

Expresiones regulares en SDL Trados Studio

Esta nota explica que las expresiones regulares, conocidas como *regex*, pueden parecer complejas a primera vista, pero en realidad son un lenguaje más, una serie de elementos combinados en patrones para expresar un significado. Su uso en herramientas de traducción asistida por computadora como SDL Trados Studio permite al traductor optimizar el tiempo mediante búsquedas y filtrados de nivel avanzado.

| Por la Traductora Nora Díaz

Las expresiones regulares, también conocidas como *regex*, son poderosas herramientas que, al usarse en SDL Trados Studio, nos permiten, entre otras cosas, potenciar búsquedas y operaciones de filtrado mediante el empleo de una secuencia de caracteres que representa un patrón de búsqueda. Veamos un ejemplo:

`\d+\.\d{2}`

Esta *regex* se interpreta de la siguiente manera:

`\d+` Uno o más dígitos

`\.` Un punto

`\d{2}` Exactamente dos dígitos

Al aplicarse como expresión de búsqueda, por un lado, nuestra *regex* encontraría números como 9.24, 12.75 y 126.98, ya que todos ellos coinciden con el patrón: uno o más dígitos seguidos por un punto seguido por exactamente dos dígitos. Por otro lado, no encontraría números como 57,89 o 1.4, ya que no coinciden con el patrón.

Las expresiones regulares no se limitan a encontrar patrones numéricos. Veamos el siguiente ejemplo:

`[A-Z][a-z]+`

Esta *regex* se interpreta así:

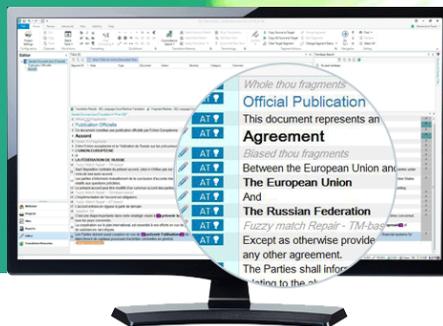
`[A-Z]` Una letra mayúscula

`[a-z]+` Una o más letras minúsculas

Con esta expresión regular, encontraremos cualquier palabra que tenga mayúscula inicial.

Componentes de las expresiones regulares

Tras entender lo que podemos lograr con las expresiones regulares, la siguiente pregunta lógica es ¿cuáles son los componentes que podemos utilizar en una *regex*? Hay mucha información acerca de este tema en diversos sitios web y libros sobre expresiones regulares, pero debemos tomar en cuenta que la gran mayoría de esa información está dirigida a programadores y no a traductores, por lo que puede parecer abrumador tratar de



encontrar lo que es útil para nuestra labor entre todo ese mar de explicaciones.

Para simplificar esta tarea, he creado una sencilla hoja de referencia que contiene todos los elementos necesarios para crear expresiones regulares durante nuestro trabajo de traducción.

Metacaracteres y caracteres especiales

Las expresiones regulares utilizan ciertos caracteres con un significado especial. Por lo tanto, cuando sea necesario utilizarlos con su significado literal, habrá que «escapar» el metacarácter. Esto se hace mediante el carácter de escape, la barra invertida \. Por ejemplo, el signo de interrogación de cierre se utiliza con un significado especial en *regex*, por lo que, cuando se requiera un signo de interrogación de cierre literal, deberá escaparse de la siguiente manera: \?. En la hoja de referencia, pueden verse todos los metacaracteres que requieren el carácter de escape cuando deben interpretarse literalmente.

Regular Expressions for Translators

Cheat Sheet

by Nora Díaz

Metacharacters	Quantifiers	Groups and ranges
^ [. \$ () ? * () \ + < >	* 0 or more + 1 or more ? 0 or 1 {2} Exactly 2 {2,} 2 or more {2,4} 2, 3 or 4 Adding a ? to a quantifier makes it lazy.	. Any character [a]b a or b (...) Group {?:} Passive (non-capturing) group [abc] Range (a or b or c) [^abc] Not a or b or c [a-z] Lowercase letter from a to z [A-Z] Uppercase letter from A to Z [0-9] Digit from 0 to 9
\ Escape character	Greedy (default) Match as many occurrences of an expression as possible.	String replacement \$n nth capturing group \$2 "xyz" in (abc{xyz}) \$1 "xyz" in (?abc){xyz} \$0 Entire string
Special characters \n New line \t Tab	Lazy (?) Match as few occurrences of an expression as possible.	Character classes \s White space \S Not white space \d Digit \D Not digit \w Word char. [A-Za-z0-9_] \W Not word character
Anchors ^ Start of segment \$ End of segment \b Word boundary \B Not word boundary	Lookaround (?=...) Lookahead (?!...) Negative lookahead (?<...) Lookbehind (?<!) Negative lookbehind	

Dado que la explicación de todos los elementos de la hoja de referencia sería demasiado extensa para este artículo, me limitaré por ahora a explicar el uso de metacaracteres y caracteres especiales.

Veamos ahora cómo funcionan en la práctica. Consideremos el siguiente texto:

Ahorro	\$100
Ahorro	£200
Ahorro	\$400
Ahorro	£300

Expresiones regulares en SDL Trados Studio

Supongamos que necesitamos encontrar todos los casos donde hay un tabulador seguido del signo de pesos. Con una *regex*, podemos hacerlo fácilmente:

```
\t\$\
```

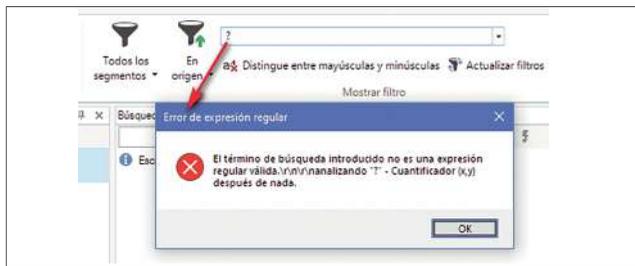
Para validar nuestra *regex*, basta con utilizar uno de los probadores disponibles en línea, como regextester.com o regextorm.net, donde podemos introducir tanto la expresión regular como el texto de prueba.



<https://www.regextester.com/>

Aplicación en SDL Trados Studio

SDL Trados Studio permite el uso de expresiones regulares en sus funciones de búsqueda, filtrado, segmentación y verificación. De hecho, las expresiones regulares son obligatorias en el filtro de pantalla básico, por lo que, al intentar buscar alguno de los metacaracteres sin el carácter de escape, no obtenemos el resultado deseado o incluso recibimos mensajes de error, como en el siguiente caso:

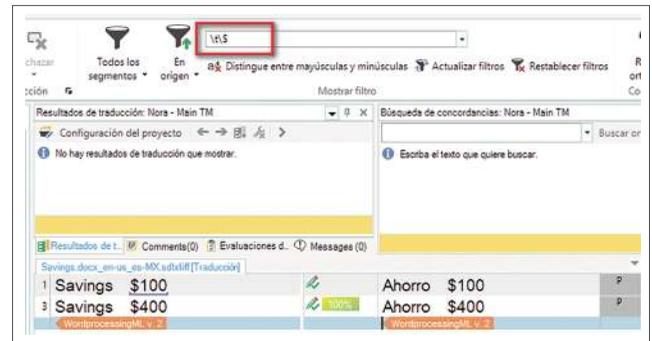


Veamos ahora el uso de nuestra *regex* de ejemplo en el filtrado y en la búsqueda en SDL Trados Studio. Este será nuestro texto de muestra:

Id	Segmento	Traducción	Calificación	Segmento	Traducción
1	Savings	\$100	100%	Ahorro	\$100
2	Savings	\$200	100%	Ahorro	£200
3	Savings	\$400	100%	Ahorro	\$400
4	Savings	£300	100%	Ahorro	£300

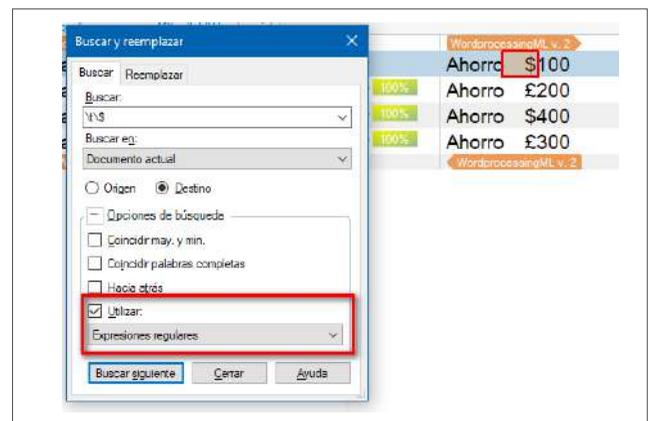
Filtrado

Al aplicar la *regex* `\t\$\` en el filtro, veremos únicamente los segmentos que coincidan con el patrón, es decir, los que tengan un tabulador seguido por un signo de pesos.



Búsqueda

A diferencia del filtro, la función de búsqueda no utiliza expresiones regulares de manera predeterminada, por lo que debemos activar la opción marcando la casilla correspondiente. Una vez hecho esto, podemos utilizar cualquier tipo de *regex* en el campo **Buscar**. Para nuestro ejemplo, la búsqueda de un tabulador seguido por un signo de pesos es como en la siguiente imagen:



Como hemos visto, las expresiones regulares pueden parecer complejas a primera vista, pero en realidad son simplemente un lenguaje más, una serie de elementos combinados en patrones para expresar un significado. Su uso en herramientas de traducción asistida por computadora como SDL Trados Studio le permite al traductor optimizar el tiempo mediante búsquedas y filtrados de nivel avanzado. ■