

Práctica LEX

Miguel Lentisco Ballesteros

8 de enero de 2018

1. Problema

El problema que he planteado para resolver con LEX surge de un problema real que he tenido yo mismo. Planteo la situación: te has descargado una serie o película (usualmente en un idioma extranjero) pero no vienen con subtítulos integrados y los necesitas, vas a tu página usual y te los descargas. La cosa se complica cuando empiezas a notar alguno de los siguientes fallos:

- Los subtítulos están retrasados o adelantados.
- Los subtítulos tienen o le faltan comentarios extras.
- Algunos subtítulos están mal escritos.
- Algunos subtítulos no están en orden.

Claro, puedes modificar tú mismo el archivo pero tienes que empezar a modificar uno por uno, cambiar números... un lío total. Y si intentas cambiar los tiempos desde tu reproductor empieza a dar problemas (y es temporal la solución) y si el tiempo es muy grande acaba por colgarse el programa.

2. Solución

La solución es muy fácil, tenemos un analizador que, usando el formato típico de los subtítulos, va a ir guardando cada subtítulo y sus tiempos de inicio y fin. Para ello usamos expresiones regulares para ir cogiendo lo que nos interesa, para hacerlo más fácil y añadirle un poco de complejidad he añadido el uso de condicionales. Esto facilita el hecho de que las líneas de texto de subtítulos pueden tener cualquier cosa y se consigue leer toda la información sin problemas sin mezclar con el resto de metainformación del

fichero.

Así el programa solo hay que introducir el nombre el fichero de subtítulos y se despliega un menú al usuario en el que se puede hacer todo lo que el usuario desee para modificar los subtítulos. Al final se creará un archivo llamado 'subtitulosModificados.srt' en el que vendrán hechas todas las modificaciones que el usuario haya hecho.

3. Uso

Estando en la misma carpeta donde se adjunta todo, el ejemplo de uso sería 'make ejemplo'. O en general se podría ejecutar 'modificarSubtitulos x' donde x es el nombre del archivo de subtítulos.

En la carpeta 'subtitulos' se incluye una colección de subtitulos ya descargados con la que se puede probar el programa.

4. Aclaraciones

Intenté ver como podía ampliar y descargar el archivo de subtítulos desde el programa, pero vi que no era una cosa tan fácil de sacar desde el .html debido a que la página redirecciona a varias páginas y encima para descargarte el archivo no es de una página en concreta sino que en un enlace tienes que pulsar en un botón (usa un script) lo cual ya dejaba de poder simplificarse en buscar una url en concreto con LEX y no sabía como poder sacarlo.

Para compensar metí las condicionales al hacer las regex en LEX para darle un poco más de complejidad y equilibrar un poco la dificultad que se planteaba al hacerlo.

5. Atención

Se asume que los cambios que hace el usuario son correctos (tiempos, orden, eliminación...), en cualquier caso el archivo original se mantiene intacto luego en el peor de los casos se puede empezar de nuevo.

Mucho cuidado con la codificación, puede haber problemas si se usa codificación con saltos de línea de Windows o Linux (en Linux funciona, en cualquiera de los casos abrir con un procesador de textos y cambiar los saltos a Linux).