ISA542300 Software Quality Assurance

Homework 2.B – Group Exercises (70%)

1. *Concordance.cc* is a C++ program to make concordances for text files, whose usage is explained in Fig. 1. A sample input/output set is presented in Fig. 2. Although *concordance* worked very well, it has been recently hacked and seeded roughly 63 faults. According to the hacker, some faults are obvious while some may be hard to identify and even cause security problems. Your task is to form an inspection team and identify potential faults.
   (a) Generate a complete inspection report which includes all the information in Appendix A (Formal Technical Review Forms.pdf). The grading will depend on the completeness of the report, the total defects found, and the number of identified defects unfound by other groups. **(50%)**
   (b) In addition to the fields in Appendix A, add the inspection metrics including *estimated total defects*, *yield*, *defect density*, *inspection rate*, and *defect finding efficiency* at the end of the inspection report. For each metric, briefly show the calculation. **(20%)**

```
This is Concordance, a program to make concordances for text files.
$Id: concordance.cc,v 0.5 1996/11/26 23:49:48 ralph Exp ralph $
Copyright (C) 1996 Ralph L. Meyer, Spotswood, NJ 08884.
Concordance is licensed to users only under the conditions stated in the GNU public
license that comes with the distribution of this program.
Concordance comes with ABSOLUTELY NO WARRANTY; for details see the GNU public
license that accompanies this program.  You are welcome to redistribute this
software under certain conditions; for details see the GNU public license. Send
BUG reports to: meyer@princeton.edu.

USAGE:
concordance [-p[q][n:num] or -l[q][n:num]] or -s[q][num]] filename [outfile]
filename is the usual /filepath/fname to the file for which you wish to make a
concordance.  -p, -l, and -s are switches by which you can tell the program whether
you wish to concordance words in the file by the page on which they are found,
by the line on which they are found, or by the stanza in which they are found.
The program considers a page to be delimited by ASCII char 12 (decimal) - FF,
the formfeed character.  Every formfeed character passed will increment the page
counter by 1 page.  To indicate stanzas, stanza numbers in the file must have the
form: number>. i.e.: 1>  for stanza 1, 2> for stanza 2, and so forth.  Stanza numbers
must precede the stanza they delimit.  Line numbers are incremented by the Line
```

feed character, ASCII character 10 (decimal). Default operation (no switches indicated on the command line) is to indicate the page number in the file on which a word will be found. Note that only one page/line/stanza switch at a time is allowed! The addition n:number can be appended to any of the type-of-count switches and the beginning number of the page or line will be changed to the number following the n: (n:num has no effect on stanza counting). If an outfilename is used, there MUST be a switch indicating the count type on the command line. The auxiliary switch 'q' causes only the copyright to be output to stdout. Without 'q' concordance outputs information on the filenames in which it has stored results.

Example:

concordance -pqn:40 MY.TXT MYCORD [Enter]

...would cause the text file MY.TXT to be concordanced by pages with the first page considered to be page 40. No messages would be output to screen but the opening copyright notice. Output would be put in the files MYCORD.wrd and MYCORD.abc . Concordance creates 2 files, one named 'filename.wds', which contains a concordance of all the words in the file 'filename'. The words are listed alphabetically. The number preceding the word is the word's length. The number immediately following the word indicates the number of times that word was used in the document filename. The succeeding numbers indicate the page/line/stanza numbers in which the word was found.

In addition to producing the file 'filename.wds', concordance also outputs a document named 'filename.abc' that contains a listing of the letters of the English alphabet, and the numerals 0-9, whose usage in the document 'filename.ext' has been counted and graphed.

After running concordance, look for these two files named for the file that was concordanced or the name you gave the outfile. They may be examined using the command 'cat', 'less', or by using your favorite text editor.

Please Note that very long documents may need to be separated into pieces to be concordanced. Larger documents should be broken into smaller pieces, each piece concordanced, and the concordances from each piece concatenated.

To get this help screen again, type 'concordance -? [Enter]' at the shell prompt, see the file README, or check the man pages.

Fig. 1. The usage of the concordance program.

Input: README

```
This is a README file.
What is your name?
My name is Susan.
Nice to meet you.
Nice to meet you, too!
```

```
How are you?
I am fine. Thank you. And you?
Fine, thanks.
```

Output 1: README.wds

```
Concordance for file: README

Note: Concordancer considers a word a set of alphanumeric characters
     bounded by whitespace at each end of the set.

Total number of words in file: 34
Initial page of file was page 1
Last page of file was page 1

Word         Number of   Page number
Len  Word:       uses:   Locations:
 1 a.................1 - 1
 2 ............am.....1 - 1
 3 ............and.....1 - 1
 3 ............are.....1 - 1
 4 ............file.....1 - 1
 4 ............fine.....2 - 1
 3 ............how.....1 - 1
 1 ............i.....1 - 1
 2 ............is.....3 - 1
 4 ............meet.....2 - 1
 2 ............my.....1 - 1
 4 ............name.....2 - 1
 4 ............nice.....2 - 1
 6 .........readme.....1 - 1
 5 ..........susan.....1 - 1
 5 ..........thank.....1 - 1
 6 .........thanks.....1 - 1
 4 ............this.....1 - 1
 2 ............to.....2 - 1
 3 ............too.....1 - 1
 4 ............what.....1 - 1
 3 ............you.....5 - 1
 4 ............your.....1 - 1
```

Output 2: README.abc

```
a ...11   9.7% aaaaaaaaaa
 b ....0     0%
 c ....2   1.8% cc
 d ....2   1.8% dd
 e ...14    12% eeeeeeeeeeeee
 f ....3   2.7% fff
 g ....0     0%
 h ....5   4.4% hhhhh
 i ...10   8.8% iiiiiiiii
 j ....0     0%
 k ....2   1.8% kk
 l ....1  0.88% l
 m ....7   6.2% mmmmmmm
 n ...10   8.8% nnnnnnnnn
 o ...11   9.7% ooooooooo
 p ....0     0%
 q ....0     0%
 r ....3   2.7% rrr
 s ....7   6.2% sssssss
 t ....9     8% ttttttttt
 u ....7   6.2% uuuuuuu
 v ....0     0%
 w ....2   1.8% ww
 x ....0     0%
 y ....7   6.2% yyyyyyy
 z ....0     0%
 0 ....0     0%
```

```
 1 ....0     0%
 2 ....0     0%
 3 ....0     0%
 4 ....0     0%
 5 ....0     0%
 6 ....0     0%
 7 ....0     0%
 8 ....0     0%
 9 ....0     0%

Total Characters: 113
```

Fig. 2. The usage of the concordance program.

Notice

COPY will get zero!!

Deadline: May, 14, 23:59

The following situation will surely lead to the deduction of your score:

- Not following the format
- Wrong-spelling in the report

If you have any questions, please feel free to contact TAs ☺