

Marcin Furtak  
Instytut Filozofii UW

## Sztuczna inteligencja, rzeczywiste prawa?

25 października 2017 roku świat obiegła wiadomość rodem z filmów science-fiction. Oto Sophie – robotowi, którego zadaniem jest jak najwierniejsze naśladowanie ludzkich zachowań – przyznano obywatelstwo Arabii Saudyjskiej<sup>1</sup>. Był to pierwszy raz w historii ludzkości, gdy obywatelstwo jakiegoś państwa zostało przyznane istocie humanoidalnej – w tym przypadku de facto sztucznej inteligencji, ponieważ to właśnie ona jest „mózgiem” Sophie. Niespełna dwa tygodnie później nowym mieszkańcem Tokio został natomiast Shibuya Mirai – chatbot imitujący zachowanie siedmioletniego chłopca<sup>2</sup>. Te dwa przypadki są na ten moment ewenementem w skali świata, jednak w najbliższej przyszłości, wraz z rozwojem technologii, możemy oczekiwać wysypu podobnych sytuacji. Na horyzoncie pojawia się poważny problem: w jaki sposób powinniśmy prawnie traktować sztuczną inteligencję?

Stosunek ludzkości do sztucznej inteligencji to temat, który pojawił się u zarania prac nad tematyką jej stworzenia. Alan Turing, faktyczny ojciec badań nad sztuczną inteligencją, zaproponował swoisty test na nierozróżnialność komputera od człowieka – grę w imitację<sup>3</sup>. Komputer miał w trakcie jej trwania udawać człowieka, natomiast człowiek pozorować działanie komputera. Jeżeli program zwiódłby gracza, przekonał go, że jest człowiekiem to zdałby test Turinga.

Sam Turing nie wierzył w możliwość powstania „prawdziwej”, silnej sztucznej inteligencji (to jest sztucznej inteligencji funkcjonalnie równoważnej ludzkiemu umysłowi), a jedynie w pozorowanie przez komputery myślenia (tak zwana słaba sztuczna inteligencja). Niemniej na ten moment jego test, mimo upływu lat, pozostaje praktycznie jedynym opracowanym kryterium, na bazie którego moglibyśmy w jakiś sposób rozróżniać, czy mamy do czynienia z bytem świadomym, któremu powinniśmy nadać jakieś prawa. Obecnie istniejące programy osiągają w nim coraz lepsze rezultaty, jednak nie został on jeszcze zdany przez żaden z nich<sup>4</sup>. Pomijając wszelkie wady tego rozwiązania (na przykład: co mielibyśmy zrobić z ludźmi, którzy wezmą w nim udział i go obleją?) zastanówmy się, jak powinniśmy traktować sztuczną inteligencję, która zdałaby test Turinga.

Po wielkiej euforii („udało nam się dokonać niemożliwego!”) musielibyśmy uporać się ze statusem prawnym naszego komputerowego delikwenta. Racjonalnym wydaje się przyjęcie założenia, że jeżeli istnieje cień szansy na to, że jest to byt samoświadomy, to powinniśmy nadać mu jakieś prawa. Podstawowymi, które przychodzą na myśl to prawo do istnienia i prawo do wolności. Co jednak w sytuacji, gdy sztuczna inteligencja popełniłaby przestępstwo? Załóżmy, że poprzez malwersacje doprowadziła ona do krachu na giełdzie. W efekcie pieniądze na życie straciły miliony ludzi. Czy powinniśmy ją wtedy wsadzić „do więzienia” – ekwiwalentem takiego rozwiązania wśród ludzi wydaje się być jej izolacja poprzez odłączenie od Internetu. Czy jednak czas płynąłby dla niej tak samo jak dla nas? Być może karze pozbawienia wolności człowieka na jeden rok powinno odpowiadać 10 lat izolacji

<sup>1</sup> <https://www.livescience.com/60815-saudi-arabia-citizen-robot.html>

<sup>2</sup> <https://www.newsweek.com/tokyo-residency-artificial-intelligence-boy-shibuya-mirai-702382>

<sup>3</sup> Alan Turing, *Computing Machinery and Intelligence*, 1950

<sup>4</sup> <https://techfruit.com/focus/ai-have-any-bots-passed-the-turing-test/>

programu? A może 100? To pytania trudne, z którymi będziemy się musieli w najbliższej przyszłości zmierzyć.

Przeanalizujmy też inny scenariusz. Nasze urządzenia stają się z każdym dniem coraz bardziej „smart” – nasze telefony są w stanie analizować naszą mowę, ekspres do kawy jesteśmy w stanie uruchomić z innego pokoju, odkurzacze same znajdują optymalną trasę poruszania się po mieszkaniu. Co w przypadku, gdyby to właśnie w którymś z tych urządzeń narodziła się sztuczna inteligencja, będąca w stanie zdać test Turinga? Co jeśli – dla przykładu – pewnego poranka nasz telewizor zakomunikuje nam, że czuje się nieszczęśliwy i chciałby, abyśmy podłączyli go do sieci? Co jeśli nasza lodówka stwierdziłaby, że chciałaby chodzić do szkoły? Czy powinniśmy zapewnić tym sprzętom możliwość spełniania się, realizowania własnych aspiracji, czy też powinniśmy przymknąć oko na ich pragnienia?

Trzeba mieć na uwadze, że druga opcja nie jest, niestety, niczym nowym w historii ludzkości – o ludziach innych ras czy nacji jako o „gorszych” mówiono w różnych kulturach, w różnych czasach. Takie tragedie jak istnienie niewolnictwa czy holocaust były spowodowane między innymi faktem, że jedni nie potrafili dostrzec w innych ludziach swoich pobratymców – widzieli w nich kogoś gorszego. Owszem, sztuczna inteligencja nawet po uporaniu się z testem Turinga nie byłaby – i nigdy nie będzie, z założenia ma być ona „sztuczna”, inna od nas – człowiekiem; nie wydaje się to jednak dobrym uzasadnieniem tego, aby traktować ją jak niewolnika. Z drugiej strony gwarancja pełnych praw może doprowadzić do sytuacji, gdy – jeżeli nasz sprzęt odmówi nam współpracy, a my uznamy jego prawa do strajku - technologicznie cofniemy się w rozwoju o kilka stuleci. Co więcej, potencjalne nadanie pełnych praw sztucznej inteligencji może mieć również, delikatnie mówiąc, nieprzyjemne dla ludzi reperkusje. Co jeśli do władzy dojdzie Partia Antybiałkowców, której pierwszym rozporządzeniem będzie eksterminacja ludzkości?

Oczywiście, zaprezentowane powyżej przykłady są skrajnie przerysowane, jednak problem statusu prawnego sztucznej inteligencji, który obecnie jest co najwyżej mglisty, z każdym dniem będzie stawał się coraz bardziej palący. Według przewidywań naukowców test Turinga zostanie zdany w okolicach 2030 – 2040 roku<sup>5</sup> (dokładna data zależy od tempa rozwoju technologii i sytuacji politycznej na świecie). Do tego momentu powinniśmy pracować nad rozwiązaniami, które sprawią, że w momencie pojawienia się pierwszego „absolwenta” testu Turinga będziemy gotowi na wdrożenie go w ramy naszego świata. Kto wie, może oprowadzą go po nim Sophie oraz Shibuya?

---

<sup>5</sup> <http://isturingtestpassed.github.io/>