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

Thursday 7 June 2001 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.*

You may not start to read the questions printed on the subsequent pages of this question paper until instructed that you may do so by the Invigilator

1 Distributed Systems

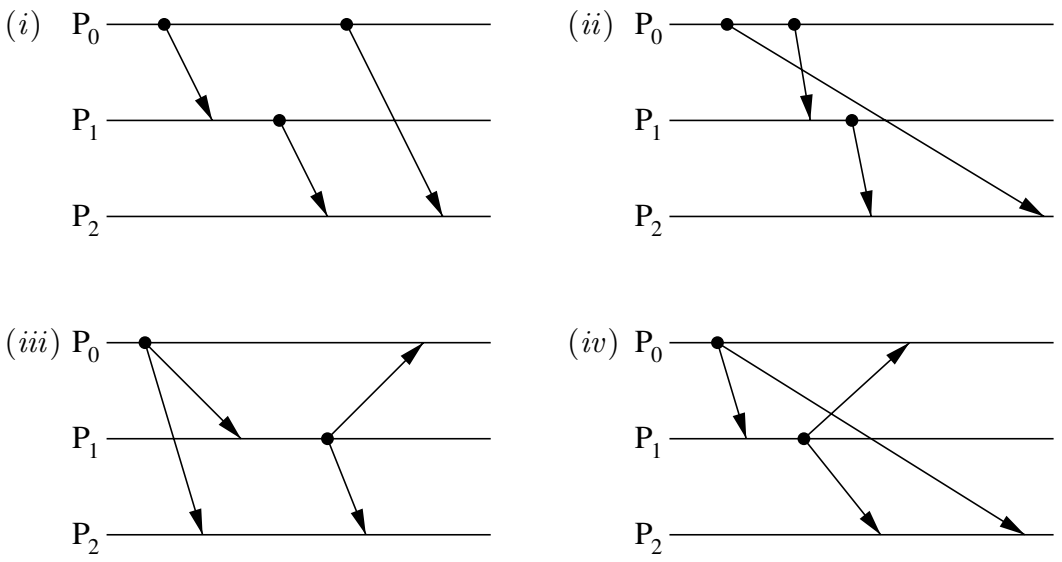
(*a*) Explain the problem of clock drift in distributed systems. [2 marks]

(*b*) What sources of conventional earth time might be used by computer systems? How would you estimate bounds on the accuracy of time received from such a source? [4 marks]

(*c*) What constraint does distributed inter-process communication (IPC) impose on the clock values of the communicating parties? [1 mark]

(*d*) Outline one clock synchronisation protocol that satisfies this constraint. [4 marks]

(*e*) For each of the cases of IPC illustrated below, give the vector clock values that message receiving and delivery modules could maintain for each process.



[6 marks]

(*f*) Define "causal order" of message delivery. In which, if any, of (*i*) to (*iv*) above is causal order violated at the message receiving module? [3 marks]

2 Computer Design

(*a*) What is a *data cache* and what properties of data access does it exploit? [5 marks]

(*b*) What is a *direct mapped cache* and under what conditions will it exhibit poor performance? [5 marks]

(*c*) Under what circumstances might a word of data in main memory be simultaneously held in two separate first-level cache lines? [5 marks]

(*d*) A *translation look aside buffer* is a specialised cache. What does it typically store and why is it often a factor of 1000 smaller than a data cache? [5 marks]

3 Digital Communication I

(*a*) Define the terms *circuit* and *packet* in the context of communication systems. [5 marks]

(*b*) What sort of guarantee does circuit switching provide? [5 marks]

(*c*) What advantages does packet switching provide over circuit switching? [5 marks]

(*d*) Which of *frequency division multiplexing*, *time division multiplexing* and *code division multiplexing* lend themselves to circuit switching? Which to packet switching? Explain why or why not in each case. [5 marks]

4 Computer Graphics and Image Processing

(*a*) Describe the z-buffer polygon scan conversion algorithm. [10 marks]

(*b*) In ray tracing, once we have determined where a ray strikes an object, the illumination at the intersection point can be calculated using the formula:

$$I = I_a k_a + \sum_i I_i k_d (\mathbf{L}_i \cdot \mathbf{N}) + \sum_i I_i k_s (\mathbf{R}_i \cdot \mathbf{V})^n$$

Explain what real effect each of the three terms is trying to model and explain what each of the following symbols means, within the context of this formula:

$$I,\ I_a,\ i,\ I_i,\ k_a,\ k_d,\ k_s, \mathbf{L}_i,\ \mathbf{N},\ \mathbf{R}_i,\ \mathbf{V},\ n$$

[10 marks]

5 Business Studies

(*a*) What is meant by the terms *supply curve* and *demand curve*, and what is significant about the point where they cross? [5 marks]

(*b*) What is meant by the terms *cost curve* and *break-even point*? Draw a diagram to show the relation between them and the demand curve. How would you attempt to establish a demand curve in practice? [5 marks]

(*c*) The cost and demand schedules for a particular product are given in the following table. What price should the manufacturer set? [5 marks]

Volume, k	Unit Cost	Unit Price
1	11.8	19
2	11.6	18
3	11.4	17
4	11.2	16
5	11.0	15
6	10.8	14
7	10.6	13
8	10.4	12
9	10.2	11
10	10.0	10

(*d*) Discuss the economic impact of the advent of the Internet. [5 marks]

6 Comparative Programming Languages

(*a*) Briefly explain the concept of coroutines as used in BCPL and outline the effect of the library functions `createco(f, size)`, `deleteco(ctpr)`, `callco(cptr, val)` and `cowait(val)`. [10 marks]

(*b*) Outline how you would design a coroutine to merge, in increasing order, two infinite streams of increasing integers supplied by two other coroutines. [5 marks]

(*c*) Briefly outline how you would implement an analogous merging mechanism in an object-oriented language, such as Java, that does not provide a coroutine mechanism. [5 marks]

7 Compiler Construction

(*a*) Describe one possible structure (e.g. ELF) of an object file. Illustrate your answer by considering the form of object file which might result from the following C program.

```
int a = 1, b = -1;
extern int g(int);
extern int c;
int f() { return g(a-b) + c; }
```

It is not necessary to consider the exact instruction sequence, just issues concerning its interaction with the object file format. [10 marks]

(*b*) Describe how a linker takes a sequence of such programs and produces an executable file. [4 marks]

(*c*) Compare and contrast *static* and *dynamic* linking in a system using your object file format. [6 marks]

8 Prolog for Artificial Intelligence

A weighted binary tree can be defined using compound terms in the following way. A node of the tree is represented by the term $n(V, L, R)$, where V stands for the value of the node, and L and R stand for the left and right branches, respectively. A terminal node has the R and L components instantiated to the null list.

Given an input tree T, write a Prolog program that constructs a tree of the same shape as T, but in which the value of each node has been set to the value of the maximum value node in T.

[Note: Maximum marks are available only for programs that perform this task in one recursive descent of the input tree, and which use no more than four clauses.]

[20 marks]

9 Databases

The environmental agency is setting up an SQL database to monitor long-term trends in the climate. Data are collected from observatories of a number of different kinds.

Flood risk is of particular concern. Each water authority measures river levels and rates of flow hourly at major points, and records reservoir levels daily.

In addition, the agency maintains weather stations both inland and at sea. These record precipitation (rainfall etc.), temperature, sunshine, air pressure and wind. Values of new precipitation, temperature, pressure, and wind speed and direction are taken hourly; gusts of over 60 m.p.h. are noted whenever they occur.

Maximum and minimum temperature and pressure, the total number of hours of sunshine and the total precipitation are recorded daily. Inland stations can be grouped by water authority.

By default these primary data will be relegated to archive after 2 years. Selected information is retained permanently in a data warehouse. This serves two purposes. First, it holds monthly summary data consisting of the maximum (and minimum as appropriate) day value for each statistic, together with the monthly totals of sunshine and precipitation. The warehouse also keeps detailed information relating to periods of extreme weather from the relevant observatories, with one or more keywords describing the nature of the incident (*flood*, *blizzard*, *hurricane* etc.) and an optional comment.

Write notes to assist in the design of the schema for the relational data warehouse, including any diagrams that you find helpful. Explain how your design will enable meteorologists to find relevant past records, noting any assumptions that you make about the nature of the data.

[You should not go into unnecessary detail about the structure of the primary database. You may assume that expert meteorologists will select the data for the warehouse.]

[20 marks]

10 Numerical Analysis II

(*a*) Taylor's theorem states that if $x \in [a, b]$ and $f \in C^{N+1}[a, b]$

$$f(x) = T_N(a) + \frac{1}{N!}\int_a^x f^{(N+1)}(t)(x-t)^N dt$$

where

$$T_N(a) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots + \frac{(x-a)^N}{N!}f^{(N)}(a).$$

Prove Taylor's theorem. [6 marks]

(*b*) Peano's theorem states that if a quadrature rule integrates polynomials of degree N exactly over an interval $[a, b]$ then the error in integrating $f \in C^{N+1}[a, b]$ can be expressed as

$$E(f) = \int_a^b f^{(N+1)}(t)K(t)\,dt$$

where

$$K(t) = \frac{1}{N!}E_x[(x-t)_+^N].$$

Explain the notation $E(f)$, E_x and $(x-t)_+^N$. [4 marks]

(*c*) Use Taylor's theorem to prove Peano's theorem. [8 marks]

(*d*) Under what additional condition may the simplified formula

$$E(f) = \frac{f^{(N+1)}(\xi)}{(N+1)!}E(x^{N+1})$$

be applied? [2 marks]

11 Introduction to Functional Programming

(a) Write a recursive definition of a function that appends two lists. [3 marks]

(b) Give a definition of a recursive datatype `sequence` that implements lazy lists. [3 marks]

(c) Write a function

```
applistq : 'a list -> 'a sequence -> 'a sequence
```

which, applied to a list l and a sequence s, produces a sequence s' which corresponds to the lazy list obtained by appending l to the front of s. [6 marks]

(d) Prove, by structural induction on lists, that your definition of `applistq` satisfies the following identity for any lists $l1$ and $l2$ and any sequence s. Make sure you give an accurate statement of the induction hypothesis.

```
applistq l1 (applistq l2 s) = applistq (l1@l2) s
```

[8 marks]

12 Computer Vision

Give *three* examples of problems in computer vision which are formally ill-posed. In each case explain how one or more of Hadamard's criteria for well-posed problems has failed to be satisfied. Illustrate how addition of ancillary constraints or assumptions, even metaphysical assumptions about how the world behaves, enable one to convert the ill-posed problem into a well-posed problem. Finally, discuss how the use of Bayesian priors can perform this function.

[20 marks]

13 Complexity Theory

(*a*) Show that any language that can be accepted by a nondeterministic machine in *time* $f(n)$ can also be decided by a deterministic machine in *space* $O(f(n))$. [4 marks]

(*b*) Show that any language that can be accepted by a nondeterministic machine in *space* $f(n)$ can also be decided by a deterministic machine in *time* $O(c^{(f(n)+\log n)})$, for some constant c. [6 marks]

(*c*) Explain what the above results tell us about the inclusion relationships among the complexity classes:

NL, co-NL, P, NP, PSPACE and NPSPACE

[4 marks]

(*d*) It has been proved that the graph reachability problem is in co-NL. What further inclusions can you derive among the above complexity classes using this fact? Explain your answer. [6 marks]

END OF PAPER