CPSC 535 Advanced Algorithms
Project 2: It's a Small World

Prof. Doina Bein, CSU Fullerton

dbein@fullerton.edu

Introduction

In this project you will design and implement one algorithm related to strings. You will design the algorithm, describe the algorithm using clear pseudocode, and implement it using C/C++/C#/Java/Python, compile, test it, and submit BOTH the report (as a PDF file) and the files. The execution should take less than one hour for each input example.

Calculating shortest connection between movie casts

This is a spin-off of the popular game "Six Degrees of Kevin Bacon" (https://en.wikipedia.org/wiki/Six_Degrees_of_Kevin_Bacon#:~:text=Six%20Degrees%20of%20Kevin%20Bacon%20or%20Bacon%27s%20Law%20is%20a,ultimately%20leads%20to%20prolific%20American) invented in 2019. Let's say that you have casts for two movies, and you want to find out what is the shortest connection between the two casts.

We assume that the name of each actor is a string, thus a cast is a set of strings, ordered in alphabetical order. But it does not matter whether they are listed alphabetically or not, but for simplicity, let's list them in alphabetical order.

The input will be a positive integer $n > 2$, and a list of n casts from which the first two sets are more significant, CAST[0] and CAST[1]. If the two casts CAST[0] and CAST[1] have at least one string in common, then the shortest connection is 1. If the two casts CAST[0] and CAST[1] do not have any string in common, then look for another cast in the list of n casts, let's called it tempCast, such that CAST[0] and tempCast have a string in common, and CAST[1] and tempCast have a string common, then the shortest connection is 2. Else the shortest connection is greater than 2 or there is no connection.

Example 1:
Input: n=6 and
CAST[0] = {"Carrie-Anne Moss", "Gloria Foster", "Hugo Weaving", "Joe Pantoliano","Keanu Reeves", "Laurence Fishburne", "Marcus Chong"}
CAST[1] = { "Andre Braugher", "Beau Garrett", "Chris Evans", "Doug Jones", "Ioan Gruffudd", "Jessica Alba", "Julian McMahon", "Kerry Washington", "Laurence Fishburne", "Michael Chiklis" }
CAST[2] = {"Ewan McGregor", "Ian McDiarmid", "Jake Lloyd", "Liam Neeson", "Natalie Portman"}
CAST[3] = { "Geoffrey Rush", "Jack Davenport", "Johnny Depp", "Jonathan Pryce", "Keira Knightley", "Orlando Bloom" }
CAST[4] = { "Angela Bassett", "Chadwick Boseman", "Danai Gurira", "Daniel Kaluuya", "Forest Whitaker", "Letitia Wright", "Lupita Nyong'o", "Martin Freeman", "Michael B. Jordan", "Sterling K. Brown", "Winston Duke"}

CAST[5] = { "Andrew Borba", "Anne Hathaway", "Bill Irwin", "Casey Affleck", "Collette Wolfe", "David Oyelowo", "Francis X. McCarthy", "Jessica Chastain", "John Lithgow", "Matthew McConaughey", "Michael Caine", "Wes Bentley", "William Devane"}

Output: shortest connection = 1, actor ="Laurence Fishburne"

Example 2:
Input: n=7 and
CAST[0] = {"Carrie-Anne Moss", "Gloria Foster", "Hugo Weaving", "Joe Pantoliano","Keanu Reeves", "Laurence Fishburne", "Marcus Chong"}
CAST[1] = { "Andrew Borba", "Anne Hathaway", "Bill Irwin", "Casey Affleck", "Collette Wolfe", "David Oyelowo", "Francis X. McCarthy", "Jessica Chastain", "John Lithgow", "Matthew McConaughey", "Michael Caine", "Wes Bentley", "William Devane"}
CAST[2] = { "Geoffrey Rush", "Jack Davenport", "Johnny Depp", "Jonathan Pryce", "Keira Knightley", "Orlando Bloom" }
CAST[3] = { "Angela Bassett", "Chadwick Boseman", "Danai Gurira", "Daniel Kaluuya", "Forest Whitaker", "Letitia Wright", "Lupita Nyongo", "Martin Freeman", "Michael B. Jordan", "Sterling K. Brown", "Winston Duke"}
CAST[4] = { "Abraham Attah", "Asa Butterfield","Anne Hathaway", "Chloe Grace Moretz", "Daniel Radcliffe", "Jeff Goldblum", "Keanu Reeves", "Tom Holland"}
CAST[5] = { "Andre Braugher", "Beau Garrett", "Chris Evans", "Doug Jones", "Ioan Gruffudd", "Jessica Alba", "Julian McMahon", "Kerry Washington", "Laurence Fishburne", "Michael Chiklis" }
CAST[6] = { "Ewan McGregor", "Ian McDiarmid", "Jake Lloyd", "Liam Neeson", "Natalie Portman"}

Output: shortest connection = 2, cast = { "Abraham Attah", "Asa Butterfield","Anne Hathaway", "Chloe Grace Moretz", "Daniel Radcliffe", "Jeff Goldblum", "Keanu Reeves", "Tom Holland"}

Example 3:
Input: n=7 and
CAST[0] = { "Ewan McGregor", "Ian McDiarmid", "Jake Lloyd", "Liam Neeson", "Natalie Portman"}
CAST[1] = { "Andrew Borba", "Anne Hathaway", "Bill Irwin", "Casey Affleck", "Collette Wolfe", "David Oyelowo", "Francis X. McCarthy", "Jessica Chastain", "John Lithgow", "Matthew McConaughey", "Michael Caine", "Wes Bentley", "William Devane"}
CAST[2] = { "Geoffrey Rush", "Jack Davenport", "Johnny Depp", "Jonathan Pryce", "Keira Knightley", "Orlando Bloom" }
CAST[3] = { "Angela Bassett", "Chadwick Boseman", "Danai Gurira", "Daniel Kaluuya", "Forest Whitaker", "Letitia Wright", "Lupita Nyongo", "Martin Freeman", "Michael B. Jordan", "Sterling K. Brown", "Winston Duke"}
CAST[4] = { "Abraham Attah", "Asa Butterfield","Anne Hathaway", "Chloe Grace Moretz", "Daniel Radcliffe", "Jeff Goldblum", "Keanu Reeves", "Tom Holland"}
CAST[5] = { "Andre Braugher", "Beau Garrett", "Chris Evans", "Doug Jones", "Ioan Gruffudd", "Jessica Alba", "Julian McMahon", "Kerry Washington", "Laurence Fishburne", "Michael Chiklis" }
CAST[6] = {"Carrie-Anne Moss", "Gloria Foster", "Hugo Weaving", "Joe Pantoliano","Keanu Reeves", "Laurence Fishburne", "Marcus Chong"}

Output: shortest connection > 2 or no connection

The variables n and CAST must be read from a file since it is too tedious to read them from the user.

Please check that your program executes correctly on the given examples.

Please create more examples on your own and test them. The IMDB dataset are available as text files at https://www.imdb.com/interfaces/#plain . You will need to download relevant files which will provide information needed for the cast of movies. Consider only English Language movies for this problem.

You may want to consider fewer than the total number of actors that show up in the cast credits but you can of course consider them all.

Grading rubric

The total grade is 35 points. Your grade will be comprised of three parts, Form, Function, and Report:

- Function refers to whether your code works properly (18 points).
- Form refers to the design, organization, and presentation of your code. The instructor will read your code and evaluate these aspects of your submission (6 points):
  - README.md completed clearly = 2 points
  - Style (whitespace, variable names, comments, helper functions, etc.) = 2 points
  - Craftsmanship (appropriate handling of encapsulation, memory management, avoids gross inefficiency and taboo coding practices, etc.) = 2 points
- Report (11 points) divided as follows:
  - Summary of report document (2 points)
  - Pseudocode of the algorithm (6 points)
  - Four screenshots: one for the group members and three for the three sample input examples (1 point each, total 3 points)

Obtaining and Submitting Code

This document explains how to obtain and submit your work:
	GitHub Education / Tuffix Instructions

Here is the invitation link for this project:
	https://classroom.github.com/a/j2lqL3wQ

Implementation

You are provided with the following files.
1. README.md contains a brief description of the project, and a place to write the names and CSUF email addresses of the group members. You need to modify this file to identify your group members.
2. LICENSE contains a description of the MIT license.

What to Do

First, add your group members' names to README.md. Then write clear pseudocode for the rewriting algorithm, describe what parameters are needed to execute your program, and submit it as a PDF report.Your report should include the following:

1. Your names, CSUF-supplied email address(es), and an indication that the submission is for project 1.
2. A full-screen screenshot with your group member names shown clearly. One way to make your names appear in Atom is to simply open your README.md.
3. The pseudocode for the algorithm
4. A brief description on how to run the code.
5. Three snapshots of code executing for the three given examples

Then implement your algorithm in C/C++/C#/Java/Python. Submit your PDF by committing it to your GitHub repository along with your code.