

UNIVERSITY OF COLOMBO SCHOOL OF COMPUTING

IS1104 - Application Laboratory

Activity 10



Sed command practice

SED stands for stream editor. This stream-oriented editor was created exclusively for executing scripts. Thus, all the input you feed into it, passes through and goes to STDOUT and it does not change the input file.

General syntax for sed: **/pattern/action**

pattern : regular expression

action : one of the commands given as follows.

- **p** : prints the line
- **d** : deletes the line
- **s/pattern1/pattern2/** : substitutes the first occurrence of pattern1 with pattern2

Delete lines with sed(d)

- `$ cat /etc/passwd | sed 'd'` : Delete all lines in file
- `$ cat /etc/passwd | sed '1d' |more` : Delete first line
- `$ cat /etc/passwd | sed '1, 5d' |more` : Delete first 5 lines
- `$ cat /etc/passwd | sed '4, +5d' |more` : Delete 5 lines after 4th line

Substitution command(s)

- Substitute the **first occurrence** on a line of the string **root** with the string **amrood**.
`$ cat passwd | sed 's/root/amrood/'`
- Global substitution, replace all occurrences.
`$ cat passwd | sed 's/root/amrood/g'`
- Substitute the string sh with the string quiet only on line 10
`$ cat passwd | sed '10s/sh/quiet/g'`
- Using regular expressions
`$ cat passwd | sed '/sh$/d'`
 - o `/a.c/` : Matches lines that contain strings such as a+c, a-c, abc, match, and a3c
 - o `/a*c/` : Matches the same strings along with strings such as ace, yacc, and arctic
 - o `/[tT]he/` : Matches the string The and the
 - o `/^$/` : Matches blank lines
 - o `/^.*$/` : Matches an entire line whatever it is
 - o `[a-z]` : Matches a single lowercase letter
 - o `[a-zA-Z]` : Matches a single letter
- Ampersand referencing (&) : & represents the contents of the pattern that was matched.

- o `$ sed 's/^[0-9][0-9]/(&)/g' number`

Problem

1. Use the file `sed.html` and create a new file with following changes.
 - I. Lines with `<article>` and `</article>` should be deleted.
 - II. Replace `<title>` with **Title:** and replace `</title>` with nothing.
 - III. Replace all `<para>` and `</para>` tags with the null string. If the resulting line is empty, delete the line. (You may need to use curly braces to make this happen.)
 - IV. Replace all `<emphasis>` and `</emphasis>` tags with asterisks. Thus:
"This is a `<emphasis>great</emphasis>` bargain" will become "This is a *great* bargain"
 - V. Replace the word **web** with **Web** everywhere.
 - VI. Replace lines starting with `<listing>` by **---begin listing**
 - VII. Replace lines starting with `</listing>` by **---end listing**
 - VIII. Between the `<listing>` and `</listing>`, do these things (you must use curly braces to do this!):
 - a. Replace all occurrences of `<` with `<`
 - b. Replace all occurrences of `>` with `>`
 - c. Replace all occurrences of `&` with `&`.Note: you must do these operations in the order shown above; otherwise, you will get the wrong results!

Note: The `&` is a special character. So you have to use backslash escaping in order to use them in expressions.