

Jowita Hącia

(studentka II roku wydziału MiNI Politechniki Warszawskiej)

Czy świadomość można zdefiniować i/lub wymodelować komputerowo?

Stworzenie modelu komputerowego wymaga uprzedniego przyjęcia definicji obiektu czy zjawiska, którego model zamierzamy stworzyć. Spróbujmy zatem odpowiedzieć na niełatwe pytanie, czym jest świadomość.

Można by stwierdzić, że świadomość jest wiedzą o stanie otoczenia, w którym się aktualnie znajdujemy oraz stanie naszego organizmu. Tak pojmowaną świadomość posiada każda istota żywa. Obejmuje ona informacje o potrzebach organizmu, panujących warunkach atmosferycznych, obiektach znajdujących się w pobliżu etc. Jest więc nierozdzielnie związana z odbieraniem bodźców zewnętrznych i wewnętrznych poprzez organy zmysłów i stanowi podwalinę instynktu przetrwania. To ona determinuje działania takie, jak szukanie ciepłego schronienia podczas mrozu, przyjemnego chłodu w upalny dzień czy strawy w momencie odczuwania głodu.

Powyższy opis nie oddaje jednak w pełni istoty świadomości, stanowi jedynie jej prymitywny rdzeń, korzeń, podstawę do rozwijania obserwacji samego siebie. O ile zawarta w nim charakterystyka dotyczyła zachowań instynktownych, bezmyślnych, opartych na prostym szablonie bodźca i reakcji na niego, schemacie [*potrzebuję – biorę*], o tyle to rozszerzenie pojęcia świadomości, o którym teraz mowa, wiąże się z poznaniem własnego celu i przemyśleniem, jakie skutki podejmowane działanie będzie miało dla nas i otoczenia. Można powiedzieć, że jest to ta część naszego umysłu, naszego jestestwa, która powstrzymuje przed danym zachowaniem, kiedy przynosząc chwilową korzyść i zaspokojenie nam samym, może ono okazać się szkodliwe dla innych lub też zgubne dla nas, kiedy rozpatrywane długoterminowo. Jest to zatem pierwiastek, który powstrzymuje przed bezmyślnym czynieniem szkody, gonieniem za zaspokojeniem chwilowych potrzeb czy zachcianek, a który związać można by z pojęciami takimi jak moralność, etyka, czy choćby empatia.

Złożenie tych dwóch aspektów nadal jednak nie wystarcza, by zdefiniować świadomość i przywodzi na myśl system wyuczonych zachowań, zdając się łączyć niektóre z założeń funkcjonalizmu i behawioryzmu.

Ostatnim w moim mniemaniu brakującym ogniwem w łańcuchu świadomości jest jeszcze coś, co nazwać można by samoświadomością. Owa trzecia składowa dotyczy świadomości myślenia, świadomości własnej motywacji, kierujących nami uczuć, świadomości, dlaczego podejmujemy takie, a nie inne działanie. Zakłada ona, że nie istnieje wyuczona, w pewnym stopniu już niezależna od nas reakcja na dany bodziec, wykształcona na skutek ciągu podobnych doświadczeń, lecz że nasze postępowanie jest elastyczne i zmienia się czy też każdorazowo ma szansę na metamorfozę przy towarzyszącej zmianie najmniej nawet na pozór istotnego elementu otoczenia. Innymi słowy, ta warstwa świadomości zmusza do każdorazowego przemyślenia podejmowanej czynności niezależnie od analogii sytuacji ją wywołującej, niepodawania się rutynowym zachowaniom. Jest tym, co każe oddać życie za drugą osobę lub ideę, tym, co wywołuje zachowanie z pozoru nielogiczne i pozbawione sensu, być może szkodliwe dla nas samych, bo wkracza w sferę emocji, niezwykle indywidualną i ulotną, a często niezrozumiałą dla nas samych.

Dopiero te trzy „warstwy” zrównoważone ze sobą i ze sobą współistniejące uznać można za świadomość. Żadna z nich nie powinna przysłaniać innych, żadna też nie powinna zostać zaniedbana.

Nieprawidłowe funkcjonowanie pierwszej z nich prowadzić może do autodestrukcji – nie mając pełnego obrazu otoczenia i siebie w tym otoczeniu, łatwo przeoczyć jakieś zagrożenie zewnętrzne. Istnieją również upośledzenia układu nerwowego (HSAN - dziedziczne neuropatie czuciowe i autonomiczne), którym towarzyszy brak odczuwania bólu, co stanowi rodzaj zagrożenia wewnętrznego, choćby przez wzgląd na nieodczuwalnie objawów chorobowych.

Wówczas funkcje upośledzonej części przejmują pozostałe „części” świadomości z „działającymi” elementami tej pierwszej – na podstawie rozumowej oceny zaistniałej sytuacji i możliwych do zbadania warunków, można dane zdarzenie zakwalifikować jako jedno z dotychczas poznanych, niejako dopasować do zdobytych doświadczeń i podjąć odpowiednie dla zaistniałej sytuacji kroki.

Pominięcie drugiej i trzeciej „warstwy” skutkuje z kolei dehumanizacją, której typowym przejawem jest bezwolne poddanie się własnym żądzom i instyktom. Dlatego ważne jest przenikanie się trzech wymienionych aspektów świadomości.

Ponadto świadomość dotyczy przeszłości, teraźniejszości i przyszłości – każdej z nich w takim samym stopniu. Skupiając się zanadto na przeszłości lub przyszłości, zaniedbujemy chwilę obecną i sprawiamy, że obecność „tu i teraz” staje się obecnością pustą i podlegającą instyktom. Do teraźniejszości odnosi się tylko pierwsza, podstawowa część świadomości – pozostałe składowe dotyczą innego czasu i miejsca. Zamiast poddawać analizie aktualnie

podejmowaną decyzję, reagujemy tylko na najprostsze bodźce, całość sytuacji wręcz na siłę dopasowując do tej, na której jesteśmy skupieni lub w ogóle tej całości nie obejmując. Stajemy się zombie z pustym wzrokiem, obecnym ciałem, lecz nie duchem. Podporządkowujemy się schematom. Widzimy, nie widząc. Widzimy człowieka, nie dostrzegając w nim istoty ludzkiej. Widzimy człowieka, dostrzegając w nim kogoś innego – kogoś, kim nie jest, a kogo nam jedynie przypomina. Sądzymy go, sądząc tak naprawdę kogoś innego. Nie mamy zatem pełnego obrazu sytuacji, nie jesteśmy jej w pełni świadomi. Znajdujemy się w gabinecie luster, usilnie próbując odnaleźć znane już odbicia. Dlatego też niezwykle ważny jest podział świadomości między to, co było, co jest, a co ma się za chwilę wydarzyć.

Spróbujmy przeanalizować próbę wymodelowania każdego z wyżej wymienionych aspektów świadomości. Model świadomości siebie w danym otoczeniu nie przedstawia dużej trudności. Wymaga jedynie zastosowania odpowiedniej ilości czujników i mierników, które swoje wskazania przesyłałyby maszynie, by ta z kolei na podstawie tak zebranych danych wybierała odpowiadający danemu stanowi otoczenia algorytm zachowania. Dla tej warstwy problemem zdaje się jedynie fakt, że maszyna sama nie jest w stanie określić, czy jest jej za zimno, czy za gorąco, czy czegoś jej brakuje, czy też przeciwnie – występuje nadmiar jakiegoś czynnika lub zasobu. Potrzebuje jakichś odgórnie, zewnętrznie ustalonych wartości granicznych, na podstawie których mogłaby dopiero stwierdzić, jakie są jej potrzeby i ‘odczucia’, a przecież odczuwanie chociażby temperatury jest sprawą indywidualną i zmienną w zależności od stanu zdrowia czy nawet upływu czasu. Model tej części jest zatem możliwy do skonstruowania, lecz ogranicza się do schematycznego modelu świadomości „zombie”.

Druga warstwa świadomości – świadomość skutków danego działania zdaje się już nieco bardziej skomplikowana, ale nadal możliwa do zaimplementowania. Sprowadza się do wewnętrznego modelu samego siebie oddziałującego z wewnętrznym modelem otoczenia. Dzięki wcześniej wspomnianym receptorom do maszyny trafiają informacje o aktualnym stanie otoczenia. Na tym etapie potrzebne jest także przewidywanie stanu otoczenia po wybraniu danego działania odpowiadającego aktualnemu stanowi. Maszyna powinna zatem móc wykonać symulację na wewnętrznym, dopasowanym do chwili obecnej modelu otoczenia, zanim zostanie wykonana przezeń operacja w otoczeniu prawdziwym.

Pozostaje jeszcze trzeci aspekt świadomości. I o ile maszyna może wiedzieć, co nakazało jej postąpić w dany sposób, o tyle zawsze będzie to albo programista, albo wiedza o skutkach uprzednio podjętych działań w

analogicznej sytuacji. W moim mniemaniu samoświadomości nie można stworzyć komputerowo, ma ona bowiem związek z uczuciami i odczuciami.

Jak już wcześniej wspomniałam, „odczucia” maszyny są zależne od jej ustawień (np. programistycznych), a te określa jej twórca. Od czego zależą zatem jej uczucia i czy w ogóle możemy mówić o uczuciach urządzenia elektronicznego? Nie ukrywajmy, mój kalkulator – choćby miał ręce, nogi i głowę na karku – raczej mnie nie pokocha. Może być mi posłuszny, jednak nie będzie to wynikało z jego woli, a z samych założeń jego istnienia. Możemy oczywiście zaprogramować maszynę tak, by nas „lubiła”, a w wyniku określonych sytuacji „lubić” przestawała. Nie będzie to jednak miało dużego związku z lubieniem w znanym ludziom rozumieniu, będzie to jedynie umiejętność gromadzenia danych warunkujących przychylną bądź nieprzychylną wobec pewnych jednostek.

Suma summarum – choć wydaje mi się, że można stworzyć lepszy lub gorszy model świadomości człowieka, nie dopuszczam jednak, aby maszynę można było w ludzką świadomość wyposażać ze względu na niemożliwość pełnego jej zdefiniowania. Procesy „myślone” maszyny będą zawsze oparte na algorytmach – czy to popartych, czy nie popartych logiką – ale nigdy na uczuciach, jakich mogą doświadczać ludzie.

Komentarz wykładowcy

Zamiast typowego komentarza przedstawiam niżej pewne ciekawsze wyjątki z korespondencji e-mailowej, która towarzyszyła ocenianiu pracy (skrót J.H. odnosi się do autorki eseju, czyli Jowity Hąci; skrót P.S. do wykładowcy, czyli Pawła Stacewicza).

P.S.:

- Pisze Pani w eseju: „*Nie ukrywajmy, mój kalkulator - choćby miał ręce, nogi i głowę na karku - raczej mnie nie pokocha*”.

Tego oczywiście Pani nie życzę; ale... gdyby był to bardzo skomplikowany kalkulator, zdolny rozwijać się, doskonalić, nabywać doświadczenia i formować w ten sposób jakoś własne „maszynowe ja”... Kto wie, czy coś by tam w nim nie powstało na kształt odczuwania, przeżywania itp.

- Pisze Pani także: „*nie dopuszczam jednak, aby maszynę można było w ludzką świadomość wyposażać ze względu na niemożliwość pełnego jej zdefiniowania.*”

Taka argumentacja chyba nie jest przekonująca. Niemożność zdefiniowania świadomości dziś nie wyklucza sztucznego wytworzenia czegoś, co będzie wykazywało wszelkie znamiona ludzkiej świadomości. Być może też owo coś (np. jakaś maszyna, nawet i taka, która oparta jest na żywych komórkach, czyli maszyna po części biologiczna) pozwoli nam zdefiniować świadomość; będzie podstawą dobrej definicji.

J.H.:

- Być może mam małą wyobraźnię lub powątpiewam w zdolności ludzkiego umysłu, jednak sądzę, że nie znając dokładnego „biologicznego” źródła świadomości (co więcej, pokusiłabym się o stwierdzenie, że takie źródło w ogóle nie istnieje w sposób materialny), nie jesteśmy w stanie jako ludzkość „dać” maszynie świadomość, jaką sami posiadamy.

P.S.:

- Nawias wewnątrz powyższej uwagi sugeruje, że nie ma tu Pani ostatecznej pewności. I słusznie. „Nie-istnienie w sposób materialny” może oznaczać co najmniej dwie rzeczy:

a) istnienie w sensie jakiegoś zbioru relacji, wiążących elementy materialne; można by rzec: istnienie formalne; algorytm np. jest takim bytem;

b) istnienie duchowe, zupełnie rozłączne z istnieniem materialnym, niepoznawalne za pośrednictwem czynników materialnych; takim bytem byłaby np. kartezjańska dusza.

J.H.:

- Pisząc w eseju, że nie można maszyny „w ludzką świadomość wyposażyć”, miałam na myśli, że być może faktycznie możliwe jest, aby na zewnątrz maszyna wyglądała na świadomą, ale wewnętrzne mechanizmy różnić się jednak będą od naszych wewnętrznych procesów. A taka świadomość, jaka wyewoluuje w maszynie, niekoniecznie będzie świadomością stricte ludzką, nie powinna więc służyć jako podstawa do definiowania świadomości w ogóle, a jedynie świadomości maszyn.

P.S.:

- Bardzo celne spostrzeżenie.

Komentarz byłby taki: jeśli uznamy maszynę za świadomą, to znaczy, że ustaliliśmy wcześniej jakieś ogólne kryterium świadomości; a jeśli tak, to analizując działanie maszyny, możemy zbliżyć się do jakiejś cząstkowej chociaż definicji świadomości, czyli czegoś, co spełnia ustalone wcześniej kryterium.

(...)