

K. J. Somaiya Institute of Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

Nov – Dec 2023 (B.Tech) Program: Computer Engineering Examination: LY Semester: VII Scheme -II Course Code: CEDLC7034 and Course Name: Natural Language Processing				
Date: 29/11/2023		Duration: 2.5 Hours		Max. Marks: 60
Instructions: (1) All questions are compulsory. (2) Draw neat diagrams wherever applicable. (3) Assume suitable data, if necessary.				
Q. No.	Question	Max. Marks	CO	BT level
Q 1	Solve any six questions out of eight:	12		
i)	Identify and Describe the ambiguity in following sentence "I saw a bat"	2	CO1	Ap
ii)	Compare Inflectional morphology with derivational morphology.	2	CO2	U
iii)	What is stemming? Apply Porter stemming algorithm on following words to get the stems: (i) Tapping (ii) Smiling (iii) Computerization	2	CO3	Ap
iv)	Which techniques are used to resolve ambiguities in Natural Language?	2	CO4	U
v)	What is text coherence?	2	CO5	R
vi)	What is POS tagging? POS tag the following sentence: "Ram went to school after two months".	2	CO2	Ap
vii)	Identify number of bigram and trigram in the following sentence. "A big bug bite a little beetle but the little beetle bit the big bug back".	2	CO3	Ap
viii)	How Semantic Analysis is different than Pragmatic Analysis?	2	CO5	U
Q.2	Solve any four questions out of six.	16		
i)	Explain syntactic and semantic level of language understanding in natural language processing. Give example for each level	CO1	4	U
ii)	Explain rule based POS tagging with example.	CO2	4	Ap
iii)	Explain with suitable example lexeme and their relations.	CO3	4	U
iv)	Describe attachment for a fragment of English.	CO4	4	U

K. J. Somaiya Institute of Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

v)	Describe any two types of referring expressions and any two types of referents.	CO5	4	U
vi)	Design a finite transducer with E-insertion orthographic rule that parses from surface level "foxes" to lexical level "Fox+N+PL" Using FST.	CO2	4	Ap
Q.3	Solve any two questions out of three	16		
i)	Explain in detail information retrieval system. What are the challenges of it w.r.t. Indian regional language	CO1	8	U
ii)	Apply POS tagging using HMM on the given corpus and find the probabilities <s> start and </s> end of the sentence Corpus :Training data : <s>Book a car. </s> <s>Park the car. </s> <s>The book is in the car. </s> <s>The car is in a park.</s> Test data: <s>The park is a book</s>	CO3	8	Ap
iii)	Explain hierarchical discourse structure and reference resolution.	CO5	8	U
Q.4	Solve any two questions out of three.	16		
i)	Explain N-gram language model. Consider the following corpus and estimate the Bigram and tri gram models probabilities. <s> start and </s> end of the sentence <s>He said thank you.</s> <s>He said bye as he walked through the door.</s> <s>He went to San Diego.</s> <s>San Diego has nice weather.</s> Test data : It is raining in San Francisco.	CO2	8	Ap
ii)	Explain semantic analysis in natural language processing.	CO4	8	U
iii)	Explain Porter's stemming algorithm.	CO3	8	U
