

## Argumenty studentów wydziału FIZYKI (2016/17) na debatę pt. „**Czy maszyny autonomiczne muszą być świadome?**”

### ◆ Argumenty „**ZA**” – muszą być świadome

[Łukasz Szelański]

Definicja pojęć (źródło słownik języka polskiego PWN):

Autonomia – samodzielność i niezależność w decydowaniu o sobie.

Świadomość – zdawanie sobie sprawy z czegoś, charakterystyczna dla człowieka zdolność poznawania i oceniania.

1. Maszyna autonomiczna, czyli taka, która będzie w pełni niezależnie decydować o sobie **musi mieć możliwość poznania otaczającego ją świata**. Bez poznania mechanizmów świata, bez wiedzy o funkcjonowaniu praw fizyki, o funkcjonowaniu społeczeństwa, maszyna nie może decydować o sobie. Bez świadomości, nie może być autonomiczności. (Wniosek na podstawie definicji pojęć)
2. Komputer IBM Deep Blue w 1997 roku pokonał mistrza świata w szachach. Jeżeli Deep Blue nie byłby świadomy, nie potrafiłby poznawać i oceniać, czy mógłby wygrać? Nie.
3. Człowiek aby podjąć jakąś niezależną i samodzielną decyzję, musi najpierw zbadać dany problem. Poznaje jego pochodzenie, możliwości jego rozwiązania i dopiero podejmuje decyzję. Człowiek dochodzi do autonomiczności, poprzez świadomość. Człowiek jako twórca maszyny autonomicznej musi nauczyć ją autonomiczności, odpowiednio napisać kod. **Ponieważ człowiek dochodzi do autonomiczności poprzez świadomość, nie będzie umiał wykreować maszyny, która będzie osiągała autonomiczność inaczej niż poprzez świadomość.**

### ◆ Argumenty „**PRZECIW**” – nie muszą być świadome

[Magdalena Kaja]

Jeśli przyjąć taką definicję autonomii:

Autonomia - samodzielność i niezależność w decydowaniu o sobie, TO:

1. Maszyna nie musi być świadoma do tego, aby podejmować samodzielne decyzje. Jeśli za samodzielność przyjmiemy to że ma ona swoją bazę danych i nikt z zewnątrz już w nią nie ingeruje, to tylko ona sama na podstawie swoich algorytmów decyduje co zrobić.
2. **Świadomość jest czymś „wyższym”**, zawiera ona w sobie więcej elementów. Maszyna, aby być autonomiczną, nie potrzebuje mieć świadomości - podejmowane przez nią decyzje nie muszą być świadome, aby być samodzielne i niezależne.
3. **Jeśli da się stworzyć maszynę autonomiczną, to nie będzie ona posiadała świadomości, bo świadomość jest zarezerwowana dla istot żywych.** Może ona posiadać pewne jej aspekty, ale nie będzie posiadała ona uczuć ani emocji, które są jednym z elementów tworzących świadomość. Czy bez tego maszyna może odczuwać aspekt

przyjemności/przykrości (jedne z własności modalności świadomości wg. Searle'a)?  
Zdecydowanie nie.

### [Dariusz Szalkowski]

1. W pierwszej kolejności zastanówmy się, co można oznaczać autonomiczność maszyny. Termin „autonomiczny” związany jest z niezależnością. Owa niezależność może pojawiać się w na różnych etapach jej działania:

- a) stawiania sobie celów
- b) rozwiązywania określonego problemu

Bez trudu umiemy stworzyć program będący częściowo autonomiczny, spełniający tylko drugie z wymienionych kryteriów. Przykładem może być sieć neuronowa, która rozwiązuje zadany przez twórcę problem np. rozpoznawanie obiektów/mowy. Maszyna taka w procesie interakcji z otoczeniem uczy się rozwiązywać zadany problem bez konieczności zmiany swojego kodu źródłowego. Program rozwiązuje tutaj problem w sposób nieprzewidywany przez twórcę, który określił jedynie „szkielet” działania algorytmu lub funkcję celu. **Maszyna w tym przypadku jest nadal jednak tylko narzędziem w ręku człowieka, które w wyniku wprowadzenia pewnych reguł działania wpadło na innowacyjne rozwiązanie.**

2. Problem stawiania sobie własnych celów (pewna inwencja twórcza) wymagałby modyfikacji funkcji celu, czyli samoprogramowania się maszyny. Pewną heurystyką oddającą pomysłowość maszyn mogłoby być wprowadzenie losowości do ich działania, co imitowałoby niezależność i nieprzewidywalność (czyli autonomię). **Proces ten w żadnym stopniu nie gwarantuje pojawienia się „świadomości maszyny”, wolnej woli, odczuwania emocji.** Może przypominać jedynie mechanizm występujący w mózgu człowieka - nagłe pojawienie się w głowie pomysłu z tzw. podświadomości. Maszynę taką należałoby wyposażyć w kryteria odrzucania i akceptacji pomysłów oraz ich sortowania, co byłoby namiastką własnego systemu i hierarchii wartości. Przypominałoby to algorytm genetyczny oparty na ewolucji (czyli pewnych mutacjach).

3. Ciekawym, choć kontrowersyjnym krokiem jest integracja elementów mechanicznych z biologicznymi. Znane są już przypadki budowy robota, którego części mechaniczne podłączone są do komórek nerwowych szczura, które nim sterują. **Przejawia on zatem pewną autonomię, ale brak świadomości.** Pozostaje jednak pytanie czy taki „twór” można nazwać maszyną, bowiem miejsce procesora zabrały komórki układu nerwowego.

### [Antoni Firej]

1. W rozumieniu świadomości prezentowanym przez Kartezjusza, świadomość jest przymiotem jedynie ludzkim. **Zwierzęta czy rośliny są autonomiczne, działają według własnego instynktu, ale świadomości nie mają.** Podobnie byłoby w przypadku maszyn.

2. Należy rozgraniczyć stopnie autonomii maszyny. Czy maszyna, która jest zdolna do samodzielnego zwiększania swojej wiedzy, jest już autonomiczna? **Jeśli tak, to należy pamiętać, że nie musi ona „zdawać sobie sprawy” ze swojego statusu, tj. wiedzieć, że jest maszyną.** Mając zdolność nauki wykracza już poza pierwotny kod twórcy, staje się od niego niezależna. Co z kwestią zależności od zasilania przez człowieka? W razie buntu maszyn, zawsze można je wyłączyć. Maszyna w pełni samodzielna musi móc sama pozyskiwać energię. Za to wciąż nie musi być świadoma swojej pozycji. Ostatecznie, czy zdolność do

samostanowienia, określania własnych celów, wymaga pełnej świadomości? Twórca wciąż ma możliwość zakodowania, ukrycia w maszynie tajnych poleceń, których nie będzie mogła zignorować. Za przykład może posłużyć tu armia klonów z Gwiezdných Wojen. Jednostki te miały pojęcie na temat swojego istnienia, pozycji w społeczeństwie, były samowystarczalne i samodzielnie określały swoje cele. Jednocześnie miały zakodowany zestaw poleceń, których nie mogły zignorować i o którego istnieniu nie miały pojęcia, do czasu wywołania komendy. Nie były w pełni świadome. Przykładem takich zasad mogą też być prawa robotyki Asimova.

3. W nawiązaniu do poprzedniego punktu, brak podstawowych zasad, które określają działanie maszyny, zapobiegają jej niewłaściwemu działaniu, sprawia że nie jest już ona maszyną, a jednostką w pełni niezależną. **Maszyna autonomiczna musi być w pewnym stopniu ograniczona, aby pozostać maszyną.**

[Piotr Paszke]

1. Na drodze ku świadomości maszyn czyhają poważne niebezpieczeństwa. Przede wszystkim, może się okazać, że pokonujemy ją szybciej, niż się spodziewaliśmy: sztuczna inteligencja, uzyskawszy zdolność do reprodukcji własnego kodu i korzystająca z informacji zwrotnej ze świata zewnętrznego, będzie doskonalić się z każdą minutą. Może wtedy wyewoluować w korzystnym lub niekorzystnym z naszego punktu widzenia kierunku poza naszą kontrolą.

2. Maszyny jako urządzenia, które ludzie stworzyli w celu realizacji jakiegoś zadania muszą realizować cel, do których zostały przeznaczone. Nawet jeśli będą autonomiczne, jeśli potraktujemy świadomość jako rozumienie sensu istnienia, **poczucie tej świadomości dla maszyny jest zupełnie niepotrzebne, ponieważ może tylko skomplikować realizację celu dla którego ta maszyna została stworzona.**

3. Kiedy robot działa wyłącznie na poziomie motorycznym, wszystko przebiega szybko i sprawnie. Kiedy musi się na coś „skupić”, czasami popełnia błędy. Ten sam efekt można zauważyć u sportowców. Jeśli wykonują daną czynność odruchowo, korzystają z podstawowego systemu powiązanego bezpośrednio z ich układem motorycznym. Po latach treningów najczęściej wychodzi im to świetnie. **Jednak gdy zaczynają się świadomie zastanawiać nad tym, co robią, efekt jest gorszy.** Robot z własną świadomością będzie działać identycznie.

4. Maszyna stworzona przez człowieka musi być całkowicie kontrolowana przez swojego twórcę. Człowiek, który sam napisał algorytm, który nie ulega zmianom pod wpływem samodoskonalenia się, może w pełni kontrolować zachowania maszyny realizującej ten algorytm. **Natomiast świadomy kod, który sam się doskonali, może popełnić błąd w ocenie sytuacji, jak każda inna świadoma istota** i w pewnym momencie wymknąć się spod kontroli człowieka, realizując niepoprawne zadania.

5. Teoretycznie można takiej maszynie zamontować wyłącznik (przycisk bezpieczeństwa), jednak świadomy robot szybko sobie poradzi z takim zabezpieczeniem. Świadomy robot - czyli taki, który „czuje” swoją egzystencję - uzna takie zabezpieczenie za zagrożenie dla samego siebie.

[Przemysław Kopka]

1. **Zdolność do energetycznego samozasilania, nie wymaga świadomości.** Proste formy życia posiadają wbudowane mechanizmy, gwarantujące niezależność energetyczną i chęć przetrwania.

2. Mówi się, że czasami warto przespać się z problemem. Następnego dnia, kiedy wstajemy rano, przychodzi nam do głowy pomysł an rozwiązanie zagadki. **Wydaje się, że istnieje mechanizm w naszym mózgu, który w trakcie snu „nieświadomie” rozwiązuje nasz problem.** Być może w taki sposób mogłaby działać maszyna autonomiczna.

**[Łukasz Kurcius]**

1. "Bunt maszyn". Świadome maszyny autonomiczne mogą z czasem nauczyć się pewnych rzeczy (uczucie w stosunku do innej maszyny/człowieka, obrony własnego istnienia), których rezultatem będzie wniosek, że człowiek jest ich wrogiem.

2. Dlaczego maszyny miałyby być świadome? By wykonywać lepiej swoją pracę? Uważam, że **brak świadomości nie powodowałby zmniejszenia ich skuteczności**, jeśli byłyby odpowiednio kontrolowane przez człowieka.

3. Aspekt społeczny. Społeczeństwo **mogłoby wpaść w panikę**. Maszyny odbierały by masowo pracę ludzi. Ludzie bali by się „buntu maszyn”.