

Sztuczna Inteligencja

Język naturalny i modele językowe

Włodzisław Duch

Katedra Informatyki Stosowanej UMK

Google: Włodzisław Duch

Co było:



- Reprezentacja wiedzy
- Systemy ekspertowe
- Predykcyjne metody uczenia maszynowego
Analiza języka w stylu GOFAI ma tylko historyczne znaczenie, jesteśmy w epoce dużych modeli językowych (LLM) i multimodalnych.
- Czym się zajmuje NLP
- Języki formalne i ich gramatyki
- Rozpoznawanie mowy
- Przykłady starszych programów
- Tłumaczenie maszynowe
- Sieci neuronowe i algorytmy językowe.
- Transformery
- Inne techniki.

Natural Language Processing

- NLP to dział AI, w skład którego wchodzi:
- Teoria gramatyk i języków formalnych
- Szukanie dobrych reprezentacji wiedzy zawartej w tekstach
- Cele: wspomaganie prac z tekstami, tworzenie słowników, programów nauczających języki obce, dialog z komputerem w języku naturalnym (ludzkim), pisanym i mówionym;
- analiza sensu/gramatyki istniejących tekstów, streszczenia;
- komentowanie tekstu, wydobywanie informacji – szukarki!
- Tłumaczenie maszynowe, analiza i synteza mowy.
- Wymagania sprzętowe i programowe – wielkie, ale nie dla drugiej dekady XXI wieku!

Podstawowe pojęcia

Nauczyć się języka nie jest łatwo ludziom ani maszynom, ale postępy w ostatnich latach są ogromne.

Syntaktyka języka: zajmuje się formatem, regułami budowy i formalną składnią.

Gramatyka jest zbiorem reguł syntaktycznych.

Rozbiór gramatyczny (parsing) – uwidacznia nieliniowe powiązania części zdania, do pewnego stopnia widoczne w tradycyjnym rozborze (podmiot, orzeczenie).

Semantyka języka: znaczenie (sens) wyrażeń i konstrukcji w danym języku – zrozumiały tylko w kontekście, często szerokim (społecznym).
To najtrudniejsza część NLP.

Historia NLP

1950, Wczesne wysiłki: identyfikacji słów kluczowych, prymitywne programy do tłumaczenia maszynowego.

1957, N. Chomsky, Syntactic Structures, formalizacja języków.

1960, Usprawnienia: baza danych, indeksacja, dopasowanie do wzorców wypowiedzi, podstawowe analizy gramatyczne.

1970, Reprezentacja logiczna, analiza semantyczna, rola kontekstu, rozszerzone sieci semantyczne (ATM).

1975, Systemy wykorzystujące bazy wiedzy, wnioskujące, rozumujące.

1980, Programowanie logiczne, gramatyki ograniczeń, modele planowania do rozumienia i generacji tekstu.

1990, Formalizacja sieci semantycznych, modele statystyczne, metody koneksjonistyczne, integracja mowy i tekstu, obszerne językowe bazy danych, pragmatyka języka.

2000, Gramatyki probabilistyczne (data oriented parsing)

2010, Deep learning i sieci neuronowe z rekurencją.

- 2018, **Bidirectional Encoder Representations from Transformers (BERT)**
- 2022, Szeroki dostęp do Large Language Models: ChatGPT, Bing, Bard ...

Gramatyki i języki formalne

- Jak opisać język, proces analizy od dźwięków i grafemów do znaczenia słów i komunikacji?
- *Język formalny*: zbiór łańcuchów symboli utworzonych ze skończonego zbioru symboli końcowych
- Gramatyka:
 - 1. *Kategorie syntaktyczne*, takie jak <zdanie> i <fraza>, czyli zmienne.
 - 2. *Symbole końcowe* lub terminalne języka, np. słowa języka naturalnego, łączone w zdania.
 - 3. Symbole nieterminalne, zmienne syntaktyczne, używane tymczasowo.
- Język \Leftrightarrow zbiór wszystkich zdań akceptowalnych przez jakąś gramatykę. Zdania określone są przez:
 - 3. *Symbol początkowy* S , zdania generowane są z S (np. symbolu nieterminalnego).
 - 4. *Reguły przepisania lub produkcji*, relacje pomiędzy łańcuchami symboli; pozwala to tworzyć frazy.

Gramatyki - przykłady

Bardzo prosta gramatyka: $S \rightarrow NP VP$

$\langle ZDANIE \rangle \rightarrow \langle FRAZA Z RZECZOWNIKIEM \rangle NP$
 $\langle FRAZA Z CZASOWNIKIEM \rangle VP$

$\langle FRAZA Z RZECZOWNIKIEM \rangle \rightarrow \langle RZECZOWNIK \rangle$

$\langle RZECZOWNIK \rangle \rightarrow \text{pies}$

$\langle RZECZOWNIK \rangle \rightarrow \text{kot}$

$\langle FRAZA Z CZASOWNIKIEM \rangle \rightarrow \text{biegnie}$

Gramatyka wyrażeń logicznych - mamy symbole logiczne: P, Q, \wedge ;

konstrukcja poprawna $P \wedge Q$,

konstrukcja niegramatyczna: $P Q \wedge$

Gramatyki formalne są przydatne do tworzenia języków programowania.

Trzeba zdefiniować reguły przypisania symboli.

Typy gramatyk (Chomsky 1957)

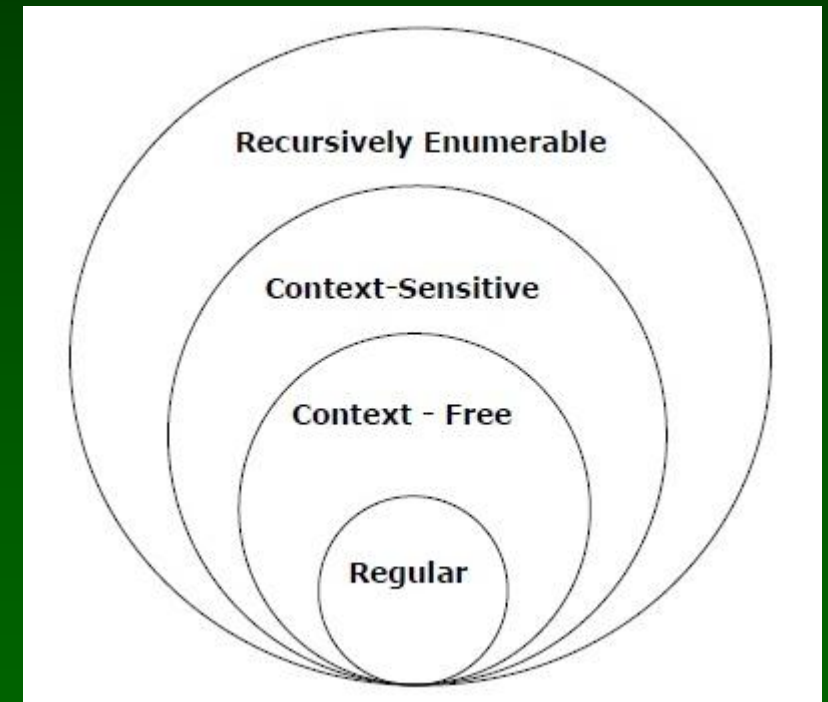
- 4 typy gramatyk formalnych, z różnymi ograniczeniami na reguły gramatyczne, od całkiem ogólnych do coraz bardziej ograniczonych, rozpoznawalne przez automaty skończone różnego typu.

Typ 0, Recursively Enumerable, to gramatyki rekurencyjnie przeliczalne, bez ograniczeń na postać reguł, generowane przez maszynę Turinga

Typ 1 Context-sensitive, gramatyki kontekstowo=wrażliwe.

Typ 2 Context-free grammar, gramatyki bezkontekstowe.

Typ 3 Regular grammar, gramatyki i języki regularne.



Typ 0

- Typ 0, recursively enumerable, czyli rekurencyjnie przeliczalna, to gramatyka najbardziej ogólna.
- Nie ma ograniczeń co do reguł przepisania, $X \rightarrow Y$ dla dowolnych ciągów X, Y , np. $S \rightarrow ACaB$.
- Generuje wszystkie możliwe języki formalne rozpoznawalne przez maszyny Turinga, obejmuje wszystkie typy gramatyk: regularne, kontekstowe i bezkontekstowe.
- Język rekurencyjnie przeliczalny jest rekurencyjnie przeliczalnym podzbiorem w zbiorze wszystkich możliwych słów w alfabecie języka.
- Język rekurencyjnie przeliczalny to język formalny, dla którego istnieje maszyna Turinga (lub inna obliczalna funkcja), która odróżnia wszystkie prawidłowe ciągi języka.
- Język rekurencyjnie przeliczalny to język formalny, dla którego istnieje maszyna Turinga (lub inna obliczalna funkcja), która zaakceptuje dowolny ciąg tego języka jako dane wejściowe.
- Maszyna Turinga może zatrzymać się, odrzucić lub zapętlić, jeśli przedstawić jej ciąg, który nie należy do języka rekurencyjnie przeliczalnego.

Typ 1

- **Typ 1, context-sensitive**, czyli **gramatyka kontekstowa**.
- Regułę przepisania $X \rightarrow Y$ można stosować tylko w danym kontekście.
- Reguły przepisania są typu $aXb \rightarrow aYb$, gdzie: X jest symbolem nieterminalnym, a, b, Y są dowolnymi ciągami symboli terminalnych lub nieterminalnych, Y jest ciągiem niepustym, a, b mogą być puste.
Np. $aXc \rightarrow aadc$, można zamienić $X \rightarrow ad$ tylko w kontekście $a .. c$.
- Gramatyka generuje język kontekstowy, rozpoznawalny przez automat liniowo ograniczony (jego pamięć jest funkcją liniową długości wejścia).
- Definiuje się też języki w oparciu o gramatyki lewo-kontekstowe i prawo-kontekstowe, ograniczając reguły do $aA \rightarrow ac$ albo $Ab \rightarrow cb$.
- Przykłady generacji języków kontekstowo-wrażliwych są tu.
- Słowo => wiele reprezentacji. Sieci neuronowe konstruuja reprezentacje słów w danym zdaniu w określonym kontekście. Model LSTM może utworzyć zanurzenie (embedding) kontekstowe.
- Embeddings from Language Model (ELMo): sekwencja słów => sekwencja wektorów. Z tego algorytmu rozwinęły się transformery.

Typ 2

Typ 2, gramatyka bezkontekstowa (context-free, CFG).

Skończony zbiór reguł przepisania $X \rightarrow Y$, gdzie X jest pojedynczym, niekończącym symbolem, a Y dowolnym zbiorem symboli terminalnych i nieterminalnych.

Języki bezkontekstowe odpowiadają niedeterministycznym automatom ze stosem, dają się przedstawić graficznie za pomocą struktur drzewiastych.

Np: gramatyka $(\{a,b\},\{S\},\{S \rightarrow aSb \mid \emptyset\})$ generuje język $a^n b^n$

Język palindromów nad alfabetem $\{a,b\}$ można wygenerować za pomocą gramatyki:

$(\{a,b\},\{S\},\{S \rightarrow aSa \mid bSb \mid a \mid b \mid \emptyset\},S)$

Np. $S \rightarrow aSa \rightarrow baSab \rightarrow abaSaba \rightarrow abaaaba \}$

Problem: jedno słowo może mieć tylko jedno znaczenie.

Typ 3, języki regularne

Typ 3, gramatyka regularna (liniowa) dopuszcza następujące reguły przypisania:

$X \rightarrow a$; lub $X \rightarrow a Y$ lub $X \rightarrow \emptyset$.

X, Y pojedyncze zmienne nieterminalne, a to pojedyncza zmienna końcowa.

Tworzy wyrażenia regularne, realizowane przez automaty skończone.

Wyrażenia regularne to zdania języków regularnych, pozwalają na identyfikację wzorców tekstu, są efektywne algorytmy by określić, czy szukany ciąg znaków pasuje do wzorca. Jest wiele narzędzi i bibliotek dla wszystkich używanych obecnie języków programowania. W praktyce wyrażenia regularne mogą nie być częścią języka regularnego ze względu na referencje wsteczne (np. w Perl).

Język typu p = język generowany przez gramatykę p .

Gramatyki typów wyższych to podzbiory typów niższych, języki też.

Typy 2 i 3 mają silne ograniczenia, ale 2 jest popularna w analizie języka naturalnego, daje spore możliwości.

Typ 3 nadaje się tylko do języków sztucznych, ale jest szybka i dlatego wykorzystywana w procesach szukania.

Analiza syntaktyczna

Zadanie: wydobyć powiązania elementów zdania, zrozumieć funkcję jaką w tym zdaniu pełnią. Używane gramatyki narzucają ograniczenia na rozbiór zdania.

Drzewo syntaktyczne tworzone jest przez parser korzystając z leksykonu.

Dla języków naturalnych potrzebna jest przynajmniej gramatyka bezkontekstowa.

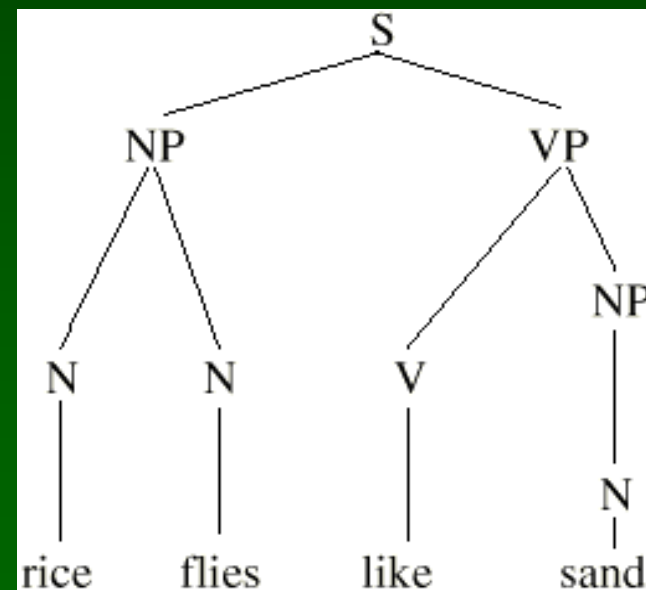
Leksykon zawiera informację o kategoriach syntaktycznych słów, gramatyka określa ich możliwe powiązania.

Drzewo dla zdania:

S = Rice flies like sand.

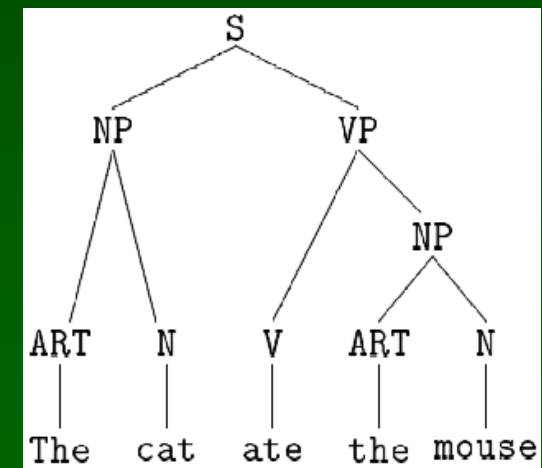
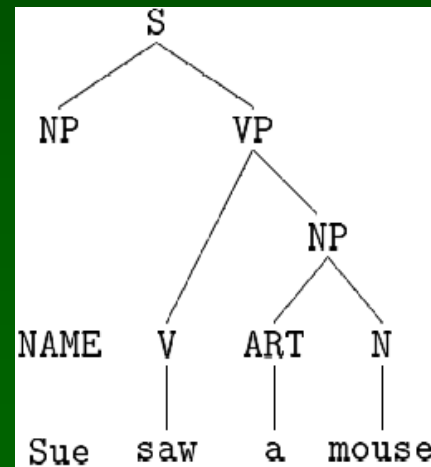
„Ryżowe muszki”, czy „ryż sypie się”?

Do tego potrzebna jest analiza kontekstu, nie wystarcza formalna analiza.



Przykład: gramatyka bezkontekstowa

Sentence	$S \rightarrow NP VP$
Verb Phrase	$VP \rightarrow V NP$
Noun Phrase	$NP \rightarrow NAME$ $NP \rightarrow ART N$
Article	$ART \rightarrow a \mid the$
Verb	$V \rightarrow ate \mid saw$
Noun	$N \rightarrow cat \mid mouse$
Name	$N \rightarrow Sue \mid Zak$



Inne rodzaje gramatyk

Jest wiele różnych rodzajów gramatyk, np:

Gramatyka generatywna – tworzy strukturalny opis zdania, pozwala rozpoznać i stworzyć nieskończoną liczbę zdań o prawidłowej strukturze.

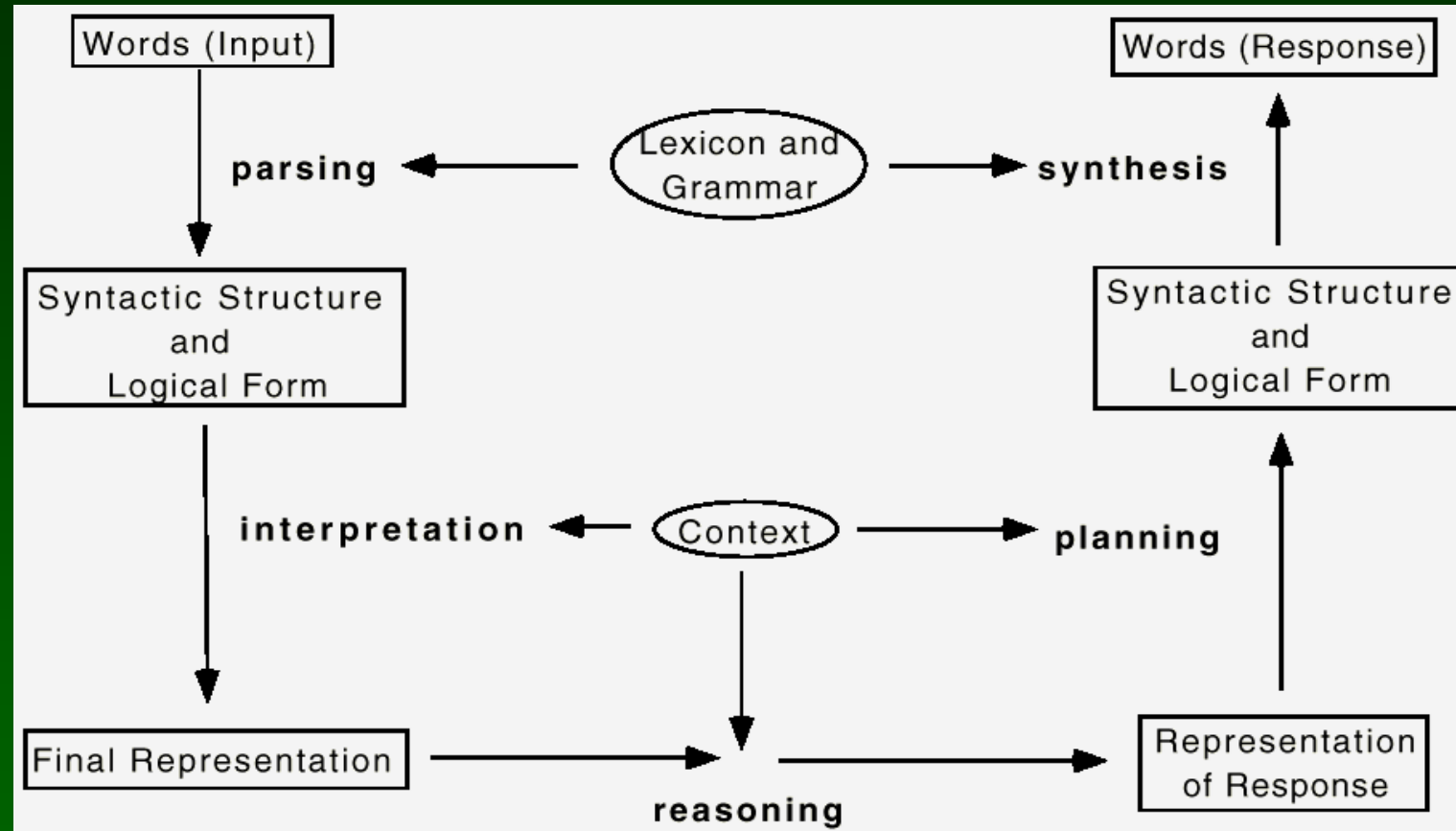
Gramatyka transformacyjna – głębokie znaczenie zdań jest niezmiennicze ze względu na transformacje „powierzchniowe” zdania.

Gramatyki systemiczne – część funkcjonalnej lingwistyki systemicznej, sieci systemowe, opisujące możliwości wyboru powiązanych znaczeń za pomocą „leksykogramów”, słów w powiązaniu z regułami.

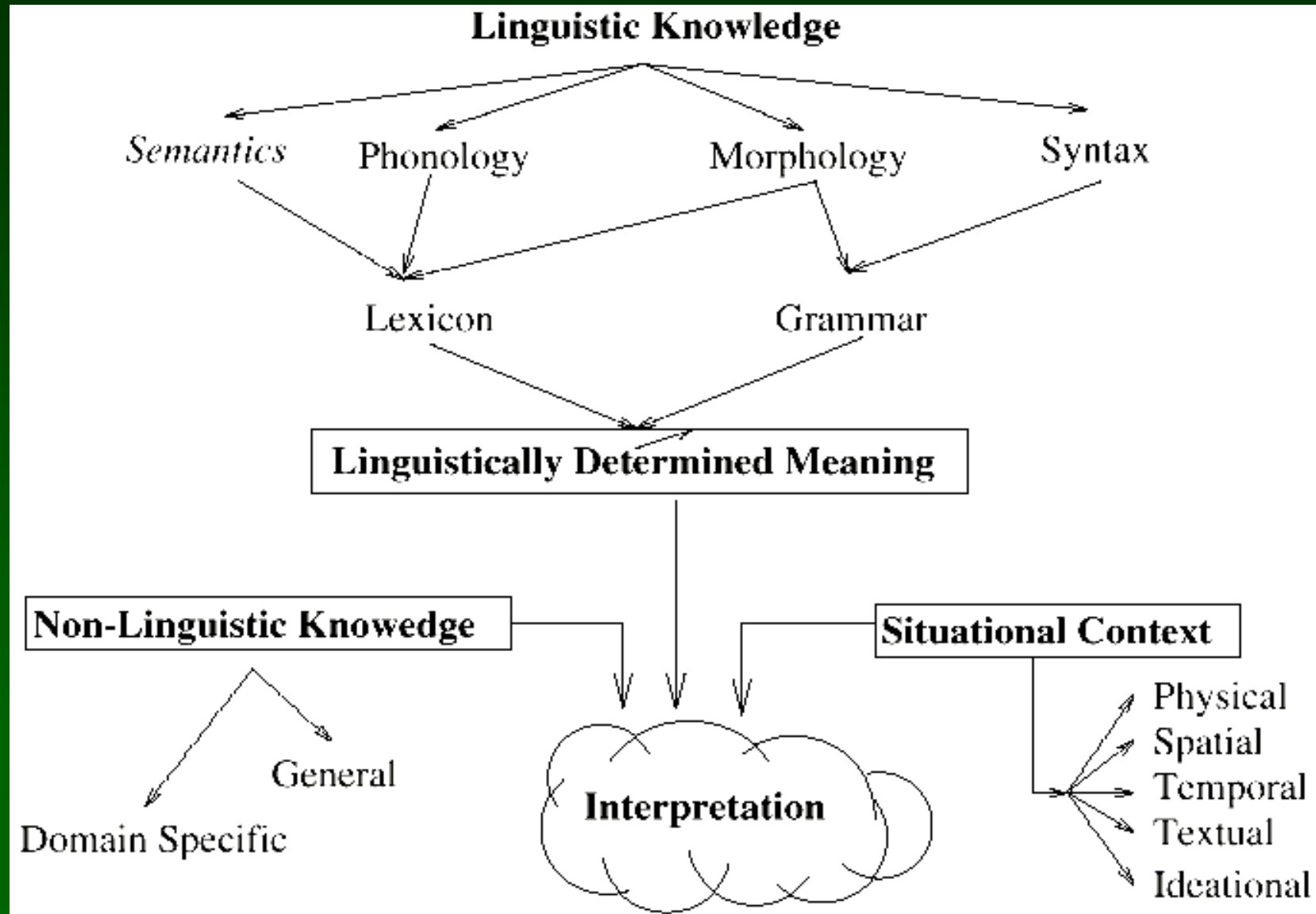
Gramatyki „przypadków” (case grammars), czyli „semantycznie istotnych powiązań syntaktycznych”; np. agent, obiekt, źródło, cel.

Parsery (analizatory składniowe), to programy do analiz struktur syntaktycznych. Różne strategie i rodzaje gramatyk w efekcie generują różne drzewa syntaktyczne.

Architektura systemów NLP



Język i interpretacja



5 kroków analizy

- **Fonologiczna** – fonetyka zajmuje się strukturą dźwięków; fonemy, formantami – istnieje międzynarodowy alfabet fonetyczny; fonologia zajmuje się organizacją fonemów w wypowiedziach.
- **Morfologiczna** – struktura słów, fleksja, rdzenie, leksykony wykorzystujące regularności języka na poziomie słów.
- **Syntaktyczna** – konstrukcja zdań, rozbiór gramatyczny.
- **Semantyczna** – sens, jednoznaczność, wybór słów, zależnie od kontekstu; markery semantyczne i organizacja leksykonu.
- **Pragmatyczna** – sposób użycia wyrazów, relacje zaimki-fraza; relacje między frazami, relacje przyczynowe, rozumienie całości, założenia.

Kroki analizy językowej.

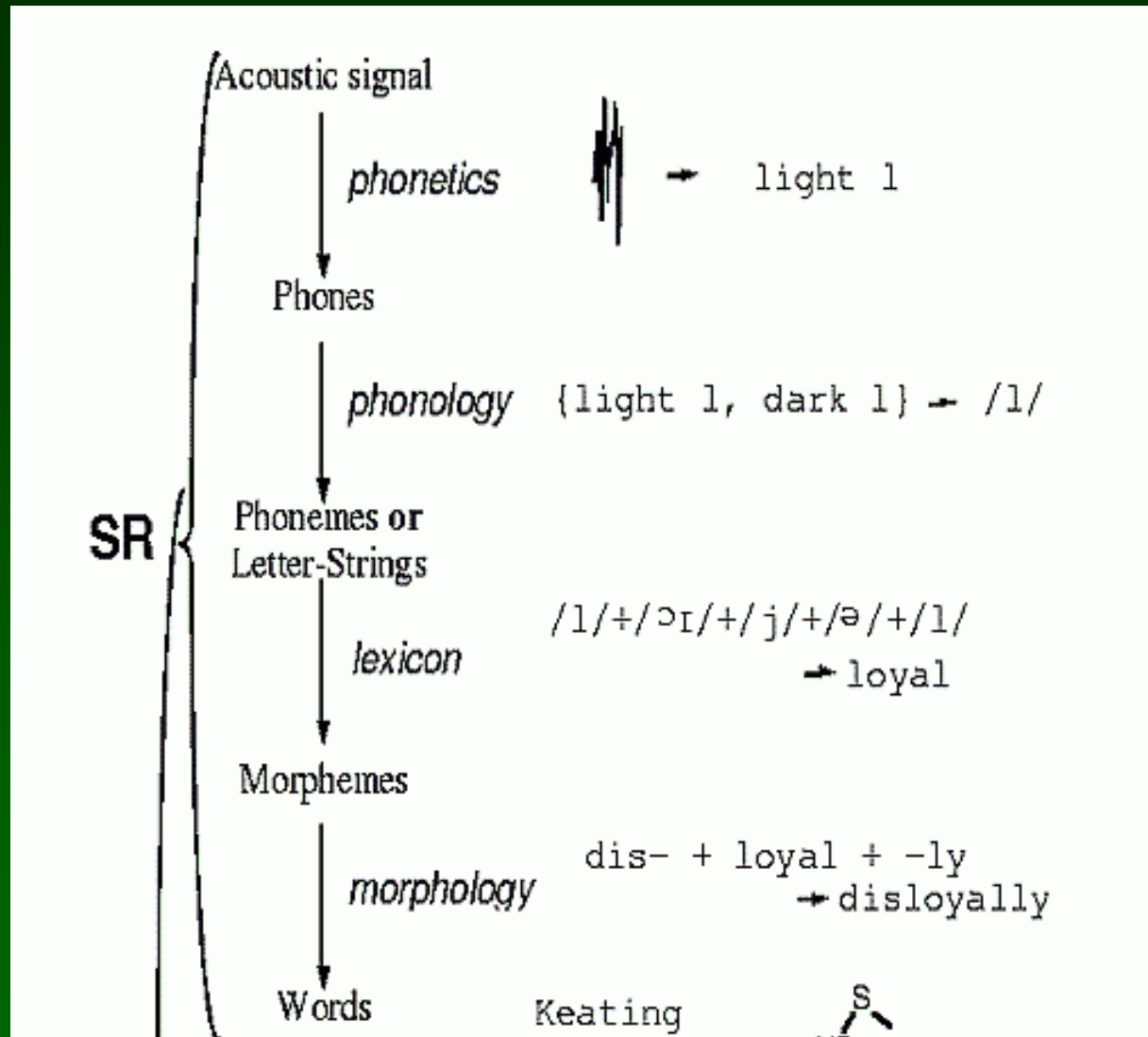
Speech recognition, czyli rozpoznawanie mowy, odbywa się tylko na poziomie fonetycznym + morfologicznym, ale korzysta ze słownika by wybrać bardziej prawdopodobną interpretację rozpoznawanych dźwięków.

Wymaga to dużych baz nagrań wypowiedzi różnych osób, i zaawansowanych algorytmów analizy sygnałów mowy.

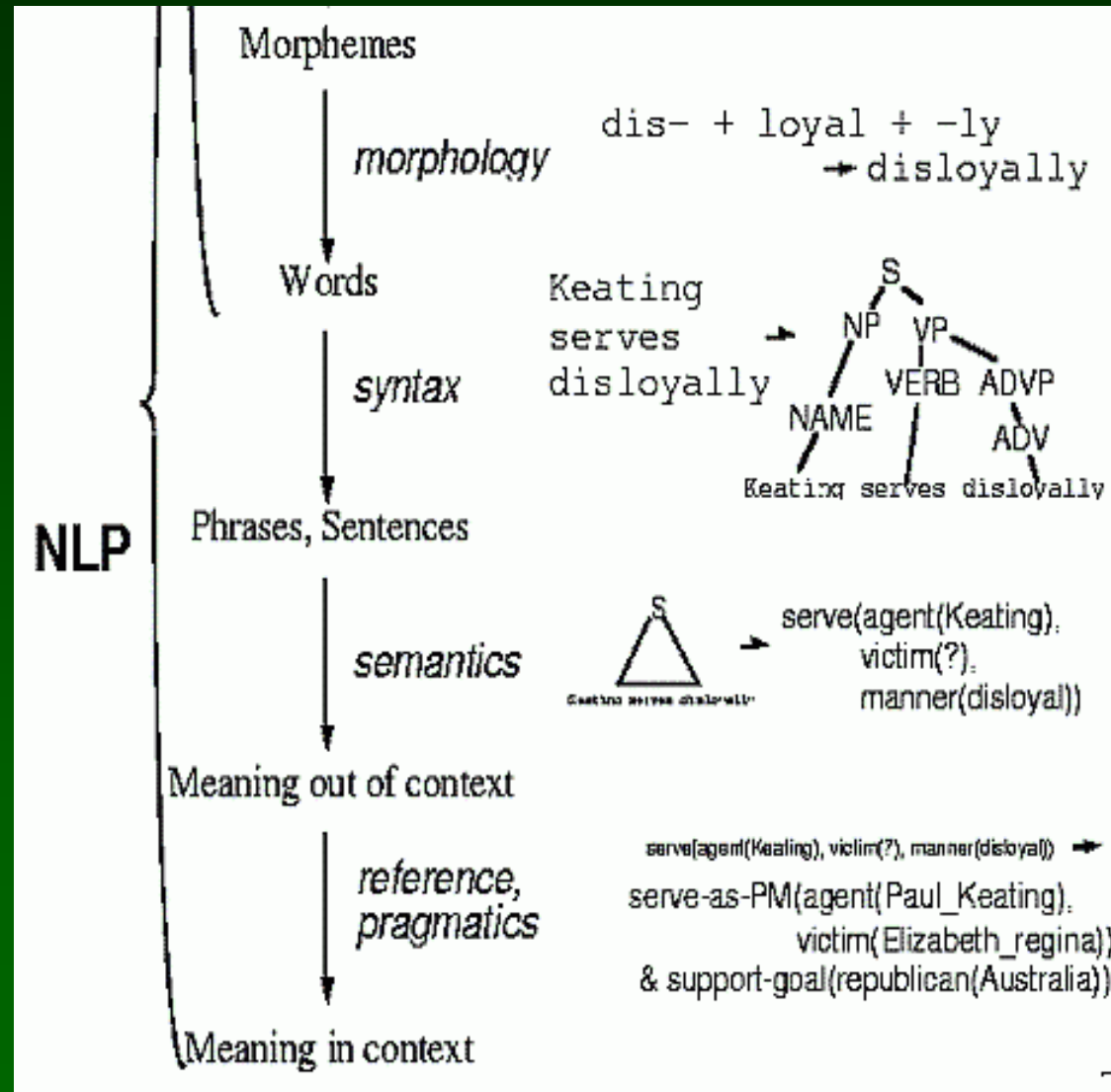
Natural Language Processing (NLP), czyli analiza języka naturalnego, zaczyna się od poziomu

- morfologicznego, obejmuje też poziom
- syntaktyczny,
- Semantyczny,
- pragmatyczny.

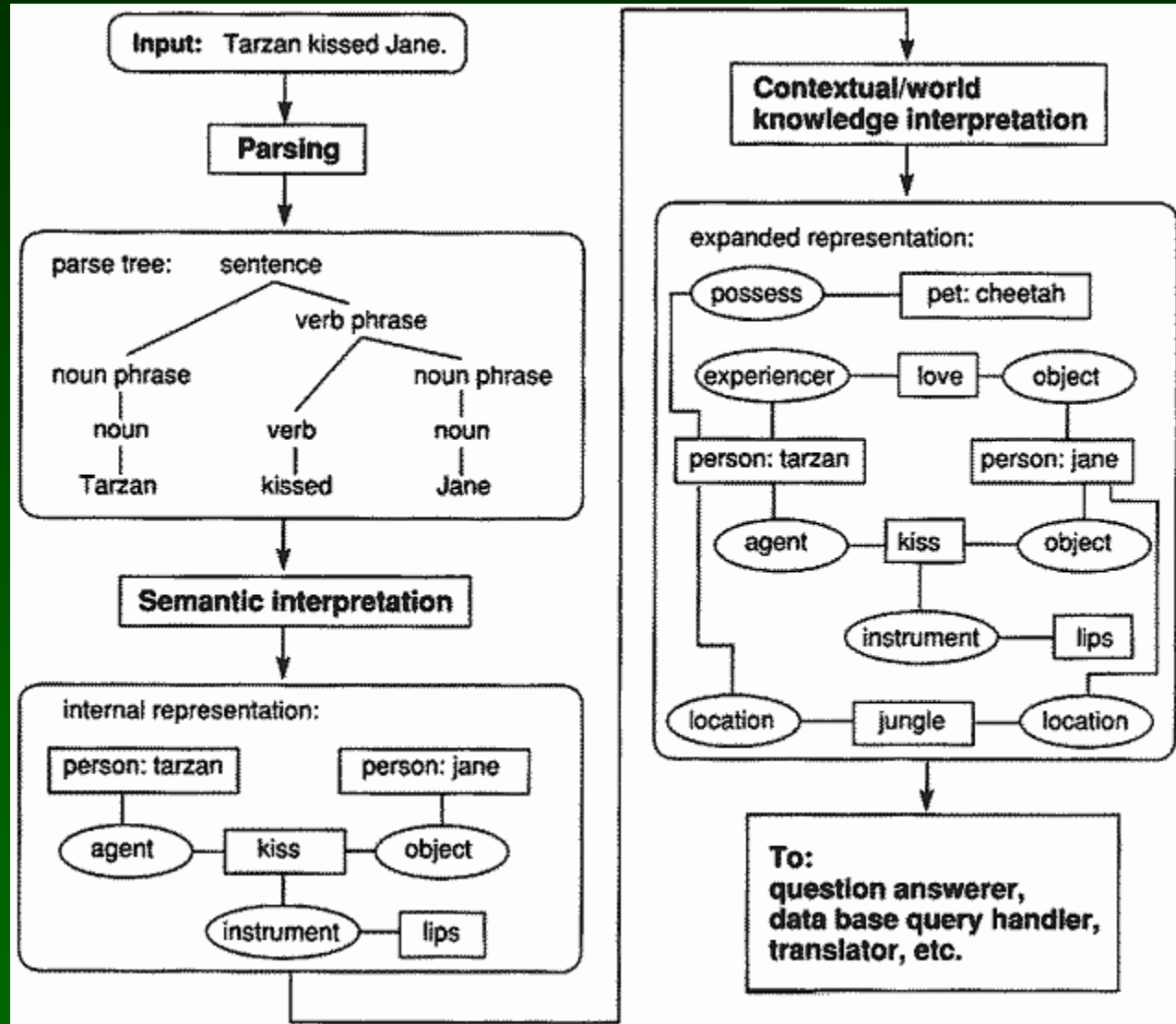
Schemat rozpoznawania mowy



Schemat rozpoznawania



Kroki interpretacyjne



Przykłady programów NLP

Dobrze jest wiedzieć, jakie pomysły wypróbowano. Wczesne programy:

SAD-SAM (R. Lindsay 1963)

Syntactic Appraiser and Diagrammer-Semantic Analyzing Machine, ok. 1700 słów, odpowiada na pytania dotyczące stopnia pokrewieństwa, buduje drzewa genealogiczne, gramatyka bezkontekstowa.

S: Joe bawi się w ogródku ciotki Jane.

Q: Kim jest Joe dla Jane?

BASEBALL (B. Green, 1963)

Baza danych o spotkaniach baseballowych, odpowiadająca na pytania w rodzaju:

Ile gier grali Yankees w lipcu?

Z jakiego miasta jest drużyna Boston?

Przykłady: SIR

SIR (B. Raphael 1968), Semantic Information Retrieval

24 struktury językowe, wyrażające relacje zbiorów, relacje przestrzenne i ilościowe, szablony typu:

* is *, * is part of *, how many * does * have, what is the * of * ?

* = rzeczowniki, lub „wszystkie, dwa, każdy, ... „.

SIR wyciąga proste wnioski z podanych faktów, domagając się ich doprecyzowania:

Każdy chłopiec to człowiek

ROZUMIEM

Palce to część ręki

ROZUMIEM

Człowiek ma 2 ręce

ROZUMIEM

Przykłady: SIR

Ile palców ma John?

POWYŻSZE ZDANIE NIE JEST JEDNOZNACZNE ALE ZAKŁADAM, ŻE ma OZNACZA ma jako części.

NIE WIEM CZY palce TO CZĘŚĆ John

John to chłopiec

ROZUMIEM

Ile palców ma John?

POWYŻSZE ZDANIE NIE JEST JEDNOZNACZNE ALE ZAKŁADAM, ŻE ma OZNACZA ma jako części.

ILE PALCÓW MA RĘKA ?

Każda ręka ma 5 palców.

POWYŻSZE ZDANIE NIE JEST JEDNOZNACZNE ALE ZAKŁADAM, ŻE ma OZNACZA ma jako części.

ODPOWIEDŹ: MA 10

Kto jest prezydentem USA?

NIEZNANA FORMA PYTANIA

Przykłady, STUDENT

STUDENT (D. Bobrow 1968)

Analizuje i rozwiązuje zadania „z treścią” z matematyki, np.:

Jeśli liczba klientów Tomka jest dwa razy większa od kwadratu 20 procent liczby ogłoszeń, które daje, a liczba jego ogłoszeń wynosi 45, ilu Tomek ma klientów?

MTS, Machine Translation System (Wilks 1973)

Korzysta z pośredniej reprezentacji semantycznej, formuł logicznych i wzorców zdań. Ok. 60 elementów podzielonych na 5 klas.

Np klasa *działania* zawiera elementy *przyczyna* i *przemieszczenie*.

„Pić” - działanie ożywionych podmiotów na materiał płynny,

ale: „Samochód pije benzynę”.

Słownik zawiera tłumaczenia wyrazów i stereotypy ich użycia.

Przykłady, LUNAR

LUNAR (Webber 1972) – próbki skał z księżyca

System do wyszukiwania informacji z bazy danych o skałach z wyprawy Apollo-11. Zdania angielskie przekłada na formuły logiczne i używa rachunku predykatów.

Analiza gramatyczna: gramatyki ATN, ok. 3500 słów, oraz reguły semantyczne interpretujące pytania do bazy danych. Radzi sobie z czasami, modalnością i różnymi subtelnościami.

SHRDLU (T. Winograd 1972) – świat klocków, planowanie.

Język aktywuje odpowiednie procedury w słuchaczu.

Reprezentacja proceduralna: wiedza ukryta jest w strukturze programu symulującego robota układającego klocki na stole.

Robot komentuje swoje działania, planuje, ma model swojego mikroświata, odpowiada na pytania.

Nie da się tak działać w szerszym kontekście.

ELIZA

Program ELIZA/Doctor (J. Weizenbaum, MIT 1966).

<http://psych.fullerton.edu/mbirnbaum/psych101/Eliza.htm>

Komputerowy psychoterapeuta ze szkoły Rogers'a (niedyrektywnej); wychwytuje słowa kluczowe, pozwala się wygadać; jest sporo realizacji.

Sekretarka Weizenbauma domagała się prywatnej rozmowy z programem, bo ją „rozumiał”.

Psychiatra dr Colby zasugerował upowszechnienie tego programu w celach leczniczych oraz pisanie programów w celu symulacji pacjentów.

Jest takich programów kilka, ciekawe rezultaty daje konfrontacja komputerowego doktora z pacjentem.

- Jeśli pacjent jest katatoniczny lub w śpiączce to wystarczy program:

Start / Stop

Przykład: ELIZA, naiwne dialogi

Wszyscy mężczyźni są tacy sami.

W jakim sensie

zawsze czegoś od nas chcą.

Czy możesz podać konkretny przykład?

Na przykład mój chłopak mnie tu przysłał.

Twój chłopak cię tu przysłał.

On mówi, że jestem przygnębiona większość czasu.

Przykro mi słyszeć, że jesteś przygnębiona

to prawda. Nie jestem szczęśliwa.

Czy myślisz, że przyjdzie tutaj tobie pomoże?

Na pewno pomoc mi się przyda.

W jaki sposób można by ci pomóc?

Może mogłabym nauczyć się lepiej żyć ze swoją matką.

Opowiedz mi o twojej rodzinie

Program jest bardzo przyjacielski i chętnie słucha zwierzeń a ludzie się na to nabierają.

Margi

MARGI (Meaning, Analysis, Response Generation, Inference) to przykładu intuicyjnego rozumienia języka.

„Konceptyjna zależność” (conceptual dependency, CD), wewnętrzna reprezentacja znaczenia niezależna od formy.

11 działań, np. transakcja to brać, dawać, przekazać itp.

Związki pomiędzy pojęciami mają postać grafu, kto co komu. Stany obiektów opisane są numerycznie, np. zdrowie (od - 10 martwy do +10), zdrowie umysłowe (+/- 10) i zdrowie fizyczne.

CD używa łańcuchów przyczynowych działań, np.:

- Akcje mogą zmienić stany obiektów.
- Stany mogą umożliwić akcje.
- Stany mogą uniemożliwić akcje.
- Stany mogą zainicjować zdarzenia umysłowe.
- Zdarzenia umysłowe mogą być przyczyną działań.

Margi i SAM/PAM

Informacja w zdaniu jest ujawniana w reprezentacji wewnętrznej.

Np. „Jan je lody łyżeczką” to

„Jan, połyka, lody, łyżeczka, lody na łyżeczce, porusza, usta”.

Pozwala to na parafrazowanie zdań i odpowiadanie na pytania.

SAM (Script Applier Mechanism)

PAM (Plan Applier Mechanism)

Programy demonstrujące użycie scenariuszy i planów w rozumieniu historyjek, oparte na gramatykach CD.

W rozumieniu opowiadań może być ważny kontekst całości.

Scenariusze to stereotypy zachowań, używane do antycypacji zdarzeń.

System potrafi prowadzić sprawne wnioskowanie, np.: z informacji „kotlet był spalony” wnioskuje, że klient był niezadowolony i prawdopodobnie go nie zjadł ani nie dał napiwku.

Daje się to zastosować w opisie schematycznych sytuacji.

Problemy NLP

- Problem: brak wiedzy w programach! Trudno jest zrozumieć sens zdania, jeśli się mało wie.
- Reprezentacja znaczenia słów i fraz wymaga reprezentacji wiedzy.
- Konieczna jest nauka znaczenia fraz z przykładów kontekstowych.
- Rozstrzygnięcie, do czego lub kogo odnosi się dany fragment tekstu jest trudne i wymaga zrozumienia kontekstu (problem co-referencji).
- Potrzebne jest dynamiczne dopełnianie fraz na podstawie modelu świata, relacji pomiędzy pojęciami. Można to osiągnąć przez predykcję maskowanych fragmentów tekstu.
- Automatyczna konstrukcja baz danych i uczenie się z tekstów w języku naturalnym wymaga rozumienia tekstu.

Dwa podejścia:

- Budowa wielkich systemów opartych na wiedzy zwerbalizowanej, wykorzystanie metod wnioskowania i klasycznych form reprezentacji wiedzy, np. WATSON, CyC.
- Budowa wielkich sieci neuronowych, które rozpoznają kontekst, dobrze odpowiadają na liczne pytania, ale nie ma w nich głębszego zrozumienia.

Synteza i analiza mowy

Synteza jest stosunkowo prosta, ale stworzenie naturalnie brzmiących głosów udało się dopiero po 2020 roku.

Przykłady mowy syntetycznej.

Zastosowania: przesyłanie informacji tekstowej, synteza po stronie użytkownika, czytanie książek.

Rozpoznawanie mowy:

speaker-dependent, czyli system nauczony rozpoznawania mowy indywidualnej osoby radzi sobie dobrze;

- Na poziomie fonologicznym było to 50-60%, ale nowe modele zagnieżdżania sygnałów w przestrzeniach wektorowych, takie jak Word2Vec, osiągają ponad 90%.
- Na poziomie rozpoznawania słów dla mowy ciągłej i niezależnej od osoby (speaker-independent) komercyjne systemy osiągają 95-98% dokładności, zależnie od warunków. Szumy w pomieszczeniu obniżają dokładność. Dla modeli trenowanych głosem określonej osoby to >98%.
- Testy zrozumiałości mowy robi się na poziomie całych zdań z redundantnym kontekstem.

Praktyczne rezultaty

Dyktowanie tekstu działa w wielu językach w telefonach. WER, czyli Word Error Rate określa się uwzględniając braki i dodatkowe wyrazy:

Word Error Rate = (Substitutions + Insertions + Deletions) / Number of Words

Microsoft w 2017 roku osiągnął błąd na poziomie ekspertów w transkrypcji rozmowy telefonicznej. W obecnie dostępnych programach WER ~5%, ale w rzeczywistych rozmowach przez kiepskie linie telefoniczne to bardziej ponad 10% błędów.

IBM [Embedded ViaVoice](#) i ich Super Human Speech Recognition Program, automatycznie generuje podpisy na filmach. Na YouTube to teraz standardowa funkcja.

[WATSON API](#) nowe rekordy dokładności, 5.1% błędów dla mowy ciągłej, z użyciem sieci rekurencyjnej LSTM. Ludzie robią w podobnych warunkach około 5% błędów, ale powtarzają, dopytują się, zgadują, więc komunikacja jest sprawna.

[IBM WATSON Speech to Text](#) ma liczne zastosowania: podpisy pod wideo, call centers, dyktafony.

W kilku modelach samochodów działają systemy rozpoznawania mowy dla ograniczonej liczby poleceń (<100), pomimo hałasu.