Allon Finezilber

Michael Ferrera

CSC 214

Assignment 3 – Indexer

In this program we traverse through a directory reading all files and subdirectories inside of it using the built in library FTW(). During the traversal we read contents of the files in the directory and tokenize the contents of the file using the tokenize() function and then indexing tokens in ascending order keeping track of how many times a specific token was used in each file it exists inside and then storing and sorting inside of a hashtable.

We also utilitzed our hashtable using UTHash with its uthash.h file. The functions runs at $O(n * 2t + f^2)$ runtime where n represents the number of files and directories in the specified directory. t represents the number of tokens that are contained in the file which are then added and printed from the hash table. f in this case represents the number of nodes which keep track the files and their counters.

USAGE:

1) Run Make to make the executable

2) ./index (file to write to) (file or directory to tokenize/traverse)