

## Wstęp

Dobra, staromodna sztuczna inteligencja (Good Old-Fashioned AI) - w skrócie GOFAI - to etykieta używana do określenia klasycznej, symbolicznej sztucznej inteligencji. Termin „AI” jest czasami używany w znaczeniu wyłącznie GOFAI, ale jest to błąd. Sztuczna inteligencja obejmuje także inne podejścia, takie jak koneksjonizm (których jest kilka odmian: patrz rozdział 5), programowanie ewolucyjne oraz robotyka usytuowana i ewolucyjna. Rzeczywiście, większość pracy w sztucznym życiu (A-Life) mieści się w szeroko rozumianej sztucznej inteligencji, pomimo tendencji A-Lifers do dystansowania się od niej. Tutaj jednak zajmujemy się wyłącznie symboliczną sztuczną inteligencją. Zarówno technologiczna, jak i psychologiczna sztuczna inteligencja wykorzystuje pełen zakres metodologii sztucznej inteligencji, w tym GOFAI. Kierują się jednak różnymi motywacjami. Celem tego pierwszego jest budowanie użytecznych systemów komputerowych, wykonujących lub pomagających w zadaniach, które ludzie chcą wykonać. Celem tej ostatniej – którą można nazwać także psychologią obliczeniową – jest rozwój wyjaśniających teorii umysłu. Czasami (według „silnej” sztucznej inteligencji) ma ona również na celu zbudowanie systemów komputerowych, które same w sobie są naprawdę inteligentne. W związku z tym psychologiczna sztuczna inteligencja z większym prawdopodobieństwem podnosi pytania interesujące z filozofii umysłu

## Opisane GOFAI

Metodologia GOFAI wykorzystuje zaprogramowane instrukcje działające na formalnych reprezentacjach symbolicznych. Dobrze pasuje do binarnej, szeregowej natury komputera cyfrowego von Neumanna. Od połowy lat pięćdziesiątych do połowy osiemdziesiątych XX wieku było to dominujące (choć nie jedyne) podejście w sztucznej inteligencji. Funkcjonalizm w filozofii umysłu został opracowany z myślą o tych programach i/lub abstrakcyjnych maszynach Turinga. Symbol GOFAI jest elementem w języku formalnym (języku programowania). Podobnie jak symbole matematyczne czy logiczne, symbole GOFAI – i złożone z nich programy – można uznać za struktury czysto formalne (bez znaczenia). W praktyce jednak są one zwykle interpretowane przez użytkownika w kategoriach określonej treści semantycznej: werbalnej, numerycznej, wizualnej, słuchowej i tak dalej. Symbole atomowe można łączyć, tworząc symbole złożone, zgodnie z określonymi formalnymi zasadami manipulacji symbolami. To, co dzieje się po uruchomieniu programu GOFAI, polega na tym, że różnego rodzaju symbole, o różnej strukturze, są budowane, przechowywane, odzyskiwane, porównywane i przekształcane. Krótko mówiąc, obliczenia GOFAI obejmują konstrukcję i transformację symbolicznych struktur danych. Kluczowymi koncepcjami tworzącymi programy GOFAI są wyszukiwanie heurystyczne i planowanie. Te ściśle powiązane idee zostały zapoczątkowane w latach pięćdziesiątych XX wieku. Są one nadal kluczowe dla dzisiejszej pracy GOFAI. Problem lub zadanie GOFAI jest reprezentowane jako przestrzeń poszukiwań: zbiór możliwości (określony przez skończony zbiór reguł generatywnych), w którym leży rozwiązanie – i w którym należy je znaleźć. Przykłady obejmują zbiór legalnych ruchów w szachach lub dopuszczalne ciągi słów przy określonej gramatyce i słownictwie. Konkretnie lokalizacje w przestrzeni poszukiwań nie muszą być wcześniej szczegółowo określone, ale mogą zostać wygenerowane przez program. W wielu przypadkach systematyczne przeszukiwanie przestrzeni metodą „brutalnej siły” jest niepraktyczne, ponieważ istnieje zbyt wiele możliwości, aby można je było rozpatrywać indywidualnie. Co więcej, często nie ma całkowicie niezawodnej zasady wybierania wszystkich (i tylko) „obiecujących” kandydatów. Jeśli tak, wyszukiwanie musi opierać się na heurystyce: praktycznych zasadach, które zwykle są pomocne w prowadzeniu programu w kierunku rozwiązania, a nie w ślepych zaułkach, ale które nie gwarantują rozwiązania problemu. (Porównaj „Chroń swojego hetmana” w szachach: jest to bardzo przydatna heurystyka, ale czasami wskazane jest poświęcenie swojego hetmana.) Planowanie to technika GOFAI, w której problem jest analizowany przez program jako hierarchiczna struktura celów, celów cząstkowych, i podcele. W przypadku paradygmatu na

początku zostaje określony cel końcowy, a zadaniem programu jest zmniejszanie różnic pomiędzy stanem bieżącym a stanem docelowym, aż różnice te znikną. (Ogólna idea, że celowe zachowanie jest zakorzenione w redukcji różnic, została odziedziczona z cybernetyki, badania idcentury nad systemami sprzężenia zwrotnego, takimi jak homeostaza biologiczna i pociski kierowane; ale cybernetyka nie określiła strukturalnych reprezentacji stanów docelowych) Program jest dostarczany przed rozpoczęciem z szeregiem możliwych różnic, listą akcji (operatorów), które mogą wyeliminować różne różnice, listą warunków wstępnych, które muszą być spełnione, jeśli określony operator ma zostać użyty, oraz heurystyką dla uporządkowanie działań, gdy w danej sytuacji możliwe jest więcej niż jedno działanie. Rzeczywiście, większość zaangażowanej „inteligencji” leży w wyborach działań, operatorów i heurystyk określonych przez programistę. Jeśli program po uruchomieniu uzna, że do osiągnięcia bieżącego celu potrzebny jest określony operator, może być konieczne utworzenie nowego celu częściowego, aby spełnić odpowiednie warunki wstępne. Proces ten można powtarzać na nieskończenie wielu poziomach hierarchicznych. W wielu przypadkach bardzo szczegółowe planowanie można pozostawić do momentu wykonania, tak aby – do pewnego momentu – można było uporać się z nieprzewidywanymi czynnikami. W najwcześniejszych badaniach GOFAI hierarchia celów była jasno wyrażona i „z góry określona”. Innymi słowy, program byłby w danym momencie skupiony na jednym celu i przełączał się na inny tylko wtedy, gdyby (1) aby go osiągnąć, trzeba było ustalić nowy cel częściowy lub (2) cel pierwotny został osiągnięty – w takim przypadku program „wyskoczy” do następnego poziomu docelowego lub zakończy się, jeśli osiągnięty zostanie najwyższy poziom. Późniejsza metodologia sztucznej inteligencji uczyniła hierarchię celów ukrytą w dużym zestawie logicznie niezależnych reguł warunku-działania. (Poszczególne reguły nazywano „produkcjami”, więc zbudowane z nich programy nazywano systemami produkcyjnymi.) Warunek i Akcja w regule mogły być pojedynczymi elementami lub połączeniem (lub rozłączeniem) kilku, a nawet wielu – różne przedmioty. Na przykład jedna z reguł w programie do rozwiązywania alfabetycznie zakodowanych problemów z dodawaniem, taka jak „DONALD + GERALD = ROBERT”, brzmi następująco: „JEŚLI bieżącym celem jest obliczenie zbioru liter, WTEDY znajdź jedną z liter (za pomocą reguły produkcji zwanej FINDLETTER) ORAZ ustaliliśmy (nowy) cel, jakim jest jej ocena. To, co wcześniej było przedstawiane jako oddzielne listy heurystyk, działań i warunków wstępnych, teraz zostało milcząco uwzględnione w produkcjach. Zatem Warunek może zawierać informację, że taki a taki cel został już ustalony. Podobnie Akcja może obejmować element wyznaczający nowy cel lub cel częściowy. Zapewniono techniki rozwiązywania konfliktów, aby poradzić sobie z przypadkami, w których więcej niż jeden Warunek został spełniony w tym samym czasie. Na przykład warunek określający dłuższą, bardziej inkluzywną koniunkcję miałby pierwszeństwo przed warunkiem, który określał tylko jedną koniunkcję. Zaletą tego ogólnego podejścia było to, że w dowolnym momencie można było dodać nową regułę (pod warunkiem, że była zgodna z istniejącym zestawem reguł), bez konieczności przepisywania złożonej procedury hierarchicznej. Ponadto metodologia ta umożliwiła przerwanie: natychmiastową zmianę celu, spowodowaną nagłymi zmianami w środowisku lub w innym miejscu działania programu. W ten sposób wyszukiwanie GOFAI stało się mniej elastyczne i „jednomyślne” niż na początku. Rozwój systemów eksperckich – na przykład do diagnostyki medycznej lub doradztwa w zakresie poszukiwań geologicznych – w dużym stopniu zależał od systemów produkcyjnych. Ogólnie rzecz biorąc, system ekspertowy to program, który reprezentuje wiedzę ludzkiego eksperta w postaci zbioru reguł JEŻELI-TO i który może być wykorzystany do udzielania porad osobom niebędącym ekspertami w danej dziedzinie. Na przykład zasada diagnozy medycznej mogłaby brzmieć: „JEŚLI te objawy zostaną zaobserwowane u pacjenta ORAZ pacjent jest dorosłym mężczyzną, WTEDY wywnioskowaj, że za chorobę odpowiedzialna jest ta bakteria ORAZ zaleć ten lek, podawany w takiej dawce przez taki okres czasu”. czas jako leczenie.” Badacze zajmujący się systemami ekspertowymi wprowadzili techniki wnioskowania dotyczące „łączenia w przód” i „łańcucha wstecz” w systemach produkcyjnych. Pierwsza metoda jest stosowana metodą oddolną, w celu wyciągnięcia wniosków z

danych. To drugie jest używane od góry do dołu, aby znaleźć dowody na przypuszczenie i/lub wyjaśnić użytkownikowi radę już udzieloną przez program poprzez podsumowanie wcześniej wydanych reguł. Ogólnie rzecz biorąc, przetwarzanie „oddolne” rozpoczyna się od zbioru odmiennych elementów danych i próbuje w jakiś sposób nadać im ogólny sens – być może za pomocą technik GOFAL (takich jak systemy produkcyjne) lub być może poprzez koneksjonizm PDP (równoległe przetwarzanie rozproszone) lub inne formy samoorganizacji. Z kolei przetwarzanie „od góry do dołu” rozpoczyna się od jakiegoś celu lub struktury wyższego poziomu i wykorzystuje je do interpretacji – a czasem do zmiany – danych wejściowych na niższych poziomach. Na przykład program wizualny GOFAL wyszukujący kostki może zdecydować się na „halucynację” brakującego rogu wejścia, pokazując niekompletny rysunek sześciianu. Jak mogą sugerować pojęcia wyszukiwania, planowania i heurystyki, programy GOFAL często symulują świadome rozważania ludzkiego myślenia na wysokim poziomie. Dzieje się tak częściowo dlatego, że mogą reprezentować twierdzenia o określonej treści semantycznej, a częściowo dlatego, że mogą obejmować odgórne procesy kontroli wykonawczej i samokontroli. Nawet jeśli modelują procesy nieświadome, zazwyczaj przedstawiają je jako podobne do świadomych rozważań. Na przykład rozpoznawanie wizualne w GOFAL jest często uważane za analogiczne do rozwiązywania problemów w logice. A logiczne rozwiązywanie problemów, a fortiori, jest postrzegane jako logika od samego początku. Jednakże rodzaj stosowanej logiki jest różny. Może to być rachunek zdań, wyrażony przez proste bramki logiczne realizujące podstawowe operacje logiczne lub przez planistów hierarchicznych, takich jak ci naszkicowani powyżej. Częściej jest to logika predykatowa. W takim przypadku można ją uzupełnić różnymi logikami modalnymi lub specjalnymi metodami mającymi na celu zwalczanie eksplozji kombinatorycznej – takimi jak dowodzenie twierdzeń o rozdzielczości (Robinson 1968). Lub może to być system produkcyjny, którego pary Warunek-Działanie opierają się na logice reguł JEŻELI-TO (po 1943 r.). Korzenie GOFAL w logice zapewniają mu zarówno mocne, jak i słabe strony. Kolejną pododmianą GOFAL jest programowanie ewolucyjne (Holland 1975; Ray 1992). Dokładniej, programowanie ewolucyjne zostało pierwotnie opracowane w ramach GOFAL i jest często używane w kontekście GOFAL – chociaż można go również wykorzystać do rozwijających się sieci koneksjonistycznych. Tutaj tak zwane „algorytmy genetyczne” dokonują losowych zmian we własnych zasadach programu, inspirowane procesami mutacji punktowych i krzyżowania w biologii. (W przypadku mutacji punktowej zmieniany jest pojedynczy element ciągu symboli kodującego regułę, tak że ABCDE może dać ABFDE; w przypadku krzyżowania zestaw sąsiadujących jednostek w jednym ciągu reguł – porównywalny z sąsiednimi genami na chromosomie – zostaje zamieniony dla zestawu sąsiednich jednostek w innym ciągu, tak że ABCDEFGH i PQRSTUUV dają ABCSTUFGH i PQRDEFUUV.) W każdym „pokoleniu” jeden lub dwa nowo wyklute programy, które odniosły większy sukces, wybierane są na rodziców (bezpłciowych lub seksualnych) dla następnego pokolenia. Zwykle wybór odbywa się automatycznie za pomocą predefiniowanej funkcji fitness; ale można to zrobić interaktywnie, przez człowieka. Tam, gdzie problem jest wystarczająco dobrze poznany, aby automatycznie rozpoznać niewielką poprawę wydajności, metodologia ta umożliwi optymalizację danego zadania. Na przykład, jeśli ktoś potrzebuje programu, który posortuje listę liczb w porządku rosnącym lub ułoży nazwy w porządku alfabetycznym, wówczas podejście ewolucyjne będzie w stanie znaleźć najbardziej wydajny algorytm, podczas gdy programista-człowiek nie będzie w stanie tego zrobić.

### **Mocne i słabe strony GOFAL**

Głównymi atutami GOFAL są jego możliwości modelowania hierarchii i porządku sekwencyjnego, umożliwiające precyzję rozwiązywania problemów oraz reprezentowanie określonej treści zdań. Każda z tych cech była zilustrowana w pionierskim programie SHRDLU (Winograd 1972), który przykuł uwagę wielu psychologów i filozofów nieobeznanych wcześniej z AI. W przeciwieństwie do poprzedzających go prostych programów dopasowujących słowa kluczowe, SHRDLU potrafił analizować zdania o

znacznej złożoności, takie jak „Ile jajek użyłbyś do ciasta, gdybyś nie dowiedział się, że przepis twojej matki jest błędny?” Angażując się w „rozmowę”, parser SHRDLU był ściśle zintegrowany z analizą semantyczną, rozwiązywaniem problemów i wiedzą o świecie. Na przykład, aby określić, do czego odnosi się wyrażenie mówiącego przez człowieka „piramida”, należy zadać sobie pytanie, czy w scenie występuje więcej niż jedna piramida, a jeśli tak, czy w rozmowie wspomniano już o jakiejś konkretnej piramidzie. Ogólnie rzecz biorąc, SHRDLU – podobnie jak dzisiejsze programy GOFAL – może naprzemiennie wykonywać różne zadania i/lub bazy wiedzy, aby rozwiązać pozornie „jeden” problem. Pytanie „Ile jajek...?” przykład ilustruje znaczenie hierarchii i porządku sekwencyjnego: Analiza składni po prostu nie mogłaby zostać przeprowadzona bez prawidłowego ich wykonania. Precyzja też jest kluczowa – bo pojedyncza litera może czasami mieć ogromne znaczenie. (Porównaj telegram o treści „Nasz syn nie żyje” z telegramem o treści „Twój syn nie żyje”). Jednak z precyzją można przesadzić. Wczesne programy GOFAL były notorycznie kruche w tym sensie, że brakujące i/lub sprzeczne dane skutkowałyby bezsensowną reakcją komputera – jeśli nie zatrzymałby się całkowicie. (Nawet SHRDLU był mniej potężny, niż się wydawało, ponieważ jego „rozmowa” nie była ciągłą wymianą, ale została połączona z wieloma różnymi okazjami.) Szczególnym przypadkiem kruchości jest problem z ramą. Termin ten jest powszechnie używany do określenia dwóch trudności: po pierwsze, wiedzy, które aspekty sytuacji zmieniają się w wyniku określonego działania, a które nie; po drugie, rozumowanie z niepełną wiedzą, wynikającą z naszej nieuniknionej niewiedzy na temat faktów rzeczywistego świata i niejasności pojęć w języku potocznym. Na przykład zagadkę, w jaki sposób osieroceni odkrywcy mogą przepłynąć przez rzekę w dżungli, można rozwiązać nie (jak można było przewidzieć) przez mozolne budowanie tratwy, ale przez nieoczekiwane uratowanie ich przez helikopter pracującej w pobliżu ekipy filmowej. Roboty GOFAL próbowali rozróżnić stany rzeczy zmienne i niezmiennie w odniesieniu do planowanych działań, ale często pomijano pewne istotne przypadki. A logika dowodzenia twierdzenia o rozdzielczości – która dowodziła twierdzeń w logice predykatów, pokazując, że ich negacje są niespójne, przy założeniu, że nie-nie-X implikuje X – była tak bezlitosna, że groziła podważeniem każdego rzeczywistego planu lub „zdrowego rozsądku” – rozwiązywacz problemów. Opracowano różne typy logiki „niemonotonicznej”, w której zdanie można uważać za prawdziwe, dopóki nie zostanie wykazane, że jest fałszywe. Mimo to te „domyślne założenia” musiały zostać określone wcześniej. (Wymagania związane z problemem ramy zachęciły później do rozwoju robotyki usytuowanej. Roboty GOFAL zaplanowały wcześniej swoje działania, kierując się od góry do dołu złożonym modelem świata, który wymagał mozolnej aktualizacji, gdyby świat nagle się zmienił, oraz założeniami o tym, co by się stało, gdyby zostało podjęte to czy tamto działanie. Natomiast roboty usytuowane nie opierały się na modelach świata wewnętrznego. Zamiast tego działały oddolnie, reagując bezpośrednio, w prosty „odruchowy” sposób na określone sygnały środowiskowe, gdy tylko napotkały je; gdyby świat nagle się zmienił, nowe, istotne reakcje zostałyby wywołane jako coś oczywistego.) Niektórzy krytycy GOFAL zdają się wierzyć, że nieustanna kruchość jest dla niego niezbędną, ale tak nie jest. Istnieje kilka sposobów, dzięki którym można do pewnego stopnia uniemożliwić programowi GOFAL wyciąganie nieodwracalnych wniosków, które mogą okazać się fałszywe. Aby zmniejszyć kruchość systemów GOFAL, opracowano specjalne środki zapobiegawcze, których szczególnym przypadkiem są wspomniane powyżej założenia domyślne, oraz różnorodne współdziałające tymczasowe (tj. tymczasowe) reprezentacje. Jak widzieliśmy, zidentyfikowano techniki rozwiązywania konfliktów, które umożliwiają programowi wybór pomiędzy dwoma (lub większą liczbą) indywidualnie dozwolonych działań. Prawdą jest jednak, że symboliczne programy AI są zazwyczaj kruche. Nie pokazują one „naturalnie” tak zwanej łagodnej degradacji, w której niedoskonałości danych prowadzą do proporcjonalnie niedoskonałych, ale często akceptowalnych wyników. Główną zaletą systemów koneksjonistycznych PDP jest to, że mają tę właściwość po prostu ze względu na sposób działania: nie musi być ona specjalnie wbudowana. Jasność ma również swoje wady. Jeśli ktoś chce, aby program GOFAL czegoś się nauczył, musi zapewnić cechy definiujące to „coś”. Często nie jest to możliwe. Możesz

na przykład nauczyć program rozpoznawania kotów. Ale czy potrafisz zdefiniować „kota”? Czy potrafisz w ogóle zidentyfikować i wymienić istotne cechy widoczne na zdjęciach tuzina kotów? Z pewnością programy GOFAI mogą czasem nauczyć się nowych koncepcji: na przykład skuteczniejszej metody diagnozowania określonej choroby. Jednak cechy definiujące te pojęcia muszą zostać wcześniej wyraźnie określone, nawet jeśli są one jedynie ujęte w dłuższej liście cech. Nie jest tak (choć często się uważa), że programy GOFAI zawsze działają odgórnie, a nie oddolnie. Często jednak angażują w nie jakiegoś dyrektora centralnego lub kontrolę wysokiego szczebla. Jest to mniej oczywiste w systemach produkcyjnych, gdzie hierarchia sterowania nie jest wyraźnie określona i gdzie przerwania mogą przenieść uwagę systemu na inny obszar. Jednakże umiejętność modelowania hierarchii w zasadzie często prowadzi w praktyce do modelowania zjawisk hierarchicznych – takich jak planowanie złożone czy język naturalny. Z biegiem lat stawało się coraz bardziej jasne, że w niektóre formy inteligencji (zarówno w obrębie jednostek, jak i grup) nie angażuje się żadna władza centralna. Raczej inteligencja i kontrola są rozproszone po całym systemie. W takich przypadkach tradycyjne podejście GOFAI należy zmodyfikować lub nawet całkowicie porzucić. Modyfikacja ma miejsce na przykład wtedy, gdy programuje się wiele w dużej mierze autonomicznych systemów („agentów”), których zbiorowe zachowanie wynika z wielu interakcji, a nie z działania jakiegoś menedżera wysokiego szczebla. Agenci mogą być wirtualnymi „softbotami” lub robotami fizycznymi, a ich interakcje mogą obejmować komunikację, negocjacje, współpracę, licytację, a nawet targowanie się. Porzucenie ma miejsce na przykład wtedy, gdy zamiast GOFAI używany jest koneksjonizm PDP. Systemy PDP opisano w dalszej części. W tym przypadku istotne jest to, że – poza unikaniem narzucania odgórnego centralnego kontrolera – metodologia PDP pozwala uniknąć pewnych innych słabości GOFAI. Przede wszystkim PDP unika kruchości, która zagraża podejściu symbolicznemu. Zamiast wyciągać jasne wnioski za pomocą logicznych środków, program PDP osiąga stan równowagi, w którym jednocześnie spełniona jest większość (potencjalnie sprzecznych) ograniczeń. Można zatem do pewnego stopnia dopuścić zanieczyszczone i/lub częściowo brakujące dane. Innymi słowy, tolerancja na szum i kompletność wzorców, które są problematyczne dla GOFAI, wynikają „naturalnie” z projektu sieci PDP. Na przykład „niekompletny” lub niekompletny rysunek/zdjęcie sześcianu zostanie rozpoznany przez system PDP jako przedstawiający sześcian, nawet jeśli programista nie spodziewał się bałaganu i/lub niekompletności, a tym bardziej przewidywał to w mniejszym stopniu celowo. Podobnie wzorców (koncepcji), które nie są dostępne dla osób uczących się GOFAI i których ludzie nie podejrzewali, którzy zbudowali system, można nauczyć się na przykładzie. (Dlatego PDP jest używany w systemach eksperckich, które identyfikują subtelne wzorce w informacjach finansowych, takie jak ruchy na giełdzie.) Jednakże PDP nie jest panaceum. Brakuje mu kluczowych atutów GOFAI w zakresie modelowania wielopoziomowej hierarchii, porządku sekwencyjnego i relacji wnioskowanych między określonymi treściami zdań. Co więcej, nie nadaje się dobrze do zadań, w których wymagana jest absolutna precyzja, a nie pełna wdzięku degradacja. Rzeczywiście, większość badań koneksjonistycznych prowadzonych w ciągu ostatnich dwudziestu lat miała na celu zdobycie właśnie tych mocnych stron – jak dotąd z jedynie ograniczonym sukcesem. Niektórzy badacze stosują zatem systemy „hybrydowe”, w których GOFAI łączy się z koneksjonizmem PDP. Przykładem takiego podejścia jest model działania i błędów działania opracowany przez Donalda Normana i Timothy’ego Shallice’a, odpowiednio psychologa poznawczego i neurologa klinicznego, oraz współpracujących z nimi badaczy sztucznej inteligencji. W skrócie, ten „System Nadzorującej Uwagi” wykorzystuje koneksjonizm do modelowania rozpoznawania percepcyjnego i pamięci/przypominania skojarzeniowego, a GOFAI do modelowania świadomego wyboru i planowania. Obydwa te aspekty mają kluczowe znaczenie dla (w dużej mierze) działań człowieka, a także – według niektórych wiodących ekspertów w dziedzinie sztucznej inteligencji – także dla skutecznej robotyki (Sahota i Mackworth 1994). Ponieważ oba aspekty są kluczowe, wynika z tego, że robotyka usytuowana nie zawsze może zastąpić GOFAI. Nadzieja wyrażona przez Rodneya Brooksa (1991), że sztuczną inteligencję można w pełni zlokalizować i porzucić

GOFAI, jest iluzoryczna. Prawdą jest, że „bezpośrednia” reakcja na otoczenie może zapewnić zaskakujący stopień porządku w zachowaniu. Tę kwestię poczynił w latach sześćdziesiątych jeden z arcykapitanów GOFAI w odniesieniu do ścieżki mrówki kroczącej po zagraconym terenie. Poruszanie się owadów może podlegać modelowaniu czysto sytuacjonistycznemu, opierając się wyłącznie na wbudowanych reakcjach odruchowych na sygnały środowiskowe. Jednak gdy w grę wchodzi zachowanie na poziomie ludzkim, czasami niezbędne jest planowanie i wewnętrzna reprezentacja. Prawdą jest również, że „wcielenie” fizycznego robota pokonującego świat materialny oferuje problemy i rozwiązania wykraczające poza te typowe dla „bezcieleśnego” GOFAI. Robot jest bowiem rzeczą materialną poruszającą się w świecie rzeczywistym, a nie programem jedynie reprezentującym ten świat, zatem napotyka fizyczne przeszkody – ale i możliwości – które w przypadku zaprogramowanej symulacji nie powstają. Dlatego zaproponowano zastąpienie GOFAI przez GOFAIR: GOFAI i Robotics. GOFAIR jest podwójnie hybrydowy. Z jednej strony duża część reakcji robota jest umiejscowiona; z drugiej strony metody koneksjonistyczne łączy się z GOFAI. Ważne jest jednak to, aby GOFAI – choć uzupełniony – został utrzymany. Nie jest to zaskakujące, ponieważ mocne i słabe strony GOFAI i PDP, a także klasycznego GOFAI i robotyki usytuowanej w dużej mierze się uzupełniają. Motto współczesnej sztucznej inteligencji, zarówno psychologicznej, jak i technologicznej, mogłoby brzmieć: „Niech zakwitnie sto kwiatów – a w każdym razie cztery”. Te cztery to podejścia GOFAI, koneksjonizm, zlokalizowana sztuczna inteligencja i systemy dynamiczne (z których wszystkie czasami wykorzystują obliczenia ewolucyjne).

### **GOFAI jako filozofia**

Kiedy filozof John Haugeland wymyślił etykietę GOFAI, jego słowo „staromodny” sugerowało, że ten typ sztucznej inteligencji został zastąpiony. I w pewnym sensie tak było: w przypadku PDP koneksjonizm narastał i wkrótce zachwycił filozofów (Clark 1989) – a także dziennikarzy. Jednak Haugeland najwyraźniej nic o tym nie wiedział, gdyż w jego książce nawet nie wspomniano o koneksjonalizmie. Jego krytyka GOFAI jako *passé* opierała się raczej na filozofii fenomenologii, podobnie jak dwadzieścia lat wcześniej Hubert Dreyfus (1965). Haugeland argumentował, że dwadzieścia lat krytyki filozoficznej pokazało, że GOFAI jest fałszywy. „Fałszywe”, a nie tylko niepraktyczne lub ograniczone. Jego definicja nie ograniczała się bowiem do wyboru metodologii sztucznej inteligencji (obliczeń symbolicznych). Określała także zaangażowanie w to, co John Searle (1980) nazwał silną sztuczną inteligencją: pogląd, że odpowiednio zaprogramowany komputer będzie naprawdę inteligentny – a co za tym idzie, naprawdę emocjonalny i naprawdę świadomy. Haugeland uważał, że silna sztuczna inteligencja została obalona na różne sposoby zarówno przez Searle’a, jak i Dreyfusa (a także przez jego własne argumenty). Każdy, kto nadal to akceptował, był nie tylko w błędzie, ale także spóźniony (tj. staromodny). Nie jest jasne, czy wszyscy badacze GOFAI są zaangażowani w silną sztuczną inteligencję. (Według definicji Haugelanda nie wszyscy badacze GOFAI są badaczami GOFAI.) Ale niektórzy z pewnością nimi są. Na przykład Allen Newell i Herbert Simon wyraźnie określili formę silnej sztucznej inteligencji w swojej teorii systemów symboli fizycznych, czyli PSS (Newell i Simon 1963; Newell 1980). Twierdzili, że intencjonalność (tj. znaczenie lub „istość”) osiąga się zarówno w umysłach, jak i komputerach, poprzez wdrożenie pewnych typów formalnych obliczeń w PSS. Postrzegano je jako „niezbędne i wystarczające środki ogólnego inteligentnego działania”. Innymi słowy, umysł (lub umysł/mózg) jest PSS. Ich zaangażowanie w silną sztuczną inteligencję było krystalicznie jasne. Programy sztucznej inteligencji Newella i Simona określały formalne systemy obliczeniowe w rodzaju zdefiniowanym przez Alana Turinga w 1936 roku. Jako psychologowie byli jednak zainteresowani obliczeniami, które reagują na świat i są w stanie kierować w nim zachowaniem. Zatem w przeciwieństwie do samego Turinga zdefiniowali obliczenia w kategoriach przyczynowych. Mówili, że symbol to wzór fizyczny mający skutki przyczynowe. Znaczenie symbolu to zbiór zmian, na które pozwala system przetwarzania informacji, wywołać albo w odpowiedzi na jakiś obiekt lub proces

(na zewnątrz lub wewnątrz samego systemu). Analogicznie przyczynowo zdefiniowano pojęcia takie jak reprezentacja, interpretacja, desygnacja, odniesienie, nazewnictwo, reprezentowanie i około. Krótko mówiąc, ich teoria semantyczna przedstawiała znaczenie jako obliczeniowe, a obliczenia jako zamierzone. To właśnie ich oświadczenie o silnej sztucznej inteligencji Searle atakował w swoim słynnym artykule na temat chińskiego pokoju. Na wypadek, gdyby niektórzy czytelnicy nie zetknęli się jeszcze z tym przykładem, można go tutaj krótko przedstawić. Searle wyobraża sobie, że siedzi w pokoju bez okien, ze szczeliną, przez którą od czasu do czasu wrzucane są kartki papieru z „zawijasami” i „zawijasami”. Na stole leży pudełko kartek z podobnymi bazgrołami o różnych kształtach; istnieje także zbiór zasad mówiący, że jeśli zostanie przekazany zawijas, Searle powinien znaleźć blingle-blungle i rozdać go lub być może przejść przez długą sekwencję parowania bazgrołów przed podaniem jakiegoś kształtu. Nieznane Searle-in-the-roomowi, bazgroły są napisane po chińsku, zbiór zasad to program sztucznej inteligencji służący do odpowiadania na pytania w języku chińskim, a Chińczycy poza pokojem szczęśliwie używają Searle’a do odpowiadania na pytania na ten czy inny temat. Kluczową kwestią, mówi Searle, jest to, że wszedł do pokoju, nie rozumiejąc języka chińskiego, i niezależnie od tego, jak długo tam pozostanie, nadal nie zrozumie ani słowa, kiedy wyjdzie. Jego wniosek przeciwny GOFAI jest taki, że same formalne obliczenia (co robi Searle-in-the-room) nie są w stanie wygenerować znaczenia ani intencjonalności. Jak to ujął, programy składają się „w całości ze składni, bez semantyki”. Tak silna sztuczna inteligencja jest niemożliwa. Kluczowym założeniem Searle’a było to, że systemy GOFAI są definiowane w kategoriach abstrakcyjnych (nieinterpretowanych) obliczeń Turinga. Wynika z tego, stwierdził, że takie programy to zwykłe tasowanie bezsensownych kształtów (tj. symboli formalnych). Kształty te można w zasadzie interpretować jako (odwzorowujące) wiele różnych czynności: matematykę, muzykę, projektowanie pułapek na myszy i tak dalej. Natomiast w samym programie wybór pomiędzy „znaczeniami” jest arbitralny. Wszelkie znaczenie, jakie wydaje się mieć, pochodzi całkowicie od nas. Zatem chociaż nie możemy powstrzymać się od mówienia o komputerach w kategoriach zamierzonych, nie zasługują one „z natury” na takie przypisania. Silna sztuczna inteligencja jest iluzją. Jeśli chodzi o „słabą” sztuczną inteligencję (tj. sztuczną inteligencję wykorzystywaną do formułowania, rozwijania i testowania teorii psychologicznych), Searle wątpił, czy mózgi na ogół realizują formalne obliczenia. Jeśli tego nie robią, psychologia obliczeniowa oparta na GOFAI nie będzie w stanie nawet zacząć wyjaśniać naszego życia psychicznego – chociaż mogłyby to zrobić teorie psychologiczne oparte na koneksjonistycznej sztucznej inteligencji. (Kiedy Searle definiował słabą sztuczną inteligencję, użył słowa „AI” do określenia tylko GOFAI; renesans koneksjonizmu jeszcze nie nastąpił.) Jednak nawet jeśli nasze mózgi faktycznie realizują formalne obliczenia, stwierdził, argument Chińskiego Pokoju pokazuje, że potrzeba było czegoś więcej, aby intencjonalność. I tym „czymś” były siły przyczynowe neuroproteiny. Był to bardzo kiepski argument, który większość komentatorów grzecznie zignorowała; gdyż to, w jaki sposób neuroproteina, uważana za substancję biochemiczną, może ugruntować intencjonalność, pozostaje filozoficzną zagadką. Pokój Chiński dał początek kwitnącemu przemysłowi filozoficznemu, którego młyny wciąż wesoło kręcą. Wielu czytelników zgodziło się, że argument Searle’a dotyczący pustej symboliki dowodził dokładnie tego, co twierdził, że udowadnia, podczas gdy równie wielu uważało go za zasadniczo błędny. Co więcej, ci, którzy uważali to za błędne, przedstawiali różne wersje tego, co było w nim złego. To nie miejsce na podsumowanie przedłużającej się debaty. Ale należy wspomnieć o dwóch kwestiach. Po pierwsze, można argumentować, że Searle zaatakował człowieka ze słomy. Newell i Simon, a także ich współpionierzy GOFAI, Marvin Minsky i John McCarthy, niewątpliwie wierzyli, że systemy sztucznej inteligencji o określonej złożoności będą naprawdę inteligentne. Jednak wbrew temu, co zakładał Searle, nie mieli na myśli abstrakcyjnych, niezinterpretowanych obliczeń Turinga. Większość filozofów uważa za oczywiste, że wszelkie obliczenia są obliczeniami Turinga i jako takie są semantycznie puste. Jeśli tak jest, to teoria PSS utrzymuje, że przyczynowość stanowi podstawę interpretacji obliczeń po ich wdrożeniu, a nie jest zaangażowana w definiowanie obliczeń jako takich. Jednakże naszkicowana

powyżej teoria semantyczna Newella i Simona sugeruje, że nie zajmowali się oni przede wszystkim obliczeniami jako nieinterpretowanymi obliczeniami Turinga. (W efekcie ich podejście przyczynowe milcząco oferowało to, co Searle nazwał „odpowiedzią robota”). Co więcej, inni naukowcy zajmujący się sztuczną inteligencją również przedstawili przyczynowe/zamierzone opisy obliczeń, czasami wyraźnie zaprzeczając, że sztuczna inteligencja zajmuje się głównie obliczeniami Turinga. Ich definicje są wprawdzie znacznie mniej jasne niż Turinga, a w jednym przypadku także metafizyczne. Nie wchodząc w szczegóły, chodzi o to, że praktykujący naukowcy zajmujący się sztuczną inteligencją nie myślą o obliczeniach tylko w sposób, w jaki zrobił to Turing w swojej pracy z 1936 roku. Wynika z tego, że znany atak „z całą składnią i bez semantyki” na GOFAL nie jest tak wyraźnie celowy, jak się zwykle zakłada. Po drugie, niektórzy filozofowie argumentowali, że intencjonalność jest zakorzeniona w naszej historii ewolucyjnej. Jeśli tak jest, to jedynie rozwinięte systemy/roboty AI (w tym przykłady oparte na GOFAL) mogą w ogóle być kandydatami do posiadania intencjonalności. Niektóre znaczenia przypisywane wyewoluowanym robotom nie są „arbitralne”, jak chciałby Searle, ale są zakorzenione w ich szczególnej historii ewolucyjnej. Na przykład minisieć w „mózgu” robota może być prawdopodobnie postrzegana jako detektor orientacji, który ewoluował w celu nawigacji w środowisku zadaniowym. Prawdą jest jednak, że to, czy znaczenia te – choć nie arbitralne – są znaczeniami rzeczywistymi, pozostaje dyskusyjne. Jak zauważono powyżej, argument dotyczący pustego programu nie był jedynym powodem, dla którego Haugeland odrzucił GOFAL. Jeszcze ważniejszy w jego oczach (i Dreyfusa) był brak ucieleśnienia programów GOFAL. Cecha ta opierała się na kartezjańskim oddzieleniu umysłu, ciała i świata – z jego obrazem życia psychicznego jako czegoś w rodzaju sztafety, w której pałeczkę przekazuje szereg analitycznie odrębnych etapów percepcji, myślenia i działania motorycznego. Obraz ten przesiąknął (większość) psychologii eksperymentalnej i neurofizjologii na długo przed pojawieniem się GOFAL. Ale GOFAL uczynił to jeszcze bardziej oczywistym, na przykład w programach analizy scen z lat 60. i 70. XX wieku, które badały „wizję” jako materię bezcielesną i quasi-intelektualistyczną. Haugeland preferuje podejście fenomenologiczne, w którym cielesna obecność i działanie w świecie materialnym (i społecznym) są źródłem wszelkiego znaczenia i świadomości. Podejście to, z towarzyszącym naciskiem na systemy dynamiczne, zostało wprowadzone do kognitywistyki przez Francisco Varełę. Dziś krytyka fenomenologiczna ma znacznie większy wpływ w tej dziedzinie. Nazwisko Martina Heideggera jest obecnie bardziej godne uwagi niż Bertrand Russell, którego prace dotyczące logiki (i logikistycznej filozofii języka) odegrały tak ważną rolę w tworzeniu symbolicznej sztucznej inteligencji. Ale ci dwaj filozofowie nie mogliby być bardziej różni. W wyniku tej zasadniczej zmiany GOFAL wyszło teraz jeszcze bardziej z mody niż w czasach świetności PDP. Z zainteresowań filozoficznych za ważne nadal uważa się prace GOFAL dotyczące robotyki i poznania rozproszonego. To, czy o jakimkolwiek robocie, GOFAL czy nie, można właściwie powiedzieć, że jest „wcielony”, jest bardzo istotną kwestią, w którą nie możemy się tutaj zagłębiać. Ale widzieliśmy już, że robot może zostać umiejscowiony w świecie fizycznym i rzucać mu wyzwania w sposób, w jaki nie są to typowe (intelektualne/intelektualistyczne) programy GOFAL. Niektórzy czytelnicy mogą uważać, że stwierdzenie, że GOFAL „wyszło z mody”, nie jest wystarczające. Przecież modę można ożywić i podziwiać na nowo. Jednak wiele osób uważa, że GOFAL na zawsze wypadło z łask, żeby nie powiedzieć, że umarło. Dlaczego tak myślą? Czy mają rację?

### **Mit porażki GOFAL**

Widzieliśmy, że mocne i słabe strony GOFAL i koneksjonizmu oraz GOFAL i robotyki usytuowanej uzupełniają się. Jednak w opinii ogółu społeczeństwa (i niektórych badaczy sztucznej inteligencji) częściej zauważa się słabości GOFAL niż mocne strony. Rzeczywiście wielu komentatorów uważa za oczywiste, że GOFAL poniosło porażkę i należy to odpowiednio spisać na straty. Na przykład Dreyfus, który około czterdzieści lat temu (1965) przewidział upadek GOFAL, obecnie uważa się za usprawiedliwionego. Jego wykład na Uniwersytecie w Houston w 1998 r. nosił tytuł „Dlaczego



symboliczna sztuczna inteligencja zawiodła” i ten jednoznaczny sąd wyraził już w druku: „Tradycja racjonalistyczna została w końcu wystawiona na próbę empiryczną, która zakończyła się niepowodzeniem” (Dreyfus i Dreyfus 1988, s. 34). Pogląd ten rozpowszechnił się tak szeroko, że wiele osób uważa go obecnie za banał. Z pewnością GOFAI nie spełniło swojego wczesnego szumu. Na przykład przewidywanie Simona z lat 50. XX w., że w 1967 r. zapewni to tytuł mistrza świata w szachach, nie sprawdziły się. (Deep Blue pokonał Gary’ego Kasparowa trzydzieści lat później dzięki specjalnym chipom umożliwiającym przewidywanie ośmiu ruchów w przód). Wiele innych wczesnych przewidywań, takich jak niemal bezbłędne tłumaczenie maszynowe, również się nie potwierdziło. Nowsze przykłady szumu wokół GOFAI obejmują pewne przewidywania dotyczące wyłonienia mistrza świata w piłce nożnej w robotyce do roku 2050. Moim zdaniem takie obietnice nie zostaną spełnione przez setki lat, jeśli w ogóle. Ale to nie oznacza „porażki”. Sukcesu w tej dziedzinie nie należy bowiem oceniać poprzez skupianie się na jej najbardziej ekstrawaganckich obietnicach – zwłaszcza że wielu badaczy sztucznej inteligencji, w przeszłości i obecnie, nie podzielało tych rozbudzonych szumem nadziei. Nie należy go też oceniać pod kątem zdania testu Turinga. Nie było to nigdy właściwym kryterium sukcesu w sztucznej inteligencji – ani też nie było takie w zamierzeniu samego Turinga (Boden 2006, rozdz. 16.ii.c). Zatem faktu, że nie została uchwalona, nie ma ani tu, ani tam. Niemniej jednak wiele osób nadal uważa, że jest to bardzo istotne. Na przykład wykład Dreyfusa „Porażka” rozpoczął się od odniesienia do przewidywań Turinga dotyczących bliskiego podobieństwa człowieka do komputera do roku 2000. Nasza ocena powinna raczej opierać się na tym, czy GOFAI poczynił znaczące postępy w realizacji celów wyrażonych w mniej nieostrożnych chwilach. I tutaj musimy przypomnieć rozróżnienie między technologiczną i psychologiczną sztuczną inteligencją. Jeśli chodzi o technologiczne GOFAI, zarzut niepowodzenia jest absurdalny. Sam problem ramy, choć wprawdzie nierozwiązany (i prawdopodobnie nierozwiązywalny) w ogólnym przypadku, został rozwiązany do celów praktycznych w wielu różnych kontekstach. Systemy ekspertowe zostały opracowane w wielu dziedzinach. Wykorzystuje się je m.in. do celów handlowych, administracyjnych, edukacyjnych, medycznych i wojskowych. Niektóre z nich są zdolne do bardzo złożonego planowania, czasami skalującego do dziesiątek tysięcy kroków. Planowanie GOFAI jest nawet ważne na przykład w grach wideo i hollywoodzkich animacjach, aby zapobiec wpadaniu na siebie postaci z wirtualnej rzeczywistości. Co więcej, upragniona sieć semantyczna – zwykle uważana za futurystyczny projekt z zakresu informatyki, a nie sztucznej inteligencji (patrz poniżej) – będzie wymagać postępu w skomputeryzowanych ontologiach, a jest to temat zapoczątkowany przez GOFAI. Systemy eksperckie, gry VR i wyszukiwarki internetowe, takie jak Google, ilustrują fakt, że w społeczeństwach przemysłowych istnieje wiele niewidzialnych zastosowań sztucznej inteligencji, z których opinia publiczna nawet nie jest świadoma. Zapaleni gracze wideo nie wiedzą, że w ich ulubionym systemie znajduje się planista GOFAI. Pracownicy call center często oferują porady pochodzące z systemu eksperckiego GOFAI, których nie widzi klient po drugiej stronie linii telefonicznej. Wiele urządzeń domowych, od samochodów po kuchenki, opiera się na wbudowanej technologii GOFAI (i koneksjonistycznej). Niewidzialność to tylko jeden z powodów, dla których sukcesy GOFAI pozostają w dużej mierze niedoceniane. Innym jest nieidentyfikowalność. Wiele aspektów sztucznej inteligencji odniosło taki sukces, że ludzie (w tym inni informatycy) uważają je jedynie za część głównego nurtu informatyki. Należą do nich techniki komputerowe, które obecnie są uważane za oczywiste, takie jak dzielenie czasu, komputery osobiste z systemem Windows i myszami oraz programowanie obiektowe. Zapomina się o ich korzeniach w sztucznej inteligencji. Pod tym względem sztuczną inteligencję można porównać do filozofii. Odważnie zadaje pytania, na które nie ma odpowiedzi, prawie nie można ich zadać, ale kiedy znajdzie rzetelny sposób na udzielenie na nie odpowiedzi, ponownie określa się je mianem pytań do „szacunkowej” nauki. Krótko mówiąc, nie można utrzymać zarzutu, że GOFAI poniosło porażkę jako technologia. Ale co z GOFAI jako przedsiębiorstwem psychologicznym? Jeśli chodzi o słabą sztuczną inteligencję, nie ma wątpliwości, że doprowadziło to do wielu odkryć

psychologicznych. Niektóre prace oparte na GOFAI z pewnością doprowadziły do tego, co filozof Karl Popper nazwał „obaleniem” pozornie obiecujących „przypuszczeń”. Na przykład GOFAI okazał się nieodpowiedni do modelowania widzenia na niskim poziomie i radzenia sobie z wieloma częściowo sprzecznymi ograniczeniami. Ale to jest normalne: dialektyka przypuszczeń/obaleń jest tym, co napędza naukę do przodu. Co więcej, psychologia obliczeniowa oparta na GOFAI nie spotkała się wyłącznie z obaleniem. Doprowadziło to do nowych danych psychologicznych i/lub głębszego zrozumienia teoretycznego na wiele tematów. Należą do nich emocje, hipnoza i psychopatologia, a także zjawiska poznawcze, takie jak rozumowanie, rozwiązywanie problemów, błędy w działaniu i język. Sztuczna inteligencja w ogóle – w tym GOFAI – jest nawet ważniejsza niż konkretne dane i teorie – skłoniła psychologów do docenienia wcześniej nieoczekiwanej subtelności i mocy obliczeniowej ludzkiego umysłu. Poeci i powieściopisarze (a także Zygmunt Freud) od dawna intuicyjnie wyczuwali to bogactwo. Jednak teoretycy przedobliczeniowi nie mieli pojęcia, do jakiego stopnia nie udało im się tego uchwycić. Silna sztuczna inteligencja to inna sprawa. Jak zauważono w poprzedniej części, pierwotna definicja GOFAI obejmowała zaangażowanie w silną sztuczną inteligencję. W tym przypadku istnieją dwa ogólne powody, dla których nie można twierdzić, że GOFAI odniosło sukces – chociaż to, czy „poniosło porażkę”, jest kwestią bardziej delikatny osąd. Z jednej strony pogląd (wysuwany na przykład przez Newella i Simona), że obliczenia symboliczne mogą wyjaśnić/wdrażać wszystkie aspekty inteligencji, jest błędny. Obliczenia koneksjonistyczne są również kluczowe. Zauważyliśmy jednak, że GOFAI ma mocne strony, z którymi inne metodologie sztucznej inteligencji nie mogą jeszcze konkurować i które są bardzo istotne dla niektórych aspektów ludzkiego umysłu. Jeśli zatem silna sztuczna inteligencja (szeroko rozumiana jako twierdzenie, że jakiś system obliczeniowy może być rzeczywiście inteligentny) jest prawdziwa, wówczas GOFAI może stanowić część naprawdę inteligentnego systemu. Z drugiej strony nie jest jasne, czy silna sztuczna inteligencja jest poprawna. Rzeczywiście, jest to równie kontrowersyjne jak zawsze – nie tylko dlatego, że wymagałoby naturalistycznego (naukowego) wyjaśnienia intencjonalności, czyli znaczenia. Filozofowie, którzy pragną znaleźć takie wyjaśnienie, nie są zgodni co do tego, jak tego dokonać. Moim zdaniem najbardziej obiecująca próba ma charakter ewolucyjny; ale nawet jego zwolennicy przyznają, że ich stanowisko jest pod wieloma względami sprzeczne z intuicją. Co więcej, wielu filozofów – w tym większość zwolenników fenomenologii – twierdzi, że naturalistyczne ujęcie intencjonalności jest w zasadzie niemożliwe. Nawet Hilary Putnam, której filozofia funkcjonalizmu (1960) przyczyniła się do powstania silnej sztucznej inteligencji opartej na GOFAI, akceptuje obecnie wersję tego poglądu i w związku z tym nazywa kognitywistykę „science fiction”. Podstawowym twierdzeniem antynaturalistów jest to, że ludzki język, znaczenie i świadomość są źródłem wszystkich naszych koncepcji i całej naszej wiedzy (w tym nauki), tak że same te zjawiska nigdy nie mogłyby zostać wyjaśnione przez naukę. Z tego punktu widzenia silna sztuczna inteligencja jest filozoficznie wywrócona do góry nogami. Schizma naturalistyczna/antynaturalistyczna jest najgłębszym podziałem w filozofii. Doprowadziło to do obelżywych „wojen naukowych”, w których nauka jako całość była zjadliwie atakowana i zaciekle broniona. Biorąc pod uwagę jej roszczenia do wyjaśniania zjawisk psychicznych, psychologiczna sztuczna inteligencja – niezależnie od tego, czy jest to GOFAI, czy nie – jest pierwsza na linii ognia. Większość naukowców to realisci: wierzą, że prawdziwy świat istnieje niezależnie od ludzi i ma swoje własne właściwości, które większość ludzkiej myśli, a zwłaszcza nauki, próbuje opisać. Realisci zazwyczaj oskarżają antynaturalistów o bycie irracjonalnym, samobójczym i niezdolnym do wyjaśnienia wielu sukcesów nauki. Moim zdaniem realisci mają rację. Żadna ze stron nie ma jednak argumentów za przełamaniem. Z pewnością ta filozoficzna batalia nie zostanie szybko rozstrzygnięta. Podsumowując: postrzegany jako technologia, GOFAI jest niezwykle imponującym, choć w dużej mierze niewidocznym (a zatem nierozpoznanym) sukcesem. Od systemów ekspertowych po Internet – bez nich codzienne życie w społeczeństwach przemysłowych byłoby zupełnie inne. Uważana za psychologię, jest częściowo skuteczna, ale wymaga uzupełnienia innymi metodologiami sztucznej

inteligencji. Rzuciło światło na bardzo szeroki zakres zjawisk psychologicznych, od codziennego poznania (w tym użycia języka) po błędy w działaniu, hipnozę i halucynacje. Uważana za filozofię jest częściowo błędna i niektórzy (tj. antynaturaliści) powiedzieliby, że zasadniczo podąża ona błędną drogą. Jeśli jednak ten ostatni zarzut jest prawdziwy, wszystkie inne formy AI/A-Life – a nawet neuronaukowe wyjaśnienia umysłu – upadają wraz z nim. Innymi słowy, osoby sprzeciwiające się GOFAI ze względów filozoficznych często posługują się argumentami, które – choć nie zawsze są z tego świadome – podają w wątpliwość także zasadność euronauki