University of Waterloo
CS240 Spring 2022
Programming 2

Due Date: Tuesday, July 26 at 5:00pm

The integrity of the grade you receive in this course is very important to you and the University of Waterloo. As part of every assessment in this course you must read and sign an Academic Integrity Declaration before you start working on the assessment and submit it **before the deadline of July 26** along with your answers to the assignment; i.e. **read, sign and submit P02-AID.txt now or as soon as possible**. The agreement will indicate what you must do to ensure the integrity of your grade.

Note: All code submitted must be your own. Also, at this point, you are expected to be able to debug your code yourself.

The Academic Integrity Declaration must be signed and submitted on time or the assessment will not be marked.

Please read `https://student.cs.uwaterloo.ca/~cs240/s22/assignments.phtml#guidelines` for guidelines on submission. Submit the file `phone.cpp` to Marmoset.

Late Policy: Programming Questions are due at 5:00pm and **no lates are accepted**. Submissions after the deadline will not be accepted but may be reviewed (by request) for feedback purposes only.

Problem 1 [20 marks]

In this programming question you are asked to implement a phone directory that permits lookups in both directions, by using the following two hash tables:

- D_{name} implements *hashing with chaining* and uses names as keys
- D_{phone} implements *cuckoo hashing* and uses phone numbers as keys

Your C++ program must provide a main function that accepts the following commands from stdin. You may assume that all inputs are valid and will begin with `r` and end with `x`; also, all inserts will be successful.

- `i name number` - inserts an entry with the given name and the given phone number into both hash tables.

- `l number` - prints the name associated with phone number followed a newline, or "not found" followed by a newline if there is no such phone number.
- `s name` - prints all phone numbers associated with the name, in lexicographic order. Each pair of phone numbers is separated by a space and the last is followed only by a newline. If no such name exists,"not found" followed by a newline is printed.
- `rh i` - rehashes dictionary i where $i = 0$ means D_{name} and $i = 1$ means D_{phone}. When rehashing, use the smallest prime that is at least 2M + 1 as the new table-size. The `stub.cpp` file provides a method to find this.
- `p i` - prints dictionary i according to the specifications below, where $i = 0$ means D_{name} and $i = 1$ means D_{phone}.
- `r` - initializes or resets both hash tables to be empty.
- `x` - terminates the program.

Details for implementing Dictionary D_{name}

- Start with an empty hash table of size $M = 11$ and use a hash function `hash0` provided in the `stub.cpp` file.

- A name is a string of arbitrary length, with characters in $\{A-Z, a-z\}$ that does not contain spaces. To convert a name into a key you must apply the flattening technique and Horner's Rule to avoid overflow (Module 7 Slide 24); use $\mathbf{R = 255}$.

- To facilitate printing (in searches by name), order the chains so that phone numbers can be printed lexicographically with a single traversal of the chain (with no sorting required). Instead of simply inserting at the front of the chain, maintain a chain sorted by number (smallest number first) and then insert the new number into its appropriate location in the chain. Also, do not apply the MTF heuristic. The chains must be implemented as linked lists. You may use `forward_list` from the STL or implement your own (its okay to reuse your implementation from the first programming question).
- To print D_{name}, print M + 1 numbers, where a pair of numbers is separated by a space and the last number is followed only by a newline. The first number is M itself. The next M numbers are the lengths of the M buckets of the hash table.

Details for implementing Dictionary D_{phone}

- Start with two initially empty hash tables, each of size $\mathbf{M = 11}$, using the hash functions `hash0` and `hash1` provided in the `stub.cpp` file. Implement the two hash tables using vectors.

- A phone number has the form $(abc)def - ghij$ where each of $a - j$ are in $\{0, \ldots, 9\}$ where digits may be repeated. To convert a phone number into a key, remove all non-digits and interpret the result as a number in base-10 (use a `long` to store the key).
- When rehashing, insert the items into the new tables by traversing the old first table, in order, starting at index 0 followed by similarly traversing the second table.
- To print D_{phone}, print 2M + 1 numbers, where a pair of numbers is separated by a space and the last number is followed only by a newline. The first number is M itself. The next M numbers are either 0 or 1, depending on whether the slots in the first table are empty (0) or not (1). The next M numbers are either 0 or 1, depending on whether the slots in the second table are empty or not.

Further Implementation Details

- Table sizes will never exceed `INT_MAX`.
- Phone numbers are unique. Names are not.
- You may use `string`, `stringstream`, `forward_list` and `vector` from the STL.

Place your entire program in the file `phone.cpp` and **clearly label (with comments) each of the hash table implementations and operations so it is easy for the markers to find them in your file**.