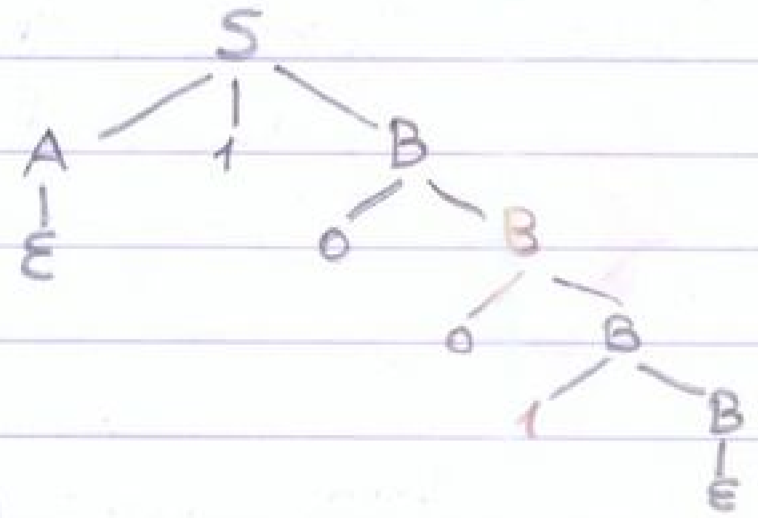
Prática 8 - Context-Free grammars (CFGs)

① a) Leftmost: $S \Rightarrow A1B \Rightarrow 1B \Rightarrow 10B \Rightarrow 100B \Rightarrow 1001B \Rightarrow 1001$

Rightmost: $S \Rightarrow A1B \Rightarrow A10B \Rightarrow A100B \Rightarrow A1001B \Rightarrow A1001 \Rightarrow 1001$

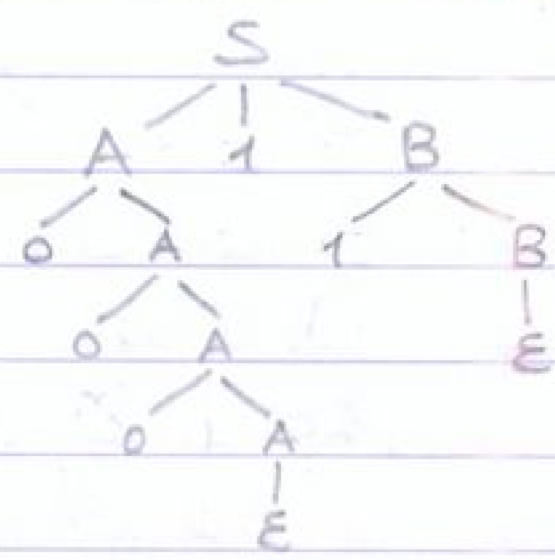
árvore de análise:



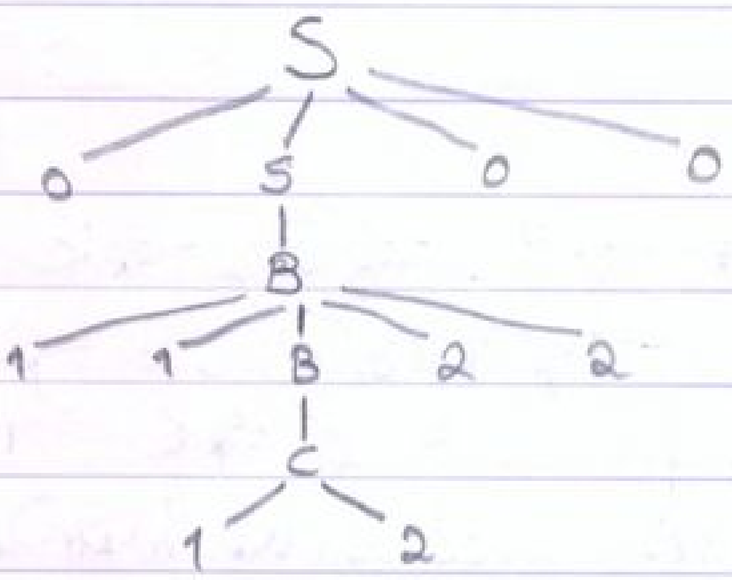
b) Leftmost: $S \Rightarrow A1B \Rightarrow 0A1B \Rightarrow 00A1B \Rightarrow 000A1B \Rightarrow 0001B \Rightarrow 00011B \Rightarrow 00011$

Rightmost: $S \Rightarrow A1B \Rightarrow A11B \Rightarrow A11 \Rightarrow 0A11 \Rightarrow 00A11 \Rightarrow 000A11 \Rightarrow 00011$

árvore de análise:



② a)



b) Linguagem das sequências iniciadas com zero ou mais 0s e terminadas com o dobro do número de 0s com que iniciaram, e que no meio tem um número ímpar de 1s seguido do mesmo número de 2s.

L(G) é o conjunto das cadeias da forma $0^m 1^n 2^n 0^{2m}$, com $m \geq 0$ e n ímpar:

$L = \{0^m 1^n 2^n 0^{2m} \mid$ com $m \geq 0$ e n ímpar$\}$.

c) 1: → Provar que as palavras aceites por G pertencem a L:

- $C \rightarrow 12$: 12 é a única palavra produzida por C e $\in L$.
- $B \rightarrow 11B22 \mid C$ (cadeias da forma $1^n 2^n$, com n ímpar).

caso base para B: $B \rightarrow C$ e a cadeia para o caso base é 12 e $\in L$.

passo indutivo para B: $B \rightarrow 11B22$. Seja w uma cadeia da linguagem produzida por B, então a próxima cadeia é dada por: 11w22

E como w é da forma 1^r2^r, com r ímpar, $\in L$, então $11w22$ corresponde a $1^{r+2}2^{r+2}$, que $\in L$.

• $S \rightarrow 0S00 \mid B$ (cadeias da forma $0^m1^r2^r0^{2m}$ com $m \geq 0$ e r ímpar).

caso base para S: $S \rightarrow B$ e cadeias em B são da forma 1^r2^r, com r ímpar e $\in L$.

passo indutivo para S: $S \rightarrow 0S00$ Seja w uma cadeia da linguagem e produzida por S, então a próxima cadeia produzida por S é dada por: $0w00$. Como w é da forma $0^m1^r2^r0^{2m}$, com $m \geq 0$ e r ímpar, $\in L$, então $0w00$ corresponde a $0^{m+1}1^r2^r0^{2(m+1)}$ que $\in L$.

Depois, devemos provar que todas as palavras de L são aceitas pela gramática G.

Indução em m:

base: $m=0$, $w=1^r2^r$, r ímpar, que para $r=1$ é derivada com $S \Rightarrow B \Rightarrow C \Rightarrow 12$ e para $r>1$ $S \Rightarrow B \Rightarrow 11B22$ (podendo derivar B recursivamente e no final usar $B \Rightarrow C \Rightarrow 12$.

passo indutivo: $0^{m+1}1^r2^r0^{2(m+1)} = 00^m1^r2^r0^{2m}00$

como pela hipótese $w = 0^m1^r2^r0^{2m}$ é aceita pela gramática e temos uma forma de produzir $0w00$ usando $S \rightarrow 0S00$, então as palavras aceitas por L por indução em m são também aceitas por G.

Indução em r:

base: $r=1$, 0^m120^{2m} $w=0^m120^{2m}$, $m \geq 0$, que para $m=0$ é derivada por $S \Rightarrow B \Rightarrow C \Rightarrow 12$, e para $m \geq 1$ é derivada com derivações recursivas em $S \Rightarrow 0S00$ terminadas com a derivação $S \Rightarrow B \Rightarrow C \Rightarrow 12$ para produzir $0^{m+1}120^{2m+2}$ e para $r>1$ $S \Rightarrow B \Rightarrow 11B22$ (podendo derivar B recursivamente e no final usar $B \Rightarrow C \Rightarrow 12$).

passo indutivo: $0^m1^{r+2}2^{r+2}0^{2m} = 0^m111^r2^r220^{2m}$.

($w_{m+1} = w_1 11 w_2 22 w_3$, em que $w_1 = 0^m$, $w_2 = 1^r2^r$, e $w_3 = 0^{2m}$)

como por hipótese $w_m = 0^m1^r2^r0^{2m}$ ($w_m = w_1w_2w_3$) é aceita pela gramática e temos uma forma de produzir $w_1 11 w_2 22 w_3$ usando $S \Rightarrow 0S00$ (podendo derivar B recursivamente e no final usar $S \Rightarrow B$) e depois usar $B \Rightarrow 11B22$ (podendo derivar B recursivamente e no final usar $B \Rightarrow C \Rightarrow 12$), então as palavras aceitas por L por indução em r são também aceitas por G.

Logo, está provado por indução.

(n → comprimento)

3 a) Caso base: $n = 1 \Rightarrow w = a$ ou $w = b$, e w não tem ba como subcadeia.

Passo indutivo: Por hipótese, w não tem ba como subcadeia.

As produções $S \rightarrow aS \mid Sb$ podem formar as cadeias aw ou wb.

Se w não tem ba, então aw e wb também não tem ba como subcadeia. Logo, está provado por indução.

b) Sequências de zero ou mais as seguidos de sequências de zero ou mais bs, mas que tem de haver pelo menos um a ou um b.

Formalmente: $L = \{a^n b^m, n \geq 0, m \geq 0$, n e m não são simultaneamente $0\}$.

4 Caso base: Seja a expressão regular do alfabeto $\{a, b\}$ $R = a$, isto é, o caso base corresponde às expressões regulares compostos por apenas um símbolo. Esta expressão é uma CFG: $S \longrightarrow a$.

Passo indutivo: Temos que mostrar que todos os operadores das expressões regulares são conversíveis em CFGs.

Sejam as expressões regulares A e B. que, pela hipótese, são conversíveis em CFGs.

Concatenação: $S \longrightarrow AB$ (CFG para a RE que resulta da concatenação de A e B).

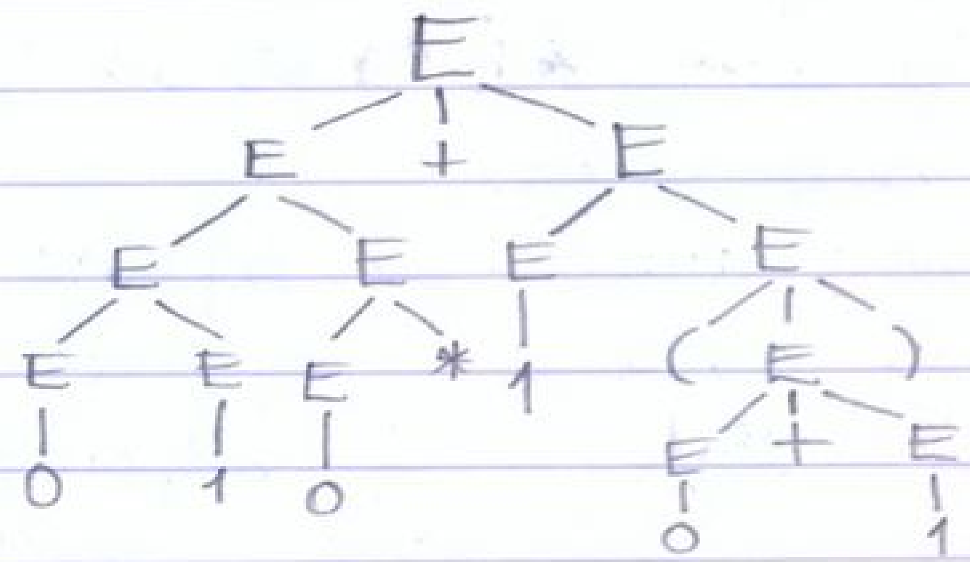
União: $S \longrightarrow A \mid B$ (CFG para a RE que resulta da união de A e B)

Fecho: $S \longrightarrow \varepsilon \mid AS$ (CFG para A^*)

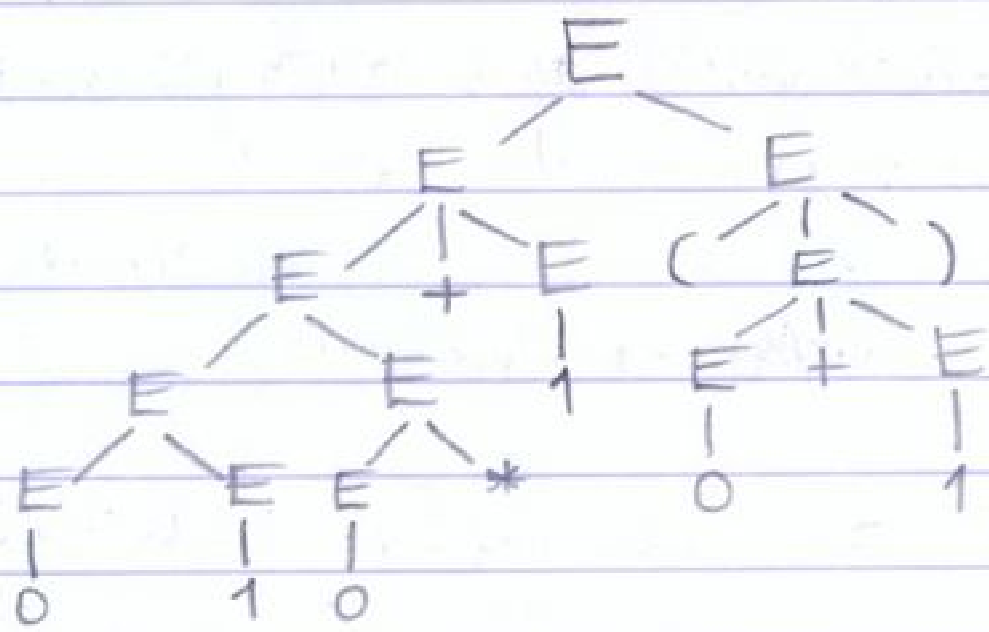
Parêntesis: $S \longrightarrow (S) \mid A$ (CFG para $(^*A)^*$)

Logo, todas as expressões regulares são conversíveis em CFGs.

⑤ a) $E \rightarrow 0 \mid 1 \mid e \mid \emptyset \mid EE \mid (E) \mid E+E \mid E^*$



b)



Sim. Esta árvore representa a mesma string e pode ser obtida por derivação mais à esquerda.

c) $E \rightarrow E+T \mid T$
$T \rightarrow TF \mid F$
$F \rightarrow F^* \mid (E) \mid 0 \mid 1 \mid \emptyset \mid e$

Não. Só existe uma derivação mais à esquerda.

⑥ a) ItemLista → <LI>Doc</LI>

b) Elemento → Texto | <P>Doc | <B>Doc<\B> | <OL>Lista<\OL> | <UL>Lista<\UL>

c) Elemento → (...) | <TABLE>Tabela<\TABLE>
Tabela → PrimeiraLinha RestLinhas
PrimeiraLinha → <TR>Cabecalho<\TR>
Cabecalho → <TH>Doc<\TH>Cabecalho | ε
Linha → <TD>Doc<\TD> Linha | ε
RestLinhas → <TR>Linha<\TR>RestLinhas | ε.

⑦ $w = \overbrace{\cdots\cdots}^{n}$ derivação de w: $G \rightarrow \underbrace{\cdots}_{m}$

Para cada derivação é adicionado um nó, portanto a árvore de sintaxe terá n nós finais mais m nós internos.

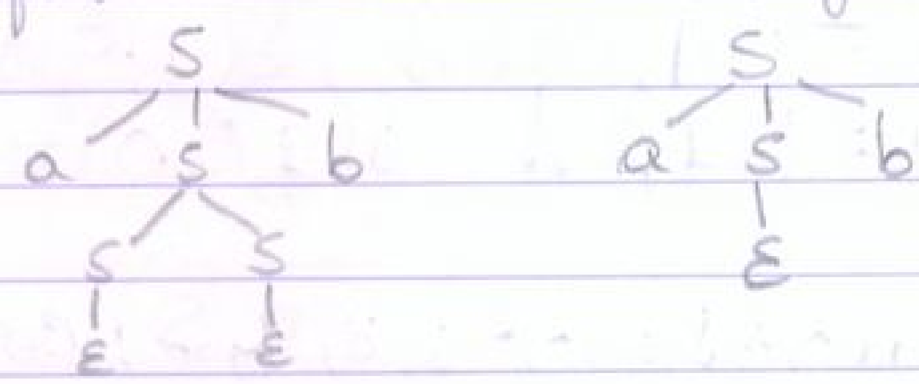
8) a) caso base: ε (0 as e 0 bs). Assume-se que S tem o mesmo nº de as e bs.
passo indutivo:

- Para SS: Pela hipótese, conclui-se que SS tem o mesmo número de as e bs, pois apenas é duplicada a quantidade de as e bs em S.
- Para ε: Não é afetado o número de caracteres existentes.
- Para aSb: Mantém-se a igualdade entre o número de as e bs, porque apenas se acrescenta um a e um b.
- Para bSa: O mesmo que para aSb.

Logo, fica provado por indução.

b) $S \Rightarrow SS \Rightarrow SaSb \Rightarrow SaaSbb \Rightarrow Saabb \Rightarrow bSaaabb \Rightarrow$
$\Rightarrow baaabb$

c) A gramática G é ambígua. Exemplo: ab



9) a) $L \rightarrow 0A$
$A \rightarrow 0A \mid 00A1 \mid \varepsilon$

b) $M \rightarrow Ab$
$A \rightarrow Bb \mid a \mid \varepsilon$
$B \rightarrow aA \mid aaA \mid \varepsilon$

c) $N \rightarrow A1$
$A \rightarrow B1 \mid 0A1$
$B \rightarrow 1C$
$C \rightarrow 1C0 \mid \varepsilon$

10) Linguagem com o mesmo número de 0s e 1s, independentemente da ordem.

11) Sim, é ambígua. Exemplo: 0101. 2 árvores de sintaxe para a mesma string, logo é ambígua.