

UNIVERSIDAD TECNOLÓGICA NACIONAL



FACULTAD REGIONAL BUENOS AIRES

SINTAXIS Y SEMÁNTICA DE LOS LENGUAJES

Compilador de lenguaje Micro

Autores:

Franco BATTAGLIA

Lautaro MARTÍNEZ LARROULET

Kevin VARGAS NAVIA

Guillermo VERDILE

Profesor:

Dr. Oscar BRUNO

23 de octubre de 2017

Índice

I	Especificación del lenguaje Micro	1
1.	Definición formal	2
1.1.	Gramática léxica	3
1.2.	Gramática sintáctica	3
II	Analizador lexicográfico	4
1.	Marco teórico	4
2.	Herramienta utilizada	4
3.	Estructura de la entrada	4
4.	Aplicación a Micro	5
III	Analizador sintáctico	6
1.	Marco teórico	6
2.	Herramienta utilizada	6
3.	Estructura de la entrada	6
4.	Aplicación a Micro	7
IV	Analizador semántico y proceso de síntesis	9
1.	Consideraciones	9
2.	Implementación en Micro	9
3.	Conclusión	13

Prefacio

El código correspondiente a este trabajo, con documentación de apoyo, se encuentra almacenado en <https://github.com/francobatta/compiladorMicro>. Allí existen dos branches de código distintas: **master** almacena los analizadores léxico, sintáctico y semántico; **síntesis** agrega la parte correspondiente al código. La bibliografía consultada fue:

- *Muchnik, Jorge Daniel. Sintaxis y semántica de los lenguajes: desde el compilador.* - 1º ed. - Buenos Aires: Centro de Estudiantes de Ingeniería Tecnológica - CEIT, 2010.
- Documentación de Bison <https://www.gnu.org/software/bison/manual/>

Parte I

Especificación del lenguaje Micro

El lenguaje sobre el cual trabajamos se llama Micro y fue creado por Charles Fischer. Micro posee una estructura muy simple que permite realizar de forma simple la construcción de un compilador básico que lo represente. Informalmente, Fischer describe al lenguaje Micro como:

- El único tipo de dato es entero.
- Todos los identificadores son declarados implícitamente y con una longitud máxima de 32 caracteres.
- Los identificadores deben comenzar con una letra y están compuestos de letras y dígitos.
- Las constantes son secuencias de dígitos (números enteros).
- Hay dos tipos de sentencias:
 - Asignación
ID := Expresión;
Expresión es infija y se construye con identificadores, constantes y los operadores + y -; los paréntesis están permitidos.
 - Entrada/Salida
leer (lista de IDs);
escribir (lista de Expresiones);
- Cada sentencia termina con un “punto y coma” (;). El cuerpo de un programa está delimitado por inicio y fin.
- inicio, fin, leer y escribir son palabras reservadas y deben escribirse en minúscula.

1. Definición formal

No obstante, para realizar la construcción de un compilador se necesita una definición formal. La misma consiste, en el caso de Micro, de una gramática *léxica* (que define los posibles lexemas que se pueden encontrar en un programa Micro) y una gramática *sintáctica* (que permite precisar las construcciones válidas en Micro).

Para la implementación del compilador es muy importante la colaboración de ambas especificaciones: la sintáctica recibirá los lexemas de la léxica, para aplicarlos a las reglas de construcción propias. En nuestro caso, la especificación formal sintáctica también se encarga de realizar comprobaciones semánticas, y por último, de realizar una síntesis del código fuente Micro a código C.

1.1. Gramática léxica

<token> → uno de <identificador><constante><palabraReservada>
 <operadorAditivo><asignación><carácterPuntuación>
 <identificador> → <letra>{<letra o dígito>}
 <constante> → <dígito>{<dígito>}
 <letra o dígito> → uno de <letra><dígito>
 <letra> → una de **a-z A-Z**
 <dígito> → uno de **0-9**
 <palabraReservada> → una de **inicio fin leer escribir**
 <operadorAditivo> → uno de **+ -**
 <asignación> → **:=**
 <carácterPuntuación> → uno de **(, ;**

Los lexemas que constituyen un programa fuente en este LP Micro son: identificadores, constantes, las cuatro palabras reservadas ya mencionadas, paréntesis izquierdo y derecho, el “punto y coma” (para terminar cada sentencia), la “coma” (para formar una lista), el símbolo de asignación (:=), y los operadores de suma (+) y de resta (-).

1.2. Gramática sintáctica

<programa> → **inicio** <listaSentencias> **fin**
 <listaSentencias> → <sentencia>{<sentencia>}
 <sentencia> → <identificador> := <expresión>; | **leer** (<listaIdentificadores>; |
escribir (<listaExpresiones>;)
 <listaIdentificadores> → <identificador> {, <identificador>}
 <listaExpresiones> → <expresión> {, <expresión>}
 <expresión> → <primaria>{<operadorAditivo><primaria>}
 <primaria> → <identificador> | <constante> | (<expresión>)

Tabla de Tokens (a modo informativo)

En el programa fuente	Nombre del Token
Inicio	INICIO
Fin	FIN
Leer	LEER
Escribir	ESCRIBIR
:=	ASIGNACION
(PARENIZQUIERDO
)	PARENDERECHO
,	COMA
;	PUNTOYCOMA
+	SUMA
-	RESTA

Parte II

Analizador lexicográfico

Un analizador léxico, analizador lexicográfico o *Scanner* es la primera fase de un compilador consistente en un programa que produce una salida compuesta de *tokens* (componentes léxicos) o símbolos.

1. Marco teórico

El análisis léxico es realizado por un módulo llamado Scanner. Este analizador lee, uno a uno, los caracteres que forman un programa fuente, y produce una secuencia de representaciones de Tokens. Es muy importante tener en cuenta que el Scanner es una rutina que produce y retorna la representación del correspondiente token, uno por vez, en la medida que es invocada por el *Parser*. Existen dos formas principales de implementar un Scanner:

1. A través de la utilización de un programa auxiliar tipo lex, en el que los datos son Tokens representados mediante Expresiones Regulares (método usado en este trabajo),
2. Mediante la construcción de una rutina basada en el diseño de un apropiado AFD (Autómata Finito Determinístico).

2. Herramienta utilizada

Para realizar el analizador léxico se usó el programa flex¹, que representa una alternativa de código abierto derivada del software lex. Flex recibe un archivo de extensión .l que especifica, de forma abstracta, la gramática léxica deseada, y construye como salida un archivo llamado lex.yy.c que contiene la gramática en forma de código fuente C.

Para compilar un archivo de entrada usando flex, basta con escribir `flex «ruta entrada»` desde la línea de comando del sistema.

3. Estructura de la entrada

La estructura del archivo de extensión .l usado como entrada se divide en tres secciones separadas por el símbolo `%%`. Las secciones representan, en el orden correspondiente:

1. Inclusión de **macros** y **librerías** de C, declaración de variables.
Definición de **tokens** en la forma `«nombre símbolos_asociados»` separados por un renglón.
2. Definición de **reglas**: meta expresiones regulares formadas por tokens y operadores asociados a las metaER, acompañadas por porciones de código C asociadas a la detección de dichas metaER. Son de la forma `«regla {código_C}»` separadas por un renglón.
3. Inclusión de **funciones** y **procedimientos** escritos en C, sobre todo útiles para ser llamados en la sección 2 para brindar información complementaria sobre la detección de tokens. Nota: una vez hecha la especificación sintáctica, esta sección deja de tener sentido y se desarrolla en dicha especificación.

¹Fast lexical analyzer generator. Descarga para Windows en <http://gnuwin32.sourceforge.net/packages/flex.htm>

4. Aplicación a Micro

Teniendo en cuenta toda la información mencionada anteriormente (haciendo hincapié en la gramática léxica de la sección 1.1), se escribió esta especificación léxica:

```

1 %option noyywrap /* Avisa que solo leera un archivo */
2 %{
3 #include <stdio.h>
4 #include <stdlib.h>
5 #include <string.h>
6 #include "sintactico.tab.h"
7 %} /* Finaliza inclusion de librerias*/
8 LETRA [a-zA-Z]
9 DIGITO [0-9]
10 OPERADOR_ADITIVO [+|-]
11 ASIGNACION :=
12 LETRA_DIGITO ({LETRA}|{DIGITO})
13 SEPARADOR ([ \t\n])+ /* Para ignorar espacios en blanco*/
14 void yyerror(char *);
15 %% /* Finaliza declaracion de tokens*/
16 inicio {return INICIO;}
17 fin {return FIN;}
18 leer {return LEER;}
19 escribir {return ESCRIBIR;}
20 "(" {yyval.cadena= strdup(yytext);return PI;}/* yyval=strdup... sirve para que se
    pueda realizar el proceso de sintesis*/
21 ")" {yyval.cadena= strdup(yytext);return PD;}/* yyval guarda el contenido de la
    deteccion para su futuro uso */
22 "," {return COMA;}
23 ";" {return PUNTOCOMA;}
24 {DIGITO}+ {yyval.cadena= strdup(yytext);return CONSTANTE;}
25 {LETRA}{LETRA_DIGITO}{0,31} {yyval.cadena= strdup(yytext);return IDENTIFICADOR;} /*
    Limite de 32 caracteres para los ID */
26 {OPERADOR_ADITIVO} {yyval.cadena= strdup(yytext);return OP_ADITIVO;}
27 {ASIGNACION} {return ASIGNACION;}
28 {SEPARADOR} {} /* Ignora espacio en blanco */
29 . {yyerror("Caracter Desconocido, ")} /* El punto detecta cualquier input. Debido a la
    naturaleza secuencial del analizador, se pone a lo ultimo*/
30 %% /* Finaliza reglas.*/
31 /* Parte en C implementada en especificacion sintactica */

```

lexico.l

Esta especificación no contiene mayores complicaciones, sino que intenta asemejarse a la especificación abstracta lo más posible. En particular notar que las palabras reservadas se escriben como reglas directamente. La tercera parte del archivo está vacía debido a que se implementa directamente en la especificación sintáctica (que posee estructura similar). Debido a esto, el archivo .l se desprende de toda comprobación ajena a la léxica: los roles semánticos y de síntesis estarán ligados, veremos más adelante, al analizador sintáctico.

Parte III

Analizador sintáctico

Un analizador sintáctico (o Parser) es un programa informático que analiza una cadena de símbolos de acuerdo a las reglas de una gramática formal. El parser procesa los tokens provistos por el analizador léxico para construir la estructura de datos.

Nota importante: en esta documentación elegimos separar las explicaciones del analizador sintáctico de los procesos semántico y de síntesis (si bien el código no realiza estas distinciones), para facilitar la comprensión y separación de la gramática sintáctica respecto de las funciones de C que les producen efecto.

1. Marco teórico

El análisis sintáctico es realizado por un módulo llamado Parser. Este analizador procesa los tokens que le entrega el Scanner hasta que reconoce una construcción sintáctica que requiere un procesamiento semántico.

Entonces, invoca directamente a la rutina semántica que corresponde. Algunas de estas rutinas semánticas utilizan, en sus procesamientos, la información de la representación de un token, como veremos más adelante. Nótese que no existe un módulo independiente llamado Analizador Semántico, sino que el análisis semántico se va realizando, en la medida que el Parser lo requiere, a través de las rutinas semánticas.

Las rutinas semánticas realizan el Análisis Semántico y también producen una salida en un lenguaje; esto último forma parte de la etapa de Síntesis del compilador.

2. Herramienta utilizada

Para realizar el analizador sintáctico (y luego veremos semántico y de síntesis) se usó la herramienta bison² que convierte la descripción formal de un lenguaje, escrita como una gramática independiente de contexto GIC, en un programa en C que realiza análisis sintáctico. Esta especificación se recibe en forma de archivo con extensión .y y construye como salida un archivo de extensión .tab.c, que al ser compilado en C constituye el analizador sintáctico correspondiente.

Para compilar un archivo de entrada usando bison, basta con escribir `bison -d «ruta entrada»` desde la línea de comando del sistema.

3. Estructura de la entrada

La estructura del archivo de extensión .y (similar a aquella del archivo .l) usado como entrada se divide en tres secciones separadas por el símbolo `%%`. Las secciones representan, en el orden correspondiente:

1. Inclusión de **macros** y **librerías** de C, declaración de variables.
Declaración de los tokens (especificados en léxico) en la forma `%token «nombre token»` separados por un renglón.
2. Definición de **reglas**: gramática formada por composición de tokens, acompañadas por porciones de código C asociadas a la detección de dichas estructuras. Son de la forma `«gramática {código_C}»` separadas por un renglón. Con fines didácticos, en esta parte del trabajo no se muestran dichas reglas.

²Descarga para Windows en <http://gnuwin32.sourceforge.net/packages/bison.htm>

3. Inclusión de **funciones** y **procedimientos** escritos en C, que proporcionan un menú al analizador sintáctico, y le proveen funcionalidades adicionales como detección de archivos por línea de comando o reporte de errores detallado. Más adelante se incluirán funciones de carácter semántico y de síntesis en esta sección.

4. Aplicación a Micro

Teniendo en cuenta la información de la Parte I (sobre todo la gramática sintáctica provista), se escribió este analizador sintáctico:

```

1  %{
2  /* Analizador Sintactico */
3  #include <stdio.h>
4  #include <stdlib.h>
5  #include <conio.h>
6  #include <string.h>
7  extern int yylex(void);
8  extern char *yytext;
9  extern FILE *yyin;
10 void yyerror(char *s);
11 void yyerrors(int tipo, char *s);
12 %} // Finaliza inclusion de librerias
13 %error-verbose // Reporte detallado de errores
14 %start Programa
15 %token INICIO
16 %token FIN
17 %token IDENTIFICADOR
18 %token LEER
19 %token ESCRIBIR
20 %token ASIGNACION
21 %token PUNTOCOMA
22 %token CONSTANTE
23 %token COMA
24 %token PI
25 %token PD
26 %token OP_ADITIVO
27 %% // Finaliza listado de tokens lexicos
28 // Inicio de gramatica sintactica
29 Programa: INICIO ListaDeSentencias FIN {YYACCEPT;}
30 ListaDeSentencias: Sentencia
31 | ListaDeSentencias Sentencia
32 Sentencia: IDENTIFICADOR ASIGNACION Expresion PUNTOCOMA
33 | LEER PI ListaIdentificadores PD PUNTOCOMA
34 | ESCRIBIR PI ListaExpresiones PD PUNTOCOMA
35 ListaIdentificadores: IDENTIFICADOR
36 | ListaIdentificadores COMA IDENTIFICADOR
37 ListaExpresiones: Expresion
38 | ListaExpresiones COMA Expresion
39 Expresion: Primaria
40 | Expresion OP_ADITIVO Expresion
41 Primaria: IDENTIFICADOR
42 | CONSTANTE
43 | PI Expresion PD
44 // Fin de gramatica sintactica
45 %%
46 // Menu del analizador sintactico creado
47 void yyerror(char *s)
48 {
49 printf("Error %s",s);

```

```
51 }
52 int main(int argc, char **argv)
53 {
54     if (argc>3 || argc<1)
55     {printf("Error, cantidad incorrecta de parametros");
56      getch();
57      return 1;}
58     else if (argc==2) // Permite pasaje de archivo por linea de comando
59     {yyin=fopen(argv[1], "rt");
60      printf("Archivo %s abierto\n", argv[1]);}
61     else
62     {printf("Por favor ingrese programa a ser analizado\n");
63      yyin=stdin;} // Analiza input del usuario
64     switch(yyparse()){
65     case 0:
66         printf("Compilado Correctamente\n Presione una tecla para salir..."); break;
67     case 1:
68         puts("\nCompilacion abortada"); break;
69     case 2:
70         puts("Memoria insuficiente"); break;
71     } // Manejo del comando yyparse segun documentacion de bison
72     getch();
73     return 0;
74 }
```

sintactico_noSemantico_noSintesis.y

Análogamente a la parte léxica, esta especificación sintáctica es muy parecida a la provista por la bibliografía y no plantea cambios radicales. La inclusión del comando `%error-verbose` da un interesante complemento al reporte de errores por parte del analizador sintáctico. El manejo de la función `yyparse()` surge de la documentación de bison, que incluiremos como bibliografía.

Parte IV

Analizador semántico y proceso de síntesis

El análisis semántico, del cual ya hemos hablado anteriormente, es la tercera fase de análisis que existe en un compilador. Por un lado, el análisis Semántico está inmerso dentro del análisis Sintáctico, y, por otro lado, es el comienzo de la etapa de Síntesis del compilador.

La etapa de síntesis constituye la traducción del archivo fuente ingresado correctamente en una especificación de bajo nivel para una máquina virtual. En este trabajo elegimos realizar una síntesis (traducción) a lenguaje C, dado que es muy fácil compilarlo posteriormente.

1. Consideraciones

Debido a que Micro es un lenguaje de carácter didáctico, no posee muchas restricciones semánticas: por ejemplo, las variables son declaradas implícitamente. Es por ello que nos fue posible generar, con una cantidad relativamente pequeña de funciones C, comprobaciones semánticas importantes para el funcionamiento del compilador.

2. Implementación en Micro

```

1  %{
2  /* Analizador Sintactico */
3  #include <stdio.h>
4  #include <stdlib.h>
5  #include <conio.h>
6  #include <string.h>
7  extern int yylex(void);
8  extern char *yytext;
9  extern FILE *yyin;
10 void yyerror(char *s);
11 typedef struct Nodo{
12     char cadena[33];
13     struct Nodo* siguiente;
14 } Nodo;
15 typedef struct Lista{
16     Nodo* cabeza;
17 } Lista;
18 Lista* inicializacion();
19 Nodo* crearNodo(char *a);
20 void insertarIdentificador(Lista *lista, char *a);
21 void yyerrors(int tipo, char *s);
22 void inicializarArchivo(FILE *f);
23 int estaEnLista(Lista *lista, char *a);
24 void agregarIdentificadorArchivo(FILE *f, char *s, char *a);
25 void agregarLeerArchivo(FILE *f, char *a);
26 void agregarEscribirArchivo(FILE *f, char *a);
27 void nuevoAgregarIdentificadorArchivo(FILE *f, char *s);
28 void finalizarArchivo(FILE *f);
29 Lista *lista;
30 FILE *f;
31 %} // Fin prototipo funciones
32 %union {
33     char * cadena;
34 } // Declara tipo de dato char como <cadena>

```

```

35 %error-verbose
%start Programa
37 %token INICIO
%token FIN
39 %token <cadena>IDENTIFICADOR
%token LEER
41 %token ESCRIBIR
%token ASIGNACION
43 %token PUNTOCOMA
%token <cadena>CONSTANTE
45 %token COMA
%token <cadena>PI
47 %token <cadena>PD
%token <cadena>OP_ADITIVO
49 %type <cadena>Expresion
%type <cadena>Primaria
51 %% // Fin declaracion tokens
Programa: INICIO ListaDeSentencias FIN {YYACCEPT;}
53 ListaDeSentencias: Sentencia
|ListaDeSentencias Sentencia
55 Sentencia: IDENTIFICADOR ASIGNACION Expresion PUNTOCOMA {if(estaEnLista(lista,$1)==0){
yyerrors(1,$1);YYABORT;}
else{insertarIdentificador(lista,$1);
agregarIdentificadorArchivo(f,$1,$3);}}
57 // Rutina de asignacion
|LEER PI ListaIdentificadores PD PUNTOCOMA
59 |ESCRIBIR PI ListaExpresiones PD PUNTOCOMA
ListaIdentificadores: IDENTIFICADOR {if(estaEnLista(lista,$1)==1){insertarIdentificador
(lista,$1);nuevoAgregarIdentificadorArchivo(f,$1);}
61 agregarLeerArchivo(f,$1);}
// Rutina de lectura
63 | ListaIdentificadores COMA IDENTIFICADOR {if(estaEnLista(lista,$3)==1){
insertarIdentificador(lista,$3);nuevoAgregarIdentificadorArchivo(f,$3);}
agregarLeerArchivo(f,$3);}
65 ListaExpresiones: Expresion {agregarEscribirArchivo(f,$1);}
|ListaExpresiones COMA Expresion {agregarEscribirArchivo(f,$3);}
67 // Rutina de escritura
Expresion: Primaria
69 |Expresion OP_ADITIVO Expresion {$$=strdup(strcat(strcat($1,$2),$3))}
Primaria: IDENTIFICADOR {if(estaEnLista(lista,$1)==1){yyerrors(0,$1);YYABORT;}}
71 | CONSTANTE
| PI Expresion PD {$$=strdup(strcat(strcat($1,$2),$3))}
73 %%
void yyerror(char *s)
75 {
printf("Error %s",s);
77 }
// Inicio funciones lista
79 Lista* inicializacion(){
Lista *lista=(Lista *)malloc(sizeof(Lista));
81 lista->cabeza = NULL;
return lista;
83 }
Nodo* crearNodo(char *a){
85 Nodo* nodo = (Nodo *)malloc(sizeof(Nodo));
strncpy(nodo->cadena,a,sizeof(nodo->cadena));
87 nodo->siguiente=NULL;
return nodo;
89 }
// Fin funciones lista
// Inicio funciones semanticas
91 void insertarIdentificador(Lista *lista,char *a){

```

```

93     Nodo* nodo=crearNodo(a);
94     if(lista->cabeza== NULL){
95         lista->cabeza=nodo;
96     }
97     else{
98         Nodo* aux=lista->cabeza;
99         while(aux->siguiente){
100             aux = aux->siguiente;
101         }
102         aux->siguiente=nodo;
103     }
104 }
105 int estaEnLista(Lista *lista, char *a){
106     Nodo *aux=lista->cabeza;
107     if(aux==NULL){
108         return 1;
109     }
110     else{
111         while(aux!=NULL){
112             if(strcmp(aux->cadena,a)==0)
113                 return 0;
114             else
115                 aux=aux->siguiente;
116         }
117         return 1;
118     }
119 }
120 // Fin funciones semanticas
121 int main(int argc, char **argv)
122 {
123     int ret;
124     lista=inicializacion();
125     inicializarArchivo(f);
126     if (argc>3 || argc<1)
127     {printf("Error, cantidad incorrecta de parametros");
128     getch();
129     return 1;}
130     else if (argc==2)
131     {yyin=fopen(argv[1], "rt");
132     printf("Archivo %s abierto\n", argv[1]);}
133     else
134     {printf("Por favor ingrese programa a ser compilado\n");
135     yyin=stdin;}
136     switch(yparse()){
137     case 0:
138         printf("Compilado Correctamente\n Presione una tecla para salir...");finalizarArchivo(
139         f); break;
140         // Rutina Fin
141     case 1:
142         ret = remove("salidaEnC.c"); // Borrar archivo salida ante error
143         puts("\nCompilacion abortada"); break;
144     case 2:
145         puts("Memoria insuficiente"); break;
146     }
147     getch();
148     return 0;
149 }
150 void yyerrors(int tipo, char *s){
151     switch(tipo){
152     case 0:
153         printf("Identificador %s no declarado ",s);break;
154     case 1:

```

```

155     printf("Identificador %s repetido",s);break;
156 }
157 // Inicio funciones de sintesis
void inicializarArchivo(FILE *f){
159     f=fopen("salidaEnC.c","w");
    fprintf(f,"#include <stdio.h>\n");
161     fclose(f);
    f=fopen("salidaEnC.c","a+");
163     fprintf(f,"#include <stdlib.h>\n");
    fprintf(f,"int main(){\n");
165     fclose(f);
}
167 void agregarIdentificadorArchivo(FILE *f,char *s,char *a){
    f=fopen("salidaEnC.c","a+");
169     fprintf(f,"int %s = %s;\n",s,a);
    fclose(f);
171 }
void nuevoAgregarIdentificadorArchivo(FILE *f,char *s){
173     f=fopen("salidaEnC.c","a+");
    fprintf(f,"int %s;\n",s);
175     fclose(f);
}
177 void agregarLeerArchivo(FILE *f,char *a){
    f=fopen("salidaEnC.c","a+");
179     fprintf(f,"scanf(\"%d\",&s);\n",a);
    fclose(f);
181 }
void agregarEscribirArchivo(FILE *f,char *a){
183     f=fopen("salidaEnC.c","a+");
    fprintf(f,"printf(\"%d\\n\",%s);\n",a);
185     fclose(f);
}
187 void finalizarArchivo(FILE *f){
    f=fopen("salidaEnC.c","a+");
189     fprintf(f,"\\nreturn 0;\n");
    fclose(f);
191 }
// Fin funciones de sintesis

```

sintactico.y

Reglas Agregamos piezas de código C para realizar distintas acciones. El bison usa una notación posicional de izquierda a derecha para determinar el token sobre el cual se opera: \$\$ quiere decir “resultado”, y \$N indica según el número $N = 1, 2, 3 \dots$ la posición del token deseado. Por ejemplo, en la línea 55 se llama al token \$1 para comprobar si se encuentra en la lista, entendiendo que \$1 en aquel caso es un identificador.

Listas La colección de identificadores resulta ser una lista implementada en C: para ello, desde la línea 79 hasta la 89 se pueden ver las funciones de creación de lista.

Semántica Elegimos armar listas con funciones de C para poder almacenar identificadores y poder compararlos entre ellos; así, por ejemplo, es posible determinar si una rutina escribir(...) intenta escribir un identificador no inicializado anteriormente. Otra comprobación interesante es evitar la doble asignación := de una variable, o chequear si una rutina leer(...) declara de forma implícita o si recibe una variable inicializada anteriormente. Desde la línea 92 hasta la 119 se encuentran las funciones que realizan comprobaciones semánticas con ayuda de las reglas.

Síntesis Las funciones de síntesis se encuentran desde la línea 158 hasta la última línea del programa. Las mismas se encargan de administrar un archivo de texto de salida, al cual se le va agregando la información pasada por parámetros de funciones. El llamado a estas funciones se realiza dentro de cada regla (por ejemplo, en la línea 65 se llama a la función `agregarEscribirArchivo(f, ...)`) y de forma secuencial, es decir, luego de cada token se ejecuta la síntesis correspondiente. Una ventaja de esta implementación es que la síntesis desconoce los pormenores del análisis sintáctico y semántico, ya que la misma es llamada dentro de estructuras de control de las reglas.

3. Conclusión

La construcción de este compilador fue un proceso gradual, en el cual nos fue muy práctico desarrollar en primer lugar el analizador sintáctico. Debido a la naturaleza de Micro, este análisis impone restricciones fuertes para los futuros errores semánticos. El análisis semántico fue sencillo de aplicar, pero resultó necesario crear estructuras más complejas (implementación en listas). Por último, la síntesis fue un proceso independiente del análisis y un agregado de interés para crear una salida funcional en C.