

Sintaxis y Semántica de los lenguajes Trabajo práctico Anual.

GRUPO 6

Índice

1 Creando un lenguaje básico	2
1 Creando un lenguaje básico	3
2 Gramática léxica: Categorías léxicas	3
2 Gramática léxica: Categorías léxicas	3
3 Gramática léxica: Autómata finito y expresiones regulares	4
3 Gramática léxica: Autómata finito y expresiones regulares	5
4 Gramática de estructura de frases: Categorías sintácticas	6
4 Gramática de estructura de frases: Categorías sintácticas	7
5 Gramática de estructura de frases: BNF	8

Creando un lenguaje básico

Tipo de datos

Decidimos crear un lenguaje que posea el tipo de dato entero.

Convención predefinida de los identificadores

- La longitud de los identificadores, es decir, el número finito de caracteres individuales permitidos en un identificador, es de 16.
- Los identificadores comenzarán con una letra mayúscula, y están compuestas únicamente por letras, queda totalmente prohibido la utilización de dígitos o caracteres especiales al momento de definir una variable.

Extra: Las variables se declaran de forma implícita, ya que el único tipo de dato es entero.

Sentencias definidas

La convención de fin de sentencia es el punto “.”, el cuerpo del programa está delimitado por las palabras reservadas **Inicio** y **Fin**.

1) Asignación

ID = Expresión.

Siendo ID un identificador de tipo entero, y la Expresión se construya a partir de identificadores, operadores aritméticos y constantes enteras. Se permite el uso de paréntesis.

2) Entrada/Salida

Capturar(ID).

Imprimir(ID).

3) Declaración

Declarar Identificador.

Palabras reservadas

Decidimos tener 4 palabras reservadas, **Inicio**, que dará apertura al programa, **Fin**, indicando que el programa terminó, **Capturar** para obtener un dato de consola, **Imprimir** para mostrar un resultado en consola, y **Declarar** para reservar un espacio en memoria a un identificador. Se puede apreciar que todas las palabras reservadas siguen la convención de los identificadores, empezando con una letra mayúscula.

Alfabeto utilizado

$\Sigma = \{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z, A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, 1, 2, 3, 4, 5, 6, 7, 8, 9, 0, ., , , (, =, *, +, -, / \}$

Gramática léxica: Categorías léxicas

Componentes léxicos o **tokens**: Identificador, palabra reservada, operador, asignación, constante entera, carácter puntuación.

1) Identificadores

$\langle \text{Identificador} \rangle ::= \langle \text{letra mayuscula} \rangle \langle \text{palabra} \rangle$

$\langle \text{letra mayuscula} \rangle ::= A \mid B \mid C \mid D \mid \dots \mid Z$

$\langle \text{palabra} \rangle ::= \langle \text{letra} \rangle \mid \langle \text{letra} \rangle \langle \text{palabra} \rangle$

$\langle \text{letra} \rangle ::= a \mid b \mid c \mid d \mid \dots \mid z \mid A \mid B \mid C \mid D \mid \dots \mid Z$

2) Palabras reservadas

$\langle \text{palabra reservada} \rangle ::= \text{Inicio} \mid \text{Fin} \mid \text{Capturar} \mid \text{Imprimir} \mid \text{Declarar}$

3) Operador

$\langle \text{operador} \rangle ::= * \mid + \mid - \mid /$

4) Asignación

$\langle \text{asignación} \rangle ::= =$

5) Constante entera

$\langle \text{constante entera} \rangle ::= \langle \text{constante entera} \rangle \langle \text{digito} \rangle \mid \langle \text{digito} \rangle$

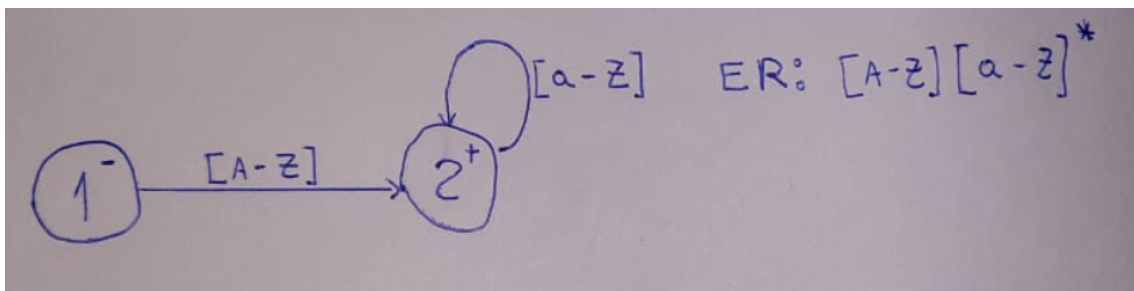
$\langle \text{digito} \rangle ::= 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$

6) Carácter puntuación

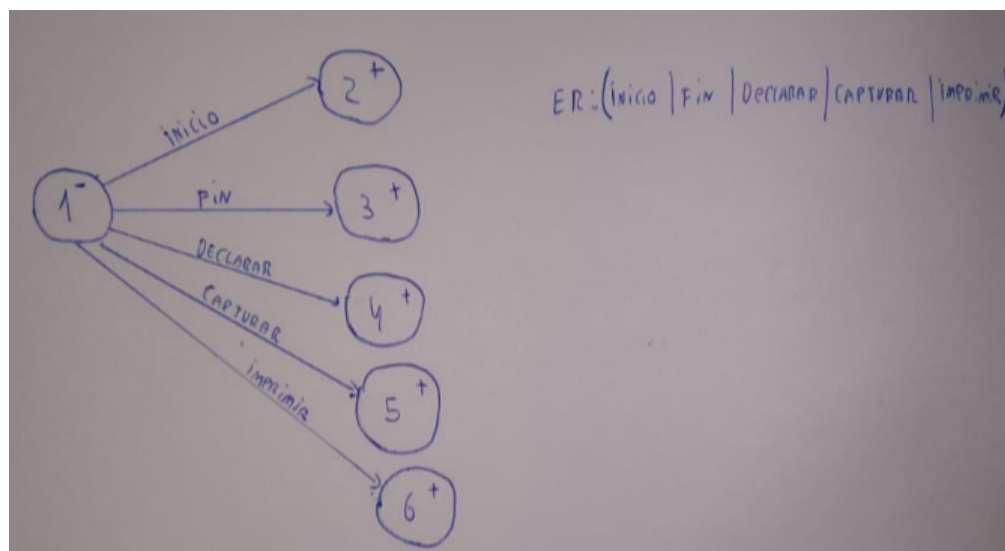
$\langle \text{carácter puntuación} \rangle ::=) \mid (\mid .$

Gramática léxica: Autómata finito y expresiones regulares.

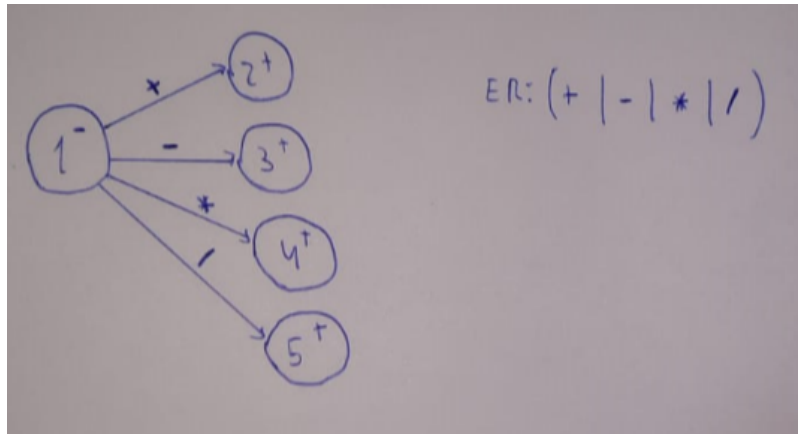
1) Token identificador (Alfabeto {a,b,c,...,z,A,B,C,...,Z})



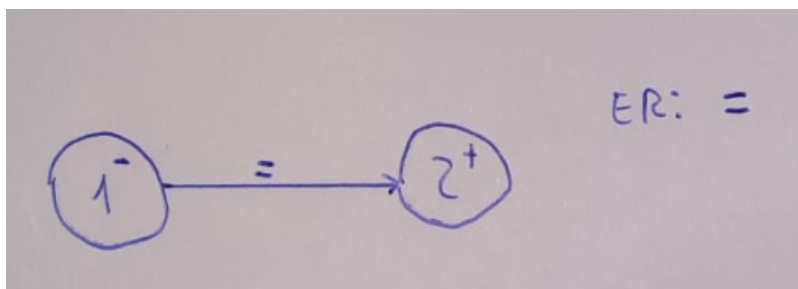
2) Token palabra reservada (Alfabeto {Inicio, Fin, Capturar, Imprimir, Declarar})



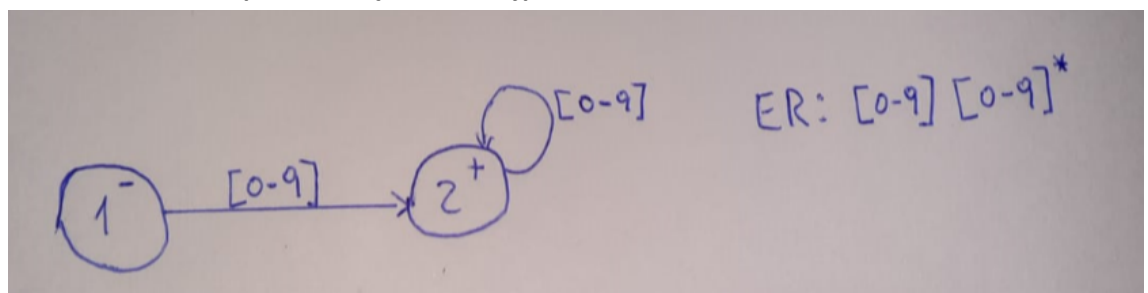
3) Token operador (Alfabeto {+, -, /, *})



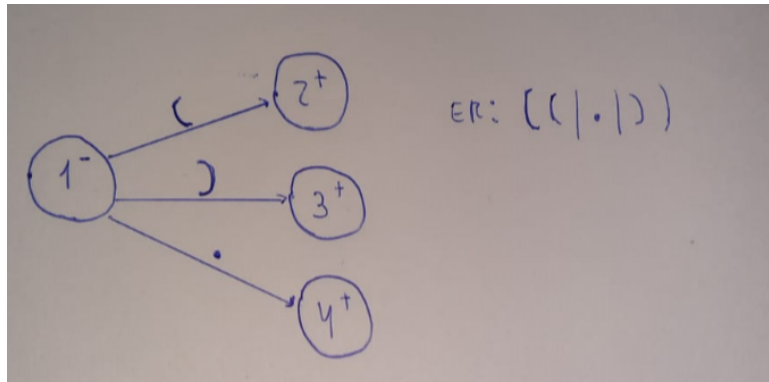
4) Asignación (Alfabeto {=})



5) Constante entera (Alfabeto {0,1,2,...,9})



6) Carácter puntuación (Alfabeto { }, (, . })



Gramática de Estructura de Frases: Categorías Sintácticas.

Analizaremos la sintaxis de los constructos , las cuales son:

1) EXPRESIONES

Expresion -> Termino | Termino Operador Termino

Termino -> Constante | (Expresion)

Siendo **Expresion** el Axioma.

Siendo **Operador** parte de las categorías léxicas, cuyo token es **Operador**.

Siendo **Constante** parte de las categorías léxicas, cuyo token es **Constante entera**.

Siendo (y) parte del alfabeto y componentes del token **carácter puntuación**.

2) DECLARACIONES

Declaracion -> Declarar Identificador

Siendo **Declaracion** el Axioma.

Siendo Declarar un elemento del alfabeto, y elemento de la categoría léxica **palabra reservada**.

Siendo **Identificador** parte de las categorías léxicas, cuyo token es **Identificadores**.

3) SENTENCIAS

- Estructura del programa

Un programa está compuesto por declaraciones, asignaciones, y sentencias:

Programa -> Inicio RestoPrograma

RestoPrograma -> Fin | SentAsignacion. RestoPrograma |

Declaracion. RestoPrograma | SentEntrada. RestoPrograma |

SentSalida. RestoPrograma

Siendo **Inicio** y **Fin** partes del **alfabeto** además de **palabras reservadas**.

Siendo '.' parte del alfabeto y componente del token **carácter puntuación**.

Siendo **Declaracion** una categoría sintáctica definida anteriormente.

Siendo **SentAsignacion**, **SentEntrada** y **SentSalida** sentencias definidas a continuación.

- Asignación

SentAsignacion -> Identificador = Expresion

Siendo **Identificador** parte de las categorías léxicas, cuyo token es **Identificadores**.

Siendo = parte del alfabeto y componente del token **Asignacion**.

Siendo **Expresion** una categoría sintáctica definida anteriormente.

- Entrada

SentEntrada -> Capturar(Identificador)

Siendo **Identificador** parte de las categorías léxicas, cuyo token es **Identificadores**.

Siendo (y) parte del alfabeto y componentes del token **carácter puntuación**.

Siendo **Capturar** parte del alfabeto y componente del token **palabras reservadas**.

- Salida

SentSalida -> Imprimir(Identificador)

Siendo **Identificador** parte de las categorías léxicas, cuyo token es **Identificadores**.

Siendo (y) parte del alfabeto y componentes del token **carácter puntuación**.

Siendo **Imprimir** parte del alfabeto y componente del token **palabras reservadas**.

Gramática de Estructura de Frases: BNF

1) Expresiones

<expresion> ::= <termino> | <termino> <operador> <termino>

<termino> ::= <constante entera> | (<expresion>)

2) Declaraciones

<declarar> ::= Declarar <identificador>

3) Sentencias

Estructura del programa

<programa> ::= Inicio <resto del programa>

<resto del programa> ::= Fin |

<sentencia asignacion> <punto> <resto del programa> |

<declaracion> <punto> <resto del programa> |

<sentencia entrada> <punto> <resto del programa> |

<sentencia salida> <punto> <resto del programa>

<punto> ::= .

Asignacion

<sentencia asignacion> ::= <identificador> <asignacion> <expresion>

Entrada

<sentencia entrada> ::= Capturar(<identificador>)

Salida

<sentencia salida> ::= Imprimir(<identificador>)