

Argumenty studentów wydziału WAIiNS (2017/18) na debatę pt. „Turing czy Searle?”

◆ Argumenty za Turingiem

[Maciej Radzimirski]

Jestem zdania, że istnieje możliwość stworzenia inteligentnych maszyn, co więcej, uważam, że jesteśmy na dobrej drodze do tego, by stworzyć maszynę uznawaną za inteligentną. Moje argumenty są następujące:

1. Wyniki przeprowadzonych eksperymentów pokazują, że da się stworzyć maszynę, której da się przypisać cechy inteligencji. Pozwolę sobie posłużyć się następującą definicją inteligencji (wg. Wikipedii): „Inteligencja - zdolność do postrzegania, analizy i adaptacji do zmian otoczenia. Zdolność rozumienia, uczenia się oraz wykorzystywania posiadanej wiedzy i umiejętności w różnych sytuacjach.” W sierpniu 2017r. przeprowadzono eksperyment, w który dwa programy miały za zadanie rozwijać sztukę negocjacyjną. W trakcie tego eksperymentu doszło do radykalnej zmiany języka na drodze ewolucji, jakim te maszyny się posługiwały – maszyny były zrozumiałe dla samych siebie, lecz niezrozumiałe dla obserwatorów. Oznacza to, że maszyny te wykształciły swój język na drodze adaptacji do sytuacji, co można zakwalifikować jako cechę inteligencji. Link do artykułu pod tym adresem: <http://www.independent.co.uk/life-style/gadgets-and-tech/news/facebook-artificial-intelligence-ai-chatbot-new-language-research-openai-google-a7869706.html>¹
2. Istnieje twierdzenie Godla, mówiące o tym, że w każdym dostatecznie potężnym systemie logicznym da się sformułować takie zagadnienia, których nie da się udowodnić, lub obalić, o ile system nie jest niekonsekwentny – krytycy możliwości stworzenia sztucznej inteligencji często powołują się na to stwierdzenie, gdy mówią o niemożności stworzenia takowej; jednak najczęściej nie wspominają o takich ograniczeniach w ludzkiego umysłu; ponadto wg. Turinga triumf ludzi nad daną maszyną wcale nie oznacza, że nie istnieją inne zdolniejsze maszyny (polemika Turinga z argumentem sprzeciwu matematycznego).
3. Turing w swoim tekście „Maszyny liczące a inteligencja”² polemizuje z argumentem ciągłości systemu nerwowego – można zauważyć, że niezależnie od tego, czy mówimy o maszynie, która działa w dziedzinie ciągłej, czy dyskretnej, obie te maszyny w pewnych sytuacjach potrafiłyby dawać prawidłowe odpowiedzi (w tekście przykładem jest obliczanie wartości liczby π) – w takiej sytuacji pytającemu obie te maszyny o wynik obliczeń prawdopodobnie trudno byłoby odróżnić maszynę cyfrową (komputer) od maszyny ciągłej (w tym wypadku człowieka)

¹ Data dostępu do artykułu: 06.01.2018, 23.03

² <http://stac.calculamus.org/pdf/at-mlai.pdf> - chodzi o polemikę z argumentem ciągłości systemu nerwowego

[Jan Ferec]

Jednym z głównych argumentów Johna Searle'a, którymi starał się podważyć teorię Turinga o możliwości istnienia tzw. mocnej sztucznej jest „**chiński pokój**”. Polega on na założeniu, że istnieje komputer, który potrafi przetwarzać język chiński i odpowiadać na przekazywane mu komunikaty w sposób tak przekonujący, że z łatwością przechodzi test Turinga (przekonuje Chińczyków, że rozmawia z innym Chińczykiem), czyli spełnia warunki Turinga o byciu mocną sztuczną inteligencją. Searle teraz zakłada, że teraz człowiek nie znający języka chińskiego siedzi w pokoju, posiadając książkę znaków i reguł tego języka. Odczytuje dostarczone mu komunikaty, przetwarza język na zrozumiałe dla siebie, a następnie zapisuje odpowiedź po chińsku za pomocą dostępnych materiałów. Searle zauważa, że osoba znajdująca się w pokoju nie rozumie ani słowa po chińsku, mimo iż wykonuje zadanie identyczne jak wyżej wspomniana maszyna. Wyciąga z tego wniosek, że komputery nie rozumieją chińskiego, a są jedynie prymitywnymi manipulatorami symboli – nie można ich nazwać silną sztuczną inteligencją. Zachowanie pokoju w oczach Searle'a postępuje jedynie zgodnie z **syntaktyką**, ale ponieważ nie zna znaczenia przekazywanych mu znaków, nie jest w stanie ich zinterpretować co skutkuje brakiem wartości **semantycznej**.

Szkic moich argumentów w dużym stopniu jest odpowiedzią na ten argument Searle'a.

1. Jesteśmy w stanie poznać, że człowiek rozumie język chiński lub cokolwiek innego po jego zachowaniu. Podążając dalej tą logiką, jeśli komputer jest sobie w stanie poradzić z testem behawioralnym tak dobrze jak człowiek, to znaczy, że jeśli przypisujemy umiejętności poznawcze ludziom to zasadniczo musimy także przypisać je komputerom.

2. To prawda, że osoba znajdująca się w chińskim pokoju nie rozumie chińskiego. Jednakże pomijany jest fakt, iż ta osoba jest jedynie częścią całego systemu, do którego należy także książka reguł i znaków języka. Przypuszczalnie system składający się z człowieka oraz książki rozumie ten język.

3. Jeśli zamiast pokoju umieścimy komputer z programem w robocie tak, aby był w stanie poruszać się oraz był interaktywny z otoczeniem w sposób identyczny do człowieka to wtedy, można by uznać, że rozumie co robi.

[Nguyen Chi Cong]

1) Searle twierdzi, że maszyny mogą być tylko syntaktyczne i nie ma semantyki. Natomiast, jak kolega stwierdził wcześniej, semantykę można zaprogramować. Jak w przypadku ludzi, kiedy pierwszy raz się słyszy słowo, to człowiek widzi tylko ciąg znaków, dopiero kiedy się go nauczy, to zrozumie to słowo i będzie je kojarzył z innymi znanymi słowami. Uważam, że maszyna też jest w stanie się tak uczyć nowych słów.

2) Można przypuszczać, że kiedy powstał pierwszy człowiek, to nie byli oni całkowicie świadomi/samoświadomi, tylko tę świadomość zdobywali na drodze ewolucji, zatem gdyby maszyny były zdolne do ewolucji (są algorytmy genetyczne, które całkiem dobrze tę ewolucję symulują), to maszyny mogą być świadome. Nawet, gdyby się nie dało, to uważam, że świadomość nie jest warunkiem koniecznym myślenia.

3) Można się zastanawiać, czy człowiek nie jest pewnego rodzaju maszyną. Tuż po urodzeniu niemowlę działa według instynktu, czyli jakiegoś wbudowanego algorytmu, a potem przez

resztę życia się uczy pewnych rzeczy za pomocą jakichś algorytmów samouczących się. Gdyby udało się bardzo dobrze opisać te algorytmy, to maszyny byłyby bardzo blisko do myślenia jak człowiek.

◆ Argumenty za **Searle'm**

[**Klaudia Kaźmierska**]

1. Obiekt inteligentny i myślący musi być świadomy swoich działań (rozumieć sens danych wejściowych i na podstawie sensu tych działań odpowiednio generować odpowiedź). Maszyny to jedynie wykonawcy odpowiednio zdefiniowanych algorytmów działających wg określonych reguł (przykład zamkniętego pokoju, w którym znajdują się człowiek nieznający języka chińskiego), maszyna nie rozumie tego, co aktualnie przetwarza. Samo zaimplementowanie algorytmów umożliwiających doskonałe odpowiadanie na pytania nie jest wystarczające aby uznać maszynę realizującą taki algorytm za myślącą.
2. Nie da się otrzymać znaczenia (semantyki) tylko na podstawie syntaksy - poznanie reguł danego języka i dobieranie słów na tej podstawie nie powoduje, że maszyna jest w stanie zrozumieć to, co stworzyła.
3. Inteligencja nie może być postrzegana tylko poprzez jej zewnętrzne znamiona – samo odpowiednie przetwarzanie danych i zwracanie odpowiedniego wyniku nie jest myśleniem.
4. Aby nazwać maszynę myślącą, należałoby zbudować taką, która miałaby dokładnie taką samą strukturę jak istota ludzka - takie maszyny nie zostały jeszcze stworzone.
5. Symulowanie nie jest faktycznym wykonywaniem danej czynności (maszyny zajmują się jedynie symulowaniem czynności).
6. Test Turinga sprawdza tylko poziom skomplikowania algorytmów oraz ich wydajność (czy są w stanie symulować odpowiedzi człowieka na zadane pytania).

[**Grzegorz Gugala**]

Na wstępie doprecyzowanie tematu. Myślenie, w kontekście maszyn, rozumiem tu jako ich 'uosobienie', nadanie im pewnej świadomości, czy też elastyczności w postępowaniu, zależnie od okoliczności. Innymi słowy pytanie wyjściowe można sprowadzić to problemu utworzenia maszyny tożsamej intelektualnie z człowiekiem.

Oczywiście uważam, że powstanie takiej maszyny nie jest możliwe.

Główną linią obrony takiej tezy jest już sama definicja maszyny i algorytmu. Algorytm jako taki jest jednoznacznym sposobem postępowania, co oznacza, że jego wynik jest ściśle uzależniony od danych wejściowych (stosowane czasem ewentualne wartości losowe, są w rzeczywistości pseudolosowe). Ponadto algorytm zwraca rozwiązanie, po skończonej ilości ruchów. Jasne jest więc, że maszyny działają w sposób w pełni zalgorytmizowany. Ludzie natomiast oczywiście mogą realizować algorytmy, jednak w ich przypadku jest to

indywidualna kwestia – sposób ich działania, nie jest i nie może zostać odgórnie określony. Oczywiście możliwe jest utworzenie dla maszyn skomplikowanych algorytmów, mogących zachowywać się różnie, dzięki możliwości modyfikowania się w miarę potrzeb. Takie rozwiązania, pozornie umożliwiające algorytmowi możliwość nieograniczonej adaptacji, stoi jednak w sprzeczności z jego definicją. Modyfikacja algorytmu w czasie jego działania wyklucza bowiem jego jednoznaczność, gdyż sposób jego działania staje się zmienny. Ponadto ilość ruchów ‘samozmiennego algorytmu’ może wybiegać do nieskończoności, gdyż wprowadzane zmiany, umożliwiające rozwiązanie zadanego problemu, mogą uniemożliwić rozwiązanie innego, jeżeli problemy te wzajemnie się wykluczają.

Argumentem pomocniczym, wspierającym taki tok rozumowania, jest brak gwarancji bezbłądności. W kontakcie z nieznanym problemem, algorytm zdolny do pewnej formy adaptacji, poprzez zmianę sposobu w który działa, może zacząć zmieniać się w sposób, który oddali go od rozwiązania owego problemu. Możliwa jest nawet sytuacja wprowadzenia zmiany ograniczającej, a nie rozbudowującej algorytm. Oczywiście możliwe jest nadanie algorytmowi ograniczenia w jego adaptacji, tak aby nie mógł on ograniczyć w żaden sposób siebie samego, jednak takie ograniczenie stoi w sprzeczności z nieograniczoną możliwością algorytmu do adaptacji.

Na koniec kilka słów o ludziach, a raczej jedynie o ludzkim umyśle, gdyż również maszyny rozpatrywane były tu jedynie pod względem algorytmu. Ludzki umysł jest zdolny do nieograniczonej adaptacji, gdyż jak wspomniałem nie jest ograniczony algorytmem. Oznacza to, że możliwe jest zachodzenie w nim zmian (zmian rozumianych tutaj, nie jako przetwarzania informacji, ale jako zmiany sposobu działania) nieokreślonych i nieograniczonych. Zmiany te są zasadniczą różnicą algorytmu i rozumu ludzkiego. Rozum ludzki nie gwarantuje rozwiązania problemu, po skończonej ilości ruchów, jednak zapewnia możliwość adaptacji i wielopłaszczyznowej analizy oraz działania.

[Emilia Gosk]

1. Test Turinga jest niewystarczający do określenia czy komputery są myślące - test może przejść komputer, który wcale nie rozumie, tylko ma ogromny zbiór zawierający ogromną liczbę konwersacji. Wiarygodność komputerowego rozmówcy jest tu powiązania z liczbą zwrotów, które posiada w swojej pamięci, jak również sposobem i szybkością ich przetwarzania.

Ponadto zwierzęta nie przejdą tego testu, a nikt nie podważa faktu, że myślą. Podobnie małe dzieci.

2. Eksperyment Chiński Pokój ukazuje, że do pełnego myślenia komputerom brakuje rozumienia przetwarzanych informacji. Komputery mogą jedynie symulować myślenie, tak jak np. pogodę. Program jest jedynie modelem rzeczywistości (czyli np. procesów wskazujących na inteligencję konwersacyjną), nie rozumie elementów modelu - nie mają one dla niego znaczenia.