New **Y**ork **U**niversity
Computer Science Department
Data Structures
Dr. Anasse Bari

Homework Five: Binary Search Trees

Deadline: *See NYUclasses for the deadline, 15% off per day after the deadline (4 days maximum).*

Learning Objectives:

- Learning how to design Binary Search Trees

Read the guidelines bellow carefully to avoid receiving a zero grade on the HW:

- You must use generic types when implementing BST for this HW.
- Read attached Pdf file for additional design tips for this homework.
- Please follow the instruction as they are in the HW (DO NOT change the file format or the output format otherwise it will result into a null grade)
- Attach the Java source files and include them into HW's zip file. The file name should be **YourLastName_HW5.zip**
- Make an archive (zip file or compressed file) with all the **java files (the .java files NOT the .class files)** and post it on NYU Classes
- You must comment you code (basic comments explaining the role of a class, a method or variables used in your submission)
- Compile and run the program before you submit.
- It is your responsibility to make sure if the Zip files has your actual latest files. You may send the file to yourself by email to double check that is the actual file before you upload on NYU classes.
- **If the graders cannot open the file, you will receive a grade of zero.**
- **If you send the .class files instead of the .java files (source files) you will receive a zero.**
- An act of cheating will be severely addressed with an immediate zero on the homework and a report to the academic advisor and the administration.
- You will automatically lose 50% of the points for an exercise if the program does not compile and run correctly.
- **Plagiarized assignments will get a ZERO grade**. You cannot change the variable names of other student's solution and submit it as yours. The program structure of other students must not match yours. Every student must come up with his/her own solution. Any cheating (e.g. copying from internet without citing sources) is a serious violation of the University student code.
- **Homeworks sent by *email* to the instructor or to the graders will NOT be reviewed and will not be accepted.**

Problem Statement

This assignment involves modeling a community of people using the data structures we covered in class to be able to answer certain questions about a community and its members.

Each member x of the community is a person that has a unique SSN, a name (first and last name), a mother, a father, and a list of people (from within the community) that x considers to be his/her friends.

Note that if x considers y a friend, it does not necessarily follow that y considers x a friend. The kinds of queries that your program should be able to answer are:

For a given person x, find all the children of x;

For a given person x, find all the half-siblings of x;

For a given person x, find all the full-siblings of x;

For a given person x, find the friends of x that consider x as their friend as well;

For a given person x, find all the Persons that consider x as their friend;

Who has the most mutual friends;

The precise types of questions your program should answer are detailed below.

The Input:

Your program will take two input files from the command prompt, the first file containing the information about the community, and the second file containing the queries that your program should answer.

The first input file, which has the information about the community, will be such that each person has one "paragraph", with at least one blank between successive paragraphs. The format of the paragraph of a person is:

FIRST NAME: word // just one word representing the first name

LAST NAME: word // just one word representing the last name

SSN: number // social security number (any unsigned integer)

FATHER: number // representing the SSN of the father

MOTHER: number // representing the SSN of the mother

FRIENDS: a coma-separated list of SSNs of this person's friends

Note: The list after FRIENDS is the list of the SSNs of the people that the person in question considers to be his/her friends. Those SSNs are not necessarily sorted.

The second input file will have a sequence of queries, one query per line, where each query is any of the following:

NAME-OF SSN

MOTHER-OF SSN

FATHER-OF SSN

HALF-SIBLINGS-OF SSN

FULL-SIBLINGS-OF SSN

CHILDREN-OF SSN

MUTUAL-FRIENDS-OF SSN

INVERSE-FRIENDS-OF SSN

WHO-HAS-MOST-MUTUAL-FRIENDS

Note that SSN in the above queries stands for an unsigned int.

The output:

Your program must take as input the two input files above, in that order. That is, we should be able to compile your program into, then call it from the argument list on Eclipse or you can use the filereader and read communityFileName.txt and queriesFileName.txt

Make sure that the files are within your source code file so you do not have to specify the full path of the input files.

Your program must then output as many lines as there are query lines in Eclipse: the kth output line should be the answer to the kth query line in file queriesFileName.

Your output line answer to a query should begin by repeating the exact same query, followed by a colon (:), followed by the query-answer. The query-answer for the first three types of query should be the two words representing the first and last name of the answer-person. The query-answer for the next five types of query should be a list of first name and last name pairs, separated by commas, and ordered so that their SSNs are in increasing order.

For example, a possible output for "CHILDREN-OF 3155" can be:

CHILDREN-OF 3155: John Smith, Adam Smith, Susan Smith

And for the query "MUTUAL-FRIENDS-OF 3155", a possible answer can be:

MUTUAL-FRIENDS-OF 3155: John Redman, Janet Fisher, Joe Fisher

The answer to INVERSE-FRIENDS-OF SSN query should consist of the people that consider the person with the specified SSN to be their friend.

The answer to a WHO-HAS-MOST-MUTUAL-FRIENDS query should be the full name of the person who has the most mutual friends; in case of a tie, choose the one with the smallest SSN.