

National Institute of Technology Calicut
Department of Computer Science and Engineering
Fourth Semester B. Tech.(CSE)-Winter 2021-22
CS2094D Data Structures Laboratory
Assignment #5 - Part A

Submission deadline (on or before): 04.04.2022, 9:00 AM

Policies for Submission and Evaluation:

- You must submit your assignment in the Eduserver course page, on or before the submission deadline.
- Ensure that your programs will compile and execute without errors using gcc compiler.
- During the evaluation, failure to execute programs without compilation errors may lead to zero marks for that evaluation.
- Your submission will also be tested for plagiarism, by automated tools. In case your code fails to pass the test, you will be straightaway awarded zero marks for this assignment and considered by the examiner for awarding F grade in the course. Detection of ANY malpractice related to the lab course can lead to awarding an F grade in the course.

Naming Conventions for Submission

- Submit a single ZIP (.zip) file (do not submit in any other archived formats like .rar, .tar, .gz). The name of this file must be

ASSG<NUMBER>_<ROLLNO>_<FIRST-NAME>.zip

(Example: *ASSG1_BxyyyyCS_LAXMAN.zip*). DO NOT add any other files (like temporary files, input files, etc.) except your source code, into the zip archive.

- The source codes must be named as

ASSG<NUMBER>_<ROLLNO>_<FIRST-NAME>_<PROGRAM-NUMBER>.c

(For example: *ASSG1_BxyyyyCS_LAXMAN_1.c*). If you do not conform to the above naming conventions, your submission might not be recognized by our automated tools, and hence will lead to a score of 0 marks for the submission. So, make sure that you follow the naming conventions.

Standard of Conduct

- Violation of academic integrity will be severely penalized. Each student is expected to adhere to high standards of ethical conduct, especially those related to cheating and plagiarism. Any submitted work MUST BE an individual effort. Any academic dishonesty will result in zero marks in the corresponding exam or evaluation and will be reported to the department council for record keeping and for permission to assign F grade in the course. The department policy on academic integrity can be found at: http://cse.nitc.ac.in/sites/default/files/Academic-Integrity_new.pdf.

QUESTIONS

1. Write a program that implements the DISJOINT-SET data structure using rooted forests. Also, write functions to implement the **ranked union and path compression** heuristics on your data structure. Your program should compute the efficiency of the DISJOINT-SET data structure FIND operation by applying neither, either or both of the heuristics. The efficiency is calculated by counting the total number of data accesses performed over the course of the program. Your program must support the following functions:

- MAKESET(x) - creates a singleton set with element x .
- FIND(x) - finds the representative of the set containing the element x .
- UNION(x, y) - merges the sets containing elements x and y into a single set. The representative of the resultant set is assigned with FIND(x), unless the ranked union heuristic is used and the ranks of both FIND(x) and FIND(y) are different. Otherwise, the representative is assigned in accordance with the ranked union heuristic.

Note that looking up an element in the data structure must be done in $\mathcal{O}(1)$ time.

Input Format:

- The input consists of multiple lines, each one containing a character from {'m', 'f', 'u', 's'} followed by zero, one or two integers separated by single space. The integer(s), if given, is in the range 0 to 10000.
 - Call the function MAKESET(x) if the input line contains the character 'm' followed by an integer x . Print -1 if x is already present in some set, and the value of x , otherwise.
 - Call the function FIND(x) if the input line contains the character 'f' followed by an integer x . Print the value of FIND(x) if x is present, and -1 if x is not present.
 - Call the function UNION(x, y) if the input line contains the character 'u' followed by space separated integers x and y . Print -1, without terminating, if either x or y isn't present in the disjoint set. Print FIND(x) itself if FIND(x)=FIND(y). Otherwise, print the representative of the resultant set. The representative of the resultant set is assigned with FIND(x), unless the ranked union heuristic is used and the ranks of both FIND(x) and FIND(y) are different. Otherwise, the representative is assigned in accordance with the ranked union heuristic.
 - If the input line contains the character 's', print the number of data accesses performed by the FIND function by each of the data structures over the course of the program and terminate.

Output Format:

- The output consists of multiple lines of single space separated columns. The columns correspond to the following disjoint-set data structures:
 - a. with neither ranked union nor path compression applied.
 - b. with only ranked union applied.
 - c. with only path compression applied.
 - d. with both ranked union and path compression applied.
- Each line in the output contains the output of the corresponding line in the input, after applying to the respective data structures.
- The last line of the output contains the number of data accesses performed by the FIND function by each of the data structures over the course of the program.

Sample Input

```
m 1
m 2
m 3
m 4
```

m 5
m 6
m 7
m 8
m 9
u 1 2
u 3 4
u 5 6
u 7 8
u 9 8
u 6 8
u 4 8
u 2 8
f 9
m 10
u 10 9
s

Sample Output

1
2
3
4
5
6
7
8
9
1 1 1 1
3 3 3 3
5 5 5 5
7 7 7 7
9 7 9 7
5 5 5 5
3 5 3 5
1 5 1 5
1 5 1 5
10
10 5 10 5
38 32 33 30