

ROBERT POCZOBUT

CZY ŚWIADOMOŚĆ JEST NIEZALEŻNA OD SUBSTRATU?

Naturalne jest przypuszczenie, że świadomość powstaje na mocy organizacji funkcjonalnej mózgu. Z tego punktu widzenia chemiczny, a faktycznie też kwantowy substrat mózgu nie ma znaczenia dla powstania świadomości. Liczy się abstrakcyjna organizacja przyczynowa w mózgu, organizacja, która może być realizowana w wielu substratach fizycznych.

David CHALMERS (2010, 411)

Some future computer simulations might turn out to be conscious, but the very different mechanisms and energy operation of computers and brains make it equally plausible that simulations of thinking will not achieve consciousness, just as simulations of ocean waves are not wet.

Paul THAGARD (2022, 85)

WSTĘP

Ważnym i aktualnym zagadnieniem filozofii umysłu i filozofii sztucznej inteligencji jest problem substratu świadomości. W odróżnieniu od badań nad jej neuronalnymi korelatami w dyskusjach na temat substratu świadomości stawia się pytania na wyższym poziomie ogólności: Czy możliwa jest sztuczna (maszynowa) świadomość? Czy prawa przyrody limitują zakres substratów, które mogą stanowić fizyczne podłoże świadomości? Czy każda forma świadomości musi być u swych podstaw biologiczna? Jaka organizacja funkcjonalna jest konieczna

Dr hab. ROBERT POCZOBUT, Prof. UwB — Uniwersytet w Białymstoku (UwB), Wydział Filozofii, Zakład Epistemologii i Kognitywistyki, kierownik; absolwent KUL; adres do korespondencji: Plac NZS 1, 15-420 Białystok; e-mail: r.poczobut@uwb.edu.pl; ORCID: <https://orcid.org/0000-0003-2591-7002>.

do wytworzenia stanów świadomych? Czy organizacja funkcjonalna jest zależna od substratu?

Dyskusje na temat substratu świadomości przypominają pod wieloma względami spory dotyczące natury życia. Zwolennicy teorii substratowych za cechę istotną procesów życiowych uznawali ich materialny substrat (na przykład biopolimery, takie jak białka i kwasy nukleinowe). Zakładano, że substrat jest konstytutywny dla procesów życiowych i nie jest możliwa ich realizacja poza określonym substratem. Natomiast ujęcia funkcjonalistyczne, związane z cybernetyką i teorią informacji, w wyjaśnieniach natury życia abstrahowały od substratu, akcentując rolę prawidłowości funkcjonalnych oraz przetwarzania informacji w układach o wysokim poziomie złożoności. Funkcjonalisci dopuszczali możliwość realizacji procesów życiowych w substratach nieorganicznych, w szczególności możliwość wytworzenia sztucznego życia (OLSON 1997; BEDAU 1998). Analogonem tych dyskusji jest współczesny spór o to, czy świadomość jest niezależna od substratu¹.

Celem artykułu jest analiza krytyczna wybranych stanowisk broniących niezależności świadomości od substratu, w konsekwencji zaś dopuszczających możliwość świadomości maszynowej. W sekcji pierwszej określam pojęcie świadomości, którego dotyczy dyskusja, oraz dokonuję krytycznej prezentacji stanowiska Davida Chalmersa. W sekcji drugiej analizuję zasadę niezależności od substratu w ujęciu Maxa Tegmarka. W sekcji trzeciej przedstawiam i poddaję krytycznej ocenie Paula Thagarda argument przeciwko zasadzie niezależności od substratu. W zakończeniu formułuję pytania i problemy wymagające dalszej dyskusji.

SUBSTRAT ŚWIADOMOŚCI A ORGANIZACJA FUNKCJONALNA

W literaturze filozoficznej i naukowej można spotkać twierdzenie, że świadomość to zdolność do automonitoringu, dzięki której odpowiednio zorganizowane układy uzyskują dostęp do informacji na temat swoich stanów wewnętrznych.

¹ Problem niezależności świadomości od substratu (fizycznego, materialnego, biologicznego, cielesnego) jest przedmiotem dyskusji także na gruncie stanowisk dualistycznych, zakładających funkcjonalną, a często także egzystencjalną niezależność świadomości od procesów zachodzących w mózgu i ciele człowieka. Interesującej wersji dualizmu, dopuszczającego częściową autonomię świadomości wobec jej fizycznego (cielesnego) podłoża, broni Stanisław Judycki w pracy *Świadomość i pamięć. Uzasadnienie dualizmu antropologicznego* (JUDYCKI 2004). W artykule omawiam problem niezależności świadomości od substratu w kontekście obliczeniowej teorii świadomości oraz świadomości maszynowej.

Taką zdolnością dysponują organizmy biologiczne wyposażone w układy nerwowe oraz maszyny o hierarchicznych architekturach z pętlami informacyjnymi i mechanizmem sprzężeń zwrotnych. Organizacja takich systemów umożliwia im dostęp do informacji o ich stanach wewnętrznych, dzięki czemu dysponują one zdolnością do samoregulacji i kontroli zachowania. Gdyby klasę takich zdolności uznać za cechy definicyjne świadomości, należałoby przyznać jej cechę niezależności od substratu (HOLLAND 2003; KRAUS i MAIER 2020).

Odpowiedź na pytanie, czy świadomość jest niezależna od substratu, zależy również od tego, jak rozumiemy świadomość. W literaturze zagadnienia opisane wyżej zdolności określa się mianem „świadomości funkcjonalnej” (BOŁTRUĆ 2013). W powyższym sensie za świadome moglibyśmy uznać zarówno układy biologiczne, jak i maszyny cybernetyczne, w których stwierdzamy występowanie takich zdolności. Pojęcie świadomości funkcjonalnej pomija jednak istotne cechy, które wiążemy ze świadomością, takie jak doznawanie wrażeń i przeżywanie stanów wewnętrznych. Maszyny z wbudowanym systemem automonitoringu i kontroli zachowania, lecz pozbawione przeżyć świadomych, są systemami typu *zombie*.

W związku z tą dystynkcją w tradycji analitycznej wprowadzono pojęcie świadomości fenomenalnej. Polega ona na zdolności do posiadania subiektywnych przeżyć — prostych doznań zmysłowych, jak odczucie bólu czy doznanie zieleni, a także złożonych przeżyć świadomych towarzyszących emocjom, myśleniu czy aktywności twórczej (CHALMERS 2010a). Powstaje pytanie: Czy świadomość fenomenalna jest niezależna od substratu? Czy maszyny mogą mieć subiektywne przeżycia świadome, choćby w postaci prostych doznań zmysłowych? Opinie filozofów i naukowców są podzielone (HAIKONEN 2007; 2009). Aby konkluzywnie odpowiedzieć na to pytanie, musielibyśmy wiedzieć: Na mocy jakich mechanizmów w organizmach biologicznych powstają przeżycia świadome? Czy takie mechanizmy można odtworzyć w systemach sztucznych? W dalszym ciągu, mówiąc o świadomości, będę miał na uwadze świadomość fenomenalną.

Wszystkie znane systemy, którym przypisujemy przeżycia świadome, są organizmami biologicznymi. Z tego powodu świadomość fenomenalna wydaje się szczególnie mocno związana z substratem biochemicznym. Inne procesy poznawcze, jak uczenie się czy pamięć, bez większego oporu traktujemy jako niezależne od substratu. Zdaniem Maxa Tegmarka do realizacji procesu poznawczego nie jest konieczny określony rodzaj materii, lecz jej organizacja funkcjonalna. Pamięć, obliczanie, uczenie się i inteligencja są niezależne od fizycznych szczegółów substratu, dzięki czemu jest możliwe istnienie sztucznej inteligencji oraz maszyn realizujących procesy poznawcze. Jak pisze Tegmark (2019, 94): „Sprzęt jest materią, a oprogramowanie – wzorcem. Niezależność obliczeń od substratu

oznacza, że możliwe jest istnienie sztucznej inteligencji: inteligencja nie wymaga ciała, krwi ani atomów węgla”. Czy podobnie jest ze świadomością?

David Chalmers, zastanawiając się, jak systemy fizyczne wywołują świadomość, zwraca uwagę, że kluczowa jest ich organizacja funkcjonalna, a nie substrat, z którego są zrobione. Jego zdaniem: „Dowolny system o organizacji funkcjonalnej odpowiedniego rodzaju jest świadomy bez względu na to, z czego się składa” (CHALMERS 2010a, 513). Zdaniem funkcjonalistów implementacja odpowiedniej architektury obliczeniowej wystarcza do osiągnięcia organizacji umożliwiającej wytworzenie świadomości. Systemy świadome są układami fizycznymi o strukturze przyczynowej odzwierciedlającej formalną strukturę obliczeń. Sama struktura obliczeń (oprogramowanie jako obiekt abstrakcyjny) nie wystarcza do powstania świadomości. Dopiero jej implementacja w postaci układu konkretnych procesów fizycznych powiązanych przyczynowo umożliwia powstanie świadomości fenomenalnej. Innymi słowy, organizacja funkcjonalna układu określa podstawowy mechanizm przyczynowy, który wywołuje przeżycia świadome. Zadaniem neuroobliczeniowej teorii świadomości jest odkrycie organizacji funkcjonalnej mózgu, która odpowiada za wytwarzanie świadomości. Zadaniem silnej sztucznej inteligencji jest odtworzenie tej organizacji w systemie sztucznym.

Aby dokładniej zrozumieć stanowisko Chalmersa, należy odróżnić: (i) abstrakcyjny poziom obliczeń, programów, maszyn Turinga, automatów komórkowych etc., (ii) poziom procesów poznawczych i przeżyć świadomych, fizycznie ucieleśnionych i pozostających w związkach przyczynowych z innymi obiektami w świecie, (iii) relację implementacji między abstrakcyjnym poziomem obliczeniowym a układami fizycznymi realizującymi procesy obliczeniowe w formie sekwencji procesów przyczynowych. Jak pisze Chalmers (2010a, 518): „Nieformalnie mówimy, że układ fizyczny implementuje obliczenie, gdy struktura przyczynowa układu odzwierciedla formalną strukturę obliczeń. Oznacza to, że układ implementuje obliczenie, jeśli istnieje sposób odwzorowania stanów układu na stany obliczenia, tak, że fizyczne stany powiązane przyczynowo są odwzorowane na odpowiednio formalnie powiązane stany formalne”.

Obliczenia dopuszczają możliwość wielorakiej realizacji w tym sensie, że mogą być implementowane w różnych strukturach fizycznych – na przykładach w komputerach, robotach, mózgach. Co ważne, obliczeniowe wyjaśnienia procesów fizycznych są możliwe w takim zakresie, w jakim procesy te implementują określone klasy obliczeń (CHALMERS 1994). Same obliczenia nie są w stanie wywołać żadnego procesu fizycznego, w szczególności przeżyć świadomych. Mogą to jednak uczynić fizyczne implementacje obliczeń. Zdaniem Chalmersa (2010a, 516,

521): „Teza o mocnej AI wyrażona w kategoriach obliczania mówi nam, że implementacja stosownych obliczeń wystarczy do powstania świadomości. [...] Jeśli mówimy, że układ fizyczny implementuje dane złożone obliczenie P, to mówimy coś bardzo istotnego o strukturze przyczynowej tego układu, coś, co może być bardzo użyteczne do wyjaśnień poznawczych, a być może także do rozumienia podstaw świadomości”.

Kluczową rolę w rozważaniach Chalmersa na temat świadomości odgrywa zasada niezmienności organizacyjnej, którą wyraża sformułowanie: „Dla każdego układu fizycznego, w którym powstaje przeżycie świadome, istnieje pewna organizacja funkcjonalna F realizowana przez system, taka, że jest konieczne przyrodniczo, iż każdy system, który realizuje F , będzie mieć identyczne przeżycia świadome” (CHALMERS 2010a, 455). Na mocy tej zasady, gdyby zbudowano szczegółowy model obliczeniowy ludzkiego mózgu, obejmujący opis jego organizacji funkcjonalnej na odpowiednim poziomie dokładności, a następnie dokonano implementacji tego modelu w innym układzie fizycznym odtwarzając w nim organizację funkcjonalną ludzkiego mózgu, spowodowałoby to wytworzenie przeżyć świadomych nieodróżnialnych od przeżyć pierwowzoru. Oczywiście, przy założeniu, że zostałyby odwzorowane wszystkie relewantne elementy organizacji funkcjonalnej mózgu wystarczające do wywołania przeżyć świadomych (CHALMERS, 2010a, 423).

Dokładnie to samo dotyczy organizacji funkcjonalnej prostszych mózgów, takich jak mózgi myszy, ryb czy ślimaków. Zakładając, że ich posiadacze dysponują prostszym repertuarem przeżyć świadomych, odtworzenie organizacji funkcjonalnej takich mózgów automatycznie spowodowałoby powielenie ich przeżyć. Zasadne jest pytanie: Czy istnieje jakiś próg związany z minimalną organizacją funkcjonalną, po którego przekroczeniu pojawiają się przeżycia świadome? Jak pisze Chalmers (2010a, 483): „Gdy schodzimy na skali od ryb i ślimaków przez proste sieci neuronowe aż do termostatów, w którym momencie świadomość powinna zaniknąć? Fenomenologia ryb i ślimaków prawdopodobnie nie będzie prymitywna, ale stosunkowo złożona, odzwierciedlając różne rozróżnienia, które mogą one poczynić. Zanim fenomenologia całkowicie zniknie, przypuszczalnie stanie się jakąś maksymalnie prostą fenomenologią”. Wbrew jednak sugestii o istnieniu progu złożoności, po którego przekroczeniu pojawia się świadomość minimalna jako cecha emergentna, Chalmers dopuszcza możliwość, że świadomość jest własnością fundamentalną, powszechnie występującą w przyrodzie².

² Jest to jedna z wersji pan(proto)psychizmu, będąca następstwem przyjęcia zasady podwójnego aspektu informacji. Chalmers określa swoje stanowisko także mianem „naturalistycznego dualizmu własności”, z uwagi na nieredukowalny do stanów fizycznych status przeżyć świadomych. To zagadnienie to wykracza poza ramy problemowe artykułu, w którym skupiam się na funkcionali-

Głównym założeniem tezy Chalmersa o niezależności świadomości od substratu jest przekonanie, że świadomość powstaje na bazie organizacji funkcjonalnej, rozumianej jako abstrakcyjny wzorzec oddziaływań przyczynowych między częściami systemu oraz między systemem a środowiskiem. Używając innej terminologii, można powiedzieć, że implementacja organizacji funkcjonalnej świadomości stanowi jej mechanizm. Świadomość jest niezależna od substratu, ponieważ niezależna od substratu jest organizacja funkcjonalna, która wywołuje świadomość: „Dana organizacja funkcjonalna może być realizowana przez różne układy fizyczne. Na przykład organizacja realizowana przez mózg na poziomie neuronowym może być w zasadzie realizowana przez układ fizyczny. Opis organizacji funkcjonalnej mózgu abstrahuje od natury fizycznej wchodzących w grę części oraz sposobu implementacji związków przyczynowych. Liczy się wyłącznie istnienie części i stosunki zależności między ich stanami” (CHALMERS 2010a, 412).

W pracach Chalmersa nie znajdziemy szczegółowego opisu organizacji funkcjonalnej mózgu wystarczającej do wywołania stanów świadomych. Najprawdopodobniej wciąż nikt nie dysponuje kompletnym opisem takiej organizacji. Chalmers zdaje sobie sprawę, że organizację mózgu można opisywać na wielu poziomach. Przy czym nie jest oczywiste, że poziom obwodów (sieci) neuronowych jest właściwym poziomem opisu. Bardzo możliwe, że w wytwarzaniu stanów świadomych uczestniczą struktury wyższego rzędu, a także struktury molekularne. Autor *Świadomego umysłu* wprowadza w związku z tym pojęcie organizacji funkcjonalnej wysokiej dokładności oraz formułuje niemal tautologiczne twierdzenie: układy mające dokładnie taką samą organizację funkcjonalną (funkcjonalnie izomorficzne) będą posiadały takie same stany świadome, niezależnie od substratu, w którym taka organizacja funkcjonalna zostałaby zrealizowana³.

Twierdzenie o funkcjonalnie izomorficznych układach świadomych zrealizowanych w różnych substratach nabrałoby treści empirycznej, gdyby podano chociaż jeden przykład takich układów. Jak dotychczas takiego przykładu nie udało się wskazać. Nie wiadomo, na przykład, czy organizacja funkcjonalna ludzkiego mózgu jest możliwa do zrealizowania w substracie krzemowym. Czy związki

stycznym aspekcie jego poglądów. Chyba nikt inny tak precyzyjnie nie sformułował założeń nie-redukcjonistycznego funkcjonalizmu (maszynowego) w teorii świadomości fenomenalnej, jak zrobił to Chalmers w swojej klasycznej pracy.

³ Nie jest oczywiste, czy teza o zależności świadomości od organizacji funkcjonalnej jest niesprzeczna z panpsychizmem. Musielibyśmy przyjąć, że wszędzie w przyrodzie zrealizowane są organizacje funkcjonalne umożliwiające istnienie jakiejś formy przeżyć świadomych. Znacznie bardziej intuicyjna jest interpretacja w duchu teorii emergencji – dopiero po przekroczeniu określonego progu złożoności organizacji funkcjonalnej w organizmie lub w układzie sztucznym mogą pojawić się przeżycia świadome. Problem w tym, że wciąż nie potrafimy wskazać takiego progu.

krzemem dają takie możliwości strukturalno-funkcjonalne, jak związki węgla? Zastanawiając się nad tą kwestią, Chalmers (2010a, 431) zauważa:

Może krzem po prostu nie ma zdolności do wykonywania funkcji, jakie realizują neurony w mózgu, więc żaden układ krzemowy nie nadawałby się do tego zadania. Nie jest jasne, czy istnieje podstawa dla tego zarzutu; mamy już protezy rąk i nóg, a trwają prace nad prototecznymi oczami; więc dlaczego nie prototetyczne neurony? W każdym razie, nawet gdyby krzemowy, izomorficzny funkcjonalnie odpowiednik był niemożliwy [...], argument na rzecz zasady niezmienności pozostawałby tak samo w mocy. Zasada niezmienności mówi tylko, że jeśli istnieje izomorficzny funkcjonalnie odpowiednik świadomego systemu, to będzie on miał tego samego rodzaju przeżycia świadome. Jeśli krzemowe odpowiedniki izomorficzne są niemożliwe, ocena systemów krzemowych jest tu po prostu nieistotna.

Stanowisko Chalmersa budzi szereg wątpliwości. Na zakończenie tej sekcji sformułuję pięć uwag o charakterze krytyczno-polemicznym:

(i) Wciąż brakuje świadectw empirycznych, że istnieje choćby jeden niebiologiczny substrat, w którym dałoby się zrealizować organizację funkcjonalną mózgu wystarczającą do pojawienia się przeżyć świadomych. W konsekwencji racjonalne jest przyjęcie hipotezy alternatywnej: substrat biologiczny jest jedynym substratem, w którym taka organizacja funkcjonalna może być zrealizowana.

(ii) Argument odwołujący się do przykładu protez narządów biologicznych oraz implantów domózgowych, nie jest konkluzywny, ponieważ w takich wypadkach nie mamy do czynienia z organizacją funkcjonalną wystarczającą do wywołania przeżyć świadomych. Z kolei eksperyment myślowy odwołujący się do wymiany wszystkich neuronów na ich krzemowe zamienniki, opisuje scenariusz, którego nie jesteśmy w stanie ocenić jako możliwy do zrealizowania⁴.

(iii) W rozważaniach Chalmersa brakuje analizy, w jaki sposób natura substratu limituje klasę organizacji funkcjonalnych możliwych do zrealizowania w tym substracie. Zależność od substratu może być znacznie silniejsza, niż sugeruje Chalmers. Powyższy problem wymaga szczegółowych badań empirycznych i jest niemożliwy do rozstrzygnięcia w trybie analizy pojęciowej.

(iv) Chalmers nie wykazał, że przeżycia świadome zależą tylko i wyłącznie od organizacji funkcjonalnej układu, któremu przysługują. Przyjął takie założenie i wyprowadził z niego daleko idące konsekwencje. Jednak świadomość może być zależna również od charakterystyk wykraczających poza organizację funkcjonalną.

⁴ Tym bardziej dotyczy to takich scenariuszy jak transfer świadomego umysłu do maszyny czy hipoteza, że żyjemy w symulacji komputerowej (BOSTROM 2003; CHALMERS 2010b; KURZWEIL 2012). Por. uwagi krytyczne Paula Thagarda omówione w dalszej części artykułu. Interesujące uwagi na temat empirycznych podstaw przypisywania świadomości komputerom zawiera (GAMEZ 2012).

(v) Organizacje funkcjonalne w systemach sztucznych mogą prowadzić do wytworzenia zupełnie innych form świadomości. Jeśli świadomość jest silnie zależna od substratu, wykluczona jest możliwość wytworzenia świadomego izomorfa w innym substracie. Jednak nie przekreśla to możliwości istnienia różnych form świadomości w różnych substratach.

2. ZASADA NIEZALEŻNOŚCI OD SUBSTRATU A ŚWIADOMOŚĆ – UJĘCIE MAXA TEGMARKA

W pracach na temat inteligencji, procesów poznawczych i świadomości Max Tegmark przypisuje szczególną rolę zasadzie niezależności od substratu (TEGMARK 2014; 2015; 2019)⁵. Wiele procesów makroskopowych ma cechy niezależne od podłoża. Przykładem są fale mające takie cechy jak prędkość, długość czy częstość, które można ująć za pomocą równania opisującego makroskopowy wzorzec, w znacznym stopniu niezależny od mikroskopowych szczegółów substratu, w którym się one rozchodzą się (parametrem zależnym od substratu jest szybkość rozchodzenia się fali, różna w różnych substratach). Tegmark (2019, 92-93) w następujący sposób charakteryzuje zasadę niezależności od substratu:

(i) Niezależność od substratu nie polega na tym, że makroskopowy wzorzec istnieje bez substratu. Jakiś substrat jest konieczny, nie ma wzorca bez substratu. Chodzi o to, że większość mikroskopowych cech substratu nie ma znaczenia z punktu widzenia cech samego wzorca. Jak pisze Tegmark (2019, 92): „Nie można przeprowadzić obliczeń bez materii, ale każda materia będzie dobra, o ile da się z niej utworzyć bramki NAND, sieć neuronów lub inną strukturę umożliwiającą uniwersalne obliczenia”.

(ii) Zjawiska niezależne od substratu mają cechy i dynamikę, które nie przysługują substratowi. Dzięki temu „uzyskują one własne życie” (fala przemieszczająca się po jeziorze ma cechy i dynamikę, których nie mają cząsteczki wody; fala wywołana przez kibiców na stadionie ma cechy i dynamikę, której nie mają kibice; układ neuronów przejawia cechy i dynamikę, której nie mają neurocy, etc.). Badając makroskopowe wzorce koncentrujemy się na cechach, które nie mają odpowiedników na poziomie substratu. Zjawiska niezależne od substratu określa się także mianem „zjawisk emergentnych”. Zdaniem Tegmarka są one powszechne w przyrodzie.

⁵ W odpowiedzi na pytanie, jakie pojęcie lub termin naukowy powinien być bardziej znany we współczesnym świecie („What Scientific Term or Concept Ought to Be More Widely Known?”), Tegmark (2017) wskazał na „niezależność od substratu” („Substrate-Independence”).

Zasada niezależności od substratu odgrywa fundamentalną rolę w rozumieniu wielu zjawisk fizycznych, takich jak informacja, obliczanie, pamięć, uczenie się, inteligencja czy świadomość. Informację możemy kodować w różnych substratach, o ile są one odpowiednio zorganizowane: „każdy kawałek materii może stanowić substrat dla pamięci, pod warunkiem, że ma wiele różnych stanów stabilnych” (TEGMARK 2019, 110). W kodowaniu informacji istotny jest makroskopowy wzorzec, a nie szczegóły substratu, w którym jest on zrealizowany. Dzięki temu „informacja może żyć własnym życiem niezależnie od jej fizycznej realizacji” (TEGMARK 2019, 82). Można ją przechowywać i przetwarzać w układach biologicznych i sztucznych, jeśli dysponujemy stanami pamięci i funkcjami umożliwiającymi ich przekształcanie. Mając odpowiednio złożone funkcje (algorytmy, programy), możemy budować inteligentne maszyny realizujące złożone cele. Materia staje się inteligentna, jeśli wykazuje dynamikę umożliwiającą magazynowanie i przekształcanie informacji zgodnie z określonymi funkcjami. Układy fizyczne mogą realizować każde dobrze zdefiniowane obliczenie (PICCINI 2015) i mogą to robić na wiele różnych sposobów: „obliczenia są niezależne od substratu, tak jak informacje: mogą żyć własnym życiem, niezależnie od ich fizycznej realizacji” (TEGMARK 2019, 91).

Podstawową zdolnością układu inteligentnego jest uczenie się, które polega na reorganizacji wewnętrznego porządku, co daje mu nowe możliwości obliczeniowe. Sztuczne sieci neuronowe potrafią obliczyć każdą funkcję z dowolną dokładnością reorganizując wyrażane liczbowo wagi połączeń synaptycznych. Układy fizyczne mogą spełniać prawa fizyki, kodować i przetwarzać informacje, a także pamiętać, obliczać i uczyć się — i nie muszą to być układy biologiczne (TEGMARK 2019, s. 109). Jedną z konsekwencji niezależności od substratu jest to, że takie zjawiska jak inteligencja wydają się niefizyczne: „[inteligencja] sprawia wrażenie tak niefizycznej, ponieważ jest niezależna od substratu, żyjąc własnym życiem, które nie zależy od fizycznych szczegółów ani ich nie odzwierciedla. Krótko mówiąc, obliczanie jest wzorcem w czasoprzestrzennym układzie cząstek, i to nie cząstki, ale ich wzorzec ma znaczenie! Materia się nie liczy” (TEGMARK 2019, 93).

Pozór niefizyczności jest jeszcze silniejszy w wypadku świadomości. Tegmark posługuje się pojęciem świadomości fenomenalnej w sensie Chalmersa. Bycie świadomym polega na posiadaniu subiektywnych doznań zmysłowych oraz przeżyć towarzyszącym wyższym procesom poznawczym. Z punktu widzenia fizyki istotne jest pytanie, jakiego rodzaju układy są świadome. Wiemy z doświadczenia, że organizacja mózgu umożliwia posiadanie przeżyć świadomych, mimo że zdecydowana większość procesów przetwarzania informacji w mózgu ma cha-

rakter nieświadomy. Naukowa teoria świadomości powinna umożliwiać predykcję, które układy fizyczne są świadome oraz w jakim dokładnie stanie fizycznym powinny się znajdować, by mieć świadomość. Dzięki mocy predykcyjnej teoria świadomości będzie empirycznie testowalna (TEGMARK 2019, 368–373).

Badania dotyczące neuronalnych korelatów świadomości koncentrują się na obwodach neuronowych. Tymczasem ogólna teoria świadomości poszukuje fizycznych korelatów świadomości rozumianych jako wzorce przetwarzania informacji, które mogą być świadomie doświadczane. Taka teoria powinna mieć formę matematyczną, jej zaś równania powinny przewidywać i wyjaśniać wyniki eksperymentów związanych ze świadomością w mózgach i maszynach. Jak pisze Tegmark: „Jeśli teoria potrafi poprawnie przewidzieć, co jest świadome, a co nie, odwołując się tylko do fizycznych elementów składowych, takich jak cząstki elementarne i pola sił, to może tworzyć ona przewidywania nie tylko dla mózgow, ale także dla wszelkich innych układów materii, w tym dla systemów sztucznej inteligencji w przyszłości” (TEGMARK 2019, s. 384).

W przekonaniu Tegmarka świadomość może zostać wytworzona przez procesy fizyczne umożliwiające zaawansowane przetwarzanie informacji, pomimo tego, że własności świadomości wykraczają poza własności części układów, w których jest ona wytwarzana. Świadomość, jak wiele innych zjawisk fizycznych, jest zjawiskiem emergentnym, którego cechy wykraczają poza właściwości części układów, którym ona przysługuje (TEGMARK 2014; 2015). Pojedyncze atomy, molekuly czy neurony nie są świadome, ale ich odpowiednio zorganizowane układy mogą mieć przeżycia świadome. Emergencja jest powszechna w świecie fizycznym, zaś świadomość jest jej spektakularnym przykładem (TEGMARK 2019 385).

Tegmark, podobnie jak Chalmers i zwolennicy obliczeniowych teorii świadomości, wiąże pojęcia świadomości i informacji. Jego zdaniem najciekawszą matematyczną koncepcję świadomości zaproponował Giulio Tononi w formie teorii zintegrowanej informacji. Zgodnie z tą teorią, aby informacja mogła być świadomie przetwarzana, musi osiągnąć odpowiedni poziom integracji, symbolicznie oznaczony za pomocą parametru Φ . Zdaniem Tononiego należy tak określić wartość Φ , by za pomocą odpowiednich metod pomiarowych była możliwa identyfikacja świadomego przetwarzania informacji⁶.

Teoria Tononiego stwarza pewne nadzieje. Jak dotychczas udało się z wykorzystaniem jej detektora świadomości odróżnić nieświadomą aktywność mózgu

⁶ Teoria zintegrowanej informacji jest rozwijana od przeszło dwudziestu lat. Narosła wokół niej olbrzymia literatura naukowa i filozoficzna. Z punktu widzenia problemów poruszanych w tym artykule szczególnie ważna jest praca „Integrated Information Theory: From Consciousness to Its Physical Substrate” (TONONI, MASSIMINI i KOCH 2016), w której autorzy bronią częściowej niezależności świadomości od substratu.

podczas snu głębokiego i stanu śpiączki farmakologicznej od stanu świadomego na podstawie różnych poziomów integracji informacji przetwarzanej w tych stanach. Udało się również zidentyfikować świadome przetwarzanie informacji u pacjentów w zespole zamknięcia. Zgodnie z teorią Tononiego obecne komputery nie mają takiego poziomu zintegrowanej informacji, by można je było uznać za świadome, co jest zgodne z powszechną intuicją. Natomiast problemem tej teorii jest to, że na jej podstawie musielibyśmy uznać za świadome układy fizyczne odznaczające się wysokim poziomem zintegrowanej informacji, które intuicyjnie wydają się całkowicie nieświadome⁷. Zdaniem Tegmarka teoria zintegrowanej informacji powinna być nadal rozwijana. Niezależnie od niej warto jednak tworzyć alternatywne teorie fizycznych korelatów świadomości w układach biologicznych i sztucznych (TEGMARK 2019, 392–395).

Jak na razie nie dysponujemy teorią określającą warunki wystarczające świadomego przetwarzania informacji. Możemy natomiast formułować warunki konieczne. Tegmark wskazuje na cztery takie warunki, które określa jako: (i) zasadę informacji (układy świadome muszą dysponować odpowiednio dużą pojemnością pamięci), (ii) zasadę dynamiki (układy świadome muszą znajdować się w wielu stanach dynamicznych), (iii) zasadę niezależności (układy świadome powinny przetwarzać informację w sposób względnie niezależny od środowiska), (iv) zasadę integracji (elementy układu świadomego muszą pozostawać w silnych oddziaływaniach, aby zapewnić wysoki poziom zintegrowanej informacji). Ogólna teoria świadomości powinna uwzględnić powyższe zasady oraz przybrać postać rygorystycznej teorii matematycznej, którą można testować eksperymentalnie (TEGMARK 2019, 390–391).

Powracając do zagadnienia niezależności świadomości od substratu, Tegmark podkreśla, że jeśli proces przetwarzania informacji spełnia określone warunki, może dać początek świadomości jako zjawisku emergentnemu. Samo przetwarzanie informacji jest już procesem niezależnym od substratu, świadomość zaś znajduje się na jeszcze wyższym poziomie niezależności, z czym wiąże się silne, choć iluzoryczne poczucie, że jest ona czymś niefizycznym:

⁷ Niektóre wypowiedzi Giulio Toniego i Christopa Kocha mogą sugerować, że są zwolennikami panpsychizmu. Jest to jednak interpretacja niezgodna z duchem teorii zintegrowanej informacji, która ma odróżniać świadome przetwarzanie informacji od przetwarzania nieświadomego. Bardziej trafna jest interpretacja emergentystyczna i funkcjonalistyczna, co wykazują: Ignacio Cea (2020) i Niccolò Negro (2022). Zgodnie z funkcjonalizmem emergencyjnym świadomość wymaga określonego poziomu integracji przetwarzanych informacji, który można osiągnąć w układach biologicznych i sztucznych.

Świadomość jest zjawiskiem fizycznym odczuwanym niefizycznie, ponieważ ma taki sam charakter jak fale i obliczenia: ma właściwości niezależne od swojego konkretnego podłoża fizycznego. Wynika to logicznie z koncepcji świadomości jako informacji. Konstatacja ta prowadzi do radykalnej idei, którą bardzo lubię: jeśli świadomość to sposób, w jaki odczuwa się informację, gdy jest przetwarzana w określony sposób, to musi być niezależna od podłoża; liczy się tylko struktura przetwarzania informacji, a nie struktura materii dokonującej tego przetwarzania. Innymi słowy, świadomość jest podwójnie niezależna od podłoża. (TEGMARK 2019, 389–390).

Tegmark twierdzi, że świadomość jest zjawiskiem fizycznym odczuwanym w sposób, który sprawia pozór niefizyczności. Jest ona fizyczna jak inne zjawiska emergentne niezależne od substratu – fale, informacje, pamięć, obliczanie, uczenie się i inteligencja. Świadomość jest zjawiskiem emergentnym, występującym na poziomie układów o wysokiej złożoności oraz integracji wewnętrznej, takich jak mózgi biologiczne (świadomość naturalna) i systemy sztucznej inteligencji (świadomość maszynowa). Chalmers i Tegmark dopuszczają możliwość istnienia świadomej sztucznej inteligencji. Przy tym Chalmers ogranicza się do stwierdzenia samej możliwości, Tegmark zaś formułuje śmiało hipotezy na jej temat:

(i) przeżycia świadome wysoko rozwiniętej sztucznej inteligencji daleko wykraczałyby poza to, czego doświadczają ludzie – takie systemy mogłyby mieć wewnętrzne i zewnętrzne sensory dające dostęp do jakości nieosiągalnych dla człowieka,

(ii) ponieważ sygnały elektromagnetyczne w sztucznym systemie rozchodzą się z prędkością światła, miałyby one znacznie więcej przeżyć w jednostce czasu niż układy biologiczne, w których przepływ sygnałów jest powolny,

(iii) zdecydowana większość informacji przetwarzanych w systemach sztucznych miałaby charakter nieświadomy,

(iv) z uwagi na łatwość komunikacji oraz wymiany informacji i oprogramowania między systemami sztucznymi, ich świadomość byłaby w znacznie mniejszym stopniu zindywidualizowana niż u ludzi,

(v) systemy sztuczne podejmując decyzje miałyby wewnętrzne poczucie wolności wyboru, nawet w sytuacji, gdyby to był proces deterministyczny. Obliczeniowy mechanizm poczucia wolnej woli jest taki sam w systemach naturalnych i sztucznych. Jak pisze Tegmark (2019, 400): „ich subiektywne doświadczenie wolnej woli polega po prostu na tym, jak czują owe obliczenia od wewnątrz; nie znają wyniku obliczeń dopóty, dopóki obliczenia nie są skończone. To właśnie mamy na myśli mówiąc, że obliczenie jest decyzją”.

Jeden z problemów, z którego Tegmark zdaje sobie sprawę, polega na oddzieleniu uniwersalnych cech świadomości od cech specyficznych, związanych z okre-

ślonym substratem. Czy przeżycia doświadczane przez ludzi, takie jak ból, strach, pożądanie, wstyd, duma czy poczucie sensu istnienia, mogą być doświadczane także przez systemy sztuczne? Wydaje się to wątpliwe z powodu innej formy ucieleśnienia (fizycznej realizacji). Czy wobec tego nie należałoby wiązać z każdym typem substratu cech świadomości charakterystycznych dla danego gatunku? Byłyby to cechy świadomości silnie zależne od substratu, formy ucieleśnienia, historii osobniczej i gatunkowej oraz środowiska, w którym rozwija się świadomość. Oprócz cech gatunkowych należałoby wyróżnić cechy uniwersalne przysługujące każdej formie świadomości. Istniejące teorie świadomości nie pozwalają na przeprowadzenie takiej dystynkcji. Co gorsza, jedyną formą świadomości, jaką znamy z doświadczenia wewnętrznego, jest nasza własna świadomość.

W rozważaniach Tegmarka występuje niejasne, a zarazem kluczowe dla jego stanowiska twierdzenie. Chodzi o przytoczoną wcześniej charakterystykę świadomości jako sposobu, w jaki odczuwa się informację, gdy jest ona przetwarzana w określony sposób. O ile sposób przetwarzania informacji konieczny do wywołania przeżyć świadomych, Tegmark charakteryzuje w dość ogólnikowy sposób, to pojęcie odczuwania informacji nie doczekało się żadnej eksplikacji. Obecność tej charakterystyki w definicji świadomości powoduje, że staje się ona kolista — odczuwanie jest formą świadomości. Powstaje pytanie: Skąd się bierze owo odczuwanie informacji, od czego zależy i jak powstaje? Być może, podobnie jak w teorii Chalmersa, odczuwanie jest aspektem informacji przetwarzanej w określony sposób, a nie czymś od niej odrębnym. Gdyby było inaczej, teoria Tegmarka zakładałaby to, co ma wyjaśniać. Zagadnienie to wymaga dalszych analiz i dyskusji.

3. PAULA THAGARDA ARGUMENT PRZECIWKO TEZIE O NIEZALEŻNOŚCI OD SUBSTRATU

Paul Thagard sformułował interesujący argument podważający: (i) funkcjonalistyczną tezę o niezależności inteligencji od substratu, (ii) twierdzenie, że żyjemy w symulacji komputerowej, (iii) koncepcję zakładającą możliwość przeniesienia umysłu do komputera. Argument Thagarda (2022) stawia również pod znakiem zapytania możliwość budowy silnej sztucznej inteligencji, o ile tę możliwość wyprowadza się z tezy o niezależności inteligencji (informacji, obliczeń, myślenia, świadomości) od substratu⁸.

⁸ Thagard zwraca uwagę, że preferowany przez niego termin „niezależność od substratu” stał się popularny w ostatnich latach i jest wykorzystywany zarówno przez naukowców jak filozofów. Jeszcze do niedawna, w nawiązaniu do prac Hilarego Putnama i Jerry’ego Fodora, używano terminu

Zdaniem Thagarda funkcjoniści ignorują fakt, że energia jest konieczna dorealizacji procesów obliczeniowych w układach biologicznych i w komputerach. Zarówno mózgi, jak i urządzenia elektroniczne potrzebują energii, by wykonywać pracę. W wypadku mózgu niezbędna jest energia chemiczna (neurony są zasilane głównie przez glukozę), w wypadku zaś komputerów energia elektryczna. W odróżnieniu od obiektów abstrakcyjnych, takich jak maszyny Turinga, których nie dotyczą fizyczne ograniczenia, działanie mózgów i komputerów podlega wielorakim limitacjom związanym z czasem, przestrzenią i zasobami energetycznymi. Tymczasem zdaniem zwolenników tezy o niezależności od substratu nie ma znaczenia, jaki rodzaj energii jest wykorzystywany w układach biologicznych i w maszynach.

Argument przeciwko tezie o niezależności od substratu ma następującą postać:

- (i) przetwarzanie informacji w świecie fizycznym jest zależne od energii,
- (ii) energia jest zależna od materialnych substratów,
- (iii) zatem: przetwarzanie informacji jest zależne od materialnych substratów,
- (iv) zatem: teza o niezależności od substratu jest fałszywa (Thagard 2022, 73).

Thagard poświęca wiele miejsca uzasadnieniu przesłanek (i)-(ii). W szczególności analizuje zależność między informacją a energią/substratem w układach biologicznych. Energia jest niezbędna, by tworzyć, utrzymywać i przetwarzać informacje w organizmach; pozostaje też bezpośrednio związana z materialnym substratem, z którego zbudowane są organizmy. Jak pisze Thagard (2022, 75):

Energia i informacja są współzależne w tym sensie, że organizm potrzebuje energii, by pozyskiwać i wykorzystywać informacje, także w celu zdobywania pokarmu, który jest źródłem energii. Wszystkie organizmy oraz środowiska, w których one funkcjonują, są zależne od równowagi między energią i informacją. Informacja to nie tylko ustrukturowane dane, kody lub teksty; musi ona być przechowywana, przetwarzana, wysyłana i uzyskiwana w procesach, które wymagają energii”.

Przetwarzanie informacji wymaga energii, ta zaś jest zależna od materialnego substratu. Określenie „zależna” wskazuje, że nabywanie, przetwarzanie i wykorzystywanie informacji to procesy realizowane przez mechanizmy biologiczne składające się z części pozostających w relacjach przyczynowych, które wymagają odpowiednich zasobów energetycznych.

Tezę o zależności procesów informacyjnych od energii/substratu Thagard ilustruje również na przykładzie działania komputerów neuromorficznych i uczenia maszynowego. Ta zależność jest pomijana na gruncie matematycznej teorii obli-

„wieloraka realizacja”. Z nowszych opracowań zagadnienia wielorakiej realizacji na uwagę zasługują: PICCININI 2015; POLGER i SHAPIRO 2016; AIZAWA 2018.

czeń badającej obiekty abstrakcyjne, takie jak maszyny Turinga, które różnią się pod ważnymi względami od urządzeń obliczeniowych. Analiza obliczeń neuro-morficznych i uczenia maszynowego wskazuje na analogiczną zależność między informacją a energią/substratem jak w układach biologicznych. W obu wypadkach przetwarzanie informacji wymaga działania mechanizmów zależnych od energii materialnego substratu, w którym i dzięki któremu przetwarzanie jest realizowane (THAGARD 2022, 76–78).

Zwolennicy zasady niezależności informacji i obliczeń od substratu, a tym samym niezależności pamięci, uczenia się i świadomości, odwołują się do tezy Churcha-Turinga, która głosi, że każde efektywnie wykonalne obliczenie może zostać przeprowadzone przez maszynę Turinga, każdy zaś proces obliczeniowy może zostać zrealizowany przez każdy uniwersalny komputer (COPELAND 2017; TEGMARK 2019). Zdaniem Thagarda teza Churcha-Turinga może sugerować, że obliczanie jest niezależne od substratu, ponieważ te same obliczenia mogą zostać zrealizowane niezależnie od zastosowanej metody obliczeń. Jednak teza ta nie mówi niczego na temat fizycznych maszyn obliczeniowych oraz warunków, jakie muszą być spełnione, by takie maszyny mogły działać. Z tego powodu teza Churcha-Turinga nie stanowi argumentu na rzecz niezależności obliczeń od substratu. Pomija ona kluczową kwestię warunków fizycznych niezbędnych do realizacji procesów obliczeniowych przez urządzenia fizyczne, których działanie jest zawsze ograniczone czasem, przestrzenią, prawami przyrody i dostępem do energii (THAGARD 2022, 79).

Inny argument na rzecz niezależności od substratu wskazuje na możliwość stopniowego zastąpienia neuronów w mózgu syntetycznymi odpowiednikami — bez jakiegokolwiek zmiany na poziomie świadomości. Gdyby to było możliwe, substrat rzeczywiście nie miałby znaczenia. Neurony jednak nie są izolowanym jednostkami — wchodzi w złożone interakcje wymagające energii oraz substancji chemicznych umożliwiających te interakcje. Nie ma racjonalnych podstaw, by wierzyć, że możliwe jest zachowanie funkcji naturalnej sieci neuronowej po zastąpieniu jej elementów składowych syntetycznymi zamiennikami o innych właściwościach energetyczno-substratowych. Problematiczne jest już samo wszczepienie elementów syntetycznych do sieci zbudowanej z neuronów i substancji chemicznych. Krzemowe neurony miałyby problem z synchronizacją stanów z neuronami biologicznymi; nie wiadomo, na czym polegałyby interakcje i wymiana sygnałów między neuronami naturalnymi i sztucznymi, ani w jaki sposób te substratowo heterogeniczne grupy neuronów byłyby zasilane (energia chemiczna i energia elektryczna mają szereg odmiennych parametrów). Zdaniem Thagarda (2022, 80): „Opowieść o stopniowym zastąpieniu jest atrakcyjna dla tych, którzy już wierzą,

że substrat nie jest ważny dla myślenia i obliczania. Jednak bliższe przyjrzenie się mechanizmom, które działają w różnych systemach, pokazuje, że faktyczne zastąpienie jest problematyczne z powodu ograniczeń związanych z czasem, przestrzenią i energią. Abstrakcyjna możliwość stopniowego zastąpienia nie zwiększa wiarygodności tezy o niezależności od substratu”.

Stanowisko Thagarda wprowadza do dyskusji na temat substratu świadomości ważny wątek. Nie zawiera jednak argumentu definitywnie wykazującego fałszywość tezy o jej niezależności od substratu. Na zakończenie tej części rozważań sformułuję trzy uwagi krytyczno-polemiczne pod adresem zaprezentowanego stanowiska:

(i) Thagard nie twierdzi, że zasada niezależności od substratu jest zawsze fałszywa, a jedynie, że może być fałszywa w odniesieniu do niektórych przypadków. Problem polega na rozstrzygnięciu, czy niezależność świadomości od substratu rzeczywiście pomija uwarunkowania substratowo-energetyczne. Wydaje się, że zarówno Chalmers jak Tegmark zgodziliby się, że nie każdy substrat spełnia warunki energetyczne umożliwiające realizację procesów informacyjno-obliczeniowych (w tym świadomości). Potrzebujemy szczegółowej wiedzy, jakie substraty umożliwiają spełnienie warunków energetycznych koniecznych do realizacji procesów umysłowo-poznawczych. Gdyby okazało się, że wyłącznie substrat biologiczny, oparty na związkach węgla, stwarza warunki energetyczne konieczne do wytworzenia świadomości, teza o niezależności świadomości od substratu uległaby osłabieniu. Tego jednak wciąż nie wiemy i dlatego nie można wykluczyć, że inne rodzaje substratów także spełniają warunki konieczne do wytworzenia świadomości.

(ii) Zwolennicy tezy o niezależności świadomości od substratu zgodziliby się, że procesy informacyjne zachodzące w świecie fizycznym są zależne od energii, która jest niezbędna do zmiany stanów układu informacyjnego. Informacja, obliczenia, pamięć i uczenie się mogą zachodzić tylko w substracie o określonej organizacji i potencjale energetycznym. Przy czym możliwy jest więcej niż jeden rodzaj substratu, w którym realizowane są procesy informacyjne (PICININI 2015; AIZAWA 2018). Teorie obliczeniowe opisujące i wyjaśniające świadomość w kategoriach przetwarzania informacji opierają się na tezie o niezależności obliczeń od substratu — jeśli obliczenia są niezależne od substratu, to obliczeniowo rozumiana świadomość fenomenalna dziedziczy tę własność. Ponadto, energia, chociaż jest konieczna do działania układu świadomego, nie jest kategorią wyjaśniającą naturę świadomości. Zależność świadomości od energii/substratu pod pewnymi względami nie wyklucza jej niezależności pod innymi względami. Zwłaszcza, że energia występuje w różnych formach (w pewnym sensie również podlega

wielorakiej realizacji), które mogą zostać wykorzystane przez różne systemy informacyjne — naturalne, sztuczne i hybrydowe.

(iii) Neurobiolodzy odkrywający mózgowy mechanizm świadomości dopuszczają możliwość, że zostaną one zrealizowane również w systemach sztucznych, o ile mechanizmy te są implementacją organizacji funkcjonalnej wystarczającej do wywołania stanów świadomych. Stanislas Dehaene (2023) określa warunki, jakie muszą zostać spełnione, aby możliwe było wytworzenie sztucznej świadomości mającej poczucie wolnej woli. Autor zgrubnie charakteryzuje architekturę takiego systemu wzorowaną na organizacji mózgu oraz formułuje następującą konkluzję: „Przyznaję, że powyższe koncepcje są szkicowe. Przełożenie ich na szczegółowy schemat będzie wymagać wielkiej pracy. Niemniej co do zasady nie widzę powodu, aby nie miało to prowadzić do stworzenia sztucznej świadomości” (Dehaene 2023, s. 353). Tego typu opinie kształtują klimat intelektualny naszych czasów.

ZAKOŃCZENIE

W dyskusjach na temat niezależności świadomości od substratu pojawiają się dwa, najczęściej łączone, chociaż w pewnym zakresie niezależne od siebie wątki. Pierwszy dotyczy częściowej autonomii świadomości względem substratu — niezależnie od tego, z jakim substratem mamy do czynienia. Przeżycia świadome, jako makroskopowe zjawiska emergentne, mają dynamikę oraz cechy, które nie przysługują częściom jej mechanizmu. Rozumiana w ten sposób autonomia nie przekreśla tego, że istnienie i aktywność świadomości są wielorako zależne od energii/substratu. Niezależność pod pewnymi względami nie wyklucza zależności pod innymi względami.

Drugi wątek dotyczy świadomości maszynowej. Współczesna debata na ten temat przybrała zawrotne rozmiary (KRAUSS i MAIER 2020). Mimo to wciąż brakuje argumentów rozstrzygających. Gdyby przeżycia świadome mogły istnieć tylko w układach biologicznych, nie przekreślałoby to ich autonomii we wcześniej określonym znaczeniu. Gdyby zaś zbudowano świadomość niebiologiczną, zasada niezależności od substratu uzyskałaby nowy wymiar i uzasadnienie. Jednak i w tym wypadku niezależność byłaby tylko częściowa, z uwagi na nieuniknione uwarunkowania fizyczne. Ponadto, świadomość maszynowa znacznie różniłaby się od ludzkiej pod względem ilościowym (zakresem i szybkością przetwarzania) i jakościowym (sposobem doznawania i przeżywania).

Klimat intelektualny naszych czasów trafnie oddaje metafora Hansa Moraveca, zgodnie z którą krajobraz ludzkich kompetencji jest stopniowo zatapiany przez szybko podnoszącą się wodę symbolizującą ekspansję sztucznej inteligencji. Kolejne kompetencje poznawcze, niedawno zarezerwowane wyłącznie dla ludzi, stają się udziałem maszyn. Możliwe, że w niedalekiej przyszłości zostaną „zatopione” także kompetencje związane ze świadomością (TEGMARK 2019, 76). Jeśli rzeczywiście sprawy idą w tym kierunku, możemy pytać: Czy warto dążyć do wytworzenia świadomych maszyn? Czy do realizacji ludzkich potrzeb nie wystarczą inteligentne narzędzia pracujące w trybie *zombie*? Powstanie świadomych maszyn wiązałoby się z nadaniem im statusu moralnego i prawnego. Nie traktowalibyśmy ich jak narzędzi, lecz uznałibyśmy za osoby. Niewykluczone, że sztuczna świadomość rozwinięłaby jakąś formę emocji, a w swoich decyzjach kierowałaby się zaszczerpionymi lub rozwiniętymi wartościami. Wówczas różnica między inteligencją naturalną a sztuczną przestałaby być tak wyraźna. Jak twierdzi Frank Wilczek (2020, 85), laureat Nagrody Nobla w dziedzinie fizyki: „wszelka inteligencja jest maszynowa, a inteligencję naturalną od sztucznej odróżnia nie to, czym są, ale jak są zrobione”. Odnosząc to twierdzenie do świadomości powiedzielibyśmy, że świadomość biologiczna jest artefaktem ewolucyjnym wytworzonym w laboratorium Natury. Świadomość maszynowa byłaby zaś artefaktem technologicznym wyhodowanym w laboratorium sztucznej inteligencji i robotyki kognitywnej. Dyskusja pozostaje otwarta.

REFERENCJE

- AIZAWA, Kenneth. 2018. „Multiple Realization and Multiple ‘Ways’ of Realization: A Progress Report”. *Studies in the History and Philosophy of Science* 68: 3–9.
- BEDAU, Mark E. 1998. „Four Puzzles about Life”. *Artificial Life* 4: 125–140.
- BOLTRUĆ, Piotr. 2013. „BICA jako szansa stworzenia świadomych maszyn”. *Przegląd Filozoficzny. Nowa Seria* 22, nr 2 (86): 185–196.
- BOSTROM, Nick. 2003. „Are We Living in a Computer Simulation?”. *The Philosophical Quarterly* 53: 243–255.
- CEA, Ignacio. 2020. „Integrated information theory of consciousness is a functionalist emergentism”. *Synthese* 8 (1–2): 2199–2224.
- CHALMERS, David. 1994. „On implementing a computation”. *Minds and Machines* no. 4: 391–402.
- CHALMERS, David. 2010a. *Świadomy umysł. W poszukiwaniu teorii fundamentalnej*. Tł. Marcin Miłkowski. Warszawa: Wydawnictwo Naukowe PWN.
- CHALMERS, David. 2010b. „The Singularity: A Philosophical Analysis”. *Journal of Consciousness Studies* 17: 7–65.
- COPELAND, B. Jack. 2017. „The Church-Turing thesis”. W: *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/church-turing/>.

- DEHAENE, Stanislas. 2023. *Mózg i świadomość. Odczytywanie kodu naszych myśli*. Tłum. Dariusz Rossowski Kraków: Copernicus Center Press.
- GAMEZ, David. 2012. „Empirically grounded claims about consciousness in computers”. *International Journal of Machine Consciousness* 4(2): 421–438.
- HAIKONEN, Pentti O. 2007. *Robot Brains: Circuits and Systems for Conscious Machines*. Hoboken, NJ: Wiley & Sons.
- HAIKONEN, Pentti O. 2009. „Qualia and conscious machines”. *International Journal of Machine Consciousness* 1, no. 2: 225–234.
- HOLLAND, Owen. 2003. *Machine Consciousness*. Exeter: Imprint Academic.
- JUDYCKI, Stanisław. 2004. *Świadomość i pamięć. Uzasadnienie dualizmu antropologicznego*, Lublin: Towarzystwo Naukowe KUL.
- KRAUSS, Patrick, i Andreas MAIER. 2020. „Will We Ever Have Conscious Machines? *Frontiers in Computational Neuroscience* 14, 556544. DOI: <https://doi.org/10.3389/fncom.2020.556544>.
- KURZWEIL, Ray. 2012. *How to Create a Mind*. New York: Viking.
- NEGRO, Niccolò. 2022. „Emergentist Integrated Information Theory”. *Erkenntnis* 89 (5): 1–23. DOI: <https://doi.org/10.1007/s10670-022-00612-z>.
- OLSON, Eric T. 1997. „The Ontological Basis of Strong Artificial Life”. *Artificial Life* 3: 29–39.
- PICCINI, Gualtiero. 2015. *Physical Computation: A Mechanistic Account*. Oxford: Oxford University Press.
- POLGER, Thomas, i Lawrence SHAPIRO. 2016. *The Multiple Realization Book*. Oxford: Oxford University Press.
- Pokropski M. 2022. *Mechanisms and Consciousness*, Routledge: New York, London.
- TEGMARK, Max. 2014. „Consciousness is a state of matter, like a solid or gas”. *New Scientist* 222 (2964): 28–31.
- TEGMARK, Max. 2015. „Consciousness as a state of matter”. *Chaos, Solitons & Fractals*. 76: 238-270.
- TEGMARK, Max. 2017. „Substrate-Independence”. [Odpowiedź na pytanie: What Scientific Term or Concept Ought to Be More Widely Known?]. *Edge*, dostęp 20.03.2024, <https://www.edge.org/response-detail/27126>.
- TEGMARK, Max. 2019. *Życie 3.0. Człowiek w erze sztucznej inteligencji*. Tłum. Tomasz Krzyszoń, Warszawa: Prószyński i S-ka.
- THAGRAD, Paul. 2022. „Energy Requirements Undermine Substrate Independence and Mind-Body Functionalism”. *Philosophy of Science* 89: 70-88. DOI: <https://doi.org/10.1017/psa.2021.15>.
- TONONI, Giulio, Melanie BOLDY, Marcello MASSIMINI i Christoph KOCH. 2016. „Integrated Information Theory: From Consciousness to its Physical Substrate”. *Nature Reviews Neuroscience* 17(7): 450–461.
- WILCZEK, Frank. 2020. „Jedność inteligencji”. W: John Brockman, red. *Człowiek na rozdrożu. Sztuczna inteligencja: 25 punktów widzenia*. Tłum. Marcin Machnik. Gliwice: Helion, s. 85-95.

CZY ŚWIADOMOŚĆ JEST NIEZALEŻNA OD SUBSTRATU?

Streszczenie

Celem artykułu jest prezentacja, porównanie oraz analiza krytyczna dwóch współczesnych wersji zasady niezależności świadomości od substratu. Pierwsza z nich pochodzi od Davida Chalmersa, druga od Maxa Tegmarka. Obaj autorzy dopuszczają możliwość istnienia sztucznej świadomości

oraz zakładają, że przeżycia świadome mogą zostać zrealizowane na bazie niebiologicznego substratu. Chalmers jednak broni stanowiska antyfizykalistycznego, a nawet pewnej wersji panprotopsychizmu. Tegmark zaś jest zwolennikiem fizykalistycznie interpretowanej teorii emergencji. W końcowej części artykułu przedstawiam oraz oceniam Paula Thagarda argument przeciwko zasadzie niezależności od substratu. Konkluzja, jaką formułuję w zakończeniu, głosi, że wciąż brakuje nam danych empirycznych, by rozstrzygnąć toczący się spór.

Słowa kluczowe: świadomość; zasada niezależności od substratu; świadomość maszynowa; funkcjonalizm; obliczeniowe teorie świadomości

IS CONSCIOUSNESS SUBSTRATE INDEPENDENT?

S u m m a r y

The goal of the article is to present, compare, and critically analyze two contemporary versions of the principle of consciousness independence from substrate. The first one comes from David Chalmers, and the second from Max Tegmark. Both authors allow for the possibility of the existence of artificial consciousness and assume that conscious experiences can be realized on a non-biological substrate. However, Chalmers defends an anti-physicalist position, and even a certain version of panprotopsychism. Meanwhile, Tegmark advocates for a physicalist interpretation of the theory of emergence. In the final part of the article, I present and evaluate Paul Thagard's argument against the substrate independence principle. The conclusion I formulate at the end asserts that we still lack empirical data to settle the ongoing dispute.

Keywords: consciousness; the principle of substrate independence; machine consciousness; functionalism; computational theories of consciousness

Information about the Author: ROBERT POCZOBUT, PhD habil., Prof. at UwB — University of Białystok (UwB), Faculty of Philosophy, Department of Epistemology and Cognitive Science, head; graduate of KUL; correspondence address: Plac NZS 1, 15-420 Białystok; e-mail: r.poczobut@uwb.edu.pl; ORCID: <https://orcid.org/0000-0003-2591-7002>.