

Inteligentne Systemy Obliczeniowe

Streszczenie wykładu
(do egzaminu)

2011 / 2012

Cel wykładu

Głównym celem wykładu jest zapoznanie słuchaczy z dziedziną tak zwaną sztuczną inteligencją, co obecnie nabiera coraz większego znaczenia.

W szczególności, omówione zostaną sztuczne sieci neuronowe, algorytmy genetyczne, systemy rozmyte, systemy ekspertowe.

Prezentowane jest też zastosowanie cennej wiedzy z innych dziedzin nauki, psychologii poznawczej i neurofizjologii.

Różne systemy złożone o dużym poziomie równoległości.

Ogólnie, tematyka wykładu nie ma się skupiać na prezentowaniu konkretnych szczegółowych rozwiązań, tylko bardziej na prezentowaniu szerszego podejścia i przyszłości inteligentnych systemów, aby zdobyta wiedza była możliwie cenna w przyszłości.

Jednak nie wykluczone są także próby wspólnego opracowania wybranych inteligentnych programów, rozwiązań lub eksperymentów, które mogą być, między innymi, tematem ciekawych publikacji.

Czy sztuczna inteligencja jest osiągalna ?

(tak ...? dlaczego nie ...? jak ?)

Marvin Lee Minsky

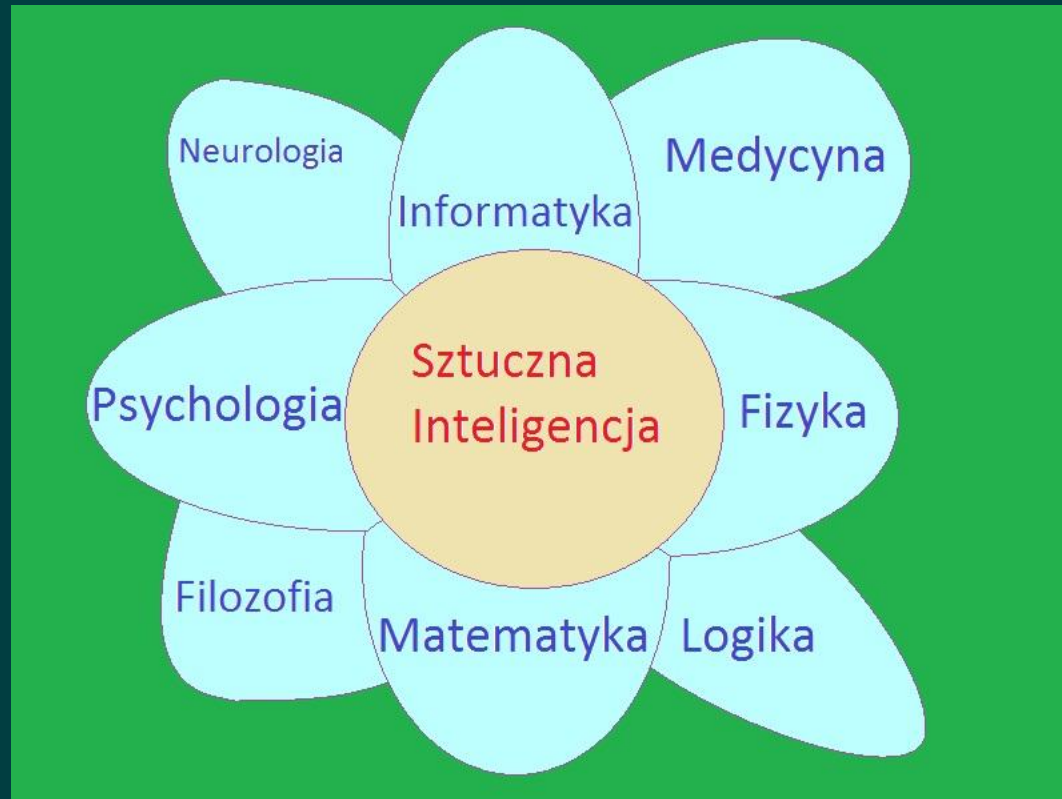
Marvin Lee Minsky (ur. 1927), zajmuje się głównie naukami kognitywnymi, zwłaszcza Sztuczną Inteligencją, teorią języka, filozofią. Współzałożyciel laboratorium **AI** w **MIT**. Autor bardzo wielu prac, publikacji, eksperymentów.

Sztuczna Inteligencja jest dziedziną wiedzy, która postawiła sobie za cel i przedmiot badań – maszyny, które potrafiłyby rozwiązywać zadania, przy rozwiązywaniu których człowiek korzysta ze swojej inteligencji. Zadaniem takimi mogą być, na przykład, rozumowanie czy podejmowanie decyzji.

Przy czym, dziedzina ta odgrywa olbrzymią rolę przy zrozumieniu tego, jak mając do czynienia z takimi zadaniami funkcjonuje ludzki mózg i umysł. A przez to, ta dziedzina jest zależna od wielu innych dziedzin.

Sztuczna Inteligencja

Matematyka, Informatyka, Psychologia,
Neurofizjologia, Medycyna, Fizyka,
Biologia, Logika, Filozofia, Robotyka.



Nauki Kognitywne

Termin „Nauki Kognitywne” odnosi się do interdyscyplinarnego stadium dotyczącego nabywania i użycia wiedzy.

W stadium to wkład wnoszą: Sztuczna Inteligencja, psychologia, filozofia, matematyka, informatyka, logika, meta-matematyka, lingwistyka, neurofizjologia, nauki o mózgu, pedagogika, antropologia, robotyka.

Nauki kognitywne rozwinęły się głównie dzięki osiągnięciom: wynalazkowi komputera, programom imitującym i modelującym myślenie i psychologii poznawczej, teorii lingwistyki gramatyki generatywnej i jej pochodnym.

Nauki kognitywne stanowią syntezę wiedzy istotowej, na jakiej opiera się ludzkie poznanie, zajmując się procesami przetwarzania informacji przez człowieka i komputerowym modelowaniem tychże procesów. Istnieje pięć głównych pól badawczych w naukach kognitywnych: reprezentacja wiedzy, język, uczenie się, myślenie i percepcja.

Nauki kognitywne są studium inteligencji i systemów inteligentnych, ze szczególnym odniesieniem do zachowania inteligentnego jako procesu dającego się obliczyć (???)

Podstawowe znane techniki sztucznej inteligencji

- * Oparte logice formalnej.
- * Sieci Neuronowe.
- * Algorytmy genetyczne.
- * Programowanie ewolucyjne.
- * Oparte na teorii zbiorów rozmytych.
- * Oparte na funkcji wiarygodności *BEL*.
- * Systemy wielo-agentowe.
- * Systemy ekspertowe.
- * Inne systemy złożone (układy dynamiczne).
- * Systemy hybrydowe.



Trochę historii

Obserwacja całego procesu zwykle daje więcej niż tylko skupienie uwagi wokół sytuacji obecnej.

Widzimy wtedy problemy jakie zostały pokonane, sposoby rozwiązań, a także trudności które stwarzały przestoje.

Warto znać przeszłość i teraźniejszość, aby ekstrapolować przyszłość.

Kurt Godel

Szokujące odkrycie meta-matematyczne !

„ *nie wszystko da się udowodnić* „

Twierdzenie Godla głosi, że dla dowolnego systemu formalnego, który oznaczamy jako ***M***, zawierającego część arytmetyki liczb naturalnych jest możliwe skonstruowanie (w języku systemu ***M***) takiego zdania, które nie da się udowodnić, a także nie da się udowodnić negacji tego zdania.

Alan Turing – proponuje **Maszynę Turinga**.

Zadania Turinga na drugą połowę XX wieku

- * Gry !! Umiejętność gry (szachy, Go, itd.).
- * Dowodzenie twierdzeń matematycznych (logika).
- * Rozpoznawanie obrazów (dźwięków, sygnałów).
- * Uczenie maszynowe.
- * Analiza mowy i języka, tłumaczenie (translatory).
- * Rozpoznawanie analogii, podobieństw.



Heurystyka

heuresis – odnaleźć, odkryć , *heureka* – znalazłem.

Heurystyka w informatyce, to metoda znajdowania rozwiązań, dla której nie ma gwarancji znalezienia rozwiązania poprawnego, optymalnego.

Rozwiązań tych używa się, na przykład, wtedy gdy pełny poprawny algorytm jest zbyt kosztowny, nieznany, albo nawet niemożliwy, lub gdy nawet ocena sytuacji lub rozwiązania jest niemożliwa, niezdefiniowana. W ten sposób, często otrzymujemy rozwiązanie przybliżone, a czasami metoda nawet nie jest skuteczna, tak więc nie jest pewna na 100%.

Czyli mówiąc prosto, jest to działanie „na chybił trafił”, „na oko”, a nie przy pomocy poprawnie wyprowadzonego algorytmu, który ma doprowadzić zawsze do prawidłowego rozwiązania. Z takimi metodami mamy bardzo często do czynienia w dziedzinie sztucznej inteligencji. Często nawet taka metoda staje się źródłem definicji (pojęcia, zakresy, klasyfikacji, celu zadania) i tym samym już przestaje być heurystyką.

Zestaw bardzo ogólnych zasad działania, opiera się także na znajdowaniu różnych zależności między faktami znanymi lub przewidywanymi, uplastycznia proces.

Podstawowe algorytmy programów grających (analiza drzewa , strategia min / max)

Obecnie wszystkie powszechnie znane teorie komputerowej gry w szachy opierają się na tak zwanym algorytmie „ Min / Max ”. Max przy własnym ruchu i Min podczas ruchu przeciwnika.

Jest tak, między innymi dlatego, że liczba możliwych partii szachowych jest tak duża, że żaden współczesny komputer nie jest na tyle szybki aby można było przeanalizować wszystkie możliwe pozycje szachowe (oprócz końcówek).

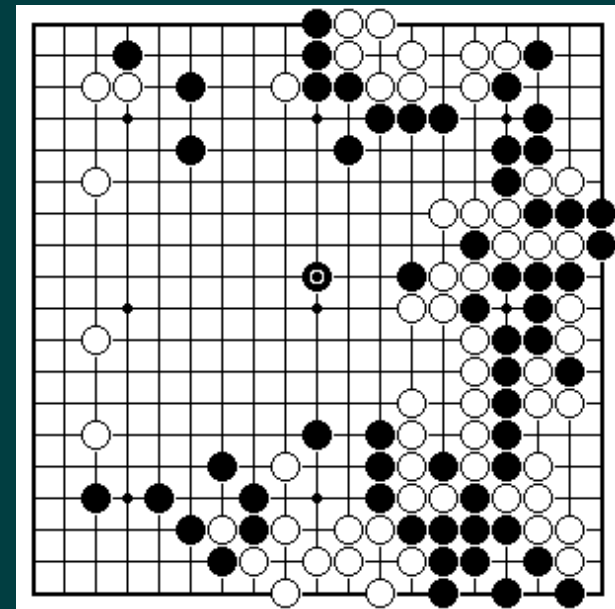
W najprostszej wersji program tworzy drzewo wszystkich możliwych ruchów do pewnej głębokości, wykonuje wirtualnie te ruchy i określa wartość pozycji po wykonaniu tych ruchów, i na podstawie tych wartości wybiera optymalną strategię.

Inne gry (warcaby , shogi , go , brydż , poker)

Shogi



GO



Jeszcze większym wyzwaniem jest stworzenie programów do gry w Shogi i GO. Złożoność „obliczeniowa” jest jeszcze większa i te gry wymagają jeszcze więcej twórczego myślenia, wyobraźni, intuicji, których jak widać maszynom jeszcze brakuje.

Ogólny schemat gry „N” (*dylematu więźnia*)

<div><div><div>B</div><div>A</div></div></div>		A	
		Współpracuj	Atakuj
B	Współpracuj	R R	T S
	Atakuj	S T	P P

T - Pokusa zdrady (gdy drugi współpracuje).

R - Nagroda za współpracę.

P - Kara za zdradę (gdy drugi też zdradzi).

S - Krzywda oszukanego.

Jeśli : $P < S$
to jest to
model gry
w tchórza

$$T > R > P > S$$

$$2R > T + S$$

$$T > R > P > S$$

To warunki konieczne, aby gra spełniała warunki dylematu więźnia. Te warunki zapewniają, że oszukiwanie jest w każdej sytuacji bardziej opłacalne niż współpraca i jednocześnie że obaj tracą gdy obaj oszukują. W przypadku iterowanego dylematu więźnia, powinien zachodzić dodatkowy warunek:

$$2R > T + S$$

Bo jeśli nie jest on spełniony, współpraca graczy nie daje im większych zysków niż naprzemienne oszukiwanie się.

Na podstawie sprawdzonych strategii, stwierdzono cechy :

Przyjazność , Mściwość , Wybaczenie , Brak zazdrości.

A dla gry w tchórza – chaotyczność, nieprzewidywalność !

Ciekawe spostrzeżenia związane z teoriami ewolucyjnymi.

Dylemat podróżnika

Gra zaproponowana w 1994 roku. Dwóch graczy stara się o wypłatę do wysokości 100 pkt. Obaj jednocześnie podają kwoty, „niższa wygrywa”. Jeśli są równe to wypłata jest dokładnie taka. Jeśli nie, to „wygrywa” ten co podał niższą, dostaje gratyfikację (2 pkt.), a ten drugi dostaje tą kwotę, ale minus 2 pkt. Sytuacja nieco podobna, współpraca się opłaca (ale ludzie grają inaczej).

Wojna na wyczerpanie

Dwóch graczy rywalizuje o nagrodę o wartości V . Podczas jednej partii zwiększają koszt (x), aż do czasu gdy któryś z graczy się podda, koszt (a). A wszyscy gracze płacą kwotę zaproponowaną przez przegrywającego. Jeśli równe, to obaj płacą $V/2 - a$. Dokładnie, wygrywający dostaje $V - a$, tym czasem przegrywający płaci a (dostaje $-a$). Czyli jakby model aukcji. Jest to najbardziej złożona gra, nie daje się opisać macierzą.

Dowodzenie Twierdzeń Logicznych Matematycznych

Metody wnioskowania

Zmienne : p, q, r, s , operatory : $\neg, \vee, \wedge, \rightarrow, \leftrightarrow$; nawiasy : $(,)$;

Metoda podstawiania :

Jeśli $\forall p : \phi(p)$ prawdziwe, to : $\phi(x)$ prawdziwe.

Metoda odrywania :

$(p \rightarrow q)$ i p , to : q ;

Metoda łańcuchowa :

$(p \rightarrow q)$ i $(q \rightarrow r)$, to : $(p \rightarrow r)$;



Metoda rezolucji

Metoda Rezolucji opracowana Alana Robinsona w 1965 roku.

Zakłada się że optymalny proces dowodzenia jest „nie wprost”.

A więc : „reductio ad absurdum”.

Reguła Rezolucji :

$$\left((\neg \alpha \vee \beta) \wedge (\alpha \vee \gamma) \right) \rightarrow (\beta \vee \gamma) ;$$

Formuły kolidujące \rightarrow Rezolwenta.

Klauzula – alternatywa złożona z formuł atomowych (literałów +/-).

Klauzula Horna – taka, która zawiera tylko jeden literał pozytywny.
(tym samym stanowi równoważność implikacji kilku argumentowej).

$$(\neg \alpha \wedge (\alpha \vee \gamma)) \rightarrow \gamma ;$$

Oraz wyprowadzanie klauzuli pustej : $(\neg \alpha \wedge \alpha) \rightarrow \mathbf{\varnothing} ;$

Przeszukiwanie Przestrzeni Stanów

Przestrzeń stanów

Czwórka uporządkowana : $[N , A , S , GD]$

N – Zbiór wierzchołków, czyli stany procesu.
odpowiadają stanom badanego procesu.

A – Zbiór krawędzi, czyli kroki procesu.

S – Niepusty podzbiór N , opisujący stany początkowe.

GD – Niepusty podzbiór N , opisujący stany docelowe.

Które są zadawane poprzez :

własności stanów.

własności ścieżek.



Podstawowe problemy przeszukiwania przestrzeni stanów

- * Czy metoda gwarantuje znalezienie rozwiązania ?
- * Czy algorytm zakończy się w każdym przypadku ?
(czy może też wpaść w pętlę nieskończoną)
- * Czy rozwiązanie będzie optymalne ?
Jaką mamy gwarancję że rozwiązanie jest optymalne ?
- * Jaka będzie złożoność czasowa i pamięciowa ?
Złożoność obliczeniowa całego procesu ?
- * Czy można zredukować złożoność obliczeniową ?

Sformułowanie zadania dla algorytmów przeszukiwania p.s.

- * Definicja stanu przestrzeni (state space).
- * Stan początkowy problemu (initial state).
- * Zbiór dopuszczalnych operatorów, akcji, funkcji.
(operator set , action set , successor function).
- * Zbiór stanów docelowych (goal states).
lub funkcja weryfikacji celu (goal test).
- * Funkcja kosztu ścieżki (path cost).
z reguły jest związana z głębokością przeszukiwania.

Własności przestrzeni stanów w reprezentacji grafowej

- * Powtarzające się stany, czyli cykle (podczas procesu).
- * Identyfikacja typu grafu reprezentującego p.s.
(drzewa , acykliczne drzewa skierowane , podgrafy).
- * Rozmiar przestrzeni stanów (ograniczanie złożoności).
- * Dodatkowe funkcje oceny położenia (odległość od celu).
- * Czy znamy pozycję celu ? Ile jest takich pozycji ?

Generalny podział typów poszukiwań

Przeszukiwanie ślepe

Przeszukiwanie heurystyczne

Inny podział, ze względu na naturę problemu.

Przeszukiwanie podstawowe

Przeszukiwanie z przeciwnikiem

Przeszukiwanie z ograniczeniami (więzami).



Przeszukiwanie z przeciwnikiem

A więc, sytuacja podobna jak w grach (na przykład szachy).

Podstawową metodą jest strategia min / max.

Szuka się maksimum funkcji heurystycznej dla naszego ruchu, przy minimum funkcji dla ruchów przeciwnika.

W tym przypadku funkcją h jest ocena pozycji w grze.

Metody są analogiczne do poprzednich, z tym że bardziej złożone są metody liczenia kosztów i wartości funkcji h .

Jeszcze inna jest sytuacja w grach negocjacyjnych.

Wtedy funkcja h musi być wielowymiarowa (f_1, f_2).

Rozpoznawanie Obrazów

Metody rozpoznawania obrazów (klasyfikacji).

Metody minimalno-odległościowe.

Metody wzorców.

Metody aproksymacyjne.

Metody probabilistyczne

Metody ciągowe.

Metody drzewowe.

Metody grafowe.

Metody oparte na sieciach neuronowych.

Metody oparte na algorytmach ewolucyjnych.

Metody oparte na teorii zbiorów rozmytych.

Metody oparte na rachunku relacji.

Metody hybrydowe.



Rozpoznawanie obrazów w przestrzeni cech

Ogólnie, odwzorowanie F dzielimy na trzy funkcje

$$c_i = D(K(R(\text{obraz})))$$

$$F = D \circ K \circ R$$

Podejmowanie
decyzji

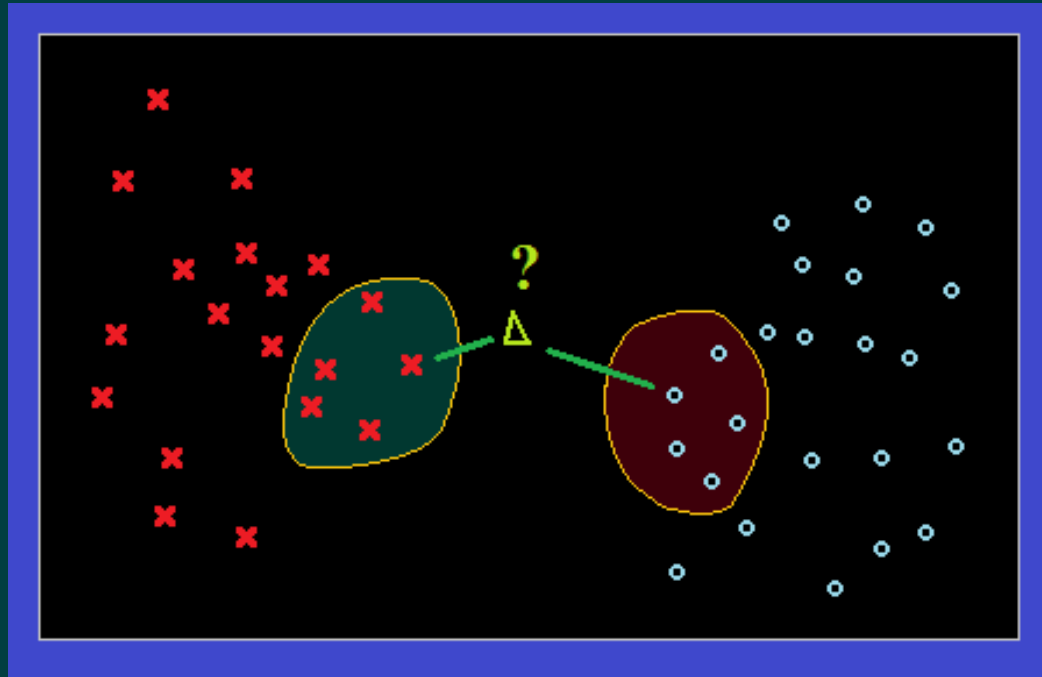
Obliczanie
funkcji
klasyfikujących

Recepcja
obliczanie cech

Bardzo dużo zależy od wyboru cech, czyli procesu recepcji.
Istotna jest także redukcja przestrzeni cech.



Metoda k najbliższych sąsiadów (k - nearest neighbours)
 k -nn.



Z każdej klasy wybiera się k -najbliższych sąsiadów (badanego obiektu) i na ich podstawie ocenia się do której klasy badany obiekt należy, zwykle stosując metodę odległościową. Metoda często stosowana, bardzo chwalona.
+ Wcześniej, proces recepcji (obliczanie wartości cech).



Metody syntaktyczne

Poprzednio opisane metody, nawet te bardzo zaawansowane matematycznie, nie są skuteczne podczas bardziej złożonych sytuacji, zwłaszcza podczas rozpoznawania hierarchicznego, lub choćby gdy liczba klas jest bardzo duża. W takich przypadkach często używane są metody wielopoziomowe, syntaktyczne.

Syntaktyka (składnia) – to dział gramatyki naukowej, zajmujący się budową wypowiedzi. Gałąź semiotyki logicznej zajmująca się funkcjami syntaktycznymi, relacjami między wyrażeniami logicznymi, a innymi wyrażeniami językowymi (analogicznie jest w rozpoznawaniu obrazów).

Semiotyka – syntaktyka, semantyka, pragmatyka (w logice).

Metody te są opisywane językiem automatów i gramatyki (złożone).

Pamięć

Procesy
uczenia się

Czym jest proces uczenia ?

w dziedzinie informatyki i matematyki :

- * Korygowanie informacji już znanej
- * Powiększanie zbioru znanej informacji
- * Modyfikacja procedur działania

W teorii, uczenie maszynowe ma prowadzić do :

- * Tworzenie klasyfikatorów (rozpoznawanie nowych obiektów).
- * Formułowania coraz lepszych reguł decyzyjnych.
- * Tworzenie coraz lepszych reguł proceduralnych (działanie).
- * Tworzenia nowych pojęć, także tych bardziej abstrakcyjnych.
- * Wykrywania nieznanych prawidłowości w danych
- * Przyswajania nowych pojęć i struktur (uogólnienia i analogie).
- * Modyfikowania, uogólniania i precyzowania danych
- * Zdobywanie wiedzy poprzez interakcję z otoczeniem
- * Formułowanie wiedzy zrozumiałej dla człowieka, dla maszyny.

Różne typy wiedzy :

(Prof. Włodzisław Duch).

Wiedza o obiektach. (jak: psy, ptaki, skrzydła, domy, ludzie).
(jest to najprostszy rodzaj pamięci, także relacje obiektów).

Wiedza o zdarzeniach. (działanie, ptak macha skrzydłami).
(samochód jedzie, samolot lata, ptak też i macha skrzydłami).

Wiedza „jak”, umiejętności. (wiemy jak postępować ...).
(to zwykle typ wiedzy praktycznej, pisanie, naprawianie czegoś).

Meta-wiedza – czyli wiedza o samej wiedzy.

(jest to zwykle opis z punktu widzenia wyższego poziomu, reguły).
(meta-wiedza jest głównie charakterystyczna dla nauki).

Przekonania. Wiedza oparta na podobieństwach, emocjach.
(wiedza nie poparta żadnymi dowodami, przesłankami).

Psychologia jednak dzieli zwykle pamięć na nieco inne kategorie. (*).

- * Sieci semantyczne
- * Ramy (klatki, procedury demoniczne)
- * Syntaktyczna analiza wzorców
- * Generacja reprezentacji strukturalnych
- * Analiza reprezentacji strukturalnych
- * Interpretacja reprezentacji strukturalnych
- * Indukcja gramatyki generatywnej
- * Wielowymiarowe gramatyki generatywne
- * Systemy regułowe , inne wnioskujące.

Ontologia

W informatyce, termin ontologia pojawił się po raz pierwszy w 1969, ale dopiero pod koniec lat dziewięćdziesiątych termin ten nabrał rozgłosu w związku z tworzeniem nowoczesnych modeli reprezentacji wiedzy (sieci semantyczne, strukturalne, regułowe, atrybutowe, ramowe) głównie w dziedzinie sztucznej inteligencji. Ostatnio pojęcie ontologii staje się coraz ważniejsze w dziedzinie sztucznej inteligencji i nauk kognitywnych.

Ontologia

- * Ontologia musi uchwycić rzeczywistość na różnych poziomach atomizacji, jak również relacje pomiędzy tymi warstwami.
- * Ontologia jest oderwana od teorii poznania (epistemologii), powiązana jest z obiektem, a nie jego subiektywnym odbiorem.
- * Ontologia nie stanowi listy, katalogu czy taksonomii obiektów, stwarza natomiast formalne przesłanki, wedle których takowe mogą być budowane.
- * Obecnie uznany jest brak możliwości stworzenia jednej ogólnej ontologii, istnienie wielu ontologii.
- * W przeciwieństwie do nauki, relacje między obiektami nie są ujęte funkcyjnie (zależności nie są ilościowe).
- * Nauka rozpoczyna proces od mierzenia i predykcji, ontologia zaś od budowania taksonomii.

Psychologia Poznawcza

Percepcja

Wyobrażenia

Pamięć

Uwaga

Uczenie się

Myślenie , rozumowanie

Rozwiązywanie problemów

Podejmowanie decyzji

Język i komunikacja

Spostrzeganie

Konstruktywizm – doświadczalna rzeczywistość jest tworzona w naszym umyśle. Jak jest naprawdę, tego nie możemy wiedzieć. Wiemy tylko tyle, że stymuluje do tworzenia obrazów, wrażeń w naszym umyśle.

Podejście ekologiczne – spostrzegany świat, jest właśnie taki, jaki jest w swej istocie. Żywe organizmy w procesie ewolucji, tak dostroiły swoje analizatory, że są w stanie odbierać cechy fizyczne w taki sposób, w jaki występują one w przyrodzie.

► **Myślenie**

✠ **Rozumowanie**

❖ **Rozwiązywanie problemów**

► Myślenie

Myślenie to proces łączenia elementów poznawczej reprezentacji świata (obrazów, pojęć, sądów) w dłuższe ciągi. Reprezentacje poznawcze mogą być trwałe lub nietrwałe.

Trwałe – pojęcia, sądy, fakty, znane relacje.

Nietrwałe – spostrzeżenia, wyobrażenia.

Jednak zawsze są strukturami poznawczymi, a nie procesami. Jeśli owe struktury poznawcze ulegają dalszemu przetwarzaniu, to mamy do czynienia z procesem myślenia.

✠ Rozumowanie

Rozumowaniem nazywamy proces umysłowy, polegający na formułowaniu wniosków na podstawie przesłanek.

Przesłanki nie muszą być wcale pewne ani udowodnione, czasami podstawą rozumowania stają się hipotetyczne założenia albo potoczne obserwacje.

Rozumowanie dzielimy na **dedukcję i indukcję**.

Dedukcja – cała wiedza jest zawarta w przesłankach, trzeba ją tylko przedstawić w postaci twierdzenia.

Indukcja – Na podstawie skończonej liczby przypadków formułujemy ogólniejsze wnioski, ogólniejszą teorię.



❖ Rozwiązywanie problemów

Problem to rozbieżność między stanem wyjściowym a pożądanym lub narzuconym stanem docelowym (której nie można zredukować rutynowo).

Rozwiązywaniem problemu nazywamy aktywność, która zmierza do zredukowania rozbieżności między stanem aktualnym (wyjściowym) a stanem pożądanym docelowym.

- * Istota i typy sytuacji problemowych.
- * Konstrukcja modelu umysłowego (prosta , złożona).
- * Etapy rozwiązywania.
- * Cztery fazy : preparacja, inkubacja, iluminacja, elaboracja.
- * Algorytmy i heurystyki.

Sieci Neuronowe

Badanie Układu Nerwowego

Odkrycie komórek nerwowych w mózgu

Badanie potencjałów elektrycznych komórki

Odkrycie transmisji synaptycznej

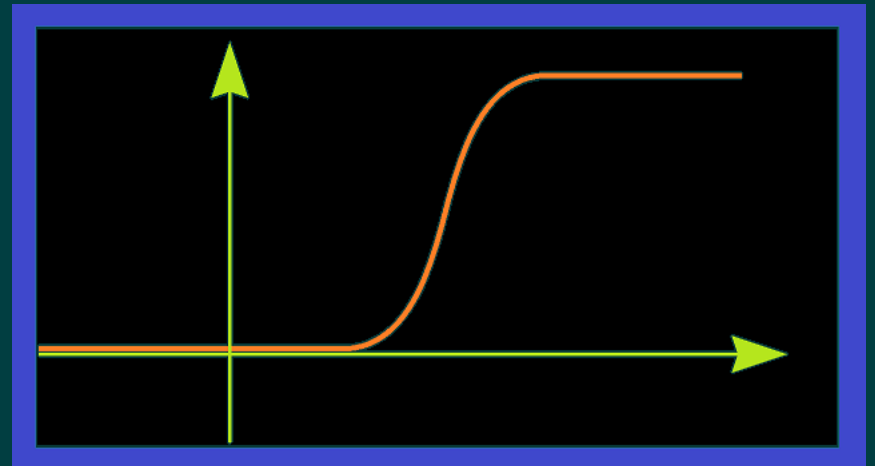
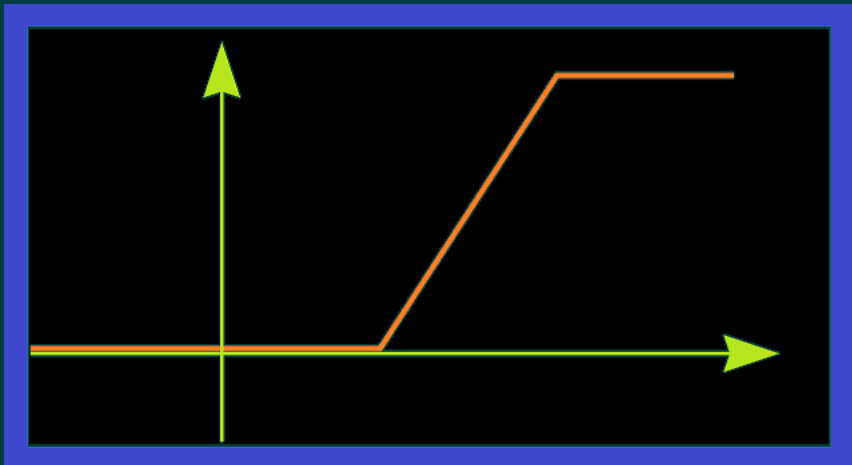
Badanie inicjacji potencjału czynnościowego

Reguła **Hebba**

Powstawanie następnego
potencjału czynnościowego
jest modelowane funkcją aktywacji f .

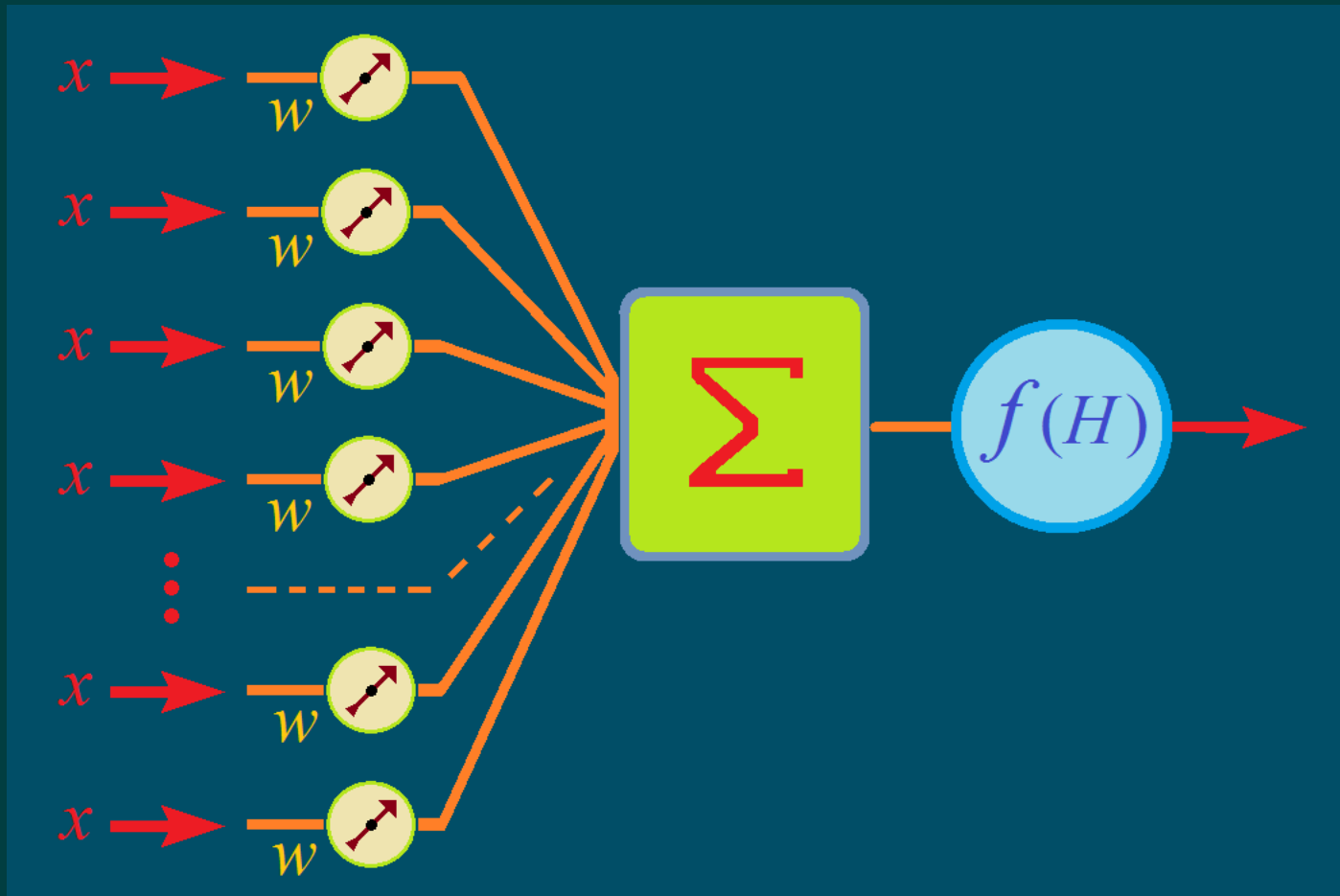
$$y = f(H) = f\left(\sum w_i x_i\right) ;$$

Stosowane są różne modele funkcji aktywacji.
Najczęściej są to różne funkcje sigmoidalne :



Perceptron , Rosenblatt (1968).

Perceptron też można traktować jako model neuronu.
Jednak istotną nowością są metody uczenia takiej struktury.



Reguła Hebba !!

Dwie komórki (lub systemy komórek), które aktywują się wielokrotnie w tych samych chwilach, mają tendencję do „stowarzyszania się”, tak żeby aktywność jednej ułatwiała aktywność drugiej.

Okazało się, że jest to przyczyną zmian własności transmisyjnych synaps. Gdy aktywują się jednocześnie, to powstaje wzmocnienie dodatnie, a gdy się aktywują oddzielnie to wzrasta hamowanie.

Modelowanie reguły Hebba : $\Delta w_{ij} = \eta x_i x_j$;



Podstawowa zasada działania (dynamika) dla sieci Hopfielda synchronicznej.

i – indeks komórki (neuronu) w sieci.

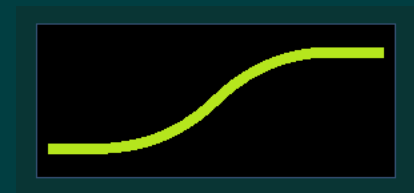
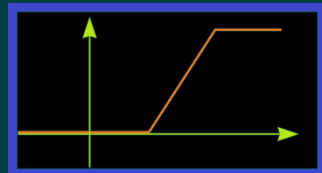
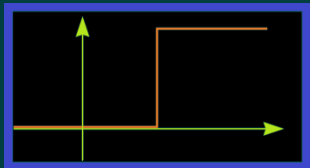
k – indeks czasowy, numer kroku czasowego.

x_{ik} – stan i -tej komórki w sieci, w k -tym kroku czasowym.
(reprezentuje częstość potencjałów czynnościowych).

f – funkcja aktywacji.
(reprezentuje proces inicjacji potencjału czynn.).

$$x_{i,k+1} = f \left(\sum w_{ij} x_{jk} \right) ;$$

Funkcja aktywacji f może być skokowa, liniowa, lub różniczkowalna.



Zasada dla sieci synchronicznej.

Wszystkie neurony zmieniają stan jednocześnie.

i – indeks komórki (neuronu) w sieci.

k – indeks czasowy, numer kroku czasowego.

x_{ik} – stan i -tej komórki w sieci, w k -tym kroku cz.

w_{ij} – waga połączenia komórki i -tej z j -tą. ($j \rightarrow i$).
(reprezentuje częstość potencjałów czynnościowych).

f – funkcja aktywacji.

(reprezentuje proces inicjacji potencjału cz.).

$$x_{i,k+1} = f \left(\sum w_{ij} x_{jk} \right) ;$$

Wzór dotyczy sieci Hopfilda, ale może być stosowany dla innych struktur, tylko trzeba zaznaczyć, że wtedy indeks j numeruje neurony połączone z i -tym neuronem w podsieci (neuronu i -tego).

Dla sieci asynchronicznych dynamika musi być określona w bardziej złożony sposób

Wprowadzamy pojęcie konfiguracji sieci $u(k)$.
Konfiguracja – to stany wszystkich neuronów całej sieci w ustalonym kroku czasowym $[x_{1k}, x_{2k}, \dots, x_{nk}]$.

Zmiana konfiguracji jest określona odwzorowaniem F .

$$u(k+1) = F(u(k)) ;$$

W przypadku dynamiki sieci asynchronicznej, należy najpierw wybrać neuron lub grupę neuronów, które właśnie w tym kroku (k) mają zmienić swój stan.

Dla sieci asynchronicznych dynamika musi być określona w bardziej złożony sposób

$$u(k+1) = F(u(k))$$

Określenie neuronu (lub grupy neuronów), zmieniają stany.
Indeks $i = p(k)$; gdzie p jest ciągiem naturalnym, przyporządkowującym krokom czasowym k indeksy neuronów, może to być permutacja z funkcji $k \bmod n$.
Dynamikę sieci określamy następująco :

$$x_{i,k+1} = f \left(\sum w_{ij} x_{jk} \right) / \quad \text{dla } i = p(k) ;$$

$$x_{i,k+1} = x_{i,k} / \quad \text{w pozostałych przypadkach} ;$$

Dynamika sieci asynchronicznych

Neurony mogą być wybierane losowo !!

$$u(k+1) = \underline{F}(u(k))$$

Jeszcze inna dynamika sieci :

Indeks i jest wybierany losowo z prawdopodobieństwem $\mathbf{P}(j=i)$, które zwykle jest równe $1/n$ (równe szanse).

$$x_{i,k+1} = f\left(\sum w_{ij} x_{jk}\right) \quad / \quad \text{dla } i = p(k) ;$$

$$x_{i,k+1} = x_{i,k} \quad / \quad \text{w pozostałych przypadkach ;}$$

W takim przypadku, już trzeba użyć łańcuchów Markowa !!
Model sieci staje się już modelem częściowo losowym.

Energia Sieci Hopfielda

$$E = \frac{1}{2} \sum_{i,j} w_{ij} x_i x_j \quad ;$$

Jeżeli uwzględnimy indeks czasowy k , to wzór będzie :

$$E_k = \frac{1}{2} \sum_{i,j} w_{ij} x_{ik} x_{jk} \quad ;$$

Wraz z „ruchem sieci”, Energia sieci stale maleje.

Należy porównać dwie sytuacje (przed i po zmianie stanu).

Niech x_{ik} będzie już wartością stanu po zmianie, czyli zgodnie z wzorem zmiany, to inna wartość zwiększa energię.

Metoda wstecznej propagacji błędu

Polega na minimalizacji odległości – wzorców od badanych obiektów na wyjściu sieci. W ten sposób oblicza się wagi.

Zwykle ten proces polega na metodzie gradientowej, a wtedy warunkiem koniecznym jest różniczkowalna funkcja aktywacji neuronów. Obliczane są pochodne cząstkowe.

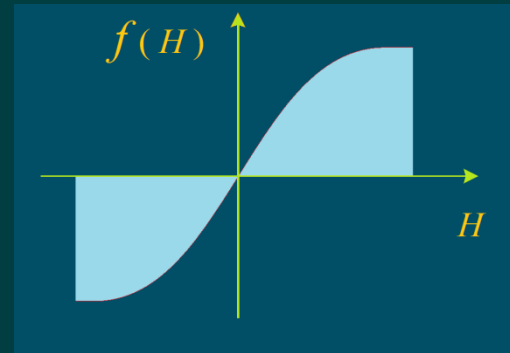
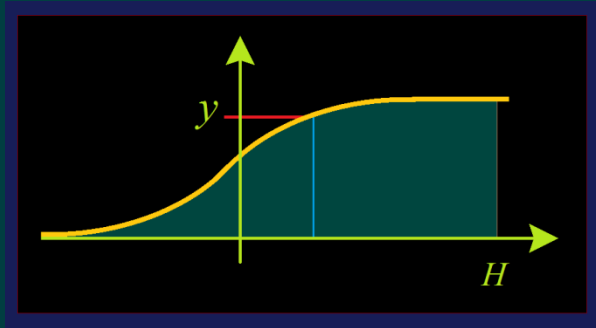
$$D(w) = \|x_j - q_j\| \quad ;$$

$$\Delta w_{ij} = - \delta \cdot \partial D / \partial w_{ij} \quad ;$$

W sieciach wielowarstwowych powstają złożone wzory

Maszyna Boltzman

Przykład sieci probabilistycznej. Funkcja aktywacji wyznacza prawdopodobieństwo zmiany stanu.



Stany neuronów zmieniają się zgodnie rozkładem prawdopodobieństwa określonym funkcją f .

Wartości stanów mogą być tylko dodatnie (z zerem) lub dodatnie i ujemne. Neurony są wybierane losowo.

Klasyfikacja architektur podstawowych znanych sieci neuronowych

Sieci dzielimy na jednokierunkowe i rekurencyjne.

Wyróżniamy też sieci jednowarstwowe i wielowarstwowe.

Stosowane są także inne nietypowe, złożone struktury.

(takie jak sieci lokalne, częściowo lokalne, mniejsze struktury łączone w całość lub sieci lokalne z dodatkową siecią długich połączeń).

Zupełnie innym typem sieci, są sieci samoorganizujące się, które zmieniają swoją architekturę w trakcie uczenia się.

W przypadku sieci rekurencyjnych potrzebna jest dynamika.

Dynamika – sieci synchroniczne i asynchroniczne.



Podstawowe metody uczenia SN

- * Reguła Hebba
- * Wsteczna propagacja błędu
- * Metoda konkurowania neuronów
- * Metody probabilistyczne
- * Inne metody hybrydowe



Problemy filozoficzne SI

❖ Problem chińskiego pokoju - Johna Searle'a.

❖ Stanowisko Patrycji i Paula Churchland.

Możliwe jest, że wielka maszyna równoległa
może generować świadomość .. !! ??

Jest to nadal nie rozwiązana hipoteza !

Na koniec warto przytoczyć stanowisko Penrose'a.

Roger Penrose

(Fizyk, matematyk, filozof, badacz AI)

Myślenie zawsze polega na obliczeniach, a w szczególności świadome doznania powstają wskutek realizacji odpowiedniego procesu myślowego.

Świadomość jest cechą fizyczną działającego mózgu. Wprawdzie wszystkie fizyczne procesy w mózgu można symulować obliczeniowo, symulacjom obliczeniowym jednak nie towarzyszy świadomość.

Odpowiednie procesy fizyczne w mózgu powodują powstanie świadomości, ale tych procesów nie można nawet symulować.

Świadomości nie można wyjaśnić w żaden fizyczny, obliczeniowy czy inny naukowy sposób.

Egzamin

Cztery zadania
na jednej kartce

Każde zadanie po 10 punktów

Koniec