COMPUTER SCIENCE TRIPOS Part IA

Monday 31 May 1999 1.30 to 4.30

Paper 1

Answer **two** *questions from Section A, and* **one** *question from* **each** *of Sections B, C, D and E.*
Submit the answers in six **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.*

SECTION A

1 Foundations of Computer Science

A *permutation* of a list is any list that can be derived from it by re-arranging the elements. Write an ML function that returns the list of all the permutations of its argument. Explain your code clearly and carefully.

For example, applied to the list [1,2,3], your function should return the list whose elements are [1,2,3], [2,1,3], [2,3,1], [1,3,2], [3,1,2] and [3,2,1].

You may assume that the elements of the argument are distinct. The elements of the result may appear in any order.

[10 marks]

2 Discrete Mathematics

Let $M_n = 2^n - 1$ be the n^{th} Mersenne number.

Show that M_n can be prime only if n is. [5 marks]

Let $\Delta_m = m(m+1)/2$ be the m^{th} triangular number and recall that a perfect number is one equal to the sum of its factors (including 1 but excluding the number itself).

Suppose that $p = M_n$ is prime. Show that Δ_p is a perfect number. [5 marks]

3 Programming in Java

In a large project it is possible that two programmers, working on different parts of the code, happen to select the same name for something that they define. If this accident leads to some third programmer accessing the wrong one of these two versions the results could be unsatisfactory. Explain the steps that the designers of the Java language have taken to reduce the probability of such clashes. [10 marks]

4 Operating Systems

An operating system uses a single queue round-robin scheduling algorithm for all processes. You are told that a *quantum* of three time units is used.

What can you infer about the scheduling algorithm? [1 mark]

Why is this sort of algorithm suitable for a multi-user operating system? [1 mark]

The following processes are to be scheduled by the operating system.

Process	Creation Time	Required Computing Time
P_1	0	9
P_2	1	4
P_3	7	2

None of the processes ever blocks. New processes are added to the tail of the queue and do not disrupt the currently running process. Assuming context switches are instantaneous, determine the *response time* for each process. [6 marks]

Give one advantage and one disadvantage of using a small quantum. [2 marks]

SECTION B

5 Foundations of Computer Science

"We face the Year 2000 crisis because programmers did not apply the principles of data abstraction." Discuss. [4 marks]

Your employer asks you to implement a dictionary. The pattern of usage will consist of taking an empty dictionary and performing many insertions and lookups. You must choose one of three data structures. Each requires $O(\log n)$ time for the lookup and update operations, where n is the number of items in the dictionary. They are (1) binary search trees, which take $O(\log n)$ time in the average case; (2) balanced trees, which need complicated algorithms but take $O(\log n)$ time in the worst case; (3) self-adjusting trees, which take $O(\log n)$ amortised time in the worst case.

Explain the differences between the three notions of $O(\log n)$ time. Argue that any of the three data structures might turn out to be the best, depending upon further details of the application. If no further details are available, which of the three is the safest choice? [8 marks]

An algorithm requires $T(n)$ units of space given an input of size n, where the function T satisfies the recurrence

$$\begin{aligned} T(1) &= 1 \\ T(n) &= T(n/2) + n \qquad (n > 1). \end{aligned}$$

Express the algorithm's space requirement using O-notation, carefully justifying your answer. [8 marks]

6 Foundations of Computer Science

Describe ML's facilities for treating functions as data, giving examples of their use in programs. Illustrate your answer by discussing the function `foldr`:

```
fun foldr f ([], e) = e
  | foldr f (x::xs, e) = f(x, foldr f (xs,e));
```

[7 marks]

You have been asked to implement a data structure to represent family relationships. For each person, it should record his or her name, mother, father, and children. As a first attempt, you have been given the following `datatype` declaration:

```
datatype person = Person of string * person * person * person list;
```

Identify two problems with this declaration that make it unusable. Modify the declaration to correct these problems. [6 marks]

Consider the following, simpler data structure for associating a person with his or her children:

```
datatype famtree = B of string * famtree list;
```

Write an ML function that takes two arguments: a predicate P over family trees (a function of type `famtree->bool`) and a family tree t. The result should be the list of all subtrees of t (possibly including t itself) satisfying the predicate. For full credit, give due attention to efficiency. [7 marks]

SECTION C

7 Discrete Mathematics

Define Euler's totient function $\phi(n)$. [2 marks]

Prove the Fermat–Euler Theorem that $a^{\phi(n)} \equiv 1 (\mathrm{mod}\ n)$ for appropriate a. [8 marks]

Deduce a theorem of Fermat about $a^{p-1} - 1$ for a prime number p. [2 marks]

Given a prime, p, with $p \neq 2$ and $p \neq 5$, show that there are infinitely many natural numbers, each of which has 9s as all its digits and which is divisible by p. [8 marks]

8 Discrete Mathematics

Let Ω be a universal set and define a relation between subsets A, $B \subseteq \Omega$ by $A \cong B \Leftrightarrow \exists$ a bijection $f : A \to B$. Prove carefully that $\cong$ is an equivalence relation. [6 marks]

What does it mean to say that a set is *countable*? [2 marks]

State without proof the Schröder–Bernstein theorem concerning the existence of a bijection between two sets. [2 marks]

Show that the integers and the rational numbers are countable but that the real numbers are uncountable. [6 marks]

An ML program consists of a finite sequence of characters drawn from a finite alphabet. Show that the set of ML programs is countable. [4 marks]

SECTION D

9 Programming in Java

Explain how to set up a new sort of Java Exception called (say) "MyException". Describe the syntax associated both with causing such exceptions to be activated and with responding to occurrences of them. Do any special issues arise if an exception is to be used as a way of exiting through multiple layers of procedures, for instance if the exception will be activated from deep within a nest of calls of a pair of mutually recursive methods? [10 marks]

There are Java classes (each with a constructor) called `File`, `FileInputStream`, `InputStreamReader`. In the last of these there is a method `read()` which returns the integer code of the next Unicode character from the stream it represents. All the IO functions can raise some sub-class of `IOException`. Write code that counts the number of characters and lines in a file and displays those to `System.out`. [10 marks]

10 Programming in Java

In the Discrete Mathematics course you learned that RSA encryption involved having a public key (N, e) where N is the product of two secret primes P and Q and e is an exponent. To encrypt a message that is represented by a number m you just compute $m^e \bmod N$.

The Java `BigInteger` class contains (among others) methods called `add`, `subtract`, `multiply`, `divide` and `remainder`.

The class `String` has a method `charAt` that allows you to extract a character at a given position, and `length` to tell you how long the string is. Casting a character to an integer yields its character code.

Supposing you are given a `BigInteger` that represents N and an integer for e, and not using any built-in Java methods for raising numbers to powers, write code that

(a) takes a string and encodes it as an integer; if the string contains characters c_0, $c_1 \ldots$ the integer required will be $c_0 + Kc_1 + K^2c_2 + \cdots$ with the constant K set to 2^{16} so that the full Unicode character set can be accommodated; [7 marks]

(b) encodes this number (assuming it is less than N) using the RSA method; [7 marks]

(c) creates an encoded string by viewing the integer as if it was written $d_0 + Ld_1 + L^2d_2 + \cdots$ with $L = 26$ and then representing each d_1 as a lower-case letter so that the 26 possible values are all accounted for. [6 marks]

SECTION E

11 Operating Systems

Most operating systems provide each process with its own *address space* by providing a level of indirection between virtual and physical addresses.

Give *three* benefits of this approach. [6 marks]

Are there any drawbacks? Justify your answer. [2 marks]

A processor may support a *paged* or a *segmented* virtual address space.

(*a*) Sketch the format of a virtual address in each of these cases, and explain using a diagram how this address is translated to a physical one. [8 marks]

(*b*) In which case is physical memory allocation easier? Justify your answer. [2 marks]

(*c*) Give *two* benefits of the segmented approach. [2 marks]

12 Operating Systems

File systems comprise a *directory service* and a *storage service*.

What are the two main functions of the directory service? [2 marks]

What is a directory *hierarchy*? Explain your answer with the aid of a diagram. [2 marks]

What information is held in file *meta-data*? [4 marks]

What is a *hard link*? Does file system support for hard links place any restrictions on the location of file meta-data? [2 marks]

What is a *soft* (or *symbolic*) link? Does file system support for soft links place any restrictions on the location of file meta-data? [2 marks]

Describe with the aid of a diagram a Unix *inode*. You should explain the motivation behind *indirect blocks*, and how they are used when accessing a file. [8 marks]