

U.S.N.								
--------	--	--	--	--	--	--	--	--

B.M.S. College of Engineering, Bengaluru-560019

Autonomous Institute Affiliated to VTU

January / February 2025 Semester End Main Examinations

Programme: B.E.

Semester: VII

Branch: Computer Science and Engineering

Duration: 3 hrs.

Course Code: 22CS7PENLP

Max Marks: 100

Course: Natural Language Processing

Instructions: 1. Answer any FIVE full questions, choosing one full question from each unit.
2. Missing data, if any, may be suitably assumed.

			UNIT - I		
			<i>CO</i>	<i>PO</i>	Marks
	1 a)	Analyze the effectiveness of grammar-based versus statistical language models in NLP applications. Which one would be more effective for information retrieval in multilingual scenarios?	CO1	PO1	10
	b)	Imagine you are developing a chatbot for a government service that needs to operate in Hindi and English. What challenges in language and grammar would arise? Suggest a solution using NLP applications and justify it.	CO1	PO1	10
OR					
	2 a)	Evaluate the challenges faced when processing Indian languages for NLP. How do they differ from challenges in other languages?	CO1	PO1	10
	b)	Consider the corpus given below. <s>I am Henry</s> <s>I like College</s> <s>Do Henry like college</s> <s>Henry I am</s> <s>Do I like Henry</s> <s>Do I like college</s> <s>I do like Henry</s> Predict the next word using trigram. i) do I like ____ ii) do I like college ____ Show the detailed steps of solving.	CO2	PO2	10
			UNIT - II		
	3 a)	Explain how regular expressions are used to identify specific word patterns in text.	CO2	PO2	10

Important Note: Completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages. Revealing of identification, appeal to evaluator will be treated as malpractice.

	b)	Describe the challenges in spelling error detection for morphologically rich languages.	CO2	PO2	10																												
		OR																															
4	a)	Compare the effectiveness of using regular expressions versus finite-state automata for tokenization tasks.	CO2	PO2	5																												
	b)	Recall the key principles of context-free grammar in syntactic parsing	CO2	PO2	5																												
	c)	Given the grammar and its probabilities, analyze if the following statement is ambiguous by designing appropriate parse trees. Also calculate the probabilities of both the parse trees. Justify which parse tree is more probable. <i>“Astronomers saw stars with ears”</i>	CO2	PO2	10																												
		<table border="1"> <thead> <tr> <th>Rule</th> <th>Probability</th> <th>Rule</th> <th>Probability</th> </tr> </thead> <tbody> <tr> <td>$S \rightarrow NP VP$</td> <td>1.0</td> <td>$NP \rightarrow NP PP$</td> <td>0.4</td> </tr> <tr> <td>$PP \rightarrow P NP$</td> <td>1.0</td> <td>$NP \rightarrow$ astronomers</td> <td>0.1</td> </tr> <tr> <td>$VP \rightarrow V NP$</td> <td>0.7</td> <td>$NP \rightarrow$ ears</td> <td>0.18</td> </tr> <tr> <td>$VP \rightarrow VP PP$</td> <td>0.3</td> <td>$NP \rightarrow$ saw</td> <td>0.04</td> </tr> <tr> <td>$P \rightarrow$ with</td> <td>1.0</td> <td>$NP \rightarrow$ Stars</td> <td>0.18</td> </tr> <tr> <td>$V \rightarrow$ saw</td> <td>1.0</td> <td>$NP \rightarrow$ Telescope</td> <td>0.1</td> </tr> </tbody> </table>	Rule	Probability	Rule	Probability	$S \rightarrow NP VP$	1.0	$NP \rightarrow NP PP$	0.4	$PP \rightarrow P NP$	1.0	$NP \rightarrow$ astronomers	0.1	$VP \rightarrow V NP$	0.7	$NP \rightarrow$ ears	0.18	$VP \rightarrow VP PP$	0.3	$NP \rightarrow$ saw	0.04	$P \rightarrow$ with	1.0	$NP \rightarrow$ Stars	0.18	$V \rightarrow$ saw	1.0	$NP \rightarrow$ Telescope	0.1			
Rule	Probability	Rule	Probability																														
$S \rightarrow NP VP$	1.0	$NP \rightarrow NP PP$	0.4																														
$PP \rightarrow P NP$	1.0	$NP \rightarrow$ astronomers	0.1																														
$VP \rightarrow V NP$	0.7	$NP \rightarrow$ ears	0.18																														
$VP \rightarrow VP PP$	0.3	$NP \rightarrow$ saw	0.04																														
$P \rightarrow$ with	1.0	$NP \rightarrow$ Stars	0.18																														
$V \rightarrow$ saw	1.0	$NP \rightarrow$ Telescope	0.1																														
		UNIT - III																															
5	a)	Demonstrate how to resolve an ambiguous sentence using word sense disambiguation techniques.	CO1	PO1	10																												
	b)	Elaborate the steps of centering algorithm. Apply the same for reference resolution on the following phrases U1: Ramu borrowed a book from the library U2: He gave it to Ramesh U3: He read it.	CO1	PO1	10																												
		OR																															
6	a)	Discuss the procedure for analyzing the discourse. Consider the following phrases and write the discourse structure. S₁ – Ram went to the bank to deposit money. S₂ – He then took a train to Shyam's cloth shop. S₃ – He wanted to buy some clothes. S₄ – He does not have new clothes for party. S₅ – He also wanted to talk to Shyam regarding his health	CO1	PO1	10																												

	b)	Illustrate the importance of cohesion in discourse processing with an example	CO1	PO1	10
		UNIT - IV			
7	a)	Outline the various tasks carried out by each phase of input / output of NLG with a neat diagram	CO2	PO2	10
	b)	Classify the different types of machine translation approaches.	CO2	PO2	10
		OR			
8	a)	Given the discourse “Savitha sang a song. The song was good although some people do not like it” draw the knowledge base and the discourse plan. Describe the discourse planning task in NLG	CO2	PO2	10
	b)	Give a systemic grammar for the sentence “Monica wrote a letter”. Elaborate on the three layers of representation.	CO2	PO2	10
		UNIT - V			
9	a)	Illustrate the importance of stemmers in improving information retrieval systems.	CO3	PO3	10
	b)	Classify the classical models of Information Retrieval (IR).	CO3	PO3	10
		OR			
10	a)	Give a design of how a basic information retrieval process is carried. Illustrate the process of indexing, stemming and elimination of stop words.	CO2	PO2	10
	b)	Compare the features of FrameNet and WordNet in the context of lexical resources.	CO3	PO3	10
