

Natural Language **Processing**



Introduction

- ☐ **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**
- ☐ **NLP is very important in today's world of web. On the internet vast amount of information available as text, so we would like to process and use this information using NLP**
- ☐ **The field of NLP involves making computers to perform useful tasks with the natural languages humans use.**



Introduction

- ❑ **The input/output of a NLP system can be:**
 - **written text**
 - **speech**
- ❑ **To process written text, we need:**
 - **lexical, syntactic, semantic knowledge about the language**
 - **discourse information, real world knowledge**
- ❑ **To process spoken language, we need everything required to process written text, plus the challenges of speech recognition and speech synthesis.**



Goals of NLP

☐ Science Goal:

- Understand the way language operates. (How language is produced , how is it understood by an intelligent entity)

e.g. How do I process the following sentence?

I went to the bank to withdraw some money.

☐ Engineering Goal:

- Build systems that analyse and generate language; reduce the man-machine gap



Components of NLP

☐ Natural Language Understanding

- Mapping the given input in the natural language into a useful representation.
- Different level of analysis required:
 - morphological analysis,
 - syntactic analysis,
 - semantic analysis,
 - discourse analysis, ...

☐ Natural Language Generation

- Producing output in the natural language from some internal representation.



Components of NLP

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**
- The NLU is harder than NLG**



Why NL Understanding is hard?

- ❑ **Natural language is extremely rich in form and structure, and very ambiguous.**
 - **How to represent meaning,**
 - **Which structures map to which meaning structures.**
- ❑ **One input can mean many different things. Ambiguity can be at different levels.**
 - **Lexical (word level) ambiguity -- different meanings of words**
 - **Syntactic ambiguity -- different ways to parse the sentence**
 - **Interpreting partial information -- how to interpret pronouns**



Why NL Understanding is hard?

- **Contextual information -- context of the sentence may affect the meaning of that sentence.**
- ❑ **Many input can mean the same thing.**



NLP Terminology

☐ Phonology:

- Concerns how words are related to the sounds that realize them

☐ 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.



NLP Terminology

☐ 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.