One and a half hours

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

Algorithms and Imperative Programming

Date: Tuesday 13th January 2015

Time: 09:45 - 11:15

Please answer any TWO Questions from the THREE Questions provided

Use a SEPARATE answerbook for each QUESTION

This is a CLOSED book examination

The use of electronic calculators is permitted provided they are not programmable and do not store text

1. Algorithm design.

   For each of the computational tasks (a), (b) and (c) below,

   (i) describe an algorithm for the task. Your description may be a program in a standard language, in pseudocode, or a clear and precise step-by-step description. You should explain your algorithm and why it works. Marks are awarded for a correct algorithm. Some marks are also awarded for efficiency: the more efficient your algorithm is, the more marks it will be awarded.

   (ii) give the worst-case time complexity of your algorithm in terms of the size of the input and the number of operations required. You may use Big-Oh notation. Explain how you calculated your answer.

   a) The two words "street" and "tester" are *anagrams* - one is formed from the other by a rearrangement of its letters. Considering words as lists of letters, the task is to determine whether or not any two such lists are anagrams of each other. (6 marks)

   b) List symmetric difference: Given two lists of integers, compute a list of integers which consists of those integers which are in one or other of the two given lists, but *not* in both (the order of the result list does not matter and, if numbers appear several times in the lists, the symmetric difference need not reflect this multiplicity - though it may). For example, given lists $[2,5,3,8,2,4,7]$ and $[6,7,2,4,9,1]$, one possible symmetric difference list is $[5,3,8,6,9,1]$. (7 marks)

   c) Choosing k-items from a list of n distinct integers *at random and without repetition* (i.e. an item must not be chosen more than once). Assume you are given a function *random(m)* which chooses an integer at random between 1 and m. (7 marks)

2. a) What is the *worst case* time complexity of each of the following sorting algorithms in terms of the number of items N to be sorted? Give your answer in Big-Oh notation.

   (i) insertion sort,
   (ii) merge sort,
   (iii) bucket sort,
   (iiv) quick sort,
   (v) radix sort,
   (vi) selection sort.

   (3 marks)

   b) Carefully explain the meaning of each of the following terms

   (i) stable sorting
   (ii) in-place sorting
   (iii) distribution-based sorting
   (iv) radix sorting

   (7 marks)

   c) Comparison-based sorting algorithms can be used to sort numbers into an arbitrary order (not necessarily the standard numerical order) by writing a suitable compare(a,b) function for comparison of two integer arguments, a and b.

   i) Write pseudocode for such a compare(a, b) function that would result in sorting numbers so that exact multiples of 8 appeared first, exact multiples of 4 (but not 8) appeared next, and all other numbers appeared in normal ascending numerical order (assuming the compare function was called by a sorting algorithm sorting in ascending order). Note: The multiples of 8 should be all considered equal to each other. The multiples of 4 should also be considered equal to each other. The function should return 0 if the arguments are equal, -1 if a is less than b, and 1 if a greater than b. (7 marks)

   ii) If your compare function is called by merge sort during the final merge phase on the two lists of numbers below, what would the result be? Assume that the merge sort is stable, so that items which are treated as equal by the compare function are ordered so that those in the left list appear first.

   $$8,20,4,2,7,19 \qquad | \qquad 4,12,13,14,17,17$$

   (3 marks)

3. a) Look at the following pairs of functions. In each case state whether (asymptotically) $f(n)$ grows faster, $g(n)$ grows faster, or they are of the same order. (Remember: asymptotically means for sufficiently large n).

   For example, for the pair $f(n) = 5n$, and $g(n) = 10n$, the answer would be: THE SAME (since they are of the same *order*, both $O(n)$) .

   While, for the pair $f(n) = 100$, and $g(n) = n$, the answer would be $g(n)$ because $g(n)$ grows faster with n than $f(n)$ does. (Notice: $f(n)$ does not grow with n at all).

   NOTE: "grows faster" means has *higher* complexity.

   i) $f(n) = n^{1.5} + n$, and $g(n) = 6n^2 + 2$ (1 mark)
   ii) $f(n) = n \log_2 n$, and $g(n) = \sqrt{n}$ (1 mark)
   iii) $f(n) = 2^{\sqrt{n}}$, and $g(n) = 5n^{14}$ (1 mark)
   iv) $f(n) = \log \log n$, and $g(n) = \log n$ (1 mark)

b) Which of the following statements (A-E) are true about the four functions given below? In each case, indicate *all* the statements that apply.

   A. The function has exponential growth
   B. The function is constant, it does not depend on n
   C. The function is $O(n^2)$ (Big-Oh of n^2, it grows at the same order or more slowly than n^2)
   D. The function has linear complexity, it is $\Theta(n)$
   E. The function is $\Omega(n^5)$ (it grows at the same order or faster than n^5)

   i) $6n \log n$ (2 marks)
   ii) $3.73n + 44$ (2 marks)
   iii) 1.5^n (2 marks)
   iv) $3n^2 + 15n + 2$ (2 marks)

c) In the following cases, state which would be the most appropriate type of complexity analysis to perform: worst-case OR average-case OR amortized. State only ONE answer in each case and EXPLAIN your answer.

   i) A company has invented a new web search engine and wishes to analyse how quickly it returns results for a set of common search queries. (2 marks)
   ii) A cloud computing company hosting an algorithm for weather forecasting and needs to guarantee to compute the next national daily forecast from pressure and other observations data in under 4 hours. (2 marks)
   iii) A database is sorted the first time a query is made, if previously unsorted. We want to analyse how long a number of consecutive queries would take to perform using this database system. (2 marks)
   iv) A pilot is flying a plane and his inputs on the control stick are converted into wing surface movements by calculations made in software. The stability of the plane depends on fast responses; we want to analyse if the plane is safe. (2 marks)

END OF EXAMINATION