

Podstawy inteligencji ogólnej: bezpośrednia droga do sztucznej inteligencji ogólnej

Wprowadzenie

Tu omówiono koncepcję „sztucznej inteligencji ogólnej” (AGI) – jej naturę, znaczenie i najlepsze sposoby jej osiągnięcia. Nasz model teoretyczny zakłada, że inteligencja ogólna składa się z ograniczonej liczby odrębnych, lecz wysoce zintegrowanych, podstawowych komponentów funkcjonalnych. Pomyślnie wdrożenie tego modelu doprowadzi do powstania wysoce adaptacyjnego systemu ogólnego przeznaczenia, który może samodzielnie zdobywać niezwykle szeroki zakres specyficznej wiedzy i umiejętności. Co więcej, będzie mógł dzięki temu doskonalić własne zdolności poznawcze samokształcenie. Wierzmy, że przy właściwym projekcie obecna technologia sprzętu/oprogramowania jest odpowiednia do projektowania praktycznych systemów AGI. Poniżej opisano naszą obecną realizację funkcjonalnego prototypu. Idea „wywiadu ogólnego” jest dość kontrowersyjna; Nie angażuję się tutaj zasadniczo w tę debatę, ale raczej przyjmuję istnienie takich zdolności, które nie są specyficzne dla domeny, jako oczywiste. Należy również zauważyć, że ten esej koncentruje się przede wszystkim na zdolnościach poznawczych niskiego poziomu (tj. mniej więcej na poziomie zwierzęcym). Funkcjonalność wyższego poziomu, choć stanowi integralną część naszego modelu, jest adresowana jedynie peryferyjnie. Wreszcie, niektóre szczegóły algorytmiczne zostały pominięte ze względu na własność.

Inteligencja ogólna

Inteligencję można po prostu zdefiniować jako zdolność jednostki do osiągania celów – przy większej inteligencji do radzenia sobie w bardziej złożonych i nowych sytuacjach. Złożoność sięga od spraw trywialnych – termostaty i mięczaki (które w większości kontekstów nawet nie uzasadniają określenia „inteligencja”) – po fantastycznie złożone; autonomiczne systemy kontroli lotu i ludzie. Adaptowalność, czyli umiejętność radzenia sobie ze zmieniającymi się i nowatorskimi wymaganiami, również obejmuje szerokie spektrum: od sztywnych, wąsko specyficznych dla danej dziedziny, po wysoce elastyczne, ogólnego przeznaczenia. Co więcej, elastyczność można zdefiniować w kategoriach zakresu i trwałości – jak bardzo i jak często się zmienia. Wdrukowanie jest przykładem ograniczonego zakresu i dużej trwałości, podczas gdy innowacyjne, abstrakcyjne rozwiązywanie problemów znajduje się na drugim końcu spektrum. Chociaż podmioty o wysokiej zdolności adaptacyjnej i elastyczności są wyraźnie lepsze – mogą potencjalnie nauczyć się osiągać każdy możliwy cel – istnieje wysoka cena za wydajność: na przykład, gdyby Deep Blue został również zaprojektowany do nauki języka, bezpośrednio linie lotnicze ruchu ulicznego i przeprowadzania diagnoz medycznych, nie pokonałby mistrza świata w szachach (przy wszystkich innych czynnikach niezmiennych). Inteligencja ogólna obejmuje podstawowe, niezależne od dziedziny umiejętności niezbędne do zdobycia szerokiego zakresu wiedzy specyficznej dla danej dziedziny (danych i umiejętności) – tj. możliwość nauczenia się czegokolwiek (w zasadzie). Mówiąc dokładniej, ta zdolność uczenia się musi być autonomiczna, ukierunkowana na cel i wysoce adaptacyjna:

Autonomiczny. Uczenie się zachodzi zarówno automatycznie, poprzez kontakt z danymi zmysłowymi (bez nadzoru), jak i poprzez dwukierunkową interakcję ze środowiskiem, obejmującą eksplorację i eksperymentowanie (samo nadzór).

Ukierunkowany na cel. Uczenie się jest ukierunkowane (autonomicznie) na osiągnięcie różnorodnych i nowych celów i celów cząstkowych – niezależnie od tego, czy są one „wpisane na stałe”, określone zewnątrz, czy też wygenerowane samodzielnie. Ukierunkowanie na cel oznacza również bardzo selektywne uczenie się i pozyskiwanie danych (z niezwykle bogatego w dane, hałaśliwego i złożonego środowiska).

Adaptacyjny. Uczenie się ma charakter kumulatywny, integrujący, kontekstowy i dostosowuje się do zmieniających się celów i środowiska. Ogólna adaptacyjność nie tylko radzi sobie ze stopniowymi zmianami, ale także zasiewa i ułatwia nabywanie zupełnie nowych umiejętności.

Ogólne zdolności poznawcze ostro kontrastują z nieodłącznymi specjalizacjami, takimi jak rozpoznawanie mowy lub twarzy, bazy danych/ontologie wiedzy, systemy ekspertowe lub algorytmy wyszukiwania, regresji lub optymalizacji. Pozwala jednostce na zdobycie praktycznie nieograniczonego zakresu nowych, wyspecjalizowanych umiejętności. Cechą ogólnie inteligentnego systemu nie jest posiadanie dużej wiedzy i umiejętności, ale umiejętność ich zdobywania i doskonalenia – oraz umiejętność ich odpowiedniego zastosowania. Ponadto wiedzę należy zdobywać i przechowywać w sposób odpowiedni zarówno do charakteru danych, jak i celów i zadań. Na przykład, biorąc pod uwagę prawidłowy zestaw podstawowych możliwości, system AGI powinien być w stanie nauczyć się rozpoznawać i kategoryzować szeroką gamę nowych wzorców percepcyjnych, które są nabywane za pomocą różnych zmysłów, w wielu różnych środowiskach i kontekstach. Ponadto powinien być w stanie samodzielnie uczyć się odpowiednich, ukierunkowanych na cel reakcji na takie konteksty wejściowe (biorąc pod uwagę pewien mechanizm sprzężenia zwrotnego). Uważamy, że ta koncepcja ma zastosowanie nie tylko w przypadku inteligencji ludzkiej na wysokim poziomie, ale także w przypadku zdolności zwierzęcych na niższym poziomie. Stopień „ogólności” (tj. zdolności adaptacyjnych) zmienia się w kontinuum, od genetycznie „zakodowanych na stałe” reakcji (brak zdolności adaptacyjnych), do wysokiego poziomu elastyczności zwierzęcia (znaczną zdolność uczenia się, jak na przykład u psa), i wreszcie do samoświadoma ogólna zdolność uczenia się człowieka.

Podstawowe wymagania dotyczące inteligencji ogólnej

Inteligencja ogólna, jak opisano powyżej, wymaga szeregu nieredukowalnych cech i możliwości. Aby proaktywnie gromadzić wiedzę z różnych (i/lub zmieniających się) środowisk, wymagane jest:

1. zmysły w celu uzyskania cech „świata” (wirtualnego lub rzeczywistego);
2. spójny sposób przechowywania zdobytej w ten sposób wiedzy; i
3. adaptacyjne mechanizmy wyjściowe/uruchamiające (zarówno statyczne, jak i dynamiczne).

Wiedza taka wymaga także automatycznego dostosowywania i bieżącej aktualizacji; nowa wiedza musi być odpowiednio powiązana z istniejącymi danymi. Co więcej, postrzegane byty/wzorce muszą być przechowywane w sposób ułatwiający tworzenie i uogólnianie koncepcji. Skutecznym sposobem przedstawienia złożonych relacji między cechami jest kodowanie wektorowe. Wszelkie praktyczne zastosowania AGI (a już na pewno wszelkie zastosowania w czasie rzeczywistym) muszą z natury umożliwiać przetwarzanie danych tymczasowych jako wzorców w czasie – a nie tylko jako wzorców statycznych o wymiarze czasowym. Co więcej, AGI muszą radzić sobie z danymi z różnych sond zmysłowych (np. wizualnych, słuchowych i danych) oraz radzić sobie z takimi atrybutami, jak: zaszumione, skalarne, niewiarygodne, niekompletne, wielowymiarowe (zarówno wymiarowe w przestrzeni/czasie, jak i posiadające dużą liczbę jednoczesnych funkcji) itp. Rozmyte dopasowywanie wzorców pomaga radzić sobie ze zmiennością wzorców i szumem. Kolejnym istotnym wymogiem inteligencji ogólnej jest radzenie sobie z nadmiarem danych. Rzeczywistość przedstawia znacznie więcej cech i szczegółów, niż jest to (kontekstowo) istotne lub które można z pożytkiem przetworzyć. Dlatego też system musi mieć jakąś kontrolę nad tym, jakie dane wejściowe zostaną wybrane do analizy i uczenia się – zarówno pod względem jakich danych, jak i stopnia szczegółowości. Zmysły („sondy”) są potrzebne nie tylko do selekcji i skupienia, ale także do ugruntowania pojęć – nadania im (opartego na rzeczywistości) znaczenia. Chociaż dane wejściowe muszą być poważnie ograniczone poprzez skupienie i selekcję, niezwykle ważne jest również uzyskanie wielu widoków rzeczywistości – danych z

różnych ekstraktorów cech lub zmysłów. Pod warunkiem, że te różne wzorce wejściowe są odpowiednio powiązane, mogą pomóc w zapewnieniu sobie kontekstu, ułatwieniu rozpoznania i dodaniu znaczenia. Oprócz możliwości wyczuwania za pomocą wielu adaptacyjnych grup wejściowych i sond, AGI musi także być w stanie oddziaływać na świat – czy to w celu eksploracji, eksperymentowania, komunikacji, czy wykonywania przydatnych działań. Mechanizmy te muszą zapewniać zarówno statyczne, jak i dynamiczne dane wyjściowe (stany i zachowanie). One także muszą być przystosowawcze i zdolne do uczenia się. U podstaw całej tej funkcjonalności leży przetwarzanie wzorców. Co więcej, nie tylko odczuwanie i działanie opierają się na ogólnych wzorcach, ale także wewnętrzna aktywność poznawcza. W rzeczywistości nawet abstrakcyjne myślenie, język i rozumowanie formalne na wysokim poziomie – umiejętności wykraczające poza zakres naszego obecnego projektu – są „tylko” rozwinięciami tego zagadnienia wyższego rzędu.

Zalety ogólnego charakteru inteligencji

Zalety ogólnej inteligencji są niemal zbyt oczywiste, aby zasługiwały na wymienienie ich; ilu z nas marzyłoby o rezygnacji ze zdolności do adaptacji i uczenia się nowych rzeczy? W kontekście sztucznej inteligencji zagadnienie to nabiera nowego znaczenia. Istnieje niewyczerpane zapotrzebowanie na skomputeryzowane systemy, które mogą pomóc ludziom w wykonywaniu złożonych zadań, które są wysoce powtarzalne, niebezpieczne lub wymagają wiedzy, zmysłów lub zdolności, których ich użytkownicy mogą nie posiadać (np. wiedzy eksperckiej, pamięci „fotograficznej”, pokonywania niepełnosprawności, itp.). Zastosowania te obejmują niemal wszystkie dziedziny działalności człowieka. Obecnie potrzeby te zaspokajają przede wszystkim systemy zaprojektowane specjalnie dla każdej domeny i zastosowania (np. systemy ekspertowe). Problemy związane z kosztami, czasem realizacji, niezawodnością i brakiem możliwości dostosowania się do nowych i nieprzewidzianych sytuacji poważnie ograniczają potencjał rynkowy. Adaptacyjna technologia AGI opisana w tym artykule obiecuje znacząco zmniejszyć te ograniczenia i otworzyć te rynki. W szczególności oznacza to:

- Aby systemy mogły uczyć się (i być nauczane) szerokiego spektrum danych i funkcjonalności
- Potrafią dostosować się do zmieniających się danych, środowisk i zastosowań/celów
- Można to osiągnąć bez zmian w programie – możliwości się uczy, a nie koduje

Mówiąc dokładniej, technologia ta może potencjalnie:

- Znacząco zmniejsz „kruchość” systemu poprzez dopasowywanie wzorców rozmytych i uczenie się adaptacyjne – zwiększając odporność w obliczu zmieniających się i nieprzewidzianych warunków lub danych.
- Ucz się autonomicznie, automatycznie gromadząc wiedzę o nowych środowiskach poprzez eksplorację.
- Umożliwiają przeszkolenie operatorów systemów w zakresie identyfikowania nowych obiektów i wzorców; reagować na sytuacje w określony sposób i nabywać nowe zachowania.
- Wyeliminuj programowanie w wielu aplikacjach. Systemy można stosować w wielu różnych środowiskach i przy różnych parametrach, po prostu w drodze samokształcenia.
- Ułatwienie łatwego wdrażania w nowych domenach. Ogólny silnik inteligencji z podłączanymi niestandardowymi sondami wejściowymi/wyjściowymi umożliwia szybką i niedrogą realizację specjalistycznych aplikacji

Z punktu widzenia projektowania AGI ma tę zaletę, że cały wysiłek można skupić na osiągnięciu najlepszych ogólnych rozwiązań – rozwiązując je raz, a nie raz dla każdej konkretnej domeny. AGI ma oczywiście także ogromne implikacje ekonomiczne: ponieważ systemy AGI zdobywają większość swojej wiedzy i umiejętności (oraz dostosowują się do zmieniających się wymagań) w sposób autonomiczny, czas realizacji programowania i koszty można radykalnie skrócić, a nawet wyeliminować. Fakt, że obecnie nie istnieją (sztuczne!) systemy posiadające takie możliwości, wydaje się sugerować, że osiągnięcie tych celów jest bardzo trudne (lub niemożliwe). Wierzę jednak, że podobnie jak w przypadku innych przykładów odkryć i wynalazków człowieka, rozwiązanie będzie wydawać się dość oczywiste z perspektywy czasu. Sztuka polega na właściwym wybraniu kilku krytycznych opcji rozwoju

Skróty do AGI

Często wyjaśniając sztuczną inteligencję ogólną niewtajemniczonym słyszy uwagę, że z pewnością wszyscy w AI pracują nad osiągnięciem ogólnej inteligencji. To pokazuje, jak głęboko błędnie rozumiana jest inteligencja. Chwilą prawdy jest, że ostatecznie tak się stanie w przypadku konwencjonalnych (specyficznych dla danej dziedziny) wysiłków badawczych które zbiegają się z wynikami AGI, bez celowych wskazówek prawdopodobnie tak się stanie długi i nieefektywny proces. Inteligencja wysokiego poziomu musi być adaptacyjna, musi być ogólna – choć niewiele robi się, aby określić, co jest ogólną inteligencją a jest, czego wymaga i jak ją osiągnąć. Oprócz zrozumienia ogólnej inteligencji, projektowanie AGI wymaga również oraz docenienie różnic pomiędzy sztucznymi (syntetycznymi) i biologiczną inteligencją oraz pomiędzy zaprojektowanymi i rozwiniętymi systemami. Nasze szczególne podejście do osiągnięcia AGI opiera się na obszernej analizie tych problemów oraz na stopniową ścieżkę rozwoju mającą na celu minimalizację wysiłku rozwojowego (czas i koszt), złożoność techniczna i ogólnie ryzyko projektu. W szczególności skupiamy się na inżynierii szeregu funkcjonalności (ale o niskiej rozdzielczości/pojemności) prototypy sprawdzające koncepcję. Wydajność zagadnienia specyficznym związane z komercjalizacją są przydzielone do odrębnego opracowania utworu. Co więcej, nasze początkowe wysiłki koncentrują się na identyfikacji i najpierw wdrażając najbardziej ogólne i podstawowe elementy, wychodząc poznanie wysokiego poziomu, takie jak myślenie abstrakcyjne, język i logika formalna, pod zagospodarowanie terenu (więcej na ten temat). Skupiamy się również bardziej na selektywnym, uczenie się bez nadzoru, dynamiczne, przyrostowe i interaktywne; na hałaśliwym, złożonym, dane analogowe; oraz na integracji cech jednostki i atrybutów koncepcji w jednej sieci kompleksowej. Chociaż nasz projekt może nie być jedynym, w którym poruszana jest ta konkretna kwestia, jasne jest, że zdecydowana większość obecnych prac nad sztuczną inteligencją zasadniczo podąża tą ścieżką. Nasza praca koncentruje się na:

- ogólne, a nie specyficzne dla danej domeny zdolności poznawcze;
- nabyta wiedza i umiejętności w porównaniu z załadowanymi bazami danych i zakodowanymi umiejętnościami;
- dwukierunkowa interakcja w czasie rzeczywistym w porównaniu z przetwarzaniem wsadowym;
- uwaga adaptacyjna (koncentracja i selekcja) w porównaniu z wcześniej wybranymi danymi ludzkimi;
- podstawowa obsługa wzorców dynamicznych w porównaniu do danych statycznych;
- uczenie się bez nadzoru i samonadzoru a uczenie się pod nadzorem;
- adaptacyjne, samoorganizujące się struktury danych w porównaniu ze stałymi sieciami neuronowymi lub bazy danych;

- kontekstowe, ugruntowane koncepcje w porównaniu z zakodowanymi na stałe koncepcjami symbolicznymi;
- wyraźne projektowanie funkcjonalności zamiast jej rozwijania;
- projekt koncepcyjny a inżynieria odwrotna;
- ogólna weryfikacja koncepcji w porównaniu z rozwojem konkretnych rzeczywistych aplikacji;
- Poznanie na poziomie zwierzęcym a myśl abstrakcyjna, język i logika formalna.

Przyjrzyjmy się każdemu z tych wyborów bardziej szczegółowo.

Ogólne, a nie specyficzne dla domeny zdolności poznawcze

Korzyści wymienione w poprzedniej sekcji wynikają z faktu, że ogólnie inteligentne systemy mogą ostatecznie nauczyć się dowolnej specjalistycznej wiedzy i umiejętności – dowodem jest ludzka inteligencja! Odwrotna sytuacja oczywiście nie jest prawdą. Kompletna, dobrze zaprojektowana zdolność AGI do zdobywania możliwości specyficznych dla domeny jest ograniczona jedynie pojemnością przetwarzania i przechowywania. Co więcej, znaczna część jego nauki będzie autonomiczna – bez nauczycieli, a już na pewno bez wyraźnego programowania. Podejście to wdraża (i wykorzystuje) istotę „nasionowej sztucznej inteligencji” – systemów z ograniczonym, ale starannie wybranym zestawem podstawowych, początkowych możliwości, które pozwalają im (w procesie „bootstrappingu”) radykalnie zwiększyć swoją wiedzę i umiejętności poprzez samodzielne - ukierunkowane uczenie się i adaptacja. Koncentrując się na starannym projektowaniu nasion inteligencji, a następnie pielęgnowaniu ich do dojrzałości, w zasadzie inteligencję ładuje się. W naszym projekcie AGI samodoskonalenie przybiera dwie różne formy/fazy

1. Kodowanie podstawowych umiejętności, które pozwalają systemowi na zdobycie dużej ilości konkretnej wiedzy.
2. System osiąga wystarczającą inteligencję i zrozumienie koncepcyjne własnego projektu, aby umożliwić mu celowe ulepszanie własnego projektu.

Zdobyta wiedza i umiejętności a załadowane bazy danych i zakodowane umiejętności

Jedną z kluczowych miar ogólnej inteligencji jest jej zdolność do zdobywania wiedzy i umiejętności, a nie to, ile posiada. Wiele wysiłków związanych ze sztuczną inteligencją koncentruje się na gromadzeniu ogromnych baz wiedzy i kodowaniu ogromnych ilości określonych umiejętności. Jeśli AGI będzie możliwa – a dowody wydają się przytłaczające – wówczas znaczna część tego wysiłku pójdzie na marne. AGI nie tylko będzie w stanie pozyskać te dodatkowe umiejętności (w dużej mierze) samodzielnie, ale co więcej, będzie w stanie aktualizować i doskonalić swoją wiedzę. Nie tylko zaoszczędzi to początkowego gromadzenia i przygotowania danych, a także programowania, ale także radykalnie ograniczy konserwację.

Ważną cechą naszych projektów jest to, że nie istnieją tradycyjne

bazy danych zawierające wiedzę ani programy kodujące wyuczone umiejętności: cała zdobyta wiedza jest integrowana w adaptacyjną centralną sieć wiedzy/umiejętności. Wzorce reprezentujące wiedzę są powiązane w sposób ułatwiający konceptualizację i wrażliwość na kontekst. Naturalnie taka konstrukcja jest potencjalnie znacznie mniej podatna na kruchość i bardziej odporna na uszkodzenia. Dwukierunkowa interakcja w czasie rzeczywistym a przetwarzanie wsadowe Systemy uczenia adaptacyjnego muszą mieć możliwość dwukierunkowej interakcji ze środowiskiem – wirtualnym lub rzeczywistym. Muszą zarówno wyczuwać dane, jak i na bieżąco działać/reagować. Wiele systemów AI

uczy się w trybie wsadowym i ma niewielką zdolność do uczenia się przyrostowego lub nie ma jej wcale. Takie systemy nie mogą łatwo dostosować się do zmieniającego się środowiska lub wymagań – w wielu przypadkach nie są w stanie przystosować się poza początkowy zestaw szkoleniowy bez przeprogramowania lub ponownego szkolenia. Oprócz percepcji i uczenia się w czasie rzeczywistym inteligentne systemy muszą także potrafić działać. Wymagane są trzy różne obszary możliwości działania:

1. Działanie na „świat” – czy to w celu komunikowania się, nawigacji, eksploracji, czy też manipulowania jakąś zewnętrzną funkcją lub urządzeniem w celu osiągnięcia celów.
2. Kontrolowanie lub modyfikowanie wewnętrznych parametrów systemu (takich jak szybkość uczenia się lub tolerancja na hałas itp.) w celu ustawienia lub ulepszenia funkcjonalności.
3. Kontrolowanie parametrów wejściowych systemu, takich jak ostrość, selekcja, rozdzielczość (ziarnistość), a także dostosowywanie parametrów ekstrakcji cech.

Uwaga adaptacyjna (koncentracja i selekcja) a wstępnie wybrane dane ludzkie

Jak wspomniano wcześniej, rzeczywistość przedstawia znacznie większą obfitość, szczegółowość i złożoność danych zmysłowych, niż jest to wymagane do wykonania dowolnego zadania lub które można przetworzyć. Tradycyjnie problem ten rozwiązywano poprzez staranny wybór i formatowanie danych przed wprowadzeniem ich do systemu. Chociaż taka pomoc człowieka może poprawić wydajność w określonych zastosowaniach, często nie zdajemy sobie sprawy, że ta dodatkowa inteligencja tkwi w człowieku, a nie w oprogramowaniu. Zewnętrzne wskazówki i szkolenia mogą oczywiście przyspieszyć naukę; jednakże systemy AGI muszą z natury być zaprojektowane tak, aby same zdobywały wiedzę. W szczególności muszą kontrolować, jakie dane wejściowe są przetwarzane – gdzie konkretnie uzyskać dane, jak bardzo szczegółowo i w jakim formacie. W przypadku braku tej możliwości system albo zostanie przytłoczony nieistotnymi danymi, albo, odwrotnie, nie będzie w stanie uzyskać kluczowych informacji lub uzyskać ich w wymaganym formacie. Naturalnie takie mechanizmy skupiania się na danych i selekcji same w sobie muszą mieć charakter adaptacyjny.

Podstawowa obsługa wzorców dynamicznych w porównaniu z danymi statycznymi

Przetwarzanie wzorców czasowych to kolejny podstawowy wymóg interaktywnej inteligencji. Opierają się na tym co najmniej trzy aspekty AGI: percepcja musi uczyć się/rozpoznawać dynamiczne byty i sekwencje, działanie zwykle obejmuje złożone zachowanie, a poznanie (przetwarzanie wewnętrzne) jest z natury tymczasowe. Pomimo tej oczywistej potrzeby wewnętrznej obsługi wzorców dynamicznych wiele systemów sztucznej inteligencji przetwarza jedynie dane statyczne; Sekwencje czasowe, jeśli w ogóle są obsługiwane, są często konwertowane („spłaszczane”) zewnątrz, aby wyeliminować wymiar czasu. Przetwarzanie wzorców czasowych w czasie rzeczywistym jest technicznie dość trudne, nic więc dziwnego, że większość projektów stara się tego unikać.

Uczenie się bez nadzoru i samonadzoru a uczenie się pod nadzorem

Systemy autoadaptacyjne, takie jak AGI, wymagają wszechstronnych możliwości uczenia się bez nadzoru. Takie niezależne od nauczyciela nabywanie wiedzy i umiejętności można podzielić na dwie szerokie kategorie: bez nadzoru (oparte na danych, oddolne) i samonadzorowane (odgórne, zorientowane na cel). Idealnie byłoby, gdyby te dwa sposoby uczenia się płynnie integrowały się ze sobą – i oczywiście także z innymi, nadzorowanymi metodami. W tym przypadku, podobnie jak w przypadku innych wyborów projektowych, ogólne systemy adaptacyjne są trudniejsze do zaprojektowania i dostrojenia niż te bardziej wyspecjalizowane, niezmiennie. Widzimy to szczególnie wyraźnie w przytłaczającym skupieniu się na propagacji wstecznej w rozwoju sztucznych sieci

neuronowych (ANN). Stosunkowo niewiele badań ma na celu lepsze zrozumienie i ulepszenie przyrostowego, autonomicznego uczenia się. Nasz własny projekt kładzie duży nacisk na te aspekty. Adaptacyjne, samoorganizujące się struktury danych a stałe sieci neuronowe lub bazy danych Kolejnym podstawowym wymogiem narzuconym przez uczenie się w czasie rzeczywistym zorientowanym na dane/cel jest posiadanie elastycznej, samoorganizującej się struktury danych. Z jednej strony reprezentacja wiedzy musi być wysoce zintegrowana, z drugiej strony musi być w stanie dostosować się do zmieniającej się gęstości danych (i innych właściwości) oraz do zmieniających się celów lub rozwiązań. Nasz AGI koduje całą zdobytą wiedzę i umiejętności w zintegrowanej strukturze przypominającej sieć. To centralne repozytorium charakteryzuje się elastyczną, dynamicznie samoorganizującą się topologią. Zdecydowana większość innych projektów sztucznej inteligencji opiera się albo na luźno powiązanych obiektach danych lub agentach, albo na stałych topologiach sieci i predefiniowanych ontologiach, hierarchiach danych lub układach baz danych. Często poważnie ogranicza to ich zdolność do samouczenia się, zdolność adaptacji i solidność lub tworzy ogromne wąskie gardła w komunikacji lub inne narzuty na wydajność.

Kontekstowe, ugruntowane koncepcje kontra zakodowane na stałe koncepcje symboliczne

Koncepcje są prawdopodobnie najważniejszym aspektem projektowania AGI; w rzeczywistości można powiedzieć, że „inteligencja wysokiego poziomu to inteligencja pojęciowa”. Podstawową cechą koncepcji jest ich zdolność do reprezentowania ultrawysokowymiarowych zbiorów rozmytych, które są osadzone w rzeczywistości, a jednocześnie płynne w odniesieniu do kontekstu. Innymi słowy, kodują powiązane zbiory złożonych, spójnych, wielowymiarowych wzorców, które reprezentują cechy bytów. Pojęcia uzyskują swoje podłoże (a tym samym swoje znaczenie) dzięki wzorom wynikającym z cech odczuwanych bezpośrednio z bytów istniejących w rzeczywistości. Ponieważ pojęcia są definiowane przez zakresy wartości w obrębie każdego wymiaru cechy (czasami w złożonych relacjach), niezbędny jest pewien rodzaj rozmytego dopasowywania wzorców. Ponadto zakres pojęć musi być płynny; muszą być wrażliwe i dostosowywać się zarówno do kontekstu środowiskowego, jak i celów. Tworzenie koncepcji autonomicznych jest jednym z kluczowych testów inteligencji. Wiele systemów sztucznej inteligencji opartych na zakodowanych na stałe lub zdefiniowanych przez człowieka koncepcjach nie zdaje tego podstawowego testu. Co więcej, systemy, które nie czerpią swoich koncepcji poprzez interaktywną percepcję, nie są w stanie oprzeć swojej wiedzy na rzeczywistości, a zatem brakuje im kluczowego znaczenia. Wreszcie struktury pojęciowe, których aktywacji nie można modulować kontekstem i stopniem dopasowania, nie są w stanie uchwycić subtelności i płynności inteligentnego uogólnienia. W połączeniu te ograniczenia okaleczą każdego aspirującego AGI.

Jawne projektowanie funkcjonalności zamiast jej rozwijania

Projektowanie drogą ewolucji jest wyjątkowo nieefektywne – czy to w przyrodzie, czy w informatyce. Co więcej, rozwiązania ewolucyjne są na ogół nieprzejryste; zoptymalizowany tylko pod kątem określonej „funkcji kosztu”, a nie pod kątem zrozumiałości, modułowości lub łatwości konserwacji. Co więcej, uczenie się ewolucyjne wymaga również większej ilości danych lub prób, niż jest dostępnych w codziennym rozwiązywaniu problemów. Programowanie genetyczne i ewolucyjne ma swoje zastosowania – to potężne narzędzia, które można wykorzystać do rozwiązania bardzo specyficznych problemów, takich jak optymalizacja dużych zbiorów zmiennych; jednakże generalnie nie są one odpowiednie do tworzenia dużych systemów infrastruktury. Sztucznie rozwijająca się inteligencja ogólna bezpośrednio wydaje się szczególnie problematyczna, ponieważ nie jest znana żadna funkcja mierząca taką zdolność na pojedynczym kontinuum – a przy braku takiego kierunku ewolucja nie wie, co zoptymalizować. Jednym ze sposobów poradzenia sobie z tym problemem jest próba wydobycia inteligencji ze złożonej ekologii konkurujących ze sobą agentów – zasadniczo odtwarzającej naturalną ewolucję. Ogólnie rzecz biorąc, wydaje się, że techniki programowania genetycznego są odpowiednie,

gdy skończą się konkretne pomysły inżynierskie. Oto krótkie podsumowanie zalet wyraźnie zaprojektowanej funkcjonalności:

- Projekty mogą bezpośrednio wykorzystywać i kodować wiedzę i spostrzeżenia projektanta.
- Projekty posiadają przejrzystą dokumentację projektową.
- Projekty mogą być o wiele bardziej modułowe – mniejsze zapotrzebowanie na wiele funkcjonalności i duża współzależność podsystemów niż w systemach rozwiniętych.
- Systemy mogą mieć bardziej schemat blokowy i logiczną konstrukcję – ewolucja nie przewiduje przyszłości.
- Można je zaprojektować z pomocą przy debugowaniu – ewolucja tego nie potrzebowała
- Połączenie tych funkcji ułatwia zrozumienie, debugowanie, interfejs i – co ważne – umożliwia wielu zespołom jednoczesną pracę nad projektem.

Projekt koncepcyjny a inżynieria odwrotna

Oprócz uniknięcia niedociągnięć technik ewolucyjnych, projektowanie i konstruowanie inteligentnych systemów w oparciu o wymagania funkcjonalne ma również wiele zalet, zamiast próbować kopiować ewolucyjny projekt mózgu. Jak lotnictwo obszernie pokazało, znacznie łatwiej jest budować samoloty niż odtwarzać ptaki metodą inżynierii odwrotnej – znacznie łatwiej jest uzyskać lot za pomocą ciągu niż trzepocząc skrzydłami. Podobnie przy tworzeniu sztucznej inteligencji sensowne jest wykorzystanie naszych ludzkich atutów intelektualnych i inżynierskich – ignorowanie parametrów projektowych charakterystycznych dla systemów biologicznych, zamiast próbować kopiować projekty natury. Projekty zaprojektowane specjalnie w celu osiągnięcia pożądanej funkcjonalności są znacznie łatwiejsze do zrozumienia, debugowania, modyfikowania i ulepszania. Ponadto wykorzystanie znanych i istniejących technologii pozwala nam najlepiej wykorzystać istniejące zasoby. Po co więc ograniczać się do jednego rozwiązania dotyczącego inteligencji stworzonego przez ślepego, nieświadomego Zegarmistrza z własnym planem (przetrwanie w środowisku ewolucyjnym bardzo różniącym się od dzisiejszego)? Inteligentne maszyny zaprojektowane od podstaw nie niosą ze sobą bagażu ewolucyjnego ani dodatkowej złożoności związanej z epigenetyką, reprodukcją i zintegrowaną samonaprawą biologicznych mózgów. Oczywiście nie oznacza to, że możemy się czegośkolwiek nauczyć poprzez badanie mózgów, po prostu nie musimy ograniczać się do biologicznej wykonalności naszych projektów. Nasz (obecnie) jedyny działający przykład inteligencji ogólnej wysokiego poziomu (mózg) zapewnia kluczowy koncepcyjny model poznania i może wyraźnie zainspirować wiele specyficznych cech projektowych. Oto kilka pożądanych cech poznawczych, które można uwzględnić w projekcie AGI, a które nie istniałyby (a w niektórych przypadkach nie mogłyby) istnieć w mózgu poddanym inżynierii wstecznej:

- Skuteczniejsza kontrola neurochemii („stanów emocjonalnych”).
- Dobór odpowiedniego stopnia logicznego myślenia kontra intuicja.
- Skuteczniejsza kontrola nad skupieniem i uwagą.
- Możliwość uczenia się natychmiast, na żądanie.
- Bezpośrednie i szybkie połączenie z bazami danych, Internetem i innymi maszynami – potencjalnie natychmiastowy dostęp do całej dostępnej wiedzy.
- Opcjonalna pamięć „fotograficzna” i przywoływanie („odtworzenie”) wszystkimi zmysłami!

- Lepsza kontrola nad zapamiętywaniem i zapominaniem (zamrażanie ważnej wiedzy i możliwość oduczenia się).
- Umiejętność dokładnego cofania się i przeglądania procesów myślowych i decyzyjnych (odtworzenie i badanie ścieżek logicznych).
- Wzory, węzły i łącza można łatwo oznaczać (oznaczać) i kategoryzować.
- Możliwość optymalizacji projektu pod kątem dostępnego sprzętu zamiast konieczności dostosowywania się do wymagań mózgu.
- Umiejętność wykorzystania najlepszych istniejących algorytmów i technik oprogramowania – niezależnie od tego, czy są one biologicznie wiarygodne.
- Specjalnie zaprojektowany AGI (w przeciwieństwie do mózgow) może mieć prostą ścieżkę zwiększania prędkości/pojemności.
- Możliwość kompleksowej integracji z innymi systemami AI (takimi jak systemy ekspertowe, robotyka, wyspecjalizowane preprocesory zmysłów i rozwiązania problemów).
- Możliwość konstruowania AGI, które są wysoce zoptymalizowane dla określonych domen.
- Dane węzła, łącza i parametrów wewnętrznych są dostępne jako „dane wejściowe” (pełna introspekcja).
- Dostępne są specyfikacje projektowe (dla projektanta i samego AGI!)
- Projekt początkowej sztucznej inteligencji: maszynę można z natury zaprojektować tak, aby łatwiej była zrozumieć i udoskonalić swoje własne funkcjonowanie, w ten sposób podnosząc inteligencję na coraz wyższy poziom.

Ogólny dowód słuszności koncepcji a rozwój konkretnych rzeczywistych aplikacji

Zastosowanie danych zasobów do minimalistycznych projektów sprawdzających koncepcję zwiększa prawdopodobieństwo wytyczenia szybkiej, bezpośredniej ścieżki do ostatecznego celu. Skoro naszym celem jest sztuczna inteligencja ogólna wysokiego poziomu, nie ma sensu marnować zasobów na sprawy nieistotne. Oprócz skupiania naszych wysiłków na możliwości samodzielnego zdobywania wiedzy, zamiast jej przechwytywania i kodowania, naszym celem jest także przyspieszenie postępu w kierunku pełnego AGI poprzez redukcję kosztów i złożoności poprzez:

- Koncentrowanie się na prototypach potwierdzających słuszność koncepcji, a nie na wykonaniu komercyjnym. Obejmuje to pracę przy niskiej rozdzielczości i objętości danych oraz odłożenie optymalizacji. Skalowalność jest rozważana jedynie na poziomie teoretycznym i niekoniecznie wdrażana.
- Praca z radykalnie ograniczonymi zdolnościami zmysłowymi i motorycznymi. Fakt, że osoby głuche, niewidome i poważnie sparaliżowane mogą osiągnąć wysoką inteligencję (Helen Keller, Stephen Hawking) wskazuje, że nie jest ona niezbędna do rozwoju AGI.
- Radzenie sobie ze złożonością poprzez chęć eksperymentowania i wdrażania słabo poznanych algorytmów – tj. stosując podejście inżynierskie. Korzystanie z samodostrajających się pętli sprzężenia zwrotnego w celu zminimalizowania wolnych parametrów.

- Nie dawanie się zwieść próbom dorównania wydajnością projektów specyficznych dla danej dziedziny – skupienie się bardziej na sposobie osiągania możliwości (np. wyuczona konceptualizacja zamiast zaprogramowanych lub ręcznie określonych koncepcji), a nie na czystej wydajności.
- Tworzenie i testowanie w środowiskach wirtualnych, a nie w wdrożeniach fizycznych. Większość aspektów AGI można w pełni ocenić bez narzutów (czasu, pieniędzy i złożoności) związanych z robotyką.

Poznanie na poziomie zwierzęcym a myśl abstrakcyjna, język i logika formalna Istnieje wiele dowodów na to, że osiągnięcie poznania na wysokim poziomie wymaga jedynie niewielkich ulepszeń strukturalnych w stosunku do możliwości zwierząt. Odkrycia w psychologii poznawczej wskazują, że uogólnione przetwarzanie wzorców jest podstawowym mechanizmem funkcjonowania wszystkich wyższych poziomów. Z drugiej strony, stosunkowo niewielkie różnice między wyższymi zwierzętami i ludźmi potwierdzają także badania genetyki, harmonogramu ewolucji i psychologii rozwojowej. Podstawowym wyzwaniem AGI jest osiągnięcie solidnej, adaptacyjnej zdolności uczenia się pojęciowego u wyższych naczelnych i małych dzieci. Jeśli celem jest inteligencja na poziomie ludzkim, wówczas podążanie za robotyką, językiem lub logiką formalną (na tym etapie) jest kosztownym elementem ubocznym – niezależnie od tego, czy jest to motywowane niezrozumieniem problemu, czy też względami komercyjnymi lub „politycznymi”.

Streszczenie

Chociaż nasz projekt w dużej mierze opiera się na badaniach prowadzonych w wielu wyspecjalizowanych dyscyplinach, jest to jedno z niewielu wysiłków poświęconych integracji takiej interdyscyplinarnej wiedzy z konkretnym celem, jakim jest rozwój ogólnej sztucznej inteligencji. Jesteśmy głęboko przekonani, że wiele kwestii poruszonych powyżej ma kluczowe znaczenie dla wczesnego osiągnięcia naprawdę inteligentnych systemów uczenia się adaptacyjnego.

Podstawowe zdolności poznawcze

Inteligencja ogólna wymaga szeregu podstawowych zdolności poznawczych. Przy pierwszym przybliżeniu musi być w stanie:

- zapamiętywać i rozpoznawać wzorce reprezentujące spójne cechy rzeczywistości;
- powiązać takie wzorce poprzez różne podobieństwa, różnice i skojarzenia;
- uczyć się i wykonywać różnorodne czynności;
- oceniać i kodować informacje zwrotne z systemu celów;
- samodzielnie dostosowuje parametry sterowania systemem;

Jak wspomniano wcześniej, funkcjonalność ta musi obsługiwać bardzo szeroką gamę typów i charakterystyk danych (w tym czasowych) i musi działać interaktywnie, w czasie rzeczywistym. Poniższy rozszerzony opis opiera się na naszej konkretnej implementacji; jednakże wymienione funkcje byłyby generalnie wymagane (w jakiejś formie) w każdej implementacji sztucznej inteligencji ogólnej. Uczenie się wzorców, dopasowywanie, uzupełnianie i przywoływanie wzorców Podstawowa metoda przyswajania wzorców polega na zastrzeżonej adaptacji leniwego uczenia się. Nasze implementacje przechowują wzorce (statyczne i dynamiczne) z adaptacyjnymi, rozmytymi tolerancjami, które następnie określają sposób przetwarzania podobnych wzorców. Nasz algorytm rozpoznawania dopasowuje wzorce na zasadzie konkurencyjnej „zwycięzca bierze wszystko” jako zbiór lub agregat

podobnych wzorców lub w drodze wymuszonego wyboru. Oferuje również nieodłączną obsługę uzupełniania wzorców i przywoływania (w stosownych przypadkach).

Gromadzenie i zapominanie danych

Ponieważ nasz system uczy się wzorców przyrostowo, potrzebne są mechanizmy konsolidacji i usuwania nadmiaru danych. Wykryte wzorce (lub wzorce podrzędne), które mieszczą się w dynamicznie ustawionej tolerancji szumu/błędu istniejących, są automatycznie konsolidowane przez mechanizm przypominający Hebb'a, który nazywamy „podsuwaniem”. Algorytm ten gromadzi również pewne informacje statystyczne. Z drugiej strony usuwane są wzorce, które okazują się nieistotne (według różnych kryteriów).

Kategoryzacja i grupowanie

Wzorce cech kodowane wektorowo są pozyskiwane w czasie rzeczywistym i przechowywane w wysoce adaptacyjnej strukturze sieci. To centralne, samoorganizujące się repozytorium automatycznie grupuje dane w hiperwymiarowej przestrzeni wektorowej. Zdolność naszego algorytmu dopasowywania do przywoływania wzorców według dowolnego wymiaru zapewnia nieodłączne wsparcie dla elastycznej, dynamicznej kategoryzacji. Dodatkowe mechanizmy kategoryzacji ułatwiają grupowanie wzorców według dodatkowych parametrów, powiązań lub funkcji.

Hierarchie i skojarzenia wzorców

Wzorce cech percepcyjnych nie są odosobnione – wywodzą się ze spójnej rzeczywistości zewnętrznej. Kodowanie relacji między wzorcami spełnia kluczowe funkcje dodawania znaczenia, kontekstu i przewidywania. Nasz system wychwytuje niskopoziomowe, oparte na percepcji skojarzenia wzorców, takie jak: sekwencyjne lub przypadkowe w czasie, pobliskie w przestrzeni, powiązane przez grupę cech lub modalność zmysłów. Dodatkowe relacje są kodowane na wyższych poziomach sieci, łącznie z warstwami uruchamiającymi. Ta ogólna struktura przypomina nieco „sieć dualną” opisaną przez.

Gruntowanie wzoru i rozprowadzanie aktywacyjne

Podstawową funkcją łączy asocjacyjnych jest łączenie węzłów powiązanych z prime4. Pomaga to ujednoznaczyć dopasowanie wzorca i wybrać alternatywy kontekstowe. W przypadku, gdy aktywacja jest szczególnie silna, a aktywność percepcyjna niska, zapisane wzorce zostaną „rozpoznane” spontanicznie. Zarówno zakres, jak i szybkość zaniku takiego rozprzestrzeniania się aktywacji są kontrolowane adaptacyjnie. Dynamika ta łączy się z pierwotną aktywacją opartą na percepcji, tworząc pamięć krótkotrwałą systemu.

Wzory działania

Obwody o działaniu adaptacyjnym służą do sterowania parametrami w trzech następujących obszarach:

1. zmysły, w tym regulowane ekstraktory cech, mechanizmy skupiania i selekcji;
2. siłowniki wyjściowe do nawigacji i manipulacji;
3. metapoznanie i kontrola wewnętrzna;

Różne stany i zachowania akcji (sekwencje akcji) dla każdego z tych wyjść sterujących można utworzyć w czasie projektowania (za pomocą skryptu konfiguracyjnego) lub uzyskać interaktywnie. Uczenie się w czasie rzeczywistym odbywa się albo poprzez bezpośrednie nauczanie, albo autonomicznie, poprzez losową eksplorację. Po nabyciu działania te można powiązać z konkretnymi bodźcami percepcyjnymi

lub całymi kontekstami poprzez różne mechanizmy reakcji na bodziec. Te łączy S-R (zarówno aktywacyjne, jak i hamowanie) są dynamicznie modyfikowane poprzez ciągłe uczenie się przez wzmacnianie.

Kontrola metapoznawcza

Oprócz adaptacyjnej percepcji i funkcjonalności działania, projekt AGI musi także umożliwiać szerokie monitorowanie i kontrolę ogólnych parametrów i funkcji systemu (w tym poznawczych strategii behawioralnych przypominających emocje). Każdy złożony interaktywny system uczenia się zawiera wiele kluczowych parametrów kontrolnych, takich jak tolerancja na hałas, tempo uczenia się i eksploracji, priorytety i zarządzanie celami, oraz a mnóstwo innych. System musi nie tylko być w stanie adaptacyjnie kontrolować wiele interaktywnych wektorów, ale musi także odpowiednio zarządzać swoimi różnymi funkcjami poznawczymi (takimi jak rozpoznawanie, przypominanie, działanie itp.) i trybami (takimi jak eksploracja, ostrożność, uwaga itp.) , dynamicznie oceniając je pod kątem skuteczności. Nasz projekt radzi sobie z tymi wymaganiami za pomocą wysoce adaptacyjnej „sondy” do introspekcji/kontroli.

Inteligencja na wysokim poziomie

Nasz model AGI zakłada, że do poznania wyższego poziomu nie są potrzebne żadne dodatkowe podstawowe funkcje. Myślenie abstrakcyjne, język i myślenie logiczne są rozwinięciami podstawowych umiejętności. Ta kontrowersyjna kwestia zostanie omówiona w dalszej części.

AGI w przygotowaniu

Funkcjonalny prototyp sprawdzający koncepcję jest obecnie w fazie rozwoju w Adaptive A.I. Inc. ma na celu ucieleśnienie wszystkich wyżej wymienionych wyborów, wymagań i funkcji. Nasza ścieżka rozwoju była następująca:

1. ramy rozwoju;
2. rdzeń pamięci i struktura interfejsu;
3. indywidualne podstawowe komponenty poznawcze;
4. zintegrowane poznanie niskiego poziomu;
5. zwiększony poziom funkcjonalności (na czym się obecnie skupiamy).

Oprogramowanie składa się ze szkieletu silnika AGI z następującymi podstawowymi komponentami:

- zestaw wtykowych, programowalnych (wirtualnych) czujników i elementów wykonawczych (zwanymi „sondami”);
- centralny magazyn/silnik wzorców zawierający wszystkie algorytmy danych i poznawcze;
- konfigurowalny, dynamiczny świat wirtualny 2D oraz różne narzędzia szkoleniowe i diagnostyczne.

Konstrukcja silnika AGI opiera się na spostrzeżeniach z szerokiego zakresu badań z zakresu nauk kognitywnych – w tym informatyki, neuronauki, epistemologii , nauk o kognitywistyce i psychologii – i uwzględnia wnioski z nich. Szczególnie silne wpływy obejmują: systemy ucieleśnione , reprezentację zakodowaną wektorowo , adaptacyjne, samoorganizujące się sieci neuronowe (zwłaszcza rosnący gaz neuronowy), uczenie się bez nadzoru i samonadzoru, uczenie się percepcyjne, i logika rozmyta. Chociaż nasz projekt obejmuje kilka nowatorskich i zastrzeżonych algorytmów, naszą kluczową innowacją jest szczególnie wybór i integracja ustalonych technologii i wcześniejszych spostrzeżeń.

Architektura i cechy konstrukcyjne silnika AGI

Nasz silnik AGI (który zapewnia tę podstawową zdolność poznawczą) można logicznie podzielić na trzy części:

- Rdzeń poznawczy;
- logika sterowania/interfejsu;
- sondy wejściowe/wyjściowe.

Ta „architektura agenta umiejscowionego” odzwierciedla znaczenie posiadania systemu AGI, który może dynamicznie i adaptacyjnie współdziałać z otoczeniem. Z perspektywy teorii umysłu uznaje zarówno kluczową potrzebę ugruntowania pojęć (za pośrednictwem zmysłów), jak i absolutną potrzebę uczenia się przez doświadczenie i samokontrolę. Wymienione poniżej komponenty zostały specjalnie zaprojektowane z uwzględnieniem funkcji wymaganych do adaptacyjnej inteligencji ogólnej w (ostatecznie) rzeczywistych środowiskach. Między innymi radzą sobie z dużą różnorodnością i ilością danych statycznych i dynamicznych, radzą sobie z niejasnymi i niepewnymi danymi oraz celami, sprzyjają spójnym, zintegrowanym reprezentacjom rzeczywistości, a przede wszystkim promują adaptacyjność.

Rdzeń poznawczy

Jest to centralne repozytorium wszystkich statycznych i dynamicznych wzorców danych – w tym wszystkich wyuczonych stanów poznawczych i behawioralnych, skojarzeń i sekwencji. Wszystkie dane są przechowywane w jednej, zintegrowanej strukturze węzłów i łączy. W projekcie wprowadzono innowacje w zakresie specyficznego kodowania „rozmycia” wzoru (oprócz innych atrybutów). Rdzeń pozwala na kilka typów węzłów/połączeń o różnej dynamice, co pomaga zdefiniować strukturę poznawczą sieci. Topologia sieci jest dynamicznie samoorganizująca się – cecha zainspirowana projektem „Growing Neural Gas” [10]. Umożliwia to dostosowanie gęstości sieci do rzeczywistych wymagań dotyczących danych i/lub celów. Różne adaptacyjne parametry lokalne i globalne dodatkowo definiują strukturę i dynamikę sieci w czasie rzeczywistym.

Logika sterowania i interfejsu

Ogólny system kontroli koordynuje cykl wykonawczy sieci, steruje różnymi algorytmami poznawczymi i porządkowymi oraz kontroluje/dostosowuje parametry systemu. Za pośrednictwem menedżera interfejsu przekazuje również dane i informacje sterujące do i z sond.

Oprócz obsługi „nakrętek i śrub” związanych z wykonaniem programu i komunikacji oraz zarządzania różnymi algorytmami poznawczymi, system kontroli obejmuje także monitorowanie i kontrolę metakognitywną. Jest to zasadniczo poznawczy aspekt emocji; takie stany jak ciekawość, nuda, przyjemność, rozczarowanie itp.

Sondy

Menedżer interfejsu umożliwia dynamiczne dodawanie i konfigurację sond. Kluczowe cechy konstrukcyjne architektury sondy obejmują możliwość posiadania programowalnych ekstraktorów cech, zmienną rozdzielczość danych oraz mechanizmy fokusu i selekcji. Takie mechanizmy selekcji danych są niezbędne dla ogólnej inteligencji: nawet umiarkowanie złożone środowiska zawierają bogactwo danych, które znacznie przekracza zdolność dowolnego systemu do użytecznego przetwarzania. System obsługuje bardzo szeroką gamę typów danych i steruje wymaganiami dotyczącymi sygnałów – w tym dla danych wizualnych, dźwiękowych i surowych (np. Baza danych,

Internet, klawiatura), a także różnymi elementami wykonawczymi. Nowatorska „sonda systemowa” zapewnia systemowi monitorowanie i kontrolę jego stanów wewnętrznych (forma metapoznania). Dodatkowe sondy – albo niestandardowe interfejsy z innymi systemami, albo dodatkowe rzeczywiste czujniki/elementy wykonawcze – można łatwo dodać do systemu.

Środowisko programistyczne, język i sprzęt

Kompletny silnik AGI wraz z powiązаныmi programami wsparcia są zaimplementowane w (obiektywnym) języku C# w ramach platformy .NET firmy Microsoft. Obecne testy pokazują, że praktyczną wydajność prototypu (weryfikacja koncepcji) można osiągnąć na jednym, konwencjonalnym komputerze PC (2 GHz, 512 MB). Nawet implementacja, która nie jest dostosowana do wydajności, może przetwarzać kilka złożonych wzorców na sekundę w bazie danych zawierającej setki tysięcy przechowywanych funkcji.

Od algorytmów do inteligencji ogólnej

W tej sekcji omówiono niektóre z naszych krótkoterminowych badań i rozwoju; ma na celu zilustrowanie naszej oczekiwanej ścieżki w kierunku znaczącej inteligencji ogólnej. Chociaż praca ta ledwo zbliża się do wyższego poziomu poznania zwierząt (przekracza go w niektórych aspektach, ale znacznie odbiega od innych, takich jak umiejętności czuciowo-motoryczne), uważamy, że jest to kluczowy krok w udowodnieniu ważności i praktyczności naszego modelu. Co więcej, rzeczywista osiągnięta funkcjonalność powinna być wysoce konkurencyjna, jeśli nie wyjątkowa, w zastosowaniach, w których głównymi wymaganiami są znaczna autonomiczna adaptacyjność i selekcja danych, brak kruchości, dynamiczne przetwarzanie wzorców, elastyczne uruchamianie i samonadzorowane uczenie się. Ogólna inteligencja nie obejmuje jednego, genialnego wynalazku lub cechy konstrukcyjnej; zamiast tego wyłania się z synergicznej integracji szeregu podstawowych składników. Od strony strukturalnej system musi integrować bodźce zmysłowe, pamięć i siłowniki, natomiast od strony funkcjonalnej różne możliwości uczenia się, rozpoznawania, przypominania i działania muszą płynnie działać w szerokim zakresie wzorców statycznych i dynamicznych. Ponadto te zdolności poznawcze muszą mieć charakter pojęciowy i kontekstowy – muszą umieć uogólniać wiedzę i interpretować ją na różnych płaszczyznach. Kluczowy kamień milowy w naszym projekcie został osiągnięty, gdy rozpoczęliśmy testowanie zintegrowanej funkcjonalności podstawowych komponentów poznawczych w ramach naszego ogólnego szkieletu AGI. Do testowania spójnego funkcjonowania całego systemu wykorzystuje się szereg specjalnie opracowanych, wysoce konfigurowalnych narzędzi testowych. Na tym etapie większość prac rozwojowych/testów AGI odbywa się przy użyciu naszych środowisk wirtualnych, sterowanych przez niestandardowe skrypty. Uzupełnieniem tego zautomatyzowanego szkolenia i oceny są ręczne eksperymenty w wielu różnych środowiskach i zastosowaniach. Doświadczenie zdobyte w wyniku tych testów pomaga udoskonalić złożoną dynamikę oddziałujących algorytmów i parametrów. Jedną z ogólnych trudności związanych z rozwojem AGI jest określenie bezwzględnych miar sukcesu. Jednym z powodów jest to, że dziedzina ta wciąż dopiero się rodzi, w związku z czym nie ma uzgodnionych definicji, nie mówiąc już o testach lub miarach ogólnej inteligencji niskiego poziomu. W miarę realizacji naszego projektu spodziewamy się opracowania coraz skuteczniejszych protokołów i wskaźników oceny zdolności poznawczych. Ocena wydajności naszego systemu opiera się na następującym opisie: Inteligencja ogólna obejmuje zdolność do zdobywania (i dostosowywania) wiedzy i umiejętności wymaganych do osiągnięcia szerokiego zakresu celów w różnorodności domen. W tym kontekście:

- Akwizycja obejmuje wszystkie poniższe czynności: automatyczne, poprzez dane wejściowe (na podstawie funkcji/danych); wyraźnie nauczone; odkryte w drodze eksploracji lub eksperymentów; procesy wewnętrzne (np. stowarzyszenie, kategoryzacja, statystyki itp.).

- Adaptacja oznacza, że nowa wiedza jest odpowiednio zintegrowana.
- Wiedza i umiejętności odnoszą się do wszelkiego rodzaju danych i zdolności (stanów i zachowań), które system nabywa w krótkim lub długim okresie.

Nasz początkowy protokół oceny AGI ma na celu uwzględnienie szerokiego spektrum domen i celów poprzez symulację przykładowych aplikacji w wirtualnych światach 2D. W szczególności testy te powinny ocenić stopień, w jakim podstawowe umiejętności działają jako zintegrowana, wzajemnie wspierająca się całość – i to bez interwencji programisty! Poniżej znajdują się trzy przykłady.

Przykładowe domeny testowe dla początkowych kryteriów wydajności

Adaptacyjny monitor bezpieczeństwa

System ten skanuje monitory wideo i panele alarmowe nadzorujące bezpieczny obszar (powiedzmy fabrykę, budynek biurowy itp.) i odpowiednio reaguje na nietypowe warunki. Należy zauważyć, że jest to nieco podobne do aplikacji do monitorowania lokalizacji w MIT. Ta symulacja wymaga środowiska wizualnego, które zawiera wiele szczegółów, ale ma jedynie ograniczoną aktywność dynamiczną – jest to jego normalny stan (zielony). Istnieją dwa poziomy nieprawidłowości: (i) niewielkie lub znane zakłócenia (żółty); (ii) poważne lub nieznanne zakłócenia (kolor czerwony). System musi początkowo nauczyć się normalnego stanu poprzez proste naświetlenie (automatyczne skanowanie otoczenia) w różnych rozdzielczościach (szczęgoty). Musi także nauczyć się warunków „żółtych”, pokazując mu kilka próbek (niektóre w wysokiej rozdzielczości). Wszystkie pozostałe stany muszą wyświetlać wartość „czerwoną”. Standardowa operacja polega na ciągłym skanowaniu otoczenia w niskiej rozdzielczości. W przypadku wykrycia jakichkolwiek nietypowych warunków system musi nauczyć się przełączać na wyższą rozdzielczość, aby rozróżnić kolor „żółty” od „czerwonego”. System musi dostosować się do zmian w środowisku (i zupełnie innych środowiskach) poprzez proste szkolenie w zakresie ekspozycji.

Asystent wzroku

System steruje ruchomym „okiem” (za pomocą poleceń głosowych), które umożliwia identyfikację (za pomocą głosu) co najmniej stu różnych obiektów na świecie. Trener będzie dynamicznie uczył system nowych nazw, skojarzeń i poleceń związanych z ruchem oczu. Sonda wizualna może wybierać spośród różnych scen (symulujących pomieszczenia) i skupiać się na różnych częściach każdej sceny. Sceny przedstawiają obiekty o różnych atrybutach: kolorze, rozmiarze, kształcie, różnej dynamice itp. (i ich kombinacje) na różnym tle. Początkowe szkolenie będzie polegało na dołączeniu prostych poleceń dźwiękowych umożliwiających manewrowanie „okiem” oraz skojarzeniu etykiet słownych z wybranymi obiektami. System musi wtedy niezawodnie wykonywać polecenia głosowe i odpowiadać odpowiednią identyfikacją (jeśli występuje). Dodatkową funkcjonalnością może być skanowanie przez system różnych scen w stanie bezczynności i automatyczne raportowanie wybranych ważnych obiektów. Identyfikacja obiektu musi obejmować szerokie spektrum różnych kombinacji atrybutów i tolerancji. System musi łatwo uczyć się nowych scen, obiektów, słów i skojarzeń, a także dostosowywać się do zmian którejkolwiek z tych zmiennych.

Odkrywca labiryntu

Podmiot (wirtualny) bada średnio złożone środowisko. Odkrywa, jakie typy obiektów pomagają, a jakie utrudniają jego cele, jednocześnie ucząc się poruszać po tym dynamicznym świecie. Można go również wyszkolić do wykonywania określonych zachowań. Wirtualny świat jest wypełniony dużą liczbą różnych obiektów (patrz poprzedni przykład). Ponadto niektóre z tych obiektów poruszają się w przestrzeni z różną prędkością i dynamiką i mogą być stałe i/lub nieruchome. Grupy różnych rodzajów obiektów mają wstępnie przypisane atrybuty, które wskazują negatywnie lub pozytywnie. Silnik AGI

kontroluje kierunek i prędkość jednostki w tym wirtualnym świecie. Jego celem jest nauka poruszania się po obiektach nieruchomych i negatywnych, aby wiarygodnie dotrzeć do ukrytych pozytywów. System można także wytrenować w zakresie reagowania na polecenia operatora w celu wykonywania zachowań o różnym stopniu złożoności (na przykład czynności przypominających „sztuczki”, których można nauczyć psa). Ten „Eksplorator labiryntu” można łatwo skonfigurować do wykonywania dość skomplikowanych zadań.

W stronę zwiększonej inteligencji

Oczywiście opisane powyżej zadania same w sobie nie stanowią przełomu w badaniach nad sztuczną inteligencją. Zostały już osiągnięte wielokrotnie. Jednak to, co uważamy za znaczące i wyjątkowe, to osiągnięcie tych różnych zadań bez programowania lub parametryzacji specyficznego dla zadania. Nie chodzi o to, co się robi, ale jak to się robi. Rozwój wykraczający poza te podstawowe testy sprawdzające słuszność koncepcji będzie postępował w dwóch kierunkach: (1) w celu znacznego zwiększenia rozdzielczości, objętości danych i złożoności w aplikacjach podobnych do testów; (2), aby dodać funkcjonalność wyższego poziomu. Oprócz prac mających na celu dalszy rozwój i sprawdzanie naszego ogólnego modelu inteligencji, można wprowadzić także wiele praktycznych ulepszeń. Obejmowałyby one wdrożenie wersji wieloprotocessorowych i sieciowych oraz integrację naszego systemu z bazami danych lub z innymi istniejącymi technologiami sztucznej inteligencji, takimi jak systemy eksperckie, rozpoznawanie głosu, robotyka czy moduły sensoryczne ze specjalistycznymi ekstraktorami cech. Zdecydowanie najważniejszy z tych przyszłych zmian dotyczy umiejętności na wyższym poziomie. Oto częściowa lista elementów działania, z których wszystkie wywodzą się z podstaw niższego poziomu:

- rozpowszechnianie aktywacji i zachowywanie kontekstu przez dłuższy okres;
- wspierać bardziej złożone wewnętrzne wzorce czasowe, zarówno w celu lepszego rozpoznawania i przewidywania, jak i sekwencji poznawczych i działań;
- wewnętrzne sprzężenie zwrotne aktywacji do przetwarzania bez wprowadzania danych;
- dedukcja, osiągana poprzez selektywną aktywację koncepcji;
- zaawansowana kategoryzacja według dowolnych wymiarów;
- uczenie się bardziej złożonych zachowań;
- tworzenie koncepcji abstrakcyjnych i scalonych;
- ustrukturyzowane przyswajanie języka;
- zwiększona świadomość i kontrola stanów wewnętrznych (introspekcja);
- Uczenie się logiki i innych metod rozwiązywania problemów.

Inne badania

Istnieje wiele różnych podejść do sztucznej inteligencji; niektóre różnice są proste, inne zaś subtelne i opierają się na trudnych kwestiach filozoficznych. W związku z tym dokładne umiejscowienie naszej pracy w porównaniu z pracą innych jest trudne i w istocie otwarte na debatę. Nasz pogląd, że „inteligencja jest właściwością podmiotu, który wchodzi w dwukierunkową interakcję ze środowiskiem zewnętrznym”, technicznie rzecz biorąc, plasuje nas w obszarze „systemów agentowych” [27]. Jednak nasz nacisk na koneksjonistyczne, a nie klasyczne podejście do modelowania poznawczego, umieszcza naszą pracę w obszarze „kognitywistyki ucieleśnionej”. Chociaż nasze podejście jest podobne do innych badań w zakresie kognitywistyki ucieleśnionej, pod pewnymi względami nasze cele są zasadniczo różne. Kluczową różnicą jest nasze przekonanie, że podstawowy zestaw zdolności

poznawczych współpracujący ze sobą jest wystarczający do wytworzenia ogólnej inteligencji. Stanowi to wyraźny kontrast w stosunku do innych badaczy kognitywistyki ucieleśnionej, którzy uważają, że inteligencja jest z konieczności specyficzna dla zbioru problemów w danym środowisku. Innymi słowy, wierzą, że autonomiczni agenci zawsze istnieją w niszach ekologicznych. W związku z tym koncentrują swoje badania na budowaniu bardzo ograniczonych systemów, które skutecznie radzą sobie z niewielką liczbą problemów w określonym ograniczonym środowisku. Prawie wszystkie prace w tej dziedzinie są zgodne z tą zasadą, aby zapoznać się z kilkoma dobrze znanymi przykładami. Ich stanowisko zaprzecza faktowi, że ludzie posiadają ogólną inteligencję; potrafimy skutecznie radzić sobie z szeroką gamą problemów, które znacznie wykraczają poza wszystko, co można nazwać naszą „niszą ekologiczną”. Być może najbliższym naszym projektem, ściśle związanym z obszarem kognitywistyki ucieleśnionej, jest projekt Cog na MIT. Celem projektu jest zrozumienie dynamiki interakcji międzyludzkich poprzez konstrukcję przypominającego człowieka robota, wyposażonego w górną część tułowia, głowę, oczy, ramiona i dłonie. Chociaż projekt ten jest znacznie bardziej ambitny niż inne projekty pod względem poziomu i złożoności dynamiki i możliwości systemu, system nadal jest zasadniczo skoncentrowany na niszy (elementarne interakcje społeczne i fizyczne między ludźmi) w porównaniu z naszymi własnymi wysiłkami w zakresie ogólnej inteligencji. Prawdopodobnie najbliższym naszym dziełem w tym sensie, że ma również na celu osiągnięcie inteligencji ogólnej, a nie niszowej, jest projekt Novamente pod kierownictwem Bena Goertzela. (Projekt był wcześniej znany jako Webmind) Novamente opiera się na hybrydzie niskopoziomowej dynamiki podobnej do sieci neuronowej do rozprzestrzeniania aktywacji i torowania koncepcji, w połączeniu z konstrukcjami semantycznymi wysokiego poziomu reprezentującymi różnorodne relacje logiczne, przyczynowe i przestrzenno-czasowe. Chociaż semantyka stanu wewnętrznego systemu jest stosunkowo łatwa do zrozumienia w porównaniu z podejściem ściśle koneksjonistycznym, klasyczne elementy projektu systemu otwierają drzwi do wielu podstawowych problemów, które nękały klasyczną sztuczną inteligencję przez ostatnie pięćdziesiąt lat. Na przykład semantyka wysokiego poziomu wymaga złożonej metalogiki zawartej w zakodowanym na stałe rozumowaniu wysokiego poziomu i innych systemach poznawczych wysokiego poziomu. Te systemy wysokiego poziomu zawierają znaczącą, ukrytą semantykę, która może nie być osadzona w interakcji ze środowiskiem, ale jest raczej zakodowana na stałe przez projektanta, co powoduje problemy z uziemieniem symboli. Stosunkowo stałe, wysokopoziomowe metody reprezentacji i manipulacji wiedzą, jakie pociąga za sobą to podejście, są również podatne na problemy z „układem odniesienia” i „kruchością”. Jak przyjęliśmy, w ściśle ucieleśnionym podejściu do nauk kognitywnych cała wiedza wywodzi się z interakcji agent-środowisko, unikając w ten sposób długotrwałych problemów klasycznej sztucznej inteligencji. Projekt MAGNUS Igora Aleksandra również uwzględniał wiele kluczowych koncepcji AGI, które zidentyfikowaliśmy, ale został poważnie ograniczony przez klasyczne podejście do sztucznej inteligencji i maszyny skończonej. Valeriy Nenov i Michael Dyer z UCLA wykorzystali „masowo” równoległy sprzęt (maszynę połączeniową CM-2) do wdrożenia wirtualnego, interaktywnego projektu percepcyjnego zbliżonego do naszego, ale o sztywniejszej, wstępnie zaprogramowanej strukturze. Niestety, od tego ambitnego i przełomowego dzieła porzucono. Projekt był prawdopodobnie poważnie utrudniony przez ograniczony (wówczas) sprzęt. Oddalając się od kognitywistyki ucieleśnionej w kierunku czysto klasycznych badań nad inteligencją ogólną, być może najbardziej znanym systemem jest projekt Cyc, nad którym pracuje. Zasadniczo Lenat postrzega inteligencję ogólną jako „zdrowy rozsądek”. Ma nadzieję osiągnąć ten cel poprzez dodanie wielu milionów faktów o świecie do ogromnej bazy danych. Po wielu latach pracy i milionach dolarów w funduszach nadal jest przed nami długa droga, ponieważ sama liczba faktów, które ludzie znają na temat świata, jest naprawdę zdumiewająca. Wątpimy, czy bardzo duża baza danych podstawowych faktów wystarczy, aby dać komputerowi dużą inteligencję ogólną – brakuje mechanizmów autonomicznego zdobywania wiedzy. Będąc klasycznym podejściem do sztucznej inteligencji, wiąże się to również z podstawowymi problemami klasycznej sztucznej

inteligencji wymienionymi powyżej. Na przykład ponownie pojawia się problem uziemienia symboli: jeśli fakty dotyczące kotów i psów zostaną po prostu dodane do bazy danych, z której komputer może korzystać, nawet jeśli nigdy nie widział zwierzęcia ani nie miał z nim kontaktu, czy te koncepcje naprawdę mają znaczenie dla systemu? Chociaż jego projekt również twierdzi, że dąży do „inteligencji ogólnej”, w rzeczywistości bardzo różni się od naszego, zarówno pod względem podejścia, jak i trudności, jakie napotyka. Analiza ciągłych niepowodzeń sztucznej inteligencji w przewyżczeniu jej długotrwałych ograniczeń pokazuje, że nie chodzi o to, że wypróbowano sztuczną inteligencję ogólną i zakończyło się to niepowodzeniem, ale raczej o to, że dziedzina ta została w dużej mierze porzucona – czy to ze względów teoretycznych, historycznych czy powodów komercyjnych. Z pewnością niewiele uwagi poświęca się naszemu szczególnemu podejściu, opisanemu szczegółowo w poprzednich sekcjach.

Szybka ścieżka AGI: dlaczego tak rzadko?

Powszechne zastosowanie sztucznej inteligencji utrudnia szereg podstawowych ograniczeń, które nękały tę dziedzinę od samego początku, a mianowicie:

- koszty i opóźnienia związane z programowaniem indywidualnych aplikacji na zamówienie;
- niezdolność systemów do automatycznego uczenia się na podstawie doświadczenia lub możliwości uczenia się/szkolenia przez użytkownika;
- problemy z niezawodnością i wydajnością spowodowane „kruchością” (niezdolność systemów do automatycznego dostosowywania się do zmieniających się wymagań lub danych spoza wcześniej określonego zakresu);
- ich ograniczona inteligencja i zdrowy rozsądek.

Najbardziej bezpośrednią drogą do rozwiązania tych długotrwałych problemów jest koncepcyjne zidentyfikowanie podstawowych cech wspólnych całej inteligencji wysokiego poziomu oraz zaprojektowanie systemów wyposażonych w tę podstawową funkcjonalność w sposób wykorzystujący siłę ludzką i technologiczną. Ogólna inteligencja jest kluczem do stworzenia solidnych systemów autonomicznych, które potrafią się uczyć i dostosowywać do szerokiego zakresu zastosowań. Jest to także kamień węgielny samodoskonalenia, czyli Seed AI – wykorzystywania podstawowych umiejętności do ładowania umiejętności wyższego poziomu. W tym rozdziale zidentyfikowano podstawowe elementy ogólnej inteligencji, a także kluczowe kwestie dotyczące skutecznego rozwoju sztucznej odmiany. Podkreślono fakt, że bardzo niewielu badaczy faktycznie podąża tą najbardziej bezpośrednią drogą do AGI. Jeżeli podejście opisane powyżej jest tak obiecujące, to dlaczego poświęcono mu tak mało uwagi? Dlaczego prawie nikt nad tym nie pracuje? Krótka odpowiedź: Spośród wszystkich osób pracujących w dziedzinie zwanej AI:

- 80% nie wierzy w koncepcję inteligencji ogólnej (zamiast tego w duży zbiór konkretnych umiejętności i wiedzy).
- Spośród tych, którzy to robią, 80% nie wierzy, że sztuczna inteligencja na poziomie ludzkim jest możliwa – kiedykolwiek lub przez długi, długi czas.
- Spośród tych, które to robią, 80% pracuje nad projektami AI dotyczącymi konkretnej dziedziny ze względów komercyjnych lub akademicko-politycznych (rezultaty są bardziej natychmiastowe).
- Spośród pozostałych 80% ma słabe ramy koncepcyjne...

Choć powyższe jest karykaturą, zawiera w sobie więcej niż ziarno prawdy. Ogromna liczba badaczy odrzuca ważność i znaczenie „inteligencji ogólnej”. Dla wielu kontrowersje w psychologii (takie jak te wywołane przez The Bell Curve) sprawiają, że jest to temat niepopularny, jeśli nie tabu. Inni, uwarunkowani dziesięcioleciami pracy dziedzinowej, po prostu nie dostrzegają korzyści wynikających z Seed AI – jednorazowego rozwiązania problemów. Spośród tych, którzy w zasadzie nie sprzeciwiają się ogólnej inteligencji, wielu nie wierzy, że AGI jest możliwe – za ich życia lub kiedykolwiek. Niektórzy zajmują to stanowisko, ponieważ sami próbowali i ponieśli porażkę „w młodości”. Inni uważają, że AGI nie jest najlepszym podejściem do osiągnięcia „AI” lub zupełnie nie wiedzą, jak to zrobić aby się tym zająć. Bardzo niewielu badaczy faktycznie badało ten problem z naszej perspektywy (ogólnej inteligencji/nasionowej sztucznej inteligencji). Niektórzy faktycznie próbują dokonać inżynierii wstecznej mózgu – jedna funkcja na raz. Są też tacy, którzy mają zastrzeżenia moralne lub się tego boją. Oczywiście wielu jest tak skupionych na konkretnych, wąskich aspektach inteligencji, że po prostu nie skupiają się na patrzeniu na szerszą perspektywę – zostawiają to innym, aby to urzeczywistniło. Należy również zauważyć, że często istnieje silna presja finansowa i instytucjonalna, aby rozwijać wyspecjalizowaną sztuczną inteligencję. Wszystko to razem tworzy dynamikę, w której prawdziwa sztuczna inteligencja nie jest „modna” – nie cieszy się szacunkiem, finansowaniem i wsparciem – co jeszcze bardziej zmniejsza liczbę przyciąganych do niej osób! To powinno być więcej niż wystarczające powody, aby wyjaśnić brak postępu w AGI. Ale jest coraz gorzej. Badacze faktycznie próbujący zbudować systemy AGI są dodatkowo utrudniani przez niezliczoną ilość błędnych przekonań, złych wyborów i braku zasobów (funduszy i badań). Wiele kwestii technicznych omówiono już wcześniej (patrz sekcje 3 i 7), ale warto wspomnieć o kilku innych:

Epistemologia

Modele AGI mogą być tak dobre, jak ich podstawowa teoria wiedzy – natura wiedzy i jej związek z rzeczywistością. Świadomość, że inteligencja wysokiego poziomu opiera się na koncepcyjnej reprezentacji rzeczywistości, leży u podstaw decyzji projektowych, takich jak adaptacyjne kodowanie wektorów rozmytych i interaktywne podejście ucieleśnione. Inne konsekwencje to potrzeba skupienia się i selekcji opartej na zmysłach oraz aktywacja kontekstowa. Centralne znaczenie wysoce zintegrowanej sieci wzorców – szczególnie włączając sieci dynamiczne – staje się oczywiste po zrozumieniu relacji pomiędzy bytami, atrybutami, koncepcjami, działaniami i myślami. Te i kilka innych spostrzeżeń stanowią podstawę do rozwiązywania problemów związanych z uziemieniem, kruchością i zdrowym rozsądkiem. Wreszcie, nadal istnieje wiele niepotrzebnego zamieszania dotyczącego relacji między pojęciami i symbolami. Dynamiką, która w dalszym ciągu utrudnia sztuczną inteligencję, jest utrzymująca się schizma między tradycjonalistami a koneksjonistami. To niestety pomaga utrwalić fałszywą dychotomię pomiędzy wyrazistymi symbolami/schematami i niezrozumiałe wzorce.

Teoria umysłu

Kolejnym problemem jest niechlujne sformułowanie i słabe zrozumienie kilku kluczowych pojęć: świadomości, inteligencji, woli, znaczenia, emocji, zdrowego rozsądku i „quali”. Fakt, że setki badaczy sztucznej inteligencji uczestniczy co roku w konferencjach, podczas których główni prelegenci ogłaszają, że „nie rozumiemy świadomości (ani qualiów czy czegośkolwiek innego) i prawdopodobnie nigdy tego nie rozumiemy”, wskazuje, jak powszechny jest ten problem. Charakterystyka świadomości jako „słowa walizkowego” dokonana przez Marviną Minsky’ego jest poprawna. Po prostu się rozpakujemy

To! Tego typu błędy często powodują, że badania przebiegają stycznie w stosunku do ustalonych celów długoterminowych. Dwa przykłady to nadmierny nacisk na biologiczną wykonalność i przekonanie, że

ucieleśniona inteligencja nie może być wirtualna i że należy ją zaimplementować w fizycznych robotach.

Psychologia poznawcza

Jest rzeczą oczywistą, że właściwe zrozumienie pojęcia „inteligencja” jest kluczem do jej opracowania. Oprócz epistemologii, kilka dziedzin psychologii poznawczej ma kluczowe znaczenie dla odkrycia jej znaczenia. Niezrozumienie inteligencji doprowadziło do kosztownych rozczarowań, takich jak ręczne gromadzenie ogromnych ilości w dużej mierze bezużytecznych danych (wiedzy bez znaczenia), wysiłki mające na celu osiągnięcie inteligencji poprzez łączenie mas głupich agentów lub próby uzyskania znaczącej rozmowy z izolowanej sieci symboli.

Koncentracja na projekcie

Nieliczne projekty, które faktycznie realizują AGI w oparciu o stosunkowo solidne modele, niosą ze sobą jeszcze jedno ryzyko: łatwo mogą stracić koncentrację. Czasami względy komercyjne wpływają na kierunek projektu, podczas gdy inne są spychane na boczny tor przez (stosunkowo) nieistotne kwestie techniczne, takie jak próba osiągnięcia nierealistycznie wysokiego poziomu wydajności, skupienie się na biologicznej wykonalności projektu lub próba wdrożenia funkcji wysokiego poziomu przed czasem. Jasno wytyczona ścieżka rozwoju prowadząca do inteligencji na poziomie ludzkim może służyć jako potężne antidotum na utratę z oczu „szerszego obrazu”. Wizja tego, jak dostać się „stąd” do „tam”, również pomaga utrzymać motywację w tak trudnym przedsięwzięciu.

Wsparcie badawcze

AGI wykorzystuje, a dokładniej, integruje dużą liczbę istniejących technologii AI. Niestety, wiele najważniejszych obszarów jest niestety niedostatecznie zbadanych. Należą do nich:

- uczenie się przyrostowe w czasie rzeczywistym, bez nadzoru/samonadzoru (w porównaniu z propagacją wsteczną);
- zintegrowane wsparcie dla wzorców czasowych;
- dynamicznie adaptacyjne topologie sieci neuronowych;
- samodostrajanie parametrów systemu, integrujące autoadaptację oddolną (opartą na danych) i odgórną (opartą na celu/metapoznaniu);
- sondy wykrywające z autoadaptacyjnymi ekstraktorami cech.

Naturalnie, te same ograniczenia powodują zmniejszenie wsparcia dla badań AGI.

Koszt i trudność

Osiągnięcie wysokiego poziomu AGI będzie trudne. Nie będzie to jednak tak trudne, jak sądzi większość ekspertów. Kluczowym elementem teorii „prawdziwej sztucznej inteligencji” (i jej wdrożenia) jest skupienie się na zasadniczych elementach inteligencji. Załączkowa sztuczna inteligencja staje się problemem możliwym do rozwiązania – pod pewnymi względami znacznie prostszym niż inne główne cele sztucznej inteligencji – poprzez wyeliminowanie ogromnych obszarów o trudnej, ale nieistotnej złożoności sztucznej inteligencji. Kiedy już uruchomimy kluczową podstawową funkcjonalność, większość wymaganej dodatkowej „inteligencji” (umiejętności) zostanie nauczona lub wyuczona, nie zaprogramowana. Powiedziawszy to, jestem przekonany, że skalowanie systemu do poziomu ludzkiego możliwości przechowywania i przetwarzania będzie wymagało bardzo znacznych zasobów. Jednakże

znacznie bardziej umiarkowane początkowe prototypy postępują jako dowód słuszności koncepcji dla AGI, potencjalnie zapoczątkowując dużą liczbę nowych, praktycznych zastosowań.

Wniosek

Zrozumienie ogólnej inteligencji i identyfikacja jej podstawowych komponentów jest kluczem do tworzenia systemów AI nowej generacji – systemów, które są znacznie tańsze, a jednocześnie znacznie wydajniejsze. Oprócz skupienia się na ogólnych zdolnościach uczenia się, podejście przyspieszone powinno również szukać ścieżki najmniejszego oporu – takiej, która wykorzystuje mocne strony inżynierii ludzkiej i dostępną technologię. Czasami wiąże się to z wybraniem rzadziej uczęszczanej drogi AI. Wierzę, że opisany powyżej model teoretyczny, komponenty poznawcze i ramy, w połączeniu z innymi moimi strategicznymi decyzjami projektowymi, stanowią solidną podstawę do osiągnięcia praktycznych możliwości AGI w dającej się przewidzieć przyszłości. Pomyślnie wdrożenie w znaczący sposób rozwiąże wiele tradycyjnych problemów związanych ze sztuczną inteligencją. Potencjalne korzyści obejmują:

- minimalizacja początkowego programowania specyficznego dla środowiska (poprzez konfigurację samoadaptacyjną);
- znaczne ograniczenie bieżących zmian w oprogramowaniu, ponieważ duża ilość dodatkowej funkcjonalności i wiedzy będzie zdobywana autonomicznie w drodze samonadzoru;
- znaczne zwiększenie zakresu zastosowań, ponieważ użytkownicy uczą i szkolą dodatkowe możliwości; I
- poprawiona elastyczność i solidność wynikająca ze zdolności systemów do dostosowywania się do zmieniających się wzorców danych, środowisk i celów.

AGI obiecuje wnieść istotny wkład w tworzenie oprogramowania i systemów robotycznych, które będą bardziej użyteczne, inteligentne i przyjazne dla człowieka. Wydaje się, że nadszedł czas na poważną inicjatywę na tej nowej ścieżce rozwoju ludzkości, która jest teraz dla nas otwarta.