ESCOM

Clasificación de Gramáticas

Teoría Computacional
Prof. Luis Enrique Hernández Olvera

ESCOM

Recordando

- La **gramática** es el **estudio de las reglas y principios** que regulan el uso de las lenguas y la organización de las palabras dentro de una oración. También se denomina así al "**conjunto de reglas y principios que gobiernan el uso de un lenguaje**" así, cada lenguaje tiene su propia gramática.

Teoría Computacional
Prof. Luis Enrique Hernández Olvera 2

Recordando

Una Gramática G se representa con una cuádrupla:

G=(N,Σ,S,P)

Donde:

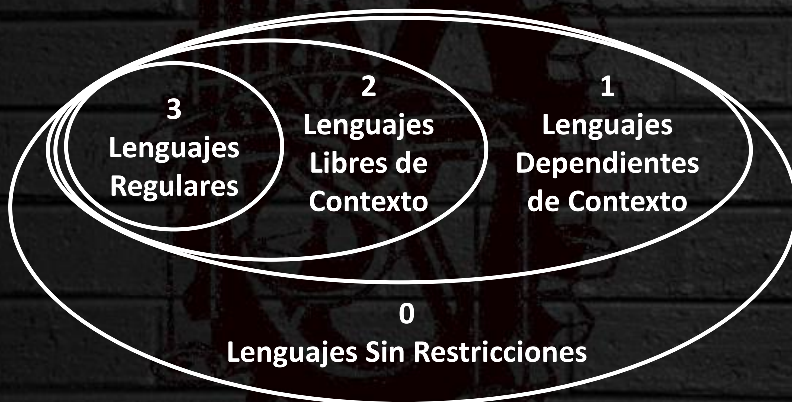
- N es una colección finita de no terminales.
- Σ es un alfabeto (Conjunto de terminales).
- S es un no terminal llamado **Símbolo inicial.**
- P es una colección finita de **reglas de sustitución** llamadas **producciones**.

Clasificación de Gramáticas

- En 1956, Noam Chomsky clasificó las gramáticas en cuatro tipos de lenguajes y esta clasificación es conocida como la jerarquía de Chomsky, en la cual cada lenguaje es descrito por el tipo de gramática generado. Estos lenguajes sirven como base para la clasificación de lenguajes de programación.

Clasificación de Gramáticas

- Los cuatro tipos son:



Clasificación de Gramáticas

- En función de la forma de sus producciones, se puede caracterizar qué tan compleja es una gramática formal.
  - **Gramáticas Tipo 0** *(sin restricciones)*
  - **Gramáticas Tipo 1** *(dependientes de contexto)*
  - **Gramáticas Tipo 2** *(independientes o libres de contexto)*
  - **Gramáticas Tipo 3** *(gramáticas regulares)*

Gramáticas Tipo 3 *(gramáticas regulares)*

- Estas gramáticas se clasifican en los dos grupos siguientes:
  - **Gramáticas lineales por la izquierda**, cuyas reglas de producción pueden tener una de las formas siguientes:
    - **A → a**
    - **A → Va**
    - **S → λ**
  - **Gramáticas lineales por la derecha**, cuyas reglas de producción tendrán la forma:
    - **A → a**
    - **A → aV**
    - **S → λ**

Gramáticas Tipo 3 *(gramáticas regulares)*

- Los lenguajes representados por este tipo de gramáticas se denominan **lenguajes regulares**.
- Los lenguajes regulares se utilizan para definir **estructura léxica** de los lenguajes de programación. **Definen la sintaxis** de los **identificadores**, **números**, **cadenas** y otros **elementos básicos** del lenguaje.
- Estos lenguajes son todos los lenguajes que pueden ser reconocidos por los autómatas finitos.

Gramáticas Tipo 3 *(gramáticas regulares)*

Para una gramática de la forma G = (N, Σ, S, P)

G_1 = ({A, B}, { 0, 1}, A, { $A \rightarrow B1 \mid 1$, $B \rightarrow A0$})

Gramática lineal por la izquierda que describe el lenguaje:

L_1 = { 1, 101, 10101, ... } = $\{1(01)^n \mid n = 0, 1, 2, ...\}$

G_2 = ({ 0, 1}, {A, B}, A, { $A \rightarrow 1B \mid 1$, $B \rightarrow 0A$})

Gramática lineal derecha que genera el mismo lenguaje que la gramática anterior.

Gramáticas Tipo 2 *(independientes o libres de contexto)*

- Generan los lenguajes libres de contexto. Están definidas por reglas de la forma:

$$\mathbf{A} \rightarrow \gamma$$

  - **A** es un no terminal
  - γ es una cadena de terminales y no terminales (incluye λ).
- Se denominan independientes de contexto porque A puede sustituirse por γ independientemente de las cadenas por las que esté acompañada.
- Estos lenguajes son todos los lenguajes que pueden ser reconocidos por los autómatas de pila.

Gramáticas Tipo 2 *(independientes o libres de contexto)*

- Los lenguajes independientes de contexto constituyen la base teórica para la sintaxis de la mayoría de los lenguajes de programación. Definen la sintaxis de las declaraciones, las proposiciones, las expresiones, etc.
- Sea la gramática G = ({S}, {a, b}, S, { S $\rightarrow$ aSb | ab}).
- La derivación de la palabra aaabbb será:
- S $\Rightarrow$ aSb $\Rightarrow$ aaSbb $\Rightarrow$ aaabbb
- Puede verse que el lenguaje definido por esta gramática es $\{a^n b^n \mid n=1, 2, ...\}$

Gramáticas Tipo 1 *(dependientes de contexto)*

- Generan los lenguajes dependientes de contexto. Contienen reglas de producción de la forma:

$$\alpha A \beta \rightarrow \alpha \gamma \beta$$

  - **A** es un no terminal
  - α, β y γ **son** cadenas de terminales y no terminales.
  - α y β pueden ser vacíos, pero γ ha de ser distinto del vacío.

- Se denominan gramáticas dependientes del contexto, porque, como se observa, A puede ser sustituido por γ si está acompañada de α por la izquierda y de β por la derecha.
- Estos lenguajes son todos los lenguajes que pueden ser reconocidos por autómatas lineales acotados (Maquina de Turing Determinista).

Gramáticas Tipo 1 *(dependientes de contexto)*

- G = ({S, B, C}, {a, b}, S,P), donde P es:
  - $S \rightarrow aSBc \mid aBC$
  - $bB \rightarrow bb$
  - $bC \rightarrow bc$
  - $CB \rightarrow BC$
  - $cC \rightarrow cc$
  - $aB \rightarrow ab$

Gramáticas Tipo 0 *(sin restricciones)*

- Incluyen todas las gramáticas formales.
- El más general, al que pertenece la semántica de los lenguajes naturales y artificiales.
- A estos lenguajes no se les impone restricción alguna.
- Estos lenguajes son todos los lenguajes que pueden ser reconocidos por una *máquina de Turing.*

Clasificación de Gramáticas

- Todo lenguaje de tipo 3 es de tipo 2, todo lenguaje de tipo 2 es de tipo 1, y todo lenguaje de tipo 1 es de tipo 0.
- Se dice que un lenguaje es de tipo k [k = 0, k = 1, k = 2, k = 3] cuando existe una gramática de tipo k que genera ese lenguaje.

Clasificación de Gramáticas

- Para clasificar una gramática hemos de analizar una a una todas sus reglas de producción obteniendo el tipo de cada una de ellas. La clasificación de la gramática será la correspondiente al tipo de la producción de menor clasificación

Gramática	Lenguaje	Reglas de Producción	Si $\mu \rightarrow \omega$, relación entre $\|\mu\|$ y $\|\omega\|$	Solución
Tipo-0	Recursivas	Sin restricciones		Máquinas de Turing
Tipo-1	Dependiente de contexto	$\alpha A\beta \rightarrow \alpha\gamma\beta$	$\|\mu\| \leq \|\omega\|$	Autómatas lineales acotados
Tipo-2	Independiente de contexto	$A \rightarrow \gamma$	$\|\mu\| = 1$	Autómatas de pila
Tipo-3	Regular	$A \rightarrow aB$ $A \rightarrow a$	$\|\mu\| = 1$	Autómatas finitos, regulares

Ejercicios.

- Clasificar las siguientes gramáticas dadas sus reglas de producción.

(a)	(b)	(c)	(d)	(e)	(f)	(g)
Z → yX X → y X → λ yX → yx	X → xZyW yW → yx Z → vy	E → E+T E → E-T E → T T → T*F T → T/F T → F F → (E) F → id	*S → aAbc* *Ab → bA* *Ac → Bbcc* *bB → bbaA* *A → aa* *B → bb*	A → bC A → bBC bB → bCa C → b C → λ yCc → yCc	S → aAB|A A → cBd B → e|fS C → gD|hDt D → x|y|z	S → aS S → aN N → bN N → bM N → b M → c