Unit 1

Introduction to Natural Language Processing

Prof. Pranjal Pandit
CSE AI

Unit 1: Introduction to Natural Language Processing

- **Introduction: Scope, Applications and Challenges.**
- **Brief history and evolution of NLP: Programming languages Vs Natural Languages, Are Natural Languages Regular? Finite automata for NLP**
- **Stages of NLP**
- **NLP Basics of text processing: Tokenization, Stemming, Lemmatization, Part of Speech Tagging;**
- NLP Components: Syntax, Semantics, and Pragmatics;
- **Introduction to Regular Expressions and their Applications in NLP**.

NLP Components: Syntax, Semantics, and Pragmatics

Syntax

- **Definition**: The study of the structure and grammatical rules of language. It focuses on how words are arranged to form valid sentences.
- **Key Techniques**:
  - **Parsing**: Analyzing the grammatical structure of a sentence.
    - Example: Identifying subject, verb, and object in "The cat sat on the mat."
  - **Part-of-Speech Tagging**: Assigning grammatical labels (e.g., noun, verb) to each word in a sentence.
    - Example: "Cats [Noun] love [Verb] milk [Noun]."
- **Applications**:
  - Grammar checking
  - Sentence structure analysis for translation or summarization

Semantics

- **Definition**: The **study of meaning** in language. It focuses on understanding the meaning of words, phrases, and sentences.
- **Key Techniques**:
  - **Word Sense Disambiguation**: Determining the correct meaning of a word in context.
    - Example: "Bank" (a financial institution) vs. "bank" (riverbank).
  - **Named Entity Recognition (NER)**: Identifying entities like names, locations, and dates.
    - Example: **"Google" → Organization, "Paris" → Location.**
  - **Semantic Parsing**: Mapping natural language to a formal representation of meaning.
- **Applications**:
  - Sentiment analysis
  - Chatbots and virtual assistants
  - Text understanding for search engines

Pragmatics

- **Definition**: The study of language in context. It deals with how meaning changes based on situational and conversational context.
- **Key Techniques**:
  - **Contextual Understanding**: Interpreting sentences **based on prior knowledge** or surrounding text.
    - Example: "Can you pass the salt?" is understood as a polite request, not a question.
  - **Coreference Resolution**: Identifying when different words refer to the same entity.
    - Example: In "**John loves his dog. He takes it for walks**," "He" refers to "John" and "it" to "dog."
  - **Speech Act Analysis**: Classifying utterances based on intent (e.g., request, command, or question).
- **Applications**:
  - Conversational AI
  - Context-aware recommendation systems
  - Dialogue systems in customer support

Contextual Understanding Language is often ambiguous without context. Pragmatic analysis considers the situational, cultural, and conversational context to deduce meaning.

- Example:
  - "It's cold in here."
    - Literal meaning: The temperature is low.
    - Pragmatic interpretation: A request to close the window or turn up the heat.

Speech Acts Introduced by philosopher J.L. Austin, speech acts classify utterances based on their function or intent, such as making a statement, asking a question, or giving a command.

- Example:
  - "It's cold in here."

- **Types of Speech Acts**:
  - **Locutionary**: The literal meaning (e.g., "The door is open").
  - **Illocutionary**: The intended function (e.g., a request to close the door).
  - **Perlocutionary**: The effect on the listener (e.g., someone closes the door).

Implicature Coined by philosopher H.P. Grice, implicature refers to meanings that are implied but not explicitly stated.

- Example:
  - **"Are you coming to the party?"**
    - Response: "I have to work early tomorrow."
    - Implication: The speaker is unlikely to attend the party.

- **Deixis** Words or phrases **(deictic expressions**) whose **meanings depend on the context,** such as "here," "now," "this," or "that.“

- Example:
  - "I’ll meet you there tomorrow."
    - Requires context to understand "I," "you," "there," and "tomorrow."

Presuppositions Assumptions that a speaker considers to be true or known by the listener.

- Example:
  - "Have you stopped smoking?"
    - Presupposes that the listener used to smoke.

- **Politeness and Social Context** Pragmatics considers social norms, politeness strategies, and the relationship between speakers.

- Example:
  - "Could you please pass the salt?" is more polite than "Pass the salt.“

Pragmatic analysis is vital for creating AI systems that understand and respond naturally, making human-computer interactions more intuitive and effective.