GRAMÁTICA

DEF: Uma GRAMÁTICA é uma quádrupla (T, V, S, R) em que

- T : conjunto finito de símbolos - TERMINAIS
- V : conjunto finito de símbolos - VARIÁVEIS
- $S \in V$: símbolo inicial
- R : conjunto de REGRAS

$T \cap V = \emptyset$

DEF: Uma REGRA tem o formato $u \to v$ em que

- u e v são palavras do alfabeto : $u, v \in (T \cup V)^*$
- u contém pelo menos um símbolo de V : $u \notin T^*$

DEF: DADA $G=(T,V,S,R)$ E $x,y \in (T \cup V)^*$,

DIZEMOS QUE y PODE SER DERIVADO EM UM PASSO A PARTIR DE x,

DENOTADO POR $x \Rightarrow y$, SE EXISTE REGRA $u \to v \in R$ T.Q.

$$x = x' u x'' \quad \text{E} \quad y = x' v x''$$

⤳ A DERIVAÇÃO (EM UM PASSO) SUBSTITUI UMA SUBPALAVRA.

DEF: UMA DERIVAÇÃO $x \Rightarrow^* y$ É UMA SEQUÊNCIA DE DERIVAÇÕES EM UM PASSO

↳ POSSIVELMENTE ZERO

$$x \Rightarrow y_1 \Rightarrow y_2 \Rightarrow \cdots \Rightarrow y$$

OBS: $x \Rightarrow^* x$

DEF: A linguagem gerada por G é o conjunto de palavras

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

↳ apenas com símbolos terminais.

DEF: Dizemos que uma gramática $G = (T, V, S, R)$ é regular se cada elemento de R é de um dos formatos a seguir

1. $X \to aY$ ⟶ no máximo uma variável

↳ $\in T$

2. $X \to a$

3. $X \to \varepsilon$

sempre uma variável isolada

em que $X, Y \in V$ e $a \in T$

EX: $G = (T, V, S, R)$, $T = \{0, 1\}$, $V = \{X, Y, Z\}$, $S = X$

$$R = \begin{cases} X \rightarrow 0X & 1 \\ X \rightarrow 1Y & 2 \\ Y \rightarrow 0 & 3 \\ Y \rightarrow 1X & 4 \\ Y \rightarrow 1Z & 5 \\ Z \rightarrow \varepsilon & 6 \end{cases}$$

$$\begin{array}{ll} X \rightarrow 1Y' & 2' \\ Y' \rightarrow 0 & 3' \\ Y' \rightarrow 1X & 4' \\ Y' \rightarrow 1Z & 5' \end{array}$$

QUAL É A LINGUAGEM GERADA POR G?

$$0011 \in L(G): X \overset{1}{\Rightarrow} 0X \overset{1}{\Rightarrow} 00X \overset{2}{\Rightarrow} 001Y \overset{5}{\Rightarrow} 0011Z$$

$$\overset{6}{\Rightarrow} 0011$$

GRAMÁTICAS REGULARES E AUTÔMATOS FINITOS

OBJ: MOSTRAR QUE AS LINGUAGENS GERADAS POR GRAMÁTICAS REGULARES SÃO PRECISAMENTE AS LINGUAGENS REGULARES.

1. DADA GRAM. REG. G, VAMOS CRIAR UM AFND QUE ACEITA $L(G)$

2. DADO AFD A VAMOS CRIAR GRAMÁTICA REGULAR G T.Q. $L(G) = L(A)$.

1) $G = (T, V, S, R)$ REGULAR

VAMOS CONSTRUIR UM AFND $A = (\Sigma, Q, q_0, F, \Delta)$ T.q. $L(A) = L(G)$

- $\Sigma = T$
- $Q = V \cup \{q_f\}$
- $q_0 = S$
- $F = \{q_f\}$
- Δ DEFINIDO POR

a) SE $X \to aY \in R$, ADICIONAMOS Y à $\Delta(X, \underset{\in T}{a})$

b) SE $X \to a \in R$, ADICIONAMOS q_f à $\Delta(X, a)$

c) SE $X \to \varepsilon \in R$, ADICIONAMOS q_f à $\Delta(X, \varepsilon)$

• Δ DEFINIDO POR

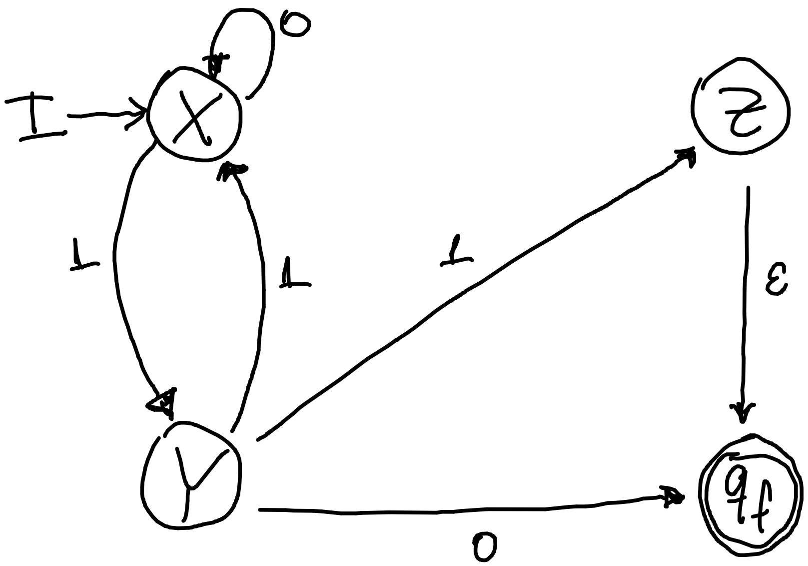
a) SE $X \to aY \in R$, ADICIONAMOS Y à $\Delta(X, a)$

b) SE $X \to a \in R$, ADICIONAMOS q_f à $\Delta(X, a)$

c) SE $X \to \varepsilon \in R$, ADICIONAMOS q_f à $\Delta(X, \varepsilon)$

Ex: $R = \{ X \xrightarrow{a} 0X,\ X \xrightarrow{a} 1Y,\ Y \xrightarrow{b} 0,\ Y \xrightarrow{a} 1X,\ Y \xrightarrow{a} 1Z,\ Z \xrightarrow{c} \varepsilon \}$

$S = X$



0011

$X \Rightarrow 0X \Rightarrow 00X \Rightarrow 001Y \Rightarrow 0011Z \Rightarrow 0011$

$(X, 0011) \vdash (X, 011) \vdash (X, 11) \vdash (Y, 1) \vdash (Z, \varepsilon) \vdash (q_f, \varepsilon)$

2) DADO AFD $A = (\Sigma, Q, q_0, F, \delta)$

VAMOS CONSTRUIR GRAM. REG. $G = (T, V, S, R)$ T.Q. $L(G) = L(A)$

- $T = \Sigma$
- $V = Q$
- $S = q_0$
- R :

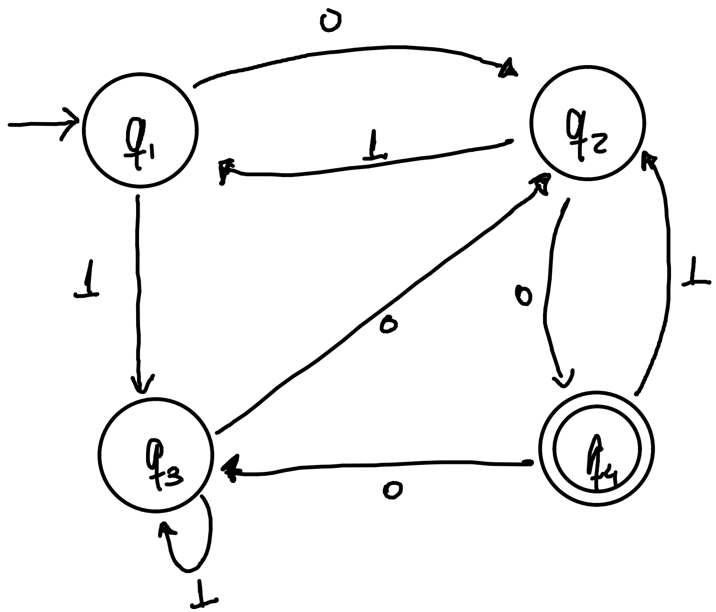
a) SE $q' = \delta(q, a)$, ADICIONAMOS $q \to a q'$ A R

b) SE $q \in F$, ADICIONAMOS $q \to \varepsilon$ A R

a) SE $q' = \delta(q, a)$, ADICIONAMOS $q \to a q'$ A R

b) SE $q \in F$, ADICIONAMOS $q \to \varepsilon$ A R

EX :



$T = \Sigma = \{0, 1\}$

$V = Q = \{q_1, q_2, q_3, q_4\}$

$S = q_0 = q_1$

$$
R = \begin{array}{c|ll}
 & a & b \\
1 & q_1 \to 0 q_2, & q_1 \to 1 q_3 \\
2 & q_2 \to 0 q_4, & q_2 \to 1 q_1 \\
3 & q_3 \to 0 q_2, & q_3 \to 1 q_3 \\
4 & q_4 \to 0 q_3, & q_4 \to 1 q_2 \\
 & q_4 \xrightarrow{c} \varepsilon &
\end{array}
$$

01100 :		
$(q_1, 01100)$	q_1	
$\vdash (q_2, 1100)$	$\Rightarrow 0 q_2$	1a
$\vdash (q_1, 100)$	$\Rightarrow 01 q_1$	2b
$\vdash (q_3, 00)$	$\Rightarrow 011 q_3$	1b
$\vdash (q_2, 0)$	$\Rightarrow 0110 q_2$	3a
$\vdash (q_4, \varepsilon)$	$\Rightarrow 01100 q_4$	2a

$\xRightarrow{c} 01100$

TEOREMA: AS SEGUINTES AFIRMAÇÕES SÃO EQUIVALENTES

1) L É UMA LINGUAGEM REGULAR
2) L É ACEITO POR UM AFD
3) L É ACEITA POR UM AFND
4) L PODE SER GERADA POR UMA EXPRESSÃO REGULAR
5) L PODE SER GERADA POR UMA GRAMÁTICA REGULAR

GRAMÁTICA LIVRE DE CONTEXTO

⤳ REGRAS MAIS FLEXÍVEIS QUE A GRAMÁTICA REGULAR

(GLC)

DEF: DIZEMOS QUE UMA GRAMÁTICA $G = (T, V, S, R)$ É LIVRE DE CONTEXTO SE TODAS AS SUAS REGRAS SÃO DO TIPO

$$X \longrightarrow w$$

COM $X \in V$ E $w \in (T \cup V)^*$

LEMBRE-SE QUE GRAM. REG:

$X \rightarrow aY$

$X \rightarrow a$

$X \rightarrow \varepsilon$

OBS: TODA GRAMÁTICA REGULAR É LIVRE DE CONTEXTO

NEM TODA GRAMÁTICA LIVRE DE CONTEXTO É REGULAR

EX: $R = \{ S \rightarrow 0S,\ S \rightarrow 1,\ X \rightarrow XX \}$

EX: G_1 com $T=\{0,1\}$, $V=\{S\}$ e $R=\{S\to 0S1,\ S\to\varepsilon\}$

$$L(G_1)=\{0^n1^n : n\in\mathbb{N}\}$$

EX: G_2 com $T=\{0,1\}$, $V=\{S\}$ e $R=\{S\to X1X,\ X\to 0X,\ X\to\varepsilon\}$

$$L(G_2)=\{0^n10^m : n,m\in\mathbb{N}\} = 0^*10^*$$

(LLC)

DEF: Uma linguagem L é livre de contexto se existe GLC G t.q. $L(G)=L$.

EX: $L(G_1)=\{0^n1^n : n\in\mathbb{N}\}$ é uma linguagem livre de contexto que não é regular

$0^{n_0}1^{n_0}$ $\rightsquigarrow$ 1) $0^{n_0+m}1^{n_0}$ $\notin L(G_1)$

2) $\cancel{0^{n_0}1^{n_0+m}}$

$w=xyz$

$x\underline{y}^kz\in L$

3) $\cancel{0^{m_1}0^{m_2}1^{m_3}0^{m_c}1^{m_3}1^{m_4}}$

$000\underbrace{0\,1\,1}_{w}1\,1$

$00\,00\,11\,0011\,11$

TEO: A classe das linguagens regulares é um subconjunto próprio da classe das LLCs

EX: G_2 com $T = \{0, 1\}$, $V = S$ e $R = \{S \xrightarrow{(1)} X1X,\ X \xrightarrow{(2)} 0X,\ X \xrightarrow{(3)} \varepsilon\}$

$$L(G_2) = \{0^n 1 0^m : n, m \in \mathbb{N}\} = 0^* 1 0^*$$

$$0^2 1 \in L(G_2)$$

$$S \Rightarrow \underline{X}1X \overset{(2)}{\Rightarrow} 0\underline{X}1X \overset{(2)}{\Rightarrow} 00\underline{X}1X \overset{(3)}{\Rightarrow} 001\underline{X} \overset{(3)}{\Rightarrow} 001$$

$$\Rightarrow X1\underline{X} \overset{(3)}{\Rightarrow} \underline{X}1 \overset{(2)}{\Rightarrow} 0\underline{X}1 \overset{(2)}{\Rightarrow} 00X1 \overset{(3)}{\Rightarrow} 001$$

↳ indica em qual variável aplicamos a regra

GRAMÁTICAS QUE NÃO SÃO LIVRES DE CONTEXTO

EX: $T = \{a, b, c\}$, $V = \{S, X, Y\}$

$$R = \left\{ \begin{array}{ll} 1\ S \to abc & 5\ S \to aXbc \\ 2\ Xb \to bX & 6\ Xc \to Ybc^2 \\ 3\ bY \to Yb & 7\ aY \to a^2 \\ 4\ aY \to a^2X & \end{array} \right\}$$

TEMOS $a^2b^2c^2 \in L(G)$

$$S \overset{5}{\Rightarrow} a\underline{Xb}c \overset{2}{\Rightarrow} ab\underline{Xc} \overset{6}{\Rightarrow} a\underline{bY}bc^2$$

$$\overset{3}{\Rightarrow} \underline{aY}b^2c^2 \overset{7}{\Rightarrow} a^2b^2c^2$$

TEMOS QUE $a^nb^nc^n \in L(G)\ \forall n$

CONJ. $L(G) = \{a^nb^nc^n : n \in \mathbb{N}\}$

EX: GRAMÁTICA QUE GERA UMA FÓRMULA LEGÍTIMA,

$G_{exp} = (T, V, S, R)$ COM $T = \{id, +, *, (,)\}$, $V = \{E\}$, $S = E$

REGRAS :

$E \to E + E$

$E \to E * E$

$E \to (E)$

$E \to id$

$3 \cdot x + 7$

$3 \cdot y + 7$

$7 \cdot y + 3$

→ NÃO PRECISAMOS CONHECER AS VARIÁVEIS, APENAS SUAS POSIÇÕES

→ NÃO QUEREMOS EXPRESSÕES DO TIPO $\underline{(+\underline{id})\,\underline{id}\,*} \in T^*$

→ O MESMO SÍMBOLO id PARA VARIÁVEIS

REGRAS :

$E \rightarrow E + E$

$E \rightarrow E * E$

$E \rightarrow (E)$

$E \rightarrow id$

AMBÍGUA

$7 + 8 \cdot 52$

⇝ DERIVAR $id + id * id$

$E \Rightarrow E + E \Rightarrow E + E * E \Rightarrow id + E * E \Rightarrow id + id * E \Rightarrow id + id * d$

$E \Rightarrow E * E \Rightarrow E + E * E \Rightarrow id + E * E \Rightarrow id + id * E \Rightarrow id + id * d$

REGRAS :

$$E \to (E + E)$$

$$E \to (E * E)$$

$$E \to id$$

AMBÍGUA

$$\underline{7 +}\ 8 \cdot 52$$

~> DERIVAR $\cancel{id + id * id}$ ~> $((id + id) * id)$ ou $(id + (id * id))$

$$E \Rightarrow E + \underline{E} \Rightarrow \underline{E} + E * E \Rightarrow id + \underline{E} * E \Rightarrow id + id * \underline{E} \Rightarrow id + id * d$$

$$E \Rightarrow \underbrace{\underline{E} * E} \Rightarrow \underbrace{E + E} * E \Rightarrow id + \underline{E} * E \Rightarrow id + id * \underline{E} \Rightarrow i \quad d * d$$

DEF: SEJA G UMA GLC E $w \in L(G)$.

UMA DERIVAÇÃO MAIS À ESQUERDA DE w EM G É AQUELA NA QUAL EM CADA PASSO, A VARIÁVEL À QUAL A REGRA FOI APLICADA É A VARIÁVEL MAIS À ESQUERDA DA PALAVRA.

ANALOGAMENTE, PODEMOS DEFINIR DERIVAÇÃO MAIS À DIREITA

EX: PALÍNDROMOS

DEF: UM PALÍNDROMO É UMA PALAVRA w CUJO REFLEXO w^R É IGUAL A w.

EX: 101 , 00 111001001111

OBS:

1) w É UM PALÍNDROMO SE COMEÇA E TERMINA COM O MESMO SÍMBOLO σ, E

2) $w = \sigma w' \sigma$ EM QUE w' É UM PALÍNDROMO

$G_{Pal} = (T, V, S, R)$ COM $T = \{0, 1\}$, $V = \{S\}$, E

$$R = \{S \to 0S0,\ S \to 1S1,\ S \to \varepsilon,\ S \to 0,\ S \to 1\}$$

COMPACTA : $S \to 0S0 \mid 1S1 \mid \varepsilon \mid 0 \mid 1$

OBS: COMO GERAR APENAS PALÍNDROMOS DE COMPRIMENTO PAR?

FECHAMENTO LLC

SUPONDO $V_1 \neq V_2$

1. <u>UNIÃO</u>: SEJAM L_1 E L_2 DUAS LLCs.

SEJA $G_i = (T_i, V_i, S_i, R_i)$ A GRAMÁTICA QUE GERA L_i, $i=1,2$.

TOME $G_U = (T_U, V_U, S_U, R_U)$ EM QUE

$T_U = T_1 \cup T_2$

$V_U = V_1 \cup V_2 \cup \{S_U\}$

$R_U = R_1 \cup R_2 \cup \{S_U \to S_1 \mid S_2\}$

SE $w_i \in L_i$. ENTÃO $S_i \overset{R_i}{\Rightarrow}{}^{*} w_i$. LOGO $S_U \Rightarrow S_i \overset{R_i}{\Rightarrow} w_i$, $i=1,2$

LOGO $L(G_U) = L_1 \cup L_2$

2.

CONCATENAÇÃO

$G_\bullet = (T_\bullet, V_\bullet, S_\bullet, R_\bullet)$

SUPONDO $V_1 \neq V_2$

$T_\bullet = T_1 \cup T_2$

$V_\bullet = V_1 \cup V_2 \cup \{S_\bullet\}$

$R_\bullet = R_1 \cup R_2 \cup \{S_\bullet \to S_1 S_2\}$

$w_1 \in L_1$

$w_2 \in L_2$

$w_1 w_2 \in L_\bullet = L_1 L_2$

3.

ESTRELA DE KLEENE

$G_* = (T_*, V_*, S_*, R_*)$

$T_* = T_1$, $V_* = V_1 \cup S_*$

$R_* = R_1 \cup \{S_* \to S_1 S_* \mid \varepsilon\}$

$w \in L_1$ $\quad w_1, w_2, \ldots, w_k \in L$

$w_* \in L_{1*}$ $\quad w_1 w_2 \cdots w_k \in L_*$

$\varepsilon, w, ww, w \cdots w$

LEMBRE-SE QUE $\{a^n b^n c^n : n \in \mathbb{N}\}$ NÃO É LLC

4. <u>INTERSEÇÃO</u>: NÃO É VERDADE QUE SE L_1 E L_2 SÃO LLC,

ENTÃO $L_1 \cap L_2$ É LLC.

EX: $L_1 = \{a^n b^n c^m : n, m \in \mathbb{N}\}$
$L_2 = \{a^m b^n c^n : n, m \in \mathbb{N}\}$ SÃO LLC

$L_1 \cap L_2 = \{a^n b^n c^n : n \in \mathbb{N}\}$ QUE NÃO É LLC

OBS: JÁ SABEMOS QUE $L_{ab} = \{a^n b^n : n \in \mathbb{N}\}$ E $L_{bc} = \{b^n c^n : n \in \mathbb{N}\}$
SÃO LLC.

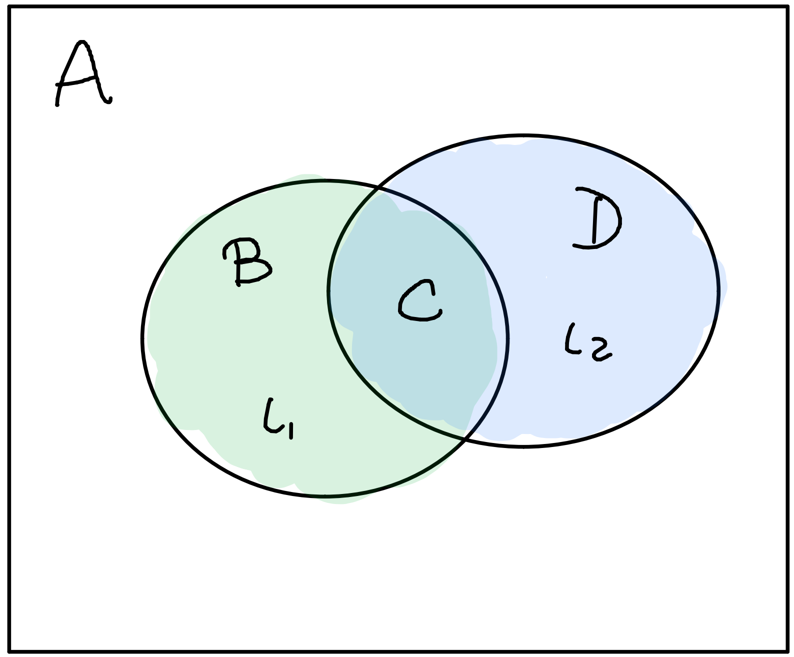
LOGO $L_{ab} \cdot \{c\}^*$ E $\{a\}^* \cdot L_{bc}$ SÃO LLC.

COMPLEMENTO: TAMBÉM NÃO VALE.

LEMBRE-SE QUE $L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$

SUPONHA QUE O COMPLEMENTO DE UMA LLC É LLC, ENTÃO

1) $\overline{L_1}, \overline{L_2}$ É LLC
2) LOGO, A UNIÃO $\overline{L_1} \cup \overline{L_2}$ É LLC
3) FINALMENTE $\overline{\overline{L_1} \cup \overline{L_2}}$ É LLC



$\{0,1\}^*$

$\overline{L_1} = A \cup D$

$\overline{L_2} = A \cup B$

$\overline{L_1} \cup \overline{L_2} = A \cup B \cup D$

$\overline{\overline{L_1} \cup \overline{L_2}} = C = L_1 \cap L_2$

DIFERENÇA: TAMBÉM NÃO É VERDADE.

SUPONHA QUE PARA TODAS LLCs L_3 E L_4 TEMOS QUE $L_3 \setminus L_4$ É LLC.

TOME $L_3 = \Sigma^*$, QUE É REGULAR E, PORTANTO, É LLC.

E NOTE QUE $\overline{L_4} = L_3 \setminus L_4$

Árvores de Análise Sintática

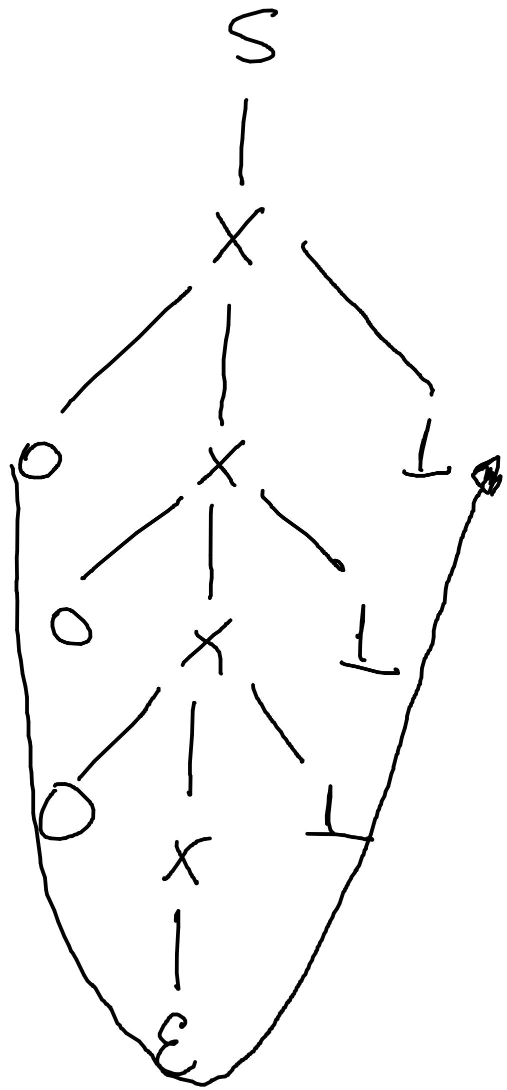
Ex: $S \rightarrow X$

$X \rightarrow 0X1 \mid \varepsilon$

$\omega = 000111$

$S \Rightarrow X \Rightarrow 0X1 \Rightarrow 00X11$

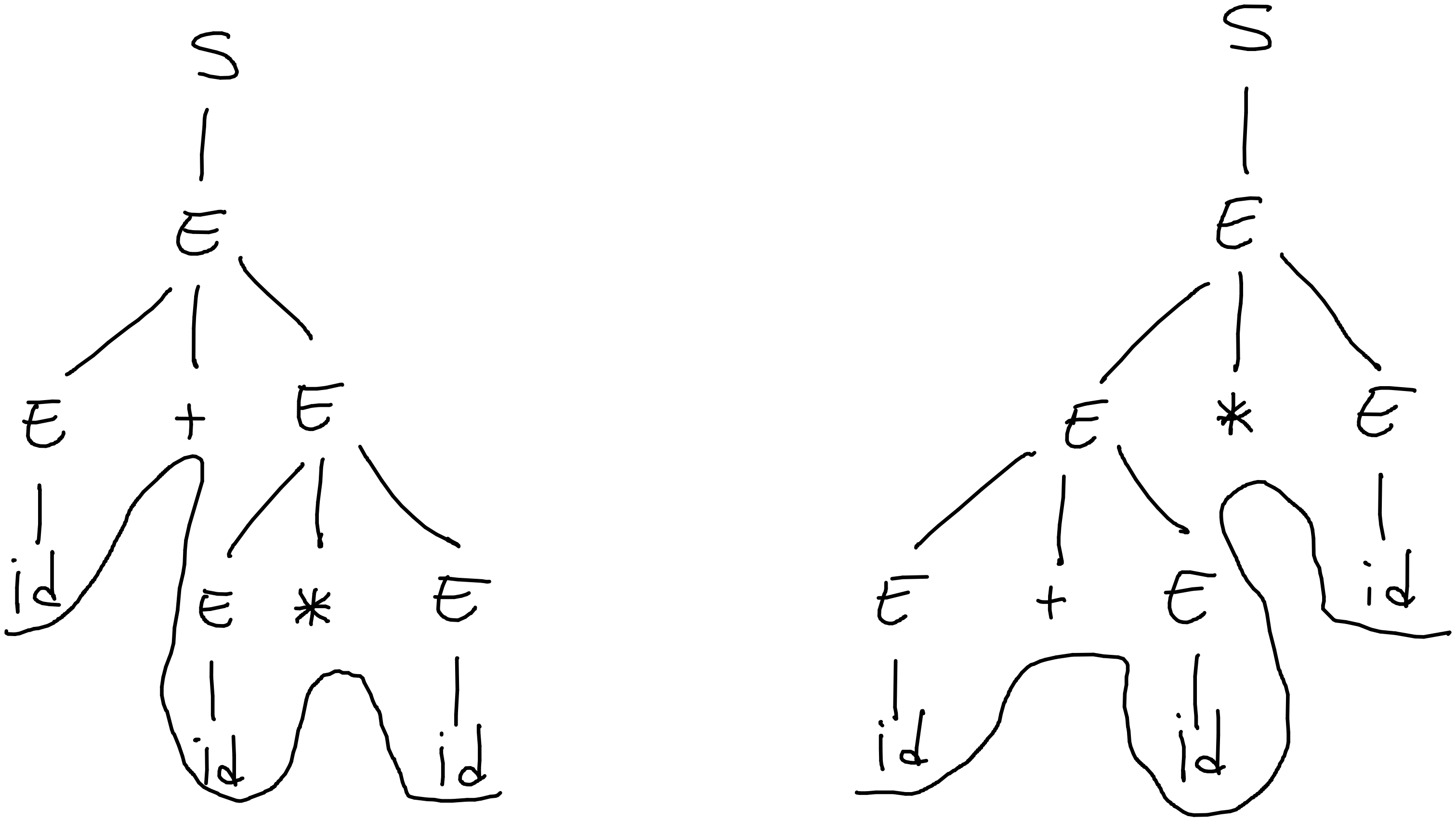
$\Rightarrow 000X111 \Rightarrow 000111$



Ex: $S \to E$

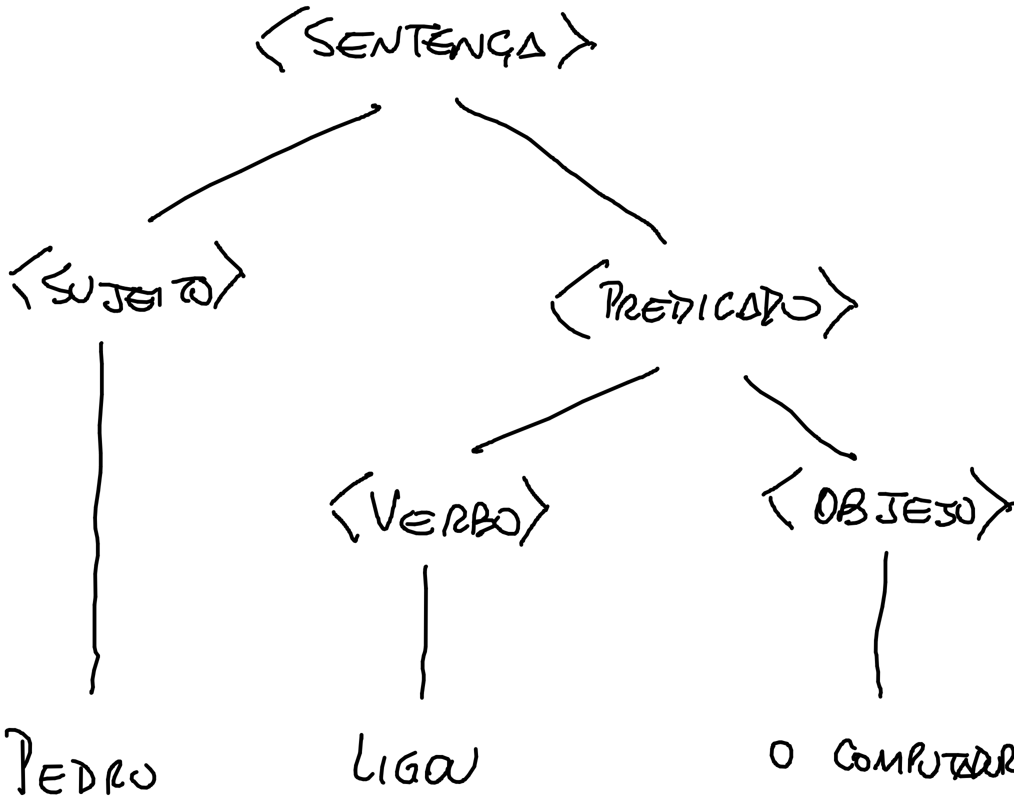
$E \to E+E \mid E*E \mid id$

$w = id + id*id$



- DIAGRAMAR UMA LINGUAGEM.
- GRAMÁTICAS GERAM FRASES COM SINTAXE CORRETA
- GOSTARÍAMOS DE ANALISAR A ESTRUTURA DE UMA FRASE E DETERMINAR O SEU SIGNIFICADO

EX: $\underbrace{\text{PEDRO}}_{\text{SUJEITO}}\ \underbrace{\underbrace{\text{LIGOU}}_{\text{VERBO}}\ \underbrace{\text{O COMPUTADOR}}_{\text{OBJETO}}}_{\text{PREDICADO}}$



- CADA BIFURCAÇÃO CORRESPONDE A UMA REGRA

<SENTENÇA> ⟶ <SUJEITO><PREDICADO>
<SUJEITO> ⟶ PEDRO
<PREDICADO> ⟶ <VERBO><OBJETO>
<VERBO> ⟶ LIGOU
<OBJETO> ⟶ O COMPUTADOR

DEFINIÇÃO FORMAL

- ÁRVORE : GRAFOS SEM CICLOS
- AQUI : "ORIENTADOS" , SEM SETAS, "DE CIMA PARA BAIXO"

1) HÁ ÚNICO VÉRTICE CHAMADO DE RAIZ

2) OS SUCESSORES DE UM VÉRTICE SÃO TOTALMENTE ORDENADOS
"DA ESQUERDA PARA DIREITA"

↳ PORQUE QUEREMOS ESCREVER DA ESQUERDA P/ DIREITA

DEF: SEJA V' UM VÉRTICE DA ÁRVORE. HÁ UM ÚNICO CAMINHO QUE LIGA ELE À RAIZ.

SEJA V UM VÉRTICE NESTE CAMINHO

- V É ASCENDENTE DE V', E V' É DESCENTENTE DE V
- SE V E V' ESTÃO SEPARADOS POR UMA ÚNICA ARESTA ENTÃO V É PAI DE V' E V' É FILHO DE V
- FILHOS DO MESMO PAI SÃO IRMÃOS
- VÉRTICE SEM FILHOS É UMA FOLHA
- VÉRTICE QUE NÃO É FOLHA É INTERIOR

⇝ A ORDENAÇÃO ENTRE IRMÃOS PODE SER ESTENDIDA PARA UMA ORDENAÇÃO DAS FOLHAS.