# CS241 : Assignment III

Prof. Hemangee K. Kapoor
Dr. Aryabartta Sahu
Department of CSE,
IIT Guwahati

September 20, 2018

---

This part of the assignment will make you understand the basics of the regular expression with the use of sed and grep/egrep commands.

**Bash Commands with egrep**

**Problem 1** Write a single bash shell statement that will perform the operation requested below. Each solution must be a one-line shell statement, but you may use input/output redirection operators such as >, <, | etc.
Write a command by using pipes to combine the "ls" and "grep"

1. to list all the directories and files in the current directory that have "rwx" permissions for user, group, and "rw" permission for other.
2. to count all the files/folders that are last modified in the month of August.
3. to display all the files/folders of size in terms of MBs or more than that.
4. to list and count all the file/folders that are not modified in 'Sep'.

**Problem 2** For rest of the problems, the given **words.txt** file is used as an input file to the grep or egrep command. Use single grep/egrep and the appropriate regular expression.

1. Write a command to list the output the name of words that contain the letter 'z' followed by any single character, followed by a letter 'a'.
2. Write a command to list all the words that contain "apple", "papaya" and "kiwi" as a string or substring.
3. Write a command to list all the words start with "abid".
4. Write a command to list all the words that end with "ever".
5. Write a command to list all the words that contain 4 vowels (a, e, i, o and u) in a consecutive manner (Note: the vowel sequence may be repeated).
6. Write a command to list all the words that are 8 characters long and print it in reverse alphabetic order.

7. Write a command to list all the words that start with letter 'm' or 'n' and contain a string "oon".

**Bash Commands with Sed**

In each of the following problems, use the SED (Stream EDitor) command to answer the following queries. SED is used to perform a basic transformation on a stream of data which is provided either from a file or from a pipe. For some of the following problems, you may need to combine grep/egrep and sed using |. Each problem should use at most one call to sed, but you may use input/output redirection operators such as >, <, and | to combine it with other commands as per the requirement.

**Problem 3**

1. By using sed, write a command to output the content of file email_old.txt to the new file email_new.txt with all "ernet.in" replaced with "ac.in".

2. Write a sed command to output the content of the file QuizProgram.java with all occurrences of word public replaced with private (Note that the word contain the substring "public" is not replaced and changed with the command you used. For eg: the word publicly in the file remains to be intact.)

3. Suppose a file msg.txt is given which are having too many exclamation symbols (!). Use a sed command output a content of a file that replaces one or more occurrences of ! with the single period. (Note that the multiple occurrences of !!!!!! is also to be replaced with the single period(.).)

4. For a given Appointment_dates.txt file American time format is given, (such as May 12, 2010) where the European uses different time format (such as 12 May 2010). Write a sed command to change the content of time format in the file from the American to European.

**Awk Command Related Questions**

Instructions: Download the “sample.txt” file from the course website and save that to your home directory. Each row of the “sample.txt” consists of 11 fields with delimited by space, and each field represents the following; Field-1 : timestamp, field-2 : job-id, field-3 : task index within job, field-4 : machine-id, field-5 : event type, field-6 : scheduling class, field-7 : priority, field-8 : CPU usage, field-9 : RAM usage, field-10 : disk usage and field-11 : machine constraint. All the questions require this file, and you are required to use “awk” commands to solve these questions.

1. Write an awk command to find the total number of lines in the file “sample.txt” without using NR of awk or wc command?

2. Write a command to find the distinct number of tasks of the job whose job-id is “515042969” (Each job consists of more than one tasks).

3. Write the command to find the total memory usage of the job whose job-id is “515042969” (Hint: One job has different tasks, and you need to find the sum of the memory usage of all it’s tasks).

4. What is the command to print the records of the “sample.txt” file, whose machine constraint is “1” ?

5. How many records are present in the “sample.txt” file, whose CPU usage is in-between 0.001 to 0.009?

6. Write a command to find the length of the longest line present in the file “sample.txt”.

7. Write command to display the (job-id, CPU usage, RAM usage, and disk usage ) of records from line 100 to 200 of the “sample.txt” file.

8. Write a command to split the “sample.txt” file, so that all the records whose priority value is less than and equal to “0” are in the file "sample0.txt", and the rest are in the file "sample1.txt".

9. Write a command to split the file “sample.txt” into multiple files at every 100th line . i.e., first 100 lines into File1, next 100 lines into File2 and so on.

10. Suppose you want to insert a new column (say the serial number) before the 1st column of the file “sample.txt”. Write a command to perform that.