

Turing vs. Nietoperze:

Test Turinga a problem innych umysłów

Jakub Mańczak

2025-06-09

WSTĘP, PRZYBLIŻENIE SEDNA SPRAWY

Uczenie się i zapamiętywanie rzeczy oderwanych od kontekstu, takich z którymi się nie utożsamiamy, lub których nie mamy jak połączyć z czymś, co już wiemy jest trudne - przynajmniej dla mnie.

Problem innych umysłów zdecydowanie taką trudną rzeczą nie jest: za enigmatyczną nazwą kryje się bowiem nic innego, jak dzieciennie proste spostrzeżenie, że każdy człowiek widzi świat tylko z własnej perspektywy. Nie wiem jak Szanowny Czytelnik - ale ja zdecydowanie pamiętam, jak mając zaledwie garstkę lat w drodze do szkoły snulem teorie, jakoby cały świat był „tylko dla Mnie”, a inni ludzie tak naprawdę nie istnieli - zaraz obok myśli o możliwości dowiedzenia się przez nauczycielkę o braku pracy domowej poprzez czytanie w moich myślach. Takie poglądy, uważane raczej za eksperyment myślowy bardziej niż poważne teorie, nazywa się w filozofii solipsystycznymi. A takich teorii mówiących o tym, skąd wiemy o innych ludziach różne rzeczy jest więcej - tak samo jak i odpowiedzi nie zniżających się tak drastycznie do negowania istnienia innych - i zebrane są właśnie pod nazwą problemu innych umysłów (Morton, 2002).

Ale jak się okazuje, problem nie kończy się na ludziach. Teraz postrzeganie maszyn jako myślące, albo myślącopodobne, widać jak na dłoni - jednak Turing (1972) myślał o tym już całe 75 lat temu, w roku 1950, gdy wydał oryginalną, angielską wersję artykułu. Zaproponował on wtedy koncept gry w naśladowanie, nad którą nadbudował *Test Turinga*. Miał on badać hipotetyczne trudności w odróżnieniu maszyn od ludzi w scenariuszu, w którym jako pomoc w odgadywaniu dostępne są jedynie ich wypowiedzi pisemne. A więc osoba dokonująca rozpoznania dokonuje je w pełni na podstawie tego, co byt maszynowy lub organiczny uzewnętrznia - w sposób, nad którym ma pełną kontrolę i który, przynajmniej współcześnie, nie jest uzależniony od możliwości kosmetycznych. Maszyna wygrywa, jeżeli zmyli człowieka dokonującego rozpoznania.

A skoro człowiek myli maszynę z innym człowiekiem, co do którego z grzeczności zakładamy że myśli, to czy maszyna też myśli? Turing (1972) argumentuje, że tak - lub przynajmniej zbija argument przeciwko wykorzystaniu jego gry w dowodzie na iluzję maszynowego myślenia, równając sprzeciw do przyjmowania pozycji solipsystycznej - absurdalnej, więc błędnej.

TEORIE BEHAWIORALNE

Gdy Morton (2002) w rozdziale siódmym, w punkcie trzecim swojego „Przewodnika po teorii poznania” wprowadza teorie behawioralne, rozpoczyna od wytyczenia ich przewodniej myśli:

Aby uzyskać przekonania dotyczące ludzkich przeżyć, musimy nauczyć się słów określających tego rodzaju przeżycia.

Zakłada on, że nasze przekonania o cudzych stanach mentalnych opierają się na obserwacji zachowań oraz postępującym procesie nauki tego, jak przypisywać znaczenia reakcjom i wypowiedziom. Pokrywa się to z jedynym, jak się wydaje (a zatem firmowanym przez Turinga), sposobem na zbliżenie maszyn do ludzi: odcięciu się od ich procesów wewnętrznych czy budowy i skupieniu się na tym, co uzewnętrzniane. Zatem jest to teoria zgodna z testem Turinga.

Ponadto, opisuje proces zapoznawania się z tymi przekonaniem i dodaje nawet, że niekiedy dzieci mające problemy z nauczeniem się zasad tych powszechniejszych z przekonani mówią problemy z diagnozą własnych przeżyć - co pokrywa się dość przyjemnie z docześnie odczuwanymi limitacjami najnowszych zawodników gry w naśladowanie: LLMów.

OGRANICZENIA TEORII BEHAVIORALNEJ

Jednakże teorie behawioralne mają swoją reputację (nienajlepszą) i to nie bez powodu. Przykładowo, Morton zauważa kilka problemów z behawioryzmem.

Po pierwsze, zachowanie może nie być tożsame z faktycznym stanem mentalnym osoby je uzewnętrzniającej - może być powierzchowne lub mylące, celowo lub przez przypadek. W tym celu Morton wprowadza dodatkowe kryteria do przekonań o stanach mentalnych - oprócz prezentowania zachowań typowych dla stanu, osoba nie powinna prezentować zachowań typowych dla udawania danego stanu. Test Turinga nie rozróżnia poprawności czegokolwiek - nie bawi się w szukanie prawdy, a jego celem nie jest odsłanianie rzeczywistych stanów umysłu.

Po drugie, niektórych stanów osoby je żywiące wcale nie muszą sygnalizować. Jest to niewykrywalne w teście Turinga, jako że podmiot ma pełną kontrolę nad uzewnętrznieniem na papierze lub dalekopisie swoich stanów mentalnych, a zatajanie prawdziwych może być w takim teście przydatne. Morton przywołuje przykład z subiektywnym odczuciem bólu i nań tolerancji; jeżeli mamy wysoką tolerancję, możemy nic nie poczuć gdy inna osoba może zwijać się z bólu. Wiąże się to poniekąd z trzecim powodem:

Qualia; subiektywne doznania mogą skutkować różnymi reakcjami na ten sam bodziec zewnętrzny, co rujnuje teorię o zbliżonych do jednolitych reakcjach. Maszyna podczas Testu Turinga, celem naśladowania ludzi ogółem, a nie konkretnego człowieka, zapewne obierze najbardziej przeciętną reakcję na dany bodziec. Jako że qualia są niekomunikowalne, niekomunikowalna jest również różnica w ich postrzeganiu w ramach testu.

TE TYTUŁOWE NIETOPERZE

Przeciw Turingowi dodatkowo zdaje się stawać Nagel (1996), który w swoim artykule prezentuje tezę klóącą się ze sprowadzeniem wprost naśladowania człowieka do postępowania jak człowiek, a więc myślenia. Na przykładzie fizjologa wiedzącego wszystko o nietoperzach, postuluje on, że nawet taki człowiek (a co dopiero przeciętny zjadacz chleba!) wyobrażając sobie jak by to było być nietoperzem, nie czyni tego w oderwaniu od własnego doświadczenia ludzkiego - tj. może on sobie wyobrazić jedynie jak to jest być człowiekiem w ciele nietoperza, a nie jak to jest być nietoperzem we własnym ciele. Prowadząc analogię wprost z tego przykładu można by zbudować argument, że komputer nie myśli, lub nie myśli jak człowiek, bo nigdy nie osiągnie wiedzy jak to jest być człowiekiem w ciele człowieka - jedynie jak to jest go udawać.

PODSUMOWANIE

Test Turinga wydaje się być obarczony dużymi problemami, jeśli chodzi o pełnoprawne, filozoficzne wnioski. Podobnie jak podejścia behawiorystyczne, wnioski są zbyt wąskie aby zaakceptować pełny pakiet jego przesłanek - Test Turinga pozostaje jednak wartościowym tematem. Jako narzędzie myślowe pozwala uwidocznić, jak uzależnieni jesteśmy od teorii niepełnych, a także pozwala zauważyć rozmycie między myśleniem, a jedynie symulowaniem myślenia - u maszyn jak i u ludzi. Tym samym test nie daje nam jednoznacznych odpowiedzi, ale zmusza do dalszych pytań.

BIBLIOGRAFIA

- Morton, A. (2002). Wiedza o umysłach. W Adam Morton, *Przewodnik po teorii poznania*. SPACJA.
- Nagel, T. (1996). Jak to jest być nietoperzem?. *Przegląd Filozoficzny – Nowa seria*, 5, 129–141.
- Turing, A. (1972). Maszyny liczące a inteligencja. W E. Feigenbaum i J. Feldman (red.), *Maszyny matematyczne i myślenie* (s. 24–47). PWN.