

3. Jerarquía de Chomsky

3.1 Gramáticas regulares: RGs

3.1.1 Gramáticas lineales: por la derecha, por la izquierda

Definición: Sea $G = (N, \Sigma, P, S)$ una gramática, decimos que G es:

a) lineal por la derecha: si $\forall p \in P$,

"p" es de la forma $\begin{cases} A \rightarrow \alpha B \\ A \rightarrow \alpha \end{cases}$

b) lineal por la izquierda: si $\forall p \in P$,

"p" es de la forma $\begin{cases} A \rightarrow B\alpha \\ A \rightarrow \alpha \end{cases}$

Ejemplo: $G = (\{\text{Digito}\}, \{0, 1, \dots, 9\}, \{\text{Digito} \rightarrow 0|1|1-19\}, \text{Digito})$

es lineal por la derecha.

3.1.2 Gramáticas regulares: RGs

Definición: Una gramática $G = (N, \Sigma, P, S)$ se dice regular si:

i) G es lineal por la derecha.

ii) G es lineal por la izquierda.

NOTA: En adelante identificaremos por defecto

Gramática regular \equiv Gramática lineal por la derecha

3.2 Gramáticas de contexto libre: CFGs.

3.2.1 Definición de CFG. Lenguajes de contexto CFLs: CFLs

Definición: Una gramática $G = (N, \Sigma, P, S)$ es de contexto-libre si $\forall p \in P$, " p " es de la forma $A \rightarrow \alpha$.

Definición: Una CFG $G = (N, \Sigma, P, S)$ se dice lineal si:

$$\forall p \in P \left\{ \begin{array}{l} \text{i) } p \equiv A \rightarrow uBv \\ \text{ii) } p \equiv A \rightarrow w \end{array} \right.$$

NOTA: Si $u = \epsilon$ (resp. $v = \epsilon$) es lineal por la derecha (resp. por la izquierda).

3.2.2 Árboles de derivación, fronteras.

Definición: Un árbol orientado etiquetado T es un árbol de derivación para una CFG $G(S) = (N, \Sigma, P, S)$ si:

- i) La raíz de T está etiquetada por S .
- ii) Si T_1, \dots, T_k son subárboles de los descendientes directos de la raíz de T , y la raíz de T_i está etiquetada por S_i ; entonces $S \rightarrow S_1 \dots S_k$ es una producción en P .
- iii) T_i debe ser:
 - a) Un árbol de derivación para $G(S_i)$, $G(S_i) = (N, \Sigma, P, S_i)$ si $S_i \in N$
 - b) Un nodo simple etiquetado S_i si $S_i \in \Sigma$.

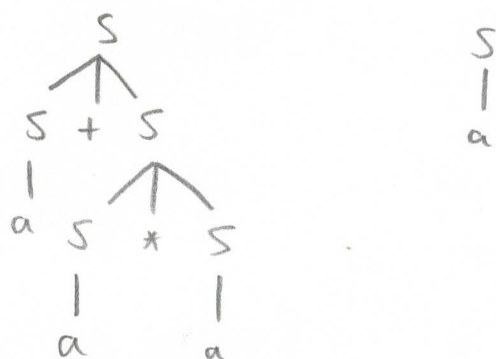
iv) En el caso de que:

a) T_1 sea el único subárbol de la raíz de T

b) T_1 está etiquetado por ε

entonces $S \rightarrow \varepsilon \in P$.

Ejemplo: Con la gramática de las expresiones aritméticas



Definición: La frontera de un árbol de derivación es la cadena obtenida concatenando las etiquetas de las hojas (los terminales del árbol) de izquierda a derecha.

Ejemplo: En el ejemplo anterior, las fronteras son respectivamente: $a + a * a$
 a

Definición: Sea T un árbol, y sean " x " e " y " dos nodos del mismo. Decimos que " y " desciende directamente de " x " si " y " es un hijo de " x " en T .

NOTACIÓN: $x \Gamma y$

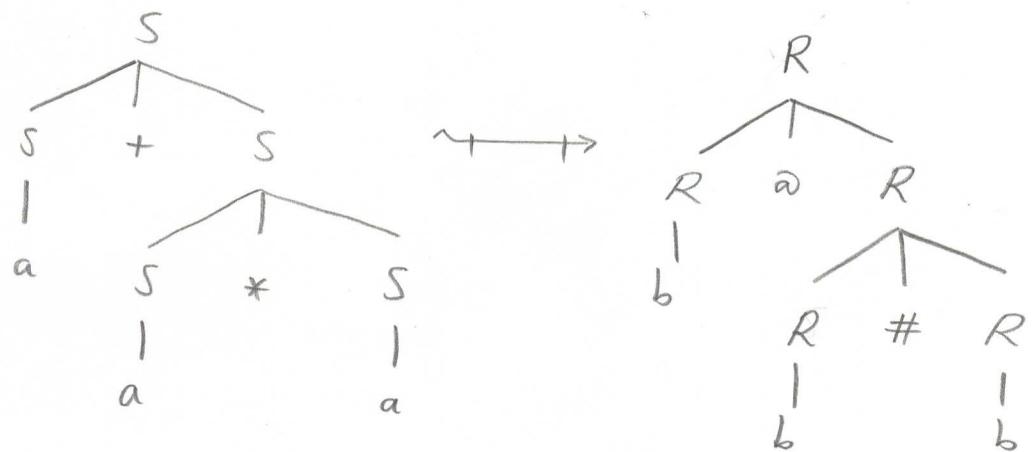
3.2.3 Isomorfismos estructurales entre árboles.

Definición: Dos árboles T, T' son estructuralmente isomórficos si $\exists T \mapsto T'$ tal que:

$$x \Gamma y \text{ (resp. } x L y) \Leftrightarrow x' \Gamma y' \text{ (resp. } x' L y')$$

NOTA: Intuitivamente T y T' son idénticos salvo por su etiquetaje.

Ejemplo:



3.2.4 Derivaciones por la derecha (canónicas) y por la izquierda.

Definición: Sea $G = (N, \Sigma, P, S)$ una CFG y sea $\alpha A \beta \Rightarrow \alpha \gamma \beta$ una derivación directa en G , diremos que se trata de una derivación por la derecha (resp. por la izquierda)

si:

- $A \rightarrow \gamma \in P$

- $\beta \in \Sigma^*$ (resp. $\alpha \in \Sigma^*$).

NOTACIÓN: Utilizaremos la notación

$$\alpha A \beta \Rightarrow_{rm} \alpha \gamma \beta \quad (\text{resp. } \alpha A \beta \Rightarrow_{lm}^* \alpha \gamma \beta).$$

En ambos casos, podemos considerar los conceptos análogos en derivaciones indirectas, que notaremos:

$$i) \quad \alpha A \beta \Rightarrow_{rm}^+ \alpha \gamma \beta \quad (\text{resp. } \alpha A \beta \Rightarrow_{lm}^* \alpha \gamma \beta)$$

$$ii) \quad \alpha A \beta \Rightarrow_{rm}^+ \alpha \gamma \beta \quad (\text{resp. } \alpha A \beta \Rightarrow_{lm}^+ \alpha \gamma \beta)$$

Ejemplo: Considerando las expresiones aritméticas, tenemos que:

$$S \xRightarrow{4}_{lm} (a+a), \text{ puesto que } S \Rightarrow_{lm} (S) \Rightarrow_{lm} (S+S) \Rightarrow_{lm} (a+S) \Rightarrow_{lm} (a+a)$$

$$S \xRightarrow{4}_{rm} (a+a), \text{ puesto que } S \Rightarrow_{rm} (S) \Rightarrow_{rm} (S+S) \Rightarrow_{rm} (S+a) \Rightarrow_{rm} (a+a)$$

NOTA: En adelante consideraremos que la notación $\alpha A \beta \Rightarrow^+ \alpha \gamma \beta$ expresa un orden prefijado de derivación (el canónico), salvo que se indique lo contrario.

3.2.5 CFGs ambiguas y no ambiguas.

Definición: Una CFG $G = (N, \Sigma, P, S)$ es ambigua si $\exists x \in L(G)$ tal que \exists dos derivaciones canónicas $S \Rightarrow^* x$.

G no ambigua

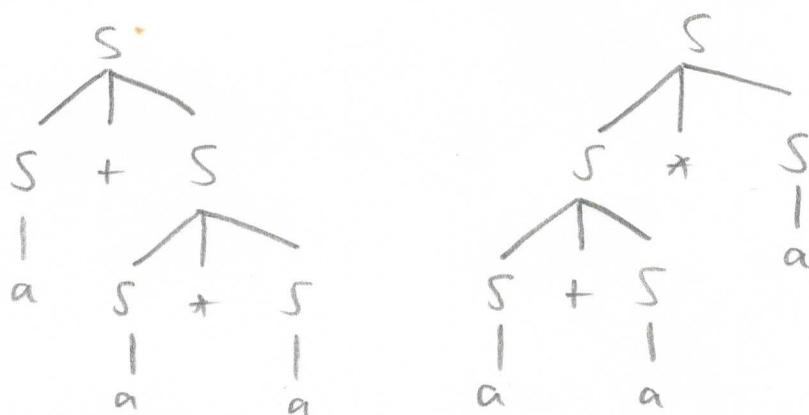
Lema: Sea $G = (N, \Sigma, P, S)$ una CFG, entonces:

G es no ambigua $\Leftrightarrow \forall T, T'$ árboles de derivación de S :
 $fr(T) = fr(T') \Rightarrow T = T'$

NOTA: $fr(T) \equiv \text{frontera}(T)$

demo trivial.

Ejemplo: La gramática anteriormente considerada para las expresiones aritméticas, es ambigua:



Definición: Un CFL L es no ambiguo si $\exists G = (N, \Sigma, P, S)$
 CFG no ambigua / $L = L(G)$.

Ejemplo: El lenguaje L de las expresiones aritméticas es no ambiguo puesto que: $L = L(G)$ donde G es una CFG no ambigua dada por las reglas:

$$\begin{array}{lll}
 E \rightarrow E + T & T \rightarrow T * F & F \rightarrow (E) \\
 \mid T & \mid F & \mid a
 \end{array}$$