Nick Manolov

Programming Assignment Four

"Simulating Datamarts"

04/05/2017

Nick Manolov

C202

Dr. Hettiarachchi

April 5, 2017

Spell Checker

This assignment was confusing at first but one we figured out that the dictionary was held in an array of LinkedLists everything else was a breeze. The main idea of the project was to load a dictionary into an array of linkedLists by filtering the first character of each word by it's ASCII value. This was done by subtracting 97(ASCII value of the letter A). Using this method an alphabetical array of LinkedLists was built.

The main idea of the project was to use the above mentioned library to compare the words from another file and see which words exist in the dictionary. If they did exist, that means the word is spelled correctly and the appropriate variable that kept count was updated. The tricky part of this process was filtering out the right information. The text file"oliver.txt" was full of numbers and special characters. To get around this we used a regular expression which ignored everything but words.

We saw a couple of things that we would have done differently if we had time. One thing that we would have written differently would have been the matrix population. We felt like if we were to use a different data structure such as a binary search tree the efficiency of our algorithm would improve dramatically.

In conclusion we filtered a huge data file to find certain information that we used to compare against a dictionary and give back results such as how many words were found, not found, how many comparisons were made, and the average of comparisons.

```
run:
There were 914054 found words in this file
The amount of comparisons of words found: 3247710180
There were 64537 words in this file, that were not found.
The number of comparisons for each word not found was: 479317049
The number of words found to comparisons of words ratio was 3553
The number of words not found to comparison to total words ratio was 7427
BUILD SUCCESSFUL (total time: 23 seconds)
```