

# Sztuczna Inteligencja

## Big CyC

Włodzisław Duch

Katedra Informatyki Stosowanej UMK

Google: Włodzisław Duch

# CYC



- Zdrowy rozsądek
- CYC - podstawy
- CYC - ontologia podstawowa
- CYC – możliwości i zastosowania

# Zdrowy rozsądek



Najbardziej ambitny regułowy system ekspertowy.

Realizowany pod kierunkiem D. Lenata w MCI od 1984 roku,  
od 1995 w komercyjnej firmie Cycorp, [www.cyc.com](http://www.cyc.com)

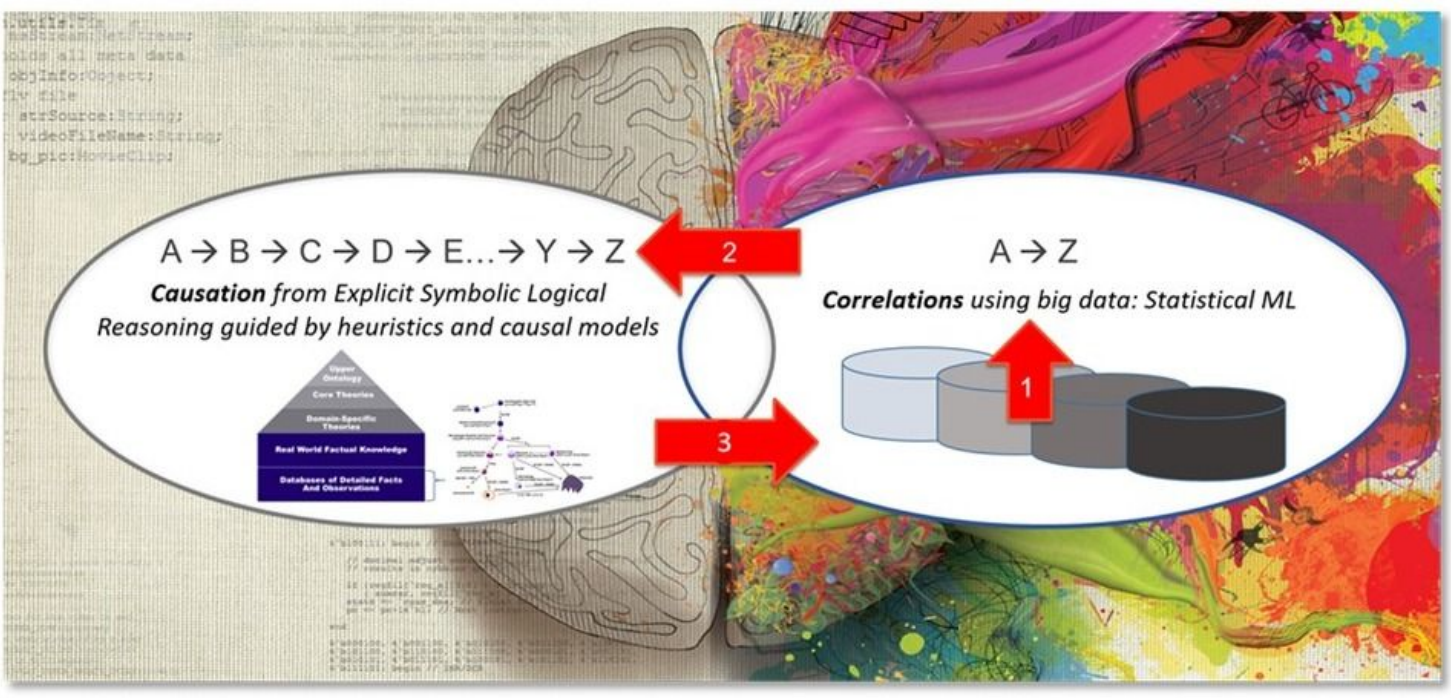
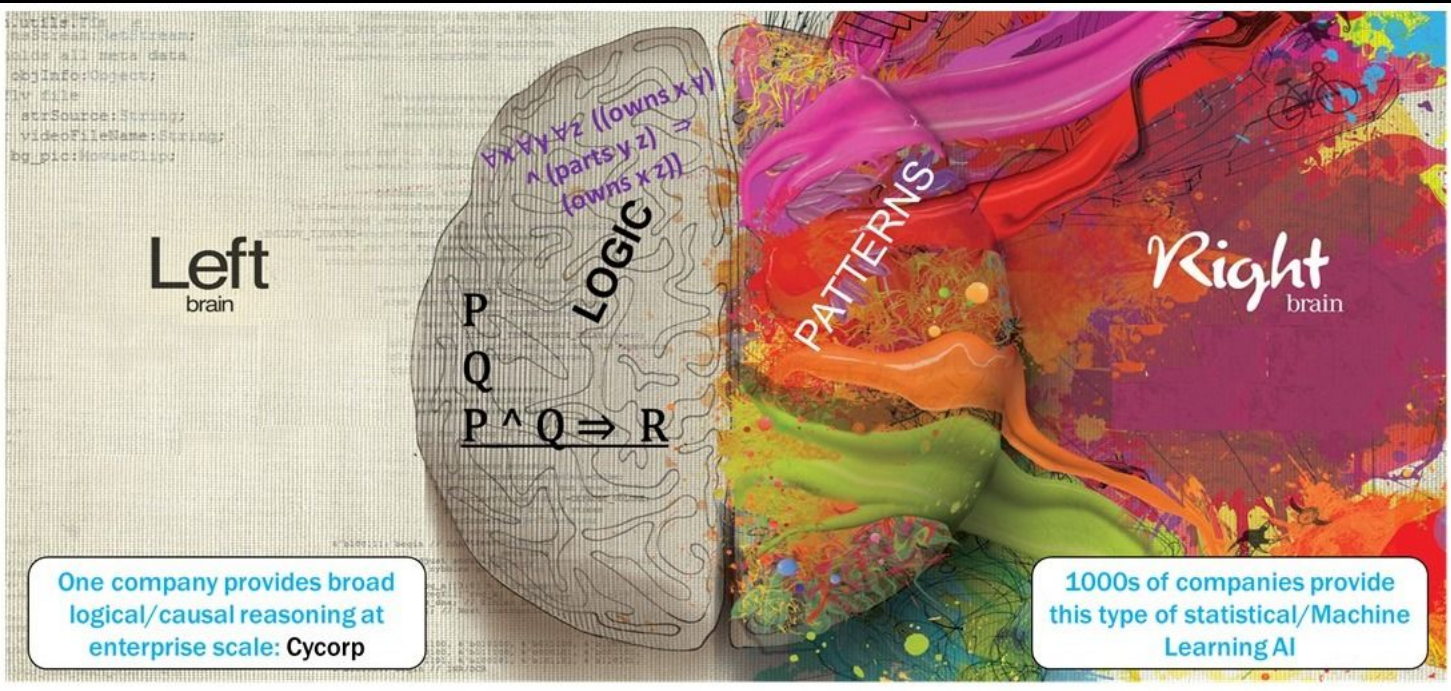
Zdrowy rozsądek wymaga milionów reguł! Czy to realne?

Test ostateczny dla technologii ES - czy tak złożony system będzie działał?

Początkowe oceny - 5% szans na użyteczność.

„Kruchość” systemów ekspertowych:

- niewiele systemów ma ponad 1000 reguł;
- którą regułę/ramę użyć?
- formy pytań muszą być ściśle określone.



# Zdrowy rozsądek cd



Ludzie nie dają się łatwo zbić z tropu.

„Szeroka” wiedza  $\Leftrightarrow$  odpowiedzi szybkie = rozumowania płytkie.  
Potrzebujemy wielu sposobów kojarzenia faktów (reasoners).

Baza wiedzy - fakty, heurystyki, wiedza specyficzna.

System-partner człowieka?

Musi znać nastawienia ludzi, światopogląd, mikroteorie zależne od kontekstu.

Czy znane sposoby reprezentacji wiedzy wystarczą?

Reprezentacja podstawowych pojęć - ontologia podstawowa.

Czas, przestrzeń, ruch, substancja ... wspólna wiedza ES.

Knowledge-based systems - podstawa porozumiewania się.

# Zdrowy rozsądek cd.



System CYC zawiera w wersji podstawowej ponad milion reguł.

Baza wiedzy powstała w oparciu o analizy ekspertów.

Bootstrapping - analiza tekstów w oparciu o zgromadzoną wiedzę pozwoli na automatyzację procesu.

Język reprezentacji wiedzy deklaratywnej + procedury manipulowania faktami.

Język opisuje obiekty, zdarzenia, nastawienia.

Język reprezentacji CycL rozwijał się wraz z KB.

Oparty na ramach + rachunek predykatów + uzupełnianie zmiennych domyślnych.

Nowe pojęcia mogą wymagać rozszerzenia CycL - „ekspresywność” języka nie jest dostateczna.

Obecnie CycL jest stabilny.

# CyCL 1

Ramy - dziedziczenie dla specjalnych połączeń, *isa*, *instance*

„Wszystkie ptaki mają dwie nogi”.

Szufladka *nogi* w ramach typu *ptak*, ma wartość „2”.

„Wszyscy przyjaciele królika lubią marchewkę”.

Ramki z *królik* w szufladce *przyjaciel* dziedziczą *marchewka* jako zawartość szufladki *lubi*.

Dziedziczenie przez łańcuch relacji:

„wszyscy krewni przyjaciół królika mają długie uszy”.

Szufladki *uszy* krewnych przyjaciół królika  *długie*.

# Cycl 2

*Constraint language* - opis przez specyfikację ograniczeń.

„Norbert nie lubi ludzi, którzy programują w Fortranie”.

Nie zapełnienia szufladek stwierdzeniem

„Norbert ich lubi/nie lubi”

Zmienna ograniczająca *fortran-constraint*, w szufladce „lubi” ramy Norbert.

**Język specyfikacji ograniczeń** pozwala na realizację pełnej logiki pierwszego rzędu, uzupełnienie ram.

Większa sprawność rozumowania opartego o ramy niż rozumowania wymagającego logicznych wniosków.



# CyCL - mikroteorie

Mechanizmy wnioskowania: uniwersalno-proceduralne.

Identyfikacja często używanych klas wnioskowań

=> dodaje się sprawne mechanizmy.

Specyficzne mechanizmy wnioskowania, zależnie od dziedziny.

Wnioskowanie i tworzenie „mikroteorii”, np. opis posługiwania się pieniędzmi, robienie zakupów, jazda autem.

1. Pieniądze daje się w zamian za towary, rzeczy, usługi lub jako dar.

2. Każdy czynnik jest związany z pewną sumą pieniędzy.

3. Opłaty do 10\$ dokonywane są zwykle gotówką, powyżej 50\$ czekiem lub kartą kredytową.

.....

17. Kiedy coś kupujesz jednym z podzdarzeń jest płacenie sprzedawcy lub w kasie.

# Spójność wiedzy

CYC zawiera > 1100 różnych mechanizmów wnioskowania. (reasoners, inference engines), łączących 1.5 mln pojęć.

Dla każdego z nich usiłuje zachować spójność wiedzy.

- Czy w szufladkach są dopuszczalne wartości ?
- Jak rozstrzygać sprzeczności?

Testy rozumienia: analiza opowiadań i pytania.

Czas stosowania domyślnego rozumowania zależy od reguły.

Reguły *if-needed rule* używane tylko gdy się do nich odwołujemy; reguły *if-added rule* oceniające wszystko, co się da.

*Forward rules*, reguły wyprzedzające - dużo czasu, mało przydatne.

Np. własności znajomego Jana „programuje w Fortranie” to nazwisko znajomego może się pojawić na liście osób, których Jan nie lubi.

# Poziomy reprezentacji

Podsystem utrzymywania spójności CYC (truth maintenance system). Dodaje nowe fakty w nocy lub wolnego czasu.

Poziom epistemologiczny (EL) - abstrakcyjny poziom reprezentacji, wykorzystuje język specyfikacji ograniczeń, wymaga abstrakcyjnego rozumowania logicznego.

Poziom reprezentacji heurystycznej (HL) - deklarowane fakty, efektywne heurystyczne sposoby rozumowania.

Zbiór mechanizmów wnioskowania skojarzony jest z każdą szufladką. Pytanie proste - sprawdź zawartości szufladki.

Pytanie złożone - wnioski w celu zapełnienia szufladki.

Strategia kontrolna wnioskowania - również w oparciu o ramy.

Schematy wnioskowania w ramach, kopie zmodyfikowane stosuje się do nowych przypadków

Kompilator języka CycL tworzy kod w Lispie.

# Asercje

Wiedza deklaratywna, wnioskowanie częściowo proceduralne.  
CycL wykorzystuje logiczne stwierdzenia (*assertions*).

CYC ma >10.000 predykatów, ok. 25 mln asercji, prostych faktów, reguł i sposobów wnioskowania; stwierdzenie może wiązać się z 10-100 innymi, stąd liczba powiązań w tym systemie sięga milionów.

Pierwsze > 1.5M reguł/powiązania odnosi się do globalnej ontologii.  
Jak reprezentować podstawowe pojęcia w bazie wiedzy?

Opis + relacje najczęstszych sposobów użycia danego pojęcia.  
Zamiast zredukowanego opisu próbuje się opisać wszystkie sytuacje.

CYC to pierwszy program wykorzystujący globalną ontologię, czyli klasyfikację bytów.

Kwestie ontologii dyskutowane były od zarania filozofii ...

# Ontologia podstawowa

Wszystko jest rodzajem rzeczy, konkretnej lub abstrakcyjnej.  
Rzeczy indywidualne i kolekcje, zbiory rzeczy.

Indywidualne, np. *Jan, Polska, nos Jana*, mogą mieć części.

Kolekcje, np. *osoba, naród, nos*, mogą mieć podzbiory.

Rzeczy *nienamacalne* nie mają masy: *zdarzenia, liczby, prawa*.

Rzeczy *namacalne* mają masę: *ciało człowieka, jabłko czy kurz*.

Rzeczy złożone - cechy namacalne i nienamacalne,

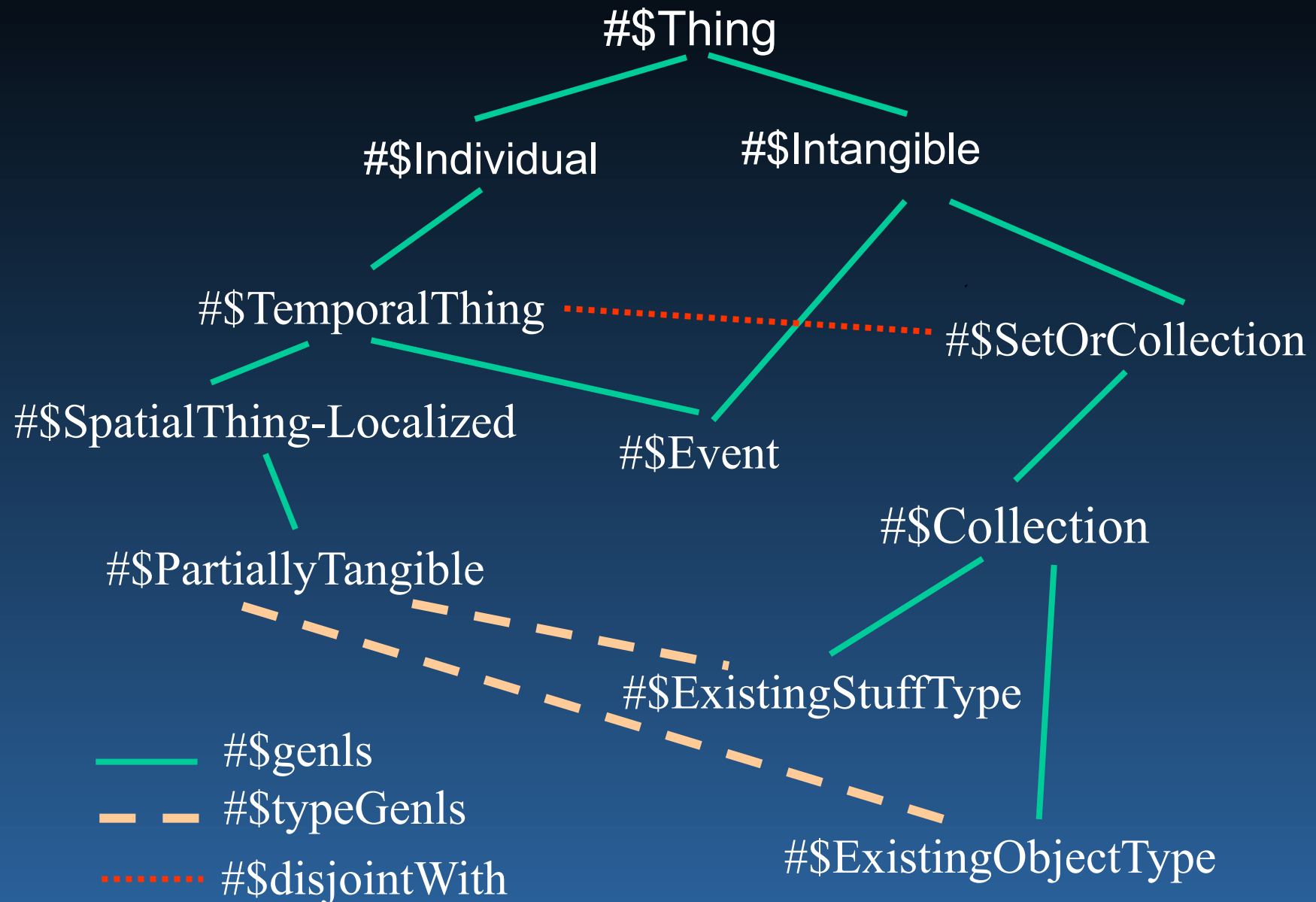
np. *osoba* ma *ciało* i *umysł*.

Substancja to *ObiektIndywidualny*, pocięta zachowuje własności;

Własności zewnętrzne (*extrinsic*)

wewnętrzne (*intrinsic*), zarówno części jak i całego obiektu.

# Kolekcje na najwyższych poziomach



# Ontologia cd.

Zdarzenia: rzeczy dziejące się w czasie.

Procesy: zdarzenia, które po podzieleniu zachowują własności.

*Spacer* jest procesem, ale *kilometrowy-spacer* nie jest.

Zdarzenia: wł. temporalne, interwały czasowe i zbiory interwałów.

*Slot*, czyli *szufladka*, to podklasa nienamacalnego.

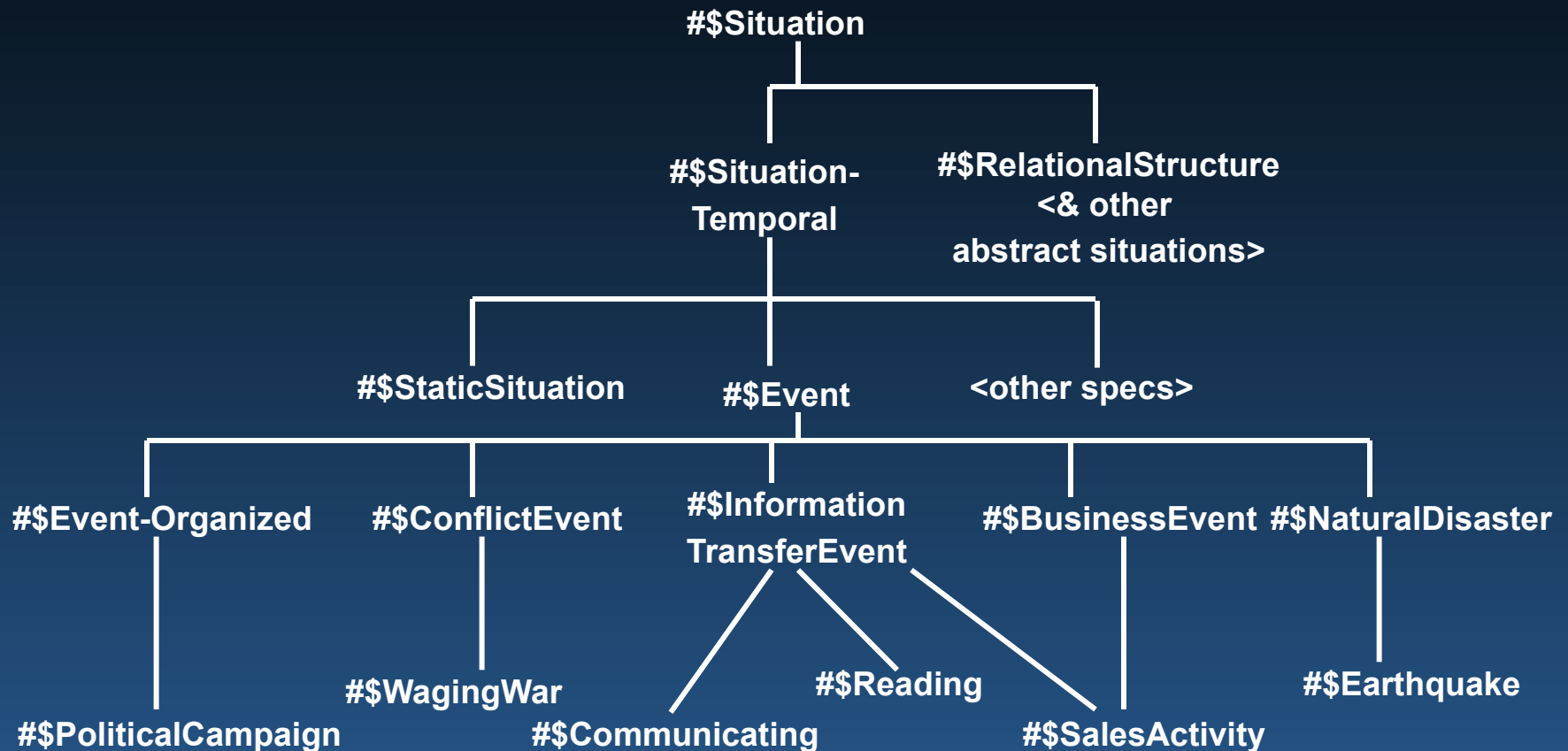
Szufladki definiujące, buchalteryjne (zapisujące informacje o pochodzeniu), ilościowe itp.

**Agent**: obiekt złożony, zbiór inteligentnych istot, ludzie, firmy, programy.

Agent ma przekonania (*beliefs*), może też przypisywać przekonania innym.

Przekonania nie zawsze są prawdziwe, trzeba odróżnić lokalne przekonania agentów od własnej wiedzy o świecie.

# Częściowa hierarchia zdarzeń





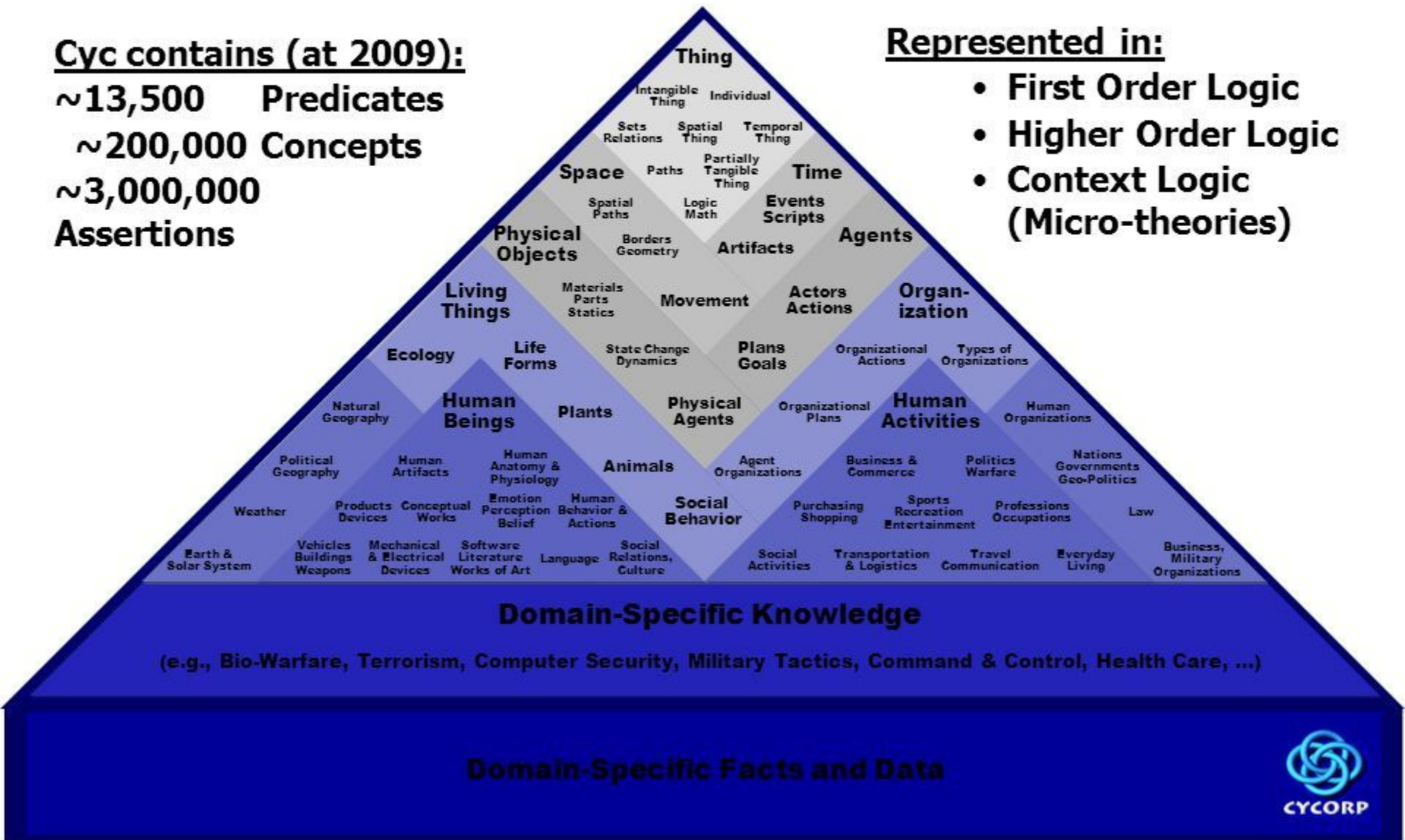
# Cyc Ontology & Knowledge Base

**Cyc contains (at 2009):**

~13,500 Predicates  
 ~200,000 Concepts  
 ~3,000,000 Assertions

**Represented in:**

- First Order Logic
- Higher Order Logic
- Context Logic (Micro-theories)



# Ontologia cd.

„Upper Cyc ontology” to 6000 najbardziej ogólnych pojęć podstawowych „human concensus reality”, dostępne w Internecie.

Szczegółowe struktury dodawane są do bazy wiedzy Cyc od 1984 roku. Dodatkowa wiedza z dziedziny, do której stosuje się system Cyc dodawana jest przez użytkownika.

Podstawowa ontologia jest: uniwersalna, niezbędna do skonstruowania innych pojęć, zawiera rozróżnienia konieczne i wystarczające dla większości zastosowań.

„Konieczne” - warto je zrobić z teoretycznego i pragmatycznego punktu widzenia.

„Wystarczające” - wystarczająco dużo rozróżnień by umożliwić zastosowania: dzielenie się wiedzą, rozstrzygnięcie niejednoznaczności języka naturalnego, szukanie wiedzy w bazach danych.

Budowa ontologii to duża gałąź AI, strona pokazująca ontologię CyC: The [UpperOntology of the Cyc Project](#) używa [hiperbolicznego drzewa](#).

# Terminologia Cyc

Term - stała, np., # $\$$ Skin, nazwa pojęcia w KB.

Stała = zbiór (np. zbiór wszystkich ludzi), obiekt indywidualny (konkretną osobę), słowo, kwantyfikatory (np. 'istnieje'), relacje (predykat, funkcję, szufladkę, atrybut), itp.

Podstawowa ontologia - szczyt wielkiej góry pojęć KB.

Opis stałej: nazwa pojęcia + komentarz pomagający zrozumieć znaczenie i sposób użycia tej stałej.

Informacja dotycząca hierarchii, *isa*, zbiory i nadzbiory.

Hiperpołączenia do innych pojęć, np. # $\$$ TactileSensor.

Pojęcie „skóra” reprezentuje jeden z obiektów, które zdefiniowane są jako zbiór kawałków, czyli jako kolekcję.

# Terminologia cd.

Fizjologia: definicja stałej #Skin, czyli „Skóra”.

(Kawałek) skóry służy jako zewnętrzna ochrona i czujniki dotykowe pokrywające (część) ciała zwierzęcia. Skóra jest zbiorem wszystkich kawałków skóry. Niektóre przykłady to:

#TheGoldenFleece („Złote Runo”, reprezentujące całą skórę zwierzęcia) i (#BodyPartFn #YulBrynner #Scalp) (reprezentujący małą część jego skóry).

isa: #AnimalBodyPartType

genls: #BiologicalLivingObject #AnimalBodyPart #SheetOfSomeStuff  
#VibrationThroughAMediumSensor #TactileSensor

#Skin należy („isa”, czyli jest elementem) do zbioru „części zwierząt”, przykłady kojarzące się ze skórą - skóra na łysinie znanego aktora, oraz nadzbiory (genls, czyli jest podzbiorem) zawierające skórę. X = skóra może być „płatem czegoś” lub „czujnikiem wrażliwym na dotyk”.

# Terminologia cd.

#\$Collection - zbiory, klasy obiektów o wspólnych własnościach. Definicje kolekcji są hierarchiczne, tylko specyficzne generalizacje.

Zmniejsza to bazę faktów, wydłuża rozumowanie.

#\$Relation, zbiór wszystkich relacji obejmujących funkcje i predykaty. Należy określić typ argumentów, np. *isa* odnosi się do kolekcji.

#\$mother : <Animal> <FemaleAnimal>

(#\$mother ANIM FEM) oznacza że #\$FemaleAnimal FEM jest rodzaju żeńskiego parent of the #\$Animal ANIM.

isa: #\$FamilyRelationSlot #\$BinaryPredicate

Komentarz wyjaśnia, że definicja dotyczy tylko biologicznego sensu słowa matka, a nie metaforycznego, dotyczy tylko zwierząt.

Cycl: wiele sposobów ograniczania rodzaju argumentów predykatów.

# Terminologia cd.

Wykorzystanie relacji: matką osoby  $X$  jest osoba  $Y$

Wyprowadzenie ogólnych reguł, np.

Jeśli matką  $X$  jest  $Y$  to  $Y$  urodziła się przed  $X$  i  $Y$  była prawdopodobnie żywa w momencie narodzin  $X$ .

Funkcje w Cyc są podobne do predykatów, zwracają dowolne stałe.

Np. nazwa osobnika rodzaju żeńskiego danego gatunku:

`#$FemaleFn ($Jeleń) = łania.`

Pozwala to uniknąć definicji pojęcia „łania”.

Wartości i argumenty funkcji ograniczone są przez predykaty:

`#$resultIsa`, „rezultat należy do”

`#$resultGenl`, „rezultat jest podzbiorem”.

Wynik działania `#$FemaleFn` musi należeć do zbioru `#$FemaleAnimal`, musi też być podzbiorem zwierząt.

# Przykład

Animals sleep at home.

```
(ForAll ?x (ForAll ?S (ForAll ?PLACE  
(implies (and  
(isa ?x Animal)  
(isa ?S SleepingEvent)  
(performer ?S ?x)  
(location ?S ?PLACE))  
(home ?x ?PLACE))))))
```

Jeśli x jest zwierzęciem i  
jeśli jest wykonawcą „sleeping event”, czyli śpi,  
to miejscem tego spania jest dom x.

Konstrukcje te nie są naturalne, ale w planach jest tłumaczenie z  
angielskiego na CyCo-logiczny.

# CYC - rozwój

CYC jak umysł, ciągle się rozwija.

W sumie > 2000 osobolat pracy, kosztował ok. 60 M\$, już w 2002 roku zawierał 300.000 pojęć + 3 miliony reguł; >25 mln reguł w 2020 roku.

Początkowo napisany był w Lispie na stacjach Symbolics, teraz jest w C pod Unix, MacOS i Windows.

Skalowalność systemu, przetwarzanie rozproszone.

Stworzenie ontologii - wysiłek wielu osobo-wieków.

Nowe ontologie, np. militarne, polityczne, ekonomiczne, naukowe.

Testowanie wnioskowania na dziesiątkach milionów przykładów.

Kategorie i funkcje, które wydawały się przydatne potem odrzucono, np.

`#$CompletelyTangibleObject` czy `#$TransfersThrough`);

niektóre kategorie niezbyt przyjemnie wyglądające, np.

`#$SiblingDisjointCollection` czy `#$SkolemFunction` okazały się bardzo przydatne i przetrwały wszystkie testy użyteczności.



# OpenCyc

Otwarta, publicznie dostępna wersja systemu (od 2002 roku) o nazwie OpenCyc, składająca się w v4.0 (ostaniej) z:

- 239.000+ pojęć i ponad 2 milionów faktów dotyczących tych pojęć, częściowo definiujących, ograniczających i definiujących wzajemne relacje, których jest 26.000+.
  - Modułu wnioskującego Cyc Inference Engine.
  - Przeglądarki bazy wiedzy Cyc Knowledge Base Browser.
  - Narzędzi do akwizycji wiedzy ze specjalistycznych dziedzin przez dialog z ekspertem w języku naturalnym.
  - Dokumentacji systemu, specyfikacji języka CycL, specyfikacji interfejsu użytkownika Cyc API i innych.
  - Kilku programów demonstracyjnych.
  - Prawie cały CyC jest publicznie dostępny, fragmenty pozostały komercyjne.
- 4/2006, powstała CyC Foundation i publiczny serwer, ale od 2017 roku jest to system komercyjny.

# Research Cyc

Cyc Foundation udostępniał wersję systemu CyC do badań o nazwie ResearchCyc, składająca się (w 2009 roku, v1.1) z:

- Ponad 500.000+ pojęć i 5 milionów faktów dotyczących tych pojęć, częściowo definiujących, ograniczających i definiujących wzajemne relacje, których jest 26.000+.
- Modułu wnioskującego Cyc Inference Engine.
- Przeglądarki bazy wiedzy Cyc Knowledge Base Browser.
- Narzędzi do akwizycji wiedzy ze specjalistycznych dziedzin przez dialog z ekspertem w języku naturalnym.
- Dokumentacji systemu, specyfikacji języka CycL, specyfikacji interfejsu użytkownika Cyc API i innych.
- Kilku programów demonstracyjnych.

Dostęp wymaga licencji, która jest darmowa, ale od 2019 roku firma przestała udostępniać tą wersję, obiecując, że przygotują nową.

Fundacja Cyc zapowiada też Cyclopedię pozwalającą na przeszukiwanie Wikipedii używając pojęć ontologii CyC, ale projekt ten nie został zrealizowany.

# CYC - zastosowania.



- Analiza, rozumienie i generacja tekstów w językach naturalnych.
- Tłumaczenie maszynowe, rozstrzyganie niejednoznaczności.
- Semantyczna integracja baz danych, tworzenie tezaurusów tech/med.
- Semantyczne szukanie informacji (WWW!), szukanie wiedzy.
- Automatyczne anotacje, np. znaczników XML.
- Sprawdzanie spójności wiedzy, integracja heterogenicznych baz danych.
- Symulacje wykorzystujące ograniczenia zawarte w ontologii.

# CYC – zastosowania cd.



- Dzielenie się wiedzą przez pracujące niezależnie grupy.
- Sprzedawanie przez Internet usług i dóbr.
- Budowa i wykorzystywanie modeli użytkownika jakiegoś systemu.
- Modelowanie użytkowników programów i urządzeń technicznych – HCI, ergonomia.
- Inteligentne interfejsy programowe, reagujące na intencje użytkownika.
- Filtrowanie informacji, np. poczty elektronicznej.
- Symulacja inteligentnych zachowań postaci w grach komputerowych.
- Inteligentna symulacja wirtualnej rzeczywistości.

# Cyc-NL

NLP wymaga wiedzy!

Jan widział jak samolot leciał nad miastem.

Jan widział góry gdy leciał nad miastem.

Skąd wiemy, że „leciał” odnosi się raz do samolotu, raz do Jana?

Cyc wie, że samoloty latają a góry nie - bez takiej wiedzy rozstrzygnięcie co czy kto leciał nie jest możliwe.

Cyc-NL:

- leksykon,
- parser syntaktyczny,
- interpreter semantyczny.

# Cyc-NL

- Umożliwia analizę zdań złożonych i niejednoznacznych, zdań modalnych, negacji i zdań kwantyfikujących na kilku hierarchicznych poziomach, najczęściej jest to meta-rozumowanie heurystyczne.
- Komunikacja z Cyc za pomocą Cyc-NL pozwoli na używanie Cyc bez znajomości języka CycL.
- Cyc-NL: interfejsy baz danych w języku naturalnym, wykorzystania w tłumaczeniu maszynowym i analizie mowy ciągłej.
- CyCorp twierdzi, że ekspersyjność Cyc-NL jest taka jak języka angielskiego.

Gra FACTory pomagała CyC określać prawdziwość faktów, zbierając odpowiedzi ludzi określała stopień prawdziwości różnych faktów.

GECKA: Game Engine for Commonsense Knowledge Acquisition,  
Games with a purpose,

# MS Probase + ConceptGraph



A screenshot of a Bing search result for the query "American politicians birthday". It displays a table of American politicians with columns for Name, U.S. Vice President, Birthdate, Country, Order of office, and Birthplace.

Name	U.S. Vice President	Birthdate	Country	Order of office	Birthplace
38	Richard Nixon	January 9, 1913	USA	38	Yorba Linda, California
39	Hubert H. Humphrey	April 27, 1911	USA	39	Waukegan, Illinois
40	Walter Mondale	January 1, 1927	USA	40	Weston, Ohio, New York
41	George H. W. Bush	July 12, 1912	USA	41	Milton, Massachusetts
42	George W. Bush	July 7, 1946	USA	42	New Haven, Connecticut
43	Donald Trump	June 14, 1945	USA	43	Queens, New York
44	Benjamin Franklin	January 17, 1706	USA	44	Philadelphia, Pennsylvania
45	John Adams	October 3, 1735	USA	45	Boston, Massachusetts
46	Thomas Jefferson	April 13, 1743	USA	46	Shadwell, Virginia

UNDERSTANDING

悟 & 懂

Terms	Entity	Attribute
China	<input checked="" type="radio"/>	<input type="radio"/>
Russia	<input checked="" type="radio"/>	<input type="radio"/>
India	<input checked="" type="radio"/>	<input type="radio"/>
Unknown type	<input type="radio"/>	<input type="radio"/>

I think you are talking about

UNDERSTANDING

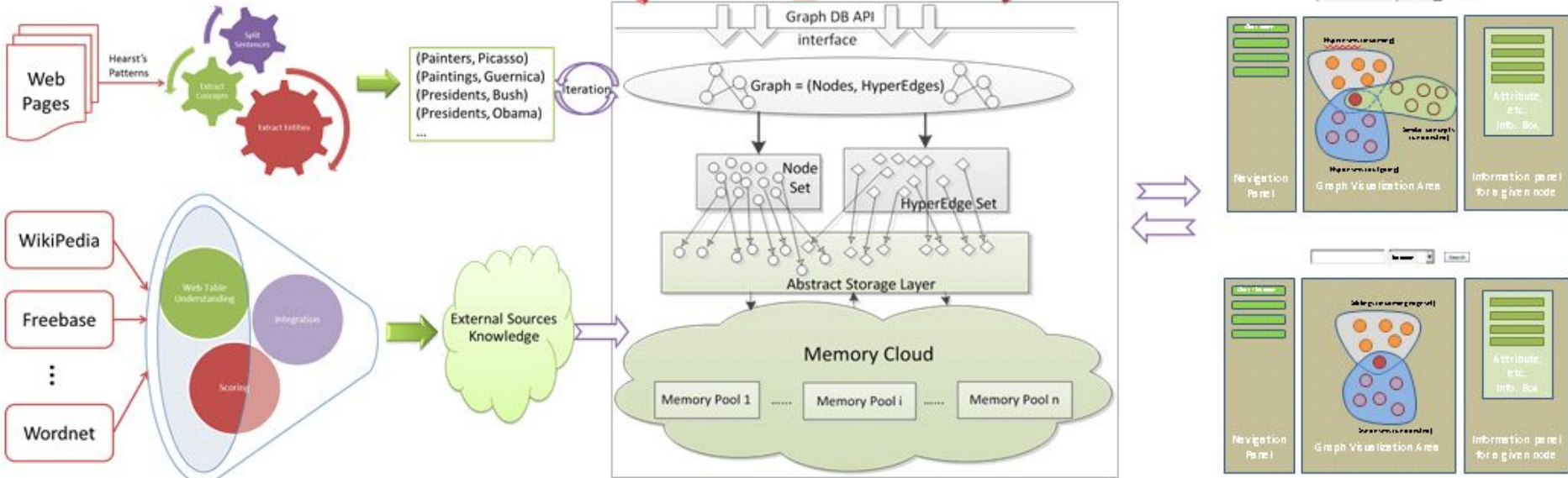
悟 & 懂

Terms	Entity	Attribute
China	<input checked="" type="radio"/>	<input type="radio"/>
Russia	<input checked="" type="radio"/>	<input type="radio"/>
India	<input checked="" type="radio"/>	<input type="radio"/>
Unknown type	<input type="radio"/>	<input type="radio"/>

I think you are talking about

Applications

Infrastructure



Trinity: Graph Database

# Open-Domain Taxonomy

name	# of concepts	# of isA pairs
<b>Freebase</b>	<b>1,450</b>	<b>24,483,434</b>
<b>WordNet</b>	<b>25,229</b>	<b>283,070</b>
WikiTaxonomy	111,654	105,418
YAGO	352,297	8,277,227
DBPedia	259	1,900,000
<b>ResearchCyc ≈</b>	<b>120,000</b>	<b>&lt; 5,000,000</b>
KnowItAll	N/A	< 54,753
TextRunner	N/A	< 11,000,000
OMCS	173,398	1,030,619
NELL	123	< 242,453
<b>Probase</b>	<b>2,653,872</b>	<b>20,757,545</b>

Przeanalizowano około 326 mln zdań z korpusu 1.7 mld stron WWW.  
143 mln par zdań typu isA, pozwoliło na odkrycie ponad 9 mln hiperonimów i ponad 11 mln hiponimów (pojęć nadrzędnych i podrzędnych w taksonomii).  
Po odrzuceniu pojęć, reprezentowanych przez pojedyncze przypadki zostało ok. 2.7 mln pojęć.



# Heterogeniczne bazy danych

W bazach danych jest informacja szczegółowa, kilka reguł może zawierać wiedzę pozwalającą streścić całą informację.

Proces ten może wymagać rozumienia nazw atrybutów.  
Np. atrybut „miasto” odnosi się do pojęcia #City.

Opis struktury bazy danych w terminach CycL pozwala na zadawanie złożonych pytań bazie danych.

Np. info. o ludziach, ich zawodach i miejscu zatrudnienia w jednej bazie + informacje o firmach i ich adresach, to można zadać pytanie:

Wypisz osoby z wyższym wykształceniem mieszkające nad Wisłą.

Cyc wie, że pracownicy uczelni, zakładów badawczych, firm prawniczych, klinik medycznych mają zwykle wyższe wykształcenie, że ludzie zwykle mieszkają w pobliżu miejsca pracy oraz że „nad Wisłą” oznacza w pobliżu rzeki może odpowiedzieć na takie pytanie. Korzystając z Cyc-NL można je zadać bezpośrednio w języku angielskim, bez pisania programu w SQL.

# Bazy danych

Wiedza Cyc pozwala na formułowanie meta-pytań w stosunku do baz danych, używając hierarchii pojęć zawartej w ontologii.

Możliwa jest konstrukcja raportów na podstawie pytań niezależna od struktury czy opisu samej bazy danych, na podstawie pytań.

Wiedza pozwala na powiązanie dostępnych nazw z poszukiwanymi.

Np. szukamy zdjęcia „kogoś w niebezpiecznej sytuacji”.

Cyc wie, że „strzelający terrorysta” lub „żołnierz w akcji” pasuje.

Cyc + Cyc-NL w systemie wyszukiwawczym dla setek tysięcy zdjęć. Opisy zdjęć => formuły w języku CycL.

Cyc-NL tłumaczy z ang. na CycL. Np. „przestraszona osoby”:

(#\$and (\$isa ?x #\$Person)

(\$feelsEmotion ?x #\$Fear #\$High))

# CYC - zastosowania cd.

Cycorp udostępnia Cyc w całości lub części, np. w systemach SAIC, Teknowledge i innych firm.

Nowa koncepcja łączy się z innymi, podaje się stwierdzenia łączące ją z innymi koncepcjami - jeśli jest ich dostatecznie dużo wystarczy płytke wnioskowanie.

Nowe koncepcje muszą być prawidłowo usytuowane w stosunku do podstawowych by dwa systemy można było zintegrować.

Wiedza CYC ma strukturę sformalizowanych powiązań, mikroteorii. Różne poziomy szczegółowości, różnice punktów widzenia, różnice kulturowe, czasu w którym powstała ocena itp.

10 różnych modeli ekonomicznych USA jako mikroteorie, których koncepcje odnoszą się do podstawowej ontologii. Te same terminy wewnątrz mikroteorii mają różne definicje, podlegają innym regułom, ale zbliżone odniesienia do podstawowych koncepcji ontologicznych.

# CYC - friends.

2 systemy CYC tworzą „The Cyclic Friends Network”, protokół wymiany informacji w sieci (KQML).

Wspólna wiedza podstawowa, The Core Cyclic KB, 0.5 Ma.

Wiedza specjalistyczna, np. geografia i polityka.

Rozwiązują problemy przekraczające możliwości każdego z nich.

Przykładowy problem:

„Które państwa Bliskiego Wschodu mogą zaatakować inne kraje ?”

Odpowiedź: „Irak może zaatakować Kuwejt.”

Fakty: Irak ma autorytarny rząd. Irak ma silną armię. Kuwejt ma słabą armię. Irak graniczy z Kuwejtem. Irak leży na Bliskim Wschodzie.

Jeśli państwo ma rząd autorytarny i silną armię oraz graniczy z państwem, które jest militarne słabe, to jest duża szansa, że je zaatakuje.

# CYC - friends 1

„Wymień jakichś demokratycznie wybranych przywódców krajów leżących na północ od równika”.

Odpowiedź: „John Major” , ... bo mamy fakty:

Jeśli kraj ma demokratyczny rząd to przywódca został demokratycznie wybrany.  
Wielka Brytania ma demokratyczny rząd.

John Major jest premierem rządu W. Brytanii.

Jeśli rejon A jest częścią rejonu B a rejon B jest na północ od C to rejon A jest na północ od C.

Europa leży na półkuli północnej.

Półkula północna leży na północ od równika.

Wielka Brytania leży w Europie.

Odpowiedzi wymagają wiadomości geograficznych + politycznych.

Rozumowanie ogólne (G.B. na północ od równika) i proceduralne.

# CYC - friends 2

Agent Cyc mający trudności rozgląda się za pomocą - innym agentem.

Jeśli nie potrafi go znaleźć to zwraca się do Cyc-brokera informacji.

Proces konsultacji może wymagać kooperacji wielu agentów.

Prototyp rozproszonej architektury i 3 agentów:

GeoAgent, PolAgent i EcoAgent - znają się na geografii, polityce i ekonomii.

Automatyczna rozbudowa bazy wiedzy Cyc w określonych domenach.

**Baza informacji o filmach** (Cardiff Database) zawiera informacje o aktorach, reżyserach, producentach, scenariuszach.

Struktura nie jest regularna, dokumenty w html.

Agent-ekspert filmowy odwołuje się do tej bazy a inne systemy Cyc, potrzebujące takiej wiedzy, odwołują się do tego agenta.

„W jakich filmach grał Roland Reagan?”

CycL # $\$$ actedInMovie # $\$$ RonaldReagan ?x

Dane giełdowe, profile firm, informacje z CIA World Fact Book.

Niestety nie widać nowszych eksperymentów w tym kierunku.

# HPKB



HPKB, High Performance Knowledge Bases, inicjatywa DARPA.

Projekt rozpoczęty w 1997 roku, ma znaczenie historyczne.  
Rapid Knowledge Formation (DARPA 1999-2013), Digital Aristotle (project Halo), collaborative Assistant for Rapid Knowledge Formation.

Bazy wiedzy przydatne do problemów strategii militarnej i politycznych sytuacji kryzysowych.

Problemy stanowiące wyzwanie dla systemów wspomaganie decyzji.

Wspomaganie podejmowania decyzji, analiz sytuacji międzynarodowej, przewidywanie kryzysów, ostrzeganie przed kryzysami, wspomaganie kształcenia politologów i wspomaganie modelowania historii kryzysów.

- Scenariusz kryzysu w zatoce Perskiej.
- Scenariusze kryzysów na Bliskim Wschodzie.
- Zarządzanie jednostkami wojskowymi na polu bitwy.

# HPKB - Q53

(Q53) What risks would Iran face in closing the Strait?

Odpowiedzi:

- Economic sanctions from {Saudi Arabia, GCC, US, UN}
- The closure of the Strait of Hormuz would violate an international norm promoting freedom of the seas and would jeopardize the interests of many states.
- In response, states might act unilaterally or jointly to impose economic sanctions on Iran to compel it to re-open the Strait.
- The UN Security Council might authorize economic sanctions against Iran.
- Limited military response from {Saudi Arabia, GCC, US, others}...



# HPKB - TQO212

(TQO212a) During August 10 1998 in the 1998 Iranian-Taliban Crisis, how might Iran's tensions with the Taliban be affected by the hypothetical event in which the Taliban release citizens of Iran?

Answer:

- Tensions would decrease.
- Because the abduction of the Iranian diplomats and citizens led to an escalation of tensions between Iran and the Taliban, their safe return would likely have the effect of reducing hostilities.

Source(s):

- 1998 Iranian-Taliban Crisis Historical Case.
- Definition of escalation.

# HPKB - TQO212b

(TQO212b)

During August 10 1998 in the 1998 Iranian-Taliban Crisis, how might Iran's tensions with the Taliban be affected by the hypothetical event in which the Taliban attempt to pressure Iran into Iran recognize Taliban as the official government of Afghanistan?

Odpowiedzi:

- Tensions would increase.
- Because relations between Iran and the Taliban are poor prior to the crisis, and since the Taliban pose a threat to Iran, Iran would likely resent any Taliban efforts to compel Iran into granting them diplomatic recognition during a crisis situation. The Taliban's pressure tactics and the Iranian response would serve to increase tensions between them.

Source(s):

- 1998 Iranian-Taliban Crisis Historical Case.

# HPKB - TQO212

(TQO212a)

During August 10 1998 in the 1998 Iranian-Taliban Crisis, how might Iran's tensions with the Taliban be affected by the hypothetical event in which the Taliban release citizens of Iran?

Odpowiedzi:

- Tensions would decrease.
- Because the abduction of the Iranian diplomats and citizens led to an escalation of tensions between Iran and the Taliban, their safe return would likely have the effect of reducing hostilities.

Source(s):

- 1998 Iranian-Taliban Crisis Historical Case.
- Definition of escalation.

# HPKB - TQO212b

(TQO212b)

During August 10 1998 in the 1998 Iranian-Taliban Crisis, how might Iran's tensions with the Taliban be affected by the hypothetical event in which the Taliban attempt to pressure Iran into Iran recognize Taliban as the official government of Afghanistan?

Odpowiedzi:

- Tensions would increase.
- Because relations between Iran and the Taliban are poor prior to the crisis, and since the Taliban pose a threat to Iran, Iran would likely resent any Taliban efforts to compel Iran into granting them diplomatic recognition during a crisis situation. The Taliban's pressure tactics and the Iranian response would serve to increase tensions between them.

Source(s):

- 1998 Iranian-Taliban Crisis Historical Case.

# CycSecure

Program symulujący ataki na sieci komputerowe.

Program składa się z:

- programu symulującego sieć komputerową na podstawie analizy parametrów rzeczywistej sieci.
- bazy wiedzy o zagadnieniach bezpieczeństwa sieci, ciągle uzupełnianej o nowe informacje
- analizatora planów ataku, odgrywającego rolę hackera dokonującego ataków na symulowaną sieć.

Zalety: nie zakłóca działania sieci;

pozwała na testowanie planów, sekwencji działań;

działa systematycznie, automatyzuje przyrost nowej wiedzy

Inne historie [wykorzystania CyC](#) z sukcesem

# Cyc Answers

Program zarządzający wiedzą i odpowiadający na pytania.

- Odpowiada sensownie na pytania odwołując się do swojej szerokiej wiedzy o świecie.
- Monitoruje stan swojej bazy wiedzy szukając niespójności, braków wiedzy.
- Integruje bazy danych i bazy wiedzy.
- Analizuje pytania na które nie zna odpowiedzi i podpowiada, jakiej wiedzy mu brakuje.
- Wspomaga wyszukiwanie informacji dla pytań, na które nie ma bezpośredniej odpowiedzi.
- Wykorzystuje różne profile użytkowników, odpowiadając w różny sposób, w zależności od użytkownika.

Np: jakie są tam możliwości wypoczynku? Zależy kto pyta ...

# Research Cyc

Od 2006 roku dostępna jest pełna wersja dla celów badawczych.

System programów Research Cyc Release 1.1 (2017) zawiera:

- Ponad 630,000 pojęć w ontologii ogólnej.
- Ponad 7 mln asercji (faktów i reguł), ponad 38,000 relacji częściowo definiujących (wiążących, ograniczających) koncepcje.
- Skompilowana wersja Inference Engine, Knowledge Base Browser.
- Funkcje NLP i CycL-to-English.
- Narzędzia do formułowania zapytań tłumaczące na CycL.
- Szablony do wprowadzania wiedzy.
- Ontology Exporter do tworzenia plików ontologii w formacie OWL.
- Dokumentacja i tutoriale.
- Specyfikacja języka CycL, CycL-to-Lisp, CycL-to-C, itd.
- Specyfikacja Cyc API do tworzenia aplikacji ResearchCyc.

# Inne zastosowania



Ponad 600 licencji, wiele zastosowań wojskowych. Lista zastosowań.

Wiele rozwiązań wymagających integracji (cross-vertical business challenges).

- Cross-Vertical Solutions
- Energy    Pre-Spud Well Planning    Virtual Drilling Advisor
- Virtual Completions Advisor
- Healthcare    Patient Care and Throughput Monitor
- Financial Services    Virtual Financial Analysis
- Compliance Monitoring and Reporting
- Conversational Training Tool
- Resource Planner    Supply Chain Optimization
- Enterprise Knowledge Layer
- >200 szpitali
- Autonomous human-like cognitive Machine Reasoning\_





# Przykładowe pytania

- Dlaczego Watson może się przydać onkologom?
- Po co stworzono CYC?
- Jaki zmierzyć jego wielkość?
- Jakie są podstawowe zasady konstrukcji CYC?
- Co wiesz o języku CyCL?
- Jakie mechanizmy wnioskowania w nim zaimplementowano?
- Co to jest ontologia? Globalna ontologia?
- Opisać podstawową ontologię CYC.
- Co daje się zrobić z CYC?
- Jakiego rodzaju zastosowania CYC umożliwia?