Computer Science 250

Project 1: Word Lengths in a File

**Due:** Monday, Feb. 11, at the beginning of class

**Reminders:**

* You are not allowed to work with, consult, or seek help from anyone except the tutor or instructor as you write your program.
* You must use "ice" to develop your program, along with the emacs editor and gcc compiler. Neither the tutor nor I will give any assistance in developing or debugging unless you are using ice with those development tools.
* Your program will be evaluated using the Linux environment and the gcc compiler.

Write a program that prints a histogram of the lengths of words in a file. For the purposes of this program, a "word" is a sequence of non-whitespace characters preceded or followed by whitespace. The whitespace characters are the space, newline, tab, vertical tab, formfeed, and carriage return characters: ' ', '\n', '\t', '\v', '\f', and '\r'. You can use the isspace() function from ctype.h to test for these characters in an if statement. Isspace() will return true if its character parameter is one of those six whitespace characters, and false otherwise.

The name of the input file will be given on the command line. You should assume that the input file contains at least one word, and that no word in the file has more than 15 characters.

When you are printing the histogram, the length of each bar in the histogram should correspond to the percentage of words that have that length. One-hundred percent should be indicated by 60 asterisks, with other percentages represented proportionally.

As an example, assume that the file example.txt has the following contents:

The quick brown fox jumped over a lazy dog.

Now is the time

for all good persons

to come to the aid

of their country.

rubber baby buggy bumpers

toy boat - toy boyt - troy boyt

Here is the output your program should give when run on this file:

 1: \*\*\*\*

 2: \*\*\*\*\*\*

 3: \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

 4: \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

 5: \*\*\*\*\*\*

 6: \*\*\*

 7: \*\*\*

 8: \*

 9:

10:

11:

12:

13:

14:

15:

Use good programming practices (meaningful identifier names, proper indentation, proper use of constants, comments, etc.) when writing your code. Points will be deducted for code that does not follow good programming practices, even if it gives the correct output.

**What to turn in:** When you are ready to turn your program in, print out a hard copy of the source file to submit. Then, email me a copy of the source file as an email attachment. Make sure the subject of your email is "CS 250 – Project 1 - <your last name, your first name>". Both the hard copy and email must be received by the due date to be considered an on-time submission.

**How your program will be evaluated:** Your program will be graded on three criteria

1. The quality of your code.
2. The correctness of your algorithms.
3. Whether or not your program gives the correct output on a series of test cases. Make sure to thoroughly test your code on a variety of different input files. When testing, try to come up with input files that are legal, but test all the different things that could happen in your code.