UNIT-3

2 MARKS

1. Define a context-free grammar (CFG) formally as a 4-tuple. Name each component.

2. Write a context-free grammar rule for an English sentence (S).

3. What is a treebank? Give one major use of treebanks in natural language processing.

4. Convert the production rule A → B C D into Chomsky Normal Form (CNF). Show the steps.

6. Explain the difference between **constituency parsing** and **dependency parsing**.

7. Give two distinct parse trees for the sentence *"Visiting relatives can be boring"*

8. Describe the **CYK algorithm** in one sentence. What property must the grammar have for CYK to be applied?

9. What is **shallow parsing** (also called chunking)? Give an example of a chunk (e.g., a noun phrase) from the sentence *"The boy ate a pizza"*.

10. In a probabilistic context-free grammar (PCFG), how is the probability of a parse tree computed? Write the general formula.

12. Define **feature structure unification**.

13. Which property must the grammar have for CYK to be applied?

14. Write the general formula for Probabilistic context Free Grammar (PCFG)

5 MARKS

1. Define a context-free grammar (CFG) formally.
2. Write a CFG that generates simple English declarative sentences with noun phrases and verb phrases.
3. Explain how treebanks are used to extract probabilistic context-free grammar (PCFG) rules.
4. Convert the following CFG into Chomsky Normal Form.
5. Draw a dependency parse tree for the sentence *"She quickly ate the delicious pizza with a fork"*.
6. Given the following CNF grammar and the input sentence **"the cat chased the dog"**, run the CYK algorithm.
   S → NP VP
   NP → Det N
   VP → V NP
   Det → the
   N → cat | dog
   V → chased
7. Compute the Viterbi (most probable) parse for the sentence **"the cat saw a dog"**. Show the probability table and back pointers, and give the resulting parse tree with its probability.
8. Convert the following simple CFG into GNF:

$S \rightarrow AB$
$A \rightarrow aA \mid a$
$B \rightarrow bB \mid b$

9. How is a treebank constructed? Discuss two major challenges in treebank annotation, such as consistency and cross-linguistic variation.
10. Define feature structure unification. Unify the following two feature structures: [CAT V, AGR [NUM sg, PER 3], SUBJ [AGR ?]] and [CAT NP, AGR [NUM sg, PER 3]].
11. Describe the **probabilistic CYK algorithm** step by step. In the initialization step, how are probabilities for terminals assigned?
12. Construct a Context-Free Grammar (CFG) for simple English sentences of the form:
    Subject + *Verb* + *Object*.
    Apply your grammar to generate and parse the sentence: *"The boy eats apples"*.
13. Explain the concept of Treebanks and illustrate how a parsed sentence is stored in a treebank with a suitable example.
14. Explain the CYK algorithm with an example.
15. Given the CFG Rule
    S->NP VP
    VP->V NP PP
    Convert the grammar into Chomsky Normal Form (CNF) and justify each transformation step.
16. Compare constituency parsing and dependency parsing.
17. Explain about feature structure unification with an example.
18. Identify syntactic ambiguity in the sentence "I saw the man with a telescope". Construct two different parse trees and analyze the ambiguity
19. Explain about Heads and Head Dependencies with an example.
20. Apply shallow parsing (chunking) on the sentence "The quick brown fox jumps over the lazy dog". Identify the noun and verb phrases.
21. Explain feature structures and demonstrate unification using two feature sets (e.g., agreement in number and gender).