

Maria PIESKO

## O SUBTELNEJ RÓŻNICY MIĘDZY SŁABĄ A MOCNĄ WERSJĄ KONCEPCJI SZTUCZNEJ INTELIGENCJI NA PRZYKŁADZIE TEKSTU TURINGA

Słynny artykuł Alana Turinga z 1950 roku *Computing Machinery and Intelligence* uznawany jest powszechnie za kamień węgielny badań nad sztuczną inteligencją<sup>1</sup>. Zazwyczaj również uważa się jego autora za jednego z głównych przedstawicieli tzw. *silnej SI*<sup>2</sup>. Czy jednak wspomniany artykuł zawiera wystarczające przesłanki dla takiej tezy — czy jest on pisany w duchu *silnej* koncepcji SI?

Dyskutowany jest w nim problem „Czy maszyny mogą myśleć?”. Wydaje się więc, że sposób rozstrzygnięcia tej kwestii przez Turinga powinien pozwolić na łatwe zakwalifikowanie go do zwolenników *słabej* lub *mocnej* wersji SI. Spróbujmy zatem przeanalizować jego odpowiedź.

Logika tekstu przypomina odrobinę schemat scholastycznych wywodów. Wyraźnie rozpada się on na trzy części: postawienie kwestii

---

\*UWAGA: Tekst został zrekonstruowany przy pomocy środków automatycznych; możliwe są więc pewne błędy, których sygnalizacja jest mile widziana (obi@opoka.org). Tekst elektroniczny posiada odrębną numerację stron.

<sup>1</sup>Określenie „sztuczna inteligencja” weszło na stałe do słownika w 1956 r., kiedy to McCarthy zaproponował taki temat konferencji (ang. *Artificial Intelligence*). Dalej w tekście, jak przyjęto w literaturze przedmiotu, będę używała skrótu SI.

<sup>2</sup>Choć określenie to weszło do powszechnego użycia, gwoli ustalenia uwagi, za Searle’em przez koncepcję *silnej SI* rozumiem będziemy pogląd, zgodnie z którym *umysł jest tym samym dla mózgu, czym program dla komputera* (Searle 1995). Oczywiście sam termin „sztuczna inteligencja” jak i jej klasyfikacje powstały znacznie później niż omawiany artykuł. Turing po raz kolejny stworzył bardzo zaawansowaną koncepcją czegoś, co wcześniej nie istniało.

i wstępne jej rozwiązanie, dyskusję argumentów i wreszcie pełną prezentację poglądów autora. Wydaje się, że w miarę rozwoju argumentacji Turing coraz bardziej radykalizuje swoje stanowisko.

Podana na początku propozycja „gry w udawanie”, która miałaby być kryterium „myślenia” maszyn, jest powszechnie znana, choćby z polemiki jaką podjął Searle (por. np. Searle, 1995; lub Kloch, 1998) Przypomnijmy krótko: gra polega na tym, że jej uczestnik, rozmawiając **pisemnie** z człowiekiem i komputerem, nie widząc ich, nie słysząc itp., ma na podstawie otrzymywanych odpowiedzi odgadnąć, który z rozmówców jest człowiekiem. Jeśli maszynie uda się oszukać zadającego dowolne pytania gracza, to stanowić to będzie wystarczający dowód jej zdolności myślenia.

Jak sądzę, taka propozycja kryterium może być przyjmowana przez zwolenników tak *mocnej*, jak i *słabej* wersji SI, ale bliższa jest duchowi tej ostatniej. Podobnie bowiem jak w koncepcji *słabej SI* nie stwierdza się tu wcale, że mózg człowieka jest swego rodzaju maszyną wykonującą program „umysłu”, a jedynie rozważa możliwość, czy komputery będą mogły wykonywać czynności znamionujące wedle powszechnej opinii istoty inteligentne.

Podobne motywy znaleźć można także w krytyce argumentów przeciw myśleniu maszyn. Kiedy dyskutowany jest, jak się wydaje, jeden z kluczowych problemów — świadomość, Turing podkreśla, że wcale nie neguje faktu istnienia „zagadki świadomości”. Zwraca jedynie uwagę, że z jednej strony weryfikacja, czy ktoś posiada świadomość, jest niewykonalna<sup>3</sup>, z drugiej zaś strony rozwiązanie tejże „zagadki” nie wydaje się wcale konieczne dla odpowiedzi, czy maszyny mogą myśleć (zwłaszcza przy użyciu kryterium „gry w udawanie”).

Powyższa analiza mogłaby być przytoczona jako przykład stanowiska *słabej SI*. Wydaje się jednak, że zupełnie inaczej brzmi większość

---

<sup>3</sup>Konsekwentne żądanie weryfikacji musiałyby doprowadzić do soplizmu. *W myśl najbardziej radykalnej wersji [...], pewność co do tego, że jakaś maszyna myśli, można by osiągnąć tylko dzięki temu, że byłoby się tą maszyną i miało poczucie myślenia. Można by zakomunikować światu to poczucie, ale nikt nie musiałby brać tego na serio* (Turing, 1995, s. 284). Możemy być pewni jedynie tego, że sami posiadamy świadomość.

pozostałych argumentów. Przede wszystkim całość rozważań wydaje się mieć za podstawę porównanie zdolności ludzkich i możliwości komputerów<sup>4</sup> : ludzie bywają oryginalni, ale i maszyny potrafią zaskakiwać; istnieją problemy nierozstrzygalne dla komputerów, ale ktoś zagwarantuje, że ludzie potrafią je rozwiązać. Co więcej, choć można znaleźć człowieka mądrzejszego od każdego istniejącego w danym momencie komputera, to zdaniem Turinga, można skonstruować komputer, który będzie rozwiązywać dany problem lepiej niż ten człowiek.

W tych porównaniach Turing wydaje się coraz bardziej skłaniać ku widzeniu człowieka jak dość skomplikowanej maszyny. Koronny przykład takiego ujęcia pojawia się w rozważaniach „argumentu z nieformalności zachowania”, gdzie Turing zwraca uwagę, że oprócz zasad jakimi ludzie kierują się w postępowaniu (*rules of conduct*), które rzeczywiście trudno byłoby zaprogramować ze względu na ich różnorodność i nieokreśloność, istnieć mogą także prawa regulujące nasze zachowanie (*laws of behaviour*). Tych ostatnich nie jesteśmy zwykle świadomi, a być może stanowią one skończony zbiór, który dałby się sformalizować i przenieść na maszynę.

Znak równości między maszyną a człowiekiem, który zdaje się wyłaniać z tych rozważań, poprzez można i takim cytatem: *Któż może być pewny, że jego „twórcze dokonanie” nie wzrosło po prostu z nasienia, jakie zasiała w nim edukacja, albo nie jest następstwem kornej eksploatacji powszechnie znanych zasad ogólnych* (A.M. Turing, 1997, s. 289). W dodatku według Hartee’go nic nie stoi na przeszkodzie, by możliwe było zaprojektowanie maszyn z wbudowanym odruchem warunkowym i możliwością uczenia się. Turing nie tylko zgadza się z takim stanowiskiem, ale w zakończeniu artykułu podaje zarys projektu, jak należałoby się zabrać do takiego zadania. Chodziłoby przede wszystkim o napisanie programu, bo nie wydaje się, by skonstruowanie sprzętu porównywalnego z możliwościami naszego mózgu nastę-

---

<sup>4</sup>Używam słowa „komputer” rzecz jasna w obecnym znaczeniu. Może warto jednak zwrócić uwagę, że w tekście artykułu Turinga słowo „komputer” oznacza jeszcze człowieka wykonującego obliczenia (ang. „compute”).

czało większe trudności (sic!)<sup>5</sup>. Program powinien uwzględniać etap „dziecka” i „edukacji”, by doprowadzić do stworzenia systemu dorównującego zdolnościom intelektualnym dorosłego człowieka.

Wydaje się zatem, że istotnie Turing jest klasycznym przedstawicielem nurtu *silnej SI*, a przytoczone powyżej ostrożne sformułowania były jedynie wybiegiem mającym na celu uniknięcie zbyt ostrej krytyki. Należałoby przyjąć za A. Hodgesem (Hodges, 1998), że mimo pierwotnego przekonania o zasadniczych różnicach między ludźmi a maszynami Turing ok. roku 1941 zmienił swe poglądy i utożsamiał umysł z programem. Być może.

Być może jednak warto przynajmniej zastanowić się nad innym tropem. Jeszcze we wstępnej analizie „gry w udawanie” Turing wysuwa przypuszczenie, że maszyny mogłyby myśleć w inny sposób, niż czynią to ludzie. Zostawia je jednak bez dyskusji, zauważając, że także wówczas taki komputer mógłby z powodzeniem grać w „grę w udawanie”.

Turing krytykuje przywiązywanie zbyt wielkiej wagi do podobieństwa między neuronami a urządzeniami elektronicznymi, polegającego na wykorzystywaniu sygnałów elektrycznych. Jednak twierdzi, że można by się spodziewać znalezienia jakiegoś podobieństwa, ale na innym poziomie: *powinniśmy szukać raczej matematycznych analogii funkcjonalnych* (Turing, 1995, s. 277). A stąd znów krok tylko do teorii głoszącej, że najistotniejsza jest *informacja* a jakiego nośnika użyto do jej zapisu, czy jest to białko czy krzem, nie ma znaczenia.

Proponuję jednak, żeby spojrzeć na te zagadnienia z innej jeszcze strony. Zobrazujmy to podejście porównaniem: podobnie jak samolot lata, choć nie jest ptakiem, tak i komputery mogłyby „myśleć” nie będąc tym samym ludźmi. Jak sadzę, warto zwrócić uwagę na kilka istotnych aspektów tej analogii.

---

<sup>5</sup>Oszacowanie „parametrów” ludzkiego mózgu zamieszczone w omawianym artykule wypada niezwykle skromnie — jakieś  $10^{10}$ - $10^{15}$  bitów *I incline to lower values and believe, that only very small fraction is used for the higher types of thinking* (Turing, 1950, s. 455). Takżę prędkość przesyłania sygnału w neuronach jest niewielka w porównaniu z elementami elektronicznymi.

## 1. HOMO SAPIENS

Pytamy, czy komputer mógłby wykonywać to, co od początku stanowiło powód do dumy gatunku ludzkiego — myśleć. Czyż odpowiedź na to pytanie, nie wymaga zrozumienia, jak to się dzieje, że *Homo Sapiens* jest właśnie człowiekiem *myślącym*? (podobnie jak rozwój odpowiedniego działu fizyki był konieczny dla skonstruowania samolotów). Wydaje się, że w otaczającej nas rzeczywistości jesteśmy pod tym względem wyjątkowi i choć np. mechanizmy widzenia można było dość dobrze symulować dzięki analizie sposobu spostrzegania zwierząt, to wydaje się, że to co dla myślenia wydaje się najistotniejsze i najciekawsze zaobserwować możemy tylko na przykładzie człowieka.

Jeśli przyjmiemy, że najważniejsze w myśleniu jest poznawanie, rozumienie rzeczywistości, to naturalną konsekwencją tego będzie odzielenie „myślenia” jako sposobu poznawania świata od „sfery duchowej” człowieka — zagadnień wolności, odpowiedzialności, ale także emocji<sup>6</sup>. Nie wiadomo jednak, czy podział taki jest wykonalny.

Tak czy inaczej, to w czym podkreśliliśmy swą odrębność gatunkową to właśnie myślenie — *sapere*. Także Turing dostrzega, że źródłem zastrzeżeń co do możliwości intelektualnych maszyn może być właśnie przekonanie o naszej wyjątkowości w świecie — jeśli powstałby myślący komputer, to co stałoby się z tożsamością, odrębnością człowieka?

## 2. RACJONALNOŚĆ

Możliwość wznoszenia się samolotu w powietrze zależy od realizacji pewnych warunków aerodynamicznych. W przypadku „czynności” tak subtelnej jak myślenie, wkracamy w dziedzinę rozwiązywania problemów, a dalej w dziedziny logiki i matematyki. Z punktu widzenia zaproponowanej analogii można dostrzec, jak ważne dla problemu „myślenia maszyn” jest to, czy nasz kontakt poznawczy ze

---

<sup>6</sup>Znamienne jest, że w swoim artykule Turing niemal zupełnie pomija tego rodzaju zagadnienia.

światem ma charakter racjonalny a dalej — formalizowalny. Z jednej strony jest to inne oblicze postawionego wyżej pytania, czy myślenie da się wypreparować z całej sfery wolitywno-emocjonalnej, ale z drugiej strony jest to również pytanie o racjonalność świata. Może to stąd wynika obserwowalne w rozwoju badań nad SI coraz silniejsze poszukiwanie zrozumienia nie tylko samego procesu myślenia, ale i sposobu reprezentacji wiedzy, ogarnięcia ogromnego całokształtu danych — *rzeczywistości* — w racjonalne, sformalizowane i rządzące się precyzyjnymi regułami systemu informacji.

Można spojrzeć na to zagadnienie jeszcze z innej strony: często zwolennikom SI, a już z całą pewnością *silnej SI*, zarzuca się poglądy „deterministyczno-redukcjonistyczne”. Sądzę, że warto zwrócić uwagę na zakres na jaki ów domniemany redukcjonizm miałby się rozciągać: w skrajnej wersji dotyczyłby oczywiście budowy mózgu, który dałoby się opisać w postaci skończonej teorii i następnie zgodnie z jej prawami zbudować.

Drugim polem do formalizacji byłaby cała rzeczywistość z wyłączeniem mózgu. Wtedy, jeśli przez myślenie próbować by rozumieć właśnie sposób „przyswajania świata”, „tłumaczenia go na swoje, na formułę dla siebie zrozumiałą”, to poszukiwania rozmaitych sposobów reprezentacji wiedzy, byłoby tym samym próbą „przerzucenia mostu do rzeczywistości” niezależnie od naturalnego (wypracowanego przez ewolucję) sposobu poznawania świata przez człowieka. W takim przypadku możliwe stałoby się wypracowanie innego „mechanizmu myślenia” — komputer mógłby rozumieć świat na swój własny, sobie właściwy sposób. Ciekawe, że byłoby to jeszcze jedno wykroczenie poza Kantowskie kategorie. Przecież zadając różne rodzaje logik „rozumowania” komputerowi, moglibyśmy w różny sposób modelować obraz świata, jaki one narzucają.

Idąc dalej tym tropem można pokusić się jeszcze o taką refleksję: najprawdopodobniej różne są sposoby komunikacji (szeroko rozumianej) człowieka ze światem. Istotny jej fragment stanowi niewątpliwie warstwa posługująca się językiem matematyczno-logicznym. To ona pozwoliła na rozwój nowożytnej nauki, a następnie fenomenalny po-

stęp technologiczny. Bardzo prawdopodobne, że właśnie ze względu na specyfikę „języka”, ten fragment ludzkiego myślenia uda się sformalizować i wtedy ta warstwa rzeczywistości pozwoliłaby na ujęcie w ramy programu komputerowego, podobnie jak możliwe okazało się jej poznanie *naukowe*<sup>7</sup>.

Pozostaje jeszcze jedno pytanie i jedno zastrzeżenie:

**Pytanie:** Czy uda się kiedyś tę warstwę racjonalno-formalizowalną rozciągnąć, utożsamić z całością rzeczywistości?

**Zastrzeżenie:** To jednak ludzie tworzą programy. I nawet w przypadku programowania genetycznego i sieci neuronowych pomysłodawcą całej idei jest człowiek. Nie chodzi tutaj po prostu o dokładne przewidzenie sposobów działania i efektów procedury, ale o pomysł samego mechanizmu. Wracamy więc do Kantowskich ograniczeń. Potrafimy o tyle tylko stworzyć nowe sposoby widzenia rzeczywistości, o ile potrafimy je za pomocą naszych ograniczonych i specyficznych mózgów wymyślić.

Ale nawet tu dopatrzeć się można pewnej ukrytej analogii pomiędzy człowiekiem a komputerem. I tak np. spadkobiercy idei Kanta mogą się upierać, że człowiek o tyle tylko jest w stanie przekroczyć niektóre swoje ograniczenia poznawcze, o ile pozwalają na to pewne bardziej zasadnicze prawa rządzące jego myśleniem. Podobnie można spoglądać na komputery, upierając się przy tym, że nawet programy zdolne do tworzenia nowych programów, mogą to czynić jedynie dlatego, że zostały tak pomyślane przez człowieka

### 3. SYMULACJA — REALIZACJA

Jeden z głównych zarzutów w stosunku do „gry w udawanie” i możliwości „intelektualnych” maszyn, zilustrowany przez Searle’a słynnym eksperymentem „Chińskiego Pokoju”, polega na tym, jakoby komputer jedynie symulował myślenie a nie myślał naprawdę. Należy

---

<sup>7</sup>Oczywiście wcale nie twirdzę, jakoby wszystko co „naukowe” nawet w scisłym tego słowa sensie (ang. *science*) było tym samym formalizowalne i przekładalne na program komputerowy.

wobec tego zapytać, czym różni się udawanie, samo naśladowanie, od realizacji. Powróćmy do porównania z samolotem — chyba zasadniczą różnicę pomiędzy symulacją lotu a rzeczywistym unoszeniem się w powietrzu determinuje „kontakt z rzeczywistością”. Prawdziwy samolot znajduje się w przestrzeni powietrznej o konkretnych współrzędnych czasowo-przestrzennych i określonych parametrach fizycznych. W przypadku symulacji (albo samolotu — zabawki) wszystkie te warunki są jedynie symulowane — zastępowane przez coś innego, np. wirtualne środowisko w komputerze (albo rękę i wyobraźnię bawiącego się dziecka). Podobnie o automatycznym pilocie nie twierdzimy, że symuluje jedynie lot, ale że naprawdę steruje samolotem, gdy wyznacza on zachowanie rzeczywistej maszyny w rzeczywistym środowisku.

Z tego punktu widzenia pomysł Turinga wydaje się jeszcze bardziej atrakcyjny: „myślenie” komputera w „grze w udawanie” dokonuje się w interakcji z ludźmi, w ramach rzeczywistej rozmowy. Więc może nie tylko operacjonizm stanowił inspirację eksperymentu zaproponowanego przez Turinga<sup>8</sup>.

Mam nadzieję, że wskazane powyżej zależności pozwalają dostrzec, jak subtelna różnica dzieli zwolenników *mocnej* i *słabej* wersji SI. O ile ktoś nie zadeklaruje swej przynależności do jednej z tych dwóch grup, albo nie dookreśli swych poglądów na temat relacji pomiędzy umysłem i mózgiem a programem i komputerem, niezwykle trudno jest ocenić, czy jego dociekania na temat myślących maszyn stawiają je na równi z ludźmi.

Niezależnie od tego, czy będzie on uważał mózg za skomplikowaną maszynę, czy będzie podkreślał wyjątkowość „ducha” człowieka, projektując programy SI, będzie musiał w jakiś sposób wzorować się i stale się odnosić do tego jak myśli człowiek i jak działa jego mózg.

---

<sup>8</sup>Sądzę, że ciekawie byłoby też porównanie sposobu ujmowania kwestii SI przez Turinga i Searle’a przez prymat koncepcji języka Wittgensteina. Wręcz narzuca się skojarzenie poglądów Searle’a z ujęciem *Traktatu logiczno-filozoficznego* w kontraście do pokrewieństwa pomysłu „gry w udawanie” z „grami językowymi” z *Dociekań...* Jak każde narzucające się podobieństwo, może być ono jedynie powierzchowne; zbadaniu tej kwestii należałoby poświęcić osobno uwagę.



Niezależnie od swoich — redukcjonistycznych lub nie — poglądów, próbując skonstruować maszyny „rozumiejące” w jakiś sposób świat, będzie usiłował odnaleźć prawidłowości i sformalizować przynajmniej fragment rzeczywistości, a także sam proces wyciągania konsekwencji.

Niezależnie od tego, czy będzie uważał, że komputer myśli rzeczywiście, czy jedynie operuje symbolami, by móc wykorzystać te możliwości maszyn, musi zweryfikować ich działanie w prawdziwym środowisku.

Sądzę, że nawet radykalnie brzmiące pytanie Turinga — któż może wiedzieć, czy i oryginalność człowieka i całe jego zachowanie nie mogłyby być zdeterminowane przez pewien zespół praw wrodzonych i nabytych na drodze edukacji? – można zarówno dopowiedzieć redukcjonistyczną wizję człowieka, jak i odczytać jako postulat nie „składania broni” i poszukiwania sposobów, reguł, które pozwoliłyby zaprojektować maszyny „myślące”. Reinterpretację pozostałych przytoczonych wyżej przykładów pozostawiam Czytelnikowi.

Ostatecznie wydaje się, że wybór między *słabą* i *mocną* koncepcją SI nie może dokonać się w oparciu o same wyniki badań ani informatycznych ani neurofizjologicznych. Sądzę, że problem ten przynależy do tej klasy zagadnień, do rozwiązania których nauka musi zaprząć filozofię. Mam nadzieję, że udało mi się wskazać, że opowiedzenie się za *słabą* lub *mocną* wersją SI zależy nie tylko i nie tyle od osiągnięć technologicznych, co od zrozumienia, kim jest człowiek. A to wykracza poza możliwości (współczesnej?) nauki.

### LITERATURA CYTOWANA

#### **Hodges A.**

[1998] *Turing*, Warszawa: Wydawnictwo Amber.

#### **Hodges A.**

[2002] *Enigma. Życie i śmierć Alana Turinga*, Warszawa: Prószyński i S-ka.

#### **Marciszewski W.**

[1998] *Sztuczna Inteligencja*, Kraków: Znak.

**Searle J.R.**

[1995] *Umysł, mózg i nauka*, Warszawa: PWN.

**Turing A.M.**

[1950] *Computing Machinery and Intelligence*, *Mind* X, Vol. LIX, No 236, ss. 413-440.

**Turing A.M.**

[1995] *Maszyna licząca i inteligencja*, [w:] B. Chwedeńczuk (wybór i tłum.), *Filozofia umysłu*, Warszawa: Fundacja ALETHEIA — Wyd. SPACJA (tłum. Turing, 1950).