

*Marcin Miłkowski*

## Mechanizmy reprezentacyjne i abstrakcje\*

Pojęcie reprezentacji uważa się za jeden z podstawowych terminów teoretycznych w kognitywistyce. Pozostaje jednak kontrowersyjne. W psychologii często wystrzegano się postulowania reprezentacji: wymienić wystarczy radykalne odmiany behawioryzmu<sup>1</sup> czy psychologii ekologicznej<sup>2</sup>. Niedawno pojawiły się nowe radykalne głosy krytyczne. Zwolennicy układów dynamicznych i ucieleśnienia sądzą, że pojęcie reprezentacji można całkowicie wyeliminować z nauk o zachowaniu i poznaniu<sup>3</sup>. Przedstawiają argumenty na rzecz tezy, że pojęcie to w neuronaukach i kognitywistyce bywa tylko listkiem figowym; na przykład proste detektory cech i mechanizmy śledzące nie gwarantują istnienia autentycznych reprezentacji<sup>4</sup>. Przy

\* Praca nad niniejszym artykułem była finansowana przez Narodowe Centrum Nauki w ramach programu OPUS, nr grantu 2011/03/B/HS1/04563.

<sup>1</sup> W mniej radykalnych jego odłamach postulowano różnego rodzaju stany odpowiadające w pewnej mierze kognitywistycznemu pojęciu reprezentacji. Por. I. Kustrzeba, *Znak i znaczenie w psychologii behawiorystycznej*, Państwowe Wydawnictwo Naukowe, Warszawa 1971.

<sup>2</sup> Z antyreprezentacjonizmem w psychologii ekologicznej łączy się zazwyczaj Gibsona (np. J.J. Gibson, *The Ecological Approach to Visual Perception*, Psychology Press, Hove 1986), ale wielu innych przedstawicieli tego nurtu postulowało reprezentacje (U. Neisser, *Cognition and reality. Principles and implications of cognitive psychology*, W.H. Freeman, San Francisco 1976). W pracach Gibsona można też poszukiwać podstaw dosyć oryginalnej teorii reprezentacji, jak argumentują Bickhard i Richie (M.H. Bickhard, D.M. Richie, *On the nature of representation. A case study of James Gibson's theory of perception*, Praeger, New York 1983).

<sup>3</sup> A. Chemero, *Radical embodied cognitive science*, MIT Press, Cambridge, Mass. 2009; F.C. Garzon, *Towards a General Theory of Antirepresentationalism*, „The British Journal for the Philosophy of Science”, 2008, t. LIX, nr 3, s. 259-292; D.D. Hutto, *Cognition without Representation?*, w: *Understanding representation in the cognitive sciences*, red. A. Riegler, M. Peschl i A. von Stein, Kluwer Academic/Plenum Publishers, New York 1999, s. 57-74; D.D. Hutto, E. Myin, *Radicalizing Enactivism. Basic Minds without Content*, MIT Press, Cambridge Mass. 2013; F. Keijzer, *Representation and Behavior*, MIT Press, Cambridge, Mass. 2001.

<sup>4</sup> W.M. Ramsey, *Representation Reconsidered*, Cambridge University Press, Cambridge 2007.

tym zwolenników reprezentacjonizmu oskarża się o nadmierną liberalność, gdyż przypisują oni reprezentacyjny charakter takim systemom, jak regulator Watta: reprezentacji w takim nadmiernie liberalnym sensie jest po prostu za dużo, aby to pojęcie się nie zbanalizowało<sup>5</sup>. William Ramsey podkreśla, że poprawna teoria reprezentacji powinna przede wszystkim odpowiedzieć na pytanie, jaką *reprezentacyjną* rolę odgrywa postulowana reprezentacja w systemie poznawczym. Niestety większość istniejących koncepcji semantyki znaturalizowanej na to pytanie nie odpowiada.

W tym artykule naszkicuję ogólne mechanicystyczne ramy pojęciowe służące do zawężenia pojęcia reprezentacji bez jednoczesnego rozstrzygnięcia, które cechy mechanizmów reprezentacyjnych występują w realnych systemach poznawczych. Mówiąc nieco inaczej, zamiast definicyjnie rozstrzygać kwestie empiryczne, pozostawię szereg pytań otwartych, zaznaczając, że konkretne odpowiedzi mogą być podane dopiero po przeprowadzeniu badań, nie zaś na drodze analizy pojęciowej. Podaję tu jedynie ramy pojęciowe czy też *schemat* mechanizmu reprezentacyjnego, którego szczegóły poznać można tylko mając odpowiednie świadectwa empiryczne. Schemat ten można wypełniać w różny sposób – zgodnie z wieloma istniejącymi wyjaśnieniami procesów poznawczych i nowymi obserwacjami<sup>6</sup>.

Filozoficzne ujęcia reprezentacji często nie unikają rozstrzygnięcia kwestii związanych z badaniami empirycznymi, przez co można z nich wyprowadzić pewne testowalne konsekwencje. I tak z teorii reprezentacji Jerry'ego Fodora, postulującego istnienie języka myśli opartego na amodalnych symbolach (czyli symbolach niereprezentowanych w sposób związany z określoną jakością zmysłową)<sup>7</sup>, Zenon Pylyshyn wyprowadzał wniosek, że nie mogą istnieć obrazy umysłowe, czyli reprezentacje analogiczne (choć, trzeba przyznać, niebagatelną rolę w jego argumentacji odgrywała również drobiazgową, krytyczną analizę teorii obrazów umysłowych)<sup>8</sup>.

<sup>5</sup> T. Van Gelder, *What might cognition be, if not computation?*, „The Journal of Philosophy” 1995, t. XCII, nr 7, s. 345-381; W. Bechtel, *Representations and Cognitive Explanations. Assessing Dynamicalist's Challenge in Cognitive Science*, „Cognitive Science” 1998, t. XXII, nr 3, s. 295-318; N. Stepp, A. Chemero, M.T. Turvey, *Philosophy for the Rest of Cognitive Science*, „Topics in Cognitive Science” 2011, t. III, nr 2, s. 425-437; K.S. Nielsen, *Representation and dynamics*, „Philosophical Psychology” 2010, t. XXIII, nr 6, s. 759-773.

<sup>6</sup> Proste przykłady z badań analizowałem w: M. Miłkowski, *Explaining the Computational Mind*, MIT Press, Cambridge, Mass. 2013.

<sup>7</sup> J.A. Fodor, *The Language of Thought*, wyd. I, Thomas Y. Crowell Company, New York 1975.

<sup>8</sup> Z.W. Pylyshyn, *Spór o wyobraźnię: medium analogowe czy wiedza ukryta?*, w: *Psychologia poznawcza w trzech ostatnich dekadach XX wieku*, red. Z. Chlewiński, przeł. J. Suchecki, Gdańskie Wydawnictwo Psychologiczne, Gdańsk 2007, s. 366-408. Upraszczam oczywiście sprawę, gdyż Pylyshyn przedstawiał wiele innych argumentów, ale samą propozycję istnienia reprezentacji innych niż amodalne symbole uznawał za groteskową.

Czy jednak spór o wyobrażenia umysłowe można rozstrzygnąć definicją, wykluczając pojęciowo możliwość istnienia reprezentacji analogicznych? Wówczas zależy on wyłącznie od wyboru aparatu teoretycznego, przez co staje się niezależny od faktów. A przecież kwestia ta co do zasady jest empiryczna, nawet jeśli w latach 70. nie było możliwości technicznych i eksperymentalnych, aby ją rozstrzygnąć. Rzecz jasna, zwolennik Fodora czy Pylyshyna może odrzec, że nie sposób systematycznie rozróżnić kwestii czysto empirycznych od zagadnień teoretycznych (odwołując się do znanych argumentów Quine'a o niemożności rozróżnienia zdań analitycznych od syntetycznych). Nawet jeśli Quine ma rację, to i on nie neguje, że jest różnica między sprawdzaniem zdania „ $2 + 2 = 4$ ” i zdania „Szympanś Waldek iskał się za uchem”. Prawdziwości drugiego zdania nie chcemy przesądzać na drodze definicyjnej.

Aby nie rozstrzygać definicjami pytań, na które co do zasady może odpowiedzieć eksperyment, koncepcja reprezentacji powinna pozostawać maksymalnie neutralna tam, gdzie wchodzi w grę empiria. Dlatego też będę pokazywać, że bronione tu ujęcie nie przesądza wcale (albo przesądza w stopniu stosunkowo minimalnym), która z teorii psychologicznych jest właściwa. Uczynię to w dalszej części artykułu na przykładzie współczesnych sporów o naturę reprezentacji abstraktów.

Nie przedstawiam tu też analizy pojęcia reprezentacji, odpowiadającej całkowicie jego faktycznemu użyciu. Jest tak dlatego, że jest ono stosowane w wielu bardzo różnych rozumieniach, co prowadzi do jego chwiejności i wieloznaczności, a co za tym idzie – do sporów czysto werbalnych. Chociaż teorię mechanizmów reprezentacyjnych można uznać za nieco może rozwlekłą definicję regulującą, to termin „reprezentacja” będę traktować jako techniczny, nie przywiązując zbyt dużej wagi do tego, czy odpowiada on typowym użyciom w języku potocznym czy też w literaturze fachowej. To po prostu osobna kwestia.

## 1. *Mechanicizm i reprezentacje*

Mechanicizm, którego narzędzia pojęciowe tutaj stosuję, jest obecnie najbardziej dyskutowaną koncepcją wyjaśniania w filozofii nauki, stosowaną zwłaszcza do analizy przypadków w biologii, neurologii i kognitywistyce<sup>9</sup>. Przedstawiciele tego nurtu bronią różnych definicji pojęcia mecha-

---

<sup>9</sup> P. Machamer, L. Darden, C.F. Craver, *Myslenie w kategoriach mechanizmów*, przeł. W.M. Hensel, „Przegląd Filozoficzno-Literacki” 2011, nr 2-3(31), s. 145-173; S.S. Glennan, *Rethinking Mechanistic Explanation*, „Philosophy of Science” 2002, t. LXIX, nr S3, s. S342-S353; W. Bechtel, *Mental Mechanisms*, Routledge, New York 2008; W. Bechtel, *Mechanism and Biological Explanation*, „Philosophy of Science” 2011, t. LXXVIII, nr 4, s. 533-557; W. Bechtel, A. Abrahamsen, *Explana-*

nizmu, lecz podstawowa idea pozostaje stosunkowo jasna i wyraźna: mechanizmy to złożone struktury funkcjonalne, zawierające zorganizowane składniki i procesy (lub działania) oddziałujące wzajemnie i prowadzące wspólnie do powstania funkcji danej struktury. Wyjaśnianie mechaniczne jest przyczynowe, podobnie jak oddziaływania składników mechanizmu. Warto odróżnić pełen (ze względu na przyczynowo istotne czynniki) opis mechanizmu od opisu niepełnego. Pełność opisu jest niezbędna, aby mógł on służyć jako całkowite wyjaśnienie przyczynowe mechanizmu. Carl Craver wyróżnia dwa rodzaje niepełnych opisów mechanizmów: *skizce*, w których pełno luk (zapełnianych terminami-»wytrychami«), oraz *schematy*, będące skróconymi reprezentacjami abstrahującymi od pewnych szczegółów organizacji. W schemacie szczegóły nie są wypełniane »zapchajdziurami«. Poniżej zaprezentuję właśnie schemat mechanizmu reprezentacyjnego: nie zawiera on ogólnikowych określeń w rodzaju „kodowanie”, „reprezentowanie”, „transformacja”, lecz stanowi raczej szablon do tworzenia dokładniejszych reprezentacji faktycznych mechanizmów<sup>10</sup>.

Mechanicyści podkreślają, że ich ramy pojęciowe szczególnie dobrze nadają się do analizy obiektów teoretycznych przyjmowanych w naukach szczegółowych. Użycie tych ram można uzasadniać, wskazując na fakt, że reprezentacja to obiekt teoretyczny w naukach o zachowaniu, mózgu i poznaniu, a do tych nauk do tej pory z powodzeniem stosowano analizy mechanicznych. Jednakże mechanicyzm jest szczególnie przydatny właśnie w omawianym przypadku. Po pierwsze, wyjaśnienia mechaniczne wymagają wskazania (co najmniej jednej) dyspozycji (*capacity*) mechanizmu, czyli *explanandum*, które jest wyjaśniane przyczynowo. Wpisuje się to w program naturalizowania reprezentacji, gdyż wymaga kompletnego przyczynowego wyjaśnienia tej dyspozycji – w tym wypadku dyspozycji do reprezentowania. Po drugie, w wyjaśnianiu mechanicznym podkreśla się istotność organizacji systemu, w którym postuluje się istnienie mechanizmu. Zamiast abstrahować od roli reprezentacji w organizmie, który z niej korzysta, mechanicyzm wymaga określenia tej roli. Reprezentacje nie mogą wystąpić sa-

---

*tion. A mechanist alternative*, „Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences” 2005, t. XXXVI, nr 2, s. 421-441; C.F. Craver, *Explaining the Brain. Mechanisms and the mosaic unity of neuroscience*, Oxford University Press, Oxford 2007; G. Piccinini, *Computing Mechanisms*, „Philosophy of Science” 2007, t. LXXIV, nr 4, s. 501-526; M. Miłkowski, *Explaining the Computational Mind...*, *op. cit.*

<sup>10</sup> Przedstawionemu niżej schematu mechanizmu reprezentacyjnego nie należy mylić z abstrakcyjną reprezentacją jego organizacji. Schemat ten bowiem nie może posłużyć do satysfakcjonującego wyjaśniania bez podania dodatkowych szczegółów. Mówiąc nieco inaczej, nie jest on abstrakcją w sensie Levy’ego i Bechtela (*Abstraction and the Organization of Mechanisms*, „Philosophy of Science” 2013, t. LXXX, nr 2, s. 241-261).

modzielnie, nie będąc elementem układu złożonego. Z tego też względu mechanistyczne ramy pojęciowe wymagają podania poprawnej odpowiedzi na pytanie Ramseya o rolę reprezentacji w poznaniu.

Choć mechanizmy reprezentacyjne mogą służyć do wyjaśniania działania większych mechanizmów (odgrywając rolę *explanansu*), skupię się na przypadku, w którym występują jako *explanandum*. Dyspozycję mechanizmu wyróżnia się w ścisłej zależności od tego, jak postawione jest pytanie o wyjaśnienie. Innymi słowy, przypisanie dyspozycji (funkcji) mechanizmowi zależy od zainteresowania teoretyka: dyspozycję do wyjaśnienia można pojmować różnie w zależności od tego, co ma być wyjaśnione. We wszystkich takich wyjaśnieniach wspólne będzie jednak to, że mechanizm reprezentacyjny ma dyspozycję do udostępniania informacji w systemie poznawczym<sup>11</sup>. W takim przypadku informacja staje się *semantyczna* w tym sensie, że zmienia gotowość systemu działania w taki a taki sposób (mówiąc nieco dokładniej, warunkowe prawdopodobieństwo działania zmienia się odpowiednio w zależności od zmiany informacji<sup>12</sup>). W niektórych ujęciach mechanizmów ich dyspozycje ujmują się w kategoriach funkcji, a w zależności od tego, jak silne jest to pojęcie funkcji, koncepcja mechanizmu reprezentacyjnego może mniej lub bardziej przypominać wymagania zwykle przyjmowane przez teleosemantyków, a mianowicie, że informacja ma funkcję (w niebanalnym sensie) bycia semantyczną<sup>13</sup>.

Dwie tezy wiążą się z ujęciem dyspozycji mechanizmu reprezentowania w kategoriach zmiany gotowości do działania na podstawie informacji dostępnych dla systemu. Po pierwsze, reprezentacje są przede wszystkim związane z działaniem, choć związek ten nie oznacza, że wszelkie reprezentacje służą po prostu do sterowania czynnościami motorycznymi systemu czy też aktywacji jego efektorów. Może istnieć treść reprezentacji niewykorzystywana w działaniu; zmienia się bowiem tylko *gotowość* do działania. Wydaje się, że inne koncepcje reprezentacji związanych z działaniem w podobny sposób starają się oddalić zarzut, że zbyt silnie powiązано reprezentację z motorycznością<sup>14</sup>. Pojęcie działania należy rozumieć liberalnie, aby obejmowało też poznanie. Po drugie, w bronionej

<sup>11</sup> M. Miłkowski, *Explaining the Computational Mind...*, *op. cit.*

<sup>12</sup> D.M. MacKay, *Information, mechanism and meaning*, M.I.T. Press, Cambridge 1969.

<sup>13</sup> R.G. Millikan, *Language, thought, and other biological categories. New foundations for realism*, The MIT Press, Cambridge, Mass. 1984; F.I. Dretske, *Błędna reprezentacja*, w tym numerze PF-L; R. Cao, *A teleosemantic approach to information in the brain*, „Biology & Philosophy” 2011, t. XXVII, nr 1, s. 49-71; C. Price, *Functions in mind: a theory of intentional content*, Clarendon Press, Oxford-New York 2001.

<sup>14</sup> Zarzut ten formuluje A. Clark, *Visual Experience and Motor Action. Are the Bonds Too Tight?*, „Philosophical Review” 2001, t. CX, nr 4, s. 495-519. Analogiczną do mojej odpowiedź dają m.in.

koncepcji istotną rolę odgrywa pojęcie informacji. Choć istnieje wiele matematycznych miar informacji, nie należy ich mylić z samym pojęciem informacji<sup>15</sup>. Minimalnym warunkiem istnienia informacji jest istnienie nośnika fizycznego, który ma co najmniej dwa stopnie swobody (lub dwa poziomy głąbi logicznej) relewantne dla zachowania systemu. Innymi słowy, system musi reagować inaczej na co najmniej dwa różne stany nośnika fizycznego. Wprowadzone tutaj nieformalnie pojęcie informacji jest równoważne tzw. informacji strukturalnej (zwanej przez MacKaya *logonem*); wprowadzając nadawcę i odbiorcę, kanał i niepewność odbiorcy, można też zastosować miary Shannona. Dla moich celów w tym artykule nie jest to potrzebne, mimo że te miary *można* stosować, o ile tylko podsystem otrzymuje informacje. Pominę tutaj jednak szczegóły techniczne.

## 2. Podstawowe zdolności mechanizmu reprezentacyjnego

Wstępnego określenia dyspozycji mechanizmu reprezentacyjnego nie należy mylić z pełną teorią reprezentacji. Mechanizm taki ma inne istotne dyspozycje wskazywane w wyjaśnieniach odwołujących się do pojęcia reprezentacji. Zazwyczaj wprowadza się je w celu mówienia o desygnatach (lub referentach, a także zakresie, ekstensji lub denotacji), do których odnosi się reprezentacja, a także po to, aby mówić o charakterystyce desygnatów (czyli intensji bądź treści). Co więcej, informacje nie mogą po prostu »tkwić« w mechanizmie; systemowi musi w jakiś sposób na nich zależeć. Istnieją więc trzy inne dyspozycje tego mechanizmu:

(a) odnoszenie się do przedmiotu reprezentacji (jeśli przedmiot istnieje);

(b) wskazywanie charakterystyki przedmiotu;

(c) ocena wartości epistemicznej informacji o przedmiocie.

Dwie pierwsze dyspozycje – (a) i (b) – przypominają tradycyjne pojęcia zakresu i treści nazwy, stosowane w takiej lub innej postaci w teoriach znaczenia przynajmniej od czasów Johna Stuarta Milla<sup>16</sup>, dyspozycja (c)

---

M.L. Anderson i G. Rosenberg, *Content and action: The guidance theory of representation*, „Journal of Mind and Behavior” 2008, t. XXIX, nr 1-2, s. 55-86.

<sup>15</sup> D.M. Mackay, *Information, mechanism and meaning*, *op. cit.*

<sup>16</sup> G. Frege, *Pisma semantyczne*, przeł. B. Wolniewicz, Państwowe Wydawnictwo Naukowe, Warszawa 1977; R. Carnap, *Meaning and necessity*, University of Chicago Press, Chicago 1947; J. Mill, *System logiki dedukcyjnej i indukcyjnej*, przeł. C. Znamierowski, Państwowe Wydawnictwo Naukowe, Warszawa 1962; R. Cummins, M. Roth, *Meaning and Content in Cognitive Science*, w: *Prospects for Meaning*, red. R. Schantz, de Gruyter, Berlin & New York 2012; R.G. Millikan, *Language, thought, and other biological categories. New foundations for realism...*, *op. cit.*; R. Cummins, *Representations, targets, and attitudes*, MIT Press, Cambridge, Mass. 1996. W polskiej literaturze tego rodzaju roz-

wiąże mechanizm reprezentacyjny z działaniem podmiotu lub systemu, który z tego mechanizmu korzysta. Warto w tym miejscu zaznaczyć, że choć tradycyjna teza, że treść w pełni wyznacza zakres pojęcia, została słusznie skrytykowana przez Hilary'ego Putnama<sup>17</sup>, nie znaczy to, że zakres nie może zostać wyznaczony przez treść przynajmniej częściowo (o ile istnieją desygnaty, czyli elementy denotacji).

Czy nie można byłoby pozbyć się wyznaczania charakterystyki przedmiotu z listy dyspozycji mechanizmu reprezentacyjnego? Czy nie jest to mnożenie bytów ponad konieczność? Tak mogliby sądzić przedstawiciele przyczynowej teorii odniesienia, którzy starają się zrezygnować w ogóle z residuów koncepcji treści czy intensji. Weźmy przykład. Argumentowano, że pojedyncze stany neuronalne traktowane jako wskaźniki określonych zewnętrznych przedmiotów nie są pełnymi reprezentacjami, gdyż nie charakteryzują tych przedmiotów<sup>18</sup>. Jeśli nawet koncepcja stanów neuronalnych jako prostych wskaźników pozbawionych treści jest przekonująca – wszak znaki naturalne zwykle uznaje się za nieposiadające intensji – to rozproszone struktury nerwowe (reprezentacje populacyjne) nie mogą być rozumiane jako proste zbiorowisko pojedynczych wskaźników. Zresztą i w wypadku jednego wskaźnika można mieć wątpliwości<sup>19</sup>, gdyż nawet on niesie pewne informacje o przedmiocie, czyli pewna charakterystyka przedmiotu, choć rzeczywiście minimalna, jest dostępna.

Nie bronię tutaj też deskrypcyjnej teorii znaczenia, zgodnie z którą reprezentacja zawsze wymaga opisu przedmiotu; charakterystyka przedmiotu może być równie dobrze rozumiana w kategoriach np. »plików umysłowych« Récanati<sup>20</sup>. Koncepcja mechanizmów reprezentacyjnych jest ekumeniczna; sprawą empirycznych badań jest rozstrzygnięcie, czy dany system korzysta z nośnika o bogatej strukturze, czy też nie; nie można tego rozstrzygnąć w zaciszu gabinetu.

Za uwzględnieniem charakterystyki przedmiotu jest jeszcze jeden argument. Jak wskazywał Daniel Dennett, nieprzezroczyłość odniesienia

---

różnienia bronił przede wszystkim Kazimierz Twardowski, por. K. Twardowski, *O treści i przedmiocie przedstawień*, w: *Wybrane pisma filozoficzne*, przeł. I. Dąbbska, Państwowe Wydawnictwo Naukowe, Warszawa 1965, s. 3-91.

<sup>17</sup> H. Putnam, *Znaczenie wyrazu „znaczenie”*, w: *Wiele twarzy realizmu i inne eseje*, przeł. A. Grobler, PWN, Warszawa 1998, s. 93-184.

<sup>18</sup> R. Cummins, *Representations, targets, and attitudes...*, *op. cit.*; W.M. Ramsey, *Representation Reconsidered...*, *op. cit.*

<sup>19</sup> Por. A. Morgan, *Representations gone mental*, „Synthese” 2013, t. CXCI, nr 2, s. 213-244. Morgan dokładniej omawia zarzuty Cumminsa i Ramseya, starając się je zbić. Tu zostawiam sprawę na marginesie.

<sup>20</sup> F. Récanati, *Mental Files*, Oxford University Press, Oxford 2012.

(czy też intensjonalność) stanowi jeden z probierzy eksplanacyjnej przydatności reprezentacji<sup>21</sup>. Taką nieprzezroczystość łatwo wyjaśnić, wskazując, że bierze się z zastosowania różnych charakterystyk przedmiotu reprezentacji. Bez owej charakterystyki – choćby w minimalnej formie plików mentalnych w sensie Récanatiego czy etykiety stosowanej do rozróżniania przedmiotów – intensjonalność pozostaje dosyć tajemnicza.

Trzecia dyspozycja mechanizmu reprezentacyjnego, tj. ocena wartości epistemicznej informacji o przedmiocie, może wydawać się zupełnie nie na miejscu i zupełnie niezwiązana z reprezentowaniem. Niekiedy rzeczywiście warto abstrahować od wartości epistemicznych. Na przykład zwykle nie uwzględnia się tego rodzaju czynników, opracowując gramatykę formalną, a przynajmniej nie czyni się tak bezpośrednio (choć w niektórych teoriach gramatyk formalnych wykorzystywanych w psychologii istotną kwestią jest praktyczna obliczalność obliczeniowego modelu gramatyki). W wielu teoriach w naukach o zachowaniu i poznaniu proces reprezentowania jest jednak ujmowany w ramach, w których uwzględnia się wartość reprezentacji dla danego organizmu. I tak etologiczny opis zwierzęcia rozpoznającego określony zapach jako wskaźnik obecności drapieżnika odwołuje się przynajmniej *implicite* do adaptacyjnej wartości uniknięcia spotkania z owym drapieżnikiem.

Różne wartości mogą pozostawać ze sobą w pewnym napięciu, a wybór jednej może pociągać odrzucenie drugiej, przez co prawdziwość, dokładność czy pewność nie mogą być traktowane jako jedyne wartości epistemiczne. Dla bobra korzystniejsze może być błędne reprezentowanie pewnego dźwięku jako wskaźnika niebezpieczeństwa niż zignorowanie go<sup>22</sup>. Mówiąc inaczej, teorie dotyczące zachowania, mózgu i poznania zwykle pokazują, dlaczego w ogóle określone systemowi poznawczemu »zależy na« danej reprezentacji; wiąże się to z wartością adaptacyjną samego mechanizmu reprezentacyjnego (nie zaś pojedynczej reprezentacji przezeń generowanej). Wartość adaptacyjna nie wyznacza jednak jednoznacznie, jakie wartości epistemiczne – takie jak praktyczna obliczalność, ogólność, szczegółowość, dokładność, prawdziwość czy pewność – są najistotniejsze z punktu widzenia organizmu. Adaptacyjność może pozostawać w konflikcie z prawdziwością (jak w przykładzie z bobrem), a ogólność ze szczegółowością. Co więcej, zignorowanie pewnych informacji może być dla organizmu korzystne, gdyż są one, tak czy inaczej, nieistotne dla działań organizmu. Na przykład mapa miasta, nawet jeśli jest tworzona na podstawie zdjęć satelitarnych, nie przedstawia samochodów parkujących

<sup>21</sup> D.C. Dennett, *Content and Consciousness*, Routledge and Kegan Paul, London 1969.

<sup>22</sup> R.G. Millikan, *Biosemantics*, „Journal of Philosophy” 1989, t. LXXXVI, nr 6, s. 281-297.



przy ulicach w momencie wykonania zdjęć. Dla kartografa samochody są szumem zakłócającym właściwe informacje o topografii. Jednakże dla inżyniera ruchu lub projektanta parkingów zdjęcia samochodów są kluczowe, a budynki mogą być szumem. Mówiąc inaczej, na podstawie tych samych dostępnych informacji wejściowych można utworzyć różne reprezentacje, wykorzystując różne składniki wejścia. Z tego też względu maksymalna dokładność, polegająca na zachowaniu wszystkich szczegółów, praktycznie nigdy nie jest korzystna dla systemu poznawczego. Dokładność nie jest więc absolutną wartością epistemiczną.

Podałem co prawda pewne przykłady wartości epistemicznych, ale nie wszystkie z nich można efektywnie czy praktycznie oceniać (np. prawdziwość nie może być zawsze wyznaczona obliczeniowo ze względu na znane twierdzenia o niezupełności). Dla mózgu zwierzęcia zbyt kosztowne mogłoby być wyznaczanie złożoności obliczeniowej algorytmów przetwarzania informacji; szybszą i oszczędniejszą metodą dającą podobny pożądany efekt – a więc unikanie zbyt czasochłonnego przetwarzania – byłoby przetrwanie przetwarzania informacji po pewnym czasie (co może objawiać się w zachowaniu w postaci „znudzenia” itp.). Zamiast zaś oceniać prawdziwość czy weredyeczność jako taką, można wykrywać rozbieżności między dwiema informacjami.

Wszystkie wspomniane wyżej zdolności mechanizmu reprezentacyjnego są ściśle ze sobą powiązane. Informacje semantyczne zmieniające gotowość systemu do działania, które mogą być w postaci charakterystyki przedmiotu lub po prostu wskazywać przedmiot, podlegają ocenie epistemicznej. Jak mechanizm może w stabilny sposób przejawiać wszystkie te zdolności? Otóż ocena epistemiczna jest szczególnie istotna. Aby ocenić informacje już występujące w systemie, mechanizm musi być w stanie porównać dwa zbiory charakterystyk przedmiotu. Innymi słowy, ocena epistemiczna wymaga więcej niż tylko sprzężenia zwrotnego ujemnego w mechanizmie przetwarzania informacji: sprzężenie modyfikuje jedynie wartość wejściową systemu. Wymagane jest coś innego, a mianowicie wykrywanie błędu przez sam system. Idea, że błąd wykrywalny przez system poznawczy stanowi próbiez reprezentacyjności, nie jest bynajmniej nowa<sup>23</sup>.

Proste systemy z ujemnym sprzężeniem zwrotnym, takie jak często analizowany w dyskusji o reprezentacjach regulator Watta, *nie* są mechanizmami reprezentacyjnymi w znaczeniu wprowadzonym w tym tek-

---

<sup>23</sup> Najbardziej wyrazistym jej orędownikiem jest Mark Bickhard. Por. *idem*, *Representational content in humans and machines*, „Journal of Experimental & Theoretical Artificial Intelligence” 1993, t. V, nr 4, s. 285-333; *Levels of representationality*, „Journal of Experimental & Theoretical Artificial Intelligence” 1998, t. X, nr 2, s. 179-215. Zob. też jego artykuł *Model interaktywistyczny* w tym numerze PF-L.

ście. Przypomnijmy. Regulator Watta to mechaniczne urządzenie służące do stabilizacji prędkości maszyny parowej, a jego działanie wyjaśnia się zazwyczaj w kategoriach teorii sterowania; jest też używany jako metafora dynamicznych koncepcji poznania<sup>24</sup>. Regulator ten należy do klasy odśrodkowych regulatorów obrotów i stanowił ogromne osiągnięcie w czasach rewolucji przemysłowej. Zawiera on dwie ciężkie kule zamontowane na ramie napędzanej przez silnik, a siła odśrodkowa powoduje ich podnoszenie lub opadanie, co z kolei prowadzi do zamknięcia lub otwarcia zaworu dławiącego dopływ pary. Innymi słowy, zachodzi ujemne sprzężenie zwrotne między prędkością wykrywaną przez regulator odśrodkowy a silnikiem. Nawet jeśli analizować regulator Watta w kategoriach przetwarzania informacji<sup>25</sup>, sama struktura tego systemu nie zawiera żadnych mechanizmów przetwarzających błąd (mówiąc językiem teorii sterowania, informacja o „uchybie” nie jest przetwarzana). Zamiast *oceniać* i *modyfikować* wcześniejsze informacje *wyjściowe* („sygnał” otwarcia lub zamknięcia zaworu), po prostu zmienia działanie silnika przez modyfikację *wejścia* (podniesienie lub opuszczenie ciężarków). Z tego też względu regulator Watta nie wykrywa żadnych błędów w informacjach: chociaż z punktu widzenia obserwatora mógłby błędnie reprezentować stan sterowanego silnika parowego, błąd nie jest wykrywany przez regulator. Co więcej, w działaniu regulatora występuje nieusuwalny bez dodatkowej kompensacji uchyb ustalony (w nowszych układach regulujących uchyb ten jest eliminowany przez człon całkujący), przez co regulacja prędkości może być skuteczna w dosyć ograniczonym zakresie.

Wykrywanie błędów przez system może się różnie urzeczywistniać; na przykład wówczas, gdy system ma dwa niezależne źródła informacji<sup>26</sup> lub gdy wcześniejsze informacje wejściowe służą do przewidywania stanu przyszłego, a następnie są porównywane z owym stanem, kiedy jest on już aktualny. Tę drugą możliwość, choć pozornie zbliżoną do ujemnego sprzężenia zwrotnego, należy od niego odróżnić, gdyż przetwarzanie informacji następuje na wielu poziomach. Jest to typowe na przykład dla współczesnych koncepcji przetwarzania predykcyjnego<sup>27</sup>. Idea tego rodzaju przetwa-

---

<sup>24</sup> T. Van Gelder, *What might cognition be, if not computation?...*, *op. cit.*

<sup>25</sup> Czyni tak np. W. Bechtel w swoim artykule *Representations and Cognitive Explanations. Assessing Dynamicist's Challenge in Cognitive Science...*, *op. cit.* Dokładniejszą analizę zawiera artykuł: K.S. Nielsen, *Representation and dynamics...*, *op. cit.*

<sup>26</sup> F.I. Dretske, *Błędna reprezentacja*, *op. cit.*

<sup>27</sup> A. Clark, *Whatever next? Predictive brains, situated agents, and the future of cognitive science*, „The Behavioral and Brain Sciences” 2013, t. XXXVI, nr 3, s. 181-204; K. Friston i S. Kiebel, *Predictive Coding: A Free-Energy Formulation*, w: *Predictions in the Brain: Using Our Past to Generate a Future*, red. M. Bar, Oxford University Press, Oxford 2011, s. 231-246.

rzania pojawiła się u zarania cybernetyki<sup>28</sup> i można ją już dostrzec w biosemiotyce Uexkülla<sup>29</sup>, który opisywał ją na swoich diagramach. Tworzony jest model przyszłego stanu systemu, a następnie ten model jest porównywany z aktualnym stanem systemu w odpowiednim momencie; z technicznego punktu widzenia model staje się kolejnym źródłem informacji, choć zwykle zależnym statystycznie. Pełen opis mechanizmu reprezentacyjnego musi dokładnie opisywać strukturę podsystemu oceniającego, określając szczegółowo, w jaki sposób mechanizm wykrywa rozbieżność między dwoma informacjami wejściowymi lub między modelem a informacją wejściową.

Ponieważ mechanizmy reprezentacyjne muszą przetwarzać informacje (np. żeby je oceniać), wiele aspektów ich funkcjonowania można ująć w kategoriach obliczeniowych; dlatego też reprezentowanie wymaga obliczania – lecz nie sprowadza się do obliczania, gdyż pojęcia takie jak „działanie” po prostu nie są obliczeniowe<sup>30</sup>.

Podsumujmy. Koncepcja mechanizmów reprezentacyjnych określa ich zdolności, a choć są one wprowadzone w dużej mierze abstrakcyjnie, to pozwalają wskazać podstawowe cechy takich mechanizmów. Nie tylko przetwarzają i przechowują informacje modyfikujące ich gotowość do działania, lecz także są (przynajmniej czasami) w stanie odnosić się do przedmiotów, identyfikować je i (z konieczności) oceniać wartość przechowywanej informacji. Dzięki temu mechanizmy reprezentacyjne mogą wykryć, że się mylą (oceniając wartość epistemiczną), a przynajmniej w niektórych przypadkach mogą też błędnie identyfikować swoje przedmioty z powodu nieprzezroczystości odniesienia. Obie te własności, a więc wykrywanie błędów przez system (podkreślane przez Bickharda i Andersona) oraz nieprzezroczystość odniesienia, czyli intensjonalność (akcentowana przez Dennetta) postulowano już wcześniej jako podstawę istotności treści jako treści. Mówiąc krótko, na wyzwanie Ramseya jest odpowiedź: reprezentacja ma swoistą rolę w mechanizmie reprezentacyjnym.

### 3. Reprezentacje abstrakcyjne

Sz szczególnie interesującym przypadkiem do analizy mogą być reprezentacje abstrakcyjne. Jest tak z kilku powodów. Po pierwsze, moja koncepcja

---

<sup>28</sup> G. Boccignone, R. Cordeschi, *Predictive brains. Forethought and the levels of explanation*, „Frontiers in Psychology” (listopad) 2012, t. III, s. 2-3.

<sup>29</sup> J. von Uexküll, *Istota żywa jako podmiot: wybór pism Jakoba Johannesesa von Uexkülla*, red. A. Pobjewska, przeł. A. Pobjewska i M. Pólrola, Studio Wydawnicze KARTA, Łódź 1998.

<sup>30</sup> Dokładniejszą argumentację na rzecz tej tezy przedstawia M. Miłkowski, *Explaining the Computational Mind...*, *op. cit.*

wiąże reprezentowanie z działaniem, a związek abstraktów z działaniem może wydawać się co najmniej wątpliwy. Po drugie, naturalistyczne koncepcje reprezentacji często mają niemałe trudności z treścią reprezentacji abstrakcyjnych. Popularna teza empiryzmu genetycznego – że wszystkie reprezentacje pochodzą z doświadczenia zmysłowego – ma przynajmniej na pierwszy rzut oka trudność ze wskazaniem, jakie doświadczenie zmysłowe wiąże się z takimi pojęciami, jak PRAWDA czy SPRAWIEDLIWOŚĆ DYSTRYBUTYWNA<sup>31</sup>. Po trzecie wreszcie, w badaniach nad poznaniem – w psychologii poznawczej, neurosemantyce, lingwistyce poznawczej, psycholingwistyce itd. – toczy się żywa debata na temat natury reprezentacji abstrakcyjnych<sup>32</sup>. W debacie tej współcześnie często odrzuca się istnienie amodalnych reprezentacji symbolicznych, które uznawano za szczególnie dobrze nadające się do powiązania z abstraktami. Zaznaczyłem jednak na wstępie, że moje ujęcie będzie neutralne względem faktów empirycznych; warto będzie zatem zobaczyć, czy nie przesądzam z góry na rzecz jednego ze stanowisk w debacie.

W semiotyce zwykle przez „nazwy abstrakcyjne” rozumie się nazwy desygnujące abstrakty, a przez „nazwy konkretne” – desygnujące konkrety, a rozróżnienie abstraktów i konkretnych jest uznawane za dychotomię. W kognitywistyce terminy „abstrakcyjny” i „konkretny” stosowane są zwykle jako stopniowalne, gdzie „abstrakcyjność” rozumie się jako stopień oderwania od sfery fizycznej (termin ten rzadko bywa wprowadzany ściślej; ponieważ stosuję go w znaczeniu stosowanym w psychologii i kognitywistyce, nie będę go definiował w sposób regulujący, lecz pozostanę przy tej dosyć niewyraźnej definicji). Warto zauważyć, że rozróżnienie to nie pokrywa się z rozróżnieniem między reprezentacjami analogicznymi i symbolicznymi (wprowadzam je w następnej części artykułu) ani też z rozróżnieniem między reprezentacjami postrzeżeniowymi (modalnymi) i niepostrzeżeniowymi (amodalnymi)<sup>33</sup>. Reprezentacja analogiczna, taka jak obraz, może być bardziej lub mniej abstrakcyjna; szkicowa karykatura różni się od kolorowego zdjęcia właśnie stopniem abstrakcyjności.

---

<sup>31</sup> W tym artykule stosuję powszechną w kognitywistyce konwencję, aby wyróżniać umysłowe pojęcia nazwami w wersalikach (np. PRAWDA), w odróżnieniu od ich nazw językowych (np. „prawda” w języku polskim).

<sup>32</sup> Debata ta ma zresztą dosyć głębokie korzenie w historii psychologii, w której od dawna łączy się badania nad specyfiką ludzkiego działania z analizą charakteru reprezentacji abstrakcyjnych i konkretnych. Por. M. Maruszewski, *Uwagi o badaniach psychologicznych nad specyfiką ludzkiego działania*, w: *Z problematyki psychologii i teorii poznania*, red. C. Nowiński, Państwowe Wydawnictwo Naukowe, Warszawa 1958, s. 101-197.

<sup>33</sup> A. Paivio, *Mental representations. A dual coding approach*, Clarendon Press, New York, Oxford 1986, s. 17.

Warto jednocześnie podkreślić, że w badaniach psychologicznych, kognitywistycznych czy neurosemantycznych nie ma zgody nawet co do tego, czy reprezentacje abstrakcyjne istnieją, nie mówiąc już o tym, jaka jest ich natura. Nie sposób nawet podać ich dokładniejszej charakterystyki, nie przesądzając na mocy definicji, że określona empiryczna teoria reprezentacji abstrakcyjnych jest fałszywa. To jednak kłóciłoby się z metodologiczną zasadą, którą wprowadziłem w części pierwszej tego artykułu. Niestety analiza mechanicystyczna musi cechować się pewnego rodzaju ascezą, jeśli mamy stosować ją jednakowo do różnych teorii. Musimy mieć wszelako pewność, czy te teorie są teoriami przynajmniej z grubsza tego samego zjawiska, a nie zajmują się po prostu różnymi pytaniami, stosując jedynie podobne słowa do ich stawiania. Mam wrażenie, że teorie, przynajmniej wskazane w następnej części artykułu, zajmują się w pewnej mierze tym samym. Zamiast wprowadzać definicję konotacyjną tego czegoś, po prostu wskażę na przykłady (będzie to zatem definicja enumeracyjna niezupełna): kiedy piszę o teorii prawdy Tarskiego, zapewne mam pojęcie PRAWDA; kiedy analizuję argumenty Johna Rawlsa, posługuję się pojęciem SPRAWIEDLIWOŚĆ DYSTRYBUCYJNA; kiedy piszę ten artykuł, używam pojęcia MECHANICYZM. Wszystkie te pojęcia – czymkolwiek by pojęcia były – są abstrakcyjne, w odróżnieniu od pojęć konkretnych, takich jak OŁÓWEK, PODSTAWKA POD SZKLANKĘ, KROWA, RYŚ czy SŁUCHAWKI. Wymieniając jedynie przykłady, nie przesądzam, że rozróżnienie między tymi reprezentacjami abstrakcyjnymi i konkretnymi ma charakter dychotomiczny, gdyż w wielu teoriach w psychologii poznawczej bynajmniej tego charakteru nie ma; nie eksplikuję też, na czym polega abstrakcyjność pojęcia PRAWDA w odróżnieniu od konkretności pojęcia KROWA, bo teoretycy mają tutaj różne koncepcje, a moja analiza jest metateoretyczna, a przez to – minimalistyczna i nieco ascetyczna.

Można oczywiście wątpić, czy istnieje jakiś szczególny problem związany z reprezentacjami abstrakcyjnymi. Takie stanowisko zajmuje dwóch teoretyków kojarzonych z diametralnie różnymi poglądami: Jerry Fodor i Robert Rosen. Zdaniem Fodora, dla teorii treści w psychologii nie istnieje żadna interesująca różnica między reprezentacjami abstrakcyjnymi i konkretnymi: „Wszystkie predykaty wyrażają własności, a wszystkie własności są abstrakcyjne”<sup>34</sup>. Dlatego też semantyka wyrażenia „koń” i wyrażenia „cnota” nie różni się w żaden interesujący sposób – oba odnoszą się do własności (w teorii Fodora na mocy nomicznej relacji bycia przyczyną wystąpienia tego wyrażenia a odpowiednimi egzemplifikacjami tych

---

<sup>34</sup> J.A. Fodor, *A theory of content and other essays*, MIT Press, Cambridge, Mass. 1992, s. 111.

własności). Argument Fodora jest jednak o tyle nietrafny, o ile chcemy powiedzieć, że istnieją reprezentacje, które odnoszą się nie tylko do własności. Fodor z góry przesądza, że zwolennik np. reizmu nie może mieć racji, gdyż każde jego pojęcie będzie wyrażać własność abstrakcyjną; innymi słowy, teoria semantyczna Fodora przesądza kwestie ontologiczne, co jest co najmniej dziwaczne, zwłaszcza że niepoparte żadną dokładniejszą argumentacją. Przesądzanie o fałszywości pewnych tez ontologicznych na mocy określonej teorii znaczenia to stawianie wozu przed koniem: ponieważ Fodor przyjmuje stanowisko semantyki znaturalizowanej, osobliwe jest, że chce przesądzić, iż w ogóle o konkretach mówić nie można. A przecież w naukach o konkretach mówi się dosyć często. Czyżby Fodor chciał powiedzieć, że np. geolog mówiący o konkretnej próbce minerału zawsze mówi o własnościach i uniwersaliach?

O dziwo, nieco podobne do Fodora stanowisko ma Robert Rosen, biolog teoretyczny znany m.in. z dosyć szczegółowej i formalnej analizy relacji modelowania; jego zdaniem, każdy model jest abstrakcją, gdyż niektóre stopnie swobody systemu modelowanego nie są obecne w systemie modelującym<sup>35</sup>. Obserwacje z tego względu też są abstrakcjami, gdyż nigdy nie prowadzą do zamodelowania wszystkiego, lecz pomijają część dostępnych informacji. Za Rosenem można by powiedzieć, że w modelu próbki minerału w geologii pominięte zawsze zostaną pewne własności tejże próbki (np. własność taka, że wspomina o niej autor niniejszego artykułu). W tym sensie model powstaje na drodze abstrahowania – wybiera się tylko niektóre stopnie swobody rzeczywistości uwzględniane w modelu. Nie znaczy to jednak, że odnosimy się wyłącznie do abstraktów: sama reprezentacja opiera się na abstrahowaniu, a więc jest abstrakcją, ale niekoniecznie jej desygnat, który może być konkretem. Rosen nie twierdzi więc, w przeciwieństwie do Fodora, że modele zawsze odnoszą się do własności, które są abstrakcyjne, lecz że samo modelowanie jest abstrahowaniem. O tyle też nie przesądza tak wiele w swojej teorii modelowania: reizm nie staje się przy jej przyjęciu automatycznie fałszywy.

A może istnieją tylko reprezentacje konkretne? Tak, zdawałoby się, mógłby twierdzić rzecznik poglądu, że poznanie ma naturę cielesno-zmysłową. Jednakże reprezentacje abstrakcyjne należy traktować poważnie, nawet jeśli ob staje się przy bardzo ucieleśnionej koncepcji poznania. Otóż badania dotyczące rozumienia języka naturalnego wskazują, że istnieją różnice w przetwarzaniu informacji słów odnoszących się do abstraktów i słów odnoszących się do konkretów. Mowa o tzw. efekcie konkretno-

---

<sup>35</sup> R. Rosen, *Anticipatory systems. Philosophical, mathematical, and methodological foundations*, wyd. II, Springer, New York 2012, s. 17.

ści: u osób zdrowych przetwarzanie wyrazów desygnujących konkrety jest szybsze niż przetwarzanie wyrazów desygnujących abstrakty, co zostało potwierdzone licznymi eksperymentami<sup>36</sup>. Co ciekawe, u osób cierpiących na demencję semantyczną obserwuje się efekt odwrotny: łatwiej posługują się wyrazami abstrakcyjnymi<sup>37</sup>. Z tego też względu istnieje psychologiczna różnica, której nie powinno się zaciierać w koncepcji reprezentacji. Wiadomo więc, że istnieje pewne zjawisko, które teorie empiryczne chcą zbadać. Pora więc zobaczyć, jak się to ma do mechanicyzmu.

#### 4. Czy koncepcja mechanizmów reprezentacyjnych radzi sobie z abstraktami?

We współczesnej debacie kognitywistycznej można wyróżnić podstawowe cztery stanowiska w sprawie reprezentacji abstrakcyjnych<sup>38</sup> (poza klasycznym stanowiskiem związanym z amodalnymi symbolami):

1. Stanowisko George'a Lakoffa, zgodnie z którym wszystkie pojęcia są metaforyczne i związane z ludzką cielesnością, w tym także pojęcia abstrakcyjne<sup>39</sup>. Lakoff musi jednak odróżnić pojęcia mniej i bardziej cielesne;

---

<sup>36</sup> F. Jessen *et al.*, *The concreteness effect: evidence for dual coding and context availability*, „Brain and Language” 2000, t. LXXIV, nr 1, s. 103-112. Rzecz jasna przetwarzanie języka naturalnego to tylko pośredni dowód na istnienie co najmniej dwóch różnych typów reprezentacji umysłowych, desygnujących odpowiednio abstrakty i konkrety. W neurosemantyce zwykle przechodzi się gładko od analizy procesów odbioru i nadawania komunikatów językowych do analizy stosujących u ich podłoża reprezentacji. Jednakże to przejście można kwestionować; można też twierdzić, że myślenie abstrakcyjne wcale nie wymaga opanowania języka – to też kwestia do rozstrzygnięcia empirycznie, a nie czysto teoretycznie.

<sup>37</sup> T. Shallice, R.P. Cooper, *Is there a semantic system for abstract words?*, „Frontiers in Human Neuroscience” 2013, t. VII, s. 175.

<sup>38</sup> D. Pecher, I. Boot, S. Van Dantzig, *Abstract Concepts. Sensory-Motor Grounding, Metaphors, and Beyond*, w: *The Psychology of Learning and Motivation*, red. B. Ross, Academic Press, Burlington 2011, vol. 54, s. 217-248; K. Wiemer-Hastings, L.W. Barsalou, *Situating Abstract Concepts*, w: *Grounding cognition. The role of perception and action in memory, language, and thought*, red. D. Pecher i R. Zwaan, Cambridge University Press, New York 2005, s. 129-163. Pomijam tu stosunkowo niedawno zaproponowaną teorię, zgodnie z którą pojęcia abstrakcyjne mają większą walencję emocjonalną, gdyż ta teoria nie mówi nic o swoistym znaczeniu tych pojęć. Por. S.-T. Kousta *et al.*, *The representation of abstract words: why emotion matters*, „Journal of Experimental Psychology. General” 2011, t. CXL, nr 1, s. 14-34. Emocjonalne zabarwienie koreluje nie tylko z abstrakcyjnością, ale także z rozmiarem desygnatu (B. Yao *et al.*, *Semantic Size of Abstract Concepts. It Gets Emotional When You Can't See It.*, „PloS one” 2013, t. VIII, nr 9, s. e75000).

<sup>39</sup> G. Lakoff, *Kobiety, ogień i rzeczy niebezpieczne: co kategorie mówią nam o umyśle?*, red. E. Tabakowska, przeł. M. Buchta, A. Kotarba i A. Skucińska, Towarzystwo Autorów i Wydawców Prac Naukowych Universitas, Kraków 2011.

w przeciwnym razie jego teoria będzie wykluczać występowanie efektu konkretności.

2. Nieco pokrewne stanowisko Arthura Glenberga, zgodnie z którym wszystkie pojęcia mają naturę ucieleśnioną. Glenberg twierdzi, że ucieleśnienie to przejawia się w postaci bliskiego związku z działaniem; jego badania dowodzą, że rozumienie nawet abstrakcyjnych zdań aktywuje reakcje motoryczne<sup>40</sup>.

3. Teorię symulacyjno-percepcyjną Lawrence'a Barsalou, zgodnie z którą znaczeniem pojęcia jest postrzeżeniowa symulacja sytuacji, w których pojęcie się stosuje<sup>41</sup>. Barsalou analizuje np. pojęcie PRAWDA w postaci serii wyobrażeń sytuacji kogoś, kto komunikuje drugiej osobie, że jakieś zdanie jest prawdziwe, a ta druga osoba weryfikuje percepcyjnie, czy to, co wyraża dane zdanie, rzeczywiście ma miejsce.

4. Konceptje pluralistyczne, które postulują wiele kodów czy formatów reprezentacji, w tym kod werbalny bądź symboliczno-amodalny, mający odpowiadać reprezentacjom abstrakcyjnym<sup>42</sup>.

Nie jest w tym miejscu istotne, która z tych teorii jest faktycznie prawdziwa; wiele z nich może wydawać się mało przekonujących. Interesuje mnie pytanie, czy i w jaki sposób koncepcja mechanizmów reprezentacyjnych może posłużyć do analizy, a następnie uszczegółowienia wskazanych stanowisk w odniesieniu do odpowiednich zjawisk poznawczych. Ze względu na bliski związek z działaniem jest dosyć ewidentne, że stanowisko Glenberga da się łatwo połączyć z proponowaną przeze mnie koncepcją. Co jednak z innymi?

Mechanistyczna parafraza stanowiska Lakoffa nie wydaje się niemożliwa logicznie<sup>43</sup>, lecz jest o tyle kłopotliwa, że Lakoff odrzuca istotność

<sup>40</sup> A.M. Glenberg *et al.*, *Processing abstract language modulates motor system activity*, „Quarterly Journal of Experimental Psychology” 2008, t. LXI, nr 6, s. 905-19.

<sup>41</sup> L.W. Barsalou, *Perceptual symbol systems*, „The Behavioral and Brain Sciences” 1999, t. XXII, nr 4, s. 577-609; dyskusja s. 610-60; L.W. Barsalou, *Abstraction in perceptual symbol systems*, „Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences” 2003, t. CCCLVIII, nr 1435, s. 1177-87; K. Wiemer-Hastings i L.W. Barsalou, „Situating Abstract Concepts”, *op. cit.*

<sup>42</sup> A. Paivio, *Mental representations. A dual coding approach...*, *op. cit.*; G. Dove, *Beyond perceptual symbols. A call for representational pluralism*, „Cognition” 2009, t. CX, nr 3, s. 412-31; T. Maruszewski, *Psychologia poznania: umysł i świat*, Gdańskie Wydawnictwo Psychologiczne, Gdańsk-Sopot 2011; K. Sakreida *et al.*, *Are abstract action words embodied? An fMRI investigation at the interface between language and motor cognition*, „Frontiers in Human Neuroscience” 2013, t. VII, s. 125; B. Straube *et al.*, *Supramodal neural processing of abstract information conveyed by speech and gesture*, „Frontiers in Behavioral Neuroscience” 2013, t. VII, s. 120.

<sup>43</sup> Lakoff walczy z dosyć karykaturalną wersją komputacjonizmu niezgodnego z ucieleśnieniem (jest to o tyle osobliwe, że robotycy nie byłoby komputacjonistami); a ponieważ to jednak inny kom-



pojęcia odniesienia i flirtuje z idealistycznym podejściem do poznania. Wbrew temu, co sądzi, ucieleśnienie poznania jako takie wydaje się zgodne z obiektywistycznym podejściem do procesów poznawczych. Dokładne omówienie poglądów Lakoffa nie jest tutaj możliwe, lecz w każdym razie możliwa jest mechanistyczna analiza stanowiska Glenberga, bliskiego w wielu miejscach Lakoffowi. Tak czy inaczej, analizy Lakoffa nie zawierają wszystkich szczegółów niezbędnych do pełnego wyjaśnienia przyczynowego, które jest wymagane przez mechanicyzm; pozostaje głównie na poziomie analizy lingwistycznej (często popieranej jedynie świadectwami anegdotycznymi, a nie rzetelną analizą korpusową), pomijając mechanizmy psychologiczne i neuronalne.

Podejście neoempirystyczne Barsalou nie wiąże reprezentacji tak blisko z motoryką, jak u Glenberga, lecz też wydaje się zupełnie zgodne z przyjmowaną tu ramą pojęciową: informacje semantyczne będą po prostu miały kod percepcyjny (związany z określoną modalnością zmysłową), regulując działanie systemu poznawczego.

Mniej oczywista jest zgodność ze stanowiskami pluralistycznymi, postulującymi co najmniej dwa kody reprezentacji. Przyjmijmy, że mamy do czynienia ze stanowiskiem, które postuluje istnienie (1) reprezentacji analogicznych i (2) symbolicznych; pierwsze opierają się na stosunku homomorfizmu między nośnikiem reprezentacji a jej desygnatem, a drugie na zastosowaniu arbitralnego symbolu w podobny sposób, jak stosuje się funkcję do argumentów<sup>44</sup>. Jednocześnie, co podkreśla Aaron Sloman, musi istnieć proces obliczeniowy, który odpowiednio reaguje na strukturę nośnika. Jest dosyć jasne, że reprezentacje analogiczne zawierają informacje semantyczne w sensie zdefiniowanym wcześniej, o ile tylko system z nich korzystający reaguje inaczej na różne stany ich nośników. Co jednak z kodami symbolicznymi? Symbole oczywiście także niosą informacje, chociaż nie na mocy relacji homomorfizmu między ich nośnikiem a desygnatami. Jeśli jednak wziąć pod uwagę fakt, że symbole należy analizować *wraz* z obsługującymi je procesami obliczeniowymi, to z pewnością będziemy mieć do czynienia z informacjami strukturalnymi: od struktury składniowej odpowiednich wyrażeń symbolicznych (mających różną głębię logiczną, czyli stopnie swobody) będzie zależeć zachowanie systemu

---

putacjonizm niż broniony przeze mnie, nie jest to zbyt wielkim problemem. Wynika to u niego zresztą z przyjmowania Putnamowskiej analizy pojęcia obliczania, z którą wielokrotnie już polemizowano (por. M. Miłkowski, *Explaining the Computational Mind...*, *op. cit.*; J. Buechner, *Godel, Putnam, and Functionalism*, MIT Press, Cambridge, Mass. 2008; D.J. Chalmers, *A Computational Foundation for the Study of Cognition*, „Journal of Cognitive Science” 2011, nr 12, s. 325-359).

<sup>44</sup> A. Sloman, *The computer revolution in philosophy: philosophy, science, and models of mind*, Humanities Press, Atlantic Highlands N.J. 1978, rozdz. 7.

interpretującego je. Skoro tak, to relacje przepływu informacji można będzie odkryć także w wypadku kodu symbolicznego, co zresztą sugerowała w latach 80. semantyka kauzalno-informacyjna Dretskego<sup>45</sup> czy Fodora. W odróżnieniu od tych stanowisk nie twierdzę, że informacja zawarta w symbolach stanowi ich znaczenie; wymagam jedynie, aby ta informacja była semantyczna, tzn. regulowała gotowość systemu do działania.

Cóż to dokładniej oznacza? Otóż rozmaicie zakodowane – lub sformułowane – informacje mogą równie dobrze służyć jako nośniki informacyjne; koncepcja mechanizmów reprezentacyjnych żadnego tu nie uprzywilejowuje; nie ma większego znaczenia, czy mówimy o symbolach, czy o modelach analogowych<sup>46</sup>. Nic dziwnego, że eksperymentalne rozstrzygnięcie, czy mamy do czynienia z reprezentacją analogiczną, czy symboliczną, napotyka ogromne trudności, gdyż niosą informacje w podobny sposób<sup>47</sup>. Różnią się jednak wyraźnie procesami stosowanymi do przetwarzania takich informacji: symbole, jeśli informują o rzeczywistości fizycznej, muszą w istocie być z nią powiązane na mocy jakichś prawidłowości fizycznych. Rzecz jasna, procesy obliczeniowe korzystające z reprezentacji symbolicznych i te korzystające z analogicznych mogą różnić się znacznie złożonością obliczeniową. To właśnie – jak się zdaje – ma miejsce w przypadku ludzi i może wyjaśniać wspomniane efekty konkretności i odwrócony efekt konkretności, związane zapewne z generowaniem reprezentacji analogicznych i symbolicznych (zgodnie z teoriami pluralistycznymi) podczas odbioru komunikatów językowych. W świetle analizy mechanicznej okazuje się, że teoria pluralistyczna musiałaby postulować uprzywilejowanie lub mniejszą złożoność procesów związanych z reprezentacjami analogicznymi niż w przypadku procesów korzystających z reprezentacji symbolicznych. Tego rodzaju różnicę – w złożoności przetwarzania – muszą przyjąć zresztą także zwolennicy bardziej ucieleśnionych ujęć abstraktów, tacy jak Glenberg czy Barsalou.

Informacje – niezależnie od ich formatu – modyfikują gotowość systemu do działania i na mocy procesów przepływu informacji pozwalają na wnioskowanie o strukturach będących u ich źródła. Weźmy prosty

---

<sup>45</sup> F.I. Dretske, *Knowledge and the Flow of Information*, wyd. II, MIT Press, Cambridge, Mass. 1982.

<sup>46</sup> Obie analizować można bowiem przy użyciu pojęcia informacji semantycznej w sensie MacKaya. Warto zauważyć, że posługuję się tu pojęciem symbolu stosunkowo wąsko, wskazując na nieanalogiczny charakter samego jego nośnika. Wydaje się, że pojęcie symbolu np. u H. Pattee jest bliższe mojemu pojęciu informacji semantycznej. Por. H.H. Pattee, *Physical and Functional Conditions for Symbols, Codes, and Languages*, „Biosemiotics” 2008, t. I, nr 2, s. 147-168.

<sup>47</sup> J.R. Anderson, *Arguments concerning representations for mental imagery*, „Psychological Review” 1978, t. LXXXV, nr 4.

przykład: taniec pszczół jest nośnikiem informacji o położeniu nektaru (istnieje matematyczne odwzorowanie cech nektaru w analogiczną reprezentację w postaci tańca), jak i odpowiednio modyfikuje gotowość pszczół-obszerników tańca do działania. Treść tego dynamicznego znaku nie jest oderwana tu od jego efektywności, czyli tego, jak reguluje on działanie pszczół.

W przypadku reprezentacji abstrakcyjnych w ujęciu pluralistycznym (i klasycznym, czysto symbolicznym) można mówić o nośniku informacji semantycznych, który nie informuje o żadnym czasoprzestrzennym przedmiocie, ale ma strukturę, która pozwala na przeprowadzanie różnych operacji obliczeniowych i modyfikuje gotowość do działania. Może się tak dzieć w przypadku reprezentacji pustej, ale także skrajnie abstrakcyjnej, np. PRAWDA czy ZBIÓR. Wytworzenie takich pojęć w obliczeniowym systemie sztucznym jest łatwiejsze niż stworzenie reprezentacji konkretnych, gdyż nie wymaga zachowania związku informacyjnego ze światem rzeczywistym. Jednakże dopiero wtedy, gdy system jest w stanie wykrywać błędność takich reprezentacji, możemy mówić o tym, że są one jego reprezentacjami. System wykrywający niepoprawne zastosowanie pojęcia PRAWDA musi być w stanie dokonywać pewnych wnioskowań na temat pewnych innych (zapewne zdaniowych) reprezentacji i ich zgodności z tym, co mają reprezentować, co wymaga zdolności do wykrywania rozbieżności reprezentacji i tego, co miała reprezentować. Jest to zdolność, jak widać, dosyć złożona, ale może być zrealizowana w różny sposób. Pełne wyjaśnienie działania mechanizmu reprezentacyjnego posługującego się takim pojęciem musi spełnić wszelkie wymogi wyjaśnień mechanistycznych, co oznacza, że wymaga ono podania kompletnego modelu przyczynowego takiego mechanizmu, a to z kolei jest niemożliwe bez przeprowadzenia stosunkowo skomplikowanych badań empirycznych (wykracza to więc poza ramy tego artykułu).

## 5. Podsumowanie

W tym artykule przedstawiono koncepcję mechanizmów reprezentacyjnych i krótko odniesiono ją do problematyki reprezentacji abstrakcyjnych. Argumentowałem, że koncepcja ta nie przesądza z góry kwestii empirycznych, gdyż jest stosunkowo neutralna, a jedynie wymaga stosowalności pojęć informacyjno-obliczeniowych, a także pojęcia odniesienia, treści i wartościowania. Wskazując na wagę wykrywania błędności przez system, podkreśliłem, że dopiero istnienie mechanizmów ewaluacyjnych jest probierzem reprezentacyjnego charakteru nośnika informacji semantycznych w danym systemie poznawczym.

Ze względu na ograniczenia objętości niektóre zagadnienia musiałem potraktować skrótowo, co może pozostawiać pewne uczucie niedosytu. Jest to cena prezentacji w miarę pełnego obrazu koncepcji wraz z jej zastosowaniem. Niestety dokładne studia przypadków reprezentacji abstrakcyjnych to temat na inny artykuł. Sądzę jednak, że mechanistycznie można analizować większość kognitywistycznych badań na temat reprezentacji, w tym pluralistyczne i czysto symboliczne (tradycyjne), a nie tylko wyrażenie związane z motoryką, co sugerowałby związek reprezentacji z działaniem, jaki postuluje się w koncepcji mechanistycznej.

### *Representational Mechanisms and Abstractions*

In this paper, I outline a theory of action-oriented representational mechanisms, which is a conceptual framework used to analyze mental representations in cognitive science. The framework is based on neo-mechanism, and it insists that representational explanations are mechanistic explanations, which are causal and appeal to functional mechanisms. Representations are always posited as related to capacities of a cognitive system. This means that they are not just a vehicle of semantic information, but also control the behavior of the system in a representational manner; in other words, they are action-oriented. Representations that are not poised to control any processes are denied to be representations. However, it may be asked whether the framework does not presuppose that certain psychological theories – the ones that posit representations not related directly to action – are incorrect. One example of such representations is abstract representations such as TRUTH or CONDITIONAL PROBABILITY. My aim is to show that the framework does not require narrowly conceived motor operations as immediate outcomes of representational processing. I shortly review psychological theories of abstract representations to show that they can be compatible with the framework of representational mechanisms.