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

Thursday 8 June 2000 1.30 to 4.30

Paper 13 (Paper 4 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 Distributed Systems

In a distributed electronic conference application each participant has a replica of a shared whiteboard. Only one user at a time may write to the whiteboard, after which the user's update is propagated to all conference members. A dedicated process manages each whiteboard replica. Define and discuss a protocol that the replica managers could use to achieve this mutual exclusion. [10 marks]

A service database is replicated within a widely distributed system to achieve high availability and low access latency. Within a hierarchy of replicas a set of top-level primary servers maintain strong consistency of their replicas. Describe how this strong consistency could be achieved. [10 marks]

2 Computer Design

Why are the following statements fallacies?

(*a*) MIPS is an accurate measure for comparing performance among computers. [5 marks]

(*b*) A benchmark is a typical program which accurately predicts the performance of all other applications. [5 marks]

(*c*) Complex instruction set computers minimise the semantic gap between machine code and high-level languages, thereby making applications run more quickly. [5 marks]

(*d*) Data caches always improve processor throughput. [5 marks]

3 Digital Communication I

Compare circuit switching and packet switching, paying attention to channel characteristics and resource efficiency. [7 marks]

What is *wave division multiplexing* (WDM)? Is it more like circuit switching or packet switching and why? [7 marks]

Wave length conversion is the process, either optical or optical–electronic–optical, of receiving a signal on one wavelength and transmitting on another.

How does wave length conversion ease the problem of routing optical carriers in a network? [3 marks]

"The huge capacity of WDM systems will mean that IP becomes redundant." Discuss. [3 marks]

4 Computer Graphics and Image Processing

Give an algorithm for drawing the part of a circle which lies in the first octant. Assume that the circle has integer radius and is centered at the origin. Assume that you have a function *setpixel*(x, y) which turns on pixel (x, y). [10 marks]

Derive a matrix, or a product of matrices, to perform a clockwise 2D rotation of arbitrary angle, θ, about an arbitrary point, (x_c, y_c). [4 marks]

Provide an algorithm to ascertain whether the Bezier curve defined by $P_1P_2P_3P_4$ lies within some tolerance, ϵ, of the straight line segment, $\overline{P_1P_4}$, which joins the Bezier curve's end points. Your algorithm must return *false* if the Bezier curve is outside the tolerance; it must return *true* if the curve is well inside the tolerance; it may return either *true* or *false* if the curve is inside, but not well inside, the tolerance. [6 marks]

5 Business Studies

What is meant by *a critical path*? [5 marks]

The village bakery has asked you to advise them about setting up a web site, including a trading function.

Draw up a project plan, illustrated by a GANTT chart, and indicate the critical path. [5 marks]

Make an estimate of the costs involved, and estimate how much working capital you would need. [5 marks]

What other advice would you give them? [5 marks]

6 Comparative Programming Languages

Give a brief summary of the main syntactic constructs found in the programming language Smalltalk. Other languages often have the conditional constructs if–then–else and while. Show how these two constructs can be defined in Smalltalk. [8 marks]

Illustrate the use of Smalltalk by showing how you would define a method to compute the factorial of an integer. [8 marks]

Although Smalltalk was originally designed to be an interpretive language, modern implementations are dramatically more efficient. Briefly outline what techniques might have been used to make this improvement. [4 marks]

7 Compiler Construction

Describe how a parse tree can be translated into a sequence of assembly language instructions based on a pattern matching graph derived from a set of tree rewriting rules where each rule has a cost and a corresponding fragment of code. Illustrate your answer using the following rules:

```
Ri = Kk                    LDI  Ri,Kk     Cost 2
Ri = add(Ri,Kk)            ADDI Ri,Kk     Cost 3
Ri = add(Ri,Rj)            ADD  Ri,Rj     Cost 3
Ri = add(Ri,add(Rj,Kk))    ADD  Ri,Rj,Kk  Cost 4
```

applied to the following parse tree:

```
add(K1,add(add(K2,add(K3,K4)),add(K5,K6)))
```

[15 marks]

Discuss the advantages and disadvantages of this approach to code generation. [5 marks]

8 Prolog for Artificial Intelligence

One of the regulations of the International Rugby Board (IRB) states that for a player to be eligible to play for a given country, the player's father or mother or grandfather or grandmother must have been born in that country. Assume that there is a complete genealogical database consisting of Prolog clauses of the form person(P, B, F, M), where P is a person's name, B is the country of P's birth, F is their father's name and M is their mother's name. For example, the clause

person(bruce, australia, rhodri, bronwyn).

might appear in such a database. Further assume that names in the database are constructed so as to refer uniquely to individuals. Write Prolog clauses defining the predicate eligible such that goals of the form eligible(P,C) succeed if and only if person P is eligible to play for country C according to the above regulation. [10 marks]

Given a list of players on a given country's team, define a predicate checkteam that will check each member of the team for eligibility according to the eligible predicate, and furthermore check that each player appears on the list only once. The checkteam goal will fail if any player is ineligible or if any player is listed more than once. [10 marks]

9 Databases

Describe the basic architecture of the ODMG standard for Object Data Management. [10 marks]

What support is provided for transactions? What locking modes are available, and how are they used by the database runtime systems? [4 marks]

The query language OQL is recognised as a standard by the Object Management Group (OMG). To what extent is it similar to SQL, and in what ways does it differ? [6 marks]

10 Numerical Analysis II

Explain the terms (a) *positive definite*, (b) *positive semi-definite* for a symmetric matrix $\mathbf{A}$. If a square matrix $\mathbf{B}$ is non-singular, which of the properties (a) or (b) most accurately describes $\mathbf{B}^T\,\mathbf{B}$? What if $\mathbf{B}$ is singular? [4 marks]

State *Schwarz's inequality* for the product $\mathbf{AB}$. In what way is this modified for the product $\mathbf{Ax}$, where $\mathbf{x}$ is a vector? What are the *singular values* of $\mathbf{A}$, and how are they related to the l_2 norm of $\mathbf{A}$? In the *singular value decomposition* $\mathbf{A} = \mathbf{UWV}^T$, what is $\mathbf{W}$? [5 marks]

Let $\hat{\mathbf{x}}$ be an approximate solution of $\mathbf{Ax} = \mathbf{b}$, and write $\mathbf{r} = \mathbf{b} - \mathbf{A}\hat{\mathbf{x}}$, $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$. Find an expression which is an upper bound for the relative error $||\mathbf{e}||/||\mathbf{x}||$ in terms of computable quantities. Explain how this result may be interpreted if the l_2 norm is used. [8 marks]

Suppose $\mathbf{A}$ is a 5×5 matrix and $\mathbf{Ax} = \mathbf{b}$ is to be solved by singular value decomposition. If *machine epsilon* $\simeq 10^{-15}$ and the singular values of $\mathbf{A}$ are $1, 10^{-6}, 10^{-10}, 10^{-17}, 0$ write down the generalised inverse $\mathbf{W}^+$ that you would use. [3 marks]

11 Introduction to Functional Programming

The following is a recursive definition of a datatype `ltree`, which is intended to represent binary trees in which data is stored only at the leaves, not at internal nodes.

```
datatype 'a ltree = Empty
                  | Leaf of 'a
                  | Branch of ('a ltree) * ('a ltree);
```

(a) Write a simple recursive function

```
elems: ('a ltree) -> ('a list)
```

which gives a list of the data elements stored in a tree. [4 marks]

(b) Write an iterative version of this function

```
elemsi: ('a ltree * 'a list) -> ('a list)
```

which does not require appending of lists, and which satisfies the equality:

$$\texttt{elemsi}(t, l) = \texttt{elems}(t)@l$$

You do not have to prove the equality. [6 marks]

(c) Given the datatype of sequences:

```
datatype 'a seq = Nil
                | Cons of 'a * (unit -> 'a seq)
```

write a function `appendq:(('a seq)*('a seq))-> ('a seq)` for appending two sequences. [4 marks]

Use this to define a function `elemsq:('a ltree)-> ('a seq)` which, given a tree, produces a lazy list of the data elements stored in it. [6 marks]

12 Computer Vision

Define the *Correspondence Problem*, detailing the different forms that it takes in stereo vision and in motion vision.

(*a*) In each case, explain why the computation is necessary. [5 marks]

(*b*) What are the roles of space and time in the two cases, and what symmetries exist between the stereo vision and the motion vision versions of the Correspondence Problem? [5 marks]

(*c*) How does the complexity of the computation depend on the number of underlying features that constitute the data? [5 marks]

(*d*) Briefly describe at least one general approach to an efficient algorithm for solving the Correspondence Problem. [5 marks]

13 Complexity Theory

State the hierarchy theorems for time and space. [4 marks]

A linear time reduction from a language L_1 to L_2 is a reduction that can be computed by a deterministic Turing machine in time $O(n)$.

A class of languages $\mathcal{C}$ is closed under linear time reductions if whenever $L_2 \in \mathcal{C}$ and L_1 is linear-time reducible to L_2, then $L_1 \in \mathcal{C}$.

For *each* of the following complexity classes (a) to (d), say

- whether it is closed under linear time reductions
- whether it contains problems that are complete under linear time reductions

Give full justification for your answers.

(a) DSPACE(n^2) [4 marks]

(b) L [4 marks]

(c) P [4 marks]

(d) NP [4 marks]

END OF PAPER