

„SZTUCZNA INTELIGENCJA” ZMIERZCHEM HEGEMONII LUDZKIEGO UMYSŁU?

Jeszcze kilkanaście lat temu oglądając filmy z gatunku science-fiction byliśmy zdumieni wizją przyszłości kreowaną przez ich twórców. Wydawało się, że wyobrażenia przyszłych losów świata były zdecydowanie przesadzone. Możemy wspomnieć choćby podbój galaktyk w „Gwiezdnym wojnach” George’a Lucasa, istnienie maszyn – zabójców w „Terminatorze” Jamesa Camerona czy imponujące, futurystyczne konsole komputerów w serialu „Star Trek” Gene’a Roddenbery’ego itp. Wielu z nas sądziło wówczas, że tego świata z pewnością nie ujrzymy w przyszłości. A jednak dzisiaj – pod koniec 2013 roku – korzystamy z wielu urządzeń i gadżetów, których istnienie, a zarazem funkcjonalność – jeszcze jakiś czas temu mogły być dla nas jedynie pobożnym życzeniem. Żadną nowością nie jest dzisiaj dotykowy ekran w telefonie komórkowym czy w komputerze, telewizja w systemie trójwymiarowym również przestała być luksusem, a niektórzy, wreszcie, budują „inteligentne” domy – „naszpikowane” wieloma systemami, które czynią je bardziej funkcjonalnymi, wygodnymi i bezpiecznymi.

Faktem jest, że pod względem techniki – ludzkość poczyniła ogromny postęp. Dzisiaj w wykonywaniu różnych prac zastępują człowieka o wiele bardziej niezawodne, wytrzymałe i niezwykle wydajne maszyny, komputery i inne urządzenia. Zaskakują nas, często sprawiając wrażenie działania inteligentnego, a naukowcy idą znacznie dalej, próbując stworzyć sztuczny umysł. Pozostaje jednak pytanie: czy dzisiaj powinniśmy się obawiać tego, że w niedalekiej przyszłości władzę nad światem przejmą świadome maszyny? Czy to oznacza, że „Sztuczna Inteligencja” położy kres ludzkiej dominacji wśród istot żyjących we wszechświecie, a zrealizuje się jeden z mrocznych scenariuszy filmów fantasy?

Inspiracją niniejszego opracowania są klasyczne, a wręcz „kultowe” dla filozofii przyrody oraz filozofii nauki teksty J. R. Searle’a: „Umysł, mózg i nauka” (1980)¹ oraz A. M. Turinga: „Maszyny liczące a inteligencja” (1950)². Celem tego szkicu jest przybliżenie znaczenia „Sztucznej Inteligencji” w kontekście postawionego w tytule problemu.

¹ J. R. SEARLE, *Umysł, mózg i nauka*, przeł. J. Bobryk, Warszawa 1995.

² A. M. TURING, *Maszyny liczące a inteligencja*, tłum. D. GAJKOWICZ, w: *Maszyny matematyczne i myślenie*, red. E. FEIGENBAUM, J. FELDMAN, Warszawa 1972.

„Sztuczna Inteligencja” jest rozumiana jako kierunek badań naukowych, budzący od samego początku wiele kontrowersji, pomimo tego, iż jest w pewnym stopniu ugruntowany logicznie. Badania nad „Sztuczną Inteligencją” prowadzi wiele dziedzin wiedzy, takich jak: logika, informatyka, psychologia, matematyka, fizyka oraz rzecz jasna - filozofia. Synkretyczny charakter badań tłumaczy spore zainteresowanie naukowców i wciąż otwarty konflikt pomiędzy zwolennikami i przeciwnikami jej istnienia. Sam termin „Artificial Intelligence” (z ang. *Sztuczna Inteligencja*) wprowadził w 1955 r. amerykański naukowiec – John McCarthy, początkowo jako nazwę dla konferencji naukowej w Dartmouth College. Podejmowana problematyka była na tyle interesująca, że w ciągu kilku lat zwołano Międzynarodową Radę do Spraw „Sztucznej Inteligencji”. W dobie coraz szybszego rozwoju technologicznego, możliwość zbudowania komputera bardziej inteligentnego od człowieka wydawała się tylko kwestią czasu. Czy sztuczny umysł może być naprawdę inteligentny? Czy systemy formalne (oparte na wiedzy) są dobrym kandydatem dla stworzenia takich urządzeń?

Idea myślącej maszyny wydaje się dla nas mało prawdopodobna, jednak istnieją poglądy aprobujące taki stan rzeczy. Wielu badaczy uznaje komputer z uruchomionym nań oprogramowaniem za „umysł w sensie ścisłym, w którym zachodzą stany intelektualne”³. Z tezy tej wynika konieczność rozumienia przez komputery znaczenia ciągów symboli, samoistnie przez nie generowanych. Jednak, aby przekonanie to było prawdziwe, potrzeba jego intersubiektywności. Maszyny muszą zrozumieć materiał poznawczy, by ich „inteligencja” była osiągalna. Przeciwnicy „myślenia maszyn” uważają, że programy formalne nie umożliwiają rozumienia znaczenia znaków. W jaki więc sposób symbole w systemach formalnych nabierają znaczenia? Czy maszyna uzbrojona w algorytmy i wcześniej zdefiniowane sposoby segregacji informacji może nosić miano „Sztucznej Inteligencji?”

Zdaniem J. Życińskiego, przypisywanie maszynom przekonań i zdolności myślenia prowadzi do wizji świata bliskiej starogreckim mitom, w których wybuchy wulkanów, wyładowania atmosferyczne, czy inne kataklizmy uważano za przejaw uczuć, gniewu lub agresji okazywanych przez siły przyrody⁴. Maszyna nie może dorównać w swej głębi człowiekowi, gdyż jej nie posiada. Autor „Granice racjonalności” jako jeden z pierwszych polskich naukowców odniósł się do wyważonych poglądów amerykańskiego filozofa języka – Johna Searle’a i jego słynnego argumentu „Chińskiego Pokoju”, jednego z najistotniejszych passusów jego dzieła. Znaczenie argumentu potwierdzają liczne pozytywne reakcje środowisk naukowych. Publikacja wywołała serię licznych odpowiedzi uczonych z kręgu „AI”. „Chiński

³ J. KLOCH, *Świadomość komputerów?*, Tarnów 1996, s. 13.

⁴ Por. J. ŻYCIŃSKI, *Granice racjonalności. Eseje z filozofii nauki*, Warszawa 1993, ss. 234-240.

Pokój” był sprzeciwem wobec tzw. mocnego programu sztucznej inteligencji, zakładającego rozumowanie komputera na sposób ludzki.

Słynny eksperyment naukowy wykazał „błądność twierdzenia o posiadaniu przez komputer stanów umysłowych i możliwości wyjaśnienia funkcji ludzkiego umysłu przy pomocy programów komputerowych”⁵. Searle utrzymywał stałą nierealność zrozumienia czegokolwiek przez zaprogramowany komputer ze względu na brak intencjonalności – tj. własności stanów umysłowych, przez które są one ukierunkowane ku, lub wokół tego, co się dzieje w świecie. To właśnie intencjonalność jest własnością umysłu, odróżniającą ludzkie myślenie od działania programu komputerowego. Według Searle’a, jest to fakt empiryczny, biorąc pod uwagę relacje przyczynowe między mózgiem a procesami umysłowymi. W relacji komputer – program, intencjonalność nie występuje⁶. Intencjonalność systemu przetwarzania informacji to innymi słowy uznanie, że symbole, którymi operuje umysł, odnoszą się do czegoś; wymagane jest uznanie relacji pomiędzy obiektami w świecie rzeczywistym i w świecie umysłu. Maszyna nie jest w stanie wytworzyć, zidentyfikować ani nawet określić takich relacji. Interesującym jest, że dotychczasowe osiągnięcia informatyków, wychodząc naprzeciw oczekiwaniom – umiejętnie imitują sposób działania ludzkiego umysłu, a zwłaszcza procesy rozwiązywania problemów.

Kamieniem milowym w konstruowaniu myślących maszyn jest dzieło A. M. Turinga, eksperyment będący próbą formalnego zdefiniowania „Sztucznej Inteligencji”. Abstrakcyjna „maszyna Turinga” początkowo była w stanie wykonywać tylko jedną zaprogramowaną matematyczną operację, czyli algorytm, np.: podnieść liczbę do kwadratu, podzielić, dodać, odjąć, itd. Jednak stworzył on też uniwersalną maszynę, która w zależności od instrukcji zapisanej na taśmie, miała wykonywać dowolną operację. Tym samym udowodnił, że nie istnieje algorytm dotyczący nierozstrzygalności każdego twierdzenia. A zatem nawet „uniwersalna maszyna Turinga” nie była w stanie określić nierozstrzygalnych stwierdzeń⁷.

Znacznym problemem w pracach nad prostymi programami komputerowymi, np.: do gry w szachy, warcaby czy brydża był początkowo fakt, że dla maszyny wszystkie ruchy były jednakowo dobre. Program przegrywał zakładając, że przeciwnik zrobi ruch najgorszy z możliwych (co w rzeczywi-

⁵ J. KLOCH, tamże, s. 21.

⁶ Por. J. R. SEARLE, *Czy komputery mogą myśleć?*, w: tenże, *Umysł, mózg i nauka*, dz. cyt., ss. 13-15 (za dokumentem tekstowym pobranym z sieci).

⁷ Warto zajrzeć do kilku pozycji odnoszących się bezpośrednio do pracy A. M. Turinga. Por. J. D. BOLTER, *Człowiek Turinga: kultura Zachodu w wieku komputera*, przeł. T. GOBAN-KLAS, Warszawa 1990; A. HODGES, *Enigma: życie i śmierć Alana Turinga*, przeł. W. BARTOL, Warszawa 2002; P. ŁUPKOWSKI, *Test Turinga: perspektywa sędziego*, Poznań 2010.

stości raczej się nie zdarza). Podobnie, mało prawdopodobne jest, że zrobi ruch najlepszy z możliwych – potrzebny był zatem kompromis. Z tego zapotrzebowania narodziła się koncepcja maszyn przystosowujących się do warunków na bieżąco, tj. uczących się. W późniejszym czasie powstało kilka dość obiecujących programów, jednak ostatecznie okazało się, że sprawność ich działania zależała wyłącznie od umiejętności programisty opisującego problem. Kolejny raz przekonano się, że maszyna sama w sobie nie potrafi myśleć. W tym miejscu należałoby powrócić do argumentu „Chińskiego Pokoju”.

Na samym początku jego analizy, niezbędne jest wyjaśnienie pojęcia „rozumienia”. Po odrzuceniu złożonych analiz językowych, Searle przystępnie tłumaczy „rozumienie” jako pewien zakres działania podmiotu. „Rozumiem opowiadanie po angielsku, w mniejszym stopniu po francusku, w ogóle zaś po chińsku (...) W sensie dosłownym, zaprogramowany komputer rozumie tyle, co samochód lub kalkulator, a mianowicie - dokładnie nic”⁸. Głównym celem argumentu było wykazanie niedoskonałości programu komputerowego do wytworzenia intencjonalności. W uzasadnieniu tej tezy, Searle zastąpił program człowiekiem i postawił pytanie: co by się stało, gdyby umysł człowieka pracował tak, jak by odpowiadał na pytania? Po przeanalizowaniu tego eksperymentu doszedł do wniosku, że człowiek, który w ogóle nie znał języka chińskiego, nie potrafił go odróżnić od innych. Korzystając z dostępnych dla niego instrukcji, poprawnie odpowiadał na stawiane mu w tym języku pytania. Zupełnie inaczej rzecz się miała, gdy odpowiadał on w ojczystym języku. Dla kogoś z zewnątrz, udzielałby on wówczas właściwych odpowiedzi w obu językach. Zachodzi tu jednak istotna różnica między daniem odpowiedzi po chińsku i po angielsku. W wypadku tego drugiego języka, zachodziło pełne rozumienie, podczas gdy odpowiedzi w języku chińskim były „produkowane” bez zrozumienia, na zasadzie dobierania jednych niezinterpretowanych znaków do drugich, według ściśle określonych instrukcji (programu)⁹. W tym przypadku człowiek zachowywał się jak komputer – wykonywał jedynie operacje na chińskich symbolach, zastępując czy naśladowując konkretny program.

W ciągu następnych lat „Chiński Pokój” został modyfikowany przez autora, w celu uproszczenia struktury logicznej argumentu, zachowując swoją zasadniczą treść. Eksperyment ten wykazał, że z czysto formalnych operacji na symbolach (a właśnie w taki sposób działają programy komputerowe), nie wynika treść myślowa. Ponadto stał się bodźcem do powstania innych eksperymentów myślowych, takich jak „Koreański Pokój” czy „Chińska Sala Gimnastyczna”. Naukowcy doszli do wniosku, że aby

⁸ J. KLOCH, tamże, s. 22.

⁹ Por. tamże, s. 23-24.

maszyna mogła być inteligentna, jej sztuczne systemy muszą naśladować system biologiczny mózgu. Jednak nawet zakładana przez „Chińską Salę Gimnastyczną” ludzka sieć neuropodobna nie doprowadzi do interpretacji semantycznej zadania, nie utworzy umysłu.

Jak trudnym zadaniem jest stworzenie AI, wiedzą współcześni naukowcy. Coraz bardziej zaawansowane badania ludzkiego mózgu dają więcej odpowiedzi, jednak nadal nie wystarczają do stworzenia choćby częściowej „imitacji” mózgu, w postaci elektronicznego systemu operacyjnego. Programy ani nie kształtują myślenia, ani do niego nie wystarczają. System zdolny do spowodowania umysłu musiałby mieć siłę przyczynową (przynajmniej) równoważną mózgowi. Ponadto sposób, w jaki mózg ludzki wytwarza zjawiska umysłowe, nie może polegać jedynie na wykonywaniu programu komputerowego¹⁰. Ambicja stworzenia „Sztucznej Inteligencji” jeszcze przez pokolenia będzie spędzać naukowcom sen z powiek, gdyż jak na razie nie zapowiada się spektakularny przełom w tej dziedzinie. Maszyna może być najwyżej modelem umysłu, ale nigdy nim samym.

Bezapelacyjnie jesteśmy świadkami epokowego skoku rozwoju nauk przyrodniczych. W ostatnich dziesięcioleciach naukowcy dokonali wielu imponujących odkryć, a cenne jest dla nas to, że uczeni za każdym razem sygnalizowali pragnienie pokonywania kolejnych granic. Współczesny świat, w którym na co dzień żyjemy, wypełniony jest komputerami, maszynami liczącymi, „nafaszerowany” elektroniką dla naszego dobra, dla wygodniejszego życia. Mimo wszystko dobrze wiemy, że nawet szczyt osiągnięć współczesnej nauki i techniki – urządzenie skonstruowane przez człowieka, zawsze będzie jedynie pochodną jego umysłu. Choć zauważamy pewne symptomy i zwiastuny „Sztucznej Inteligencji” lub chcąc być bardziej precyzyjni – jej odblaski i przejawy, nie powinniśmy ulegać niepokojom, podobnie jak apokaliptycznym scenariuszom, ferującym ludzkiej cywilizacji. Nieznana przyszłość, w której myślące maszyny będą wykorzystywać gatunek ludzki, chociażby dla pozyskania energii – jak widzieli to Larry i Andy Wachowscy w „Matrixie” – na pewno nam nie grozi. Przedstawione argumenty dotyczące problemu intencjonalności i innych aktualnych trudności – wyzwania „Sztucznej Inteligencji” dowodzą z jednej strony, że jesteśmy na dobrej drodze poszukiwań, z drugiej jednak - jawi się niepewność, jaki duży dystans pozostał do pokonania.

¹⁰ Por. tamże, s. 38.