

# Mémo expressions régulières

Inspiré de [cheatography.com/1/cs/5/](http://cheatography.com/1/cs/5/)

Métacaractères	
<code>^</code>	Début de chaîne ou début de ligne
<code>\$</code>	Fin de chaîne ou fin de ligne
<code>.</code>	N'importe quel caractère sauf retour à la ligne
<code> </code>	Disjonction (a b pour a ou b)
<code>[ ]</code>	Marques de début et fin de classe
<code>( )</code>	Marques de début et fin de groupe
Ces caractères peuvent être déspecialisés avec le car. <code>\</code>	

Groupes et ensembles	
<code>(...)</code>	Groupe
<code>(abc)+</code>	abc répété 1 à n fois
<code>\n</code> ou <code>\$n</code>	nième groupe
<code>(a b c)+\1</code>	ex : a b c a b c
<code>[...]</code>	Ensemble
<code>[abc]</code>	a ou b ou c
<code>[^abc]</code>	Non a ou b ou c
<code>[a-z]</code>	Lettre min. de a à z
<code>[0-9]</code>	Entier (0 à 9) ( <code>\d</code> )
<code>[a-z0-9_]</code>	Lettre min. de a à z ou entier ou <code>_</code>
<code>[a-z'-]</code>	Lettre min. de a à z ou apostrophe ou tiret

Opérateurs de répétition	
<code>*</code>	0 à n fois
<code>+</code>	1 à n fois
<code>?</code>	0 ou 1 fois
<code>{3}</code>	exactement 3 fois
<code>{3,}</code>	3 fois ou plus
<code>{3,5}</code>	3 à 5 fois
Les opérateurs portent sur le car. ou le groupe qui précède	
Les opérateurs ont par défaut un comportement gourmand ("greedy"), ajoutez un <code>?</code> pour les rendre "ungreedy"	

Ensembles POSIX	
<code>[:upper:]</code>	Lettres maj.
<code>[:lower:]</code>	Lettres min.
<code>[:alpha:]</code>	Lettres
<code>[:alnum:]</code>	Car. alpha numériques
<code>[:digit:]</code>	Entiers
<code>[:xdigit:]</code>	Entiers en hexadécimal
<code>[:punct:]</code>	Ponctuation
<code>[:space:]</code>	Car. d'espace
<code>[:cntrl:]</code>	Car. de contrôle
<code>[:print:]</code>	Car. imprimables et espaces
<code>[:word:]</code>	Entiers, lettres et underscore

Caractères spéciaux	
<code>\n</code>	Nouvelle ligne
<code>\r</code>	Retour à la ligne
<code>\b</code>	Borne de mot
<code>\t</code>	Tabulation
<code>\x{0281}</code>	Car. Unicode U+0281 (ⵁ)

Ensembles Unicode	
<code>\p{Letter}</code>	Lettres (tous alphabets) ( <code>\p{L}</code> )
<code>\P{Letter}</code>	Non lettres
<code>\p{Lu}</code>	Lettres maj.
<code>\p{Ll}</code>	Lettres min.
<code>\p{P}</code>	Ponctuation
<code>\p{Arabic}</code>	Car. de l'alphabet arabe
<code>\p{Cyrillic}</code>	Car. de l'alphabet cyrillique
<code>\p{InTamil}</code>	Car. du bloc Unicode Tamoul (U+0250–U+02AF)

Formation regex
LabEx EFL - opération LR-1 S. El Ayari, C. Plancq, A. Roulois <a href="http://ressources.labex-efl.org">ressources.labex-efl.org</a>

Sites de référence
<a href="http://regular-expressions.info">regular-expressions.info</a> <a href="http://rubular.com">rubular.com</a> <a href="http://regexcrossword.com">regexcrossword.com</a>

Regex sous Ms Word
<a href="#">documentation</a>