

Procesamiento de Lenguajes (PL)

Curso 2018/2019

Práctica 4: traductor ascendente usando bison/flex

Fecha y método de entrega

La práctica debe realizarse de forma individual, como las demás prácticas de la asignatura, y debe entregarse a través del servidor de prácticas del DLSI **antes de las 23:59 del 28 de abril de 2019**. Los ficheros fuente “plp4.l”, “plp4.y” y “comun.h” se comprimirán en un fichero llamado “plp4.tgz”, sin directorios.

Al servidor de prácticas del DLSI se puede acceder de dos maneras:

- Desde la web del DLSI (<http://www.dlsi.ua.es>), en el apartado “Entrega de prácticas”
- Desde la URL <http://pracdlsi.dlsi.ua.es>

Una vez en el servidor, se debe elegir la asignatura PL y seguir las instrucciones.

Descripción de la práctica

La práctica consiste en implementar un traductor ascendente utilizando las herramientas **bison** y **flex**, versiones modernas de **yacc** y **lex**, para el mismo proceso de traducción de la práctica 3.

Gramática

La gramática que debe utilizarse como base para diseñar el traductor es la siguiente:

S	\rightarrow	program id pyc B
D	\rightarrow	var L endvar
L	\rightarrow	$L V$
L	\rightarrow	V
V	\rightarrow	id dosp C pyc
C	\rightarrow	$A C$
C	\rightarrow	P
A	\rightarrow	array cori R cord of
R	\rightarrow	R coma G
R	\rightarrow	G
G	\rightarrow	numentero ptopto numentero
P	\rightarrow	pointer of P
P	\rightarrow	$Tipo$
$Tipo$	\rightarrow	integer
$Tipo$	\rightarrow	real
B	\rightarrow	begin D SI end
SI	\rightarrow	SI pyc I
SI	\rightarrow	I
I	\rightarrow	id asig E
I	\rightarrow	write pari E pard
I	\rightarrow	B
E	\rightarrow	E opas T
E	\rightarrow	T
T	\rightarrow	T opmul F
T	\rightarrow	F
F	\rightarrow	numentero
F	\rightarrow	numreal
F	\rightarrow	id

Observa que, para facilitar el diseño del traductor (especialmente en lo relacionado con los atributos heredados y su implementación en traductores ascendentes), hay muchos no terminales que presentan recursividad por la izquierda, que **no** debes eliminar.

Mensajes de error

Los errores léxicos y semánticos deben generar exactamente el mismo mensaje que en la práctica 3, con el mismo formato. En el caso de los mensajes de error sintáctico, el mensaje de error debe ser simplemente:

```
Error sintactico (fila,columna): en 'lexema'
```

Esto se debe a que las tablas generadas por **bison** son difíciles de interpretar y por tanto, por simplificar, no es necesario mostrar los tokens que se esperaban en el lugar del token incorrecto.

En la página web de la asignatura se dejará un fichero con una función para emitir mensajes de error léxico y sintáctico.

Notas técnicas

1. Los programas **yacc** (**bison**) y **lex** (**flex**), a partir de una especificación léxica y un ETDS, generan un traductor ascendente basado en un analizador sintáctico LALR(1), escrito en C/C++. Junto con este enunciado se publicará una hoja técnica acerca de estas herramientas, con un ejemplo práctico.
2. Como en las prácticas anteriores, se publicará un autocorrector para facilitar la tarea de depurar la práctica.
3. La práctica debe estar contenida en estos ficheros (como en el ejemplo de la hoja técnica): **comun.h**, **plp4.1** y **plp4.y**
4. El proceso de traducción es el mismo que en la práctica 3, para lo que es necesario utilizar una tabla de símbolos (con básicamente dos funciones, **anyadir** y **buscar**)
5. Para comprobar que después de un programa correcto no aparecen más tokens en el fichero, es decir, que el siguiente token es el del final del fichero, es necesario hacer una comprobación en la acción semántica situada al final de la regla:

$$S \longrightarrow \text{program id pyc } B$$

En dicha acción semántica se debe llamar directamente al analizador léxico:

```
int token = yylex();
```

Si el valor de **token** es 0, el fichero termina correctamente. Si es cualquier otro valor, hay que generar un error sintáctico con la llamada **yyerror("")**

6. Recuerda que el lenguaje fuente no es *case sensitive*, por lo que por ejemplo la palabra reservada **“var”** puede aparecer como **“VAR”**, **“Var”** o **“VaR”**, o cualquier combinación de mayúsculas y minúsculas. La expresión regular que reconoce ese lenguaje sería:

```
[Vv] [Aa] [Rr]
```