

**M.Sc DEGREE THIRD SEMESTER EXAMINATION IN COMPUTER SCIENCE  
(SOFT COMPUTING), JANUARY 2021  
19-323-0325 NATURAL LANGUAGE PROCESSING**

**Time: 3 Hours**

**Maximum Marks: 50**

(Answer ANY FIVE questions)  
(Each Question carries 10 Marks)  
(Regular)

- I**

**a** Do you consider Python a more suitable language for NLP tasks than C ? Justify your answer  
(5Marks)

**b** Suppose you are hired by an airline company to create an NLP based passenger query answering system. Explain the role and type of corpus to be used in your work ,and how you would go about the creation of such a system.  
(5Marks)
- II**

**a** Discuss with examples the different categories of data attributes used in NLP.  
(6Marks)

**b** Consider the sentence” He plays cricket “. Demonstrate how you would create CFG and derive the above sentence.  
(4Marks)
- III**

Discuss in detail about different types of ambiguities and strategies to resolve them  
(10Marks)
- IV**

**a** Outline the differences between stemming and lemmatisation including code in python.  
(6Marks)

**b** In each of the following systems what type of preprocessing will be needed.  
(i) Grammar correction system.  
(ii) Sentiment analysis.  
(4Marks)
- V**

**a** Distinguish between bottom up and top down parsers.  
(4Marks)

**b** Demonstrate how the probability of a tree is calculated in probabilistic context free grammar.  
(6Marks)
- VI**

**a** Explain BOW model with an example  
(5Marks)

**b** Write a note on word2Vec model.  
(5Marks)
- VII**

With a neat diagram explain the architecture and working of a rule based system for NLP  
(10Marks)

[illegible]