CS202, Fall 2024

Homework 1 - Binary Search Trees

Due: 23:59, 05/03/2025

Before you start your homework, please read the following instructions carefully:

FAILURE TO FULFIL ANY OF THE FOLLOWING REQUIREMENTS WILL RESULT IN A GRADE SCORE OF 0 (zero) WITHOUT ANY CHANCE OF REDEMPTION.

- Make sure to abide by the Honor Code for Assignments.
- Store your solutions in a folder named **studentID_name_surname_hw1**, and then convert it to a ZIP archive. Use the Moodle submission form to upload your archive.
- Your ZIP archive should contain **only** the following files:
    - **studentID_name_surname_hw1.pdf**, the file containing the answers to Questions 1, and 3. Your answer to Question 1 **MUST BE HANDWRITTEN** on paper and then scanned.
    - In this assignment, you must have separate interface and implementation files (i.e., separate .h and .cpp files) for your class. Your class name MUST BE **BST** and your file names MUST BE **BST.h** and **BST.cpp**. Note that you may write additional class(es) in your solution.
    - **analysis.h** and **analysis.cpp** files related to the time analysis of Question 3.
    - Do not forget to put your name, student ID, and section number in all of these files. Add comments on your implementation well. Add a header (see below) to the beginning of each file:

```
/**
 * Title: Binary Search Trees
 * Author : Name & Surname
 * ID: 12345678
 * Section : 1
 * Homework : 1
 * Description : description of your code
 */
```

    - Do not put any unnecessary files such as the auxiliary files generated from your preferred IDE.
    - Please do not use **Turkish** letters in your file and folder names.
- Your code must **compile**.
- Your code must be **complete**.

- You **ARE NOT ALLOWED** to use any data structure or algorithm-related function from the C++ standard template library (STL) or any other external libraries.
- The code (main function) given in Question 2 is just an example. We will test your implementation using different scenarios, which will contain different function calls. Thus, do not test your implementation only by using this example code. We recommend you to write your own driver files to make extra tests. However, you **MUST NOT** submit these test codes (we will use our own test code). In other words, do not submit a file that contains a function called **main**.
- We will test your code using the **Google test library**. Therefore, the output message for each operation in Question 2 MUST match the format shown in the output of the example code.
- Your code must run on the **dijkstra.cs.bilkent.edu.tr** server. You can test your implementation using the tests from '**/home/cs/klea.zambaku/cs202/'.**
- For any questions related to the homework contact your TA: *klea.zambaku@bilkent.edu.tr*

LATE SUBMISSION: Late submissions are not allowed and will receive a ZERO grade.

Question 1 - 30 points

A. [5 points] Insert 24, 31, 5, 6, 7, 1, 21, 12, 25 and 23 into an empty binary search tree in the given order. Show the resulting BST after every insertion.
B. [15 points] What are the preorder, inorder, and postorder traversals of the BST you have after (A)? Write only the number sequence.
C. [5 points] Delete 24, 1, 7 and 5 from the BST you have after (A) in the given order. Show the resulting BST after every deletion.
D. [5 points] What is the 6th key in the preorder traversal of a binary search tree if its postorder traversal is as follows?

Postorder: 4, 9, 7, 21, 20, 24, 65, 32, 12

Question 2 - 50 points

Implement a pointer-based Binary Search Tree whose keys are integers. Your solution must be implemented in a class called **BST**. Below is the required public part of the **BST class**. The interface for the class must be written in a file called **BST.h** and its implementation must be written in a file called **BST.cpp**. You can define additional public and private member functions and data members in this class. You can also define additional classes in your solution.

```
class BST{
    public:
    BST(int keys[], int size);
    ~BST();
    void insertKey(int key);
    void deleteKey(int key);
    void keyLevel (int key);
    void displayInorder();
    void displayPostorder();
    void findFullBTLevel();
    void lowestCommonAncestor(int A, int B);
    void maximimSumPath();
    void maximumWidth();
}
```

The member functions are defined as follows (Each function executed properly is worth 5 points):

BST: Constructor. Reads the integer keys of the array one by one, and inserts them into the binary search tree one by one. You can assume that there are no identical keys in the array.
insertKey: Inserts a key to the BST, if it does not exist in the BST before. If the key already exists, a message saying so should be displayed.
deleteKey: Deletes a key from the BST, if it exists in the BST. If the key does not exist, a message saying so should be displayed. If the BST is empty, a message should be displayed.
keyLevel: Finds the level of the key.
displayInorder: Displays an inorder traversal of the BST.
displayPostorder: Displays a postorder traversal of the BST.
findFullBTLevel: Finds the maximum level at which the tree is a full binary tree, and displays the level number. Assume that root is level 1 in all the level-related questions.
lowestCommonAncestor: Finds the common Ancestor of keys A and B that has the highest level. Assume that a node A is the ancestor of itself.
maximumSumPath: Finds the path from root to a leaf node which has the maximum sum of the keys included in the path, and displays the keys of the path from its root to the leaf. You can assume that there is only one path with the maximum Sum path in the tree.
maximumWidth: Finds the level of the tree which has the maximum number of nodes and displays its keys from left to right. In case of multiple maximum widths select the lowest level.

Figure 1 is an example of a BST tree used in our test example.

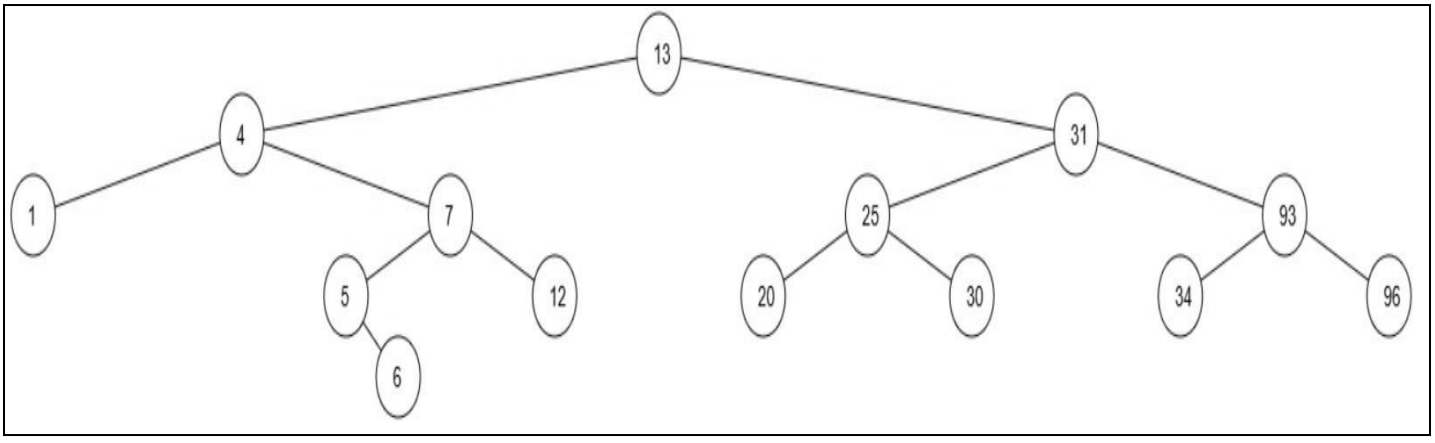

Figure 1

Example test code:

```
#include "BST.h"

int main(){
    int keys[] = {13, 4, 31, 7, 93, 12, 1, 34, 25, 5, 6, 96, 20, 30};
    int size = 14;
    BST bst (keys, size);
    bst.findFullBTLevel();
    bst.lowestCommonAncestor(1,6);
    bst.lowestCommonAncestor(5,25);
    bst.maximimSumPath();
    bst.maximumWidth();
    bst.keyLevel(7);
    bst.insertKey(15);
    bst.insertKey(34);
    bst.deleteKey(25);
    bst.deleteKey(0);
    bst.displayInorder();
    bst.displayPostorder();

    return 0;
}
```

Example test output:

```
BST with size 14 created.
Full binary tree level is: 3
Lowest common ancestor of 1 and 6 is: 4
Lowest common ancestor of 5 and 25 is: 13
Maximum sum path is: 13, 31, 93, 96
Maximum width level is: 5, 12, 20, 30, 34, 96
The level of key 7 is: 3
Key 15 is added!
Key 34 is not added! It already exists in the BST.
Key 25 is deleted.
Key 0 is not deleted. It does not exist in the BST.
Inorder display is: 1, 4, 5, 6, 7, 12, 13, 15, 20, 30, 31, 34, 93, 96
Postorder display is: 1, 6, 5, 12, 7, 4, 15, 30, 20, 34, 96, 93, 31, 13
```

Note: This example is for illustration purposes, DO NOT submit a file that contains a function called **main**. Check the google tests to see how the functions will be called.

Question 3 - 20 points

In this question, you will analyze the time performance of the pointer-based implementation of binary search trees. Write a function, **void timeAnalysis()** and add your code to **analysis.h** and **analysis.cpp** files. Your implementation must follow the best-practices; i.e., the function interfaces must be defined in the header (.h) file, and the function implementations must be in the C++ file (.cpp). This function generates a list of 10,000 random numbers and begins inserting them into an empty pointer-based Binary Search Tree (BST). After every 1000 insertions, displays:

a) the time taken for those insertions (using the clock function from the **ctime** library to calculate elapsed time)
b) the height of the tree.

After running your program, you are expected to prepare *a single-page report* about the experimental results obtained from the timeAnalysis function. With the help of a spreadsheet program (Google Sheets, Matlab, or other tools):

1. [10 points] Plot the BST height versus the number of nodes in the BST after each period of insertions to the BST. What does this function look like? Is it a linear function or a logarithmic function? Is this the expected behavior? Why?

2. [10 points] Plot the elapsed time versus the number of nodes in the tree after each period of insertions to the BST. How would the time complexity of your program change if you inserted sorted integers into it instead of randomly generated ones?