

This program is a general textual search application. There are three input files that included document file, index file and stop word file. The last two files are optional. The index file saved sections' index of the input document. The stop word file contains some words from document need to be ignored when searching some words.

There are three types of data type which are implement to the Section, Line, and Word. The Word is a single linked list, which contains the column number of the first character at current line, the string content, string size, Boolean to indicate a stop word and the next Word. The Line is a single linked list which contains line number, first Word at this line and the next line. The Section is a double linked list which contains first line, title of current section, next section and previous section.

This program will load the input file and save single word into Word. Create a new Line when finished current line. Create a new Section when the current section is finished. This program will load the first section, the first line in this section and the first word in this line if there is no a section name is indicated. So that the worst case is $O(n)$ where n is the number of words in this document. The reason of using double linked list is that program needs to get the first line number and the first line number of the next section to determine how many lines this section has. It is better if the stop words using trie to store stop words because it is faster to match current word and stop words.