CS2302 Data Structures
Fall 2012
Lab 5
Hash Tables

Deadlines: Code - Friday, November 2,
Report - Monday, November 5,
Demo: Tuesday, November 6, 8:00-12:00

Instructions

In this lab you will compare the efficiency of binary search trees and hash tables with chaining for search and storage.

You will generate semi-random sequences of integers and then determine the integer that appears the largest number of times in the sequence, breaking ties arbitrarily.

Each experiment to be performed is explained by the following pseudocode:

For i = 0 to n-1
  • generate random number k in the 0 to 50,000 range
  • search for k in the data structure (tree or hash table)
  • if k is present (that is, k has appeared before in the sequence), increase its counter by one
    else create a new counter for k, initialize it to 1 and insert object containing k and its counter to data structure

Find the element with the largest occurrence counter value and display the element and its number of occurrences

You may modify the code provided in the class webpage to generate your random numbers. Perform experiments with different values of n (for example, 1000, 10000, 100000, and 1000000) and compare the times it takes to process the arrays with each of the implementations. Use the hash function \( h(k) = k \mod 2311 \).

You are not allowed to use the built-in implementations of these data structures; you have to implement your own. As usual, write a report explaining your work, as described in the syllabus.