

# 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  $u \rightarrow v$  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  $u \rightarrow v \in R$  T.Q.

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

$\leadsto$  A DERIVAÇÃO (EM UM PASSO) SUBSTITUI UMA SUBPALAVRA.

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

$\hookrightarrow$  POSSIVELMENTE ZERO

$$x \Rightarrow y_1 \Rightarrow y_2 \Rightarrow \dots \Rightarrow y$$

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.  $X \rightarrow aY$  ↳ NO MÁXIMO UMA VARIÁVEL  
↳  $a \in T$
2.  $X \rightarrow a$  SEMPRE UMA VARIÁVEL ISOLADA
3.  $X \rightarrow \epsilon$

EM QUE  $X, Y \in V$  e  $a \in T$

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

$R = \left\{ \begin{array}{ll} X \rightarrow 0X & 1 \\ X \rightarrow 1Y & 2 \\ Y \rightarrow 0 & 3 \\ Y \rightarrow 1X & 4 \\ Y \rightarrow 1Z & 5 \\ Z \rightarrow \varepsilon & 6 \end{array} \right.$

Qual é a linguagem gerada por G?

$0011 \in L(G) : X \xrightarrow{1} 0X \xrightarrow{1} 00X \xrightarrow{2} 001Y \xrightarrow{5} 0011Z \xrightarrow{6} 0011$

# GRAMÁTICAS REGULARES E AUTÔMATOS FINITOS

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

1. DADO 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.

$$L(A) = L(G)$$

•  $\Sigma = T$

•  $Q = V \cup \{q_f\}$

•  $q_0 = S$

•  $F = \{q_f\}$

•  $\Delta$  DEFINIDO POR

a) se  $X \rightarrow aY \in R$ , ADICIONAMOS

$$Y \xrightarrow{\Delta} \Delta(X, a)$$

b) se  $X \rightarrow a \in R$ , ADICIONAMOS

$$q_f \xrightarrow{\Delta} \Delta(X, a)$$

c) se  $X \rightarrow \varepsilon \in R$ , ADICIONAMOS

$$q_f \xrightarrow{\Delta} \Delta(X, \varepsilon)$$

$\varepsilon^T$   
 $\downarrow$

•  $\Delta$  DEFINIDO POR

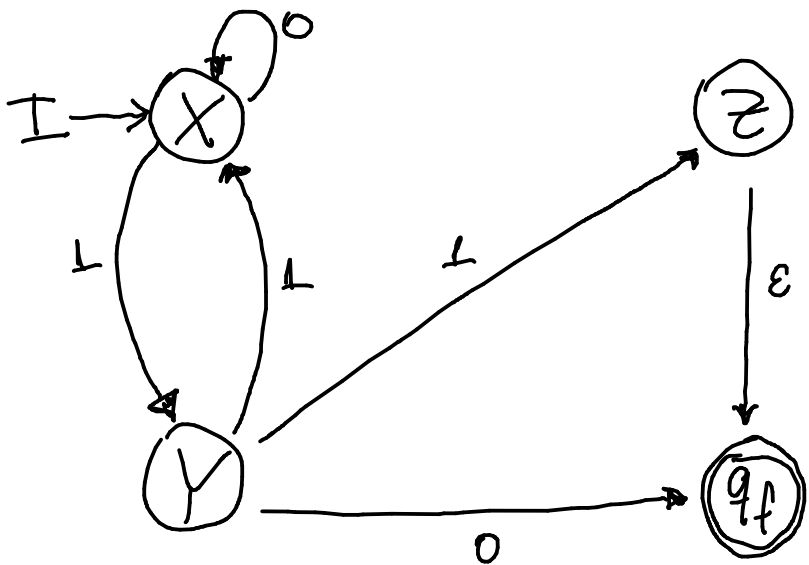
a) SE  $X \rightarrow aY \in R$ , ADICIONAMOS  $Y \in \Delta(X, a)$

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

c) SE  $X \rightarrow \epsilon \in R$ , ADICIONAMOS  $q_f \in \Delta(X, \epsilon)$

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

$S = X$



$X \Rightarrow 0X$	$(X, 011)$
$\Rightarrow 00X$	$\vdash (X, 11)$
$\Rightarrow 001Y$	$\vdash (Y, 1)$
$\Rightarrow 0011Z$	$\vdash (Z, \epsilon)$
$\Rightarrow 0011$	$\vdash (qf, \epsilon)$

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

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

•  $T = \Sigma$

•  $V = Q$

•  $S = q_0$

•  $R$  :

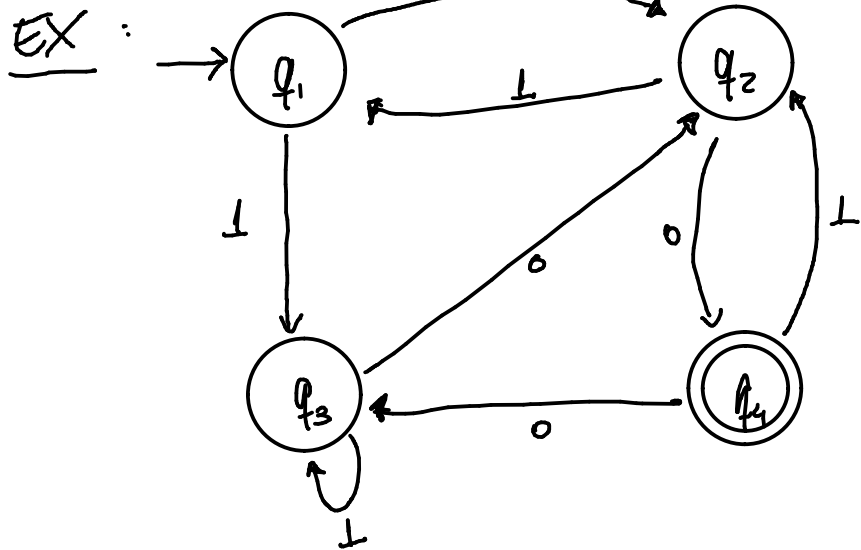
a) SE  $q' = \delta(q, a)$ , ADICIONAMOS  $q \rightarrow aq'$  A  $R$

b) SE  $q \in F$ , ADICIONAMOS  $q \rightarrow \epsilon$  A  $R$



a) se  $q' = \delta(q, a)$ , ADICIONAMOS  $q \rightarrow a q' \Delta R$

b) se  $q \in F$ , ADICIONAMOS  $q \rightarrow \epsilon \Delta R$



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

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

$$S = q_0 = q_1$$

$$R = \begin{matrix} & a & b \\ 1 & \left\{ \begin{array}{l} q_1 \rightarrow 0q_2, q_1 \rightarrow 1q_3 \\ q_2 \rightarrow 0q_4, q_2 \rightarrow 1q_1 \\ q_3 \rightarrow 0q_2, q_3 \rightarrow 1q_3 \\ q_4 \rightarrow 0q_3, q_4 \rightarrow 1q_2 \end{array} \right. & \\ & & q_4 \xrightarrow{\epsilon} \epsilon \end{matrix}$$

$$01100 : (q_1, 01100)$$

$$\vdash (q_2, 1100)$$

$$\vdash (q_1, 100)$$

$$\vdash (q_3, 00)$$

$$\vdash (q_2, 0)$$

$$\vdash (q_4, \epsilon)$$

$q_1$

$$\Rightarrow 0q_2$$

1a

$$\Rightarrow 01q_1$$

2b

$$\Rightarrow 011q_3$$

1b

$$\Rightarrow 0110q_2$$

3a

$$\Rightarrow 01100q_4$$

2a

$$\xrightarrow{\epsilon} 01100$$

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

$$X \rightarrow w$$

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

Lembre-se que

Gram. REG:

$$X \rightarrow aY$$

$$X \rightarrow a$$

$$X \rightarrow \epsilon$$

OBS: TODA GRAMÁTICA REGULAR É LIVRE DE CONTEXTO

NEM TODA GRAMÁTICA LIVRE DE CONTEXTO É REGULAR

EX:  $R = \{ S \rightarrow OS, S \rightarrow I, X \rightarrow XX \}$

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

$$L(G_1) = \{0^n 1^n : n \in \mathbb{N}\}$$

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

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

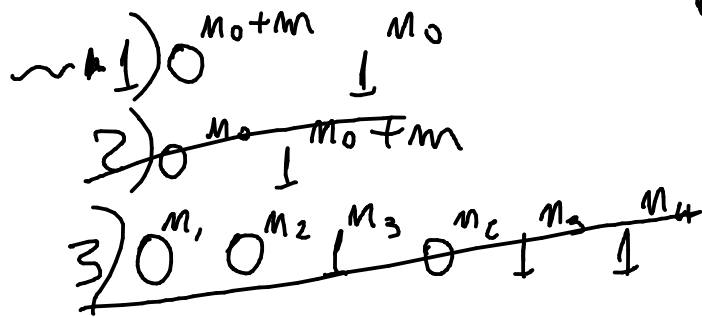
(LLC)

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

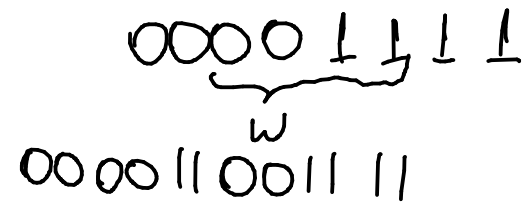
EX:  $L(G_1) = \{0^n 1^n : n \in \mathbb{N}\}$  é uma linguagem livre de contexto

que **NÃO É REGULAR**

$w = xyz$   
 $x y^k z \in L$



$\notin L(G_1)$



Teo: A classe das linguagens regulares é um subconjunto próprio

DA classe das LLCs

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

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

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

$$S \xrightarrow{(1)} X \perp X \xrightarrow{(2)} 0X \perp X \xrightarrow{(2)} 00X \perp X \xrightarrow{(3)} 00 \perp X \xrightarrow{(3)} 00 \perp$$

$$\Rightarrow X \perp X \xrightarrow{(3)} X \perp \xrightarrow{(2)} 0X \perp \xrightarrow{(2)} 00X \perp \xrightarrow{(3)} 00 \perp$$

↳ INDICA EM QUAL VARIÁVEL APLICAMOS A REGRA

# GRAMÁTICAS QUE NÃO SÃO LIVRES DE CONTEXTO

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

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

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

$$\begin{aligned} S &\stackrel{5}{\Rightarrow} a\underline{X}bc \stackrel{2}{\Rightarrow} ab\underline{X}c \stackrel{6}{\Rightarrow} ab\underline{Y}bc^2 \\ &\stackrel{3}{\Rightarrow} \underline{aY}b^2c^2 \stackrel{7}{\Rightarrow} a^2b^2c^2 \end{aligned}$$

TEMOS QUE  $a^m b^m c^m \in L(G) \quad \forall m$

CONJ.  $L(G) = \{a^m b^m c^m : m \in \mathbb{N}\}$

EX: GRAMÁTICA QUE GERA UMA FÓRMULA LEGÍTIMA,

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

REGRAS :

$E \rightarrow E + E$

$E \rightarrow E * E$

$E \rightarrow (E)$

$E \rightarrow id$

$3 \cdot x + 7$

$3 \cdot y + 7$

$7 \cdot y + 3$

→ NÃO PRECISAMOS CONHECER AS VARIÁVEIS, APENAS SUAS POSIÇÕES

→ NÃO QUEREMOS EXPRESSÕES DO TIPO  $(+id) id * E T^*$

→ O MESMO SÍMBOLO  $id$  PARA VARIÁVEIS

REGRAS :

$$E \rightarrow E + E$$

$$E \rightarrow E * E$$

$$E \rightarrow (E)$$

$$E \rightarrow id$$

AMBIGUA

$$\underline{\underline{7 + 8.52}}$$

~> DERIVAR id + id \* id

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

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



REGRAS :

$$E \rightarrow (E + E)$$

$$E \rightarrow (E * E)$$

$$E \rightarrow id$$

AMBÍGUA

~> DERIVAR ~~id + id \* id~~ ~>  $\left( (id + id) * id \right) \text{ ou } \left( id + (id * id) \right)$  7 + 8.52

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

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

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: LOL , 00 111 001 00111

OBS:

- 1)  $w$  É UM PALÍNDROMO SE COMEÇA E TERMINA COM O MESMO SÍMBOLO  $\sigma$ , 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

$R = \{ S \rightarrow 0S0, S \rightarrow 1S1, S \rightarrow \epsilon, S \rightarrow 0, S \rightarrow 1 \}$

COMPACTA :  $S \rightarrow 0S0 \mid 1S1 \mid \epsilon \mid 0 \mid 1$

OBS: COMO GERAR APENAS PALÍNDROMOS DE COMPRIMENTO PAR?

# FECHAMENTO LLC

SUPONDO  $V_1 \neq V_2$

1. UNIÃO : 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

$$T_U = T_1 \cup T_2$$

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

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

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

$$\text{LOGO } L(G_U) = L_1 \cup L_2$$

2.  
CONCATENAÇÃO

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

SUPONDO  $V_1 \neq V_2$

$$T_{\bullet} = T_1 \cup T_2$$

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

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

$$w_1 \in L_1$$

$$w_2 \in L_2$$

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

3.  
ESTRELA DE KLEENE

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

$$T_{*} = T_1, \quad V_{*} = V_1 \cup S_{*}$$

$$R_{*} = R_1 \cup \{S_{*} \rightarrow S_1 S_{*} \mid \epsilon\}$$

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

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

$$\epsilon, w, ww, w \dots w$$

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

4. INTERSECÇÃO : NÃO É VERDADE QUE SE  $L_1, L_2$  SÃO LLC,  
ENTÃO  $L_1 \cap L_2$  É LLC.

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

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

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

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

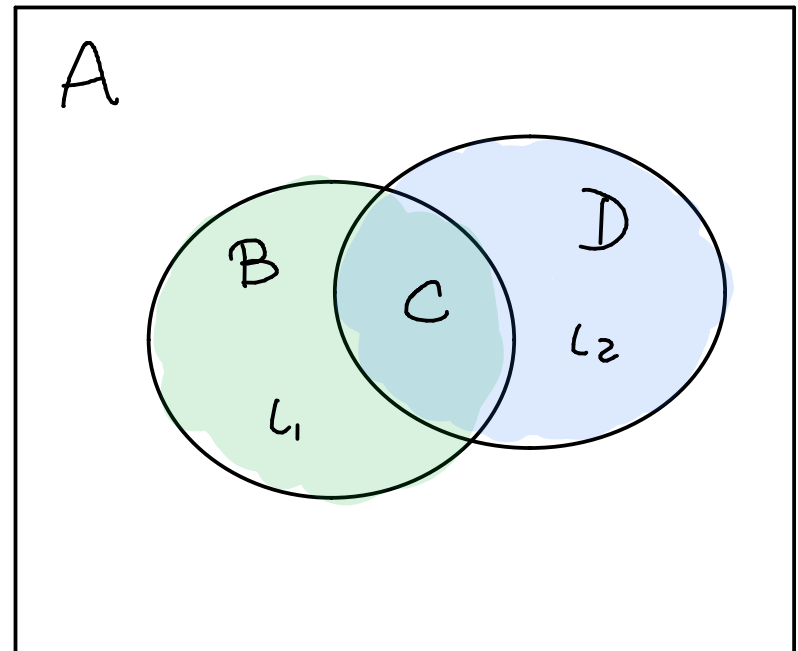
LOGO  $L_{ab} \cdot \{c\}_*$  E  $\{a\}_* \cdot L_{bc}$  SÃO LLC.

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  $L_3 \setminus L_4 \in \text{LLC}$ .

TOME  $L_3 = \sum^*$ , QUE É REGULAR e, PORTANTO, É LLC.

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



# ÁRVORES DE ANÁLISE SINTÁTICA

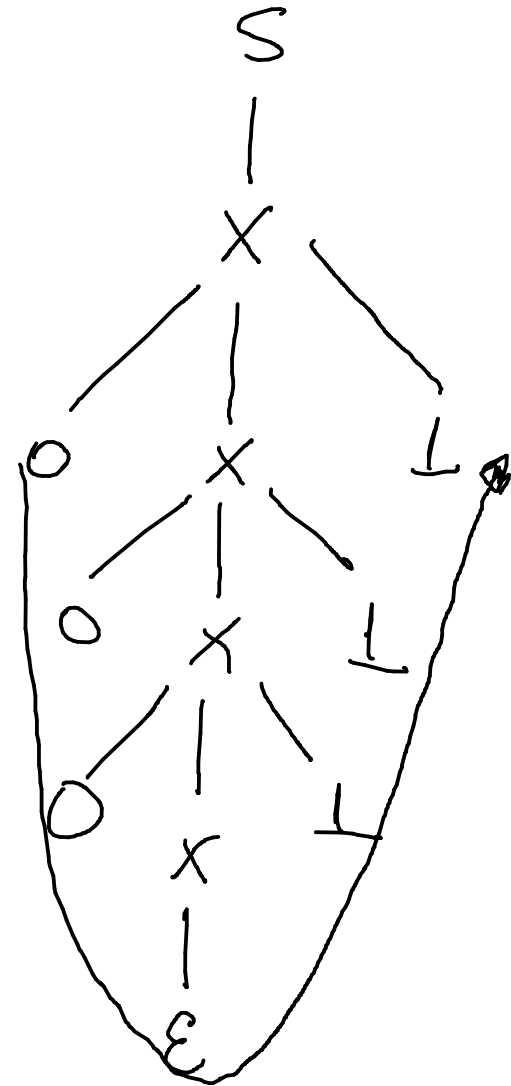
EX:  $S \rightarrow X$

$X \rightarrow OX \mid \epsilon$

$w = 000111$

$S \Rightarrow X \Rightarrow OX \Rightarrow OOX11$

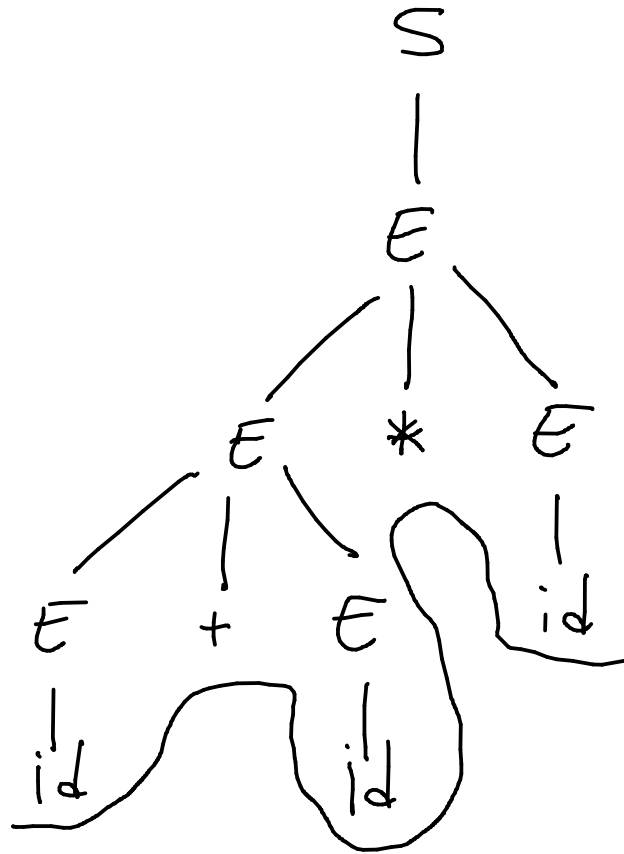
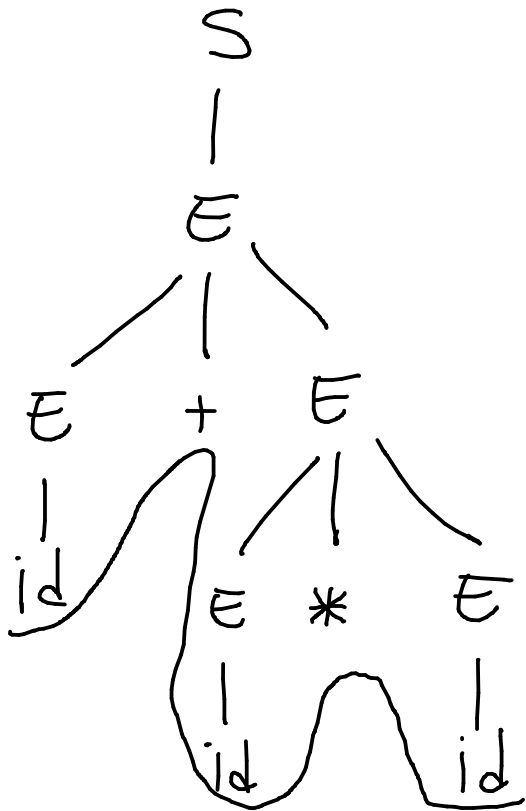
$\Rightarrow OOX111 \Rightarrow OOX111$



Ex:  $S \rightarrow E$

$E \rightarrow E + E \mid E * E \mid id$

$w = id + id * id$





## DEFINIÇÃO FORMAL

- **ÁRVORE**: GRAFOS SEM CICLOS

- Aqui: "ORIENTADOS", SEM SETAS, "DE CIMS PARA BAIXO"

1) Há ÚNICO VÉRTICE CHAMADO DE RAÍZ

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 RAÍZ.

SEJA  $v$  UM VÉRTICE NESTE CAMINHO

- $v$  É ASCENDENTE DE  $v'$ , E  $v'$  É DESCENDENTE 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.

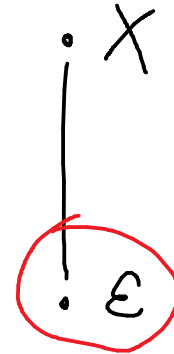
# DEFINIMOS ÁRVORES DE ANÁLISE SINTÁTICA RECURSIVAMENTE

FIXE GRAMÁTICA  $G = (T, V, S, R)$

BÁSICAS:  $\sigma \in T$ ,



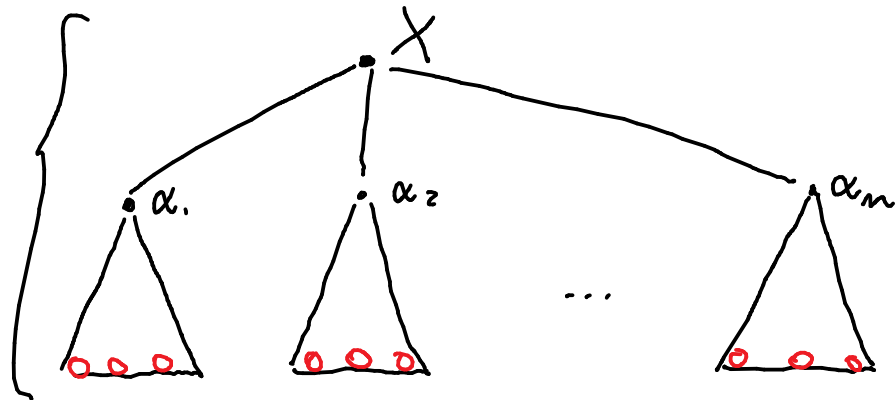
$X \in V \in X \rightarrow \epsilon \in R$



REGRAS DE COMBINAÇÃO:  $T_1, \dots, T_m$  ÁRVORES COM RAÍZES  $\alpha_1, \dots, \alpha_m$ , RESP.

SE  $X \rightarrow \alpha_1 \alpha_2 \dots \alpha_m \in R$ , ENTÃO

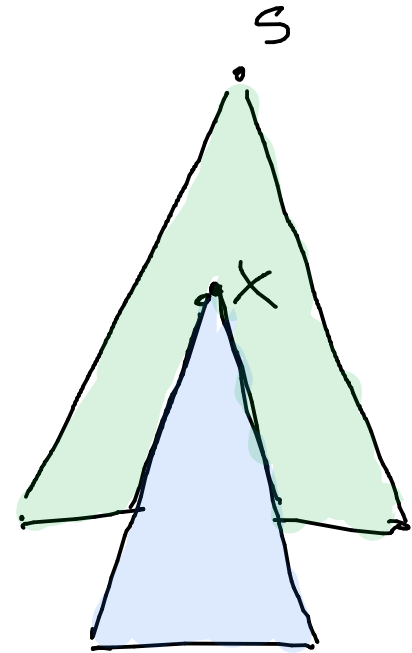
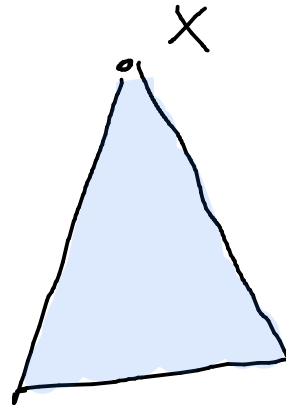
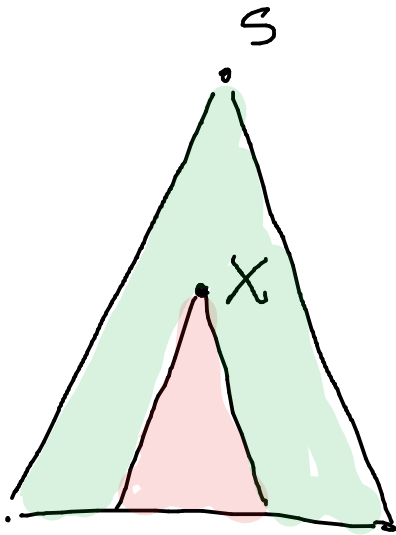
É UMA  
 $X$ -ÁRVORE



É ÁRVORE, COM **REGRAS ASSOCIADAS**

$X \rightarrow \alpha_1 \alpha_2 \dots \alpha_m$

- EM PARTICULAR, PELAS REGRAS BÁSICAS, AS FOLHAS DE UMA ÁRVORE DE ANÁLISE SINTÁTICA SÃO APENAS ELEMENTOS EM  $T \cup \{\epsilon\}$
- E OS VÉRTICES INTERNOS SÃO VARIÁVEIS.



- OBS: SE UMA ÁRVORE  $\mathcal{T}$  POSSUI UM VÉRTICE ROTULADO POR  $X$ , PODEMOS SUBSTITUIR TODA SUBÁRVORE DE  $\mathcal{T}$  ENRAIZADA EM  $X$ , POR OUTRA ÁRVORE ENRAIZADA EM  $X$ .

# COLHEITA

COLHEITA OU RESULTADO DE UMA ÁRVORE É UMA FUNÇÃO  $C$  QUE LEVA ÁRVORES EM PALAVRAS

## BÁSICAS

$$C \left( \begin{array}{c} \bullet \\ \sigma \end{array} \right) = \sigma$$

$$C \left( \begin{array}{c} \bullet X \\ | \\ \bullet E \end{array} \right) = E$$

## COMBINAÇÃO

$$C \left( \begin{array}{c} \text{+} \\ || \\ X \\ / \quad | \quad \backslash \\ \alpha_1 \quad \alpha_2 \quad \alpha_m \\ \triangle T_1 \quad \triangle T_2 \quad \dots \quad \triangle T_m \\ \circ \quad \circ \quad \circ \quad \dots \quad \circ \quad \circ \quad \circ \end{array} \right) = C \left( \begin{array}{c} \bullet \alpha_1 \\ \triangle T_1 \\ \circ \quad \circ \quad \circ \end{array} \right) C \left( \begin{array}{c} \bullet \alpha_2 \\ \triangle T_2 \\ \circ \quad \circ \quad \circ \end{array} \right) \dots C \left( \begin{array}{c} \bullet \alpha_m \\ \triangle T_m \\ \circ \quad \circ \quad \circ \end{array} \right)$$

$$C(T) = C(T_1) C(T_2) \dots C(T_m)$$

- COMO AS FOLHAS ESTÃO ORDENADAS, NÃO HÁ **AMBIGUIDADE** NA COLHEITA.



DEF: UMA ÁRVORE COM RAIZ  $X$  É UMA  $X$ -ÁRVORE

DEF: SE  $w \in L(G)$ , ENTÃO UMA ÁRVORE DE DERIVAÇÃO  $P/w$  É UMA  $S$ -ÁRVORE  $T$  CUJA COLHEITA É  $w$ , I.E.,  $C(T) = w$

DEF: SEJA  $w \in (TUV)^*$ ,  $w = \alpha_1 \alpha_2 \dots \alpha_m$ .

UMA  $w$ -FLORESTA É UMA SEQUÊNCIA ORDENADA  $T_1, \dots, T_m$ ,  
DE ÁRVORES, EM QUE  $T_i$  É UMA  $\alpha_i$ -ÁRVORE

• COMO OBTER UMA DERIVAÇÃO A PARTIR DE UMA ÁRVORE?

ENTRADA : UMA  $X$ -ÁRVORE  $\mathcal{T}$ , COM  $X \in \mathcal{T}$

SAÍDA : UMA DERIVAÇÃO MAIS À ESQUERDA DE  $C(\mathcal{T})$

1) INICIALIZAÇÃO  $\mathcal{F} = \mathcal{T}$

2) SE  $\mathcal{F}$  É UMA  $w$ -FLORESTA, IMPRIMA  $w$ .

3) SE  $w \in T^*$ , ENTÃO PARE

4) CASO CONTRÁRIO, SEJA  $v$  A RAIZ DA PRIMEIRA ÁRVORE DE  $\mathcal{F}$  QUE É ROTULADA POR UMA VARIÁVEL

FAÇA  $\mathcal{F} = \mathcal{F} \setminus v$

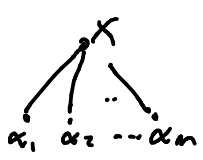


PROPOSIÇÃO: SEJA  $X$  UMA VARIÁVEL DE UMA GLC.

SE EXISTE DERIVAÇÃO  $X \Rightarrow \Delta^* w$ , ENTÃO  $w$  É  
A COLHEITA DE UMA  $X$ -ÁRVORE.

PROVA: INDUÇÃO NO NÚMERO  $P$  DE PASSOS DA DERIVAÇÃO.

BASE: UM ÚNICO PASSO.  $X \Rightarrow \Delta w$ . ENTÃO  $X \rightarrow w \in R$ .

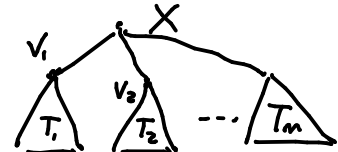
LOGO, SE  $w = \alpha_1 \dots \alpha_m$ , ENTÃO  É UMA  $X$ -ÁRVORE  
CUJA COLHEITA É  $w$ ,

PASSO INDUTIVO: SUPONHA QUE TODA  $P$ -DERIVAÇÃO  $Y \rightarrow w'$   
PODE SER "TRANSFORMADA" EM UMA  $Y$ -ÁRVORE CUJA COLHEITA É  $w'$

SEJA  $X \Rightarrow \Delta v_1 \dots v_m$  O PRIMEIRO PASSO DE  $X \Rightarrow \Delta^* w$

ISSO QUER DIZER QUE  $w = w_1 \dots w_m$  E QUE EXISTEM  
DERIVAÇÕES  $v_i \Rightarrow \Delta^* w_i$  PARA TODO  $i$ .

MAS  $v_i \Rightarrow \Delta^* w_i$  É UMA  $r$ -DERIVAÇÃO COM  $r \leq P$ . PELA HI,  
EXISTE  $v_i$ -ÁRVORE  $T_i$  CUJA COLHEITA É  $w_i$ ,

LOGO,  É UMA  $X$ -ÁRVORE CUJA COLHEITA É  $w$ .  $\square$

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

- 1) EXISTE ÁRVORE DE DERIVAÇÃO CUJA COLHEITA É  $w$
- 2) CADA ÁRVORE DE DERIVAÇÃO CUJA COLHEITA É  $w$  CORRESPONDE A UMA ÚNICA DERIVAÇÃO MAIS À ESQUERDA DE  $w$ .

# AMBIGUIDADE

• ÁRVORES NOS AJUDAM A INTERPRETAR UMA FRASE CORRETAMENTE.

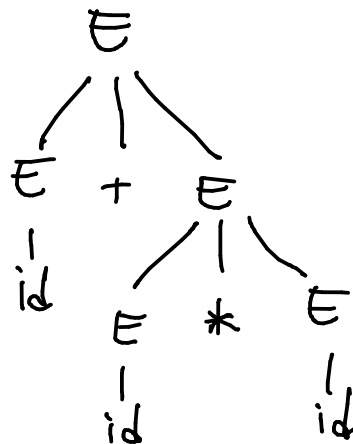
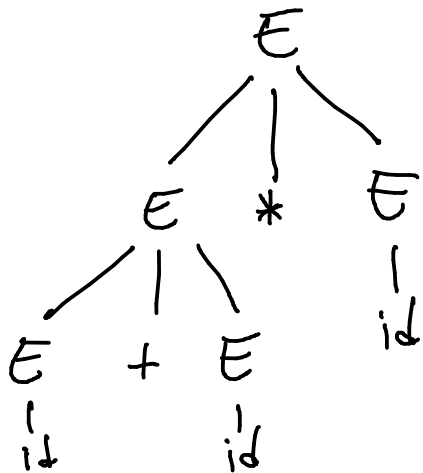
• ALGUMAS FRASES POSSUEM INTERPRETAÇÕES DISTINTAS

EX: A SEGUIR VEIO UMA MÃE COM UMA CRIANÇA EMPURRANDO UM CARRINHO.

QUEM ESTÁ EMPURRANDO O CARRINHO?

DEF: A GRAMÁTICA  $G$  É AMBIGUA SE EXISTE UMA PALAVRA  $w \in L(G)$  QUE ADMITE DUAS ÁRVORES DE DERIVAÇÃO DISTINTAS.

EX:



EX:  $T = \{a, +, -\}$ ,  $V = \{S, A\}$ , -

$$R = \begin{cases} S \rightarrow A \\ A \rightarrow A+A \mid A-A \mid a \end{cases}$$

EX: SEJAM  $L_1$  E  $L_2$  DUAS LINGUAGENS COM GRAMÁTICAS

$$G_1 = (T_1, V_1, S_1, R_1) \quad \text{E} \quad G_2 = (T_2, V_2, S_2, R_2)$$

SEJA  $U = L_1 \cup L_2$  A LINGUAGEM COM GRAMÁTICA

$$G_U = (T_1 \cup T_2, V_1 \cup V_2, S_U, R_1 \cup R_2 \cup \{S_U \rightarrow S_1 \mid S_2\})$$

SE  $L_1 \cap L_2 \neq \emptyset$ ,  $G_U$  É AMBÍGUA.

SE  $w \in L_1 \cap L_2$ , COMO  $S_1 \Rightarrow^* w$  E  $S_2 \Rightarrow^* w$ , ENTÃO

$S_U \Rightarrow S_1 \Rightarrow^* w$  E  $S_U \Rightarrow S_2 \Rightarrow^* w$  SÃO DERIVAÇÕES

QUE PRODUZEM ÁRVORES DE ANÁLISE SINTÁTICA DIFERENTES.

EX: A GRAMÁTICA  $G_{EXP}''$  COM REGRAS

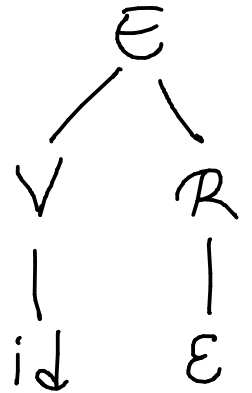
$$\begin{cases} S \rightarrow E \\ E \rightarrow (E)R \mid \underline{VR} \\ R \rightarrow +E \mid *E \mid \varepsilon \\ V \rightarrow \underline{id} \end{cases}$$

- PODEMOS DEDUZIR A DERIVAÇÃO A PARTIR DA PALAVRA DADA.

PROVA: INDUÇÃO NO COMPRIMENTO DA PALAVRA  $w$ .

SE  $|w| = 1$ , ENTÃO  $w = id$

$$\begin{aligned} S &\Rightarrow E \Rightarrow VR \Rightarrow idR \Rightarrow id \\ &\Rightarrow VE \Rightarrow id \end{aligned}$$



SE  $|w| = n$  E  $w = idw'$ , ENTÃO  $|w'| < n$ , E HA' UMA ÚNICA DERIVAÇÃO  $R \Rightarrow w'$

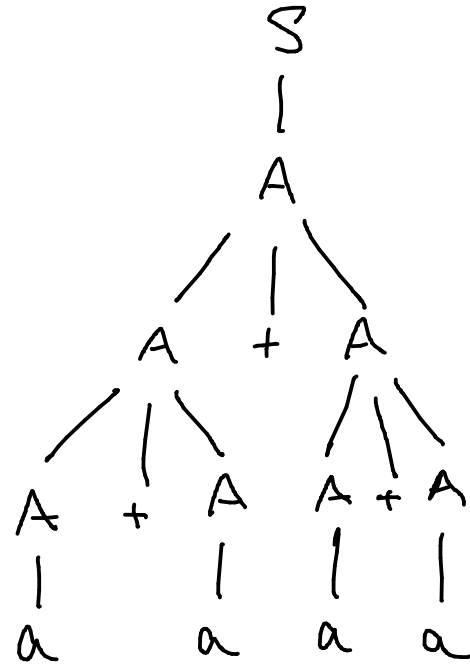
SE  $|w| = n$ , E  $w = (w_1)w_2$  E  $|w_1|, |w_2| < n$   
E TEMOS  $E \Rightarrow^* w_1$  E  $R \Rightarrow^* w_2$



$$\text{EX: } L = \{ a (+a)^n : n \geq 0 \} = \{ a, a+a, a+a+a, \dots \}$$

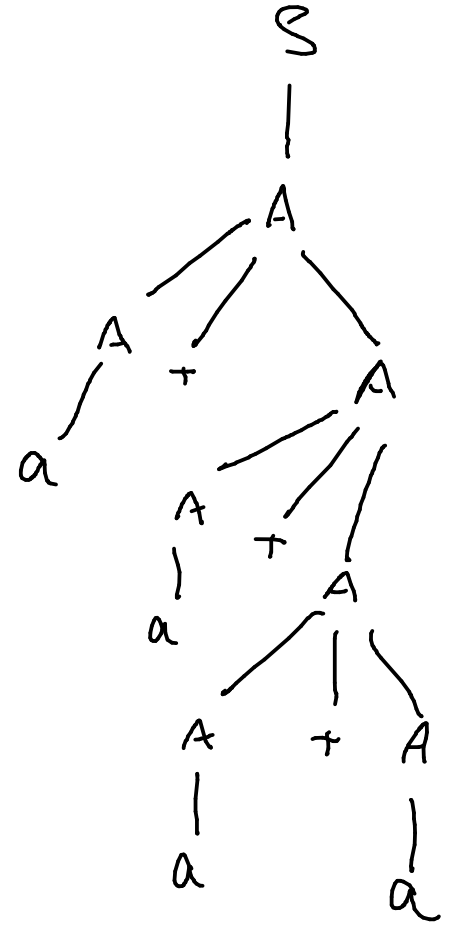
$$R_1 = \begin{cases} S \rightarrow A \\ A \rightarrow A+A \mid a \end{cases}$$

É AMBIGUA



$$R_2 = \begin{cases} S \rightarrow A \\ A \rightarrow A+a \mid a \end{cases}$$

NÃO É AMBIGUA.





Obs: O problema de decidir se uma gramática é ambígua é indecidível.

- Uma linguagem  $L$  é inerentemente ambígua se toda gramática  $G$  t.q.  $L = L(G)$  é ambígua

$$\text{EX. } \{a^n b^m c^m d^n : n, m > 0\} \cup \{a^n b^n c^m d^m : n, m > 0\}$$

- SEMPRE É POSSÍVEL OBTER UMA GRAMÁTICA NÃO AMBÍGUA PARA UMA LINGUAGEM REGULAR.

Alg: DADA UMA GRAM. REG.

DEVOLVE GRAMÁTICA  $G'$  NÃO AMBÍGUA T.q.  $L(G') = L(G)$

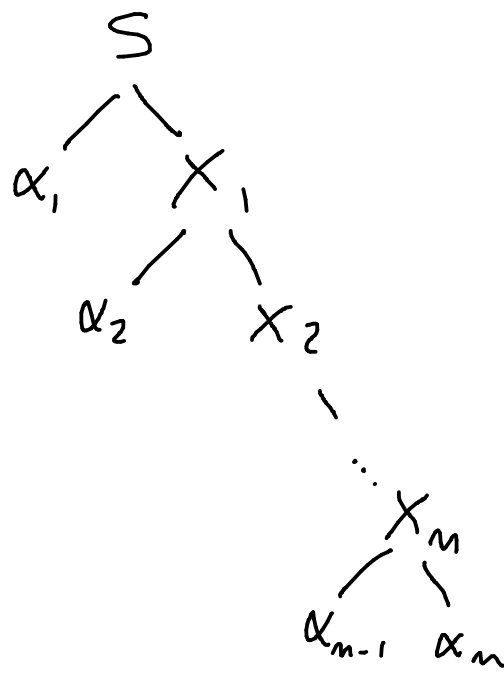
1) DETERMINAR AFND  $M$  PARA  $L(G)$

2) DETERMINAR APD  $M'$  PARA  $L(G)$

3) DETERMINAR GRAMÁTICA  $G'$  OBTIDA DE  $M'$ .

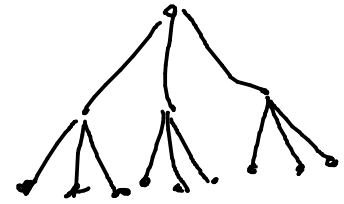


$wX$



# BOMBAMENTO LLC

• Como provar que uma linguagem NÃO É LLC?



ÁRVORE 3-ÁRIA  
COMPLETA COM  
ALTURA 2

EX:  $L_{abc} = \{a^m b^m c^m : m \geq 0\}$  NÃO É LLC

→ VAMOS provar que toda LLC possui uma propriedade testável!

DEF: DIZEMOS QUE UMA ÁRVORE É **m-ÁRIA** SE CADA VÉRTICE  
POSSUI NO MÁXIMO **m** FILHOS.

→ A **ALTURA** DE UMA ÁRVORE É O COMPRIMENTO DO CAMINHO  
MAIS LONGO DE SUA RAÍZ ATÉ UMA FOLHA

COM ALTURA **h**

→ TODO NÓ  
POSSUI **m**  
FILHOS

DEF:  $f(h)$  = NÚMERO DE FOLHAS DE UMA ÁRVORE **m-ÁRIA COMPLETA**

$$f(0) = 1, f(1) = m \quad \begin{array}{c} \diagup \\ \diagdown \end{array}, f(2) = m^2 \quad \begin{array}{c} \diagup \diagdown \\ \diagup \diagdown \end{array}, \dots, f(h) = m^h$$

DEF: A **AMPLITUDE**  $\alpha(G)$  DE UMA GLC É O COMPRIMENTO MÁXIMO DAS PALAVRAS QUE APARECEM À DIREITA DE UMA REGRA

EX:  $X \rightarrow 0X1 \mid 0 \mid 1 \mid \varepsilon$  TEM AMPLITUDE 3.

OBS: SE  $G$  POSSUI AMPLITUDE  $\alpha$ , ENTÃO TODAS AS SUAS ÁRVORES SÃO  $\alpha$ -ÁRIAS

LEMA 1: SEJA  $G$  UMA GLC. SE  $X$  É UMA VARIÁVEL DE  $G$ , E  $w$  É A COLHEITA DE UMA  $X$ -ÁRVORE DE ALTURA  $h$ , ENTÃO

$$|w| \leq \alpha(G)^h$$

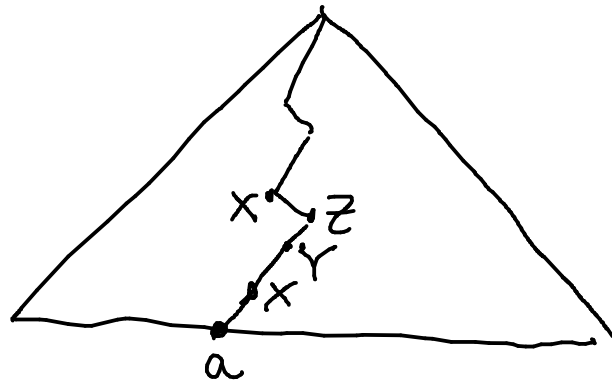
↳ ISSO IMPLICA QUE SE A COLHEITA DE UMA ÁRVORE É MUITO GRANDE, ENTÃO A ALTURA DA ÁRVORE TAMBÉM É.

LEMA: SEJA  $G$  UMA GLC. SE  $X$  É UMA VARIÁVEL DE  $G$ , E  $w$  É A COLHEITA DE UMA  $X$ -ÁRVORE DE ALTURA  $h$ , ENTÃO

$$|w| \leq \alpha(G)^h$$

↳ ISSO IMPLICA QUE SE A COLHEITA DE UMA ÁRVORE É MUITO GRANDE, ENTÃO A ALTURA DA ÁRVORE TAMBÉM É.

↳ SE A ALTURA DA ÁRVORE FOR MAIOR QUE O NÚMERO DE VARIÁVEIS DA GRAMÁTICA, UM CAMINHO MÁXIMO (NA ÁRVORE) CONTÉM DOIS VÉRTICES ROTULADOS COM A MESMA VARIÁVEL.



OU SEJA, AS REGRAS ASSOCIADAS A ESSES VÉRTICES TÊM A MESMA VARIÁVEL À ESQUERDA.

LEMA DO BOMBAMENTO: SEJA  $G$  UMA GRAMÁTICA LIVRE DE CONTEXTO.

EXISTE INTEIRO  $p$  (QUE DEPENDE DE  $G$ ) TAL QUE SE  $w \in L(G)$  E  $|w| \geq p$ , ENTÃO EXISTE DECOMPOSIÇÃO DE  $w$  COMO  $w = uvxyz$ ,

PARA A QUAL

1)  $vy \neq \epsilon$

2)  $|vxy| \leq p$

3)  $uv^mxy^mz \in L(G)$ , PARA TODO  $m \in \mathbb{N}$ .

$$\alpha(G)^k \geq |w| \geq p = \alpha(G)^{k+1}$$

PROVA: TOME  $p = \alpha(G)^{k+1}$ , EM QUE  $k$  É O NÚMERO DE VARS DE  $G$ .

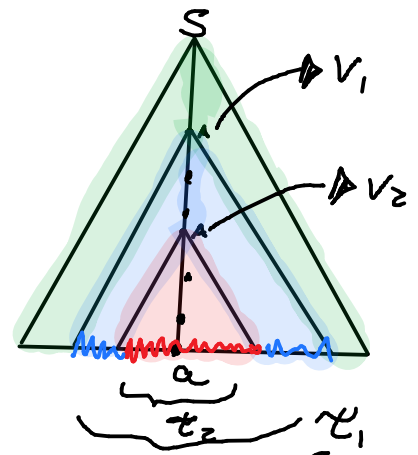
SEJA  $w \in L(G)$  COM  $|w| \geq p$ , E SEJA  $T$  UMA ÁRVORE DE DERIVAÇÃO  $P/w$ . PELO LEMA 1, A ALTURA DE  $T$  É PELO MENOS  $k+1$ . HÁ UM CAMINHO DA RAÍZ DE  $T$  ATÉ UMA FOLHA QUE REPETE VARIÁVEL

1)  $v \neq \epsilon$

2)  $|vxy| \leq p$

3)  $uv^mxy^mz \in L(G)$ , PARA TODO  $m \in \mathbb{N}$ .

$X_2$  ~~~~~  
 $X_1$  ~~~~~



PROVA: Tome  $p = \alpha(G)^{k+1}$ , em que  $k$  é o número de vars de  $G$ .

Seja  $w \in L(G)$  com  $|w| \geq p$ , e seja  $T$  uma árvore

\* com o menor número de vértices possível

de derivação  $p/w$ . \* Pelo lema 1, a altura de  $T$  é pelo menos  $k+1$ . Há um caminho da raiz de  $T$  até uma folha que repete variável, digamos  $A$ .

Sejam  $v_1$  e  $v_2$  os dois últimos vértices que repetem variável

Seja  $\psi_1$  a árvore com raiz  $v_1$ , com colheita  $X_1$

$\Rightarrow X_2$  é uma subpalavra de  $X_1$

$\Rightarrow X_1 = v X_2 y$

$\Rightarrow$  Como  $\psi_2$  não repete variável, a altura de  $\psi_2$  é no ma

SEJAM  $v_1$  E  $v_2$  OS DOIS ÚLTIMOS VÉRTICES QUE REPETEM VARIÁVEL.

SEJA  $\mathcal{T}_i$  A ÁRVORE COM RAÍZ  $v_i$ , COM COLHEITA  $X_i$ ;

$\Rightarrow X_2$  É UMA SUBPALAVRA DE  $X_1$ ,

$\Rightarrow X_1 = v X_2 y$  (EXCETO A)

$\Rightarrow$  COMO  $\mathcal{T}_1$  NÃO REPETE VARIÁVEL, A ALTURA DE  $\mathcal{T}_1$  É NO MÁXIMO  $k+1$

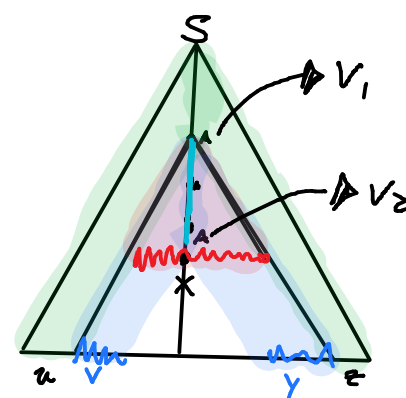
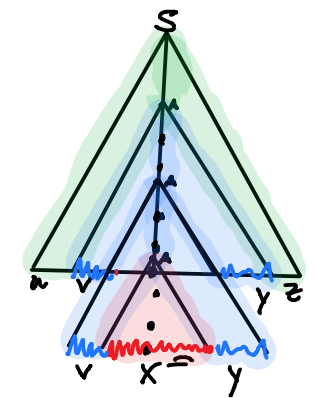
ISSO QUER DIZER QUE A COLHEITA DE  $\mathcal{T}_1$  ( $X_1$ ) TEM

COMP. NO MÁXIMO  $\alpha(G)^{k+1}$ , OU SEJA  $|X_1| = |v X_2 y| \leq p$

$\Downarrow$   
ISSO PROVA O ITEM 2.

$\Rightarrow$  SE  $|v y| = 0$ , ENTÃO  $v y = \epsilon$ , CONTRADIÇÃO COM A MINIMALIDADE DE T.

$\hookrightarrow$  ISSO PROVA O ITEM 1.



$\Rightarrow$  SUBSTITUINDO  $\mathcal{T}_1$  NO LUGAR DE  $\mathcal{T}_2$  REPETIDAMENTE OBTENEMOS UMA ÁRVORE DE DERIVAÇÃO  $\mathcal{T}' / u v^n X y^n z v_m$

$\hookrightarrow$  ISSO PROVA O ITEM 3





• Como aplicar

- (\*)
- 1)  $vy \neq \epsilon$
  - 2)  $|vxy| \leq p$
  - 3)  $zuv^mxy^mz \in L(G)$ , PARA TODO  $m \in \mathbb{N}$ .

1) SUPÕE QUE  $L \in \bar{L}C$ .

2) TOMA QUALQUER GLC  $G \neq L$

3) USA O LEMA DE BOMBAMENTO EM  $G$ . OBTENOS  $P$

4) ESCOLHE PALAVRA  $w \in L$  COM  $|w| \geq p$

3) MOSTRAR QUE NÃO É POSSÍVEL BOMBEAR.

EX:  $L_{abc} = \{a^n b^n c^n : n \geq 0\}$

SUPÕNH QUE  $L_{abc} \in \bar{L}C$ , E SEJA  $G_{abc}$  UMA GLC  $\neq L_{abc}$ .

PELO L.B. EXISTE  $P$  COM AS PROP. (\*).

TOME  $w = a^P b^P c^P$ , CLARAMENTE  $|w| = 3P > P$

PELO L.B.  $w = uvxyz$  COM (\*)

EX.  $L_{abc} = \{a^m b^n c^m : m \geq 0\}$

SUPONHA QUE  $L_{abc} \in \text{LLC}$ , E SEJA  $G_{abc}$  UMA GLC P/  $L_{abc}$ .

PELO L.B. EXISTE  $P$  COM AS PROP. (\*).

TOME  $w = a^P b^P c^P$ . CLARAMENTE  $|w| = 3P > P$

PELO L.B.  $w = uvxyz$  COM (\*)

$$w = \underbrace{a \dots a}_{vxy} \underbrace{b \dots b}_{vxy} \underbrace{c \dots c}_{vxy}$$

ANÁLISE DE CASO  $i+k > 0$  POR (1)

1) SE  $vxy = a^i a^j a^k$ , ENTÃO POR (3)  $w' = uv^2 xy^2 z \in L$ ,  
 MAS  $w' = a^{P+i+k} b^P c^P \notin L$ , UMA CONTRADIÇÃO.

ANALOGAMENTE, VERIFICAMOS O CASO  $vxy \subseteq b^P$  OU  $vxy \subseteq c^P$

2) SE  $vxy = a^i b^j$ , TOME  $w' = uv^2 xy^2 z$

TEMOS  $w' = a^{P'} b^{P''} c^P$  COM  $P' > P$  OU  $P'' > P$

$$\frac{a \dots a}{v} \frac{b \dots b}{x} \frac{b \dots b}{y}$$

ANALOGAMENTE, VERIFICAMOS O CASO  $w' = b^i c^j$

(1)  $vy \neq \epsilon$

(\*) (2)  $|vxy| \leq P$

(3)  $uv^m x y^m \in L(G)$ , PARA TODO  $m \in \mathbb{N}$ .

LOGO,  $L_{abc}$  NÃO É LLC.

EX:  $L_{\text{PRIMOS}} = \{0^q : q \in \text{PRIMOS}\}$

Pelo L.B. EXISTE  $p$  t.q. (\*)

SEJA  $m$  UM PRIMO MAIOR QUE  $p$

$$w = 0^m \in L_{\text{PRIMOS}}$$

Pelo L.B.  $w = uvxyz$  COM  $|uxy| \leq p$ ,  $xy \neq \epsilon$

SEJAM  $i, j$  t.q.  $v = 0^i$  e  $y = 0^j$

Pelo L.B.  $w_m = u v^m x y^m z \in L_{\text{PRIMOS}} \forall m$

TOME  $n = m + 1$

$$w_m = u \underset{0^i}{v^{m+1}} \underset{0^j}{x y^m} z = 0^{m + \underline{m \cdot i} + \underline{m \cdot j}} = 0^{\overbrace{m(1+i+j)}^{x^0 \neq 0}} \quad \left\{ \begin{array}{l} \text{NÃO É PRIMO} \end{array} \right.$$

COMO  $m(1+i+j)$  NÃO É PRIMO,  $w_m \notin L_{\text{PRIMOS}}$ , UMA CONTRADIÇÃO  
Logo,  $L_{\text{PRIMOS}}$  NÃO É LLC. □

# AUTÔMATOS DE PILHA

• Um AUTÔMATO DE PILHA NÃO-DETERMINÍSTICO É UM AUTÔMATO COM UMA PROPRIEDADE EXTRA: MEMÓRIA INFINITA.

• MEMÓRIA: PILHA  $\rightsquigarrow$  O ÚLTIMO ELEMENTO A SER COLOCADO É O PRÓXIMO A SER LIDO.

LISTA :  $(a_1) a_2 \dots a_n$

$\text{POP}(\text{LISTA}) = a_1$

PILHA :  $a_1 a_2 \dots (a_n)$

$\text{POP}(\text{PILHA}) = a_n$

$\rightsquigarrow$  DISCOS PERFORADOS EM UMA HASTE.

$\rightsquigarrow$  LEMBRAR É COLOCAR UM DISCO NA HASTE E UM DISCO SÓ PODE SER REMOVIDO SE ESTIVER NO TOPO DA PILHA.

EX:  $L_1 = \{vcv^R : v \in \{a,b\}^*\}$

$abaa\ caaba \in L_1$

$abcab \notin L_1$

ETAPA 1: SE LÊ  $a$ , EMPILHA  $a$   
SE LÊ  $b$ , EMPILHA  $b$

ETAPA 2: SE LÊ  $c$ , MUDA DE ATITUDE

ETAPA 3: COMPARA O QUE ESTÁ NO TOPO  
DA PILHA COM O SÍMBOLO LIDO.

SE SÃO IGUAIS, OK.

SE SÃO DIFERENTES, PARA A COMPUTAÇÃO

EX:  $L_2 = \{vv^R : v \in \{a,b\}^*\}$

NO NÃO HÁ SÍMBOLO PARA INDICAR A MUDANÇA DE ATITUDE.

NO UM AUTÔMATO  $\forall L_2$  DEVE ADIVINHAR A HORA DE MUDAR DE ATITUDE.

