Inteligentne Systemy Obliczeniowe

dr inż. Piotr Wąsiewicz

Samodzielna Pracownia Sztucznej Inteligencji

Instytut Systemów Elektronicznych, Politechnika Warszawska

http://staff.elka.pw.edu.pl/~pwasiewi

pwasiewi@elka.pw.edu.pl

Literatura

1. M.J. Kasperski, "Sztuczna Inteligencja", Helion, 2003
2. J.J. Mulawka, "Systemy Ekspertowe", PWN, 1996
3. P. Cichosz, "Systemy uczące się", WNT, 2000
4. L. Bolc, W. Borodziewicz, M. Wójcik "Podstawy przetwarzania informacji niepewnej i niepełnej", seria Współczesna Nauka i Technika - Informatyka, PWN, 1991
5. R. Rychlik, M. Wójcik, "Od logiki do reprezentacji wiedzy", WNT - Informatyka, PWN
6. A. Skowron, "Podstawy Sztucznej Inteligencji", WNT - Informatyka, PWN
7. L. Bolc, Zaremba, "Wprowadzenie do uczenia się maszyn", PWN, 1993
8. S. Russel, P. Norvig, "Artificial Intelligence: A Modern Approach", Prentice-Hall, 1995

Definicje Sztucznej Inteligencji

Sztuczna Inteligencja - nauka o budowaniu maszyn robiących rzeczy, które, jako skonstruowane przez człowieka, wymagałyby inteligencji (M. Minsky)

Stworzenie obrazu myślącej ludzkiej istoty - stworzenie człowieka elektronicznego

Sztuczna Inteligencja - jest radykalnym wyrazem możliwości komputera cyfrowego, jest pochwałą nowej technologii

Pierwsze kroki

Pierwsze kroki w definiowaniu właściwości umysłu:

- Arystoteles - człowiek jest zwierzęciem wyposażonym w *logos* tnz. mówienie lub pojmowanie, czy też myślenie logiczne, łaciński odpowiednik *ratio* oznacza już tylko obliczanie.
- Kartezjusza *res cogitans* tzn. software (umysł) oraz *res extensa* tzn. hardware (ciało). Zwierzęta są maszynami jak mechaniczne lalki.
- Kartezjusza *sensus communis* zmysł wspólny według Lema inteligencji nie da się wytworzyć w zamkniętym środowisku (mózg w słoju tylko śniłby).
- La Mettrie - umysł konsekwencją skomplikowania materii, ale sam język odróżnia człowieka od zwierząt.
- Dalgarna "Sztuka znaków" mówi o uniwersalnym języku wszystkich ludzi np. jeśli n oznacza "żywą istotę", e - "zwierzę", k - "czworonoga", to $neke$ jest odpowiednikiem słowa "koń", $neki$ to "osioł" itd.

Rozwój koncepcji

Kolejne etapy realizacji sztucznego umysłu:

- Leibniz - maszyna myśląca w sensie Leibniza systematyzuje wiedzę tworząc odpowiedni język. Ma pamięć, sensory i uczy się (dodaje nowe obiekty) oraz przeprowadza dowody.
- Lady Ada Lovelace - maszyna licząca jak mechaniczne krosno "tka" wzory i nigdy nie wychodzi poza program.
- Boole tworzy swoją algebrę operującą na liczbach w systemie dwójkowym i staje się ona podstawą najbardziej uniwersalnego języka na świecie, ale języka maszyn.
- Godel ogłasza że dla dowolnego systemu formalnego, który oznaczymy jako M, zawierającego część arytmetyki liczb naturalnychjest możliwe skontsruowanie w języku systemu M takiego zdania, które nie tylko nie da się udowodnić w tym M, ale jego negacja pozostanie także bez dowodu.
- Turing tworzy model teoretyczny każdego komputera tzw. *Maszynę Turinga* podobną do rybosomu organelli komórkowej czytającej nić rna (dna) i tworzącej odpowiednie łańcuchy peptydów (białka).

Test Turinga: nagroda Loebnera

Warunki testu:

- Mamy co najmniej dwóch ludzi i maszynę myślącą zdolną do konwersacji.
- Jeden z ludzi jest sędzią i nie widzi maszyny ani innych ludzi, z którymi porozumiewa się tylko słowami.
- Jeśli sędzia po dowolnej ilości jakichkolwiek pytań, dojdzie do wniosku, że rozmawia tylko z ludźmi, to wspomniana maszyna jest maszyną myślącą w sensie Turinga.
- Nagroda Loebnera jest otrzymywana za najbardziej "ludzki" komputer ang. most human computer. Test Turinga ograniczony jest do jednej dziedziny wiedzy.

Wydruk dialogu z programem ELIZA

Patient: Men are all alike.

ELIZA: In what way?

Patient They're always bugging us about something or other.

ELIZA: Can you think of a specific example?

Patient: Well, my boyfriend made me come here.

ELIZA: Your boyfriend made you come here?

Patient: He says I'm depressed much of the time.

ELIZA: I am sorry you are depressed.

Współczesne technologie i idee

Współczesne idee:

- Von Neumann tworzy model sekwencyjnej maszyny zwanej *maszyną von Neumanna*, gdzie program i dane są przechowywane w stałej pamięci w postaci binarnej.
- Lem przytacza przykład planety gramofonu z gigantyczną pamięcią znającą odpowiedź na wszystkie pytania (np. *Deeper Blue* z programem gigantem tworzonym 6 lat, który wygrał z Kasparowem), ale z drugiej strony wspomina o opowiadaniu Dnieprowa, gdzie tranzystory zastępują ludzie i podczas procesu tłumaczenia pojedyńczy tranzystor nic nie wie o samym tłumaczeniu, gdyż przekazuje pojedyńcze litery i operacje na nich tzw. późniejszy argument *Chińskiego Pokoju Searle'a*.
- Jednak mimo nieświadomości neuronów mózg ludzki ma świadomość jako całość tzw. późniejszy argument *Jasnego Pokoju* (poruszanie magnesem w pokoju nie generuje światła, ale może jest zbyt wolne, to samo stosujemy do obliczeń i komputerowych architektur czyli magnetyzm i elektryczność jako programy wystarczą razem z energią jako składnią zdań do udowodnienia istoty światła tutaj inteligencji.

Współczesne technologie i idee

Współczesne idee:

- Ashby stwierdza, że wszystkie układy przekazujące informację w ilości powyżej 10^8 bitów na sekundę powinny być świadome. Dzisiaj istnieją już takie komputery, ale daleko im jeszcze do inteligencji człowieka.
- Adleman w laboratorium inżynierii genetycznej konstruuje komputer oparty na DNA i jego przetwarzaniu. W pojedyńczej probówce mieści się około 10^{13} odcinków DNA z zakodowaną informacją przetwarzaną masywnie równolegle.
- Bariera obliczeń - problemy NP-trudne (Nondeterministic polynomial). Maszyna deterministyczna i niedeterministyczna, a równoległość i czas obliczeń. Pamięć asocjacyjna.
- Problemy NP-zupełne - zadanie spełnialności zdania logicznego (funkcji logicznej).

Maszyna myśląca

Stanowiska według Rogera Penrose'a (od "silnej AI" do "słabej AI"):

- $\mathcal{A}$ Myślenie to po prostu obliczenia, a świadome doznania to wynik tych obliczeń (test Turinga wystarczający).
- $\mathcal{B}$ Symulacje komputerowe świadomości nie mają nic wspólnego z samą świadomością.
- $\mathcal{C}$ Procesów fizycznych w mózgu nie da się zasymulować (brak dokładniejszych lub nowych praw fizyki).
- $\mathcal{D}$ Świadomości nie da się wyjaśnić w żaden obliczeniowy i naukowy sposób (agnostycyzm).

Definicje Sztucznej Inteligencji

Celem Sztucznej Inteligencji jest stworzenie systemów:

myślących jak ludzie tzn. formułujących w podobny sposób myśli np. GPS.	*myślących rozumnie* tzn. formułujących myśli z pomocą komputerowych modeli np. systemy ekspertowe.
działających jak ludzie tzn. o reakcjach wyglądających tak samo np. Eliza.	*działających rozumnie* tzn. podających suboptymalne, satysfakcjonujące rozwiązania np. algorytmy genetyczne.

Sztuczny człowiek?

Inteligentny jak człowiek program powinien

- komunikować się np. po angielsku,
- gromadzić wiedzę,
- wysnuwać na jej podstawie wnioski,
- korzystając z doświadczenia dostosowywać się do zmieniających się warunków uzupełniając wiedzę nowymi wnioskami
- oraz wykorzystywać zaawansowane systemy robotyki i wizji.

Podstawowe zagadnienia sztucznej inteligencji

- Programy do prowadzenia dialogu z maszyną np. ELIZA
- Programy do rozwiązywania problemów np. GPS
- Systemy ekspertowe
- Pozyskiwanie wiedzy
- Uczenie się maszyn

Zestawienie najważniejszych osiągnięć w okresie rozwoju metod sztucznej inteligencji

Okres	Kluczowe osiągnięcia
Lata przed II wojną światową	Logika formalna, psychologia poznawcza
Lata powojenne 1945-1954	Powstanie komputerów, rozwój cybernetyki
Rozpoczęcie badań w dziedzinie sztucznej inteligencji 1955-1970	Rozwój komputerów, LISP, sformułowanie programu ogólnego rozwiązywania problemów
Badania w dziedzinie rozwiązywania problemów 1961-1970	Heurystyki, robotyka, programy do gry w szachy
Systemy oparte na bazach wiedzy 1971-1980	MYCIN, HEARSAY II, MACSYMA, EMYCIN, Prolog
Po 1981 r. liczne zastosowania praktyczne	PROSPECTOR, nie zrealizowany japoński projekt komputerów piątej generacji, powstanie wielu firm zajmujących się zastosowaniem sztucznej inteligencji

Podział dziedzin sztucznej inteligencji

- rozwiązywanie problemów i strategie przeszukiwań
- teoria gier
- automatyczne dowodzenie twierdzeń
- przetwarzanie języka naturalnego (włączając przetwarzanie mowy)
- systemy ekspertowe
- robotyka
- procesy percepcji (wizja, słuch, dotyk)
- uczenie się maszyn
- wyszukiwanie informacji (inteligentne bazy danych)
- programowanie automatyczne

Zastosowania Sztucznej Inteligencji

- Program szachowy z komputera Deep Blue pokonał mistrza świata Gary Kasparova.
- Program PEGASUS rezerwuje miejsca w amerykańskich liniach lotniczych słuchając poleceń klientów.
- Program ALVINN może w każdych warunkach atmosferycznych kierować ciężarówką np. przejechał nią z Washingtonu do San Diego.
- Inteligentne programy rozpoznają twarze np. w bankach, odręczne pismo, sprawdzają lub projektują układy elektroniczne np. EURISKO, rekonstruują projekty architektów, szuka złóż geologicznych np. PROSPECTOR, interpretuje związki chemiczne np. SCANMAT, DENDRAL.
- Programy zwane systemami ekspertowymi pomagają lub są lepsze w diagnozach lekarskich np. MYCIN, CADUCEUS, CASNET, Intellipath, Pathfinder; konfigurują sprzęt komputerowy np. XCON; pomagają w podejmowaniu finansowych decyzji znajdując zdefraudowane, nietypowe lub błędne transakcje np. AMEX credit check.
- Programy mogą udowadniać matematyczne twierdzenia, tłumaczyć na języki obce np. Altavista, planować procesy produkcyjne, operacje w trudnych warunkach np. DART.

Pojęcia podstawowe

Symbol - encja reprezentująca element ze zbioru znaczeń zdefiniowanych a priori

Dane - zapisany zbiór symboli

Informacja - dane z przypisanym znaczeniem

Pojęcie - zbiór encji z jakiegoś powodu zunifikowany

Język - zbiór pojęć i reguł do tworzenia opisu rzeczywistości

Opis - wyrażenie w pewnym języku charakteryzujące obiekt lub zbiór obiektów

Wiedza - zorganizowana, uogólniona i/lub abstrakcyjna informacja

Definicje wiedzy

Wiedza deskrypcyjna to - *opisy* obiektów, ich klasy

Wiedza preskrypcyjna to - *procedury* opisujące dopuszczalne operacje, jakie można dokonać na relacjach, funkcjach tzw. przepisy

Wiedza to zbiór *faktów, reguł, domniemań* (ang. believes - fakty i reguły nie w pełni wiarygodne), *heurystyk*

Wiedza może być *prywatna* (np. inżyniera architekta), *publiczna* (ogólnodostępna), *ściśle tajna*

Wiedza może być *płytka* (opiera się na rozpoznaniu np. stylu architektury danego budynku), *głęboka* (sięga głębiej, opiera się na regułach np. dokładne poznanie wymiarów i materiałów użytych w konstrukcji budynku)

Książki to wiedza starego typu w formie *pasywnej*. Zanim zostanie ona użyta, musi być pobrana, a następnie odpowiednio zinterpretowana po czym trzeba zadecydować jak ją wykorzystać do efektywnego rozwiązywania problemu.

Rodzaje wiedzy

Wiedza	Zakres	Cel(sposób)	Ważność
Pies jest ssakiem	specific	descriptive	certain
Pies ma cztery łapy	specific	descriptive	uncertain
Aby wykazać, że X jest psem należy pokazać, że rodzice X są psami	specific	prescriptive	certain
Aby udowodnić $P(X)$, wykaż, że $\neg P(X)$ jest niemożliwe ($\neg\forall x\ P(X) \Leftrightarrow \exists x\ \neg P(X)$)	general	prescriptive	uncertain
Rzeczy - obiekty rzeczywiste są obserwowalne	general	descriptive	uncertain (nie oznacza to, że zawsze je widać)

Reprezentacje wiedzy

Reprezentacja proceduralna - polegająca na określeniu zbioru procedur, których działanie reprezentuje wiedzę o dziedzinie np. $V = \frac{4}{3}\pi r^3$

Reprezentacja deklaratywna - polegająca na określeniu zbioru specyficznych dla rozpatrywanej dziedziny faktów, stwierdzeń i reguł (np. katalog rzeczy)

Zaletą reprezentacji proceduralnej jest wysoka efektywność reprezentowania procesów.
Zaletą reprezentacji deklaratywnej jest to, że jest ona bardziej "oszczędna" (każdy fakt lub reguła zapisywany tylko raz) i łatwiejsza w formalizacji.
Jako rozwiązanie optymalne można uznać reprezentację łączącą w sobie cechy reprezentacji proceduralnej i deklaratywnej np. ramy, języki obiektowe.

Metody reprezentacji wiedzy

- Zastosowania logiki (rachunek zdań, rachunek predykatów, syntaktyka, semantyka)
- Zapis twierdzeń, zapis reguł w systemach ekspertowych (schemat rezolucji na klauzulach Horna, wnioskowanie w przód i wstecz)
- Wiedza nieprecyzyjna (teoria Bayesa, współczynniki niepewności w systemie MYCIN, teoria Dempstera-Shafera)
- Teoria zbiorów przybliżonych (tablice warunkowo-działaniowe, relacje nierozróżnialności, klasyfikacje, aproksymacja dolna i górna, reguły pewne i możliwe)
- Teoria zbiorów rozmytych (funkcja przynależności, liczby rozmyte, relacje rozmyte)
- Sieci semantyczne
- Algorytmy genetyczne i sieci neuronowe

Język logiki I rzędu: syntaktyka

Wyrażenia języka składają się z term i formuł.

1. termy - encje, obiekty
   - symbole stałych (zwykle z początku alfabetu): $a, b, \ldots$
   - symbole zmiennych
   - n-argumentowe symbole funkcyjne $f(t_1, \ldots, t_n)$, gdzie $t_1, \ldots, t_n$ to termy

   np. $a, f(g(x, b), c)$ to termy zamknięty (bez zmiennych) oraz otwarty (ze zmiennymi), termy mogą być z indeksami: a_1, f_3^n, gdzie n to ilość argumentów funkcji

Język logiki I rzędu: syntaktyka

2. formuły - fakty zachodzące w świecie

- formuły atomowe - symbole relacji 0-argumentowych zwanych stałymi zdaniowymi oraz relacje n-argumentowe oznaczane P tzn. $P(t_1, \ldots, t_n)$, gdzie t_n to termy dla $n \geqslant 1$
- formuły - z formuł atomowych, $\neg$, $\Rightarrow$, $\forall$

i) α, β

ii) $(\neg\alpha)$

iii) $(\alpha \Rightarrow \beta)$

iv) $(\forall x\ \alpha)$, gdzie x jest zmienną

Język logiki I rzędu: syntaktyka

Relacja n-argumentowa oznaczana literą P jest zwana predykatem np. predykat P związany jest z pojęciem jakiejś konkretnej rzeczy tzn. $P(a)$ np. jest symbolem rzeczy osoby oznaczonej termem a np: Joanny.

Literałem pozytywnym jest α, a negatywnym $\neg\alpha$.

Język logiki I rzędu: syntaktyka

Korzystając z podanych definicji tworzenia formuł rozszerza się zbiór spójników i kwantyfikatorów języka poprzez następujące definicje:

- $(\alpha \vee \beta) = ((\neg\alpha) \Rightarrow \beta)$
- $(\alpha \wedge \beta) = (\neg((\neg\alpha) \vee (\neg\beta)))$
- $(\exists x\ \alpha) = (\neg(\forall x(\neg\alpha)))$

Symbol $\exists$ jest kwantyfikatorem szczegółowym (egzystencjalnym), $\alpha \vee \beta$ - alternatywą formuł, $\alpha \wedge \beta$ - koniunkcją formuł.

Język logiki I rzędu: syntaktyka

Formuła zamknięta zwana także zdaniem lub formułą zdaniową jest formułą bez zmiennych wolnych (zmiennych nie związanych z kwantyfikatorem $\forall$ lub $\exists$) w przeciwieństwie do formuły otwartej np. $P(x, y) \Rightarrow \exists x \forall z P(x, y)$ jest formułą otwartą.

W celu poprawienia czytelności można pomijać także nawiasy kierując się następującą listą - od najmocniej do najsłabiej wiążących - spójników i kwantyfikatorów: $\neg\ \forall\ \exists\ \wedge\ \vee\ \Rightarrow$

Przykłady zdań logicznych

Piotr nie jest wysoki.

$\neg$wysoki(Piotr)

Na stole leży tylko owoc.

$\forall x$ na$(x,$stół$) \Rightarrow$ owoc(x)

Liczba całkowita może być parzysta i nieparzysta.

$\forall x$ całkowita$(x) \Rightarrow$ (parzysta$(x) \vee$ nieparzysta(x))

Wszyscy studenci są zdolni.

$\forall x$ student$(x) \Rightarrow$ zdolny(x)

Każdy na świecie student jest zdolny.

$\forall x$ student$(x) \wedge$ zdolny(x)

Co niektóry student jest zdolny.

$\exists x$ student$(x) \wedge$ zdolny(x) OK!

$\exists x$ student$(x) \Rightarrow$ zdolny(x) Zła składnia!

Przykłady formuł logicznych

Każdy delfin jest ssakiem.

$\forall x$ delfin$(x) \Rightarrow$ ssak(x)

Istnieje ssak, który znosi jaja.

$\exists x$ ssak$(x) \wedge$ znosi_jaja(x) OK!

$\exists x$ ssak$(x) \Rightarrow$ znosi_jaja(x) Zła składnia!

Każdy ogrodnik lubi słońce.

$\forall x$ ogrodnik$(x) \Rightarrow$ lubi$(x,$ słońce$)$

Wszystkie czerwone grzyby są trujące.

$\forall x($grzyb$(x) \wedge$ czerwony$(x)) \Rightarrow$ trujący(x)

Żaden czerwony grzyb nie jest trujący.

$\neg\exists x$ czerwony$(x) \wedge$ grzyb$(x) \wedge$ trujący(x)

$\forall x($grzyb$(x) \wedge$ czerwony$(x)) \Rightarrow \neg$trujący(x)

Są dokładnie dwa czerwone grzyby.

$\exists x \forall y$ grzyb$(x) \wedge$ czerwony$(x) \wedge$ grzyb$(y) \wedge$ czerwony$(y) \wedge \neg(x = y) \wedge \forall z($grzyb$(z) \wedge$ czerwony$(z)) \Rightarrow ((x = z) \vee (y = z))$

Przykłady zdań logicznych

Każdy lubi kogoś.

$\forall x \exists y$ lubi(x, y)

Ktoś lubi każdego.

$\exists x \forall y$ lubi(x, y)

Możesz kochać niektórych ludzi cały czas.

$\exists x \forall t($osoba$(x) \wedge$ czas$(t)) \Rightarrow$ można_kochać(x, t)

Można kochać każdego człowieka przez pewien okres czasu.

$\forall x \exists t($osoba$(x) \wedge$ czas$(t)) \Rightarrow$ można_kochać(x, t)

Istnieje student, który interesuje się co najmniej dwoma różnymi przedmiotami wykładanymi na jego wydziale.

$\exists x($student$(x) \wedge \exists y \exists z (y \neq z \wedge$wykładany$(y,$ wydział$(x)) \wedge)$wykładany$(z,$ wydział$(x)) \wedge$ interesuje_się$(x, y) \wedge$ interesuje_się$(x, z)))$

Semantyka

Semantyka logiczna zwana *teorią modeli* opisuje związki pomiędzy językiem, a fragmentem lub fragmentami "świata rzeczywistego". W logice fragmenty takie nazywane są strukturami.

Semantyka - struktura logiczna

$$S = (D, \mathbf{F}, \mathbf{R}, C),$$

gdzie $D \neq 0$ zwany jest dziedziną struktury, a jego elementy - obiektami struktury, $\mathbf{F}$ jest zbiorem funkcji $D^n \rightarrow D$, (R) jest zbiorem relacji w D^M, zaś C jest funkcją realizacji języka (interpretacją), która:

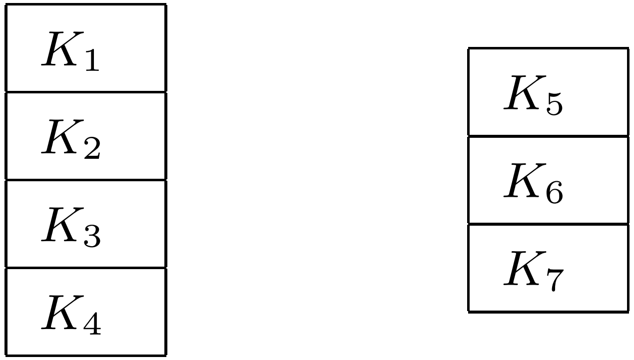
- każdemu symbolowi stałej przyporządkowuje jakiś obiekt z D,
- każdemu symbolowi funkcji n-argumentowej przyporządkowuje funkcję z $\mathbf{F}$,
- każdemu symbolowi predykatowemu przypisuje relację ze zbioru $\mathbf{R}$,
- każdej stałej zdaniowej przyporządkowuje wartości logiczne $\mathbf{0}, \mathbf{1}$, którym przypisuje się odpowiednio wartości *fałszu i prawdy*.

Przez $|P|_S$ oznacza się relację ze zbioru $\mathbf{R}$, którą funkcja C przyporządkowuje symbolowi P, tzn. $|P|_S = C(P)$.

KLOCKI

$D = (K_1, \ldots, K_7)$

góra$(K_2) = K_1$ dół$(K_1) = K_2$



góra: $D \rightarrow D$ na: $D^2 \rightarrow \{0, 1\}$

dół: $D \rightarrow D$ nad: $D^2 \rightarrow \{0, 1\}$

$$S = (D, \mathcal{F}, \mathcal{R}, C)$$

$$\Rightarrow D, \mathcal{F}, \mathcal{R}$$

$$\Downarrow$$

$C(a) = K_1, \ldots$

$C(g) = K_7$

$C(q) =$ góra

$C(h) =$ dół

$C(P) =$ na

$C(Q) =$ nad

$\Leftarrow$

<u>stałe</u>: a, b, c, d, e, f, g

<u>funkcje</u>: q, h

<u>symbole predykatowe</u>: P, Q

<u>zmienne</u>: x, y

Semantyka - interpretacja zmiennych

Każdą funkcję ω, która symbolowi zmiennej x przyporządkowuje pewien obiekt z D nazywa się wartościowaniem zmiennych w $S = (D, \mathbf{F}, \mathbf{R}, C)$, a ich zbiór to Ω_S. Interpretację termu t w S przy wartościowaniu ω oznaczamy jako $I^S_\omega(t)$.

i) $I^S_\omega(t) = C(t)$, jeżeli t jest symbolem stałej;

ii) $I^S_\omega(t) = \omega(t)$, jeżeli t jest symbolem zmiennej;

iii) $I^S_\omega(f(t_1, \ldots, t_n)) = C(f)(I^S_\omega(t_1), \ldots, I^S_\omega(t_n))$;

Semantyka - interpretacja formuł

i) $I_\omega^S(\alpha) = C(\alpha)$, jeśli α jest stałą zdaniową;

ii) $I_\omega^S(P(t_1, \ldots, t_n)) = C(P)(I_\omega^S(t_1), \ldots, I_\omega^S(t_n))$;

iii) $I_\omega^S(\alpha) = \mathbf{1} - I_\omega^S(\beta)$, jeśli α ma postać $\neg\beta$;

iv) jeżeli $\alpha = (\beta \Rightarrow \gamma)$, to $I_\omega^S(\alpha) = \mathbf{1}$, jeśli $I_\omega^S(\beta) = \mathbf{0}$ lub $I_\omega^S(\gamma) = \mathbf{1}$, zaś $I_\omega^S(\alpha) = \mathbf{0}$ w przeciwnym razie

v) jeżeli $\alpha = \forall x \beta$, to $I_\omega^S(\alpha) = \mathbf{1}$, jeśli dla $\forall \omega'$ $\omega' \in \omega[x]$ jest $I_\omega^S(\beta) = \mathbf{1}$, zaś $I_\omega^S(\alpha) = \mathbf{0}$ w przeciwnym razie, gdzie $\omega[x]$ to zbiór wartościowań dla wszystkich zmiennych, oprócz co najwyżej zmiennej x.

Spełnialność formuł

- α jest *spełniona (prawdziwa) w S dla* $\omega \in \Omega_S \Leftrightarrow I_\omega^S(\alpha) = \mathbf{1}$
- α jest spełnialna, jeżeli $\exists S \exists \omega$ dla $\omega \in \Omega_S$, dla których $I_\omega^S(\alpha) = \mathbf{1}$
- *Spełnialność* formuł zdaniowych zależy jedynie od struktury
- α jest spełniona $\Rightarrow S$ jest <u>modelem (semantycznym)</u> formuły α
- Modelem zbioru formuł jest struktura, która jest modelem dla każdej formuły z tego zbioru.

Semantyczna konsekwencja

α jest *semantyczną konsekwencją* zbioru formuł $\Phi \Leftrightarrow$ dowolny model zbioru Φ jest modelem formuły α, co zapisujemy

$$\Phi \models \alpha$$

Dla $\Phi = 0$ α jest zawsze prawdziwa i zwana *tautologią*

$$\models \alpha$$

α jest *równoważna* $\beta \Leftrightarrow \alpha \models \beta$ i $\beta \models \alpha$
α jest *modelowo równoważna* $\beta \Leftrightarrow \alpha \models \beta$

Przykładowa formuła świata klocków

$$\alpha = (\neg P \vee Q) \quad \Leftrightarrow \quad (\neg Q \Rightarrow \neg P)$$
$$\alpha = (P \Rightarrow Q) \quad \Leftrightarrow \quad (\neg Q \Rightarrow \neg P)$$

$$\begin{aligned}
{}^{1)}\alpha(P=1,Q=1) \quad &= \quad 1 \\
(\neg 1 \vee 1) \quad &\Leftrightarrow \quad (\neg 1 \Rightarrow \neg 1) \\
(0 \vee 1) \quad &\Leftrightarrow \quad (0 \Rightarrow 0) \\
1 \quad &\Leftrightarrow \quad 1 \\
{}^{2)}\alpha(P=1,Q=0) \quad &= \quad 1 \\
(\neg 1 \vee 0) \quad &\Leftrightarrow \quad (\neg 0 \Rightarrow \neg 1) \\
(0 \vee 0) \quad &\Leftrightarrow \quad (1 \Rightarrow 0) \\
0 \quad &\Leftrightarrow \quad 0
\end{aligned}$$

$$\begin{aligned}
{}^{3)}\alpha(P=0,Q=1) \quad &= \quad 1 \\
(\neg 0 \vee 1) \quad &\Leftrightarrow \quad (\neg 1 \Rightarrow \neg 0) \\
(1 \vee 1) \quad &\Leftrightarrow \quad (0 \Rightarrow 1) \\
1 \quad &\Leftrightarrow \quad 1 \\
{}^{4)}\alpha(P=0,Q=0) \quad &= \quad 1 \\
(\neg 0 \vee 0) \quad &\Leftrightarrow \quad (\neg 0 \Rightarrow \neg 0) \\
(1 \vee 0) \quad &\Leftrightarrow \quad (1 \Rightarrow 1) \\
1 \quad &\Leftrightarrow \quad 1
\end{aligned}$$

Wnioskowanie

Wnioskowanie to inaczej *inferencja*.

Regułą wnioskowania nazywamy dowolną operację, która skończonemu ciągowi formuł $\alpha_1, \ldots, \alpha_n$ nazywanych *przesłankami* (ang. premises), przyporządkowuje formułę β, nazywaną *wnioskiem* (ang. conclusion), co zapisuje się jako:

$$\frac{\alpha_1, \ldots, \alpha_n}{\beta}$$

Reguły wnioskowania

Reguła *modus ponens* zwana także regułą odrywania ma postać

$$\frac{\alpha, \alpha \Rightarrow \beta}{\beta}$$

Reguła *uogólniania* ma postać

$$\frac{\alpha}{\forall x \;\; \alpha}$$

Proces inferencji

Formuła α jest *wyprowadzalna* ze zbioru formuł Φ za pomocą reguł ze zbioru $\mathcal{R} \Leftrightarrow \exists$ ciąg formuł $\beta_1, \ldots, \beta_k$ taki, że:

i) $\alpha = \beta_k$;

ii) $\forall (i \leqslant k)$ $\beta_i \in \Phi$ lub β_i jest wnioskiem reguły należącej do $\mathcal{R}$ z pewnych formuł z $\{\beta_1, \ldots, \beta_{k-1}\}$.

Ciąg formuł $\beta_1, \ldots, \beta_k$ nazywamy *dowodem* formuły β_k z Φ z zastosowaniem reguł wnioskowania z $\mathcal{R}$.

Teoria

Teoria jest sformalizownym opisem świata rzeczywistego i składa się z *języka* czyli zbioru formuł oraz struktury dedukcyjnej: zbioru *aksjomatów logicznych*, zbioru *aksjomatów specyficznych* i zbioru *reguł wnioskowania*.

Aksjomaty logiczne

Aksjomaty logiczne muszą być *tautologiami*.

Niech α, β, γ będą dowolnymi formułami teorii. Typowe aksjomaty logiczne teorii pierwszego rzędu to:

1. $\alpha \Rightarrow (\beta \Rightarrow \alpha)$
2. $(\alpha \Rightarrow (\beta \Rightarrow \gamma)) \Rightarrow ((\alpha \Rightarrow \beta) \Rightarrow (\alpha \Rightarrow \gamma))$
3. $(\neg\beta \Rightarrow \neg\alpha) \Rightarrow ((\neg\beta \Rightarrow \alpha) \Rightarrow \beta)$
4. $\forall x\ \alpha(x) \Rightarrow \alpha(t)$,
   gdzie t jest termem, $\alpha(x)$ - formułą, zaś $\alpha(t)$ - formułą $\alpha(x)$ po zastąpieniu każdego wolnego wystąpienia zmiennej x termem t, ponadto $\exists$ zmienna z termu $t \Rightarrow \forall$ wolne wystąpienie zmiennej x nie leży w zasięgu działania kwantyfikatorów $\forall z$ lub $\exists z$.
5. $\forall x(\alpha \Rightarrow \beta) \Rightarrow (\alpha \Rightarrow \forall x\ \beta)$,
   jeśli α nie ma wolnych wystąpień zmiennej x

Reguły wnioskowania, a teoria

Jako reguły wnioskowania przyjmuje się regułę *modus ponens* i *uogólniania*.

Możliwy jest dobór innych aksjomatów logicznych i reguł wnioskowania np. system Gentzena.

Skoro aksjomaty logiczne oraz reguły wnioskowania są ustalone, to teorię określa się lub też w praktyce utożsamia ze zbiorem aksjomatów specyficznych.

Aksjomaty specyficzne

Aksjomaty specyficzne są formułami, które arbitralnie zostały uznane przez twórców teorii za *prawdziwe*, a opisujące cechy świata rzeczywistego np. formuła

$$(\forall x)(człowiek(x) \Rightarrow śmiertelny(x))$$

opisuje fakt, że każdy człowiek jest śmiertelny.

Twierdzenie

α jest *wyprowadzalna w teorii* T lub jest *jej twierdzeniem* $\Leftrightarrow$ α jest wyprowadzalna za pomocą reguł wnioskowania tej teorii z formuł pochodzących ze zbiorów jej aksjomatów logicznych i specyficznych, co zapisujemy tak

$$T \vdash \alpha$$

W przypadku rachunku predykatów (brak aksjomatów specyficznych) zapis jest następujący:

$$\vdash \alpha$$

Niesprzeczność teorii

Teoria jest *niesprzeczna* $\Leftrightarrow \forall\alpha\ \neg\alpha$ i α nie są jednocześnie twierdzeniami tej teorii. Można wykazać, że

- Teoria jest niesprzeczna $\Leftrightarrow \exists$ model tej teorii
- Dla dowolnej niesprzecznej teorii istnieje *przeliczalny* model.

α jest twierdzeniem niesprzecznej teorii $\Leftrightarrow$ α jest prawdziwa w dowolnym modelu tej teorii, co formalnie można zapisać

$$T \vdash \alpha \Leftrightarrow T \models \alpha$$

Zupełność teorii

Teoria jest *zupełna* $\Leftrightarrow$ $\forall$ zamkniętej formuły α tej teorii, albo $T \vdash \alpha$ albo $T \vdash \neg\alpha$. Taka teoria opisuje wszystkie informacje związane z reprezentowanym przez nią światem.

Rozstrzygalność teorii

Teorię nazywa się *rozstrzygalną*, jeżeli można w skończonej liczbie kroków stwierdzić, czy dowolna formuła należąca do języka tej teorii jest, czy też nie jest jej twierdzeniem.

Teorię nazywa się *półrozstrzygalną*, jeżeli można w skończonej liczbie kroków udowodnić każde twierdzenie tej teorii. Nie ma jednak gwarancji, na efektywne określenie, czy dana formuła nie jest twierdzeniem w T. Teorie I rzędu w ogólnym przypadku nie są rozstrzygalne, lecz są półrozstrzygalne. Niemniej jednak istnieją pewne rozstrzygalne klasy formuł, np: formuły z predykatami jednoargumentowymi lub poprzedzone tylko kwantyfikatorami ogólnymi lub poprzedzone tylko kwantyfikatorami egzystencjonalnymi.

Monotoniczność

Zbiór twierdzeń teorii I rzędu zwiększa się wraz ze wzrostem aksjomatów specyficznych. Własność ta nazywa się *monotonicznością*.

Standaryzacja formuł

Standaryzacja polega na przekształceniu formuł wyjściowych w formuły, które cechują się tym, że

1. wszystkie kwantyfikatory wyprowadzane są na początek formuły - *postać preneksowa normalna*;
2. kwantyfikatory egzystencjalne zostają wyeliminowane - *postać normalna Skolema* $F_S \models F$;
3. wyrażenie pod kwantyfikatorami jest koniunkcją alternatyw.

Z koniunkcji alternatyw przechodzi się do ich zbioru. Jeśli alternatywa jest złożona tylko z formuł atomowych pozytywnych i negatywnych, to nazywa się *klauzulą*.

$\alpha \Leftrightarrow \beta \equiv (\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha)$

$\alpha \Rightarrow \beta \equiv \neg\alpha \vee \beta$

$\neg(\neg\alpha) \equiv \alpha$

$\neg(\alpha \vee \beta) \equiv \neg\alpha \wedge \neg\beta$

$\neg(\alpha \wedge \beta) \equiv \neg\alpha \vee \neg\beta$

$\neg\forall x\alpha \equiv \exists x\neg\alpha$

$\neg\exists x\alpha \equiv \forall x\neg\alpha$

$\alpha \wedge (\alpha \vee \gamma) \equiv (\alpha \wedge \beta) \vee (\alpha \wedge \gamma)$

$\alpha \vee (\alpha \wedge \gamma) \equiv (\alpha \vee \beta) \wedge (\alpha \vee \gamma)$

CNF: $\neg p \vee \cdots \vee p$

DNF: $\neg p \wedge \cdots \wedge p$ - fałsz!

$Q \in \{\exists, \forall\}$

$Qx\alpha \vee \beta \equiv Qx(\alpha \vee \beta)$

$Qx\alpha \wedge \beta \equiv Qx(\alpha \wedge \beta)$,

gdzie β bez wolnych zmiennych

$\forall x\alpha \wedge \forall x\beta \equiv \forall x(\alpha \wedge \beta)$

$\exists x\alpha \vee \exists x\beta \equiv \exists x(\alpha \vee \beta)$

$Qx\alpha \equiv Qx\alpha[x/y]$,

gdzie y bez wolnych zmiennych,

a x z wolnymi zmiennymi

PNF: $Q_1x_1Q_2x_2\ldots Q_nx_n\mu$,

gdzie μ jest koniunkcją alternatyw

Klauzula Horna

Klauzulę postaci

$$\neg\beta_1 \vee \ldots \vee \neg\beta_m \vee \gamma_1 \vee \ldots \vee \gamma_n$$

nazywa się *klauzulą Horna* $\Leftrightarrow n = 0$ lub $n = 1$ dla $m \geqslant 0$.
Inny zapis to

$$\beta_1 \wedge \ldots \wedge \beta_m \Rightarrow \gamma$$

Schemat rezolucji

<u>Schematem rezolucji</u> (Robinson - 1965) nazywa się *regułę inferencyjną*

$$\frac{A \vee B, C \vee \neg B}{A \vee C},$$

gdzie A, B, C są formułami, $A \vee C$ jest *rezolwentą binarną* klauzul wejściowych, a *klauzula pusta (NIL)* nie jest spełniona w żadnej strukturze.

Do danego zbioru klauzul Φ dołącza się zbiór klauzul modelowo równoważnych negacji formuły $\neg\alpha$, którą zamierzamy udowodnić. Potem stosuje się wielokrotnie schemat rezolucji. Uzyskanie rezolwenty równej NIL oznacza, że zbiór $\Phi \cup \{\neg\alpha\}$ jest sprzeczny i α nie jest twierdzeniem.

$K_{1,4}$	$K_{2,4}$	$K_{3,4}$	$K_{4,4}$
$K_{1,3}$ Kuchnia $K_{1,3} \vee \neg K_{1,3}$	$K_{2,3}$	$K_{3,3}$	$K_{4,3}$
$K_{1,2}$ Robot $c_{1,2}$	$K_{2,2}$ $\neg c_{2,2}$	$K_{3,2}$	$K_{4,2}$
$K_{1,1}$ $\neg c_{1,1}$	$K_{2,1}$ $\neg c_{2,1}$	$K_{3,1}$	$K_{4,1}$

R_1: *modus ponens lub rezolucja*

$\neg K_{1,1} \wedge \neg K_{1,2} \wedge \neg K_{2,1}$

eliminacja $\wedge$: $\neg K_{1,1}, \neg K_{1,2}, \neg K_{2,1}$

$R_2 : \neg K_{1,1}, \neg K_{2,1}, \neg K_{2,2}, \neg K_{3,1}$

$R_3 : \neg K_{2,1}, \neg K_{1,2}, \neg K_{2,2}, \neg K_{3,2}, \neg K_{2,3}$

$R_4 : K_{1,3} \vee K_{1,2} \vee K_{2,2} \vee K_{1,1}$

rezolucja $\Longrightarrow$

c - zapach kawy w pokoju (i, j)

k - kuchnia w pokoju (i, j)

Wiedza: $\neg c_{1,1}$, $\neg c_{2,1}$, $\neg c_{2,2}$, $c_{1,2}$

Korzystając ze zmysłu zapachu robot znajduje kuchnie!

$R_1 \quad \neg c_{1,1} \Rightarrow (\neg K_{1,1} \wedge \neg K_{1,2} \wedge \neg K_{2,1})$

$R_2 \quad \neg c_{2,1} \Rightarrow (\neg K_{1,1} \wedge \neg K_{2,1} \wedge \neg K_{2,2} \wedge \neg K_{3,1})$

$R_3 \quad \neg c_{2,2} \Rightarrow (\neg K_{2,1} \wedge \neg K_{1,2} \wedge \neg K_{2,2} \wedge \neg K_{3,2} \wedge \neg K_{2,3})$

$R_4 \quad c_{1,2} \Rightarrow (K_{1,2} \vee K_{2,2} \vee K_{1,1} \vee K_{1,3})$

$R_4 \vee \neg K_{1,1} : K_{1,3} \vee K_{1,2} \vee K_{2,2}$

$R_4 \vee \neg K_{1,1} \vee \neg K_{2,2} : K_{1,3} \vee K_{1,2}$

$R_4 \vee \neg K_{1,1} \vee \neg K_{2,2} \vee \neg K_{1,2} : K_{1,3}$