2005 Paper 4 Question 1

Compiler Construction

(a) A (phrase-structured) grammar is often defined to be a 4-tuple (N, T, R, S) where R is a set of production rules. Explain what the other components of the 4-tuple are. Explain also the (most general form of) production rules, how these are conventionally restricted and why one might wish to restrict them.

[6 marks]

(b) Give a grammar which is ambiguous. [2 marks]

(c) Give a grammar which is not a regular grammar but which generates a regular language containing an infinite number of strings. [2 marks]

(d) Is it possible to write a grammar which generates the strings $\{aa, aaa, aaaaa, \ldots, a^p, \ldots\}$ where p is prime? (A general argument for or against suffices.) [2 marks]

(e) It is desired to construct a simple "pocket-calculator" program using yacc and lex (or other similar automated tools of your choice) which can parse strings such as "1+(10-5-3)*5+2=" and print the result, 13 in this case. Outline the overall structure of your program components. Give full details of the input to yacc and lex (or equivalent). (Precise syntactic details are not important, but your answer should show an understanding of the principles involved.)

[8 marks]