

Programming Assignment #5 – Due Wed, 5 Dec

Problem Statement

Described below are three common file-processing programs that illustrate the power of using a “scripting” language rather than a general purpose programming language like C/C++ or Java. You are to use Python to implement the solutions. You may not invoke the corresponding Unix commands that are available on Unix/Linux that perform these functions, but you can use built-in modules in Python. Each Python program should have the extension “.py”.

Task #1: Printing the contents of a file

In Unix/Linux, the **cat** command is used to print a file to the standard output (see the **cat** man page on Linux). Applying **cat** to the `/etc/passwd` file on a Unix system reads the file and prints it line-by-line. Implement a similar utility called **cat.py**, which optionally takes a command line option ‘-n’ that numbers the lines starting at 1 and an option ‘-s’ that prints the file single-spaced (i.e., does not print consecutive blank lines).

```
linux% python cat.py /etc/passwd
nobody:*:-2:-2:Unprivileged User:/nohome:/noshell
root:*:0:0:System Administrator:/var/root:/bin/tcsh
daemon:*:1:1:System Services:/var/root:/noshell
smb:*:25:25:Sendmail User:/private/etc/mail:/noshell
www:*:70:70:World Wide Web Server:/Library/WebServer:/noshell
mysql:*:74:74:MySQL Server:/nohome:/noshell
sshd:*:75:75:sshd Privilege separation:/var/empty:/noshell
unknown:*:99:99:Unknown User:/nohome:/noshell
```

Task #2: Counting the number of line, words and characters in a file

In Unix/Linux, there is a utility called **wc** which is used to count the number of lines, words, and characters in a text file (see the **wc** man page on Linux). Implement a Python script **wc.py** that operates the same way as **wc** and supports the command line options **-l** for lines, **-w** for words, and **-c** for characters, or a combinations such as **-lw**. If no command line options are given, the default is to print the number of lines, words and characters followed by the file name.:

```
linux% python wc.py /etc/passwd
 16   72  722 /etc/passwd
linux% python wc.py -l /etc/passwd
 16 /etc/passwd
linux% python wc.py -w /etc/passwd
 72 /etc/passwd
linux% python wc.py -c /etc/passwd
722 /etc/passwd
linux% python wc.py -lw /etc/passwd
 16   72 /etc/passwd
```

Note: In Python, **sys.argv** is a list where **argv[0]** is the name of the Python script that was executed, **sys.argv[1]** is the first argument to the script if any, **sys.argv[2]** the second if any, etc. If no arguments are provided, then the list consists of just the name of the script. To use **sys.argv**, import the **sys** module at the start of your program. The **getopt** module is also extremely useful in processing command line arguments. See **help(“sys”)** and **help(“getopt”)**.

EXTRA CREDIT: Sorting a file based on a key value and delimiter

Assuming a file consisting of multiple string fields per line followed by some type of delimiter (e.g., tab, space, ‘:’, ‘;’, etc.) where any one of the fields can be treated as a key on which to sort the lines of the file. Implement a Python script that sorts the non-comment lines of the file based on a designated key field and delimiter and writes the resulting sorted file to the standard output. Comment lines are skipped in the sorting. A comment line begins with a special character denoted by the command line flag `-c`. For example, `-c“#”` sets the comment character to “#” so all lines beginning with that character are ignored. Implement a Python script `sort.py` that sorts a file into ascending order based on a key field. For example, `sort.py` applied to the `/etc/passwd` file sorted into ascending order with field 1 as the key and the ‘:’ character as the delimiter is:

```
linux% python sort.py -k1 -t":" -c"#" /etc/passwd
daemon*:1:1:System Services:/var/root:/noshell
mysql*:74:74:MySQL Server:/nohome:/noshell
nobody*:-2:-2:Unprivileged User:/nohome:/noshell
root*:0:0:System Administrator:/var/root:/bin/tcsh
smmsp*:25:25:Sendmail User:/private/etc/mail:/noshell
sshd*:75:75:sshd Privilege separation:/var/empty:/noshell
unknown*:99:99:Unknown User:/nohome:/noshell
www*:70:70:World Wide Web Server:/Library/WebServer:/noshell
```

Similarly, `sort.py` should be able to sort on a different field as the key, with the `-r` option used to indicate that `sort.py` should print the results in reverse (descending) order:

```
linux% python sort.py -r -k3 -t":" -c"#" /etc/passwd
unknown*:99:99:Unknown User:/nohome:/noshell
sshd*:75:75:sshd Privilege separation:/var/empty:/noshell
mysql*:74:74:MySQL Server:/nohome:/noshell
www*:70:70:World Wide Web Server:/Library/WebServer:/noshell
smmsp*:25:25:Sendmail User:/private/etc/mail:/noshell
daemon*:1:1:System Services:/var/root:/noshell
root*:0:0:System Administrator:/var/root:/bin/tcsh
nobody*:-2:-2:Unprivileged User:/nohome:/noshell
```

Submission Requirements:

You are to hand in at the start of class on the due date a source code listing of your program with your name and comments included. You are also to submit an electronic copy of your source program to the course TA using the **turnin** program on CS Department machines BEFORE class begins on the due date. Late assignments will not be accepted.

Honor Policy:

This assignment is a personal learning opportunity that will be evaluated based on your ability to think independently, work through a problem in a logical manner and implement a software program on your own. You may however discuss verbally or via email the general nature of the conceptual problem to be solved with your classmates, the course TA or the course instructor, but you are to complete the actual programming for this assignment without resorting to help from any other person or other resources that are not authorized as part of this course. If in doubt, ask the course instructor. *Any indication of improper cooperation or the use of non-approved course materials will result in a grade of zero being assigned for this project and referral to the College for possible disciplinary action with respect to your continued participation in this course. **By submitting a solution to this assignment under your name, you are asserting that this assignment was done by you under this honor policy.***