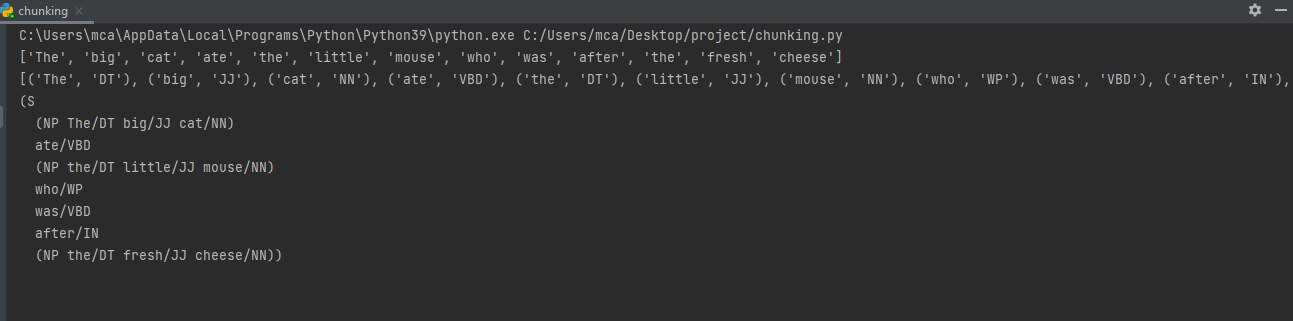
Date:23-02-2022

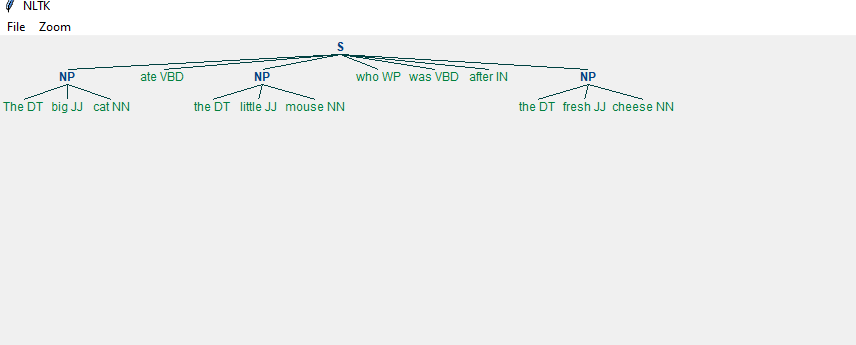
Program No:22

Aim: write a python program for natural language processing which performs chunking

import nltk  
new="The big cat ate the little mouse who was after the fresh cheese"  
new\_tokens=nltk.word\_tokenize(new)  
print(new\_tokens)  
new\_tag=nltk.pos\_tag(new\_tokens)  
print(new\_tag)  
grammer="NP:{<DT>?<JJ><NN>}"  
chunkParser=nltk.RegexpParser(grammer)  
chunked=chunkParser.parse(new\_tag)  
print(chunked)  
chunked.draw()

OUTPUT





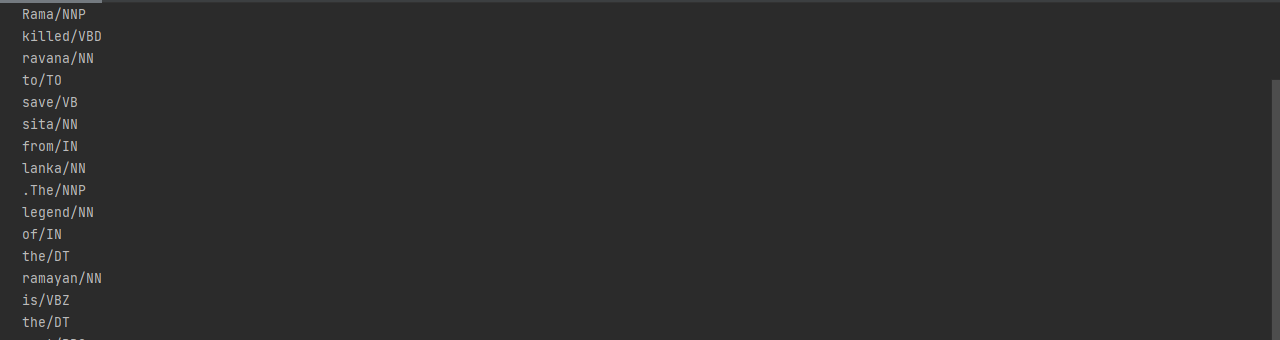
Date:23-02-2022

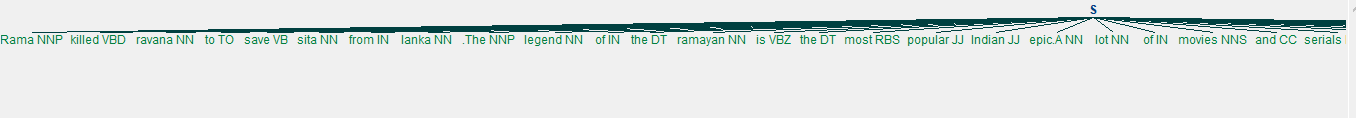
Program No:23

Aim: write a python program for natural language processing which performs chunking

import nltk  
nltk.download('averaged\_perceptron\_tagger')  
sample\_text=""" Rama killed ravana to save sita from lanka .The legend of the ramayan is the most popular  
Indian epic.A lot of movies and serials have already been shot in several languages here in india based on the Ramayana  
"""  
tokenized=nltk.sent\_tokenize(sample\_text)  
for i in tokenized:  
 words=nltk.word\_tokenize(i)  
 tagged\_words=nltk.pos\_tag(words)  
 chunkGram=r"""VBnb: {}"""  
 chunParser=nltk.RegexpParser(chunkGram)  
 chunked=chunParser.parse(tagged\_words)  
 print(chunked)  
 chunked.draw()

OUTPUT





Date:23-02-2022

Program No:24

Aim: write a python program for natural language processing which performs scraping

import scrapy  
class Quotesspider(scrapy.Spider):  
 name='quotes'  
 def start\_requests(self):  
 urls=['http://quotes.toscrap.com/page/1/'  
 'http://quotes.toscrap.com/page/2/']  
 def parse(self, response):  
 page=response.url.split("/")[-2]  
 filename='quotes-%s.html'% page  
 with open(filename,'wb') as f:  
 f.write(response.body)  
 self.log('saved files '% filename)