7. (a) What is the need of POS (Part-of Speech) Tagging in NLP?

   (b) How can you estimate N-Gram conditional probability?

   (c) Draw the parse tree of the sentence : "John learned the lessons. "
   Consider the rules of the sentence as follows for Top down Parsing :
   S-> N VP
   NP ->DET N | NP PP
   VP->V NP

4 + 4 + 4 = 12

Group – E

8. (a) Design a systemic grammer to build a multiple layer analysis of the following sentence.
   (i) The document will be saved by the system
   (ii) Will the document be saved by the system?

   (b) What is the need of sentiment analysis? Briefly give the steps of sentiment analysis.

7 + 5 = 12

9. (a) (i) Random sequence of letters with probabilities are
   P=1/8, t=1/4, k=1/8, a=1/4, i=1/8, u=1/8
   Calculate the entropy of each letter.
   (ii) Probability distribution of a random variable X and two approximate distributions m1 and m2 are as follows:

X	x1	x2	x3	x4
p	0.3	0.2	0.2	0.3
m1	0.2	0.4	0.1	0.3
m2	0.3	0.5	0.1	0.1

   Calculate the entropy of X and cross-entropy of m1 and m2.

   (b) (i) What is the need of discourse analysis in natural language processing?
   (ii) How to evaluate trigram system under Information Theory? Explain with example.

(2 + 4) + (2 + 4) = 12

NATURAL LANGUAGE PROCESSING
(CSEN 4161)

Time Allotted : 3 hrs **Full Marks : 70**

Figures out of the right margin indicate full marks.

Candidates are required to answer Group A and any 5 (five) from Group B to E, taking at least one from each group.

Candidates are required to give answer in their own words as far as practicable.

Group – A
(Multiple Choice Type Questions)

1. Choose the correct alternative for the following: **10 × 1 = 10**

   (i) Methods for statistical NLP is/are
   (a) Viterbi decoding (b) N-gram language modeling
   (c) Statistical indexing (d) both (a) and (b).

   (ii) The grammar that consist rules with a single symbol on the left hand side of the rewrite rules is
   (a) CFG (b) CSG
   (c) Naive Bayes (d) HMM.

   (iii) One of the main challenge/s of NLP Is
   (a) Handling Ambiguity of Sentences (b) Handling Tokenization
   (c) Handling POS-Tagging (d) (a), (b) and (c).

   (iv) In linguistic morphology, _____ is the process for reducing inflected words to their root form.
   (a) rooting (b) stemming
   (c) text-Proofing (d) both (a) & (b).

   (v) For a document *d* and a class *c, Bayes' Rule Applied to Documents and Classes is given by*
   (a) $P(c|d)=\frac{P(d|c)P(c)}{P(d)}$ (b) $P(c|d)=\frac{P(d|c)P(d)}{P(d)}$
   (c) $P(c|d)=\frac{P(d|c)P(d)}{P(d)P(c)}$ (d) $P(c|d)=P(d|c)P(d)$.

   (vi) Many words have more than one meaning. We have to select the meaning which makes the most sense in context. This can be resolved by
   (a) Fuzzy Logic (b) Word Sense Disambiguation
   (c) Shallow Symantic Analysis (d) (a), (b) and (c).

(vii) The optimum parameters are the ones for which each feature's predicted expectation equals its empirical expectation. The optimum distribution is/are
(a) Always unique (but parameters may not be unique)
(b) Always exists (if feature counts are from actual data)
(c) Both (a) and (b)
(d) (a) but not (b).

(viii) Word probability is calculated by
(a) likelyhood probability (b) bayes rule
(c) joint probability (d) none of (a), (b) and (c).

(ix) Methods of Dependency Parsing are
(a) Dynamic programming (like in the CKY algorithm)
(b) Constraint Satisfaction
(c) (a), (b), Deterministic parsing
(d) None of the above.

(x) Viterbi algorithm is used in
(a) speech processing (b) language processing
(c) both (a) and (b) (d) none of (a) & (b).

Group – B

2. (a) State the importance of Regular Expression(RE) in NLP. Give one example of an application using RE.
Describe the class of strings matched by the following regular expression:
(i) [a-zA-Z]+
(ii) [A-Z][a-z]*.

(b) Write the algorithm to convert Context Free Grammar (CFG) to Chomsky Normal Form (CNF).
Convert the following CFG into CNF
S → ASA | aB
A → B | S
B → b | ε

(2 + 2 + 2) + (2 + 4) = 12

3. (a) What is Finite State Transducer? Explain probability language modelling with examples.

(b) Using a Bigram model find the probability of the following
(i) P("versatile actress whose")
(ii) P("versatile across whose").

Where the Sentence given is "a stellar and versatile actress whose combination of sass and glamour…" .The counts from the Corpus of Contemporary American English with add-1 smoothing given are

(i) P(actress|versatile)=.000021 P(whose|actress) = .0010
(ii) P(across|versatile) =.000021 P(whose|across) = .000006.

(3 + 3) + 6 = 12

Group – C

4. (a) Explain morphological Parsing with Finite State Transducers for the following example

Lexical :

	c	a	t	+N	+PL			

Surface:

	c	a	t	s				

(b) Explain text Classification using Naive Bayes classifier .

(3 + 3) + 6 = 12

5. (a) What is the importance of smoothing in NLP? Explain Laplace smoothing.

(b) Define Hidden Markov Model (HMM). Write an algorithm to predict next word of a sentence using Bi-gram model (give suitable examples). How will you compute the accuracy of that model?

(2 + 2) + (2 + 4 + 2) = 12

Group – D

6. (a) What are the basic requirements of a meaning representation? Explain canonical form and expressiveness with suitable examples.

(b) (i) Consider the following grammar,
S->NP,VP
VP-> V,NP
NP-> Det,N
NP -> Det, ADJ, N
N-> cat|rat
V-> chased
ADJ -> large
Det -> Article
Using the above grammar, draw the parse tree for the sentence "The large cat chased the rat" and show all the steps clearly.

(ii) Briefly describe Supervised Learning based wordsense disambiguation approach with suitable example.

(1 + 2 + 2) + (4 + 3) = 12