CS132 – Computer Science II

Spring 2024

Mr. MacKay

**Project 3** – due Wednesday, May 1st, at the **beginning** of CLASS.

**Description**: A *text concordance* is an alphabetical listing of all the distinct words in a document. The basic storage structure for a *concordance* is an array of Binary Search Trees, one tree for words starting with A, one for words starting with B, etc.

* You will write a program that reads a text file and constructs the *concordance* that contains the distinct words in the file **and**, for each word, the line number of the first occurrence, and how many there are in the file then allows the user to search the *concordance* for a particular word, and to print out a list of all the distinct words in the *concordance.*
* Words that are upper case (or have uppercase letters in them) are treated the same as the word written with all lowercase letters.
* Words in the list will NOT have ANY non-alphabetic characters at the beginning or end of the word
  + i.e. the word **hello!!** and the word **---Hello** will both be saved as the word **hello**
* The program will ask the user to input the file name at run-time
* The program will ask the user for words to search for until the user indicates they are done
* After searching for words, the program will display a list of the distinct words in alphabetical order, the line number of their first occurrence, and how many there are in the file

You will email me a copy of all your files