Name: Ryan Doherty

Section: 004

Github URL: <https://github.com/RD345/CS435-Project-1>

URLs/ISBNs for ALL consulted websites/textbooks:

* Geeksforgeeks.org
* W3schools.org
* Stackoverflow.com
* docs.python.org

CS 435 – Project 1: Trees

**Due Dates:**

* **Parts 1-3 due 11:59pm, February 25th.**
* **Parts 4-7 due 11:59pm, March 4th.**
* **Part 8 due 11:59pm, March 12th.**

For this project, you will be responsible for uploading all code in a Github repository. Please print this and turn in all written answers here. **Do not turn in written code here! Code must be submitted via a link to a Github repository!**

Sresht will personally review all of the code you submit, and leave comments on what you can improve on. For Project 1, you will be required to amend your code and push it back up to address those code comments. For Project 2, you will be required to give your peers code review yourself! So please pay attention to what kinds of comments I am making in code review.

Note: For this project, you may use whatever programming language you want. However, we have implemented this project in Java, Python, and C++, so using one of those languages would make it significantly easier to get help from me and the TAs.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Part: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Total |
| Points: | 20 | 10 | 5 | 20 | 15 | 15 | 0 | 15 | 100 |
| Score: |  |  |  |  |  |  |  |  |  |

# 1. (20 points) Binary Search Trees

## (2 points) In your own words, list the properties of a Binary Search Tree (BST).

A binary search tree is a type of tree which possess three main properties: one, it has two children max; two: the left child is less than the root, which in turn is less than the right child; and three: it has no duplicate values. All trees consist of nodes, which have store data, typically a value or set of values, and either have children, or no children, which makes them a leaf. The node from which all other nodes are descendants is known as the root.

## (2 points) What are the respective asymptotic worst-case run-times of each of the following operations of a BST? Give a Θ bound if appropriate. Justify your answers. You do NOT need to do a line-by-line analysis of code.

### Insert

### Delete

### Find-next

### Find-prev

### Find-min

### Find-max

1. Worst case is O(n), where n = the number of nodes in the tree. The worst case is when every node must be visited, such as in a tree where the height = # of nodes (skewed tree).
2. Worst case is O(n), for the same reasons as insert.
3. Same
4. Same
5. Same

## Same (8 points) (You must submit code for part of this question!) Use the framework below to describe how you would recursively implement the following methods of a Binary Search Tree. Afterwards, submit an implementation of all of the methods in your GitHub. Note that the Rec suffix simply means that the function is recursive.

### insertRec

### deleteRec

### findNextRec

### findPrevRec

### findMinRec

### findMaxRec

**You must submit answers to all of the below questions for full credit!!** Feel free to keep the answers as short or as long as you need to - you definitely don’t need to write more than a couple of sentences for each one.

### Repeat the question in your own words. List assumptions you make about the requirements.

1. Create a binary search tree, with methods to insert, find next, find previous, find minimum, and find maximum. Make all methods recursive in their implementations.

### Enumerate edge cases that you will want to consider. In this assignment, you don’t have to write code to address those as long as you call them out here.

1. This won’t account for anything that’s not integers, and it won’t account for duplicates values.

### Illustrate examples of input and output.

### Come up with an algorithm for each method. You don’t need to submit anything for this part, but I strongly recommend doing this before starting to code.

### Identify whether there are any issues with performance or space in your algorithm, and if so, iterate on it until it’s as optimized as possible. If you’re stuck here, please ask for help during Office Hours or on Slack!!!

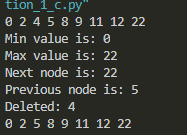
1. Since this algorithm is recursive, it uses implicit space, which will grow as the number of elements grows. If there is a recursion limit, that can also cause issues.

### (*You must submit code for this question!*) Translate your algorithm into real code. For this exercise, please do so in an IDE and upload it to Github. Add the suffix *Rec* to all of your recursive methods. For example, your recursive implementation of insert should be called insertRec().

1. Submitted on GitHub (question\_1\_c.py).

### What are problems/trade-offs with your current method? How might you optimize it to prevent those issues? You don’t have to optimize them here, but you must enumerate them.

1. Although recursive is easier to implement, it is also more costly than an iterative solution.



## (8 points) (*You must submit code for this question!*) Submit an implementation of the following iterative methods in a Binary Search Tree. You do not need to submit written answers to the framework from above, but it might be useful for you to consider the answers to those questions when writing code. Note that the Iter suffix simply means that the function is iterative. Keep in mind that an iterative solution cannot make a single recursive call!

### insertIter

### deleteIter

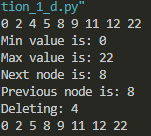
### findNextIter

### findPrevIter

### findMinIter

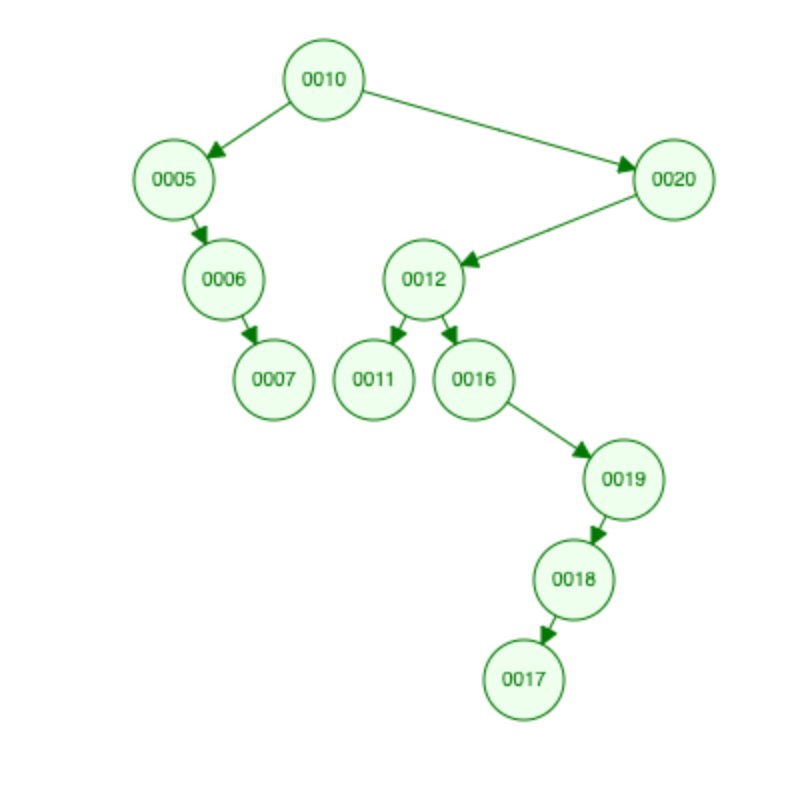
### findMaxIter

A: Submitted on GitHub (question\_1\_d.py).



# 2. (10 points) Sort It!

## (1 point) In the following BST, what is the sorted order of elements, from lowest value to highest value? Write your answer as a comma-separated list.



A: 0005, 0006, 0007, 0010, 0011, 0012, 0016, 0017, 0018, 0019, 0020

## (4 points) In your own words, describe an algorithm that uses the properties of a BST to take in a list of unsorted elements and output a list of sorted elements.

Using the properties of a BST can make creating a sorted list very easy. Since we know that: left < root < right, we can simply do an In-order traversal to print a list of increasing value. If we start with an unsorted list, we just feed that data into a new binary tree, then perform the In-order traversal.

## (5 points) (*You must submit code for this question!*) Implement the algorithm that you described above in sort().

A: Submitted on GitHub (question\_2\_c.py).



# 3. (5 points) Arrays of Integers

## (2 points) (*You must submit code for this question!*) Implement a function getRandomArray(n) where the output is an array of size n, and contains distinct random numbers (in other words, no two numbers in the array should be the same number). Math.rand() might be useful here.

A: submitted on GitHub (question\_3\_a.py).

Ran with 20 elements:



## (3 points) (*You must submit code for this question!*) Implement a function getSortedArray(n) where the output is an array of size n. The 0*th* element should be equal to n, the 1*st* element should be equal to n-1, and so on.

A: submitted on GitHub (question\_3\_b.py).

Ran with 20 elements:



This concludes the set of problems that must be completed and turned in by 11:59pm on Tuesday, February 25*th*. The rest of this project must be completed and turned in by 11:59pm on Wednesday, March 4*th*.

# 4. (35 points) Balanced Binary Search Trees

## (2 point) In your own words, list the properties of a Balanced Binary Search Tree (BBST). Use terminology discussed in class.

A balanced BST is one in which every branch of the tree is within +-1 (absolute value of 1) of each other in height. It also inherits the values that any other BST has.

## (3 points) What are the respective asymptotic worst-case run-times of each of the following operations of a BBST? Give a Θ bound if appropriate. Do not forget to include the complexity of the rebalancing operation where needed. Justify your answers. You do NOT need to do a line-by-line analysis of code.

Where n = number of nodes in the BBST:

### Insert:

O(log n), as the number of nodes to search is cut in half at each level.

### Delete:

O(log n), for the same reasons as in i.

### Find-next:

Same as above.

### Find-prev:

Same as above.

### Find-min:

θ(n), where n = the height of the tree. The min will always be in the leftmost node.

### Find-max:

θ(n), where n = the height of the tree. The max will always be in the rightmost node.

## (15 points) You must submit code for this question!) Submit an implementation of the following iterative methods in an AVL Tree. You do not need to submit written answers to the framework from above, but it might be useful for you to consider the answers to those questions when writing code. Note that the Iter suffix simply means that the function is iterative. Keep in mind that an iterative solution cannot make a single recursive call! You will need to use your implementation of Node from the previous question.

### insertIter

### deleteIter

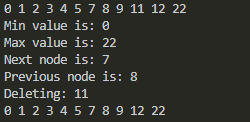
### findNextIter

### findPrevIter

### findMinIter

### findMaxIter

A: Submitted on GitHub (question\_4\_c.py).



*From here, we will prove the efficiency of a balanced binary search tree as it compares to an unbalanced binary search tree using by building trees using a list of integers.*

# 5. (15 points) Constructing Trees

## (5 points) (*You must submit code for this question!*) Use your recursive implementation of your BST and your iterative implementation of your AVL Tree from Parts 1 and 2 to construct trees using getRandomArray(10,000). Both trees must be made from the same array. In other words, do not call the method twice - store the output of the method from getRandomArray(10,000) once and use it to construct both trees.

A: Submitted on GitHub (question\_5\_a.py).

## (5 points) Did you run into any issues? Test your code on a smaller input (say, getRandomArray(10), and see if you’re still running into the same error. If it works on inputs of size 10 but not size 10,000, your code is probably fine and this is expected! Explain why you’re running into issues (or might run into issues), using concepts we covered in class.

A: The code ran fine in both cases, even with 10,000 elements in the input arrays. However, the recursive implementation might fail if the recursion goes too deep, as the implicit space will eventually overflow. I was able to get this to run with 100,000, and then I tried 1,000,000 elements as well. The 100k run took about a minute, and the 1 million run took well over half an hour, but was still going when I stopped it. I was watching in task manager as it ran, and the RAM usage of python slowly crept up the entire time starting at 43.4MB, and going up to 46.3MB by the time I stopped it.



## (5 points) (*You must submit code for this question!*) Use your iterative implementations of your AVL Tree and BST from Parts 1 and 2 to construct trees from the input of your implementation of getRandomArray(10,000). Both trees must be made from the same array. In other words, do not call the method twice - store the output of the method from getRandomArray(10,000) once and use it to construct both trees.

A: Submitted on GitHub (question\_5\_c.py).

# 6. (15 points) Compare Implementations

## (5 points) (You must submit code for this question!) Modify your iterative implementations of your methods in AVLTree and BinarySearchTree by keeping track of how many times you traverse one level down in the tree. In other words, if you go from a node to its child, add 1 to the counter.

A: Modified all three trees including recursive. Submitted on GitHub

## (5 points) (*You must submit code for this question!*) Construct a BST and AVLTree iteratively using getRandomArray(10000). Compare how many levels we have to traverse in the two trees. You can include a screenshot of your code’s output or write it out by hand.

A: Submitted on GitHub (question\_6\_b.py).



## (5 points) (*You must submit code for this question!*) Construct a BST and AVLTree iteratively using getSortedArray(10000). Compare how many levels we have to traverse in the two trees. You can include a screenshot of your code’s output or write it out by hand.

A: Submitted on GitHub (question\_6\_c.py).



# 7. (13 points) Extra Credit

*Note: Extra credit problems require significantly more independent research and effort than the for-credit problems. They will also not off-set poor comprehension of previous parts of this project. It is strongly recommended to save these for last.*

## (3 points) (*You must submit code for extra credit on this question!*) Use time packages in your respective language to quantify (in milliseconds/picoseconds) how much longer it takes to run 10,000 inserts and 10,000 deletes on a Binary Search Tree versus a Balanced Binary Search Tree.



\*Delete is implemented as well, but has a bug.

## (10 points) *Warning: This problem will be very time consuming!!* (*You must submit code for extra credit on this question!*)Use [https://www.geeksforge](https://www.geeksforgeeks.org/red-black-tree-set-1-introduction-2/)eks.[org/red-black-tree-set-1-introduction-2/](https://www.geeksforgeeks.org/red-black-tree-set-1-introduction-2/) as a guide to learning about Red-Black trees (or R/B Trees), which are another self-balancing tree. Implement a R/B tree that supports the same features as the AVL tree you implemented, and compare run-times in milliseconds.

This concludes the set of problems that must be completed and turned in by 11:59pm on Wednesday, March 4*th*.

# 8. (15 points) Code Review

After submitting your code, Sresht will be individually reading and reviewing your code. Comments will be added to your code. At some point before March 12*th*, it is your responsibility to amend your code and push up a new commit to the same repository. If you sufficiently address all points made in code review, you will receive full credit on this part of the project