Artificial Intelligence & Natural Language Processing

Components of NLP

Prepared by,

P.Ramya

Assistant Professor/CSE-IoT

SNS College of Engineering

Components of NLP

•Natural Language Processing (NLP) refers to AI method of communicating with an intelligent systems using a natural language such as English.

•Processing of Natural Language is required when you want an intelligent system like robot to perform as per your instructions, when you want to hear decision from a dialogue based clinical expert system, etc.

•The field of NLP involves making computers to perform useful tasks with the natural languages humans use. The input and output of an NLP system can be −

　•Speech

　•Written Text

P.Ramya/AI & NLP/19SB601/Components of NLP

Contd…

- There are two components of NLP as given −
- Natural Language Understanding (NLU)
- Natural Language Generation (NLG)

- Natural Language Understanding (NLU)
- Understanding involves the following tasks −
- Mapping the given input in natural language into useful representations.
- Analyzing different aspects of the language.

P.Ramya/AI & NLP/19SB601/Components of NLP

Natural Language Generation (NLG)

Natural Language Generation (NLG)

It is the process of producing meaningful phrases and sentences in the form of natural language from some internal representation.

It involves −

•Text planning − It includes retrieving the relevant content from knowledge base.

•Sentence planning − It includes choosing required words, forming meaningful phrases, setting tone of the sentence.

•Text Realization − It is mapping sentence plan into sentence structure.

P.Ramya/AI & NLP/19SB601/Components of NLP

Contd…

NL has an extremely rich form and structure.

It is very ambiguous. There can be different levels of ambiguity −

•**Lexical ambiguity** − It is at very primitive level such as word-level.

•For example, treating the word "board" as noun or verb?

•**Syntax Level ambiguity** − A sentence can be parsed in different ways.

•For example, "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** − Referring to something using pronouns. For example, Rima went to Gauri. She said, "I am tired." − Exactly who is tired?

•One input can mean different meanings.
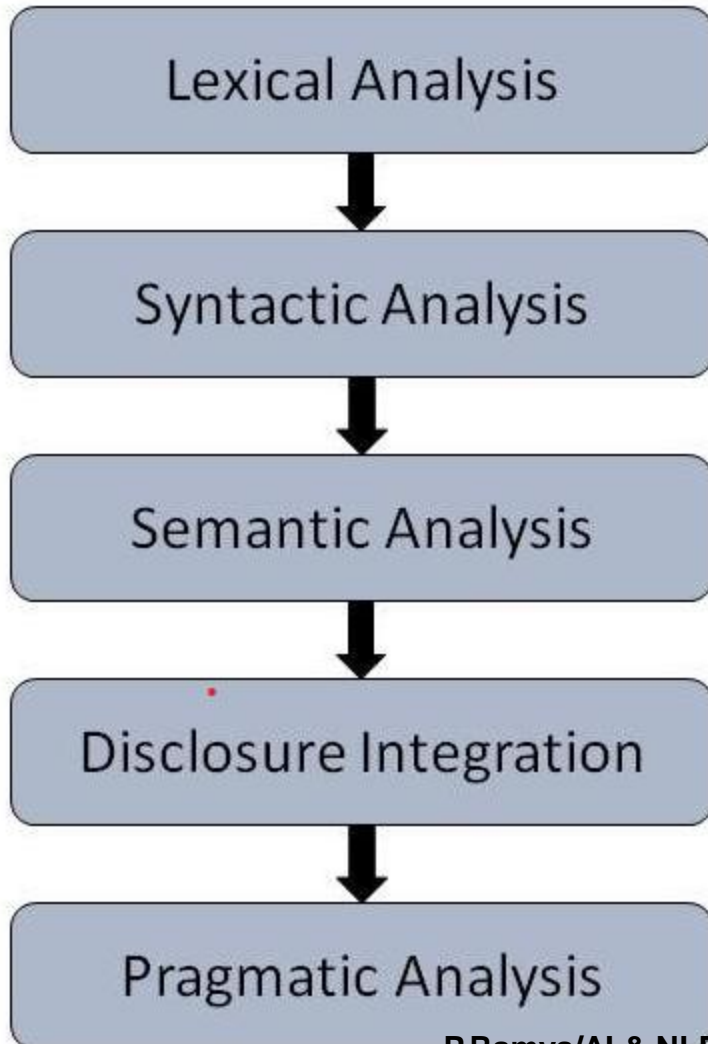
•Many inputs can mean the same thing.

NLP TERMINOLOGIES

•**Phonology** − It is study of organizing sound systematically.

•**Morphology** − It is a study of construction of words from primitive meaningful units.

•**Morpheme** − It is primitive unit of meaning in a language.

•**Syntax** − It refers to arranging words to make a sentence. It also involves determining the structural role of words in the sentence and in phrases.

•**Semantics** − It is concerned with the meaning of words and how to combine words into meaningful phrases and sentences.

•**Pragmatics** − It deals with using and understanding sentences in different situations and how the interpretation of the sentence is affected.

•**Discourse** − It deals with how the immediately preceding sentence can affect the interpretation of the next sentence.

•**World Knowledge** − It includes the general knowledge about the world.

P.Ramya/AI & NLP/19SB601/Components of NLP

NLP TERMINOLOGIES

Lexical Analysis

↓

Syntactic Analysis

↓

Semantic Analysis

↓

Disclosure Integration

↓

Pragmatic Analysis

There are general five steps −

•**Lexical Analysis** − It involves identifying and analyzing the structure of words. Lexicon of a language means the collection of words and phrases in a language. Lexical analysis is dividing the whole chunk of txt into paragraphs, sentences, and words.

•**Syntactic Analysis (Parsing)** − It involves analysis of words in the sentence for grammar and arranging words in a manner that shows the relationship among the words. The sentence such as "The school goes to boy" is rejected by English syntactic analyzer.

P.Ramya/AI & NLP/19SB601/Components of NLP