Reg.No.:												

VIVEKANANDHA COLLEGE OF ENGINEERING FOR WOMEN
[AUTONOMOUS INSTITUTION AFFILIATED TO ANNA UNIVERSITY, CHENNAI]
Elayampalayam – 637 205, Tiruchengode, Namakkal Dt., Tamil Nadu.

Question Paper Code: 130014

B.E. / B.Tech. DEGREE END-SEMESTER EXAMINATIONS – NOV. / DEC. 2025
Seventh Semester
Computer Science and Technology
U19CTV35 – NATURAL LANGUAGE PROCESSING
(Regulation 2019)

Time: Three Hours Maximum: 100 Marks

Answer ALL the questions

Knowledge Levels (KL)	K1 – Remembering	K3 – Applying	K5 - Evaluating
	K2 – Understanding	K4 – Analyzing	K6 - Creating

PART – A

(10 x 2 = 20 Marks)

Q.No.	Questions	Marks	KL	CO
1.	State the conditions under which two Regular Expressions can be considered equivalent.	2	K1	CO1
2.	Mention how the concept of "morpheme boundary" influences Finite-State Morphological Parsing.	2	K2	CO1
3.	Estimate the interpolated bigram probability using weights $\lambda_1 = 0.6$ and $\lambda_2 = 0.4$ for unigram and bigram models respectively.	2	K2	CO2
4.	Given an HMM with transition probability P(VB|NN)=0.3 and emission probability P(eat|VB)=0.6, calculate P(NN,VB) for a two-word sequence "dog eat".	2	K2	CO2
5.	Given the PCFG rule probabilities:S → NP VP [0.7] ,NP → N [0.6],VP → V [0.8]. Calculate the probability of the sentence structure S → NP VP → N V.	2	K2	CO3
6.	Describe the role of sub-categorization in English syntax within context-free grammar.	2	K1	CO3
7.	Interpret the first-order predicate logic formula:$\exists x(Cat(x) \wedge Black(x)$ for this statement.	2	K2	CO4
8.	State one use of lexeme senses in improving machine translation accuracy.	2	K2	CO4
9.	How does dialogue act recognition benefit conversational agents in natural language understanding?	2	K2	CO5
10.	A statistical machine translation system gives the following phrase translation probability: Phrase A → Phrase A' with 0.7 Phrase B → Phrase B' with 0.5. Calculate the probability of translating the two-phrase sentence "A B" as "A' B'" assuming independent.	2	K2	CO5

PART – B

(5 x 13 = 65 Marks)

Q.No.		Questions	Marks	KL	CO
11.	a)	Discuss in detail about the role and implementation of finite state automata and finite-state transducers in morphological parsing and also describe how these models handle both regular and irregular morphological phenomena with examples of inflectional and derivational morphology processing using finite-state methods.	13	K1	CO1
		(OR)			
	b)	Elaborate on the design and working of the Porter Stemmer algorithm with its impact on natural language processing tasks, highlighting both its strengths and limitations. Additionally, compare the Porter Stemmer with other stemming or morphological analysis techniques.	13	K2	CO1
12.	a)	Apply smoothing techniques such as back off and deleted interpolation to overcome data sparsity in N-gram models while comparing the core principles and advantages of rule-based, stochastic, and transformation-based POS tagging methods.	13	K3	CO2
		(OR)			
	b)	Given a vocabulary of size V=100 and the following counts from a corpus:	13	K3	CO2

Bigram	Count
The book	50
The pen	30
The page	20
The (unigram)	200
Vocabulary size(V)	100

1. Calculate the MLE bigram probabilities $P(w_i \mid the)$ for "book", "pen", and "page".
2. Calculate the Add-One (Laplace) smoothed bigram probabilities $P(w_i \mid the)$ for the same words.
3. Using smoothed probabilities, compute the entropy H of the distribution $P(w \mid the)$ over the word "book", "pen", "page", and all other unseen words (assume uniform probability for unseen words).

13.	a)	Evaluate context-free grammars (CFGs) in relation to English syntax by constructing a grammar for a simple English sentence with noun, verb, and prepositional phrases, drawing its parse tree, analysing how top-down and Earley parsing handle ambiguity, then judging the strengths and weaknesses of both methods in natural language processing tasks.	13	K2	CO3
		(OR)			

	b)	Consider the simplified CFG rules: S → NP VP NP → DT NN | NP PP VP → V NP | VP PP | V PP → P NP Construct the parse tree for the sentence "The girl painted the picture on the wall." Provide a detailed step-by-step derivation. Explain how sub-categorization frames for verbs (e.g., which complements they require) influence the correctness of parse trees in this grammar. Compare and contrast how top-down parsing and Earley parsing would process this sentence, particularly in resolving attachment ambiguity with the prepositional phrase "on the wall."	13	K2	CO3
14.	a)	Analyze the importance of lexemes and their senses in semantic understanding by creating an example of word sense disambiguation (WSD) using a lexeme with multiple senses, examining the impact of accurate WSD on information retrieval systems, and assessing semantic approaches that improve retrieval performance.	13	K3	CO4
		(OR)			
	b)	Compare the advantages and disadvantages of top-down parsing versus bottom-up parsing techniques in natural language processing. Provide examples of scenarios where one method is preferred over the other.	13	K3	CO4
15.	a)	Explain different machine translation approaches, including transfer metaphor, interlingua, and statistical methods, with an example scenario showing how they yield varied translations while handling linguistic diversity and how discourse-level phenomena affect translation coherence	13	K2	CO5
		(OR)			
	b)	Compare the advantages and disadvantages of top-down parsing versus bottom-up parsing techniques in natural language processing with example scenarios where one method is preferred over the other.	13	K2	CO5

PART – C

(1 x 15 = 15 Marks)

Q.No.		Questions	Marks	KL	CO
16.	a)	Analyze the architecture and working of a conversational agent with emphasis on discourse management, dialog act interpretation, natural language generation, illustrating how discourse planning improves text coherence and user interaction, examining challenges in coherent multi-turn dialogue, then judging solutions for ambiguity, reference resolution, and context retention.	15	K4	CO5
		(OR)			

b)	Examine the architecture and key components of a statistical machine translation system by illustrating how transfer metaphor and interlingua approaches differ when translating between two languages, assessing the strengths and limitations of statistical methods, then judging strategies that address discourse-level coherence, idiomatic expressions, and translation quality.	15	K4	CO5