THE UNIVERSITY
of LIVERPOOL

JANUARY 2002 EXAMINATIONS

Bachelor of Science : Year 3

EFFICIENT SEQUENTIAL ALGORITHMS

TIME ALLOWED : Two Hours and a half

INSTRUCTIONS TO CANDIDATES

Credits will be given for the best **FOUR** answers only. Each question is marked out of 25.

If you attempt to answer more than the required number of questions (in any section), the marks awarded for the excess questions will be discarded (starting with your lowest mark).

THE UNIVERSITY
of LIVERPOOL

1. Dynamic programming: the Longest Common Subsequence (LCS) problem.

   (a) Define the Longest Common Subsequence Problem. **(6 marks)**

   (b) State the theorem characterising the structure of the longest common subsequences of two given strings. **(6 marks)**

   (c) Describe, using pseudo-code, the dynamic programming algorithm for computing the length of the longest common subsequence of two strings. **(7 marks)**

   (d) Given the strings $X \equiv$ (A C B C D B) and $Y \equiv$ (A B C B), compute table lcs for $0 \leq i \leq 6$ and $0 \leq j \leq 4$. **(6 marks)**

2. String algorithms. Knuth-Morris-Pratt algorithm (KMP).

   (a) Define the notion of prefix function π of a pattern P. **(6 marks)**

   (b) Define the prefix function π for the pattern $P \equiv$ 112112112. **(7 marks)**

   (c) List the sequence of values taken by the "state" variable q in the pseudo-code of KMP when the pattern $P \equiv$ abac is matched against the text $T \equiv$ ababacabab-cababccbabacba. (Hint: You will need to compute π first). **(12 marks)**

3. Huffman codes.

   (a) Given the following table

Symbol	a	b	c	d	e	f	g
Frequency	0.26	0.254	0.118	0.116	0.127	0.063	0.062

   compute the Huffman code for the alphabet $\Sigma = \{$ a, b, c, d, e, f, g$\}$. Describe the code using a tree whose leaves are pairs (symbol,codeword). **(9 marks)**

   (b) Prove one of the following properties of the Huffman coding algorithm:

   greedy-choice An optimal solution can be reached by making a locally optimal choice at each step.

   optimal-substructure An optimal solution is formed by combining optimal solutions to subproblems.

   (9 marks)

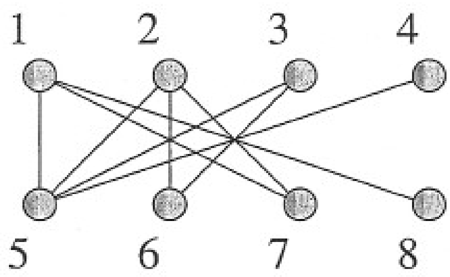
   (c) Give a description, using pseudo-code, of a decoder for the Huffman code built in part (a). You may assume both the encoded and the decoded texts are stored in arrays. **(7 marks)**

4. Graph algorithms. Relationship minimal vertex covers maximum matchings in bipartite graphs.

   (a) Define the following graph theoretic notions: *vertex cover*, *matching*, *maximal matching*, *maximum matching*. **(7 marks)**

   (b) Give a pseudo-code description for the hungarian algorithm to find a maximum matching in a bipartite graph. **(7 marks)**

   (c) Run the hungarian algorithm on the graph in the figure below:



   assuming the starting maximal matching is $M = \{\{1,5\},\{2,6\}\}$. Show the forests F built at every "augmentation" iteration. List the edges in the maximum matching returned by the hungarian algorithm. **(9 marks)**

   (d) What is a minimum vertex cover in the same graph? **(2 marks)**

5. Space complexity.

   (a) Define the term *crossing sequence*. **(6 marks)**

   (b) Explain when two crossing sequences are said to be compatible with respect to a portion of tape y. **(8 marks)**

   (c) State Hopcroft and Ullman's theorem on the space efficient simulation of time bounded Turing machines. **(5 marks)**

   (d) Prove Hopcroft and Ullman's theorem by describing the space efficient simulation based on crossing sequences. **(6 marks)**