

Introduction aux Expressions rationnelles

Rémi Laurent
remi.laurent@fsugar.be



<http://www.fsugar.be/>

7 mars 2008

Une notation permettant de décrire des chaînes de caractères :

- de repérer des schémas donnés de chaîne de caractères
- de les manipuler
- (d'impressionner|de traumatiser) (sa copine|ses copains|son patron)

- \wedge le début d'une chaîne
- $\$$ la fin d'une chaîne
- \cdot un caractère quelconque
- a^* zéro ou plusieurs 'a'
- b^+ un ou plusieurs 'b'
- $\cdot?$ zéro ou une fois un caractère quelconque

- verbes en 'er' - er\$
donner, casser, ...
- les dérivés de 'absol' - ^absol
absolument, absolution, ...
- les mots terminés par XXXIIIXnt - .*II.nt\$
allant, aillent, verrouillent, ...

```
dpkg -purge
```

```
$ dpkg -l | grep ^rc  
# Trouver les paquets désinstallés mais dont  
# il reste des fichiers de configuration
```

$fo\{2\} - foo$ $fo\{2\} - foo$
 $fo\{1,\}$ $fo, foo, fooo, \dots$
 $fo\{0,2\}$ f, fo, foo

Boitam.eu

```
$ grep '^.*eu[xt]$\ ' /usr/share/dict/french
# Se faire sa liste de nom de domaines en .eu
# ayant du 'sens' ...
# boitam.eu n'en fera malheureusement pas partie
```

Grep (KDE|Gnome) Games

```
$ apt-cache search games | grep '^[gk].*'
```

```
# Trouver des jeux sous Gnome ou KDE
```

```
# PS: la méthode n'est pas infaillible évidemment
```

```
Grep -v (KDE|Gnome) Games | grep -v library
```

```
$ apt-cache search games | \  
  grep '^([gk].*)' | grep -v '^lib'  
  
# Les jeux ni sous Gnome, ni sous KDE  
# et sans les bibliothèques ^lib  
#  
# grep(1) en ce qui concerne '-v'
```

Alphanumérique

```
$ grep '[a-zA-Z0-9]' un_fichier_texte
```

```
# les lignes qui ne contiennent  
# que des lettres ou chiffres
```

Egrep

```
$ ls / | grep -E '^m(edia|root)'
```

```
# on cherche une chaîne qui commence
```

```
# par 'm' suivi de 'edia' OU 'root'
```

```
#
```

```
# donc on cherche ce qui commence
```

```
# par 'media' OU 'mroot'
```

Egrep & |

```
$ ls / | grep -E '^media|root'
```

```
# on cherche une chaîne qui commence
```

```
# par 'media' OU 'root'
```

Sed - Email obfuscation

```
$ echo "remi.laurent@gmail.com" | \  
  sed 's/@/ at /' | sed 's/\./ p0iNt /'  
# --> remi p0iNt laurent at gmail.com
```

```
$ echo "remi.laurent@gmail.com" | \  
  sed -e 's/@/ at /' -e 's/\./ p0iNt /g'  
# --> remi p0iNt laurent at gmail p0iNt com
```

```
$ echo "remi p0iNt laurent at gmail p0iNt com" | \  
  sed -e 's/ at /@/ig' -e 's/ point /./gi'  
# --> remi.laurent@gmail.com
```

Sed - Email mess

```
$ echo "remi.laurent@gmail.com" | \  
  sed -r -e 's/gmail/googlemail/' \  
        -e 's/(.*)\.(.*)@(.)/\2.\1@3/'  
  
# foo.bar@gmail.com -> bar.foo@googlemail.com  
# foo.bar@foobar.com -> bar.foo@foobar.com  
# gmail.baz@fgmail.com -> baz.googlemail@gmail.com
```

rename(1)

Rename

```
# Knoklezout_au_taqet-Jamendo-David_TMX.mp3  
  
rename 's/^(.*)-Jamendo-(.*)\.mp3/$2_-_$1.mp3/' *.mp3  
  
# --> David_TMXr_-_Knoklezout_au_taqet.mp3
```

- Perl, Ruby, Python, et aussi PHP, Java, C, C++,...
- sed, awk, grep, vim, rename,...
- apache & lighttpd (mod_rewrite, vhosts, ...), adblock (firefox),
...

- Une introduction à Perl et une introduction au regexp
<http://sylvain.lhullier.org/publications/perl.html> (FR)
- Le livre de référence (édition 1) de Ruby avec une petite partie sur les regexp
<http://rubycentral.com/book/language.html> (EN)
- Une approche plus théorique et (donc) moins marrante avec Python
<http://www.amk.ca/python/howto/regex/> (EN)
- Une introduction rapide réalisée dans notre FSUG (merci à Krunch)
<http://www.fsugar.be/docu/regex/regex.html> (FR)

Questions ?

Workshop / Atelier