Computer Engineering Department

TED University

CMPE 252 C Programming, Spring 2020

HOMEWORK 2

14.04.2020, Tuesday

DUE DATE: 28.04.2020, Tuesday , 23.59

Send your C code file via Moodle.

This is an individual work. No team work is allowed.

Similarity check will be applied to submitted codes.

A NEW SOFTWARE TOOL FOR QUERYING CITY DATA?

In the airline company that you are working as an intern, you became famous by your success because the software tool (HW#1) that you have developed which reads the world cities database and successfully returns sorted list of nearest cities in a given range was very successful and you got compliments for your work done !

Therefore, another department of this airline company comes to you and asks that they also have a world cities database file and tell you that their upper management frequently asks them to prepare reports about populations of cities by querying from this database for the given countries with their names or sometimes only a part of the country name is given. To prepare these reports, they tell you that they manually look through the list one by one which results in office work of day and night, long overtimes and they sometimes can not even go to their home for sleeping.

You tell them, “don’t worry, the software tool that I will develop will generate your reports in less than one second. Let’s get to work”. They are too happy and told you that your program is simply required to;

- read their input World Cities Database File,
- search the city records which belong to the countries queried by a given search string,
- output those cities which are SORTED alphabetically in ASCENDING order by the City Names along with their corresponding information to an output file which is exactly in the same format with the input World Cities Database file,

REQUIREMENTS

1. **When you check the World Cities Database file they provided, you get surprised since the file they are using is different than the previous assignment so that;**
   - **The file is a CSV – Comma Separated File [**https://en.wikipedia.org/wiki/Comma-separated_values**] where each column of data are separated by a ‘,’ (comma) delimiter.** *(When you ask why they are using this format, they reply that this type of file can be opened by MS EXCEL software which they can work on more efficiently.)*
   - **The file looks like as below when you open it in a Text Editor.**

```
ID,CITY_NAME,CNTRY_NAME,STATUS,FIPS_CNTRY,POP,POP_RANK,POP_CLASS
1,Kamphaeng Phet,Thailand,Provincial capital,TH,58787,6,50;000 to 99;999
2,Cabinda,Angola,Provincial capital,AO,66020,6,50;000 to 99;999
3,Phichit,Thailand,Provincial capital,TH,35760,7,Less than 50;000
4,Onjiva,Angola,Provincial capital,AO,10169,7,Less than 50;000
5,Kimberley,South Africa,Provincial capital,SF,142089,5,100;000 to 249;999
6,Long Xuyen,Vietnam,Provincial capital,VM,158153,5,100;000 to 249;999
7,Krasnoyarsk,Russia,Provincial capital,RS,927200,3,500;000 to 999;999
8,Nakhon Sawan,Thailand,Provincial capital,TH,-999,7,Less than 50;000
9,Hlotse,Lesotho,Provincial capital,LT,47675,7,Less than 50;000
10,Benguela,Angola,Provincial capital,AO,151226,5,100;000 to 249;999
```

The first line of the file is the header line which includes the titles of each column of the data file which are:

ID: Unique identifier of the cities in ascending order.

CITY_NAME : Name of the City (as you may notice, some city names are composed of more than one words separated with blank characters, e.g. "Kamphaeng Phet" in Thailand)

CNTRY_NAME : Name of the Country that the City belongs to. (as you may notice, some country names are composed of more than one words e.g. "South Africa")

STATUS: Status information of the city such as whether it is a "provincial capital", "national capital", "other" etc.

FIPS_CNTRY: International character code (a 2-character code) compliant to Federal Information Processing Standard (FIPS).

POP: Population of the city

POP_RANK: Population rank of the city

POP_CLASS : column providing information about the population categorization of the city.

2. **Your program shall be executed from the command line as below using 3 command line arguments:**

`>program.exe WorldCitiesHw2.CSV <CountryNameSearchString> <OutputCSVFileName>`

where, each command line argument is described as below:

`WorldCitiesHw2.CSV` : 1st argument is the world cities database filename which is provided.

`<CountryNameSearchString>` : 2nd argument is the search string for the country name(s) that are searched to query the requested city records.

The search string can be composed of:

- **An exact name of a country e.g. `THaiLand` (case <u>insensitive</u>) which queries that country (i.e. Thailand) only OR**
- **A substring concatenated with * character e.g. `tur*` (used for querying all the country names starting with "tur" substring, such as "Turkey", "Turkmenistan", etc. *(notice that the search was case–insensitive (lowercase or uppercase letters should be considered equal when searching))***
- **`*` character only, which means, you need to search all countries.**

Note that <u>no other options</u> are required to be implemented such as `*land` or `T*Y` , which means countries ending with "land" or starting with "T"and ending with "Y" substrings. You don't need to implement such other cases except the 3 cases given above.

`<OutputCSVFileName>` *:* 3rd argument is the Output CSV filename whose <u>format should be exactly the same as</u> the input World Cities Database file, <u>including the header line</u>, containing the same comma separated records for the resultant city records queried by their country names.

e.g.: below are some sample cases where the program can be executed from command line prompt :

```
>.\HW2SOL.exe WorldCitiesHw2.csv THaiLand Out1-THaiLand.csv

>.\HW2SOL.exe WorldCitiesHw2.csv tH* Out2-tHstar.csv

>.\HW2SOL.exe WorldCitiesHw2.csv * Out3-star.csv

>.\HW2SOL.exe WorldCitiesHw2.csv a* Out4-Astar.csv

>.\HW2SOL.exe WorldCitiesHw2.csv xxx* Out5-nocntry.csv
```

Below screenshot shows <u>a part</u> of the Out1-THaiLand.csv file corresponding to the executing the program with parameters given above in the first line.

```
ID,CITY_NAME,CNTRY_NAME,STATUS,FIPS_CNTRY,POP,POP_RANK,POP_CLASS
2536,Ang Thong,Thailand,Provincial capital,TH,-999,7,Less than 50;000
33,Bangkok,Thailand,National and provincial capital,TH,5104476,1,5;000;000 and greater
187,Buriram,Thailand,Provincial capital,TH,-999,7,Less than 50;000
39,Chachoengsao,Thailand,Provincial capital,TH,49741,7,Less than 50;000
2528,Chainat,Thailand,Provincial capital,TH,15469,7,Less than 50;000
185,Chaiyaphum,Thailand,Provincial capital,TH,58350,6,50;000 to 99;999
```

Note that; CodeBlocks IDE debugger does not work correctly with * character as a standalone 2nd argument. You may need to call that test case as below using double quotations if you are using CodeBlocks:

```
>.\HW2SOL.exe WorldCitiesHw2.csv "*" Out3-star.csv
```

3. **WHAT WILL YOUR PROGRAM COMPUTE?**

- **Read the input World Cities Database File records whose filename is given as the 1st argument to your executable from the command prompt line.**

- **Search the city records which belong to the countries queried by the 2nd Argument which is called <CountryNameSearchString>**

- **Output those cities which are SORTED alphabetically in ASCENDING order by the City Names along with their corresponding information to the Output CSV File, whose name is provided as the 3rd Argument.**

4. **SAMPLE OUTPUTS.**
   Along with this HW2 description, you will be given 5 debug test case output files (*Out1-THaiLand.csv, Out2-tHstar.csv, Out3-star.csv, Out4-Astar.csv, Out5-nocntry.csv*) generated for your own debugging purposes. Please make sure that your output files are "exactly the same" as these output files corresponding to below sample execution of our ground truth software so that you will be sure that your solution and the output format is correct.

```
.\MYHW2SOL.exe WorldCitiesHw2.csv THaiLand Out1-THaiLand.csv
.\MYHW2SOL.exe WorldCitiesHw2.csv tH* Out2-tHstar.csv
.\MYHW2SOL.exe WorldCitiesHw2.csv "*" Out3-star.csv
.\MYHW2SOL.exe WorldCitiesHw2.csv a* Out4-Astar.csv
.\MYHW2SOL.exe WorldCitiesHw2.csv xxx* Out5-nocntry.csv
```

5. **You can use any available string sorting algorithm in your source code in the literature, but your source code is expected to be comprised by only 1 source file (so all your functions including the sorting function should be in the same file).**

! GRADING

For grading, your executables will be executed using ;

- **5 debug test case parameters given along with the HW description,**
- **5 more other interesting test cases that will not be provided to you,**

and exactly the same output with the ground truth solution will be expected from your software to get full grade.

HINTS

- ***In this homework, you are mainly expected to use your skills from Chapter 8 "Strings" of your book. Therefore, please first go over the chapter contents and related material (including slides, lecture recordings, gnu c manual and other online material etc.) before designing your solution.***
- ***Check out string.h library functions, especially you may find strcpy, strcmp, strncmp, strtok, strstr functions and similar others useful, if applicable to your solution.***
- ***Besides, check also character functions available in ctype.h.***
- ***You may also consider using fgets, fputs functions under stdio.h, if applicable to your solution.***
- ***You can use any string sorting algorithm, of course your solution should also continue tracking the line numbers or indexes of those cities in the database file, so that you will be able to output that whole line to the output file.***
- ***Don't forget to also write the header line to the output file.***

Have Fun!.