

**B.TECH. SEM -VII INFO. TECH. 2014 COURSE (CBCS) : WINTER -  
2017**

**SUBJECT: ELECTIVE – III: NATURAL LANGUAGE PROCESSING**

Day: **Friday**  
Date: **19/01/2018**

Time: **02.30 PM TO 05.30 PM**  
Max. Marks: **60**

**W-2017-2306**

**N.B.:**

- 1) All questions are **COMPULSORY**.
- 2) Figures to the right indicate **FULL** marks.
- 3) Assume suitable data, if necessary.
- 4) Draw neat and labeled diagram **WHEREVER** necessary.

---

**Q.1** Write a note on Natural language processing (NLP) application. **[10]**

**OR**

**Q.1** Explain bigram language model also explain equation for bigram probability estimation. **[10]**

**Q.2** Write Finite State Automata for time of day expressional be eleven o' clock, twelve thirty, midnight, or a quarter to ten and others. **[10]**

**OR**

**Q.2** Use Penn Treebank tag set to tag each word in the following (ignore punctuation) **[10]**

- a) It is nice night.
- b) This crap game is over a garage in fifty-second street.
- c) .... Nobody ever takes the paper she sells.....

**Q.3** Explain stochastic part of speech tagging using Viterbi algorithm with an example. **[10]**

**OR**

**Q.3** What is inflectional and derivational morphology? Give examples. **[10]**

**Q.4** Define context free grammar. Consider following context free grammar: **[10]**

$S \rightarrow SS + | SS^* / a.$

And the string  $aa + a^*$

- i) Give left most derivation
- ii) Give right most derivation
- iii) Give parse tree for string.

**OR**

**Q.4** Write a note on probabilistic parsing. **[10]**

**Q.5** What are different ways to improve user queries in information extraction? **[10]**

**OR**

**Q.5** What is vector space model? Explain stages of it in detail. **[10]**

**Q.6** How is linguistic data managed with the help of GATE? **[10]**

**OR**

**Q.6** Write note on semantic web technologies. **[10]**