Group: Douglas Rudolph, Nandan Thakkar
Professor: Dr. Francisco
Course: Systems Programming - CS:214
Due Date: March 29th, 2017
Assignment: Assignment 2

Design:

The goal for my indexer was modularize the code base. The code base for my project consists of 5 files - each of which handles their own part of the project

Files and Descriptions:

1. **modules.c**
    i. *Modules.c* is a file that containers all of the helper functions. It contains functions like *split*(a function that will split out garbage tokens with actual tokens), *addToken*(a function that will add the important data from a file to a linked list of files), *readFile*(a function to read through a file and store is as a string, to then be *split*), and *destroyTable*(goes through the global hash table and deletes all token references).
2. **modules.h**
    i. *Modules.h* is a header file that stores all the functions as well as global global hash table to where all the tokens get stored
3. **driver.c**
    i. *Driver.c* is a file that has two functions - *main and onStartUp. Main* simply runs *onStartUp,* and *onStartUp* handles all the error checking with the data that is passed in within terminal.

Time Complexity:

Reading in a file takes O(M*3N). 1 iteration to *find* the size of the file, 1 iteration to *rewind* back to beginning of the file, and then 1 iteration to then *read* the data to the end of the file times the files. (M is the amount of files and N is the amount of iterations per file)Then generalized to **O(M*N)**.

Writing the data to the *File_HashTable* Takes **O(N)** due to the fact that each file has its tokens stored during the process of reading a file.

Sorting the tokens from a file takes **O(N^2)** due to the fact that I am using insertion sort due to time constraints.

Counting the amount of unique tokens takes **O(N)** because the data is already sorted, thus finding the groups of unique tokens can be done in one iteration.

Writing to a file takes **O(N)**, because we have to go through all tokens that were read in and write the XML tags.