# Natural Language Processing

Paper Code: CS-821/6
UG 8th Sem

Dr. Samit Biswas, *Assistant Professor*

Department of Computer Sc. and Technology

Indian Institute of Engineering Science and Technology, Shibpur

*Email: samit@cs.iiests.ac.in*

# Plan for Today

- An Introduction To Natural Language Processing
- Course Overview
- Some NLP Application

**What is Natural Language Processing (NLP)?**

- The process of computer analysis of input provided in a human language (natural language), and conversion of this input into a useful form of representation.

**Forms of Natural Language**

- The input/output of a NLP system can be:
  - Written Text
  - Speech
  - NLP is mostly concerned with **written text** (not 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.

## Knowledge of Language

- **Phonetics or Phonology:** study of linguistic sounds.
- **Morphology:** study of meaningful components of words.
- **Syntax:** study of structural relationships between words.
- **Semantics:**
  - study of meaning/ context-independent meaning.
- **Pragmatics:**
  - study of how language is used to accomplish goals and intentions of the speaker.
- **Discourse:**
  - study of linguistic units larger than a single utterance.

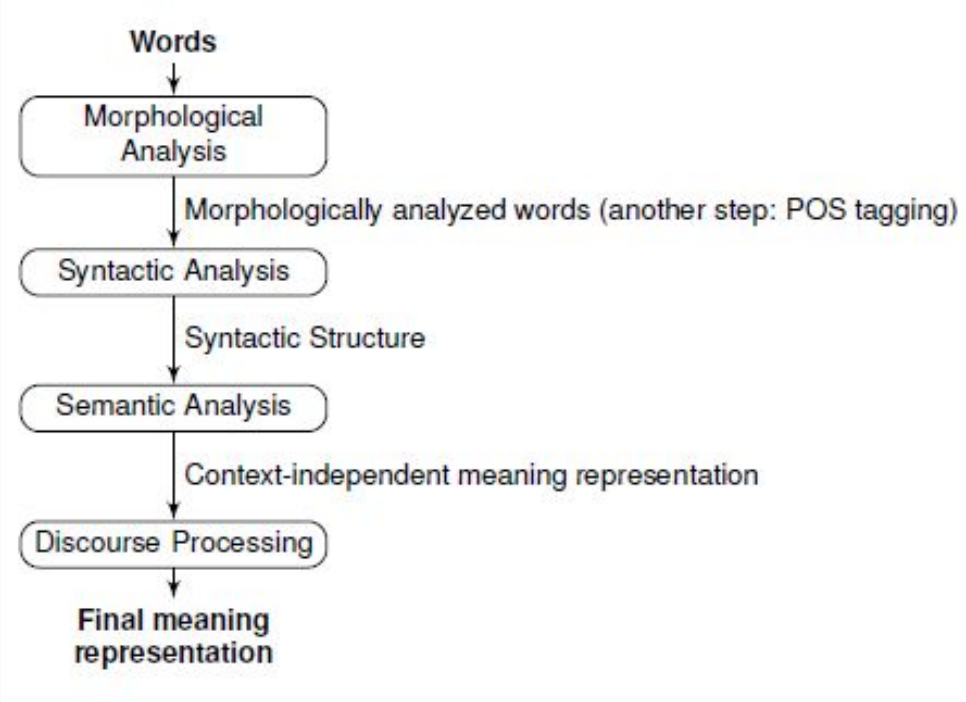
# Models / Algorithm

- State machines
- Rule Systems
- Logic
- Probabilistic Models
- Vector Space Models

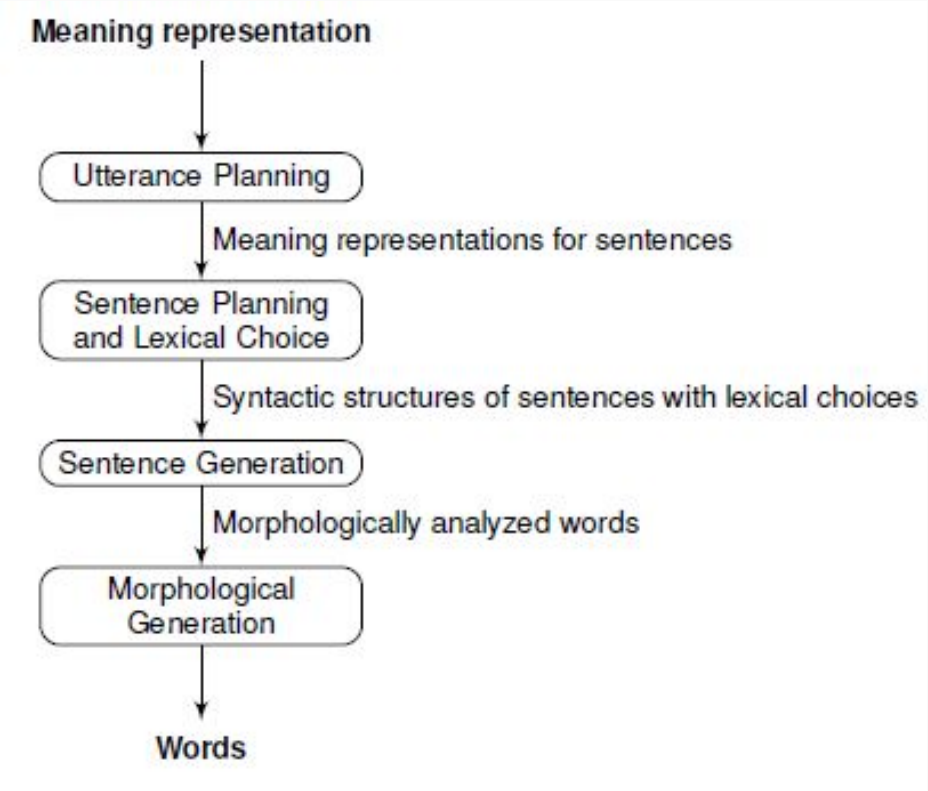
## Components of NLP

- **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.
  - Different level of synthesis required.

# Natural language Understanding



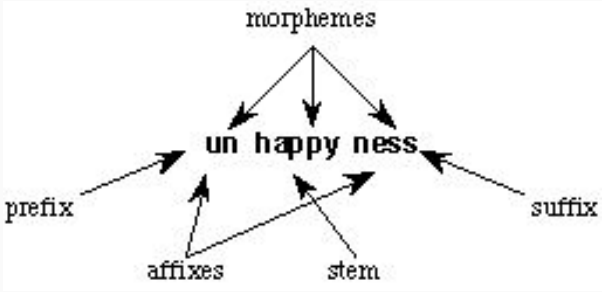
# Natural Language Generation



# Morphological Analysis

**Morphological Analysis**

- Morphology is a subdiscipline of linguistics that studies word structure. Analyzing words into their linguistic components (morphemes).
- **“minimal unit of meaning” - “the minimal unit of grammatical analysis”.**
- **Consider a word like: “unhappines”:**



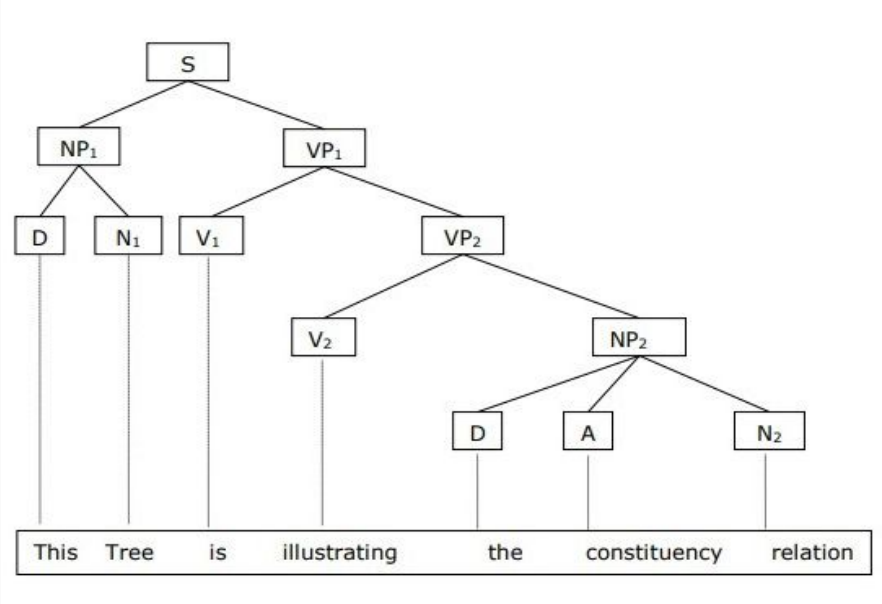

- **There are three Morphemes:**
  **un** means “not”
  **ness** means “being in a state or condition”
  **Happy** is a free morpheme

# Syntactic Processing / Analysis (Parsing)

- recognising higher level units of structure that allow us to compress our description of a sentence
- Goal of syntactic analysis (parsing):
  - Detect if a sentence is correct
  - Provide a syntactic structure of a sentence
- Parsing is the task of uncovering the syntactic structure of language and is often viewed as an important prerequisite for building systems capable of understanding language.
- Syntactic structure is necessary as a first step towards semantic interpretation

# Example: Syntactic Processing / Analysis (Parsing)

- The basic clause structure is understood in terms of Noun Phrase (NP) and Verb Phrase (VP).
- We can write the sentence **“This tree is illustrating the constituency relation”** as follows -

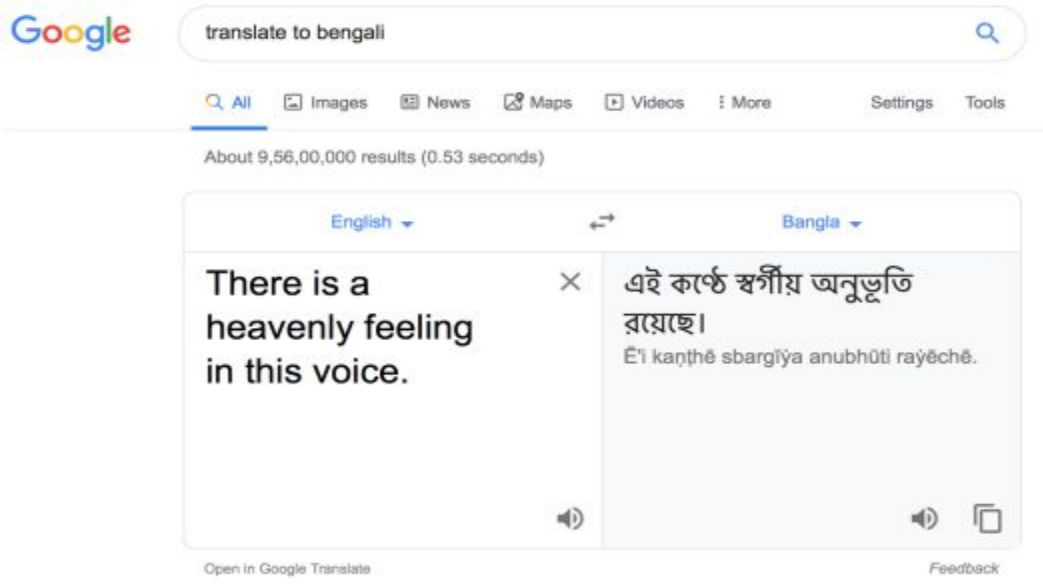


# Semantic Analysis

- The work of semantic analyzer is to check the text for meaningfulness.
- Semantics and its understanding as a study of meaning covers most complex tasks like:
  - finding synonyms,
  - word sense disambiguation,
  - constructing question-answering systems,
  - translating from one NL to another,
  - populating base of knowledge.

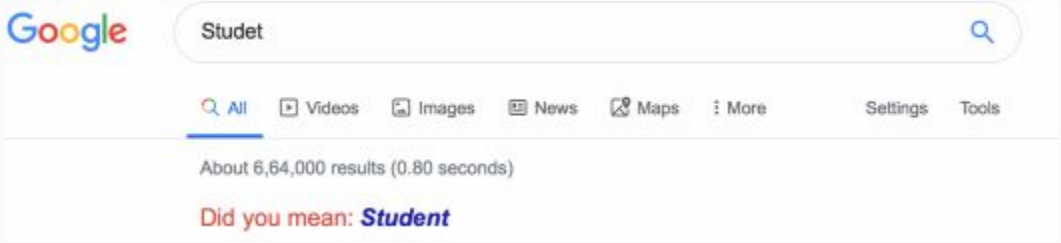
# Some NLP Applications

- **Machine Translation:** Translation between two natural languages.
  **Example:** Google Translator



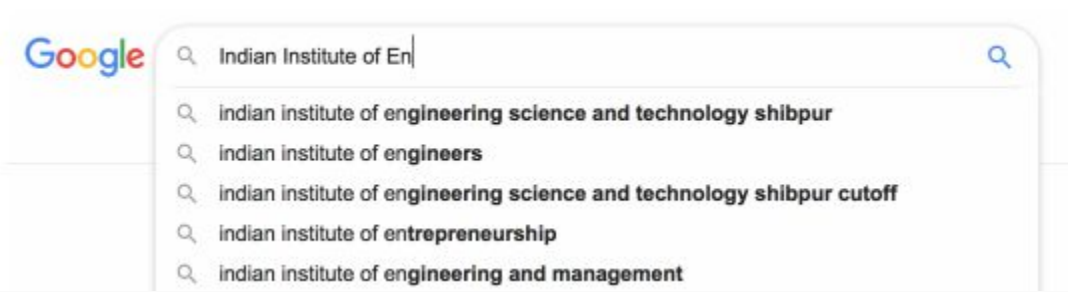
# Some NLP Applications ...

- **Information Retrieval:** Web search (uni-lingual or multi-lingual).
- **Query Answering/Dialogue:** Natural language interface with a database system, or a dialogue system.
- **Report Generation:** Generation of reports such as weather reports.
- **Some Small Applications:** Grammar Checking, Spell Checking, Spell Corrector.



# Some NLP Applications ...

- Automatic query completion



# Some NLP Applications ...

- **Information Extraction**
- **Some other applications:** Email filters, Digital phone calls, Data analysis, Text analytics.

# References

**References**


- Daniel Jurafsky & James H. Martin, “*Speech and Language Processing*”, Prentice Hall
- Diana Maynard, Kalina Bontcheva, Isabelle Augenstein, “*Natural Language Processing for the Semantic Web*”, A Publication in the Morgan & Claypool Publishers series.
- Steven Bird, Ewan Klein and Edward Loper, “*Natural Language Processing with Python*”, By O‘Reilly.
- Christopher Manning and Hinrich Schtze, “*Foundations of Statistical Natural Language Processing*”, The MIT Press.

# Thank You


Contacts:

Dr. Samit Biswas
Department of CST,
IIEST, Shibpur

samit@cs.iiests.ac.in
https://www.iiests.ac.in/IIEST/Faculty/cs-samit