

Marcin Miłkowski

epistemologia znaturalizowana

O filozofach analitycznych opowiada się od lat serią dowcipów z cyklu „Dowód X-a, że *p*”. Na przykład dowód Donalda Davidsona, że *p*, jest następujący: „Załóżmy, że *p*”.

Pójdę przykładem Davidsona i założę, że słuszne jest podejście naturalistyczne w filozofii. Choćby to założenie było fałszywe, być może cały tekst okaże się interesującym eksperymentem myślowym. W eksperymencie pokażę, jak epistemologia znaturalizowana może skonstruować najprostszą, znaturalizowaną wersję systemu poznawczego: coś w rodzaju robotycznego *ego* transcendentálnego. Niestety, nie załączę do tego tekstu dokumentacji technicznej.

Naturalizacja epistemologii

Punktem wyjścia epistemologii znaturalizowanej jest konstatacja, że jedyną rzetelną metodą uzyskiwania wiedzy jest metoda naukowa (co nie znaczy, że nauki są pozbawione ograniczeń poznawczych)¹. Gdyby bowiem istniały lepsze i rzetelniejsze metody, stałyby się one wnet naukowe. Jeśli do tej pory się nimi nie stały, to widocznie na to nie zasługują.

To jednak nie oznacza potępienia w czambuł wszystkich innych stanowisk. Swoistością prac naturalistów są próby przyswajania, czyli naturalizowania, innych podejść: próby przełożenia ich na praktykę naukową. Dlatego też naturalista może znaturalizować *ego* transcendentálne, czyli zespół koniecznych i wystarczających warunków procesu poznania.

Ponieważ epistemologia dotyczy procesów poznawczych, które są dostępne empirycznie, metodą ich badania nie jest tylko formalizacja, lecz też obserwacja i tworzenie teorii. Dlatego właśnie Quine'owska epistemologia znaturalizowana dąży do wyeliminowania spekulacji filozoficznych na rzecz psychologii poznawczej (Quine [1986]). Idzie o zastąpienie apriorycznych konstrukcji teoriami empirycznymi. Oczywiście, ta-

kie *dictum* może oznaczać koniec epistemologii jako dziedziny filozoficznej. Naturaliści zwykle dodają: być może to nie będzie filozofia, ale dlaczego musimy tkwić przy tradycyjnych podziałach dziedzin wiedzy? Niektórzy nawet dostrzegają w dociekaniach Husserla czy Heideggera sporo treści pożywnych dla empiryków w psychologii (np. Clark [1997]).

Najbardziej przekonująco tego rodzaju epistemologia brzmi w kontekście praktycznych badań kognitywistów. I gdy Quine pozostawał jeszcze pod wpływem behawioryzmu, to dzisiejsi kognitywiści nie boją się postulować procesów umysłowych i tworzyć na ich temat teorii. Dlatego też epistemologia znaturalizowana dostarcza siatki pojęciowej kognitywistyce, a kognitywistyka - problemów epistemologii znaturalizowanej.

Normatywne aspekty funkcji poznawczych

Tradycyjnym zarzutem przeciwko epistemologii znaturalizowanej jest to, że nie jest w stanie ona dostarczyć tego, co daje zwykła epistemologia: zasad normatywnych. Jak nadać normatywny wyraz robotycznemu *ego*? Wielu naturalistów się z tym zarzutem zgodzi, choć już nie Quine. Dla radykalnego Quine'a epistemologia jest, jak każda dziedzina normatywna, po prostu rodzajem inżynierii. Jak się zdaje, Quine sądził, że zasady normatywne, czy to etyczne, czy to epistemologiczne, dają się sprowadzić do wiedzy praktycznej, którą opracowują inżynierowie.

Sądzę, że da się pogłębić ideę Quine'a, jeśli inżynierię będziemy rozumieć nieco głębiej niż autor *Słowa i przedmiotu*. Nauki inżynierskie słabo się mieszczą w tradycyjnym podziale na nauki empiryczne i dedukcyjne. Weźmy chociażby informatykę, gdzie mamy rozważania czysto matematyczne i teoretyczne, quasi-empiryczne badania związane z modelowaniem komputerowym lub empirycznym badaniem budowanych robotów, wreszcie zwykłą technikę budowania

różnych maszyn liczących. Taką swoistość dzielają różne dziedziny inżynieryjne, w tym konglomerat nauk kognitywnych.

Zdaniem Daniela Dennetta, nie tylko kognitywistyka, ale i biologia stanowią dziedziny w istocie inżynieryjne, choć biologia nie zajmuje się artefaktami (Dennett [1995]). Jednak artefakty i wytwory doboru naturalnego mają cechę wspólną: jest to konstrukcja czy plan konstrukcyjny, który można opisać w sposób funkcjonalny, przyjmując tak zwane „nastawienie funkcjonalne” (*design stance*). Z owego nastawienia funkcjonalnego opisywać można także wewnętrzne funkcje elementów architektury poznawczej, jak i zewnętrzne funkcje systemu wyposażonego w aparat poznawczy.

Otóż na poziomie funkcjonalnego opisu elementów architektury do czynienia możemy mieć z normatywnością. Normatywnością w tym sensie, że można wyróżnić normalny stan, w którym architektura funkcjonuje poprawnie, oraz stany anormalne. Podobnie na poziomie opisu osobowego do czynienia można mieć z dysfunkcjami i stanami pożądanymi.

To bardzo słaby rodzaj normatywności. Nie ma nic wspólnego z normatywami, którymi rzekomo kierują się podmioty poznające, typu „nie popełniaj nigdy sprzeczności” (szczegółowa krytyka idei zasad normatywnych w epistemologii znajduje się w Knowles [2003]).

A zatem epistemologia znaturalizowana będzie opisywać minimalny system poznawczy w kategoriach funkcjonalnych, a nie czysto fizykalnych czy intencjonalnych. W intencjonalnych kategoriach nie ma sensu go opisywać, gdyż to one mają w ostatecznym rozrachunku zostać wyjaśnione. Jest to zatem pewnego rodzaju redukcjonizm. Nie jest to jednak redukcjonizm skrajnie radykalny, gdyż uznaje się, że poziom opisu fizykalnego jest niespecjalnie interesujący dla epistemologii. Między poziomem funkcjonalnym a fizykalnym zachodzą zależności nieliniowe, uniemożliwiające prostą redukcję; ogólnie rzecz biorąc można powiedzieć, iż poziom funkcjonalny będzie emergentny względem fizykalnego.

Konstrukcja i funkcje

Aby modelować funkcje umysłowe minimalnego systemu poznawczego, trzeba wiedzieć, co to znaczy realizować funkcję i jakie funkcje chcemy modelować. W obu wypadkach można dać

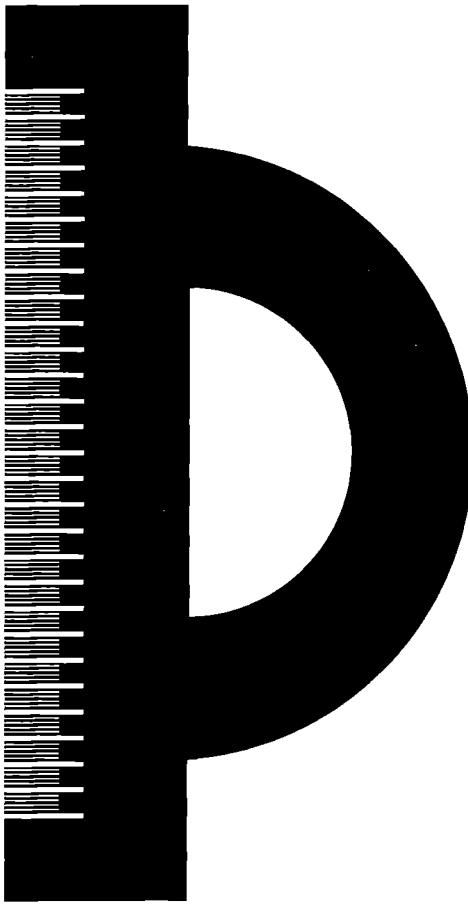
odповідź zawężającą, jak i szeroką. Z jednej strony za funkcję można uznawać wszelką rolę przyczynowo-skutkową (Cummins [1975]). Wówczas można mówić o funkcjonowaniu elektronów w atomach czy zwyczajów w społecznościach. Można także zawęzić pojęcie do funkcji, które istnieją dlatego, że występowały one u organizmów będących przodkami danego organizmu (biologiczno-etiologiczne pojęcie, Millikan [1984]). To ujęcie nieco za wąskie – u Millikana nie można mówić o funkcji prototypów. Nieco bardziej użyteczne jest odniesienie funkcji do konstrukcji (*design*) czy planu konstrukcyjnego. Otóż funkcje mają te elementy złożonego układu, które zostały dobrane zgodnie z planem konstrukcyjnym albo przez dobór naturalny, albo przez świadomego konstruktora. W pierwszym wypadku planem konstrukcyjnym jest po prostu genom, w drugim – zamiysł.

Podobnie różnorako definiować można zakres modelowanych funkcji. Z jednej strony można starać się modelować ludzi, co oczywiście jest ogromnie trudne ze względu na złożoność. Z drugiej strony, można próbować analizy prostych organizmów lub prostych artefaktów. W tym drugim wypadku nie wiadomo, czy w takim prostym systemie ujawnią się wszystkie interesujące nas cechy dowolnych systemów poznawczych. Ponieważ jednak ego transcendentalne to idealizacja, a nie empiryczne ja jakiegoś podmiotu, nie będziemy się tym zbytnio frasować.

Architektury poznawcze

W tradycyjnej epistemologii sporo mówiło się o władzach poznawczych. Dziś to pojęcie niemożliwe nawet wśród tradycjonalistów. Ale kognitywiści chętnie mówią o funkcjach poznawczych, modułach czy innych elementach organizacyjnych umysłu. Wszystkie te hipotezy określa się często metaforycznie mianem „architektury poznawczej”.

Kant żywił pewną predylekcję do symetrii w „architektonice”. Dzisiejsi filozofowie zatrucili może pociąg do czterech serii trójczłonowych podziałów, lecz posługują się podobną metaforą. Pojęcie architektury używane jest jednak także w informatyce, gdzie nie jest tylko i wyłącznie metaforycznie. Opisuje ono strukturę danego komputera; mówi się o architekturach związanych z pewnymi typami procesorów, oprogramowaniem itd. Przez analogię w kognitywistyce wprowadzono pojęcie architektury



systemu poznawczego, analizowanego na wielu poziomach, też podobnie jak w informatyce.

Architektura jest to więc organizacyjna struktura umysłu. Różnego rodzaju architektury postulowano zrazu w badaniach nad sztuczną inteligencją - pojęcie ma więc informatyczny rodowód. Postulowano architektury umożliwiające przetwarzanie symboli fizycznych i architektury neuropodobne, rozproszone i autonomiczne. Na tej podstawie tworzone były programy, które miały wykazywać się pewną inteligencją.

Modelowanie w sztucznej inteligencji w dużej mierze jest intuicyjne (zwłaszcza w nurcie symbolicznym) lub mocno abstrakcyjne (w nurcie sieci neuronowych). Istnieją też inne sposoby postulowania wewnętrznej architektury systemu: na podstawie istniejących kompetencji poz-

nawczych postuluje się struktury wewnętrzne wystarczająco złożone, aby mogły realizować te kompetencje. Na podstawie wiedzy o ograniczeniach historii ewolucyjnej gatunku ludzkiego można wyróżnić grupę warunków, które musiał faworyzować dobór naturalny, a mianowicie zdolności sprzyjających przetrwaniu i reprodukcji. Będzie to więc stosunkowo mała podatność na uszkodzenia (np. w walce z drapieżnikami); możliwość korzystania z ograniczonych świadectw i wnioskowania na ich podstawie. Mówiąc ogólnie, można nałożyć na postulowaną architekturę ograniczenia związane z rzeczywistością realizowanymi zadaniami poznawczymi, np. z posługiwaniem się językiem. Będziemy więc mieli do czynienia z wnioskowaniem o architekturze na podstawie kompetencji osobowych, a jednocześnie subsobowych ograniczeń związanych z budową biologiczną organizmu.

Wszystkie te ograniczenia nie wskazują jednoznacznie na jedną, określoną architekturę ego transcendentального, lecz raczej na cały ich szereg. Pod względami poznawczymi jednak będą to architektury równoważne.

Najprostsze systemy poznawcze

Modelowanie najprostszyc systemów poznawczych - ściśle ujmowanie ich funkcji, a więc i konstrukcji - w dużej mierze wiąże się z formalną analizą pojęciową, czyli próbą formalnego uchwycenia istoty procesu poznawczego. Mniejszy będzie tu związek z istniejącymi nurtami w badaniach nad sztuczną inteligencją.

Punktem wyjścia mogą być najogólniejsze rozważania. Do miana systemu poznawczego pretendować może system, który rozpoznaje stany środowiska (system poznawczy dokonujący wyłącznie dedukcji z „wrodzonych” założeń wydaje się nieco bardziej ograniczony, choć być może różnica nie jest taka ostra). Rozpoznanie stanu środowiska to rozpoznanie wzorca występującego w środowisku. Do rozpoznania wzorca konieczny jest algorytm, który definiuje, czym jest wzorec. A zatem najprostszyc system poznawczy musi dysponować mechanizmem, który umożliwi realizację algorytmów rozpoznawania wzorców; być może nie wzorców wszelkiego rodzaju, ale przynajmniej jakiegoś rodzaju wzorców. Wzorec poza tym można zdefiniować jako nielosowy ciąg wartości. Ciągi nielosowe cechują się mniejszą złożonością niż ciągi stochastyczne (miarą złożoności, tzw. złożoności

Kołmogorowa-Solomonoffa-Chaitina jest stopień kompresowalności ciągów, por. Chaitin [1975]). Najprostszy system poznawczy będzie to zatem taki układ fizyczny, który dzięki swojej budowie może realizować algorytm. Wymaga to istnienia odpowiedniej struktury - czyli architektury, gdyż nie każda struktura jest w stanie realizować dowolny algorytm (por. np. Chalmers [1994]).

Ten system poznawczy jest na razie konstrukcją pojęciową, lecz dokładne określenie minimalnego algorytmu, który może rozpoznawać wzorzec, będzie początkiem inżynierskiej specyfikacji, mogącej posłużyć do zbudowania prototypu takiego systemu (prototyp równie dobrze może być wirtualny i modelowany komputerowo).

Stopniowe komplikowanie tego modelu może okazać się drogą do odpowiedzi na tradycyjne pytania stawiane przez filozofów. Na samym początku odpowiedzi będą mało satysfakcjonujące, bo czegoś spodziewać się po artefakcie o inteligencji nieco mniejszej od kompetencji poznawczych prostej bakterii? A jednak będzie to robotyczne *ego* transcendentalne, zrealizowane fizycznie, lecz opisane na poziomie funkcjonalnym. Minimalne *ego* robota musi być w stanie rozpoznawać występujący faktycznie wzorec. Być może tropizmy bakterii można by postrzegać właśnie tak, choć bakterie są znacznie bardziej złożone, gdyż poza tym potrafią sterować swoim zachowaniem. Minimalne *ego* ma raczej poziom umysłowy godny organelli komórkowych.

Naturalizacja epistemologii i filozofii

Znaturalizowana epistemologia znakomicie dopełnia się z kognitywistyką. Być może inne rodzaje epistemologii nadają się do innych celów. Nie czuję się jednak na siłach powiedzieć, czy i jaka epistemologia jest epistemologią pleno titulo, jaka może być filozofią pierwszą i jaka mogłaby wystąpić jako prolegomena do metafizyki.

Wydaje mi się, że zarysowany wyżej program naturalistyczny jest wykonalny. Jest to oczywiście jedynie program, a jego realizacja wymaga szczegółowego rozpisania wymagań, jakie musi spełnić architektura systemu poznawczego, najpierw minimalnego, a później bardziej złożonego. Narzuca się pytanie, czy taka działalność w ogóle jest filozofią. Naturalista może repli-

kować, że skoro udaje się naturalizować także Heideggera, to i wcześniejsi filozofowie faktycznie uprawiali proteorię empiryczną, choć robili to na modłę pana Jourdain. Inny z kolei, o bardziej matematycznej duszy, odwoła się do idei Johna McCarthy'ego, który postrzega epistemologię korzystającą z osiągnięć sztucznej inteligencji jako dyscyplinę formalną i matematyczną (McCarthy [1989]). A od dziedzin matematycznych do idei tradycyjnej filozofii i epistemologii dosyć blisko. Mam wszelako wrażenie, że jeśli nie dokona się systematycznej obrony idei naturalizmu, w którym podział na dyscypliny aprioryczne i aposterioryczne jest nieco zatarty, to wszelkie odpowiedzi będą brzmiały nieczysto. Ale tu nie miejsce na pełną legitymizację programu naturalistycznego. Dlatego przyjąłem takie, a nie inne założenie, że p.

Bibliografia

- Chaitin, G. J., 1975, *Randomness and Mathematical Proof*, „Scientific American” 232 nr 5, s. 47-52.
- Chalmers, D.J., 1994, *On Implementing a Computation*, „Minds and Machines”, 4, s. 391-402.
- Clark, A., 1997, *Being There. Putting Brain, Body, and World Together Again*, MIT Press, Cambridge, Mass.
- Cummins, R., 1975, *Functional Analysis*, „The Journal of Philosophy”, 72, nr 20, s. 741-765.
- Dennett, D., 1995, *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, Simon & Schuster, Nowy Jork.
- Knowles, J., 2003, *Norms, Naturalism and Epistemology: The Case for Science without Norms*, Palgrave Macmillan.
- Mccarthy, J., 1989, *Artificial Intelligence, Logic and Formalizing Common Sense*, w: *Philosophical Logic and Artificial Intelligence*, (red.) Richmonda Thomasona, Kluwer Academic: Dordrecht.
- Millikan, R.G., 1984, *Language, Thought, and Other Biological Categories. New Foundations for Realism*, MIT Press, Cambridge (Mass.).
- Quine, W.V., 1986, *Epistemologia znaturalizowana*, w: W.V. O. Quine, *Granice wiedzy i inne eseje filozoficzne*, tłum. B. Stanosz, Warszawa, s. 106-125.

¹ To jest moje założenie, że p.