







LPIC-1. Examen 101. Objectius 103.7

LPI 103.7 Fer cerques en fitxers de text utilitzant expressions regulars

http://acacha.org/mediawiki/index.php/LPI_103.7

Objectius

103.7. Fer cerques en fitxers de text utilitzant expressions regulars	
	<ul style="list-style-type: none"> ▪ Objectiu: Els candidats han de ser capaços de manipular fitxers i dades de text utilitzant expressions regulars. Aquest objectiu inclou la creació d'expressions regulars simples que continguin diversos elements de notació. També inclou l'ús d'eines d'expressions regulars per a fer cerques al sistema de fitxers o en els continguts d'un fitxer. ▪ Pes: 2
	<p>Àrees Clau de Coneixement:</p> <ul style="list-style-type: none"> ▪ Crear expressions regulars simples que continguin diversos elements de notació. ▪ Utilitzar eines d'expressions regulars per a fer cerques a tot el sistema de fitxers o al contingut d'un fitxer.
	<p>La següent és una llista parcial de fitxers, termes i utilitats utilitzades:</p> <ul style="list-style-type: none"> ▪ grep ▪ egrep ▪ fgrep ▪ sed ▪ regex(7)
	<p>Apunts: LPI 103.7. Fer cerques en fitxers de text utilitzant expressions regulars</p>



Expressions regulars

- ♦ **Hi ha dos tipus d'expressions regulars:**
 - ♦ **Bàsiques:** Els caràcters `?`, `+` i `|` i els símbols `(` i `)` perden el seu significat especial i compten com a caràcters normals i corrents. Alguns programes que suporten expressions regulars bàsiques, permeten que aquests símbols tinguin el seu significat especial si s'escapen amb el caràcter barra invertida (`\`), per exemple `\?`.
 - ♦ **Extenses:** Els exemples i explicacions d'aquesta wiki són per a expressions extenses. Les diferències són subtils entre els dos tipus d'expressions però els fonaments són les mateixos.



Expressions regulars

- ♦ El tipus d'expressió regular ha utilitzar depèn del programa, trobarem programes que només accepten un dels dos tipus i altres que accepten tots dos.
- ♦ **NOTA:** També es possible que trobeu programes que utilitzen una variant pròpia (similar però amb certs matisos) de les expressions regulars estàndard.
- ♦ La expressió regular més simple conté el text amb el que volem treballar. Per exemple
 - Ethernet
- ♦ Aquesta expressió regular coincidirà amb qualsevol text que contingui els caràcters Ethernet.



Expressions regulars

- ♦ **La potència de les expressions regulars la trobem en l'ús de caràcters especials**
 - ♦ **Expressions amb claudàtors ([])**
 - Els caràcters que estiguin entre claudàtors (també anomenats parèntesis rectangulars), constitueixen expressions regulars que coincideixen amb qualsevol del caràcter de l'interior dels claudàtors.
 - ♦ Per exemple, la expressió regular:

```
p[aeiou]t  
Coincidirà amb els següents texts:  
pat  
pet  
pit  
pot  
put
```



Expressions regulars

- ◆ Expressions amb rang [valorinicial-valorfinal]

[2-6]x

Coincideix amb els texts:

2x

3x

4x

5x

6x

- ◆ Qualsevol caràcter simple (.)
- ◆ Inici i final de línia:
 - El caràcter ^ representa l'inici de línia.
 - El caràcter dol·lar (\$) representa el final de la línia.
 - El caràcter \ s'anomena carat en anglès



Expressions regulars

♦ Operadors de repetició

- ♦ **Asterisc (*)**: Indica zero o més ocurrències. Asterisk en anglès.
- ♦ **Sumna (+)**: coincideix amb una o més ocurrències.
- ♦ **Interrogant (?)**: o cap ocurrència o una ocurrència.
- ♦ Una de les expressions regulars més utilitzades és:
 - **.***: Coincideix amb qualsevol combinació de caràcters
- ♦ **Múltiples conjunts de caràcters (|)**
- ♦ **Parèntesi: (cotxe|moto|camió)**



Expressions regulars

- ♦ Escapar caràcters especials (\)
 - 192\.168\.1\.1
- ♦ Es pot consultar més informació sobre les expressions regulars al manual:

```
$ man regex
REGEX(7)
  Linux Programmer's
Manual                                REGEX(7 )

NAME
    regex - POSIX.2 regular expressions

DESCRIPTION
    Regular expressions ("RE"s), as defined in
    POSIX.2, come in two forms: modern REs (roughly those
    of egrep; POSIX.2 calls these "extended" REs) and
    obsolete REs (roughly ...
```




grep

- ♦ **És una de les comandes més útils i versàtils de Linux.**
 - ♦ Mostra només les línies que compleixen amb un patró. La sintaxi és:

```
$ grep [opcions] regexp [fitxers]
```
 - ♦ On regexp és una expressió regular. Grep treballa per defecte amb expressions regulars bàsiques però podem indicar que treballi amb expressions regulars extenses amb l'opció -e (**ordre egrep**)



grep

- ♦ El resultat de l'execució de la comanda és el nom del fitxer que conté el text buscat i la línia que conté aquest text. Per exemple:

```
$ grep -r "Linux" /  
/sbin/wpa_action:# On Debian GNU/Linux systems, the text of the  
GPL license can be  
...
```

- ♦ Opció recursiva. També existeix la comanda **rgrep** que és equivalent a `grep -r`.
- ♦ Si volem que el text de cerca proporcionat no sigui interpretat com a una expressió regular, hem d'utilitzar l'opció `-F` (`--fixed-strings`). La comanda **fgrep** és equivalent a `-F`. També es coneix com a **fast grep**.



grep

♦ Mostrar els números de línia

- ♦ Buscar un text en múltiples fitxers i mostra la línia i el fitxer on s'ha trobat:

```
$ sudo grep -n -r "udpcast_rec_cmd" /opt/drbl
```

- ♦ Treure els comentaris d'un fitxer:

```
$ cat /etc/apache2/apache2.conf | grep -v '#'
Si a més volem treure les línies en blanc:
$ cat squid.conf | grep -v '^#\|^$\|^;'
```

- ♦ Buscar paràmetres als manuals

```
$ man ls | grep "\-all"
```



Reconeixement 3.0 Unported

Sou lliure de:



copiar, distribuir i comunicar públicament l'obra



fer-ne obres derivades

Amb les condicions següents:



Reconeixement. Heu de reconèixer els crèdits de l'obra de la manera especificada per l'autor o el llicenciador (però no d'una manera que suggereixi que us donen suport o rebeu suport per l'ús que feu l'obra).

- Quan reutilitzeu o distribuïu l'obra, heu de deixar ben clar els termes de la llicència de l'obra.
- Alguna d'aquestes condicions pot no aplicar-se si obteniu el permís del titular dels drets d'autor.
- No hi ha res en aquesta llicència que menyscabi o restringeixi els drets morals de l'autor.

Advertiment

Els drets derivats d'usos legítims o altres limitacions reconegudes per llei no queden afectats per l'anterior
Això és un resum fàcilment llegible del text legal (la llicència completa).

<http://creativecommons.org/licenses/by/3.0/deed.ca>

LPIC-1. Examen 101. Objectiu 103.7

ICE-UPC



Autor: Sergi Tur Badenas