
Reprezentacje w systemach klasycznych i koneksjonistycznych

MARCIN MIŁKOWSKI

Instytut Filozofii i Socjologii PAN

Zakład Logiki i Kognitywistyki

Streszczenie. Autor artykułu broni tezy, że niektóre systemy obliczeniowe mogą mieć własności semantyczne. Wskazana została klasa systemów obliczeniowych, w których reprezentacje mogą mieć przynajmniej dwie własności: własność odnoszenia się do obiektów (desygnowanie) i własność wspomaganie rozpoznawania obiektów oznaczanych przez daną reprezentację (konotowanie). Autor argumentuje także, że własności semantyczne reprezentacji nie zależą wyłącznie od architektury systemów obliczeniowych, w których te reprezentacje występują. Konkretna architektura obliczeniowa nie jest czynnikiem kluczowym, a bodaj najmniej istotne są same rodzaje struktur danych, które mają mieć własności desygnowania czy konotowania. Własność desygnowania czy konotowania nie musi być zlokalizowana w samych reprezentacjach, może być własnością wyższego rzędu, powstającą w mechanizmie wyższego poziomu. Własności semantyczne reprezentacji mogą być wielorako realizowane. Systemy klasyczne, koneksjonistyczne czy też hybrydowe mogą równie dobrze mieć własności semantyczne, jak ich nie mieć.

1. Wstęp

Czy reprezentacje w systemach obliczeniowych mogą mieć własności semantyczne? I czy tzw. klasyczne podejście do sztucznej inteligencji (GOFAI, *Good-Old Fashioned AI*) różni się pod tym względem od koneksjonizmu?

W tym artykule bronić będę tezy, że niektóre systemy obliczeniowe mogą mieć własności semantyczne¹. Wbrew wielu głosom sceptycznym

¹Abstrakcyjna struktura systemu obliczeniowego, np. w postaci zapisu programu komputerowego, o ile nie odpowiada prawidłowościom przyczynowym (czyli jest realizowana przez komputer), nie może mieć żadnych własności semantycznych. Wątpliwe jest też, aby sam taki zapis mógł uchodzić

(Searle 1995, Searle 1999, Harnad 1990) wskażę klasę systemów obliczeniowych, w których reprezentacje mogą mieć przynajmniej dwie własności: własność odnoszenia się do obiektów (desygnowanie) i własność wspomaganie rozpoznawania obiektów oznaczanych przez daną reprezentację (konotowanie²). Mówiąc inaczej, reprezentacje w tych systemach obliczeniowych mają zarówno ekstensję, jak i intensję. Odpowiadam więc pozytywnie na pierwsze pytanie.

Na drugie pytanie odpowiedź będzie negatywna. Własności semantyczne reprezentacji nie zależą wyłącznie od architektury systemów obliczeniowych. Konkretna architektura obliczeniowa nie jest czynnikiem kluczowym, a bodaj najmniej istotne są same rodzaje struktur danych, które mają mieć własności desygnowania czy konotowania. W istocie bowiem własność desygnowania czy konotowania nie musi być zlokalizowana w samych reprezentacjach, może być własnością wyższego rzędu, powstającą w mechanizmie wyższego poziomu. Sprowadzanie własności semantycznych do wewnętrznych własności indywidualnych struktur danych jest błędem lokalizacji funkcjonalnej (por. Wimsatt 2007; Bechtel i Richardson 1993). Będę więc twierdził, że własności semantyczne reprezentacji mogą być wielorako realizowane. Systemy klasyczne, koneksjonistyczne czy też hybrydowe mogą równie dobrze mieć własności semantyczne, jak ich nie mieć.

W dyskusjach nad zdolnościami semiotycznymi sztucznych systemów poznawczych – a w szczególności systemów obliczeniowych – często porusza się kwestię internalizmu i eksternalizmu semantycznego. Zagadnienia te dotyczą głównie konotacji (treści) wyrażań, nie zaś ich odniesienia. „Odniesienie” natomiast jest analizowane stosunkowo jednomyślnie w ramach przyczynowej teorii odniesienia. Ze względu na brak miejsca będę abstrahować od obszernych dyskusji między internalistami a eksternalistami. Przedstawię jednak analizę mechanizmów semantycznych, która bierze pod uwagę zarówno środowisko wewnętrzne mechanizmu (endostrukturę), jak i jego otoczenie zewnętrzne (egzostrukturę). Jeśli internalizm ma więc pociągać za sobą metodologiczny solipsyzm, to moja analiza będzie z nim niezgodna: z solipsystyczną

za *system obliczeniowy*. Niżej mam na myśli wyłącznie fizycznie zrealizowane systemy obliczeniowe, kiedy mówię po prostu o systemach obliczeniowych.

²Przez „konotację” rozumie się zazwyczaj albo (1) zbiór konstytutywnych cech desygnatów; (2) charakter aspektowy czy też tryb prezentacji desygnatów; (3) warunki umożliwiające wyróżnianie desygnatów (np. w postaci listy cech konstytutywnych). W tym tekście termin ten jest używany w znaczeniu (3). Nie zajmuję się też trybem prezentacji, jeśli nie jest on wykorzystywany do identyfikacji desygnatów.

analizą niezgodne jest bowiem uwzględnianie egzostruktury systemu. Argumenty za tego rodzaju skrajnym internalizmem są jednak wątle (por. McClamrock 1995, rozdział 2).

2. Przetwarzanie informacji a ekstensja reprezentacji

Teza, że systemy sztuczne mogą zawierać reprezentacje, które odnoszą się do obiektów, jest stosunkowo mało kontrowersyjna – spór dotyczy przede wszystkim tego, czy te reprezentacje mogą mieć intensję, nie zaś tylko ekstensję (Chalmers 1992). Do innych przedmiotów mogą odnosić się chociażby oznaki — po prostu powiązane przyczynowo z tym, na co wskazują. Oznakę, np. dym oznaczający ogień, można wytworzyć sztucznie i celowo, więc system obliczeniowy nie byłby niczym nadzwyczajnym. System zawierający oznaki i z nich korzystający będzie miał reprezentacje, które ewidentnie się do czegoś odnoszą. Oba warunki — zawieranie znaków o charakterze oznak i korzystanie z nich — spełniają wszystkie systemy uzyskujące i przetwarzające informacje, w których istnieją mechanizmy reprezentacyjne³.

Systemy obliczeniowe rozumiem jako pewnego rodzaju systemy przetwarzania informacji (nie rozstrzygając w tym miejscu kontrowersyjnej kwestii, czy jest możliwe istnienie przetwarzania informacji na drodze nieobliczeniowej; jeśli nie jest, są to klasy równoważne). Nie wszystkie te systemy muszą przetwarzać informacje z otoczenia; jeśli informacje te docierają do systemu w postaci innych reprezentacji, to ekstensja reprezentacji systemu zależy od tego, czy przed dotarciem do systemu odnosiły się one do czegokolwiek. Krótko mówiąc, systemy korzystające z informacji już przetworzonych niekoniecznie mają reprezentacje niepuste; ich semantyczne własności zależą bezpośrednio od własności informacji wejściowych. Jeśli reprezentacje są wprowadzane przez człowieka, który np. opisuje za ich pomocą fikcyjny świat wirtualny, to nie będą one miały ekstensji w postaci obiektów empirycznych. Jednak ważniejsze jest pytanie, w jaki sposób powstają reprezentacje obliczeniowe odnoszące się do rzeczywistych obiektów — i to bez oczywistej interwencji człowieka. Aby na to pytanie odpowiedzieć, należy zanalizować samo pojęcie przetwarzania informacji.

Pojęcie „informacja” jest stosowane często i w bardzo różnych kontekstach (Poczobut 2005). Dlatego też nie sposób mówić o „informacji jako takiej”, trzeba wybrać jakąś określoną koncepcję. Postaram się jednak nakreślić ogólną ideę, z którą zgodna będzie większość dzisiej-

³W poniższych rozważaniach nad przetwarzaniem informacji wykorzystuję analizy z Miłkowski (2008).

szych ujęć kwantytatywnych. Będzie to analiza w stylu mechanistycznym⁴. Skupiam się na przetwarzaniu informacji z otoczenia — uzyskane wnioski można potem uogólnić także na mechanizmy, które operują na istniejących już nośnikach informacji, nie przetwarzając uzyskanych samodzielnie informacji o otoczeniu. Przetwarzanie informacji o otoczeniu wymaga jego właściwej reprezentacji w systemie poznawczym. W przeciwieństwie do wielu teoretyków (Harnad 1990) nie uważam jednak, aby interakcje sensomotoryczne z otoczeniem były warunkiem koniecznym referencyjności reprezentacji w systemie — są one warunkiem wystarczającym, lecz system może także zawierać reprezentacje dotyczące własnych stanów wewnętrznych lub generować nowe informacje na podstawie dostarczonych już informacji w postaci innej niż percepcyjna (zawartych w odpowiednim nośniku).

Przetwarzanie informacji jest procesem, który zachodzi (przynajmniej) w systemach wykorzystujących informacje do sterowania własnym zachowaniem. Nie zachodzi ono tam, gdzie mamy do czynienia jedynie z zakodowaną informacją, np. wydrukowana w podręczniku tabliczka mnożenia ani wzór nie przetwarzają informacji. Podobnie słoje w drewnie, choć stanowią oznaki i są wywoływane przez środowisko zewnętrzne, a więc informują o nim, nie przetwarzają informacji jako takie. Tabliczki mnożenia, wzory i słoje drewna to jedynie *nośniki informacji*. Przetwarzanie bowiem — w przeciwieństwie do samego zawierania lub przechowywania informacji — polega na tym, iż istnieje niezawodny mechanizm dokonujący przekształcenia informacji, czyli mechanizm działający stosunkowo stabilnie w danym środowisku. Informacje te muszą do tego mechanizmu dotrzeć, a następnie zostać przesłane dalej, czyli po przetworzeniu muszą być dostępne dla samego systemu, a zatem docierać do innych jego składników.

Wspomniany mechanizm musi przetwarzać informacje pochodzące z danej sytuacji — czyli informacje już w niej zawarte⁵. Każda sytuacja (rozumiana abstrakcyjnie, niekoniecznie jako fizyczny stan rzeczy) może zostać opisana przy użyciu ciągu informacji. Mówiąc poglądowo, można zadać pewien ciąg pytań rozstrzygnięcia na temat danej sytuacji, a ciąg odpowiedzi — „tak” lub „nie” — to ciąg informacji w niej zawartych (w ilościowych teoriach informacji taka minimalna odpowiedź zwana

⁴Wyjaśnienie mechanistyczne rozumiem podobnie jak Bechtel i Richardson (1993) lub Craver (2006). Wyjaśnienia mechanistyczne do obliczeniowości odnosi m.in. G. Piccinini (2007).

⁵Analiza pojęcia informacji jest szczególnie płodna w kategoriach semantyki sytuacyjnej. Por. Israel i Perry (1990). Tezę, że informacja jest zawarta w sytuacjach, uzasadnia Collier (1990).

jest niekiedy „logonem”; nie idzie tu o bit w sensie informatycznym). Uznaję więc, że struktura sytuacji sama w sobie zawiera informacje, które zostają ujawnione w jej opisie. Pewne sytuacje można wskazać w sposób jednoznaczny przy użyciu ciągów skończonych o małej długości; inne są znacznie bogatsze informacyjnie. Długość i złożoność tych ciągów charakteryzuje się matematycznie. Przykładem może być algorytmiczna teoria informacji, rozwijana przez Kołmogorowa, a następnie przez Chaitina; pozwala ona rozróżnić stochastyczne ciągi informacji od niestochastycznych (algorytmicznych)⁶. Opisy, w których widać maksima regularności, a które jednocześnie pozwalają jednoznacznie identyfikować egzemplarze lub typy obiektów, to opisy, których poszukuje nauka. One bowiem chwytają rzeczywiście zachodzące zjawiska. Optymalnie skompresowane, a więc algorytmiczne informacje o sytuacji opisywanej jako typ lub jako konfiguracja typów są informacjami na temat rodzajów naturalnych.

Proste oznaki — takie jak kreski na podziałce termometru spirytusowego za oknem — mogą służyć do tworzenia opisu sytuacji (w tym wypadku opisu w kategoriach analogowych, w ciągłych wartościach oznaczających temperaturę). Istnieje wówczas pewien mechanizm opisujący sytuacje czy też generujący informacje, czyli mechanizm semiotyczny. O przetwarzaniu w całym znaczeniu tego słowa nie ma mowy, gdyż uzyskane informacje nie są poddawane dalszej obróbce, nie są w żaden sposób gromadzone, selekcjonowane i nie służą do modyfikacji zachowania. Nie są dostępne dla innych komponentów systemu, jakim jest termometr. Dostęp do wskazań temperatury ma tylko użytkownik termometru. Termometr pozostaje jedynie nośnikiem informacji, a nie systemem, który je przetwarza. Można by co najwyżej powiedzieć, z przymrużeniem oka, że to człowiek plus termometr okienny stanowią mechanizm przetwarzający informacje o temperaturze powietrza za oknem.

Przetwarzanie informacji wymaga istnienia mechanizmu, który nie tylko generuje odpowiednie opisy sytuacji (wydobywa informacje zawarte w sytuacjach), ale także z nich korzysta. Praktycznym kryterium pozwalającym odróżnić mechanizmy informacyjne — wydobywające informacje, semiotyczne — od przetwarzających informacje jest to, czy istnieje zysk predykcyjny z opisu w kategoriach informacyjnych w stosunku do opisu czysto mechanicznego. Jeśli opis można wyciąć brzytwą Ockhama, nie tracąc żadnych istotnych przewidywań zachowania syste-

⁶Najbardziej przystępne wprowadzenie do algorytmicznej teorii informacji daje Chaitin (1975).

mu, to znaczy, że mamy do czynienia wyłącznie z mechanizmem informacyjnym, nie zaś przetwarzającym informacje. To kryterium określam jako praktyczne, gdyż jest zawodne i ma naturę epistemiczną. Możemy przecież nie dostrzec jakiegoś aspektu złożonego systemu, dzięki któremu jest on nie tylko systemem semiotycznym czy informacyjnym, ale także systemem przetwarzającym informacje⁷. Druga praktyczna reguła dotyczy magazynowania informacji — tylko mechanizmy, które nie reagują bezpośrednio na bodźce, lecz wykorzystują informacje o nieobecnych już sytuacjach, mogą uchodzić za mechanizmy reprezentacyjne. Mechanizmy reprezentacyjne są najbardziej interesującą klasą mechanizmów przetwarzających informacje z otoczenia. Istniejące w przyrodzie systemy poznawcze, jak się zdaje, wszystkie korzystają z różnego rodzaju mechanizmów reprezentacyjnych⁸. Ponieważ reprezentacje są generowane z informacji wydobytych z sytuacji, mają one ekstensję — jest nią mianowicie ten stan rzeczy, który te informacje opisują. Nie przesądzam w tym miejscu, jaką strukturę będą miały desygnaty reprezentacji: informacje, choć sprowadzalne do pytań rozstrzygnięcia na temat sytuacji, mogą mimo to mieć różny charakter. Zależy on od tego, jakie stawiane jest pytanie, czyli jaki aspekt sytuacji wchodzi w grę.

Jedyną niezawodną kryterium przetwarzania informacji to istnienie odpowiedniej budowy — architektury — wewnętrznej powiązanej z odpowiednim kontekstem działania mechanizmu przetwarzającego informacje. Dopiero zbadanie funkcjonowania systemu, możliwie wyczerpujące, daje odpowiedź, czy system przetwarza informacje — uzyskuje

⁷W wypadku formalnych systemów przetwarzających informacje, tj. systemów nie korzystających z informacji zawartych w rzeczywistych sytuacjach, pominąć można etap semiotyczny, czyli opisywania sytuacji za pomocą nośnika informacji — w tym sensie, że nie korzysta się tu z oznak niosących informacje o środowisku, tylko z innych nośników informacji (w takim formalnym systemie informacje zawarte na nośniku informacji nie muszą być rozumiane, w przeciwieństwie do pozostałych wypadków, jako z konieczności prawdziwe: nośniki nie są w tym wypadku z konieczności oznakami, mogą być też znakami konwencjonalnymi). Kryterium odróżniania takiego systemu (zrealizowanego fizycznie) od wszelkich innych systemów mechanicznych pozostaje analogiczne: jeśli działanie maszyny można opisać mechanicznie w prostszy sposób (tj. uzyskać większą kompresję informacji na temat działania maszyny, ujmując rzecz w kategoriach Chaitina), to znaczy, że nie należy jej zaliczać do klasy urządzeń przetwarzających informacje.

⁸Pojęcie reprezentacji należy tu rozumieć szeroko; nie idzie tu o reprezentację symboliczną czy językową (tzw. reprezentacje subsymboliczne w sieciach neuronalnych też zaliczam do reprezentacji).

je, następnie przenosi do kolejnych komponentów, a wreszcie selekcjonuje itd. — czy też nie. Skoro wszelkie obliczanie jest przetwarzaniem informacji, to każda architektura obliczeniowa spełnia warunek przetwarzania informacji. Nie każda jednak przetwarza informacje wydobyte z sytuacji — ekstensja reprezentacji zależy więc od powiązania przyczynowego ich ze światem, nie zaś od tego, czy system jest koneksjonistyczny, czy klasyczny. Kluczowe jest oddziaływanie z otoczeniem, z którego wydobywane są informacje, lub ekstensjonalność informacji już zakodowanych (w wypadku systemów nieprzetwarzających informacji ze świata bezpośrednio).

Kryterium budowy systemu jednak jest nieoperacyjne — nie istnieje żaden ogólny algorytm wykrywania dowolnego rodzaju informacji, gdyż taki algorytm mógłby posłużyć do rozstrzygnięcia problemu, czy dany ciąg informacji jest algorytmiczny (niestochastyczny), czy też nie (algorytmiczne ciągi informacji są pewnym rodzajem informacji). Ponieważ ten problem jest nierozstrzygalny (Chaitin 1987), to nie możemy z góry zawsze wiedzieć, co jest relewantną informacją, a co nie — nie możemy więc ocenić działania każdego systemu.

Oprócz szczegółowych wskaźników istnienia procesów przetwarzania informacji istnieje szereg kryteriów ogólnych stosowanych w metodologii naukowej, takich jak prostota opisu, wartość eksplanacyjna i predykcyjna; spójność opisu i realistyczne ujęcie interakcji przyczynowych, nie zaś ich ignorowanie. Mechanizm przetwarzający musi być względnie izolowany, aby można było go wyróżnić z otoczenia; jeśli system ma cechy umożliwiające mu przetwarzanie informacji ze środowiska, to interakcje wewnątrzsystemowe muszą być częstsze niż interakcje na jego granicach⁹. Bezpośredni styk z otoczeniem mają tylko wejścia i wyjścia systemu. Co więcej, realizator mechanizmu musi być opisywalny mechanistycznie, w kategoriach strukturalno-funkcjonalnych¹⁰. Mechanizm przetwarzania jest „zadany” przez strukturę mechanizmu i poddaje się też wyjaśnieniu adaptacjonistycznemu, gdy w grę wchodzi systemy biologiczne.

⁹Mówiąc nieco precyzyjniej, strukturę systemu można odkryć, badając istotne statystycznie regularności w jego funkcjonowaniu. Granice składników najłatwiej wyznaczać przez różnice względnych częstości interakcji, lecz odkrycie relacji między składnikami wymaga zastosowania bardziej skomplikowanych miar statystycznych. Jeśli system podlega prawidłościom nie tylko statystycznym, lecz znanym skądinąd prawom ścisłym, przy jego indywidualności i wyznaczeniu jego struktury wewnętrznej można je oczywiście także wykorzystać.

¹⁰Analogiczne kryteria postulowałem w odniesieniu do procesów algorytmicznych (obliczeniowych). Por. Miłkowski (2007) i Miłkowski (2009).

Przedstawiona wyżej, uproszczona wersja semantyki informacyjnej wypływająca z analizy mechanizmów reprezentacyjnych wystarcza, aby pokazać, że systemy obliczeniowe — zawierające mechanizmy reprezentacyjne — mogą mieć stany wewnętrzne, które mają ekstensję. Sposób realizacji tych zdolności, czyli architektura sprzętowa czy też programowa, nie jest w tym wypadku kluczowy, o ile tylko system spełnia ogólne warunki opisane powyżej. Ponieważ w szczególności zarówno systemy klasyczne, o ile tylko wyposażone są w odpowiednie detektory lub korzystają z już zakodowanych informacji (o niepustej ekstensji), jak i systemy koneksjonistyczne czy też rozwiązania hybrydowe mogą przetwarzać informacje we wspomnianym sensie, uzasadnione jest powiedzenie, że mamy do czynienia z wieloraką realizacją własności odnoszenia się. W różny bowiem sposób można i przetwarzać informacje, i je reprezentować (kodować w systemie).

Systemy samodzielnie wydobywające informacje zawarte w sytuacjach i je przetwarzające można uznać za autonomiczne pod względem semantycznym, w odróżnieniu od większości systemów obliczeniowych, które są nieautonomiczne i do których dane wprowadzane są przez ludzi. W takim wypadku własności semantyczne tych informacji nie zależą od systemu (są wtórne wobec istniejących uprzednio własności semantycznych).

Powyższe ujęcie uwzględnia możliwość błędnej reprezentacji: błędna reprezentacja może pojawić się na dowolnym etapie działania mechanizmu i polega po prostu na dysfunkcji odpowiednich mechanizmów cząstkowych. Ponieważ większość systemów przetwarzających informacje korzysta z szeroko rozumianych procedur heurystycznych (Wimsatt 2007), istnieją sytuacje, w których będą ujawniać się systematyczne błędy i dysfunkcje w przetwarzaniu informacji z sytuacji.

3. Konotacje reprezentacji obliczeniowych

Zdecydowanie bardziej kontrowersyjna jest kwestia istnienia konotacji dla reprezentacji obliczeniowych. Z jednej strony można w ogóle kwestionować istnienie konotacji — zwłaszcza w sensie zbioru wystarczających i koniecznych warunków stosowania danej reprezentacji. Jest to typowe zwłaszcza dla eksternalizmu (por. np. McClamrock, rozdział 8). Z drugiej strony jest ewidentne, że Searle przede wszystkim konotację czy też treść reprezentacji ma na myśli, kiedy neguje istnienie zdolności semantycznych w systemach obliczeniowych.

Dwa różne zagadnienia dotyczą konotacji. Pierwsze dotyczy tego, w jaki sposób w ogóle może pojawić się konotacja reprezentacji — jeśli

system nie ma żadnego zewnętrznego „wspomagania” (np. w postaci kultury i języka naturalnego). Innymi słowy, jest to pytanie o redukowalność treści reprezentacji do innych niesemantycznych własności¹¹.

Drugie pytanie jest następujące: czy systemy obliczeniowe mogą mieć zdolności rozpoznawania desygnatów swoich reprezentacji? Treść reprezentacji ma bowiem umożliwić im prawidłową kategoryzację, a następnie zastosowanie reprezentacji określonego rodzaju. Na to drugie pytanie najczęściej odpowiada się, że konieczne będą własności sensomotoryczne, typowe raczej dla robotów, nie zaś dla systemów obliczeniowych.

Zajmę się przede wszystkim drugim zagadnieniem, gdyż pierwsze wymagałoby dokładniejszego omówienia problemu ugruntowania symboli, na co nie mam tutaj po prostu miejsca. Zresztą znane nam biologiczne systemy korzystające z reprezentacji wyposażonych w treść nie generują tej treści wyłącznie same z siebie, tylko w ramach społecznego podziału pracy językowej. Dlatego też redukcja treści semantycznych do własności niesemantycznych, choćby była ciekawym teoretycznie problemem, pozostaje kwestią stosunkowo akademicką, gdy mowa o systemach biologicznych czy wzorowanych na biologicznych.

Postulowanie zdolności sensomotorycznych jako warunku koniecznego rozpoznawania desygnatów dlatego wydaje się tak oczywiste, że kategoryzacja otoczenia wymaga aktywnej eksploracji (zdolności motoryczne) i detekcji obiektów (zdolności sensoryczne). Jednak zakłada się wówczas *de facto* empirystyczną koncepcję treści umysłowej: pochodzą one wszystkie z wrażeń zmysłowych, bezpośrednio lub przez kombinację. Nie odbiega to więc daleko od Hume'owskiej koncepcji idei, które albo wywodzą się bezpośrednio z doświadczenia, albo są konstruowane z innych wrażeń zmysłowych. Nie trzeba długo szukać, aby przekonać się, że nie wszystkie treści reprezentacji w ogóle mają jakikolwiek aspekt sensoryczny czy też motoryczny. Wystarczy wskazać na reprezentacje matematyczne czy normatywne. Trudno twierdzić, że treść pojęcia „dobro” da się wyrazić w skończonym zbiorze koniecznych i wystarczających cech sensorycznych i motorycznych. Choć odnosimy pojęcie dobra do różnych empirycznie dostępnych przedmiotów, to nie jest ono po prostu redukowalne do poziomu detekcji sensoryczno-motorycznej. Być może skomplikowany robot byłby

¹¹Przegląd odpowiedzi na to pytanie w odniesieniu do tzw. problemu ugruntowania symboli przedstawiają Floridi i Taddeo 2005. Ich własne rozwiązanie tego problemu stanowi pozytywną odpowiedź na pytanie o możliwość redukcji własności semantycznych do niesemantycznych, por. Floridi i Taddeo 2007.

w stanie znaleźć statystyczny wzorec we wszystkich dobrych jabłkach, samochodach i uczynkach, ale naprawdę jałowe wydaje się poszukiwanie sensorycznych aspektów treści pojęcia liczby 2 czy zbioru. Treść tych pojęć bowiem nie sprowadza się całkowicie do aspektów sensomotorycznych. Aby móc utrzymać tezę, że treść reprezentacji to tyle, co warunki rozpoznawania ich desygnatów sformułowane w kategoriach sensomotorycznych, trzeba poczynić założenia klasycznego empiryzmu genetycznego. Łącznie z rozróżnieniem pojęć analitycznych (np. matematycznych) i syntetycznych, co po krytyce Quine'a jest dosyć trudne¹².

W wypadku systemów, które posługują się wyłącznie reprezentacjami typu sensorycznego, takie empirystyczne rozwiązanie będzie zapewne wystarczające (jako pewna heurystyka). Jednak system obliczeniowy korzystający z pojęć matematycznych lub wspomagający ferowanie sądów normatywnych (np. wspomagający decyzje bioetyczne) nie zmieści się na prokrustowym łożu semantyki empiryzmu logicznego.

Zresztą nawet przedstawiciele empiryzmu logicznego nie przyjmowali tak skrajnych założeń. Wystarczy odwołać się do klasycznego artykułu Carnapa (1955), który pisał o „pojęciu intensji dla robota”. Otóż zdaniem Carnapa, intensję pojęcia można wyrazić w kategoriach stanów organów sensorycznych lub opisu podawanego na wejścia językowe danego robota. Odpowiednie pobudzenia organów lub analiza opisu mają prowadzić do tego, że robot, napotkawszy określony predykat, określa, czy predykat oznacza obiekt wykryty przez organ sensoryczny lub opisany językowo — wydaje wówczas werdykt pozytywny, negatywny lub wstrzymuje się od głosu. Carnap, w przeciwieństwie do wielu innych autorów, nie przesądza, jaką strukturę ma sama treść konotacyjna — nie musi być ona ujęta np. w kategoriach cech konstytutywnych i konsekwentnych. Mówiąc krótko, Carnap dopuszcza też możliwość określania intensji predykatu wyłącznie w kategoriach opisu językowego, bez jakichkolwiek wejść sensorycznych. Zdaje się jednak, że sądzi, że brak wejść sensorycznych świadczyć może o pustości reprezentacji (np. w wypadku pojęcia „jednorożec”).

We współczesnej filozofii języka w rozmaity sposób próbuje się przewycięzać koncepcje empiryzmu. Są podejścia holistyczne, gdzie treść jest definiowana przez sieć relacji między reprezentacjami (konotacja jest rozumiana np. jako tzw. rola pojęciowa). Są też próby eliminacji

¹²Pojęcia normatywne można by próbować sprowadzać, jak emotywizm, do emocji, pozbawiając je treści. Wydaje mi się oczywiste, że to pomysł chybiony: emotywizm nie potrafi po prostu zdać sprawy z sensu terminów etycznych.

tego pojęcia i zastąpienia go np. użyciem wyrażen. Aby nie rozstrzygać na rzecz któregoś z tych podejść, będę proponował inną strategię. Wróć do systemów przetwarzających informacje wydobyte z sytuacji i przeprowadzę eksperyment myślowy podobny do proponowanego przez Carnapa.

Pojęcie konotacji lub intensji historycznie wprowadzono po to, aby móc pokazać, że równozakresowe nazwy nie mogą być podstawiane w rozumowaniach *salva veritate*. Załóżmy więc, że system generuje dwie złożone reprezentacje — struktury danych — i używa ich w jakiejś szerszej reprezentacji opisowej (np. w zdaniu oznajmującym lub jako część diagramu; nie rozstrzygam o postaci reprezentacji). Problemem, który system musi być w stanie rozwiązać, np. na drodze rozumowania, jest określenie, że te dwie reprezentacje, choć mają różną treść, są jednak równozakresowe. Treść w tym wypadku będzie rozumiana jako zespół warunków (w skrajnym, ale rzadkim wypadku: koniecznych i wystarczających), które spełnia kategoria desygnatów reprezentacji. Te warunki służą do rozpoznawania desygnatów, ale niekoniecznie na drodze detekcji sensomotorycznej. Jeśli desygnat nie jest przedmiotem podlegającym detekcji sensorycznej, bo jest np. wewnętrznym stanem obliczeniowym (liczba dwa) lub abstraktem (zbiór), system będzie traktował reprezentację nie jako wydobytą z otoczenia fizycznego — gdyż nie jest w stanie odnieść jej do wejść systemu — lecz jako element sytuacji wewnętrznej systemu, o ile system posługuje się takimi konstrukcjami, jak zbiór czy liczby. Jeśli zaś desygnat nie istnieje ani jako sytuacja w otoczeniu systemu, ani jako jego sytuacja wewnętrzna, będzie traktowany formalnie, jako reprezentacja o nieznannej ekstensji. Operacje obliczeniowe, przekształcające formalną strukturę warunków mogą doprowadzić do wskazania, czy to na drodze faktycznego próbkowania sytuacji, czy też przez zastosowanie bazy wiedzy, że intensje, choć różne, wyznaczają ten sam zakres.

Odnieśmy to do systemów klasycznych i koneksjonistycznych. Jak zauważył Chalmers (1992), klasyczne systemy sztucznej inteligencji operują raczej na atomowych reprezentacjach bez żadnej struktury wewnętrznej. Te reprezentacje są powiązane różnymi relacjami, lecz to nie wystarcza, jak twierdzi Chalmers, aby można było mówić o intensjonalności tych reprezentacji. To pojedyncza bowiem reprezentacja musi mieć na tyle złożoną strukturę, aby można było mówić, że zawiera także cechy konotacyjne. Zdaniem Chalmersa, taką strukturę — strukturę subsymboliczną — mają systemy koneksjonistyczne, gdzie reprezentacje mają postać węzłów sieci neuronowych, w których istnieją warstwy ukryte. To jedyna nadzieja, jego zdaniem, na prawdziwą realizację konotacji w systemach obliczeniowych, lecz — co podkreśla —

na razie trudno powiedzieć, jak by to mogło w szczegółach wyglądać. Ujęcie Chalmersa obarczone jest grzechem, o którym wspomniałem na początku artykułu: autor ten sądzi, że intensja reprezentacji musi być cechą wewnętrzną samej struktury danych, która w określonym systemie obliczeniowym realizuje taką reprezentację. Wydaje się, że nieistotna jest sama struktura wewnętrzna samej reprezentacji. Idzie o to, czy jest ona stosowana w systemie, w którym może mieć treść w sensie posiadania warunków wskazywania desygnatów. Na przykład pojedynczy zapalony bit w systemie komputerowym może reprezentować liczbę naturalną 1, a to na mocy struktur relacyjnych w procesach, które będą operować na tym bicie. Z kolei skomplikowany ciąg liter wytworzony przez 300-krotne powtarzanie struktury „abcabc...” będzie pozbawiony intensji w systemie rozpoznającym zdania napisane w języku polskim. Innymi słowy, tam, gdzie reprezentacje mogą służyć do kategoryzacji, tam mają też treść.

Co ciekawe, również systemy korzystające wyłącznie z informacji już przetworzonych mogą mieć tego rodzaju reprezentacje, choćby te przetworzone informacje same nie odnosiły się do żadnych desygnatów. Wydaje się, że to wynik dosyć intuicyjny (wbrew sugestiom Searle’a i Harnada): puste pojęcia mogą mieć rozbudowane intensje. Choćby wujowie centaurów byli zakresowo identyczni ze stopami centaurów, to jednak treściowo się różnią. Takie różnice w treści są w stanie wychwycić nawet proste mechanizmy obliczeniowe, korzystające ze skomputeryzowanych sieci semantycznych. Nie muszą być robotami: detekcja sensomotoryczna dotyczy tylko niektórych reprezentacji, których warunki odnoszenia są do wyrażenia w takich kategoriach. Przypisanie istnienia treści reprezentacji podlega przy tym analogicznym warunkom, jak w wypadku przetwarzania informacji: musi być potrzebne z punktu widzenia predykcji, wyjaśniania i opisu itd.

4. Podsumowanie

Twierdzę, że systemy obliczeniowe mogą zawierać reprezentacje odnoszące się do przedmiotów i że mogą dysponować strukturami danych wskazującymi warunki prawidłowego odnoszenia się do przedmiotów (choćby owe przedmioty były czysto intencjonalne). Tylko w wypadku odniesienia konieczne jest powiązanie mechanizmów obliczeniowych — w ramach systemów przetwarzania informacji — ze światem zewnętrznym (lub wewnętrznym, jeśli system przetwarza informacje na swój własny temat). Intensje mogą istnieć bardziej niezależnie od niego, choć mogą być też określone w kategoriach wymagających przyczynowego kontaktu ze światem. Na przykład intensja nazwy własnej musi

określać indywiduum, którego dotyczyła określona procedura nazywania w przeszłości: określenie to jednak polega po prostu na prześledzeniu historii indywiduum od momentu nazwania, a nie znalezieniu go choćby przez najbardziej jednoznaczny opis.

W ramach tego ogólnego ujęcia widać, że pod tym przynajmniej względem nie ma żadnej interesującej różnicy między systemami typu klasycznego (GOFAI), koneksjonistycznego, hybrydowego czy np. populacjami algorytmów ewolucyjnych. Wszystkie te architektury mogą mieć zdolności semantyczne, o ile tylko będą przetwarzać informacje z sytuacji oraz operować na warunkach odnoszenia się reprezentacji. Własności semantyczne, mówiąc krótko, są wielorako realizowalne.

Literatura

Bechtel, W., Richardson, R. C. (1993). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Princeton: Princeton University Press.

Carnap, R. (1955). Meaning and Synonymy in Natural Languages. *Philosophical Studies*, VI, 3, s. 33–47.

Chaitin, G. J. (1975). Randomness and Mathematical Proof. *Scientific American*, 232, 5, s. 47–52.

Chaitin, G. J. (1987). *Algorithmic Information Theory*. Cambridge: Cambridge University Press.

Chalmers, D. (1992). Subsymbolic computation and the Chinese Room. W: Dinsmore, J. (red.) *The Symbolic and Connectionist Paradigms: Closing the Gap*. Lawrence Erlbaum.

Collier, J. (1990). Intrinsic Information. W: Hanson, P. (red.) *Information, Language, and Cognition*. University of British Columbia Press, s. 390–409.

Craver, C. F. (2006). When Mechanistic Models Explain. *Synthese* 153, 3, s. 355–376.

Feynman, R. (1982). Simulating Physics with Computers. *International Journal of Theoretical Physics* 11, 21.

Floridi, L., Taddeo, M. (2005). The Symbol Grounding Problem: a Critical Review of Fifteen Years of Research. *Journal of Experimental and Theoretical Artificial Intelligence* 17, 4, s. 419–445.

Floridi, L., Taddeo, M. (2007). A Praxical Solution of the Symbol Grounding Problem. *Minds and Machines* 17, 4, s. 369–389.

Harnad, S. (1990). The Symbol Grounding Problem. *Physica D* 42, s. 335–346.

Israel, D. J., Perry J. (1990). What Is Information? W: P. Hanson (red.) *Information, Language, and Cognition*. University of British Columbia Press, s. 1–19.

McClamrock, R. (1995). *Existential Cognition. Computational Minds in the World*. Chicago, London: University of Chicago Press.

Milkowski, M. (2007). Is computationalism trivial? W: Dodig Crnkovic,

G., Stuart S. (red.) *Computation, Information, Cognition — The Nexus and the Liminal*. Cambridge Scholars Publishing, s. 236–246.

Miłkowski, M. (2008). Postulowanie utajonych funkcji umysłu: realizm kontra antyrealizm. W: Wróbel, Sz. (red.) *Utajone funkcje umysłu*. Poznań-Kalisz: Wydział Pedagogiczno-Artystyczny UAM w Kaliszu, s. 15–37.

Miłkowski, M. (2009). O tzw. metaforze komputerowej. *Analiza i Egzystencja* 9, s. 163–185.

Piccinini, G. (2007). Computational Explanation and Mechanistic Explanation of Mind. W: de Caro, M., Ferretti, F., Marraffa, M. (red.) *Cartographies of the Mind: The Interface between Philosophy and Cognitive Science*. Dordrecht: Springer, s. 23–36.

Poczobut, R. (2005). Od informacji fizycznej do informacji fenomenalnej. W: Heller, M., Mączka, J. (red.) *Informacja a rozumienie*. Kraków: PAU-OBI, s. 177–193.

Searle, J. R. (1995). Umysły, mózgi i programy, przeł. B. Chwedeńczuk. W: Chwedeńczuk, B. (red.) *Fragmenty filozofii analitycznej, tom II. Filozofia umysłu*. Warszawa: Fundacja Aletheia, s. 301–324.

Searle, J. R. (1999). *Umysł na nowo odkryty*, przeł. T. Baszniak. Warszawa: PIW.

Wimsatt, W. C. (2007). *Re-Engineering Philosophy for Limited Beings*. Cambridge, Mass.: Harvard University Press.