PROCESADORES DEL LENGUAJE

Práctica Final: Frontend y Backend



Grupo: 113-leal-zhu ID de grupo de prácticas: 81

Nombre de los participantes:

Liang Ji Zhu Ignacio Leal Sánchez

Correos electrónicos de los participantes:

100495723@alumnos.uc3m.es 100495680@alumnos.uc3m.es

Madrid, 9 de mayo de 2025

Índice

1.	Introducción	2
2.	Transparencia	3
3.	Traducción Frontend	3
	3.1. Variables Globales	3
	3.2. Función MAIN, Funciones Genéricas y Variables Locales	4
	3.3. Impresión de Cadenas	5
	3.4. Operadores, Precedencia, Asociatividad y Vectores	6
	3.5. Estructuras de Control	8
4.	Traducción Backend	9
	4.1. Variables Globales	9
	4.2. Función MAIN y Funciones Genéricas	10
	4.3. Impresión de Cadenas	10
	4.4. Operadores, Precedencia y Asociatividad	11
	4.5. Estructuras de Control	12
5.	Testing	12
	5.1. Variables Globales	12
	5.2. Función MAIN	13
	5.3. PUTS	14
	5.4. PRINT Simple	14
	5.5. Operadores	15
	5.6. WHILE	16
	5.7. IF-ELSE	17
	5.8. FOR	18
	5.9. Funciones (Lisp)	19
	5.10. Variables Locales (Lisp)	20
	5.11. Vectores (Lisp)	21
6	Conclusiones	23

1. Introducción

La presente memoria documenta el desarrollo integral de un traductor frontend y un traductor backend. El trabajo plantea la construcción de éstas en dos fases:

- Frontend: traduce el lenguaje C, de notación infija, a código Lisp, de notación prefija.
- Backend: convierte dicho código Lisp a notación postfija, Forth, siguiendo la filosofía de máquina de pila integrada en gforth.

Se sifuió un ciclo incremental para la realización del proyecto, en cada iteración se añadió una nueva producción. Esta estrategia permitió aislar errores tempranos y dosificar la complejidad, ya que los apartados iniciales eran más asequibles y los últimos exigían mayor autonomía y un diseño gramatical más depurado.

Se puede apreciar 3 etapas de implementación aproximadas:

- Etapa 1: se implementaron las declaraciones básicas, la función main y las primeras sentencias de E/S (puts, printf).
- Etapa 2: se integraron operadores con su precedencia y las estructuras de control while e if/else.
- Etapa 3: se añadieron funciones con retorno estructurado y vectores.
- Etapa 4: desarrollo del backend y pruebas funcionales con la directiva
 //② (main), completando el flujo C → Lisp → Forth.

El analizador léxico se mantuvo sin modificaciones; toda la lógica nueva se centró en la gramática y las acciones semánticas, añadiendon los tokens necesarios.

Cada incremento semanal concluyó con una regresión completa sobre el conjunto de casos de prueba —tanto los oficiales como los diseñados ad-hoc—y con la documentación de las decisiones en esta memoria. Así, el resultado final es un traductor robusto cuya construcción refleja paso a paso la ruta marcada por las pautas del enunciado, desde las bases del frontend hasta los detalles finales del backend.

2. Transparencia

En este apartado constatamos que el uso de la IA generativa ha sido utilizado para la corrección de errores en las declaraciones de C, aunque ha sido intercalado con la asistencia de los profesores que proponen respuestas más coherentes y simples. Y ha sido imposible usar la inteligencia artificial en más partes del trabajo debido a su incapacidad para comprender los requisitos específicos del problema, y por ello dando respuestas de nula utilidad.

Asimismo, hacemos constar que la carga de trabajo se ha repartido de forma justa entre los dos. A lo largo de todo el período —y, muy especialmente, durante la fase final— la dedicación fue dinámica: cada miembro asumió más tareas cuando su disponibilidad lo permitía y cedió responsabilidad cuando sus circunstancias lo exigían, garantizando en todo momento un equilibrio global.

3. Traducción Frontend

3.1. Variables Globales

Las variables se declaran como enteros o como arrays en los que se puede dar valor o no, en este caso se usará el valor 0 para la traducción a lisp. Se permite que se declaren múltiples variables en línea separadas con ",.º en cascada.

```
axioma:
                        var_global funcion
                        r_axioma
  r_axioma:
                        axioma
  var_global:
                        declaracion ';'
                                          var_global
12 declaracion:
                        INTEGER
                                  IDENTIF valor_global
     r_declaracion
                                    IDENTIF '[' NUMBER ']'
                        | INTEGER
13
                           r_declaracion
14
16 valor_global:
```

3.2. Función MAIN, Funciones Genéricas y Variables Locales

Se definen las funciones de dos maneras: las funciones genéricas y el main. El main es una función diferente ya que es obligatorio que exista esta función. Dentro del cuerpo de las funciones pueden declararse variables, que tanto en C como en Lisp serán variables locales. Se emplea la misma traducción de las globales, pero en este caso deben generarse dentro del cuerpo de la función. Por lo que también permitimos la declaración de variables locales que serán la variable seguida del nombre de la función en la que es variable local.

```
'(' argumento ')' '{'
                       IDENTIF
  funcion:
     var_local cuerpo '}' funcion
                       | funcion_principal
                       MAIN
                            '(' argumento ')' '{' var_local
5 funcion_principal:
      cuerpo '}'
                       INTEGER valor resto_argumento
  argumento:
                       | valor resto_argumento
10
11
                       STRING
13 valor:
                         expresion
14
15
resto_argumento:
                       ',' argumento
20
```

```
var_local:
                       declaracion_local ';' var_local
25
26 declaracion_local:
                       INTEGER
                                 IDENTIF valor_local
     r_decl_local
                        | INTEGER
                                   IDENTIF '[' NUMBER ']'
27
                           r_decl_local
28
  valor_local:
                        /* lambda */
                         '=' NUMBER
31
                        ',' IDENTIF valor_local r_decl_local
33 r_decl_local:
                        ',' IDENTIF '[' NUMBER ']'
                           r_decl_local
35
36
```

3.3. Impresión de Cadenas

La impresión de cadenas de texto se hace imprimiendo o un string o una variable. No existe la forma de impresión formateada como en C. Por ello cualquier impresión de variables en C se traducirá a la impresión de variables individuales en Lisp.

```
sentencia: asignacion
| '0' expresion
| PUTS '(' STRING ')'
| PRINTF printf
| llamada
| printf: '(' STRING r_printf ')'
| r_printf: ',' expresion r_printf
| ',' STRING r_printf
| ',' STRING r_printf
| ',' STRING r_printf
| ',' STRING r_printf
```

3.4. Operadores, Precedencia, Asociatividad y Vectores

Aquí hemos cambiado la forma de operar respecto a la versión anterior. Para garantizar la mejor precedencia sin ninguna posibilidad de errores hemos dividido las diferentes operaciones o comparaciones en No Terminales específicos de forma que la precedencia pueda ser cumplida en todo momento. Para declarar un Vector podemos usar una sintaxis de Lisp extendiendo la que usamos para variables y podrán declararse como variables globales o locales.

```
var_global:
                       declaracion ';'
                                         var_global
 declaracion:
                       INTEGER
                                 IDENTIF valor_global
     r_declaracion
                       | INTEGER
                                   IDENTIF '[' NUMBER ']'
                           r_declaracion
                       // Vector Global
  /* ==== Operadores, precedencia y asociatividad ==== */
                       IDENTIF '=' expresion
11 asignacion:
                       | vector '=' expresion
12
                       // Vector Local
15
16 expresion:
                       logical_or
17
19 llamada:
                       IDENTIF '(' argumento ')'
20
22 logical_or:
                       logical_and
                       | logical_or OR logical_and
                       igualdad
25 logical_and:
                        | logical_and AND igualdad
26
28 igualdad:
                       relacional
                        | igualdad EQ relacional
                        | igualdad NE relacional
32 relacional:
                       aditivo
```

```
relacional '<' aditivo
33
                          relacional '>' aditivo
34
                          relacional LE aditivo
                          relacional GE aditivo
37
38 aditivo:
                        multiplicativo
                          aditivo '+' multiplicativo
39
                          aditivo '-' multiplicativo
40
41
42 multiplicativo:
                        unario
                          multiplicativo '*' unario
                          multiplicativo '/' unario
44
                          multiplicativo '%' unario
45
46
  unario:
                        operando
                         '!' unario
48
                          '+' operando %prec UNARY_SIGN
49
                          '-' operando %prec UNARY_SIGN
50
51
  operando:
                        IDENTIF
                          IDENTIF '(' argumento ')'
                          NUMBER
55
                          '(' logical_or ')'
56
                           vector
57
58
                        IDENTIF '[' logical_or ']'
  vector:
60
```

Se declararon *tokens* para los operadores lógicos y de comparación con el objetivo de mejorar la legibilidad del código y establecer un acoplamiento más claro entre las fases de análisis léxico y análisis sintáctico.

En particular, los operadores de comparación fueron definidos como nonassoc, ya que, a diferencia de operadores como la suma o la multiplicación, las comparaciones (por ejemplo, ==, !=) no son asociativas. Esto significa que una expresión como:

$$a == b == c$$

no tiene un significado bien definido en C. Existen dos posibles formas de interpretarla:

$$(a == b) == c || a == (b == c)$$

Ambas interpretaciones pueden generar ambigüedad semántica. Al declarar estos operadores como nonassoc, el *parser* rechaza expresiones como a == b == c y obliga al programador a escribirlas de manera explícita, por ejemplo:

- (a == b) && (b == c)
- Uso de paréntesis para aclarar el orden de evaluación

Este enfoque permite capturar errores de lógica o errores tipográficos en tiempo de compilación, evitando comportamientos ambiguos o inesperados en tiempo de ejecución.

3.5. Estructuras de Control

Las estructuras de control se encuentran dentro del cuerpo de una función y no acaban en ";". Estas son el IF-ELSE, FOR y WHILE. Cuentan con una estructura de cuerpo diferente ya que un return de alguna de estas estructuras se interpreta diferente ya que no es un return al final de la función.

```
cuerpo:

sentencia ';' cuerpo
| sentencia ';'

estructura cuerpo
| estructura
| RETURN expresion ';'

estructura:

cuerpo_estructura '}'
```

```
| IF '(' expresion ')' '{'
                           cuerpo_estructura '}'
                        | IF '(' expresion ')' '{'
                           cuerpo_estructura '}' ELSE '{'
                           cuerpo_estructura '}'
                        | FOR '(' declaracion_for ';'
10
                           expresion ';' asignacion ')' '{'
                           cuerpo_estructura '}'
11
                       INTEGER
13 declaracion_for:
                                 IDENTIF valor_for
     r_declaracion_for
                               IDENTIF valor_for
14
                           r_declaracion_for
  valor_for:
                          '=' NUMBER
17
18
                            ',' IDENTIF valor_for
  r_declaracion_for:
     r_declaracion_for
20
  cuerpo_estructura:
                       sentencia ';'
                          estructura
23
                          sentencia ';' cuerpo_estructura
24
                          estructura cuerpo_estructura
25
                          RETURN expresion ';'
26
```

4. Traducción Backend

Una vez acabado con la implementación del frontend que traducía de C a Lisp, vamos a realizar el correspondiente backend, de Lisp a Forth. Para ello, hemos decidido dividir la implementación en las siguientes fases:

4.1. Variables Globales

Tal y como traducimos el lisp todas las variables que pueden haber en este código serán globales. Simplemente cambiará el nombre si son específicas a una función en específico. Como ya se unificaron las formas de declarar variables esta parte de la gramática es relativamente simple.

```
var_global: declaracion
var_global declaracion

declaracion:

'(' SETQ IDENTIF logical_or ')'

;
```

4.2. Función MAIN y Funciones Genéricas

La definición de funciones se puede unificar para que haya un no terminal que acoge ambos dos tipos de posibles funciones el main que sigue siendo obligatorio y las otras que siguen siendo opcionales. Esta es una forma más simple que la utilizada en el Frontend y por la traducción de Lisp no se nos requerirá el uso de argumentos en estas funciones.

4.3. Impresión de Cadenas

Esta función ya fue muy alterada para el cambio de C a Lisp y aquí la traducción es bastante simple ya que el programa puede necesitar imprimir strings o variables/resultados de operaciones.

```
sentencia: '(' PRINT STRING ')'
| '(' PRINC logical_or ')'
| '(' PRINC STRING ')'
| '(' PRINC STRING ')'
| ';
```

4.4. Operadores, Precedencia y Asociatividad

Estos son todos los operadores y comparadores que permiten que se sigan las reglas de asociatividad que creamos en la primera gramática. Esta parte de la gramática es muy similar a la del Frontend ya que el cambio de lenguaje no realiza ningún cambio estructural muy grande. Lo que sí que tuvimos en cuenta fue en cuanto a las operaciones en Lisp, que al llevar paréntesis hicimos que las operaciones tuvieran el mismo nivel de precedencia.

```
logical_or:
                       logical_and
                       / '(' OR logical_or logical_and ')'
  logical_and:
                       igualdad
                         '(' AND logical_and igualdad ')'
  igualdad:
                       relacional
                       / '(' '=' igualdad relacional ')'
                         '(' NE igualdad relacional ')'
10
  relacional:
                       operacion
                       / '(' '<' relacional operacion ')'
                       / '(' '>' relacional operacion ')'
                         '(' LE relacional operacion ')'
                         '(' GE relacional operacion ')'
15
16
  operacion:
                       unario
17
                         '(' '+' operacion operacion ')'
18
                         '(' '-' operacion operacion ')'
                         '(' '*' operacion operacion ')'
                         '(' '/' operacion operacion ')'
21
                         '(' MOD operacion operacion ')'
22
23
24 unario:
                       operando
                       | '(' NOT unario ')'
                         '+' operando %prec UNARY_SIGN
26
                         '(' '-' operando %prec UNARY_SIGN
27
                       IDENTIF
  operando:
                         NUMBER
                         '(' logical_or ')'
31
32
```

4.5. Estructuras de Control

Hemos logrado que en esta gramática las sentencias y estructuras de control estén unificadas gracias a la estructura de lisp y a la forma tan estandarizada de definir el código.

```
sentencia: '(' PRINT STRING ')'

| '(' PRINC logical_or ')'
| '(' PRINC STRING ')'
| '(' SETF IDENTIF logical_or ')'
| '(' SETQ IDENTIF logical_or ')'
| '(' LOOP WHILE logical_or DO lista_sentencia ')'
| '(' IF logical_or sentencia ')'
| '(' IF logical_or sentencia ')'
| '(' PROGN lista_sentencia ')'
| '(' PROGN lista_sentencia ')'
| '(' PROGN lista_sentencia ')'
```

5. Testing

En esta sección se presentan las pruebas realizadas para verificar el correcto funcionamiento del traductor en sus distintas etapas. Cada prueba incluye el código fuente en C, su correspondiente traducción a Lisp y finalmente la traducción a Forth.

5.1. Variables Globales

```
#include <stdio.h>
int a;
int b = 10, c, d = 5;
main() {
    puts("Variables globales");
}

//@ (main)

Variables globales

(setq a 0)
(setq b 10)
(setq c 0)
(setq d 5)
```

```
15 (defun main ()
           (print "Variables globales")
17 )
18 (main)
20 "Variables globales"
22 variable a
23 0 a !
24 variable b
25 10 b !
26 variable c
27 0 c !
28 variable d
29 5 d!
30 : main ." Variables globales";
31 main
33 Variables globales
```

5.2. Función MAIN

```
| #include <stdio.h>
2 int x;
4 main() {
   x = x + 1;
    printf("%d\n", x);
7 }
8 //@ (main)
12 (setq x 0)
(defun main ()
        (setf x (+ x 1))
        (princ x)
16
17 )
18 (main)
19
20 1
```

5.3. PUTS

```
| #include <stdio.h>
2 int main_var;
3 main() {
     puts("Hola, mundo!");
6 //@ (main)
8 Hola, mundo!
10 (setq main_var 0)
11 (defun main ()
         (print "Hola, mundo!")
13 )
14 (main)
16 "Hola, mundo!"
17
18 variable main_var
19 0 main_var !
20 : main ." Hola, mundo!" ;
23 Hola, mundo!
```

5.4. PRINT Simple

```
#include <stdio.h>
int val = 42;
main() {
    printf("%d", val);
```

5.5. Operadores

```
| #include <stdio.h>
2 int a = 1;
_{3} int b = 0;
_{4} int c = 2;
5 main() {
      // mezcla &&, ||, ==, !=, <, >, <=, >=, %, +, -, *, /
      if ((a + c * 3) % 2 == 1 && b != 0 || a <= c) {
           puts("OK");
      }
9
10 }
11 //@ (main)
13 OK
15 (setq a 1)
16 (setq b 0)
17 (setq c 2)
18 (defun main ()
```

```
(if (or (and (= (mod (+ a (* c 3)) 2) 1) (/= b 0)
              ) (<= a c))
          (progn (print "OK")))
21 )
22 (main)
24 "OK"
26 variable a
27 1 a !
28 variable b
29 0 b !
30 variable c
31 2 c !
|32|: main a 0 c 0 3 * + 2 mod 1 = b 0 0 = 0= and a 0 c 0 <=
    or if
         ." OK"
34 then;
35 main
37 OK
```

5.6. WHILE

```
| #include <stdio.h>
2 int cnt = 0;
3 main() {
     while (cnt < 3) {
         printf("%d\n", cnt);
         cnt = cnt + 1;
     }
8 }
9 //@ (main)
12 1
13 2
_{15} (setq cnt 0)
16 (defun main ()
(loop while (< cnt 3) do
     (progn (princ cnt)
18
```

5.7. IF-ELSE

```
| #include <stdio.h>
_{2} int v = 5;
3 int res;
4 main() {
     if (v \% 2 == 0) {
        puts("Par");
     } else {
        puts("Impar");
     }
     if (v > 10) {
        puts("Grande");
12
13 }
14 //@ (main)
15
16 Impar
18 (setq v 5)
(setq res 0)
20 (defun main ()
      (if (= (mod v 2) 0)
21
```

```
(progn (print "Par"))
           (progn (print "Impar")))
           (if (> v 10)
           (progn (print "Grande")))
26 )
27 (main)
29 "Impar"
31 variable v
32 5 v !
33 variable res
34 0 res !
| 35 | : main v @ 2 mod 0 = if
          ." Par"
37 else
          ." Impar"
38
39 then
_{40}|v @ 10 > if
          ." Grande"
42 then;
43 main
45 Impar
```

5.8. FOR

```
16 (setq i 0)
(setq suma 0)
18 (setq n 5)
19 (defun main ()
           (setq i 0)
           (loop while (< i n) do
           (progn (setf suma (+ suma i)))
           (setf i (+ i 1)))
           (princ suma)
25
26)
  (main)
31 variable i
32 0 i !
33 variable suma
34 0 suma !
35 variable n
36 5 n !
37 : main 0 i !
38 begin
           i @ n @ <
39
40 while
           suma @ i @ + suma !
42 i @ 1 + i !
43 repeat
_{44} suma 0 . ;
47 10
```

5.9. Funciones (Lisp)

```
#include <stdio.h>
square (int v) {
    return v * v;
}
sum (int x, int y) { return x + y; }
main() {
    int r1, r2;
```

```
r1 = square(7);
      r2 = sum(3, 4);
      printf("%d %d\n", r1, r2);
      // llamada como procedimiento (valor ignorado)
      square(5);
13 }
14 //@ (main)
16 49 7
18 (defun square (v )
          (* v v)
20 )
21 (defun sum (x y)
          (+ x y)
23 )
25 (defun main ()
          (setq r1 0)
          (setq r2 0)
          (setf r1 (square 7 ))
          (setf r2 (sum 3 4 ))
          (princ r1)
          (princ r2)
          (square 5)
  (main)
```

5.10. Variables Locales (Lisp)

```
#include <stdio.h>
main() {
    int g = 1;
    int a = 4, b;
    a = a + 1;
    b = a * 2;
    printf("%d %d\n", a, b);
}
//@ (main)
```

5.11. Vectores (Lisp)

```
| #include <stdio.h>
2 int arr [5];
3 main() {
      int i;
      for (i = 0; i < 5; i = i + 1) {
          arr[i] = i * 2;
      printf("%d %d %d %d %d\n",
             arr[0], arr[1], arr[2], arr[3], arr[4]);
10 }
11 //@ (main)
13 0 2 4 6 8
15 (setq arr (make-array 5))
  (defun main ()
          (setq i 0)
          (setq i 0)
          (loop while (< i 5) do
19
          (progn (setf (aref arr i) (* i 2)))
          (setf i (+ i 1)))
          (princ (aref arr 0))
          (princ (aref arr 1))
          (princ (aref arr 2))
```

```
(princ (aref arr 3))
(princ (aref arr 4))
(princ (aref arr 4))
(main)
(princ (aref arr 4))
(main)
```

6. Conclusiones

La presente memoria ha descrito en detalle el diseño, la implementación y la validación de un traductor de dos fases para un subconjunto de C: un frontend que convierte C en Lisp y un backend que convierte Lisp en Forth. A lo largo del proyecto se han abordado todos los objetivos formativos propuestos en la asignatura: la definición de gramáticas jerárquicas, la gestión de la semántica mediante código diferido, la generación de código intermedio y final, y la validación con conjuntos de pruebas diseñados ad hoc y suministrados por el profesorado.

El ciclo incremental de desarrollo permitió aislar y corregir errores de forma temprana en cada iteración. Gracias a esta estrategia, pudimos empezar por las declaraciones básicas y la impresión de cadenas, añadir luego operadores y estructuras de control, y finalmente incorporar funciones con retorno y vectores antes de abordar el backend. La generación de los ficheros trad.y y back.y, así como el conjunto de scripts de prueba automatizados, garantizó la robustez del traductor y facilitó la comparación de los comportamientos nativos en C, Lisp y Forth.

Como principales retos formativos cabe destacar la resolución de conflictos de gramática (shift/reduce y reduce/reduce), la correcta gestión de ámbitos en variables locales versus globales, y la inserción de instrucciones de salida (bye en GForth) para evitar bloqueos en los scripts automatizados. Estos desafíos han reforzado el entendimiento de la teoría de compiladores y la importancia de un buen diseño gramatical y semántico.

Para trabajos futuros, sería interesante ampliar el subconjunto de C soportado (por ejemplo, punteros, estructuras y llamadas con múltiples argumentos) o una implementación propio de un analizador léxico en *flex*.

En definitiva, este proyecto no solo cumple los requisitos de la práctica, sino que sienta las bases para desarrollos posteriores en el ámbito de los procesadores de lenguaje y la generación de compiladores de propósito educativo o industrial.