

Décrire et manipuler un document numérique

Cours 4 : Correction des exercices du slide

Loïc Grobol

2022-01-31

Les quantificateurs

Exercice : que reconnaissent les expressions suivantes ?

1. `pommes?`

Les chaînes de caractères `pomme` et `pommes`.

2. `a?typique`

Les chaînes de caractères `typique` et `atypique`.

3. `lo*1`

Les chaînes de caractères `l1`, `lo1`, `lool...` pour n'importe quel nombre de `o`.

4. `loo*1`

5. `lo+1`

Même chose pour ces deux expressions : es chaînes de caractères `lo1`, `lool...` pour n'importe quel nombre de `o`.

Accolades

Exercice : écrire une expression régulière qui reconnaît exclusivement `lool`, `loool` et `loooo1`.

`lo{2,4}1`

Les opérateurs

Ensembles de caractères

Exercice : Que reconnaît :

1. `[ab] [ab]`

`aa, ab, ba et bb.`

2. `[MmRr]aison`

`Maison, maison, Raison et raison.`

Écrire une expression régulière qui reconnaît exclusivement :

1. `moi, toi et soi.`

`[mts]oi`

2. `ci, si, ce et se.`

`[cs] [ie]`

3. Les nombres à deux chiffres.

`[1-9] [0-9]`

Alternatives

Exercice : que reconnaissent les expressions suivantes ?

1. `vert(es) ?`

`vert et vertes`

2. (anti|pro)nucléaire

antinucléaire et pronucléaire

3. Wiki[a-z]+

Wikia, Wikib, Wikic, ..., Wikiz.

4. Wiki[a-z]*

Wiki, Wikia, Wikib, Wikic, ..., Wikiz.

Caractères spéciaux

Exercices : à quoi correspondent ces expressions ?

1. Nom\tPrénom

Nom Prénom, le blanc étant une tabulation.

2. \bcap\b

Le mot **cap**, non-inclus dans un autre mot.

3. \bcap

Un mot commençant par **cap**

4. premier paragraphe\ndeuxième paragraphe

premier paragraphe
deuxième paragraphe

Déspecialisation

Dans le fichier Père Duchesne :

1. Recherchez `^.évolution`

21 matchs

2. Recherchez `r?évolution`

215 matchs

3. Combien y a-t-il de mots...

3.1 de 3 lettres finissant par `r` ?

`\b\w\wr\b` : 1137 matchs

- Attention à bien activer le mode Unicode
- Sur Regex101, il faut augmenter le temps d'exécutions (il faut environ 3s et la limite par défaut est 2s).

3.2 de 4 lettres au plus commençant par `r` ?

`\br\w{0,3}\b` : 449 matchs

3.3 commençant par `a` et finissant par `er` ?

`\ba\w*er\b` : 348 matchs

3.4 contenant la séquence `er` en fin de ligne ?

`er$` : 460 matchs

Exercices

Faire les quatre premiers niveaux sur : <https://alf.nu/RegexGolf>

1. Warmup

`foo`

2. Anchors

`ick$`

3. It never ends

`fu\b`

4. Ranges

`^[a-f]+$`

5. Backref

`(.{3,}).*\1`