UNIVERSITY OF LONDON
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2002

MEng Honours Degree in Information Systems Engineering Part IV
MSci 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

Friday 10 May 2002, 14:30
Duration: 120 minutes

Answer THREE questions

Paper contains 4 questions
Calculators not required

1a Briefly describe the different levels at which language can be analysed, indicating the formalism and programming style appropriate at each level.

b The famous "Eliza" conversing program does not address the usual levels of analysis. What methods does it use, and what are their advantages? In what situations are similar methods appropriate?

c How can statistics be used in natural language processing? Suggest two uses for a large corpus of processed text.

The three parts carry, respectively, 60%, 20%, 20% of the marks.

2a What are the components of a definite clause grammar (DCG)? How does it differ from a context free grammar? What is the advantage of a DCG?

b Briefly describe how each constituent of a DCG is translated into Prolog.

c Show how a DCG can be used to express the features of the sentence,

He catches them in big nets

and thus eliminate grammatically incorrect variants.

The three parts carry, respectively, 50%, 20%, 30% of the marks.

3a Show how a lambda expression can be used to represent the parameters of meaning for open category words like verbs and nouns in composing a sentence. Illustrate you answer with a semantic tree for *fido chases felix.*

b In which way are the standard quantifiers of predicate logic a mismatch for the composition of sentences with determiners or specifiers in traditional grammar? (Illustrate your answer). Explain how generalised quantifiers can be used to overcome this mismatch.

c Using an appropriate grammar, provide a parse of the sentence:

Every course has an exam

Show how lambda expressions and generalised quantifiers combine to compose a logical form.

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

4a Briefly explain, using examples where appropriate:

i) the advantages of using feature structures for a unification grammar, over a definite clause grammar.

ii) the advantages of bottom-up over top-down parsing for natural language processing.

iii) The advantages of chart parsing over conventional top down parsing, and how these advantages are achieved.

In which sense can each of these advantages be realised in practice?

b Briefly explain what the *interlingua* project hoped to gain by an approach to machine translation seeking a universal set of feature attributes. What are the wider problems of machine translation which contribute to the failure of such a programme? How can artificial intelligence planning contribute to their solution? Provide an example.

Parts a *and* b *carry, respectively, 60% and 40% of the marks.*