

# Redukcjonizm, kognitywistyka, transhumanizm\*

Jerzy Bobryk\*\*

Instytut Psychologii PAN, Warszawa

REDUCTIONISM, COGNITIVE SCIENCE, TRANSHUMANISM

The word „reductionism“ is commonly used in the methodology of human sciences; however this word is ambiguous: “Reductionism is a position that holds that theories or things of one sort, can exhaustively account for the theories or things of another sort” [The MIT Encyclopedia of Cognitive Science]. Consequently, there are two main basic meanings of the notion of “reductionism”: [1] reductionism understood as theoretical reductionism, and [2] reductionism understood as ontological reductionism. In cognitive science we have “the computational theory of mind”, which is based on the assumption (or the doctrine) that all mental states and processes are computational processes. “Computation” is usually understood as the process of manipulation of physical (token) symbols according to previously set rules. The computational theory of mind, which is a theoretical reductionism, may be transformed into an ontological reductionism. Transhumanism is an international philosophical movement associated with the theory (or the belief) that human species can evolve beyond its current biological limitations. A speculative technique commonly known as “mind uploading” – namely a transfer of the human mind into computer software and/or human brain into electronic hardware – has been considered as a way to human superintelligence and immortality.

**Key words:** cognitive science, mind, reductionism, transhumanism

## WSTĘP

Niniejszy tekst poświęcony jest niektórym istotnym, bo mającym charakter kulturowo-cywilizacyjny, następstwom przyjęcia postawy, którą nazywamy „redukcjonizmem”. „Redukcjonizm” nie tylko zakorzenił się w metodologii nauki, ale także przeniknął do archetypów myślenia potocznego, schematów narracji w sztuce i kulturze masowej. Będę tu zatem bronił stanowiska, że w pewnych warunkach skutki przyjęcia postawy redukjonistycznej mogą wykraczać daleko poza obszar nauki. Inaczej mówiąc, będę starał się pokazać przykłady, w których skutki postawy redukjonistycznej widać zarówno w różnych sferach praktycznego działania jak i sposobach naukowego i codziennego myślenia. Postaram się także odpowiedzieć na pytanie, jakie to warunki muszą być spełnione, by doszło do tak silnego oddziaływania postawy redukjonistycznej oraz jakie mogą być skutki takiego rozszerzenia redukjonizmu.

## CO TO JEST REDUKCJONIZM

Chociaż „redukcjonizm” jest słowem często używanym w metodologii nauki i w niesystematycznej potocznej metanaukowej refleksji, to jednak denotacja tego terminu jest niejednoznaczna, a jego definicje czy określenia nie

zawsze są całkiem jasne. I tak, na przykład, w *Słowniku filozoficznym* (Lacey, 1999, s. 252) czytamy:

„Redukcjonizm – tendencja do redukowania pewnych pojęć, zarówno codziennych jak obiekt fizyczny, jak i teoretycznych w naukach szczegółowych, takich jak elektron, do pojęć rzekomo prostszych lub bardziej podstawowych albo łatwiej uchwytnych empirycznie, np. ktoś mógłby utrzymywać, że można się obyć bez słowa „elektron” i mówić tylko o śladach pary w komorze pęcherzykowej [...] Redukcjonizm jest szczególną cechą fenomenalistów i innych empirystów oraz – w filozofii nauki – pozytywistów”.

Zaś w *Słowniku filozofii* (Andruszkiewicz, 2004, s. 433) mamy:

„Redukcjonizm [...] koncepcja opierająca się na metodzie wyjaśniania czegoś niezrozumiałego, przez sprowadzanie do tego, co jest pod danym względem zrozumiałe”.

Z kolei Adam Urbanek (1987, s. 564) określa redukjonizm jako:

„postulat filozoficzny i metodologiczny dowodzący, że obiekty i procesy złożone lub odnoszące się do nich teorie mogą być sprowadzone (zredukowane) do podstawowych i elementarnych zjawisk lub też ogólnych pojęć teoretycznych...”

\*\* Korespondencję dotyczącą artykułu można kierować na adres: Jerzy Bobryk, Instytut Psychologii PAN, ul. Jaracza 1, 00-378 Warszawa. jkbobryk@op.pl

\* Praca naukowa finansowana w ramach programu Ministra Nauki i Szkolnictwa Wyższego pod nazwą „Narodowy Program Rozwoju Humanistyki” w latach 2012-2016.



Podobnie widzi redukcjonizm Artur Reber (2000, s. 616):

„Redukcjonizm w szerokim sensie: pogląd filozoficzny, który utrzymuje, że złożone zjawiska można najlepiej zrozumieć przez analizę składnikową, rozkładając je na proste podstawowe elementy. Istotą zjawiska redukcjonizmu stanowi zatem przekonanie, że pełniejszy wgląd w się przez schodzenie z analiza na głębszy [podr. J.B.], bardziej podstawowy poziom...”

*Wielka Encyklopedia PWN* (2004) odróżnia dwa znaczenia terminu, podając:

„Redukcjonizm:

1. w ogólnym znaczeniu – pogląd, zgodnie z którym zjawiska i procesy złożone oraz rządzące nimi swowiste [...] prawa można wyjaśnić na podstawie zjawisk i procesów prostszych oraz odpowiadających im mniej skomplikowanych praw [...]
2. w ściślejszym znaczeniu – pogląd, który głosi, że złożone zjawiska, procesy i prawa można sprowadzić do prostszych (a nawet elementarnych) na tej zasadzie, iż różnice między nimi są jedynie ilościowe, a każda właściwość dowolnie skomplikowanego układu jest warunkowana wyłącznie właściwościami jego elementów, części składowych [...]; redukcjonizm w tym znaczeniu jest zasadniczo sprzeczny z rezultatami współczesnej nauki i jej tendencjami rozwojowymi.”

Jak widzimy określenia czym jest redukcjonizm balansują między jego rozumieniem epistemologiczno-metodologicznym, zakładającym, że jest to jedynie nazwa arbitralnie wybranego (naukowego lub potocznego) opisu zjawisk; a rozumieniem ontologicznym, w którym twierdzi się coś o istocie czy naturze zjawisk opisywanych lub badanych.

## RODZAJE REDUKCJONIZMU

Jak to powiedziano wyżej, używając terminu „redukcjonizm” najczęściej mamy na myśli jedną z trzech interpretacji tego terminu. Często jest to redukcjonizm ontologiczny, jednak tylko nieco mniej popularne są dwa pozostałe (zresztą nie zawsze dobrze odróżniane od siebie) znaczenia tego terminu, czyli redukcjonizm metodologiczny i redukcjonizm teoretyczny. *Encyklopedia Filozofii* (Honderich, 1998, s. 780) w haśle „redukcjonizm” stwierdza:

„Redukcjonizm ontologiczny odnosi się do przeświadczenia, że cała rzeczywistość składa się z niewielkiej liczby istotności czy substancji [...] Redukcjonizm metodologiczny twierdzi w sferze nauki, że „małe jest piękne”, czyli, że najlepszą strategią naukową zawsze jest próba formułowania wyjaśnień w terminach coraz mniejszych istotności [...] Redukcjonizm teoretyczny stanowi pytanie o relacje między następującymi po sobie teoriami”.

Warto tu jednak podkreślić, że podane wyżej rozróżnienia nie zawsze są do końca klarowne. Może dlatego

autorzy różnych definicji redukcjonizmu bardzo często skupiają się na funkcjach i celach redukcjonizmu. *Oksfordzki słownik filozoficzny* (Blackburn, 1997, s. 340) stwierdza:

„Redukcjonizm utrzymuje, że fakty czy przedmioty, które wydają się niezbędne, jeśli twierdzenia jakiejś dyscypliny mają okazać się prawdziwe, można zastąpić jakimiś innymi faktami czy przedmiotami [podr. J.B.]. Redukcjonizm stanowi jedną z odpowiedzi na pytanie o związek pomiędzy różnymi naukami. Ktoś mógłby opowiedzieć się za redukcją biologii do chemii [...]; ktoś mógłby również twierdzić, że chemia redukuje się do fizyki [...]. Ściśle biorąc, redukcjonizm nie jest rodzajem sceptycyzmu (nie wyklucza bowiem, że twierdzenia należące do redukowanej dyscypliny są prawdziwe i że możemy o tym wiedzieć; w istocie jednym z typowych celów redukcji jest pokazanie, w jaki sposób się do niej dochodzi).”

Trzeba tu chyba nadmienić, że poza nie zawsze starannie odróżnianym od siebie „redukcjonizmem ontologicznym”, „redukcjonizmem metodologicznym” i „redukcjonizmem teorii” mówi się niekiedy o „redukcjonizmie cech”. Znanym przykładem tego ostatniego jest redukcja takiej cechy, jaką jest temperatura ciała, do innej cechy, czyli szybkości poruszania się jego cząstek.

Nie wdając się tu w żadne próby podsumowania problemu redukcjonizmu – na które nie ma miejsca w pojedynczym artykule – chcę tu podkreślić, że każda forma redukcjonistycznego opisu zjawisk empirycznych ma skutki dwojakiego rodzaju:

- przedstawia nam uproszczony obraz badanych zjawisk;
- determinuje to, jak opisane redukcjonistycznie zjawiska są traktowane w praktyce<sup>1</sup>.

Następne rozdziały niniejszego tekstu pokazują właśnie skutki drugiego typu.

## REDUKCJONIZM W PSYCHOLOGII POZNAWCZEJ I KOGNITYWISTYCE

Najbardziej znaną formą redukcjonizmu w psychologii był behawioryzm, często rozumiany jako redukcjonizm ontologiczny – redukujący procesy psychiczne/świadomość do reakcji fizycznych – chociaż klasycy behawioryzmu często nie negowali istnienia świadomości i jedynie twierdzili, że nie można jej naukowo badać. Wiele nieporozumień budził i budzi też równie popularny i bardziej współczesny redukcjonizm neurofizjologiczny. Ciągłe trwają spory dotyczące relacji między zjawiskami neurofizjologicznymi i zjawiskami psychicznymi:

- 1) Można twierdzić, że te pierwsze są wprawdzie przyczyną i warunkiem koniecznym zaistnienia tych drugich, jednak zjawiska psychiczne stanowią swobodną odrębną i emergentną sferę rzeczywistości.

<sup>1</sup> Jeśli opiszemy chorobę psychiczną jako skutek zmian w mózgu chorego, raczej nie będziemy stosować psychoterapii jako jedynej metody leczenia.

2) Albo też można uznać, że – z ontologicznego punktu widzenia – zjawiska neurofizjologiczne i zjawiska psychiczne, to dokładnie jedno i to samo.

Stanowisko drugie, zwane teorią identyczności (procesów umysłowych i procesów mózgowych) wydaje się być najbardziej czystą i najczęściej przyjmowaną w psychologii (w myśleniu potocznym) odmianą redukcjonizmu. Ostatnio spór dotyczący relacji pomiędzy zjawiskami psychicznymi i procesami mózgowymi, w związku z pojawieniem się *cognitive science* (CS), przybrał nową formę, gdyż *cognitive science* sama odpowiedzialna jest za utrwalenie nowej formy redukcjonizmu, zwanej „redukcjonizmem maszynowym”. *Cognitive science* (CS) miała być dyscypliną naukową integrującą wcześniejsze osiągnięcia wszelkich innych nauk badających świadomość i procesy poznawcze. *Cognitive science* wprowadziła do języka naukowego i potocznego tak zwaną metaforę komputerową i związaną z nią „obliczeniową teorię umysłu”, co połączyło dość ściśle psychologię poznawczą i filozofię umysłu z teorią i praktyką sztucznej inteligencji. Obliczeniowa teoria umysłu przyjmuje założenie, że wszelkie nawet najbardziej złożone procesy poznawcze dadzą się zredukować do „obliczania”, które za A. Turingiem definiowane jest jako manipulowanie (czyli zapisywanie, usuwanie) fizycznie ucieleśnionymi symbolami, według z góry ustalonych reguł. Obliczeniowa teoria umysłu skłoniła badaczy do stosowania programów komputerowych umożliwiających badanie umysłu ludzkiego poprzez komputerowe symulacje procesów umysłowych (a dokładniej, jakoś tam hipotetycznie założonych, a potem przełożonych na programy komputerowe operacji poznawczych). Jeśli takie symulacje komputerowe są skuteczne, to znaczy, jeśli są w stanie wykonać funkcje, które wykonuje ludzki umysł, to można założyć, że takie symulacje są adekwatnym modelem ludzkich procesów umysłowych.

W pierwszych dziesięcioleciach rozwoju CS stworzono szereg programów radzących sobie (gorzej lub lepiej) z zadaniami, z jakimi radzą sobie ludzkie umysły (rozpoznanie wzorów, wyszukiwanie i porządkowanie danych, sprawdzanie poprawności formuł logicznych i wzorów matematycznych, etc.). Spierano się wtedy, jak wiadomo, przede wszystkim o to, czy tego typu programy komputerowe są w stanie zastąpić człowieka w wykonywaniu wszystkich, czy też może tylko niektórych czynności poznawczych. Wielu uczonych (np. Marcin Minsky, Hans Moravec) z kręgu CS uważało i uważa, że już w najbliższej przyszłości, odpowiednio zaprogramowane, maszyny (komputery) będą w stanie wykonać wszystkie wykonywane przez człowieka umysłowe operacje. Najwięksi optymiści uważali, że sztuczna inteligencja nie tylko szybko dorówna inteligencji naturalnej, ale też ostatecznie przekroczy jej możliwości. Inni (np. John Seale, Hubert L. Dreyfus) uważali, że istnieją tu (i zawsze będą istnieć) nieprzekraczalne dla „komputerowych umysłów” granice. Mówiąc inaczej, twierdzili oni, że tak zwana sztuczna inteligencja nigdy nie dorówna inteligencji naturalnej. Chociaż w trakcie rozwoju CS zmieniły się niektóre założenia teorii i praktyki sztucznej inteligencji

– zaczęto w komputerach symulować nie tyle przebieg procesów umysłowych, co strukturę połączeń neuronalnych i działanie mózgu – to, jak do tej pory przynajmniej, spór dotyczący relacji między działaniem ludzkiego umysłu i działaniem komputera nie został rozstrzygnięty ani na poziomie praktycznym, ani na poziomie ustaleń teoretycznych. Inne argumenty pokuszające „redukcjonistyczność” kognitywizmu podają gdzie indziej (Bobryk, 2010, 2011, rozdział 5).

Zupełnie odrębnym zagadnieniem jest nurt badań – także zaliczanych do szeroko pojmowanej kognitywistyki – w których bada się bezpośrednio pracę mózgu. Tu jednak należałoby rozstrzygnąć problem relacji między umysłem i mózgiem, oraz problem lokalizacji. Nie jest to problem rozstrzygnięty we współczesnej nauce. Tym problemem zajmowałem się nieco gdzie indziej (Bobryk 1996, 2001). Jak bardzo jest to problem złożony i ciągle daleki od rozwiązania pokazują, między innymi prace Johna Searle’a (1984, 2004).

## TRANSHUMANIZM

Współcześnie najwięksi optymiści przewidujący przyszłość rozwoju sztucznej inteligencji należą do kręgu „transhumanistów”. Transhumanizm jest obecnie silnie rozwijającym się i zyskującym na popularności filozoficznym (a może nawet kulturowym) ruchem, propagującym przekonanie, że nasz gatunek może i powinien ewoluować przekraczając ograniczenia biologiczne. W latach 80. zeszłego wieku zaczęła się organizować i działać na terenie Uniwersytetu Kalifornijskiego w Los Angeles grupa badaczy nazywających się transhumanistami. Wprawdzie zespoły badawcze i związane z nimi instytucje (katedry uniwersyteckie, instytuty naukowe, czasopiśma, etc.) rozsiane są obecnie w wielu miejscach naszego globu, jednak Kalifornia wydaje się ciągle być centrum całego ruchu.

Przewidywana, a raczej postulowana, przez transhumanistów<sup>2</sup>, ewolucja naszego gatunku ma wyjść poza ograniczenia ewolucji biologicznej dzięki współczesnym i przyszłym osiągnięciom nauki i techniki, skutkiem których będzie istotna modyfikacja ludzkiego ciała oraz umysłu. W tej kwestii transhumaniści optują za tak zwaną zasadą wolności morfologicznej (More i Vita-More, 2013). Stwierdza się tam (More i Vita-More, 2013, np. s. 55), że jako ludzie mamy prawo do dowolnych modyfikacji naszych ciał i umysłów, jeśli tylko zmiany te poprawią zdrowie, satysfakcję z życia i ogólne możliwości człowieka jako jednostki i jako gatunku. Ostatecznym

<sup>2</sup> Zamiast terminu „transhumanizm”, używa się niekiedy innych nazw. Najczęściej pojawia się słowo: „posthumanizm”. Trzeba tu jednak dodać, że „posthumanizm” może się kojarzyć z „postmodernizmem” lub innymi formami dyskredytacji czy krytyki filozofii Oświecenia. Transhumanisci (More, W: More i Vita-More, 2013, s. 10) jednak stwierdzają, że ich filozofia „broni zasadniczych idei Oświecenia i jego ideałów” (w org. *continues to champion the core of the Enlightenment ideas and ideals*), przede wszystkim racjonalizmu, metod naukowych, praw człowieka i jego wolności.

celem transhumanistów (np. More i Vita-More, 2013; Herbert, 2014; Hansel i Grassie, 2011) jest ludzka nieśmiertelność, która, ich zdaniem, będzie możliwa dzięki przeniesieniu ludzkich umysłów do komputerów i oczywiście dzięki zastąpieniu biologicznych ciał odpowiednio skonstruowanymi maszynami.

Jak widać, transhumaniści nie mają najmniejszych wątpliwości w kwestii równoważności (żeby nie powiedzieć: tożsamości) sztucznych i naturalnych intelektów, czyli komputerowych symulacji ludzkich umysłów i tego, co jest symulowane. Taka postawa – czyli przyjęcie, że model i to, co jest modelowane to właściwie jedno i to samo – jest bardzo interesująca z czysto logicznego punktu widzenia. Ponadto oznacza ona, że filozoficzne spory na temat, czym jest umysł, transhumaniści uznali za ostatecznie i przekonująco rozstrzygnięte. Choć w gruncie rzeczy nie wiadomo, które ze współczesnych stanowisk filozofii umysłu przyjęto w transhumanizmie. Nie całkiem, ale nieomal serio można tu powiedzieć, że transhumaniści „wynaleźli” nowy rodzaj „redukcjonizmu”. Ten „nowy redukcjonizm” polega na abstrahowaniu od faktu nierozstrzygnięcia ważnych z naukowego punktu widzenia sporów i postępowaniu (a przynajmniej myśleniu) w taki sposób, jakby były one rozstrzygnięte.

To co jeszcze niedawno traktowane było jako metafora i czysta fantazja transhumaniści potraktowali jako poważny i możliwy do realizacji program. Transhumanizm, z jego eschatologicznymi rojeniami, jest więc rodzajem religii. Każdą religię należy traktować serio, ideały każdej religii są mniej lub bardziej intensywnie wprowadzane w życie. Nie chodzi tu jednak o to, czy i jaki procent z nas stanie się wyznawcami tej religii. Problem z jej wierzeniami i ideałami nie polega też na tym jak szybko i jak bardzo będzie się rozwijać inżynieria sztucznej inteligencji. Problem leży raczej w braku rozumienia funkcji i istoty tworzonej przez nas technologii i w – teoretycznie znanym nam doskonale i wielokrotnie wskazywanym – nienadążaniu refleksji humanistycznej za rozwojem technicznym.

### **CZYM JEST SZTUCZNA INTELIGENCJA, JAKIE PERSPEKTYWY I ZAGROŻENIA STWARZA JEJ ROZWÓJ**

Kultura popularna coraz mocniej i skuteczniej przyzwyczajają nas do myśli o równości sztucznej i naturalnej inteligencji<sup>3</sup>. Niemożliwość zauważenia – tak w nauce jak i poza nią – fałszywości tej tezy ma swoje źródła we współczesnych programach edukacyjnych, które, jako takie, są ucieleśnieniem dominujących wartości i celów współczesności. Współcześnie wszystko jest towarem i wszystko ma przynosić wymierzalne ekonomicznie zyski. Wiedza wartościowa to wiedza użyteczna. Do war-

---

<sup>3</sup> Transhumaniści idą tu znacznie dalej i przewidują rychłe nadejście „eksplozji inteligencji”: Sztuczna inteligencja wkrótce przewyższy możliwości oraz sprawność inteligencji ludzkiej i stworzy jeszcze doskonalszą inteligencję, ta z kolei powoła do życia inteligencję jeszcze doskonalszą, etc.

tościowej wiedzy humanistycznej nie należy zatem antropologia, większość działów socjologii, niektóre działy psychologii, historia, filozofia razem z etyką i estetyką (z wyjątkiem może logiki).

Zagadnienie relacji pomiędzy sztuczną i naturalną inteligencją jest problemem tyleż filozoficznym, co i praktycznym. Nie rozstrzygnięto sporów teoretycznych na temat istoty umysłów, koniecznych i wystarczających warunków, jakie musi spełniać dany obiekt, by można było przypisać mu świadomość i zdolność do myślenia, nie ustalono jakie są związki pomiędzy procesami mózgowymi a procesami umysłowymi. Wszystkie te stare, ale ciągle trwające spory teoretyczne nie przeszkadzają konstruktorom sztucznej inteligencji w tworzeniu i doskonaleniu maszyn będących w stanie wykonywać niektóre ludzkie czynności poznawcze.

Pomijając uzasadniony, jak sądzą niektórzy, czy nieuzasadniony, jak utrzymują inni, optymizm na temat możliwości i przyszłych osiągnięć sztucznej inteligencji, zastanówmy się nad tym, co już się zdarzyło, czyli nad teraźniejszymi i przeszłymi relacjami pomiędzy naturalnym i sztucznym intelektem.

Mówiąc najogólniej, sztuczną inteligencją jest wszystko co ma wspierać, a nawet w pewnych okolicznościach zastąpić człowieka w jego czynnościach poznawczych. Wbrew pozorom historia (tak zdefiniowanej) sztucznej inteligencji nie zaczęła się w XX wieku, historia ta jest niemal tak długa, jak długa jest historia ludzkiej cywilizacji. Zaczyna się wraz z wynalezieniem pisma i liczydeł. Pismo – obrazkowe w stopniu niedoskonałym, alfabetyczne w stopniu bardziej zaawansowanym – wspiera ludzi w wykonywaniu różnych czynności pamięci. Psychologowie tradycyjnie uważają, że są to czynności trzech kategorii: czynności (procesy) zapamiętywania, przechowywania i odtwarzania (przypominania sobie). Pismo nie tylko utrwała pewne wiadomości czy informacje (zapobiega ich zapomnieniu), ale także pozwala wyjść człowiekowi poza naturalne ograniczenia pojemności jego pamięci. Wynalezienie i doskonalenie pisma było więc poprawianiem i rozszerzaniem ludzkich możliwości umysłowych. Ale pismo razem ze znakami matematycznymi i innymi językami formalnymi wspomagało też ludzkie procesy myślowe. Taka pisemna eksterioryzacja ludzkiego myślenia pozwala ludziom na przyjrzenie się ich własnej myśli. To co pomyślała i zapisała jednostka, mogła sama ocenić krytycznie, a następnie skorygować lub udoskonalić. Mogła też podlegać ocenie i krytyce ze strony innych. Pismo to nie tylko kolektywna pamięć, ale także kolektywne myślenie.

Kolejnym krokiem w uzewnętrznianiu, rozszerzaniu i doskonaleniu ludzkiego intelektu było wynalezienie i doskonalenie arytmometrów. Historia arytmometru to historia automatyzacji uzewnętrznionego ludzkiego myślenia, a przynajmniej tej jego części, która związana jest z posługiwaniem się liczbami. Rozwój arytmometrów i języków formalnych pozwolił ludziom (a przynajmniej



niektórym z nich) na uświadomienie<sup>4</sup> sobie faktu, że uzewewnętrznione, symbolizowane i zautomatyzowane procesy intelektualne, to nie tylko rachowanie czy posługiwanie się liczbami.

I tu zaczyna się pierwszy problem, czyli pytanie, czy wszystkie ludzkie procesy intelektualne można uzewewnętrznic, zautomatyzować, a następnie delegować do maszyn, czy tylko można tak postąpić z niektórymi z nich? Jasna odpowiedź na to (ale tak a nie inaczej postawione) pytanie rozstrzygnęłaby kwestie typu "czy komputer myśli?"

Twórcy sztucznej inteligencji (a przynajmniej optymiści) uważają, że mamy już gotową odpowiedź. Ale nawet jeśli odpowiedź na to pytanie brzmi "tak", wcale to nie oznacza równości sztucznej i naturalnej inteligencji, jako że musimy jeszcze rozważyć skutki i funkcje automatyzacji czegokolwiek, w tym skutki i funkcje automatyzacji ludzkich czynności.

Zastanówmy się jednak najpierw nad możliwością przeniesienia ludzkich umysłów do maszyn cyfrowych, przeanalizujmy ten cel transhumanizmu w świetle współczesnych teorii umysłu. Jedną z tych teorii została już wspomniana powyżej, mam tu na myśli teorię identyczności. Jeżeli przyjmiemy za zwolennikami tej teorii, że ludzki umysł to po prostu mózg, to oczywiście można wyobrazić sobie połączenie mózgu z jakimś sztucznym (lub naturalnym ale innym) ciałem. Nie gwarantuje to nieśmiertelności, bo tkanka mózgowa, jak każda inna musi umrzeć.

Transhumaniści mają jednak na myśli wymodelowanie w komputerze lub jakiejś innej maszynie struktury połączeń neuronalnych danego biologicznego mózgu i traktują to jako przedłużenie tego (tego co wymodelowano) życia. Z czysto logicznego punktu widzenia, trzeba powiedzieć, że jakkolwiek kopia nie jest oryginałem, a więc skopiowanie człowieka nie jest przedłużaniem jego życia, lecz powołaniem do życia innej jednostki<sup>5</sup>.

Jednak problem istoty ludzkiego umysłu i jego związków z ciałem jest bardziej skomplikowany. Kolejną współczesną teorią umysłu jest teoria powstała w ramach logicznego (analitycznego) behawioryzmu (Ryle, 1970; Wittgenstein, 1972). Odrzuca się tam jakiegokolwiek substancjalne<sup>6</sup> pojmowanie umysłu, twierdząc, że to, co nazywamy umysłem, nie jest żadną substancją duchową ani materialną (nie jest zatem mózgiem biologicznym ani cyfrowym) lecz sposobem używania ciała albo sumą dyspozycji do określonego działania. W świetle tej teorii niczego przenieść z ludzkiej głowy do maszyny nie można, bo inne ciało to inne sposoby i zasady działania, a więc inna jednostka.

<sup>4</sup> Tę świadomość wyraża przede wszystkim Turingowska definicja „obliczania” (jako procesu manipulowania fizycznie ucieleśnionymi symbolami według z góry ustalonych reguł) i koncepcja tzw. „maszyny Turinga”.

<sup>5</sup> To samo dotyczy ewentualnego tworzenia klonów człowieka.

<sup>6</sup> Substancja jest tu pojmowana jako coś co istnieje samo przez się, a więc nie jest cechą ani właściwością, sumą właściwości czegoś innego.

Istnieje też współcześnie inna odpowiedź na pytanie „czym jest umysł?”<sup>7</sup>; odpowiedź wypracowana w stanowisku, które jest syntezą filozofii analitycznej i tzw. fenomenologii egzystencjalnej. Współcześnie reprezentuje to podejście John Searle (1984, 1995), a jego prekursorem był np. Kazimierz Twardowski (1965). Odpowiedź na pytanie „czym jest umysł?” wynika z przedstawionej przez tych autorów teorii działania (*action* – Searle, 1985) albo inaczej teorii czynności (Twardowski, 1965).

Wszystkie ludzkie czynności (Twardowski, 1965), albo działania, mają charakter psychofizyczny. Czynności ludzkich nie wykonuje żadna tkanka ani organ, lecz cały człowiek (Twardowski, 1965, s. 95-96), „działania znamionują dwa składniki, składnik umysłowy i składnik fizyczny” (Searle, 1995, s. 57). Tego, co umysłowe nie da się oddzielić od tego co fizyczne, niezależnie od tego kto lub co jest podmiotem działania<sup>8</sup>. Umysłowy i fizyczny składnik albo aspekt działania, to dwie nierozdzielne strony tego samego procesu (aktu, czynności). John Searle (1984, s. 28-41, 1995, s. 25-37) jest znanym przeciwnikiem tezy o myśleniu komputerów, a dokładniej poglądu, że maszyny cyfrowe mogą wykonywać wszystkie funkcje, jakie wykonuje naturalny umysł człowieka<sup>9</sup>.

Zdaniem Searla’a (1984, s. 68; 1995, s. 63) wszystkie umysłowe aspekty albo składniki ludzkiego działania mają swoje podstawy (istnieją dzięki) temu co nieumysłowe i nieintencjonalne: „The whole network of intentionality only functions against a background of human capacities that are not themselves mental states.” – „Całość stanów intencjonalnych funkcjonuje na podstawie ludzkich zdolności, które same w sobie nie są stanami umysłowymi” (Searla’a 1984, s. 68, 1995, s. 63). Tak więc to co umysłowe – w terminologii Johna Searla’a: intencjonalne – nie może istnieć niezależnie od tego co fizyczne, jakkolwiek to co umysłowe jest nową jakością, i nie da się terminów mentalistycznych zastąpić terminami fizykalistycznymi. Niemniej jednak, to co jest materialne determinuje (jakkolwiek w sposób szczególny) to co nazywamy umysłowym. Zmiana ciał fizycznych jest w sposób konieczny zmianą umysłów „nadbudowanych” na tych ciałach. Umysłu z jednego do innego ciała przenieść się nie da, chociaż umysł i ciało można zmieniać, doskonalić. Doskonallymy nasze ciała i umysły, ucząc się nowych czynności. Zwiększamy repertuar naszych czynności i doskonallymy nasz sposób działania w świecie, wymyślając i doskonallymac narzędzia, które usprawniają nasze działanie. Narzędzia te dodane do biologicznych systemów funkcjonalnych pozwalają nam na lepsze wykonywanie naszych czynności, a także umożliwiają wykonanie czynności, których nie bylibyśmy w stanie wykonać, mając do dyspozycji tylko to, w co wyposażyla nas natura. Mając noktowitzory widzimy w ciemności lepiej niż koty, mając

<sup>7</sup> Nie omawiam tu wszystkich współczesnych teorii umysłu. Ich szerszy przegląd w Bobryk (1996).

<sup>8</sup> W tej kwestii K. Twardowski i J. Searle różnią się nieco w swoich poglądach (patrz Bobryk, 1996, 2001).

<sup>9</sup> Szerzej na ten temat, np. Bobryk (2010).

broń, zabijamy skuteczniej niż jakiegokolwiek drapieżniki, mając technikę elektroniczną, wykonujemy szybciej, pewniej i skuteczniej nasze czynności umysłowe. Sztuczna inteligencja nie jest podmiotem myślenia, tak jak dźwignia czy wielokrążek nie jest Herkulesem, jakkolwiek wszystkie te narzędzia zwiększają ludzkie możliwości. Wynalezienie takiego lub innego narzędzia (także sztucznej inteligencji) wymaga inteligencji naturalnej, ale nie gwarantuje rozsądnego stosowania tego narzędzia. Nawet najbardziej wyrafinowane narzędzia muszą być mądrze stosowane. Sztuczna inteligencja (podobnie jak wszelka automatyka) zwiększa ludzkie możliwości poznawcze, ale nie gwarantuje rozsądnego zachowania ludzi, chociaż w pewnych warunkach może uniemożliwić człowiekowi robienie głupstw (wynikających z nieuwagi, roztargnienia, brak doświadczenia, wiedzy, etc.).

Najwięcej głupstw można zrobić nie rozumiejąc istoty narzędzia, jego ograniczeń, jego funkcji. Dlatego przyspieszony i ciągle przyśpieszający rozwój techniczny może być niebezpieczny, jeśli źle się rozumie istotę techniki i jej relacji z człowiekiem. Technika cyfrowa wspierająca ludzi w wykonywaniu przez nich różnorodnych funkcji umysłowych, wymaga nie tylko rozumienia i znajomości tej techniki, ale także rozumienia przez ludzi własnych umysłów. Na nic się przydadzą pletwy temu, który nie umie pływać, a komputer temu, kto ma problemy z myśleniem. Nie można zaniedbywać filozofii w epoce „myślących” maszyn. Komputerowa metafora umysłu nie przestała być metaforą, przeto nie można jej zrozumieć dosłownie. „Kamień spadający z serca” nie przytłucze nogi, tak jak „pożar krwi” nie spali niczego, zaś komputer nie rozwiąże ludzkich problemów, z którymi, z lepszym lub gorszym skutkiem, boryka się filozofia i psychologia. W świecie ludzkim wszystko ma swoje funkcje, cele i granice, dotyczy to także najbardziej wyrafinowanych narzędzi i idei, chociaż nie zawsze o tym pamiętamy.

Aby wykorzystać możliwości komputera, trzeba rozumieć jego istotę, funkcje i ograniczenia. Podobnie, aby postawa redukcjonistyczna zwiększała nasze horyzonty poznawcze, pomagała a nie przeszkadzała w zdobywaniu wiedzy, trzeba rozumieć ograniczenia i istotę redukcjonizmu. Redukcjonizm w nauce odsłania nam tylko część prawdy, dotyczy to także tak zwanego redukcjonizmu komputerowego przyjętego w kognitywistyce. Pod pewnymi względami, w pewnych warunkach umysł ludzki działa podobnie jak komputer, jednak umysł ludzki (jako taki i ukształtowany naturalnie) nie jest komputerem, a komputer – jakkolwiek szybko by nie działał i jakkolwiek by nie był zaprogramowany – nie jest umysłem. Pytanie „czym jest umysł” jest trudniejszym pytaniem niż „czym jest komputer”<sup>10</sup>. Jeszcze trudniejszym jest pytanie „jakie w gruncie rzeczy są i jakie być powinny relacje między naturalną i sztuczną inteligencją”. Nie jest to pytanie techniczne lecz filozoficzne. Aby stworzyć sobie szansę od-

powiedzi, trzeba przyjąć odpowiednią perspektywę teoretyczną.

Perspektywy teorii działania i teorii czynności najtrafniej, moim zdaniem, pokazują relacje między sztuczną a naturalną inteligencją.

Jak powiedziane zostało wyżej człowiek wykonuje swoje czynności z użyciem prostych lub bardzo wyrafinowanych narzędzi. Czyni tak niezależnie od tego czy chcielibyśmy nazwać te czynności „umysłowymi” czy „fizycznymi”. Zresztą w świetle teorii Searle’a i Twardowskiego wszystkie czynności mają aspekt psychofizyczny (mają składnik fizyczny i umysłowy). Jakkolwiek w mojej świadomości i w moich przeżyciach „fizyczność” w jednym wypadku wydaje się być dominująca, w innym zaś na plan pierwszy wysuwa się „psychiczność”. Jeżeli popycham popsuty samochód, to nazwę to czynnością fizyczną, jeżeli wyobrażam sobie popychanie samochodu, to powiem, że wykonuję umysłową czynność, jednak w obu wypadkach zachodzą jakieś procesy fizyczne i akty psychiczne. Nie ma natomiast żadnej ludzkiej psychiki w komputerach, jakkolwiek ich użycie przez człowieka może zmieniać (zarówno ograniczać jak i rozszerzać czy „poprawiać”) ludzką psychikę, bo ludzka psychika zależy od ludzkich ciał, ale też od stworzonych przez ludzi narzędzi.

## LITERATURA

- Andruszkiewicz, A. (2004). *Słownik filozofii*. Warszawa: Świat Książki.
- Blackburn, S. (1997). *Oksfordzki słownik filozoficzny*, (przeł. Cezary Cieśliński). Warszawa: KiW.
- Bobryk, J. (1996). *Akty świadomości i procesy poznawcze*. Wrocław: Wyd. Leopoldinum.
- Bobryk, J. (2001). *Twardowski. Teoria działania*. Warszawa: Prószyński i S-ka.
- Bobryk, J. (2010). What it means to understand language: The view from nowhere and the view of the individual. W: B. Bokus (red), *Studies in the psychology of language and communication*. Warszawa: Matrix.
- Bobryk, J. (2011). *Redukcjonizm, psychologia, semiotyka*. Warszawa: Wyd. Instytut Psychologii PAN, Polskie Towarzystwo Semiotyczne.
- Hansel, G.R., Grassie, W. (2011). *Transhumanism and its critics*. Philadelphia: Metanexus.
- Herbert, D. (2014). *Becoming god. Transhumanism and the quest for cybernetic immortality*. Kitchener, Ontario: Joshua Press.
- Honderich, T. (red.). (1998). *Encyklopedia filozofii* (przeł. Jerzy Łoziński). Poznań: Zysk.
- Lacey, A.R. (1999). *Słownik filozoficzny* (przeł. Roman Matuszewski). Poznań: Zysk.
- More, M., Vita-More, N. (red.). (2013). *The transhumanist reader*. Singapore: Wiley-Blackwell.
- Reber, A. (2000). *Słownik psychologii* (przeł. Barbara Janasiewicz-Kruszyńska), Warszawa: Scholar.
- Ryle, G. (1970). *Czym jest umysł* (przeł. Witold Marciszewski). Warszawa: PWN.
- Searle, J. (1984). *Minds, brains and science*. Cambridge MA: Harvard University Press.

<sup>10</sup> Bo komputery zbudowali ludzie, a umysły ukształtowała natura.

Searle, J. (1995). *Umysł, mózg i nauka* (przeł. Jerzy Bobryk). Warszawa: PWN.

Searle, J. (2004). *Mind. A brief introduction*. Oxford: Oxford University Press.

Twardowski, K. (1965). *Wybrane pisma filozoficzne*. Warszawa: PWN.

Urbanek, A. (1987). *Redukcjonizm*. W: M. Iżewska (red.), *Filozofia a nauka: zarys encyklopedyczny*. Wrocław: Ossolineum.

Wittgenstein, L. (1972). *Dociekania filozoficzne* (przeł. Bogusław Wolniewicz). Warszawa PWN.

*Wielka Encyklopedia PWN*. (2004). Warszawa: PWN.