Ph.D. Qualify Exam., Subject: Algorithms

Four hours, Closed book.

1. Banks often record transactions on an account in order of the times of the transactions, but many people like to receive their bank statements with checks listed in order by check number. People usually write checks in order by check number, and merchant usually cash them with reasonable dispatch. The problem of converting time-of-transaction ordering to check-number ordering is therefore the problem of sorting almost-sorted input. Argue that the procedure INSERTION-SORT would tend to beat the procedure QUICKSORT on this problem. (10%)

2. Let L be a doubly linked list of length m stored in arrays *key, prev,* and *next* of length n. Suppose that these arrays are managed by ALLOCATE-OBJECT and FREE-OBJECT procedures that keep a doubly linked free list F. Suppose further that of the n items, exactly m are on list L and n-m are on the free list. Write a procedure COMPACTIFY-LIST (L,F) that, give the list L and tree list F, moves the items in L so that they occupy array positions 1, 2, ... , m and adjusts the free list F so that it remains correct , occupying array positions m + 1, m + 2, ... , n . The running time of your procedure should be $\theta(m)$, and it should use only a constant amount of extra space. Give a careful argument for the correctness of your procedure. (10%)

3. VLSI databases commonly represent an integrated circuit as a list of rectangles. Assume that each rectangle is rectilinearly oriented (sides parallel to the x- and y-axis) , so that a representation of a rectangle consists of its minimum and maximum x- and y-coordinates. Give an O (n lg n) -time algorithm to decide whether or not a set of rectangles so represented contains two rectangles that overlap. Your algorithm need not report all intersecting pairs, but it must report that an overlap exists if one rectangle entirely covers another, even if the boundary lines do not intersect. (*Hint:* Move a "sweep" line across the set of rectangles.) (10%)

4. Suppose that we have a set of activities to schedule among a large number of lecture halls. We wish to schedule all the activities using as few lecture halls as possible. Give an efficient greedy algorithm to determine which activity should use which lecture hall. (This is also known as the ***interval-graph coloring problem.*** We can create an interval graph whose vertices are the given activities

and whose edges connect incompatible activities. The smallest number of colors required to color every vertex so that no two adjacent vertices are given the same color corresponds to finding the fewest lecture halls needed to schedule all of the give activities.) (10%)

5. Suppose we have an optimal prefix code on a set $C = \{0, 1, \ldots\ldots, n - 1\}$ of characters and we wish to transmit this code using as few bits as possible. Show how to represent any optimal prefix code on C using only $2n - 1 + n\lceil \lg n \rceil$ bits. (*Hint:* Use $2n - 1$ bits to specify the structure of the tree, as discovered by a walk of the tree.) (10%)

6. There are two types of professional wrestlers : "good guys" and "bad guys." Between any pair of professional wrestlers, there may or may not be a rivalry. Suppose we have n professional wrestlers and we have a list of r pairs of wrestlers for which there are rivalries. Give an $O(n + r)$-time algorithm that determines whether it is possible to designate some of the wrestlers as good guy and the remainder as bad guys such that each rivalry is between a good guy and a bad guy. If it is possible to perform such a designation, your algorithm should produce it. (10%)

7. Professor Deaver claims that the algorithm for strongly connected components can be simplified by using the original (instead of the transpose) graph in the second depth-first search and scanning the vertices in order of *increasing* finishing times. Is the professor correct ? (10%)

8. Professor Toole proposes a new divide-and-conquer algorithm for computing minimum spanning trees, which goes as follows. Given a graph $G = (V, E)$, partition the set V of vertices into two sets V_1 and V_2 such that $|V_1|$ and $|V_2|$ differ by at most 1. Let E_1 be the sets of edges that are incident only on vertices in V_1, and let E_2 be the set of edges that are incident only on vertices in V_2. Recursively solve a minimum-spanning-tree problem on each of the two subgraphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$. Finally, select the minimum-weight edge in E that crosses the cut (V_1, V_2), and use this edge the resulting two minimum spanning trees into a single spanning tree. Either argue that the algorithm correctly computes a minimum spanning tree of G, or provide an example for which the algorithm fails. (10%)

9. Which of the following statements are true? why, why not? (10%)

(a) The lower bound of NP-complete problem is exponential if and only if NP $\neq$P is proved

(b) We have not been able to prove that efficient algorithms cannot exist for solving any NP-complete problem

10. Draw a state-transition diagram for a string-matching automaton for the pattern ababbabbababbababbabb over the alphabet $\Sigma = \{a, b\}$ (10%).