$L_{pal} = \{w \in \{0,1\}^* | w \text{ es palíndromo}\}$

No es un LR por el lema del bombeo

Definición recursiva:

Base: $\varepsilon, 0$ y 1 son palíndromos

Inductivo: si w es un palíndromo, entonces $0w0$ y $1w1$ también lo son

Base: $P \longrightarrow \varepsilon$

$P \longrightarrow 0$

$P \longrightarrow 1$

Inductivo: $P \longrightarrow 0P0$

$P \longrightarrow 1P1$

Gramática

1) $P \longrightarrow \varepsilon$

2) $P \longrightarrow 0$

3) $P \longrightarrow 1$

4) $P \longrightarrow 0P0$

5) $P \longrightarrow 1P1$

$$P \longrightarrow \varepsilon \mid 0 \mid 1 \mid 0P0 \mid 1P1$$

$$P \overset{4)}{\Rightarrow} 0P0 \overset{5)}{\Rightarrow} 01P10 \overset{2)}{\Rightarrow} 01010$$

derivación

$$P \overset{4)}{\Rightarrow} 0P0 \overset{4)}{\Rightarrow} 00P00 \overset{5)}{\Rightarrow} 001P100 \overset{4)}{\Rightarrow} 0010P0100 \overset{1)}{\Rightarrow} 0010\varepsilon0100 \Rightarrow 00100100$$

$$P \overset{5)}{\Rightarrow} 1P1 \overset{5)}{\Rightarrow} 11P11 \overset{1)}{\Rightarrow} 11\varepsilon11 \Rightarrow 1111$$

$$P \overset{1)}{\Rightarrow} \varepsilon$$

Gramáticas Libres de Contexto (GLC)
Gramáticas Independientes de Contexto (GIC)

Definición:

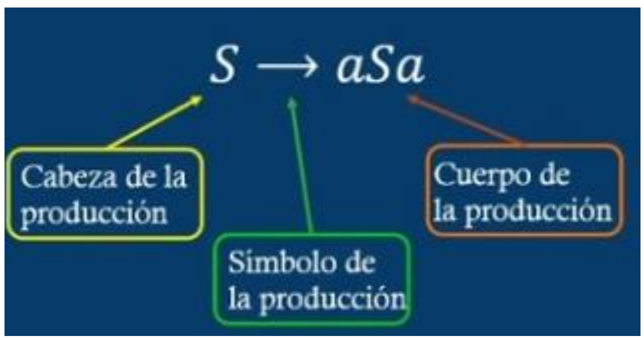
Una GLC es una 4-tupla $G = (V, T, P, S)$ donde:

i) V es un conjunto finito de variables comúnmente llamado conjunto de símbolos no terminales.

ii) T es un conjunto finito ($T \neq V$) de símbolos que forman las cadenas del lenguaje que se esta definiendo. Cada símbolo es llamado terminal. T es llamado alfabeto terminal o alfabeto de símbolos terminales.

iii) P es un conjunto finito de producciones o reglas que representan la definición recursiva del lenguaje.

iv) $S \in V$ es la variable inicial



$G_{pal} = (\{P\}, \{0,1\}, A, P)$

donde A es el conjunto de producciones:

- $P \longrightarrow \varepsilon \mid 0 \mid 1 \mid 0P0 \mid 1P1$

Ejemplo 1:

Expresión de un lenguaje de programación con los operadores $+$ y $*$
Identificadores son cadenas formadas de $a's, b's, 0's$ y $1's$ que comienzan con a o b.

$(a01b + bba * b1) * a1$ $\quad$ $(ab + a) * bb1 + a3$ $\quad$ $b + a0 + b * ba2$

$G = (\{E, I\}, T, P, E)$ $\quad$ E: Expresiones $\quad$ I: Identificadores $\quad$ con $T = \{+, *, (,), a, b, 0,1\}$

P:

$$E \longrightarrow \overset{1)}{I} \mid \overset{2)}{E + E} \mid \overset{3)}{E * E} \mid \overset{4)}{(E)}$$

$$I \longrightarrow \underset{5)}{a} \mid \underset{6)}{b} \mid \underset{7)}{Ia} \mid \underset{8)}{Ib} \mid \underset{9)}{I0} \mid \underset{10)}{I1}$$

P:

$$E \longrightarrow \overset{1)}{I} \mid \overset{2)}{E + E} \mid \overset{3)}{E * E} \mid \overset{4)}{(E)}$$

$$I \longrightarrow \underset{5)}{a} \mid \underset{6)}{b} \mid \underset{7)}{Ia} \mid \underset{8)}{Ib} \mid \underset{9)}{I0} \mid \underset{10)}{I1}$$

$$E \overset{2)}{\Longrightarrow} E + E \overset{4)}{\Longrightarrow} (E) + E \overset{3)}{\Longrightarrow} (E * E) + E \overset{1)}{\Longrightarrow} (I * E) + E \overset{5)}{\Longrightarrow} (a * E) + E \overset{1)}{\Longrightarrow} (a * I) + E$$
$$\overset{6)}{\Longrightarrow} (a * b) + E \overset{1)}{\Longrightarrow} (a * b) + I \overset{9)}{\Longrightarrow} (a * b) + I0 \overset{10)}{\Longrightarrow} (a * b) + I10 \overset{5)}{\Longrightarrow} (a * b) + a10$$

Derivaciones

Sea $\alpha A \beta$ una cadena de terminales y variables con A una variable. Si $A \longrightarrow \omega$ es una regla de la gramática, decimos que $\alpha A \beta$ infiere a $\alpha \omega \beta$ y escribimos

$$\alpha A \beta \Longrightarrow \alpha \omega \beta$$

Definición.

Escribimos $u \overset{*}{\Longrightarrow} v$ si $u = v$ o si existe una secuencia $u_1, u_2, \dots, u_k$ con $k \geq 0$ y $u \Longrightarrow u_1 \Longrightarrow u_2 \Longrightarrow \cdots \Longrightarrow u_k \Longrightarrow v$

En el ejemplo 1:

$$E \overset{*}{\Longrightarrow} (a * b) + a10$$

Derivaciones a la izquierda y derecha

$$E \underset{\text{md}}{\overset{2)}{\Rightarrow}} E + E \underset{\text{md}}{\overset{3)}{\Rightarrow}} E + E * E \underset{\text{md}}{\overset{1)}{\Rightarrow}} E + E * I \underset{\text{md}}{\overset{5)}{\Rightarrow}} E + E * a \underset{\text{md}}{\overset{1)}{\Rightarrow}} E + I * a \underset{\text{md}}{\overset{6)}{\Rightarrow}} E + b * a$$

$$\underset{\text{md}}{\overset{1)}{\Rightarrow}} I + b * a \underset{\text{md}}{\overset{5)}{\Rightarrow}} a + b * a$$

$$\boxed{E \underset{\text{md}}{\overset{*}{\Rightarrow}} a + b * a}$$

$$E \underset{\text{mi}}{\overset{3)}{\Rightarrow}} E * E \underset{\text{mi}}{\overset{1)}{\Rightarrow}} I * E \underset{\text{mi}}{\overset{5)}{\Rightarrow}} a * E \underset{\text{mi}}{\overset{4)}{\Rightarrow}} a * (E) \underset{\text{mi}}{\overset{2)}{\Rightarrow}} a * (E + E) \underset{\text{mi}}{\overset{1)}{\Rightarrow}} a * (I + E) \underset{\text{mi}}{\overset{6)}{\Rightarrow}} a * (b + E)$$

$$\underset{\text{mi}}{\overset{1)}{\Rightarrow}} a * (b + I) \underset{\text{mi}}{\overset{5)}{\Rightarrow}} a * (b + a)$$

$$\boxed{E \underset{\text{mi}}{\overset{*}{\Rightarrow}} a * (b + a)}$$

Proposición:

Si w es una cadena terminal y A una variable, entonces $A \overset{*}{\Rightarrow} w$ sii $A \underset{\text{mi}}{\overset{*}{\Rightarrow}} w$ y $A \overset{*}{\Rightarrow} w$ sii $A \underset{\text{md}}{\overset{*}{\Rightarrow}} w$

Definición:

Si $G = (V, T, P, S)$ es una GLC y si $\alpha \in (V \cup T)^*$ tal que $S \overset{*}{\Rightarrow} \alpha$, entonces α es llamada forma sentencial.

Si $S \underset{\text{mi}}{\overset{*}{\Rightarrow}} \alpha$, entonces α es llamada forma sentencial izquierda, y si $S \underset{\text{md}}{\overset{*}{\Rightarrow}} \alpha$, entonces es α es llamada forma sentencial derecha.

Lenguaje de una Gramática

Si $G = (V, T, P, S)$ es una GLC. El lenguaje de G denotado como $L(G)$ es el conjunto de las cadenas terminales que tienen derivaciones desde el símbolo inicial, es decir:

$$L(G) = \{w \in T^* | S \overset{*}{\Rightarrow} w\}$$

Definición:

Si un lenguaje L es el lenguaje de una GLC, entonces L es llamado **Lenguaje Independiente de Contexto (LIC)**.

Ejemplo 2:

Diseñar una GLC que acepte el lenguaje $L_{01} = \{0^n 1^n | n \geq 0\}$

Ejemplos de cadenas que acepta: ε, 01,0011,000111, etc.

- $S \longrightarrow \varepsilon \mid 0S1$

Ejemplo 3:

Diseñar una GLC que acepte el conjunto de paréntesis balanceados

Ejemplos de cadenas que acepta: $\varepsilon, (\ \ ), ((\ \ )(\ \ ))(\ \ ), (((\ \ )))((\ \ )(\ \ ))$

Ejemplos de cadenas que no acepta: $(, (\ \ )), ((\ \ )))(\,),))$

- $S \longrightarrow SS \mid (S) \mid \varepsilon$

Ejemplo 4:

Diseñar una GLC que acepte el lenguaje $L_1 = \{0^n 1^m 0^n \; | n, m \geq 1\}$

Ejemplos de cadenas que acepta: $010, 00100, 0110, 00111100, 00011000$

- $S \longrightarrow 0A0$

  $A \longrightarrow 0A0 \mid 1B$

  $B \longrightarrow 1B \mid \varepsilon$

Ejemplo 5:

Diseñar una GLC que acepte el lenguaje $L_2 = \{a^n b^n \, c^m d^m | n, m \geq 1\}$

Ejemplos de cadenas que acepta: $abcd, aabbcd, abcccddd, aaabbbcccccddddd$

- $S \longrightarrow AB$

  $A \longrightarrow aAb \mid ab$

  $B \longrightarrow cBd \mid cd$

Árboles de Derivación

Sea $G = (V, T, P, S)$ una GLC. Los árboles de derivación para G son árboles con las siguientes características:

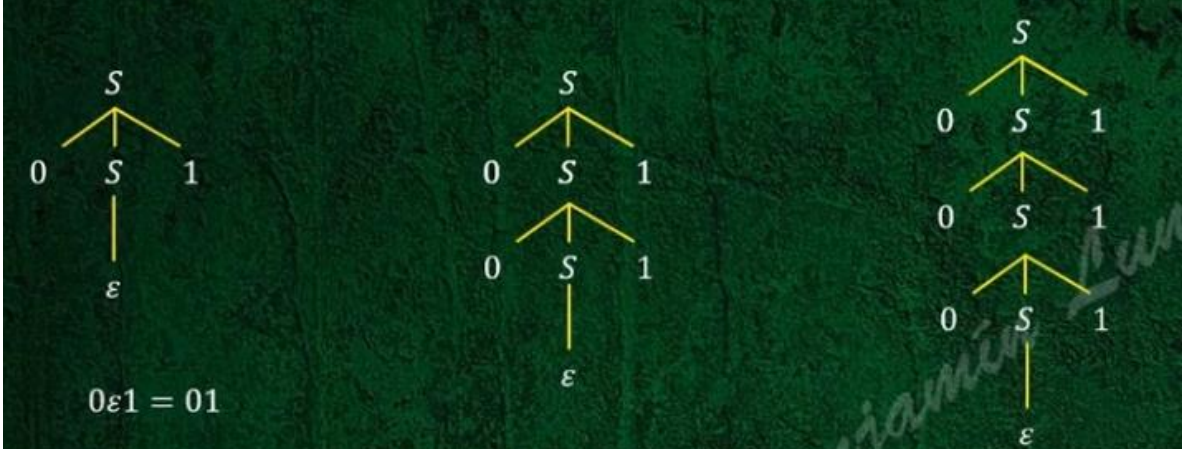
i) Cada nodo interior está etiquetado con una variable.

ii) Cada hoja está etiquetada con una variable, un terminal o ε. Si se encuentra etiquetado con ε, entonces tiene que ser hijo único de su padre.

iii) Si un nodo interior está etiquetado con A, y sus hijos están etiquetados con $X_1, X_2, \ldots, X_k$ respectivamente enumerados de izquierda a derecha, entonces $A \longrightarrow X_1 X_2 \cdots X_k$ es una producción de P.

$L_{01} = \{0^n 1^n | n \geq 0\}$

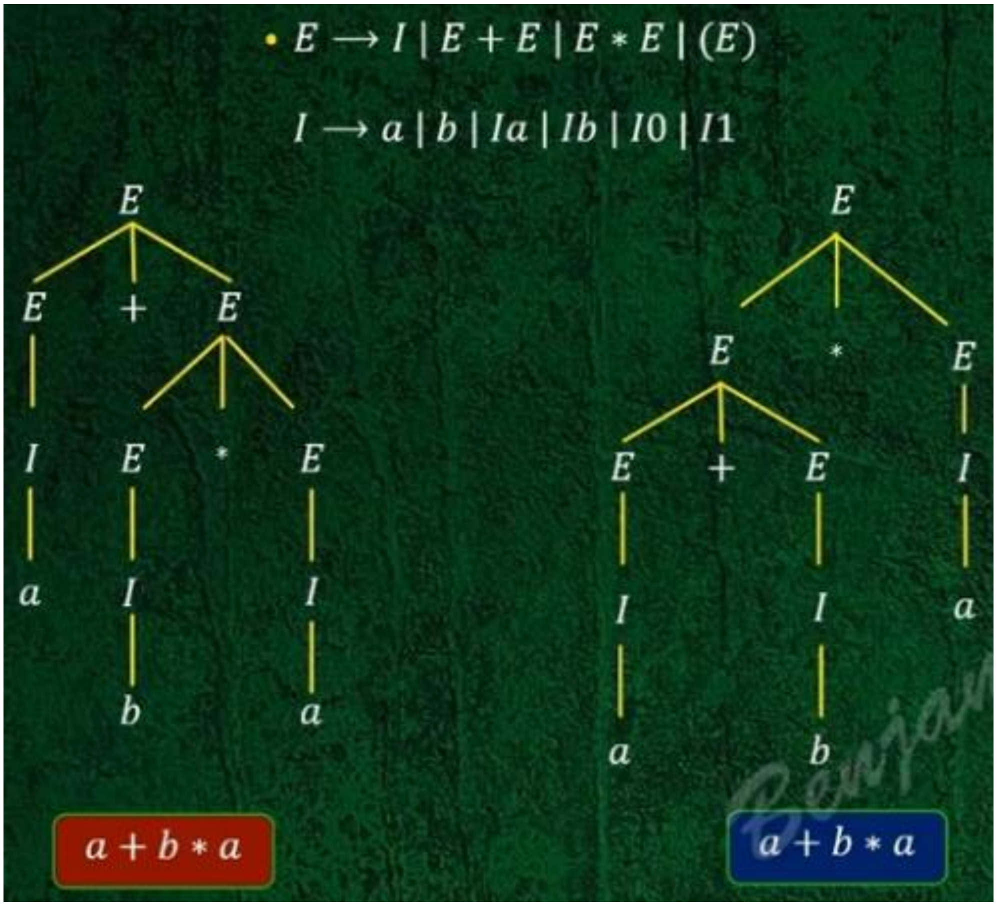
- $S \longrightarrow \varepsilon \mid 0S1$



$0\varepsilon 1 = 01$

$00\varepsilon 11 = 0011$

$000\varepsilon 111 = 000111$



Definición:

Sea $G = (V, T, P, S)$ una GLC. Se dice que G es ambigua, si existe $w \in T^*$ para la que existen dos árboles de derivación distintos con la raíz etiquetada con S cuyo resultado es w.

Si cada cadena $w \in T^*$ tiene como máximo un único árbol de derivación en la gramática, entonces la gramática no es ambigua.

observación:

No existe algoritmo alguno que decida si una gramática es ambigua o no.

Teorema:

Para toda gramática $G = (V, T, P, S)$ y toda cadena $w \in T^*$, w tiene dos árboles de derivación distintos si y solo si w tiene dos derivaciones más a la izquierda distintas desde S, si y solo si w tiene dos derivaciones más a la derecha distintas desde S.

- $E \longrightarrow I \mid E+E \mid E*E \mid (E)$

  $I \longrightarrow a \mid b \mid Ia \mid Ib \mid I0 \mid I1$

$$E \underset{\text{mi}}{\Rightarrow} E*E \underset{\text{mi}}{\Rightarrow} E+E*E \underset{\text{mi}}{\Rightarrow} I+E*E \underset{\text{mi}}{\Rightarrow} a+E*E \underset{\text{mi}}{\Rightarrow} a+I*E \underset{\text{mi}}{\Rightarrow} a+b*E$$
$$\underset{\text{mi}}{\Rightarrow} a+b*I \underset{\text{mi}}{\Rightarrow} a+b*a$$

$$E \underset{\text{mi}}{\Rightarrow} E+E \underset{\text{mi}}{\Rightarrow} I+E \underset{\text{mi}}{\Rightarrow} a+E \underset{\text{mi}}{\Rightarrow} a+E*E \underset{\text{mi}}{\Rightarrow} a+I*E \underset{\text{mi}}{\Rightarrow} a+b*E \underset{\text{mi}}{\Rightarrow} a+b*I$$
$$\underset{\text{mi}}{\Rightarrow} a+b*a$$

Definición:

Un lenguaje L es inherentemente ambiguo si todas las gramáticas que aceptan a L son ambiguas.

Forma Normal de Chomsky (FNC)

Todo LIC $L \neq \emptyset$ con $\varepsilon \notin L$ tiene una GLC G (que no tiene símbolos inútiles) en la que todas las producciones tienen una de las siguientes formas:

i) $A \longrightarrow BC$, donde $A, B, C \in V$

ii) $A \longrightarrow a$, con $A \in V$ y $a \in T$

FNC

1) Eliminar producciones ε
2) Eliminar producciones unitarias
3) Eliminar símbolos inútiles

1) Producciones ε

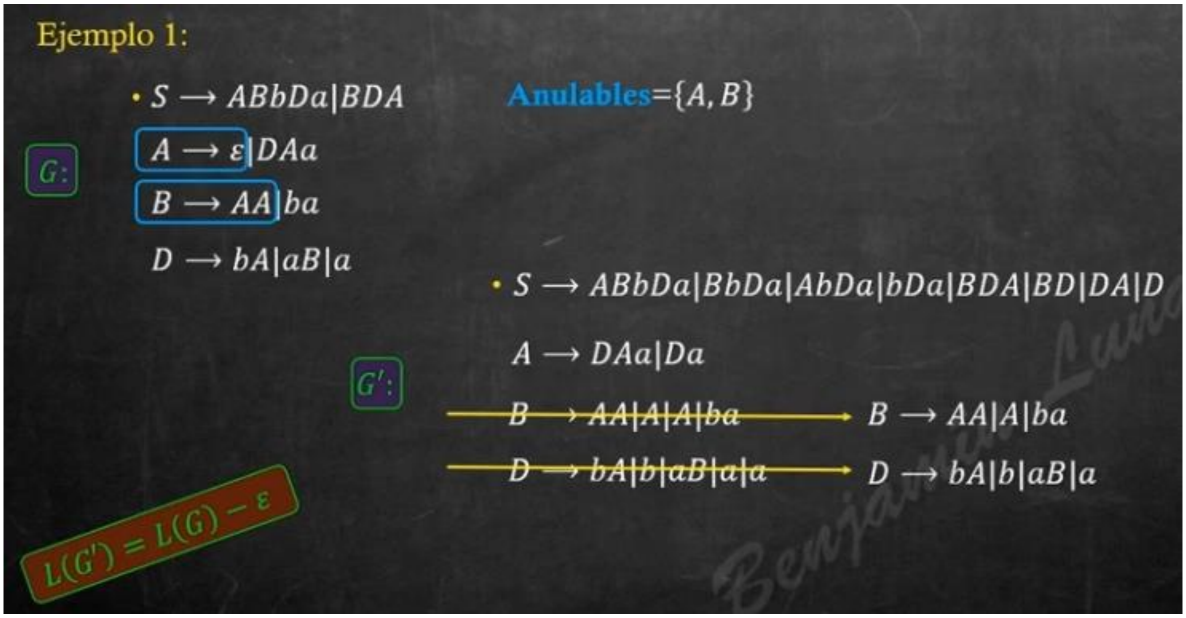
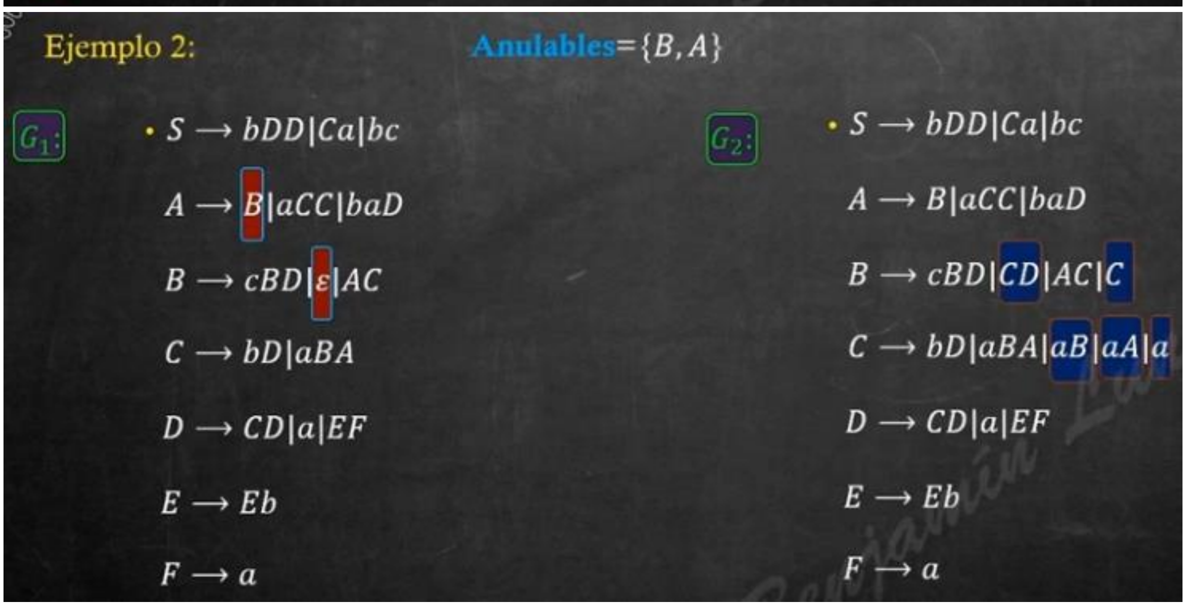
Definición 1.

Una variable A es **anulable** si $A \overset{*}{\Rightarrow} \varepsilon$.

Encontrando los símbolos anulables.

Base: Si $A \longrightarrow \varepsilon$ es una producción de G, A es anulable.

Inductivo: Si $B \longrightarrow C_1 C_2 \cdots C_k$ es una producción de G en la que todas las C_i son anulables, entonces B es anulable.


Ejemplo 1:
G:
$S \longrightarrow ABbDa|BDA$
$\boxed{A \longrightarrow \varepsilon}|DAa$
$\boxed{B \longrightarrow AA}|ba$
$D \longrightarrow bA|aB|a$
Anulables=$\{A, B\}$
G':
$S \longrightarrow ABbDa|BbDa|AbDa|bDa|BDA|BD|DA|D$
$A \longrightarrow DAa|Da$
$\cancel{B \longrightarrow AA|A|A|ba} \longrightarrow B \longrightarrow AA|A|ba$
$\cancel{D \longrightarrow bA|b|aB|a|a} \longrightarrow D \longrightarrow bA|b|aB|a$
L(G') = L(G) − ε
Benjamín Luna

Ejemplo 2:
Anulables=$\{B, A\}$
G_1:
$S \longrightarrow bDD|Ca|bc$
$A \longrightarrow B|aCC|baD$
$B \longrightarrow cBD|\varepsilon|AC$
$C \longrightarrow bD|aBA$
$D \longrightarrow CD|a|EF$
$E \longrightarrow Eb$
$F \longrightarrow a$
G_2:
$S \longrightarrow bDD|Ca|bc$
$A \longrightarrow B|aCC|baD$
$B \longrightarrow cBD|CD|AC|C$
$C \longrightarrow bD|aBA|aB|aA|a$
$D \longrightarrow CD|a|EF$
$E \longrightarrow Eb$
$F \longrightarrow a$
Benjamín Luna

2) Producciones Unitarias

Definición 2.

Una producción unitaria es una producción de la forma $A \longrightarrow B$ con $A, B \in V$.

Definición 3.

Si $A \overset{*}{\Longrightarrow} B$, entonces (A, B) es llamado un par unitario.

Encontrando pares unitarios.

Base: (A, A) es un par unitario para cada $A \in V$.

Inductivo: Si (A, B) es un par unitario y $B \longrightarrow C$ es una producción con $C \in V$, entonces (A, C) es un par unitario.

Ejemplo 3:

G_2:	Pares unitarios	producciones
$S \longrightarrow bDD \mid Ca \mid bc$	(S,S)	G_3:
$A \longrightarrow B \mid aCC \mid baD$	$(A,A)(A,B)$ (A,C)	
$B \longrightarrow cBD \mid CD \mid AC \mid C$		
$C \longrightarrow bD \mid aBA \mid aB \mid aA \mid a$	(B,B) (B,C)	
$D \longrightarrow CD \mid a \mid EF$	(C,C)	
$E \longrightarrow Eb$	(D,D)	
$F \longrightarrow a$	(E,E)	
	(F,F)	

Para cada par unitario (A, B), añadimos a G_3 todas las producciones $A \longrightarrow \alpha$, donde $B \longrightarrow \alpha$ es una producción no unitaria de G_2

$L(G_3) = L(G_2)$

Ejemplo 3:

G_2:	Pares unitarios	producciones (G_3:)
$S \longrightarrow bDD \mid Ca \mid bc$	(S,S)	$S \longrightarrow bDD \mid Ca \mid bc$
$A \longrightarrow B \mid aCC \mid baD$	$(A,A)(A,B)$ (A,C)	$A \longrightarrow aCC \mid baD$ $\mid cBD \mid CD \mid AC$ $\mid bD \mid aBA \mid aB \mid aA \mid a$
$B \longrightarrow cBD \mid CD \mid AC \mid C$	$(B,B)\,(B,C)$	$B \longrightarrow cBD \mid CD \mid AC$ $\mid bD \mid aBA \mid aB \mid aA \mid a$
$C \longrightarrow bD \mid aBA \mid aB \mid aA \mid a$	(C,C)	$C \longrightarrow bD \mid aBA \mid aB \mid aA \mid a$
$D \longrightarrow CD \mid a \mid EF$	(D,D)	$D \longrightarrow CD \mid a \mid EF$
$E \longrightarrow Eb$	(E,E)	$E \longrightarrow Eb$
$F \longrightarrow a$	(F,F)	$F \longrightarrow a$

$L(G_3) = L(G_2)$

3) Símbolos inútiles

Definición 4:

Decimos que un símbolo X es útil para una gramática $G = (V, T, P, S)$ si existe alguna derivación de la forma $S \overset{*}{\Longrightarrow} \alpha X \beta \overset{*}{\Longrightarrow} w$ con $w \in T^*$.

Si un símbolo no es útil, diremos que es inútil

Observación:

$X \in (V \cup T)$ y $\alpha X \beta$ puede ser la primera o última derivación.

Definición 5: X es generador si $X \overset{*}{\Longrightarrow} w$ para alguna cadena $w \in T^*$.

Observación: Todo símbolo terminal es un generador.

Definición 6: X es alcanzable si existe una derivación $S \overset{*}{\Longrightarrow} \alpha X \beta$ para algunos $\alpha, \beta \in (V \cup T)^*$.

Observación: Un símbolo útil es generador y alcanzable.

Símbolos inútiles

Paso 1 - Eliminar símbolos que no son generadores
Paso 2 - Eliminar símbolos que no son alcanzables

Encontrando símbolos generadores.

Base: Todo símbolo de T es generador.

Inductivo: Si $A \longrightarrow \alpha$ es una producción tal que todo símbolo de α es generador, entonces A es generador.

Encontrando símbolos alcanzables.

Base: $S \in V$ es alcanzable (S es el símbolo inicial).

Inductivo: Si A es alcanzable, entonces para todas las producciones cuya cabeza es A, todos los símbolos de los cuerpos de dichas producciones son alcanzables.

Ejemplo 4: G_3:

- $S \longrightarrow bDD|Ca|bc$

$A \longrightarrow aCC|baD|cBD|CD|AC|bD|aBA|aB|aA|a$

$B \longrightarrow cBD|CD|AC|bD|aBA|aB|aA|a$

$C \longrightarrow bD|aBA|aB|aA|a$

$D \longrightarrow CD|a|EF$

$E \longrightarrow Eb$

$F \longrightarrow a$

Generadores$=\{a, b, c, S, A, B, C, D, F\}$

No generadores$=\{E\}$

G_4:

- $S \longrightarrow bDD|Ca|bc$

$A \longrightarrow aCC|baD|cBD|CD|AC|bD|aBA|aB|aA|a$

$B \longrightarrow cBD|CD|AC|bD|aBA|aB|aA|a$

$C \longrightarrow bD|aBA|aB|aA|a$

$D \longrightarrow CD|a|\cancel{EF}$

$\cancel{E \longrightarrow Eb}$

$F \longrightarrow a$

$L(G_4) = L(G_3)$

G_4:

- $S \longrightarrow bDD|Ca|bc$

$A \longrightarrow aCC|baD|cBD|CD|AC|bD|aBA|aB|aA|a$

$B \longrightarrow cBD|CD|AC|bD|aBA|aB|aA|a$

$C \longrightarrow bD|aBA|aB|aA|a$

$D \longrightarrow CD|a$

$F \longrightarrow a$

Alcanzables$=\{S, b, D, C, a, c, B, A\}$

No alcanzables$=\{F\}$

G_5:

- $S \longrightarrow bDD|Ca|bc$

$A \longrightarrow aCC|baD|cBD|CD|AC|bD|aBA|aB|aA|a$

$B \longrightarrow cBD|CD|AC|bD|aBA|aB|aA|a$

$C \longrightarrow bD|aBA|aB|aA|a$

$D \longrightarrow CD|a$

$\cancel{F \longrightarrow a}$

G_1:

- $S \longrightarrow bDD|Ca|bc$

$A \longrightarrow B|aCC|baD$

$B \longrightarrow cBD|\varepsilon|AC$

$C \longrightarrow bD|aBA$

$D \longrightarrow CD|a|EF$

$E \longrightarrow Eb$

$F \longrightarrow a$

G_5:

- $S \longrightarrow bDD|Ca|bc$

$A \longrightarrow aCC|baD|cBD|CD|AC|bD|aBA|aB|aA|a$

$B \longrightarrow cBD|CD|AC|bD|aBA|aB|aA|a$

$C \longrightarrow bD|aBA|aB|aA|a$

$D \longrightarrow CD|a$

i) $A \longrightarrow BC$, donde $A, B, C \in V$

ii) $A \longrightarrow a$, con $A \in V$ y $a \in T$

FNC

G_5:

- $S \longrightarrow bDD|Ca|bc$

$A \longrightarrow aCC|baD|cBD|CD|AC|bD|aBA|aB|aA|a$

$B \longrightarrow cBD|CD|AC|bD|aBA|aB|aA|a$

$C \longrightarrow bD|aBA|aB|aA|a$

$D \longrightarrow CD|a$

$V_b \longrightarrow b$

$V_{bD} \longrightarrow V_b D$

$V_a \longrightarrow a$

$V_c \longrightarrow c$

- $S \longrightarrow V_{bD}D|CV_a|V_bV_c$

$A \longrightarrow V_{aC}C|V_{ba}D|V_{cB}D|CD|AC|V_bD|V_{aB}A|V_aB|V_aA|a$

$B \longrightarrow V_{cB}D|CD|AC|V_bD|V_{aB}A|V_aB|V_aA|a$

$C \longrightarrow V_bD|V_{aB}A|V_aB|V_aA|a$

$C \longrightarrow CD|a$

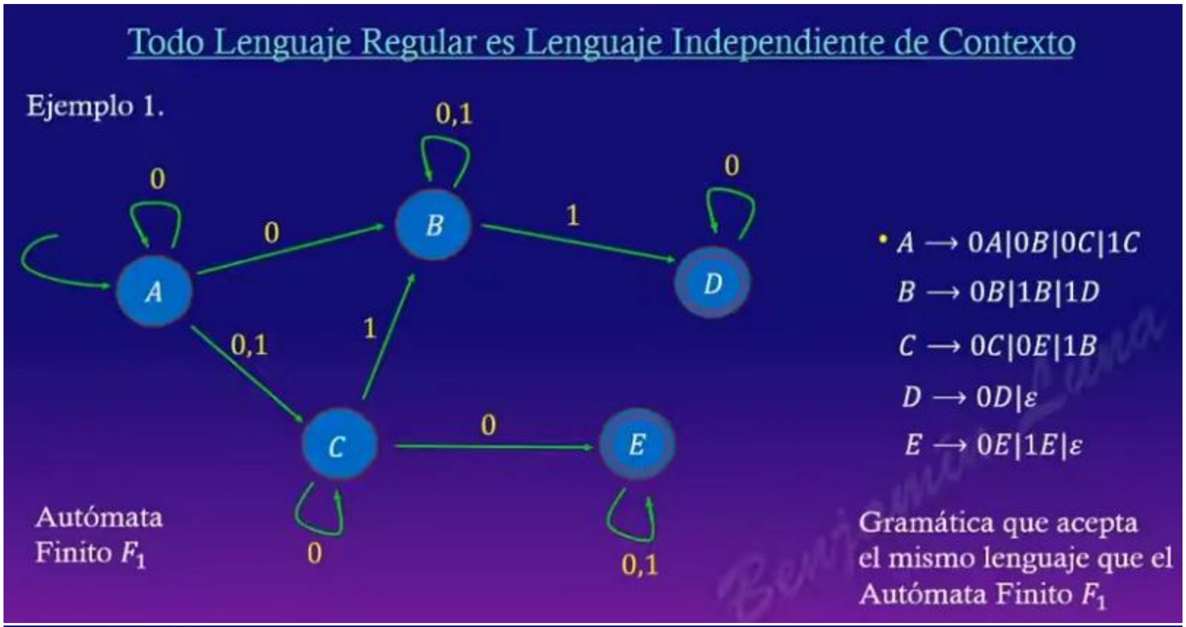
$V_{aC} \longrightarrow V_a C$

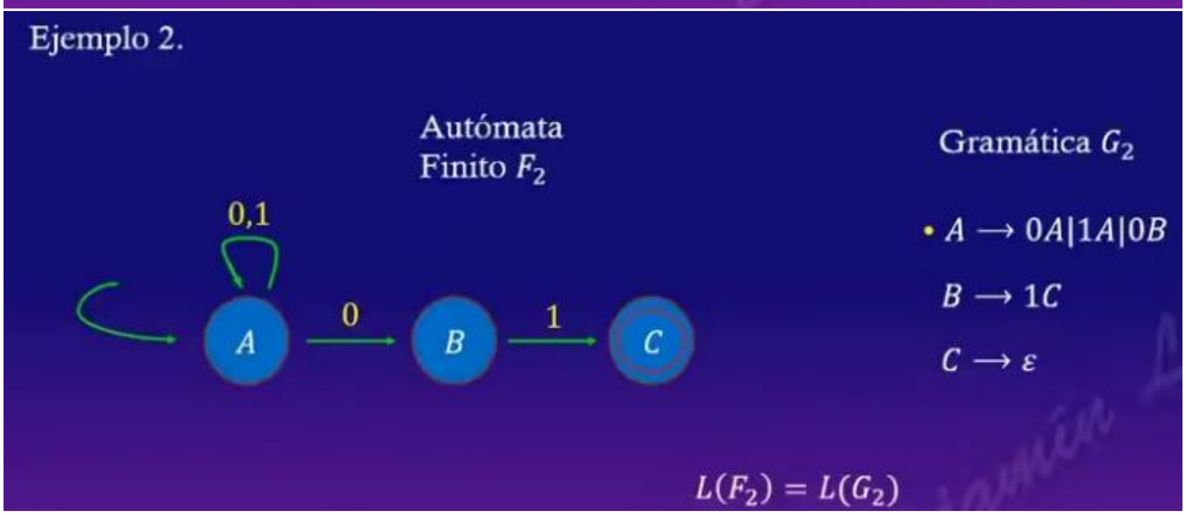
$V_{ba} \longrightarrow V_a V_b$

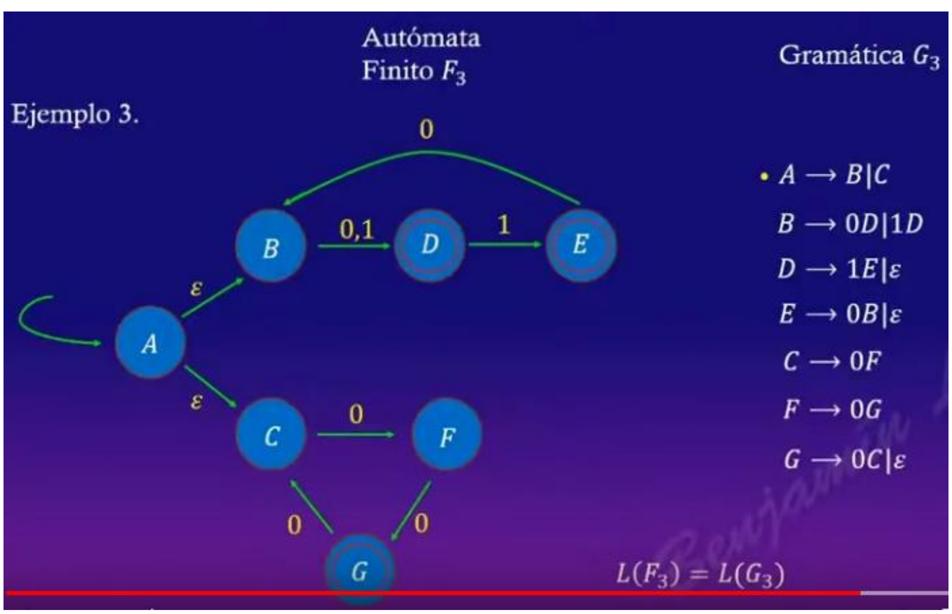
$V_{cB} \longrightarrow V_c B$

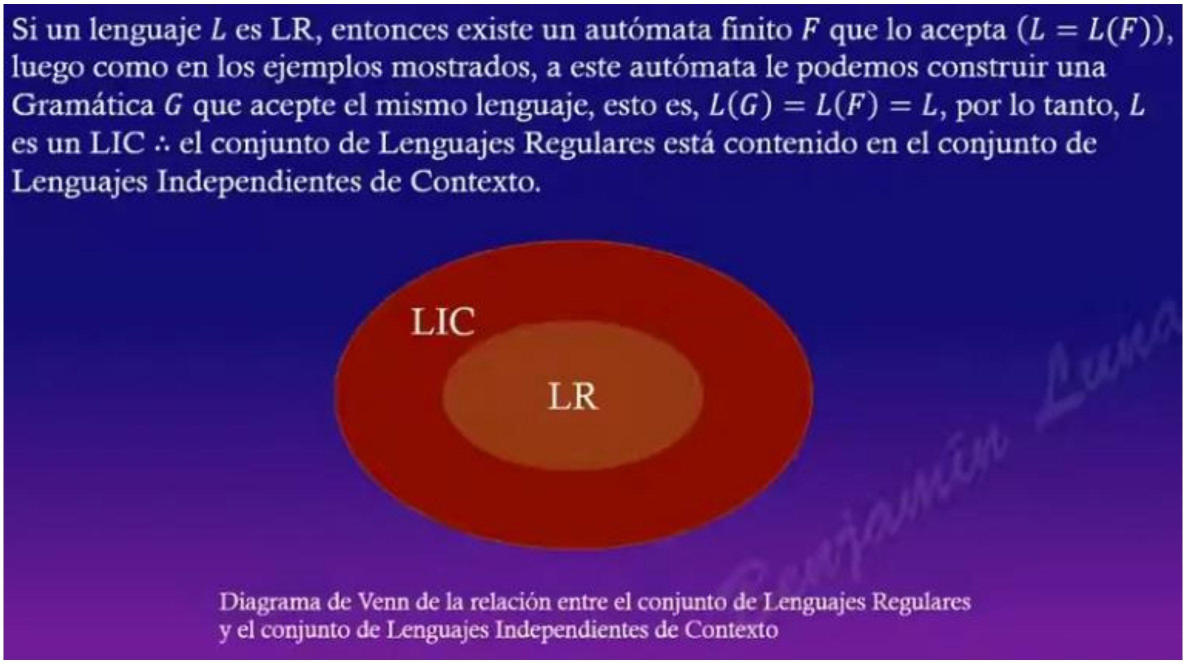
$V_{aB} \longrightarrow V_a B$

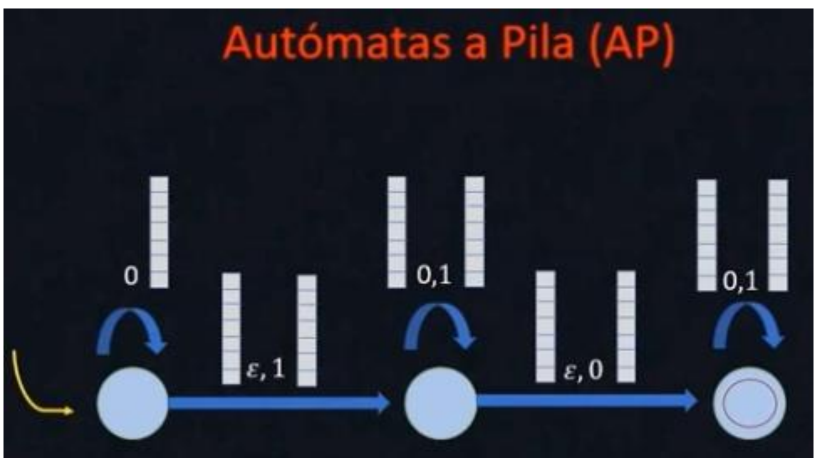
FNC


Todo Lenguaje Regular es Lenguaje Independiente de Contexto
Ejemplo 1.
0
0,1
0
0
1
A
B
D
0,1
1
C
0
E
Autómata
Finito F_1
0
0,1
$A \longrightarrow 0A|0B|0C|1C$
$B \longrightarrow 0B|1B|1D$
$C \longrightarrow 0C|0E|1B$
$D \longrightarrow 0D|\varepsilon$
$E \longrightarrow 0E|1E|\varepsilon$
Gramática que acepta
el mismo lenguaje que el
Autómata Finito F_1


Ejemplo 2.
Autómata
Finito F_2
Gramática G_2
0,1
A
0
B
1
C
$A \longrightarrow 0A|1A|0B$
$B \longrightarrow 1C$
$C \longrightarrow \varepsilon$
$L(F_2) = L(G_2)$


Autómata
Finito F_3
Gramática G_3
Ejemplo 3.
0
B
0,1
D
1
E
ε
A
ε
C
0
F
0
0
G
$A \longrightarrow B|C$
$B \longrightarrow 0D|1D$
$D \longrightarrow 1E|\varepsilon$
$E \longrightarrow 0B|\varepsilon$
$C \longrightarrow 0F$
$F \longrightarrow 0G$
$G \longrightarrow 0C|\varepsilon$
$L(F_3) = L(G_3)$


Si un lenguaje L es LR, entonces existe un autómata finito F que lo acepta ($L = L(F)$), luego como en los ejemplos mostrados, a este autómata le podemos construir una Gramática G que acepte el mismo lenguaje, esto es, $L(G) = L(F) = L$, por lo tanto, L es un LIC ∴ el conjunto de Lenguajes Regulares está contenido en el conjunto de Lenguajes Independientes de Contexto.
LIC
LR
Diagrama de Venn de la relación entre el conjunto de Lenguajes Regulares
y el conjunto de Lenguajes Independientes de Contexto


Autómatas a Pila (AP)
0
$\varepsilon, 1$
0,1
$\varepsilon, 0$
0,1


Pila
LIFO: Last In, Fist Out
X
Fondo de la Pila
Z_0

$\Sigma=\{0,1\}$
$\Gamma=\{X,Y,Z_0\}$

Ejemplo 1:

$L(P)=\{0^n 1^n \mid n>0\}$

W=0011

$(A,0011,Z_0) \vdash (A,011,XZ_0) \vdash (A,11,XXZ_0) \vdash (B,1,XZ_0) \vdash (B,\varepsilon,Z_0) \vdash (C,\varepsilon,Z_0)$

$(A,0011,Z_0) \vdash^* (C,\varepsilon,Z_0)$

W'=0111

$(A,0111,Z_0) \vdash (A,111,XZ_0) \vdash (B,11,Z_0) \vdash (C,11,Z_0)$

Descripciones instantáneas

$(A,0111,Z_0) \vdash^* (C,11,Z_0)$

Definición (autómata a Pila).

Un AP es una 7 tupla $P=(Q,\Sigma,\Gamma,\delta,q_0,Z_0,F)$ donde:

Q: Es un conjunto finito de estados

Σ: Alfabeto de entrada

Γ: Alfabeto de Pila

q_0: Estado inicial

Z_0: Símbolo del fondo de la pila

F: Conjunto de estados de aceptación

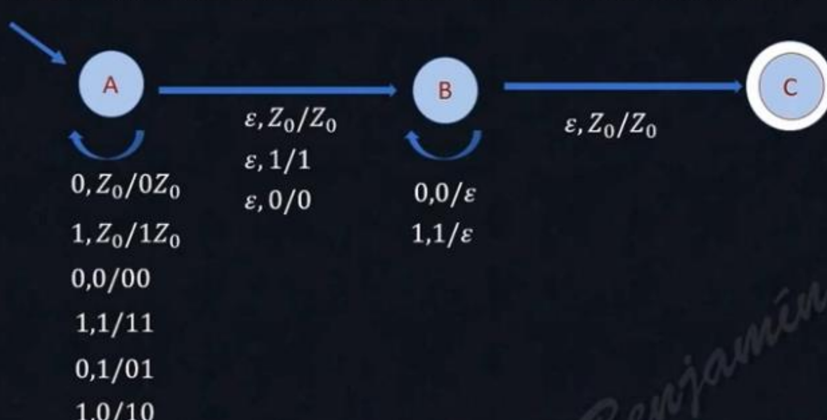
δ: Función de transición $\Sigma_\varepsilon = \Sigma \cup \{\varepsilon\}$

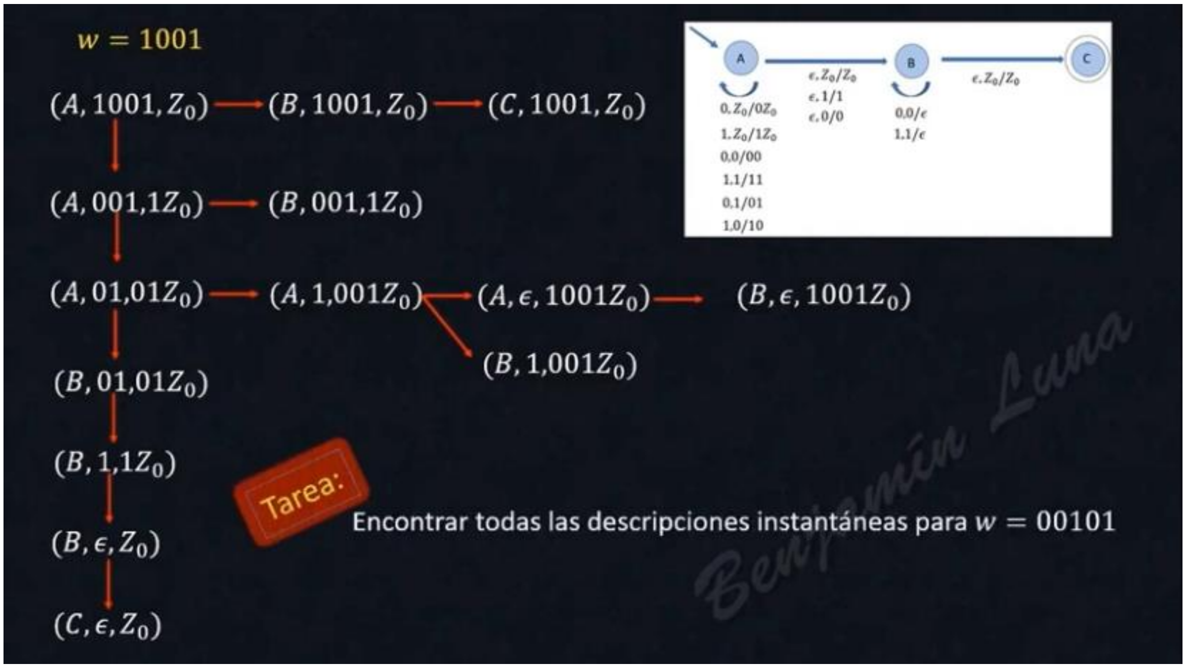
$\delta: Q \times \Sigma_\varepsilon \times \Gamma_\varepsilon \longrightarrow P(Q \times \Gamma^*)$

$\delta(q,a,X)=(p,\gamma)$

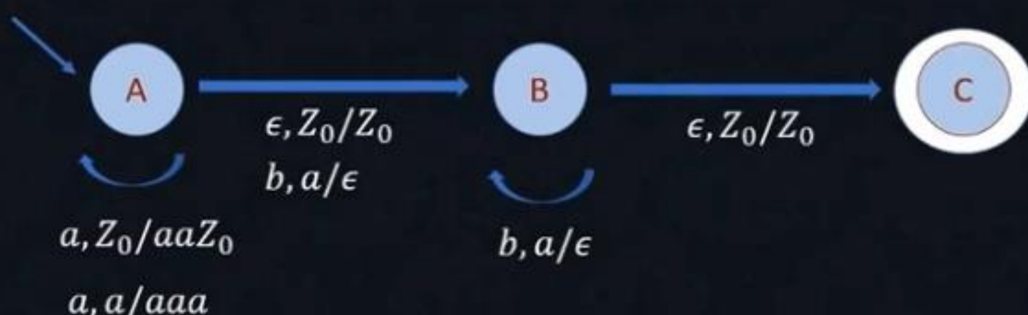
Ejemplo 3:

- Diseñar un AP que acepte el lenguaje $L=\{ww^r \mid w$ pertenece a $\{0,1\}^*\}$





Ejemplo 4: Diseñar un AP que acepte el lenguaje $L = \{a^n b^{2n} | n \geq 0\}$



Lenguaje de una Pila por estado de aceptación

Definición.

Sea $P = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ un AP. El lenguaje definido por P por estado final o de aceptación es:

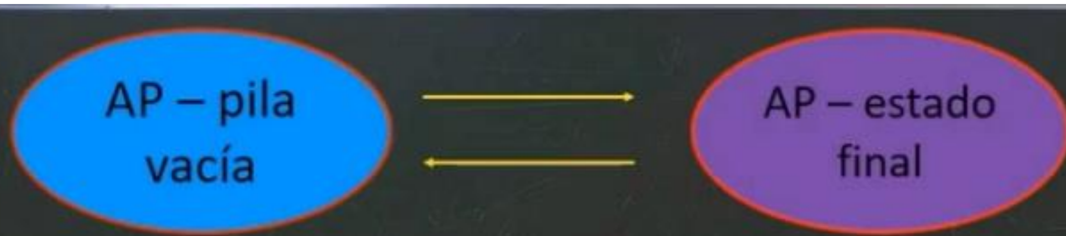
$$L(P) = \{w | (q_0, w, Z_0) \overset{*}{\vdash} (q, \epsilon, \alpha)\}$$

para algún $q \in F$ y $\alpha \in \Gamma^*$

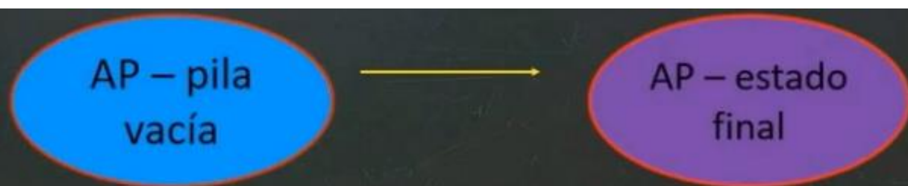
Lenguaje de una Pila

Sea $P = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ un AP. Entonces el lenguaje de P definido por

estado de aceptación o estado final	$L(P) = \{w \mid (q_0, w, Z_0) \overset{*}{\vdash} (q, \epsilon, \alpha)\}$	$q \in F$ y $\alpha \in \Gamma^*$
pila vacía	$N(P) = \{w \mid (q_0, w, Z_0) \overset{*}{\vdash} (q, \epsilon, \epsilon)\}$	$q \in Q$



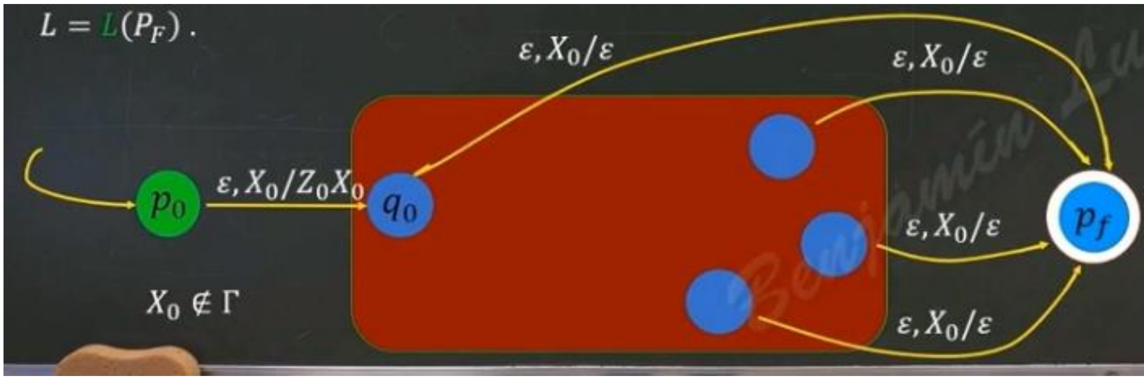
Equivalencia entre AP por pila vacía y AP por estado final

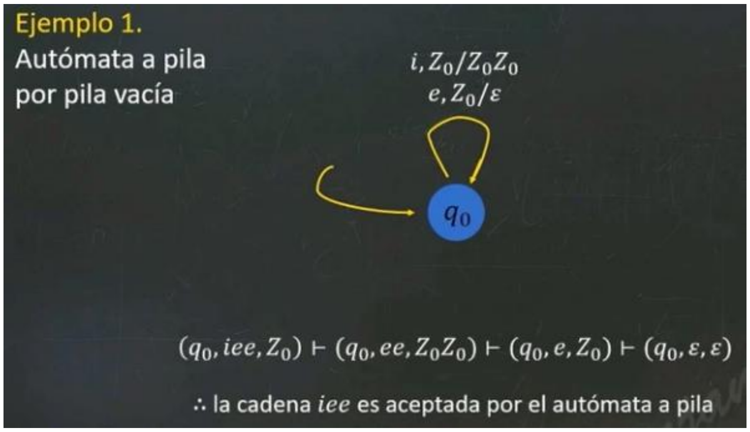


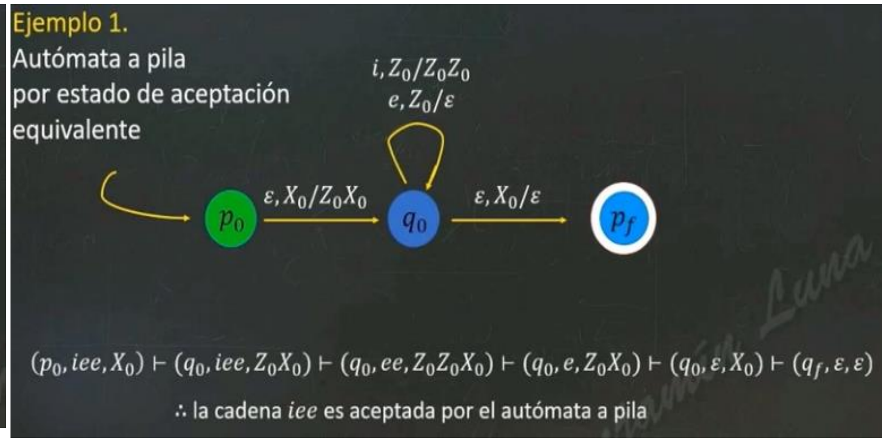
Teorema 1:

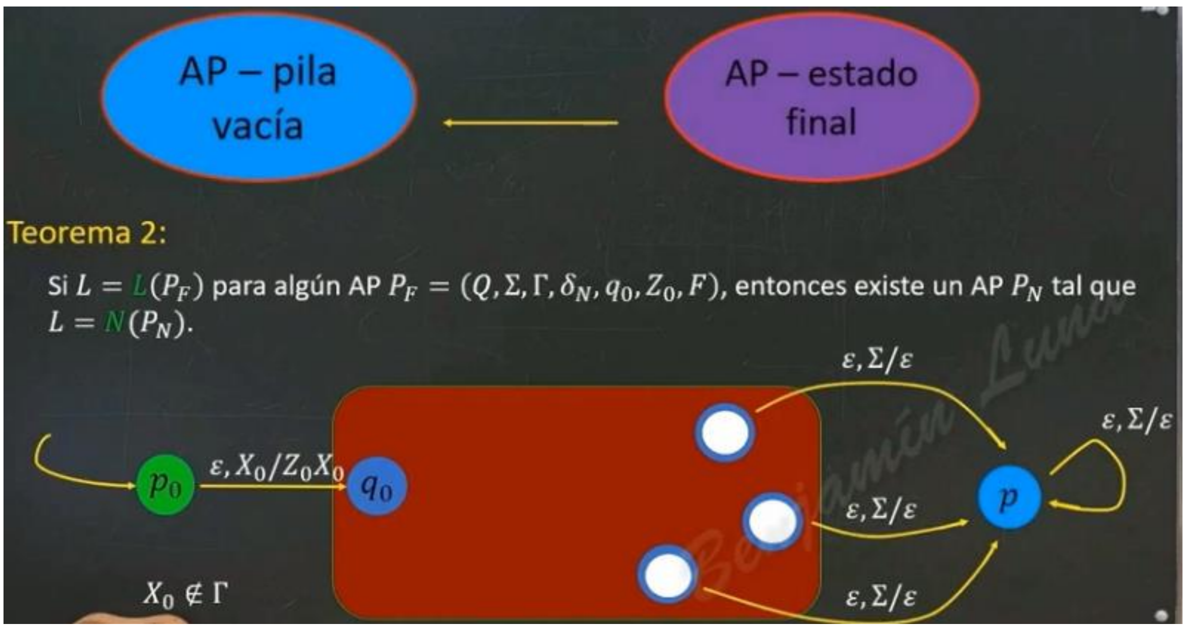
Si $L = N(P_N)$ para algún AP $P_N = (Q, \Sigma, \Gamma, \delta_N, q_0, Z_0)$, entonces existe un AP P_F tal que $L = L(P_F)$.

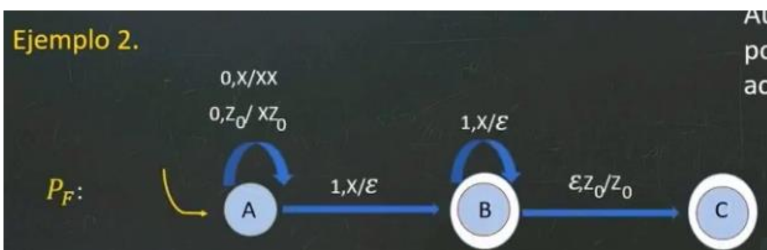
$L = L(P_F)$.

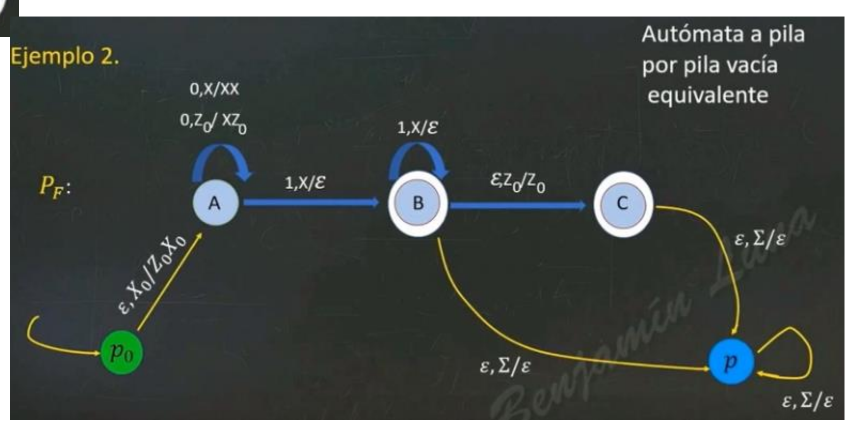



Ejemplo 1.
Autómata a pila
por pila vacía
$i, Z_0/Z_0Z_0$
$e, Z_0/\varepsilon$
q_0
$(q_0, iee, Z_0) \vdash (q_0, ee, Z_0Z_0) \vdash (q_0, e, Z_0) \vdash (q_0, \varepsilon, \varepsilon)$
∴ la cadena iee es aceptada por el autómata a pila


Ejemplo 1.
Autómata a pila
por estado de aceptación
equivalente
$i, Z_0/Z_0Z_0$
$e, Z_0/\varepsilon$
p_0
$\varepsilon, X_0/Z_0X_0$
q_0
$\varepsilon, X_0/\varepsilon$
p_f
$(p_0, iee, X_0) \vdash (q_0, iee, Z_0X_0) \vdash (q_0, ee, Z_0Z_0X_0) \vdash (q_0, e, Z_0X_0) \vdash (q_0, \varepsilon, X_0) \vdash (q_f, \varepsilon, \varepsilon)$
∴ la cadena iee es aceptada por el autómata a pila


AP – pila vacía
AP – estado final
Teorema 2:
Si $L = L(P_F)$ para algún AP $P_F = (Q, \Sigma, \Gamma, \delta_N, q_0, Z_0, F)$, entonces existe un AP P_N tal que $L = N(P_N)$.
$\varepsilon, \Sigma/\varepsilon$
$\varepsilon, \Sigma/\varepsilon$
p_0
$\varepsilon, X_0/Z_0X_0$
q_0
$\varepsilon, \Sigma/\varepsilon$
p
$X_0 \notin \Gamma$
$\varepsilon, \Sigma/\varepsilon$


Ejemplo 2.
0,X/XX
$0,Z_0/XZ_0$
1,X/Ɛ
P_F:
A
1,X/Ɛ
B
$ƐZ_0/Z_0$
C


Ejemplo 2.
Autómata a pila
por pila vacía
equivalente
0,X/XX
$0,Z_0/XZ_0$
1,X/Ɛ
P_F:
A
1,X/Ɛ
B
$ƐZ_0/Z_0$
C
$\varepsilon, X_0/Z_0X_0$
$\varepsilon, \Sigma/\varepsilon$
p_0
$\varepsilon, \Sigma/\varepsilon$
p
$\varepsilon, \Sigma/\varepsilon$

Equivalencia entre GLC y AP por pila vacía parte 1

Teorema:

Dada una GLC $G = (V, T, P, S)$, existe un AP P_N que acepta por pila vacía tal que $L(G) = L(P_N)$.

El AP $P_N = (\{q\}, T, V \cup T, \delta, q, S)$ esta dado por:

i) Para cada $X \in V$: $\delta(q, \varepsilon, X) = \{(q, \beta) | X \longrightarrow \beta \text{ } es \text{ } una \text{ } producción \text{ } de \text{ } P_N\}$.

ii) Para cada $a \in T$: $\delta(q, a, a) = \{(q, \varepsilon)\}$.

Ejemplo:

- $S \longrightarrow 0A0$

  $A \longrightarrow 0A0 \mid 1B$

  $B \longrightarrow 1B \mid \varepsilon$

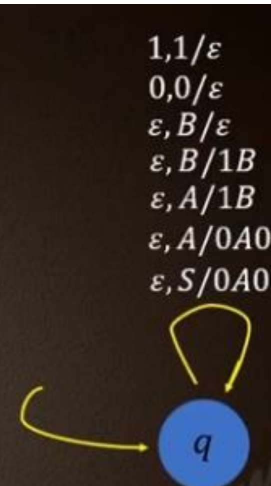
$\delta(q, \varepsilon, S) = \{(q, 0A0)\}$

$\delta(q, \varepsilon, A) = \{(q, 0A0), (q, 1B)\}$

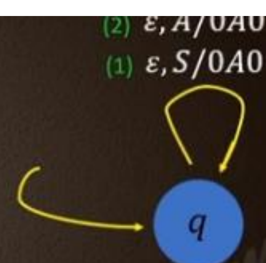
$\delta(q, \varepsilon, B) = \{(q, 1B), (q, \varepsilon)\}$

$\delta(q, 0,0) = \{(q, \varepsilon)\}$

$\delta(q, 1,1) = \{(q, \varepsilon)\}$

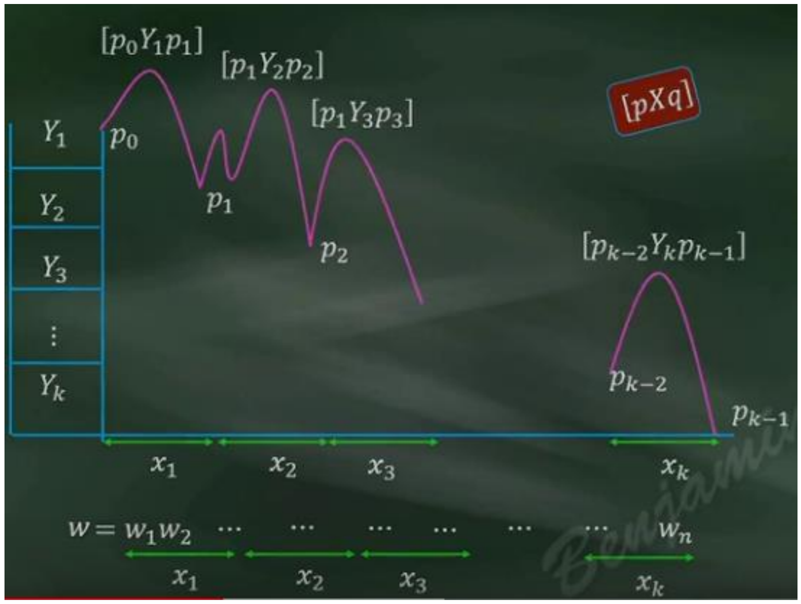


010



$(q, 010, S) \overset{(1)}{\vdash} (q, 010, 0A0) \overset{(6)}{\vdash} (q, 10, A0) \overset{(3)}{\vdash} (q, 10, 1B0)$
$\underset{(7)}{\vdash} (q, 0, B0), \underset{(5)}{\vdash} (q, 0,0) \underset{(6)}{\vdash} (q, \varepsilon, \varepsilon).$

$(q, 010, S) \overset{*}{\vdash} (q, \varepsilon, \varepsilon).$



Teorema:

Sea $P = (Q, \Sigma, \Gamma, \delta, q_0, Z_0)$ un AP que acepta por pila vacía. Entonces existe una GLC G tal que $L(G) = N(P)$.

La GLC $G = (V, \Sigma, R, S)$ está dada por:

V se forma de los siguientes elementos.

i) S es el símbolo inicial.

ii) Todos los símbolos de la forma $[pXq]$ donde p y q son estados de Q y X es símbolo de pila, $X \in \Gamma$.

i) y ii) $\Longrightarrow V = \{[pXq] | p, q \in Q, X \in \Gamma\} \cup \{S\}$

Las producciones son como sigue:

1.- para cada estado x, G contiene la producción

$$S \longrightarrow [q_0Z_0x]$$

2.- suponga que $(r, Y_1Y_2 \cdots Y_k) \in \delta(q, a, X)$ donde

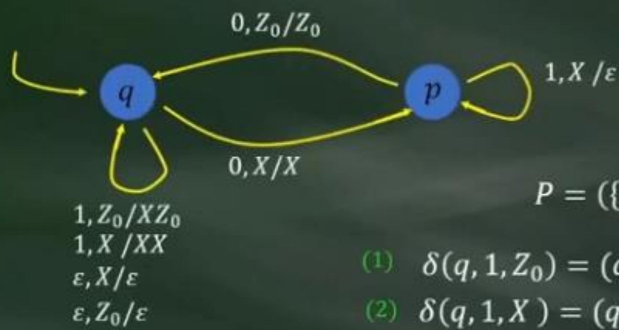
a) $a \in \Sigma$. a puede ser ε

b) $k \geq 0$. Si $k = 0$, entonces $(r, Y_1Y_2, \ldots, Y_k) = (r, \varepsilon)$.

entonces, para todas las listas de estados $r_1, r_2, \ldots, r_k$. G contiene la producción:

$$[qXr_k] \longrightarrow a[rY_1r_1][r_1Y_2r_2] \cdots [r_{k-1}Y_kr_k]$$

Ejemplo 1: Determinar la GLC equivalente al siguiente AP



$$P = (\{p, q\}, \Sigma = \{0,1\}, \Gamma = \{X, Z_0\}, \delta, q, Z_0)$$

(1) $\delta(q, 1, Z_0) = (q, XZ_0)$

(2) $\delta(q, 1, X) = (q, XX)$

(3) $\delta(q, \varepsilon, X) = (q, \varepsilon)$

(4) $\delta(q, \varepsilon, Z_0) = (q, \varepsilon)$

(5) $\delta(q, 0, Z_0) = (p, Z_0)$

(6) $\delta(q, 0, X) = (p, X)$

(7) $\delta(p, 1, X) = (p, \varepsilon)$

1.- para cada estado p, G contiene la producción

$$S \longrightarrow [q_0 Z_0 x]$$

- $S \longrightarrow [qZ_0q]$
- $S \longrightarrow [qZ_0p]$

(1) $\delta(q, 1, Z_0) = (q, XZ_0)$

(2) $\delta(q, 1, X) = (q, XX)$

(3) $\delta(q, \varepsilon, X) = (q, \varepsilon)$

(4) $\delta(q, \varepsilon, Z_0) = (q, \varepsilon)$

(5) $\delta(q, 0, Z_0) = (p, Z_0)$

(6) $\delta(q, 0, X) = (p, X)$

(7) $\delta(p, 1, X) = (p, \varepsilon)$

2.- suponga que $(r, Y_1Y_2 \cdots Y_k) \in \delta(q, a, X)$

para todas las listas de estados $r_1, r_2, \ldots, r_k$. G contiene la producción:

$$[qXr_k] \longrightarrow a[rY_1r_1][r_1Y_2r_2] \cdots [r_{k-1}Y_kr_k]$$

(1) $\delta(q, 1, Z_0) = (q, XZ_0)$

$$(r, Y_1Y_2 \cdots Y_k) \in \delta(q, a, X)$$

$$(q, XZ_0) \in \delta(q, 1, Z_0)$$

$$[qXr_k] \longrightarrow a[rY_1r_1][r_1Y_2r_2] \cdots [r_{k-1}Y_kr_k]$$

$$[qZ_0x] \longrightarrow 1[qXy][yZ_0x] \qquad \forall x, y \in Q$$

(1) $\delta(q, 1, Z_0) = (q, XZ_0)$

$$[qZ_0x] \longrightarrow 1[qXy][yZ_0x] \qquad \forall x, y \in Q$$

- $[qZ_0q] \longrightarrow 1[qXq][qZ_0q]$
- $[qZ_0q] \longrightarrow 1[qXp][pZ_0q]$
- $[qZ_0p] \longrightarrow 1[qXq][qZ_0p]$
- $[qZ_0p] \longrightarrow 1[qXp][pZ_0p]$

(6) $\delta(q,0,X) = (p,X)$

$$(r, Y_1Y_2\cdots Y_k) \in \delta(q,a,X)$$

$$(p,X) \in \delta(q,0,X)$$

$$[qXr_k] \longrightarrow a[rY_1r_1][r_1Y_2r_2]\cdots[r_{k-1}Y_kr_k]$$

$$[qXx] \longrightarrow 0[pXx] \qquad \forall x \in Q$$

- $[qXq] \longrightarrow 0[pXq]$
- $[qXp] \longrightarrow 0[pXp]$

(3) $\delta(q,\varepsilon,X) = (q,\varepsilon)$

$$(r, Y_1Y_2\cdots Y_k) \in \delta(q,a,X)$$

$$(q,\varepsilon) \in \delta(q,\varepsilon,X)$$

$$[qXr_k] \longrightarrow a[rY_1r_1][r_1Y_2r_2]\cdots[r_{k-1}Y_kr_k]$$

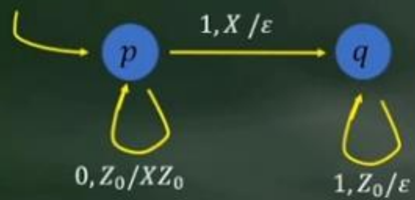
$$[qXx] \longrightarrow \varepsilon[q\varepsilon x] \qquad \forall x \in Q$$

- $[qXq] \longrightarrow \varepsilon\, [\overset{\varepsilon}{\cancel{q\varepsilon q}}] = \varepsilon$
- $[qXp] \longrightarrow \varepsilon\cancel{[q\varepsilon p]}$

- $[qXq] \longrightarrow \varepsilon$

Ejemplo 2: Determinar la GLC equivalente al siguiente AP

P:



$P = (\{p, q\}, \Sigma = \{0,1\}, \Gamma = \{X, Z_0\}, \delta, p, Z_0)$

(1) $\delta(p, 0, Z_0) = (p, XZ_0)$

(2) $\delta(p, 1, X) = (q, \varepsilon)$

(3) $\delta(q, 1, Z_0) = (q, \varepsilon)$

$L(P) = \{011\}$

- $S \longrightarrow [pZ_0q]$
- $S \longrightarrow [pZ_0p]$

(1) $\delta(p, 0, Z_0) = (p, XZ_0)$

$[pZ_0x] \longrightarrow 0[pXy][yZ_0x] \qquad \forall x, y \in Q$

- $[pZ_0q] \longrightarrow 0[pXq][qZ_0q]$
- $[pZ_0q] \longrightarrow 0[pXp][pZ_0q]$
- $[pZ_0p] \longrightarrow 0[pXq][qZ_0p]$
- $[pZ_0p] \longrightarrow 0[pXp][pZ_0p]$

(2) $\delta(p, 1, X) = (q, \varepsilon)$

$[pXx] \longrightarrow 1[q\varepsilon x] \qquad \forall x \in Q$

- $[pXq] \longrightarrow 1$

(3) $\delta(q, 1, Z_0) = (q, \varepsilon)$

$[qZ_0x] \longrightarrow 1[q\varepsilon x] \qquad \forall x \in Q$

- $[qZ_0q] \longrightarrow 1$

- $S \longrightarrow [pZ_0q]$
- $S \longrightarrow [pZ_0p]$

$[pZ_0q] = A$
$[pZ_0p] = B$

- $S \longrightarrow A$
- $S \longrightarrow B$

- $[pZ_0q] \longrightarrow 0[pXq][qZ_0q]$
- $[pZ_0q] \longrightarrow 0[pXp][pZ_0q]$
- $[pZ_0p] \longrightarrow 0[pXq][qZ_0p]$
- $[pZ_0p] \longrightarrow 0[pXp][pZ_0p]$

$[pXq] = C, [qZ_0q] = D$
$[pXp] = E$
$[qZ_0p] = F$

- $A \longrightarrow 0CD$
- $A \longrightarrow 0EA$
- $B \longrightarrow 0CF$
- $B \longrightarrow 0EB$

- $[pXq] \longrightarrow 1$

- $C \longrightarrow 1$

- $[qZ_0q] \longrightarrow 1$

- $D \longrightarrow 1$

- $S \longrightarrow A$
- $A \longrightarrow 0CD$
- $C \longrightarrow 1$
- $D \longrightarrow 1$

$L(G) = \{011\}$

Lema del Bombeo

para LIC

Teorema:

Sea $L \neq \emptyset$ un LIC. Entonces, existe una constante $n \in \mathbb{N}$, tal que si $z \in L$ con $|z| \geq$
z se puede dividir en 5 cadenas $z = uvwxy$, tal que:

$i)\ |vwx| \leq n$

$ii)\ vx \neq \varepsilon$

$iii)\ uv^iwx^iy \in L\ \forall i \geq 0$

Ejemplo 1. El lenguaje $L_{012} = \{0^m1^m2^m | m \geq 1\}$ no es un LIC.

Supongamos que L_{012} es un LIC, entonces por el lema del bombeo existe $n \in \mathbb{N}$, tal que si $z \in L_{012}$ con $|z| \geq n$, z se puede dividir en 5 cadenas $z = uvwxy$ tal que:

$i)|vwx| \leq n$ $\qquad ii)\ vx \neq \varepsilon$ $\qquad iii)\ uv^iwx^iy \in L\ \forall i \geq 0$

Sea $z = 0^n1^n2^n \in L_{012}$ con $|z| = 3n \geq n$, luego $z = uvwxy$ que cumplen $i), ii)$ y $iii)$ del lema del bombeo.

$$z = \overbrace{00\cdots0}^{n}\overbrace{11\cdots1}^{n}\overbrace{22\cdots2}^{n}$$

$$z = 00\cdots\underbrace{011}_{vwx}\cdots122\cdots2$$

Casos:

1) vwx está conformado únicamente de 0's
2) vwx está conformado únicamente de 1's
3) vwx está conformado únicamente de 2's
4) vwx está conformado únicamente de 0's y 1's
5) vwx está conformado únicamente de 1's y 2's

$I)$ Para los casos 1,2 y 3.

Supongamos sin pérdida de generalidad que vwx se compone únicamente de 0's.
entonces por $ii)$ del L.B. vx se compone únicamente de 0's y tiene al menos un 0.
Por otro lado, $uv^0wx^0y \in L_{012}$ por $iii)$ del L.B.
Sin embargo, $uv^0wx^0y = 0^k1^n2^n$ con $k < n \Rightarrow uv^0wx^0y \notin L_{012}$ $\#_c$

$II)$ Para los casos 4 y 5.

Supongamos sin pérdida de generalidad que vwx se compone únicamente de 0's y 1's.

entonces por $ii)$ del L.B. vx tiene al menos un 0 o al menos un 1.

Por otro lado, $uv^0wx^0y \in L_{012}$ por $iii)$ del L.B.

Sin embargo, $uv^0wx^0y = 0^k1^r2^n$ con $k < n$ o $r < n \Rightarrow uv^0wx^0y \notin L_{012}$

$\#_c$

De $I)$ y $II)$ se concluye que L_{012} no es un LIC. ■

Algunas propiedades de los LIC

Los Lenguajes Independientes de Contexto son cerrados bajo unión, concatenación y cerradura de Kleene.

$$G_1 = (V_1, T_1, P_1, S_1) \qquad G_2 = (V_2, T_2, P_2, S_2)$$

Unión: $G_3 = (V_1 \cup V_2 \cup \{S_3\}, T_1 \cup T_2, P_3, S_3)$

- $S_3 \longrightarrow S_1 | S_2$

concatenación: $G_4 = (V_1 \cup V_2 \cup \{S_4\}, T_1 \cup T_2, P_4, S_4)$

- $S_4 \longrightarrow S_1 S_2$

Cerradura de Kleene (de G_1): $G_5 = (V_1 \cup \{S_5\}, T_1, P_5, S_5)$

- $S_5 \longrightarrow S_1 S_5 | \varepsilon$