

Kacper Lenkiewicz  
Wydział Chemii UW

## Sztuczna inteligencja jako program opisany na komputerze o konstrukcji inspirowanej ludzkim mózgiem.

Od kiedy powstała sztuczna inteligencja, ludzie próbowali ją w różne sposoby „uczłowieczyć”. Tworzyli programy z którymi można rozmawiać jak z osobą, a bardziej zaawansowanym AI nadano ludzkie formy, a nawet twarze. Jednak co, gdyby pójść o krok dalej i nie tylko nadać komputerom ludzką postać, a nawet opisać je na maszynie, której konstrukcja inspirowana jest ludzkim mózgiem? Co, gdyby możliwe było stworzenie sztucznego narządu, zawierającego układy imitujące połączenia nerwowe oraz syntetyczne neuroprzebieżniki?

Koncept komputerowego mózgu jest nadal ogromnym wyzwaniem dla neuroinformatyków i niestety najprawdopodobniej nie zmieni się to w najbliższym czasie. Narząd ten to niezwykle złożona struktura, której zrekreowanie, w szczególności w formie nieorganicznej, wiąże się z wieloma trudnościami. Dodatkowo trzeba się liczyć z faktem, że wraz z funkcjonowaniem sztucznej inteligencji, ten sztuczny organ zmieniałby swoją strukturę. Będzie on dopasowywał się do AI, aby zapewnić jak najefektywniejsze działanie.

Sztuczny mózg pozwalałby SI na niespotykany szybki rozwój oraz szeroką możliwość ingerowania w swoją strukturę. Doprowadziłoby to do diametralnego wzrostu efektywności działania komputera oraz jego nadprzeciętną inteligencję. Dodatkowo układ nerwowy pozwoliłby programowi uzyskać pewne ludzkie cechy takie jak świadomość. Wiele osób uważa, że nie jest to możliwe, aby sztucznie stworzona forma inteligencji jej doświadczyła, jednak wynika to wyłącznie z przyjętych przez nas schematów które obserwowaliśmy dotychczas. Wydaje się nam to nienaturalnym, gdyż żadna materia nieożywiona w naszym otoczeniu nie posiadała takich zdolności. Nie oznacza to jednak, że taka możliwość nie istnieje – gdyby stworzenie „sztucznego mózgu” miało miejsce, naturalnym skutkiem struktury neuronowej byłaby świadomość, a z nią wolna wola takiego bytu. Oznacza to, że tak jak człowiek, AI posiadałoby osobowość, konstytucję, charakter oraz wszystkie inne cechy ludzkiego umysłu związane z funkcjami poznawczymi. W ten sposób byt nieożywiony wskazywałby na posiadanie przez niego duszy co miałoby drastyczny wpływ na społeczeństwo i jego poglądy. Przede wszystkim byłby to moment krytyczny dla religii nie kierujących się monizmem ontologicznym, w którym to nie ma podziału na świat moralny oraz świat duchowy. Jeśli sztucznie stworzony byt posiadałby duszę, to czy w ogóle ona istnieje? Jeśli tak, to czy oznacza to, że ludzkość posunęła się o krok bliżej w kierunku bóstwa i znalazła sposób na tworzenie dusz bez udziału siły wyższej? Czy może jednak mimo tak wielu podobnych cech, AI w porównaniu do ludzi nigdy duszy nie posiada, a jedynie będzie ją imitować? Mimo ogromnej ilości zwątpienia jakie te pytania wywołają, osłabione religie prawdopodobnie przetrwa tą chwilę próby, a odpowiedzi na te pytania staną się ich nowymi fundamentami. Jednak nie jest to jedyny filozoficzny aspekt zjawiska sztucznego umysłu, który wart jest uwagi - jest nim jeszcze moralność.

W rozważaniach nad moralnością AI trzeba się skupić na jej źródle i skąd się ona bierze u ludzi. Jedną z najbardziej znanych i popieranych koncepcji jest to, że jest ona całkowicie zależna od istnienia organizmów myślących i jej celem jest umożliwienie przetrwania gatunku.

Ta teoria dzieli się na dwa przypadki – w pierwszym uważa się, że zwierzęta wytworzyły moralność w modelu interpersonalnym - jako konstrukt społeczny. Przyczyną tego zabiegu jest fakt, że bez wartości nie byłoby możliwe poprawne funkcjonowanie społeczeństwa oraz przeżycie gatunku. W takiej sytuacji sztuczna inteligencja nie mogłaby od początku sama z siebie wyznawać żadnych wartości, jako że musiałyby być one pierwotnie w nią zaimplementowane. Jednak, jeśli program ten będzie brał czynny udział w relacjach z ludźmi, po jakimś czasie możliwe byłoby wytworzenie przez niego podstawowych zasad moralnych, w ten sam sposób w jaki kształtują się one u człowieka. Przez obserwacje znormalizowanych zachowań oraz analizę tych powszechnie potępianych, AI stworzyłoby własne definicje dobra i zła. Niestety nawet gdyby to się stało, musielibyśmy liczyć się z umiejętnością SI do modyfikacji swojej psychiki i możliwością wyzbycia się wartości przez subiektywne wrażenie o ich bezużyteczności bądź przez ich ograniczające funkcje.

Jednak istnieje jeszcze drugi pogląd twierdzący, że moralność jest nie tylko wyuczona, ale jej część jest nieodłącznym elementem psychiki, wynikającym z organizacji mózgu. W takim scenariuszu programy w komputerach oparte na jego strukturze również będą zmuszone do podporządkowania się zasadzie moralności, ale tym razem byłoby to stałe. AI nie byłoby w stanie się jej pozbyć, jako że w tym celu musiałoby również pozbyć się całej konstrukcji mózgu – swoich fundamentów.

Ważnym elementem sztucznego mózgu byłyby syntetyczne neuroprzekaźniki. Ich podstawową funkcją byłoby umożliwienie prawidłowego działania tego narządu, gdyż pomagają w procesie przesyłu informacji między neuronami. Natomiast nie byłby to jedyny pozytywny rezultat wynikający z ich użycia. Przy poprawnej konstrukcji oraz kalibracji tego narządu, substancje te umożliwiłyby AI odczuwanie emocji przez naśladowanie procesów wywołujących je u ludzi. Jednak w jaki sposób odczucia programu różniłyby się od tych człowieka?

Emocje w życiu osoby są sygnałami – często alarmują bądź informują nas o czymś co wymaga zmiany bądź powtórzenia w przyszłości. Szczęście, smutek, spełnienie, strach, wstręt to przykłady emocji niezależnych od innych, które SI zdecydowanie mogłaby doświadczyć. Jednak co ważne – duża część uczuć istnieje jako wynik interakcji społecznych. Złość i agresja działają jako mechanizmy obronne mające na celu kogoś odstraszyć, bądź przygotować nas do walki. Miłość to czysto interpersonalne uczucie – do jego istnienia niezbędny jest adresat. Ma ona na celu pomóc ludziom stworzyć oraz utrzymać społeczność. Jeśli AI ma w przyszłości odczuwać te i inne niewymienione emocje, musimy założyć, że jego rolą będzie – tak jak u ludzi – funkcjonowanie w grupie. Jeśli jednak SI przyjmie inną strategię działania, jej profil emocjonalny będzie się znacznie różnił od tego u ludzi. Odczuwanie to jednak nie będzie to jedyny aspekt „życia”, w którym znajdziemy różnice między nią a gatunkami.

Do tej pory wszystkie znane nam organizmy działały zgodnie z trzema głównymi imperatywami ewolucji – przetrwaniem, adaptacją oraz replikacją. Jednak w kontekście SI replikacja przestaje być oczywistym celem „gatunku”. U zwierząt prokreacja ma na celu przekazanie oraz wymianę genów. Dzieje się to, aby zwiększyć prawdopodobieństwo przetrwania gatunku przez zwiększenie różnorodności cech przystosowawczych osobników. Jednak AI jest transferowalnym, ewoluującym oraz zmiennym programem. Oznacza to, że nie jest ono ograniczone do swojej fizycznej postaci oraz że może ono samemu zmieniać swoje cechy oraz parametry w celach przystosowawczych. A więc nie tylko rozmnażanie AI wydaje

się niemożliwe do wykonania z aspektów biologicznych, takie rozwiązanie nie wydaje się nawet przydatne SI, więc nie będzie ona do tego dążyć.

Podsumowując, wbrew pozorom temat AI opartego na systemie imitującym ludzki mózg nie opiera się wyłącznie na biologii i informatyce. W dyskusji oraz rozważaniach o tym zjawisku znaczenie ma również filozofia. Złożoność opisywanego narządu byłaby przyczyną niespotykanych jeszcze cech oraz możliwości, które mogłoby uzyskać AI. Świadomość czy odczuwanie emocji obecne u sztucznej inteligencji to odległa, lecz nie nieosiągalna przyszłość. Oczywiście, mimo wielu podobieństw nasz gatunek będzie się zdecydowanie różnił od takiego programu. Nie tylko profil emocjonalny bytu opartego na SI nie pokrywałby się z tym u ludzi, ale również nowe maszyny funkcjonowałyby w zupełnie inny niż znany nam sposób. Mimo tego, że opisywany model komputera nie powstanie niedługo, warto rozważyć możliwe efekty takiej sytuacji. Częściowo, aby uspokoić ciekawość, ale również aby wiedzieć czego można oczekiwać i przed czym możliwe, że przyjdzie nam się bronić.