Diseño inicial del lenguaje MMPor

IG29: Compiladores e intérpretes

Segunda sesión de prácticas

1. Introducción

Este documento recoge el modelado léxico, sintáctico y semántico de un pequeño lenguaje de calculadora al que hemos dado el nombre de MMPor ya que, inicialmente, sólo es capaz de expresar sumas, restas y multiplicaciones (las operaciones *más*, *menos* y *por*) entre números enteros. Sobre este diseño, que ya te ha sido presentado en sesiones anteriores (la clase de presentación de la asignatura y la primera sesión de prácticas), deberás ir introduciendo mejoras hasta convertir tu calculadora en el intérprete de un pequeño lenguaje de programación. Estas mejoras se te irán solicitando en sucesivos boletines de prácticas.

2. Nivel léxico

Como primera *especificación léxica* de MMPor, se ha propuesto la siguiente:

Categoría léxica	Expresión regular	Atributos	Acciones
lit	`[0-9]+`	*val*	Calcular valor en *val* y emitir
opad	`\+|-`	*lex*	Copiar lexema en *lex* y emitir
opmul	`\*`	—	Emitir
blanco	`[ \t\n]+`	—	Omitir
pa	`\(`	—	Emitir
pc	`\)`	—	Emitir

3. Nivel sintáctico

El primer modelado sintáctico del lenguaje se ha llevado a cabo con la siguiente *gramática RLL(1)*:

$$\langle \mathsf{E} \rangle \rightarrow \langle \mathsf{S} \rangle \ (\ \textbf{opad}\ \langle \mathsf{S} \rangle\ )^*$$
$$\langle \mathsf{S} \rangle \rightarrow \langle \mathsf{F} \rangle \ (\ \textbf{opmul}\ \langle \mathsf{F} \rangle\ )^*$$
$$\langle \mathsf{F} \rangle \rightarrow \textbf{lit} \mid \textbf{pa}\ \langle \mathsf{E} \rangle\ \textbf{pc}$$

4. Nivel semántico

Para la *representación semántica* de las expresiones, se utilizan ASTs construidos con nodos de las clases

Clase de nodo	Significado	Atributos	Hijos
Más	Una operación de suma	—	*hi* y *hd*
Menos	Una operación de resta	—	*hi* y *hd*
Por	Una operación de multiplicación	—	*hi* y *hd*
Cte	Un literal numérico	*v*	—

donde *v* representa un valor numérico y *hi* y *hd* son los hijos izquierdo y derecho de una operación binaria.

La construcción del árbol propiamente dicha se lleva a cabo mediante el siguiente *esquema de traducción*:

$\langle \mathsf{E} \rangle \rightarrow \langle \mathsf{S} \rangle_1\ \{\ aux := \langle \mathsf{S} \rangle_1.arb\ \}$
$(\ \textbf{opad}\ \langle \mathsf{S} \rangle_2\ \{\ \text{SI (}\textbf{opad}.lex = +\text{) ENT: } aux := \text{Más(}\ aux, \langle \mathsf{S} \rangle_2.arb\ \text{) SI\_NO: } aux := \text{Menos(}\ aux, \langle \mathsf{S} \rangle_2.arb\ \text{) FIN}\ \}\ )^*$
$\{\ \langle \mathsf{E} \rangle.arb := aux\ \}$

$\langle \mathsf{S} \rangle \rightarrow \langle \mathsf{F} \rangle_1\ \{\ aux := \langle \mathsf{F} \rangle_1.arb\ \}$
$(\ \textbf{opmul}\ \langle \mathsf{F} \rangle_2\ \{\ aux := \text{Por(}\ aux, \langle \mathsf{F} \rangle_2.arb\ \text{)}\ \}\ )^*$
$\{\ \langle \mathsf{S} \rangle.arb := aux\ \}$

$\langle \mathsf{F} \rangle \rightarrow \textbf{lit}\ \{\ \langle \mathsf{F} \rangle.arb := \text{Cte(}\ \textbf{lit}.val\ \text{)}\ \}$
$\langle \mathsf{F} \rangle \rightarrow \textbf{pa}\ \langle \mathsf{E} \rangle\ \textbf{pc}\ \{\ \langle \mathsf{F} \rangle.arb := \langle \mathsf{E} \rangle.arb\ \}$