UNIVERSITY OF LONDON
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 1999

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 / I 4.24

NATURAL LANGUAGE PROCESSING
Wednesday, April 28th 1999, 10.00 – 12.00

Answer THREE questions

For admin. only:
paper contains 4 questions

1a Define the terms "well-formed substring" and "dotted production" used in chart parsing, giving examples. Explain how they help in avoiding re-evaluating the parse of different parts of a sentence.

b Draw the well-formed substring diagram (including right hand sides) produced by parsing the following sentence:

```
football crowds like goals
```

using the following grammar:

```
s --> np, vp.
np --> n ; n, np ; n, pp.
n --> [football] ; [crowds] ; [goals].
vp --> v, np ; v, pp.
v --> [like] ; [crowds].
pp --> p, np.
p --> [like].
```

c In a *separate diagram*, show all the *active* nodes that would be added to this chart in the course of parsing using a bottom-up algorithm. (hint: this will not include the same arcs/nodes as the substring diagram).

2a Show how both Definite Clause Grammars and Unification based grammars show the features in a sentence. Explain the practical differences between them with illustrations.

b Augment the following context-free grammar to show how the agreement in number and case is produced using *either* DCG *or* Unification grammars:

```
s --> np, vp.
np --> pn.
np --> d, n.
pn --> [he]; [she]; [him]; [her]; [they]; [them].
n --> [questions] ; [question].
d --> [a] ; [the] ; [many].
vp --> v, np, np.
v --> [asks] ; [ask].
```

Indicate how the rules would be represented in the *other* formalism than the one you have chosen for the *first* and *last* of these rules only.

c Show a parse tree for the following sentence showing the features for number and case using either representation:

```
she  asks them many questions
```

3a   Define the terms "sense", "reference", "coreference" and "anaphor" and explain their relevance to language discourse.

b   What syntactic constraints can be used to identify pronouns within the same sentence? What extra information is needed to extend this to other sentences?

c   Identify all coreferences in the following text:

> John asked his friend to help him with his assignment. She looked at the questions and gave him some hints. But he wanted her to do it herself. She refused and told him not to be silly. The lecturer would find out and then she would be upset.

Briefly indicate the different types of coreference that occur.

4a   Show how lambda expressions can be used to capture the meaning of a variety of word types that represent open category words. Show clearly how they disappear in the course of parsing a full sentence.

b   Explain how generalised quantifiers can be used to overcome the difference between the translation into logic of the traditional quantifiers. Show how they are represented in the lexicon.

c.   Produce a parse of the following sentence, using an appropriate grammar, showing how the lambda expressions and generalised quantifiers are both used to produce a logical form of the sentence.

> Every course has an exam

Explain any difficulties that this straightforward translation might have.

Each lettered section of each question receives approximately equal marks

End of paper