

Práctica 2

Manejo Avanzado de *Flex*

Tareas

1. Estudia la sección sobre las *condiciones de arranque* en *Intro_Flex_Bison.pdf* (páginas 7 a 9) y luego el capítulo 9 del manual de *Flex* (*flex-es-2.5.pdf*).
2. Lee la introducción de esta práctica y realiza el ejercicio propuesto.
3. Elabora la memoria de la práctica y entrégala junto con el fichero fuente según el Procedimiento de Entrega de Prácticas explicado en la Introducción a las Prácticas de la Asignatura. La fecha tope de entrega será hasta el día anterior al comienzo de la Práctica 3.

Nota: El incumplimiento de las normas de entrega se reflejará en la calificación de la práctica.

Se recuerda especialmente lo siguiente:

- Debes crear un directorio que contenga **exclusivamente** el fichero con la memoria en formato *PDF*, el fichero fuente con tu código (*.l* de *Flex*) y aquellos de prueba que hayas generado (*.html*). No utilices subdirectorios.
- En caso de que el fichero *.zip* resultante tenga un tamaño mayor de 512 KB deberás repetir la creación del directorio dividiéndolo en varios ficheros *.zip* de como máximo 512KB cada uno, en ese caso llama a los ficheros resultantes *nipPrX1.zip*, *nipPrX2.zip*, etc.

Introducción

El objetivo principal de esta práctica es aprender a desarrollar analizadores léxicos en *Flex* más sofisticados, profundizando en el manejo de lo que se conoce como *condiciones de arranque*. Las *condiciones de arranque* no son imprescindibles, ya que siempre se pueden emular utilizando código *C* en las acciones de los patrones. No obstante, su uso facilita mucho el desarrollo de los programas en *Flex*, ya que ayudan a estructurar conjuntos de patrones/acciones en función de un contexto determinado o condiciones previas.

Ejercicio 1

El objetivo central de esta práctica es utilizar las condiciones de arranque para extraer y reformatear información de páginas *HTML*. La aplicación concreta que se plantea es la extracción de información sobre horarios de autobuses de la página web del servicio de los autobuses urbanos de Zaragoza.

Para ello, la idea es generar en tiempo real un fichero *.html* con la información de un poste o parada y luego procesarlo para eliminar todos los *TAGs* propios de la sintaxis de *HTML* (delimitados por '<' y '>'). Como salida, sólo se mostrará por pantalla el número de línea, destino y tiempo estimado de llegada de los autobuses a la parada.

Para facilitar las pruebas, podeis descargar el fichero *p2_ej1.l* de la web de la asignatura. El código incluye lo necesario para la descarga de ficheros *.html* mediante una función auxiliar llamada *fich_poste(-,-)*. Dicha función auxiliar utiliza llamadas a la librería *curl* para gestionar la descarga de las páginas *HTML* dinámicamente.

Nota: Para poder compilar el ejecutable *C* en *Hendrix* es necesario añadir a la instrucción de compilación las siguientes opciones:

```
-L/usr/local/lib -lcurl -lssl -lcrypto -lldap -lrt -lsocket -lnsl -ldl -lz
```