Data Structures and Algorithms

Homework # 4

Eren Erdogan

Question 1:

Please run code, refer to instructions on how to run on Sakai. Also please view testing results txt file in folder to see test input cases and results.

No discussion was asked for this question.

Question 2:

Please run code, refer to instructions on how to run on Sakai. Also please view testing results txt file in folder to see test input cases and results.

The 2-3 and LLRB BST tree structurally give the same result but are built differently then one another. However they are so similar that a 2-3 tree can easily represented as an equivalent LLRB BST and vice versa.

Question 3:

Please run code, refer to instructions on how to run on Sakai. Also please view testing results txt file in folder to see test input cases and results.

For this question we were asked to modify our same Red-Black Tree from question 2 but to keep track of a) the cost of each insert operation during the computation and b) the number of rotations and node splits that are used to build the trees.

First we'll address the cost.

a.)

For this part we were asked to define what we used to mean cost. For my data I took the cost as the time it took for each insert operation. So for example if I were inserting 1,000,000 random data points into my Red-Black BST and it took 1.079 seconds total then the average cost per insertion would be,

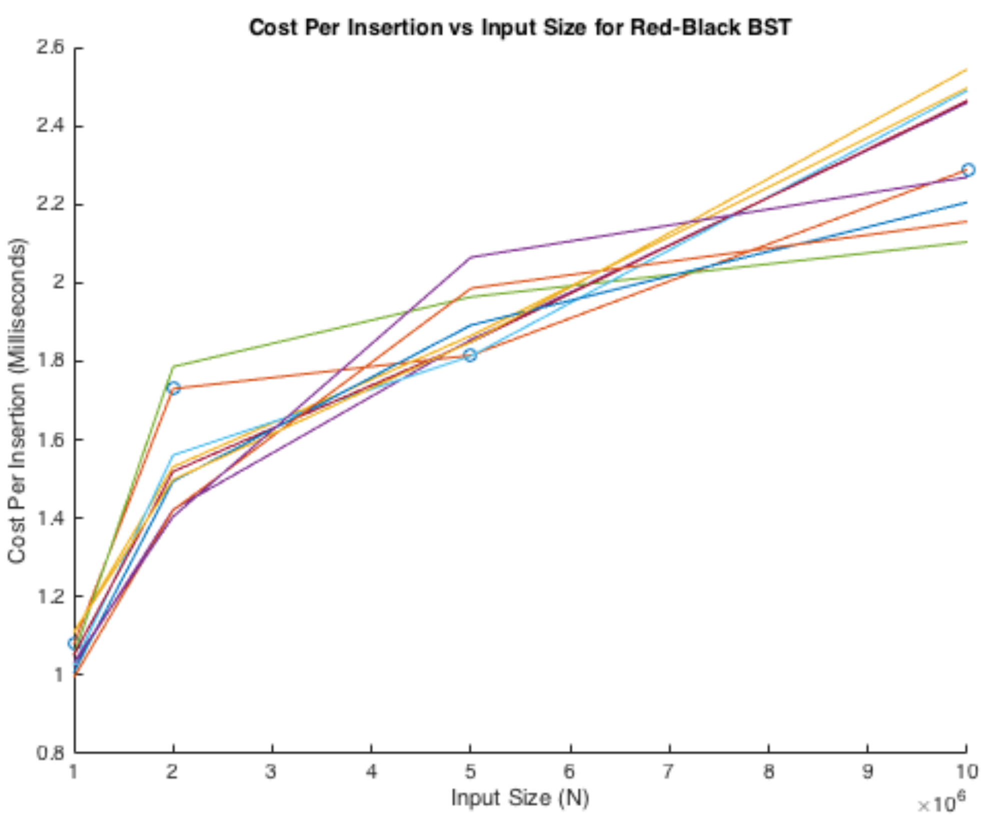
$$Cost\ Per\ Insertion = \frac{Time\ To\ Finish}{Total\ Input\ Size} = \frac{1.079\ Seconds}{1{,}000{,}000} = 1.079\ Milliseconds$$

We were also asked to implement this with different size inputs and 10 total trials. The results from the testing can be seen in the table on the next page.

Cost Per Insertion for 10 Trials (Seconds)				
	Input Size			
Trial #	**1,000,000**	**2,000,000**	**5,000,000**	**10,000,000**
1	1.079 E-6	1.73 E-6	1.815 E-6	2.288 E-6
2	1.107 E-6	1.531 E-6	1.865 E-6	2.497 E-6
3	1.023 E-6	1.422 E-6	1.855 E-6	2.459 E-6
4	1.051 E-6	1.786 E-6	1.964 E-6	2.104 E-6
5	1.024 E-6	1.561 E-6	1.813 E-6	2.490 E-6
6	1.051 E-6	1.519 E-6	1.849 E-6	2.464 E-6
7	1.007 E-6	1.494 E-6	1.892 E-6	2.205 E-6
8	0.994 E-6	1.420 E-6	1.986 E-6	2.156 E-6
9	1.110 E-6	1.498 E-6	1.849 E-6	2.544 E-6
10	1.033 E-6	1.404 E-6	2.065 E-6	2.269 E-6

We scan see that on average each trial took about the same time for each input size.

Lastly we were asked to plot our results. This can be seen in the MATLAB plot below.

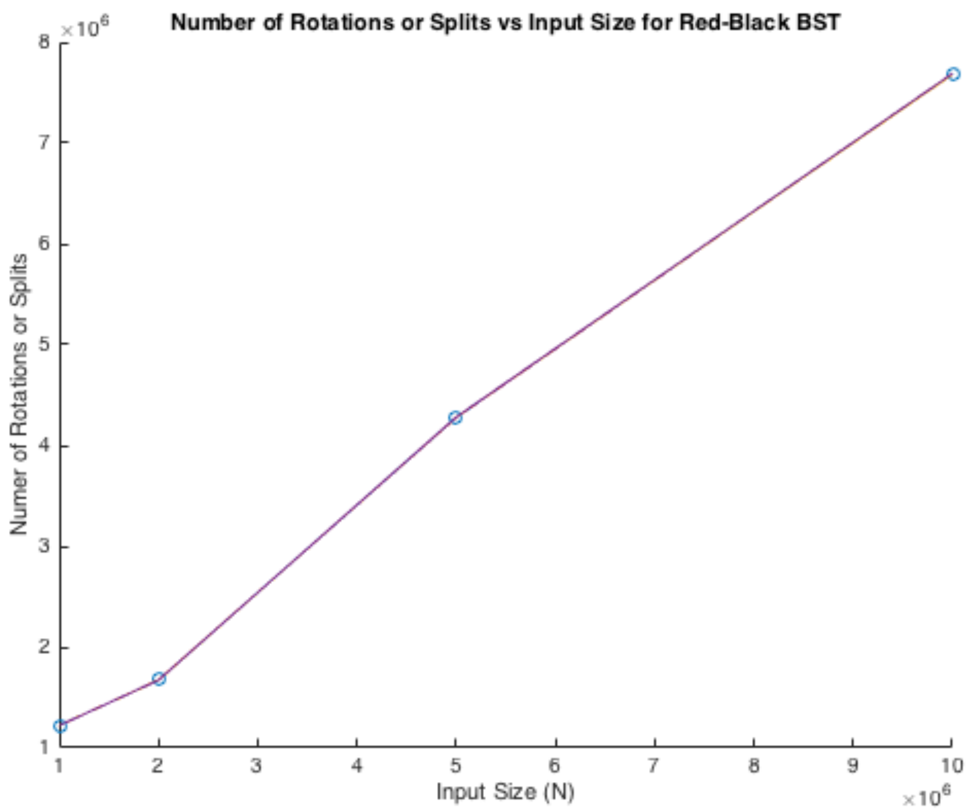


b.)

For part b of this question we were asked to do the same thing as part a but for total number of rotations and node splits that are used to build the trees.

The data and MATLAB plot for this can be seen below

Number of Rotations and Node Splits For 10 Trials				
	Input Size			
Trial #	**1,000,000**	**2,000,000**	**5,000,000**	**10,000,000**
1	1,219,754	1,674,260	4,279,165	7,685,375
2	1,221,671	1,671,961	4,278,552	7,685,541
3	1,220,913	1,670,857	4,276,384	7,688,298
4	1,223,373	1,669,200	4,281,179	7,684,226
5	1,221,978	1,671,983	4,281,575	7,684,434
6	1,221,568	1,672,830	4,276,443	7,683,570
7	1,221,066	1,671,811	4,278,841	7,686,747
8	1,222,428	1,672,351	4,279,538	7,681,721
9	1,221,579	1,669,574	4,281,459	7,680,438
10	1,221,283	1,671,499	4,281,278	7,689,667



We see that this data is so similar; each of their best fit lines over lap each other. I've also added the points for each set of data to better visualize the data on the graph. The small circles in the plot above represent these.