



C20-AIM-502

7691

BOARD DIPLOMA EXAMINATION, (C-20)
OCTOBER/NOVEMBER—2023
DAIM - FIFTH SEMESTER EXAMINATION
LANGUAGE PROCESSING

Time : 3 Hours]

[Total Marks : 80

PART—A

3×10=30

- Instructions :** (1) Answer **all** questions.
(2) Each question carries **three** marks.
(3) Answers should be brief and straight to the point and shall not exceed five simple sentences.

1. Define finite-state automata.
2. State the importance of tokenization.
3. Define N-gram.
4. List the issues in PoS tagging.
5. What is context free grammar?
6. What is syntactic parsing?
7. Define selectional restrictions.
8. Write about description logics.
9. State the importance of co-reference resolution.
10. What is British National Corpus (BNC)?

PART—B

8×5=40

- Instructions :** (1) Answer **all** questions.
(2) Each question carries **eight** marks.
(3) Answers should be comprehensive and criterion for valuation is the content but not the length of the answer.

11. (a) Explain transducers for lexicon and rules.

(OR)

(b) Explain detecting and correcting spelling errors.

12. (a) Compare Hidden Markov and Maximum Entropy models.

(OR)

(b) Explain the usage of unsmoothed and smoothed n-grams.

13. (a) Discuss dynamic programming parsing in detail.

(OR)

(b) Explain about probabilistic CYK algorithm.

14. (a) Describe the concept of thematic roles.

(OR)

(b) Explain the process of word sense disambiguation using supervised.

15. (a) Explain about the corpus WordNet.

(OR)

(b) Explain about the porter stemmer algorithm.

PART—C

10×1=10

- Instructions :** (1) Answer the following question.
(2) The question carries **ten** marks.
(3) Answer should be comprehensive and the criterion for valuation is the content but not the length of the answer.

16. Convert the following Context Free Grammar to Chomsky Normal Form (CNF).

$$S \rightarrow ASB$$

$$A \rightarrow aAS|a|\epsilon$$

$$B \rightarrow SbS|A|bb$$

★★★