

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

Monday 31 May 2004 1.30 to 4.30

Paper 10 (Paper 1 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) What is DeMorgan's theorem? [4 marks]
- (b) What are *minterms*, *essential terms* and *prime implicants*? [6 marks]
- (c) What are the prime implicants, essential terms and minimum sum of products for $f = (a.b) \oplus (c + d)$? [6 marks]
- (d) Whenever $g = a.b$ is true, if we do not care what the output of f is, determine the new minimum sum of products form for f . [4 marks]

2 Computer Design

TLBs and caches are examples of content-addressable memories (CAMs).

- (a) What is the principal difference between a CAM and a RAM? [4 marks]
- (b) What is the difference between fully associative, set associative and direct mapped lookup? [6 marks]
- (c) Why are TLBs always much smaller than caches? [4 marks]
- (d) Which of the lookup mechanisms in part (b) is usually used for a TLB and why aren't the other mechanisms usually used? [6 marks]

3 Foundations of Programming

A Java program is being developed to assist with the processing of examination marks. A test program begins as follows:

```
public class Exam
{ private static Mark[] question = {new Mark(8),
                                     new Mark(),
                                     new Mark(6),
                                     ...
```

The program makes use of a `class Mark` which begins:

```
class Mark
{ public boolean attempted;
  public int score;
  ...
```

The `Mark` array `question` has one entry for each candidate so the length of the array indicates the total number of candidates.

An entry such as `new Mark(8)` sets up a `Mark` object whose data field `attempted` is set to `true` and whose data field `score` is set to 8, indicating that the candidate attempted the question and was awarded 8 marks. It may be assumed that every score is an integer in the range 0 to 10 inclusive.

An entry such as `new Mark()` sets up a `Mark` object whose data field `attempted` is set to `false` (and whose data field `score` is arbitrary) indicating that the candidate did not attempt the question.

- (a) Supply suitable constructors for `class Mark` [4 marks]
- (b) Write two methods `int getCount()` and `double getMean()` which, when handed the actual argument `question`, return the number of candidates who attempted the question and the mean mark achieved by those candidates respectively. If no candidates attempted the question, `getMean()` should return `-1d`. [9 marks]
- (c) Write a method `int[] getRank()` which begins:

```
private static int[] getRank(Mark[] q)
{ int[] rank = {0,0,0,0,0,0,0,0,0,0,0};
```

This method should return the `int` array `rank` with each element `rank[i]` set to the number of candidates who scored more than i marks. Note that if the maximum `score` is 9 then both `rank[9]` and `rank[10]` will be zero and `rank[8]` will be the number of candidates who scored 9. [7 marks]

4 Data Structures and Algorithms

- (a) Describe how the Lempel Ziv text compression algorithm works, illustrating your answer by deriving the sequence of numbers and corresponding bit patterns it would generate when applied to a string starting with the following 24 characters:

ABCDABCDABCDABCDABCDABCD . . .

You may assume that the initial table is of size 256 (containing bytes 0 to 255) and that the codes for “A”, “B”, “C” and “D” are 65, 66, 67 and 68, respectively. [12 marks]

- (b) Estimate how many bits the algorithm would use to encode a string consisting of 1000 repetitions of the character “A”. [8 marks]

5 Comparative Programming Languages

- (a) Discuss to what extent a programmer can expect a program that conforms to a standard to generate identical results when run under different conforming compilers on different machines. [6 marks]
- (b) ALGOL 60 provided call by value and call by name, Pascal provided call by value and call by reference, and ALGOL-W provided a variety of calling methods including call by result and call by value-result. Briefly describe the calling mechanisms just mentioned and discuss why most modern programming languages provide only call by value. [8 marks]
- (c) Discuss the reasons why languages such as Fortran, Algol and PL/I designed in 1950s and 1960s are less widely used than languages designed in the last 20 years. [6 marks]

6 Operating System Foundations

(a) Describe a scheduling algorithm with the following properties:

- favours I/O-intensive processes
- responds dynamically when processes change their behaviour: e.g. enter a compute-bound or I/O-intensive phase
- has acceptable context switching overhead
- avoids indefinite overlook (starvation) of a process

[7 marks]

(b) In order to carry out its functions, a filing system holds metadata on each stored object.

(i) What is this metadata likely to comprise? [6 marks]

(ii) Describe the directory service functions of a filing system, including how the metadata is used. [7 marks]

7 Numerical Analysis I

(a) For Single Precision in the *IEEE* binary floating-point standard (*IEEE* 754) the *precision* is defined as 24, and the *exponent* requires 8 bits of storage. With reference to *IEEE* Single Precision, explain the terms *exponent*, *significand*, *precision*, *sign bit*, *normalised number*, *denormal number*. [6 marks]

(b) Explain the term *hidden bit*. What are the values of the hidden bit for normalised and denormal numbers? How is the exponent stored and why? How are the exponent, significand and sign bit arranged in memory? [4 marks]

(c) Let x^* denote the floating-point representation of a number x . Define the terms *absolute error* (ε_x) and *relative error* (δ_x) in representing x . How are ε_x and δ_x related? Define *machine epsilon* (ε_m). [3 marks]

(d) Assume $\delta_x = \delta_y = \delta_z = \varepsilon_m$. Using worst-case analysis, estimate δ_{xy} , ε_{xy} . Find an expression for δ_w where $w = z - xy$. [4 marks]

(e) Working to 4 significant decimal digits only, compute w^* when $x^* = 2.018$, $y^* = 2.008$, $z^* = 4.058$. Given $\varepsilon_m \simeq 0.5 \times 10^{-3}$, how many significant decimal digits of w^* can be relied on? [3 marks]

8 Mathematics for Computation Theory

Let A, B, C be sets. Define the *Cartesian product* ($A \times B$) and the *disjoint union* ($A + B$). [3 marks]

Let $f \subseteq (A \times B), g \subseteq (B \times C)$ be relations between A and B , B and C respectively. Define the *inverse relation* f^{-1} between B and A and the *product relation* ($f \circ g$) between A and C . [3 marks]

What conditions must be satisfied for the relation f to be a function $f : A \rightarrow B$? [2 marks]

Write $(A \rightarrow B)$ for the set of all functions from A to B . If A, B are both *finite*, $|A| = a, |B| = b$, how many elements are there in $(A \times B)$, $(A + B)$, $(A \rightarrow B)$? [2 marks]

If f and f^{-1} are both functions, we say that f is a *bijection*, and we write $A \cong B$. If A, B are both finite and $f : A \rightarrow B$ is a bijection, prove that $a = b$. (\star) [2 marks]

Establish explicit bijections between the following pairs of sets:

(a) $A \rightarrow (B \times C)$, $(A \rightarrow B) \times (A \rightarrow C)$; [3 marks]

(b) $(A + B) \rightarrow C$, $(A \rightarrow C) \times (B \rightarrow C)$. [4 marks]

If A, B, C are all finite, verify that the cardinality condition (\star) above is satisfied in each case. [1 mark]

9 Computation Theory

(a) What is Turing's Thesis? [2 marks]

(b) Explain the action of a Turing machine that is specified by a quintuplet description. [4 marks]

(c) Define the *configuration* of a Turing machine at step t , and establish equations that specify the configuration of a k -symbol Turing machine at step $(t + 1)$ in terms of the configuration at the previous step t . [6 marks]

(d) Explain how you would use your equations to simulate a specific Turing machine by a register machine whose program encodes the quintuplet description. To what extent does this support Turing's Thesis? [Explicit program for a register machine is *not* required.] [8 marks]

10 Computer Graphics and Image Processing

Describe an algorithm for performing scan conversion of a set of 3D polygons, including details of clipping, projection, and the underlying 2D polygon scan conversion algorithm. You may assume that you are given the colour of each polygon and that no lighting calculations are required. Please state any additional assumptions that you need to make. *Ray tracing* is not an acceptable answer to this question. [20 marks]

11 Introduction to Security

- (a) Explain briefly mechanisms that software on a desktop computer can use to securely generate secret keys for use in cryptographic protocols. [5 marks]
- (b) Give *two* different ways of implementing residual information protection in an operating system and explain the threat addressed by each. [5 marks]
- (c) Consider the standard POSIX file-system access control mechanism:
- (i) Under which conditions can files and subdirectories be removed from a parent directory? [2 marks]
 - (ii) Many Unix variants implement an extension known as the “sticky bit”. What is its function? [2 marks]
 - (iii) On a POSIX system that lacks support for the “sticky bit”, how could you achieve an equivalent effect? [2 marks]
- (d) VerySafe Ltd offer two vaults with electronic locks. They open only after the correct decimal code has been entered. The VS100 – a low-cost civilian model – expects a 6-digit code. After all six digits have been entered, it will either open or will signal that the code was wrong and ask for another try. The VS110 – a far more expensive government version – expects a 40-digit code. Users of a beta-test version of the VS110 complained about the difficulty of entering such a long code correctly. The manufacturer therefore made a last-minute modification. After every five digits, the VS110 now either confirms that the code has been entered correctly so far, or it asks for the previous five digits again. Compare the security of the VS100 and VS110. [4 marks]

12 Software Engineering and Design

Consider the following four documents that would have improved projects such as the Cambridge University CAPSA project or the London Ambulance Service system:

- (a) a plan for managing development risk;
- (b) a safety and reliability specification;
- (c) a user requirements document;
- (d) a test specification.

Describe *each* of these four documents, including how you would gather the data to write the document; the main sections of the document; who must read that document; who should approve the document; and how the CAPSA and London Ambulance Service projects differ in the effect this document would have had.

[5 marks each]

END OF PAPER