

Introducción a los lenguajes y las gramáticas ✓

Lenguajes naturales ✓
Lenguajes formales ✓

En el software Mathematica:

$\left\{ \begin{array}{l} \text{http://library.wolfram.com/info center/MathSource/3128} \\ \text{http://library.wolfram.com/info center/MathSource/1808} \end{array} \right\}$ ✓
→ LB ✓

Lenguajes formales y gramáticas

Definición sea A un conjunto finito no vacío. Un lenguaje formal sobre A , denotado $L(A)$, es un subconjunto de A^* , A^* constituido por todas las cadenas de caracteres, que es posible construir con cada uno de los elementos de A . Al conjunto A se le denomina "alfabeto" del lenguaje.

Un ejemplo de lenguaje formal → Los lenguajes de autómatas de estado finito sobre Σ

Definición una gramática G es una 4-tupla $G = (V, T, P, S)$ con:

- ① V un conjunto de símbolos denominados: "símbolos no terminales".
- ② T un conjunto llamado: "símbolos terminales".
- ③ P un subconjunto finito de $\mathcal{P}((V \cup T)^* - T^*) \times ((V \cup T)^*)$, llamado "conjunto de composiciones" o "producciones" de la gramática. Una composición, por lo tanto, es un par ordenado donde la primera componente es una hilera de símbolos no terminales y/o terminales, con la restricción de contener al menos un símbolo no terminal y la segunda es una hilera de símbolos terminales y/o no terminales. Una producción (α, β) se denota $\alpha \rightarrow \beta$.
- ④ S es un símbolo no terminal inicial. $V^* = \Sigma$

Ejemplo sea la gramática $G = (V, T, P, S)$ con: $V = \{S, A, B, C, D, E\}$, $T = \{a, b, c\}$, $P = \{S \rightarrow aAB, S \rightarrow aB, A \rightarrow aAC, A \rightarrow aC, B \rightarrow Dc, AD \rightarrow Ab, C \rightarrow b, CB \rightarrow CE, CE \rightarrow DE, DE \rightarrow Dc\}$ y $S = S$. Verifique que la hilera $aaabbc$ es un elemento del lenguaje de G .

Solución:

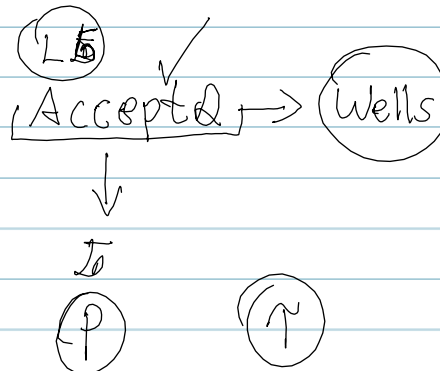
$$\begin{aligned}
 & \boxed{S \rightarrow aAB} \checkmark \\
 & \rightarrow aaACB \\
 & \rightarrow aaADc \\
 & \rightarrow aaAbc \\
 & \rightarrow aaAcbc \\
 & \rightarrow \boxed{aaabbc}
 \end{aligned}$$

$$\boxed{S \rightarrow aB} \checkmark$$

(P)

Definición Sea $G = (V, T, P, S)$ una gramática. Si $\alpha \rightarrow \beta \in P$ y $\gamma, \delta \in (V \cup T)^*$ entonces se dice que $\gamma\alpha\delta$ deriva a $\gamma\beta\delta$ y se denota: $\gamma\alpha\delta \rightarrow \gamma\beta\delta$. Además, cualquier hilera de símbolos de $(V \cup T)^*$ que se derive de $\gamma\alpha\delta$ se considera también, una derivación de $\gamma\alpha\delta$. El lenguaje generado por G , representado por $L(G)$, es el conjunto de todos los "strings" de símbolos terminales T que se derivan de S , a través de las composiciones de G .

En el software:



↓
 Ejemplo Dada la gramática $G = (V, \Sigma, P, S^*)$ con: $V = \{S, A, B, C, D\}$, $\Sigma = \{a, b, c\}$, $P = \{S \rightarrow aA, S \rightarrow aB, A \rightarrow aA, B \rightarrow Dc, C \rightarrow bC, D \rightarrow bD, D \rightarrow Dc\}$ establezca si las hilas "aa.bbbbbbb", "accccccc" y "aabbcbcc" son elementos de $L(G)$.

1) aa.bbbbbbb

$S \rightarrow aA$
 $\rightarrow aad$
 $\rightarrow aab$
 $\rightarrow aabbbbb$
 $\rightarrow aabbbbbbb$

2) accccccc

$S \rightarrow aB$
 $\rightarrow aDc$
 $\rightarrow aDcc$
 $\rightarrow aDccccc$
 $\rightarrow acccccc$

3) aabbcbcc

Por medio de software:

$G = \{S, A, B, C, D\}$
 $\Sigma = \{a, b, c\}$
 $P = \{S \rightarrow aA, S \rightarrow aB, A \rightarrow aA, B \rightarrow Dc, C \rightarrow bC, D \rightarrow bD, D \rightarrow Dc\}$

Accepted $[G, "aa.bbbbbbb"] \rightarrow \text{True}$

Accepted $[G, "accccccc"] \rightarrow \text{True}$

Accepted $[G, "aabbcbcc"] \rightarrow \text{False}$

Definición Sea $G = (V, \Sigma, P, \sigma^*)$ una gramática. Si cada composición $\alpha \rightarrow \beta_i$, $1 \leq i \leq n$, $n \in \mathbb{N}$ es sustituida por $\alpha ::= \beta_1 \mid \beta_2 \mid \dots \mid \beta_n$ y los símbolos no terminales se encierran en $\langle \rangle$ se dice que la gramática se ha expresado en notación de Backus-Naur, cuyo acrónimo es BNF.

Por ejemplo:

$$G = (V, \Sigma, P, \sigma^*)$$

$$V = \{S, A, B, C, D\}$$

$$\Sigma = \{a, b, c\}$$

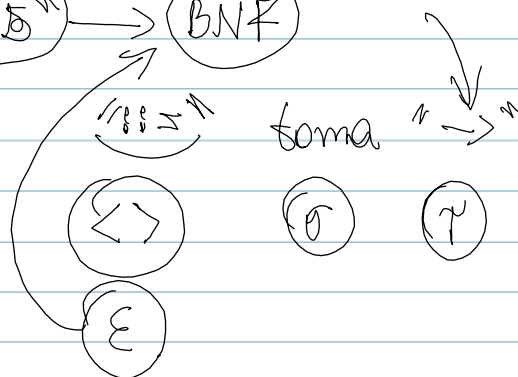
$$P = \{S \rightarrow aA, S \rightarrow aB, A \rightarrow aC, B \rightarrow Dc, C \rightarrow bC, C \rightarrow b, D \rightarrow c, D \rightarrow Dc\}$$

$$\sigma^* = S$$

En notación BNF:

$$\begin{aligned} \langle S \rangle &::= a \langle A \rangle \mid a \langle B \rangle \\ \langle A \rangle &::= a \langle C \rangle \\ \langle B \rangle &::= \langle D \rangle c \\ \langle C \rangle &::= b \langle C \rangle \mid b \\ \langle D \rangle &::= c \mid \langle D \rangle c \\ \sigma^* &= \langle S \rangle \end{aligned}$$

El paquete $\langle L_G \rangle \rightarrow \langle BNF \rangle$



BNF $\{ \{ S \rightarrow aA \mid aB, A \rightarrow aC, B \rightarrow Dc, C \rightarrow bC \mid b, D \rightarrow c \mid Dc \} \}$

$\{ \{ S, aA \}, \{ S, aB \}, \{ A, aC \}, \{ B, Dc \}, \{ C, bC \}, \{ C, b \}, \{ D, c \}, \{ D, Dc \} \}$

Definición SEA $G = (P, \Sigma, P, \sigma^*)$ una gramática y ϵ el arreglo de caracteres vacío o nulo, entonces:

1. Si toda producción de G tiene la forma: $\gamma A \delta \rightarrow \gamma B \delta$, con $A \in V$, $B \in (V \cup \Sigma)^*$ - ϵ y $\gamma, \delta \in (V \cup \Sigma)^*$, se dice que G y su lenguaje son sensibles al contexto.

2. Si toda composición de G tiene la forma: $A \rightarrow B$, con $A \in V$ y $B \in (V \cup \Sigma)^*$, G y su lenguaje son libres del contexto.

3. Si toda producción de G tiene la forma: $A \rightarrow aB$, $A \rightarrow a$, $A \rightarrow \epsilon$, con $A, B \in V$ y $a \in \Sigma$, se dice que G y su lenguaje son regulares.

Por ejemplo:

Sensible al contexto ✓
Libre del contexto ✓

Las funciones del paquete "LTL"

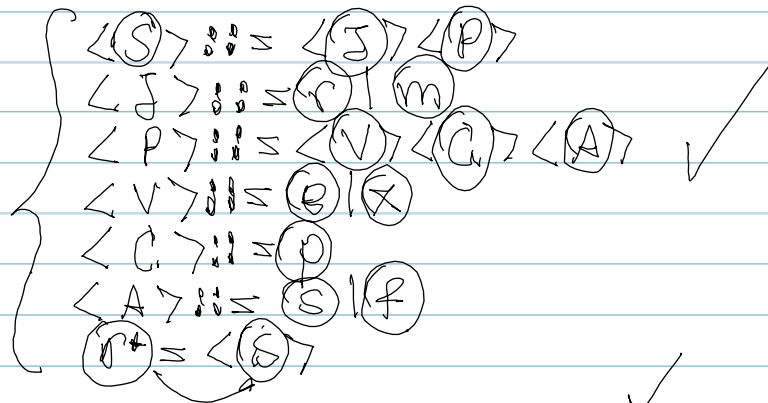
También en "LTL" :

derivations ✓ (n) (m)
words ✓ (n) (m)

Ejemplo Encuentre $L(G)$ dada la gramática libre del contexto:

$\langle \text{oración} \rangle ::= \langle \text{sujeto} \rangle \langle \text{predicado} \rangle$
 $\langle \text{sujeto} \rangle ::= \text{Roberto} \mid \text{María}$
 $\langle \text{predicado} \rangle ::= \langle \text{verbo} \rangle \langle \text{complemento} \rangle \langle \text{adverbio} \rangle$
 $\langle \text{verbo} \rangle ::= \text{estudia} \mid \text{examina}$
 $\langle \text{complemento} \rangle ::= \text{problemas}$
 $\langle \text{adverbio} \rangle ::= \text{siempre} \mid \text{frecuentemente}$
 $\Sigma^* = \langle \text{oración} \rangle$

con la intención de utilizar el comando words:



Al ensayar con los parámetros que usa words, se toma a $n=6$ y $m=4$. Luego:

$\ll L\delta \gg$

→ $\delta = \text{BNF} [d \ll s \rightarrow JP \gg, \ll J \rightarrow r \mid m \gg, \ll P \rightarrow VCA \gg, \ll V \rightarrow e \mid x \gg, \ll C \rightarrow p \gg, \ll A \rightarrow s \mid f \gg]$;
words [δ , (6), (4)]

$\{ \overline{mepf}, \overline{meps}, \overline{mxpf}, \overline{mxps}, \overline{repf}, \overline{reps}, \overline{rxpf}, \overline{rxps} \}$

La instrucción derivations del paquete $\ll L\delta \gg$:

Table [derivations [δ , (n 4), m (6)]]

$\{ d \{ JP \}, d \{ JVCA, mP, rP \},$
 $d \{ JeCA, JVcf, JVCs, JVpA, JxCA, mVCA, rVCA \},$
 $\{ Jecf, Jecs, JcpA, Jvpf, Jvps, Jxcf, Jxcs, JxpA, mecA,$
 $mVcf, mVcs, mVpA, mxCA, recA, rVcf, rVcs, rVpA, rxCA \},$
 $\{ Jepf, Jeps, Jxpf, Jxps, mecf, mecs, mepA, mVpf, mVps,$
 $mxcf, mxcs, mxpA, recf, recs, repA, rVpf, rVps,$
 $rcxf, rcxs, rxpA \},$
 $d \{ mepf, meps, mxpf, mxps, repf, reps, rxpf, rxps \} \}$

Ejemplo Verifique con ayuda de Mathematica que el lenguaje de la gramática dada a continuación, es el conjunto de los números reales.

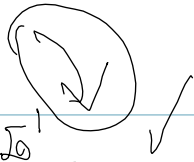
$$\begin{aligned} \langle \text{real} \rangle &::= \langle \text{real con signo} \rangle \mid \langle \text{real sin signo} \rangle \\ \langle \text{real con signo} \rangle &::= - \langle \text{parte real} \rangle \\ \langle \text{real sin signo} \rangle &::= \langle \text{parte real} \rangle \\ \langle \text{parte real} \rangle &::= \langle \text{entero sin signo} \rangle \mid \langle \text{parte decimal} \rangle \mid \langle \text{entero sin signo} \rangle \langle \text{parte decimal} \rangle \\ \langle \text{parte decimal} \rangle &::= . \langle \text{entero sin signo} \rangle \\ \langle \text{entero sin signo} \rangle &::= \langle \text{dígito} \rangle \mid \langle \text{dígito} \rangle \langle \text{entero sin signo} \rangle \\ \langle \text{dígito} \rangle &::= 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9 \\ \emptyset^* &= \langle \text{real} \rangle \end{aligned}$$

se debe reescribir la gramática:

$$\begin{aligned} \langle S \rangle &::= \langle E \rangle \mid \langle N \rangle \\ \langle E \rangle &::= - \langle P \rangle \\ \langle N \rangle &::= \langle P \rangle \\ \langle P \rangle &::= \langle E \rangle \mid \langle D \rangle \mid \langle E \rangle \langle D \rangle \\ \langle D \rangle &::= . \langle E \rangle \\ \langle E \rangle &::= \langle C \rangle \mid \langle C \rangle \langle E \rangle \\ \langle C \rangle &::= 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9 \\ \emptyset^* &= \langle S \rangle \end{aligned}$$

232.55 ✓

$$\begin{aligned} \langle S \rangle &::= \langle N \rangle \\ S &\rightarrow N \rightarrow P \rightarrow E D \\ &\rightarrow C E D \\ &\rightarrow C C E D \\ &\rightarrow C C C D \\ &\rightarrow C C C . E \\ &\rightarrow C C C . C E \\ &\rightarrow C C C . C C \\ &\rightarrow 232.55 \end{aligned}$$

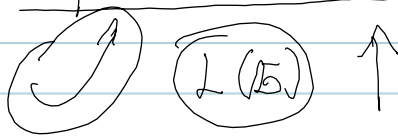


→ $\mathcal{G} = \text{BNF} [\{ "S \rightarrow E | N", "E \rightarrow -P", "N \rightarrow P", "P \rightarrow E | D | ED", "D \rightarrow .E", "E \rightarrow C | CE", "C \rightarrow 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9" \}]$;

bandera = 0;

→ For $i = 100, i \leq 119, \text{seedRandom}[i];$
 $\text{real} = \text{Random}[\text{Real}, \{0, 1000\}, \{5\}];$
 $\text{Print}["\text{En } n, \text{ real, " se obtiene: " }];$

→ $\text{v1} = \text{AcceptQ}[\mathcal{G}, \text{ToString}[\text{real}]];$
 $\text{IF} [\text{v1} == \text{False}, \text{bandera} = 1; \text{v1} = \text{real}; \text{Break}[]]; i++];$
 $\text{IF} [\text{bandera} == 0, \text{Print}["\text{La prueba fue exitosa"}]];$
 $\text{Print}["\text{La prueba no fue exitosa en el valor de " } \{P\}];$



Ejemplo sea la gramática $\mathcal{G} = (V, \Sigma, P, S)$ con: $V = \{S, A\}$,
 $\Sigma = \{a, b\}$, $P = \{S \rightarrow bS, S \rightarrow aA, A \rightarrow bA, A \rightarrow b\}$; ¿qué tipo
 de gramática es \mathcal{G} . Determine su lenguaje. ↑

\mathcal{G} es una gramática regular. Para encontrar $L(\mathcal{G})$:

→ \mathcal{G}

$\mathcal{G} = \text{BNF} [\{ "S \rightarrow bS | aA", "A \rightarrow bA | b" \}]$;

Table[derivations[\mathcal{G} , 1, 1], {m, 1}];

Table[words[\mathcal{G} , 1, 1], {n, 1}];

{daA, bS}, {ab, aA, baA, bbS},

⋮

daaaaaa, baaaaa, baabaaa, bbaabaaa, bbbabaaa, bbbbaaaa

{d}, {ab}, ...

$L(\mathcal{G}) = \{b^n a b^m / n, m \in \mathbb{N} \cup \{0\}, n \geq 0 \wedge m \geq 1\}$



Gramáticas regulares y autómatas

Teorema Sea $G = (V, \Sigma, P, \sigma^*)$ una gramática regular. El autómata de estado finito no determinístico $A = (Q \cup \{\epsilon A\}, \Sigma, \sigma, \epsilon A, \hat{A})$ con: $\epsilon A \notin (Q \cup \Sigma)$, $\hat{A} = \{\epsilon A\}$ y $\Delta(B, a) = \{D \in Q \mid B \rightarrow aD \in P\} \cup \{\epsilon A\}$, donde:

$$C = \begin{cases} \emptyset & \text{si } B \rightarrow a \notin P \\ \{\epsilon A\} & \text{si } B \rightarrow a \in P \end{cases}$$

es tal que: $A^0 = L(G)$

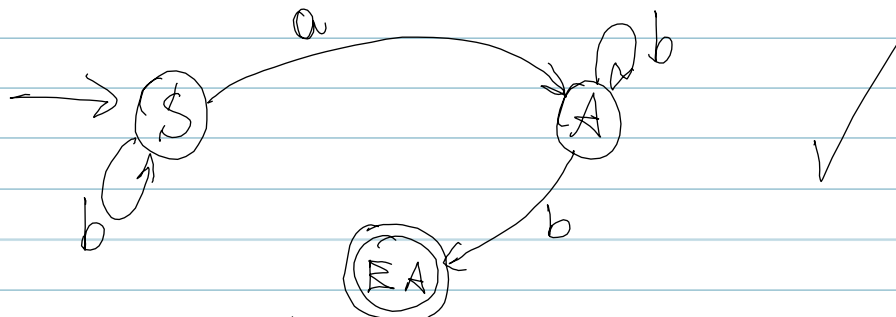
Ejemplo Determine $L(G)$ sobre la gramática regular dada, a través de un autómata de estado finito. Sea la gramática $G = (V, \Sigma, P, \sigma^*)$ con: $V = \{S, A\}$, $\Sigma = \{a, b\}$, $P = \{S \rightarrow bS, S \rightarrow aA, A \rightarrow bA, A \rightarrow b\}$, $\sigma^* = S$.

Solución:

$A = (Q \cup \{\epsilon A\}, \Sigma, \sigma, \epsilon A, \hat{A})$ donde:

$$\begin{array}{cccccc} \{S, A, \epsilon A\} & \Sigma = \{a, b\} & \sigma^* = S & \hat{A} = \{\epsilon A\} & & \\ \Delta & \begin{array}{cc} (S, a) & (S, b) \\ \{A\} & \{S\} \end{array} & \begin{array}{cc} (A, a) & (A, b) \\ \emptyset & \{A, \epsilon A\} \end{array} & \begin{array}{cc} (\epsilon A, a) & (\epsilon A, b) \\ \emptyset & \emptyset \end{array} & & \end{array}$$

Diagrama de transición:



$$A^0 = \{b^n a b^m \mid n, m \in \mathbb{N} \cup \{0\}, n \geq 0, m \geq 1\} = L(G)$$

Ejemplo Sea la gramática regular $G = (V, \Sigma, P, v)$ con: $V = \{S, A\}$,
 $\Sigma = \{a, b\}$, $P = \{S \rightarrow bS, S \rightarrow b, S \rightarrow aA, A \rightarrow aS, A \rightarrow bA, A \rightarrow a\}$.
 Halle la forma de los elementos de $L(G)$.

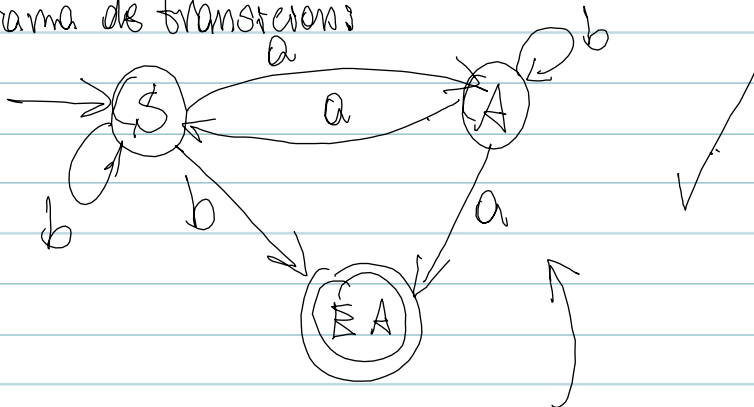
Solución

$A = (V \cup \{\epsilon\}, \Sigma, P, \hat{A})$ donde:

$\hat{A} = \{S, A, \epsilon\}$, $\Sigma = \{a, b\}$, $P = \{S \rightarrow bS, S \rightarrow b, S \rightarrow aA, A \rightarrow aS, A \rightarrow bA, A \rightarrow a\}$,
 $\Delta = \{A\}$

| | | | | | | |
|----------|----------|-------------------|-------------------|----------|-----------------|-----------------|
| | (S, a) | (S, b) | (A, a) | (A, b) | (ϵ, a) | (ϵ, b) |
| Δ | $\{A\}$ | $\{S, \epsilon\}$ | $\{S, \epsilon\}$ | $\{A\}$ | \emptyset | \emptyset |

Diagrama de transiciones:



$L(G) = \{b^n a^m \mid n, m \in \mathbb{N} \cup \{0\}, n \geq 0 \wedge m \text{ par}\} \cup$
 $\{ab^n amb^k \mid n, m, k \in \mathbb{N} \cup \{0\}, n \geq 0, m \text{ impar} \wedge$
 $k \geq 0\} \cup \{a\}$ / α es una hileras periódica que contiene a
 $b^n a b^m a$ con $n, m \in \mathbb{N} \cup \{0\}, n \geq 0, m \geq 0$

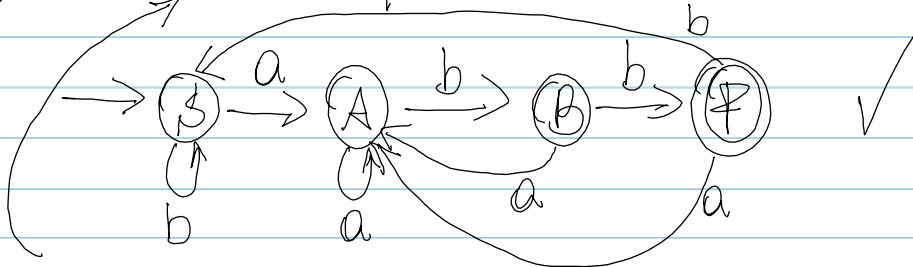
Otro teorema interesante:

$L(G) = A^0$ ✓

✓
Teorema Sea $\hat{A} = (\hat{Q}, \hat{Q}_0, \hat{Q}^*, \Delta, \hat{A})$ un autómata de estado finito determinístico. La gramática $\mathcal{G} = (\mathcal{Q}, \mathcal{T}, \mathcal{P}, \mathcal{Q}^*)$, con:
 $\mathcal{P} = \{ \hat{B} \rightarrow a\hat{D} \mid \Delta(\hat{B}, a) = \hat{D} \} \cup \{ \hat{B} \rightarrow \epsilon \mid \hat{B} \in \hat{Q}^* \}$, $\mathcal{Q}^* = \{ \hat{Q}^* \}$
 es regular, tal que: $L(\mathcal{G}) = \hat{A}^0$

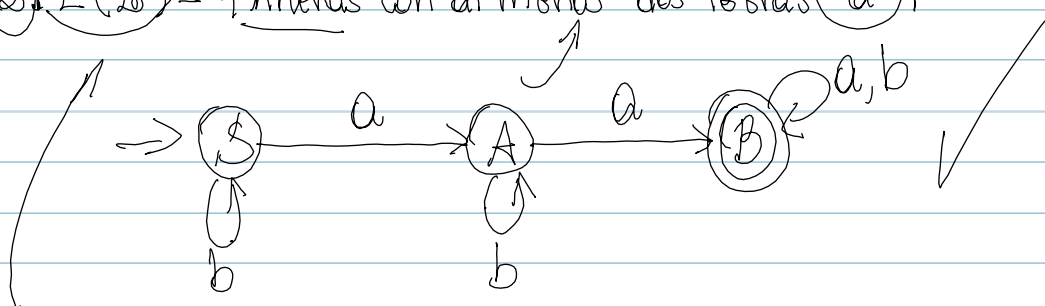
Ejemplo Diseñe una gramática regular \mathcal{G} que posea el lenguaje in dicado en cada caso.

① $L(\mathcal{G}) = \{ \text{hileras que finalicen en } \underline{abb} \}$



$\mathcal{G} = (\mathcal{Q}, \mathcal{T}, \mathcal{P}, \mathcal{Q}^*)$, $\mathcal{Q} = \{S, A, B, F\}$, $\mathcal{T} = \{a, b\}$, $\mathcal{Q}^* = \{F\}$,
 $\mathcal{P} = \{ S \rightarrow aA, S \rightarrow bS, A \rightarrow aA, A \rightarrow bB, B \rightarrow aA, B \rightarrow bF, F \rightarrow aA, F \rightarrow bS \}$

② $L(\mathcal{G}) = \{ \text{hileras con al menos dos letras } \underline{a}^n \}$



$\mathcal{G} = (\mathcal{Q}, \mathcal{T}, \mathcal{P}, \mathcal{Q}^*)$, $\mathcal{Q} = \{S, A, B\}$, $\mathcal{T} = \{a, b\}$, $\mathcal{Q}^* = \{B\}$,
 $\mathcal{P} = \{ S \rightarrow aA, S \rightarrow bS, A \rightarrow aA, A \rightarrow bA, B \rightarrow aB, B \rightarrow bB \}$

En Mathematica: $\{ \mathcal{G} \text{ RTD NDFAV}, \mathcal{G} \text{ DFATD ER} \}$ ✓