

Memoria de la Práctica de Procesadores de Lenguajes: Analizador Sintáctico

Diego José Abengózar Vilar, Alejandro García Castellanos,
Ignacio Javier Encinas Ramos

Grupo 82

November 25, 2019

Índice

1	Diseño del Analizador Sintáctico	2
1.1	Gramática	2
1.2	Autómata Reconocedor de Prefijos Viabes	3
1.2.1	Estados del autómata	4
1.3	Conflictos	4
1.4	Matriz de Transiciones	4
2	Tabla de Símbolos: Estructura e implementación	5
3	Anexo de Pruebas	5

1 Diseño del Analizador Sintáctico

1.1 Gramática

$Terminales = \{ ; \{ \} id \ ent \ cadena \ (\) \ + \ < \ ! \ = \ | \ = \ var \ int$
 $boolean \ string \ print \ input \ , \ return \ function \ if \ else \}$

$NoTerminales = \{ P \ D \ T \ F \ T1 \ A \ K \ C \ S \ L \ M \ Q \ S1 \ G \ X \ E \ U \ R \ V \ S2 \}$

$Axioma = P$

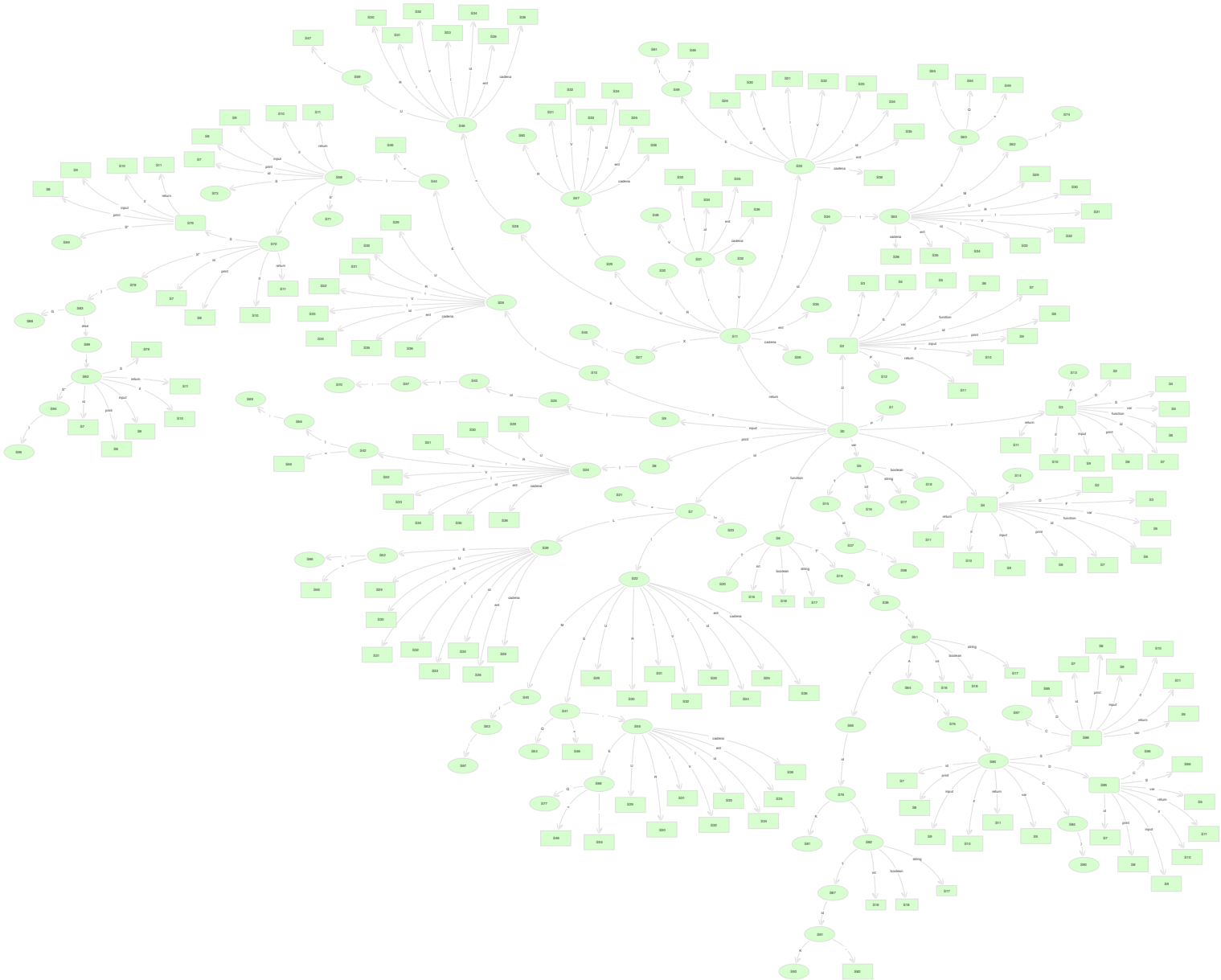
$Producciones = \{$

- $P \rightarrow D \ P$
- $P \rightarrow F \ P$
- $P \rightarrow S \ P$
- $D \rightarrow var \ T \ id \ ;$
- $T \rightarrow int$
- $T \rightarrow string$
- $T \rightarrow boolean$
- $F \rightarrow function \ T1 \ id \ (\ A \) \ \{ \ C \}$
- $T1 \rightarrow \lambda$
- $T1 \rightarrow T$
- $A \rightarrow T \ id \ K$
- $A \rightarrow \lambda$
- $K \rightarrow \lambda$
- $K \rightarrow , \ T \ id \ K$
- $C \rightarrow D \ C$
- $C \rightarrow S \ C$
- $C \rightarrow \lambda$
- $S \rightarrow id \ L \ E \ ;$
- $S \rightarrow id \ (\ M \) \ ;$
- $S \rightarrow print \ (\ E \) \ ;$
- $S \rightarrow input \ (\ id \) \ ;$
- $S \rightarrow if \ (\ E \) \ S1$
- $S \rightarrow return \ X \ ;$
- $L \rightarrow | =$
- $L \rightarrow =$
- $M \rightarrow E \ Q$
- $M \rightarrow \lambda$
- $Q \rightarrow \lambda$
- $Q \rightarrow , \ E \ Q$
- $S1 \rightarrow \{ \ S2 \} \ G$
- $S1 \rightarrow S$
- $G \rightarrow else \ \{ \ S2 \}$
- $G \rightarrow \lambda$
- $X \rightarrow E$
- $X \rightarrow \lambda$
- $E \rightarrow E < U$
- $E \rightarrow U$
- $U \rightarrow U + R$
- $U \rightarrow R$

$$\begin{aligned}
R &\rightarrow ! V \\
R &\rightarrow V \\
V &\rightarrow (E) \\
V &\rightarrow id \\
V &\rightarrow id (M) \\
V &\rightarrow ent \\
V &\rightarrow cadena \\
S2 &\rightarrow S S2 \\
S2 &\rightarrow S \\
P &\rightarrow \lambda
\end{aligned}$$

}

1.2 Autómata Reconocedor de Prefijos Viables



1.2.1 Estados del autómata

$S_0 = \{P1 \rightarrow \bullet P, P \rightarrow \bullet DP, P \rightarrow \bullet SP, P \rightarrow \bullet , D \rightarrow \bullet \text{ var } T \text{ id};, \\ F \rightarrow \bullet \text{ function } T1 \text{ id}(A)\{C\}, S \rightarrow \bullet \text{ id } L = E;, S \rightarrow \bullet \text{ id}(M); \\ S \rightarrow \bullet \text{ print}(E);, S \rightarrow \bullet \text{ input}(\text{id});, S \rightarrow \bullet \text{ return } X;\}$
 $S_1 = \{P1 \rightarrow P \bullet \}$
 $S_2 = \{P \rightarrow D \bullet P, P \rightarrow \bullet DP, P \rightarrow \bullet SP, P \rightarrow \bullet , D \rightarrow \bullet \text{ var } T \text{ id};, \\ F \rightarrow \bullet \text{ function } T1 \text{ id}(A)\{C\}, S \rightarrow \bullet \text{ id } L = E;, S \rightarrow \bullet \text{ id}(M); \\ S \rightarrow \bullet \text{ print}(E);, S \rightarrow \bullet \text{ input}(\text{id});, S \rightarrow \bullet \text{ return } X;\}$
 $S_3 = \{\}$

1.3 Conflictos

Como podemos observar en la tabla de decisión no hay ningún conflicto.

Los posibles conflictos son:

Reducción-Reducción

Podemos ver como en los posibles estados con este conflicto, $S_{x...}$, se verifica que

$\forall \{A \rightarrow \alpha \bullet, B \rightarrow \beta \bullet\} \subset S_x \Rightarrow \text{Follow}(A) \cap \text{Follow}(B) = \emptyset$ (Esto lo podemos observar al no tener dos entradas de reducción en la misma celda de cada fila de S_x)

Reducción-Desplazamiento

Podemos ver como en los posibles estados con este conflicto, $S_{x...}$, se verifica que

$\forall \{A \rightarrow \alpha \bullet b \gamma, C \rightarrow \beta \bullet\} \subset S_x \Rightarrow b \notin \text{Follow}(C)$ (Esto lo podemos observar al no tener una entrada de desplazamiento y otra de reducción en la misma celda de cada fila de S_x)

1.4 Matriz de Transiciones

MT_AFD		letra	digito	'	/	-	carácter	*	delimitador
$\rightarrow 0$	1 lee	2 C	3A	4 lee	5 lee	-1 error	-1 error	-1 error	0 lee
1	-1 error	-1 error	-1 error	-1 error	-1 error	-1 error	-1 error	-1 error	-1 error
2	9 G2	2 C	2 C	9 G2	9 G2	2 C	9 G2	9 G2	9 G2
3	10 G3	10 G3	3 B	10 G3	10 G3	10 G3	10 G3	10 G3	10 G3
4	4 C	4 C	4 C	11 G4	4 C	4 C	4 C	4 C	4 C
5	-1 error	-1 error	-1 error	-1 error	-1 error	-1 error	-1 error	6 lee	-1 error
6	6 lee	6 lee	6 lee	6 lee	6 lee	6 lee	6 lee	7 lee	6 lee
7	6 lee	6 lee	6 lee	6 lee	0 lee	6 lee	6 lee	7 lee	6 lee

MT_AFD	;	{	}	()	+	<	!	=	,
$\rightarrow 0$	12 G5	13 G6	14 G7	15 G8	16 G9	17 G10	18 G11	19 G12	20 G13	21 G14
1	-1 error	-1 error	-1 error	-1 error	-1 error	-1 error	-1 error	-1 error	8 G1	-1 error
2	9 G2	9 G2	9 G2	9 G2	9 G2	9 G2	9 G2	9 G2	9 G2	9 G2
3	10 G3	10 G3	10 G3	10 G3	10 G3	10 G3	10 G3	10 G3	10 G3	10 G3
4	4 C	4 C	4 C	4 C	4 C	4 C	4 C	4 C	4 C	4 C
5	-1 error	-1 error	-1 error	-1 error	-1 error	-1 error	-1 error	-1 error	-1 error	-1 error
6	6 lee	6 lee	6 lee	6 lee	6 lee	6 lee	6 lee	6 lee	6 lee	6 lee
7	6 lee	6 lee	6 lee	6 lee	6 lee	6 lee	6 lee	6 lee	6 lee	6 lee

2 Tabla de Símbolos: Estructura e implementación

Contiene la información de los identificadores, de los cuales se guardan los campos: lexema, tipo y desplazamiento. Para las funciones, además, se guardará el número de parámetros, su tipo, la forma de paso de parámetros y el tipo del valor de retorno.

La tabla de símbolos estará formada por dos matrices de tamaño dinámico; la primera contendrán los identificadores de ámbito global y la segunda del local. Así pues, esta segunda se creará al encontrar la declaración de una función y se borrará al acabar de ser declarada. También se utiliza un flag de declaración o uso (FlagDeclUso), un flag para saber cual es la tabla actual y dos más para el valor del desplazamiento en cada una de las tablas.

Sin embargo, en la implementación actual sólo se usa una tabla y siempre se supone que está el FlagDeclUso = Uso, pero en el caso de que no este declarada la variable se insertará en la tabla actual, ya que requerimos de la implementación del Analizador Semántico para poder saber cuando se cambia de ámbito y cuando se están declarando o usando identificadores. Así que, la acción semántica que genera los tokens de los identificadores quedaría temporalmente así:

```
G2: if (lex ∈ palRes)
    GEN_TOKEN(palRes, -)
    else if ((p:= BUSCA_TS(lex))=NULL)
        p:=INSERTAR_TS(lex)
        GEN_TOKEN(ID, p)
```

3 Anexo de Pruebas

Error 1: Número fuera de rango.

Fuente:

```
var int a;
var int b;
a = 33333;
b = a;
if (a < b) b = 1;
if (b < a) b = 8;
a = a + b;
print (a);
print (b);
```

Tokens:

```
<DEC, >
<TipoVarENT, >
<ID, 0>
<PuntoComa, >
<DEC, >
<TipoVarENT, >
<ID, 1>
```

<PuntoComa, >
<ID, 0>
<ASIG, >

Tabla Simbolos#1:

* LEXEMA: 'a'
* LEXEMA: 'b'

Error 2: Transición no prevista

Fuente:

```
var string texto; /*Comentario bueno*/  
function imprime (string msg)  
{  
    print (msg);  
}  
/ Comentario malo*/  
function pideTexto ()  
{  
    print ( 'Introduce un texto' );  
    input (texto);  
}  
pideTexto();  
var string textoAux;  
textoAux = texto;  
imprime (textoAux);
```

Tokens:

<DEC, >
<TipoVarCAD, >
<ID, 0>
<PuntoComa, >
<DECFunc, >
<ID, 1>
<ParentesisAbrir, >
<TipoVarCAD, >
<ID, 2>
<ParentesisCerrar, >
<CorcheteAbrir, >
<Print, >
<ParentesisAbrir, >
<ID, 2>
<ParentesisCerrar, >
<PuntoComa, >
<CorcheteCerrar, >

Tabla Simbolos #1:

* LEXEMA: 'texto'
* LEXEMA: 'imprime'
* LEXEMA: 'msg'

Error 3: Transición no prevista

Fuente:

```
var string texto;
function pideTexto ()
{ print ('Introduce un texto);
input (texto);
} function imprime (string msg)
{
print ('Mensaje introducido:');
print (msg);
}
pideTexto();
var string textoAux;
textoAux = texto;
imprime (textoAux);
```

Tokens:

```
<DEC, >
<TipoVarCAD, >
<ID, 0>
<PuntoComa, >
<DECFunc, >
<ID, 1>
<ParentesisAbrir, >
<ParentesisCerrar, >
<CorcheteAbrir, >
<Print, >
<ParentesisAbrir, >
<CAD, "Introduce un texto);
input (texto);
}
function imprime (string msg)
{
print (" >
<ID, 2>
<ID, 3>
```

Tabla Simbolos #1:

```
* LEXEMA: 'texto'
* LEXEMA: 'pideTexto'
* LEXEMA: 'Mensaje'
* LEXEMA: 'introducido'
```

Pruebas pasadas con éxito:

1:

Fuente:

```
var int a;
var int b;
a = 3;
```

```

b = a;
var boolean c;
c = a < b;
if (c) b = 1;
c = b < a;
if (c) b = 4;
a = a + b;
print (a);
print (b);

```

Tokens:

```

<DEC, >
<TipoVarENT, >
<ID, 0>
<PuntoComa, >
<DEC, >
<TipoVarENT, >
<ID, 1>
<PuntoComa, >
<ID, 0>
<ASIG, >
<ENT, 3>
<PuntoComa, >
<ID, 1>
<ASIG, >
<ID, 0>
<PuntoComa, >
<DEC, >
<TipoVarLOG, >
<ID, 2>
<PuntoComa, >
<ID, 2>
<ASIG, >
<ID, 0>
<MENOR, >
<ID, 1>
<PuntoComa, >
<IF, >
<ParentesisAbrir, >
<ID, 2>
<ParentesisCerrar, >
<ID, 1>
<ASIG, >
<ENT, 1>
<PuntoComa, >
<ID, 2>
<ASIG, >
<ID, 1>
<MENOR, >

```



```

<ID, 0>
<PuntoComa, >
<IF, >
<ParentesisAbrir, >
<ID, 2>
<ParentesisCerrar, >
<ID, 1>
<ASIG, >
<ENT, 4>
<PuntoComa, >
<ID, 0>
<ASIG, >
<ID, 0>
<SUMA, >
<ID, 1>
<PuntoComa, >
<Print, >
<ParentesisAbrir, >
<ID, 0>
<ParentesisCerrar, >
<PuntoComa, >
<Print, >
<ParentesisAbrir, >
<ID, 1>
<ParentesisCerrar, >
<PuntoComa, >

```

Tabla Simbolos #1:

```

* LEXEMA: 'a'
* LEXEMA: 'b'
* LEXEMA: 'c'

```

2:

Fuente:

```

var int a;
var int b;
var int c;
print ( 'Introduce el primer operando' );
input (a);
print ( 'Introduce el segundo operando' );
input (b);
function int suma (int num1, int num2)
{
return num1+num2;
}
c = suma (a, b);
print (c);

```

Tokens:

```

<DEC, >
<TipoVarENT, >
<ID, 0>
<PuntoComa, >
<DEC, >
<TipoVarENT, >
<ID, 1>
<PuntoComa, >
<DEC, >
<TipoVarENT, >
<ID, 2>
<PuntoComa, >
<Print, >
<ParentesisAbrir, >
<CAD, "Introduce el primer operando">
<ParentesisCerrar, >
<PuntoComa, >
<Input, >
<ParentesisAbrir, >
<ID, 0>
<ParentesisCerrar, >
<PuntoComa, >
<Print, >
<ParentesisAbrir, >
<CAD, "Introduce el segundo operando">
<ParentesisCerrar, >
<PuntoComa, >
<Input, >
<ParentesisAbrir, >
<ID, 1>
<ParentesisCerrar, >
<PuntoComa, >
<DECFunc, >
<TipoVarENT, >
<ID, 3>
<ParentesisAbrir, >
<TipoVarENT, >
<ID, 4>
<Coma, >
<TipoVarENT, >
<ID, 5>
<ParentesisCerrar, >
<CorcheteAbrir, >
<Return, >
<ID, 4>
<SUMA, >
<ID, 5>
<PuntoComa, >

```

<CorcheteCerrar, >
 <ID, 2>
 <ASIG, >
 <ID, 3>
 <ParentesisAbrir, >
 <ID, 0>
 <Coma, >
 <ID, 1>
 <ParentesisCerrar, >
 <PuntoComa, >
 <Print, >
 <ParentesisAbrir, >
 <ID, 2>
 <ParentesisCerrar, >
 <PuntoComa, >

Tabla Simbolos #1:

* LEXEMA: 'a'
 * LEXEMA: 'b'
 * LEXEMA: 'c'
 * LEXEMA: 'suma'
 * LEXEMA: 'num1'
 * LEXEMA: 'num2'

3:

Fuente:

```

var int a;
var int b;
a = 3;
b = a;
if (a < b) b = 1;
a = a + b;
print (a);
print (b);

```

Tokens:

<DEC, >
 <TipoVarENT, >
 <ID, 0>
 <PuntoComa, >
 <DEC, >
 <TipoVarENT, >
 <ID, 1>
 <PuntoComa, >
 <ID, 0>
 <ASIG, >
 <ENT, 3>
 <PuntoComa, >
 <ID, 1>

<ASIG, >
 <ID, 0>
 <PuntoComa, >
 <IF, >
 <ParentesisAbrir, >
 <ID, 0>
 <MENOR, >
 <ID, 1>
 <ParentesisCerrar, >
 <ID, 1>
 <ASIG, >
 <ENT, 1>
 <PuntoComa, >
 <ID, 0>
 <ASIG, >
 <ID, 0>
 <SUMA, >
 <ID, 1>
 <PuntoComa, >
 <Print, >
 <ParentesisAbrir, >
 <ID, 0>
 <ParentesisCerrar, >
 <PuntoComa, >
 <Print, >
 <ParentesisAbrir, >
 <ID, 1>
 <ParentesisCerrar, >
 <PuntoComa, >

Tabla Simbolos #1:

* LEXEMA: 'a'

* LEXEMA: 'b'