PRACTICAL NO 5

**AIM:** Write a program to create parse tree for given text in given paragraph.

**Theory:**

Natural Language Processing (NLP) is a field of study that deals with understanding, interpreting, and manipulating human spoken languages using computers.

Since most of the significant information is written down in natural languages such as English, French, German, etc. thus, NLP helps computers communicate with humans in their own languages and perform other language-related tasks.

In conclusion, NLP makes it possible for computers to read the text, hear speech, interpret and realize it, understand the sentiment, and identify important parts of a text or speech.

What is Syntax?

A natural language typically follows a hierarchical structure, and contains the following components:

Sentences

Clauses

Phrases

Words

Syntax refers to the set of rules, principles, processes that govern the structure of sentences in a natural language. One basic description of syntax is how different words such as Subject, Verbs, Nouns, Noun Phrases, etc. are sequenced in a sentence.

Some of the syntactic categories of a natural language are as follows:

Sentence(S)

Noun Phrase(NP)

Determiner(Det)

Verb Phrase(VP)

Prepositional Phrase(PP)

Verb(V)

Noun(N)

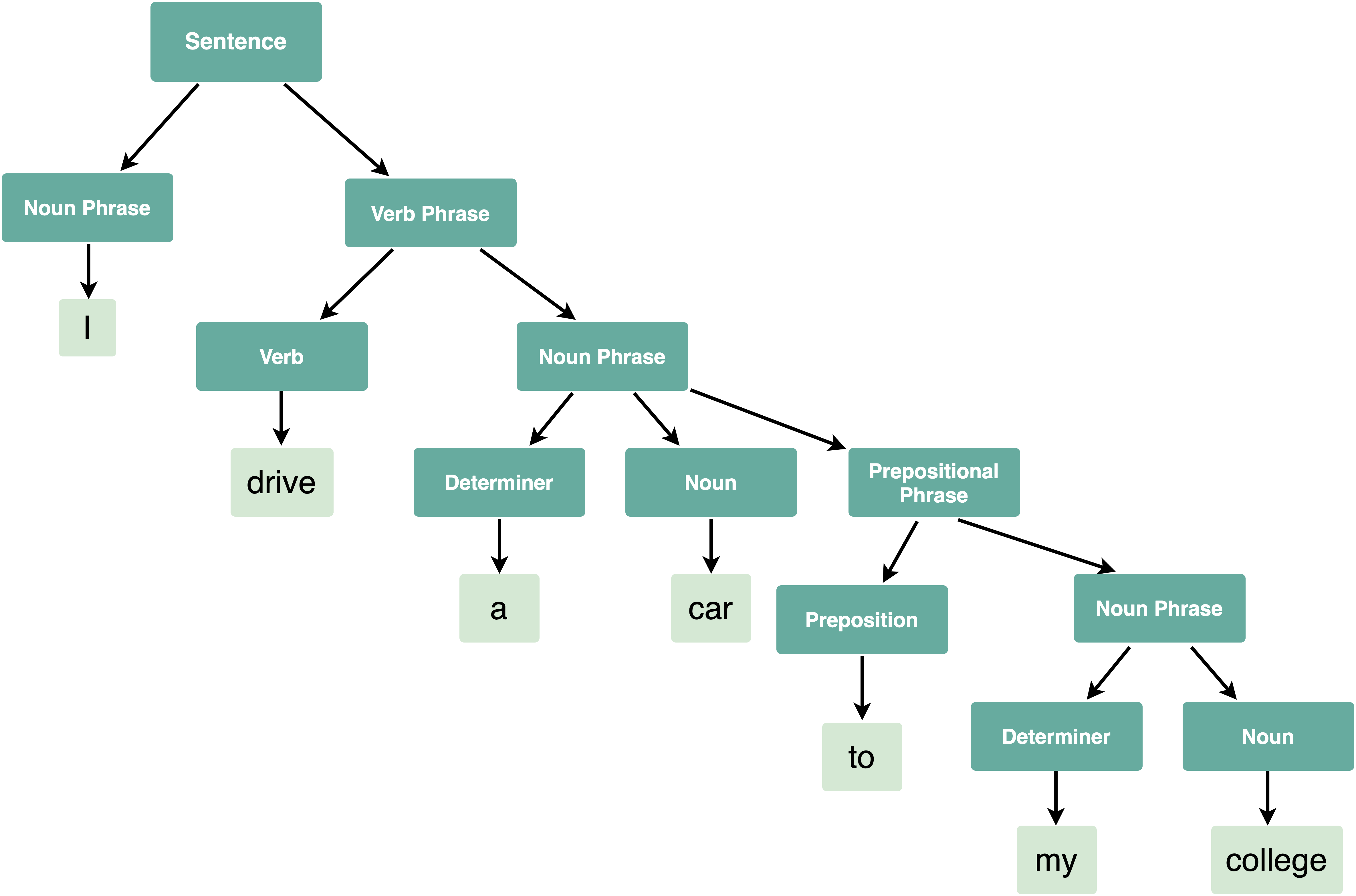
Syntax Tree:

A Syntax tree or a parse tree is a tree representation of different syntactic categories of a sentence. It helps us to understand the syntactical structure of a sentence.

Example:

The syntax tree for the sentence given below is as follows:

I drive a car to my college.

****

**CODE:**

import nltk

from nltk.tree import \*

from nltk.draw import tree

# Grammar creation

grammar = nltk.PCFG.fromstring("""

S -> NP VP [1.0]

PP -> P NP [1.0]

NP -> Det N [0.4]

NP -> Det N PP [0.2]

NP -> 'I' [0.4]

VP -> V NP [0.5]

VP -> VP PP [0.5]

Det -> 'an' [0.5]

Det -> 'my' [0.5]

N -> 'elephant' [0.5]

N -> 'pajamas' [0.5]

V -> 'saw' [1.0]

P -> 'in' [1.0]

""")

# Import example sentences to NLTK and tokenize them

str\_sentence1 = "I saw an elephant"

str\_sentence2 = "I saw an elephant in my pajamas"

print("Example sentences")

print(str\_sentence1)

print(str\_sentence2)

tokens1 = nltk.word\_tokenize(str\_sentence1)

tokens2 = nltk.word\_tokenize(str\_sentence2)

# Create the Chart and Viterbi parsers, with the input grammar

chart\_parser = nltk.ChartParser(grammar)

viterbi\_parser = nltk.ViterbiParser(grammar)

Results for the Chart Parser

print("Parse trees obtained with the Chart parser")

print("Sentence 1")

for tree in chart\_parser.parse(tokens1):

print(tree)

tree.draw()

print("Sentence 2")

for tree in chart\_parser.parse(tokens2):

print(tree)

tree.draw()

Results for the Viterbi Parser

print("Parse trees obtained with the Viterbi parser")

print("Sentence 1")

for tree in viterbi\_parser.parse(tokens1):

print(tree)

tree.draw()

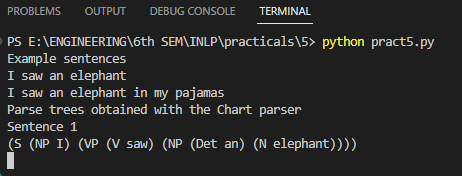
print("Sentence 2")

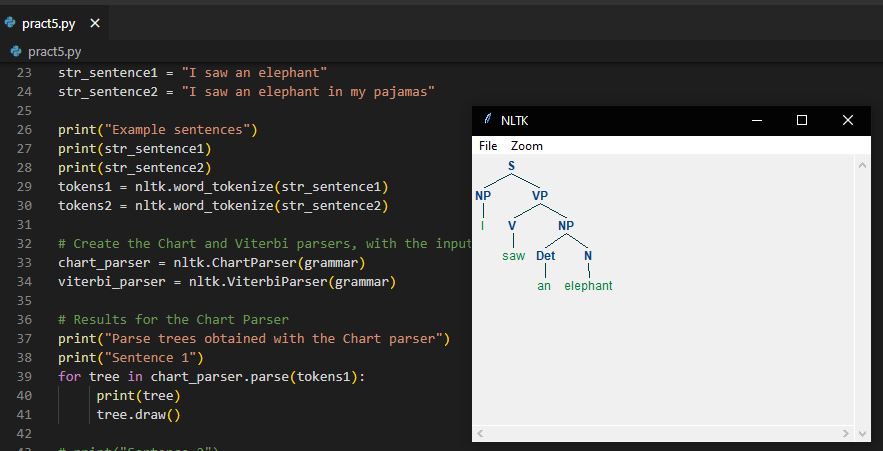
for tree in viterbi\_parser.parse(tokens2):

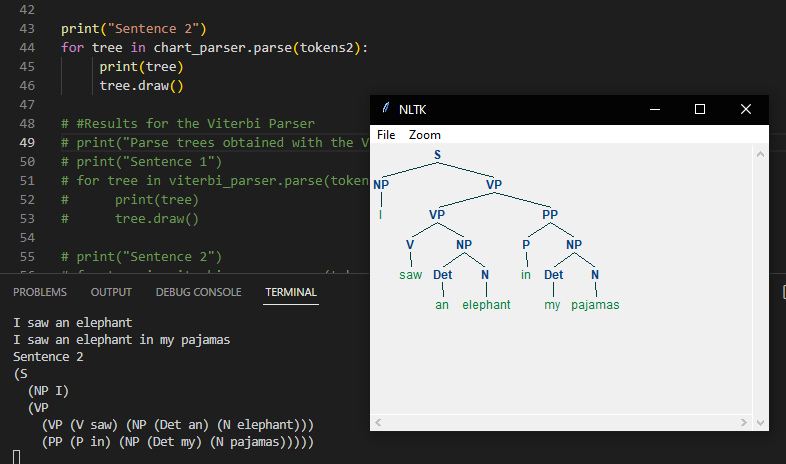
print(tree)

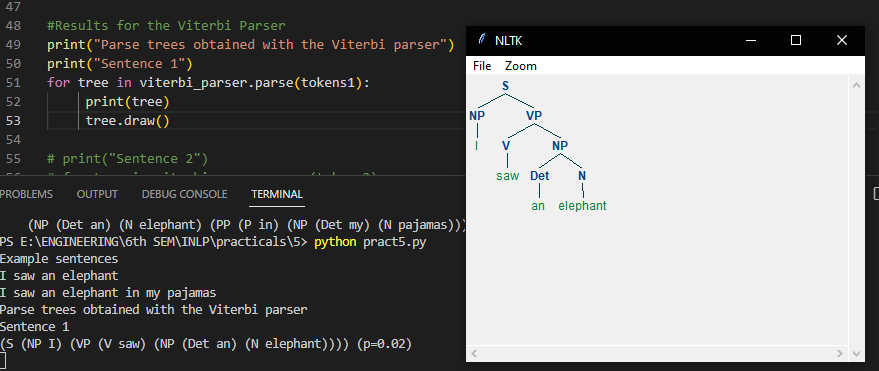
tree.draw()

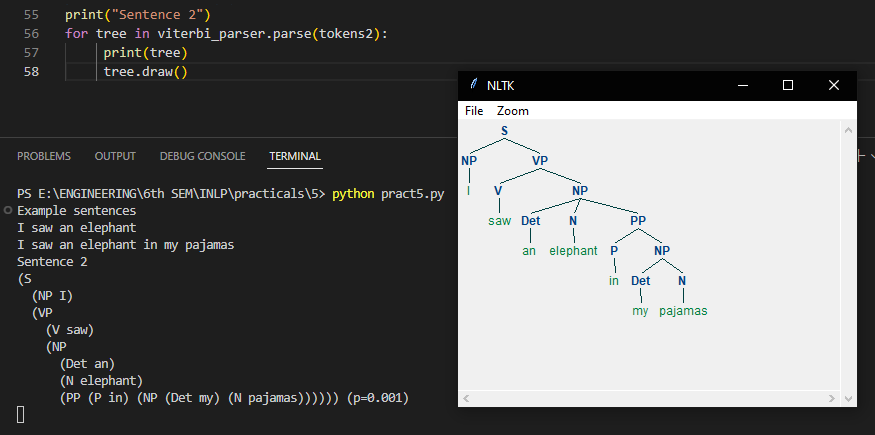
**OUTPUT:**











**CONCLUSION:** *Hence we successfully completed the practical on the parse tree for given text in given paragraph.*