CSCE 221 Cover Page
Homework Assignment # PA - 4

First Name: Dilanka Last Name: Weerasinghe UIN: 126007816

User Name: dweerasinghe E-mail address: dweerasinghe@tamu.edu

Please list all sources in the table below including web pages which you used to solve or implement the current homework. If you fail to cite sources you can get a lower number of points or even zero, read more in the Aggie Honor System Office http://aggiehonor.tamu.edu/

Type of sources					
People					
Web pages (provide URL)					
Printed material					
Other Sources					

–I DID NOT TEST THE CODE ON THE SERVER. WITH THE CURRENT MAKEFILE IT WILL RUN ON MY LOCAL MACHINE. IF IT DOES NOT WORK ON THE SERVER I CAN DEMO IT IN PERSON.

1. (PA-4) Description.

   (a) The purpose of this project is to navigate through a csv file with regex and read a roster. The roster will hold the values for the students UIN and their names, emails etc. A different index holds the quiz grades for the students. Using the power of regex and a HashTable we can obtain the grades from the csv files and stored them in a hash table. From there they can be imported into the students index at the correct location.

2. Data structures and algorityms used by the program

   (a) The most important data structures used in the program is a hash table. This specific hash table used the chaining method. With this chaining method the act of storeing values is very efficient, depending on the size of the hash table it can be very effecient for searching as well. The hash table has a hashing function that will determine the placement of the value in memory depending on its key. The way that I constructed the has table was by making a Node to hold the value of the key and a pointer to the Templated object. This Node is the basis of the element. The actual table is made up of an array of vectors.

   (b) Vectors were the next data structure I used. I thought that by using a vector iteration through the array would be simple becuase that data structure has its own iterator. This is the reason why I did not end up using a linked list. It is also easy to append another object to the end of the list with the function push_back().

(c) I dont know if an array is considered a data structure but I used it because it is an easy pointer that can be dereferenced easily and because of the properties of the hash table the size does not need to be changed.

3. Description of input and output data. List all restrictions and assumptions that you have imposed on your input data and program.

   (a) The biggest restriction that the input and ouput data places on the program is processed through the regex class. The way the regex function (not actual code funcitons) is designed is it will look for any item *. twice, then look for any item with a parentesis (*.) twice. This resulted in the regex::smatch objects to gain the line as a whole in <smatch>[0] then the UIN and the Quiz grade in [1] and [2]. This means that depending on how the input and roster data is submitted the program will have to change the regex files to accomidate for the (.csv) file. As for now it must be in {“name”,”email”,UIN,”quiz”}.

4. How have you tested your program for corrections?

   (a) There were a couple of steps that I took in order to test my program for correctness. In the hash table I made a print function to make sure that the program is actually placing the values in the correct hash. The final check was whether or not the program would give the right values to the right students in the end and that seemed to work.

5. Which C++ features or standard library classes have you used in your program?

   (a) Here are the std lib classes that were using in the making of the program

      i. string
      ii. vector
      iii. list
      iv. iostream
      v. utility
      vi. regex
      vii. fstream

   (b) Here are the ones actually required to run the program

      i. string
      ii. vector
      iii. iostream
      iv. regex
      v. fstream

6. Provide the statistics about the hash table. Are the computational results about the hashing consistent with the expected running time for the hashing algorithm? Justify your answer

   (a) So the running time of a hash table algoritym with chaining is $O(n)$ time. searching through that hash table takes $O(1) + n/m$or $O(n + m)$ where m is the number of elements being searched for. This is consistent with the computational results from my program and this is seen when looking at the print_table() function that I created. You can see all the values being placed into the hash table and it is easy to test to see how long takes if you see the print values as the program is going through.

   (b) In the program it will display the average list length, along with the min and the max. Here is the data with different “m.”

   (c)

m	average	min	max
3	5.667	5	6
5	3.4	2	5
13	1.30	0	3
17	1	0	2
19	0.89	0	3

7. Conclusion

   (a) This assignment was fun and I learned a lot! Thank you Dr.Leyk, Isaiah, Luiz and Lauren!

   (b) The hash table was easy onced I decided to use an array of vectors. The search function confused me because it returned a pointer so that we could send nullptr. This meant I needed to change my node to use a pointer to the value. The regex was confusing but after an explanation it was simple. Everything works correctly.

I certify that I have listed all the sources that I used to develop the solutions/code to the submitted work.
"*On my honor as an Aggie, I have neither given nor received any unauthorized help on this academic work.*"

Your Name (signature) Dilanka Weerasinghe Date 4/13/20