Problem Set 5, CS 5800 Spring 2017

Due: Mon, 3/27, 11AM

Problem 1. (Difficulty 4) You want to augment binary search trees to support RANK. For a key k, RANK(k) is the number of keys in the tree that are less than k.

1. At every node, in addition to the key you will store a value v so that you can answer RANK queries in time linear in the depth. Define what the value v is. Give the pseudocode for RANK.

2. Describe an algorithm for INSERT so that the values v are maintained correctly. INSERT should run in time linear in the depth of tree.

3. Explain how the values v should be updated during rotations.

Problem 2. (Difficulty 3) Prove the 2-hash claim in the case $t = 2$.

Assume that h is defined as the following, given some seed a. Let x_i and a_i represent the digits of a and x in base t, respectively. The result of the hashing function is:

$$h(x) = \sum x_i a_i \mod t$$

Problem 3. (Difficulty 3) Describe an algorithm that given an array $A[1..n]$ of integers determines if the integers are distinct. The algorithm should run in time $O(n)$ and space $O(n^2)$, and is allowed to give an incorrect answer with probability 1/1000.