

TEMA: ASP .NET – WEB FORMS
SUBTEMA: EXPRESIONES REGULARES

Expresiones Regulares

Una **expresión regular**, también conocida como **regex** o **regexp**, es una secuencia de [caracteres](#) que conforma un patrón de búsqueda. Se utilizan principalmente para la [búsqueda de patrones](#) de cadenas de caracteres u operaciones de sustituciones.

Las expresiones regulares son patrones utilizados para encontrar una determinada combinación de caracteres dentro de una [cadena de texto](#). Las expresiones regulares proporcionan una manera muy flexible de buscar o reconocer cadenas de texto.

- [] Los corchetes identifican a una clase de caracteres que siempre representa a un único carácter en un patrón de búsqueda.
 - [A]
Indica que debe localizar la letra A
 - [Aa]
Indica que debe localizar la letra A mayúscula o la a minúscula
 - [AEIOU]
Indica que debe localizar una vocal
 - [A-Z]
Indica que debe localizar cualquier letra de la A-Z
- ^ Cuando se utiliza en conjunto con los corchetes de la siguiente forma "[^ AEIOU]" permite encontrar cualquier carácter que NO se encuentre dentro del grupo indicado. En este caso Indica que debe localizar cualquier carácter excepto las vocales
- [B-DF-HJ-ÑP-TV-Z]
Indica que debe localizar cualquier consonante
- [0-9]
Indica que debe localizar cualquier dígito
- () Los paréntesis identifican un grupo de caracteres formado por uno o varios caracteres y que pueden operarse unos dentro de los otros
 - (MX|mx|Mx|mX)
Indica que debe localizar cualquier grupo (de caracteres) MX, mx, Mx, o mX
 - (0[1-9])
Indica que debe localizar cualquier grupo de 01,02,03,04,05,06,07,08,09

[0-9]{2}

Indica que debe localizar cualquier número de dos dígitos

[0-9]{1,3}

Indica que debe localizar cualquier número de uno a tres dígitos

^[A-Z]

Alt + 94

Indica que la cadena debe empezar con una letra

[A-Z]\$

Indica que la cadena debe terminar con una letra

^[A-Z].{1,}[A-Z]\$

Indica que la cadena debe empezar y terminar con una letra, y puede haber cualquier cantidad de caracteres

- {n} El carácter, la clase o el grupo anteriores deben aparecer exactamente n veces.
 - [A]{2} indica que debe localizar dos letras A contiguas
- {n,m} El carácter, la clase o el grupo anteriores aparecen como mínimo n veces y como máximo m veces.
- {n,} El carácter, la clase o el grupo anteriores aparecen como mínimo n veces o con frecuencia.
- {n} El carácter, la clase o el grupo anteriores aparecen exactamente n veces
- **La barra "|"** Sirve para indicar una de varias opciones.
Por ejemplo, la expresión regular "a|e" encontrará cualquier "a" o "e" dentro del texto. La expresión regular "este|oeste|norte|sur" permitirá encontrar cualquiera de los nombres de los puntos cardinales. La barra se utiliza comúnmente en conjunto con otros caracteres especiales.
- **^** Limita la búsqueda al inicio de una línea (otra función: elemento de negación en clases de caracteres).
- **\$** Limita la búsqueda al final de una línea.
- **.** Equivale a cualquier carácter.

- * El número del carácter, de la clase o del grupo situado antes del asterisco puede ser aleatorio (cero incluido).
- + El carácter, la clase o el grupo antes de un signo más debe aparecer como mínimo una vez.
- ? El carácter, la clase o el grupo antes del signo de interrogación es opcional y puede aparecer como máximo una vez.

- La diagonal inversa "\"

La diagonal inversa se utiliza para escapar el siguiente carácter de la expresión de búsqueda de forma que este adquiriera un significado especial o deje de tenerlo. O sea, la diagonal inversa no se utiliza nunca por sí sola, sino en combinación con otros caracteres. Al utilizarlo por ejemplo en combinación con el punto "\" este deja de tener su significado normal y se comporta como un carácter literal.

De la misma forma, cuando se coloca la barra inversa seguida de cualquiera de los caracteres especiales que se indican a continuación, estos dejan de tener su significado especial y se convierten en caracteres de búsqueda literal.

- \t — Representa un tabulador.
- \r — Representa el "retorno de carro" o "regreso al inicio" o sea el lugar en que la línea vuelve a iniciar.
- \n — Representa la "nueva línea" el carácter por medio del cual una línea da inicio. Es necesario recordar que en Windows es necesaria una combinación de \r\n para comenzar una nueva línea, mientras que en Unix solamente se usa \n y en Mac_OS clásico se usa solamente \r.
- \a — Representa una "campana" o "beep" que se produce al imprimir este carácter.
- \e — Representa la tecla "Esc" o "Escape"
- \f — Representa un salto de página
- \v — Representa un tabulador vertical
- \x — Se utiliza para representar caracteres ASCII o ANSI si conoce su código. De esta forma, si se busca el símbolo de derechos de autor y la fuente en la que se busca utiliza el conjunto de caracteres latín-1 es posible encontrarlo utilizando \xA9".

- `\u` — Se utiliza para representar caracteres Unicode si se conoce su código. `"\u00A2"` representa el símbolo de centavos. No todos los motores de Expresiones Regulares soportan Unicode. El .Net Framework lo hace, pero el EditPad Pro no, por ejemplo.
- `\d` — Representa un dígito del 0 al 9.
- `\w` — Representa cualquier carácter alfanumérico.
- `\s` — Representa un espacio en blanco.
- `\D` — Representa cualquier carácter que no sea un dígito del 0 al 9.
- `\W` — Representa cualquier carácter no alfanumérico.
- `\S` — Representa cualquier carácter que no sea un espacio en blanco.
- `\A` — Representa el inicio de la cadena. No un carácter sino una posición.
- `\Z` — Representa el final de la cadena. No un carácter sino una posición.
- `\b` — Marca la posición de una palabra limitada por espacios en blanco, puntuación o el inicio/final de una cadena.
- `\B` — Marca la posición entre dos caracteres alfanuméricos o dos no-alfanuméricos.

Notas:

^

`[A-Z]{1}`

`[0-9]{2}`

`[B-DF-HJ-NP-TV-Z]{3}`

`[0-9A-Z]{1}`

`(0[1-9]|1[0-2])(0[1-9]|1[0-9]|2[0-9]|3[0-1])`

`(AS|BC|BS|NE)`

\$

^

`[HM]`



[A-Z]{1}

[0-9]{2}

[B-DF-HJ-NP-TV-Z]{3}

[0-9A-Z]{1}

(0[1-9]|1[0-2])(0[1-9]|1[0-9]|2[0-9]|3[0-1])

(AS|BC|BS|NE)

\$