postęp w próbie naturalizacji znaczenia (a zatem w rozumieniu pierwotnej zdolności systemu do błędnego reprezentowania), zależy od tego, czy odnośne funkcje można zrozumieć w sposób bardziej naturalny. Jeśli te funkcje są (w mojej terminologii) funkcjami przypisanymi, znaczenie_f jest zanieczyszczone celami, intencjami i przekonaniem tych, którzy przypisują funkcję, z której znaczenie_f czerpie zdolności do błędnego reprezentowania⁶. Nie uda nam się uchwycić znaczenia, dopóki nie obejmie ono zdolności do błędnego reprezen-

⁵ Mam nadzieję, że jasne jest, iż nie rozważam tutaj wyrażenia „brak paliwa” (lub litery „E”), które mogłyby pojawić się na wskaźniku. Symbol ten oznacza brak – bez względu na to, co się dzieje ze wskaźnikiem paliwa. Jest to sprawa czystej konwencji. Rozważam to, co pozycja wskaźnika znaczy_n, niezależnie od tego, co zdecydujemy się wydrukować na przodzie tego urządzenia.

⁶ L. Wright nazywa te funkcje „świadomymi”; por. L. Wright, *Functions*, „The Philosophical Review” 1973, t. LXXXII, nr 2, s. 139-168.

towania. Podążymy jedynie drogą powrotną, choć bardziej pokrętną, do naszych tajemniczych zdolności do reprezentowania.

Aby lepiej uchwycić, co rozumiem przez „funkcję przypisaną”, a także to, w jaki sposób (przez nasze intencje, cele i przekonania) mamy udział w systemie mającym tę funkcję, rozważmy następujący przypadek. Czuła waga sprężynowa, skalibrowana w ułamkach gramów, została zaprojektowana i służy do określenia masy bardzo małych obiektów. Przyrząd ten, czego nie wiedzą ani jego projektanci, ani użytkownicy, jest czułym wskaźnikiem poziomu wysokości nad poziomem morza. Mógłby posłużyć jako zwyczajny wysokościomierz dzięki temu, że rejestruje zmniejszającą się wagę rzeczy w miarę wzrostu wysokości (zauważmy, że waga rzeczy jest funkcją jej wysokości nad poziomem morza) – gdyby tylko użytkownik ustalił standardową wagę i zauważał zmianę rejestracji przy zmianie wysokości. Przypuśćmy teraz, że w normalnym warunkach laboratoryjnych przyrząd funkcjonuje niepoprawnie i rejestruje 0,98 g dla przedmiotu o wadze 1 g. Czy jest to błędna reprezentacja wagi tego przedmiotu? Czy waga błędnie reprezentuje wysokość danego przedmiotu? Co znaczy napis 0,98 g? Jeżeli myślimy o znaczeniu_n, z pewnością to nie znaczy_n, że przedmiot waży 0,98 g. To nie znaczy_n też, że laboratorium znajduje się 40 000 m n.p.m. Jeśli jednak myślimy o znaczeniu_f, sensowne wydaje się twierdzenie, że wskaźnik przyrządu wskazuje (tj. znaczy_f), że przedmiot waży 0,98 g. Funkcją tego przyrządu jest wskazywanie wagi przedmiotów, a wskazuje on (jak się okazuje, błędnie), że ten przedmiot waży 0,98 g.

Ale czy to wysokość jest błędnie reprezentowana? Nie. Należy zauważyć, że przyrząd nie może błędnie reprezentować ani wysokości, ani wagi, ponieważ reprezentacja (lub błędna reprezentacja) jednego zakłada stałość (a więc niereprezentowanie) drugiego⁷. Choć przyrząd może służyć jako wysokościomierz, nie jest on wykorzystywany w ten sposób. Nie jest więc to jego funkcją. Jego funkcją jest natomiast rejestracja wagi. Tę funkcję możemy przypisać przyrządowi, gdyż można znaleźć powód jego skonstruowania i wyjaśnienie, dlaczego został on zbudowany w ten sposób. Gdyby nasze potrzeby były inne, znaczyłby on (w sensie znaczenia_f) coś innego. Ale nie były i nadal nie są inne.

Możemy czasami zmieniać funkcję przypisaną przyrządowi. Po skalibrowaniu nie używamy go na przykład do mierzenia tego, do mierzenia czego zwykle służy. Zamiast tego możemy zastosować go do znanych wielkości [pomiarowych], aby wykorzystać go jako znak (naturalny) możliwej dysfunkcji lub niedokładności samego przyrządu. W tym przypadku od-

⁷ Dzwonek do drzwi na przykład nie może znaczyć_n zarówno tego, że ktoś jest za drzwiami, jak i że wystąpiło zwarcie.

czyt 0,98 g (przy wadze 1 g) wskazuje, że sprężyna zmieniła swoje właściwości, wskaźnik jest wygięty lub jakiś inny element jest błędnie ustawiony. Znajdujemy nowe znaczenie funkcjonalne, ponieważ zmiana stanu naszej wiedzy ogólnej (zwykle na skutek odmiennych intencji i celów) wpływa na to, co zachowanie wskaźnika znaczy_n. Znaczenie_f zmienia się zgodnie z naszymi potrzebami wraz z funkcjami przypisanymi⁸.

Czasami wykorzystujemy zwierzęta w taki sam sposób, w jaki wykorzystujemy przyrządy. Psy mają czuły zmysł powonienia. Opierając się na tym fakcie, celnicy wykorzystują je do wykrywania marihuany. Gdy pies macha ogonem, szczeka lub robi coś, do czego jest przeszkolony, gdy czuje zapach marihuany, zachowanie psa służy jako znak naturalny tego, że bagaż zawiera marihuanę. Ale nie znaczy to, że zachowanie psa (czy też stanu neuronalnego, który wyzwała to zachowanie) może błędnie reprezentować zawartość tego bagażu. Zachowanie psa może spowodować przekonanie (błędne) urzędnika celnego, że marihuana jest w walizce, ale zachowanie psa znaczy to (w sensie znaczenia_f) tylko pochodnie. Jeśli pies wyjątkowo dobrze wykonuje swoje zadanie, szczekając tylko wtedy, gdy marihuana jest w walizce, można powiedzieć, że jego szczekanie wskazuje (tj. znaczy_n), że marihuana tam się znajduje. Co więcej, to znaczy_n niezależnie od czyjejs interpretacji tego jako znaczenia_n, niezależnie od użycia znaku naturalnego do naszych prywatnych śledczych celów. Gdy jednak marihuany tam nie ma, gdy pies szczeka na niewinne opakowanie ziół, szczekanie nie znaczy_n, że marihuana jest obecna. Nie znaczy_f ono w żadnym sensie, który jest niezależny od naszej interpretacji. Można oczywiście powiedzieć, co szczekanie znaczy dla nas (że marihuana jest w walizce), ale ten sposób mówienia jedynie ujawnia nasz wkład w przypisanie znaczenia zachowaniu psa. Przypisujemy to znaczenie ze względu na informację, którą chcemy uzyskać, a którą możemy zdobyć, wykorzystując psa w sposób, w jaki pies został wyszkolony w celu uzyskiwania tej informacji. Jeśli jednak odkładamy na bok nasze interesy i cele, to gdy nie ma marihuany, nie ma sensu twierdzić, że szczekanie psa znaczy, że marihuana jest w walizce. Jedynym rodzajem błędnej reprezentacji tutaj

⁸ To nie sama zmiana celu zmienia to, co coś znaczy_n (a więc znaczy_f). Zmienia to fakt, że zmiana w użyciu towarzyszy zmianą wiedzy ogólnej i znaczenie_n zmienia się wraz z zmianą wiedzy ogólnej. Jeśli na przykład *A* zależy zarazem od *B* i *C*, zmiana *A* może znaczyć_n, że *C* zmienia się, jeśli wiemy, że *B* jest stałe. Jeśli wiemy, że *C* jest stałe, może to znaczyć_n, że *B* się zmienia. Jeśli niczego nie wiemy, znaczy to jedynie, że zmienia się albo *B*, albo *C*. Znaczenie naturalne jest w tym sensie względne, ale ujednoznacznianie go (pomijając to, co wiemy, i w jaki sposób używamy danego przyrządu) nie eliminuje znaczenia naturalnego. Czyni to tylko znaczenie_n mniej zdeterminowanym. Pełniejsze omówienie tej sprawy: F.I. Dretske, *Knowledge and the Flow of Information...*, *op. cit.*, rozdz. 3.

występującej jest pochodny rodzaj błędnej reprezentacji, która wiąże się z mapami, przyrządami oraz językiem.

Dlatego też, jeżeli Z_f ma być znaturalizowanym ujęciem reprezentacji, obejmującym zdolność do błędnego reprezentowania, wówczas odnośne funkcje muszą być funkcjami naturalnymi, czyli funkcjami, które dana rzecz ma niezależnie od naszych intencji i celów. Szukamy związanych z systemem znaków naturalnych funkcji, nadających tym znakom treść, a zatem i znaczenie (tj. znaczenie_f), które nie pasożytuje na wykorzystywaniu ich przez nas w urządzeniach do gromadzenia informacji, na naszych interpretacjach⁹.

Potrzebujemy więc pewnej charakterystyki funkcji naturalnych systemu. W szczególności, skoro jesteśmy zainteresowani funkcjami, jakie wybrany system znaków naturalnych może mieć, poszukujemy tego, co znak powinien znaczyć_n, gdzie owo „powinien” odnosi się do funkcji danego znaku (lub systemu znaków) w ekonomii poznawczej organizmu. Chcemy wiedzieć, w jaki sposób pies reprezentuje zawartość bagażu – co (jeśli cokolwiek) zapach bagażu znaczy_f dla niego.

3. Potrzeby

Funkcji naturalnych oczywiście można poszukiwać w systemach biologicznych. Systemy biologiczne mają różne narządy, mechanizmy i procesy, które rozwinęły się (dojrzały i zachowały się), ponieważ odgrywały ważną rolę w gromadzeniu informacji potrzebnych gatunkom do adaptacji do otoczenia. Funkcja gromadzenia informacji niezbędnej – w większości przypadków – do spełnienia potrzeby biologicznej wtedy tylko może zostać zrealizowana w systemie, gdy system jest zdolny do zajęcia stanów służących jako znaki naturalne warunków zewnętrznych (a czasami też wewnętrznych). Jeśli zbiór fotoreceptorów, nazywany „siatkówką”, jest zdolny do wykonywania funkcji (niezależnie od tego, co dokładnie uznamy za daną funkcję), różne stany tych receptorów muszą coś znaczyć_n o charakterze i rozmieszczeniu czyjegoś otoczenia optycznego. To, co różne stany tych receptorów znaczą_f, jest (zgodnie z Z_f) wyznaczone przez dwie rzeczy: (1) funkcję wskazywania tego receptora, oraz 2) znaczenie_n różnych stanów, które umożliwiają systemowi spełnienie tej funkcji.

⁹ Uważam, że wiele naszych rozważań nad zdolnościami reprezentacyjnymi komputerów dotyczy przypisanych, a zatem wywiedzionych zdolności reprezentacyjnych. Nic nam to nie mówi o wewnętrznej zdolności maszyny do reprezentowania lub błędnego reprezentowania. Dlatego też nic nam to nie mówi o charakterze poznawczym stanów wewnętrznych. R. Cummins, moim zdaniem, świetnie to ujmuje, rozróżniając *poznanie (wersja znaczenia przypisanego) i autentyczne poznanie. Zob. R. Cummins, *The Nature of Psychological Explanation*, MIT Press, Cambridge, Mass. 1983.

W celu zilustrowania sposobu, w jaki Z_f ma mieć zastosowanie, wygodniej będzie rozważać proste organizmy z oczywistymi potrzebami biologicznymi – rzeczami lub warunkami, bez których nie mogłyby przetrwać. Jest to wygodne, gdyż omawiane podejście do problemu błędnych reprezentacji stosuje się najłatwiej do mechanizmów poznawczych wspomagających pewne podstawowe potrzeby biologiczne. Rozważenie zaś pierwotnych systemów daje nam dodatkowe korzyści wynikające z uniknięcia kolistości w analizie, która pojawiłaby się w razie odwołania do tych „potrzeb” (np. moja potrzeba korzystania z edytora tekstu), które wywodzą się z pragnień (np. mojego pragnienia, aby szybciej tworzyć teksty z mniejszą liczbą literówek). Nie możemy wprowadzać pragnień na tym etapie analizy, ponieważ pragnienia posiadają już tego rodzaju treść reprezentacyjną, którą właśnie staramy się zrozumieć.

Niektóre bakterie morskie mają wbudowane magnetytosome (tzw. magnetosomy), które działają podobnie jak igła w kompasie, ustawiając się (a w rezultacie i bakterię) równoległe do ziemskiego pola magnetycznego¹⁰. Ponieważ owe linie magnetyczne nachylają się do dołu (w kierunku północy geomagnetycznej) na półkuli północnej (do góry na półkuli południowej), bakterie znajdujące się na półkuli północnej, zorientowane przez wewnętrzne magnetosomy, wprawiają się w ruch w stronę geomagnetycznej północy. Wartość adaptacyjna magnetotaksji (tak nazywa się ten mechanizm sensoryczny) nie jest oczywista, ale uzasadnione jest przypuszczenie, że funkcjonuje ona tak, aby umożliwiać bakterii unikanie wód powierzchniowych. Ponieważ te organizmy mogą żyć tylko pod nieobecność tlenu, ruch w kierunku geomagnetycznym północnym odwodzi bakterię od bogatej w tlen wody powierzchniowej w stronę stosunkowo wolnego od tlenu osadu na dnie. Bakterie na półkuli południowej mają odwrócone magnetosomy, dzięki czemu pływają w stronę geomagnetycznego południa z tak samo korzystnymi wynikami [jak bakterie na półkuli północnej]. Przeszczepiona na północny Atlantyk bakteria południowa sama się zniszczy – płynąc w górę (w kierunku magnetycznego południa) do toksycznych, bogatych w tlen wód powierzchniowych.

Jeśli magnes prętowy zorientowany w kierunku przeciwnym do kierunku ziemskiego pola magnetycznego znajduje się w pobliżu bakterii, to mogą dać się one zwabić w stronę śmiertelnego dla siebie środowiska. Choć powróć jeszcze do tej sprawy (aby ocenić to rozumowanie), na razie wygląda na to, że jest to – najprawdopodobniej – przykład błędnej reprezentacji. Przez to, że w normalnym środowisku (habitacie) bakterii

¹⁰ Przykład zaczerpnąłem z: R.P. Blakemore, R.B. Frankel, *Magnetic Navigation in Bacteria*, „Scientific American” 1981, t. CCXLV, nr 6, s. 58-65.

wewnętrzne ukierunkowanie ich magnetosomów znaczy_n, że w danym kierunku jest relatywnie niewiele tlenu, i skoro organizm potrzebuje do przetrwania właśnie tej informacji, zasadne wydaje się, że funkcja tego mechanizmu sensorycznego służy spełnieniu tej potrzeby, uzyskaniu tej informacji, wskazaniu, że woda wolna od tlenu znajduje się w tym właśnie kierunku. Jeśli to właśnie powinno znaczyć_n, zarazem to właśnie znaczy_f. Stąd też w obecności magnezu prętowego i zgodnie z Z_f stan sensoryczny organizmu błędnie reprezentuje lokalizację wolnej od tlenu wody.

Nie oznacza to oczywiście, że bakterie mają przekonania – przekonania, że w danym kierunku jest bardzo mało tlenu lub że nie ma go wcale. Zdolność do błędnego reprezentowania jest tylko jednym z wymiarów intencjonalności, tylko jedną z właściwości, którą system reprezentacyjny musi mieć, aby kwalifikować się jako system przekonań. Żeby kwalifikować się jako przekonanie, treść reprezentacyjna musi wykazywać się (między innymi) charakterystyczną dla nastawień sądzeniowych nieprzejrzystością. Znaczenie_f (wciąż), o ile nie jest jakoś barwnie scharakteryzowane, nie wykazuje tego poziomu intencjonalności. Nasz projekt jest jednak skromniejszy. Szukamy znaturalizowanej formy błędnej interpretacji, a jeśli jeszcze nie mamy koncepcji fałszywego przekonania, to sformułowaliśmy, jak się zdaje, koncepcję fałszywej treści.

Nic, co dotąd powiedziałem, nie jest szczególnie oryginalne, oprócz kilku terminologicznych upiększeń czy nieco odmiennego sposobu rozumienia problemu. Powtarzam jedynie kroki, niektóre bardzo istotne, poczynione już wcześniej przez innych. Mam tu na myśli zwłaszcza przełomową analizę reprezentacji językowej Stampego, w której (potencjalnie błędna) treść reprezentacji jest utożsamiona z tym, co mogłoby spowodować, że reprezentacja ma dane własności w warunkach prawidłowego funkcjonowania¹¹; opracowanie przez Enę koncepcji intencjonalności stanów poznawczych na podstawie idei funkcji¹²; Fodorowskie zastosowanie teleologicznych pojęć dostarczających „językowi myśli” semantyki¹³ i imponującą analizę znaczenia u Millikan w kategoriach różnorodnych funkcji właściwych, które może mieć reprodukowalne zdarzenie (takie jak

¹¹ D.W. Stampe, *Toward a Causal Theory of Representation*, „Midwest Studies in Philosophy” 1977, t. II, nr 1, s. 42-63.

¹² B. Enę, *Intentional States of Mechanical Devices*, „Mind” 1982, t. XCI, nr 362, s. 161-183. Enę utożsamiał treść stanu funkcjonalnego z (konstrukcją) własności zdarzenia, wobec którego system ma funkcję odpowiadania.

¹³ J.A. Fodor, *Psychosemantics, or where do truth conditions come from?*, w: *Mind and Cognition*, red. W.G. Lycan, Basil Blackwell, Oxford 1990, s. 312-337.

dźwięk i gest)¹⁴. Ja sam starałem się wykorzystać (niejasne) idee funkcjonalne w mojej analizie przekonania przez określenie treści semantycznej struktury w kategoriach informacji, którą ta struktura miała przenosić (a zatem uzyskała funkcję [jej] przenoszenia)¹⁵.

4. Niezdeteminowanie funkcji

Chociaż proponowane podejście do problemu znaczenia – a zatem i błędnej reprezentacji – do pewnego stopnia zbadano, wciąż trudno traktować je jako obiecujący szkic, nie wspominając już o pełnym obrazie mylenia się w przyrodzie.

Po pierwsze, pojawia się pytanie, jak należy rozumieć zdolność systemu do błędnej reprezentacji czegoś, wobec czego nie ma on potrzeby biologicznej. Jeśli O nie potrzebuje (lub potrzebuje unikać) F , nie może (w omawianej koncepcji) być funkcją naturalną jakiegokolwiek z systemów poznawczych O , funkcją ostrzegania go o obecności (czy też nieobecności, położeniu, zbliżaniu się, tożsamości) F . A bez tego nie pojawia się możliwość błędnego reprezentowania F . Pewien wewnętrzny stan wciąż może znaczyć_n, że F było obecne (analogicznie system detektora Rovera znaczy_n, że bagaż zawiera marihuanę), ale ten stan wewnętrzny nie może znaczyć_f, że tak jest. Do tej pory zapoznaliśmy się ze sposobem, w jaki organizm może błędnie reprezentować obecność jedzenia, przeszkody, drapieżnika lub partnera seksualnego (a więc coś, wobec czego istnieje biologiczna potrzeba uzyskania lub unikania)¹⁶, lecz nie rozumiemy nadal, w jaki sposób możemy błędnie reprezentować np. otwieracze do puszek, rakiety tenisowe, tulipany czy waleta karo. Nawet gdybyśmy założyli, że nasz układ nerwowy jest na tyle skomplikowany, że wskazuje (w normalnych warunkach) na obecność rzeczy tego rodzaju, to z pewnością funkcją naturalną tych stanów neuronalnych nie mogłoby być sygnalizowanie obecności przyborów kuchennych, sprzętu sportowego, kwiatów i kart do gry, a jeszcze mniej sygnalizowanie obecności szczególnego rodzaju tych przedmiotów.

Jest to, jak sądzę, niezwykła trudność, lecz trudność do przewyciężenia. Wydaje się bowiem oczywiste, że system poznawczy mógłby rozwinąć się w celu służenia, a zatem posiadania funkcji naturalnej służenia, jakiejś biologicznej potrzebie bez ograniczania swoich operacji reprezentacyjnych

¹⁴ R.G. Millikan, *Language, thought, and other biological categories. New foundations for realism*, The MIT Press, Cambridge, Mass. 1984.

¹⁵ F.I. Dretske, *Knowledge and the Flow of Information...*, op. cit., cz. 3.

¹⁶ Coś, dla czego istnieje – w języku (wczesnego) Dennetta – „właściwa kontynuacja eferentna”: zob. D.C. Dennett, *Content and Consciousness*, Routledge and Kegan Paul, London 1969.

(jak i związanych z błędną reprezentacją) do tych właśnie potrzeb. W celu rozpoznawania swego naturalnego drapieżnika organizm mógłby rozwinąć detektory koloru, kształtu i ruchu o dość dużych zdolnościach rozróżniania. Organizm wyposażony w zdolności do rozróżniania rozmaitych kolorów, kształtów i ruchów uzyskuje więc, że tak powiem, jako dodatkową korzyść zdolność do identyfikowania (a także błędnego identyfikowania) rzeczy, wobec których nie ma potrzeb biologicznych. Stworzenie może nie potrzebować zielonych liści, ale jego potrzeba różowych kwiatów doprowadziła do rozwoju systemu poznawczego, którego poszczególne stany mogą, ze względu na związane z potrzebami znaczenie_f, znaczyć_f, że obecne są zielone liście. Być może mimo braku takiej [pierwotnej] potrzeby zacząć na tych obiektach żerować, a zatem uzyskać sposób reprezentowania ich za pomocą elementów, które już mają znaczenie_f.

Jednakże jest poważniejsze zastrzeżenie wobec takiego podejścia do problemu błędnej reprezentacji. Przyjrzyjmy się jeszcze raz bakterii. Jak ustaliliśmy, funkcją ich systemu magnetotaktycznego jest wskazywanie na środowisko pozbawione tlenu. Dlaczego jednak opisywać funkcję tego systemu w taki właśnie sposób? Dlaczego by nie powiedzieć, że funkcją tego systemu jest wskazywanie kierunku geomagnetycznego północnego? Być może, by zachować skromność, powinniśmy przypisywać temu czujnikowi funkcję wskazywania, w którym miejscu (w jakim kierunku) znajduje się magnetyczna (lecz niekoniecznie: geomagnetyczna) północ. Ten prymitywny mechanizm sensoryczny ostatecznie funkcjonuje bez zarzutu, gdy pod wpływem magnesu kieruje swojego właściciela w stronę toksycznego środowiska. W tym wypadku oczywiście coś się złego dzieje, ale nie widzę powodu, żeby winić właśnie mechanizm sensoryczny, nie widzę powodu, aby sądzić, że nie wykonuje on swojej funkcji. Ktoś mógłby równie dobrze utyskiwać, że wskaźnik poziomu paliwa nie wykonuje swojej funkcji, gdyby zbiornik na benzynę był wypełniony wodą (a w konsekwencji kierowca był w błędzie co do ilości niewykorzystanej benzyny). W tych nietypowych okolicznościach przyrząd wykonuje swoje powinności w sposób satysfakcjonujący – wskazując ilość cieczy w zbiorniku. Za to, co poszło nie tak, sam przyrząd nie odpowiada: mianowicie za załamanie zwykłych korelacji (między ilością cieczy w zbiorniku a ilością benzyny w zbiorniku), dzięki którym wskaźnik jest użyteczny jako wskaźnik poziomu paliwa i dzięki którym wskaźnik znaczy_n (w warunkach normalnych), że w zbiorniku jest paliwo. Podobnie nic złego nie dzieje się z czymś systemem percepcyjnym, gdy ktoś, spoglądając na wolniej działający zegarek, traci orientację co do godziny. Funkcją oczu jest informowanie nas, co jest na zegarku; funkcją zaś zegarka jest informowanie nas, która jest godzina. Za właściwe ujęcie tego, co trzeba wiedzieć, odpowiada

często wiele czynników. Trzeba właściwie ująć G i G musi właściwie ująć F . Dlatego nawet jeżeli F jest tym, czego ktoś potrzebuje, lub tym, o czym ktoś musi wiedzieć, funkcją systemu percepcyjnego może być tylko informowanie go o G .

Jeśli zinterpretujemy system sensoryczny bakterii w taki sposób, to jego funkcją jest dostosowanie organizmu do otaczającego pola magnetycznego. Można powiedzieć, że zadaniem północy magnetycznej jest bycie kierunkiem beztlenowej wody. Przemieszczając północną bakterię na półkulę południową, możemy spowodować katastrofę, ale nie dlatego, że przemieszczenie na inną półkulę powoduje dezorientację sensoryczną. Nie, system magnetotaktyczny funkcjonuje tak, jak powinien i jak (przypuszczalnie) wyewoluował. Co najwyżej można by sądzić, że zaszła pewna pomyłka o charakterze poznawczym (bakteria błędnie „wnioskuje” ze swoich warunków sensorycznych o kierunku wody beztlenowej). Jednakże odpowiedź tego rodzaju opiera się na błędzie *petitio principii*, gdyż zakłada, że stworzenie już posiada pojęciową lub reprezentacyjną zdolność do reprezentowania czegoś jako kierunku wody beztlenowej. Nasze pytanie dotyczy tego, czy organizm ma taką zdolność, a jeśli tak, to skąd ją ma¹⁷.

Prawdą jest, że północna bakteria nie ma potrzeby życia w klimacie północnym jako północnym właśnie. Tak więc opisanie funkcji detektorów bakterii w kategoriach roli, jaką odgrywają w identyfikowaniu północy geomagnetycznej, nie polega na opisywaniu, jak ich funkcja wiąże się ze spełnieniem potrzeb bakterii. Nie musimy jednak opisywać funkcji mechanizmu w kategoriach ostatecznych potrzeb biologicznych jego nosiciela¹⁸. Funkcją serca jest tłoczenie krwi, ale kwestia, dlaczego krew powinna krążyć, może pozostać nierozwikłaną tajemnicą.

¹⁷ Fodor (w tekście: J.A. Fodor, *Why Paramecia Don't Have Mental Representations*, „Midwest Studies In Philosophy” 1986, t. X, nr 1, s. 3-23) odróżnia organizmy, dla których reprezentacyjna teoria umysłu nie jest odpowiednia (na przykład bakteria), od tych, dla których jest odpowiednia (na przykład my), w kategoriach późniejszej zdolności do odpowiedzi na bodźce, które nie mają własności nomicznych (właściwości, które nie są wykrywalne przez przetwornik). My tak, ale nie bakteria, jesteśmy w stanie reprezentować coś takiego, jak, powiedzmy, zmięta koszula. Bycie zmiętą koszulą nie jest właściwością rzutowalną. W tym artykule Fodor nie rozważa pytania, skąd mamy tę niezwykłą reprezentacyjną zdolność (sugeruje, że wymaga to zdolności inferencyjnych). Przedstawia jedynie sposób odróżnienia nas od wielu innych percepcyjnych i *quasi*-poznawczych systemów.

Zgadzam się z Fodorem, że jest to ważne rozróżnienie, ale chcę obecnie zrozumieć, w jaki sposób system mógłby uzyskać zdolność do reprezentowania czegoś w taki właśnie sposób. Zdolność do reprezentowania czegoś jako zmiętej koszuli (w przypadku, gdy implikuje to zdolność do błędnego reprezentowania jako takiego) z pewnością nie jest wrodzona.

¹⁸ B. Enç, *Intentional States of Mechanical Devices...*, *op. cit.*, s. 168 twierdzi, że komórki fotoreceptora muszki owocówki mają m.in. funkcję umożliwiania muszce dotarcia do wilgotnych miejsc (dzięki korelacji między ciemnymi i wilgotnymi miejscami). Nie mam żadnych zastrzeżeń co do opisy-

Tak brzmi kłopotliwe dla nas pytanie: *skoro* system potrzebuje F i skoro mechanizm M pozwala organizmowi wykrywać, identyfikować czy rozpoznawać F , to powstaje pytanie, jak dany mechanizm wykonuje tę funkcję? Czy robi to, reprezentując pobliskie F -y jako pobliskie F -y czy, być może, reprezentuje je jedynie jako pobliskie G -ki, zdając zaspokojenie swoich potrzeb na naturę (korelację zachodzącą między F i G)? Opisanie mechanizmu poznawczego jako F -detektora (a, co za tym idzie, jako mechanizmu odgrywającego istotną rolę w zaspokajaniu potrzeb organizmu) nie jest jeszcze opisem funkcjonalnej struktury, dzięki której mechanizm wykonuje swoje zadanie. Wszystko, co wiemy, wiedząc, że O potrzebuje F i że m pozwala O wykrywać F , sprowadza się do tego, że M albo znaczy _{F} , że F jest obecne, albo znaczy _{F} , że występuje tam G , gdzie G jest, w naturalnym otoczeniu O , znakiem naturalnym obecności F (gdzie G znaczy _{n} , że jest F)¹⁹. Jeśli potrzebuję witaminy C , to mojemu systemowi percepcyjno-poznawczemu nie powinno się automatycznie przypisywać zdolności do rozpoznawania przedmiotów zawierających witaminę C (w znaczeniu _{F} , że zawierają witaminę C) tylko dlatego, że zaopatruje mnie on w informację niezbędną do zaspokojenia tej potrzeby. Reprezentowanie rzeczy jako pomarańczy i cytryn zupełnie wystarcza.

Problem, przed którym stajemy, polega na ujęciu zdolności do błędnego reprezentowania przez system bez sztucznego rozdmuchiwania funkcji naturalnych tego systemu. Potrzebujemy pewnego zasadniczego sposobu opisu naturalnej funkcji mechanizmu tego, co jego różne stany znaczą: nie tylko znaczą _{n} , ale co znaczą _{F} . Wydaje się trochę naciągane (przynajmniej jak na mój gust), aby opisywać mechanizm sensoryczny bakterii

wania rzeczy w ten sposób. Pozostaje jednak pytanie: w jaki sposób komórka fotoreceptora wykonuje swoją funkcję? Możemy odpowiedzieć na to pytanie bez zakładania, że jest jakiś mechanizm muszki, którego funkcją jest wskazanie stopnia wilgotności. Sensoryczny mechanizm może wykonać tę funkcję, jeśli istnieje coś, co jedynie wskazuje na natężenie światła – tj. fotoreceptor. To pozwoliłby muszce docierać do wilgotnych miejsc. Podobnie magnetotaktyczny zmysł bakterii pozwala (i – powiedzmy – posiada funkcję umożliwiającą) bakterii unikanie wody bogatej w tlen. Ale sposób, w jaki to robi (można by argumentować), polega na posiadaniu zmysłu, który wskazuje kierunek pola magnetycznego i ma funkcję wskazywania tego kierunku.

¹⁹ W ujęciu Fodora (w *Psychosemantics...*) mówi się w ten sposób jedynie, że zidentyfikowanie semantyki M (pewnej reprezentacji umysłowej) z warunkami wejścia (względny wobec określonych warunków normalności) nadal pozostawia wiele niedopowiedzeń. Można by twierdzić, że warunkiem wejścia jest brak (obecność) tlenu lub konkretna orientacja pola magnetycznego. Odwołanie do historii doboru tego mechanizmu nie decyduje o tym, co jest właściwą konkretyzacją warunków wejścia – nie mówi zatem, czy bakterie są zdolne do błędnego reprezentowania czegokolwiek. Jak sądzę, Fodor zdaje sobie sprawę z tych pozostałości niezdeterminowania i dlatego interesująco sugeruje (przypis 9), że ten problem jest analogiczny do problemów związanych z określeniem przedmiotu percepcyjnego w teorii percepcji.

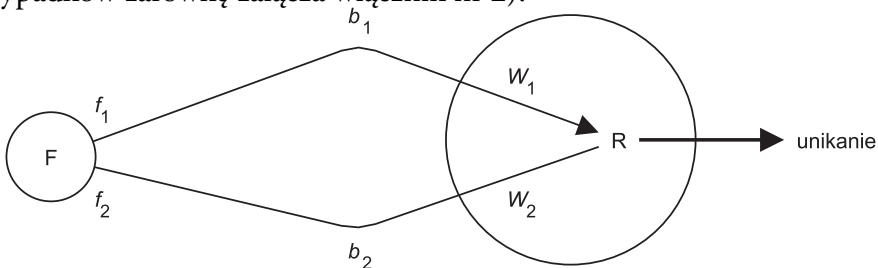
jako wskazujący miejsce, gdzie znajduje się tlen i mający funkcję wskazywania tego miejsca. Z tego właśnie powodu wydaje się, że mechanizm nie wykonuje swojej funkcji w warunkach zwodniczych (na przykład w obecności magnesu prętowego). Ostatecznie przecież jest to czujnik magneto-taktyczny, a nie chemotaktyczny. Jeśli jednak postanowimy opisać funkcję tego czujnika w zaproponowany właśnie skromniejszy sposób, to przestaje on służyć za przykład systemu ze zdolnościami do błędnego reprezentowania. Północna bakteria (przeszczepiona na półkulę południową) nie będzie błędnie reprezentować czegokolwiek, gdy pod wpływem magnetotaktycznego czujnika przesuwa się w górę (w kierunku geomagnetycznym północnym) w stronę – śmiertelnych dla niej – wód powierzchniowych. Ustawienie jej magnetosomów będzie wówczas znaczyło_n, to, co dotychczas znaczyło_n, co jest jego funkcją znaczenia_n, co powinno znaczyć_n: mianowicie, że jest to kierunek magnetycznej północy. Za klęskę odpowiadają nietypowe warunki otoczenia. Nie uratujemy nawet szczątkowej zdolności do błędnego reprezentowania, zakładając, że bakteria, pod wpływem magnesu prętowego, błędnie reprezentuje przynajmniej kierunek geomagnetycznej północy. Zatem po raz kolejny powstaje ten sam problem: dlaczego mamy zakładać, że funkcją tego mechanizmu jest wskazywanie kierunku geomagnetycznej północy, a nie po prostu kierunku otoczenia pola magnetycznego? Jeśli opisujemy funkcję tylko w drugim sensie, niemożliwe okazuje się zwodzenie organizmu i organizm nie jest zdolny do błędnego reprezentowania czegokolwiek. Dlatego wtedy stany wewnętrzne organizmu tylko znaczą_f, że pole magnetyczne wskazuje w tym kierunku, i (tak jak kompas) robią to zawsze właściwie.

5. *Determinacja funkcjonalna*

Dla jasności ograniczyłem rozważania do omawiania prostych organizmów o pierwotnych zdolnościach reprezentacyjnych. Nie powinno więc nikogo dziwić, że na tym poziomie nie można znaleźć jasnej i jednoznacznej zdolności do błędnego reprezentowania. Ta zdolność bowiem – i, jak przypuszczam, zależna od niej zdolność do przekonania – wymaga, by zdolność systemu w zakresie przetwarzania informacji odznaczała się odpowiednio dużą złożonością. Gdzieś między pojedynczą komórką a człowiekiem przekraczamy ten próg. Właśnie celem ostatniej części [artykułu] jest opisanie natury tego progu, czyli opisanie złożoności tego rodzaju, odpowiedzialnej za występujące u bardziej rozwiniętych organizmów zdolności do błędnego reprezentowania.

Załóżmy, że organizm (w przeciwieństwie do naszej bakterii) ma zdolność do wykrywania obecności pewnej toksycznej substancji *F* na dwa

sposoby. Może się tak dzieć dzięki temu, że organizm jest wyposażony w dwie modalności zmysłowe, z których każda (na swoisty dla siebie sposób) jest wrażliwa na F (lub jakiś specyficzny modalnie znak naturalny F), lub dzięki temu, że pojedyncza modalność zmysłowa wykorzystuje różne zewnętrzne znaki (lub symptomy) F . Jako przykład tego ostatniego rozważmy, jak możemy – na jeden z dwóch sposobów – wzrokowo identyfikować dęby: przez charakterystyczny kształt liści (latem) lub przez charakterystyczną teksturę i wzór kory (zimą). Mamy zatem dwa stany wewnętrzne lub dwa warunki, W_1 oraz W_2 , z których każdy powstał w efekcie innego łańcucha poprzedzających go zdarzeń, będących znakami naturalnymi obecności F . Każdy z nich znaczy_n, że F jest obecne. Założmy ponadto, że ze względu na potrzebę ucieczki od toksycznych F -ów wymienione stany wewnętrzne są związane z trzecim stanem – nazwijmy go R – który wyzwała lub uwalnia wzór zachowania unikowego. Rysunek 1 obrazuje odpowiednie fakty. Oczywiście, R jest także znakiem naturalnym F . W normalnych warunkach R nie występuje, chyba że F jest obecne. f_1 i f_2 są własnościami typowymi dla normalnych F -ów, b_1 i b_2 natomiast to bodźce proksymalne. Jeśli więc wyposażymy system w jakiś odpowiednik F (analogiczny do magnezu prętowego u bakterii), coś, co wykazuje niektóre własności rzeczywistych f (np. f_1), wywołamy łańcuch zdarzeń (b_1 , W_1 , R i unikanie), który występuje normalnie i rzeczywiście jest jedynym właściwym łańcuchem w obecności F . Jeśli spojrzymy na stan wewnętrzny R i spytamy, co on znaczy_f w tych zwodniczych warunkach, okaże się, że nie będziemy mogli powiedzieć (tak jak moglibyśmy w przypadku bakterii), iż znaczy_f cokolwiek innego niż F (tj. nie jest bardziej proksymalny niż samo F). Nawet jeśli b_1 (za pomocą W_1) wyzwała zdarzenie R , R nie znaczy_n (nie może zatem znaczyć_f), że b_1 (lub f_1) zachodzi. R jest analogiczne do żarówki podłączonej do włączników połączonych równolegle, gdzie zamknięcie któregośkolwiek z nich powoduje włączenie światła. Jeśli żarówka się świeci, to nie znaczy_n, że włącznik nr 1 jest zamknięty, nawet gdy jest to zamknięcie tego rodzaju, które powoduje, że światło się zapala. To nie znaczy_n, że tak jest, gdyż nie zachodzi regularna korelacji między świecącą żarówką a zamknięciem włącznika światła nr 1 (w 50 procentach wypadków żarówkę załącza włącznik nr 2).



Gdy system wykrywania opisany powyżej uznamy za mający funkcję umożliwiającą organizmowi wykrywanie F , wówczas wielość sposobów wykrywania F ma taką konsekwencję, że pewne stany wewnętrzne (na przykład R) mogą wskazywać (zatem znaczyć_f), że F jest obecna bez wskazywania czegokolwiek, co dotyczyłoby warunków pośrednich (tj. f_1 lub b_1), które „powiedziałyby” mu [tj. systemowi wykrywania], że F jest obecna. Problem z bakterią dotyczył znalezienia sposobu na to, by orientacja jej magnetosomu znaczyła_f, że pozbawiona tlenu woda znajduje się w pewnym kierunku, nie wykluczając arbitralnie możliwości jego znaczenia_f, że pole magnetyczne jest ustawione w tym kierunku. Teraz możemy dostrzec, że wraz z wieloma zasobami opisanymi na rysunku 1 taka możliwość może być niearbitralnie odrzucona. R nie może znaczyć_f, że f_1 lub b_1 zachodzi, ponieważ nie znaczy_n, że tak jest, nawet w warunkach optymalnych. Możemy więc sądzić, że znaleźliśmy niepochozny przypadek błędnej reprezentacji (tj. znaczenie_f R -a, że F jest obecne, gdy F nie jest obecne), którego nie można odrzucić, zastępując opis tego, co R znaczy_f, opisem, który wyeliminowałby błędną reprezentację. Groźba rozdmuchania możliwych znaczeń_f, wyrastająca z różnorodności sposobów, w jaki można by opisać naturalną funkcję systemu, została oddalona.

Wciąż jednak można by twierdzić, że nie musimy tego uznawać za autentyczny przypadek błędnej reprezentacji, jeśli chcemy uznać, że R ma znaczenie_n wyrażone w postaci alternatywy. Zapalenie żarówki (podłączonej do połączonych równolegle włączników) nie znaczy_n, że jakikolwiek włącznik jest naciśnięty, ale wskazuje, że jeden z tych włączników jest naciśnięty. Podobnie można by twierdzić, że nawet jeśli funkcją mechanizmu mającego R za swój stan końcowy jest ostrzeżenie organizmu o obecności F , mechanizm czyni to przez wskazywanie R i posiadanie funkcji wskazywania, że występuje pewien warunek wyrażony alternatywą – mianowicie f_1 lub f_2 (bądź b_1 lub b_2). Nasz hipotetyczny organizm przez pomyłkę wycofuje się z F nie dlatego, że błędnie reprezentuje odpowiednik F jako F , ale dlatego, że to, na co wskazuje poprawnie (czyli to, że substytutem f jest f_1 lub f_2), nie jest już skorelowane w normalny sposób z tym, że coś jest F .

Bez względu na to, jak bardzo wszechstronny system możemy zaprojektować, niezależnie od tego, jak wiele dróg informacyjnego dostępu możemy zapewnić organizmowi, zawsze będzie istniała możliwość opisanego jego funkcji (a zatem i znaczenia_f jego różnych stanów) jako detekcji własności proksymalnego wejścia opisanej złożoną alternatywą. Przynajmniej będzie to zawsze możliwe, jeśli mamy określony zbiór członów alternatywy, do których możemy się cofnąć.

Załóżmy jednak, że mamy system zdolny do jakiejś formy uczenia asocjacyjnego. Założmy, innymi słowy, że przez powtarzaną ekspozycję

na wb (warunkowany bodziec) w obecności F zachodzi zmiana. R (a co za tym idzie, zachowanie unikowe) może zostać wywołane przez samo wystąpienie wb . Co więcej, staje się jasne, że praktycznie nie ma ograniczeń co do rodzaju bodźca, który może nabyć tę „przenoszona” skuteczność w wyzwaniu R i dalszym zachowaniu unikowym. Prawie każde b może stać się wb , ustanawiając w ten sposób „sterowanie” R -em, funkcjonując (w „doświadczeniu” organizmu) jako znak F .

Obecnie mamy mechanizm zdolności poznawczych, który nie tylko przekształca różnorodność sensorycznych danych na wejściu (b_i) w jeden stan determinujący na wyjściu (R), ale jest też zdolny do przekształcenia charakteru odwzorowywania wielo-jednoznaczności w czasie. Jeśli ograniczymy się do danych sensorycznych na wejściu (b_i na rysunku 1), R znaczy_n jedną rzecz w t_1 (na przykład b_1 albo b_2), coś innego w chwili t_2 (na przykład b_1 lub b_2 , lub, w wyniku uczenia, wb_3), czy coś jeszcze innego w późniejszym czasie. Po prostu to, co R znaczy_n, zależy od indywidualnej historii uczenia – od tego, jakie b_i stało się wb_i dla niego. Nie istnieje niezmiennie w czasie znaczenie_n dla R ; toteż nic takiego nie mogłoby być w czasie jego funkcją wskazywania. W kategoriach b_i doprowadzających do R , R nie może mieć niezmiennego w czasie znaczenia_f.

Oczywiście, przez cały ten czas R nadal wskazuje na obecność F . Dzieje się tak, gdyż zgodnie z hipotezą każdy nowy b_i , dla którego R zostaje warunkowane, jest znakiem naturalnym F . Uczenie się to proces, w którym bodźce wskazujące na obecność F są kolejno wskazywane przez odpowiednie stany organizmu (w tym wypadku jest to R). Zatem jeśli mamy uznać te mechanizmy poznawcze za posiadające niezmienną w czasie funkcję (coś, co jest implikowane przez ciągłą – a w istocie, w wyniku uczenia się skuteczniejszą – realizację przez nie danej potrzeby), to musimy ich funkcje potraktować nie jako wskazujące na naturę proksymalnych (a nawet dystalnych) warunków, które wyzwalały pozytywne reakcje (b_i i f_i), ale jako wskazujące na warunek (F), którego te zróżnicowane bodźce są znakami. Funkcją naturalną tak opisanego mechanizmu jest zatem wskazywanie obecności F . Dlatego właśnie występowanie R znaczy_f, że F jest obecne. Ono nie znaczy_f, że b_1 i b_2 lub ... b_x zachodzi, mimo że na każdym etapie rozwoju będzie to znaczyło_n dla określonej wartości x .

Na tym poziomie złożoności system mający nie tylko wiele kanałów dostępu do tego, co powinien wiedzieć, ale także zasoby do rozbudowy swoich zasobów gromadzenia informacji, ma, moim zdaniem, autentyczną zdolność do błędnego reprezentowania. Gdy w normalnym łańcuchu znaków naturalnych pojawi się usterka, gdy np. pojawi się wb_7 (wyuczony znak F) w warunkach, w których to nie znaczy_n, że F jest obecne (w sposób, w jaki zepsuty zegar nie znaczy_n, że jest godzina 3:30), R nadal zna-

czy_f, że *F* zachodzi (choć oczywiście nie znaczy_n, że *F* zachodzi). Znaczy_f, że tak jest, ponieważ to właśnie powinno znaczyć_n, co jest jego funkcją naturalną znaczenia_n, i nie jest dostępny żaden inny warunek, w którym to może znaczyć_f²⁰.

Z angielskiego przełożyła *Krystyna Bielecka*
Przekład przejrzał *Marcin Miłkowski*

Misrepresentation

The reply to the question about the possibility of misrepresentation is more informative about the nature of representation than an account of representing correctly. One can distinguish natural and functional signs. The natural signs are signals that indicate something; for example, an expanding metal strip indicates that its temperature is changing. The relationship between this event and its sign is objective, and occurs in virtue of natural laws. For this reason, natural signs cannot misrepresent anything. Misrepresentation can occur only when a natural sign has a function – when a sign has a functional meaning. The basis of (functional) meaning is having functions, which are related to biological needs. However, it's problematic to precisely determine which functions are natural and which not.

²⁰ Jestem wdzięczny Berentowi Ençowi, Dennisowi Stampemu i Jerry'emu Fodorowi za ich pomocne uwagi, zarówno konstruktywne, jak i krytyczne, dotyczące szkicowych wersji tego eseju.