

## Práctica 6: Lenguajes formales y gramáticas

- Una gramática posee los siguientes símbolos:
  - ▶ inicial,
  - ▶ no terminales (son estructuras intermedias), y
  - ▶ terminales (son elementos del alfabeto).

**Definición:** Gramática es una tupla:  $(N, T, P, \sigma)$  donde:

- $N$  es un conjunto finito de símbolos llamados **no terminantes**.
- $T$  es un conjunto finito de símbolos, llamados **terminantes** o **alfabeto**, tal que  $N \cap T = \emptyset$
- $P$  es un conjunto finito de **reglas de producción**, donde

$$P \subseteq ((N \cup T)^* - T^*) \times (N \cup T)^*$$

Obs.:

- ▶  $(N \cup T)^* - T^*$  es el conjunto de cadenas de no terminales y terminales que contienen al menos un no terminal. Dado que  $\lambda \notin (N \cup T)^* - T^*$  no puede haber reglas del tipo  $\lambda \rightarrow \beta$ .
- ▶  $(N \cup T)^*$  es el conjunto de cadenas sobre  $N \cup T$ .
- ▶ A una producción de la forma  $(\alpha, \beta)$  la notaremos  $\alpha \rightarrow \beta$ .
- ▶ Una regla del tipo  $\alpha \rightarrow \lambda$  recibe el nombre de **regla**  $\lambda$ .

- $\sigma \in N$  es el **símbolo inicial**

$$G_1 = (N, T, P, \sigma)$$

donde:

- ▶  $N = \{\sigma\}$
- ▶  $T = \{a, b\}$
- ▶  $P$  está dado por las reglas de producción:

$$\sigma \rightarrow a\sigma bb$$

$$\sigma \rightarrow abb$$

- ▶  $\sigma$  es el símbolo inicial

**Definición 1.** Una gramática se dice:

(a) *regular* si cada producción es de la forma:  $A \rightarrow a$  o  $A \rightarrow aB$  o  $A \rightarrow \lambda$  donde  $A, B \in N$  y  $a \in T$ ,

### Gramáticas regulares (tipo 3)

- Las gramáticas **regulares** o de **tipo 3**, también llamadas **lineales**, pueden ser clasificadas como **derechas** o **izquierdas**.
- Las reglas de producción de una **gramática regular derecha** se adhieren a las siguientes restricciones:
  - ▶ El lado izquierdo debe consistir en un solo no terminal.
  - ▶ El lado derecho está formado por un símbolo terminal, que puede estar seguido (o no) por un símbolo no terminal, o la cadena vacía.
- Es decir, las producciones de una gramática regular derecha pueden tener la forma:

$$A \rightarrow a$$

$$A \rightarrow aB$$

$$A \rightarrow \lambda$$

donde  $A, B \in N$  y  $a \in T$ .

### Gramáticas regulares (tipo 3)

- Alternativamente, en una **gramática regular izquierda** las reglas de producción son de la forma:

$$A \rightarrow a$$

$$A \rightarrow Ba$$

$$A \rightarrow \lambda$$

- Por ejemplo, reglas de la forma:

$$yW \rightarrow x$$

$$X \rightarrow xZy$$

$$YX \rightarrow WvZ$$

no están permitidas en una gramática regular. (¿Por qué?)

- Se puede convertir toda gramática regular derecha en izquierda y viceversa.

- (b) *libre* (o independiente) de contexto si cada producción es de la forma  $A \rightarrow \delta$  donde  $A \in N$  y  $\delta \in (N \cup T)^*$ .

Otra definición:

- ▶ Como en las regulares, el lado izquierdo debe consistir en un solo no terminal.
- ▶ A diferencia de las regulares, no hay restricciones sobre la forma del lado derecho.

- Es decir, las reglas de producción tienen la forma:

$$A \rightarrow \delta$$

donde  $A \in N$  y  $\delta \in (N \cup T)^*$ .

- El término *independiente del contexto* hace referencia a que, debido a que en el lado izquierdo el no terminal aparece solo, la regla se puede aplicar sin importar el contexto en que aparezca dicho no terminal.
- En contraste, la regla  $xNy \rightarrow xzy$  permite reemplazar el no terminal  $N$  por  $z$  sólo cuando se encuentre en el contexto de  $x$  e  $y$ .

## Gramáticas independientes del contexto (tipo 2)

- Dado que no hay restricciones sobre la forma de  $\delta$  en  $A \rightarrow \delta$ , podría ocurrir que  $\delta$  contenga más de un no terminal, como en  $A \rightarrow zXYz$ .

- Pero ¿cuál de los no terminales reemplazamos en el paso siguiente?

- El enfoque más común es el de la **derivación por la izquierda**, que consiste en reemplazar el no terminal situado más a la izquierda.
- Análogamente, se podría aplicar derivación por la derecha, o seguir algún otro patrón.
- Resulta ser que el orden en que se apliquen las reglas no afecta la determinación de si una cadena puede ser generada por la gramática o no.
- Esto es consecuencia de que *si existe una derivación que genera una cadena, entonces también existe una derivación por izquierda que la genera*.

- (c) *sensible al contexto* si cada producción es de la forma  $\alpha A \beta \rightarrow \alpha \delta \beta$  donde  $A \in N$ ,  $\alpha, \beta \in (N \cup T)^*$  y  $\delta \in (N \cup T)^+$ .

Otra definición:

- (d) *estructurada* por frases o irrestricta si no tiene restricciones sobre la forma de sus producciones, es decir si son de la forma

$$\alpha \rightarrow \delta \quad \text{donde} \quad \alpha \in (N \cup T)^* - T^* \quad \text{y} \quad \delta \in (N \cup T)^*$$

1. Clasifique cada una de las siguientes gramáticas (dando su tipo más restrictivo):

- a)  $T = \{a, b\}$ ,  $N = \{\sigma, A\}$ , símbolo inicial  $\sigma$ , y producciones

$$\sigma \rightarrow b\sigma, \sigma \rightarrow aA, A \rightarrow a\sigma,$$

$$A \rightarrow bA, A \rightarrow a, \sigma \rightarrow b$$

Regular.

- b)  $T = \{a, b, c\}$ ,  $N = \{\alpha, A, B\}$ , símbolo inicial  $\sigma$ , y producciones

$$\sigma \rightarrow AB, AB \rightarrow BA, A \rightarrow aA,$$

$$B \rightarrow Bb, A \rightarrow a, B \rightarrow b$$

Sensible al contexto.

- c)  $T = \{a, b\}$ ,  $N = \{\sigma, A, B\}$ , símbolo inicial  $\sigma$  y producciones:

$$\sigma \rightarrow A, \quad \sigma \rightarrow AAB, \quad Aa \rightarrow ABa, \quad A \rightarrow aa,$$

$$Bb \rightarrow ABb, \quad AB \rightarrow ABB, \quad B \rightarrow b.$$

Sensible al contexto.

d)  $T = \{a, b, c\}$ ,  $N = \{\sigma, A, B\}$ , símbolo inicial  $\sigma$ , y producciones:

$$\sigma \rightarrow BAB, \quad \sigma \rightarrow ABA, \quad A \rightarrow AB, \quad B \rightarrow BA,$$

$$A \rightarrow aA, \quad Aab, \quad B \rightarrow b.$$

Independiente de contexto (libre).