Bhavin Patel
Assignment 3
Indexer

The assignment would take in 2 parameters after using make to create the exe. It would take in the output file and the input file/directory.

./index output.txt input

The analysis would work with opening the input file and it would read every word of the file that would exists in the directory and add it to the linked list which would contain all the unique words. if the word is hit again, it would increment its frequency. This would be done in O(d) where d is the number of unique words.

The linked list would consist of pointer to next node, and pointer to linked list to the filename which contain that word and the frequency of that word in that file. The linked list of the file struct would contain pointer to the next filestruct and also int frequency which would keep track of how many times that word occurred.

The linked list is added to the front of the list. Every node added.

This program would run in the number of words there are in each file * number of distinct words. Every word which is read in would have to go through the whole linked list to check if it already exists and if the filename already exist. So it would have to parse through 2 linked list and then it would be added as a new node if filename doesnt exist.

So every new word it would have to go through O(n) n = distinct words

For every non-distinct word, it would have have to go through the list of distinct words till it finds itself and then go through other linked list, so O(n+f) where n = distinct words and f = number of distinct filename nodes.

This would have to be done for every word which is read in. So this would be done depending on number of input files + how many words it contains.

Also, printing to the file would require parsing through the whole linked list. So it would be O(nf). where n = number of distinct words and f = number of file names.

If every word exists in every file then all the distinct words and its file names would be printed out.