

Universidad Rafael Landívar
Facultad de Ingeniería
Informática y Sistemas
Lenguajes Formales y Autómatas
Docente: Ing. Damaris Campos

Proyecto No.02 Fase 1 “Lenguajes Formales y Autómatas”

Palencia Aldana Kevin Rodrigo-1045324
De La Roca Martínez José Andres-1142124
Reyes Muñoz Josué David-1029024

Guatemala, 22 de octubre del 2025

Índice

I. Tokens del lenguaje y expresiones regulares	3
Palabras clave	3
Estructura “tipo XML”	3
Números enteros y reales con signo opcional	3
Error léxico (captura todo lo no reconocido).....	4
II. Método del árbol Aho–Sethi–Ullman aplicado al token numero	4
Construcción del DFA	6
Tabla de transiciones DFA final, simplificada.....	7

I. Tokens del lenguaje y expresiones regulares

Palabras clave

OPERACION_NOMBRE

(?i)\b(?:SUMA|RESTA|MULTIPLICACION|DIVISION|POTENCIA|RAIZ|INVERSO|MOD)\b

Estructura “tipo XML”

- TAG_OP_ABRE → <\s*Operacion\s*=\s*(?i)
- (?:SUMA|RESTA|MULTIPLICACION|DIVISION|POTENCIA|RAIZ|INVERSO|MOD)\s*>
- TAG_OP_CIERRA → </\s*Operacion\s*>
- TAG_NUM_ABRE → <\s*Numero\s*>
- TAG_NUM_CIERRA → </\s*Numero\s*>

Números enteros y reales con signo opcional

• NUMERO → [+ -]?(?:\d+(?:\.\d+)?)|\.\d+

Acepta: 45, -3, 7.8, +0.25, .7

1. Espacios en blanco se ignoran por el scanner

• WS → [\t\r\n]+

2. Símbolos sueltos

IGUAL → =

MENOR → <

MAYOR → >

SLASH → /

Error léxico (captura todo lo no reconocido)

DESCONOCIDO →.

II. Método del árbol Aho–Sethi–Ullman aplicado al token numero

alfabeto de clases:

- S = signo + ó - • D
- = dígito [0-9]
- P = punto.

La **ER canónica** del número es:

$$R = (S? \cdot (D^+ \cdot (P \cdot D^+)? | P \cdot D^+)) \#$$

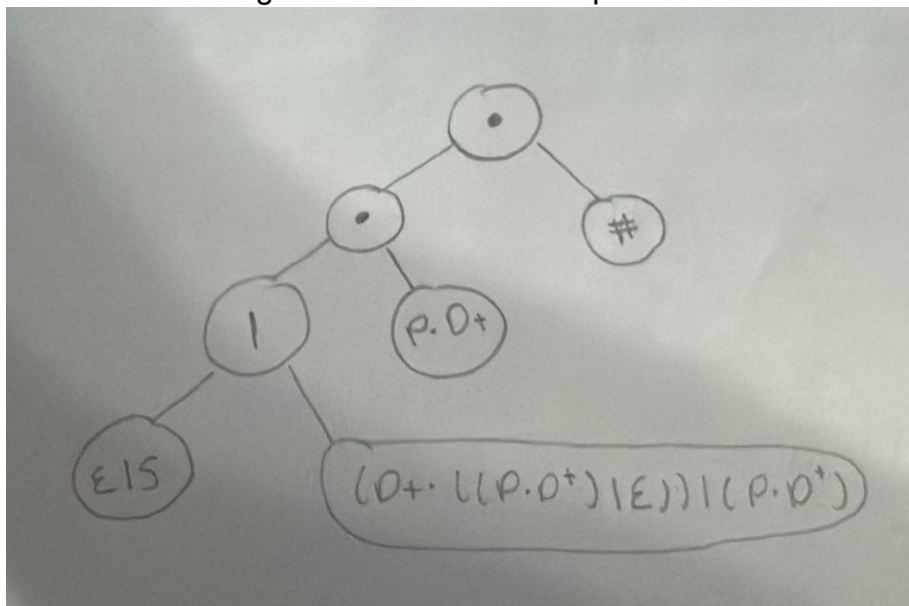
Equivalencias usadas:

- $X? = (\epsilon | X)$
- $D^+ = D \cdot D^*$

1. Árbol de la expresión:

Figura No.1

Diagrama de árbol de la expresión



Fuente: Elaboración propia (2025)

Las hojas terminales con posición son (en este orden de aparición izquierda a derecha:

- 1: S
- 2: D (el primero del D+ de la rama A)
- 3: D (el de D* en esa misma rama)
- 4: P (del P·D+ de la rama B)
- 5: D (primer D del D+ de la rama B)
- 6: D (el de D* de la rama B)
- 7: P (del P·D+ opcional después de D+ en la rama A)
- 8: D (primer D del D+ después de ese P)
- 9: D (el de D* de ese D+)
- 10: # (marcador de fin)

2. nullable, firstpos, lastpos

$\text{nullable}(\epsilon) = \text{true}$, $\text{nullable}(a) = \text{false}$

$\text{nullable}(X \cdot Y) = \text{nullable}(X) \wedge \text{nullable}(Y)$

$\text{nullable}(X|Y) = \text{nullable}(X) \vee \text{nullable}(Y)$ $\text{nullable}(X^*) = \text{true}$

$\text{firstpos}(X \cdot Y) = \text{firstpos}(X) \cup (\text{nullable}(X) ? \text{firstpos}(Y) : \emptyset)$

$\text{lastpos}(X \cdot Y) = \text{lastpos}(Y) \cup (\text{nullable}(Y) ? \text{lastpos}(X) : \emptyset)$

first/last de | es la unión; de * es el del hijo.

3. Tabla de followpos

Tabla No.01
Tabla de followpos

Posición	Símbolo	Followpos
1	S	{2,4}
2	D	{3,7,4,10}
3	D	{3,7,10}
4	P	{5}
5	D	{6,10}
6	D	{6,10}
7	P	{8}
8	D	{9,10}
9	D	{9,10}

10	#	\emptyset
----	---	-------------

Fuente: Elaboración propia (2025)

Construcción del DFA

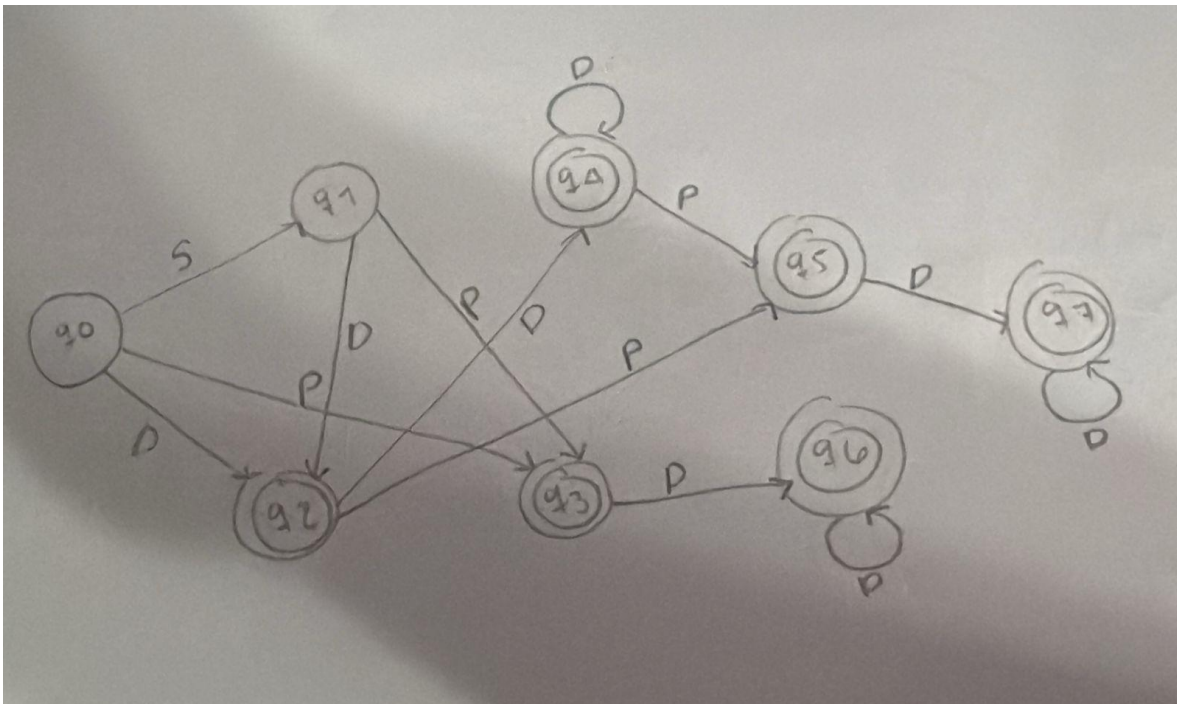
Alfabeto del autómata: {S, D, P}

Estado inicial: $I = \text{firstpos}(R) = \{1,2,4\}$

Estado de aceptación: todo estado que contenga la posición 10 (el #).

Figura No.02

DFA



Fuente: Elaboración propia (2025)

Tabla de transiciones DFA final, simplificada

Tabla No.02

Tabla de transiciones del DFA derivada del método del árbol por conjuntos de posiciones

Estado (conjunto de posiciones)	S	D	P	¿Acepta?
q0 = {1,2,4}	{2,4}	{3,7,4,10}	{5}	-
q1 = {2,4}	-	{3,7,4,10}	{5}	-
q2 = {3,7,4,10}	-	{3,7,10}	{5,8}	SI
q3 = {5}	-	{6,10}	-	SI
q4 = {3,7,10}	-	{3,7,10}	{8}	SI
q5 = {5,8}	-	{6,9,10}	-	SI
q6 = {6,10}	-	{6,10}	-	SI
q7 = {9,10}	-	{9,10}	-	SI

Fuente: Elaboración propia (2025)

Tabla No.03

Tabla final de transiciones del DFA simplificado

Estado	S	D	P	¿Acepta?
q0	q1	q2	q3	-
q1	rechazo	q2	q3	-
q2	rechazo	q4	q5	SI
q3	rechazo	q6	rechazo	SI
q4	rechazo	q4	q5	SI
q5	rechazo	q7	rechazo	SI
q6	rechazo	q6	rechazo	SI
q7	rechazo	q7	rechazo	SI

Fuente: Elaboración propia (2025)