OBJECTIVE-I Answers

Natural Language Processing | III B.Tech II Semester | A.Y. 2025-26
Subject Code: 23CS4602C | 2 Marks Each

1. Define Natural Language Processing (NLP) and interpret any two major challenges faced by early NLP systems.

Answer: Natural Language Processing (NLP) is a branch of Artificial Intelligence that enables computers to understand, interpret and generate human language.
Two challenges: (1) **Ambiguity:** A word or sentence can have multiple meanings depending on context. (2) **Language variability:** Different grammatical structures, expressions and word forms make rule-based processing difficult.

---

2. State the purpose of parameter estimation in NLP models.

Answer: Parameter estimation determines model values, such as word or n-gram probabilities, from training data. Its purpose is to help a language model predict the likelihood of words and sentences accurately. For example, maximum likelihood estimation calculates P(w) = count(w) / total word count.

---

3. Define perplexity and explain its role in evaluating language models.

Answer: Perplexity is a measure of how well a language model predicts a sequence of words. For N words, $PP(W) = P(w1, ..., wN)^{-1/N}$. It measures prediction uncertainty: **lower perplexity** generally indicates better predictions when models are evaluated on the same test data and tokenization.

---

4. Outline the need for evaluating n-gram language models.

Answer: Evaluating n-gram language models measures how accurately they predict unseen word sequences. It helps compare different n-gram orders and smoothing methods, identify problems such as data sparsity or overfitting, and select a model that generalizes well. Common evaluation measures include perplexity and cross-entropy.

---

5. Summarize the motivation for using Maximum Entropy (MaxEnt) models in language modeling.

Answer: Maximum Entropy (MaxEnt) models allow many useful, overlapping features, such as previous words and linguistic context, to contribute to word prediction. They select a probability distribution that satisfies observed feature constraints while making the fewest additional assumptions. This provides a flexible alternative to relying only on fixed n-gram counts.
