	Utech
Name :	
Roll No.:	A game of Execution and Excitors
Invigilator's Signature :	

# CS/B.TECH(CSE)/SEM-8/CS-802F/2012 2012

# NATURAL LANGUAGE PROCESSING

Time Allotted: 3 Hours Full Marks: 70

The figures in the margin indicate full marks.

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

### **GROUP - A**

# (Multiple Choice Type Questions)

1. Choose the correct alternatives for the following :

 $10 \times 1 = 10$ 

- i) The use of the period (.) is to specify
  - a) any context
- b) any number
- c) any character
- d) none of these.
- ii) Word probability is calculated by
  - a) Likelihood probability b)
- Prior probability
- c) Baye's Rule
- d) None of these.
- iii) Minimum edit distance is computed by
  - a) phonology
  - b) dynamic programming
  - c) tautology
  - d) Hidden Markov Model (HMM).

8307 [ Turn over

# CS/B.TECH(CSE)/SEM-8/CS-802F/2012

- Uiedh
- iv) The use of brackets [ ] is to specify
  - a) disjunction of characters
  - b) disjunction of numbers
  - c) words sequence
  - d) none of these.
- v) Viterbi algorithm is used in
  - a) speech processing
  - b) language processing
  - c) speech & language processing
  - d) none of these.
- vi) In deleted interpolation algorithm, which symbol is used?
  - a) γ

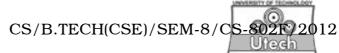
b) λ

c) o

- d) μ.
- vii) Entropy is used for
  - a) measure the information
  - b) correct the information
  - c) detect the information
  - d) handle the noise.
- viii) Phase structure Grammar is used in
  - a) Regular Grammar
  - b) Context Free Grammar ( CFG )
  - c) Context Sensitive Grammar (CSG)
  - d) none of hese.
- ix) Open class contains
  - a) Nouns

- b) Verbs
- c) both (a) & (b)
- d) none of these.
- x) Subcategorize of verbs is classified into
  - a) Transitive
- b) Intransitive
- c) both (a) & (b)
- d) none of these.

8307



#### **GROUP - B**

# (Short Answer Type Questions)

Answer any three of the following.

 $3 \times 5 = 15$ 

- 2. What is Regular Expression ? Write down the Regular Expression for the following languages :
  - a) The set of all alphabetic string
  - b) Column 1 Column 2 Column 3
  - c) 5.7 Gb.

2 + 3

3. Define two level Morphology with suitable example. Briefly describe the different types of Error Handling mechanism.

2 + 3

- 4. Write down the differences between Inflectional Morphology and Derivational Morphology with suitable example. What is stem? What is Morphemes? 3 + 1 + 1
- 5. Why POS ( Part of Speech ) Tagging is required in NLP ( Natural Language Processing ) ? Briefly compare the Top Down & Bottom Up Parsing technique. 2 + 3
- 6. Write down the concept of Feature Structure. What is Unification? What is Word Sense Disambiguation (WSD)?

3 + 1 + 1

### GROUP - C

# (Long Answer Type Questions)

Answer any *three* of the following.  $3 \times 15 = 45$ 

- 7. a) Define wordform, lemma, type, token.
  - b) Briefly describe the roles of Finite State Transducer (FST) with suitable example.
  - c) Define Prior probability and likelihood probability using Bayesian Method.
  - d) What is Confusion Matrix ? Why it is required in NLP (Natural Language Processing)? 4 + 5 + 4 + 2
- 8. a) What is Smoothing? Why it is required?
  - b) Write down the equation for trigram probability estimation.
  - c) Write down the equation for the discount  $d = c^*/c$  for add-one smoothing.
    - Do the same thing is used for Witten Bell smoothing? How do they differ? 2 + 1 + 3 + 5 + 4

# CS/B.TECH(CSE)/SEM-8/CS-802F/2012

- 9. a) Compute minimum edit by hand, figure out whether the word intention is closer to the word execution and calculate a minimum edit distance.
  - b) Estimate p(t/c) as follows (where  $c_p$  is the pth character of the word c) using Kernigham et al. four confusion matrices, one for each type of single-error.
  - c) Briefly describe Hidden Markov Model ( HMM ).
  - d) Compare open class & closed class word groups with suitable examples. 6 + 3 + 4 + 2
- 10. a) Draw tree structure for the following ATIS sentences : I prefer a morning flight.

I want a morning flight

Using  $S \rightarrow NP VP$ 

 $NP \rightarrow Pronoun$ 

| Pronoun-Noun

|Det Nominal

Nominal → | Noun Nominal

Noun

 $VP \rightarrow verb$ 

|Verb NP

|Verb NP PP

|Verb NP

- b) Write rules expressing the Verbal subcategory of English auxiliaries with example.
- c) Define predeterminers, cardianal numbers, ordinal numbers and quantifiers with suitable examples.
- d) How Transformation Based Learning ( TBL ) Rules are applied in NLP ( Natural Language Processing ) ?

5 + 3 + 4 + 3

 $3 \times 5$ 

- 11. Write short notes on any *three* of the following :
  - a) Weighted Automata
  - b) Baye's rule in noisy channel
  - c) Stochastic Part-of-Speech Tagging
  - d) HMM Tagging
  - e) Constituency & Agreement
  - f) Problems with the basic Top-Down Parser.

8307 4