

# Las Gramáticas Formales

## Como definir un Lenguaje Formal

Universidad de Cantabria

# Esquema

- 1 Motivación
- 2 Definición de Gramáticas Formales
- 3 Tipos de Notación

# Problema

Dado un lenguaje  $L$ , se nos presenta el problema de describirlo. Para resolverlo se pueden tomar dos caminos:

- Dar una forma de generar todas las palabras del lenguaje.
- Dar un algoritmo para demostrar que una palabra esta en el lenguaje.

# Problema

Dado un lenguaje  $L$ , se nos presenta el problema de describirlo. Para resolverlo se pueden tomar dos caminos:

- Dar una forma de generar todas las palabras del lenguaje.
- Dar un algoritmo para demostrar que una palabra esta en el lenguaje.

# Diferencias entre los Métodos

A simple vista, los dos métodos no son equivalentes. Podemos pensar por ejemplo en el lenguaje formado por los programas escritos en java que devuelven para cualquier entrada “Hola Mundo”.

# Diferencias entre los Métodos

También tenemos la otra cara de la moneda, saber reconocer no implica saber generar elementos del conjunto. Un ejemplo lo darían el lenguaje definido por los libros en español.

# La Idea detrás de una Gramática

Los lenguajes que vamos a tratar no son simples conjuntos de palabras. Las palabras están porque tienen una estructura, han sido generadas por unas reglas.

# Las Gramáticas Formales

## Definición (Gramáticas Formales)

*Una gramática formal es una cuaterna  $G = (V, \Sigma, Q_0, P)$ , donde:*

- *$V$  es un conjunto finito llamado alfabeto de símbolos no terminales o, simplemente, alfabeto de variables.*
- *$\Sigma$  es otro conjunto finito, que verifica  $V \cap \Sigma = \emptyset$  y se suele denominar alfabeto de símbolos terminales.*
- *$Q_0 \in V$  es una “variable” distinguida que se denomina símbolo inicial.*
- *$P \subseteq (V \cup \Sigma)^* \times (V \cup \Sigma)^*$  es un conjunto finito llamado conjunto de producciones (o, simplemente, sistema de reescritura).*



# Como Operar: Sistema de Transición

Para poder definir la dinámica asociada a una gramática, necesitamos asociarle un sistema de transición.

# Como Operar: Sistema de Transición

## Definición

Sea  $G = (V, \Sigma, Q_0, P)$  una gramática, definiremos el sistema de transición asociado  $(S_G, \rightarrow_G)$  dado por las propiedades siguientes:

- El espacio de configuraciones será dado por:

$$S_G := (V \cup \Sigma)^*.$$

- Dadas dos configuraciones  $s_1, s_2 \in S_G$ , decimos que  $s_1 \rightarrow_G s_2$  si se verifica la siguiente propiedad:

$$\exists x, y, \alpha, \beta \in S_G = (V \cup \Sigma)^*, \text{ tales que}$$

$$s_1 := x \cdot \alpha \cdot y, \quad s_2 := x \cdot \beta \cdot y, \quad (\alpha, \beta) \in P.$$

# Lenguaje generado por una Gramática

## Definición

*Dada una gramática  $G = (V, \Sigma, Q_0, P)$  y su espacio de configuraciones  $S_G$  se define el lenguaje generado por una gramática al conjunto de configuraciones  $s_1$  tales que:*

- $Q_0 \rightarrow_G s_1$ ,
- además  $s_1 \in \Sigma^*$ .

# Ejemplos

## Ejemplo

Consideremos la gramática:  $G = (V, \Sigma, Q_0, P)$ , donde

$$V := \{Q_0\}, \quad \Sigma := \{a, b\}, \quad P := \{(Q_0, aQ_0), (Q_0, \lambda)\}.$$

El sistema de transición tiene por configuraciones

$S := \{Q_0, a, b\}^*$  y un ejemplo de una computación sería:

$$aaQ_0bb \rightarrow aaaQ_0bb \rightarrow aaaaQ_0bb \rightarrow aaaa\lambda bb = aaaabb.$$

Nótese que las dos primeras veces hemos usado la regla de reescritura  $(Q_0, aQ_0)$  y la última vez hemos usado  $(Q_0, \lambda)$ .

# Ejemplos

## Ejemplo

*Utilizando la misma gramática podemos también estudiar el lenguaje generado por la gramática:*

$$L = \{a, aa, aaa, \dots\}.$$

*Para generar una palabra con  $n$  letras  $a$  seguidas simplemente hacemos*

$$Q_0 \rightarrow aQ_0 \rightarrow \dots \rightarrow \underbrace{a \dots a}_{n \text{ veces}} Q_0 \rightarrow \underbrace{a \dots a}_{n \text{ veces}}.$$

# Notación

Por analogía con el sistema de transición, se suelen usar la notación  $A \mapsto B$  en lugar de  $(A, B) \in P$ , para indicar una producción. Y, en el caso de tener más de una producción que comience en el mismo objeto, se suele usar  $A \mapsto B \mid C$ , en lugar de escribir  $A \mapsto B, A \mapsto C$ .

# Mas Ejemplos

## Ejemplo

Consideremos la gramática:  $G = (V, \Sigma, Q_0, P)$ , donde

$$V := \{Q_0\}, \quad \Sigma := \{a, b\}, \quad , P := \{Q_0 \mapsto aAb, aA \mapsto aaAb | \lambda\}.$$

Un ejemplo de una computación sería:

$$Q_0 \mapsto aAb \mapsto aaAb \mapsto aab.$$

Curiosamente, el lenguaje especificado también puede ser especificado por esta otra gramática:

$$V := \{Q_0\}, \quad \Sigma := \{a, b\}, \quad , P := \{Q_0 \mapsto b|aA, A \mapsto aA|b\}.$$

# Notación BNF

Es un modelo de notación que recuerda más los manuales de programación. En él, se introducen los siguientes cambios:

- Las variables  $X \in V$  se representan mediante  $\langle X \rangle$ .
- Los símbolos terminales (del alfabeto  $\Sigma$ ) no son modificados.
- El símbolo asociado a las producciones  $\mapsto$  es reemplazado por  $=$ .



# Ejemplo de Notación BNF

## Ejemplo

*La gramática que genera el lenguaje*

$$L = \{\lambda, a, aa, aaa, \dots\}$$

*se puede describir de la siguiente manera*

$$V = \{\langle Q \rangle\}, \quad \Sigma = \{a, b\},$$

*donde las producciones serían:*

$$\begin{aligned}\langle Q \rangle &= a\langle Q \rangle \\ \langle Q \rangle &= \lambda.\end{aligned}$$

# Notación EBNF

Es una extensión de la notación anterior. Básicamente, añade funcionalidad a la notación BNF, permitiendo repeticiones o diferentes opciones.

# Notación EBNF

Estas son las principales cambios con respecto a la notación BNF,

- Las variables  $X \in V$  no son modificadas.
- Los símbolos terminales (del alfabeto  $\Sigma$ ) se representan entre comillas simples.
- El símbolo asociado a las producciones  $\mapsto$  es remplazado por  $\therefore$ .
- Se introducen nuevos símbolos para representar repeticiones  $*$  (ninguna, una o mas repeticiones)  $+$  (una repetición al menos).
- $?$  indica que la expresión puede ocurrir o no.

# Notación EBNF

También tiene una representación gráfica

**A :**

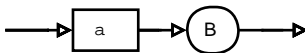


Figura: Representación A:"a"B

# Notación EBNF

**B :**

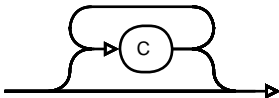


Figura: Representación  $B:C^*$

# Notación EBNF

**D :**

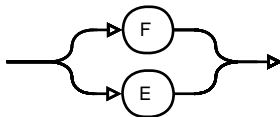


Figura: Representación D:F | E