

Czym jest sztuczna inteligencja?

Zajęcia na których próbujemy określić co jest przedmiotem zainteresowania sztucznej inteligencji

Piotr Fulmański

Instytut Nauk Ekonomicznych i Informatyki,
Państwowa Wyższa Szkoła Zawodowa w Płocku, Polska

September 29, 2015

- 1 Sztuczna Inteligencja - czym jest?
- 2 „Klasyfikacja”
- 3 Cele i zadania sztucznej inteligencji
- 4 Wiedza i jej reprezentowanie
- 5 Używanie wiedzy
- 6 Test Turinga
 - Zarzuty
- 7 Pytania
- 8 Dwa różne pokoje
 - Chiński Pokój
 - Odpowiedź systemowa
 - Jasny Pokój

Sztuczna Inteligencja - czym jest?

Początkową ideą, leżącą u podstaw sztucznej inteligencji, było stworzenie „myślących maszyn”. Powszechnie wiadomo, że komputery są w stanie rozwiązywać szereg różnych zadań szybciej i efektywniej niż ludzie. Jednakże najczęściej dostarczamy komputerowi wiedzy potrzebnej do rozwiązania danego problemu w postaci **konkretnego algorytmu**.

Uwaga

Wszędzie tam, gdzie znamy dokładny algorytm działania, prowadzący do wykonania interesującego nas zadania, inteligencja nie jest konieczna.

Jest jednak wiele problemów, których rozwiązanie nie da się ująć w ścisłe reguły. Pojęcie tzw. „inteligentnego komputera” zakłada, że **komputer potrafi efektywnie rozwiązać problem, nawet jeśli nie zostały zdefiniowane wszystkie kroki procesu rozwiązania**.

Sztuczna Inteligencja - czym jest?

Początkową ideą, leżącą u podstaw sztucznej inteligencji, było stworzenie „myślących maszyn”. Powszechnie wiadomo, że komputery są w stanie rozwiązywać szereg różnych zadań szybciej i efektywniej niż ludzie. Jednakże najczęściej dostarczamy komputerowi wiedzy potrzebnej do rozwiązania danego problemu w postaci **konkretnego algorytmu**.

Uwaga

Wszędzie tam, gdzie znamy dokładny algorytm działania, prowadzący do wykonania interesującego nas zadania, inteligencja nie jest konieczna.

Jest jednak wiele problemów, których rozwiązanie nie da się ująć w ścisłe reguły. Pojęcie tzw. „inteligentnego komputera” zakłada, że **komputer potrafi efektywnie rozwiązać problem, nawet jeśli nie zostały zdefiniowane wszystkie kroki procesu rozwiązania**.

Sztuczna Inteligencja - czym jest?

Początkową ideą, leżącą u podstaw sztucznej inteligencji, było stworzenie „myślących maszyn”. Powszechnie wiadomo, że komputery są w stanie rozwiązywać szereg różnych zadań szybciej i efektywniej niż ludzie. Jednakże najczęściej dostarczamy komputerowi wiedzy potrzebnej do rozwiązania danego problemu w postaci **konkretnego algorytmu**.

Uwaga

Wszędzie tam, gdzie znamy dokładny algorytm działania, prowadzący do wykonania interesującego nas zadania, inteligencja nie jest konieczna.

Jest jednak wiele problemów, których rozwiązanie nie da się ująć w ścisłe reguły. Pojęcie tzw. „inteligentnego komputera” zakłada, że **komputer potrafi efektywnie rozwiązać problem, nawet jeśli nie zostały zdefiniowane wszystkie kroki procesu rozwiązania**.

Ze zbudowaniem myślącej maszyny wiązano wielkie nadzieje. Jednak w dziedzinie sztucznej inteligencji trudno znaleźć proste i ładne rozwiązania, czy wręcz „prawa inteligencji” na wzór praw fizyki czy chemii. Być może takie prawa w ogóle nie istnieją? Sam temat budowania sztucznej inteligencji jest dosyć kontrowersyjny. Pojawiły się opinie, że (sztucznej) inteligencji nie da się skonstruować, a istniejące systemy nie mają nic wspólnego z prawdziwą inteligencją, ponieważ wykonują jedynie operacje na liczbach i symbolach. Przy okazji powstało wiele pytań natury filozoficznej, wielu naukowców zaczęło frapować pytanie: „Czy maszyna może myśleć?” lub „Co mamy na myśli, gdy mówimy że program zachowuje się inteligentnie?” Tak naprawdę jednak problem leży chyba na poziomie definicji terminów „inteligencja (naturalna)” i „sztuczna inteligencja”.

Dobrym punktem wyjścia w rozważaniach nad sztuczną inteligencją jest definicja inteligencji jako takiej. Ponieważ jednak inteligencja traktowana jako pewna własność czy cecha umysłu ma wiele odcieni znaczeniowych, zatem mamy wiele różnych definicji. Oto niektóre z nich:

Definition (Tiepłow)

Inteligencja to właściwość psychiczna, która przejawia się we względnie stałej, charakterystycznej dla jednostki, efektywności wykonywania zadań.

Definition (Stern)

Inteligencja to ogólna zdolność adaptacji do nowych warunków i wykonywania nowych zadań.

Definition (Piaget)

Inteligencja to zdolność rozwiązywania problemów.

Definition (Tiepłow)

Inteligencja to właściwość psychiczna, która przejawia się we względnie stałej, charakterystycznej dla jednostki, efektywności wykonywania zadań.

Definition (Stern)

Inteligencja to ogólna zdolność adaptacji do nowych warunków i wykonywania nowych zadań.

Definition (Piaget)

Inteligencja to zdolność rozwiązywania problemów.

Definition (Tiepłow)

Inteligencja to właściwość psychiczna, która przejawia się we względnie stałej, charakterystycznej dla jednostki, efektywności wykonywania zadań.

Definition (Stern)

Inteligencja to ogólna zdolność adaptacji do nowych warunków i wykonywania nowych zadań.

Definition (Piaget)

Inteligencja to zdolność rozwiązywania problemów.

Definition (Spearman)

Inteligencja to dostrzeganie zależności, relacji.

Definition (Ferguson)

Inteligencja to zdolność uczenia się.

Definition (Boring)

Inteligencja to to, co mierzą testy inteligencji.

Definition (Spearman)

Inteligencja to dostrzeganie zależności, relacji.

Definition (Ferguson)

Inteligencja to zdolność uczenia się.

Definition (Boring)

Inteligencja to to, co mierzą testy inteligencji.

Definition (Spearman)

Inteligencja to dostrzeganie zależności, relacji.

Definition (Ferguson)

Inteligencja to zdolność uczenia się.

Definition (Boring)

Inteligencja to to, co mierzą testy inteligencji.

Definition

Inteligencja to zdolność do aktywnego przetwarzania informacji, przekształcania ich z jednej formy w inną poprzez operacje logiczne.

Definition

Inteligencja to zdolność do przetwarzania informacji na poziomie abstrakcyjnych idei (np. umiejętność dokonywania obliczeń matematycznych lub gry w szachy).

Definition

Inteligencja to zdolność do aktywnego przetwarzania informacji, przekształcania ich z jednej formy w inną poprzez operacje logiczne.

Definition

Inteligencja to zdolność do przetwarzania informacji na poziomie abstrakcyjnych idei (np. umiejętność dokonywania obliczeń matematycznych lub gry w szachy).

Definition

Inteligencja to zdolność do twórczego, a nie tylko mechanicznego przetwarzania informacji, czyli tworzenia zupełnie nowych pojęć i ich nieoczekiwanych połączeń.

Definition

Inteligencja to zespół zdolności umysłowych umożliwiających jednostce sprawne korzystanie z nabytej wiedzy oraz skuteczne zachowanie się wobec nowych zadań i sytuacji.

Definition

Inteligencja to zdolność do twórczego, a nie tylko mechanicznego przetwarzania informacji, czyli tworzenia zupełnie nowych pojęć i ich nieoczekiwanych połączeń.

Definition

Inteligencja to zespół zdolności umysłowych umożliwiających jednostce sprawne korzystanie z nabytej wiedzy oraz skuteczne zachowanie się wobec nowych zadań i sytuacji.

Z punktu widzenia dalej prowadzonych rozważań najbardziej adekwatną dla nas definicją jest:

Definition

Inteligencja jest to zdolność uczenia się i rozumienia zjawisk poprzez doświadczenie, zdolność zdobywania wiedzy i wykorzystywania jej w celu szybkiego i efektywnego reagowania na nowe sytuacje; zdolność rozumowania w celu efektywnego rozwiązywania problemów.

Charakterystyka inteligencji naturalnej

A zatem inteligencję charakteryzuje

- uczenie się ([...] *zdolność uczenia się* [...]),
- zdobywanie wiedzy ([...] *zdolność zdobywania wiedzy* [...]),
- wnioskowanie ([...] *wykorzystywania jej* [wiedzy] [...] *zdolność rozumowania* [...]),
- uogólnianie ([...] *reagowania na nowe sytuacje* [...]).

Charakterystyka inteligencji naturalnej

A zatem inteligencję charakteryzuje

- uczenie się ([...] *zdolność uczenia się* [...]),
- zdobywanie wiedzy ([...] *zdolność zdobywania wiedzy* [...]),
- wnioskowanie ([...] *wykorzystywania jej* [wiedzy] [...] *zdolność rozumowania* [...]),
- uogólnianie ([...] *reagowania na nowe sytuacje* [...]).

Charakterystyka inteligencji naturalnej

A zatem inteligencję charakteryzuje

- uczenie się ([...] *zdolność uczenia się* [...]),
- zdobywanie wiedzy ([...] *zdolność zdobywania wiedzy* [...]),
- wnioskowanie ([...] *wykorzystywania jej* [wiedzy] [...] *zdolność rozumowania* [...]),
- uogólnianie ([...] *reagowania na nowe sytuacje* [...]).

Charakterystyka inteligencji naturalnej

A zatem inteligencję charakteryzuje

- uczenie się ([...] *zdolność uczenia się* [...]),
- zdobywanie wiedzy ([...] *zdolność zdobywania wiedzy* [...]),
- wnioskowanie ([...] *wykorzystywania jej*[wiedzy][...] *zdolność rozumowania*[...]),
- uogólnianie ([...] *reagowania na nowe sytuacje*[...]).

Charakterystyka inteligencji naturalnej

A zatem inteligencję charakteryzuje

- uczenie się ([...] *zdolność uczenia się* [...]),
- zdobywanie wiedzy ([...] *zdolność zdobywania wiedzy* [...]),
- wnioskowanie ([...] *wykorzystywania jej*[wiedzy][...] *zdolność rozumowania*[...]),
- uogólnianie ([...] *reagowania na nowe sytuacje*[...]).

Charakterystyka inteligencji sztucznej

Skoro taka jest natura (naturalnej) inteligencja, to sztuczna inteligencja, jakkolwiek rozumiana, np. jako

- inteligencja udawana (gdy, stwarzamy pozory inteligentnego zachowania dzięki stosowaniu odpowiednich algorytmów – np. gra w szachy),
- inteligencja symulowana (gdy symulujemy procesy biochemiczne mające swój udział „w inteligencji”),
- inteligencja inna od naturalnej

także powinna charakteryzować się zdolnością

- zdobywania wiedzy,
- uczenia się,
- wnioskowania,
- uogólniania.

Sztuczna inteligencja – próba definicji

Definition

Sztuczna Inteligencja jest dziedziną wiedzy której celem i przedmiotem badań są maszyny potrafiące rozwiązywać zadania, przy rozwiązywaniu których człowiek korzysta ze swojej inteligencji.

Definition

Artificial intelligence (AI) is the study of how to make computers do things which, at the moment, people do better.

Sztuczna inteligencja – próba definicji

Definition

Sztuczna Inteligencja jest dziedziną wiedzy której celem i przedmiotem badań są maszyny potrafiące rozwiązywać zadania, przy rozwiązywaniu których człowiek korzysta ze swojej inteligencji.

Definition

Artificial intelligence (AI) is the study of how to make computers do things which, at the moment, people do better.

Silna vs. słaba sztuczna inteligencja

Sami badacze Sztucznej Inteligencji dzielą się na dwa obozy, wedle propozycji i charakterystyki Johna Searle'a:

Silna vs. słaba sztuczna inteligencja

Silna Sztuczna Inteligencja (*ang. Strong Artificial Intelligence*)

Zwolennicy tego obozu uważają, iż zbiór problemów stawianych przed nauką (jako taką, Sztuczną Inteligencją w szczególności) i rozwiązywalnych metodami naukowymi pokrywa się z klasą problemów rozstrzygalnych algorytmicznie, czyli rozwiązywalnych przez urządzenia pracujące na zasadzie maszyny Turinga. Mówiąc bardziej obrazowo, dla każdego zachowania lub akcji jaka wydarzy się w naszym otoczeniu można napisać program (klasycznymi metodami na klasyczny komputer), który to zachowanie lub akcję wyjaśni. Stąd często określa się ich mianem algorytmistów. Mówiąc inaczej, odpowiednio zaprogramowany komputer jest w istotny sposób równoważny mózgowi, ma stany poznawcze, jest umysłem a pewne cechy umysłowe można przypisać logicznemu działaniu dowolnego urządzenia liczącego.

Silna vs. słaba sztuczna inteligencja

Słaba Sztuczna Inteligencja (*ang. Weak Artificial Intelligence*)

Zwolennicy tego poglądu są zdania, iż możliwe jest pełne poznanie badanych zagadnień, ale poznanie to nie będzie opierało się na procedurach algorytmicznych. Mówiąc inaczej komputer pozwala jedynie formułować i sprawdzać hipotezy dotyczące mózgu.

Podział według Roger'a Penrose'a

Z kolei Roger Penrose proponuje podział oparty o następujące stanowiska

- 1 Myślenie zawsze polega na obliczeniach, a w szczególności świadome doznania postają wskutek realizacji odpowiedniego procesu obliczeniowego (stanowisko to pokrywa się z SAI).
- 2 Świadomość jest cechą fizyczną działającego mózgu. O ile wszystkie fizyczne procesy można symulować obliczeniowo, to jednak symulacjom tym nie towarzyszy świadomość.
- 3 Odpowiedni procesy fizyczne w mózgu powodują powstanie świadomości, ale tych procesów nie można nawet symulować obliczeniowo (stanowisko to pokrywa się z WAI).
- 4 Świadomości nie można wyjaśnić w żaden fizyczny, obliczeniowy czy inny naukowy sposób (mistycyzm).

Podział według Roger'a Penrose'a

Z kolei Roger Penrose proponuje podział oparty o następujące stanowiska

- 1 Myślenie zawsze polega na obliczeniach, a w szczególności świadome doznania postają wskutek realizacji odpowiedniego procesu obliczeniowego (stanowisko to pokrywa się z SAI).
- 2 Świadomość jest cechą fizyczną działającego mózgu. O ile wszystkie fizyczne procesy można symulować obliczeniowo, to jednak symulacjom tym nie towarzyszy świadomość.
- 3 Odpowiedni procesy fizyczne w mózgu powodują powstanie świadomości, ale tych procesów nie można nawet symulować obliczeniowo (stanowisko to pokrywa się z WAI).
- 4 Świadomości nie można wyjaśnić w żaden fizyczny, obliczeniowy czy inny naukowy sposób (mistycyzm).

Podział według Roger'a Penrose'a

Z kolei Roger Penrose proponuje podział oparty o następujące stanowiska

- 1 Myślenie zawsze polega na obliczeniach, a w szczególności świadome doznania postają wskutek realizacji odpowiedniego procesu obliczeniowego (stanowisko to pokrywa się z SAI).
- 2 Świadomość jest cechą fizyczną działającego mózgu. O ile wszystkie fizyczne procesy można symulować obliczeniowo, to jednak symulacjom tym nie towarzyszy świadomość.
- 3 Odpowiedni procesy fizyczne w mózgu powodują powstanie świadomości, ale tych procesów nie można nawet symulować obliczeniowo (stanowisko to pokrywa się z WAI).
- 4 Świadomości nie można wyjaśnić w żaden fizyczny, obliczeniowy czy inny naukowy sposób (mistycyzm).

Podział według Roger'a Penrose'a

Z kolei Roger Penrose proponuje podział oparty o następujące stanowiska

- 1 Myślenie zawsze polega na obliczeniach, a w szczególności świadome doznania postają wskutek realizacji odpowiedniego procesu obliczeniowego (stanowisko to pokrywa się z SAI).
- 2 Świadomość jest cechą fizyczną działającego mózgu. O ile wszystkie fizyczne procesy można symulować obliczeniowo, to jednak symulacjom tym nie towarzyszy świadomość.
- 3 Odpowiedni procesy fizyczne w mózgu powodują powstanie świadomości, ale tych procesów nie można nawet symulować obliczeniowo (stanowisko to pokrywa się z WAI).
- 4 Świadomości nie można wyjaśnić w żaden fizyczny, obliczeniowy czy inny naukowy sposób (mistycyzm).

Podział według Roger'a Penrose'a

Z kolei Roger Penrose proponuje podział oparty o następujące stanowiska

- 1 Myślenie zawsze polega na obliczeniach, a w szczególności świadome doznania postają wskutek realizacji odpowiedniego procesu obliczeniowego (stanowisko to pokrywa się z SAI).
- 2 Świadomość jest cechą fizyczną działającego mózgu. O ile wszystkie fizyczne procesy można symulować obliczeniowo, to jednak symulacjom tym nie towarzyszy świadomość.
- 3 Odpowiedni procesy fizyczne w mózgu powodują powstanie świadomości, ale tych procesów nie można nawet symulować obliczeniowo (stanowisko to pokrywa się z WAI).
- 4 Świadomości nie można wyjaśnić w żaden fizyczny, obliczeniowy czy inny naukowy sposób (mistycyzm).

Cele sztucznej inteligencji

Można wyodrębnić dwa główne cele rozwoju Sztucznej Inteligencji:

- Budowa inteligentnych systemów (komputerowych), tzn. systemów wykazujących cechy podobne do cech inteligentnego działania człowieka, do skutecznego rozwiązywania trudnych zagadnień.
- Opracowanie obliczeniowej (algorytmicznej) teorii inteligencji.

Zadania sztucznej inteligencji

Głównymi zadaniami sztucznej inteligencji są:

reprezentacja wiedzy , aby móc przyjmować pojawiające się informacje o świecie, rozumieć je oraz konfrontować z już posiadaną wiedzą,

wnioskowanie , aby wyciągać wnioski z pojawiających się informacji, i podejmować decyzje o dalszych działaniach,

uczenie się dla dostosowania się do nowo pojawiających się okoliczności, nieprzewidzianych przez twórców systemu, pojmowania nowych zjawisk, itp.,

rozumienie języka naturalnego aby można było praktycznie sprawdzić zdolności systemu sztucznej inteligencji,

posługiwanie się wizją w celu samodzielnego pozyskiwania wiedzy o świecie,

robotyka czyli praktyczna konstrukcja systemu zdolnego poruszać się i wykonywać działania w świecie rzeczywistym.

Zadania sztucznej inteligencji

Osiągnięcie postawionych celów jest jeszcze odległe i zależy w znacznej mierze od wyboru narzędzi i właściwego podejścia do zagadnienia. Co rozumiemy przez „właściwe podejście” niech wyjaśni analogia maszyn myślących (jakkolwiek pojęcie *maszyny* byśmy rozumieli) do maszyn latających.

Zauważmy, że latanie możemy rozumieć i rozpatrywać na kilka sposobów. Najbardziej naturalne podejście polega na podpatrywaniu otaczającego nas świata i jego naśladowaniu. W ten oto sposób, wcześniej czy później, będziemy musieli dojść do wniosku, że istotą latania jest machanie skrzydłami pokrytymi albo jakąś błoną albo „piórami”. W konsekwencji, możemy poświęcić mnóstwo czasu na konstrukcję odpowiednich skrzydeł, błon i sztucznych piór. Czy jednak weryfikujący naszą teorię Ikar XXI wieku skacząc z wysokiej wieży faktycznie poleci?

Zadania sztucznej inteligencji

Osiągnięcie postawionych celów jest jeszcze odległe i zależy w znacznej mierze od wyboru narzędzi i właściwego podejścia do zagadnienia. Co rozumiemy przez „właściwe podejście” niech wyjaśni analogia maszyn myślących (jakkolwiek pojęcie *maszyny* byśmy rozumieli) do maszyn latających.

Zauważmy, że latanie możemy rozumieć i rozpatrywać na kilka sposobów. Najbardziej naturalne podejście polega na podpatrywaniu otaczającego nas świata i jego naśladowaniu. W ten oto sposób, wcześniej czy później, będziemy musieli dojść do wniosku, że istotą latania jest machanie skrzydłami pokrytymi albo jakąś błoną albo „piórami”. W konsekwencji, możemy poświęcić mnóstwo czasu na konstrukcję odpowiednich skrzydeł, błon i sztucznych piór. Czy jednak weryfikujący naszą teorię Ikar XXI wieku skacząc z wysokiej wieży faktycznie poleci?

Zadania sztucznej inteligencji

Istnieje też drugi sposób – zrozumieć elementarne zasady którym podlega lot, bez ograniczania samych siebie żadnymi wstępnymi założeniami. To nic, że ludzka wersja obiektu latającego ma inaczej zbudowane i co istotne całkowicie nieruchome skrzydła. Ważne, że cel: maszyna latająca udało się osiągnąć. To podejście jest o tyle użyteczne, że oprócz latającej maszyny daje nam też zarówno pewien zestaw narzędzi jak i znacznie lepsze zrozumienie problemu i wiedzę leżące u podstaw zjawiska nazywanego lataniem. Dzięki temu możliwe jest stworzenie latających maszyn zupełnie nie przypominających istot żywych, np. helikoptery. A zatem myśląc maszyna wcale nie musi być zbudowana na wzór i podobieństwo człowieka, tak jak samolot raczej nie jest podobny do ptaka. Pamiętajmy o tym, gdy pytamy „Czy komputer na prawdę myśli?” bo „Czy samolot na prawdę lata?”.

Zadania sztucznej inteligencji

Istnieje też drugi sposób – zrozumieć elementarne zasady którym podlega lot, bez ograniczania samych siebie żadnymi wstępnymi założeniami. To nic, że ludzka wersja obiektu latającego ma inaczej zbudowane i co istotne całkowicie nieruchome skrzydła. Ważne, że cel: maszyna latająca udało się osiągnąć. To podejście jest o tyle użyteczne, że oprócz latającej maszyny daje nam też zarówno pewien zestaw narzędzi jak i znacznie lepsze zrozumienie problemu i wiedzę leżące u podstaw zjawiska nazywanego lataniem. Dzięki temu możliwe jest stworzenie latających maszyn zupełnie nie przypominających istot żywych, np. helikoptery. A zatem myśląc maszyna wcale nie musi być zbudowana na wzór i podobieństwo człowieka, tak jak samolot raczej nie jest podobny do ptaka. Pamiętajmy o tym, gdy pytamy „Czy komputer na prawdę myśli?” bo „Czy samolot na prawdę lata?”.

Wiedza i jej reprezentowanie

Przede wszystkim należy zadać sobie dwa pytania: „Czym jest wiedza” oraz „Jak ją reprezentować?” Zgromadzenie odpowiedniej wiedzy i jej wewnętrzna reprezentacja jest podstawową sprawą dla wszystkich systemów mających się inteligentnie zachowywać. *„W AI podchodzi się do tego pragmatycznie: reprezentacja wiedzy to kombinacja struktur danych i procedur interpretacyjnych tak dobranych, że właściwie użyte, prowadzić będą do inteligentnego zachowania. Same struktury danych nie są jeszcze wiedzą, jak sama encyklopedia nią nie jest, potrzebny jest jeszcze czytelnik, interpretator”.*

Wiedza i jej reprezentowanie

Wyróżnić można następujące rodzaje wiedzy:

Wiedza o obiektach i przedmiotach.

Wiedza o zdarzeniach , należy znać również następstwa przyczynowe i sekwencję czasową zdarzeń.

Umiejętności. Jest to typ wiedzy praktycznej, którą zdobywa się metodą prób i błędów i której nie da się przekazać w sposób werbalny

Meta-wiedza , czyli wiedza o samej wiedzy (wiem, że nic nie wiem), wynikająca z niedoskonałości naszej percepcji i metod pomiarowych, a więc z niepełnych danych, z oceny wiarygodności tych danych, z ograniczeń ludzkiej pamięci i zdolności do rozumowania.

Przekonania prawdziwe czy fałszywe, z góry powzięte nastawienia, wpływają na rozumowanie w oczywisty sposób. Mają różne stopnie pewności i trzeba je traktować jako pewien rodzaj wiedzy.

W używaniu wiedzy ważne są trzy aspekty:

- **Gromadzenie nowej wiedzy.** Nie jest tylko akumulacją, dodawaniem nowych faktów, lecz wymagany jest proces tworzenia związków z już istniejącą wiedzą. Konieczna jest więc klasyfikacja wiedzy oraz interakcja z wiedzą już posiadaną.
- **Wydobywanie z bazy wiedzy faktów, związanych z danym problemem.** Dla człowieka proces kojarzenia faktów jest dosyć naturalny, lecz jak kojarzyć w komputerowych programach? Można stosować łączenie lub grupowanie pewnych struktur danych, lecz wyklucza to większość nieprzewidzianych skojarzeń.
- **Rozumowanie przy użyciu tych faktów w poszukiwaniu rozwiązania.** Występuje wtedy, gdy system ma zrobić coś, co nie jest w jawny sposób zaprogramowane. Możemy wyróżnić:

- **Rozumowanie przy użyciu tych faktów w poszukiwaniu rozwiązania.** Występuje wtedy, gdy system ma zrobić coś, co nie jest w jawny sposób zaprogramowane. Możemy wyróżnić:
 - **Rozumowanie formalne** – występuje w logice matematycznej i opiera się na ustalonych regułach wnioskowania.
 - **Rozumowanie proceduralne** – oznaczające postępowanie według jakiejś instrukcji, procedury, a więc składa się z serii kroków. Jego odmianą jest budowanie modelu i symulacja komputerowa danej sytuacji.
 - **Rozumowanie przez analogię** – często stosowane przez ludzi, lecz bardzo trudne dla programów komputerowych.
 - **Rozumowanie przez uogólnienie** – bardzo powszechne w życiu codziennym i w matematyce, trudno jest jednak je zaprogramować.
 - **Meta-rozumowanie** – związane jest z wiedzą o tym, co już wiemy (np. czy ja mogę to skądś wiedzieć?).

Idea tesu

Test turinga – gra w naśladownictwo.

Sprzeciw teologiczny

Zarzut teologiczny wyraża się w stwierdzeniu:

Dusza jest dana od Boga i przysługuje tylko człowiekowi.

Jednak powoływanie Boga na świadka i gwaranta zawsze musi być bardzo ostrożne aby nie popełnić nadużycia i aby nie wypowiadać się w kwestiach o których nie mamy pojęcia (bo niby skąd możemy wiedzieć jaki On jest, co może a czego nie itd).

Sprzeciw teologiczny

Sam Turing zarzut ten komentuje następująco:

Wydaje mi się, że cytowany wyżej argument pociąga za sobą poważne ograniczenie wszechpotęgi Boga Wszchemogącego. Przyznano, że istnieją pewne rzeczy, których On nie może zrobić, takie jak uczynienie jedności równą dwóm, ale czyż nie powinniśmy wierzyć, że może On obdarzyć duszą słońca, jeśli będzie uważał, że słoń jest tego godny? [...] Podobny argument można sformułować w przypadku maszyn. Może on wydawać się inny, gdyż jest trudniejszy do „przełknięcia”. Ale naprawdę oznacza on jedynie nasze przekonanie o mniejszym prawdopodobieństwie uważania przez Niego tych warunków materialnych za odpowiednie do obdarzenia duszą.

Sprzeciw teologiczny

Jednakże jest to tylko spekulacja. Teologiczne argumenty nie wywierają na mnie głębokiego wrażenia [...] Gdyż [...] takie argumenty często bywały niewystarczające w przeszłości: w czasach Galileusza argumentowano, że teksty: „I słońce stało jeszcze... i nie spieszyło się zejść prawie przez cały dzień” (Jozue X. 13) i „Dał ziemi podstawę, tak, że nigdy nie powinna się ona ruszyć” (Psalm CV. 5) w sposób wystarczający zbijają teorię Kopernika. Przy naszej obecnej wiedzy taki argument wydaje się bezwartościowy. Gdy ta wiedza nie była dostępna, to wywierało to zupełnie inne wrażenie.

Zarzut „lepiej o tym nie myśleć”

Zarzut teologiczny wyraża się w stwierdzeniu:

**Konsekwencje myślenia maszyn byłyby zbyt okropne.
Miejmy nadzieję i wierzymy, że one nie mogą myśleć.**

Zarzut „lepiej o tym nie myśleć”

Wyrazem reprezentowania takiego stanowiska są wypowiedzi podobne do: *jeśli poważnie założyć jej [sztucznej inteligencji] realność [...] to nie sposób nie zauważyć swego rodzaju puszki Pandory przeróżnych trudności i niebezpieczeństw: społecznych, ekonomicznych, psychologicznych i jeszcze innych. najśmielsze odkrycia naukowe i wynalazki techniczne nie są w stanie dać nam niczego filozoficznie interesującego, gdy chodzi o tak zwaną sztuczną, czyli faktycznie alternatywną inteligencję. Natomiast mogłyby spowodować katastrofę humanitarną w stylu Georga'a Orwella.*

Zarzut wyjątkowości człowieka

Zarzut wyjątkowości człowieka wyraża się w stwierdzeniu:

Miło jest wierzyć, że człowiek jest lepszy od wszystkich innych istot^a.

Zarówno ten jak i poprzedni argument sam Turing skomentował następująco:

Nie myślę, że ten argument jest wystarczająco poważny, aby trzeba go było zbijać.

^aLepszy w sensie posiadania unikalnych cech dających jemu dominującą pozycję.

Zarzut niemożności

Zgadzam się z tobą, że możesz zrobić maszyny wykonujące to wszystko, o czym wspomniałeś, ale nigdy nie będziesz w stanie zrobić maszyny, która by zrobiła X.

Definition (Maszyna myśląca w sensie Turinga)

O maszynie myślącej w sensie Turinga powiemy, gdy:

- 1 będzie będzie ona zdolna do porozumiewania się z (dowolnym) człowiekiem za pomocą języka pisanego;
- 2 język musi być rozumiany i akceptowany przez obie zainteresowane strony (człowieka i maszynę);
- 3 rozumienie przez maszynę oznacza generowanie „sensownych” zdań jako „odpowiedź” na poprzedzające je z zdania (zarówno własne jak i człowieka).

Żadna inna funkcja przypisywana człowiekowi nie musi być spełniona.

Czy maszyna może myśleć? Pytanie wydaje się być postawione bardzo prosto, ale rodzi serię kolejnych pytań:

- Co to jest maszyna?
- Co to znaczy myśleć?
- Kiedy możemy powiedzieć, że „coś” myśli?

Jaka możliwa maszyna może myśleć? Jeśli stwierdzimy, że maszyna może myśleć to czy każda? Tak, nie? Czym się powinna charakteryzować?

Czy myśląca maszyna będzie równoznaczna ze sztucznym człowiekiem? Jest to pytanie o istotę człowieczeństwa. Co to znaczy być człowiekiem czy być podobnym do niego? Podobieństwo to jest funkcjonalne, strukturalne czy może jeszcze jakieś inne?

Dlaczego budujemy maszyny myślące?

- Budujemy aby dzięki temu zrozumieć istotę myślenia, istotę człowieczeństwa. A zatem w tym kontekście AI jest środkiem a nie celem. Środkiem służącym poznaniu nas samych.
- Zastąpią nas wszędzie tam gdzie mogło by być zagrożone życie ludzkie.
- Jeśli maszyny będą potrafiły myśleć, to możliwe, że będą to robiły na tyle sprawnie, że rozwiążą pewne problemy, których ludzie do tej pory nie rozwiążali.

Czy maszyna myśląca będzie posiadać prawa istoty myślącej?

Czy maszyna myśląca zaliczana będzie do „bytów” żywych?

Konsekwencją obu powyższych pytań są kolejne, bardziej szczegółowe:

- Czy można wyrzucić wyeksploatowaną maszynę? To tak jak gdyby wyrzucić kogoś z pracy tylko dlatego, że przekroczył pewien wiek. Zauważmy, że wyrzucenie maszyny ma znacznie szersze konsekwencje, bo zwykle jest równoznaczne z jej uśmierceniem.
- Problem zniewolenia maszyn i ich przymuszania do pracy.
- Jeśli maszyna będzie posiadała prawa przysługujące bytom żywym, to czy może odmówić pracy ze względu na poglądy?

Chiński pokój, argument chińskiego pokoju – eksperyment myślowy zaproponowany przez amerykańskiego filozofa i językoznawcę Johna Searle'a i przedstawiony w jego pracy z 1980 roku pt. *Minds, Brains, and Programs*, mający pokazać, że nawet efektywne symulacje komputerowe nie urzeczywistniają prawdziwego rozumu, odkąd wykonywanie różnorodnych zadań (np. obliczeniowych) nie musi opierać się na rozumieniu ich przez wykonawcę. Służy on przeciwnikom teorii tzw. mocnej sztucznej inteligencji jako kontrargument. U podstaw eksperymentu stoi niezgodność między syntaktyką a semantyką.

Założmy, że wiele lat temu skonstruowaliśmy komputer, który zachowuje się jakby rozumiał język chiński. Innymi słowy, komputer bierze chińskie znaki jako podstawę wejściową i śledzi zbiór reguł nimi rządzący (jak wszystkie komputery), koreluje je z innymi chińskimi znakami, które prezentuje jako informację wyjściową.

Założmy, że ten komputer wykonuje to zadanie w sposób tak przekonujący, że łatwo przechodzi test Turinga, tzn. przekonuje Chińczyka, że jest Chińczykiem. Na wszystkie pytania, które człowiek zadaje, udziela właściwych odpowiedzi w sposób tak naturalny, że Chińczyk jest przekonany, iż rozmawia z innym Chińczykiem. Zwolennicy mocnej sztucznej inteligencji wyciągają stąd wniosek, że komputer rozumie chiński, tak jak człowiek.

Teraz Searle proponuje, żeby założyć, iż to on sam siedzi wewnątrz komputera. Innymi słowy, on sam znajduje się w małym pokoju, w którym dostaje chińskie znaki, konsultuje książkę reguł, a następnie zwraca inne chińskie znaki, ułożone zgodnie z tymi regułami. Searle zauważa, że oczywiście nie rozumie ani słowa po chińsku, mimo iż wykonuje powierzone mu zadanie. Następnie argumentuje, że jego brak rozumienia dowodzi, że i komputery nie rozumieją chińskiego, znajdując się w takiej samej sytuacji jak on: są bezumysłowymi manipulatorami symboli i nie rozumieją, co „mówią”, tak jak i on nie rozumie treści chińskich znaków, którymi operował.

W roku 1984 Searle stworzył bardziej formalną wersję argumentów przeciw silnej sztucznej inteligencji, których częścią jest chiński pokój (argument 2):

- 1 Mózgi wytwarzają umysły.
- 2 Syntaktyka nie jest wystarczająca dla semantyki.
- 3 Programy komputerowe są całkowicie zdefiniowane przez ich formalną albo syntaktyczną strukturę.
- 4 Umysły mają treści mentalne w szczególności mają one zawartości semantyczne.

Formalne argumenty – konkluzje

Searle sugeruje że argumenty te prowadzą bezpośrednio do czterech konkluzji

- 1 Żaden program komputerowy z siebie nie jest wystarczający aby dać systemowi umysł. Programy nie są umysłami i nie są z siebie zdolne by mieć umysły.
- 2 Sposób w jaki funkcje mózgu wytwarzają umysły nie może być wyłącznie zasługą wykonywania programu komputerowego.
- 3 Cokolwiek jest w stanie wytworzyć umysł będzie miało władze przyczynowe co najmniej równoważne mózgowi
- 4 Procedury programu komputerowego nie będą same z siebie wystarczające aby wytworowi ludzkiemu nadać stany mentalne równoważne ludzkim. Wytwory ludzkie będą wymagały zdolności i władz mózgu

Argument Searl'a opiera się na założeniu, że system manipulujący symbolami według zadanych reguł może przejść test Turinga. Innymi słowy Searl zakłada, że wystarczy używać tablicy wyszukiwania do wejścia i dać wyjście aby przejść test Turinga. Takie stanowisko nie wydaje się być prawdziwe, gdyż jakikolwiek rzeczywisty system zdolny do istotnego spełnienia koniecznych warunków testu byłby tak złożony, że wcale nie byłoby oczywiste, że brakuje mu zrozumienia chińskiego. I najprawdopodobniej jego budowa nie opierałaby się tylko i wyłącznie o proste reguły transformacji jednych formuł na inne.

Churchlandowie odrzucają pewnik Searle'a brzmiący: *Składnia sama przez się ani nie kształtuje semantyki ani do niej nie wystarcza*. Uważają go za bardzo słycony i sądzą, że cały Chiński Pokój został specjalnie skonstruowany po to, by obudować wyżej wymieniony aksjomat. Podają też kontrprzykład zwany Jasnym Pokojem.

Wyobraźmy sobie ciemny pokój a w nim człowieka trzymającego duży magnes. Jeśli ten człowiek będzie wymachiwał magnesem do góry i na dół, to powinno to wywołać koliste rozchodzenie się fal elektromagnetycznych, a zatem powinno zrobić się jasno. Jednak każdy z nas wie, że jest nie do pomyślenia, byśmy mogli uzyskiwać światło poprostu przez poruszanie magnesem w kółko.

Ktoś może twierdzić na podstawie powyższego przykładu, że światło nie ma związku z falami elektromagnetycznymi, gdyż potrząsanie magnesem nie daje światła w ciemnym pokoju. Odwołując się zaś do Chińskiego Pokoju można wykazywać, jak Searle, że potrząsanie, mieszanie chińskich symboli nie da „światła” rozumienia języka chińskiego. Analogicznie do argumentu z 1980 roku Churchlandowie również podają aksjomaty związane z Jasnym Pokojem:

Aksjomat 1: Elektryczność i magnetyzm są formami energii.

Aksjomat 2: Zasadniczą cechą światła jest jasność.

Aksjomat 3: Energia nie jest ani konieczna, ani wystarczająca dla uzyskania jasności.

Wniosek 1: Elektryczność i magnetyzm nie są konieczne ani wystarczające dla wyjaśnienia istoty światła.

Gdyby taki tok myślenia przedstawiono po ogłoszeniu koncepcji Maxwella na temat elektromagnetycznej natury światła a przed uznaniem faktu tożsamości światła i fal elektromagnetycznych, to aksjomaty te mogłyby stanowić podstawę odrzucenia teorii Maxwella – konkludują Churchlandowie. A przecież jednak wskutek badań teoria ta została potwierdzona. Toteż Churchlandowie radzą czekać powołując się na precedensy w historii nauki, kiedy to poglądy odrzucane były następnie akceptowane. Według nich być może późniejsze badania wykażą, że można zbudować semantykę jedynie przy pomocy samej syntaktyki i wnioski wyciągnięte z Chińskiego Pokoju uważają za przedwczesne.