Augustana Campus, University of Alberta
AUCSC 380 Operating System Concepts
Winter Term, 2019

Programming Assignment 1
UNIX Programming with the C Standard I/O Library

Due date: Friday, 2019 March 1, by midnight.

Objectives:
- To become familiar with programming in C in a UNIX environment.
- To gain experience in the use of the C standard I/O library.

Assignment:
Write a C program, `charfreq.c`, that counts the occurrences of each of the alphabetic characters in an input stream, then prints a table of character frequencies. The program should treat the uppercase and lowercase variants of each character as equivalent (i.e., both ‘`a`’ and ‘`A`’ count as the same alphabetic character). The program should also count all other characters as a group, and should maintain or calculate the total number of characters read.

Your program should use the C Standard I/O Library for input and output. It should read from each of the files specified as command-line arguments or, if no filename arguments are specified, from the standard input. If more than one file is specified on the command-line, a single set of character counts is maintained across all input files, and only one set of results is printed.

The results should be printed to the standard output in a list, with the letters ‘A’ through ‘Z’ in the leftmost column, the count of each character in the next field, and the percentage that each character count represents of the total character count in a trailing field, with percentages shown to 1 decimal place accuracy. The counts of other and total characters should be shown on the subsequent output lines. Character counts and percentages should be right-aligned in their fields. For example, the following is a sample output in the desired format; it shows the counts of the characters in a plain-text version of Herman Melville’s *Moby-Dick* (1851) from Project Gutenberg (with prefatory material omitted). This file is available in directory `/usr/local/courses/aucsc380/p1` on `byrd.augustana.ualberta.ca`.

```
A:  76197 ( 6.4%)
B:  16560 ( 1.4%)
C:  21964 ( 1.8%)
D:  37583 ( 3.2%)
E: 114704 ( 9.6%)
F:  20424 ( 1.7%)
G:  20446 ( 1.7%)
H:  61543 ( 5.2%)
I:  64250 ( 5.4%)
J:   1058 ( 0.1%)
K:   7925 ( 0.7%)
L:  41980 ( 3.5%)
```

```
       M:  22875 ( 1.9%)
       N:  64430 ( 5.4%)
       O:  68008 ( 5.7%)
       P:  16795 ( 1.4%)
       Q:   1540 ( 0.1%)
       R:  50935 ( 4.3%)
       S:  62968 ( 5.3%)
       T:  86240 ( 7.2%)
       U:  26149 ( 2.2%)
       V:   8413 ( 0.7%)
       W:  21734 ( 1.8%)
       X:   1006 ( 0.1%)
       Y:  16569 ( 1.4%)
       Z:    640 ( 0.1%)

       Other characters: 258527 (21.7%)
       Total characters: 1191463
```

Alternatively, you may format the output in a more compact three-column table:

```
   A:  76197 ( 6.4%)     J:   1058 ( 0.1%)     S:  62968 ( 5.3%)
   B:  16560 ( 1.4%)     K:   7925 ( 0.7%)     T:  86240 ( 7.2%)
   C:  21964 ( 1.8%)     L:  41980 ( 3.5%)     U:  26149 ( 2.2%)
   D:  37583 ( 3.2%)     M:  22875 ( 1.9%)     V:   8413 ( 0.7%)
   E: 114704 ( 9.6%)     N:  64430 ( 5.4%)     W:  21734 ( 1.8%)
   F:  20424 ( 1.7%)     O:  68008 ( 5.7%)     X:   1006 ( 0.1%)
   G:  20446 ( 1.7%)     P:  16795 ( 1.4%)     Y:  16569 ( 1.4%)
   H:  61543 ( 5.2%)     Q:   1540 ( 0.1%)     Z:    640 ( 0.1%)
   I:  64250 ( 5.4%)     R:  50935 ( 4.3%)
   Other characters: 258527 (21.7%)
   Total characters: 1191463
```

You may add headings, etc., as desired.

If one of the specified input files is not readable or does not exist, an error message should be written to the standard error stream, and the program should terminate with an exit status of `EXIT_FAILURE` without producing other output.

Error messages due to system errors should include the system-generated error message using the `perror` subroutine. Other error messages should be written to the standard error stream.

If the program completes successfully, it should exit with a status of `EXIT_SUCCESS`.

Lab Report:

Submit the C source code for your solution in electronic form via *eClass* by midnight on the due date specified. Assignments will be accepted after the due date and time, but late submissions will be assessed a penalty of **1% per hour** or portion of an hour.