

TD du 3 février

Yves Stadler

21 avril 2009

Exercice numéro 1 : Le téléphone pleure

- Écrire une expression régulière décrivant un numéro de téléphone français (10 chiffres, commençant par 0 puis 1 2 3 4 ou 5).
- Même, mais décrivant aussi le format international (+33 6 12 34 56 78)
- Même, mais décrivant aussi le format international complet (00 33 6 12 34 56 78)

Correction

- `^0[1-6][0-9]{8,8}$`
- `^(\\+33|0)[1-6][0-9]{8,8}$`
- `^((\\+|00)33|0)[1-6][0-9]{8,8}$`

Explications supplémentaires :

[1-6] restreint les numéros aux 01-02-03-04-05-06.

Le + est échappé pour avoir sa valeur textuelle.

8,8 permet de fixer le nombre de chiffres du "corps" à 8 exactement.

Exercice numéro 2 : L'email

- Écrire une expression régulière décrivant une adresse e-mail. (yves.stadler@codasystem.com ...)
- Même chose mais en limitant aux adresses MSN (hotmail.fr, live.fr, hotmail.com, live.com, msn.fr, msn.com)
- Même, mais décrivant aussi le format international complet (00 33 6 12 34 56 78)

Correction

- `^[a-z.]+@[a-z]+\\.[a-z]+$`
- `^[a-z.]+@(msn|live|hotmail)+\\.(com|fr)+$`

Exercice numéro 3 : Over the net

- Écrire une expression régulière décrivant une adresse d'un site internet (ex : `http://www.free.fr/`, `www.msn.fr`)
- Ajouter le protocole https

- Ajouter les ftp
- Les adresses de pages php (`http://yves.stadler.codasystem.com/rep1/login.php?a=y@y.c&p=tutu`)

Correction

- `^(http://)?(www.)?[a-z.0-9_/>+]`
- `^(https?://)?(www.)?[a-z.0-9_/>+]`
- `^((https?|ftp)://)?(www.)?[a-z.0-9_/>+]`
- `^(https?://)?(www.)?[a-z.0-9_/>+]\.php.*$`

Explications supplémentaires :

On ajoute les caractères de début et fin de phrases pour matcher strictement. On peut matcher les adresses sous toutes leurs formes (on peut ajouter encore d'autres caractères dans le corps de l'adresse comme le pourcent, mais on va faire simple). `http://` n'est pas obligatoire, ainsi que `www` (que l'on précise par simple souci de détails).

La dernière expression permet de filtrer selon les extensions `.php` (ne pas oublier l'échappement)

Exercice numéro 4 : You're in the HTML now

- Matcher une balise de lien dans un document HTML (`<a ...> lien </a>`)
- Matcher une balise qui n'est pas une ancre (le paramètre `href` de la balise à la valeur d'une adresse valide) `<a ... href="@site" ... > lien </a>`

Correction

- `<a.*>.*</a>`
- `<a.*href="(http://)?(www.)?[a-z.0-9_/>+]"*.*</a>`

Un peu plus loin

- Format date (symbolique, textuelle, ...)
- Autres idées...