1. Les expressions régulières

1.1. Présentation

Une expression régulière sert à identifier une chaîne de caractères répondant à un ou plusieurs critères comme par exemple une chaîne de caractères ne contenant des lettres.

Les expressions régulières permettent :

- de rechercher des “motifs” dans des fichiers texte comme les fichiers journaux ou les fichiers de configuration ;
- de réaliser, en une seule commande, un ensemble de tâches.
- appliquer des transformations dans du texte ;
- valider la conformité d’une donnée.

1.2. Définir des expressions régulières

Les expressions régulières utilisent divers opérateurs pour combiner des expressions de plus petite taille.

Il existe deux standards principaux en matière d'expression régulière, POSIX et Perl.

Un site pour apprendre à utiliser les expressions régulières :

- https://zestedesavoir.com/tutoriels/3651/les-expressions-regulieres-1/
- en testant les expressions régulières avec Rubilar

Utiliser les expressions régulières avec grep sous Linux :

- https://www.malekal.com/exemples-de-commande-grep-sous-linux/

1.2.1. Liste de quelques méta-caractères utilisés et leur fonction :

Méta-caractère	**Fonction**
\d	Tous les chiffres
\D	Tout sauf les chiffres
.	Tout caractère
\w	Toutes les lettres
\W	Tout sauf des lettres
[a-z]	Les lettres de a à z (attention les expressions régulières sont sensibles à la casse)
[0-9]	Les chiffres de 0 à 9
[abc]	a, b ou c
[^abc]	Ni a ni b ni c
{n}	Répéter n fois
{n,m}	Répéter de n à m fois
*	Répéter de 0 à n fois
+	Répéter de 1 à n fois
\s	Le caractère espace
\S	Tous les caractères sauf l’espace
Début de ligne ou négation	
$	Fin
\b	Début ou fin de mot
\	Caractère échappatoire
(…)	Sélection d’un groupe
|	Permettant de faire ou entre 2 sous-motifs

1.2.2. Quelques exemples d’expressions régulières et leur signification

Expression	**Signification**
test	Recherche de mots contenant test (tester, attester…)
\d{3}-\d{6}	Chaine de caractères de la forme 3chiffres-6chiffres comme 123-654321
a	Toutes les lignes commençant par la lettre a
\bf\w*\b	Tous les mots commençant par la lettre f
\d+	Tous les mots ne contenant que des chiffres (au moins un chiffre)

Expression	Signification
\w+	Tous les mots ne contenant que des lettres (au moins une lettre)
(fichier.+)\.pdf$	Va afficher toutes les lignes commençant par fichier avec l'extension .pdf mais ne va afficher que la partie avant l'extension (fichier1.pdf affichera fichier1)
ABC|XYZ	Tous les mots contenant ABC ou XYZ
\btest\b	Tous les mots « test » exactement (on ne cherche que le mot test)
(\d+)x(\d+)	Pour le texte 1920×800 va afficher deux groupes 1920 et 800

From:
https://siocours.lycees.nouvelle-aquitaine.pro/ - **Les cours du BTS SIO**

Permanent link:
https://siocours.lycees.nouvelle-aquitaine.pro/doku.php/reseau/regexp/accueil?rev=1637790491

Last update: **2021/11/24 21:48**