UNIDAD III. LENGUAJES LIBRES DE CONTEXTO.

GRAMATICA LIBRE DE CONTEXTO

Es un conjunto finito de variables, cada uno de los cuales representa un lenguaje.

Los lenguajes representados por esas variables se definen, recursivamente, unos en terminos de otros y de simbolos primitivos, llamados terminales.

Las reglas relacionadas a las variables se llaman producciones, reglas de produccion o reglas de reescritura.

Una produccion tipica establece que el lenguaje asociado con una variable dada, contiene cadenas que se forman por la concatenacion de cadenas de los lenguajes de ciertas variables.

FORMA DE BACKUS - NAUR   Producciones o Reglas de produccion

$<EXP> \to <EXP> + <EXP>$

$<EXP> \to <EXP> - <EXP>$

$<EXP> \to <EXP> * <EXP>$

$<EXP> \to <EXP> / <EXP>$

$<EXP> \to (<EXP>)$

$<EXP> \to id$

$\boxed{<EXP> \ - \ Variables}$

$\boxed{+, -, *, /, (,), id \ - \ Terminales}$

GRAMATICA LIBRE DE CONTEXTO

Una GLC se denota por la cuarteta $G = (V, T, P, S)$, *donde V es el alfabeto de variables y T es el alfabeto de terminales.*

Se asume que V y T son ajenos y que $V \cup T$ *es el vocabulario de la gramatica y sus elementos son los simbolos de la gramatica.*

S es un elemento distinguido de V y se le conoce como simbolo inicial de la gramatica.

P es el conjunto de producciones de la gramatica.

Las producciones son de la forma $A \to \alpha$ *(A produce a* α*), donde A es un elemento de V y* α *pertenece a* $(V \cup T)^*$. *A es el lado izquierdo de la produccion y* α *es el lado derecho.*

Ejemplo : Sea la gramatica $G = (\{E\}, \{+, -, *, /, (,), id\}, P, E)$

P :

$$\begin{cases} E \to E + E \\ E \to E - E \\ E \to E * E \\ E \to E / E \\ E \to (E) \\ E \to id \end{cases}$$

CONVENCION

A, B, C, D, E Y S denotan variables. S es el simbolo inicial.

a, b, c, d, e, digitos y cadenas en negritas son terminales.

u, v, w, x, y, z denotan cadenas de terminales.

α, β, γ *denotan cadenas de variables y terminales.*

Si $A \to \alpha_1, A \to \alpha_2, \ldots, A \to \alpha_k$, *entonces se puede expresar como :*

$$A \to \alpha_1 \mid \alpha_2 \mid \ldots \mid \alpha_3$$

$$\boxed{E \to E + E \mid E - E \mid E * E \mid E / E \mid (E) \mid id}$$

DERIVACIONES Y LENGUAJES

Si $A \to \beta$ *es una produccion de P y* α, γ *son cadenas de simbolos de la gramatica, entonces* $\alpha A \gamma \to \alpha\beta\gamma$, *es decir, la produccion* $A \to \beta$ *se aplica a la cadena* $\alpha A \gamma$ *para obtener la cadena* $\alpha\beta\gamma$ *en G.*

Suponiendo que $\alpha_1, \alpha_2, \ldots \alpha_m$ *pertenecen a* $(V \cup T)^*$ *con* $m \geq 1$ *y que*

$$\alpha_1 \underset{G}{\to} \alpha_2, \alpha_2 \underset{G}{\to} \alpha_3, \ldots, \alpha_{m-1} \underset{G}{\to} \alpha_m$$

entonces se dice que $\alpha_1 \overset{*}{\underset{G}{\to}} \alpha_m$ *o que* α_1 *deriva a* α_m *en G.*

De manera alternativa se dice que $\alpha \to \beta$ *si* β *se obtiene de* α *por la aplicacion de cero o mas producciones de P.*

Si α *deriva a* β *en i pasos, se escribe* $\alpha \overset{i}{\to} \beta$

El lenguaje generado por G es :

$$L(G) = \left\{ \omega \mid \omega \in T^*, S \overset{*}{\underset{G}{\to}} \omega \right\}$$

Una cadena esta en L (G) si :

I. La cadena consta unicamente de terminales.
II. La cadena es derivable a partir de S.

Un lenguaje es libre de contexto (LLC) si es un lenguaje generado por una GLC.

Una cadena de variables y terminales α*, es una forma oracional si* $S \overset{*}{\underset{G}{\to}} \alpha$

Dos gramaticas G_1 *y* G_2 *son equivalentes si generan el mismo lenguaje.*

Ejemplo:

$\boxed{G = (V, T, P, S)}$ $\boxed{V = \{S\} \quad T = \{a, b\}}$ $\boxed{P: \; S \to aSb \quad S \to ab}$

$\boxed{\text{Aplicamos la produccion 1 } (n-1) \text{ veces}}$

$$\begin{array}{c} S \to aSb \\ S \to aaSbb \\ S \to a^2Sb^2 \\ S \to a^2aSbb^2 \\ \ldots \\ S \to a^{n-1}Sb^{n-1} \\ S \to a^{n-1}abb^{n-1} \\ S \to a^nb^n \end{array}$$

$\boxed{L(G) = \{a^nb^n \mid n \geq 1\}}$

$\boxed{0^n1^n \text{ no es un lenguaje regular}}$