**UNIVERSITY OF MALTA**
**BOARD OF STUDIES FOR INFORMATION TECHNOLOGY**
**Department of Computer Science & AI**
**B.Sc. I.T.(Hons.) Year III**
**January 2000 Examination Session**

**CSA305 Natural Language Processing using Logic Programming** — **29th January 2000**

**0900-1030**

---

*Answer* **three** *questions. Each question carries 30 marks (subsequently normalised to a percentage). Be sure to provide* **adequate exemplification** *to illustrate your understanding of concepts discussed even if not explicitly asked to do so.*

1. The design of Winograd's SHRDLU program is based on three principles: (i) A computer must understand what it is talking about; (ii) all aspects of language must be integrated in trying to model language understanding; (iii) NLP requires sentences to be understood in context.

   (a) Elaborate on each of these principles with examples to show why each is important.

   **[12 marks]**

   (b) Discuss how each affected the design of Winograd's program, again giving examples.

   **[12 marks]**

   (c) What are the limitations of SHRDLU? Which of these are a consequence of the design and which are inherent to NLP?

   **[6 marks]**

2. (a) (i) Define a context-free grammar. What are the essential differences between a context-free grammar and a definite clause grammar (DCG).

   **[5 marks]**

(ii) What criteria are available for judging how well a grammar describes a language?

**[5 marks]**

(b) The following sentences all have the form

S → NG VG

where NG is a *noun group* (in this case a proper noun) and VG is a *verb group.*

John walks
John is walking
John walked
John has walked
John has been walking
John will walk
John will be walking
John would have walked
John would have been walking

(i) Complete the grammar *using only context-free grammar rules.* The grammar should distinguish the different forms of the verbs, employing different lexical categories where appropriate. N.B. The lexicon is simulated by rules which rewrite lexical categories into words, e.g. ProperNoun → "John".

**[10 marks]**

(ii) Draw the parse trees assigned by the grammar to each sentence.

**[5 marks]**

(iii) How well does the grammar describe the system that the above sentences exemplify? Your answer should employ the criteria set out in (a (ii)) of this question.

**[5 marks]**

3. (a) (i) What is meant by logical form?

(ii) What is Montague's Rule-to-Rule hypothesis?

**[4 marks]**

(b) Suggest logical forms for the following sentences, indicating any assumptions made.

(i) Mike teaches CSA305

(ii) CSA305 is a third-year course.

(iii) All courses that Mike teaches are hard.

**[6 marks]**

(c) Write a complete DCG grammar/lexicon that computes the logical form of each of the above sentences. *Hint:* the last sentence involves writing an NP rule that handles the embedded relative clause.

**[20 marks]**

4. (a) Briefly list the pros and cons of top-down versus bottom-up parsing for NLP.

**[4 marks]**

(b) (i) Give an informal description of the shift-reduce algorithm.

**[4 marks]**

(ii) How does a shift-reduce parser handle empty categories? Give an example to illustrate your answer.

**[4 marks]**

(iii) Write a program that illustrates your understanding of the left corner algorithm.

**[4 marks]**

(iv) What are the advantages of the left corner algorithm over a plain shift reduce algorithm?

**[4 marks]**

(v) What modifications must be made to the left corner algorithm for it to handle empty categories?

**[10 marks]**

---



*end of paper*