CMPT 307-08-2 Assignment 7

(From lecture on June 17, 2008)

Deadline: June 24, 5:30pm

Problem 7.1. Write the algorithm for finding the minimum and maximum elements of the array simultaneously. Verify that the number of comparisons it needs is $\lceil 3n/2 \rceil - 2$.

Problem 7.2. In the algorithm **Select**, the input elements are divided into groups of 5. Will the algorithm work in linear time if they are divided into groups of

(a) 7?

(b) 3?

In both cases find the worst-case partitioning, and build the recurrence for the worst-case running time (as we did on the lecture). Show in case (a), $T(n)$ can be upperbounded by cn, and in case (b), $T(n)$ can be lowerbounded by $dn \log n$, for suitable constants c and d.

Problem 7.3. Let $X[1 \dots n]$ and $Y[1 \dots n]$ be two arrays, each containing n numbers already in sorted order. Assume for convenience, that all $2n$ numbers are distinct. Give an $\mathcal{O}(\log n)$ algorithm to find to find the i-th smallest element of all $2n$ elements in arrays X and Y.
Hint 1: binary search
Hint 2: there are $2i$ possibilities where the i-the smallest element can occur in X and Y.

Problem 7.4. Show that if $|U| > nm$, there is a subset of U of size n consisting of keys that all hash to the same slot, so that the worst-case searching time for hashing with chaining is $\Theta(n)$. Recall that m is the size of the hash table, i.e., there are m slots.

Problem 7.5. Suppose we use a hash function h to hash n distinct keys into an array T of length m. Assuming simple uniform hashing, what is the expected number of collisions, that is what is the expected number of elements of the set

$$\{\{k, l\} : \quad k \neq l \text{ and } h(k) = h(l)\}\,?$$

Hint: use random variables X_{ij} define on the lecture.

Problem 7.6. Assume that the keys are strings with each character having p bits. Assume that we choose $m = 2^p - 1$ in the division method. Show that if string x can be derived from string y by permuting its characters, then x and y hash to the same value.