Slide 1

Gramáticas formales

$G = (\Sigma_T, \Sigma_A, P, S)$ donde

$$\begin{array}{rcll}
\Sigma_T & : & \text{alfabeto } \textit{terminal} & a, b, (,), + \dots \\
\Sigma_A & : & \text{alfabeto } \textit{auxiliar} & S, A, B, \dots \\
 & & \Sigma_T \cap \Sigma_A = \emptyset & \\
 & & \Sigma = \Sigma_T \cup \Sigma_A & X, Y \dots \\
P & \subset & \Sigma^+ \times \Sigma^* \text{ conjunto } \textbf{finito} \text{ de } \textit{reglas} & \alpha \to \beta \\
S & \in & \Sigma_A \text{ símbolo } \textit{inicial} &
\end{array}$$

$\Sigma^* : \alpha, \beta \dots$

$\Sigma_T^* : x, y, z, u, v, \dots$

$$\left.\begin{array}{l}
\alpha \to \beta_1 \\
\alpha \to \beta_2 \\
\dots \\
\alpha \to \beta_n
\end{array}\right\} \begin{array}{l}\text{se abrevia:} \\ \alpha \to \beta_1, \mid \beta_2, \dots \mid \beta_n\end{array}$$

Slide 2

Gramáticas formales. Ejemplos

$$G_1 \left\{\begin{array}{lcl} S & \to & 000S111 \\ 0S1 & \to & 01 \end{array}\right. \qquad G_2 \left\{\begin{array}{lcl} S & \to & 000S111 \\ 0S1 & \to & \varepsilon \end{array}\right.$$

$$G_3 \left\{\begin{array}{lcl} S & \to & 000S111 \\ S & \to & \varepsilon \end{array}\right. \qquad \begin{array}{l} G_4 \left\{ S \to aSbS \mid bSaS \mid \varepsilon \right. \\ G_5 \left\{ S \to S(S)S \mid \varepsilon \right. \end{array}$$

$$G_6 \left\{\begin{array}{lcl} S & \to & aSBC \mid aBC \\ CB & \to & BC \\ aB & \to & ab \\ bB & \to & bb \\ bC & \to & bc \\ cC & \to & cc \end{array}\right. \qquad \begin{array}{l} G_7 \left\{\begin{array}{lcl} S & \to & aS \mid bB \\ B & \to & bB \mid b \end{array}\right. \\ G_8 \left\{ S \to aSa \mid bSb \right. \\ G_9 \left\{\begin{array}{lcl} S & \to & aSa \mid bSb \\ B & \to & bB \mid b \end{array}\right. \end{array}$$

Slide 3

Derivaciones

Si $\alpha \rightarrow \beta$ es una regla de G, y $\gamma, \xi \in \Sigma^*$, se dice que $\gamma\alpha\xi$ *produce directamente* $\gamma\beta\xi$, (o que la segunda *deriva directamente* de la primera) *en* G, y se escribe

$$\gamma\alpha\xi \Rightarrow_{G} \gamma\beta\xi$$

$\alpha \;\Rightarrow^+\; \gamma$ es decir que $\alpha \;=\; \gamma_1 \Rightarrow\; \gamma_2 \Rightarrow\; \ldots \Rightarrow\; \gamma_n = \gamma$ para $n > 1$

$\alpha \;\Rightarrow^*\; \gamma$ es decir que $\alpha \Rightarrow^+ \gamma$ ó $\alpha = \gamma$ o sea

$$\alpha \;=\; \gamma_1 \Rightarrow\; \gamma_2 \Rightarrow\; \ldots \Rightarrow\; \gamma_n = \gamma \text{ para } n \geq 1$$

$\alpha, \beta, \gamma \in \Sigma^*$:
- B $\quad \alpha \Rightarrow^* \alpha$
- P $\quad$ Si $\alpha \;\Rightarrow^*\; \beta$ y $\beta \Rightarrow \gamma$, entonces $\alpha \;\Rightarrow^*\; \gamma$

Slide 4

Derivaciones. Ejemplos

G_1:

$$\overline{S}01S \Rightarrow \underline{00\overline{0S1}11}01S \Rightarrow 00011101S$$

$$S \Rightarrow 000S111 \Rightarrow 000000S111111 \Rightarrow 000000111111$$

G_4:

$$\overline{S}Sb \Rightarrow \underline{aSb\overline{S}}Sb \Rightarrow a\overline{S}bSb \Rightarrow ab\overline{S}b \Rightarrow abb$$

$$S \Rightarrow a\overline{S}bS \Rightarrow a\underline{bSa\overline{S}}bS \Rightarrow abSab\overline{S} \Rightarrow abSab$$

$$S \Rightarrow a\overline{S}bS \Rightarrow a\underline{bSa\overline{S}}bS \Rightarrow abSab\overline{S} \Rightarrow abSab \Rightarrow abab$$

G_6:

$$\begin{aligned}
S \quad &\Rightarrow \quad \underline{a\overline{S}BC} \Rightarrow a\underline{a\overline{S}BC}BC \Rightarrow aa\underline{aB\overline{C}}\overline{B}CBC \\
&\Rightarrow \quad aaaB\underline{BC}\overline{CB}C \Rightarrow aaaBB\overline{C}\underline{\overline{B}C}C \Rightarrow aa\overline{aB}B\underline{BC}CC \\
&\Rightarrow \quad aa\underline{a\overline{b}}\overline{B}BCCC \Rightarrow aaa\underline{b\overline{b}}\overline{B}CCC \Rightarrow aaab\underline{b\overline{b}}\overline{C}CC \\
&\Rightarrow \quad aaabb\underline{b\overline{c}}\overline{C}C \Rightarrow aaabbb\underline{c\overline{c}}\overline{C} \Rightarrow aaabbbc\underline{cc}
\end{aligned}$$

Slide 5

Lenguaje generado por una gramática

forma sentencial (**forma de frase**):

$$\alpha \in \Sigma^* \,/\, S \Rightarrow^* \alpha$$

sentencia (**frase**): (forma sentencial en Σ_T^*)

$$x \in \Sigma_T^* \,/\, S \Rightarrow^* x$$

lenguaje generado por G: conjunto de sentencias de G:

$$L(G) = \{x \in \Sigma_T^* \ / \ S \Rightarrow^+ x\}$$

$L(G_1) = \{0^{3n}1^{3n} \ / \ n \geq 1\}$	$L(G_5) =$ paréntesis equilibrados
$L(G_2) = \{0^{3n+2}1^{3n+2} \ / \ n \geq 0\}$	$L(G_6) = \{a^n b^n c^n \ / \ n \geq 1\}$
$L(G_3) = \{0^{3n}1^{3n} \ / \ n \geq 0\}$	$L(G_7) = a^* bbb^*$
$L(G_4) = \{x \in (a\|b)^+ \ / \ \|x\|_a = \|x\|_b\}$	$L(G_8) = L(G_9) = \emptyset$

Slide 6

$$L(G_4) = \{x \in (a|b)^+ \ / \ |x|_a = |x|_b\}$$

$\subseteq$: B Si $S \Rightarrow^1 x$, $|x|_a = 0 = |x|_b$ porque $x = \varepsilon$

H Si $n \geq 0$ y $S \Rightarrow^m x$ con $m \leq n$, entonces $|x|_a = |x|_b$

P Sea x tal que $S \Rightarrow^{n+1} x$. La derivación debe ser
$S \Rightarrow aSbS \Rightarrow^n ax_1bx_2 = x$ ó $S \Rightarrow bSaS \Rightarrow^n bx_1ax_2 = x$
con $S \Rightarrow^* x_1$ y $S \Rightarrow^* x_2$ en menos de n pasos (entre ambas derivaciones, n pasos, y ambas tienen al menos uno).
Por [H], $|x_1|_a = |x_1|_b$ y $|x_2|_a = |x_2|_b$.
$|x|_a = |x_1|_a + |x_2|_a + 1$ y $|x|_b = |x_1|_b + |x_2|_b + 1$
luego $|x|_a = |x|_b$

Conc. Si $S \Rightarrow^+ x$ en cualquier número de pasos, $|x|_a = |x|_b$

$\supseteq$: B Si $|x|_a = |x|_b$ y $|x| = 0$, $S \Rightarrow^* x$, porque $x = \varepsilon$ (regla 3).

H Si $n \geq 0$ y $|x|_a = |x|_b$ con $|x| = m \leq n$, entonces $S \Rightarrow^* x$

P Sea x con $|x| = n + 1$ tal que $|x|_a = |x|_b$.

Slide 7

Sea y el prefijo de x más corto que verifique $|y|_a = |y|_b$.
Siempre lo hay (a lo peor, la propia x).
x debe ser $x = ax_1$ ó $x = bx_1$ (porque no es ε).
En el primer caso, $y = ax_1b$ (porque antes del último símbolo de y, había más aes que bes y en y se han equilibrado)
$|x_1|_a = |x_1|_b$ (porque $|x_1|_a = |y|_a - 1 \ldots$)
$x = ax_1bx_2$ (x_2 es el sufijo restante de x)
$|x_2|_a = |x_2|_b$ (porque $|x_2|_a = |x|_a - |y|_a \ldots$)
$|x_1|, |x_2| < n$ (porque $|x_1| + |x_2| = |x| - 2 = n - 1$)
Por [H], $S \Rightarrow^* x_1$ y $S \Rightarrow^* x_2$
Luego $S \Rightarrow aSbS \Rightarrow^* ax_1bS \Rightarrow^* ax_1bx_2 = x$
(El segundo caso es similar)

Conc. Si $|x|_a = |x|_b$, sea cual sea la longitud de x, $S \Rightarrow^+ x$

Slide 8

Gramáticas independientes del contexto

Todas las reglas son de la forma $A \to \alpha$ con $\alpha \in \Sigma^*$ (Ej. $G_{3,4,5,7,8,9}$)

Árbol de derivación

Raíz: símbolo inicial (S)

Nodos: símbolos (de Σ) ó ε

Nodos interiores: símbolos de Σ_A (auxiliares)

Si un nodo está etiquetado con A y las etiquetas de los nodos hijos son $X_1, X_2, \ldots, X_p$, en ese orden, la regla $A \to X_1X_2 \ldots X_p$ está en P

Si un nodo está etiquetado con A y tiene un sólo hijo etiquetado con ε, la regla $A \to \varepsilon$ está en P

Resultado: etiquetas de las hojas, leídas de izquierda a derecha

Slide 9

Gramáticas i.c.: derivaciones y árboles

G_4:

$S \Rightarrow aSbS \Rightarrow abSaSbS$

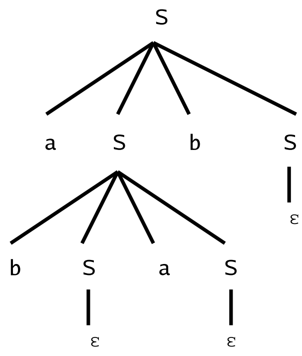
$\Rightarrow abSabS \Rightarrow abSab \Rightarrow abab$

$S \underset{mi}{\Rightarrow} aSbS \underset{mi}{\Rightarrow} abSaSbS$

$\underset{mi}{\Rightarrow} abaSbS \underset{mi}{\Rightarrow} ababS \underset{mi}{\Rightarrow} abab$

$S \underset{md}{\Rightarrow} aSbS \underset{md}{\Rightarrow} aSb$

$\underset{md}{\Rightarrow} abSaSb \underset{md}{\Rightarrow} abSab \underset{md}{\Rightarrow} abab$



Slide 10

Gramáticas i.c.: derivaciones laterales

derivación más a la izquierda:

$$xA\beta \underset{mi}{\Rightarrow} x\alpha\beta \qquad \text{(regla } A \rightarrow \alpha\text{)}$$

derivación más a la derecha:

$$\beta Ax \underset{md}{\Rightarrow} \beta\alpha x \qquad \text{(regla } A \rightarrow \alpha\text{)}$$

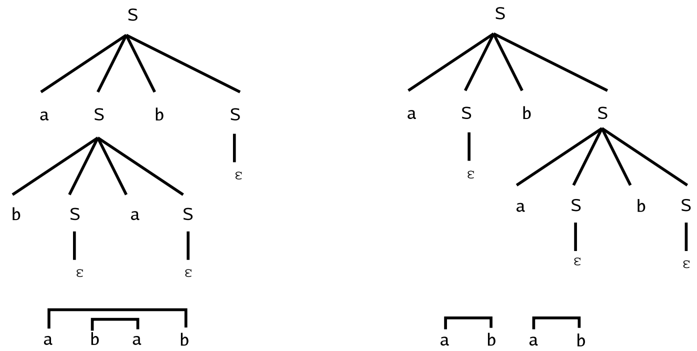
árbol de derivación $\leftrightarrow$ derivación más a la izquierda (**una**)

árbol de derivación $\leftrightarrow$ derivación más a la derecha (**una**)

árbol de derivación $\leftrightarrow$ varias derivaciones

Slide 11

Ambigüedad



$S \underset{mi}{\Rightarrow} aSbS \underset{mi}{\Rightarrow} abSaSbS$
$\underset{mi}{\Rightarrow} abaSbS \underset{mi}{\Rightarrow} ababS \underset{mi}{\Rightarrow} abab$

$S \underset{mi}{\Rightarrow} aSbS \underset{mi}{\Rightarrow} abS$
$\underset{mi}{\Rightarrow} abaSbS \underset{mi}{\Rightarrow} ababS \underset{mi}{\Rightarrow} abab$

Slide 12

Ambigüedad II

$$\left\{ S \;\rightarrow\; aSbS \,|\, bSaS \,|\, \varepsilon \right.$$

$$\left\{ \begin{array}{lcl} S & \rightarrow & aB \,|\, bA \,|\, aBS \,|\, bAS \\ A & \rightarrow & bAA \,|\, a \\ B & \rightarrow & aBB \,|\, b \end{array} \right.$$

$$\left\{ S \;\rightarrow\; S(S)S \,|\, \varepsilon \right.$$

$$\left\{ S \;\rightarrow\; (S)S \,|\, \varepsilon \right.$$

$$\left\{ \begin{array}{lcl} E & \rightarrow & E+E \,|\, E-E \\ & & |\, E*E \,|\, E/E \\ & & |\, (E) \,|\, \textbf{núm} \end{array} \right.$$

$$\left\{ \begin{array}{lcl} E & \rightarrow & E+T \,|\, E-T \,|\, T \\ T & \rightarrow & T*F \,|\, T/F \,|\, F \\ F & \rightarrow & (E) \,|\, \textbf{núm} \end{array} \right.$$

$$\left\{ \begin{array}{ll} S \rightarrow & \textbf{if } C \textbf{ then } S \\ & |\ \textbf{if } C \textbf{ then } S \textbf{ else } S \\ & |\, O \end{array} \right.$$

$$\left\{ \begin{array}{ll} S \rightarrow & P \,|\, I \\ P \rightarrow & \textbf{if } C \textbf{ then } P \textbf{ else } P \,|\, O \\ I \rightarrow & \textbf{if } C \textbf{ then } S \\ & |\ \textbf{if } C \textbf{ then } P \textbf{ else } I \end{array} \right.$$

Slide 13

Ambigüedad III

x es **ambigua** para G si (def)
existe más de un árbol de derivación en G con resultado x

x es **ambigua** para G si (def. equivalente)
existe más de una derivación m.i. en G con resultado x

G es **ambigua** si genera alguna sentencia ambigua para G

$G_1 \simeq G_2$ si (def) $L(G_1) = L(G_2)$

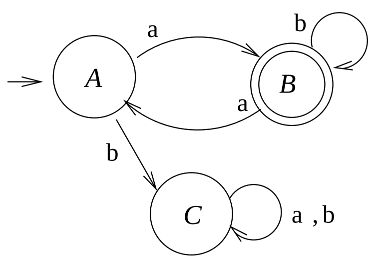
L es **inherentemente ambiguo** si (def)
toda gramática G tal que $L(G) = L$ es ambigua.

Existen lenguajes inherentemente ambiguos. Ej:

$$L = \{a^n b^n c^m d^m \,/\, n, m \geq 1\} \cup \{a^n b^m c^m d^n \,/\, n, m \geq 1\}\}$$

Slide 14

Gramáticas i.c. vs lenguajes regulares



$S \longrightarrow aB \mid bC$
$B \longrightarrow aS \mid bB \mid \varepsilon$
$C \longrightarrow aC \mid bC$

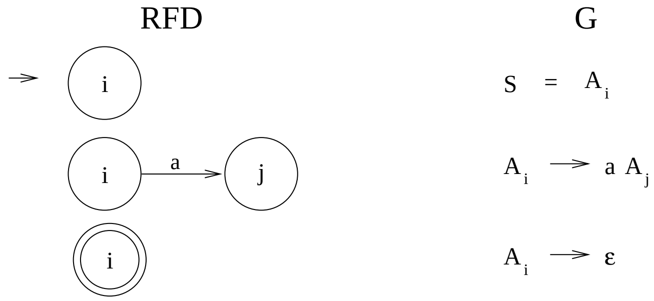
$B \longrightarrow Aa \mid Bb$
$A \longrightarrow Ba \mid \varepsilon$
$C \longrightarrow Ab \mid Ca \mid Cb$

Slide 15

Gramáticas i.c. vs lenguajes regulares

L es **independiente de contexto** si (def) existe una gramática independiente de contexto G tal que $L(G) = L$

Todo lenguaje regular es independiente de contexto



Slide 16

Eliminación de símbolos y reglas inútiles I

$A \to A$ es innecesaria

regla inútil : no existe ninguna derivación $S \Rightarrow^* x$ que la use

A **útil** : existe una derivación de la forma $S \Rightarrow^* \alpha A \beta \Rightarrow^* x$

A **terminable**: existe una derivación $A \Rightarrow^* x$

A **accesible**: existe una derivación $S \Rightarrow^* \alpha A \beta$

Si A es útil, entonces es accesible y terminable.

Si en G todo símbolo auxiliar es terminable, entonces todo símbolo accesible es útil.

Algoritmo de eliminación de símbolos inútiles:

1 Eliminar símbolos no terminables y reglas en las que aparezcan

2 Eliminar símbolos inaccesibles y reglas en las que aparezcan

Slide 17

Símbolos terminables:

marcar A si $A \to x$ con $x \in \Sigma_T^*$
repetir
    marcar B si $B \to \alpha$ y α es de terminales y/o marcados
hasta no realizar ninguna marca nueva

Símbolos accesibles:

marcar S
marcar todos los auxiliares de consecuentes de S
repetir
    para todos los auxiliares A marcados
        marcar todos los auxiliares de consecuentes de A
hasta no realizar ninguna marca nueva

Slide 18

$$\begin{array}{llll} S \to & E+T & G \to & G \mid GG \mid F \\ E \to & E \mid Z+F \mid T & T \to & T*i \mid i \\ F \to & F \mid FP \mid P & Q \to & E \mid E+F \mid T \mid Z \\ P \to & G & Z \to & i \end{array}$$

Terminables: $T, Z; E, Q; S;$ (Accesibles: $S; E, T; Z, F; P; G;$)

$$\begin{array}{llll} S \to & E+T & T \to & T*i \mid i \\ E \to & T & Q \to & E \mid T \mid Z \end{array} \qquad Z \to \quad i$$

Accesibles: $S; E, T;$

$$S \to E+T \qquad E \to T \qquad T \to T*i \mid i$$

Slide 19

Eliminación de reglas ε

A **anulable**: existe una derivación de la forma $A \Rightarrow^* \varepsilon$

Gramática **sin** ε: no tiene reglas ε o
sólo tiene $S \to \varepsilon$ y S no aparece en ningún consecuente.

Algoritmo de eliminación de reglas ε:

ConstruirP' (a partir de P):
Para cada regla $A \to \alpha_0 B_1 \alpha_1 B_2 \alpha_2 \ldots B_k \alpha_k$
con B_i anulable y α_j sin símbolos anulables
añadir a P' las reglas
$A \to \alpha_0 X_1 \alpha_1 X_2 \alpha_2 \ldots X_k \alpha_k$
para todas las combinaciones de $X_i = B_i$ ó ε
sin añadir nunca la regla $A \to \varepsilon$
Si S es anulable,
añadir como nuevo símbolo inicial S'
y las reglas $S' \to S \mid \varepsilon$

Slide 20

Símbolos anulables:

marcar A si $A \to \varepsilon$
repetir
marcar B si $B \to \alpha$ y α es de marcados
hasta no realizar ninguna marca nueva

$$\left\{\begin{array}{ll} S \to & A \mid BCa \mid aD \\ A \to & aAb \mid c \\ B \to & CD \mid b \\ C \to & cC \mid \varepsilon \\ D \to & aDd \mid \varepsilon \end{array}\right. \qquad \left\{\begin{array}{ll} S \to & A \mid BCa \mid Ba \mid Ca \mid a \mid aD \\ A \to & aAb \mid c \\ B \to & CD \mid C \mid, D \mid b \\ C \to & cC \mid c \\ D \to & aDd \mid ad \end{array}\right.$$

$$\left\{ S \to aSbS \mid \varepsilon \right. \qquad \left\{\begin{array}{ll} S' \to & S \mid \varepsilon \\ S \to & aSbS \mid abS \mid aSb \mid ab \end{array}\right.$$

Slide 21

Eliminación de reglas simples

$A \to B$ es (def.) **simple** o **de redenominación**

Entrada: una gramática *sin* ε

Para cada A construir $N_A = \{B \in \Sigma_A \,/\, A \Rightarrow^* B\}$

Para cada A

  Para cada $B \in N_A$

    Si $B \to \alpha \in P$ y no es simple, añadir $A \to \alpha$ a P'

$$\left\{\begin{array}{l} E \to E + T \,|\, T \\ T \to T * F \,|\, F \\ F \to (E) \,|\, \mathbf{n} \end{array}\right. \quad \begin{array}{l} N_E = \{E, T, F\} \\ N_T = \{T, F\} \\ N_F = \{F\} \end{array} \quad \left\{\begin{array}{l} E \to E + T \,|\, T * F \,|\, (E) \,|\, \mathbf{n} \\ T \to T * F \,|\, (E) \,|\, \mathbf{n} \\ F \to (E) \,|\, \mathbf{n} \end{array}\right.$$

Slide 22

Gramáticas regulares

Por la derecha: $A \to aB \quad A \to a$

Por la izquierda: $A \to Ba \quad A \to a$

$$\left\{\begin{array}{ll} S \to & aB \,|\, a \\ B \to & aS \,|\, bB \,|\, b \end{array}\right. \qquad \left\{\begin{array}{ll} B \to & Aa \,|\, Bb \,|\, a \\ A \to & Ba \end{array}\right.$$

Toda gramática regular por la izquierda genera un lenguaje regular.

Toda gramática regular por la derecha genera un lenguaje regular.

gramáticas	$\longleftrightarrow$	**lenguajes**
regulares	$\longleftrightarrow$	**regulares**
(por la izquierda)		
independientes del contexto	$\longleftrightarrow$	**independientes del contexto**