DASF004 Final Assignment

Instruction

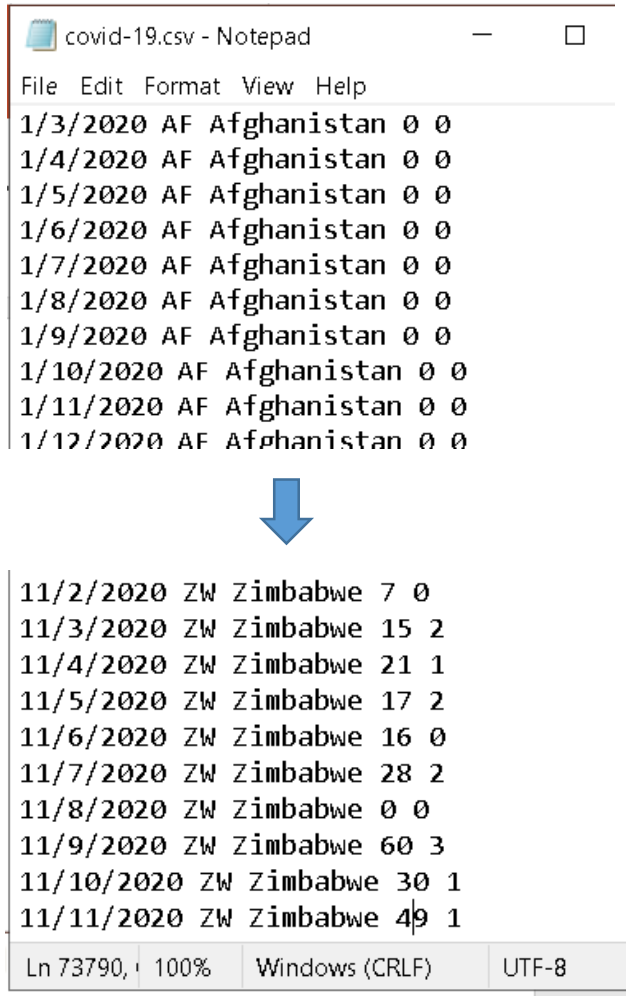
- **Read this instruction very carefully.**
- You will name the filename as *[StudentIDNumber]*.py.
  - Submit the file *[StudentIDNumber]*.py on iCampus before the deadline.
- You should submit the source code only.
- You may submit partial solution (and you'll receive partial credit).
- You may ask questions thru email at atang@skku.edu.
  - But please ask your questions early (otherwise, I may not reply in time before your deadline)
  - Do not ask question on iCampus. Ask question thru **email** only
- I will make clarifications about the problem on icampus.
- You may submit multiple versions.
  - I will grade the last version before the deadline you submitted only.
  - Work submitted after the deadline will not be graded
- Make backup!
- The deadline is Wednesday 2nd June 2021 18:00 pm.

Collaboration Policy

- The work you submit must be the work of your own.
- You are free to give or receive help when doing homework assignments, but you must follow the following restrictions:
- Only the helper can look at the code of others. Student who is receiving help must not look at the code of the helper;
- Student who is receiving help must do all the typing herself/himself. Helper must not touch the computer of the student who is receiving help; and
- All student can not post your code on the web, nor send your code to other students.

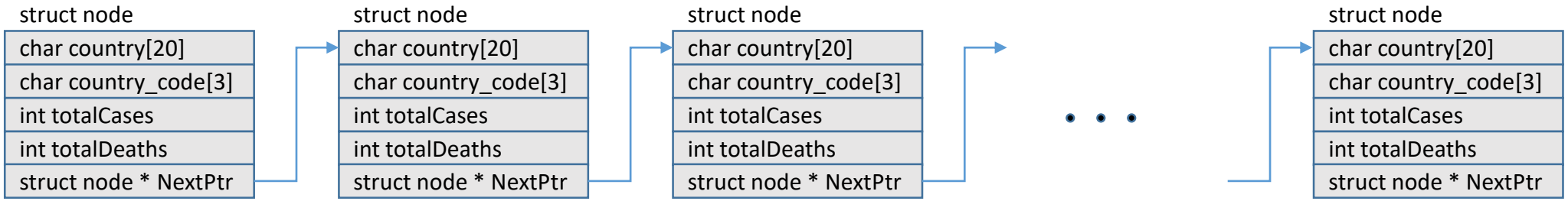
Introduction

- In this assignment, you will write a program to process a data file according to some specification.
- You are provided with a sample data file named "`covid-19.csv`".
- This file consist of covid-19 data downloaded from WHO:
  - This file is in csv format (with a space " " as common separator)
  - The file contains data from
    - Column 1 – Date
    - Column 2 – Country code
    - Column 3 – Country
    - Column 4 – New infection cases on that date
    - Column 5 – New death cases on that date
    - 3 Jan 2020 to 11 Nov 2020
    - e.g. on 3 Jan 2020, there is 0 new case and 0 new death in Afghanistan
    - This file contains over 70,000 lines of data
    - For 235 countries
    - You may also open it with Excel



The Task

- You are supplied with a code.
- You need to build your code based on the supplied code.
- What you need to do:
  - Read data from an input file (“covid-19.csv”)
  - Calculate total infection cases and total death cases for each country
  - Form a linked list as follow:



  - One node, one country
  - totalCases represents the total infection cases from day 1 to the last day in the data for one country
  - totalDeaths represents the total death cases from day 1 to the last day in the data for one country
  - NextPtr points to the node for next country

The Task (cont.)

- After you constructed the linked list, you need to:
  - Search the linked list for the country with maximum infection case, and then point the pointer "`node * maxTotalCases`" to this node
  - Search the linked list for the country with maximum infection case, and then point the pointer "`node * maxTotalDeaths`" to this node
  - Implement the 3 functions:
    - `int Length(node * chain)`
      - This function takes a pointer to a linked list as parameter, and calculate the length (i.e. number of node) in the linked list as return value.
    - `void PrintALL(node * chain)`
      - This function takes a pointer to a linked list as parameter, and then print out all the data in the linked list. This function does not have any return value.
    - `node * Find(node * chain, char * input)`
      - This function takes a pointer to a linked list and a character string as input, and then find if there ia any node that matches the country_code with the character string. The function returns a pointer pointing to the node if a match is found, and returns 0 if no found.

Assumptions

- Regarding the datafiles:
  - The data file provided is only a sample data file. Your program should work with any data file with the same format
  - You should not assume that there is 235 countries in the data files (i.e. I may test your code with data file containing 3 countries, or 300 countries)
  - You may assume that the maximum number of countries is 500
  - You should not assume that the file contains 314 days of data (i.e. I may test your code with data file containing 3 days of data, or 500 days of data)
  - The data for each country in the file is always in consecutive order (i.e. if there are 314 days of data in the file, Line 1-314 contain the data for the first country, Line 315-628 contain the data for the second country, etc).

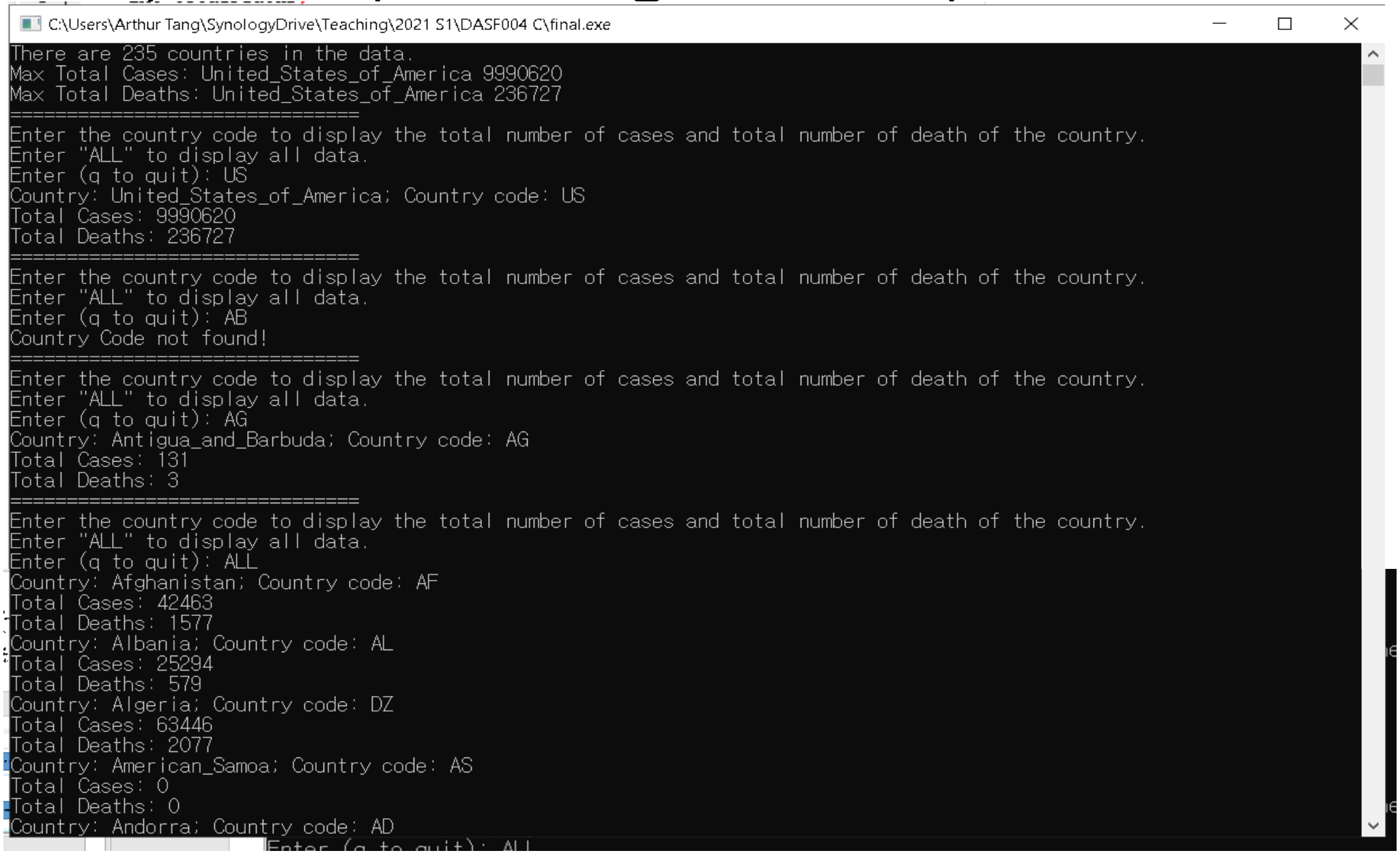
final2.cpp | FinalA.cpp

```
7       int totalCases;
8       int totalDeaths;
9       struct node * NextPtr;
10  };
11
12  // This function takes a pointer to a linked list as parameter, and calculate the length (i.e. number of node) in the linked list as return value.
13  int Length(node * chain)
14  {
15  }
16
17  // This function takes a pointer to a linked list as parameter, and then print out all the data in the linked                    any return value.
18  void PrintALL(node * chain)
19  {
20  }
21
22  // This function takes a pointer to a linked list and a character string as input, and then find if there is any node that matches the country_code with the
23  // character string. The function returns a pointer pointing to the node if a match is found, and returns 0 if no found.
24  node * Find(node * chain, char * input)
25  {
26  }
27
28  int main(void)
29  {   node chain[500];            // The linked list!
30      node * maxTotalCases=0;     // A pointer pointing at the node with max total infection cases
31      node * maxTotalDeaths=0;    // A pointer pointing at the node with max total death cases
32      node * head=&chain[0];      // A pointer pointing at the first node of the linked list
33
34      /* Fill in your code here.
35      Your code should:
36          - Read data from the data file ("covid-19.csv").
37          - Calculate the total infection cases and total death cases for each country.
38          - Construct a linked list based on the calculated results.
39      */
40
41      printf("There are %d countries in the data.\n",Length(chain));
42      printf("Max Total Cases: %s %d\n",maxTotalCases->country,maxTotalCases->totalCases);
43      printf("Max Total Deaths: %s %d\n",maxTotalDeaths->country,maxTotalDeaths->totalDeaths);
```

Make these 3 functions

Fill in your code here

Sample output using the sample data file



C:\Users\Arthur Tang\SynologyDrive\Teaching\2021 S1\DASF004 C\final.exe

```
Total Deaths: 830
Country: Viet_Nam; Country code: VN
Total Cases: 1226
Total Deaths: 35
Country: Wallis_and_Futuna; Country code: WF
Total Cases: 1
Total Deaths: 0
Country: Yemen; Country code: YE
Total Cases: 2070
Total Deaths: 602
Country: Zambia; Country code: ZM
Total Cases: 16997
Total Deaths: 350
Country: Zimbabwe; Country code: ZW
Total Cases: 8610
Total Deaths: 255
==============================
Enter the country code to display the total number of cases and total number of death of the country.
Enter "ALL" to display all data.
Enter (q to quit): ABC
Country Code not found!
==============================
Enter the country code to display the total number of cases and total number of death of the country.
Enter "ALL" to display all data.
Enter (q to quit): AQ
Country Code not found!
==============================
Enter the country code to display the total number of cases and total number of death of the country.
Enter "ALL" to display all data.
Enter (q to quit): q
```

- A set of testing input files was provided to you for testing purpose, but your program should work with any input files according to the specifications
- Submit your source code to iCampus before Wednesday 2 June 2021 18:00 pm
- You may submit multiple version, and only the latest version will be graded
- Late submission will not be graded and a zero score will be given
- In case of technical difficulties for iCampus server, you may submit it as email attachment before the deadline (atang@skku.edu)
- You should make back up to your work, just in case of accident.
- You should double check your submission. Wrong submission will not be an excuse for late submission.