LAB-I: WRITE A PROGRAM TO DRAW PARSE TREE OF A SENTENCE INPUT BY THE USER.

OBJECTIVE: To implement/draw parse tree in python

THEORY:

With the implementation of our tree data structure complete, we now look at an example of how a tree can be used to solve some real problems. In this section we will look at parse trees. Parse trees can be used to represent real-world constructions like sentences or mathematical expressions.

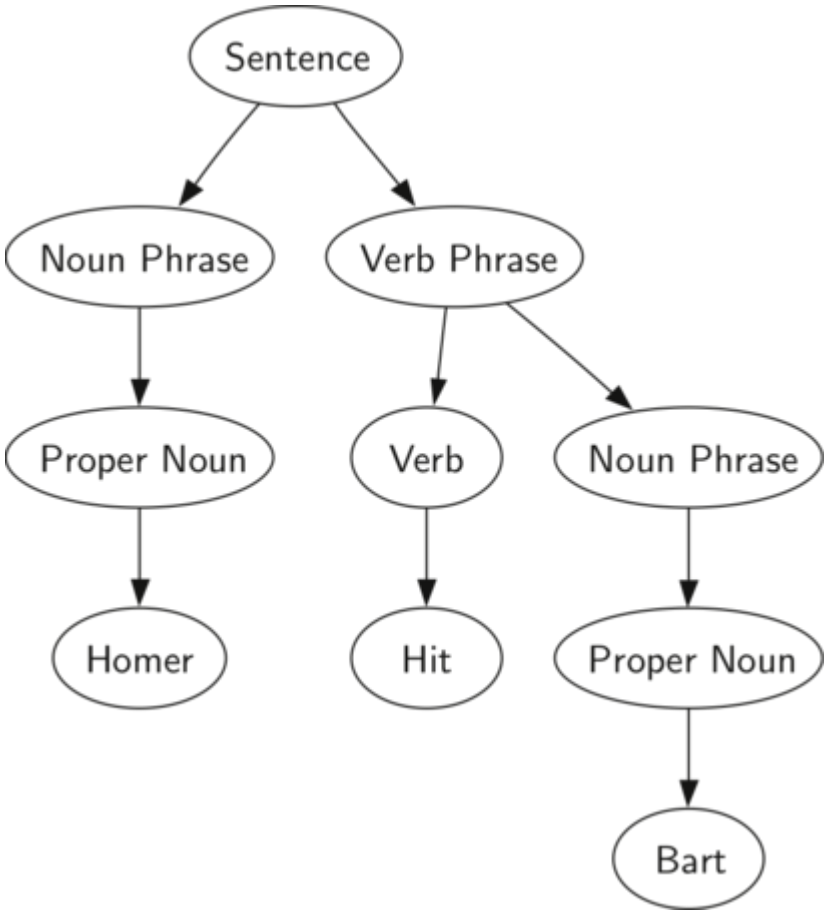


Figure 1: A Parse Tree for a Simple Sentence

Figure 1: shows the hierarchical structure of a simple sentence. Representing a sentence as a tree structure allows us to work with the individual parts of the sentence by using subtrees.

What is parsing in natural language?

Parsing, syntax analysis, or syntactic analysis is the process of analyzing a string of symbols, either in natural language, computer languages or data structures, conforming to the rules of a formal grammar. The term parsing comes from Latin pars (orationis), meaning part (of speech).

How can we represent a language in term of parse tree?

Parse Tree:

1. Parse tree is the hierarchical representation of terminals or non-terminals.
2. These symbols (terminals or non-terminals) represent the derivation of the grammar to yield input strings.

3. In parsing, the string springs using the beginning symbol.

Is natural language structured data?

Answer: False Natural language is not data that is structured. Our language is not in binary form and we do not speak in a structured manner.

What are the different types of parsers in natural language processing?
Natural Language processing provides us with two basic parsing techniques viz; Top-Down and Bottom- Up. Their name describes the direction in which the parsing process advances. We have Basic-Top-Down parsing which is the fusion of top-down and bottom-up parsing.

What is the example of Natural Language Processing?

Arguably the best-known example of NLP, smart assistants such as Siri, Alexa and Cortana have become increasingly integrated into our lives. Using NLP, they break language down into parts of speech, word stems and other linguistic features.

How is natural language processed in AI?

Natural Language Processing (NLP) is a branch of Artificial Intelligence (AI) that enables machines to understand the human language. Its goal is to build systems that can make sense of text and automatically perform tasks like translation, spell check, or topic classification.

Which symbol is used as the root of the parse tree?

The starting symbol of the grammar must be used as the root of the Parse Tree. Leaves of parse tree represent terminals. Each interior node represents productions of grammar.

Which parsing technique generates a parse tree by starting at the leaves and proceeding towards the root?

top-down parsing technique
A top-down parsing technique parses the input by starts constructing a parse tree from the root node moving down to the leaf nodes. A top-down parser uses the leftmost derivation to generate the string.

How does parsing work in natural languages?

Natural languages follow certain rules of grammar. This helps the parser extract the structure. Formally, we can define parsing as, the process of determining whether a string of tokens can be generated by a grammar.

What is a parse tree in Python?

Parse tree is the hierarchical representation of terminals or non-terminals. These symbols (terminals or non-terminals) represent the derivation of the grammar to yield input strings. In parsing, the string springs using the beginning symbol.

ALGORITHH:

Step 1: Sentence Segmentation.

Step 2: Word Tokenization.

Step 3: Predicting Parts of Speech for Each Token.

Step 4: Text Lemmatization.

Step 5: Identifying Stop Words.

Step 6: Dependency Parsing.

Step 6b: Finding Noun Phrases.

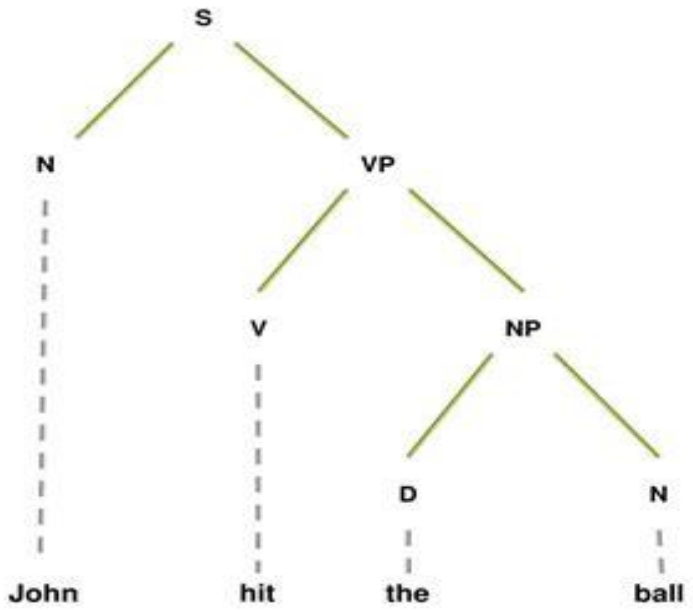
Step 7: Named Entity Recognition (NER)

PROGRAM:

```
# Import required libraries
import nltk
nltk.download('punkt')
nltk.download('averaged_perceptron_tagger')
from nltk import pos_tag, word_tokenize, RegexpParser


# Example text
sample_text = "John hit the ball"


# Find all parts of speech in above sentence
tagged = pos_tag(word_tokenize(sample_text))


#Extract all parts of speech from any text
chunker = RegexpParser("""
                        NP: {<DT>?<JJ>*<NN>} #To extract Noun Phrases
                        P: {<IN>} #To extract Prepositions
                        V: {<V.*>} #To extract Verbs
                        PP: {<p> <NP>} #To extract Prepositional Phrases
                        VP: {<V> <NP|PP>*} #To extract Verb Phrases
                """)


# Print all parts of speech in above sentence
```

```
output = chunker.parse(tagged)
print("After Extracting\n", output)


# To draw the parse tree
output.draw()
```

OUTPUT:



CONCLUSION:

Parse tree is the hierarchical representation of terminals or non-terminals. These symbols (terminals or non-terminals) represent the derivation of the grammar to yield input strings. In parsing, the string springs using the beginning symbol.