Comp/42175/Sem VII/30/11/24

Duration: 3hrs **Max Marks:80**

N.B. : (1) Question No 1 is Compulsory.
(2) Attempt any three questions out of the remaining five.
(3) All questions carry equal marks.
(4) Assume suitable data, if required and state it clearly.

1 Attempt any FOUR [20]
a What is word sense disambiguation?
b Explain reference resolution in detail
c Explain rule-based machine translation systems
d What is hybrid POS tagging?
e Differentiate between Syntactic ambiguity and Lexical Ambiguity

2 a Design FST for regular and plural nouns. [10]
b Explain the preprocessing operations in natural language processing [10]

3 a Consider the following corpus [10]
<s> a/DT dog/NN chases/V a/DT cat/NN </s>
<s> the/DT dog/NN barks/V loudly/RB </s>
<s> a/DT cat/NN runs/V fast/RB </s>
Compute the emission and transition probabilities for a bigram HMM.
Also, decode the following sentence using the Viterbi algorithm.
The cat chases the dog.
b Compare and contrast Hobbs' Algorithm and Centering Theory. [10]

4 a Explain how the supervised learning approach can be applied for word sense disambiguation [10]
b Explain the N-gram language model and its application. [10]

5 a Explain the Porter Stemming algorithm in detail. [10]
b Construct a parse tree for the following sentence using the given CFG rules: [10]
The tall girl sings.
Rules: S → NP VP
NP → Det Adj N | Det N
VP → V | V NP
Det → "the"
Adj → "tall"
N → "girl"
V → "sings"

6 a Explain text summarization in detail [10]
b Explain how Maximum Entropy is used for sequence labeling. [10]

****