

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

Nov – Dec 2024		Semester: VII
(B. Tech) Program: AI-DS Scheme II		
Regular Examination: LY	Course Name: Natural Language Processing	Max. Marks: 60
Course Code: AIC 702	Duration: 02.5 Hours	
Date of Exam: 26/11/2024		

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 two questions out of three: (05 marks each)	10		
a)	Define types and tokens. How many types and tokens are there in the following sentence: <i>"They picnicked by the pool, then lay back on the grass and looked at the stars"</i>		1	Ap
b)	<i>"Pattern Matching by Regular Expression is greedy"</i> . Justify the statement		2	An
c)	Following the CFG rules, draw a parse tree for the following sentences: 1) <i>"The woman saw a puppy"</i> 2) <i>"Sejal can bluff very well"</i>		3	An
Q 2	Solve any two questions out of three: (05 marks each)	10		
a)	Write a short note on dictionary approach used for Word Sense Disambiguation		4	U
b)	What is meant by the semantics of a natural language, and how does this differ from pragmatics?		5	U
c)	Explain in brief different steps involved in the development of Speech Recognition System		6	U
Q.3	Solve any two questions out of three. (10 marks each)	20		
a)	Analyze and identify the type of ambiguity present in the sentence, i) "Horse ran up the hill. It was very steep. It soon got very tired" ii) "The car hit the pole while it was moving." iii) "I saw a military tank." iv) "The old men and women left the room."		1	An
b)	What is the difference between stemming and lemmatization? How is lemmatization done?		2	U

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

Nov – Dec 2024		Semester: VII	
(B. Tech) Program: AI-DS Scheme II		Course Name: Natural Language Processing	
Regular Examination: LY		Max. Marks: 60	
Course Code: AIC 702		Duration: 02.5 Hours	
Date of Exam: 26/11/2024			

c)	Consider the given sentences: • Mary Jane can see Will • Spot will see Mary • Will Jane spot Mary? • Can Will spot Mary? Calculate the emission and transition probabilities. For the given sentence and probable tags: “Will can spot Mary” Will ⇒ Noun can ⇒ Modal Verb spot ⇒ Verb Mary ⇒ Noun Calculate the probability of the tags to be correct.	3	Ap
Q.4	Solve any two questions out of three. (10 marks each)	20	
a)	Explain Homonymy, Homophony, Hyponymy, Polysemy, Synonymy	4	U
b)	What are the 5 aspects of Pragmatic Analysis? Explain in brief.	5	U
c)	Given the short movie reviews, each labeled with genre: fun, couple, love, love ⇒ comedy fast, furious, shoot ⇒ action couple, fun, love, fly, fun ⇒ comedy furious, shoot, fun, fun ⇒ action fly, fast, shoot, love ⇒ action Consider a new document D: fast, couple, shoot, fly. Compute the most likely genre for D	6	Ap
