

Time: 3 Hours

Max. Marks: 80

- **Question No.1 is compulsory.**
- Solve **ANY THREE** questions from the **remaining** five questions.
- Assume suitable data wherever required, but justify the same.

1. Solve **ANY FOUR** questions from the following: (20)
 - a. List out five points that explains why natural language processing is hard.
 - b. With respect to Porter Stemmer, compute the number of vowels in the following words: brachydactyly, employability, psychology, hieroglyphology, crystallography
 - c. Explain with examples the following terms: Anaphora, Cataphora, Inference, One-anaphora, Demonstrative
 - d. Compute the probability of the occurrence of the word “that” given the previous word is “point”:

The point that I am going to point from my point of view will be the point that you will be pointing from your point of view once you understand my point. So there is no point in pointing out that point.

- e. Explain the process of tokenization as a preprocessing stage in NLP
2. (a) Explain morphological analysis with suitable examples. (10)
 - (b) Compare and contrast open class and closed class words in English language with suitable examples. (10)
3. (a) Explain measure of a word as described in Porter Stemmer. Use it to compute the measure of the following words: Supercalifragilisticexpialidocious, Pseudopseudohypoparathyroidism, Dermatoglyphics, Uncopyrightable (10)
 - (b) Illustrate parts of speech tagging technique using Hidden Markov Model with a suitable example. (10)
4. (a) Explain reference phenomenon with suitable examples. (10)
 - (b) Identify the referents and the referring expressions in the following discourse: (10)
In Roach’s observation Jim did not stop at the school buildings but continued across the sweep to the stable yard. He knew the layout of the place already. Roach decided later that he must have made a reconnaissance or studied maps. Even when Mr. Theodore waved, and Ms. Castle shrieked, he didn't stopped.
5. (a) Explain sense of a word. Illustrate word sense disambiguation task with suitable examples. (10)
 - (b) Explain term-frequency and inverse document frequency (TF-IDF) with an example (10)
6. Write short notes on **ANY FOUR** of the following: (20)
 - a. Sentiment Analysis
 - b. Named Entity Recognition
 - c. Question Answer System
 - d. Machine Translation
 - e. Text Summarization