UNIVERSITY OF YORK

BA and BSc Degree Examinations 1999-2000

DEPARTMENT OF LANGUAGE AND LINGUISTIC SCIENCE

L433: Introduction to Computational Linguistics

Time allowed: $1\frac{1}{2}$ hours

Answer ALL questions

Total marks: 106

(1) Explain in in English what is meant by the following expression.

(a) $\mathcal{L}(G) = \{w \in T^* \mid S \stackrel{*}{\Rightarrow} w\}$ (5 marks)

(b) $\alpha A \gamma \Rightarrow \alpha \beta \gamma$ (5 marks)

(2) Given the following grammar,

$$\begin{array}{rr} \mathsf{S} \rightarrow \mathsf{NP\ VP} & \textit{Fido} : \mathsf{N} \\ \mathsf{NP} \rightarrow \mathsf{N} & \textit{Kim} : \mathsf{N} \\ \mathsf{VP} \rightarrow \mathsf{V\ NP} & \epsilon : \mathsf{N} \\ & \textit{ate} : \mathsf{V} \end{array}$$

(a) Provide a bottom-up, left-to-right derivation for the string *Fido ate*, (5 marks)

(b) state what kind of derivation it is, (1 mark)

(c) provide a top-down, left-to-right derivation for the string *Fido ate Sandy*, (5 marks)

(d) state what kind of derivation it is. (1 mark)

(e) What problem would a bottom-up, left-to-right parser encounter in parsing the string *Fido ate Sandy* as defined by the above grammar, and why? (6 marks)

(3) Under what circumstances would two grammars G_1 and G_2 be said to be *strongly equivalent*? (5 marks)

(4) (a) What kind of language is $a^n b^n c^n$? (1 mark)

(b) How is this class of languages defined? (3 marks)

(c) What language is defined by the following grammar? (1 mark)

$$\mathsf{S} \rightarrow \mathsf{a\ S\ b}$$
$$\mathsf{S} \rightarrow \epsilon$$

(d) What class of languages does it belong to? (1 mark)

(e) What defines the class of regular languages? (1 mark)

(f) What is the relationship between the classes of languages referred to in questions (a), (d) and (e) above? (4 marks)

(5) (a) What is the worst case recognition time for CF-PSGs? (2 marks)

(b) What is the worst case parsing time complexity for CF-PSGs? (2 marks)

(6) It is claimed that Swiss German is an example of a non-context free language.

(a) Explain what property of Swiss German renders it non-context free. (6 marks)

(b) What is the significance of this result for the parsing of natural languages? (4 marks)

(7) Why is search necessary in parsing? (6 marks)

(8) A shift-reduce parser is named after its two basic operations.

(a) describe briefly what the function of *shift* and *reduce* in the parser are, (6 marks)

(b) show the state of the stack and buffer of a shift-reduce parse on the string *Fido ate Sandy* using the grammar in question (2) above, and (6 marks)

(c) provide the Prolog code for the procedures `shift` and `reduce`. (6 marks)

(d) Compile the shift-reduce parser over the grammar in question 2). (6 marks)

(9) (a) How does a Definite Clause Grammar (DCG) differ from a CF-PSG? (2 marks)

(b) Write a DCG for the language is $a^n b^n c^n$. (5 marks)

(c) Explain what is meant by the term *vacuous backbone* and give a short example. (6 marks)

(d) What problems does a vacuous backbone present for parsing? (5 marks)