CIS 435: Homework 5 (Due: April 6, 2004)

Solve problems 1-3 (Group 1) and either Problem 4 (Group 2) or Problem 5 (Group 3).

Problem 1. (8 points)
(a) Show that any comparison based sorting algorithm that sorts 6 keys will require at least 10 comparisons in the worst case. (2 points)
(b) Show how to sort any 5 keys with 7 comparisons in the worst case. (6 points)

Problem 2. (8 points)
Which algorithm between merge-sort and radix-sort/count-sort is asymptotically faster for each one of the following sorting problems. Justify your answer.
(a) Sorting of n keys, where each key takes values in the range $0 \ldots n^3 - 1$.
(b) Sorting of n keys, where each key takes values in the range $0 \ldots n^{3 \lg n} - 1$.
(c) Sorting of n keys, where each key takes values in the range $0 \ldots n^n - 1$.

Problem 3. (4 points)
For a set of $n = 2^m$ unordered keys we are interested in finding the 1-st, 2-nd, 4-th, ..., 2^{m-2}-th, 2^{m-1}-th order statistic. Give an algorithm whose running time is $O(n)$ in the worst case that finds all m statistics. **Hint: Be careful. If you use Select m times on the set of n keys, it takes $O(n)$ per statistic for a total of $O(nm) = O(n \lg n)$, which is too much.** (10 points)

Problem 4. (30 points)
(a) For n distinct numbers $x_1, x_2, \ldots, x_n$ with positive weights $w_1, w_2, \ldots, w_n$ such that $\sum_{i=1}^{n} w_i = 1$, the **weighted median** is the element x_k satisfying

$$\sum_{x_i < x_k} w_i \leq 1/2$$

$$\sum_{x_i > x_k} w_i \leq 1/2$$

i. Argue that the median of $x_1, x_2, \ldots, x_n$ is the weighted median of the x_i with weights $w_i = 1/n$ for $i = 1, \ldots n$. (4 points)
ii. Show how to compute the weighted median of n elements in $O(n \lg n)$ worst-case time. (6 points)
(b) Change quicksort in such a way that its worst case running time is $O(n \lg n)$. (You are only allowed to change the way the splitter is selected but you can not chnage quicksort into another algorithm). (5 points)
(c) Given a black box worst-case linear time median subroutine, give a simple, linear-time algorithm that solves the selection problem for an arbitrary order statistic. (8 points)
(d) Consider inserting keys 1, 4, 32, 18, 22, 25, into a hash table of length $m = 7$ using open addressing with the auxiliary hash function $h'(k) = k \bmod m$. Illustrate the result of inserting these keys using (i) linear probing with hash function $h(k, i) = (h'(k) + i) \bmod m$, (ii) quadratic probing with hash function $h(k, i) = (h'(k) + i^2) \bmod m$. (7 points)

Problem 5. (30 points)
Do the Programming Module that is outlined in the electronic handout available at the course web page
`http://www.cs.njit.edu/~alexg/courses/cis435/handouts/phw5.ps`
or
`http://www.cs.njit.edu/~alexg/courses/cis435/handouts/phw5.pdf`