



Veermata Jijabai Technological Institute, Mumbai 400019

Experiment No.: 02

Aim: Accept the Sentence as input and display Part of Speech Tags for the same. Also Perform dependency parsing and list out all dependency relations along with their operands/entities.

Name: Kiran K Patil

Enrolment No.: 211070904

Branch: Computer Engineering

Batch: D

Theory:

Part-of-speech tagging (POS tagging) is the process of labeling each word in a text with its grammatical category, such as noun, verb, adjective, or adverb. This technique is essential for understanding the grammatical structure of sentences and is widely used in various natural language processing applications.

1. Accepting a Sentence as Input:

- This initial step involves receiving a text or sentence as input. This input could come from a user, a file, or any other source.
- In NLP applications, accepting input text is the starting point for various language processing tasks.

2. Part of Speech (POS) Tags:

- Part of speech tagging is a linguistic analysis that assigns grammatical categories to each word in a sentence.
- Common POS tags include nouns, verbs, adjectives, adverbs, pronouns, conjunctions, and more.
- These tags provide information about the role of each word in the sentence's structure, aiding in tasks like text analysis and information retrieval.

3. Dependency Parsing:

- Dependency parsing is a more advanced NLP task that analyzes the grammatical structure and relationships between words in a sentence.
- It creates a tree-like structure that represents the dependencies and hierarchical structure of the sentence.
- The resulting parse tree reveals how words in the sentence are connected and which words depend on others.

4. Dependency Relations:

- Dependency parsing identifies specific relations between words in the sentence. For instance, it reveals which word is the subject of a verb, or which word an adjective modifies.
- These relations help in understanding the syntactic structure and meaning of the sentence.
- The relations are often represented as labeled arcs or edges in the parse tree.

5. Operands/Entities:

- In the context of dependency parsing, operands or entities are the words or phrases involved in the dependency relations.
- For example, in the sentence "The cat chases the mouse," "cat" is the subject and "mouse" is the object of the verb "chases."
- Understanding these entities is crucial for extracting structured information from text and for various NLP tasks like information retrieval.

In summary, this experiment begins by taking a sentence as input and then performs two important NLP tasks: Part of Speech tagging, which categorizes words by their grammatical roles, and Dependency Parsing, which unveils the hierarchical relationships between words and their specific dependencies. The output provides insights into the structural and grammatical aspects of the input sentence, making it useful for a wide range of NLP applications, including information extraction, text analysis, and language understanding.

Implementation:

```
# !pip install nltk
import nltk
nltk.download('punkt')
nltk.download('averaged_perceptron_tagger')
nltk.download('maxent_ne_chunker')
nltk.download('words')
nltk.download('dependency_treebank')

from nltk.tokenize import word_tokenize, sent_tokenize

text = "The quick brown fox jumping over the lazy dog"

sentences = sent_tokenize(text)

words = [word_tokenize(sentence) for sentence in sentences]

from nltk import pos_tag

pos_tags = [pos_tag(sentence) for sentence in words]

# Print the PoS tags for each word
for sentence_pos_tags in pos_tags:
    print(sentence_pos_tags)
```

```
# Dependency Parsing

import spacy
from spacy import displacy

nlp = spacy.load("en_core_web_sm")
sentence = "The quick brown fox is jumping over the lazy dog"
doc = nlp(sentence)

for token in doc:
    print("Head: {} \t Dependency: {} \t Text: {} \n".format(str(token.head.text), str(token.dep_), str(token.text)))
```

Sample sentence:

"The quick brown fox jumping over the lazy dog"

Output :

1. POS tagging :

```
[('The', 'DT'), ('quick', 'JJ'), ('brown', 'NN'), ('fox', 'NN'), ('jumping', 'NN'), ('over', 'IN'), ('the', 'DT'), ('lazy', 'JJ'), ('dog', 'NN')]
```

2. Dependency parsing

Head: fox	Dependency: det	Text: The
Head: fox	Dependency: amod	Text: quick
Head: fox	Dependency: amod	Text: brown
Head: jumping	Dependency: nsubj	Text: fox
Head: jumping	Dependency: aux	Text: is
Head: jumping	Dependency: ROOT	Text: jumping
Head: jumping	Dependency: prep	Text: over
Head: dog	Dependency: det	Text: the
Head: dog	Dependency: amod	Text: lazy
Head: over	Dependency: pobj	Text: dog

Conclusion:

In conclusion, the experiment introduced the concept of part-of-speech tagging (POS tagging), a fundamental task in natural language processing (NLP). POS tagging involves labeling each word in a text with its grammatical category, providing crucial insights into the sentence's structure and meaning. The ability to identify the part of speech of each word is invaluable for a wide range of NLP applications, from information retrieval and machine translation to sentiment analysis and text analysis. It plays a foundational role in understanding and processing human language, making it a key component in the field of NLP.