Fichas de Exercícios sobre Gramáticas sem Contexto

Resoluções/soluções para os exercícios selecionados: 1,2,3,4,6

---

1 A gramática seguinte gera a linguagem das expressões regulares 0*1(0+1)*

```
S → A1B
A → 0A | ε
B → 0B | 1B | ε
```

Obtenha derivações mais à esquerda e mais à direita e árvores de análise das seguintes cadeias:

a) 1001

b) 00011.

1a)

lm (leftmost): S ⇒ A1B ⇒ 1B ⇒ 10B ⇒ 100B ⇒ 1001B ⇒ 1001

rm (rightmost) S ⇒ A1B ⇒ A10B ⇒ A100B ⇒ A1001B ⇒ A1001 ⇒ 1001

1b)

lm (leftmost): S ⇒ A1B ⇒ 0A1B ⇒ ...(passos omitidos mas que devem ser incluidos) ⇒ 00011B ⇒ 00011

rm (rightmost) S ⇒ A1B ⇒ A11B ⇒ ...(passos omitidos mas que devem ser incluidos) ⇒ 000A11 ⇒ 00011

---

2 Seja a gramática G= (V, T, P, S) em que T = {0,1,2}, V= {S, B, C} e as produções são:

```
S → 0 S 0 0 | B
B → 1 1 B 2 2 | C
C → 1 2
```

a) Desenhe uma árvore de análise da sequência 011122200.

b) Explique em Português qual é a linguagem aceite pela gramática.

c) Prove por indução que a afirmação da alínea anterior está correcta.

2a)

```
      S
 /    |    \    \
 0    S    0    0
      |
      B
 /    |    \    \
 1    C    2    2
 /         \
 1         2
```

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

Mais formalmente: L(G) é o conjunto das cadeias da forma $0^n1^r2^r0^{2n}$, com $n \geq 0$ e r ímpar, i.e.,

$L = \{0^n1^r2^r0^{2n} \mid \text{com } n \geq 0 \text{ e r ímpar}\}$

2c)

Provar por indução que a linguagem especificada pela gramática G é $L = \{0^n1^r2^r0^{2n} \mid \text{com } n \geq 0 \text{ e r ímpar}\}$

(1) Primeiro devemos provar que as palavras aceites pela gramática G $\in$ L:

A gramática, G, inclui três variáveis, com produções:

- C $\rightarrow$ 1 2
  12 é a única palavra produzida por C e $\in$ L

- B $\rightarrow$ 1 1 B 2 2 | C (cadeias da forma 1^r2^r, com r ímpar)

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

  passo indutivo para B:
  As cadeias produzidas indutivamente com B usam a regra:
  B $\rightarrow$ 1 1 B 2 2
  Seja W uma cadeia da linguagem (i.e., $\in$ L) produzida por B, então a próxima cadeia produzida por B é dada por:
  1 1 W 2 2
  como W é da forma 1^r2^r, com r ímpar, $\in$ L, então 1 1 W 2 2 corresponde a $1^{r+2}2^{r+2}$ que $\in$ L

- S $\rightarrow$ 0 S 0 0 | B (cadeias da forma $0^{n1r}2^r0^{2n}$, com $n \geq 0$ e r ímpar)

  caso base para S: S $\rightarrow$ B

  cadeias em B são da forma 1^r2^r, com r ímpar e $\in$ L

  passo indutivo para S:

  S $\rightarrow$ 0 S 0 0
  Seja W uma cadeia da linguagem (i.e., $\in$ L) produzida por S, então a próxima cadeia produzida por S é dada por:
  0 W 0 0

como W é da forma $0^n1^r2^r0^{2n}$, com $n \geq 0$ e r ímpar, $\in$ L, então 0 W 0 0 corresponde a $0^{n+1}1^r2^r0^{2n+2} = 0^{n+1}1^r2^r0^{2(n+2}$ que $\in$ L

qed

(2) Depois devemos provar que todas as palavras de $L = \{0^n1^r2^r0^{2n} \mid \text{com } n \geq 0 \text{ e r ímpar}\}$ são aceites pela gramática:

Indução em n:

base: n=0, $w=1^r2^r$, r ímpar, que para r=1 é derivado com $S \Rightarrow B \Rightarrow C \Rightarrow 1\ 2$

e para r > 1 $S \Rightarrow B \Rightarrow 1\ 1\ B\ 2\ 2$ (podendo derivar B recursivamente e no final usar $B \Rightarrow C \Rightarrow 1\ 2$)

passo indutivo: $0^{n+1}1^r2^r0^{2(n+1)} = 00^n1^r2^r0^{2n}00$,

como por hipótese $w=0^n1^r2^r0^{2n}$ é aceite pela gramática e temos uma forma de produzir 0w00 usando $S \rightarrow 0\ S\ 0\ 0$, então as palavras aceites por L por indução em n são também aceites por G.

Indução em r:

base: r=1, 0^n120^{2n}

$w=0^n120^{2n}$, $n \geq 0$, que para n=0 é derivada por $S \Rightarrow B \Rightarrow C \Rightarrow 1\ 2$, e para $n \geq 1$ é derivada com derivações recursivas em $S \Rightarrow 0\ S\ 0\ 0$ terminadas com a derivação $S \Rightarrow B \Rightarrow C \Rightarrow 1\ 2$ para produzir $0^{n+1}120^{2n+2}$

e para r > 1 $S \Rightarrow B \Rightarrow 1\ 1\ B\ 2\ 2$ (podendo derivar B recursivamente e no final usar $B \Rightarrow C \Rightarrow 1\ 2$)

passo indutivo: $0^n1^{r+2}2^{r+2}0^{2n} = 0^n111^r2^r220^{2n}$ ($w_{n+4}=w_1 11 w_2 22 w_3$, em que $w_1=0^n$, $w_2=1^r2^r$, e $w_3=0^{2n}$)

como por hipótese $w_n=0^n1^r2^r0^{2n}$ ($w_n= w_1w_2w_3$) é aceite pela gramática e temos uma forma de produzir $w_1 11 w_2 22 w_3$ usando $S \Rightarrow 0\ S\ 0\ 0$ (podendo derivar S recursivamente e no final usar $S \Rightarrow B$) e depois usar $B \Rightarrow 1\ 1\ B\ 2\ 2$ (podendo derivar B recursivamente e no final usar $B \Rightarrow C \Rightarrow 1\ 2$), então as palavras aceites por L por indução em r são também aceites por G.

qed

---

3 Considere a CFG G definida pelas produções

$$S \rightarrow aS \mid Sb \mid a \mid b.$$

a) Prove por indução no comprimento da cadeia que nenhuma cadeia de L(G) tem **ba** como subcadeia.

b) Descreva L(G) informalmente. Justifique a resposta com base em a).

3a)

Caso base: **a** ou **b** que não têm **ba** como subcadeia

Paso indutivo:

por hipótese a cadeia *x* verifica a propriedade (não tem **ba** como subcadeia)

as produções: $S \rightarrow aS \mid Sb$ podem formar as cadeias

ax ou xb

se x não tem ba, então ax e xb também não têm **ba** como subcadeia qed

3b)

Sequências de 0 ou mais a's seguidas de sequências de 0 ou mais b's, mas que tem de haver pelo menos um a ou um b.

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

---

4 Mostre que qualquer linguagem regular é uma linguagem sem contexto. Sugestão: construa uma CFG por indução no número de operadores da expressão regular.

Pode basear-se em:

Caso base: símbolo ...

Passo indutivo: Definir produções para cada operador das expressões regulares:

Concatenação: ...

União: ...

Fecho: ...

Parêntesis: ...

---

6 Considere o seguinte fragmento de uma gramática para HTML:

Char → a | A | …	Elemento → Texto |
Texto → ε | Char Texto	<P> Doc |
Doc → ε | Elemento Doc	<B> Doc </B> |
	<OL> Lista </OL>
Lista → ε | ItemLista Lista	ItemLista → <LI> Doc

Adicione os seguintes elementos à definição:

a) Um item de uma lista deve ser fechado por uma marca </LI>.

b) Os elementos devem incluir as listas não ordenadas, com marcas <UL> e </UL>.

c) Incluir como elementos as tabelas, marcadas por <TABLE>, </TABLE>, com linhas <TR>, </TR>. A primeira linha tem componentes de cabeçalho <TH>, </TH>. As outras linhas têm componentes de dados <TD>, </TD>.

6a) Substituir ItemLista por:

ItemLista → <LI> Doc </LI>

6b) Substituir Elemento por:
Elemento → Texto | <P> Doc | <B> Doc </B> | <OL> Lista </OL> | <UL> Lista </UL>

6c) 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