Technische Universität München
Fakultät für Informatik
Lehrstuhl für Effiziente Algorithmen
Prof. Dr. Harald Räcke
Chintan Shah

Wintersemester 2014-15
Assignment 7
November 24, 2014

Efficient Algorithms and Datastructures I

Question 1 (10 Points)

An order-statistics tree is an augmented Binary Search Tree that supports the additional operations RANK(x), which returns the rank of x (i.e., the number of elements with keys less than or equal to x) and FINDBYRANK(k), which returns the kth smallest element of the tree.

Let $A[1, \cdots, n]$ be an array of n distinct numbers. If $i < j$ and $A[i] > A[j]$, then the pair (i, j) is called an inversion of A. Show how to use an order-statistics tree to count the number of inversions in A in time $O(n \log n)$.

Question 2 (10 Points)

The mean M of a set of k integers $\{x_1, x_2, \ldots, x_k\}$ is defined as

$$M = \frac{1}{k} \sum_{i=1}^{k} x_i.$$

We want to maintain a data structure D on a set of integers under the normal INIT, INSERT, DELETE, and FIND operations, as well as a MEAN operation, defined as follows:

1. INIT(D): Create an empty structure D.
2. INSERT(D, x): Insert x in D.
3. DELETE(D, x): Delete x from D.
4. FIND(D, x): Return pointer to x in D.
5. MEAN(D, a, b): Return the mean of the set consisting of elements x in D with $a \leq x \leq b$.

Describe how to modify a standard red-black tree in order to implement D, such that INIT is supported in $O(1)$ time and INSERT, DELETE, FIND, and MEAN are supported in $O(\log n)$ time.

Question 3 (10 Points)

In double hashing, if we use the hash function $h(k, i) = (h_1(k) + ih_2(k)) \bmod m$, show that when m and $h_2(k)$ have greatest common divisor $d \geq 1$ for some key k, then an unsuccessful search for key k examines $\frac{1}{d}$th of the hash table before returning to slot $h_1(k)$.

(*Note:* When $d = 1$, i.e. when m and $h_2(k)$ are relatively prime, the search may examine the entire hash table.)

Question 4 (Extra Question)
Let $U = \{0, \ldots, p-1\}$ for a prime p. For $x \in \mathbb{Z}_p$, define the hash function $h_{a,b}(x)$ as

$$h_{a,b}(x) = (ax + b \mod p) \mod n$$

Consider the class of hash functions

$$\mathcal{H} = \{h_{a,b} | a, b \in \mathbb{Z}_p\}$$

(a) Show that $\mathcal{H}$ is not universal.

(b) Show that $\mathcal{H}$ is $(1.1, 2)$ independent for p sufficiently large.

(c) Why would you not choose $\mathcal{H}$ as a class of hash functions?