



# Instituto Politécnico Nacional

## Escuela Superior de Cómputo



# *Teoría computacional*

## **Clase 11: Gramáticas**

**Solicitado:** Ejercicios 09: Gramáticas

M. en C. Edgardo Adrián Franco Martínez  
<http://computacion.cs.cinvestav.mx/~efranco>

[@efranco\\_escom](#)  
edfrancom@ipn.mx



# Contenido

- Gramática
- Elementos de una gramática
  - Ejemplo 01
  - Observaciones de las producciones
- Ejemplos de gramáticas
- Lenguaje
- Lenguaje generado por una gramática (Derivaciones)
- Importancia de las gramáticas
  - Ejemplo: Palíndroma (PAL)
  - Ejemplo: No Palíndroma (NOPAL)
- Ejercicios 09: Gramáticas



# Gramática



- 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.





“La gramática es un ente formal para especificar, de una manera finita, el conjunto de cadenas de símbolos que constituyen un lenguaje”.

- Una gramática es una cuádrupla :

$$G = ( VT , VN , S , P )$$

donde :

- $VT = \{\text{conjunto finito de símbolos terminales}\}$
- $VN = \{\text{conjunto finito de símbolos no terminales}\}$
- $S$  es el símbolo inicial y pertenece a  $VN$ .
- $P = \{\text{conjunto de producciones o de reglas de derivación}\}$





- **Todas las cadenas** del lenguaje definido por la gramática están formadas con símbolos del **vocabulario terminal VT**.
  - El vocabulario terminal se define por enumeración de los símbolos terminales.
- El **vocabulario no terminal VN** es el conjunto de **símbolos introducidos como elementos auxiliares para la definición de la gramática**, y que no figuran en las cadenas del lenguaje.
  - El vocabulario no terminal se define por enumeración de los símbolos no terminales.
- La **intersección** entre el vocabulario terminal y no terminal es el **conjunto vacío**:

$$\{VN\} \cap \{VT\} = \{\emptyset\}$$

- La **unión** entre el vocabulario terminal y no terminal es el **vocabulario** :

$$\{VN\} \cup \{VT\} = \{V\}$$



- En ocasiones es importante distinguir si un determinado vocabulario incluye o no la cadena vacía.

$$V^+ = V - \{\lambda\}$$

$$V^* = V + \{\lambda\}$$

- Las producciones **P** son **las reglas que se aplican desde el símbolo inicial** para obtener las cadenas del lenguaje.
  - El conjunto de producciones P se define por medio de la enumeración de las distintas producciones, en forma de reglas o por medio de un **metalenguaje**.
    - *p.g. BNF (Backus Naur Form) o EBNF (Extended Backus Naur Form).*



# Elementos de una gramática



- **Símbolos terminales:** son elementos del alfabeto que no se pueden transformar, por eso se llaman terminales. Normalmente se denotan por letras minúsculas.
- **Variables o símbolos no terminales:** son elementos auxiliares que permiten poner restricciones sintácticas a un lenguaje. Las variables sí se pueden transformar, utilizando las reglas, en una cadena de variables y/o terminales. Por lo general se denotan por letras mayúsculas o por la notación *<variable>*.
- **Reglas:** permiten reemplazar variables para generar oraciones válidas de un lenguaje. Puede haber varias reglas para una misma variable, en algunos casos, las distintas opciones se colocan en una sola regla con los distintos reemplazos separados por “|”. P.g.  $\alpha \rightarrow \beta \mid \gamma \mid \delta$  abrevia las tres reglas  $\alpha \rightarrow \beta$ ,  $\alpha \rightarrow \gamma$ ,  $\alpha \rightarrow \delta$ .
- **Símbolo inicial:** es el símbolo a partir del cual se generan todas las palabras válidas.



# Ejemplo 01

- $G = (VT, VN, S, P)$

- $VN = \{S, A, B\}$

- $VT = \{a, b, c\}$

- $P: S \rightarrow AccA \quad A \rightarrow BA \mid \lambda \quad B \rightarrow a \mid b \mid c$

$$w_1 = abcc \in L(G)$$

$$w_2 = acb \notin L(G)$$







- $G = ( VT , VN , S , P )$ 
    - $VN = \{ S, A, B \}$
    - $VT = \{ a, b, c \}$
    - $P: S \rightarrow AccA \quad A \rightarrow BA \mid \lambda \quad B \rightarrow a \mid b \mid c$
- $w_1 = abcc \in L(G) \quad y \quad w_2 = acb \notin L(G)$

Cadena	Regla	Derivación
<b>S</b>	$S \rightarrow AccA$	$S \Rightarrow AccA$
AccA	$A \rightarrow BA$	$\Rightarrow BAccA$
BAccA	$B \rightarrow a$	$\Rightarrow aAccA$
aAccA	$A \rightarrow BA$	$\Rightarrow aBAccA$
aBAccA	$B \rightarrow b$	$\Rightarrow abAccA$
abAccA	$A \rightarrow \lambda$	$\Rightarrow abccA$
abccA	$A \rightarrow \lambda$	$\Rightarrow \textcolor{red}{abcc}$



# Observaciones de las producciones



- Por lo general las reglas se escriben  $\alpha \rightarrow \beta$  en lugar de  $(\alpha, \beta)$ .
- Aplicar la regla  $\alpha \rightarrow \beta$  a una palabra  $a\alpha b$  produce la palabra  $a\beta b$ , por lo que las reglas pueden ser vistas como **reglas de remplazo**.



# Ejemplos de gramáticas



## Ejemplo 1

Sea la gramática :  $G = (VT, VN, S, P)$  donde  $VT = \{a, b\}$ ,  $VN = \{S\}$ , y el conjunto de producciones es :

$$S \rightarrow ab$$

$$S \rightarrow aSb$$

## Ejemplo 2

Sea la gramática  $G = (VN, VT, S, P)$  donde :

$$VN = \{ \langle \text{número} \rangle, \langle \text{dígito} \rangle \}$$

$$VT = \{ 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 \}$$

$$S = \langle \text{número} \rangle$$

Las reglas de producción  $P$  son :

$$\langle \text{número} \rangle ::= \langle \text{dígito} \rangle \langle \text{número} \rangle$$

$$\langle \text{número} \rangle ::= \langle \text{dígito} \rangle$$

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



# Lenguaje



- Se puede definir como **un conjunto de palabras de un determinado alfabeto**.
- Los lenguajes se pueden definir por enumeración de las cadenas que pertenecen a dicho lenguaje, pero esto además de ineficiente, es en muchos casos imposible (*habitualmente un lenguaje tiene un numero infinito de cadenas*).
- Así los **lenguajes se definen por las propiedades que cumplen las cadenas del lenguaje**.



# Lenguaje generado por una gramática



- El lenguaje  $L(G)$  generado por una gramática  $G$  es el conjunto de todas las **cadena**s que puede generar  $G$ . Es decir expresado formalmente :

$$L(G) = \{\eta \in VT^* / S \rightarrow \eta\}$$

- Una **cadena** pertenece a  $L(G)$  si :
  - Está compuesta de símbolos terminales
  - La cadena puede derivarse del símbolo inicial  $S$  aplicando las reglas de producción de la gramática.



# Importancia de las gramáticas



**Son una herramienta  
muy poderosa para  
describir y analizar  
lenguajes.**





# Ejemplo: Palíndroma (PAL)

- Un palíndroma (PAL) es una palabra  $w$  que cumple  $w = w^R$ , donde  $w^R$  es la misma palabra  $w$  sólo que escrita en orden inverso.
- Sea  $PAL = \{w \in (a|b)^* \mid w = w^R\}$ , es decir,  $PAL$  es el lenguaje que consiste de las palabras palíndromas sobre  $\{a, b\}$ .
- Ejemplos de palíndromas sobre  $\{a, b\}$ 
  - $a, aba, aabbaa, babaabab$ .



# Definición de *PAL*



1.  $\lambda, a, b \in PAL$
2. Si  $S \in PAL$  entonces  $aSa$  y  $bSb \in PAL$
3. Ninguna cadena pertenece a *PAL* a menos que pueda ser obtenida utilizando las reglas 1 y 2 un número finito de veces. Podemos considerar a  $S$  como una variable que representa genéricamente un elemento de *PAL* cuyo valor queremos calcular. Informalmente, las reglas que definen *PAL* pueden reescribirse como:
  - $S$  puede tomar el valor  $\lambda, a$  o  $b$ .
  - $S$  puede tomar el valor  $aSa$  o  $bSb$ , donde la nueva  $S$  debe ser calculada. Y si en lugar de escribir “*puede tomar el valor*” escribimos el símbolo  $\rightarrow$ , entonces escribiríamos ...





# Gramática que define *PAL*

1.  $s \rightarrow \lambda, s \rightarrow a, s \rightarrow b.$

Abreviando:  $s \rightarrow \lambda \mid a \mid b.$

2.  $s \rightarrow asa, s \rightarrow bSb.$

Abreviando:  $s \rightarrow asa \mid bSb.$

3. Cualquier palíndroma sobre  $\{a, b\}$  debe poder ser obtenido aplicando un número finito de veces las reglas de reemplazo 1 y 2.





$PAL = (VT, VN, S, P)$

$VT = \{a, b\}$

$VN = \{S\}$

$P =$

$L(PAL) = \{a, b, aba, bab, aabbaa, babaabab, \dots\}$

$S \rightarrow aSa$

$S \rightarrow bSb$

$S \rightarrow a$

$S \rightarrow b$

$S \rightarrow \lambda$





# Ejemplo: No Palíndroma (NOPAL)

$S \rightarrow aBb$

$S \rightarrow bBa$

$S \rightarrow aSa$

$S \rightarrow bSb$

$B \rightarrow aB$

$B \rightarrow bB$

$B \rightarrow \lambda$

$L(NOPAL) = \{ab, ba, abab, baba, aabbaabb, babaababa, \dots\}$



# Ejercicios 09 "Gramáticas"



- Diseñe gramáticas para los siguientes lenguajes:
  1. El conjunto de todas las cadenas de 0's y 1's , de tal forma que justo antes de cada 0 vaya por lo menos un 1.
  2. El conjunto de todas las cadenas de 0's y 1's que sean *palíndromos*; es decir, que la cadena se lea igual al derecho y al revés y tenga por lo menos dos 1's seguidos .
  3. El conjunto de todas las cadenas de 0's y 1's en donde 011 no aparece como subcadena.
  4. El conjunto de todas las cadenas que genere el conjunto de cadenas (sa, sasa, sasasa,...).

*\*Se entregarán antes del día **Lunes 14 de Octubre de 2013**  
(23:59:59 hora limite)*

*\*Incluir la redacción de cada ejercicio*

*\*Portada y encabezados de pagina*

