CSCE 221 Assignment 5 Cover Page

First Name Last Name UIN

User Name E-mail address

Please list all sources in the table below including web pages which you used to solve or implement the current homework. If you fail to cite sources you can get a lower number of points or even zero, read more on Aggie Honor System Office website: `http://aggiehonor.tamu.edu/`

Type of sources				
People				
Web pages (provide URL)				
Printed material				
Other Sources				

I certify that I have listed all the sources that I used to develop the solutions/codes to the submitted work.
On my honor as an Aggie, I have neither given nor received any unauthorized help on this academic work.

Your Name Date

Assignment 5 (40 pts)

Program: Due April 10 at 11:59 pm

Description: You will be given an executive file, `RBTree`, and you will use it for all testing and running in this assignment therefore you do not need to write any code here. Make sure that you use the TAMU servers (*compute*) to run this assignment.

1. How to run the program
   (a) To run all the test cases, type
   `./RBTree`
   (b) To print out the trees level by level, type
   `./RBTree -o 1`
   In place of `1` it can be any number ≤ 12 (but don't go past 4 because it will print out too much for the terminal to handle). This number is the number of tree levels to print out.
   (c) To test on a custom file, type
   `./RBTree -f file_name`
   where `file_name` must include the (relative) path to this file (say, `data-files/7p`).
2. Use the provided files to test the Red-Black and regular binary search trees created using the same input.
   (a) Use a small input (the number of nodes less than 16) to compare a Red-Black tree generated by the program with a Red-Black tree obtained by hand. Include an evidence of these comparisons (pictures, screenshots, etc).
   (b) Make a table and a plot showing the average search cost (y-axis) versus the number of tree nodes (x-axis) of all Red-Black trees created with the data used.
      i. You should produce a graph with 6 plots: plots for the linear, perfect, and random file type data for both the regular binary trees and the Red-Black trees.
      ii. Your graph plots should include the data points from all 12 input files (from 1 to 12).

Report (40 points)

Write and submit to eCampus a brief report that includes the following:

1. Provide a brief description, reason for building, and applications of the Red-Black tree data structure and its operations.
2. Provide the upper bound on individual search cost in a Red-Black and binary search tree in the worst case. Express this cost in terms of big-O notation. See the lecture notes for more details.
3. How can you justify that the computed average search costs for some Red-Black trees is higher than for perfectly balanced binary search trees? Does the formula below provide lower bound on the computed average search costs for Red-Black trees? Justify your answer.

$$\sum_{d=0}^{\log_2(n+1)-1} 2^d(d+1)/n \simeq ((n+1)\cdot \log_2(n+1)/n) - 1$$

4. Include the table and plots of computed average search costs for Red-Black and regular binary search trees discussed in the item 2 of the **Description** part, together with the comparisons with theoretical lower and upper bounds. Write your conclusion.
5. Include the testing cases for the small input data (the number of nodes less than 16) for the files selected by you in the report.
6. Summary. What have you learned by doing the assignment?