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 $u \longrightarrow v$ 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.

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

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

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

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

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 $\Delta(X, a)$ ($a \in T$)

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

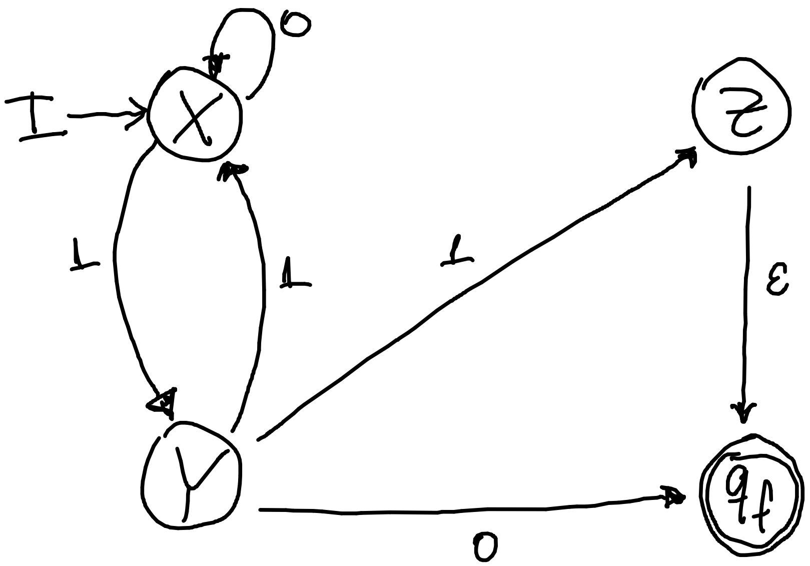
a) SE $X \to aY \in R$, ADICIONAMOS Y à $\Delta(X, a)$

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 :

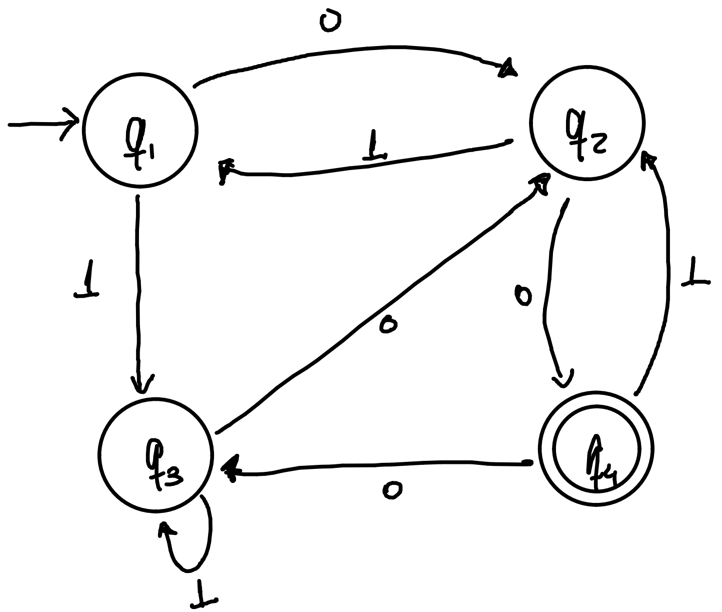
a) SE $q' = \delta(q, a)$, ADICIONAMOS $q \to a q'$ À R

b) SE $q \in F$, ADICIONAMOS $q \to \varepsilon$ À 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:



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

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

$S = q_0 = q_1$

		a	b
$R =$	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	$q_4 \to 0 q_3$,	$q_4 \to 1 q_2$
			$q_4 \xrightarrow{c} \varepsilon$

01100 :

$(q_1, 01100)$	q_1	
$\vdash (q_2, 1100)$	$\Rightarrow 0 q_2$	1a
$\vdash (q_1, 100)$	$\Rightarrow 0 1 q_1$	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

$\overset{c}{\Rightarrow} 01100$

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

$$X \longrightarrow w$$

LEMBRE-SE QUE GRAM. REG:

$X \to aY$

$X \to a$

$X \to \varepsilon$

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

OBS: TODA GRAMÁTICA REGULAR É LIVRE DE CONTEXTO

NEM TODA GRAMÁTICA LIVRE DE CONTEXTO É REGULAR

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

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

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

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

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

(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

$0^{n_0} 1^{n_0}$ $\rightsquigarrow$ 1) $0^{n_0+m} 1^{n_0}$

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

$xy^kz \in L$ 3) $\cancel{0^{n_1} 0^{n_2} 1^{n_3} 0^{n_c} 1^{n_3} 1^{n_4}}$

$\notin L(G_1)$

$\underbrace{0000\,1111}_{w}$

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

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

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

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

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 $\underline{id + id * id}$

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

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

REGRAS : $E \rightarrow (E + E)$

$E \rightarrow (E * E)$

$E \rightarrow id$

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. 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

$$T_U = T_1 \cup T_2$$

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

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

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

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

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

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

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

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

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

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.

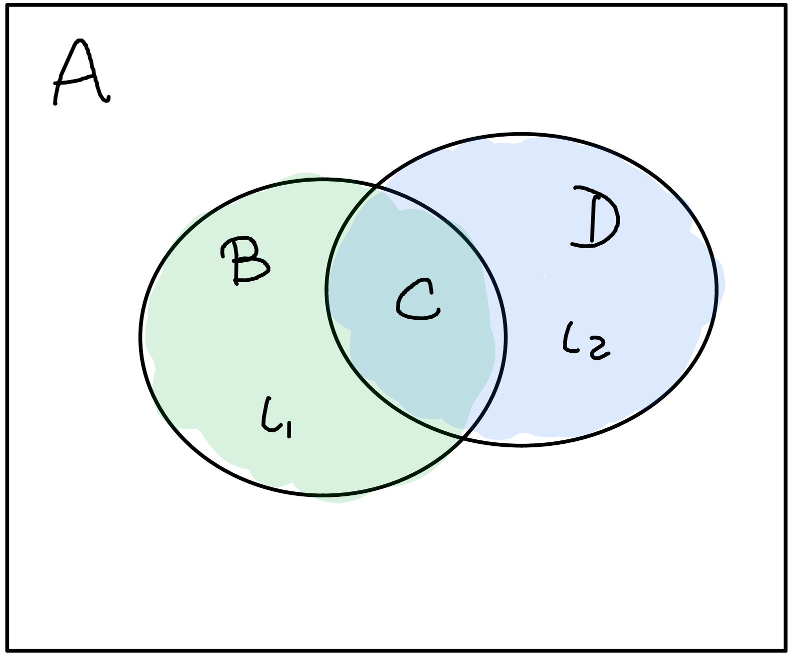
LOGO $L_{ab} \cdot \{c\}^*$ E $\{a\}^* \cdot L_{bc}$ SÃO LLC.

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

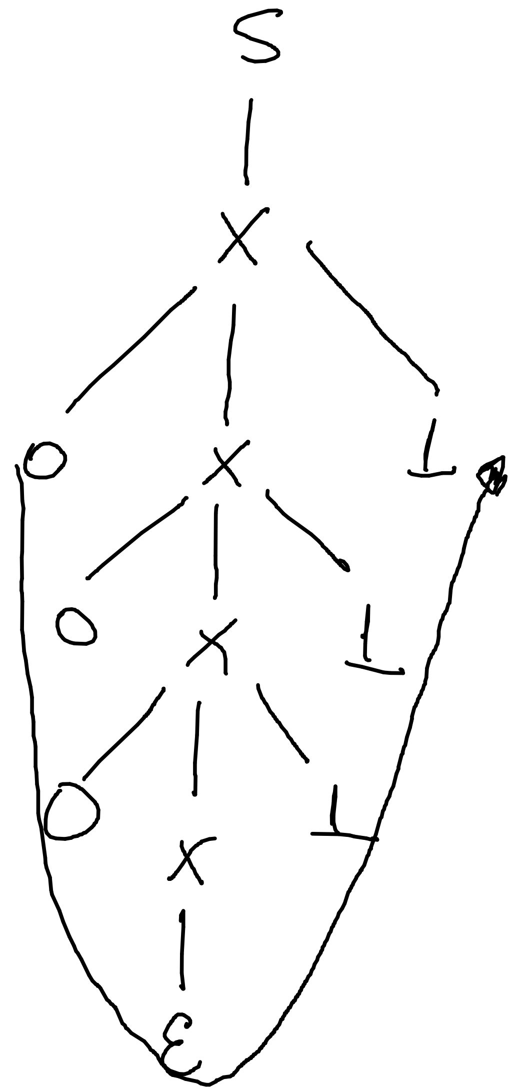
Ex: $S \to X$

$X \to 0X1 \mid \varepsilon$

$\omega = 000111$

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

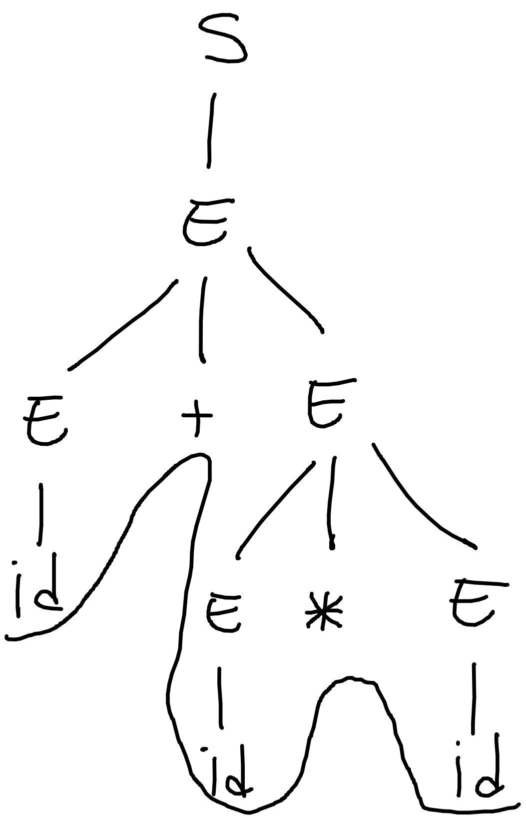
$\Rightarrow 000X111 \Rightarrow 000111$

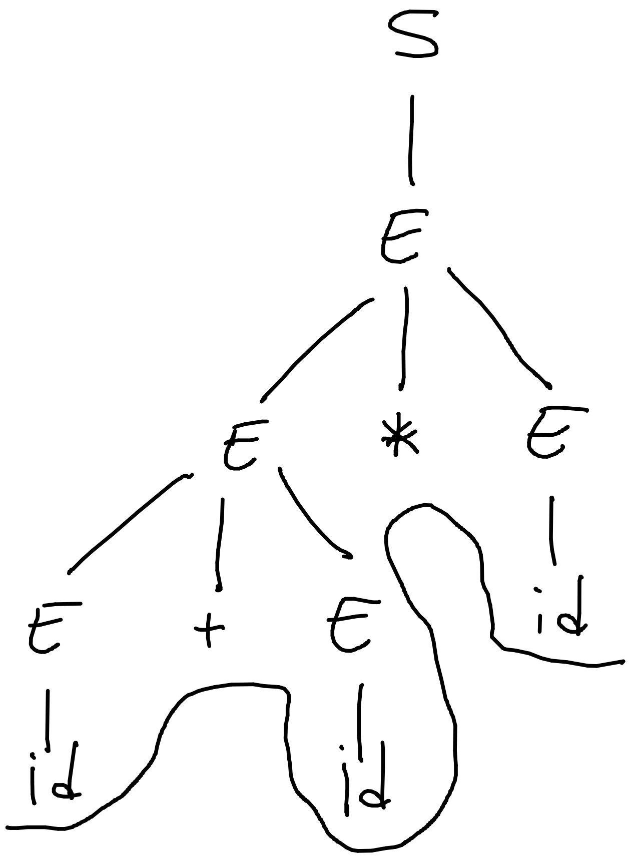


Ex: $S \to E$

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

$w = id + id*id$



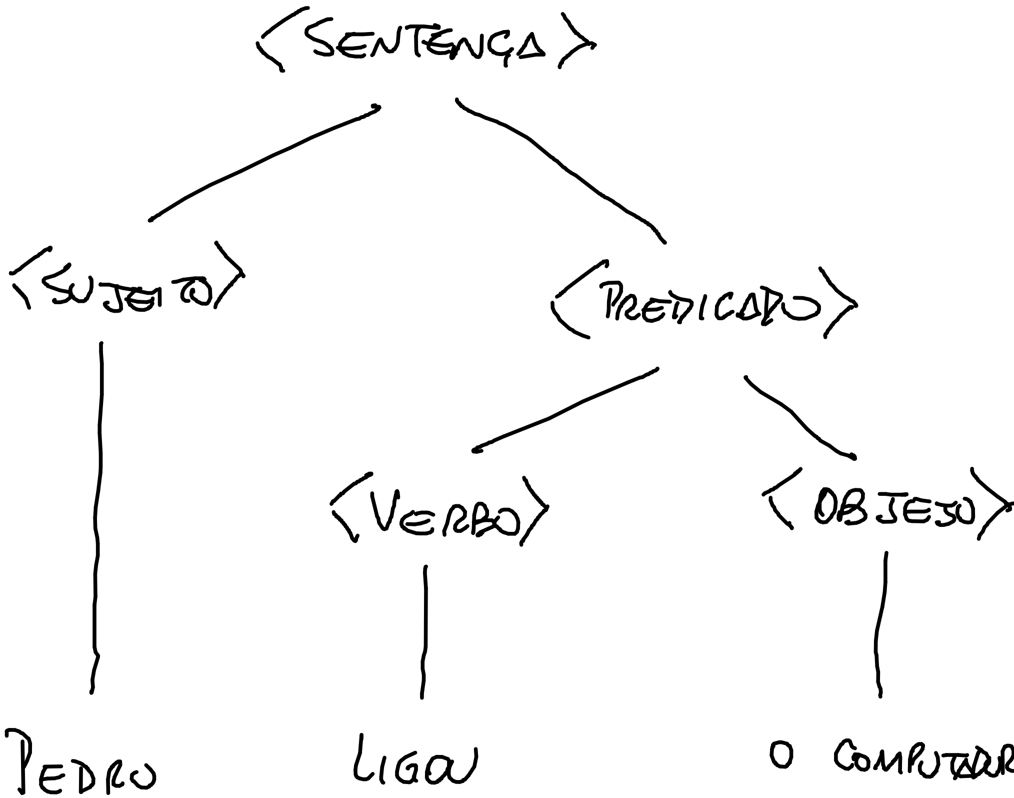


- Diagramar uma linguagem.

- 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 EQUERDA P/ DIREITA

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

SEJA V UM VÉRTICE NESTE CAMINHO

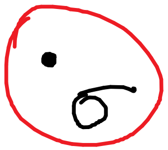
- V É ASCENDENTE DE V', E V' É DESCENTENTE DE V
- SE V E V' ESTÃO SEPARADOS POR UMA ÚNICA ARESTA ENTÃO V É PAI DE V' E V' É FILHO DE V
- FILHOS DO MESMO PAI SÃO IRMÃOS
- VÉRTICE SEM FILHOS É UMA FOLHA
- VÉRTICE QUE NÃO É FOLHA É INTERIOR

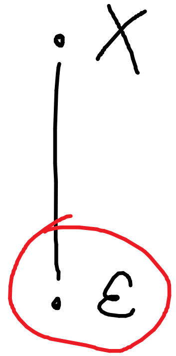
⇝ A ORDENAÇÃO ENTRE IRMÃOS PODE SER ESTENDIDA PARA UMA ORDENAÇÃO DAS FOLHAS.

DEFINIMOS ÁRVORES DE ANÁLISE SINTÁTICA <u>RECURSIVAMENTE</u>

FIXE GRAMÁTICA $G = (T, V, S, R)$

<u>BÁSICAS</u>: $\sigma \in T$, $X \in V$ e $X \to \varepsilon \in R$

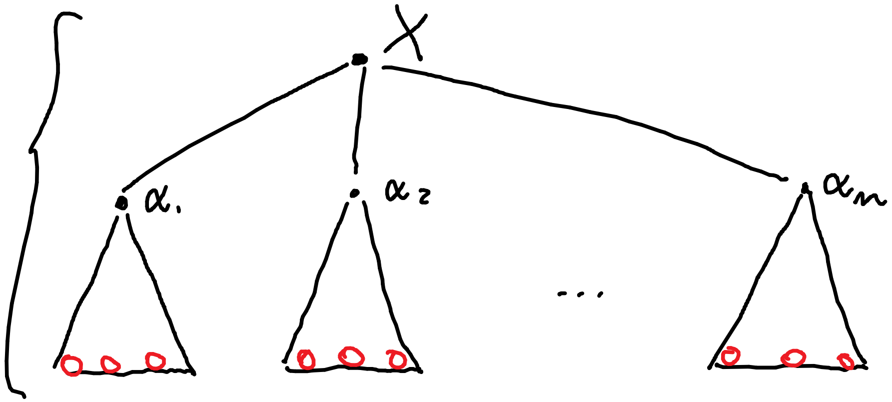




<u>REGRAS DE COMBINAÇÃO</u>: $T_1, \dots, T_n$ ÁRVORES COM RAÍZES $\alpha_1, \dots, \alpha_n$, RESP.

SE $X \to \alpha_1 \alpha_2 \cdots \alpha_n \in R$, ENTÃO

É UMA X-ÁRVORE

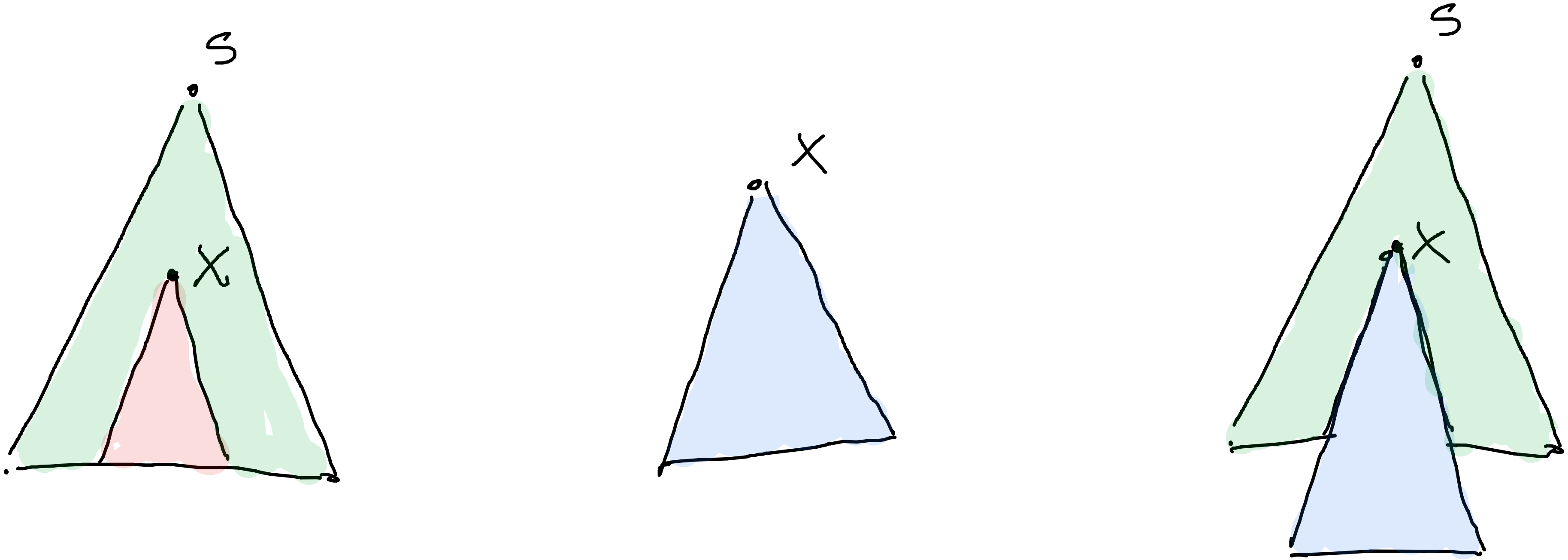


É ÁRVORE, COM REGRA ASSOCIADA

$$X \to \alpha_1 \alpha_2 \cdots \alpha_n$$

- EM PARTICULAR, PELAS REGRAS BÁSICAS, AS FOLHAS DE UMA ÁRVORE DE ANÁLISE SINTÁTICA SÃO APENAS ELEMENTOS EM $T \cup \{\varepsilon\}$

- E OS VÉRTICES INTERNOS SÃO VARIÁVEIS.



- OBS: SE UMA ÁRVORE τ POSSUI UM VÉRTICE ROTULADO POR X, PODEMOS SUBSTITUIR TODA SUBÁRVORE DE τ ENRAIZADA EM X, POR OUTRA ÁRVORE ENRAIZADA EM X.

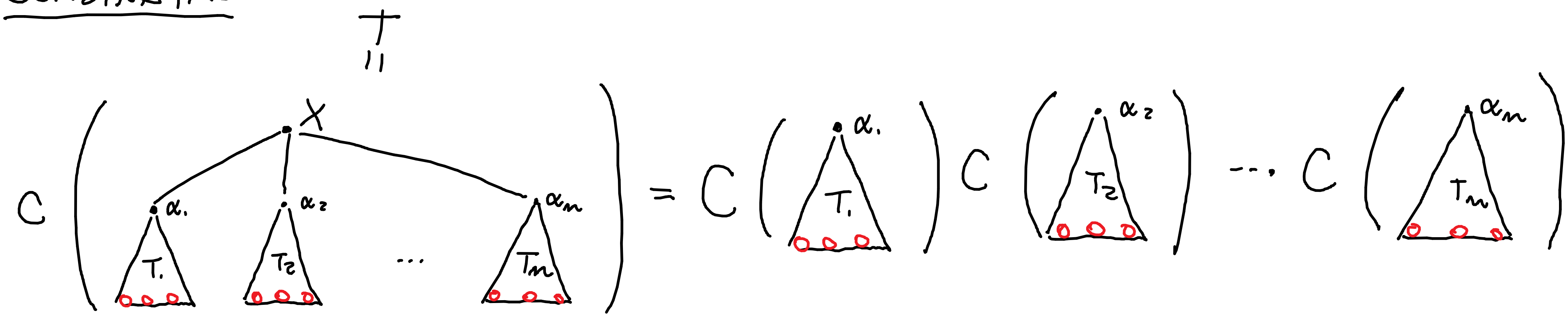
Colheita

Colheita ou resultado de uma árvore é uma função c que leva árvores em palavras

Básicas

$$c\left(\bullet_{\sigma}\right) = \sigma \qquad\qquad c\left(\begin{array}{c} \bullet\, X \\ | \\ \bullet\, \varepsilon \end{array}\right) = \varepsilon$$

Combinação



$$c(T) = c(T_1)\, c(T_2) \cdots c(T_n)$$

- 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 $w \in (T \cup V)^*$, $w = \alpha_1 \alpha_2 \cdots \alpha_m$.
UMA w-FLORESTA É UMA SEQUÊNCIA ORDENADA $T_1, \ldots, T_m$, DE ÁRVORES, EM QUE T_i É UMA α_i-ÁRVOR

- 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) INICIALIZE $\mathcal{F} = \mathcal{T}$

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

3) SE $w \in T^*$, ENTÃO PARE

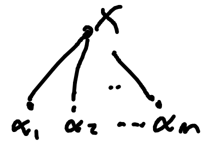
4) CASO CONTRÁRIO, SEJA v A RAIZ DA PRIMEIRA ÁRVORE DE $\mathcal{F}$ QUE É ROTULADA POR UMA VARIÁVEL

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

Proposição: Seja X uma variável de uma GLC. Se existe derivação $X \Rightarrow^* w$, então w é a colheita de uma X-árvore.

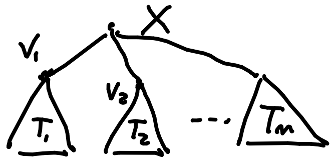
Prova: Indução no número p de passos da derivação.

Base: Um único passo. $X \Rightarrow w$. Então $X \to w \in R$.
Logo, se $w = \alpha_1 \cdots \alpha_m$, então



é uma X-árvore cuja colheita é w,

Passo indutivo: Suponha que toda p-derivação $Y \to w'$ pode ser "transformada" em uma Y-árvore cuja colheita é w'.
Seja $X \Rightarrow V_1 \cdots V_m$ o primeiro passo de $X \Rightarrow^* w$.
Isso quer dizer que $w = w_1 \cdots w_m$ e que existem derivações $V_i \Rightarrow^* w_i$ para todo i.
Mas $V_i \Rightarrow^* w_i$ 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. ∎

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

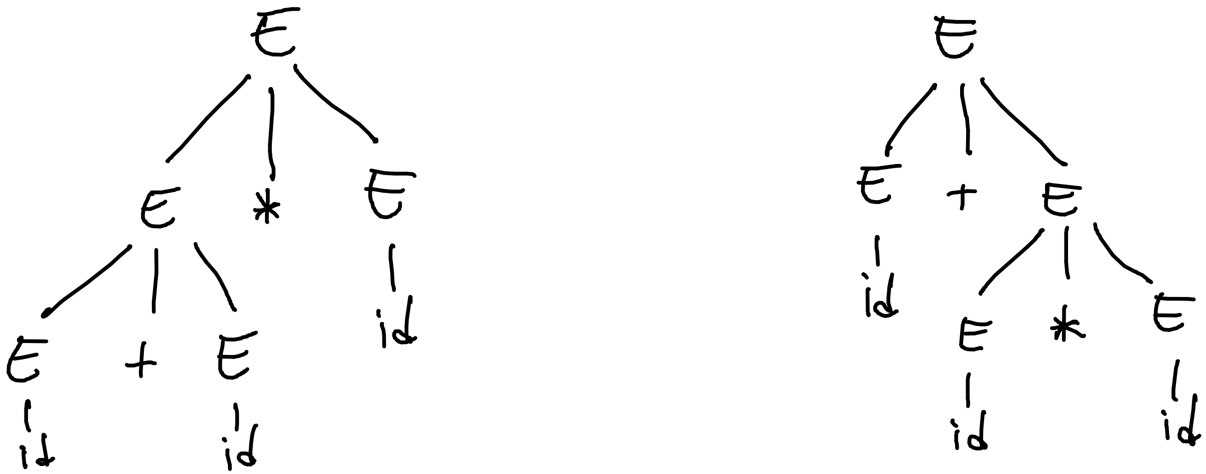
- ÁRVORES NOS AJUDAM A INTERPRETAR UMA FRASE CORRETAMENTE.
- ALGUMAS FRASES POSSUEM INTERPRETAÇÕES DISTINTAS

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

QUEM ESTÁ EMPURRANDO O CARRINHO?

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

EX:



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

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

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

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

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

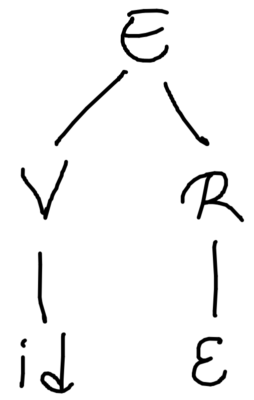
$S_U \Rightarrow S_1 \Rightarrow^* w$ E $S_U \Rightarrow S_2 \Rightarrow^* w$ SÃO DERIVAÇÕES QUE PRODUZEM ÁRVORES DE ANÁLISE SINTÁTICA DIFERENTES.

EX: A gramática G''_{EXP} com regras

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

- Podemos deduzir a derivação a partir da palavra dada.

Prova: Indução no comprimento da palavra w.



Se $|w| = 1$, então $w = id$

$$S \Rightarrow E \Rightarrow VR \Rightarrow idR \Rightarrow id$$
$$\Rightarrow V\varepsilon \Rightarrow id$$

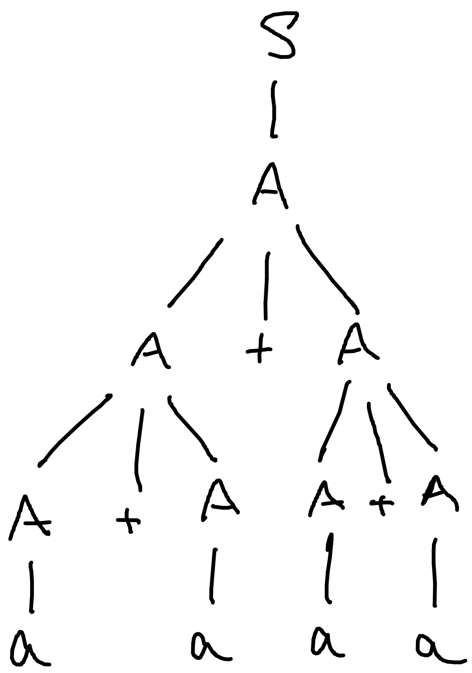
Se $|w| = n$ e $w = idw'$, então $|w'| < n$, e há uma única derivação $R \Rightarrow w'$

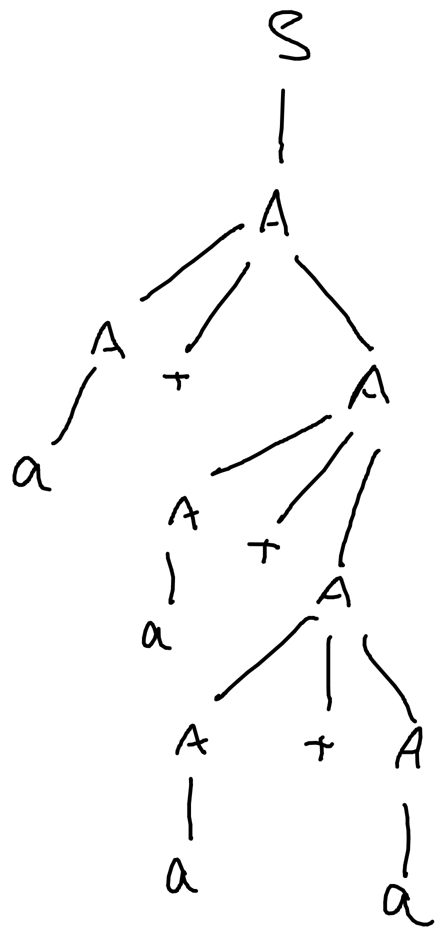
Se $|w| = n$, e $w = (w'$, então $w = (w_1)w_2$ e $|w_1|, |w_2| < n$
e temos $E \Rightarrow^* w_1$ e $R \Rightarrow^* w_2$

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

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

É AMBÍGUA





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

NÃO É AMBÍGUA.

<u>OBS:</u> 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

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 <u>AFD</u> M' PARA $L(G)$
3) DETERMINAR GRAMÁTICA G' OBTIDA DE M'.

ωX