

# WYKORZYSTANIE PROLOGA W IBM WATSON

---

---

## CO ROBI(Ł) WATSON?

- ▶ Thomas J. Watson - założyciel IBM
- ▶ IBM Watson - superkomputer, część projektu DeepQA
- ▶ DeepQA - Automatic Question Answering
- ▶ Zagrał w teleturnieju

Thomas J. Watson był założycielem firmy IBM. Na jego cześć nazwano super komputer, który odpowiada na różne pytania. I to dosłownie. IBM Watson jest częścią projektu DeepQA, który to projekt ma na celu stworzenie oprogramowania do przetwarzania języka naturalnego. Komputer przyjmuje pytanie zadane w języku angielskim, przetwarza je i odpowiada. Najlepszym przykładem jego możliwości jest wygrana w teleturnieju Jeopardy! (połączenie 1 z 10 z pokerem) Przeciwko mistrzom z poprzednich rozgrywek[1]. Co jeszcze trudniejsze Watson musiał odpowiedzieć na 2 pytania, ponieważ w Jeopardy! odpowiedź musi być podana w formie pytania.

---

## WATSON W LICZBACH

- ▶ Moc - 80 teraflopsów - 80.000.000.000.000 Flopsów
- ▶ 90 serwerów IBM Power 750
- ▶ 16 terabajtów ramu
- ▶ Zegar do 4.25 GHz

Po krótkce dużo. Komputery podobnej klasy zużywają energię na poziomie kilku mega watów. Komputer składa się z 90 serwerów IBM Power 750 mających w sobie procesor POWER7 pracujących na częstotliwości do 4.25 GHz oraz 16 terabajtów pamięci ram co pozwala przetwarzać dane z mocą 80 TeraFlopsów, czyli  $8 \cdot 10^{13}$  operacji zmiennoprzecinkowych na sekundę.

## DLACZEGO LUDZIOM TO TAK ŚWIETNIE WYCHODZI?

- ▶ Ludzki mózg ma 86 miliardów neuronów
- ▶  $10^{14}$  połączeń między nimi
- ▶ Każdy neuron jest połączony z około 1000 swoich sąsiadów
- ▶ Neurony w trakcie przepływu informacji mogą je przetwarzać
- ▶ Mimo to mózg zużywa 20 watów mocy

U człowieka informacje przetwarza mózg. Bardzo skomplikowany układ złożony z około 86 miliardów neuronów. Neurony są ze sobą wzajemnie połączone, w całym mózgu jest aż  $10^{14}$  połączeń między nimi, co daje wielkość rzędu  $10^3$  na neuron[2]. Mimo to mózg zużywa 20 W mocy. Taki wynik jest możliwy dzięki temu, że neurony są w stanie uśpienia. Aby je wybudzić potrzebny jest bodziec od sąsiedniego neuronu lub nawet kilku. Mózg jest bardzo skomplikowanym układem i nadal nie wiemy do końca jak działa. Niestety pomimo tych liczb nie potrafimy policzyć  $\ln((1726354 \cdot \pi)^2) / 31234 = 0.0009929$  z dokładnością do 4 cyfr po przecinku, oraz co więcej na przetwarzanie informacji w mózgu duży wpływ mają czynniki losowe. Komputer spisuje się znakomicie w przetwarzaniu obliczeń, a człowiek w interpretowaniu rzeczywistości.

[TVP Info](#) › [w-mozgu-odkryto-nowy-rodzaj-komunikacji-miedzy-neur...](#)

### W mózgu odkryto nowy rodzaj komunikacji między neuronami ...

19 lut 2019 - W mózgu odkryto nowy rodzaj komunikacji między neuronami ... To oznacza, że tych połączeń w mózgu Jana Kowalskiego może być mniej więcej tyle, ... O ile zapewnić im następnie pozostawanie w fizycznej bliskości.

[Pap.pl](#) › [pl](#) › [naukawpolsce](#) › [aktualnosci](#) › [news,28862,naukowcy-mo...](#)

### Naukowcy: mózg działa inaczej niż się wydawało | Nauka w ...

28 mar 2018 - Jak tłumaczą, słabe synapsy (połączenia między neuronami) uważano dotąd za mało istotne, mimo że stanowią większość. Nowy eksperyment ...

U człowieka informacje przetwarza mózg. Bardzo skomplikowany układ złożony z około 86 miliardów neuronów. Neurony są ze sobą wzajemnie połączone, w całym mózgu jest aż  $10^{14}$  połączeń między nimi, co daje wielkość rzędu  $10^3$  na neuron[2]. Mimo to mózg zużywa 20 W mocy. Taki wynik jest możliwy dzięki temu, że neurony są w stanie uśpienia. Aby je wybudzić potrzebny jest bodziec od sąsiedniego neuronu lub nawet kilku. Mózg jest bardzo skomplikowanym układem i nadal nie wiemy do końca jak działa. Niestety pomimo tych liczb nie potrafimy policzyć  $\ln((1726354 \cdot \pi)^2) / 31234 = 0.0009929$  z dokładnością do 4 cyfr po przecinku, oraz co więcej na przetwarzanie informacji w mózgu duży wpływ mają czynniki losowe. Komputer spisuje się znakomicie w przetwarzaniu obliczeń, a człowiek w interpretowaniu rzeczywistości.

---

## CO JEST TRUDNEGO W PRZETWARZANIU JĘZYKA NATURALNEGO?

- ▶ Dwuznaczność
  - ▶ Pojechał Polak do Niemiec i udawał Greka
- ▶ Kontekstowość
  - ▶ Tomek powiedział, że mam piękny sweter
  - ▶ śmiejąc się pod nosem
- ▶ Niedokładność
  - ▶ Moje auto stoi na parkingu koło AGH i jest niebieskie

Język naturalny ma kilka cech, które w znacznym stopniu utrudniają jego przetwarzanie. Sarkazm abstrakcja

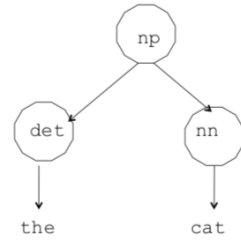
## DLACZEGO W WATSONIE UŻYWA SIĘ PROLOGA?

- ▶ Łatwy rozkład zdania na części składowe
- ▶ Łatwe tworzenie zasad języka
- ▶ Łatwe parsowanie
- ▶ Łatwe uwzględnianie relacji między słowami

W prologu w bardzo łatwy sposób można rozłożyć zdanie zapisane w języku naturalnym na części składowe. Na początku trzeba zapisać reguły danego języka. Na przykład tworzymy zasadę - term złożony - słowo "walk" to czasownik. Dzięki podstawianiu kolejnych słów - unifikacji - program może się dowiedzieć jakie słowa są jakimi częściami zdania i zrozumieć o co jest pytany. Oczywiście jest to bardzo prosty przykład, ponieważ bazy wiedzy prologa są o wiele większe i zawierają różne relacje.

## Terms

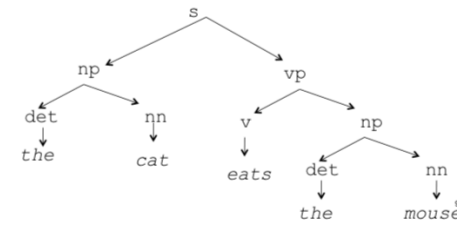
- The order is important (but free)
- Graphical representation of compound terms  
`np(det(the), nn(cat))`



8

## Terms

- Graphical representation of a compound term  
`s(np(det(the), nn(cat)), vp(v(eats), np(det(the), nn(mouse))))`



## Prolog

- And the rules (with comments!)  

```
% Agreement between two words
agree(Word1,Word2) :- number(Word1,N),
                      number(Word2, N).

% How to form noun phrase
np(Mod,Gov) :- determiner(Mod),
              noun(Gov), agree(Mod, Gov).
np(Mod,Gov) :- adjective(Mod), noun(Gov).
```
- The order of the rule is important
- We must give all the possible definitions of a rule together

9

## Prolog

- A simple program: enumerate facts and rules.  

```
determiner(the).
determiner(this).
determiner(these).
noun(cat).
noun(mice).
noun(sheep).
adjective(lazy).
number(this,singular).
number(these,plural).
number(cat,singular).
number(mice,plural).
number(sheep,_).
number(the, ).
```

10

W prologu w bardzo łatwy sposób można rozłożyć zdanie zapisane w języku naturalnym na części składowe. Na początku trzeba zapisać reguły danego języka. Na przykład tworzymy zasadę - term złożony - słowo "walk" to czasownik. Dzięki podstawianiu kolejnych słów - unifikacji - program może się dowiedzieć jakie słowa są jakimi częściami zdania i zrozumieć o co jest pytany. Oczywiście jest to bardzo prosty przykład, ponieważ bazy wiedzy prologa są o wiele większe i zawierają różne relacje.



## JAK WATSON ANALIZUJE ZDANIA?

- ▶ Rozkłada zdanie na części składowe
- ▶ Sprawdza relacje między słowami
- ▶ Buduje drzewo, które odzwierciedla
  - ▶ Budowę
  - ▶ Logikę

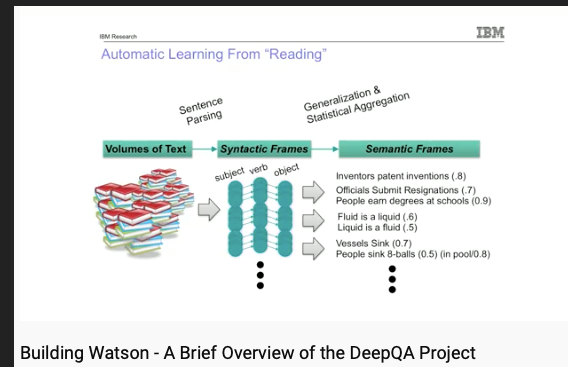
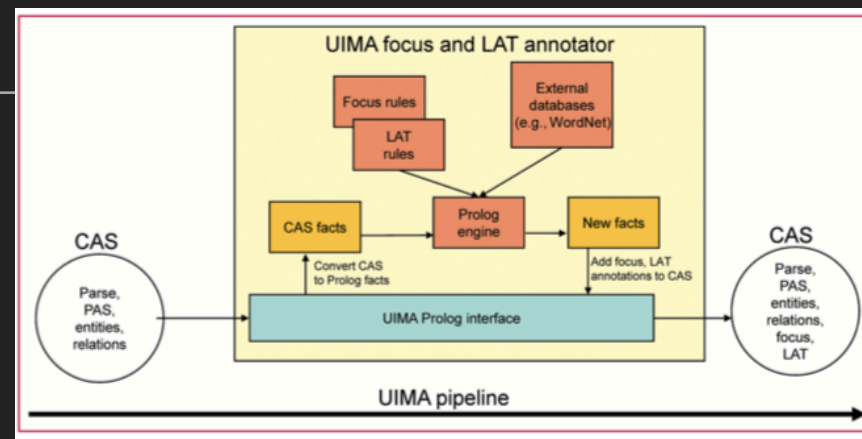
He was a bank clerk in the Yukon before he published "Songs of Sourdough" in 1907.

Watson buduje drzewo, które odzwierciedla zarówno budowę jak i strukturę logiczną zdania. Przypisuje każdemu słowu odpowiadającą mu funkcję oraz sprawdza relacje pomiędzy słowami. Prolog pod tym względem spisuję się świetnie ponieważ można łatwo wyrazić relacje i wzory relacji pomiędzy słowami.

[4]

Rozbior słowa - wezel w drzewie zawiera 1. słowo lub zestaw z argumentami logicznymi 2. liste cech 3. modyfikatory prawo i lewo stronne

UIMA- Unstructured Information Management Architecture



## Intermediate Hypotheses

- is ( "cytoplasm", "liquid" ) = 0.2
- is ( "organelle", "liquid" ) = 0.1
- is ( "vacuole", "liquid" ) = 0.2
- is ( "plasma", "liquid" ) = 0.7

He was a bank clerk in the Yukon before he published "Songs of Sourdough" in 1907.

Watson buduje drzewo, które odzwierciedla zarówno budowę jak i strukturę logiczną zdania. Przypisuje każdemu słowu odpowiadającą mu funkcję oraz sprawdza relacje pomiędzy słowami. Prolog pod tym względem spisuje się świetnie ponieważ można łatwo wyrazić relacje i wzory relacji pomiędzy słowami.

[4]

Rozbior słowa - wezeł w drzewie zawiera 1. słowo lub zestaw z argumentami logicznymi 2. liste cech 3. modyfikatory prawo i lewo stronne

UIMA- Unstructured Information Management Architecture

---

DZIĘKUJĘ ZA UWAGĘ

## ŹRÓDŁA

- ▶ [https://en.wikipedia.org/wiki/Watson\\_\(computer\)#Hardware](https://en.wikipedia.org/wiki/Watson_(computer)#Hardware)
- ▶ <https://souvenaid.pl/p/brain-nutrition>
- ▶ <https://www.youtube.com/watch?v=3G2H3DZ8rNc>
- ▶ <https://www.semanticscholar.org/paper/Natural-Language-Processing-With-Prolog-in-the-IBM-Lally-Fodor/c0ddc1e420ab0b2477865b160130335aad41dcc6>
- ▶ [https://researcher.watson.ibm.com/researcher/view\\_group\\_subpage.php?id=2159](https://researcher.watson.ibm.com/researcher/view_group_subpage.php?id=2159)
- ▶ <https://staff.fnwi.uva.nl/u.endriss/teaching/pss/slides/pss-prolog-slides.pdf>
- ▶ [http://brenocon.com/watson\\_special\\_issue/02%20question%20analysis.pdf](http://brenocon.com/watson_special_issue/02%20question%20analysis.pdf)
- ▶ [http://members.unine.ch/jacques.savoy/lectures/SemCL/PrologNLPCL.pdf?fbclid=IwAR0EDzMaxGUxJ2kYuYHIqDgKqbf1VIFOz\\_bkITHH8XUC2zrNbn5BS-7ZzSo](http://members.unine.ch/jacques.savoy/lectures/SemCL/PrologNLPCL.pdf?fbclid=IwAR0EDzMaxGUxJ2kYuYHIqDgKqbf1VIFOz_bkITHH8XUC2zrNbn5BS-7ZzSo)