

ANTROPOCENTRYCZNE KRYTERIA „OSWAJANIA” SZTUCZNEJ INTELIGENCJI

Antropocentryzm uznaje prymat człowieka i tego, co ludzkie, nad innymi aktorami (bytami) nie-ludzkimi, także technologicznymi (na przykład zrobotyzowanymi maszynami). Obejmuje następujące komponenty: (1) poznawczy, zakładający konieczność przyjmowania perspektywy ludzkiej w interpretacji każdej sytuacji; (2) ontologiczny, przyznający człowiekowi wyjątkową pozycję w hierarchii bytów; (3) metodologiczny, koncentrujący się na sprawczej roli człowieka, dopuszczający jednak ingerencję bytów nie-ludzkich; (4) aksjologiczny, zakładający nadrzędność i bezdyskusyjność wartości ludzkich.

KRYTERIUM POZNAWCZE

Antropocentryzm poznawczy podkreśla konieczność interpretowania wszelkich zagadnień z ludzkiego punktu widzenia zarówno podczas procesu rozumowania, jak i dokonywania wyborów moralnych, niezależnie od jednostkowego charakteru danej sytuacji. Generuje więc pytania o możliwości poznawcze robotów humanoidalnych, umiejętność podejmowania decyzji

i dokonywania wyborów oraz ich konsekwencje, w niniejszym artykule traktowane jako kryteria „oswajania” technologii AI.

Powyższe pytania na gruncie robotyki społecznej pojawiają się w kontekście inspiracji ludzkimi funkcjami poznawczymi, które próbuje się odtwarzać w robotach. Dotyczą one przede wszystkim integracji procesów sensomotorycznych, będących podstawą do rozumowania czy też uczenia się przez doświadczenie. Przyjęcie paradygmatu ucieleśnionego umysłu miało za zadanie przezwyciężenie impasu wynikającego z piętrzących się problemów implementacji logiki formalnej do badań nad procesami poznawczymi systemów sztucznej inteligencji, które nie przyczyniły się do znaczącego postępu w budowie inteligentnych maszyn. Robotom brakowało – i nadal brakuje – podstawowej zdroworozsądkowej wiedzy uniemożliwiającej wyciąganie podstawowych wniosków. Zdano sobie sprawę, że w codziennych praktycznych rozumowaniach niezbyt często mamy do czynienia z dowodami, które można spotkać na gruncie matematyki. Stąd też zwrot w kierunku oparcia procesów poznawczych na ucieleśnieniu sztucznej inteligencji zapoczątkował nowy etap w badaniach nad interakcją sztucznego systemu poznawczego ze środowiskiem, w którym on działa, w tym z człowiekiem (robotyka społeczna)¹⁸. Naturalnym sposobem okazało się sięganie do analogicznych rozwiązań spotykanych w świecie przyrody, jak również aplikowanie wyników badań z zakresu społecznych interakcji międzyludzkich w obszar konstruowania robotów społecznych, w tym zwłaszcza robotów humanoidalnych, które zaczęto wyposażać w liczne sensory i receptory¹⁹. Miało to ułatwić robotom zbieranie informacji dotyczących otoczenia ich działania, stanowiących reprezentację faktów, nad którą miałyby one zacząć „myśleć”, a następnie przekładać wyniki „przemysleń” na działania i próby zwerbalizowania ich w języku naturalnym. Stąd też obok badań nad procesami sensomotorycznymi można wskazać na te dotyczące przetwarzania języka naturalnego. Robot przetwarza język nie tylko w kontekście rozumienia wypowiedzi i przypisania jej odpowiedniego znaczenia, co w rezultacie ma doprowadzić do podjęcia adekwatnej decyzji i działania (należy podkreślić, że i tutaj dużą trudność sprawiają wyrażenia okazjonalne, funkcjonujące w języku potocznym, łączenie komunikacji werbalnej z pozawerbalnymi sygnałami oraz wieloznaczności dotyczące nie tylko samych słów, ale i wyrażen złożonych, na przykład amfibolii, elips, ekwiwokacji). „Zdolności komunikacyjne” robota działającego w określonym środowisku opierają się w dużej mierze na możliwości wiązania etykiet werbalnych

¹⁸ Zob. A. K i s i e l e w i c z, *Sztuczna inteligencja i logika*, Wydawnictwo Naukowe PWN, Warszawa 2017.

¹⁹ Zob. F. D e l c o m y n, *Biologically Inspired Robots*, w: *Bioinspiration and Robotics Walking and Climbing Robots*, red. M.K. Habib, I-Tech, Vienna 2007.

z przedmiotami znajdującymi się w polu jego działania. Generuje to znaczne trudności w konstruowaniu maszyn radzących sobie w dynamicznie zmiennym środowisku zarówno fizycznym, jak i społecznym. Rozwiązaniem problemu może być uczenie maszynowe (w tym: uczenie nadzorowane, półnadzorowane, nienadzorowane, uczenie na bieżąco, uczenie ze wzmocnieniem).

Z pytaniem o możliwości poznawcze robotów wiąże się problem ich autonomiczności. Ogólnie możliwości te można określić jako warunki pozwalające na sprawne działanie w środowisku fizycznym, które umożliwiają rozwiązywanie problemów bez ingerencji człowieka. Takie rozumienie autonomiczności wiąże się jednak z pewnymi ograniczeniami. O ile na gruncie rozważań antropocentrycznych za byt autonomiczny możemy przyjąć ten, który posiada możliwość rozwijania wolnej woli, co w konsekwencji czyni go niezależnym od czynników zewnętrznych w procesie podejmowania decyzji, o tyle w odniesieniu do autonomii bytów nie-ludzkich uwidacznia się pewne uproszczone jej rozumienie. Jest ono konsekwencją próby uniknięcia wikłania się w dyskusje dotyczące posiadania przez roboty wyżej wspomnianej dyspozycji w takim stopniu, w jakim przysługuje ona człowiekowi. Widoczne są w tym przypadku nawiązania do etologii i rozumienia autonomiczności poprzez odniesienie jej do sprawności działania zwierzęcia w zmiennych warunkach otoczenia. Nie pomija się w tym przypadku zagadnienia świadomości i samoświadomości, choć ich znaczenie jest mniejsze niż w kontekście dyskusji antropocentrycznych.