

Programming Assignment #7

Hashing

1 Problem Description

Hashing is a technique of mapping a large set of arbitrary data to tabular indices using a *hash function*, which represents dictionaries for large datasets. It allows lookups, updating and retrieval operation to occur in a constant time, i.e. $O(1)$.

In this programming assignment, you are asked to design your own hash function to speed up the process of finding two integers in an array whose sum is equal to the target number, instead of using linear search. Given an integer array of n integers and the target number, s , you will need to output the indices of the first two integers such that the summation of both integers is equal to s . A naive approach to solve the problem is to implement nested loops. However, the time complexity could be as large as $O(n^2)$, which is inefficient when the array contains millions of integers.

2 Input Format

The input file contains two lines. The the first line gives the target number, while the second line gives all the integer numbers in the array, which are distinct. A sample input file is given below:

sample.in	Comment
8	the target number
2 4 1 7 10 5 6 12	distinct numbers in the integer array

3 Output Format

You will need to output the indices of two numbers in the integer array whose sum is equal to the target number. A sample output file resulting from “sample.in” is given below. The number in the first line associates with the index of the first integer, while that of the second line associates with the index of the second integer.

If there exists more than one solution, you have to choose the one with minimum indices. For example, in “sample.in”, both the indices, 0 and 6, and the indices, 2 and 3, are feasible solutions. Your output must be 0 and 6, rather than 2 and 3.

Output1.txt	Comment
0	the index of the first integer in the array
4	the index of the second integer in the array

4 Command-line Parameter

In order to test your program, you are asked to add the following command-line parameters to your program:

`[executable file name] [input file name] [output file name]`

5 Submission Information

- Your program must be written in the C/C++ language and can be compiled on the Linux platform.
- The source files of your program must be named with “[your student ID].h” and “[your student ID].cpp”.
- To submit your program, please archive all source files of your program into a single zip file, named “[your student ID].zip”, and upload it to E3.

6 Due Date

Be sure to upload the zip file by “January 4, 2023”. There will be a 25% penalty per day for late submissions.

7 Grading Policy

The programming assignment will be graded based on the following rules:

- Pass sample input with compilable source code (40%)
- The hidden test cases (60%)
 - Functional correctness: 80%
 - Performance (the rank of runtime): 20%
- For each case, the runtime limit is **5 minutes**. It will be regarded as “failed” if the execution time is more than 5 minutes.

Be sure to design your own hash table in this programming assignment. You are NOT allowed to use any container in the Standard Template Library(STL), or any other open-source library, such as vector, list, map, set, and all the other container classes. The submitted source codes will NOT be graded if those open-source container classes are found in your source codes.

The submitted source codes, which are either copied from or copied by others, will NOT be graded.