University of Toronto

CSCB70S—Data Structures and Techniques, Spring 1999

Assignment 3: Simulation

Due: Monday March 22 at 10:00 am.
Worth: 15%.

The problem

The purpose of this assignment is to teach you how to simulate a complex system in C++. It will also help you to get comfortable with C++.

You are to simulate operations on elements of binary trees. We will consider only two representations for the trees: "Binary Search Trees" (BST) and "Balanced Binary Trees" (BBT). A BST is a tree with the property that at any node x, the keys on the left subtree hanging from x are smaller than the key value on node x, and the keys on the right subtree hanging from x are larger than the key value on node x. A BBT is a tree with the property that all leaves are at the last two levels of the tree, and all the leaves in the last level are as to the left as possible.

We have the following operations executed on a tree represented as a BST or as a BBT: insert a key, find a key, and delete a key. You have to write a program that simulates the operations on a tree as it receives a series of requests for insertion, retrieval and deletion of items.

Events and probabilities

The events in our simulation are operations on keys: `insertion`, some number of `find`, and a `deletion`. Each item contains a key; this key is an integer number. The keys are "randomly generated", that is, both the value of the key and the time they first appear are determined randomly. Insertions happen when a key first appears in the simulation, and each insertion schedules the appearance of the next key. For that task, after processing the insertion, a new key is created. The value of the key follows a uniform distribution on the interval a, b, while the arrival of the key follow an exponential distribution with rate s, where a, b and s are parameters of the simulation (same parameters for all keys). The insertion also determines the probability p of a "find" operation for the key ($1 - p$ is the probability of operation "delete" the key); this is done with a uniform distribution in $0, 1$.

In addition to these tasks associated with the next key, each operation on a key schedules the next operation on the same key. Thus, the original insertion of the key schedules two events: the insertion of the next key, and the following retrieval or deletion of the key now being inserted. For an individual key, operations occur with an exponential distribution with rate r (again, a parameter of the simulation being the same for all keys).

The operations `find` and `delete` do not schedule new events for other keys. An event `find` processes the search for the element in the tree and schedules the next operation for the key (another `find` or a `delete`). This is done by picking a number q in $0, 1$ uniform at random. If $q > p$, the next operation for the key is a `deletion`; otherwise, it is a `find`. An event `delete` only processes the removal of the element from the tree (no event is scheduled).

In addition to the previous events there are "pseudo-events". There are three pseudo-events: beginning of simulation, end of simulation and reporting. The beginning of simulation schedules the appearance of the first key (its time, value, and probability of event `find`). Moreover, it schedules the end of the simulation using another parameter t of the simulation. It also schedule k reports evenly spaced in time (k is a parameter of the simulation). The report events simply print statistical information, and the end of simulation terminates the execution.

You should vary the parameters of the experiment so as to see how the statistics vary. Your report should include a discussion on the results and different situations where one representation is better than the other.

Programming

Make a directory in your account:

```
% cd
% mkdir hwk3
% cd hwk3
```

and copy the skeleton file into it from:

```
http://www.cs.toronto.edu/~daniel/teaching/B70
```

You have to write a program that simulates the operations on a tree as it receives a series of requests for insertion, find and deletion of keys. The algorithms for insertion, find and deletion for both representations are in the skeleton file. You have to concentrate your efforts in the simulation part of the assignment (manipulation of events queue, generation of variables following several probability distributions, etc). Read the README file carefully: it indicates what you have to do.
You will probably find useful to base your work on the simulation example in the course notes.
Give good comments in your code, including a comment above each of the functions you write that describes the purpose of the function, the meaning of its parameters and its output, and any relevant pre- and post-conditions.

Tests and report

With your program, you must submit tests and a report. Your test data will be of two kinds. The first kind establishes correctness in the usual way. The second kind obtains experimental data. Vary the parameters of the experiment and see how the results change. You may decide for yourself what statistics to collect.
Your report should describe your program and explain your experimental results. It should be brief and should aid the marker in understanding your work and the meaning of the results.

How to hand in your assignment

Hand in the following:

- the cover page (included below); a hardcopy printout of all your programs;
- a hardcopy printout of the output of your program; and the report.

Put this into an ENVELOPE. Write the following on the outside of the envelope:

- your name (with surname underlined) and student number;
- your account name; and your tutorial section.

Your programs must also be submitted electronically by the deadline. You must submit all your files for assignment 3 by executing `exactly the following commands` from the directory that contains the source files on your account:

```
% make clean
% submit -N a3 cscb70s *
```

You will lose marks if you don't follow these submission instructions exactly.

COVER SHEET FOR ASSIGNMENT 3

Account name:

Surname:

Given Name(s):

Student Number:

Tutorial:

I declare that this assignment is my own work and is submitted in accordance with the University of Toronto Code of Behavior on Academic Matters, the Code of Student Conduct, and the guidelines for avoiding plagiarism described in the course syllabus.

I discussed this assignment with the following people:

Signature

GRADING

Program		
Operations with the events queue	______	/ 20
Probability distributions	______	/ 20
Style (including comments)	______	/ 15
Thoroughness of your own tests	______	/ 10
Report		
Content	______	/ 25
Style and readability	______	/ 10
Total	______	/ 100