UNIVERSITY OF LONDON
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2000

MEng Honours Degree in Information Systems Engineering Part IV
MEng Honours Degree in Mathematics and Computer Science Part IV
MEng Honours Degrees in Computing Part IV
MSc 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
Associateship of the City and Guilds of London Institute
This paper is also taken for the relevant examinations for the
Associateship of the Royal College of Science*

PAPER C485=I4.24

NATURAL LANGUAGE PROCESSING

Thursday 11 May 2000, 10:00
Duration: 120 minutes

Answer THREE questions

Paper contains 4 questions

1a What are the problems encountered by experiments in machine translation which have prevented its widespread use? Give examples of sentences whose translation is problematic and explain the difficulties encountered. What are sensible goals for work in machine translation today?

b A number of telephone companies are pursuing the idea of a telephone which interprets from one language to another. Explain in detail the steps that an algorithm to perform this process would need to pass through, concentrating on the central linguistic processing. Which parts of this process are achievable and which are likely to prove less tractable?

The two parts each carry 50% of the marks.

2a Show how you would use

i) a Definite Clause Grammar rule and

ii) a Unification-based grammar

to indicate the number and case aspects of the subject, verb and object in a sentence.

b Show how each of these apply to parsing the sentence:

She enjoys a good joke with him

Give two other linguistic uses of the parameters that are available in these grammars.

c What are the advantages and disadvantages of both of these formalisms compared with context-free grammars? Also what are their advantages and disadvantages when compared with each other?

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

3a Explain the phenomenon of unbounded movement using the example:

the book that she hoped for in the library was not there

What is the underlying structure of this sentence? How can this phenomenon be represented in a grammar? Explain carefully how this may be dealt with using DCGs.

b How is quantification handled in a general manner when parsing a sentence into an internal logical form?

Show how the logical form may be generated for the sentence:

Every exam has a catch

showing the parse tree annotated with all necessary variable bindings.

If there is an ambiguity in this sentence, show what it is and explain briefly how it could be treated in the same framework.

The two parts each carry 50% of the marks.

4a Define top-down and left-corner parsing. Describe carefully the way each operates and show how they differ.

b What types of grammatical productions cannot be handled by each of the above? Give examples and show how the limitations can be overcome.

c How does one measure the efficiency of a parsing system? Give reasons why many important natural language processing systems often use active chart parsers, even though they are less efficient than other methods on small examples.

The three parts carry, respectively, 30%, 35%, 35% of the marks.