

<http://dx.doi.org/10.16926/fil.2014.11.09>

Dariusz DĄBEK

Akademia im. Jana Długosza w Częstochowie

Podstawowe problemy klasycznej koncepcji sztucznej inteligencji i nauk kognitywnych – część 1

Streszczenie

Tekst artykułu wprowadza w interesujący i problematyczny świat sztucznej inteligencji, wyjaśniając podstawowe pojęcia rządzące wirtualną rzeczywistością. Naświetla kilka głównych problemów towarzyszących projektowaniu sztucznych systemów, których pokonanie wciąż stanowi barierę do skonstruowania inteligencji nie gorszej od ludzkiej.

Słowa kluczowe: sztuczna inteligencja, nauki kognitywne, inteligentny agent, problem ramy, problem ucieleśnienia, światy wirtualne.

W ramach wstępu

Obszar zagadnień sztucznej inteligencji, *Artificial Intelligence*¹, jest od każdej strony wysoce interesujący i problematyczny. Nie jest to prosta dyscyplina, ponieważ obejmuje swoim zasięgiem szereg obszarów naukowych i omawianie jej w jednym tylko kontekście prowadzi do wielu nieporozumień i fałszywych wniosków. Koniecznym wydaje się ujednoczenie szeregu pojęć i przyjęcia w badaniach jednakowych narzędzi i sposobów działania, co w przypadku AI jest wyjątkowo trudne. Już samo określenie AI jest problematyczne. Użycie sformułowania „sztuczna inteligencja” sugeruje zbudowanie sztucznego modelu inteligencji opartej na najbliższej znanej nam ludzkiej inteligencji. Zakładając, że nasze rozumowanie jest prawdziwe, natrafimy na kolejną, wyjątkowo nie-wdzięczną do rozwikłania trudność, ponieważ jak dotąd trwają spory pomiędzy

¹ *Artificial Intelligence* – termin wprowadzony i zaproponowany po raz pierwszy przez Johna McCarthy’ego (jako nazwa konferencji odbywającej się w 1956 r.).

naukowcami co do jednoznacznego stanowiska, dotyczącego tego, czym tak naprawdę jest inteligencja.

Różne nauki starają się różnie określać inteligencję w ogóle i w zależności od obszaru badań różnie ją definiują. Inaczej psychologia poznawcza, inaczej lingwistyka, inaczej neurobiologia, filozofia i inne. W przypadku braku ogólnie akceptowanej definicji inteligencji, gdzie nie ma zgody nawet co do tego, czym tak naprawdę jest inteligencja, zagadnienia z obszaru sztucznej inteligencji natrafiają na szereg poważnych trudności. Obszar badań staje się niejasny i właściwie nie do pokonania.

Podawane do tej pory bardziej intuicyjne przybliżenia rozumienia AI oparte na ludzkiej inteligencji wydają się z góry skazane na niepowodzenie. Budowane w ten sposób definicje inteligencji są raz za szerokie, a raz zbyt specjalistyczne i za wąskie. Poza tym większość badaczy tego tematu zgodnie podkreśla brak jednoznacznej odpowiedzi na pytanie: co to jest inteligencja? Mówiąc o sztucznej inteligencji opartej na zasadach działania ludzkiej inteligencji, to co najwyżej cel, ewentualnie wyidealizowane dążenie twórców inteligencji innej niż ludzka. Uściślając nieznacznie powyższe spostrzeżenia, nie należy bezpośrednio utożsamiać inteligencji naturalnej (ludzkiej) z jej sztucznym odpowiednikiem.

Jak zatem traktować określenie sztucznej inteligencji? Z jednej strony, „sztuczna” to nie to samo co naturalna po prostu, z drugiej zaś – „sztuczna” to jak najbardziej podobna do naturalnej, ale nie biologiczna. Zbudowana z innych niż biologiczna komponentów, ale dążąca do jak najwierniejszego odwzorowania (naśladowania) naturalnej. Posiadać ma jednak inne niż naturalne założenia lub sposoby działania. Z jednej strony, ma udawać i wiernie naśladować działania neuronów i pracy ludzkiego mózgu, z drugiej natomiast zaś strony ma być czymś pozabiologicznym, pozbawionym wad i ułomności. Dla jednych AI jest dyscypliną wielu sprzeczności i wykluczających się wzajemnie idei. Dla drugich, przez swą różnorodność i wieloaspektowość, jest źródłem tego, co najistotniejsze w dotarciu do inteligencji w ogóle, a poza tym wspaniałym polem do badań. Dlatego niektóre z pośród stanowisk zostaną z konieczności potraktowane marginalnie, aby móc skoncentrować się jedynie na najistotniejszych (kilku reprezentatywnych) problemach sztucznej inteligencji.

Zasadnym wydaje się zatem postawienie pytania: czy praca na styku tak wielu przecież różnych dyscyplin naukowych jest w ogóle możliwa, a jeśli tak, to jak to zrobić? Sam fakt udziału w rekonstrukcji sztucznej inteligencji wielu różnych dyscyplin naukowych, biorących udział w tak rozległym obszarze badań, jest w swym zakresie oszałamiający, ale i fascynujący. Można mieć zatem uzasadnione wątpliwości co do naukowej konstrukcji problematyki AI. Niewątpliwie jednak należy stwierdzić, że wśród zagadnień AI da się wyraźnie wyróżnić zarówno przedmiot badań, jak i specyficzny obszar zagadnień, a poszczególne problemy, jak na naukę przystało, da się ściśle i jednoznacznie określić.

Kłopoty z terminami

Poruszając problematykę sztucznej inteligencji, na pierwszy plan wysuwają się kłopoty terminologiczne. Używając do tej pory ogólnie terminu „sztuczna inteligencja”, ma się na uwadze przeważnie jej dwa rozumienia. Jedno, jako historycznie przyjętego skrótu AI do językowego określenia „sztucznej inteligencji”, albo drugie, jako zbioru zagadnień związanych ze sztuczną inteligencją.

Innym kłopotliwym określeniem jest angielskie *cognitive science*, które traktuje się w pracach naukowych jako nauki kognitywne bądź nauki kognitywistyczne (obie formy jako dopuszczalne przewijają się w naszym obszarze językowym). Kłopot polega na tym, że w literaturze anglojęzycznej *cognitive science* występuje w liczbie pojedynczej, a w naszych rozważaniach na ten temat będziemy odwoływać się do szeregu różnych nauk, dlatego aby to podkreślić, używamy liczby mnogiej. Oczywiście zakres badań obejmujących *cognitive science* jest dużo szerszy niż zagadnienia samej sztucznej inteligencji. Nie mniej jednak przedmiot badań AI w całości mieści się w przedmiocie badań *cognitive science* i z tego względu właściwe rozumienie tego określenia wydaje się dla AI kluczowe.

W polskiej literaturze naukowej przyjęło się stosować kilka nazw dla denotacji *cognitive science*. Najpopularniejsze z nich to: „nauka poznawcza” (zaproponowana przez T. Maruszewicza i K. Domańską)², „nauka o poznawaniu” (wprowadzona przez Idę Kurcz w 1987 r.)³, „nauka o poznaniu” (zaproponowana przez J. Kozińskiego w 1996 r.)⁴, „nauki kognitywne”⁵, „nauka o procesach poznawczych” (zaproponowana przez K. Gurbę w 1995 r.)⁶ i „kognitywistyka” (zaproponowana przez Z. Chlewińskiego i A. Klawitera w 1999 r.)⁷. Używa się także terminu nieprzetłumaczonego „*cognitive science*” (zaproponowane przez J. Bobryka w 1987 r., potem w 1994 r., oraz R. Piłata w 1992 r.)⁸.

Definicje *cognitive science*

W literaturze przedmiotu funkcjonuje wiele definicji *cognitive science*. Pozwalają one zwięźle określić, czym zajmuje się ta dyscyplina, jednak przyjmują

² Por. T. Maruszewski, *Psychologia poznawcza*, Polskie Towarzystwo Semiotyczne, Warszawa 1996; K. Domańska, *Metafora komputerowa w psychologii poznawczej*, [w:] *Psychologia i poznanie*, red. M. Materska, T. Tyszka, Warszawa 1997, s. 12–37.

³ Por. I. Kurcz, *Język a reprezentacja świata w umyśle*, PWN, Warszawa 1987, s. 16 i nn.

⁴ Por. J. Koziński, *Koncepcje psychologiczne człowieka*, Żak, Warszawa 1996, s. 8–32.

⁵ Por. *Encyklopedia filozofii*, red. T. Honderich, Zysk i S-ka, Poznań 1998.

⁶ Por. K. Gurba, *Sztuczna inteligencja z naturalnymi ograniczeniami*, „Znak” 1995, nr 9 (484).

⁷ Por. Z. Chlewiński, *Umysł. Dynamiczna organizacja pojęć*, Wydawnictwo Naukowe PWN, Warszawa 1999, s. 6 i nn.; A. Klawiter, *Umysł a rzeczywistość*, „Poznańskie Studia z Filozofii Humanistyki”, t. 5 (18), red. A. Klawiter, L. Nowak, P. Przybysz, Poznań 1999.

⁸ Por. J. Bobryk, *Locus umysłu*, Zakład Narodowy im. Ossolińskich, Wrocław 1987; tegoż, *Akty świadomości i procesy poznawcze*, Zakład Narodowy im. Ossolińskich, Wrocław 1994; R. Piłat, *Teorie semantyczne a informatyczne podejście w epistemologii*, [w:] *Pogranicza epistemologii*, red. J. Niżnik, IFiS PAN, Warszawa 1992.

różne punkty odniesienia; metody, przedmiot badania czy też stosunek do innych nauk. Ogólnie można stwierdzić, że jest ona interdyscyplinarną nauką zajmującą się badaniem systemów przetwarzających informację w sposób tak wysoce zorganizowany, że ich działanie może być nazwane inteligentnym. Systemy te określa się jako systemy poznawcze (*cognitive systems*). Przetwarzanie informacji przez systemy poznawcze polega na manipulowaniu reprezentacjami, które odnoszą się do świata zewnętrznego. *Cognitive science* zakłada, że proces ten daje się opisać w sposób formalny. Ma on charakter obliczeniowy (*computable*). Dzięki temu uzasadnione jest wykorzystywanie metod symulacji komputerowej do badania umysłu człowieka.

W ramach porównania właściwym wydaje się podanie kilku ugruntowanych w literaturze przedmiotu definicji. Aby zachować również i historyczny charakter badań, poniżej przytoczymy kilka starszych, ale i chyba reprezentatywnych definicji. Poniższa encyklopedyczna definicja, pochodząca z 1990 r., kładzie nacisk na wpływ innych nauk na badania prowadzone w obrębie nauki poznawczej:⁹

Termin *cognitive science* odnosi się do interdyscyplinarnego studium dotyczącego nabywania i użycia wiedzy. W studium to wkład wnoszą: nauka o sztucznej inteligencji, psychologia, lingwistyka, filozofia, antropologia, neurofizjologia [*neuroscience*] i nauki o wychowaniu. Ruch ten odznacza się szerokim zasięgiem i zróżnicowaniem, zawierając w sobie kilka punktów widzenia. Nauka poznawcza rozwinęła się dzięki trzem osiągnięciom: wynalazkowi komputerów i próbom stworzenia programów wykonujących zadania, które właściwe są ludziom; rozwojowi psychologii poznawczej [...]; oraz rozwinięciu się w dziedzinie lingwistyki teorii gramatyki generatywnej i teorii jej pochodnych. *Cognitive science* jest syntezą zajmującą się istotą wiedzy, na jakiej opiera się ludzkie poznanie, procesami przetwarzania informacji przez człowieka i komputerowym modelowaniem tychże procesów. Istnieje pięć głównych pól badawczych w nauce poznawczej: reprezentacja wiedzy, język, uczenie się, myślenie i percepcja¹⁰.

Podobna do poprzedniej jest definicja (wprowadzona 1997 r. przez P. Salo):

Z historycznego punktu widzenia *cognitive science* była całkiem nową propozycją w naukowym badaniu umysłu: był to pogląd, że ludzki umysł jest systemem manipulującym symbolami, pewnego rodzaju komputerem. Tak więc na początku *cognitive science* była fuzją nauki o sztucznej inteligencji i psychologii poznawczej. [...] Obecnie, także lingwistyka i neurofizjologia [*neural sciences*] dołączają do owej fuzji, łącząc się coraz bardziej. Dołączają także inne działy psychologii. Filozoficzna debata nad tymi zagadnieniami jest gorąca, zróżnicowana i szersza niż kiedykolwiek, nauka o komputerach [*computer science*] dostarcza podstawy do weryfikacji teorii [...]¹¹.

⁹ Tłumaczenia poniższych definicji pochodzą z dwóch źródeł: <http://www.kognitywistyka.net/kognitywistyka/definicje.html> [stan z 12.11.2014] oraz A. Chuderski, *Wykorzystanie metod sztucznej inteligencji w badaniach nad umysłem*, <http://www.kognitywistyka.net/artykuly/achwmsiwnu.pdf> [stan z 12.11.2014], w nieznanym tłumaczeniu.

¹⁰ *The Blackwell Dictionary of Cognitive Psychology*, red. M. Eysenck, Brasil Blackwell, Cambridge, MA 1990, cyt. za: <http://www.kognitywistyka.net/kognitywistyka/definicje.html> [stan z 12.11.2014].

¹¹ P. Salo, *What Is Cognitive Science?*, 1997, cyt. za: A. Chuderski, *Wykorzystanie...*

Ciekawą definicję *cognitive science* w psychologii poznawczej wprowadziła Domańska, która oprócz przedmiotu badań nauki poznawczej przedstawia także jej metodę badawczą:

Dyscyplina, która stawia sobie za zadanie wyjaśnienie przebiegu procesów poznawczych i czyni to, korzystając z metod symulacji komputerowej [...] *cognitive science* zajmuje się badaniem systemów poznawczych „w ogóle”, niezależnie od tego, czy występują u człowieka, czy w komputerze¹².

Na uwagę zasługuje również definicja zaproponowana przez Greena i in. z 1996 r.:

Zdefiniujmy zakres *cognitive science* jako interdyscyplinarne, naukowe studium umysłu. Praktyka i wiedza tej nauki są pochodną osiągnięć dyscyplin, które przyczyniły się do jej powstania: nauki o komputerach (informatyki), lingwistyki, neurofizjologii [*neuroscience*], psychologii, neuropsychologii poznawczej i filozofii. Celem jej jest zrozumienie działania umysłu w kategoriach procesów manipulacji na reprezentacjach. Umysł, a więc i podstawa inteligentnego zachowania się w świecie, jest widziany w kategoriach obliczeń [*computations*] albo przetwarzania informacji [*information-processes*]¹³.

Pojęcie agenta

Kłopotem jest również pojęcie agenta lub, ściślej, inteligentnego agenta – *intelligent agent*, wykorzystywane powszechnie w określaniu zachowań systemów sztucznej inteligencji. Nie ma jednej, precyzyjnej definicji agenta. Zwłaszcza różnica pomiędzy agentem a tzw. obiektem lub programem komputerowym jest nieprecyzyjna. W literaturze często pojawia się również pojęcie *M-agenta*, opisujące jakąś architekturę. Agent jest to element służący do budowy systemów zdecentralizowanych. Według tej koncepcji obiekt jest co prawda autonomiczny, ale nie ma inicjatywy działania, natomiast agent działa w oparciu o obserwacje otoczenia. Według innej definicji, opisującej system MRROC++, agentem jest cokolwiek, co postrzega środowisko i w jakiś sposób na nie oddziałuje. Dla dalszych rozważań przyjąć można, że agent jest systemem (programem) komputerowym usytuowanym w określonym środowisku, mogącym korzystać z określonych zasobów, zdolnym do autonomicznego działania, w celu osiągnięcia określonych celów i posiadającym motywację do działania.

Powyższa definicja nie precyzuje, co to jest środowisko oraz co to jest autonomia agenta. Można przyjąć, że agent musi posiadać jakieś czujniki, receptory, którymi jest w stanie odbierać sygnały wejściowe oraz elementy manipulacyjne (efektory), przez które jest w stanie wpływać na otaczające go środowisko.

¹² K. Domańska, *Metafora komputerowa...*, cyt. za: A. Chuderski, *Wykorzystanie...*, s. 7.

¹³ D. Grenn i in., *Cognitive Sciences. An Introduction*, Blackwell, Cambridge 1996, cyt. za: A. Chuderski, *Wykorzystanie...*, s. 7.

Inne zagadnienia

Do tej pory próbowano jedynie pobieżnie określić naturę inteligencji i omówić wstępnie tzw. paradygmat kognitywistyczny, który ciągle zbyt dominuje nauce i w ogólnym myśleniu o inteligencji. Da się zauważyć, że dotychczas w kilku miejscach zaznaczono tylko potencjalne problemy związane z tym paradygmatem. Zasadne wydaje się, aby w następnej kolejności bliżej przyjrzeć się temu, czym owe problemy naprawdę są i dlaczego pojawiają się w pierwszej kolejności. Paradygmat kognitywistyczny zaniedbuje fakt, iż inteligentni agenci, czyli ludzie, zwierzęta, roboty, są ucieleśnieni i funkcjonują w rzeczywistym świecie fizycznym, co prowadzi do istotnych braków przy próbie zrozumienia i wytłumaczenia zjawiska inteligencji.

Nakreślenie problemów paradygmatu kognitywistycznego i zrozumienie ich pochodzenia pomaga nam, z jednej strony, uniknąć popełniania wielokrotnie tych samych błędów, z drugiej – dostarcza inspiracji do szukania nowych sposobów rozwiązania problemu. Ze względu na obszerny charakter podejmowanych zagadnień zasadne wydaje się skoncentrowanie jedynie na przeglądzie kilku wybranych i reprezentatywnych przykładów. Zbiór zawartych poniżej zagadnień powinien być wystarczający, bez przytaczania bardziej obszernych szczegółów. Celem jest naszkicowanie głównych problemów, które historycznie doprowadziły pracowników naukowych do ponownego rozważenia ich podejścia do nauki o inteligencji.

Prawdziwe światy przeciwko światom wirtualnym

Klasyczne modele AI, tj. modele, które rozwinęły się w myśl paradygmatu kognitywistycznego, skupiają się na modelowaniu zaawansowanych procesów, takich jak rozwiązywanie praktycznych problemów, rozumowanie, formułowanie wniosków czy choćby (przytoczonej poniżej) praktycznej gry w szachy. W wielu z wymienionych zagadnień osiągnięto wielki postęp, gdy okazało się, że w przypadku modelowania gry w szachy komputer jest zdolny zagrać wystarczająco dobrze, by pokonać mistrza świata w tej dyscyplinie. W innych obszarach, postęp nie był tak szybki, na przykład jeśli chodzi o wizję samodzielnie działającego autonomicznego komputera. Okazało się, że wydobywanie informacji z aparatu fotograficznego w formie tablicy pikseli i przełożenie ich na całościowe obrazy przedstawiające świat stało się bardziej wymagające i skomplikowane niż tego oczekiwano. Głównym powodem tych trudności, a także powodem podstawowych problemów związanych z AI ogólnie jest to, że modele nie biorą prawdziwego świata wystarczająco pod uwagę. Wiele pracy w klasycznej sztucznej inteligencji zostało poświęcone stworzeniu albo – mówiąc ściślej – odwzorowaniu abstrakcyjnych światów wirtualnych ze szczegółowo zdefiniowanymi stanami i działaniami, które zupełnie nie przypominają prawdziwego świata.

Aby bliżej przyjrzeć się omawianemu zagadnieniu, można postarać się zobrazować je na przykładzie gry w szachy. Szachy są grą formalną. Reprezentują wirtualny świat z dyskretnymi, wyraźnie zdefiniowanymi stanami, pozycjami na planszy, możliwymi działaniami i dozwolonymi ruchami. Jest to również gra zawierająca kompletne informacje. Jeżeli znamy pozycje na planszy, możliwa jest zupełna i wystarczająca wiedza na temat tego, jak grać, ponieważ dla pewnej pozycji na planszy możliwe ruchy są dokładnie określone i ograniczone do pewnej liczby pól. Chociaż nie można znać dokładnego ruchu, który wykona nasz przeciwnik, posiadamy informację co do tego, że będzie to ruch zgodny z regułami. Zakładając oczywiście, że jeśli przeciwnik wykona niedozwolony ruch, to musiałby – zgodnie z zasadami – zaprzestać gry w szachy (nie bierzemy pod uwagę innych zachowań, niebędących częścią samej gry). Szachy są również grą statyczną, w tym sensie, że jeżeli nikt nie zrobi ruchu, nic się nie zmienia. Ponadto typy możliwych ruchów nie zmieniają się w czasie gry.

Innym podobnym pod tym względem przykładem, o nieco szerszych interakcjach z otoczeniem, może być działający z powodzeniem model automatycznego pilota. Występujące środki bezpieczeństwa w samolotach i innych pojazdach od lat sprawdzają się pod tym względem. Wystarczy przeanalizować działanie takiego układu, aby zobaczyć, że jest on pod względem budowy łądząco podobny do wcześniejszego przykładu. Różnica polega jedynie na większej liczbie nieprzewidywalnych warunków zewnętrznych, które należy wziąć pod uwagę. Systemy takie są tak mocno zintegrowane z poszczególnymi maszynami, że ze względów bezpieczeństwa muszą być cały czas rozwijane i na nowo aktualizowane (uwzględniając własności danego pojazdu, czas reakcji, dynamikę itd.).

Z drugiej strony, można przytoczyć odmienny model na przykładzie gier zespołowych. Piłka siatkowa jest wyraźnie grą nieformalną. Ma miejsce w prawdziwym świecie, gdzie nie ma żadnych wyjątkowo zdefiniowanych stanów. Świat piłki siatkowej – prawdziwy świat – jest światem ciągłym. Jako ludzie możemy opracować model gry w piłkę siatkową i model taki może być zmienny i posiadać różne stany, ale sama gra w piłkę siatkową (realna gra) już taką nie jest i stałych stanów gry nie posiada. Brak jakichkolwiek wyjątkowo zdefiniowanych stanów sugeruje również, że dwie sytuacje w prawdziwym świecie nie są nigdy identyczne. Ponadto, w kontraście do światów wirtualnych, dostępna informacja, którą agent może nabyć o prawdziwym świecie, jest zawsze niekompletna. Siatkarz nie może wiedzieć o działaniach wszystkich innych graczy równocześnie, a działania te leżą w prawie nieograniczonym zasięgu możliwości. Nie da się nawet zdefiniować tego, co może znaczyć określenie „kompletna” informacja. Całość może zostać zdefiniowana tylko w zamkniętym, formalnym świecie. Dopóki całość nie jest zdefiniowana, lepiej jest pozostać w obrębie ograniczonej liczby informacji. Siatkarz w realnej grze tylko ogranicza informację na temat ogólnej sytuacji, a właściwie ogranicza te informacje o rzeczywistym świecie, które są możliwe do zdobycia.

Możliwości zbierania informacji zawsze są ograniczone z powodu „wciele-
nia” (ang. *embodiment*), wynikającego z biologicznej struktury uczestnika gry. Pole widzenia jest ograniczone, zasięg zmysłów jest ograniczony, a system ner-
wowy i motoryczny potrzebują odpowiednio dużo czasu, by zadziałać. Ponadto
w prawdziwym świecie istnieje presja czasu i rzeczy zdarzają się, nawet jeśli my
nic nie robimy, a wszystko to dzieje się w czasie rzeczywistym. Jeżeli zatem
chcemy na przykład uniknąć potrącenia przez samochody, musimy zawsze zejść
z linii nadjeżdżającego pojazdu. Jeżeli skaczemy z muru, na nasze ciało działają
prawa fizyki (grawitacja), musimy więc zareagować szybko, aby nie ucierpieć.
Związane są z tym również inne prawa fizyki, na przykład aby wystąpiło poru-
szanie się, musi zaistnieć tarcie, ruch wymaga energii, a wszystkie fizyczne or-
ganizmy posiadają metabolizm, tak więc również potrzebują energii. To wszyst-
ko razem składa się na fizyczne zjawiska, które nie muszą być w żaden sposób
inicjowane, by zaistnieć, wystarczy zatem poprzestać na stwierdzeniu, że po
prostu istnieją (przejawiają się).

W prawdziwym świecie, każde fizyczne urządzenie ulega awariom, posiada
zakłócenia i może powodować hałas. To założenie w zasadzie sprawdza się dla
każdego systemu nerwowego (pojmowanego w sensie biologicznym). Innymi
słowy, informacje zgromadzone za pośrednictwem zmysłów zawsze mogą za-
wierać błędy. W końcu prawdziwy świat jest w sposób nieokreślony bogaty, do
poznania jest zawsze więcej niż mamy tego świadomość. Uściślając, dopóki na-
bywanie informacji zajmuje czas, trzeba ograniczyć samego siebie do wiedzy
o pewnej części prawdziwego świata. Nie zależy to od wrażliwości danego sys-
temu nerwowego czy systemu zmysłów. Biorąc pod uwagę właściwości praw-
dziwego świata i ograniczenia fizycznego agenta jakiegoś typu, prawdziwy świat
jest tylko częściowo poznawalny. To z kolei sugeruje, że prawdziwy świat
przewidywalny jest również tylko w pewnym ograniczonym zakresie.

Na koniec powyższego porównania dwóch przykładowych światów istotne
wydaje się, by zwrócić uwagę na konsekwencje przyjętej i wprowadzonej termi-
nologii „prawdziwy” i „wirtualny”. Użyto zatem pojęcia „wirtualny”, by ściśle
określić formalne światy takiego typu, jak przytoczona gra w szachy. Określenie
„wirtualny świat” (ang. *virtual world*) lub „udawany świat” (ang. *simulated
world*) często używane jest w znaczeniu sztucznego życia¹⁴ i „wirtualnej rzeczy-
wistości” (ang. *virtual reality*)¹⁵, gdzie przykładem są gry wideo. Innym przy-
kładem są sztuczne fizyczne światy Karla Simsa, w których sztuczne stworzenia
ewoluują w różnych warunkach, np. na ziemi lub w wodzie. W tych światach
można zdefiniować nowe fizyczne prawa i nowe prawa natury – to jeden z licz-

¹⁴ Por. Ch. Langton, *Artificial Life – An Overview*, Epstein, Axtell, [w:] *Growing Artificial Socie-
ties, Social Science from the Bottom Up Random*, Brookings Institution Press, MA 1996,
s. 372–377.

¹⁵ Por. R. Kalawsky, *The Science of Virtual Reality and Virtual Environments*, Addison-Wesley
Ltd, UK 1993, s. 328–342.

nych powodów, który czyni je tak fascynującymi. Na przykład jeśli w wirtualnym świecie grawitacja jest symulowana, to można tak dostosować stałą grawitacji (G), aby obserwować zmiany w zachowaniu udawanych (symulowanych) organizmów, które zamieszkują ten świat¹⁶. Z perspektywy agentów, którzy żyją w takim wirtualnym świecie, posiada on kilka cech, które staramy się wskazać w prawdziwych światach. Przykładowo z punktu widzenia agenta nieoczekiwane i nowe sytuacje zdarzają się. Często pojawiają się nowi przeciwnicy i wrogowie, którzy posiadają nieznaną moc. Jakkolwiek z punktu widzenia programisty, który utworzył wirtualny świat, te same wydarzenia nie są ani nowe, ani nieoczekiwane, gdyż on sam je zaprojektował i umieścił w systemie¹⁷.

Z powyższych rozważań i przytoczonych przykładów wyłania się następująca wniosek: prawdziwe światy znacząco różnią się od wirtualnych. Problemy klasycznego AI i nauk kognitywnych mają swoje źródło głównie w zaniedbaniu tych różnic.

Kilka dobrze znanych problemów z klasycznymi systemami

Aby jeszcze inaczej zrozumieć istotę problemu sztucznej inteligencji, zasadne wydaje się przedstawienie kilku spośród bardziej znanych problemów związanych z klasycznymi systemami AI. W dotychczasowej dyskusji używamy pojęcia „klasycznych systemów AI” na określenie czysto symbolicznych systemów, takich jak systemy ekspertowe lub tradycyjne systemy planowania, np. STRIPS (skrót od **S**tanford **R**esearch **I**nstitute **P**roblem **S**olver). Głównym zatem celem wydaje się opisanie zagadnień i problemów, które historycznie motywowały naukowców do szukania alternatyw. Wydaje się, że w sporej części społeczności naukowców zajmujących się AI panuje zgodne przekonanie o tym, że systemom klasycznym, brakuje niezawodności, trwałości i takich możliwości uogólnienia, których nie da się rozwijać w czasie rzeczywistym. Sprawia to, że słabo nadają się one do zastosowania w prawdziwym świecie. Ponadto są one w istocie sekwencyjne, to znaczy że wykonują jedną operację po drugiej. Są uruchamiane na kolejnych maszynach, podczas gdy ludzki mózg jest w swym działaniu równoległy. Należy zatem bliżej przyjrzeć się i choćby krótko zbadać każdy z tych punktów.

Niezawodność, Trwałość i Uogólnianie: tradycyjne systemy AI często nie są niezawodne, co oznacza, że brakuje im tolerancji na hałas, odporności na uszkodzenia i umiejętności działania w nowych sytuacjach i okolicznościach.

¹⁶ Por. K. Sims, *Evolving 3D Morphology and Behavior by Competition*, *Artificial Life IV*, R. Brooks and Pattie Maes, Eds., MIT Press, Cambridge, MA 1994. DOI: <http://dx.doi.org/10.1162/artl.1994.1.353>.

¹⁷ Por. R. Pfeifer, Ch. Scheier. *Understanding Intelligence*, MIT Press. Cambridge, MA, London, England 1999, s. 234 i nn.

Każdy tworzony system o konkretnej specyfice posiada tolerancję hałasu. Działa on właściwie wówczas, kiedy dane zawierają hałas, a system rejestruje je jako przypadkowe wahania się w danych. Czujniki są zawsze hałaśliwe, ponieważ są fizycznymi urządzeniami, a działania silnika są zawsze nieprecyzyjne, jest to właściwe dla fizycznych urządzeń. System musi być odporny na uszkodzenia, jeżeli nadal ma funkcjonować prawidłowo, podczas gdy niektóre z jego komponentów zostały uszkodzone.

Standardowe modele przetwarzania danych (symboli) nie są ani odporne na hałas, ani na uszkodzenia, chyba że ich programowanie ściśle przewiduje hałas i poszczególne rodzaje błędów. Najważniejszym niedoborem tradycyjnych systemów AI pod względem odporności jest ich niezdolność do właściwego funkcjonowania w nowych sytuacjach, brak im zdolności uogólniania. Jeśli ma miejsce sytuacja, która nie została uprzednio zdefiniowana i uwzględniona w programowaniu, tradycyjny system psuje się lub przestaje działać. W prawdziwym świecie zdolność uogólniania jest szczególnie ważna, nigdy bowiem nie zdarzają się dwie identyczne sytuacje.

Przetwarzanie w czasie rzeczywistym: ponieważ prawdziwy świat ma własną dynamikę, systemy muszą być w stanie zareagować szybko, aby przetrwać i poprawnie wykonać swoje zadania. Systemy bazowały zatem na klasycznym paradygmacie osadzonym w prawdziwych robotach, które są zazwyczaj powolne, ponieważ przetwarzają informacje centralnie.

Z kolei istnieje ugruntowany pogląd mówiący o tym, że ludzki mózg jest „siedzibą inteligencji”. Jeżeli wszystkie sygnały sensoryczne muszą być przetransmitowane do urządzenia centralnego w celu przetworzenia (zintegrowane z innymi sygnałami czuciowymi, mapowane na wewnętrznym obrazie sytuacji planującej sekwencję kolejnych działań) i w końcu wygenerują sygnał dla silnika w czasie rzeczywistym, odpowiedź może ledwie zostać osiągnięta bądź w ogóle nie da się jej osiągnąć.

Sekwencyjny charakter programów: architektura dzisiejszych programów AI jest zasadniczo sekwencyjna i działają one na zasadzie „krok po kroku”. Przeciwnie, przetwarzanie ludzkiego mózgu jest masywnie równoległe z działalnością występującą w wielu częściach mózgu o każdym czasie. Ten problem wynika z faktu, że aktualne technologie komputerowe są głównie oparte na architekturach typu Von Neumann, które są na poziomie przetwarzania informacji maszynami sekwencyjnymi. Należy jednak zauważyć, że na fizycznym poziomie maszyna Von Neumann jest masywnie równoległa, tak jak każdy inny system w naturze.

Inne problemy: dodatkowymi zarzutami pod adresem klasycznych systemów jest ich zorientowanie na cel, hierarchiczne zorganizowanie oraz centralne przetwarzanie informacji. Krytyczne oceny modeli AI, sformułowane do tej pory, są już historycznie ugruntowane i dobrze znane. Od połowy lat 80. ich liczba nieustannie wzrasta, ciągle jednak dotyczą podstawowych kwestii. Dowiedzione

zostało zwłaszcza to, iż tradycyjne systemy AI kuleją z powodu „problemu ramy” (*frame problem*) oraz tzw. „uziemiaenia”, brakuje im właściwości „wcielenia” i „umieszczenia”. Ma to związek z tym, że maszyny nie mają na razie szans na samoświadomość, tzn. nie zdają sobie faktu z własnego istnienia, a co za tym idzie – nie potrafią prawidłowo odnieść własnej pozycji w otoczeniu. To zagadnienie jest zbyt obszerne, aby je w tym miejscu spróbować nawet poprawnie określić.

Właściwym wydaje się zatem szersze omówienie zagadnień AI związanych ze wspomnianym „problemem ramy”, dotyczącym klasycznego rozumienia tzw. „wcielenia”, „umieszczenia”, „uziemiaenia” oraz tematów z nimi ściśle związanych. Dla lepszej orientacji w tematyce omawianego zagadnienia posłużymy się kilkoma dość obszernymi i interesującymi przykładami, które ze względu na swój charakter i objętość umieścimy w części drugiej.

Basic Problems of the Classical Concept of Artificial Intelligence and Cognitive Sciences – Part 1

Summary

The article introduces us to the interesting and complicated world of artificial intelligence and explains basic ideas which the virtual worlds are ruled by. There are a few main problems in the creation of artificial systems that have been explained. Finding solutions to them is still a barrier on the road to building the intelligence as brilliant as human intelligence.

Keywords: artificial intelligence, cognitive science, intelligent agent, frame problem, embodiment problem, virtual worlds.