EXPERIMENTATION

Algorithms & Data Structures

SEPTEMBER 8, 2018
Student's Name:Panagiotis Goulas
Student's ID: 886001

Method

The BST – Binary Search Tree is a node based binary tree data structure which allows fast search, addition and removal of items.
In the assignment, in the first stage, we read the csv file and save the data at the time. We create the binary search tree and we try to find a place to insert through it. We search for the letter and we go left if smaller and right if larger. In Stage 2 we implement a linked list and thus the search becomes more efficient. We use a data structure for our BST.

Time Complexity

We cannot talk about efficiency without talking also about time complexity. Sorted data always take less time in theory. In stage 1 the number of nodes is equal to the number of rows. The cost to search is $O(\log(n))$ as well as the cost to insert, where n is the number of nodes. In stage 2 with the implementation of the linked list we have a cost of $O(\log(m))$ with m being the names. Of course $n > m$ and thus $O(\log n) > O(\log m)$ and therefore we conclude that Stage 2 according to the theory will take less time and will indeed be more efficient.

Experimentation

Using time function in the command line we run stage 1 and stage 2 separately and we measure how much time they need to run into completion. We use the file: athlete events filtered. We randomly choose 10 names and we run stage 1 and stage 2. Our data are shown in the following table. Stage 1 on the left side and Stage 2 on the right side. Position's Name and time in seconds are used for our measurements.

Stage 1		Stage 2	
Position	Time(s)	Position	Time(s)
1	1.237	1	0.979
27001	1.237	27001	0.993
54001	1.163	54001	0.983
81001	1.288	81001	1.169
108001	1.287	108001	1.047
135001	1.293	135001	1.033
162001	1.281	162001	1.025
189001	1.266	189001	1.234
216001	1.197	216001	0.966
243001	1.163	243001	1.094

The result is something that we predicted from the theory already. Stage 2 can be completed in less time. It's more efficient and our data shows that as well.

And their bar charts are following below: