Ian Riley
03/24/16
CS 2123 - Girgis

Performance Report Homework 2

The theoretical performance, for n insertions, of a Binary Search Tree is O(nlog(n)), in the best case, and O(n^2), in the worst case. The input data sets are all around 10^4 to 10^5 in magnitude. My processor performs roughly 10^9 instructions per second. For input data sets of this magnitude, we should see each data set being executed somewhere in the ballpark of milliseconds. If we look to the empirical performance of each data set with the Binary Search Tree, we can see that the slowest trial, the processing of Huckleberry Finn which consists of roughly 116,954 words, took around 773 milliseconds. This shows that the empirical results gathered from our trials is consistent with the expected theoretical results.

The theoretical performance, for n insertions, of an AVL tree is O(nlog(n)) in the best case and in the worst case. Considering our input sizes and the approximate processing speed of my machine, we should see each data set being executed in roughly the same time as the Binary Search Tree. That is, each trial should have a time measured in milliseconds. If we look to the slowest trial, Huckleberry Finn once again, we can see that this trial took roughly 806 milliseconds which is consistent with our expectations.

The advantage of AVL trees over Binary Search Trees is that the rotations of an AVL tree allow us to avoid the worst case running time of a Binary Search Tree. Because of this, we would expect that the AVL tree would run slightly faster than the Binary Search Tree. From the results that I have gathered in these few trials, this expectation was not met. In all of the trials, the Binary Search Tree ran slightly faster than the AVL tree. I believe that this is most likely due to the input sizes. The largest data set, Huckleberry Finn, is only 10^5 in magnitude. This is a fairly small data set compared to the average speed of the modern processor. To properly test the performance advantages of AVL trees over BST, we should use data sets that measure in the millions, if not billions.