Eric Bronner
Tobias Perelstein
CS214 – Brian Russell
Pa3 – Indexer

Our Indexer is made up of three main .c files:

1. dirtest.c which includes the main method and recursively crawls through the files and directories within the starting directory,
2. tokenizer.c which tokenizes all words within non-empty files and calls the add_node() function from hashadd.c
3. and hashadd.c which converts tokenized words to nodes and adds them to the hash table. Nodes contain Linked Lists which keep track of the different files the words appear in and the number of times they appear. This file also prints out the nodes to a file.

Additionally, our hash table comes from UTHash's uthash.h file.

Functions run in O(n * 2t + f^2) time, where n is the number of files/directories in the input directory, t is the number of tokens in the file, which are all added to and printed from the hash table and f is the number of nodes that keep track of the files and counters, which are sorted.