

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

Tuesday 4 June 2002 1.30 to 4.30

Paper 11 (Paper 2 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.*

**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 Digital Electronics

- (a) Find a minimal sum of products form for each of the following partially specified boolean functions. Each partially specified function, g_i , is specified by a function f_i which is true when g_i must be true and d_i which is true when g_i may be true or false (that is, d_i represents the “don’t cares”).

$$(i) \quad f_1 = yx\bar{w} + xyz + \bar{y}\bar{x}\bar{z}w + xwz$$

$$d_1 = \bar{x}z \quad [4 \text{ marks}]$$

$$(ii) \quad f_2 = y\bar{w}\bar{z} + \bar{w}\bar{x}\bar{z} + \bar{y}\bar{w}\bar{z}$$

$$d_2 = \bar{x}w\bar{z} \quad [4 \text{ marks}]$$

- (b) What is the maximum number of product terms in a minimal sum of products form of a function of n boolean variables? [2 marks]
- (c) How do “don’t cares” arise in practice and how may they be exploited? Are there any pitfalls in using them? Illustrate your answer with examples. [10 marks]

2 Foundations of Programming

Four of the many facilities provided in Java for writing threaded code are the methods `start()`, `run()`, `wait()` and `notifyAll()`. Which of these is inherited from class `Thread`? From where are the others inherited? [4 marks]

The following class illustrates a single-value buffer.

```
class Buffer
{ private int value;
  private boolean full = false;

  public synchronized void put(int i) throws InterruptedException
  { while (this.full)
    this.wait();
    this.value = i;
    this.full = true;
    this.notifyAll();
  }

  public synchronized int get() throws InterruptedException
  { while (!this.full)
    this.wait();
    this.full = false;
    this.notifyAll();
    return this.value;
  }
}
```

It is assumed that two threads are running, one of which incorporates a loop which puts a succession of values into the buffer via the method `put()` and the other of which incorporates a loop which gets a succession of values out of the buffer via the method `get()`.

Describe the first few times round the loops of each thread in a typical run. Your account should refer to the features of Java which enable the threads to operate satisfactorily and should distinguish between the terms *blocked on synchronized* and *blocked on wait*. [9 marks]

Write a test program which incorporates two threads that exercise class `Buffer` by calling `put()` and `get()` respectively. There is no need to copy out class `Buffer` itself. [7 marks]

3 Compiler Construction

The specification for a pocket-calculator-style programming language is as follows:

- Valid inputs consist either of an Expression followed by the `enter` button or of an Expression followed by `store` Identifier `enter` ;
 - Expressions consist of Numbers and Identifiers connected with the binary operators `+` , `×` and `↑` (in increasing binding power), with the unary operators `-` and `abs`, and possibly grouped with parentheses. Unary operators bind more strongly than `+` but weaker than `×` so that $-a + b$ means $(-a) + b$ but $-a \times b$ means $-(a \times b)$;
 - Numbers consist of a sequence of at least one digit, possibly interspersed with exactly one decimal point, and possibly followed by an exponential marker “e” followed by a signed integer, e.g. `6.023e+22`. Identifiers are sequences of lower-case letters.
- (a) Give a Context-Free Grammar (Type 2 in the Chomsky Hierarchy) for the set of valid input sequences using names beginning with an upper-case letter for non-terminals. It should be complete in that you should go as far as to define e.g.

`Letter ::= a | b | ... | z` [10 marks]

- (b) Indicate, giving brief reasoning, which non-terminals are appropriate to be processed using lexical analysis and which using syntax analysis proper. [5 marks]
- (c) Give yacc or CUP input describing those elements deemed in part (b) to be suitable for syntax analysis. You need not give “semantic actions”. [5 marks]

4 Data Structures and Algorithms

- (a) Carefully describe how Shellsort works and state an estimate of its efficiency using big O notation. [8 marks]
- (b) Robert Sedgewick suggests that a good sequence of separations used in the algorithm is $\dots, 121, 40, 13, 4, 1$. Explain why this is a good sequence. Under what circumstances would you recommend a sequence that approaches 1 more rapidly? [4 marks]
- (c) Describe how radix sort from the least significant end works and suggest a data structure that could be used in its implementation. [8 marks]

5 Operating System Foundations

- (a) Write brief notes on the following:
- (i) Interrupt-based I/O. [3 marks]
 - (ii) Direct Memory Access. [3 marks]
 - (iii) File access control in Unix. [3 marks]
 - (iv) Mounting file systems in Unix. [3 marks]
- (b) You are one of the design team of the new Xinu file system. As part of this design you have proposed an inode structure similar to that in “traditional” Unix. (Blocks are assumed to be 4KB, and file pointers are 4 bytes.)
- (i) Give details of your proposal for the inode structure. [6 marks]
 - (ii) Calculate the maximum file size, in blocks. [2 marks]

6 Mathematics for Computation Theory

Give a brief explanation of each of the following terms, where relevant giving an example that either fits the term or shows how it can fail to apply:

- (a) Well-founded relation;
- (b) Power-set;
- (c) Cartesian Product;
- (d) Reflexive relation;
- (e) Partial Order;
- (f) Lexicographic Order;
- (g) Disjoint Union;
- (h) Bijection;
- (i) Equivalence Relation;
- (j) Free Algebra.

[2 marks each]

7 Computation Theory

- (a) Your mathematician friend can prove to you that there are uncountably many functions from numbers to numbers, but does not know any computation theory. Explain to her what is meant by a *partial recursive* function and by a (total) *recursive* function. How would you convince her that there must exist functions that are not recursive? [12 marks]
- (b) What does it mean for a set of numbers $S \subseteq \mathbb{N}$ to be
 - (i) *decidable*; [1 mark]
 - (ii) *recursively enumerable* (r.e.)? [2 marks]
- (c) S is called *co-r.e.* if its complement $\{x \in \mathbb{N} \mid x \notin S\}$ is r.e. Show that S is decidable if it is both r.e. and co-r.e. (Any standard results about computable functions that you use should be clearly stated.) [5 marks]

8 Numerical Analysis I

- (a) Define *absolute error* and *relative error*. How are they related? How are absolute errors combined when two numbers are added together? How are relative errors combined when two numbers are multiplied together? [5 marks]
- (b) Explain the term *loss of significance* in terms of absolute error and relative error. [1 mark]
- (c) Writing δ_x for the relative error in x , what is the worst-case *relative error* in evaluating x^2 ? What is the worst-case *absolute error*? What is the worst-case *absolute error* in evaluating $x^2 - y^2$? [3 marks]
- (d) Let δ_s, δ_c be the relative errors in the values of $\sin \theta, \cos \theta$ respectively. Find the worst-case relative errors in evaluating each of the formulae

$$\begin{aligned}\sin 2\theta &= 2 \sin \theta \cos \theta \\ \cos 2\theta &= 2 \cos^2 \theta - 1\end{aligned}$$

For what values of $\cos \theta$ does the second formula display loss of significance? [6 marks]

- (e) Consider the evaluation of $x^2 + y^2 - z^2$ in two cases
- (i) $|y| \simeq |z|$, $|x|$ is very small,
- (ii) $|x| \simeq |y| \simeq |z|$, $|x|$ is not small.

Taking each case separately, can loss of significance occur? Explain your answers. [4 marks]

- (f) How would you compute $x^2 + y^2 - z^2$ to achieve greater accuracy, especially if guard digits were in use? [1 mark]

9 Computer Graphics and Image Processing

(a) With reference to the characteristics and performance of the human visual system, provide an estimate for each of the following. In each case you are expected to justify your estimate:

- (i) the maximum resolution required by a display device;
- (ii) the maximum number of distinct intensity levels required by a display device;
- (iii) the optimal number of dimensions required to represent colour;
- (iv) the maximum refresh rate required of a CRT monitor.

[5 marks]

(b) A programmer suggests three different implementations of a polygon drawing algorithm:

- (i) standard z -buffer;
- (ii) standard A -buffer with an 8×8 mask size;
- (iii) standard z -buffer at 8×8 normal resolution followed by averaging operation which produces a normal resolution image by finding the average value of each 8×8 block.

Compare the three implementations in terms of both execution speed and resulting image quality. Which would be the best implementation to use if the average polygon covers 50 pixels? Which would be the best implementation to use if the average polygon covers 2 pixels? Which would be the best implementation to use if the display resolution was equal to the answer that you gave to (a)(i) above? [7 marks]

(c) Show how to perform 2D rotation about an arbitrary point. Provide a matrix in homogeneous coordinates for each step in the operation. [2 marks]

(d) Show how to perform 3D rotation about an arbitrary axis. Again, give matrices in homogeneous coordinates for each step in the operation. [6 marks]

10 Natural Language Processing

An open-domain question-answering system should be able to answer questions such as the following by searching a database of documents and returning an appropriate text snippet. For example:

User: What debts did Quintex leave?

System 1: About \$1.4 billion (Australian)

extracted from:

Quintex Australia Ltd. and Quintex Ltd. together have debt of about \$1.4 billion (Australian), according to two analysts at Australian brokerage firms.

and:

System 2: Around ADollars 1.5bn (Pounds 680m)

from:

Quintex group collapsed yesterday. The failure left corporate debts of around ADollars 1.5bn (Pounds 680m) and additional personal debts.

- (a) What natural language processing techniques would be needed to build such a system? [8 marks]
- (b) What problems would arise as a consequence of the open-domain requirement? [4 marks]
- (c) What problems would arise selecting an appropriate text snippet from a matching sentence? [4 marks]
- (d) How feasible would it be to have the system reliably rank such answers (in this case preferring the second)? [4 marks]

11 Software Engineering I

- (a) Compare and contrast the problems experienced in developing the London Ambulance control system with those of the Cambridge University financial accounting system (CAPSA). [6 marks]
- (b) For *each* of these systems, describe software engineering techniques that might have been used to avoid the problems. Identify advantages and disadvantages of these techniques. [14 marks]

12 Software Engineering II

- (a) Briefly describe top-down refinement, listing both advantages and drawbacks. [4 marks]
- (b) How can we achieve efficiency without sacrificing reliability or programmer productivity? [3 marks]
- (c) How can we minimise the hazards associated with memory management? [3 marks]
- (d) In the Z specification language, what are Δ and Ξ schemas? [5 marks]
- (e) Contrast a loop invariant with the Z concept of invariant. [5 marks]

END OF PAPER