Dept. of Computer Science and Engineering, National Sun Yat-sen Univ.
Second Semester of 2004 PhD Qualifying Exam   Computer Algorithms

1. Seven symbols *A*, *B*, *C*, *D*, *E*, *F*, and *G* are possibly transmitted on a communication channel where the frequencies of the transmission of these symbols are 8, 9, 10, 12, 18, 20, and 23, respectively. Please design the Huffman codes for the seven symbols and calculate the average code length. (20%)
2. A planar point (x_1, y_1) dominates (x_2, y_2) if $x_1 > x_2$ and $y_1 > y_2$. A planar point is called a maximal point if no other point dominates it. The problem: given a set of n planar points, find the maximal point(s) of the set. A straightforward solution is to compare each pair of points. However, the time complexity is $O(n^2)$. Please design a more efficient algorithm (e.g. $O(n\log n)$ or lower) to solve the problem, and show the time complexity of your method. (20%)
3. The sum-of-subsets problem is defined as: given a set of positive integers $A = \{a_1, a_2, \ldots, a_n\}$ and a constant c, determine if there exists $A' \subseteq A$ such that $\sum_{a_i \in A'} a_i = c$? For example: let $A = \{7, 5, 19, 1, 12, 8, 14\}$. $A' = \{7, 14\}$ if $c = 21$ and there is no solution if $c = 11$. The partition problem is defined as: given a set of positive integers $A = \{a_1, a_2, \ldots, a_n\}$, determine if there exists $A' \subseteq A$ such that $\sum_{a_i \in A'} a_i = \sum_{a_i \in A-A'} a_i$? For example: if $A = \{3, 6, 1, 9, 4, 11\}$, $A' = \{3, 1, 9, 4\}$ or $\{6, 11\}$. Prove that the sum-of-subsets problem $\propto$ the partition problem. (20%)
4. Given an empty stack, we perform a sequence of operations $X_1, X_2, \ldots, X_m$ on the stack where X_i consists of w (≥ 0) pops (from the stack) and one push (into the stack) for each i with $1 \leq i \leq m$. Assume that the sequence of m operations can be performed successfully (i.e., no overflow or underflow occurs) and the time consumed by each push or pop is 1. If t_i denotes the time consumed by X_i, the average time per operation is $t_{average} = \frac{1}{m}\sum_{i=1}^{m} t_i$. Please show that $t_{average} \leq 2$ by amortized analysis of the push-and-pop sequence. (20%)
5. Entity *A* and entity *B* know a large prime x and they keep x confidential. Design a randomized algorithm *D* based on the quadratic residue problem for interactive proofs such that *B* can be convinced that *A* is the real *A* with probability $1\text{-}2^{-m}$ after *D* has been performed successfully where m is a positive integer. (20%)