Departamento de Ciência de Computadores — FCUP

Modelos de Computação CC1004 — 2015/2016 ARF.

Exame – 04.07.2016 — *duração: 3h*

N.º [ ] Nome [ ]

1. Seja L a linguagem de alfabeto $\Sigma = \{a, b\}$ constituída pelas palavras que têm ab como subpalavra e número ímpar de a's antes do b mais à esquerda na palavra.

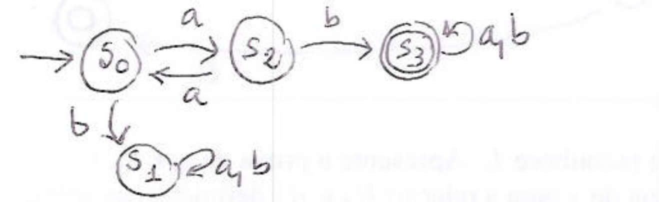
a) Apresente as regras de uma GIC G que gere L, não seja linear à direita nem à esquerda e tenha símbolo inicial K. Explique sucintamente, partindo da descrição de L.

$K \to B\,ab\,R$
$B \to \varepsilon \mid aaB$
$R \to \varepsilon \mid aR \mid bR$

As palavras de L saõ da forma $x\,ab\,y$, com $x \in \{aa\}^*$ e $y \in \{a,b\}^*$.
A variável B gera $\{aa\}^*$
A variável R gera $\{a,b\}^*$
A regra para K traduz a forma descrita acima.
A GIC não é linear à direita nem à esquerda porque a regra para K não é da forma $X \to wY$ nem $X \to w$ (nem $X \to Yw$ nem $X \to w$), com X, Y vars e $w \in \Sigma^*$.

b) Indique uma expressão regular (abreviada) que descreva L. $(aa)^*ab(a+b)^*$

c) Desenhe o diagrama do AFD mínimo que reconhece L e descreva $\mathcal{L}_s = \{x \mid x \in \Sigma^\star \text{ e } \hat{\delta}(s_0, x) = s\}$ por uma expressão regular (abreviada), para cada estado s, sendo s_0 o estado inicial.



$\mathcal{L}_{s_0}$: $(aa)^*$

$\mathcal{L}_{s_1}$: $(aa)^*b(a+b)^*$

$\mathcal{L}_{s_2}$: $(aa)^*a$

$\mathcal{L}_{s_3}$: $(aa)^*ab(a+b)^*$

d) Usando o corolário do Teorema de Myhill-Nerode, prove a correção do AFD que apresentou em **1c**).

Os estados s_0, s_1, s_2 e s_3 correspondem às classes $[\varepsilon]$, $[b]$, $[a]$ e $[ab]$ de R_L.

$\varepsilon \not R_L a$ porque para $z=b$, $\varepsilon z \notin L$ e $az \in L$.
$\varepsilon \not R_L b$ porque para $z=ab$, $\varepsilon z \in L$ e $bz \notin L$.
$b \not R_L a$ porque para $z=b$, $az \in L$ e $bz \notin L$
$ab \not R_L x$ para $x \in \{\varepsilon, a, b\}$ porque $ab \in L$ e $x \notin L$.
$ab\, R_L\, abb$ } porque $\forall z \in \Sigma^*$ $abz \in L$ e $abbz \in L$ e $abaz \in L$
$ab\, R_L\, aba$ }
$b\, R_L\, ba$ e $b\, R_L\, bb$ porque $\forall z \in \Sigma^*$ $bz \notin L$ e $baz \notin L$ e $bbz \notin L$
$aa\, R_L\, \varepsilon$ porque $aaz \in L$ sse $z \in L$.

(Continua)

A.27.

N.º | Nome

2. Sejam $r = (((\mathtt{aa})^\star) + ((\mathtt{bb})^\star))$ e $s = (((\mathtt{aa}) + (\mathtt{bb}))^\star)$ expressões regulares sobre $\Sigma = \{\mathtt{a}, \mathtt{b}\}$.

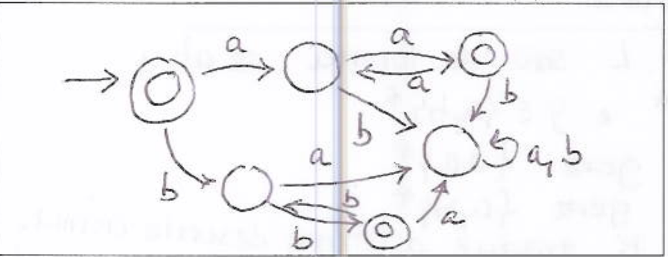
a) Apresente uma GIC não ambígua gere $\mathcal{L}(r)$.

$S \to aaA \mid bbB \mid \varepsilon$
$A \to \varepsilon \mid aaA$
$B \to \varepsilon \mid bbB$
símbolo inicial S

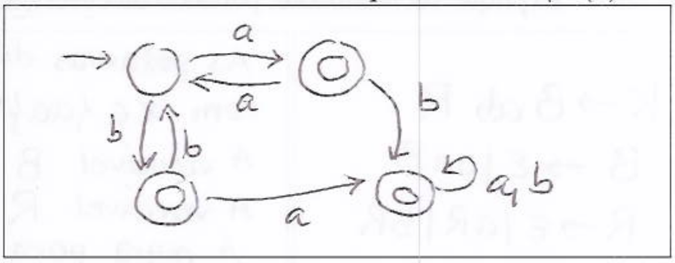
b) Apresente uma GIC não ambígua gere $\mathcal{L}(s)$.

$S \to \varepsilon \mid aa\,S \mid bb\,S$
símbolo inicial S

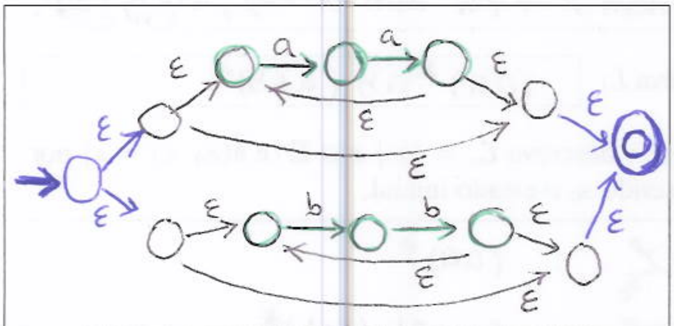
c) Desenhe o AFD mínimo que aceita $\mathcal{L}(r)$.

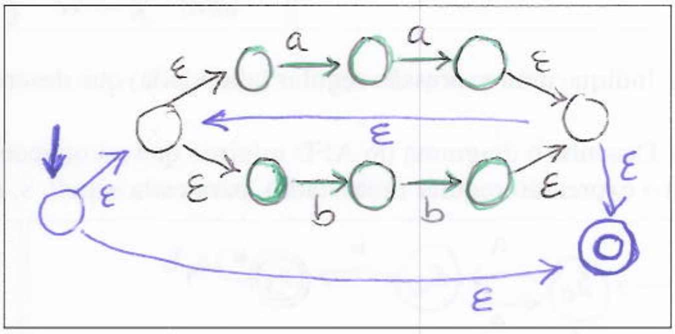


d) Desenhe o AFD mínimo que aceita $\Sigma^\star \setminus \mathcal{L}(s)$.



e) Desenhe os diagrama de transição dos autómatos finitos que resultam da aplicação do método de Thompson às expressões regulares r e s, segundo a construção dada nas aulas.





3. Seja L uma linguagem regular e seja A um AFD que reconhece L. Apresente a prova de que $C_x \subseteq [x]$, para todo $x \in \Sigma^\star$, sendo C_x e $[x]$ as classes de equivalência de x para a relação R_A e R_L definidas nas aulas.

Seja $A = (S, \Sigma, \delta, s_0, F)$. Então, $\mathcal{C}_x = \{y \mid y \in \Sigma^*, y R_A x\} =$
$= \{y \mid y \in \Sigma^*, \hat{\delta}(s_0, x) = \hat{\delta}(s_0, y)\}$ por definição de R_A.
Por definição de R_L tem-se $[x] = \{y \mid y R_L x\} = \{y \mid \forall z \in \Sigma^*\, xz \in L \Leftrightarrow yz \in L\}$.
Podemos concluir que se $y \in \mathcal{C}_x$ então $y \in [x]$ porque
$\hat{\delta}(s_0, xz) = \hat{\delta}(\hat{\delta}(s_0, x), z) \underset{y \in \mathcal{C}_x}{=} \hat{\delta}(\hat{\delta}(s_0, y), z) = \hat{\delta}(s_0, yz)$
para todo $z \in \Sigma^*$.
Logo, $\hat{\delta}(s_0, xz) \in F$ sse $\hat{\delta}(s_0, yz) \in F$, o que quer dizer que $xz \in L$ sse $yz \in L$. cqd.

(Continua)

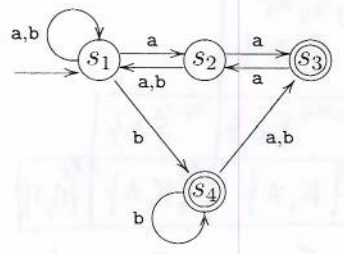
APF.

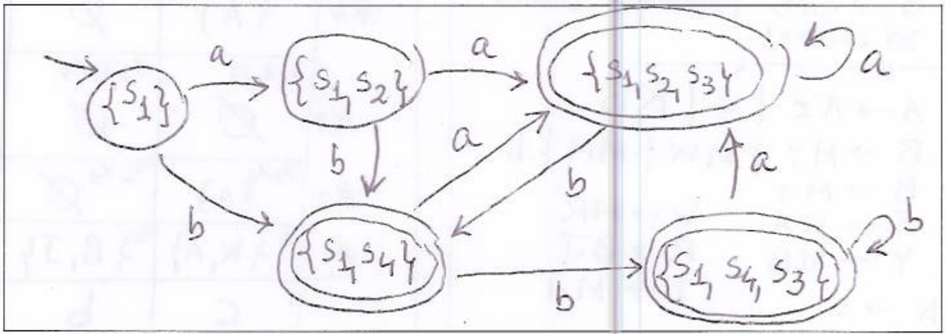
N.º [ ] Nome [ ]

Resolva apenas um dos problemas 4. e 5.

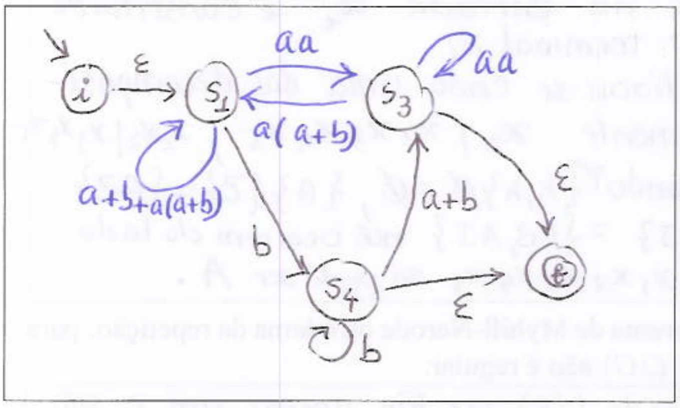
apenas um!

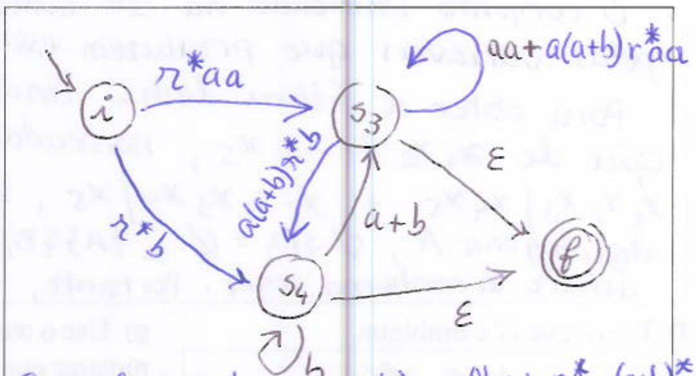
4. Desenhe o diagrama de transição do AFD equivalente ao AFND representado à esquerda que resulta da aplicação do método de conversão (baseado em subconjuntos). Os estados devem ser **obrigatoriamente** designados por subconjuntos. Crie apenas os que são acessíveis do estado inicial.





5. Considere novamente o AFND representado em **4.**. Suponha que se aplica o método de eliminação de estados e que na fase de eliminação se começa por remover s_2 e a seguir s_1. Apresente o diagrama após a remoção de s_2 e de s_1 (não simplifique as expressões intermédias).





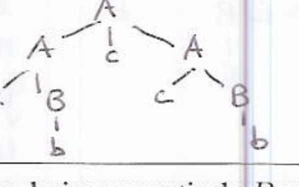
Com $r = a+b+a(a+b)$. Obs: $r^* \equiv (a+b)^*$

6. Considere a GIC $G = (\{A, B\}, \{\mathrm{a}, \mathrm{b}, \mathrm{c}\}, P, A)$, com P dado por:

$$A \rightarrow AcA \mid cB \qquad\qquad B \rightarrow aaB \mid aaBb \mid aa \mid b \mid \varepsilon$$

a) Prove que cbccb $\in \mathcal{L}(G)$, indicando uma derivação e a árvore de derivação correspondente, e complete a frase "cbccb admite [uma] ~~árvores~~ árvore de derivação e [seis] derivações distintas".

$A \Rightarrow AcA \Rightarrow cBcA \Rightarrow cbcA$
$\Rightarrow cbccB \Rightarrow cbccb$



b) Indique a forma das palavras de $\{A, B, \mathrm{a}, \mathrm{b}\}^\star$ que se podem derivar a partir de B em G, numa derivação com n passos, para $n \geq 1$ qualquer, se a regra $B \rightarrow \mathrm{a}\mathrm{a}B$ for aplicada k vezes, com $0 \leq k \leq n$. Explique.

$B \Rightarrow^n x$ sse $x = (aa)^n B$ ou $x = (aa)^n B\, b^{n-k}$ ou $x = (aa)^{n-1} y\, b^{n-k-1}$
com $y = aa$ ou $y = b$ ou $y = \varepsilon$

Caso $x = (aa)^n B$, a regra $B \to aaB$ foi aplicada n vezes.
caso $x = (aa)^n B\, b^{n-k}$ apenas se aplicou $B \to aaB$ e $B \to aaBb$, por qualquer ordem
Caso restante, aplica-se $n-1$ vezes essas regras e uma de $B \to aa$, $B \to b$, $B \to \varepsilon$

(Continua)

ADS.

N.º [ ] Nome [ ]

c) Indique uma GIC G' na forma normal de Chomsky tal que $\mathcal{L}(G) = \mathcal{L}(G')$.

Começa-se por eliminar a regra $B \to \varepsilon$ introduzindo a regra $A \to c$ e a regra $B \to aab$ (dado que $B \to aa$ já existia).

$A \to AZ \mid c \mid KB$
$B \to MY \mid MW \mid MM \mid b$
$B \to MT$
$Z \to KA$
$Y \to MB$
$W \to MR$
$R \to BJ$
$T \to MJ$
$K \to c$
$M \to a$
$J \to b$

d) Prove que cbccb $\in \mathcal{L}(G')$, aplicando o algoritmo CYK.

#5	$x_1x_2x_3x_4x_5$ $\{A\}$				
#4	$x_1x_2x_3x_4$ $\{A\}$	$x_2x_3x_4x_5$ $\emptyset$			
#3	$x_1x_2x_3$ $\emptyset$	$x_2x_3x_4$ $\emptyset$	$x_3x_4x_5$ $\{Z\}$		
#2	x_1x_2 $\{A\}$	x_2x_3 $\emptyset$	x_3x_4 $\{Z\}$	x_4x_5 $\{A\}$	
#1	x_1 $\{K,A\}$	x_2 $\{B,J\}$	x_3 $\{K,A\}$	x_4 $\{K,A\}$	x_5 $\{B,J\}$
	c	b	c	c	b
	x_1	x_2	x_3	x_4	x_5

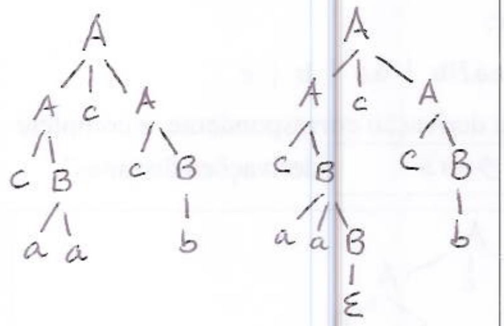
Como $A \in$ Categoria para cbccb, então cbccb $\in \mathcal{L}(G)$.

e) Explique em detalhe como se obtém a *primeira* e a *última* linha da tabela que apresentou em **6d)**.

O conjunto indicado na 1ª linha na entrada x_i é constituído pelas variáveis que produzem esse terminal x_i.
Para obter a última linha, analisou-se cada uma das decomposições de $x_1x_2x_3x_4x_5$, nomeadamente $x_1 \mid x_2x_3x_4x_5$, $x_1x_2 \mid x_3x_4x_5$, $x_1x_2x_3 \mid x_4x_5$ e $x_1x_2x_3x_4 \mid x_5$, sendo que $\{K,A\}\emptyset = \emptyset$, $\{A\}\{Z\} = \{AZ\}$ da categoria A, $\emptyset\{A\} = \emptyset$ e $\{A\}\{B,J\} = \{AB, AJ\}$ não ocorrem do lado direito de nenhuma regra. Portanto, $x_1x_2x_3x_4x_5$ só pode ser A.

f) Prove que G é ambígua.

caacc b $\in \mathcal{L}(G)$ admite duas árvores distintas



g) Use o teorema de Myhill-Nerode ou o lema da repetição. para mostrar que $\mathcal{L}(G)$ não é regular.

As palavras de $L(G)$ que têm apenas um c são da forma $c(aa)^{n-1}y\,b^{n-k-1}$, com $y = aa$ ou $y = b$ ou $y = \varepsilon$, $0 \leq k \leq n-1$.
Vamos ver que $L(G)$ não satisfaz a condição do lema da Repetição para Ling. regulares para nenhum $n > 0$. Dado $n > 0$, tomamos $z = ca^{2n}b^{n+1}$ (caso em que $y = b$, sendo o número de b's máximo). Então, quaisquer que sejam $u, v, w \in \Sigma^*$, com $uvw = z$, $|uv| \leq n$ e $v \neq \varepsilon$ tem-se $uv^0w \notin L(G)$ pois corta o c ou corta a's e fica com excesso de b's. cqd.

Resolva apenas uma das alíneas seguintes

h) Prove que a linguagem $\mathcal{L}(G)$ não é ambígua. Justifique sucintamente a correção da resposta.

i) Apresente um autómato de pilha que reconheça $\mathcal{L}(G)$ por pilha vazia. Justifique sucintamente a correção.

Use o verso da folha para responder à questão.

(Fim)