COMPUTER SCIENCE TRIPOS Part Ia

Monday 2 June 2008   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.*

> **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**

STATIONERY REQUIREMENTS
Script paper
Blue cover sheets
Tags

SPECIAL REQUIREMENTS
None

SECTION A

1 Foundations of Computer Science

(*a*) Write brief notes on the type-checking of polymorphic functions that involve equality tests. What is the type of the equality function itself?   [3 marks]

(*b*) The list l_2 is a *sublist* of l provided there exist lists l_1 and l_3 such that $l = l_1@l_2@l_3$. Write an ML function `sublist` to test whether one list is a sublist of another one. Include a clear explanation of how your code works.

[7 marks]

2 Operating Systems

Let N be the 16-bit value $1001\,0101\,0000\,0000_2$.

(*a*) What is the value of N when interpreted as:

    (*i*) An unsigned integer?   [1 mark]

    (*ii*) A sign-and-magnitude format integer?   [1 mark]

    (*iii*) A 2's complement integer?   [2 marks]

    (*iv*) A floating-point number with a 5-bit bias-15 exponent and a normalised mantissa? [State any assumptions you make.]   [4 marks]

(*b*) Imagine N has been loaded into the 16-bit register `r1`. Explain what the values of the C (carry) and V (overflow) flags would be after the CPU executes the instruction `add r0 ← r1, r1`.   [2 marks]

3 Programming in Java

For each of the following pairs of concepts that are used in Java explain similarities and differences and discuss when you would use one rather than the other. If there are important syntactic differences, method names, degrees of generalisation possible or the like, then any code fragments you give in illustration should be as short as possible to make the desired point.

(*a*) `JApplet` and simple Java applications   [4 marks]

(*b*) `Vector<String>` and arrays of strings (`String []`)   [3 marks]

(*c*) `class`, `abstract class` and `interface`   [3 marks]

4 Algorithms

(*a*) Briefly describe the update operations supported by a *priority queue*.

[2 marks]

(*b*) Explain the data structure known as a *heap* and describe how a heap can be implemented using a simple linear block of memory.   [2 marks]

(*c*) Describe, and estimate the costs of, procedures to

(*i*)   find the parent and offspring of a given node;   [2 marks]

(*ii*) insert a new item into an existing heap;   [2 marks]

(*iii*) delete the topmost item from a non-empty heap.   [2 marks]

SECTION B

5 Foundations of Computer Science

(*a*) Describe how lazy lists, which have possibly infinite length, can be implemented in ML. Illustrate your answer by presenting a function that accepts one (or more) lazy lists and produces another lazy list.   [6 marks]

(*b*) A *lazy binary tree* either is empty or is a branch containing a label and two lazy binary trees, possibly to infinite depth. Present an ML datatype to represent lazy binary trees.   [2 marks]

(*c*) Present an ML function that produces a lazy binary tree whose labels include all the integers, including the negative integers.   [3 marks]

(*d*) Present an ML function that accepts a lazy binary tree and produces a lazy list that contains all of the tree's labels.   [9 marks]

All ML code must be explained clearly.

(TURN OVER)

6   Foundations of Computer Science

A puzzle, or one-person game, can be represented in ML by two functions:

- a next-state function, which maps a state to a list of possible next states, and

- a wins function, which returns true if the given state counts as a win.

A simple example is a puzzle that has states consisting of positive integers, a next-state function that maps n to $[n+2, n+5]$, and a "wins" function that returns true if $n = 10$. We can win if we start from $n = 2$ but not from $n = 7$.

(a) Code a polymorphic datatype `'a puzzle`, to represent a puzzle by the pair of a next-state function and a wins function.   [2 marks]

(b) Briefly contrast depth-first search, breadth-first search and iterative deepening as techniques for solving such puzzles.   [6 marks]

(c) Write a function `depth` that accepts a puzzle, a state and a depth limit. It should use depth-first search to determine whether the puzzle can be solved from the given state within the given depth limit.   [6 marks]

(d) Write a function `breadth` that accepts a puzzle and a state. It should use breadth-first search to determine whether the puzzle can be solved from the given state.   [6 marks]

All code must be explained clearly. You may assume that any necessary ML data structures or functions are available.

SECTION C

7   Operating Systems

(*a*) What is *file metadata*?   [2 marks]

(*b*) Explain with the aid of a diagram how file metadata is managed in:

    (*i*)   the Unix file-system;   [4 marks]

    (*ii*) the FAT32 file-system;   [4 marks]

    (*iii*) the NTFS file-system.   [4 marks]

(*c*) A researcher suggests using non-volatile flash memory to store the NTFS log file. He believes this will *improve performance*, *reduce power consumption* and make the system *more resilient to failure*. Is he right? Briefly justify your answer in each case.   [2 marks each]

8   Operating Systems

(*a*) In general we can consider a process to move between five *process states* during its existence. Discuss, with the aid of a diagram, the circumstances in which a process will enter or leave each of these states.   [6 marks]

(*b*) Both Unix and Windows NT use *dynamic priority scheduling*. Compare and contrast their scheduling algorithms, with a particular focus on how dynamic priorities are managed.   [8 marks]

(*c*) Which data structures would you use if implementing a dynamic priority scheduling algorithm? Justify your answer.   [2 marks]

(*d*) Some industry predictions suggest that in less than five years we shall have chips with hundreds or even thousands of CPUs on them. What kinds of scheduling algorithms do you think will be appropriate for such systems? What problems do you foresee?   [4 marks]

(TURN OVER)

SECTION D

9   Programming in Java

(*a*) Show how you would create a Java array of 1000 integer values.   [1 mark]

(*b*) The values in an array could be used to represent digits in the decimal
representation of a number.   For example, the number $\frac{1}{7}$ has decimal
representation $0.142857\ldots$ and that could be stored in an array whose elements
started $\{1, 4, 2, \ldots\}$. For a number stored that way write code that multiplies
the number by a given integer, returning the whole number part of the result
and leaving the array updated to hold the fractional part of the product.
[5 marks]

(*c*) To convert a fraction to a representation base 16 (i.e. hexadecimal) you can
multiply it by 16, and the resulting integer part is the first digit of the base-16
representation. Multiplying the fraction left over by 16 gets the second digit
and so on. Write a method that accepts an array of digits (base 10) and creates
and returns a new array representing the same fraction but now base 16. Your
code should work for any length input array, not just one of length 1000, and
you may make the output array have the same length as your input array.
[6 marks]

(*d*) Suppose the input to your method in part (*c*) was of length 1000 and started
off with the decimal digits of $\frac{1}{7}$ in it. Although the initial digits in the output
array are the correct hexadecimal representation of $\frac{1}{7}$ the last few end up
looking odd. Explain.   [3 marks]

(*e*) One way to ensure that numerical results are correct is to use *interval
arithmetic*. A value is represented as a pair of arrays, one representing a
number less than (or equal to) the true value and one a value greater than
it. So if using 6 decimal places the number $\frac{1}{7}$ would be held as a pair
$\{142857, 142858\}$. If the two final digits differ by at most 1 then the smaller
of them can be viewed as fully accurate. Using this idea, write code that
accepts a fraction in decimal form and returns a vector denoting the same
value in another base n (now no longer necessarily 16) such that all the digits
in the result vector are correct. Clarity in your code is to be preferred to
performance, but if you are aware of particular ways in which the code you
present is particularly inefficient, you should note and explain them.
[5 marks]

10 Programming in Java

The following is one of the examples from the Foundations of Computer Science course:

```
exception Change;
fun change (till, 0) = []
  | change ([], amt) = raise Change
  | change (c::till, amt) =
      if amt<0 then raise Change
      else (c :: change(c::till, amt-c))
          handle Change => change(till, amt);
```

(a) Define a Java class of your own (i.e. do not use any library class that you may be aware of) to represent linked lists of integers. Provide it with methods that can be used to reverse a list and to append two lists. Comment on whether your design has led you to make the methods for append and reverse static.

[7 marks]

(b) Prepare a method called `change` that works in the same way as the ML code shown above. Provide the class that you define it in with a `main` method that uses it to try to make change for 73p using 2p, 5p and 20p coins, printing the result neatly. [7 marks]

(c) If you have lists of non-zero integers such that all the values in them are less than 256 and you have at most 8 items in any list you can pack eight 8-bit fields into a single 64-bit "long". That gives a representation that some people might expect to be faster than using lists represented by chained up instances of a class. Re-work your change-giving code based on the above idea, and modelling the use of exceptions by making the change function return `-1L` in the exceptional case. [6 marks]

(TURN OVER)

SECTION E

11 Algorithms

(*a*) What is a binary search tree *rotation*, and how are rotations useful in the creation of efficient search tree algorithms?   [2 marks]

(*b*) Write pseudocode for a recursive function `select(x, i)` which, given a binary search tree with root node x, executes a sequence of rotations to move the ith largest node to the root of the tree and returns a pointer to the new root node.   [6 marks]

(*c*) Making use of `select` to deal with tricky cases, write pseudocode for a recursive function `delete(x, k)` which deletes the node containing key value k from the tree and returns a pointer to the new root node.   [5 marks]

(*d*) Write a more efficient version of `delete` which does not use recursion or rotation to perform its work. How does its time complexity compare with your answer to part (*c*)?   [7 marks]

12 Algorithms

(*a*) Carefully describe the purpose and structure of a skip list. Provide a pseudocode definition for a skip-list node.   [7 marks]

(*b*) Write pseudocode to find an integer in a skip list.   [7 marks]

(*c*) Describe *two* situations in which you would implement a skip list in preference to a hash table.   [6 marks]

END OF PAPER