

CSC 2720: Data Structures

Lab 4

Instructor: Shiraj Pokharel

Due : @ 11:00 PM ET , Next Day after release

Answer the below questions. You may use whatever IDEs / editors you like, but you must submit your responses on iCollege as .java files. Failure to comply with this simple requirement will result in a score of Zero. Please, be careful not to be assigned a Zero score this way.

Few Rules to be followed, else will receive a score of ZERO

- (1) Your submissions will work exactly as required.
- (2) Your files shall not be incomplete or worse corrupted such that the file does not compile at all. Make sure you submit a file that compiles.
- (3) Your submission will show an output. Should you receive a Zero for no output shown do not bother to email me with "but the logic is perfect" !

Note that your program's output must **exactly** match the specs(design , style) given here for each problem to pass the instructor's test cases .

Design refers to how well your code is written (i.e. is it clear, efficient, and elegant), while *Style* refers to the readability of your code (commented, correct indentation, good variable names).

This week as well, you will just play games with Arrays. :) **Although few in number, I've received feedback from students that they see these exercises as "irrelevant" and "not making sense". Let me be polite in answering that you will not see these exercises as "irrelevant" and "not making sense" once you complete this course and transition into senior courses or work in a Tech internship or start work full time as an Engineer :)**

In user content generated web-services - lets say YouTube - the process of de-duplication is of serious importance. One straight forward reason is same video by different names is just an extra cost in data-storage. So getting rid of exact duplicate content makes financial sense. Please be reminded that exact same content with different video qualities may not be candidates for removal via de-duplication.

In today's Lab we will explore on ways to do a de-duplication of videos where video filenames are presented as integers. For the purposes of our task, we will set a very narrow criterion for de-duplication : just the filenames.

Below is how the filenames are represented
ARRAY[] = [11, 21, 21, 33, 40, 40, 40, 50, 50]

Below is the expected output after de-duplication
ARRAY[] = [11, 21, 33, 40, 50]

Please do not forget to notice the reduced size of the output array.

Also, please be reminded that you cannot use library functions in Java to perform the de-duplication operation. Doing so would straight up result in a score of Zero !

You will solve the problem in two ways:-

- (1) [50 points] Implement the function in such a way that your solution solves the problem with $O(n)$ time complexity and $O(n)$ space complexity. Here, n is the length of the list of input integers (array)
- (2) [50 points] In a separate implementation, improve the solution in (1) in such a way that your solution solves the problem with $O(n)$ time complexity but stays within the bounds of $O(1)$ space complexity. Here, n is the length of the list of input integers (array) as well.

Ungraded but Important : As a comment in your code, please write your test-cases on how you would test your solution assumptions and hence your code. Example of cases to be tested for are like : What if the array input which is expected does not exist - that is , input is a null. How should your code handle such a situation ? Maybe output some message like "Null input case, so no output"? What if the length of the array is one ?....so on and so forth. Please Remember : Although, written as comments - You will address your test cases in the form of code and not prose :)