UNIVERSITY OF LONDON
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 1998

MEng Honours Degrees in Computing Part IV
MEng Honours Degree in Information Systems Engineering Part IV
MSci Honours Degree in Mathematics and Computer Science Part IV
MSc Degree in Advanced Computing
for Internal Students of the Imperial College of Science, Technology and Medicine

This paper is also taken for the relevant examinations for the
Diploma of Membership of Imperial College
Associateship of the City and Guilds of London Institute
Associateship of the Royal College of Science

PAPER 4.85 / I4.24

NATURAL LANGUAGE PROCESSING

Tuesday, April 28th 1998, 2.00 - 4.00

Answer THREE questions

For admin. only: paper contains 4 questions

1a Most natural language sentences can have several different syntactic parses. How does a chart parser overcome the problem of representing these different parses? What are the other advantages of this approach?

b Explain the difference between a well-formed substring table and an active chart. Show how the fundamental rule operates.

c Show the active chart for the following sentence:

Jules saw that justice was important

using the following grammar:

s → np, vp.	rel → [that].
vp → v, np | v, adj | v, rel s.	d → [that] | [] .
np → d n | prop | pn.	n → [justice].
v → [saw] | [was].	prop → ['Jules'].
adj → [important].	pn → [that].

The three parts carry, respectively, 30%, 30%, 40% of the marks.

2a Explain, with examples, why commercial grammar checkers have not achieved the same level of use as spell checkers. What improvements in their performance would be required to make them useful to the average writer? Give **two** suggestions of measures that could be taken.

b Why have statistical and rule-based methods traditionally been seen as opposing approaches in natural language processing? Describe how this situation has changed in the last few years and how the two approaches can be combined.

3a Distinguish the aspects of a language known as morphology, syntax, semantics and pragmatics. Give examples of each and comment on the power of the formalism required to represent them.

b Describe a general method for handling quantifiers in the lexicon and show carefully how they combine with the rest of the sentence to generate a logical structure.

c What syntactic and semantic constraints would be used to identify anaphoric pronouns both within and between sentences. Illustrate your answer with examples.

4a What are the advantages of unification grammar (UG) compared with context-free grammars and compared with the simpler DCGs? Show briefly how to represent a UG grammar in Prolog and the steps necessary to convert it to Prolog.

b Given the following unification grammar:

$$S \rightarrow \begin{bmatrix} NP \\ num: X \\ case: nom \end{bmatrix} \begin{matrix} VP \\ [num: X] \end{matrix}$$

$$\begin{matrix} VP \\ [num: X] \end{matrix} \rightarrow \begin{matrix} V \\ [num: X] \end{matrix} \begin{matrix} NP \\ [case: acc] \end{matrix} \; PP$$

$$\begin{bmatrix} NP \\ num: X \\ case: C \end{bmatrix} \rightarrow \begin{bmatrix} Pronoun \\ num: X \\ case: C \end{bmatrix}$$

$$\begin{matrix} NP \\ [num: X] \end{matrix} \rightarrow \begin{matrix} D \\ [num: X] \end{matrix} \begin{matrix} N \\ [num: X] \end{matrix}$$

$$PP \rightarrow P \; \begin{matrix} NP \\ [case: acc] \end{matrix}$$

$$\begin{bmatrix} Pronoun \\ case: acc \\ num: pl \end{bmatrix} \rightarrow them$$

$$\begin{bmatrix} Pronoun \\ case: nom \\ num: sg \end{bmatrix} \rightarrow he$$

$$\begin{matrix} D \\ [num: pl] \end{matrix} \rightarrow two$$

$$\begin{matrix} N \\ [num: pl] \end{matrix} \rightarrow cases$$

$$V \rightarrow brought$$

$$P \rightarrow to$$

show a complete parse and syntactic structure for the following sentence:

he brought two cases to them

The two parts carry, respectively, 60%, 40% of the marks.

End of paper