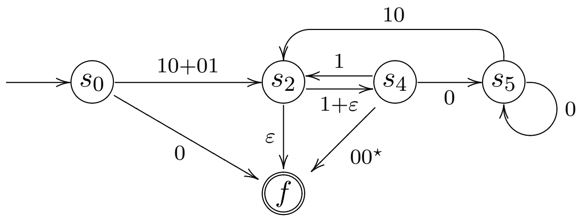
Modelos de Computação (CC218) 2013/14

Exame (Recurso) – 10.07.2014 *duração: 3h*

Uma resolução (parcial) **Cotação: 1+2+2, 3, 1+2+1+2+1.5, 1.5+1+2**

1. O diagrama seguinte foi obtido de um autómato finito A após algumas iterações do método de eliminação de estados (de Brzozowski-McCluskey).



a) Justifique que `011100101100000` $\in \mathcal{L}(A)$.

Resposta:

O método de eliminação preserva $\mathcal{L}(A)$, o que implica que, em cada iteração do método, as linguagens descritas por expressões regulares determinadas por percursos de s_0 para f estão contidas em $\mathcal{L}(A)$. Assim, podemos concluir que `011100101100000` $\in \mathcal{L}(A)$ porque o percurso

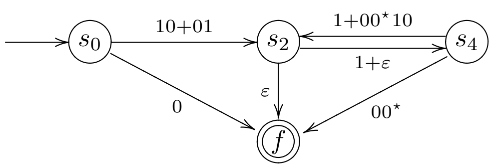
$$s_0 \xrightarrow{10+01} s_2 \xrightarrow{1+\varepsilon} s_4 \xrightarrow{1} s_2 \xrightarrow{1+\varepsilon} s_4 \xrightarrow{0} s_5 \xrightarrow{0} s_5 \xrightarrow{10} s_2 \xrightarrow{1+\varepsilon} s_4 \xrightarrow{1} s_2 \xrightarrow{1+\varepsilon} s_4 \xrightarrow{00^\star} f$$

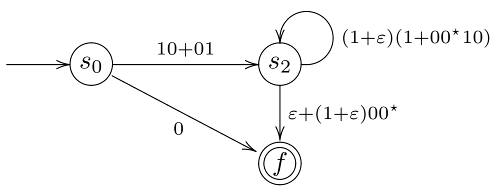
corresponde à expressão $r = (10+01)(1+\varepsilon)(1+\varepsilon)(1+\varepsilon)(0)(0)(10)(1+\varepsilon)(1+\varepsilon)(1+\varepsilon)(00^\star)$ e `011100101100000` $\in \mathcal{L}(r)$, uma vez que `011100101100000` $= (01)(1)(1)(\varepsilon)(0)(0)(10)(1)(1)(\varepsilon)(00000)$.

b) Determine uma expressão regular que descreva $\mathcal{L}(A)$. Deverá apresentar os passos intermédios.

Resposta:

Obtém-se a expressão se se prosseguir a eliminação dos estados. Se se eliminar s_5 e depois s_4, obtém-se:





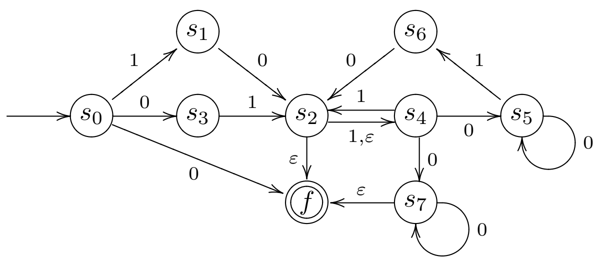
Resta eliminar s_2 no último (diagrama acima à direita): $\longrightarrow s_0 \xrightarrow{0+(10+01)((1+\varepsilon)(1+00^\star 10))^\star(\varepsilon+(1+\varepsilon)00^\star)} f$.

Conclui-se que $\mathcal{L}(A)$ pode ser descrita pela expressão regular $0 + (10 + 01)((1 + \varepsilon)(1 + 00^\star 10))^\star(\varepsilon + (1 + \varepsilon)00^\star)$.

c) Apresente o diagrama de transição de um autómato finito equivalente a A. Se não for determinístico, converta-o num determinístico equivalente. Apresente os passos principais da resolução.

Resposta:

Se se partir do autómato finito representado abaixo e se eliminar s_6, s_3, s_1, e s_7, o diagrama que se obtém é o que consta do enunciado. Assim, tal autómato finito é equivalente ao autómato finito A aí referido.

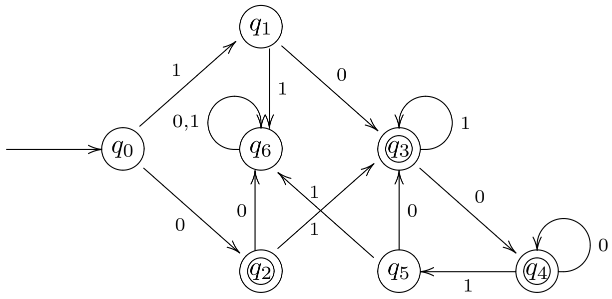


(continua)

Resposta 1c) cont.:

Como se trata de um AFND-ε e nos é pedido um AFD, vamos convertê-lo usando o método dado no curso para conversão de um AFND-ε num AFD. Nesse AFD, o estado inicial é $Fecho_\varepsilon(s_0)$ que, neste caso, é $\{s_0\}$. A função de transição δ' definida por $\delta'(E,a) = Fecho_\varepsilon\left(\bigcup_{e\in Fecho_\varepsilon(E)} \delta(e,a)\right)$, com $a \in \Sigma$, está representada na tabela. Os estados finais são os que correspondem aos conjuntos designados por q_2, q_3 e q_4.

	Símbolos / Estados	0	1
q_0	$\{s_0\}$	$\{s_3, f\}$	$\{s_1\}$
q_1	$\{s_1\}$	$\{s_2, s_4, f\}$	$\{\}$
q_2	$\{s_3, f\}$	$\{\}$	$\{s_2, s_4, f\}$
q_3	$\{s_2, s_4, f\}$	$\{s_5, s_7, f\}$	$\{s_2, s_4, f\}$
q_4	$\{s_5, s_7, f\}$	$\{s_5, s_7, f\}$	$\{s_6\}$
q_5	$\{s_6\}$	$\{s_2, s_4, f\}$	$\{\}$
q_6	$\{\}$	$\{\}$	$\{\}$



2. Resolva apenas uma das alíneas 2a) ou 2b). Se resolver ambas, só será avaliada a alínea 2b).
Se tiver dúvidas sobre a pergunta 1., deve resolver 2b).

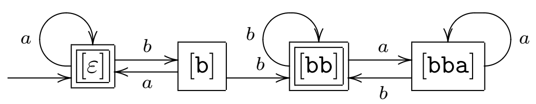
a) Na continuação de **1.**, determine o autómato finito determinístico mínimo que aceita $\mathcal{L}(A)$. Justifique.

Resposta omitida: O AFD mínimo pode ser obtido por aplicação do algoritmo de Moore ao AFD determinado em 1c). Note-se também que as questões 1c) e 2a) podiam ser resolvidas conjuntamente, se se determinasse o AFD mínimo seguindo a caraterização dada pelo teorema de Myhill-Nerode, com análise do diagrama representado no enunciado. Contudo, como os ramos desse diagrama estão etiquetados com expressões regulares, a justificação da equivalência ou não equivalência de palavras segundo R_L requer algum cuidado (e rigor).

b) Seja L a linguagem de alfabeto $\{\mathtt{a}, \mathtt{b}\}$ constituída pelas palavras que terminam b se e só se têm bb como subpalavra. Por aplicação do teorema de Myhill-Nerode, determine o autómato finito determinístico mínimo que aceita L. Justifique.

Resposta

$\delta([\varepsilon], \mathtt{b}) \stackrel{\text{def}}{=} [\mathtt{b}]$	$[\mathtt{b}] \neq [\varepsilon]$ porque $\mathtt{b} \notin L$ e $\varepsilon \in L$.
$\delta([\varepsilon], \mathtt{a}) \stackrel{\text{def}}{=} [\mathtt{a}] = [\varepsilon]$	pois $\mathtt{a}z \in L$ se e só se z não tem bb como subpalavra e z não termina em b ou z tem bb como subpalavra e z termina em b. Logo, $\mathtt{a}z \in L$ se e só se $\varepsilon z \in L$.
$\delta([\mathtt{b}], \mathtt{b}) \stackrel{\text{def}}{=} [\mathtt{bb}]$	$[\mathtt{bb}] \neq [\varepsilon]$ porque $\mathtt{bba} \notin L$ e $\varepsilon\mathtt{a} \in L$; e $[\mathtt{bb}] \neq [\mathtt{b}]$ porque $\mathtt{b} \notin L$ e $\mathtt{bb} \in L$.
$\delta([\mathtt{b}], \mathtt{a}) \stackrel{\text{def}}{=} [\mathtt{ba}] = [\varepsilon]$	$[\mathtt{ba}] = [\varepsilon]$ pois $\mathtt{ba}z \in L$ sse $z \in L$ (como acima para $[\mathtt{a}] = [\varepsilon]$).
$\delta([\mathtt{bb}], \mathtt{b}) \stackrel{\text{def}}{=} [\mathtt{bbb}] = [\mathtt{bb}]$	pois $\mathtt{bb}z \in L$ se e só se z não termina em a; o mesmo se tem para $\mathtt{bbb}z \in L$.
$\delta([\mathtt{bb}], \mathtt{a}) \stackrel{\text{def}}{=} [\mathtt{bba}]$	$[\mathtt{bba}]$ é uma nova classe pois $\mathtt{bba}z \in L$ se e só se z termina em b.
$\delta([\mathtt{bba}], \mathtt{a}) \stackrel{\text{def}}{=} [\mathtt{bbaa}] = [\mathtt{bba}]$	porque $\mathtt{bbaa}z \in L$ se e só se z termina em b.
$\delta([\mathtt{bba}], \mathtt{b}) \stackrel{\text{def}}{=} [\mathtt{bbab}] = [\mathtt{bb}]$	porque $\mathtt{bbab}z \in L$ se e só se z não termina em a.



3. Considere a gramática $G = (\{S, D\}, \{\mathtt{a}, \mathtt{b}\}, P, S)$, em que P é constituído por:

$$S \to \mathtt{ba}S \qquad D \to \mathtt{bb}D$$
$$S \to \mathtt{b}D \qquad D \to \mathtt{a}$$
$$D \to \mathtt{bb}$$

a) Defina *informalmente* a linguagem de $\{\mathtt{a}, \mathtt{b}\}^\star$ gerada a partir da variável D.

Resposta

As palavras dessa linguagem têm número par de b's (possivelmente zero), pelo menos um símbolo e no máximo um a, podendo a ocorrer apenas no fim da palavra.

b) Justifique que a linguagem $\mathcal{L}(G)$ é regular e indique uma expressão regular que a descreva. Explique como deduziu tal expressão. Comece por apresentar a noção de *linguagem gerada por uma gramática*.

Resposta

Uma linguagem é regular se e só se pode ser descrita por uma expressão regular. Neste caso, como a gramática linear à direita, podemos concluir imediatamente que tal expressão existe. De facto, $\mathcal{L}(G) = \mathcal{L}((\mathtt{ba})^\star\mathtt{b}(\mathtt{bb})^\star(\mathtt{bb}+\mathtt{a}))$.

Por definição, a linguagem gerada por G é o conjunto das palavras de $\{\mathtt{a},\mathtt{b}\}^\star$ que se podem derivar a partir de S (símbolo inicial de G). As regras para D não introduzem S na derivação e, qualquer derivação a partir de S que produza uma sequência de terminais obriga à utilização da regra $S \to \mathtt{b}D$ e à utilização de regras de D. Podemos concluir que, para cada $k \geq 1$, as sequências que se podem derivar de S em k passos sem usar regras para D são da forma $S \Rightarrow^k (\mathtt{ba})^k S$ ou $S \Rightarrow^k (\mathtt{ba})^{k-1}\mathtt{b}D$. Por outro lado, qualquer que seja $n \geq 1$, tem-se $D \Rightarrow^n w$ se e só se $w = (\mathtt{bb})^n D$ ou $w = (\mathtt{bb})^{n-1}\mathtt{a}$ ou $w = (\mathtt{bb})^n$. Assim, as sequências de terminais que se derivam de S são as da forma $(\mathtt{ba})^{k-1}\mathtt{b}(\mathtt{bb})^{n-1}\mathtt{a}$ ou $(\mathtt{ba})^{k-1}\mathtt{b}(\mathtt{bb})^n$, com $k \geq 1$ e $n \geq 1$. Logo, $\mathcal{L}(G)$ pode ser descrita pela expressão regular $(\mathtt{ba})^\star\mathtt{b}(\mathtt{bb})^\star(\mathtt{bb}+\mathtt{a})$.

c) Diga, justificando, se a gramática G é ambígua e se a linguagem $\mathcal{L}(G)$ é ambígua.

Resposta

A linguagem $\mathcal{L}(G)$ não é ambígua porque nenhuma linguagem regular é ambígua.

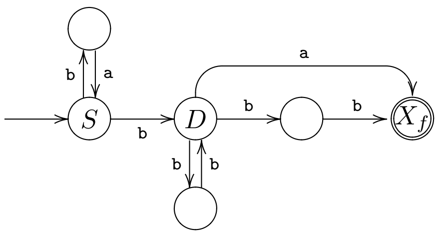
A gramática G não é ambígua porque cada palavra de $\mathcal{L}(G)$ admite uma só derivação (e, portanto, só admite uma derivação pela esquerda). Para obter $(\mathtt{ba})^{k-1}\mathtt{b}(\mathtt{bb})^{n-1}\mathtt{a}$, é necessário aplicar $S \to \mathtt{ba}S$ exatamente $k-1$ vezes, depois $S \to \mathtt{b}D$ uma vez, $D \to \mathtt{bb}D$ exatamente $n-1$ vezes e, finalmente, $D \to \mathtt{a}$. A única derivação para $(\mathtt{ba})^{k-1}\mathtt{b}(\mathtt{bb})^n$ é idêntica, diferindo apenas no último passo, em que se terá de usar $D \to \mathtt{bb}$ em vez de $D \to \mathtt{a}$.

Por definição, uma LIC é ambígua se todas as GICs que a gerem forem ambíguas. Sendo G não ambígua, concluimos que $\mathcal{L}(G)$ não é ambígua. Mas, o resultado invocado acima bastaria.

d) Determine um autómato finito que reconheça $\mathcal{L}(G)$ e não tenha transições-ε. Explique a sua correção.

Resposta

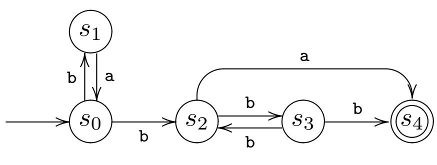
Por aplicação do algoritmo de conversão de uma gramática linear à direita para um AFND-ε, obteriamos o autómato seguinte, o qual não tem transições por ε.



e) Seja L a linguagem das palavras de alfabeto $\{\mathtt{a},\mathtt{b},\mathtt{c}\}$ da forma $x\mathtt{c}^n y$, com $|x| = |y|$, $n \geq 2$, $x \in \mathcal{L}(G)$ e $y \in \mathcal{L}(G)$. Prove que L é independente de contexto.
Note que x, y e n são quaisquer (não estão fixos).

Resposta

Para provar que L é independente de contexto basta definir uma gramática independente de contexto que a gere ou definir um autómato de pilha que a reconheça. Como já se tem um autómato finito que reconhece $\mathcal{L}(G)$, será mais fácil definir o autómato de pilha para L. Para diminuir o número de estados, partimos do AFND representado abaixo à esquerda (equivalente ao AFND indicado em 3d)). O autómato de pilha coloca um X na pilha por cada símbolo de x e, depois de processar os c's, retira um X por cada símbolo de y. Para garantir que x e y pertencem a $\mathcal{L}(G)$, imita a alteração de estados do AFND. A aceitação é por pilha vazia, s_0 é o estado inicial, Z é o símbolo inicial na pilha, e δ está representada à direita.



$\delta(s_0,\mathtt{b},\mathtt{Z}) = \{(s_2,\mathtt{XZ}),(s_1,\mathtt{XZ})\}$
$\delta(s_1,\mathtt{a},\mathtt{X}) = \{(s_0,\mathtt{XX})\}$
$\delta(s_0,\mathtt{b},\mathtt{X}) = \{(s_2,\mathtt{XX}),(s_1,\mathtt{XX})\}$
$\delta(s_2,\mathtt{b},\mathtt{X}) = \{(s_3,\mathtt{XX})\}$
$\delta(s_2,\mathtt{a},\mathtt{X}) = \{(s_4,\mathtt{XX})\}$
$\delta(s_3,\mathtt{b},\mathtt{X}) = \{(s_4,\mathtt{XX}),(s_2,\mathtt{XX})\}$
$\delta(s_4,\mathtt{c},\mathtt{X}) = \{(r_1,\mathtt{X})\}$
$\delta(r_1,\mathtt{c},\mathtt{X}) = \{(r_1,\mathtt{X}),(q_0,\mathtt{X})\}$

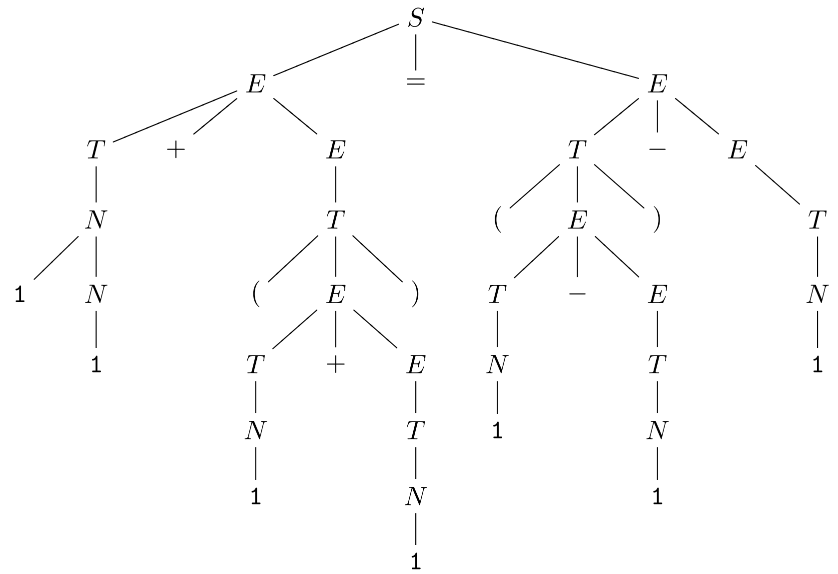
$\delta(q_0,\mathtt{b},\mathtt{X}) = \{(q_2,\varepsilon),(q_1,\varepsilon)\}$
$\delta(q_1,\mathtt{a},\mathtt{X}) = \{(q_0,\varepsilon)\}$
$\delta(q_0,\mathtt{b},\mathtt{X}) = \{(q_2,\varepsilon),(q_1,\varepsilon)\}$
$\delta(q_2,\mathtt{b},\mathtt{X}) = \{(q_3,\varepsilon)\}$
$\delta(q_2,\mathtt{a},\mathtt{X}) = \{(q_4,\varepsilon)\}$
$\delta(q_3,\mathtt{b},\mathtt{X}) = \{(q_4,\varepsilon),(q_2,\varepsilon)\}$
$\delta(q_4,\varepsilon,\mathtt{Z}) = \{(q_4,\varepsilon)\}$

4. Considere a gramática $G = (\{S,T,N,E\},\Sigma,P,S)$ com $\Sigma = \{\,\mathtt{1},\mathtt{-},\mathtt{)}\,,\,\mathtt{(},\mathtt{=},\mathtt{+}\,\}$ e P dado por:

$$\begin{aligned}
&S \to E\mathtt{=}E \;\mid\; E\mathtt{=}S\\
&E \to T \;\mid\; T\mathtt{+}E \;\mid\; T\mathtt{-}E\\
&T \to \mathtt{(}E\mathtt{)} \;\mid\; N\\
&N \to \mathtt{1} \;\mid\; \mathtt{1}N
\end{aligned}$$

a) Apresente uma árvore de derivação para a palavra $\mathtt{11+(1+1)=(1-1)-1}$ de $\mathcal{L}(G)$ e ainda duas derivações distintas que correspondam a essa árvore. Diga, justificando, se tal implica que G seja ambígua.

Resposta



Duas derivações distintas que determinam essa árvore:

$$\begin{aligned}
S \;\Rightarrow_G\;& E\mathtt{=}E \;\Rightarrow_G\; T\mathtt{+}E\mathtt{=}E \;\Rightarrow_G\; N\mathtt{+}E\mathtt{=}E \;\Rightarrow_G\; \mathtt{1}N\mathtt{+}E\mathtt{=}E \;\Rightarrow_G\; \mathtt{11+}E\mathtt{=}E \;\Rightarrow_G\; \mathtt{11+}T\mathtt{=}E \;\Rightarrow_G\\
& \mathtt{11+(}E\mathtt{)=}E \;\Rightarrow_G\; \mathtt{11+(}T\mathtt{+}E\mathtt{)=}E \;\Rightarrow_G\; \mathtt{11+(}N\mathtt{+}E\mathtt{)=}E \;\Rightarrow_G\; \mathtt{11+(1+}E\mathtt{)=}E \;\Rightarrow_G\\
& \mathtt{11+(1+}T\mathtt{)=}E \;\Rightarrow_G\; \mathtt{11+(1+}N\mathtt{)=}E \;\Rightarrow_G\; \mathtt{11+(1+1)=}E \;\Rightarrow_G\\
& \mathtt{11+(1+1)=}T\mathtt{-}E \;\Rightarrow_G\\
& \mathtt{11+(1+1)=(}E\mathtt{)-}E \;\Rightarrow_G\; \mathtt{11+(1+1)=(}T\mathtt{-}E\mathtt{)-}E \;\Rightarrow_G\; \mathtt{11+(1+1)=(}N\mathtt{-}E\mathtt{)-}E \;\Rightarrow_G\\
& \mathtt{11+(1+1)=(1-}E\mathtt{)-}E \;\Rightarrow_G\; \mathtt{11+(1+1)=(1-}T\mathtt{)-}E \;\Rightarrow_G\; \mathtt{11+(1+1)=(1-}N\mathtt{)-}E \;\Rightarrow_G\\
& \mathtt{11+(1+1)=(1-1)-}E \;\Rightarrow_G\; \mathtt{11+(1+1)=(1-1)-}T \;\Rightarrow_G\; \mathtt{11+(1+1)=(1-1)-}N \;\Rightarrow_G\\
& \mathtt{11+(1+1)=(1-1)-1}
\end{aligned}$$

$$S \;\Rightarrow_G\; E\mathtt{=}E \;\Rightarrow_G\; E\mathtt{=}T\mathtt{-}E \;\Rightarrow_G^{\star}\; \mathtt{11+(1+1)=}T\mathtt{-}E \;\Rightarrow_G^{\star}\; \mathtt{11+(1+1)=(1-1)-1}$$

Na segunda derivação, a partir de $E\mathtt{=}T\mathtt{-}E$, segue-se a sequência de regras da primeira, substituindo-se sempre a variável mais à esquerda, até obter $\mathtt{11+(1+1)=}T\mathtt{-}E$. A partir dessa forma, a derivação é exatamente igual à primeira. As duas derivações determinam a mesma árvore.

Uma gramática G é ambígua se alguma palavra de $\mathcal{L}(G)$ admitir duas árvores de derivação. Portanto, nada se pode concluir sobre a ambiguidade da gramática, pois as duas derivações pedidas no enunciado teriam de corresponder à mesma árvore. Continuariamos sem saber se existiam ou não outras árvores para essa palavra (ou se existia alguma outra palavra com mais do que uma árvore de derivação).

Se as derivações determinam a mesma árvore de derivação então, por definição de árvore de derivação, terão de ser derivações da mesma palavra. Existem muitas *derivações* que conduzem à mesma árvore (porque, em passo da derivação, podemos substituir qualquer uma das variáveis da forma obtida no passo anterior). Se substituirmos sempre a variável mais à esquerda (como na primeira derivação apresentada), a derivação diz-se *derivação pela esquerda*. Se substituirmos sempre a variável mais à direita, diz-se *derivação pela direita*. É conhecido que cada árvore de derivação determina uma só derivação pela esquerda, e vice-versa. Analogamente, cada árvore de derivação determina uma só derivação pela direita, e vice-versa.

A gramática dada não é ambígua, mas a demonstração desse resultado requer uma análise bem mais cuidada e o nível de dificuldade desse exercício não seria comparável!

b) Prove que a linguagem $\mathcal{L}(G)$ não é regular.

Resposta

A linguagem não satisfaz a condição do Lema da Repetição para linguagens regulares. Dado $n \geq 1$, tomamos a palavra $(^n\mathtt{1})^n\mathtt{=}\ \mathtt{1}$. Esta palavra pertence a $\mathcal{L}(G)$ porque se pode reescrever n vezes E em T e T em (E) para introduzir os parentesis.

Em qualquer decomposição uvw dessa palavra, tal que $|uv| \leq n$ e $v \neq \varepsilon$, a subpalavra v é uma sequência de parentesis abertos. Assim, se se tomar $i = 0$ (cortar v), a palavra uv^iw terá mais parentesis fechados do que abertos e, consequentemente, não é palavra de $\mathcal{L}(G)$. De facto, G tem apenas uma regra que pode introduzir parentesis, nomeadamente $T \rightarrow \mathtt{(}E\mathtt{)}$. Como esta regra introduz sempre um parentesis aberto e um fechado, o número de parentesis abertos é igual ao número de parentesis fechados em qualquer palavra que se possa derivar a partir de S (podendo tal número ser zero).

Resposta (alternativa)

Pelo teorema de Myhill-Nerode, a linguagem $\mathcal{L}(G)$ seria regular se e só se o conjunto das classes de equivalência de $R_{\mathcal{L}(G)}$ fosse finito.

Mas, se $n \neq m$ então $(^n$ e $(^m$ não são equivalentes segundo $R_{\mathcal{L}(G)}$, porque se z for a palavra "$\mathtt{1})^n\mathtt{=}\ \mathtt{1}$" (sem aspas), a palavra $(^n z$ pertence a $\mathcal{L}(G)$ mas $(^m z$ não pertence a $\mathcal{L}(G)$. Portanto, cada palavra $(^n$ determina uma classe de equivalência distinta, para $n \geq 1$. Logo, o conjunto das classes de equivalência de $R_{\mathcal{L}(G)}$ é infinito. Consequentemente, $\mathcal{L}(G)$ não é regular.

Resolva apenas uma das três alíneas seguintes. Se resolver mais do que uma, só a primeira é avaliada.

c) Determine uma gramática G' que seja equivalente a G e esteja na forma normal de Chomsky. Por aplicação do algoritmo CYK, mostre que $\mathtt{1{=}1{+}1}$ pertence a $\mathcal{L}(G')$ e que $\mathtt{1{=}1{+}}$ não pertence a $\mathcal{L}(G')$. Note que a segunda palavra é igual à primeira mas não tem o $\mathtt{1}$ final.

d) Por aplicação da forma normal de Greibach, determine um autómato de pilha que reconheça $\{x \mid x \in \Sigma^\star \text{ e } E \Rightarrow_G^\star x\}$. Note que não se pretende um autómato que reconheça $\mathcal{L}(G)$.

e) Seja M a linguagem das palavras de $\mathcal{L}(G)$ que são da forma $x\mathtt{=}y\mathtt{=}z$, com $x \in \{\mathtt{1}, \mathtt{+}\}^\star$, $y \in \{\mathtt{1}\}^\star$, $z \in \{\mathtt{1}, \mathtt{+}\}^\star$, e em que o número de $\mathtt{1}$'s em y é igual ao número de $\mathtt{1}$'s em x e também em z.
Defina uma máquina de Turing que reconheça M. A máquina não deve repor o estado inicial da fita. Indique o significado dos estados.

Resposta

(Resposta omitida. Exercício de valorização.)

(Fim)