Technische Universität München
Fakultät für Informatik
Lehrstuhl für Theoretische Informatik
Prof. Dr. Harald Räcke
Richard Stotz, Dennis Kraft

Winter term 2016 / 2017
Assignment 6
November 28, 2016

Efficient Algorithms and Data Structures I

*Deadline: December 5, 10:15 am in the **Efficient Algorithms** mailbox.*

Homework 1 (3 Points)

In an ancient book on squirrel mathematics, Alexander discovers Algorithm 1. The description has almost faded, but seems to be `squirrelnutco`, its subroutine being called `wicked`. The description also indicates that the algorithm takes a boolean array A and an integer k as input.

Algorithm 1: `Squirrelnutco`(A, k)

```
Algorithm squirrelnutco(A, k)
    for j = 0; j < k; j ← j + 1 do
        wicked(A)

Procedure wicked(A)
    i ← 0
    while i < A.length and A[i] == TRUE do
        A[i] ← FALSE
        i ← i + 1
    if i < A.length then
        A[i] ← TRUE
```

1. Describe briefly what the procedure `wicked` does.
2. Using a suitable potential function, show that the amortized running time of procedure `wicked` during an execution of `squirrelnutco` is $\mathcal{O}(1)$.

Homework 2 (5 Points)

Show how to maintain a dynamic set U of numbers that supports the operation MIN-DIFF, which gives the magnitude of difference of the two closest numbers in U. Make the operations INSERT, DELETE, SEARCH, and MIN-DIFF as efficient as possible, and analyze their running times.
Example: If the set $U = \{1, 50, 99, 1024, 1066, 2016\}$, then MIN-DIFF$(U)$ returns $1066 - 1024 = 42$, since 1066 and 1024 are the two closest numbers in U.

Homework 3 (6 Points)

The *Mean* MEAN and the *Mean Squared Error* MSE of a set X of k integers $\{x_1, x_2, \ldots, x_k\}$ are defined as

$$\text{MEAN}(X) = \frac{1}{k}\sum_{i=1}^{k} x_i \ ,$$

$$\text{MSE}(X) = \frac{1}{k}\sum_{i=1}^{k} (x_i - \text{MEAN}(X))^2 \ .$$

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:

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

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

2. Describe how your tree must be augmented further in order to support calculating the Mean Squared Error of a subset, i.e. MSE(D, a, b), the MSE of the set consisting of element $x \in D$ with $a \leq x \leq b$. The calculation of the MSE should take at most $\mathcal{O}(\log n)$ steps.

Homework 4 (6 Points)

App-developer Andy Arpeggio wants to store information about the users of his new app *Tremolo* in a $(a, 2a)$-tree (i.e. a B-tree). After an immense marketing campaign, his app has 1999999 users.
Andy's old computer however has limited main memory of 8192 bytes. Each tree node holds a pointer to the parent, pointers to the child nodes and the keys. All pointers and the keys are 8 bytes large.
Initially, the root resides in the main memory and cannot be evicted from main memory. The rest of the tree resides on a slow hard drive. A disk access returns a block of 4096 bytes. Andy's goal is to minimize the number of disk accesses.

1. Andy fist practices his knowledge on (a, b)-trees by inserting the keys 4,7,10,2,15,22 sequentially in a $(2, 4)$-tree. Draw the tree after each insert.

2. How should Andy choose the parameter a of the tree for his app? Explain your choice!

3. With your choice of a, what is the maximum number of disk accesses needed during the search for a key? Note that Andy uses an external search tree.

Tutorial Exercise 1

Prove that there exists a sequence of n insert and delete operations on a $(2,3)$-tree s.t. the total number of split and merge operations performed is $\Omega(n \log n)$.

```
The more you think about what the B
could mean, the more you learn about
B-Trees, and that is good.
                          - R. Bayer
```