Bloque 2 TC

Eduardo Pérez Fraguela

Def.- Gramática Regular

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

- N Conjunto finito de símbolos NO TERMINALES o variables.
- Σ Conj. finito de símbolos NO TERMINALES o alfabeto
- P Conj. finito de reglas de escritura de la forma $A \to w$, $A \in N$ y $w \in (N \cup \Sigma)^*$:
  - w. tiene como máximo un símbolo terminal
  - Si w tiene un símbolo no terminal, entonces dicho símbolo es el último de la cadena.
- S Símbolo destacado de N, símbolo inicial o axioma de la gramática.

Ejemplo:

$G=(N=\{S,A\}, \Sigma=\{a,b\}, P=\{S\to bA, A\to aaA \mid b \mid \varepsilon\}, S=S)$

NOTA: Todo lo que esté en mayúsculas es NO TERMINAL y lo que esté en minúsculas es TERMINAL. S es siempre el axioma.

$S \Rightarrow bA \Rightarrow baaA \Rightarrow baaaaA \Rightarrow \dots \Rightarrow baaaa\dots aa$ o bien $baaaa\dots aab$

$S \overset{*}{\Rightarrow} b(aa)^*(b\cup\varepsilon)$

Def.- Lenguaje de una gramática

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

Teorema

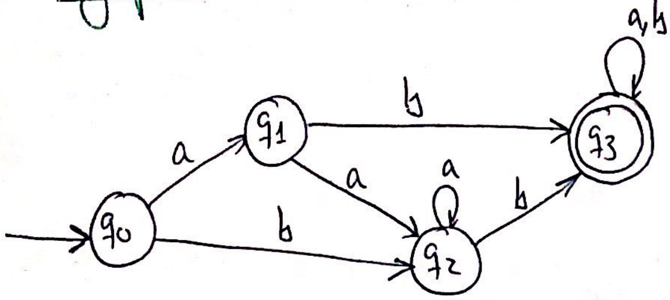
Dado cualquier L regular, $\exists$ G regular que lo genera.

Si L es regular, $\exists$ A.F $M=(Q,\Sigma,s,\delta,F)$; $L(M)=L$. Construimos $G=(N,\Sigma,P,S)$ regular; $L(G)=L$:

$N=Q \qquad \Sigma=\Sigma \qquad S=s$

$P=\{q_i \to aq_j \mid \delta(q_i,a)=q_j\}\cup\{q\to\varepsilon \mid q\in F\}$

Ejemplo:



$N=\{q_0,q_1,q_2,q_3\}$ $\Sigma=\{a,b\}$ $S=q_0$

$$P=\{q_0\to aq_1 \mid bq_2,$$
$$q_1\to aq_2 \mid bq_3,$$
$$q_2\to aq_2 \mid bq_3,$$
$$q_3\to aq_3 \mid bq_3 \mid \varepsilon\}$$

TEOREMA:

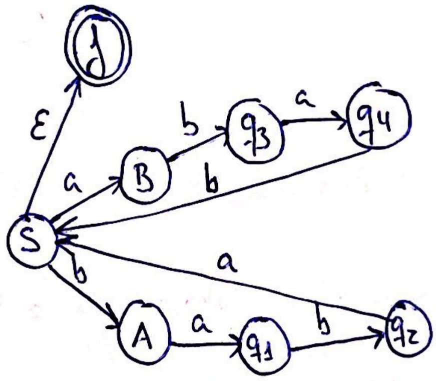
Cualquier G regular, $\exists$ un AFN M ; $L(M)=L(G)$

Ejemplo:

$S\to aB \mid bA \mid \varepsilon$

$A\to abaS$

$B\to babS$



Def.- Una G regular se dice que es:

- LINEAL POR LA IZQUIERDA si $A\to wB$
- LINEAL POR LA DERECHA si $A\to Bw$

$A,B\in N$ y $w\in\Sigma^+$

Def.- GIC | $G=(N,\Sigma,P,S)$ donde $P\subseteq N\times(N\cup\Sigma)^*$

Def.- LIC $\Rightarrow L(G)=\{w\in\Sigma^* \mid S\Rightarrow^* w\}$

NOTA: Toda G regular es GIC y todo L regular es LIC

Ejercicio: G dada por $S \to aSa \mid aSb \mid bSa \mid bSb \mid \varepsilon$ no es regular, pero $L(G)$ sí.

Diga cuál es ese lenguaje y obtenga una gramática regular G' tal que $L(G) = L(G')$

$L(G) = ((a \cup b)(a \cup b))^*$ y G' dada por $S \to aaS \mid abS \mid baS \mid bbS \mid \varepsilon$

Ejerc.: Obtener GIC que genere $\{a^m b^n \mid m \geq n\}$

$S \to aSb \mid aS \mid \varepsilon$

Ejercicio: GIC $\{a^m b^n \mid n \leq m \leq 2n\}$

$S \to ASb \mid \varepsilon$

$A \to a \mid aa$

Árboles de derivación y ambigüedad.

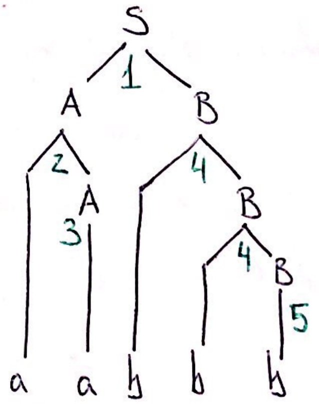
1) $S \to AB$ 2) $A \to aA$ 3) $A \to a$ 4) $B \to bB$ 5) $B \to b$

Por la izquierda: $S \overset{1}{\Rightarrow} \underline{A}B \overset{2}{\Rightarrow} a\underline{A}B \overset{3}{\Rightarrow} aa\underline{B} \overset{4}{\Rightarrow} aab\underline{B} \overset{4}{\Rightarrow} aabb\underline{B} \overset{5}{\Rightarrow} aabbb$

Por la derecha: $S \overset{1}{\Rightarrow} A\underline{B} \overset{4}{\Rightarrow} Ab\underline{B} \overset{4}{\Rightarrow} Abb\underline{B} \overset{5}{\Rightarrow} \underline{A}bbb \overset{2}{\Rightarrow} a\underline{A}bbb \overset{3}{\Rightarrow} aabbb$

Alternando: $S \Rightarrow A\underline{B} \overset{4}{\Rightarrow} \underline{A}bB \overset{2}{\Rightarrow} aAb\underline{B} \overset{4}{\Rightarrow} a\underline{A}bbB \overset{3}{\Rightarrow} aabb\underline{B} \overset{5}{\Rightarrow} aabbb$

Todas tienen el mismo árbol de derivación



La gramática es ambigua si hay 2 árboles o más de derivación

<u>Nota:</u> La ambigüedad se define para gramáticas no para lenguajes.

No obstante, $\exists$ lenguajes inherentemente ambiguos que cualquier gramática que los genere es siempre ambigua.

Ejercicio: Gramática que genere $\{a^i b^j c^k \mid i=j \text{ o } j=k\}$

$S \to AB \mid CD \quad A \to aAb \mid \varepsilon \quad B \to cB \mid \varepsilon \quad C \to aC \mid \varepsilon \quad D \to bDc \mid \varepsilon$

Forma normal de Chomsky (FNC)

GIC, G, se puede escribir otra G' equivalente con producciones de la forma

$A \to a$ o bien $A \to BC$. G' está en FNC.

Ejemplo: $S \to bA \mid aB \quad A \to bAA \mid aS \mid a \quad B \to aBBB \mid bS \mid b$

1º: $S \to C_bA \mid C_aB \quad A \to C_bAA \mid C_aS \mid a \quad B \to C_aBBB \mid C_bS \mid b \quad C_a \to a \quad C_b \to b$

Finalmente:

$S \to C_bA \mid C_aB$

$A \to C_bD_1 \mid C_aS \mid a \quad D_1 \to AA$

$B \to C_aD_2 \mid C_bS \mid b \quad D_2 \to BD_3 \quad D_3 \to BB$

$C_a \to a$

$C_b \to b$

LEMA DEL BOMBEO

Ejercicio: Demuestra mediante bombeo que $L = \{a^i b^i c^i \mid i \geq 0\}$ no es un LIC.

Suponemos que L es un LIC y sea K su cte. asociada según el lema del bombeo.

$z \in L \mid |z| \geq K$, por ejemplo $z = a^K b^K c^K$

Las posibles descomposiciones de z en uvwxy, tales que $|vwx| \leq K$ y $|v|+|x| > 0$ son:

- vwx compuesta solo por "aes". Cualquier bombeo uv^iwx^iy, $i>1$, produce cadenas con más "aes" que bes y ces, y que por tanto no pertenecen a L.
- vwx solo por "bes". Igual que el de antes pero con "bes"
- vwx solo por "ces". Igual al anterior pero con "ces".
- vwx compuesta por "aes y bes". Tres subcasos:
  - v formada solo por "aes", x solo por "bes". Al bombear aumentamos "aes" y "bes" pero nunca ~~ambas~~ "ces".
  - v formada por "aes" y "bes", x solo por "bes". Cualquier bombeo hará que se mezclen las "aes" y "bes".
  - v solo por "aes" y x por "aes" y "bes". Cualquier bombeo mezcla "aes" y "bes".
- vwx compuesto por "bes" y "ces". Similar al anterior, lo que ocurría con las "aes" y "bes" ahora ocurre con las "bes" y "ces".

La posición vwx no ~~tiene~~ puede contener "aes", "bes" y "ces" ya que tendría K "bes" y algún símbolo más, es decir, tendría longitud al menos K+2, y no verificaría la condición $|vwx| \leq K$.

Hemos analizado los casos de descomposición de z y, en todos ellos, hemos encontrado al menos un bombeo que produce cadenas fuera de L, con lo cual L no es un LIC.

Otra forma de resolverlo:

- $z = a^K b^K c^K$

- Analizamos todas sus posibles descomposiciones:

$$\underbrace{\overbrace{a \ldots \ldots a}^{K)}}_{1}\;\underbrace{\overbrace{b \ldots \ldots b}^{K)}}_{2}\;\underbrace{\overbrace{c \ldots \ldots c}^{K)}}_{3}$$

4   5

1) Cualquier bombeo $i \geq 2$ produce cadenas con más "aes" que "bes" y "ces".

2) Más "bes" que "aes" y "ces".

3) Más "ces" que "aes" y "bes".

4) $\ldots\ldots aaaaa \,|\, bbbbb$

1: v w x
2: v w x
3: v w x

4.1) Más "aes" y "bes" que "ces"

4.2)
4.3) } Lo mismo que 4.1 y "aes" y "bes" se mezclan.

5) $\ldots\ldots bbbbbccccc\ldots$

5.1) 5.2) 5.3)

5.1) Más "bes" y "ces" que "aes"

5.2)
5.3) } Lo mismo que 5.1 pero mezclado

Para todos los casos, hemos encontrado al menos un bombeo que produce cadenas $\notin L \Rightarrow L$ NO ES LIC.

Autómatas de Pila

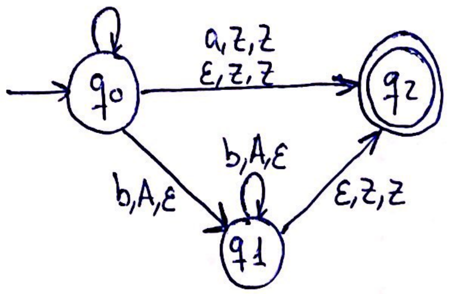
APN es $M = (Q, \Sigma, \Gamma, s, Z, F, \Delta)$

Γ es el alfabeto de los símbolos de la pila

$Z \in \Gamma$ símbolo inicial de la pila

Ejemplo: $\{a^n b^n \mid n \geq 0\} \cup \{a\}$:

$\left.\begin{array}{l} a, A, AA \\ a, Z, AZ \end{array}\right\}$ o bien a, ε, A

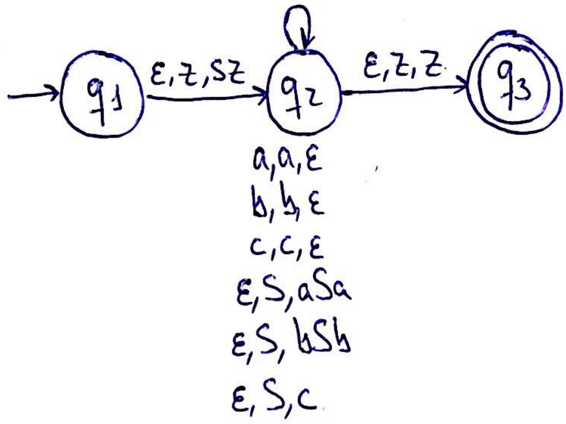


El lenguaje aceptado por un A.P está formado por las cadenas que al procesarse totalmente lo hacen llegar al estado final, pudiendo quedar vacía la pila o no.

Teorema: Dada una GIC $G = (N, \Sigma, P, S)$, siempre se puede construir un APN M tal que $L(M) = L(G)$

Ejemplo: APN a partir del GIC $S \to aSa \mid bSb \mid c$, abcba, deriva de la forma

$S \Rightarrow aSa \Rightarrow abSba \Rightarrow abcba$



$(q_1, abcba, Z)$
$\vdash (q_2, abcba, SZ)$
$\vdash (q_2, abcba, aSaZ)$
$\vdash (q_2, bcba, SaZ)$
$\vdash (q_2, bcba, bSbaZ)$
$\vdash (q_2, cba, SbaZ)$
$\vdash (q_2, cba, cbaZ)$
$\vdash (q_2, ba, baZ)$
$\vdash (q_2, a, aZ)$
$\vdash (q_2, \varepsilon, Z)$
$\vdash (q_3, \varepsilon, Z)$

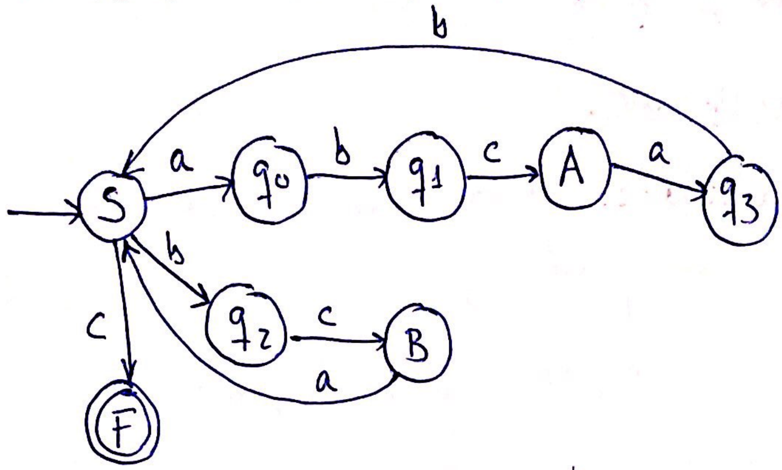
Autómata margarita, por el nº de lazos.

FNG o Forma Normal de Greibach

GIC está en FNG si sus reglas son de la forma $A \rightarrow a\alpha$, donde a es terminal y $\alpha \in N^*$.

Examen 20-21

1.- AF de $S \to abcA \mid bcB \mid c \quad A \to abS \quad B \to aS$



2.- Chomsky

$C_a D_3 \qquad C_b D_2$

$\equiv$

$C_a \to a$
$C_b \to b$
$C_c \to c$

$S \to C_a C_b C_c A \mid C_b C_c B \mid C_c$

$A \to C_a C_b S \equiv C_a D_1$

$B \to C_a S$

$D_1 \to C_b S$

$D_2 \to C_c B$

$D_3 \to C_b D_4$

$D_4 \to C_c A$

3.- $G_1: E \to E+T \mid T$
$T \to T \times F \mid F$
$F \to (E) \mid a$

$G_2: E \to E+E \mid E \times E \mid (E) \mid a \qquad w = a+a \times a$

¿G1 y G2 mismo lenguaje?

4.- GIC de $\{a^n b^m \mid n \neq m\}$

$S \rightarrow A \mid B$

$A \rightarrow aaAb \mid aaGb$

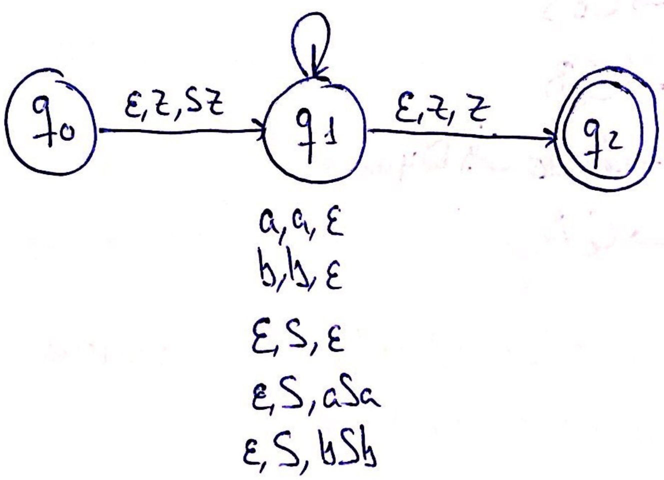
$B \rightarrow aBbb \mid aGbb$

$G \rightarrow aGb$

5.-

6.- $\{ww^I \mid w \in \{a,b\}^*\}$ AP

$S \rightarrow aSa \mid bSb \mid \varepsilon$



EX 20-21

5.-

a) GIC ¿cómo sabemos si $L(G)$ es vacío o no?

Seguir las producciones desde el axioma, si no hay ninguna regla de producción para el mismo, podremos afirmar que no genera ninguna cadena. A veces puede ser necesario aplicar transitividad, pero mientras desde S no se introduzca ningún símbolo terminal bastaría.

b) GIC ambigua ¿$L(G)$ es siempre un lenguaje ambiguo?

No, el contraejemplo G1 y G2 del ej 3.

En caso de que la gramática independiente del contexto genere un lenguaje regular, este siempre tendrá un AFN el cual se podrá trasformar en un AFD, el cual es siempre determinístico y no presenta ningún tipo de ambigüedad.

c) ¿K en bombeo para LIC?

Representa 2^n, donde n es nº de símbolos NO TERMINALES de la gramática en FNC que genera L en caso de que fuese independiente del contexto.

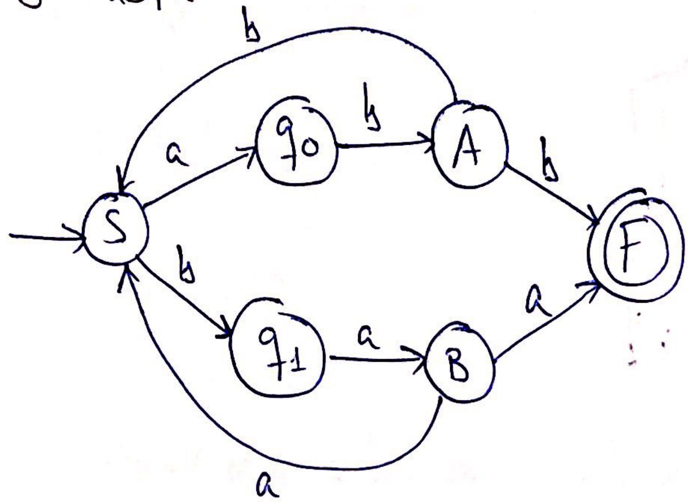
2^n sería la posible anchura de un árbol de derivación en el último nivel para una cadena que tiene todos los terminales sin repetirlos.

EXAMEN 2017-18

1.- $S \to abA \mid baB$

$A \to bS \mid b$

$B \to aS \mid a$



2.- Chomsky

$S \to C_aC_bA \mid C_bC_aB \Rightarrow C_aD_1 \mid C_bD_2$

$A \to C_bS \mid b$

$B \to C_aS \mid a$

$D_1 \to C_bA$

$D_2 \to C_aB$

$C_a \to a$

$C_b \to b$

3.- Demostrar que los LIC son cerrados para la unión.

L_1 y L_2 son LIC generados por $G_1=(N_1,\Sigma_1,P_1,S_1)$ y $G_2=(N_2,\Sigma_2,P_2,S_2)$

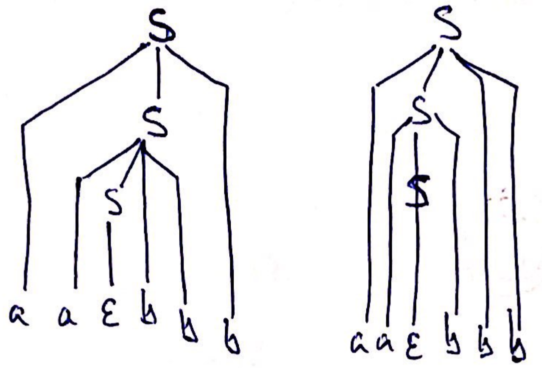
Se puede construir una nueva gramática $G=(N,\Sigma,P,S)$ tal que $L(G)=L_1 \cup L_2$.

$N=N_1 \cup N_2 \cup \{S\}$ $\Sigma=\Sigma_1 \cup \Sigma_2$ $P=P_1 \cup P_2 \cup \{S \to S_1 \mid S_2\}$ $S=S$

4.- Gramática que genere: $\{a^i b^j \mid i \leq j \leq 2i\}$ ¿es ambigua?

$$S \to aSb \mid aSbb \mid \varepsilon$$

La cadena aabbb tiene 2 árboles:



5.- Bombeo de $L = \{w \mid w \in \{a,b,c\}^*$ y aes = bes = ces sin importar el orden$\}$

Suponemos L un LIC y sea K su constante asociada según el lema del bombeo.

Tomemos una cadena $z \in L$ tal que $|z| \geq K$, $z = a^K b^K c^K$

z puede descomponerse en $uvwxy$, donde $|vwx| \leq K$ y $|v|+|x|>0$ y cualquier bombeo $uv^i w x^i y$ deberá estar en L, $\forall i \geq 0$

Las posibles descomposiciones de z son:

- vwx compuesto solo por aes. Cualquier bombeo $i \geq 2$, más aes que bes y ces $\notin L$
- vwx compuesto " " bes. " " más bes que aes y ces $\notin L$
- " " ces " " más ces que aes y bes $\notin L$
- vwx por aes y bes. " " más aes y bes que ces.
- vwx por bes y ces. " " más bes y ces que aes.
- Ninguna otra.

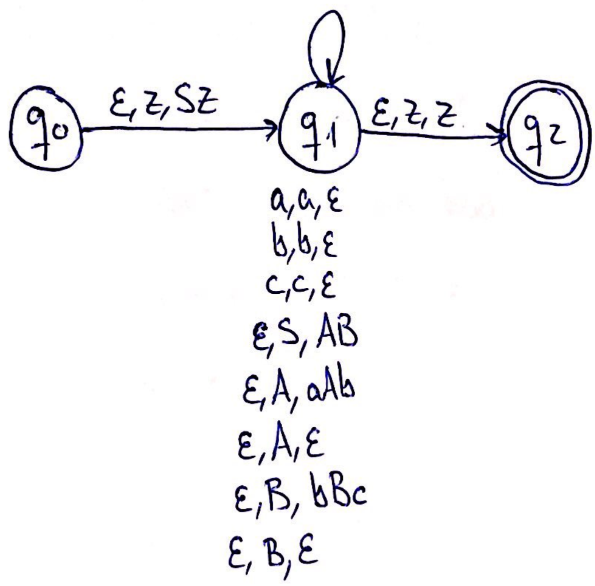
Tras analizar todos los tipos de descomposición de z y, en todos ellos, detectar al menos un bombeo que produce cadenas fuera de L, se concluye que L no es un LIC.

6.- ¿Qué es K?

La cte K del lema del bombeo es 2^n, donde n es el nº de símbolos NO terminales de la gramática en forma normal de Chomsky que generaría L, en caso de que L fuera independiente del contexto.

7.- AP $\{a^n b^{n+m} c^m \mid n,m \geq 0\}$

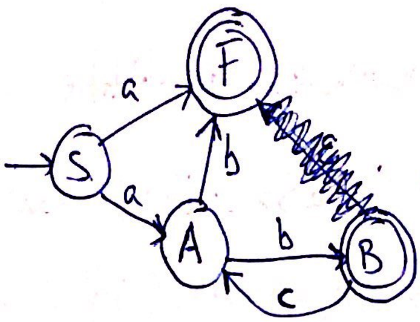
$$S \longrightarrow AB$$
$$A \longrightarrow aAb \mid \varepsilon$$
$$B \longrightarrow bBc \mid \varepsilon$$



Boletín 3 TC

41.- AFD

$S \rightarrow aA \mid a$
$A \rightarrow bB \mid b$
$B \rightarrow cA \mid \varepsilon$



42.- $\{a^n b^m \mid 0 \leq m \leq n \leq 3m\}$ Si $m=1 \Rightarrow n=2$ Si $m=5$ $n=6$

$a^2 b^1 = aab$ aaaaaabbbbb

$S \rightarrow aSb \mid aaSb \mid aaaSb \mid \varepsilon$

43.- uv, donde $u \in (a \cup b)^*$, $v \in (c \cup d)^*$ y $|u| \geq |v|$.

aac bbc aad bbd

$S \rightarrow USV \mid US \mid \varepsilon$

$U \rightarrow a \mid b$

$V \rightarrow c \mid d$

44.- GIC $\{a^i b^j \mid i \leq j \leq 2i\}$

$i=2 \Rightarrow a^2$ $j=3 \Rightarrow b^3 \Rightarrow aabbb$

$S \rightarrow aSb \mid aSbb \mid aSbbb \mid \varepsilon$

45.- $\{a^i b^j c^k \mid i+j=k\}$

$S \rightarrow aSc \mid B \mid \varepsilon$

$B \rightarrow bBc \mid \varepsilon$

46.- GIC para $\{a,b\}$ cadenas en las cuales la relación entre "aes" y "bes" es de 3 a 2 (abaab)

$$\cancel{S \rightarrow aaaSbb \mid aaSbbb \mid \varepsilon}$$

$$S \rightarrow SaSaSaSbSbS \mid SaSaSbSbSb \mid \dots \mid \varepsilon$$

47.- $\Sigma = \{p, q, r, V, F, \neg, \vee, \wedge, \Rightarrow, \Leftrightarrow, (,)\}$, L el conjunto de todas las expresiones lógicas válidas (ej: $p \wedge (q \Rightarrow F) \wedge \neg r$, etc). Escribir GIC tal que $L(G) = L$

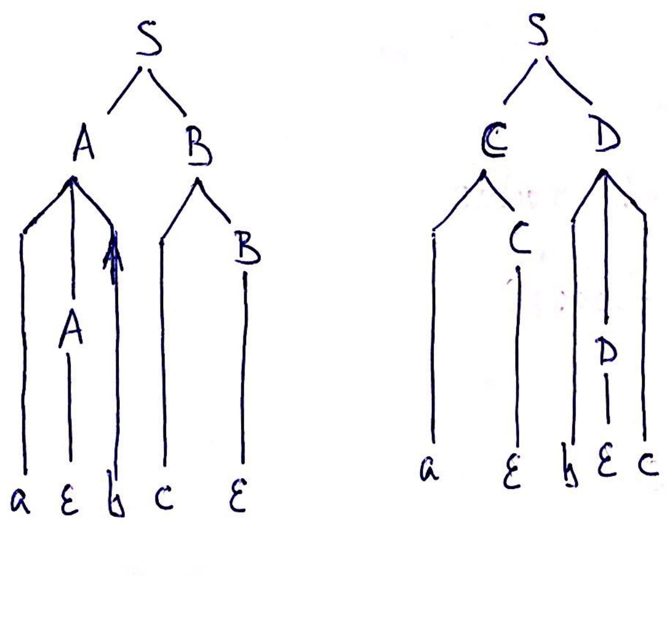
$$E \rightarrow \neg E \mid E \vee E \mid E \wedge E \mid E \Rightarrow E \mid E \Leftrightarrow E \mid (E) \mid V \mid F \mid p \mid q \mid r$$

48.- GIC ¿es ambigua? $\{a^i b^j c^k \mid i=j \text{ o } j=k\}$

$S \rightarrow AB \mid CD$

$A \rightarrow aAb \mid \varepsilon$ } $a^n b^n$ y ces libres
$B \rightarrow cB \mid \varepsilon$

$C \rightarrow aC \mid \varepsilon$ } $b^n c^n$ y aes libres.
$D \rightarrow bDc \mid \varepsilon$



Es ambigua.

49.- $L = \{a^n b^n c^m d^m \mid n,m \geq 0\} \cup \{a^n b^m c^m d^n \mid n,m \geq 0\}$, es decir, L consta de $a^*b^*c^*d^*$

- o bien #aes = #bes y #ces = #des
- o bien #aes = #des y #bes = #ces

GIC, ¿ambigua?

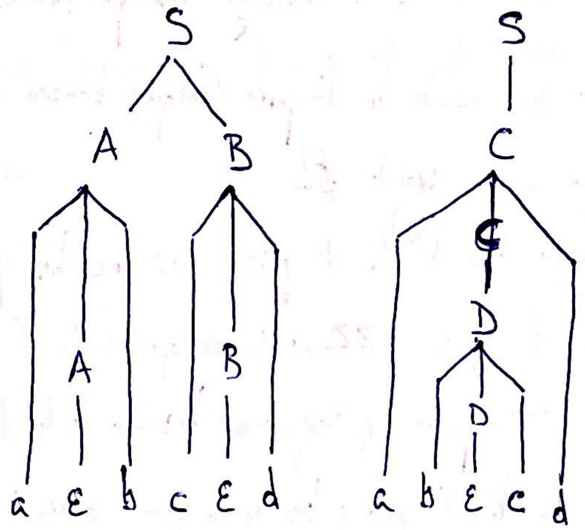
$S \to AB \mid C$

$A \to aAb \mid \varepsilon$

$B \to cBd \mid \varepsilon$

$C \to aCd \mid D$

$D \to bDc \mid \varepsilon$



L es ambigua

51.- Bombeo $\{a^i b^j c^k \mid i \leq j \leq k\}$

L es LIC y K cte. asociada según el lema del bombeo.

cadena $z \in L$ tal que $|z| \geq K$, $z = a^K b^K c^K$. Descomposiciones de z en $uvwxy$, tales que $|vwx| \leq K$ y $|v|+|x| > 0$:

- vwx compuesta solo por aes. Cualquier bombeo uv^iwx^iy, $i \geq 2$ produce cadenas con más aes que bes y ces. $\notin L$
- vwx sólo por bes. " " $i \geq 2$ más bes que ces (se puede), pero más bes que ces, entonces $\notin L$.
- vwx solo por ces. " " $i \geq 2$. Cualquier bombeo cumple. Pero $i=0$ produce una cadena con menos ces que aes y bes $\Rightarrow \notin L$.
- vwx compuesta por aes y bes, entonces:
  - subcadena v formada por aes, x solo por bes. $i \geq 2$ aumenta aes o bes o ambas, pero nunca las ces.

...

52.- Bombeo $\{ww \mid w \in \{a,b\}^*\}$ no es L.I.C

L es LIC y K la cte. asociada según el lema del bombeo.

$z \in L$ tal que $|z| \geq K \Rightarrow z = a^K b^K a^K b^K$.

Posibles descomposiciones de z en uvwxy tales que $|vwx| \leq K$ y $|v|+|x|>0$:

- vwx solo por aes de la 1ª parte. Cualquier bombeo $i \geq 2$, cadenas de más aes en la 1ª parte que en la segunda $\notin L$
- vwx solo por bes de la 1ª parte. $i \geq 2$, más bes que bes de la 2ª parte $\Rightarrow \notin L$
- vwx aes 2ª parte. $i \geq 2$, más aes que en la 1ª parte $\Rightarrow \notin L$
- vwx bes 2ª parte $i \geq 2$, más bes que en la 1ª parte $\Rightarrow \notin L$
- vwx aes de la 1ª parte y bes de la primer parte. $i=0$ desaparecen símbolos ya no obedece el formato ww
- vwx bes de la 1ª y aes de 2ª. Mismo razonamiento que antes
- vwx aes 2ª y bes 2ª. Mismo razonamiento que antes.
- Ninguna otra es posible.

En todos los casos de descomposición de z hemos detectado al menos un bombeo que produce cadenas fuera de L, L no es LIC

53.- Bombeo $\{w \mid w \in \{a,b,c\}^*$ y aes = bes = ces, sin importar el orden$\}$

L es un LIC y K es cte. asociada según el lema del bombeo.

Tomemos la cadena $z \in L$ tal que $|z| \geq K$ por ejemplo $a^K b^K c^K = z$. Las posibles descomposiciones $uvwxy$, tales que $|vwx| \leq K$ y $|v|+|x| > 0$ son:

- Cadena toda por aes. bombeo $i \geq 2$ cadenas con más aes que bes y ces $\notin L$
- " "bes" " "más bes que aes y ces $\notin L$.
- " "ces" " "más ces que aes y bes $\notin L$.
- Por cada vwx aes y bes $i \geq 2$ aumentaría aes y bes pero no ces
- vwx bes y ces $i \geq 2$. aumentaría bes y ces pero no ces
- Ninguna otra

L no es un LIC

56.- "Todo subconjunto de un lenguaje independiente del contexto es independiente del contexto".

$L_1 = \{a^i b^j c^j \mid i,j \geq 0\}$ es un LIC

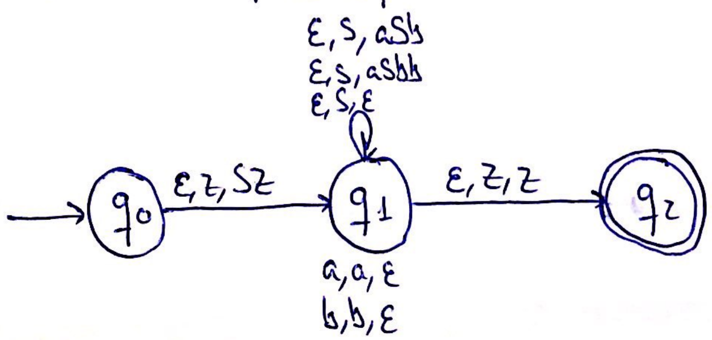
$L_2 = \{a^i b^i c^i \mid i \geq 0\}$ no es un LIC

Dado que $L2 \subseteq L1 \Rightarrow$ falso.

57.- AP $\{w \mid w \in \{a,b\}^*$ y w tiene el mismo nº de bes que de ces sin importar el orden$\}$

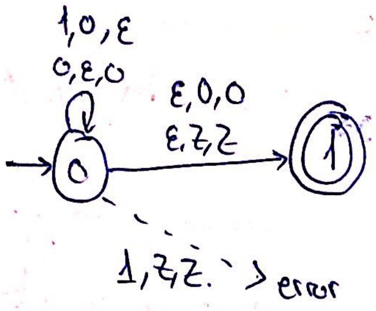
58.- AP $\{a^i b^j \mid i \leq j \leq 2i\}$

$S \rightarrow aSb \mid aSbb \mid \varepsilon$



59.- AP que acepte cadenas con ceros y unos tales que ningún prefijo tenga más unos que ceros.

- Los ceros siempre apilan ceros
- Los unos siempre desapilan un cero:
- Los unos no se admiten si la cima de la pila es Z.



60.- FNC

$$\left.\begin{array}{l} S \to aSb \mid cAd \\ A \to S \mid \varepsilon \end{array}\right\} S \to aSb \mid cSd \mid cd$$

$S \to C_a D_1 \mid C_c D_2 \mid C_c C_d$

$D_1 \to S C_b$

$D_2 \to S C_d$ — 1 NO TERMINAL → 2 NO TERMINALES

$C_a \to a$

$C_b \to b$ — 1 NO TERMINAL → 1 TERMINAL

$C_c \to c$

$C_d \to d$