Gramáticas.

U.D. Computación

DSIC - UPV

2017-18

Índice

Gramáticas Formales

Definición intuitiva

Una gramática es una forma finita de describir un lenguaje

$$G = (N, \Sigma, P, S)$$

- Se parte de un axioma S
- El objetivo es obtener palabras (elementos de Σ^*) de un lenguaje L
- Empleando reglas de reescritura (producciones, elementos de P)
- Con la ayuda de símbolos auxiliares (elementos de N)

Gramáticas Formales

Definición Formal

$$G = (N, \Sigma, P, S)$$

- N es un conjunto finito de elementos llamados *no terminales*.
- Σ es un *alfabeto*
  - $N \cap \Sigma = \emptyset$, $N \cup \Sigma = V$ (conjunto de *símbolos*)
- $P \subset V^*NV^* \times V^*$
  - Denotaremos $(\alpha, \beta) \in P$ mediante $\alpha \rightarrow \beta$
  - Si $\alpha \rightarrow \beta_1 \ldots \alpha \rightarrow \beta_n$ escribimos $\alpha \rightarrow \beta_1|\beta_2| \ldots |\beta_n$
- $S \in N$

Ejemplo

La gramática

$$G = (\{S\}, \{a, b\}, S \to aSb|\lambda, S)$$

- Define el lenguaje $L = \{a^n b^n : n \geq 0\}$
- Las palabras se generan partiendo de S y aplicando las reglas de producción
- Generación de *aabb* (por ejemplo):
  - $S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow aa\lambda bb = aabb$

Lenguaje generado por una gramática G

Dados $\alpha', \beta' \in V^*$, decimos que α' *deriva directamente* en β' ($\alpha' \Rightarrow_G \beta'$) en G si:

- $\alpha' = \gamma\alpha\delta$.
- $\beta' = \gamma\beta\delta$.
- $\alpha \rightarrow \beta \in P$

Ejemplo: en la gramática anterior $aSb \Rightarrow aaSbb$

Decimos que α *deriva* en β ($\alpha \overset{*}{\underset{G}{\Rightarrow}} \beta$) en G si existe una cadena de cero o más derivaciones directas que convierten α en β.

Ejemplo: en la gramática anterior $aSb \overset{*}{\underset{G}{\Rightarrow}} aabb$.

Lenguaje generado por una gramática G

$\alpha \in V^*$ es una *forma sentencial* si $S \overset{*}{\underset{G}{\Rightarrow}} \alpha$.
Si $\alpha \in \Sigma^*$ se dice que α es una *palabra* generada por G.

Lenguaje generado por G

$$L(G) = \{x \in \Sigma^* : S \overset{*}{\underset{G}{\Rightarrow}} x\}$$

Tipos de gramáticas: la Jerarquía de Chomsky

Imponiendo restricciones a las reglas de producción se obtienen los siguientes tipos de gramáticas:

- Tipo 3, LINEALES POR LA DERECHA (IZQUIERDA): Sus reglas son de la forma $A \rightarrow aB(A \rightarrow Ba)$ con $A \in N$, $a \in \Sigma \cup \lambda$, $B \in N \cup \lambda$.

Tipos de gramáticas: la Jerarquía de Chomsky

Imponiendo restricciones a las reglas de producción se obtienen los siguientes tipos de gramáticas:

- Tipo 3, LINEALES POR LA DERECHA (IZQUIERDA): Sus reglas son de la forma $A \to aB(A \to Ba)$ con $A \in N$, $a \in \Sigma \cup \lambda$, $B \in N \cup \lambda$.
- Tipo 2, INCONTEXTUALES o INDEP. DE CONTEXTO: Sus reglas son de la forma $A \to \alpha$ con $A \in N$, $\alpha \in V^*$.

Tipos de gramáticas: la Jerarquía de Chomsky

Imponiendo restricciones a las reglas de producción se obtienen los siguientes tipos de gramáticas:

- Tipo 3, LINEALES POR LA DERECHA (IZQUIERDA): Sus reglas son de la forma $A \rightarrow aB(A \rightarrow Ba)$ con $A \in N$, $a \in \Sigma \cup \lambda$, $B \in N \cup \lambda$.
- Tipo 2, INCONTEXTUALES o INDEP. DE CONTEXTO: Sus reglas son de la forma $A \rightarrow \alpha$ con $A \in N$, $\alpha \in V^*$.
- Tipo 1, CONTEXTUALES o DEP. DE CONTEXTO: Reglas: $\gamma A\delta \rightarrow \gamma\alpha\delta$ con $A \in N$, $\gamma, \delta \in V^*$, $\alpha \in V^+$. (se excluye $S \rightarrow \lambda$ a condición de...)

Tipos de gramáticas: la Jerarquía de Chomsky

Imponiendo restricciones a las reglas de producción se obtienen los siguientes tipos de gramáticas:

- Tipo 3, LINEALES POR LA DERECHA (IZQUIERDA): Sus reglas son de la forma $A \to aB(A \to Ba)$ con $A \in N$, $a \in \Sigma \cup \lambda$, $B \in N \cup \lambda$.
- Tipo 2, INCONTEXTUALES o INDEP. DE CONTEXTO: Sus reglas son de la forma $A \to \alpha$ con $A \in N$, $\alpha \in V^*$.
- Tipo 1, CONTEXTUALES o DEP. DE CONTEXTO: Reglas: $\gamma A\delta \to \gamma\alpha\delta$ con $A \in N$, $\gamma, \delta \in V^*$, $\alpha \in V^+$. (se excluye $S \to \lambda$ a condición de...)
- Tipo 0, SIN RESTRICCIONES:

Jerarquía de Chomsky

- Se dice que un lenguaje es de Tipo i ($i = 0, 1, 2, 3$) si es generable por una gramática de tipo i.
- Llamamos $\mathcal{L}_i$ a la familia de lenguajes de tipo i.

Jerarquía de Chomsky

Se puede demostrar que : $\mathcal{L}_3 \subset \mathcal{L}_2 \subset \mathcal{L}_1 \subset \mathcal{L}_0$

Ejemplos

Según el tipo de gramática que los genera, los lenguajes se denominan Regulares, Incontextuales,...

Ejemplos

- La gramática cuyas reglas son $S \rightarrow aA|bA|\lambda, A \rightarrow aS|bS$:
  -Es regular y
  -Genera el lenguaje de todas las palabras de longitud par sobre $\{a, b\}$
- La gramática cuyas reglas son $S \rightarrow aSb|\lambda$:
  -Es Incontextual y
  -Genera el lenguaje de todas las palabras de la forma $a^n b^n$ sobre $\{a, b\}$

Equivalencia entre Autómatas Finitos y Gramáticas Regulares

Si L es un lenguaje regular, entonces L es aceptado por un autómata finito

Sea $G = (N, \Sigma, P, S)$ una gramática lineal por la derecha tal que $L(G) = L$. Construimos un AFλ $A = (Q, \Sigma, \delta, q_0, F)$ tal que $L(A) = L(G)$.
$Q = N \cup \{X\}, X \notin N, \quad q_0 = S, \quad F = \{X\}$.

- $\forall (A_i \rightarrow a_j A_k) \in P$ se define $A_k \in \delta(A_i, a_j), \quad A_i, A_k \in N, a_j \in \Sigma \cup \{\lambda\}$.
- $\forall (A_i \rightarrow a_j) \in P$ se define $X \in \delta(A_i, a_j), \quad A_i \in N, a_j \in (\Sigma \cup \{\lambda\})$.
- $\forall a \in (\Sigma \cup \{\lambda\})$ se define $\delta(X, a) = \emptyset$.

Ejemplo

Gramáticas

U.D. Computación

Gramáticas Formales.

Tipos de gramáticas: la Jerarquía de Chomsky

Equivalencia entre AFs y Gramáticas Regulares

Gramáticas Incontextuales

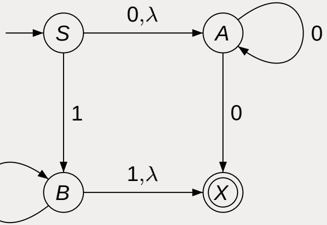
Obtención del *AF* equivalente a:

$$
\begin{aligned}
S &\rightarrow A|0A|1B \\
A &\rightarrow 0A|0 \\
B &\rightarrow 1B|1|\lambda
\end{aligned}
$$

Ejemplo

Obtención del *AF* equivalente a:

$$\begin{aligned} S &\rightarrow A|0A|1B \\ A &\rightarrow 0A|0 \\ B &\rightarrow 1B|1|\lambda \end{aligned}$$



Equivalencia entre Autómatas Finitos y Gramáticas Regulares

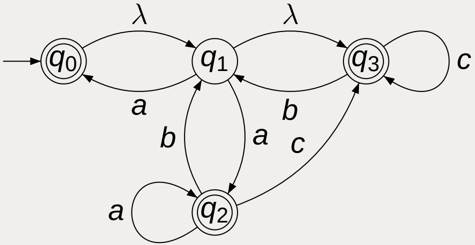
Si L es un lenguaje aceptado por un autómata finito, entonces L es regular

Sea $A = (Q, \Sigma, \delta, q_0, F)$ un AFλ tal que $L(A) = L$.
Construimos una gramática lineal por la derecha $G = (N, \Sigma, P, S)$ que genere L.
$N = Q, \quad S = q_0$.

- $\forall q' \in \delta(q, a)$, se define $(q \to aq') \in P, \quad q, q' \in Q, a \in (\Sigma \cup \{\lambda\})$.
- $\forall q \in F$ se define $(q \to \lambda) \in P$.

Ejemplo

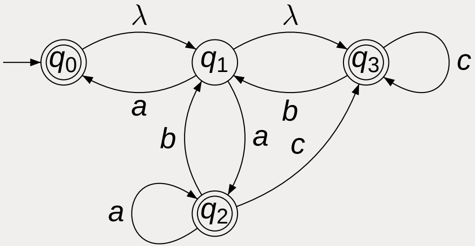
Obtención de la gramática lineal por la dch. equivalente al *AF* de la figura.



Ejemplo

Obtención de la gramática lineal por la dch. equivalente al *AF* de la figura.
Solución:

$$\begin{array}{rl} q_0 \to q_1|\lambda & q_1 \to aq_0|aq_2|q_3 \\ q_2 \to bq_1|aq_2|cq_3|\lambda & q_3 \to bq_1|cq_3|\lambda \end{array}$$



Gramáticas Incontextuales

Sea $G = (N, \Sigma, P, S)$ una gramática incontextual, es decir, todas sus reglas son de la forma $A \to \alpha$ con $A \in N$, $\alpha \in V^*$.

Árbol de derivación

Un árbol de derivación de la gramática G es un árbol tal que:

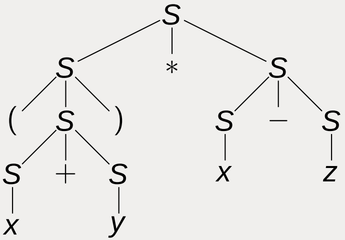
- Los nodos estan etiquetados por símbolos de $V \cup \{\lambda\}$, y la raiz esta etiquetada por S
- Si un nodo es interior, entonces esta etiquetado por un símbolo de N
- Si el nodo n esta etiquetado por A y sus sucesores, $n_1, n_2, \cdots, n_k$ estan etiquetados por $X_1, X_2, \cdots, X_k$ respectivamente, entonces $A \to X_1, X_2, \cdots, X_k \in P$
- Si un nodo tiene asignada la etiqueta λ entonces es hoja y el único sucesor de su predecesor

Ejemplo

Expresiones algebraicas sintácticamente correctas sobre las variables x, y y z:

$$S \rightarrow x \mid y \mid z \mid S + S \mid S - S \mid S * S \mid S/S \mid (S)$$

Un árbol de derivación para $(x + y) * x - z$



Forma Normal de Chomsky (FNC)

Una gramática incontextual $G = (N, \Sigma, P, S)$ está en FNC cuando todas sus reglas son de la forma:

$$A \to BC \qquad \text{con } A, B, C \in N$$
$$A \to a \quad \text{con } A \in N, \quad a \in \Sigma$$

FNC de la gramática
$S \to x \mid y \mid z \mid S + S \mid S - S \mid S * S \mid S/S \mid (S)$

$$\begin{array}{lll}
S \to\ x \mid y \mid z \mid SW_1 \mid SW_2 \mid SW_3 \mid SW_4 \mid X_oW_5 & & \\
W_1 \to\ X_+S & W_2 \to\ X_-S & W_3 \to\ X_*S \\
W_4 \to\ X_/S & W_5 \to\ SX_t & X_+ \to\ + \\
X_- \to\ - & X_* \to\ * & X_/ \to\ / \\
X_o \to\ (& X_t \to\ ) &
\end{array}$$