Printed Pages: 02 **Sub Code:KAI-052**

Paper Id:	2	3	1	9	8	0

Roll No.													

B.TECH.
(SEM V) THEORY EXAMINATION 2022-23
NATURAL LANGUAGE PROCESSING

Time: 3 Hours ***Total Marks: 100***

Note: Attempt all Sections. If you require any missing data, then choose suitably.

SECTION A

1. Attempt *all* questions in brief. 2x10 = 20

(a) Differentiate between bigram and trigram.
(b) Define the N-grams model.
(c) What are the problems with PCFG?
(d) What do you understand by ambiguity?
(e) Give a detailed account of similarities and differences among the following set of lexemes: imitation, synthetic, artificial, fake and simulated.
(f) Explain Word Sense Disambiguation.
(g) Explain Articulatory Phonetics.
(h) What are the reasons to use short-time fourier transform.
(i) List the various issues of machine translation system.
(j) What is discourse planning?

SECTION B

2. Attempt any *three* of the following: 10x3 = 30

(a) Explain the need of smoothing and if we are given the following corpus:
<s> I am Sam </s>
<s> Sam I am </s>
<s> I am Sam </s>
<s> I do not like green eggs and Sam </s>
Using a bigram language model with Laplace smoothing, what is P(Sam |am)? Include <s> and </s> in your counts just like any other token.
(b) Discuss the relative advantages and disadvantages of deep versus shallow parsing. Also list various types of parsers. Explain in detail.
(c) Discuss various relations among the word senses. Also discuss various knowledge sources in WSD.
(d) Explain speech coding and speech synthesis with pattern matching in detail.
(e) Explain different pattern comparison techniques in detail. Also elaborate on Cepstral distances, and Weighted Cepstral distances.

SECTION C

3. Attempt any *one* part of the following: 10x1 = 10

(a) Write an algorithm for parsing a finite-state transducer using the pseudo code. Explain the algorithm with an example.

(b) Write an algorithm for simple top-down parser with an example.

4. Attempt any *one* part of the following: 10 x1 = 10

(a) Given the grammar and lexicon below, show the final chart for the followingsentence after applying the bottom-up chart parser.Remember that the finalchart contains all edges added during theparsing process. You may useeither the notation from class (i.e.nodes/links) or the notation from the book todepict the chart.

S → VP
VP → Verb NP NP → NP PP NP → Det Noun

PP → Prep Noun
Det → the
Verb → Find
Prep → in
Noun → men | suits

Find the men in suits.

(b) Give an account of CYK parser

5. Attempt any *one* part of the following: 10x1 = 10

(a) Discuss various knowledge sources in WSD.

(b) Between the words *eat* and *find* which would you expect to be more effectivein selecting restriction-based sense disambiguation.

6. Attempt any *one* part of the following: 10x1 = 10

(a) Define Articulatory Phonetics. Also explain production and classification of speech sounds in detail. Give examples.

(b) Write regular expressions for the following languages.

(i) the set of all alphabetic strings;
(ii) the set of all lower case alphabetic strings ending in a b;
(iii) the set of all strings from the alphabet a,b such that each a is immediately preceded by and immediately followed by a b;

Explain how these regular expressions will be used in speech processing.

7. Attempt any *one* part of the following: 10x1 = 10

(a) Explain different feature extraction and pattern comparison techniques used in speech analysis. Also elaborate on likelihood distortions, and spectral distortion, in detail.

(b) Explain hidden markovmodel with Baum-Welch parameter re-estimation. Also elaborate on its implementation issues.