**SPRING 2012 CS 401 (Hoffman) Lab #9**

**Lab #9 will read in a file of strings and determine if the set of strings had any duplicates.  
It will then generate a histogram of the strings mapping length to frequency.**

**Background**

An unsolved problem in CS is the question of how many unique strings can be formed from a NxN grid of unique characters. Assuming the grid below consisting of 4 unique characters

**a b**

**c d**

It is possible to fom 64 unique strings. If you increase the grid size to 3x3 like this:

a b c

d e f

gh i

you can form 10,305 unique strings. The condition is that you may not jump over a cell or use a cell twice.

There is no known formula (closed form expression) to compute the number of possible unique strings. The only known way is to write a program that generates and counts them. I wrote just such a program and have captured the output of strings generated. One of the first steps to veriying my program is to verify that all the strings are unique and there are no dupes. This is where you come in.

* INPUT FILE: [2x2output.txt](http://www.cs.pitt.edu/~hoffmant/401/lab-09/2x2output.txt)
* INPUT FILE: [3x3output.txt](http://www.cs.pitt.edu/~hoffmant/401/lab-09/3x3output.txt)

The above files contain 64 and 10,305 strings respectively. Your job is to read those strings in and then print out the count of strings and how many dupes were found. Once that is done you must generate a histogram that maps the different string lengths found to their frequency of occurance.

* + You must a use the most efficient conatainer possible ( w/respect to insert/search) for the uniqueness testing.
  + For the histigram you must use a HashMap of Integer -> Integer

The correct output must match below. This will be hand graded with script assist.

C:\Users\tlh\Desktop\lab-09>java Lab9 2x2output.txt

2x2output.txt contained 64 strings. There were 0 duplicates found.

Histogram of string lengths to number of strings of that length

1 -> 4

2 -> 12

3 -> 24

4 -> 24

C:\Users\tlh\Desktop\lab-09>java Lab9 3x3output.txt

3x3output.txt contained 10305 strings. There were 0 duplicates found.

Histogram of string lengths to number of strings of that length

1 -> 9

2 -> 40

3 -> 160

4 -> 496

5 -> 1208

6 -> 2240

7 -> 2984

8 -> 2384

9 -> 784

C:\Users\tlh\Desktop\lab-09>