U.S.N.										

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

Autonomous Institute Affiliated to VTU

April 2025 Semester End Make-Up Examinations

Programme: B.E.   **Semester: VII**

Branch: Computer Science and Engineering   **Duration: 3 hrs.**

Course Code: 22CS7PENLP   **Max Marks: 100**

Course: Natural Langauge Processing

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

<div style="writing-mode: vertical">**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.</div>

			UNIT - I	CO	PO	Marks
1	a)		Outline any 5 applications of Natural language processing	CO2	PO2	**5**
	b)		Give the general structure of X-Bar. Apply X-Bar theory to the following phrase **"Dan ate fresh lobster at the pier"**	CO1	PO1	**5**
	c)		Analyze if the following phrase is ambiguous with appropriate parse tree. Give the associated phrase structured grammar as well "*I saw the man with a telescope*"	CO2	PO2	**10**
			OR			
2	a)		List out differences between grammar-based language model and Statistical Language model	CO2	PO2	**5**
	b)		Give the annotated C structure and F structure for the following phrase "*She saw stars*"	CO1	PO1	**5**
	c)		Consider the following corpus. <S> I am Ramu </S> <S> I like Office </S> <S> Do Ramu like office </S> <S> Ramu I am</S> <S> Do I like Ramu </S> <S> Do I like Office </S> <S> I do like Ramu </S> i. Construct a frequency table ii. With the corpus given above predict what follows <S> I like Ramu ____ using bigrams iii. Find out which of the two sentences is more probable <S> I like Office </S> or <S>Do I like Ramu </S>	CO2	PO2	**10**

			UNIT - II			
3	a)		Design a Minimum edit distance algorithm. Apply the same to find the minimum number of operations to be applied to convert "Execution" to "Intention"	CO1	PO1	**10**
	b)		Analyze the following grammar and show how the parsing for the sentence is carried on using Earley parser: "Take this book". Also give the parse tree. S -> NP VP | VP NP -> PRON | Det Nom Nom -> N | Nom N VP -> V | V NP Det -> the |a |this PRON -> he|she N -> book| boys| girl PRP -> with |in V -> take | takes	CO2	PO2	**10**
			OR			
4	a)		Define morphological parsing. Outline the two levels of morphological parsing with a neat diagram Find the possible sequences of states in finite state transducer for i) cats ii) foxes	CO1	PO1	**10**
	b)		Outline the CYK algorithm. Apply the same to show how the string "*a very heavy orange book*" is parsed for the following grammar NP → Det Nom Nom → AP Nom | Nom AP → Adv A | A Det → a | an Adv → very | extremely A → heavy | orange | tall | muscular Nom → book | orange | man	CO2	PO2	**10**
			UNIT - III			
5	a)		Illustrate word sense disambiguation, elaborate with an example? Analyze how lesk's algorithm and walker's algorithm will disambiguate the word 'bank' in the sentence "The bank by the river was peaceful". Assume a suitable database for the word "bank".	CO2	PO2	**10**
	b)		Apply Centering algorithm for the following discourses *"Shobha saw a beautiful Necklace at the store."* *"She showed it to Bindu"* *"She bought it."*	CO1	PO1	**10**
			OR			
6	a)		Elaborate any 5 coherence relation with example. Analyze the following and statements and give the coherence structure for the same by identifying appropriate coherence relations	CO1	PO1	**10**

			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 did not have new clothes for party. S5 − He also wanted to talk to Shyam regarding his investments			
	b)		What is reference resolution? Outline the steps in Mitkov's pronoun resolution. Analyze the following sentence and identify the antecedent indicators for the possible referents "she" and "it". "*The salesman at the shop showed a Dell laptop to Sindu. She liked it very much*"	CO2	PO2	**10**
			UNIT - IV			
7	a)		Give the Functional unification grammar for the sentence "Arun ate an apple"	CO2	PO2	**5**
	b)		How is Indian languages different from translating English Language? With suitable example outline the characteristics of Indian language	CO2	PO2	**5**
	c)		With a simple example analyse how example-based Machine translation is carried out. Outline the example-based Machine translation system with the various operations it performs	CO2	PO2	**10**
			OR			
8	a)		Outline how the input is processed through the three NLG tasks with a neat diagram. Elaborate the task carried out in each phase.	CO2	PO2	**8**
	b)		How is Rule based Machine translation different from direct Machine translation. Analyze both the models with suitable examples for each.	CO2	PO2	**10**
	c)		Analyze any two challenges in Machine translation.	CO2	PO2	**2**
			UNIT - V			
9	a)		Elaborate on WORDNET. Analyze the various applications where researchers can use this database.	CO2	PO2	**3**
	b)		Analyze how classical Information retrieval model functions in-order to retrieve documents based on user query. Elaborate any 2 Models.	CO2	PO2	**10**
	c)		Define Zipf's law. You have a document represented by 3 terms {cat, dog, mouse} with frequency being {4,1,1} respectively In the collection there are 100 documents. 15 documents contain the term "cat", 20 documents contain the word "dog" and 40 documents contain the word "mouse". Compute the tf-idf table	CO1	PO1	**7**
			OR			
10	a)		Elaborate on FRAMENET. Analyze the various applications where researchers can use this database	CO2	PO2	**3**

		b)	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**
		c)	Suppose there are 3 documents d1, d2, d3 and 5 terms {t1, t2, t3, t4, t5}, the term-document matrix is given below. Compute how these documents are clustered if clustering model is used for information retrieval. The threshold value is 0.7. Also find out the cluster vector r1 and r2. Use cosine similarity to compute the similarity matrix.	*CO1*	*PO1*	**7**

	D1	D2	D3
T1	1	1	0
T2	0	1	0
T3	1	1	1
T4	0	0	1
T5	1	1	0

******