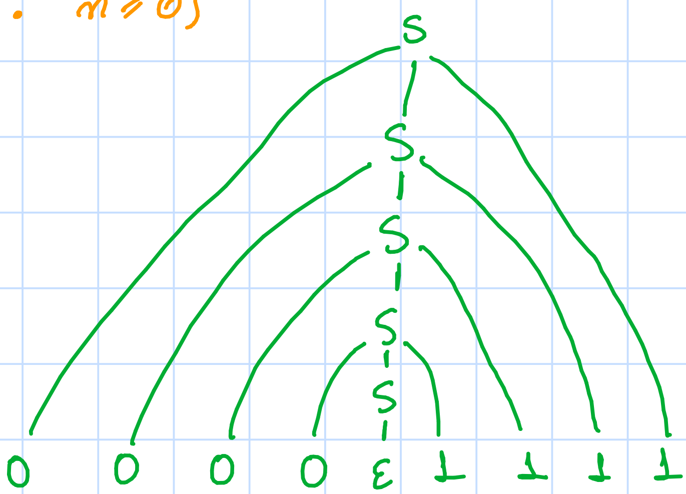
Lema do bombeamento para linguagens livres de contexto

Como as gramáticas livres de Contexto geram linguagens infinitas?

$\{0^n 1^n : n \geq 0\}$

$S \to 0S1$

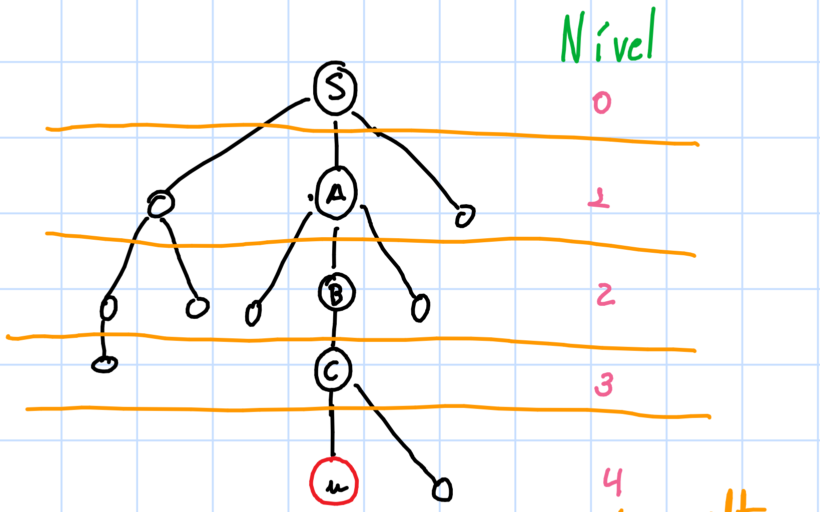
$S \to \varepsilon$



Recursão!

Qual a maior cadeia que uma gramática pode gerar sem recursão?

Seja $V = \{S, A, B, C\}$ o conjunto de variáveis.



altura máxima $\leq |V|$

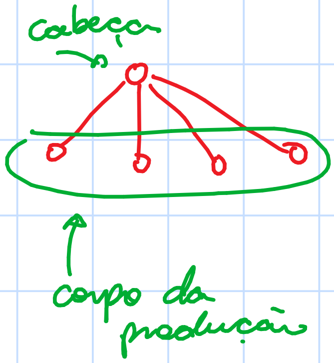
definição: A <u>altura de uma àrvore de derivação</u> é igual ao nível da folha mais profunda (níveis rotulados a partir de zero).

Fato1: Árvore de derivação sem recursão tem altura no máximo igual ao número de variáveis.

> Qual a maior cadeia que uma árvore de derivação com altura h pode gerar?

Seja G uma gramática e seja b o maior comprimento do corpo de uma produção de G.

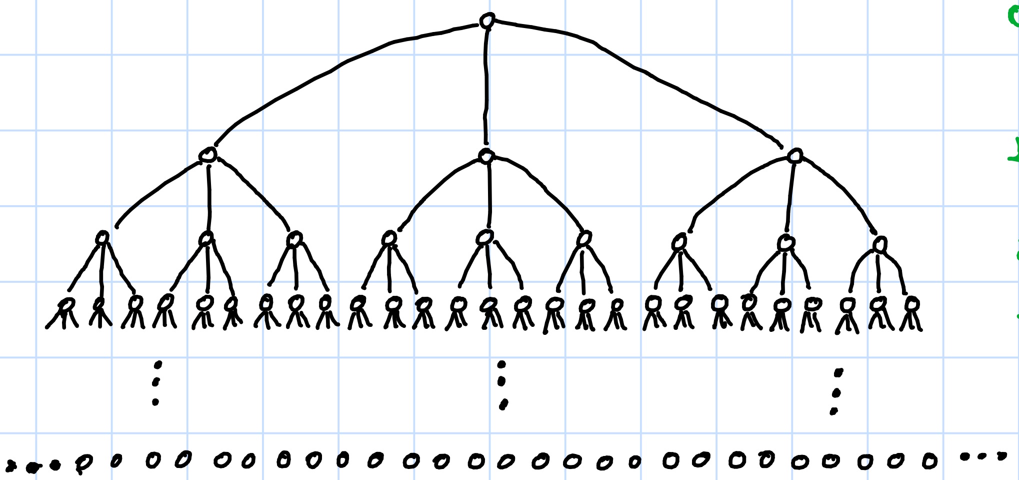
$b = 4$



$A \to 0A1A \mid A0A1$
$A \to B$
$B \to 3B \mid \varepsilon$

FATO A. Uma árvore de derivação com altura h pode gerar uma cadeia de comprimento no máximo b^h.

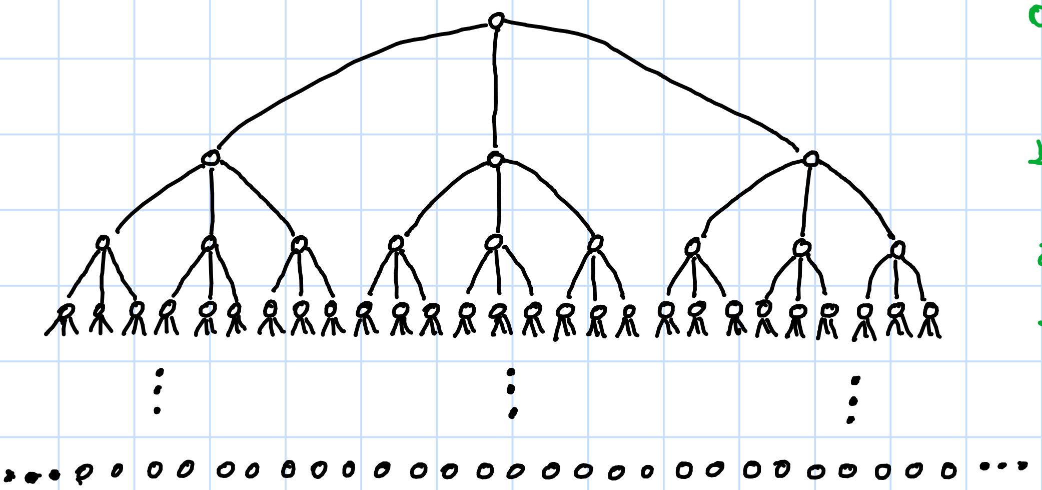
$b=3$



Nível	# nós
0	$1 = 3^0$
1	$3 = 3^1$
2	$3^2 = 9$
3	$3^3 = 27$
	$\vdots$
h	$3^h = b^h$

Fato B. Se w é uma cadeia tal que $|w| > b^h$, então a altura de uma árvore de derivação de w é maior que h.

Pelos fatos A e 1



Nível	# nós
0	b^0
1	b^1
2	b^2
3	b^3
$\lvert w\rvert$	$b^{\lvert w\rvert}$

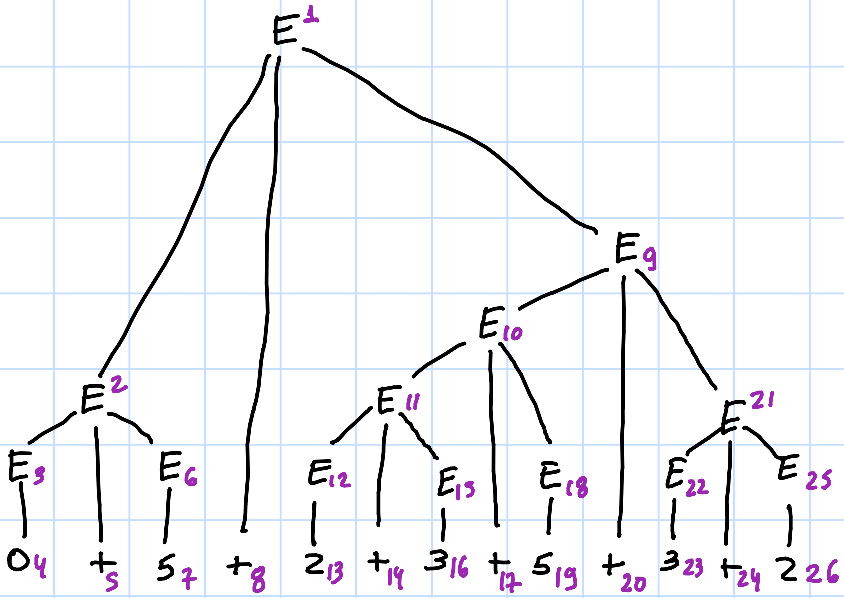
FATO: Uma gramática pode gerar, sem utilizar recursão, uma cadeia de comprimento no máximo $b^{|V|}$.

Se $|w| > b^{|V|}$, então a gramática usa recursão p/ derivar w.

Como a gramática explora a recursão?

Ex.:

$$E \to E+E \mid 0 \mid 1 \mid \cdots \mid 9$$

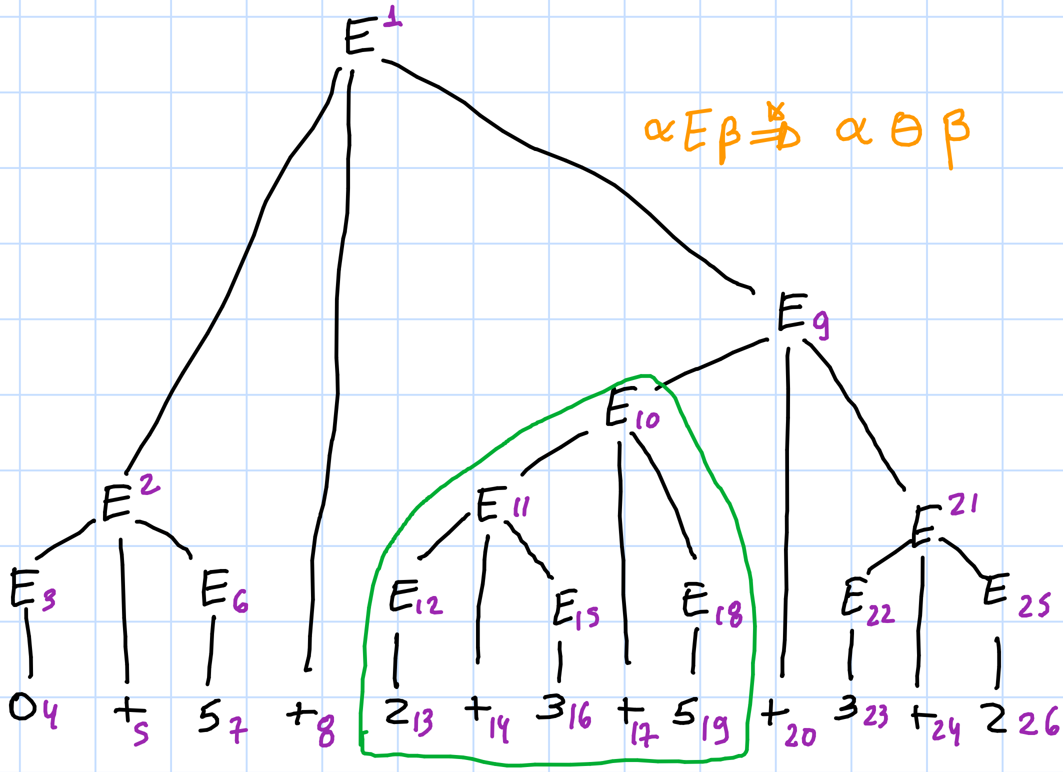


$$E_1 \Rightarrow E_2 +_8 E_9 \Rightarrow E_3 +_5 E_6 +_8 E_9 \Rightarrow 0_4 +_5 E_6 +_8 E_9 \Rightarrow 0_4 +_5 5_7 +_8 E_9$$

$$\Rightarrow 0_4 +_5 5_7 +_8 E_{10} +_{20} E_{21} \Rightarrow 0_4 +_5 5_7 +_8 E_{11} +_{17} E_{18} +_{20} E_{21} \Rightarrow$$

Uma subárvore enraizada define a derivação de um pedaço da sentença.

$\alpha E \beta \overset{*}{\Rightarrow} \alpha \theta \beta$

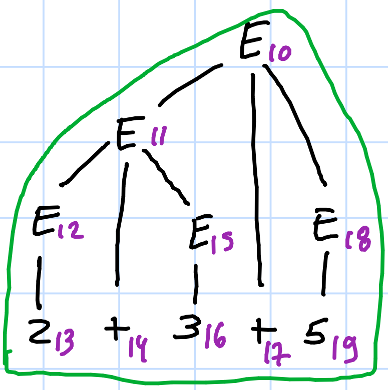


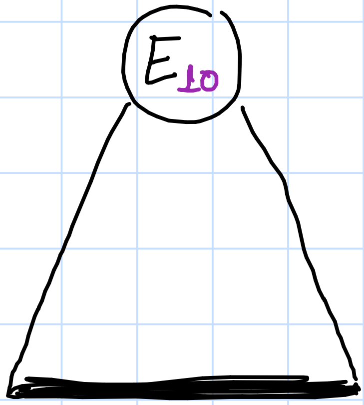
$$E_{10} \Rightarrow E_{11} +_{17} E_{18} \Rightarrow E_{11} +_{17} 5_{19} \Rightarrow E_{12} +_{14} E_{15} +_{17} 5_{19}$$

$$\Rightarrow E_{12} +_{14} 3_{16} +_{17} 5_{19} \Rightarrow 2_{13} +_{14} 3_{16} +_{17} 5_{19}$$

$$E_{10} \overset{*}{\Rightarrow} 2 + 3 + 5$$

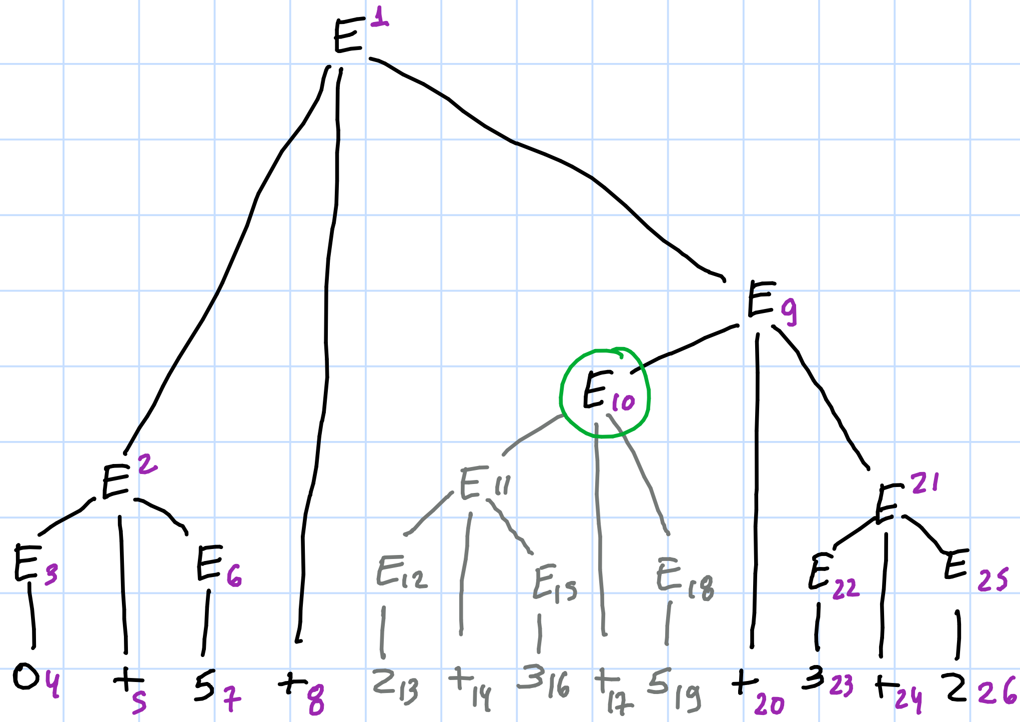
Representação esquemática





$x \in T^*$

Ao podarmos uma subárvore, obtemos uma forma sentencial



$$E_1 \Rightarrow E_2 +_8 E_9 \Rightarrow E_2 +_8 E_{10} +_{20} E_{21}$$

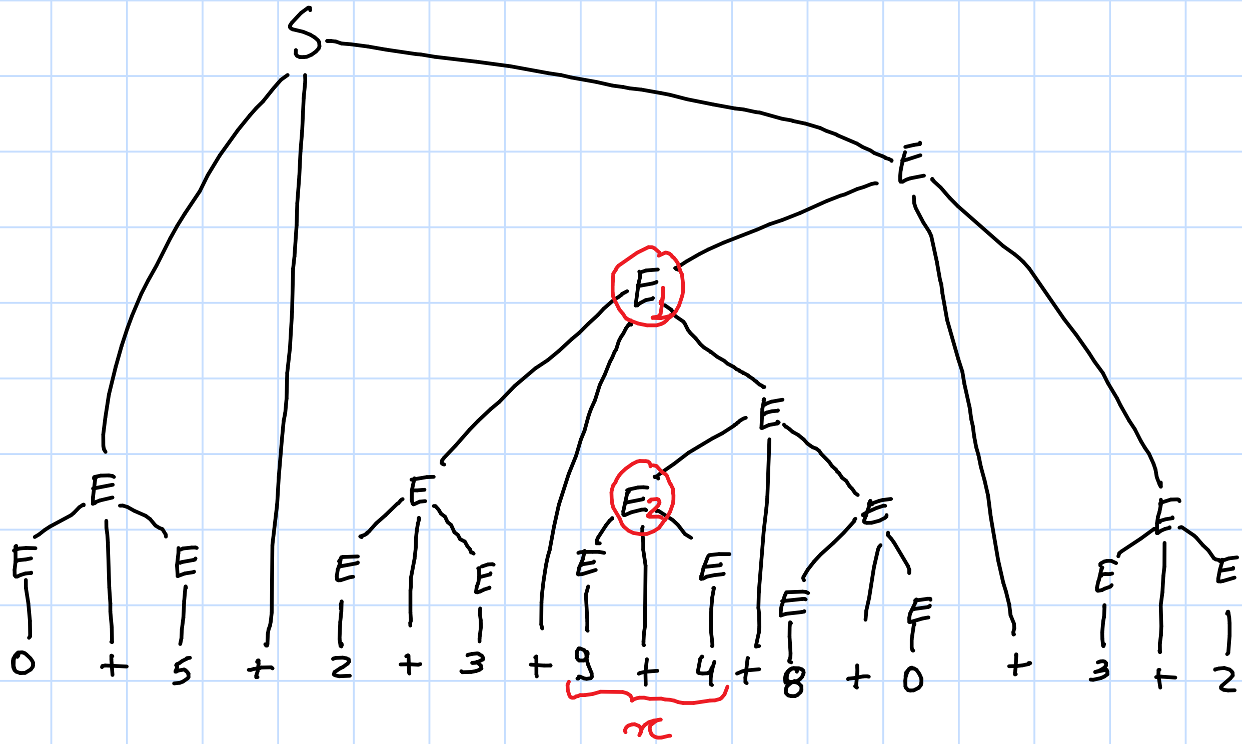
$$\overset{*}{\Rightarrow} E_2 +_8 E_{10} +_{20} 3_{23} +_{24} 2_{26}$$

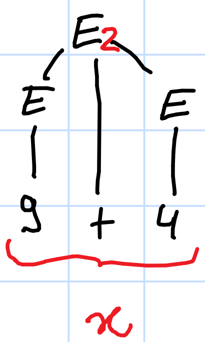
$$\overset{*}{\Rightarrow} 0_4 +_5 5_7 E_{10} +_{20} 3_{23} +_{24} 2_{26}$$

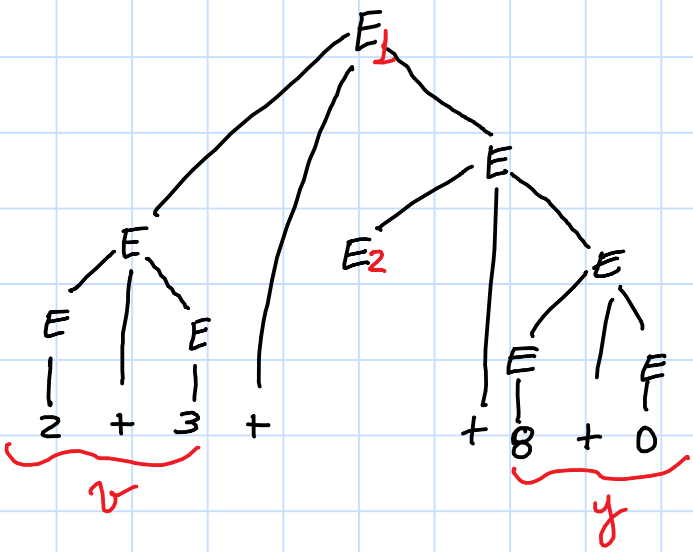
$$E_1 \overset{*}{\Rightarrow} 0_4 +_5 5_7 E_{10} +_{20} 3_{23} +_{24} 2_{26}$$

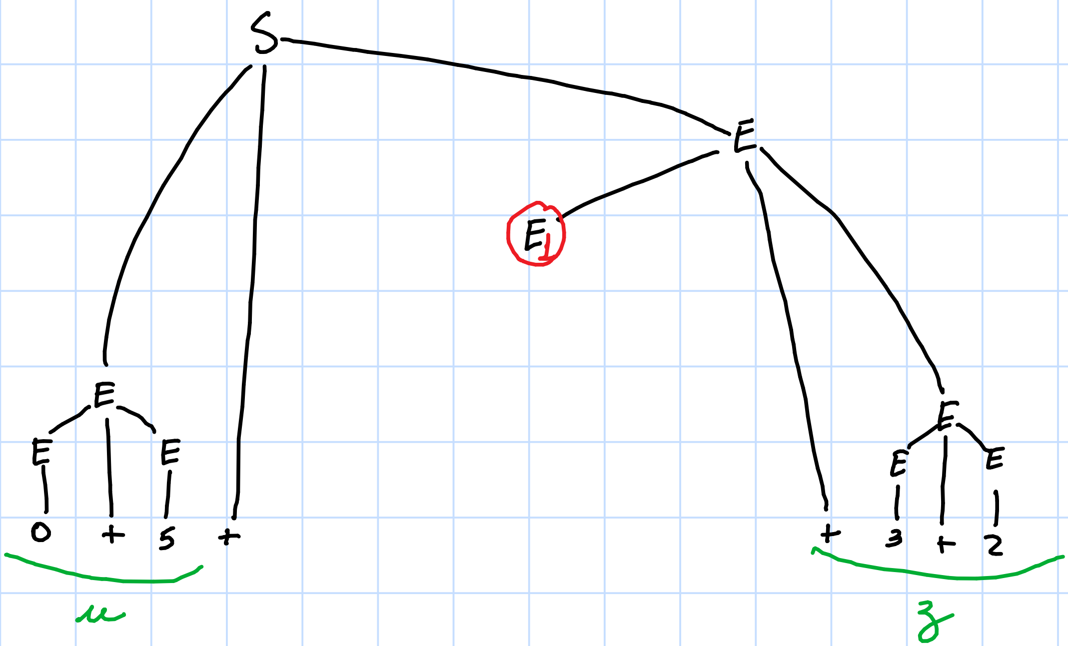
$S \to E$

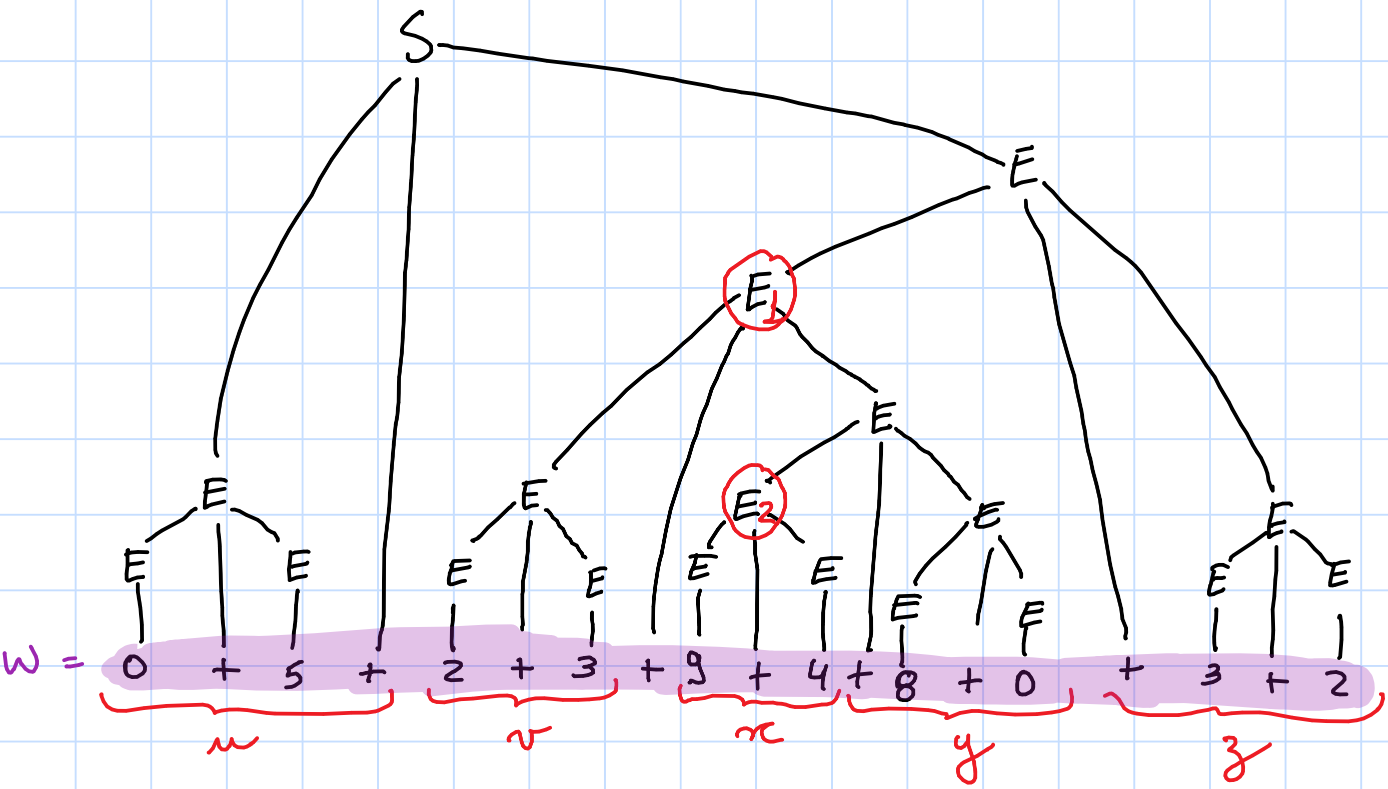
$E \to E+E \mid 0 \mid 1 \mid \cdots \mid 9$











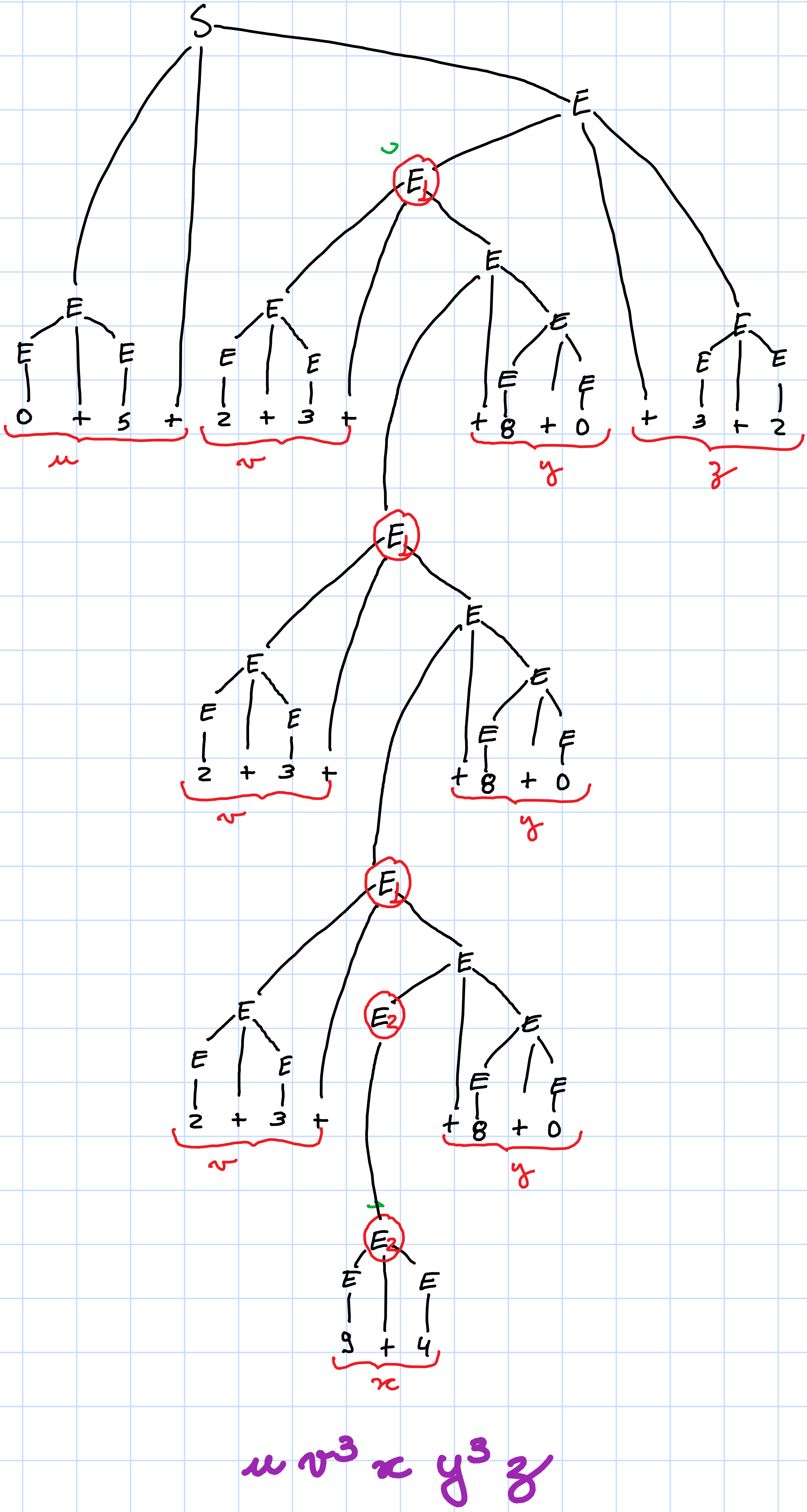
$w = uvxyz$

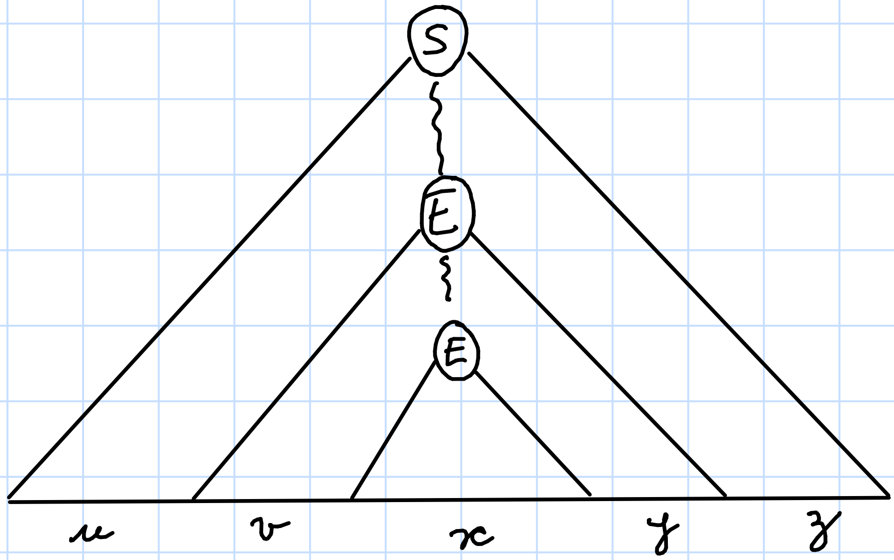
Obs:

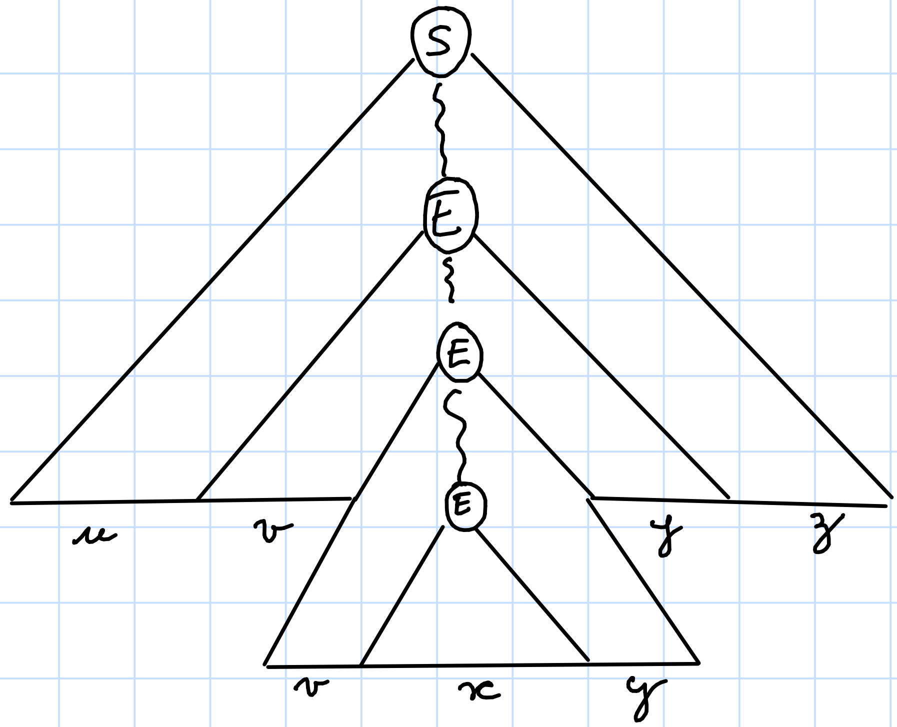
- $E \overset{*}{\Rightarrow} x$
- $E \overset{*}{\Rightarrow} vEy$
- $S \overset{*}{\Rightarrow} uEz$
- $S \overset{*}{\Rightarrow} uEz \overset{*}{\Rightarrow} uxz$
- $S \overset{*}{\Rightarrow} uEz \overset{*}{\Rightarrow} uvEyz \overset{*}{\Rightarrow} uvxyz$
- $S \overset{*}{\Rightarrow} uEz \overset{*}{\Rightarrow} uvEyz \overset{*}{\Rightarrow} uvvEyyz$

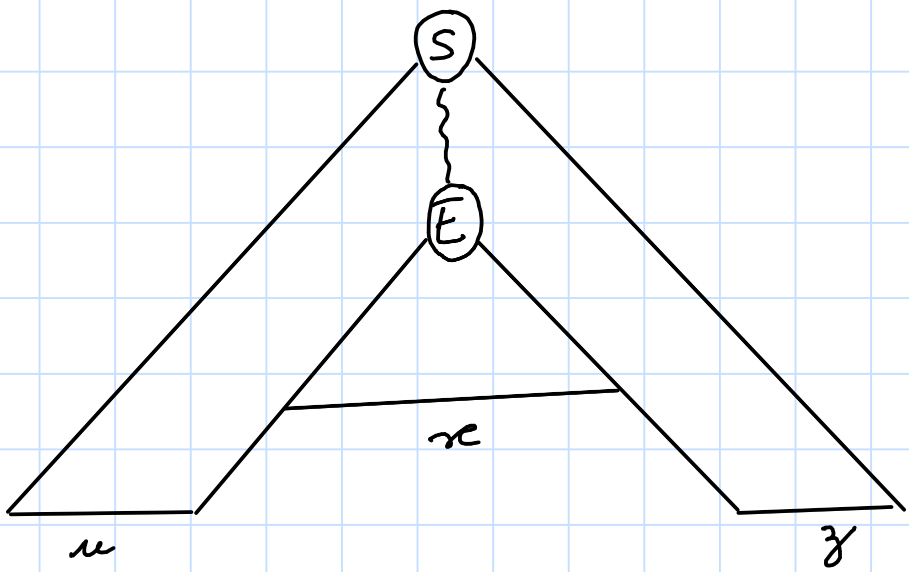
$\overset{*}{\Rightarrow} uvvvEyyyz \overset{*}{\Rightarrow} uvvvxyyyz$

$$\boxed{S \overset{*}{\Rightarrow} uv^ixy^iz \quad \text{p/ } i \geq 0}$$


S
E
E1
E2
0 + 5 +
2 + 3 +
+ 8 + 0
+ 3 + 2
9 + 4
u
v
x
y
z
uv3xy3z


S
E
E
u
v
x
y
z


S
E
E
E
u
v
y
z
v
x
y


S
E
x
u
z

Teorema

Se A é uma linguagem livre de contexto, então existe um número p (chamado de comprimento de bombeamento) tal que, se $w \in A$ e $|w| \geq p$, então w pode ser dividida em cinco partes $w = uvxyz$ satisfazendo as seguintes condições:

1. $uv^ixy^iz \in A$, para todo $i \geq 0$;
2. $|vy| > 0$;
3. $|vxy| \leq p$.

Demonstração

- Seja G uma gramática livre de contexto que gera a linguagem A.

podemos assumir que $b \geq 2$ (por quê?)

- Seja b o maior número de símbolos do lado direito de uma produção de G.

- Seja V o conjunto de variáveis de G.

- Seja $p = b^{|V|+1}$

- Seja $w \in A$ tal que $|w| \geq p$

FATO B. Se w é uma cadeia tal que $|w| > b^h$, então a altura de uma árvore de derivação de w é maior que h.

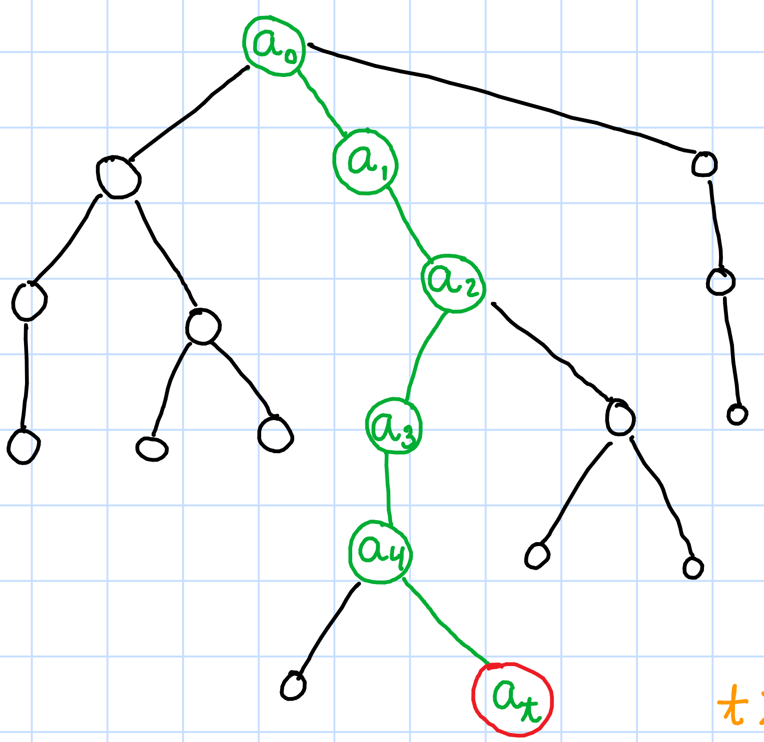
- Então a altura de qualquer árvore de derivação de ω é pelo menos $|V|+1$.

altura = # de níveis -1

qualquer árvore de derivação de w tem ao menos $|V|+2$ níveis

- Seja τ uma árvore de derivação de w com o menor número de nós.

- Seja q um nó folha de τ no maior nível de τ e seja $a_t, a_{t-1}, \dots a_0$ o caminho de q até a raiz de τ



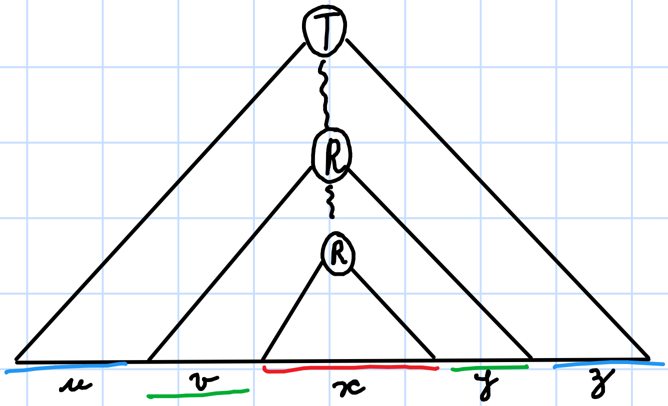
$t=5$

$t \geq |V|+2$

- Este caminho deve conter ao menos $|V|+1$ variáveis, então existe uma variável R que aparece mais de uma vez nesse caminho.

- Escolhemos R de forma que ela seja a primeira variável a se repetir nesse caminho.

- Dividimos w de acordo com a representação abaixo.



- Em outras palavras, se R_1 é a primeira ocorrência ao longo do caminho, R_2 a segunda e T é a raiz, temos que

(A) $R_1 \overset{*}{\Rightarrow} x$

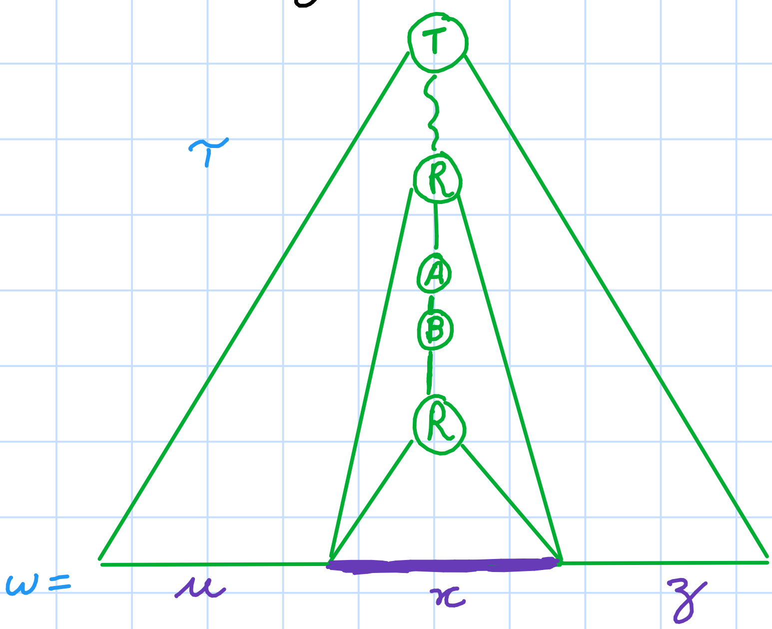
(B) $R_2 \overset{*}{\Rightarrow} v R_1 y$

(C) $T \overset{*}{\Rightarrow} u R_2 z$

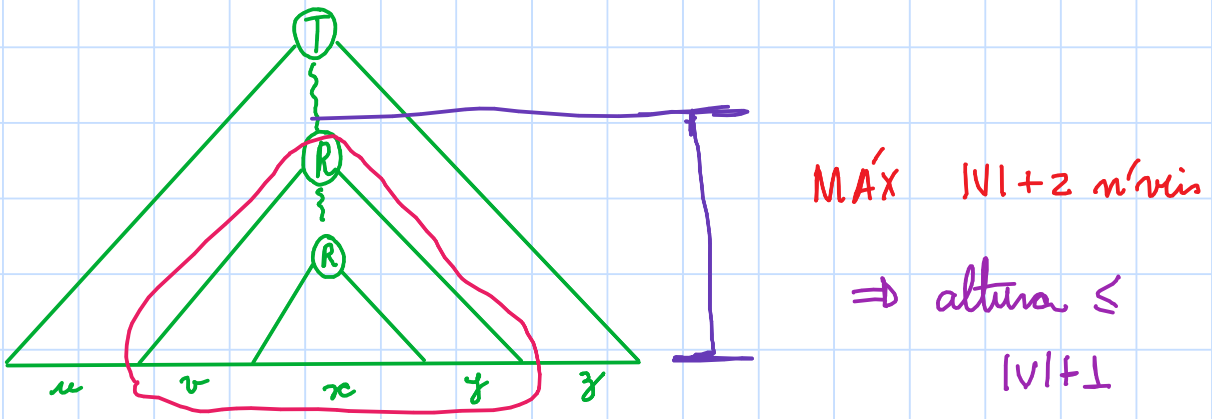
- Por (A), (B) e (C), temos que

$$T \overset{*}{\Rightarrow} u v^i x y^i z \quad \text{para } i \geq 0$$

- Se $|vy| = 0$, então ao eliminarmos o subcaminho entre R_1 e R_2, obteríamos uma árvore de derivação de w com menos nós, contrariando a escolha de τ. Então $|vy| > 0$.



- Agora, vamos analisar a condição ③



FATO A. uma árvore de derivação com altura h pode gerar uma cadeia de comprimento no máximo b^h.

$$|vxy| \leq b^{|V|+1} = p$$

■

Exercício

Use o lema do bombeamento para demonstrar que a linguagem

$$B = \{ a^n b^n c^n \mid n \geq 0 \}$$

não é livre de contexto.

Solução

- Suponha, para uma contradição, que B é uma linguagem livre de contexto.

- Seja p o comprimento de bombeamento de B.

- Seja $w = a^p b^p c^p$

- Note que $w \in B$ e $|w| \geq p$.

- Assim, w deve satisfazer as propriedades do lema do bombeamento.

- Podemos dividir w em cinco subcadeias u, v, x, y e z de modo que $w = uvxyz$ e valha o lema do bombeamento.

- Por ③, sabemos que $|vxy| \leq p$.

$$w = \quad u \quad \overbrace{v \quad x \quad y}^{\leq p} \quad z$$

$$\underbrace{aaa\cdots a}_{p}\,\underbrace{bbb\cdots b}_{p}\,\underbrace{ccc\cdots c}_{p}$$

n̄ contém:

vxy — a ou c

- Logo vxy não contém a's ou c's (ou ambos).

Caso 1. não veem a

$$\underbrace{aaa\cdots a}_{p}\overbrace{\underbrace{bbb\cdots b}_{p}\underbrace{ccc\cdots c}_{p}}^{vxy}$$

Caso 2. não veem c

$$\overbrace{\underbrace{aaa\cdots a}_{p}\underbrace{bbb\cdots b}_{p}}^{vxy}\underbrace{ccc\cdots c}_{p}$$

- Primeiro, suponha que vxy não contém a's.
- Seja $\theta = uv^2xy^2z$
- Como vxy não contém a's, sabemos que $|\theta|_a = p$

$$\underbrace{aaa\cdots a}_{p}\overbrace{\underbrace{bbb\cdots b}_{p}\underbrace{ccc\cdots c}_{p}}^{vxy}$$

- Sabemos que $|vy|>0$. Assim, adicionar uma cópia de v e y (bombear 1x) resulta na adição de pelo menos 1 símbolo b ou c.
- Assim, $|\theta|_b > p$ ou $|\theta|_c > p$ e $\theta \notin B$, uma contradição.

- Agora, suponha que vxy não contém c's

$\boxed{\text{Exercício}}$

$\square$

Roteiro para demonstrar que L não é ~~regular~~ LC

- Suponha, para uma contradição, que L é Livre de contexto.

- Seja p o comprimento de bombeamento de L
- Qualquer cadeia de comprimento maior que p pode ser bombeada
- Escolha uma cadeia $w \in L$, $|w| \geq p$, que seja conveniente

- Se $w = uvxyz$, então $uv^i x y^i z \notin L$, para algum $i \neq 1$.

$$w = \underbrace{w_1\, w_2\, w_3}_{u}\ \underbrace{w_4\, w_5}_{v}\ \underbrace{w_6\, w_7}_{x}\ \underbrace{w_8\, w_9}_{y}\ \underbrace{w_{10}}_{z}$$

$$w = \underbrace{w_1\, w_2\, w_3}_{u}\ \underbrace{w_4\, w_5}_{v}\ \underbrace{w_6\, w_7}_{x}\ \underbrace{w_8\, w_9\, w_{10}}_{y}$$

$z = \varepsilon$

- Conclua que w não pode ser bombeada

↳ L não é ~~regular~~ !
Livre de contexto.

Exercício (exemplo 2.37)

Demonstre, usando o lema do bombeamento, que a linguagem

$$C = \{ a^i b^j c^k \mid 0 \leq i \leq j \leq k \}$$

não é livre de contexto.

Demonstração

- Seja p o comprimento de bombeamento de C.
- Seja $w = a^p b^p c^p$
  - $\to w \in C$
  - $\to |w| \geq p$
- Pelo lema do bombeamento $w = uvxyz$ tq.
  1. $uv^i xy^i z \in C, \forall i \geq 0$
  2. $|vy| > 0$
  3. $|vxy| \leq p$

$w = u \; \overbrace{v \; x \; y}^{\leq p} \; z$

$w = \underbrace{aa a \cdots a}_{p} \underbrace{bbb \cdots b}_{p} \underbrace{ccc \cdots c}_{p}$

caso 1. não vem a

$\underbrace{aaa \cdots a}_{p} \overbrace{\underbrace{bbb \cdots b}_{p} \underbrace{ccc \cdots c}_{p}}^{vxy}$

caso 2. não vem c

$\overbrace{\underbrace{aaa \cdots a}_{p} \underbrace{bbb \cdots b}_{p}}^{vxy} \underbrace{ccc \cdots c}_{p}$

caso 1. não vem a $\quad vxy$

$\underbrace{aaa \cdots a}_{p} \underbrace{b \overbrace{bb}^{v} \cdots b}_{p} \underbrace{c \overbrace{c}^{y} c \cdots}_{p}$

b ou c < p

$\theta = uv^2 xy^2 z$

pode pertencer!

$C = \{ a^i b^j c^k \mid 0 \leq i \leq j \leq k \}$

$\theta' = uv^0 xy^0 z = uxz$

$|\theta'|_a = p$

$|w|_a = |w|_b = |w|_c = p$

$|w|_b = |w|_c = 2p$

$|\theta'|_b + |\theta'|_c < 2p \Rightarrow |\theta'|_b < p$ ou $|\theta'|_c < p$

$\Rightarrow \theta' \notin C$

$\theta' = a^p b^i c^j \notin C$

$i \leq p$

$j \leq p$

caso 2. não vem c

$\overbrace{\underbrace{aaa \cdots a}_{p} \underbrace{bbb \cdots b}_{p}}^{vxy} \underbrace{ccc \cdots c}_{p}$

v e y: contém b's ou a's

$C = \{ a^i b^j c^k \mid 0 \leq i \leq j \leq k \}$

$\phi = uv^2 xy^2 z$

$|vy| > 0$

$|w|_a = |w|_b = |w|_c = p$

$|\phi|_c = p$

- $|vy| > 0$ e $|vy|_c = 0 \Rightarrow |\phi|_a + |\phi|_b > 2p$

$\Rightarrow |\phi|_a > p$ ou $|\phi|_b > p$

$|\phi|_c = p$

$C = \{ a^i b^j c^k \mid 0 \leq i \leq j \leq k \}$

$aaa \cdots \underbrace{ab}_{v} \underbrace{bb}_{y} \cdots bccc \cdots c$

$\{ a^i b^j c^j : i \geq 0 \text{ e } j \geq 0 \}$

$\theta = uv^2 xy^2 z$

$\theta = aaa \cdots \underbrace{ab}_{v} \underbrace{ab}_{v} \underbrace{bb \cdots}_{y} \underbrace{\;}_{y} \quad bccc \cdots c$