

lab4 15-111 Sections G and H

Lab 4

Start: Monday 10.01.01

Due: wednesday 10.10.01

Goal: FileInput class, frequency counts && sorting

Course Web Site:

www.andrew.cmu.edu/~guna/15-111

Assignment:

We have built a FileInput class for you to use to read text files. You will use an object of this class to read the scrabble dictionary (bigdata.txt) and determine the frequency of lengths of the words in the dictionary.

Specifications:

- All code for this lab can be written in the Driver.java class. You do not have to build a new class for this lab but you can do so if you want to.

- Driver's main method must NOT contain all of the code for this lab. Method main must call other methods or pass messages to read the file, print frequency counts, and sort the frequency counts.

- You will need an array of type int to tally the frequencies of the lengths of the words in the dictionary.

- There are no words longer than 30 characters.

- The scrabble dictionary has over 170,000 words in it. Each word is on a single line of the file. You will read each word in the file into a String object and you should store the words in an array or other data structure such as a vector.

- For each word, determine the length of the word and increment the corresponding element of the frequency array.

- When the file reading is completed, you must print the frequency array and a label for each value.

You should print only non-zero values of the frequency array.

- You must sort the original array in ascending order (small to big) of the word.

- You must print the sorted array of frequencies with a label for each value again printing only non-zero values.

Practical Concerns

This lab is also due in 10 days. The purpose of this lab is to give you experience reading files using our FileInput class. It also introduces sorting for the first time this term. The sort you use is up to you. If you do not know one, your text explains three sorting algorithms to you starting on page 104. We told you the book would come in handy. .

Handin

Write the output to a file bigdDataSorted.txt

Some Last Thoughts

You will do a great deal of file reading so spend some time working with the FileInput class.

Sample output

The words in alphabetical order are:

andy

awesome

...

bind

blind

....

// print all 170,000 words

The frequencies are

length

1

2

count

12

345 etc..