

Marek Hetmański

Czy komputer jest podmiotem poznania?

Filozofia Nauki 4/3, 131-142

1996

Artykuł został zdigitalizowany i opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

Marek Hetmański

Czy komputer jest podmiotem poznania?

Pytanie tytułowe jest sformułowane w stylistyce teorii sztucznej inteligencji, lecz mimo to nie zamierzam go w tej konwencji rozstrzygać. Niemniej jednak moją intencją jest analiza związków i podobieństw między człowiekiem a maszyną liczącą, ściślej mówiąc — podobieństw między pewnymi rodzajami mechanicznego funkcjonowania organizmu ludzkiego (na poziomach procesów cielesnych, stanów umysłowych, zachowania się w określonych sytuacjach) a funkcjonowaniem komputera cyfrowego. Pytanie tytułowe musi jednak zostać przeformułowane na kilka innych, bardziej szczegółowych i sensownych, gdyż w takiej postaci mylnie sugeruje, iż komputer, na którym napisałem te słowa, jest podmiotem poznającym w takim zakresie i stopniu, jak ja — czy moi Czytelnicy. Krótko mówiąc, choć komputer spełnia niektóre z funkcji ludzkiej podmiotowości, to nie jest jednak w pełni podmiotem poznania analogicznym do podmiotu ludzkiego. Można mimo to mówić o *sui generis* podmiotowości maszyny cyfrowej, pod tym jednak warunkiem, że spojrzymy na tę własność z perspektywy cybernetyki i teorii informacji, w których zarówno człowiek, jak i cyfrowa maszyna licząca są układami cybernetycznymi, wykazującymi pewne istotne podobieństwa. Człowiek i komputer — a dokładniej, ich podukłady, ściśle określone poziomy organizacji i funkcjonowania — są podklasami ogólnej klasy maszyn. Podlegają one tym samym prawom fizycznym, a ich organizacja i działanie sprowadzają się do podobnych procedur mechanicznych.

Należy zatem postawić następujące pytania: Czy maszyna — cyfrowa maszyna licząca czyli komputer — i człowiek (jej twórca) są układami, które wykazują, mimo zasadniczych różnic w fizycznej strukturze i budowie, podobieństwa w organizacji podstawowych procesów tej struktury oraz reguł ich działania? Czy istnieje wspólny zbiór reguł działania takich układów? Czy istnieje jeden model wyjaśniający ich mechaniczne działanie? Nie są to pytania tak prowokujące, jak pytanie tytułowe i dotyczy

zagadnień o wiele starszych niż problem sztucznej inteligencji, bo sięgających siedemnastego stulecia i stawianych już przez Kartezjusza i La Mettriego.

Zauważmy, że porównanie człowieka z maszyną analogową (zegarem, centralą telefoniczną itp.) ma odmienne znaczenie niż porównanie z maszyną cyfrową, która charakteryzuje się innymi zasadami działania. Człowiek jest w pewnym stopniu maszyną analogowo-cyfrową, komputerem (von Neumann 1963, s. 80). Dotyczy to szczególnie jego ośrodkowego układu nerwowego i mózgu, które — zgodnie z coraz bardziej zaawansowanymi badaniami — wykazują analogię z budową i działaniem komputera cyfrowego. Analogia pomiędzy maszyną a człowiekiem pojawia się jednak i nabiera sugestywnego, choć mylącego znaczenia dopiero w sytuacji przyjęcia jednostronnej postawy badawczej — skrajnie redukcjonistycznej: analizującej działanie obu układów i ich zachowanie zewnętrzne z pominięciem istotnych różnic pomiędzy ich budową i tworzywem. To behawiorystyczne nastawienie właściwe teorii sztucznej inteligencji rodzi ograniczenia i nieporozumienia.

Myślenie a modele

Nauki neurofizjologiczne i psychologia eksperymentalna, z jednej strony, oraz nauki o komputerach, kognitywistyka i teoria sztucznej inteligencji, z drugiej, w ramach swoich metod i uzyskiwanych wyników realizują podobny cel — konstruują modele budowy oraz działania człowieka (jego mózgu i układu nerwowego) oraz komputera. Model jest teorią (zbiorem hipotez, założeń i wniosków do weryfikacji), ale również fizycznym konstruktem, obrazującym dane zjawisko (por. mapa, globus). W każdym wypadku model odwzorowuje z różnym stopniem dokładności faktyczną budowę i działanie badanego zjawiska; bez modelowania zjawisk przyrodniczych czy fizycznych nie byłoby ich poznania. Model jest użytecznym narzędziem poznawczym, niekiedy jednak powoduje pewne nieporozumienia.

Matematyczne i logiczne modele mózgu stworzone przez McCullocha, Pittsa, Arbibę (1968; 1977), modele cybernetyczne Wienera (1971) czy biologicznych i sztucznych homeostatów, skonstruowane przez Ashby'ego (1954, 1961), są próbą formalnego odwzorowania złożoności centralnego układu nerwowego człowieka. Dla ich konstrukcji twórcy ci świadomie odwoływali się do logicznych i algebraicznych modeli opisujących konstrukcję i zasady działania pierwszych komputerów. Tu leżą historyczne i merytoryczne przesłanki analogii człowieka i komputera, rozwijanej przez teorię sztucznej inteligencji. W tym kontekście chcę postawić zasadnicze pytanie tego tekstu. Czy model mózgu — w tym również jego stanów funkcjonalnych, stanów umysłowych — może być zbudowany przez analogię do modelu opisującego komputer cyfrowy? Czy model komputera, jakim jest maszyna Turinga, może być modelem mózgu? Pytanie to jest o wiele bardziej interesujące niż pytanie o to, kiedy komputer będzie myślał dokładnie tak, jak my. Innymi słowy, chcemy rozstrzygnąć problem, czy maszyna Turinga, będąc homomorficznym obrazem każdej maszyny cyfrowej, jest izomorficzna

z dowolnym obrazem homomorficznym mózgu i/lub umysłu. Pytanie to opiera się na założeniu, do pewnego stopnia umownym i uproszczonym, niemniej użytecznym dla niniejszych rozważań, iż obiekt X jest modelem obiektu Y , gdy homomorficzny obraz X -a — X' jest izomorficzny (tożsamy) z homomorficznym obrazem Y -a — Y' . Izomorficzność X' i Y' , która jest relacją symetryczną, jest więc warunkiem tego, że X jest modelem Y -a (Briukow, Geller 1983, s. 39). Trzeba zaznaczyć, iż X jest modelem Y -a w granicach ich wzajemnej izomorficzności, jeśli jest przez jakiś podmiot (operacyjny, poznawczy) traktowany jako model, jako odwzorowywanie.

Sformułowanie „Mózg jest komputerem” czy „Umysł jest programem komputerowym” — może mieć charakter tylko metaforyczny. Metafora umożliwia poznanie jakiegoś układu (w tym wypadku mózgu, umysłu) przez porównanie go z innym układem (maszyną cyfrową), który znamy lepiej. Nie może jednak być uznana za teorię tego układu, a tylko za źródło wartościowych hipotez na jego temat (Arbib 1977, s. 29). Przy ciągle niepełnej wiedzy o budowie i funkcjonowaniu organizmu ludzkiego — jego porównanie z maszyną cyfrową musi być stale korygowane, gdyż w przeciwnym razie nośna metafora zastąpi adekwatny i rzetelny opis złożonego zjawiska i da fałszywy jego obraz.

Człowiek i maszyna jako układy częściowo niepoznawalne

Na marginesie głównego wątku rozważań warto w tym miejscu postawić pytanie natury metodologicznej. Na jaki rodzaj wiedzy można liczyć w wypadku tworzenia modeli układów o wysokim stopniu złożoności, w dużej części nieobserwowalnych? Ashby w koncepcji «czarnej skrzynki» (1961, s. 128-169) sformułował ogólne warunki poznawalności układu częściowo nieobserwowalnego. Układ taki badamy poprzez rejestrację jego «wejść» i «wyjść», sprzężeń zwrotnych, następstwa czasowego stanów, ich zdeterminowania lub statystycznego prawdopodobieństwa, korelacji układu i obserwatora. Można o takim układzie mieć wiedzę fragmentaryczną, częściową, i która mimo to jest zupełna (w znaczeniu algebraicznym): która jest jego reprezentacją kanoniczną. Zestawienie ze sobą dwu różnych maszyn (odmiennych co do tworzywa, budowy czy przeznaczenia) możliwe jest pod warunkiem zredukowania opisu (modelu) maszyny bardziej złożonej do poziomu opisu maszyny prostszej. Porównanie maszyn w pełni izomorficznych, określonych ze względu na «podobieństwo postaci», «równość», polega na tym, że ze względu na sposób zachowania są one funkcjonalnie nieodróżnialne. Taką sytuację miał na myśli Turing w swojej koncepcji «gry w udawanie» (1995, s. 271), w której zakładał, że człowiek i komputer (pod koniec obecnego stulecia) — analizowane pod kątem udzielanych odpowiedzi na te same pytania — zachowują się w równym stopniu inteligentnie, czyli tak, jak dwie maszyny izomorficzne w znaczeniu nadanym przez Ashby'ego. Na tym założeniu oparta jest cała teoria sztucznej inteligencji. Wydaje się jednak, że nic nie uzasadnia przekonania o pełnej izomorficzności

człowieka i komputera. Pozorne podobieństwo «inteligentnych» odpowiedzi komputera i odpowiedzi człowieka na pewno nie wynika z izomorficzności obydwu układów.

Mniejszy stopień wzajemnego podobieństwa wykazują tzw. maszyny homomorficzne. Mamy tu do czynienia z przykładem redukcji opisu (modelu) maszyny złożonej poprzez niepełne rozróżnienie jej stanów, eliminację niektórych sprzężeń zwrotnych, uproszczenie korelacji między układem a obserwatorem — i zestawieniem tego opisu z maszyną prostszą, której model może być w relacji izomorficzności z tak uproszczonym opisem maszyny złożonej. Jest to podobieństwo oparte na znacznym, stopniowalnym uproszczeniu układu złożonego, które nie wyklucza mimo to zupełności wiedzy o takim układzie. Z zupełnej znajomości «czarnej skrzynki» (np. dowolnego zdeterminowanego układu fizycznego, o skończonej liczbie stanów i reguł działania), wynika, że ma ona znaną reprezentację kanoniczną i można przewidzieć skutki jej działania i stany przyszłe. W wypadku układów bardziej złożonych, takich jak organizmy biologiczne, z niepełnej ich znajomości można tylko ekstrapolować ich własności. Wraz ze wzrostem złożoności układu metoda jego opisu (modelowania) staje się coraz bardziej skomplikowana.

Operacyjny poziom działania człowieka i maszyny

Odpowiedź na pytanie, czy maszyna cyfrowa jest właściwym modelem podmiotu ludzkiego, jest następująca. Człowiek wykonując daną czynność umysłową, np. liczenie, realizuje ją przy pomocy operacji logiczno-matematycznych, które są podstawą (formą) jego psychicznego procesu myślenia; podobnie też i maszyna matematyczna, wykonując czynność obliczania, realizuje ją za pomocą podobnych operacji. Człowiek i maszyna działają na poziomie operacyjnym podobnie — realizują uniwersalne reguły logiczne, ujęte w postaci praw logiki dwuwartościowej. Obie maszyny «liczą» przetwarzając informację w sposób nieciągły (dyskretny, skokowy), konwencjonalny i skończony (Bolter 1990, s. 124). Operacje te są operacjami mechanicznymi i mogą być przedstawione w wypadku komputera — za pomocą analizy numerycznej, tj. przez zastosowanie rachunku o skończonej liczbie kroków. Rachunek na operacjach (instrukcjach) maszyny może być przedstawiony algebraicznie, dzięki czemu algebra jest modelem czynności zarówno myślenia człowieka, jak i operacji komputera cyfrowego.

Operacje logiczne człowieka są ewolucyjnym produktem jego działania i poznania, operacje logiczne jakiegokolwiek liczącej maszyny są zaś zadane przez jej konstruktora i mogą podlegać tylko częściowej «ewolucji» — samoprogramowaniu (o czym świadczy rozwijająca się niezwykle szybko poddyscyplina teorii sztucznej inteligencji — teoria sieci neuronowych). Mogą one być zapisane w określonym języku, w wypadku człowieka — w rachunku zdań, a w wypadku komputera — w zero-jedynkowym zapisie przetwarzanych symboli.

Każdy system znakowy, będący zapisem (reprezentacją) dowolnej formalnej operacji, musi być, ze względu na skuteczność swojego działania (jest to tak samo ważne w

wypadku człowieka, jak i w wypadku maszyny) zrealizowany w przedmiotowej, fizycznej sferze, musi być ucieleśniony lub — by użyć języka komputerowego — «implementowany w *hardware*». Fizycznym podłożem procesów myślowych człowieka są elektrochemiczne procesy obwodowego systemu nerwowego, głównie zaś mózgu. Neurologiczna podstawa procesów umysłowych jest faktem, który musi uwzględniać każda analiza procesu myślenia. Z racji złożoności, wielopoziomowości procesów nerwowo-mózgowych i niedostatecznej jeszcze ich znajomości — każda z nauk zajmujących się tymi zjawiskami musi określić korelacje pomiędzy treściową zawartością reguł i operacji logicznych, dotyczących poszczególnych czynności poznawczych, a ich neurologiczną reprezentacją. Na tle trudności w realizacji tego zadania w odniesieniu do człowieka — jego realizacja w wypadku komputera cyfrowego jest o wiele łatwiejsza. Zapisanie danej operacji logicznej w języku wysokiego poziomu i następnie jego realizacja w językach programowania i w języku maszynowym — pozwala na jednoznaczne skorelowanie ich z konkretnym stanem fizycznym (impulsami prądu) odpowiedniego *hardware*'u. (Istotą komputera cyfrowego jest przy tym to, że odpowiednia reguła czy program może być zrealizowany przy użyciu dowolnego, wymiennego sprzętu. Takiej własności nie ma ludzki mózg.)

Różnice w korelacji znakowej sfery umysłu i programu maszyny z ich fizyczną realizacją (wzajemne przyporządkowanie operacji logicznych odpowiednim stanom rzeczy jest zawsze pewnym uproszczeniem) nie są zresztą najważniejsze. Istotniejsza jest kwestia semantycznej realizacji operacji znakowych i zasadnicze w tym względzie różnice między umysłem a programem komputerowym. Jest to ważny problem przy rozstrzygnięciu różnic i podobieństw pomiędzy człowiekiem i maszyną. (Zagadnienie to rozważałem gdzie indziej; por. Hetmański 1995.)

Dla analizy podmiotowości człowieka i maszyny (komputera) jako układów cybernetycznych ważne jest rozpatrzenie działania operacji logicznych w obrębie samego układu, niezależnie od ich semantycznego odniesienia, a tylko z uwzględnieniem syntaktycznej organizacji i realizacji. Jeśli weźmiemy pod uwagę jedną tylko operację wykonywaną przez człowieka i maszyny, podstawową dla ich zestawienia i najczęściej analizowaną — mianowicie liczenie — to zauważymy odmienną jej realizację. Jest to odmienną strategię planowania i realizacji poszczególnych składowych operacji liczenia. Człowiek wykonując taką samą operację arytmetyczną jak maszyna — realizuje strategię ograniczającą liczbę poszczególnych działań przy jednoczesnym ich względnym skomplikowaniu i rozbudowaniu. Maszyna działa według strategii odmiennej — mnoży liczbę działań maksymalnie uproszczonych. Jest to różnica związana z różnicą czasu wykonywania tej samej operacji. Maszynowe operowanie znakami realizuje się na fizycznym podłożu elektronicznych zjawisk przebiegających w środowisku odmiennym niż fizyko-chemiczne podłoże potencjałów czynnościowych systemu nerwowego człowieka. Niezwykła prędkość pierwszych procesów jest rekompensowana w wypadku człowieka przez odmienną strategię realizacji tych samych operacji. W obydwu wypadkach podstawą wykonywania operacji jest realizacja algorytmu. Powstaje

jednak pytanie, czy algorytmizacja operacji w wypadku maszyny cyfrowej jest tej samej miary i stopnia, co w wypadku człowieka?

Algorytmizacja myślenia

Algorytmizacja działania komputerów i ludzi jest koronnym dowodem teoretyków sztucznej inteligencji na rzecz prawdziwości tezy o zasadniczym podobieństwie komputerów i ludzi. Teza ta budzi jednak duże wątpliwości. Algorytm jest systematyczną procedurą wykonania danego zadania w skończonej liczbie kroków, którą można ponadto opisać w skończony sposób. Tak rozumiana procedura jest stosowana zarówno w sytuacjach problemowych (np. dowodzenia) w naukach formalnych, jak i w sytuacjach rozwiązywania problemów życiowych. Praktyka nauki i życia stała się zagadnieniem teoretycznym z chwilą postawienia przez Hilberta pytania o możliwość sformułowania ogólnego algorytmu rozstrzygnięcia problemów matematycznych w obrębie względnie szerokiej i dobrze zdefiniowanej ich klasy. W ramach dowolnego, sformalizowanego działu matematyki należało sporządzić dostatecznie bogatą listę aksjomatów i reguł wnioskowania, tak aby można było wyrazić w nim wszystkie stosowane tam metody poprawnego rozumowania (dowodzenia). System takich aksjomatów i reguł miałby być niesprzeczny i zupełny, co znaczy, że jeśli się nim posłużymy, to będziemy mogli określić wartość logiczną dowolnego zdania, możliwego do sformułowania w języku tego systemu, w skończonej procedurze dowodzenia. Dowodzenie w obrębie takiego systemu miałoby charakter maszynowy, tj. musiałby istnieć algorytm takiej operacji. Niemożliwość realizacji programu formalizmu matematycznego Hilberta wykazał Gödel w swojej tezie o niezupełności niesprzecznego systemu aksjomatycznego, odpowiednio bogatego. Teza o niezupełności podważa również założenie o algorytmizacji procedur dowodzenia matematycznego i logicznego. W nowym świetle stawia to główne założenie teorii sztucznej inteligencji o podobieństwie, co najmniej na poziomie logiczno-matematycznym, procedur komputerowych i ludzkich.

Spór wokół podstaw aksjomatycznych systemów formalnych zrodził istotne pytanie o możliwość i istotę ogólnej procedury mechanicznej dowodzenia, która w zasadzie pozwoliłaby rozwiązać wszystkie problemy matematyczne. Pojęcie „procedury mechanicznej” zakłada wykonanie przez daną maszynę — bez względu na jej tworzywo i budowę, w skończonej liczbie dyskretnych, wewnętrznych stanów, przy użyciu nieograniczonej pamięci zewnętrznej — określonej operacji logiczno-matematycznej, oraz podanie odpowiedzi na zadane pytanie. Teoretycznym obrazem takiej maszyny jest maszyna Turinga, będąca podstawą konstrukcji każdego komputera cyfrowego. Działanie maszyny Turinga jest całkowicie zdeterminowane jej stanami wewnętrznymi i danymi zewnętrznymi. Można założyć, iż maszyna Turinga działa w otoczeniu reprezentowanym przez nieskończoną taśmę i w skończonej liczbie kroków opisywanych zawsze przez skończoną procedurę. Jest ona wówczas modelem dowolnego układu

praktyczno-poznawczego, działającego w zróżnicowanym środowisku. Powstaje ważne pytanie (Penrose 1995, s. 66), czy pewne układy fizyczne — np. mózgi ludzkie, które podlegają prawom fizycznym na równi z innymi układami — są w stanie wykonać dokładnie te same operacje logiczno-matematyczne (szerzej, poznawcze), co maszyna Turinga. Czy ich możliwości są takie same? Innymi słowy — czy one też działają przy użyciu algorytmów? Łączy się to z kolejnym pytaniem, czy algorytm, zarówno w wypadku maszyny, jak i człowieka, jest podstawą rozwiązywania wszystkich problemów, które są im dane (zadane)?

Odpowiedzi na to ostatnie pytanie udzielił sam Turing, który — paradoksalnie wobec przyjętych założeń — wykazał, że istnieją takie problemy matematyczne, dla których nie ma algorytmu, i których tym samym maszyna Turinga nie jest w stanie rozwiązać. Prawdziwość danych zdań systemu nie może być rozstrzygnięta na podstawie dowodu; więcej, samo stwierdzenie tego faktu też nie ma dowodu w postaci skończonej procedury. Przekładając tę paradoksalną własność, o której mówił Gödel, na inny język, Turing stwierdził, że maszyna Turinga nie jest w stanie dokonać obliczenia (podać skończonego dowodu) pewnej klasy problemów, odwołując się do właściwych jej zasad działania. Niektóre zadania maszyna ta wykonuje łatwo w tym sensie, iż możliwe jest podanie algorytmu zakończenia przez nią pracy na podstawie rekurencyjnej procedury dowodzenia; istnieje zatem algorytm wykonywania przez nią danego algorytmu. Istnieje jednakże całkiem okazała klasa zadań, dla których nie ma uniwersalnego algorytmu, pozwalającego sprawdzić, czy dowolna maszyna Turinga zakończy pracę. Jeśli nie ma algorytmu określającego zakończenie pracy maszyny matematycznej, to tym samym nie ma uniwersalnego algorytmu, który można by zastosować do wszystkich zadań matematycznych. Algorytmizacja w matematyce ma swoje istotne ograniczenia, choć jest procedurą szeroko stosowaną w wielu jej działach. W ten sposób Turing potwierdził prawdziwość tezy Gödla. Postawił przy tym ważne zagadnienie istoty procedury dowodzenia i obliczania oraz podobieństw i różnic tych czynności wykonywanych przez maszynę i człowieka.

Przykładem niealgorytmicznego charakteru procedur poznawczych u człowieka jest «intuicja matematyczna», «wgląd» (Penrose 1995 s. 456), tj. zdolność stwierdzenia prawdziwości danego twierdzenia bez podania jego dowodu, tj. bez odwołania się do algorytmu tej czynności. Dla wielu teoretyków i krytyków teorii sztucznej inteligencji fakty intuicyjnego poznania matematycznego są potwierdzeniem ich tezy, że nie ma podobieństwa pomiędzy działaniem umysłu człowieka a działaniem komputera, i pierwszy ma zasadniczą przewagę nad drugim (Scriven 1961, s. 125; Lucas 1961, s. 113). Dobrym wyrazem tej przewagi ma być to, że człowiek radzi sobie z problemami matematycznymi, z którymi maszyna matematyczna nie może sobie poradzić: np. z wykazaniem niesprzeczności systemu aksjomatycznego (teza Gödla). Ze względu na tę istotną różnicę nie można komputera i człowieka uznać za równorzędne podmioty poznające. Poznanie człowieka odznacza się takimi własnościami, jak intuicja, bez-

pośredniość wglądu i zupełność, które nie są procedurami poddającymi się mechanizacji i algorytmizacji.

Inteligencja sztuczna a inteligencja naturalna

Zagadnienie podmiotowości maszyn cyfrowych w ramach teorii sztucznej inteligencji jest sformułowane niewłaściwie i przez to dwuznaczne. Ta dwuznaczność ujawnia się już w nazwie samej dziedziny badań. Sugeruje ona badania inteligencji nazwanej „sztuczną” w znaczeniu skonstruowanej, odtworzonej, powielonej w maszynie na wzór inteligencji naturalnej, którą ma tylko człowiek. Jest to sugestia błędna. Po pierwsze, człowiek nie ma «naturalnej» inteligencji, gdyż rozwój tej dyspozycji poznawczej zarówno w perspektywie ontogenetycznej, jak i filogenetycznej, odbywa się na drodze zapośredniczania działania i myślenia przez obiekty świata zewnętrznego, narzędzia, sferę znakowo-symboliczną — a więc dziedzinę w dużej części «sztuczną». Inteligencja człowieka jest «naturalna» w stopniu tylko minimalnym: na poziomie odruchów i instynktów (lecz czy wówczas jest to inteligencja?); w pełni kształtuje się ona i wyraża na poziomie «sztucznym» — semantycznym, technicznym, kulturowym. Nie mając innej inteligencji niż «sztuczna», człowiek nie ma w stworzonych przez siebie komputerach (maszynach w ogóle) żadnej konkurencji. Są to tylko artefakty obdarzone przez niego namiastką ludzkiej cechy gatunkowej. Używając stylistyki teorii sztucznej inteligencji (zwłaszcza Minsky’ego) powiemy, że człowiek nie ma po stronie komputerów konkurentów, od których musiałby się czegośkolwiek uczyć lub być przez nie traktowany z pobłażaniem i wyniosłością.

Niemniej jednak problem wzrastającej roli «inteligentnych» narzędzi w życiu ich twórcy jest ważny i powinien stać się przedmiotem poważnego namysłu. Chodzi w szczególności o dokładniejszą analizę pojęcia „inteligencji”.

Założeniem teorii sztucznej inteligencji jest to, że komputer wykazuje się inteligencją, kiedy spełnia jedną z poznawczych czynności podmiotu ludzkiego, o której się powszechnie sądzi, że wymaga tej dyspozycji (Minsky 1968, s. V). Chodzi przede wszystkim o postrzeganie, rozpoznawanie dźwięków, orientację w prostym otoczeniu, imitację ruchu (są to dziedziny znaczących, trzeba przyznać, osiągnięć konstrukcyjnych), podstawowe wnioskowanie. Sformalizowanie którejsz z nich, jej wymodelowanie i implementacja na sprzęcie komputerowym oraz sprawdzenie (zweryfikowanie), czy wyniki osiągnięte w tych dziedzinach przez komputer są podobne do wyników otrzymywanych przez człowieka, ma wystarczać do przypisania danemu komputerowi zachowania inteligentnego. U podstaw tego założenia leży przekonanie, że podmiotowość poznawcza maszyny cyfrowej osiągnięta jest już w momencie realizacji (odtworzenia) sztucznymi środkami jednej z czynności inteligentnego działania człowieka. Seryjne i wypreparowane z kontekstu elementarne czynności, realizowane na coraz doskonalszym sprzęcie i przy użyciu rozbudowywanego oprogramowania, jest przez teoretyków sztucznej inteligencji klasyfikowane jako wyróżnik komputerowej podmiotowości. Od-

tworzenie i implementacja na dowolnym sprzęcie określonej (w istocie rzeczy — spreparowanej) pojedynczej czynności człowieka jest uznawane za wystarczające spełnienie warunków podmiotowości; część wystarcza za całość. Jest to zastanawiające uproszczenie i zbanalizowanie zagadnienia podmiotowości w ogóle.

W świetle bogatej literatury ogólnofilozoficznej, teoriopoznawczej i psychologicznej za główne kryterium podmiotowości człowieka należy uznać (1) silną współzależność działania praktycznego i poznawczego. Współzależność obu działań jest wzbogacona przez (2) upośrednianie ludzkiego działania i poznania przedmiotową sferą znakowo-narzędziową, która rozszerza jego zakres i skuteczność. Dopełnieniem podmiotowości jednostki, jej «różnicą gatunkową», jest (3) myślenie pojęciowe nadbudowane nad poznaniem zmysłowym oraz (4) współdziałanie z innymi podmiotami (Cackowski 1979, s. 29-89). Kolejność tych kryteriów odzwierciedla ich merytoryczną ważność oraz chronologiczny (w ontogenezie nieco odmienny niż w filogenezie) porządek. Współwystępowanie i przede wszystkim współzależność tych czynników jest zasadniczym kryterium podmiotowości człowieka.

Czy kryteria te (wszystkie razem, a nie tylko niektóre) spełnia «podmiotowość» maszyny cyfrowej? Na to pytanie nie można odpowiedzieć jednoznacznie twierdząco. Teoria sztucznej inteligencji koncentruje się na modelowaniu i symulacji poznania zmysłowego, w mniejszym stopniu pojęciowego, spełnienie zaś pozostałych kryteriów podmiotowości jest poza zasięgiem jej obecnych możliwości. Wprawdzie w jej ramach rozwija się robotyka, która zakłada odtworzenie drogi ewolucji naturalnej i koncentruje się na senso-motorycznym aspekcie działania robotów — inteligentny umysł to 'umysł w ruchu', jak twierdzi Moravec (1988 s. 6-50) — ale daleko jej do odtworzenia choćby części ewolucji umysłu. (Teoretycy robotyki zdają się przeoczać fakt, iż ewolucja inteligentnych maszyn nie jest samoistnym procesem, lecz częścią, kulminacją ewolucji kulturowo-technicznej człowieka — a nawet to, że w istocie rzeczy ewolucja inteligencji, szczególnej cechy gatunkowej człowieka, nie jest ewolucją *stricte* biologiczną.) Poza możliwością realizacji pozostaje zatem integracja poszczególnych czynności poznawczych w ramach jednorodnego, sztucznego podmiotu. Zróżnicowanie poszczególnych funkcji poznawczych człowieka i ich zwielokrotnienie może być łatwo osiągnięte przez maszynę (zarówno analogową, jak i cyfrową), podobnie jak samoregulacja i samokontrola (np. homeostat) czy orientacja w otoczeniu. Każda z tych czynności oddzielnie a nawet częściowa ich suma nie jest jednak podstawą do zakładania integracji sztucznie skonstruowanego podmiotu.

Automatyzm maszyny jako warunek jej podmiotowości

O podmiotowości maszyny można mówić w wypadku jej automatyzmu. Nie każda maszyna jest jednak automatem, tj. nie każda odznacza się wysokim stopniem złożoności, powiązaniem z otoczeniem — w tym z człowiekiem jako konstruktorem i

użytkownikiem maszyny — oraz wzrastającym stopniem samodzielności. Podmiotowość ta nie ma jednakże bezwzględnego charakteru.

Działanie każdej maszyny wymaga dwóch rodzajów energii: zasilającej i sterującej (Cackowski 1989, s. 511-524; de Latil 1958, s. 43). Oba rodzaje energii zapewniają danemu układowi realizację określonego mechanizmu, są podstawą jego działania. W wielu wypadkach energia zasilania stanowi jednocześnie czynnik sterujący działaniem układu. Wyróżnienie w danym podmiocie jego podukładu sterującego i zasilanie go przez dodatkową energię sterującą dowodzi jego złożoności, wzrastająca rola układu sterującego jest zaś objawem wzrostu podmiotowości układu. Automatyizacja mechanizmu działania danego układu (człowieka czy maszyny) następuje wraz ze sterowaniem przez ten układ własną energią sterowania, zmianami swego działania w czasie i przestrzeni oraz przekazywaniem informacji swoim organom wykonawczym (de Latil, s. 50-51).

Automatyzm każdego — naturalnego czy sztucznego — układu, jest stopniowalny i można w nim wyróżnić odpowiednie poziomy. Pierwszym stopniem automatyzmu odznacza się mechanizm, którego działanie się nie zmienia lub zmienia się przypadkowo; odpowiednikiem takiego mechanizmu w mózgu jest łuk odruchowy. Stopniem drugim automatyzmu jest mechanizm determinowania zmian w działaniu układu, powolnego uwalniania się od przypadkowości za pomocą zaczątków programu (wewnętrznego elementu maszyny), będącego «rozdzielaczem» energii sterowania. Zmiana działania danej maszyny, dokonująca się bez udziału człowieka, lecz zgodnie z kierunkiem przez niego wyznaczonym, stanowi trzeci stopień automatyzmu; osiągnęta jest ona na etapie techniki elektronicznej, a jej odpowiednikiem po stronie organizmu ludzkiego jest działanie śródmózgowia. Kolejny — czwarty — stopień automatyzmu jest organizacją działania układu przy użyciu mechanizmu sprzężenia zwrotnego, a więc korygującego wpływ danego efektora na działanie całego układu; użytkownik maszyny jest dla niej na tym poziomie «przypadkowością», a nie czynnikiem, jak dotychczas, pełnego zdeterminowania. Poziom jeszcze bardziej rozbudowanej organizacji, w której ma miejsce kontrola znacznej części czynników działania dzięki sprzężeniu zwrotnemu, jest stopniem piątym automatyzacji; jego przykładem jest homeostat skonstruowany przez Ashby'ego — maszyna działająca zasadniczo bez programu i uzyskująca równowagę przez przypadek, lecz mimo to wykazująca zdolność do jej utrzymania a nawet uczenia się i stosowania wypracowanych strategii działania, a nie tylko podlegania wpływom ze strony otoczenia. Szóstym stopniem automatyzmu odznaczają się układy zdolne do zmiany celu przez każdy efektor, jeśli wymaga tego utrzymanie ogólnej równowagi układu; układ taki (multistat) dostosowuje zarówno działanie do środków, jak i środki do celów, przy pełnej swobodzie wyboru sposobów zachowania; nieodzowna jest tu wszakże pamięć (zbędna jeszcze na poziomie poprzednim), jaką odznacza się np. człowiek.

Przedstawiona powyżej wzrastająca automatyzacja układów sztucznych i naturalnych, polega na organizacji zakładającej samosterowanie, zmianę i modyfikowanie

działania w trakcie jego realizacji oraz poszerzanie i wzbogacanie środowiska działania. Na pierwszy rzut oka maszyna jest układem, którego wzrastająca podmiotowość polega na uniezależnianiu się od człowieka, jej konstruktora, jak np. w wypadku możliwości samoprogramowania się przez sieci neuronowe. Wzrastająca podmiotowość człowieka jest natomiast procesem poszerzania działalności praktycznej i poznawczej przy użyciu rozbudowanej i złożonej sfery techniki — narzędzi i instrumentów (prostego, niesamodzielnego przedłużenia ręki) oraz maszyn o wzrastającej automatyzacji. Nie są to procesy rozbieżne czy przeciwstawne. Błędem jest rozpatrywanie maszyn (w tym zwłaszcza komputerów) jako samodzielnie działających i poznających układów, izolowanych od człowieka. Najbardziej zaawansowany technicznie układ jest tylko tworem swojego konstruktora i w podstawowych funkcjach realizuje jego program, nawet jeśli jest w stanie znacznie go modyfikować i rozwijać. Podmiotowość maszyny i podmiotowość człowieka jest nie tyle cechą ich samoistnych, niezależnych własności i paramaterów, ile raczej wynikiem ich wzajemnego sprzężenia w realizacji konkretnych celów praktyczno-poznawczych.

Mamy więc do czynienia ze szczególnym podmiotem, jakim jest układ człowiek-maszyna (nie Kartezjański czy La Mettrie'owski «człowiek-maszyna», ani też «sztuczna inteligencja» z laboratoriów MIT). Oba jej człony wykazują we wzajemnych oddziaływaniach wzrastającą podmiotowość, jako funkcję tych oddziaływań. Podmiotowość ta ma cechy procesu ewolucyjnego. Poszerzający za pomocą techniki (automatów, robotów) swoje możliwości działania i poznania człowiek oraz samoprojektująca się maszyna, złączone w dynamicznym układzie, wykazują rodzaj podmiotowości bogatszej niż podmiotowość każdego z członów tego układu wziętego oddzielnie.

Bibliografia

- Arbib M.A., *Mózg, maszyna, matematyka*, PWN, Warszawa 1968
 Arbib M.A., *Mózg i jego modele*, PWN, Warszawa 1977
 Ashby W.R., *Wstęp do cybernetyki*, PWN, Warszawa 1961
 Ashby W.R., *Design for A Brain*, Chapman and Hall, London 1954
 Bolter J.D., *Człowiek Turinga. Kultura Zachodu w wieku komputera*, PIW, Warszawa 1990
 Briukow B.W. Geller J.F., *Cybernetyka w naukach humanistycznych*, Ossolineum, Wrocław-Warszawa 1983
 Cackowski Z., *Człowiek jako podmiot działania praktycznego i poznawczego*, KiW, Warszawa 1979
 Cackowski Z., *Zasadnicze zagadnienia filozofii*, KiW, Warszawa 1989
 Hetmański M., „Epistemologiczne aspekty nauk poznawczych i sztucznej inteligencji”, [w:] A. Bronk (red.) *Filozofować dziś. Z badań nad filozofią najnowszą*, Lublin, TN KUL, 1995, s. 211-228
 Jackendoff R., *Consciousness and the Computational Mind*, MIT Press, Cambridge Mass. 1987
 Latil P. de, *Sztuczne myślenie. Wstęp do cybernetyki*, PWT, Warszawa 1958
 Lucas J.R., „Minds, Machines and Gödel”, *The Philosophy*, vol. XXXVI (1961), nr 137, s. 112-127
 Minsky M., (wyd.) *Semantic Information Processing*, MIT Press, Cambridge Mass. 1968
 Moravec H., *Mind Children. The Future of Robot and Human Intelligence*, Harvard University Press, Cambridge Mass. 1988
 Neumann J. von, *Maszyna matematyczna i mózg ludzki*, PWN, Warszawa 1963
 Penrose R., *Nowy umysł cesarza. O komputerach, umyśle i prawach fizyki*, PWN, Warszawa 1995

Scriven M., „The Compleat Robot: A Prolegomena to Androidology”, [w:] S. Hook (red.), *Dimensions of Mind*, Collier Books, New York 1960, s. 113-133

Turing A., „Maszyna licząca a inteligencja”, [w:] B. Chwedeńczuk (red.), *Filozofia umysłu*, Wydawnictwo Spacja, Warszawa 1995, s. 271-300

Wiener N., *Cybernetyka, czyli sterowanie i komunikacja w zwierzęciu i maszynie*, PWN, Warszawa 1971