



10 DE MARZO DE 2021

GRAMATICA
PROYECTO 1 REGEXIVE

ARIEL RUBELCE MACARIO CORONADO - 201905837
UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
Compiladores 1 B2



GRAMÁTICA.

Sea la gramática $G = \text{Terminales, No terminales, } P$, donde P son las Producciones, a continuación se definen los terminales y No terminales.

DECLARACIÓN DE ALFABETO Y SÍMBOLOS DEL LENGUAJE.

ALFABETO.

blancos = $[\backslash r \backslash t \backslash n]^+$

letra = $[a, b, c, d, e, f, g, h, i, j, k, l, m, n, \tilde{n}, o, p, q, r, s, t, u, v, w, x, y, z, A, B, C, D, E, F, G, H, I, J, K, L, M, N, \tilde{N}, O, P, Q, R, S, T, U, V, W, X, Y, Z.]$

entero = $[0, 1, 2, 3, 4, 5, 6, 7, 8, 9,]$

identificador = $\{letra\}(\{letra\} | _ | \{entero\})^*$

cadena = $" "[^"]^+ "$

especialC = $("\\ " n | "\\ " \backslash | "\\ " "")$

comunilinea = $("/" .* \backslash n) ("/" .* \backslash r)$

commultilinea = $("< "" ! "[^ \backslash !] * "" ! "" ">)$

a33: $!"$

a34: $" "$

a35: $"\#"$

a36: $"\$"$

a38: $"\&"$

a39: $"\backslash"$

a40: $"($

a41: $)"$

a47: $"/"$

a60: $"<"$

a61: $"="$

a64: $"@"$

a91: $"["$

a92: "\\"

a93: "]"

a94: "^"

a95: "_"

a96: "`"

SÍMBOLOS PROPIOS DEL LENGUAJE.

Unoomasveces: "+"

concatenacion: "."

ceroomasveces: "*"

disyuncion: "|"

ceroounavez: "?"

tildeseparacion: "~"

llaveapertura: "{"

llavecerradura: "}"

dospuntos: ":"

guion: "-"

mayorq: ">"

puntoycoma: ","

porcentaje: "%"

comma: ","

conj: "CONJ"

Debemos de definir todos los símbolos que nuestro lenguaje usara para poder dar lectura al archivo, también debemos de definir qué tipos de símbolos puede admitir nuestro lenguaje, en este caso lo llamamos alfabeto.

TERMINALES.

- conj, unaomasveces, concatenacion, ceroomasveces, disyuncion, ceroounavez, tildeseparacion, especialC;

- llaveapertura, llavecerradura, dospuntos, guion, mayorq, puntoycoma, porcentaje, identificador;
- entero, cadena, comma, a33, a34, a35, a36, a38, a39, a40, a41, a47, a60, a61, a64, a91, a92, a93, a94, a95, a96;

Los Terminales a que se usaron para la definición de la gramática, son aquellas palabras reservadas que el lenguaje tiene, así como los símbolos que usa y los símbolos que permite ingresar.

NO TERMINALES.

INICIO, CUERPO, CONJUNTO, CONJUNTODEF, CODIGOASCCI, EXPVALIDAR, CADENAAR EXPRESION, EXPRESIONDEF

Los No terminales son aquellos inician las producciones que deben de seguir, usan los no terminales y terminales.

PRODUCCIONES.

INICIO→

llaveapertura CUERPO porcentaje porcentaje porcentaje porcentaje EXPVALIDAR llavecerradura

Esta producción da inicio a todas las demás producciones, es la que de que el archivo tenga llave de apertura y cerradura y tenga un cuerpo dentro y por último que tenga su porcentajes.

CUERPO→ CONJUNTO CUERPO

→EXPRESION CUERPO

→EXPRESION

→CONJUNTO

Esta producción da inicio al cuerpo, donde en el cuerpo nos puede hacer llegar los conjuntos, y la expresión para el árbol binario.

CONJUNTO→ conj dospuntos identificador guion mayorq CONJUNTODEF puntoycoma

Esta producción es la que nos define nuestros conjuntos, en donde le especificamos que es lo que deben de venir, como se puede observar debe de venir la palabra reservada CONJ seguida de los dos puntos un identificador, seguidos de un guion y mayor que y finalizando con punto y coma.

CONJUNTODEF→

- identificador tildeseparacion identificador
- entero tildeseparacion entero
- CODIGOASCCI tildeseparacion CODIGOASCCI
- identificador comma CONJUNTODEF
- entero comma CONJUNTODEF
- CODIGOASCCI comma CONJUNTODEF
- identificador
- entero
- CODIGOASCCI
- especialC

Para esta producción estas definiendo que símbolos deben permite el archivo para los conjuntos, en donde puede venir una letra, numero, o un Carácter ASCII del 35 a 124, y en donde se pueden definir como carácter a carácter o colocando un rango de inicio y usando la tilde de separación con un rango de fin.

CODIGOASCCI→ a33

- a34
- a35
- a36
- a38
- a39
- a40
- a41
- a47
- a60
- a61
- a64
- a91
- a92

- a93
- a94
- a95
- a9
- naomasveces
- concatenacion
- ceroomasveces
- disyuncion
- ceroounavez
- tildeseparacion
- llaveapertura
- llavecerradura
- dospuntos
- guion
- mayorq
- puntoycoma
- porcentaje
- comma

La producción anterior nos ayuda a definir que Carácter de tipo ASCII acepta nuestros conjuntos.

EXPRESION→ identificador guion mayorq EXPRESIONDEF puntoycoma

Esta producción nos ayuda a definir la expresión para poder reconocer las cadenas que próximamente estarán en siendo evaluadas.

EXPRESIONDEF →

- disyuncion EXPRESIONDEF EXPRESIONDEF
- concatenacion EXPRESIONDEF EXPRESIONDEF
- unaomasveces EXPRESIONDEF
- ceroomasveces EXPRESIONDEF

→ ceroounavez EXPRESIONDEF

→ llaveapertura identificador llavecerradura

→ cadena

→especialC

Producción que nos sirve para reconocer las expresiones, debemos de saber como viene construida y realizar la precedencia de operadores de la mejor manera.

EXPVALIDAR→ CADENAAR EXPVALIDAR

→CADENAAR

Producción que nos ayuda a la lectura de las cadenas ingresadas por el archivo de entrada.

CADENAAR→ identificador dospuntos cadena:a puntoycoma

→ identificador dospuntos especialC:a puntoycoma

La anterior producción nos ayuda a validar que la cadena ingresada cumpla con los parámetros requeridos por el lenguaje.