CISC621 Algorithm Design and Analysis, Spring 2016
Homework set III, problems 7,8,9, due Thursday, April 14, 7:00pm

Check the "homework sheet" from the syllabus for general homework details. In particular, each homework solution is on an entirely separate (set of) sheet(s) of paper and is identified with your name(s). Do not staple solutions to two or more problems together. Submit in 201 Smith Hall directly to Fanchao Meng or place on shelf marked CISC 621.

7. [Individual Problem — CLRS problem 13.4, Treaps.] Answer parts (a) uniqueness, (b) expected height (note the relevance of Theorem 12.4), (c).Treap-Insert function..

8. [Individual Problem — Hash authentication] Hashing is done in public areas like client-server situations. If an adversary can introduce many keys hasning to the same hash table location, that can be a denial of service attack. Choosing hash functions at random from carefully designed sets of hash functions can avoid this problem.

   Let U be a set of possible keys. This is often called the *universe* of keys. Let H be a set of hash functions on U. H is called *universal* if, for every distinct pair $k, l \in U$, the number of hash functions $h \in H$ such that $h(k) = h(l)$ is at most $|H|/m$. (This is as defined in CLRS, section 11.3.3.)

   Let $M^2 = M \bigoplus M$ denote the set of all pairs (i,j) with $0 \leq i, j < m$. H is called *2-universal* if, for every distinct pair $k, l \in U$, for h chosen at random from H, the probability that $(h(k), h(l)) = (i, j)$ is equal for every pair $(i, j) \in M^2$,

   (a) Show that if a set of hash functions, H, is 2-universal, then it is universal.

   (b) Let U be the set of n-vectors whose entries are in $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$, where p is a prime number. Given n-vector $a \in U$ define

   $$h_a(x) = ((\sum_{i=1}^{n} a_i x_i) \bmod p) \bmod m, \{ \text{ dot product of a with x } \}$$

   Define $H = \{h_a | a \in U\}$. Show that H is universal, but not 2-universal.

   (c) Suppose we modify the class of hash functions slightly. Let $H' = \{h'_{ab}\}$, where $a \in U$ as before and b is in $\mathbb{Z}_p$,

   $$h'_{ab}(x) = ((b + \sum_{i=1}^{n} a_i x_i) \bmod p) \bmod m.$$

   Show that H' is 2-universal.

   This is CLRS problem 11.4. Please also absorb part (d) there, which indicates the significance of this problem. However, you are not required to submit anything concerning part (d).

9. [Group Problem — M&M]

   (a) [majority:] Given an array A of n elements, a value x is a majority element of A if more than half of the elements of A have the value x. Note that there can be at most one majority element in an array. Give a linear time algorithm for determining whether or not an array has a majority element. Assume that the only comparisons allowed between values are tests of equality. That is, you may not assume that an order relation exists between elements[1].

   (b) [mode:] The *frequency* of an element in an array is the number of times it occurs in the array. The *mode* of an array is an element with the largest frequency. Suppose that A is a sorted array of n numbers and it is known that m, the frequency of the mode, is at least $n^{1/3}$. Give and analyze an algorithm to compute the mode of A that runs in time o(n). Note that this is "little o." You may assume, if you wish, that there is no tie for the mode.

   Remark: For your algorithm, how much can you weaken the assumption that that $m \geq n^{1/3}$?

[1] Brassard and Bratley, Algorithmics, Theory and Practice, Prentice Hall, 1988