

Piotr Nayar
MIMUW

Czy komputery mogą myśleć?

Bynajmniej nie jest celem tej pracy udzielenie wyczerpującej odpowiedzi na tak postawione pytanie. Możemy jednak przez chwilę zastanowić się nad tym, co ono tak naprawdę oznacza. Po pierwsze zauważmy, że w pytaniu występują zasadniczo cztery słowa. To bardzo dobrze. Bywają przecież znacznie dłuższe pytania. Pierwsze słowo nadaje zdaniu sens pytania. Ale dalej stoi już bardzo istotne słowo *komputer*. Właściwie wiadomo, jaka jest istota urządzenia, zwanego komputerem. Komputer to urządzenie, zbudowane z myślą o przechowywaniu i przetwarzaniu informacji. Jest to pewna maszyna, w działaniu której wykorzystywane są prawa fizyki, takie jak zjawiska elektryczne, magnetyczne, a w przyszłości zapewne nawet kwantowe. Nie da się ukryć, że komputery mają coraz więcej możliwości. W miarę upływu czasu, ich zastosowania stają się coraz bardziej powszechne. Tak, to oczywiste. Można jednak bezsprzecznie stwierdzić, że komputer, w swojej budowie, jest pewną sekwencją atomów. Pod tym względem nie różni się znacząco od innych rzeczy, z którymi stykamy się na co dzień.

Ale dość już o komputerze. Teraz zastanówmy się nad znaczeniem słowa *myśleć*. To jest już bardziej skomplikowane słowo. Niemniej jednak możemy coś o nim powiedzieć. Wydaje się, że myślenie jest charakterystyczne dla osoby ludzkiej. Przynajmniej powstanie samego słowa *myśleć* było zapewne inspirowane procesem występującym u ludzi. Trudno jednak precyzyjnie powiedzieć, czym tak naprawdę jest ten proces. Na razie pozostawimy to bez komentarza.

W końcu występowanie w pytaniu słowa *mogą* sugeruje, że autor pytania zastanawia się nad tym, czy myślenie, proces charakterystyczny dla istoty ludzkiej, mógłby przysługiwać również maszynom. Wydaje się również, że autor pytania jest przekonany, że komputery jeszcze nie są w stanie myśleć, ale być może w przyszłości będą.

Przypuśćmy, że przez myślenie będziemy rozumieli proces wyciągania wniosków z pewnego zestawu przesłanek. Człowiek jest zdolny do tego rodzaju myślenia. Świadczy o tym, między innymi, rozwój nauki. Pytamy teraz, czy komputer również jest w stanie myśleć w tym sensie. Wyobraźmy sobie, że napisaliśmy program komputerowy, w którym zaimplementowaliśmy reguły dowodzenia formuł logicznych. Oczywiście da się to zrobić (dla przykładu szkolny system tabel 0-1). Przypuśćmy, że uruchomimy nasz program i każemy mu generować losowo formuły logiczne, sprawdzać ich poprawność i zapamiętywać te zdania, które były prawdziwe. Po pewnym czasie w pamięci komputera znajdzie się pewna liczba dowiedzionych formuł. Można więc powiedzieć, że komputer nauczył się czegoś, zatem komputer myślał w rozważanym sensie.

Możemy jednak nieco inaczej podejść do problemu myślenia maszyny. Otóż rozważmy następujący, słynny eksperyment Alana Turinga. Człowiek prowadzi pięciominutową konwersację z rozmówcą, przy czym nie wie, czy jego rozmówca jest komputerem, czy człowiekiem. Strony nie mają oczywiście bezpośredniego kontaktu. Po rozmowie człowiek musi powiedzieć, czy jego zdaniem rozmawiał z innym człowiekiem, czy z maszyną. Jeżeli drugi rozmówca był komputerem, ale w 50% przypadków rozmawiający z nim ludzie twierdzą, że jest człowiekiem, to mówimy, że

komputer przeszedł test. Można wtedy uznać, że maszyna myśli w pewnym szczególnym rozumieniu tego słowa. Mianowicie wówczas maszyna ma inteligencję konwersacyjną podobną do ludzkiej, staje się dla człowieka równorzędnym rozmówcą. Chociaż żaden komputer nie przeszedł jeszcze w zadowalającym stopniu testu Turinga, to jednak łatwo można sobie wyobrazić, że budowa takiego komputera jest tylko kwestią czasu. Wiarygodność komputerowego rozmówcy można bowiem powiązać z liczbą zwrotów, które posiada on w swojej pamięci, jak również ze sposobem i szybkością ich przetwarzania.

Dochodzimy tu jednak do zasadniczego problemu. Czy tego rodzaju testy oddają istotę procesu ludzkiego myślenia? Niektórzy ludzie uważają, że ludzkie myślenie również rządzi się pewnymi, z góry ustalonymi regułami. Mianowicie człowiek mówi i robi to, co wynika z aktualnego ułożenia atomów w jego mózgu. Zgodnie z takim podejściem człowiek nie różni się w sposób istotny od maszyny. Mamy więc następującą sytuację: zarówno człowiek, jak i maszyna, składają się z pewnej liczby atomów. Możemy sobie ponadto wyobrazić maszynę, która w swoich zachowaniach zewnętrznych nie będzie się różniła od człowieka.

Pytamy więc, gdzie miałyby wystąpić ta zasadnicza różnica między maszyną i człowiekiem? Otóż jest taka cecha, której maszyna może teoretycznie nie posiadać. Jest nią mianowicie świadomość. Lecz cóż to jest świadomość? Świadomość to coś takiego, co zaczynamy odczuwać, gdy budzimy się każdego dnia rano. Oczywiście trudno zdefiniować świadomość. Gdybym jednak we śnie zaczął nagle chodzić po domu, to robiłbym to nieświadomie. Łatwo więc zrozumieć, czym różni się stan wiadomości od snu. Świadomość to pewnego rodzaju wewnętrzna pewność własnego istnienia. Możemy więc tytułowe pytanie rozumieć jako pytanie o możliwość istnienia świadomych maszyn.

Niestety, zasadniczy problem polega na tym, że autor niniejszej pracy może tylko stwierdzić, że on sam jest świadomy. Jeżeli rozmawiam z drugą osobą, to, pomimo najszczerzych chęci, nie mogę być pewien, że mój rozmówca jest, tak jak ja, świadomy. Świadomość (lub nieświadomość) jest indywidualną cechą każdego człowieka, przy czym cecha ta jest w oczywisty sposób nieweryfikowalna. Mamy przecież dostęp do informacji, które przychodzą do nas za pośrednictwem zmysłów, a takie informacje dają nam wiedzę jedynie o zewnętrznych cechach istot i rzeczy. Nawet jeśli ktoś powie mi, że jest świadomy, to czyż może to być dla mnie wystarczający dowód potwierdzający jego świadomość? Oczywiście nie. Mogłem przecież zostać oszukany, a mój rozmówca zareagował całkowicie automatycznie na moje pytanie, tak jak robią maszyny. Możliwa jest też sytuacja, w której tylko ja jestem tak naprawdę świadomy, a w moim świecie istnieją same nieświadome maszyny. Być może cały świat jest jedynie snem, wytworem mojej wyobraźni.

Na pewne pytania, które możemy zadać, zwyczajnie nie ma odpowiedzi. Do takich pytań należy pytanie, czy komputery mogą być świadome. Nigdy nie będziemy w stanie stwierdzić, czy maszyny są świadome, ponieważ jest to ich osobista sprawa. Pytania o świadomość i wolną wolę są pytaniami spoza obszaru nauki. Wkraczają one w obszar osobistych, indywidualnych przekonań każdego człowieka.

Z podobną sytuacją możemy się spotkać w logice matematycznej. Mam tu na myśli twierdzenie Gödla o zupełności dedukcyjnych (czyli zawierających arytmetykę Peano) systemów algorytmicznie generowalnych. W myśl tego twierdzenia w pewnej szerokiej klasie systemów da się zawsze znaleźć zdanie o tej własności, że po pierwsze, nie da się go udowodnić w obrębie system, a po drugie, nie da się udowodnić jego zaprzeczenia. Wydaje się, że mamy tu do czynienia z silnym podobieństwem. Zdanie *komputery mogą być świadome* można wypowiedzieć

w dostępnym nam języku, ale wydaje się ono być niezależne od dostępnych człowiekowi metod poznawczych.

Ostatecznie rysuje się więc przed nami następująca wizja. Możemy na wiele sposobów rozumieć pojęcie *myśląca maszyna*. Jednak jeżeli będziemy brać pod uwagę jedynie cechy komputera, które ujawniają się w jego relacji z otoczeniem, to łatwo wyobrazimy sobie komputer, który będzie w tego rodzaju relacjach bardzo podobny do człowieka. Jeśli jednak mamy na myśli głębsze cechy, takie jak świadomość i wolna wola, to na postawione pytanie nigdy nie będziemy w stanie odpowiedzieć. Autor niniejszej pracy jest przekonany, że ludzkie poznanie ma bardzo wyraźnie nakreślone granice. Możemy oczywiście rozwijać nowe technologie, odkrywać nowe zjawiska fizyczne. Nie wolno nam jednak przeceniać swojej działalności. Żaden fizyk, ani matematyk, ani żaden człowiek na świecie, nigdy nie dojdzie do poznania sedna sprawy, czyli do istoty związku pomiędzy indywidualnym istnieniem każdego z nas, a istnieniem całego świata. I właśnie dlatego możemy ciągle wierzyć w rzeczy wzniosłe i piękne, i w to, że czeka nas lepsza *przyszłość*, niż przyszłość, nie szukając daleko, mojego ośmioletniego komputera.