

# Introduction to NLP

Steve Nouri

Head of Data Science at Nod.

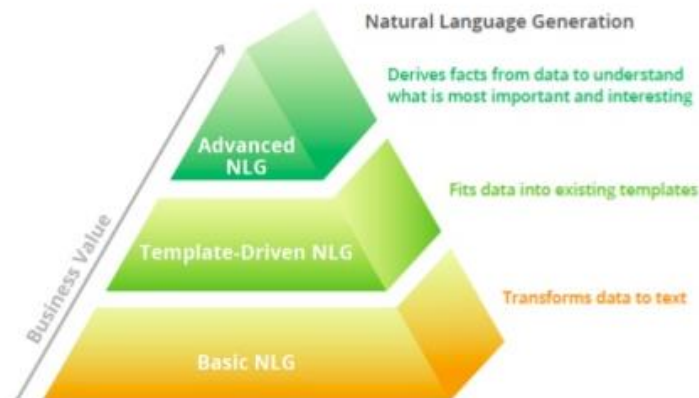
4<sup>th</sup> July, 2019

# Natural Language Processing

- Natural Language Processing (NLP) refers to AI method of communicating with an intelligent system using a natural language such as English, Spanish, Hindi etc.
- The goal of natural language processing is to allow non-programmers to obtain useful information from computing systems or give commands to the computing system using natural languages which they may speak or write.
- There is a vast store of information recorded in the Natural Language that could be accessible via computer system.

# Components of NLP

- Natural Language Understanding
  - Mapping the given input in natural language into useful representations.
- Natural Language Generation
  - Producing meaningful phrases and sentences in the form of natural language from some internal representation



## Importance

- NLP helps to make communication and handling easy between the user and computer system.
- Help to understand large social data available on the internet.
- Improve the efficiency and accuracy of documentation, and identify the most pertinent information from large databases.

# Applications of Nat. Lang. Processing

- Machine Translation
- Database Access
- Information Retrieval
  - Selecting from a set of documents the ones that are relevant to a query
- Text Categorization
  - Sorting text into fixed topic categories
- Extracting data from text
  - Converting unstructured text into structure data
- Spoken language control systems
- Spelling and grammar checkers

## Real world example

- Understand a Job Resume
- Match it with a Job Description
- Rank the resumes based on relevance
- Rank the resumes based on capability

# Ambiguity

- Ambiguity
  - Lexical ambiguity
    - Treating the word “board” as noun or verb?
  - Syntactical ambiguity
    - “He lifted the beetle with red cap”
    - Did he use cap to lift the beetle or he lifted a beetle that had red cap?
  - Referential ambiguity
    - Rima went to Gauri. She said, “I am tired.”
    - Exactly who is tired?



# Challenges

- Phrases / Idioms
  - “A perfect storm”
    - The worst possible situation
- Connecting language and machine perception
- Sentence generation
- Text summarization
- Keyword extraction



# Natural language understanding

Raw speech signal



- **Speech recognition**

Sequence of words spoken



- **Syntactic analysis** using knowledge of the grammar

Structure of the sentence



- **Semantic analysis** using info. about meaning of words

Partial representation of meaning of sentence



- **Pragmatic analysis** using info. about context

Final representation of meaning of sentence

# Natural Language Understanding

- Input/Output data
- Processing stage
- Other data used

**Frequency spectrogram**



**Word sequence**

"He loves Mary"

**Sentence structure**



**Partial Meaning**

$\exists x \text{ loves}(x, \text{mary})$

**Sentence meaning**

$\text{loves}(\text{john}, \text{mary})$

**speech recognition**

**syntactic analysis**

**semantic analysis**

**pragmatics**

freq. of diff.

sounds

grammar of

language

meanings of

words

context of

utterance

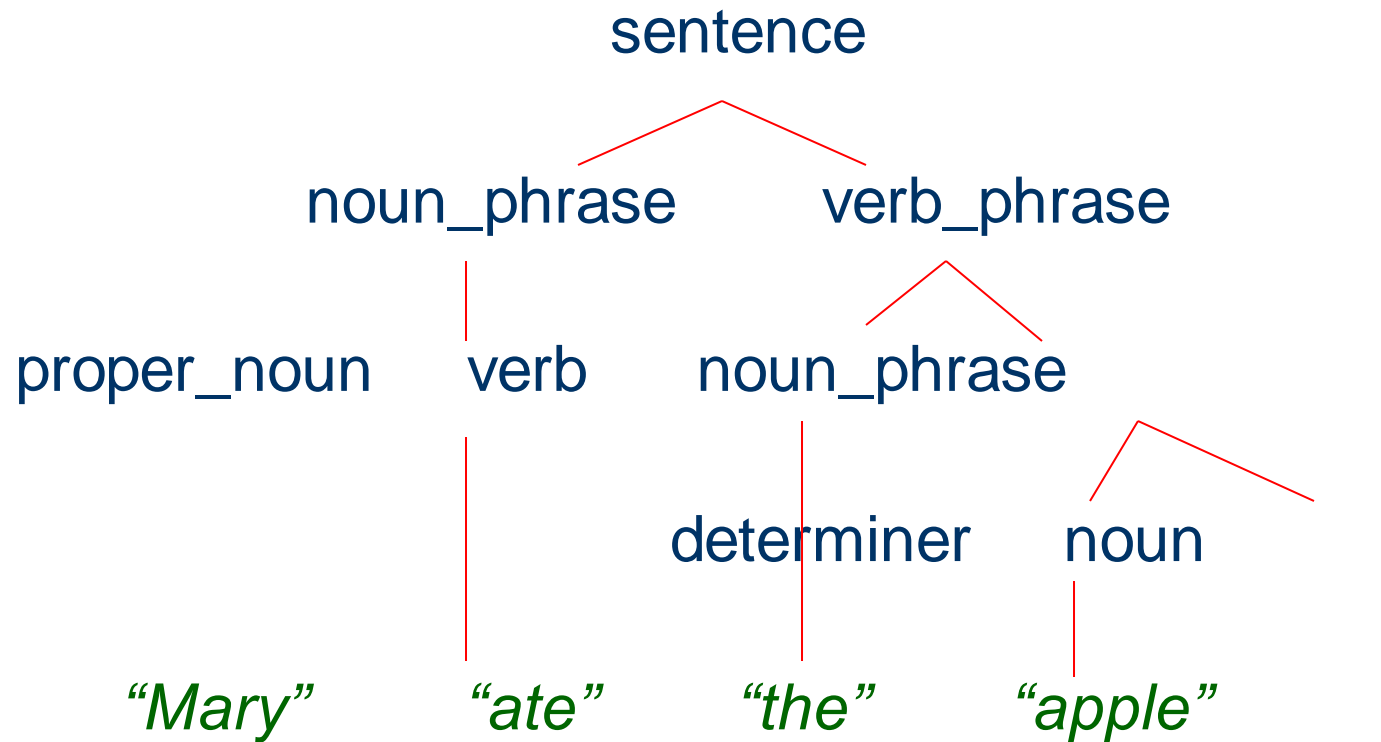
# Syntactic Analysis

- Rules of syntax (grammar) specify the possible organization of words in sentences and allows us to determine sentence's structure(s)
  - “John saw Mary with a telescope”
    - John saw (Mary with a telescope)
    - John (saw Mary with a telescope)
- Parsing: given a sentence and a grammar
  - Checks that the sentence is correct according with the grammar and if so returns a **parse tree** representing the structure of the sentence

# Syntactic Analysis - Grammar

- `sentence -> noun_phrase, verb_phrase`
- `noun_phrase -> proper_noun`
- `noun_phrase -> determiner, noun`
- `verb_phrase -> verb, noun_phrase`
- `proper_noun -> [mary]`
- `noun -> [apple]`
- `verb -> [ate]`
- `determiner -> [the]`

# Syntactic Analysis - Parsing



# Syntactic Analysis – Complications (1)

- Number (singular vs. plural) and gender
  - `sentence` → `noun_phrase (n) , verb_phrase (n)`
  - `proper_noun (s)` → [mary]
  - `noun (p)` → [apples]
- Adjective
  - `noun_phrase` → `determiner, adjectives, noun`
  - `adjectives` → `adjective, adjectives`
  - `adjective` → [ferocious]
- Adverbs, ...

## Syntactic Analysis – Complications (2)

- Handling ambiguity
  - Syntactic ambiguity: “fruit flies like a banana”
- Having to parse syntactically incorrect sentences



# Semantic Analysis – Complications

- Handling ambiguity
  - Semantic ambiguity: “I saw the prudential building flying into Boston”

# Pragmatics

- Uses context of utterance
  - Where, by who, to whom, why, when it was said
  - Intentions: *inform, request, promise, criticize, ...*
- Handling Pronouns
  - “Mary eats apples. She likes them.”
    - She=“Mary”, them=“apples”.
- Handling ambiguity
  - Pragmatic ambiguity: “you’re late”: What’s the speaker’s intention: informing or criticizing?

# Natural Language Generation

- Talking back! 😊
- What to say or text planning
  - flight(AA,london,boston,\$560,2pm),
  - flight(BA,london,boston,\$640,10am),
- How to say it
  - “There are two flights from London to Boston. The first one is with American Airlines, leaves at 2 pm, and costs \$560 ...”
- Speech synthesis
  - Simple: Human recordings of basic templates
  - More complex: string together phonemes in phonetic spelling of each word
    - Difficult due to stress, intonation, timing, liaisons between words

# Thank you!

Steve Nouri

Keep in touch and connect on

