COMPUTER SCIENCE TRIPOS Part II (General)
DIPLOMA IN COMPUTER SCIENCE

Wednesday 7 June 2000 1.30 to 4.30

Paper 12 (Paper 3 of Diploma in Computer Science)

Answer **five** *questions.*
Submit the answers in five **separate** *bundles, each with its own cover sheet. On each cover sheet, write the numbers of* **all** *attempted questions, and circle the number of the question attached.*
Write on **one** *side of the paper only.*

1 Data Structures and Algorithms

Explain what is meant by the terms *directed graph*, *undirected graph* and *bipartite graph*. [3 marks]

Given a bipartite graph, what is meant by a *matching*, and what is an *augmenting* path with respect to a matching? [4 marks]

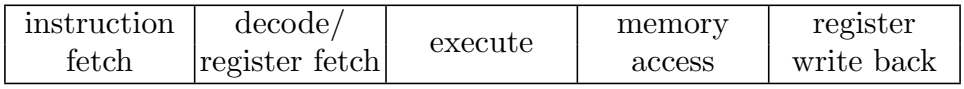
Prove that if no augmenting path exists for a given matching then that matching is maximal. [6 marks]

Outline an algorithm based on this property to find a maximal matching, and estimate its cost in terms of the number of vertices n and edges e of the given bipartite graph. [7 marks]

2 Computer Design

Modern processors exploit instruction level parallelism in order to improve throughput.

(*a*) Control-flow processors use pipelining to improve throughput. What limits throughput in the simple 5-stage pipeline depicted below? [10 marks]



(*b*) What is the principle of operation of a static data-flow processor and how does it resolve data dependencies? [10 marks]

3 Digital Communication I

What is meant by the term *flow control*? [3 marks]

What is meant by the term *credit-based* flow control? [4 marks]

What is meant by *start–stop* (or XON–XOFF) flow control? [4 marks]

A start–stop system is used on a 10 kbps link with a constant delay of 5 ms. How much buffer must a receiver keep in reserve for "stopping time" in order to prevent information loss? [3 marks]

Which system is more appropriate to use across the Internet and why? [6 marks]

4 Computer Graphics and Image Processing

Explain how a cathode ray tube (CRT) works, including details of how colour is achieved. [8 marks]

Describe a run-length encoding scheme for encoding images whose pixels have eight-bit intensity values. [8 marks]

Calculate the *best* possible compression ratio achievable with your scheme and describe the situation(s) in which this ratio would be achieved. [2 marks]

Calculate the *worst* possible compression ratio achievable with your scheme and describe the situation(s) in which this ratio would be achieved. [2 marks]

5 Business Studies

What are the differences between *profit and loss* and *cash flow* statements? [5 marks]

What are the differences between *debt* and *equity* finance? [5 marks]

What is an *option* and how might it be valued? [5 marks]

Comment on the current prices of high-tech stocks. [5 marks]

6 Comparative Programming Languages

Outline how you would implement complex numbers in C++. Your implementation should attempt to make complex numbers look as if they were built into the language by allowing new complex numbers to be declared, initialised, assigned and operated on by the normal arithmetic operators. [13 marks]

Discuss to what extent a good C++ compiler could implement your version of complex numbers as efficiently as if they had been a primitive type in the language. [7 marks]

7 Compiler Construction

Give a brief description of the main features of *either* Lex and Yacc *or* the corresponding Java tools JLex and Cup. [5 + 5 marks]

Illustrate their use by outlining how you would construct a parser for expressions composed of identifiers, integers, unary minus and binary operators $+$, $-$, $*$ and $/$. Your parser is expected to create a parse tree in a format of your choice representing the expression that is presented to it. If it helps, you may assume that expressions will be terminated by writing a semicolon after them. [10 marks]

8 Prolog for Artificial Intelligence

Consider the following problem to be solved using a Prolog program:

Given a closed planar polygon chain represented as a list of n vertices

$$\texttt{[v(}x_1\texttt{,}y_1\texttt{), v(}x_2\texttt{,}y_2\texttt{), ..., v(}x_n\texttt{,}y_n\texttt{)]}$$

compute the area of the enclosed polygon, and the orientation of the chain. The area is computed by the line integral $1/2 \int x \, \mathrm{d}y - y \, \mathrm{d}x$ where the integral is over the polygon chain. A naïve solution is given by the following program, which defines the predicate `area`. The goal `area(Chain,Area)` succeeds when `Chain` is the list of vertices, and the magnitude of `Area` is the area of the polygon bounded by the chain. The sign of `Area` is positive if the orientation of the polygon is anticlockwise and negative if it is clockwise:

```
area([X],0).
area([v(X1,Y1),v(X2,Y2)|VS],Area):-
          area([v(X2,Y2)|VS],Temp),
          Area is Temp + (X1 * Y2 - Y1 * X2) / 2.
```

Explain how vertices are processed by this procedure. [4 marks]

Why does this program execute inefficiently? [3 marks]

Write an alternative definition that is tail-recursive and makes use of accumulator variables. [10 marks]

Explain why your alternative definition executes more efficiently. [3 marks]

9 Databases

What is meant by a *functional dependency* between sets of attributes in a relational database schema? What conditions must be satisfied for a relation to be in Boyce–Codd Normal Form (BCNF)? [4 marks]

The Department of Transport is implementing plans to tax traffic congestion. From 2002, cars will carry approved radio-control units which at first will be used only to monitor vehicle movement. In controlled areas sensors identify all vehicles, recording their positions periodically. Amber signs flash when overall traffic flow drops below some threshold, and vehicles within the controlled area may be fined for lack of progress.

The owner of each vehicle has an account with the Department of Transport; owners can transfer funds to ensure that their account is in credit. Once credit is exhausted the level of fine increases by a factor 3, and a summons is sent by mail to the vehicle owner's registered address. In order to maintain proper accounts it is essential to keep an accurate record of each monitored offence.

You are employed to design the relational database that will enforce the scheme, including provision for vehicle and driver registration, monitoring of vehicle offences and management of vehicle accounts. Describe the schema you propose, stating clearly any assumptions that you make. You need not discuss the calculation of the fines due. [12 marks]

Outline the flow of information through the database. To what extent does the application require real-time transaction programming? [4 marks]

10 Numerical Analysis II

A Riemann integral over $[a, b]$ is defined by

$$\int_a^b f(x)\,dx = \lim_{\substack{n\to\infty \\ \Delta\xi\to 0}} \sum_{i=1}^{n} (\xi_i - \xi_{i-1})\, f(x_i)$$

Explain the terms *Riemann sum* and *mesh norm*. [4 marks]

With respect to an integral over $[-1, 1]$ which of the following are *not* Riemann sums? Give explanations.

(a) $0.2f(-0.9) + 0.8f(-0.1) + 0.8f(+0.1) + 0.2f(+0.9)$

(b) $0.8f(-0.9) + 0.2f(-0.1) + 0.2f(+0.1) + 0.8f(+0.9)$

(c) $0.7f(-0.6) + 0.3f(-0.4) + 0.3f(+0.4) + 0.7f(+0.6)$

(d) $0.5f(-0.7) + 0.8f(0) + 0.5f(+0.7)$

(e) $0.3f(-0.7) + 1.0f(+0.1) + 0.7f(+0.7)$

[5 marks]

Suppose $\mathbf{R}$ is a rule that integrates constants exactly over $[-1, 1]$, and $f(x)$ is bounded and Riemann-integrable over $[a, b]$. Write down a formula for the composite rule $(n \times \mathbf{R})f$ and prove that

$$\lim_{n\to\infty} (n \times \mathbf{R})f = \int_a^b f(x)\,dx$$ [6 marks]

Which of the examples (a) to (e) converge in composite form? [2 marks]

Does the rule

$$-0.5f(-1) + 1.5f(-0.4) + 1.5f(+0.4) - 0.5f(+1)$$

converge in composite form? Comment on its suitability for this purpose.
[3 marks]

11 Introduction to Functional Programming

Consider the following definitions of the functionals `foldl` and `foldr`:

```
fun foldl f e []     = e
  | foldl f e (h::t) = foldl f (f(h,e)) t;

fun foldr f e []     = e
  | foldr f e (h::t) = f(h, foldr f e t);
```

What is the type of `foldl`? [2 marks]

What is the type of the expression `foldr op/`? [2 marks]

For *each* of the following functions, write an ML definition using one of the functionals `foldl` or `foldr`.

(*a*) `product: (real list) -> real`, which given a list of real numbers gives their product.

(*b*) `exists: ('a -> bool) -> ('a list) -> bool`, which given a predicate p and a list l determines whether there is any element of l satisfying p.

(*c*) `length: ('a list) -> int` which determines the length of a list.

[9 marks]

Prove, by induction on lists, that for all lists of integers l, the following identity is true:

$$\texttt{foldl op+}\ 0\ l = \texttt{foldr op+}\ 0\ l$$

[7 marks]

12 Computer Vision

Explain the Bayesian approach to solving problems in computer vision. Explain the notion of an *Inverse Problem* and how computer vision can be regarded thereby in a formal sense as *inverse graphics*. Write down Bayes' rule in general form, and explain the interpretation of its terms as:

- probability of the image, given the object
- probability of the object, given the image

What is the role of the "prior?"

Discuss and illustrate the Bayesian approach in terms of 3D surface reconstruction, given the reflectance data in an image.

[20 marks]

13 Complexity Theory

Give precise definitions of *polynomial time reductions* and *NP-completeness.* [2 marks each]

Consider the following two decision problems on *undirected* graphs.

> **3-node-colourability**: the collection of graphs $G = (V, E)$ for which there is a mapping $\chi : V \to \{r, g, b\}$ such that if $(u, v) \in E$, then $\chi(u) \neq \chi(v)$.
>
> **3-edge-colourability**: the collection of graphs $G = (V, E)$ for which there is a mapping $\chi : E \to \{r, g, b\}$ such that if $(u, v), (u, v') \in E$, with $v \neq v'$, then $\chi(u, v) \neq \chi(u, v')$.

Show that there is a polynomial time reduction from **3-edge-colourability** to **3-node-colourability**. [8 marks]

The problem **3-edge-colourability** is known to be NP-complete. Using this information, for *each* of the following statements, state whether or not it is true. In each case, give complete justification for your answer.

(a) There is a polynomial time reduction from **3-node-colourability** to **3-edge-colourability**. [3 marks]

(b) **3-node-colourability** is NP-complete. [3 marks]

(c) **3-edge-colourability** is in PSPACE. [2 marks]

END OF PAPER