Departamento de Ciência de Computadores FCUP

Modelos de Computação (CC1004) 2014/15

1º Teste – 11.04.2015 **Uma resolução (v1)** *duração: 2h + 30m*

1. Seja $\Sigma = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$ e seja r a expressão $((\mathtt{a}((\mathtt{b}+\mathtt{c})^\star))\mathtt{a})$.

a) Baseando-se na definição de expressão regular, mostre que r é uma expressão regular sobre Σ.

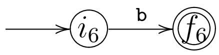
Resposta:
Uma expressão regular sobre Σ é qualquer sequência finita de símbolos de $\Sigma \cup \{\emptyset, \varepsilon, +, *,), (\}$ que se possa obter por aplicação das regras seguintes: ε, $\emptyset$, $\mathtt{a}$, $\mathtt{b}$ e $\mathtt{c}$ são expressões regulares sobre Σ; se r e s são expressões regulares sobre Σ então $(r+s)$, (rs) e $(r^\star)$ são expressões regulares sobre Σ.

Assim, r é uma expressão regular sobre Σ porque $r = (r_1 r_2)$, com $r_1 = (\mathtt{a}((\mathtt{b}+\mathtt{c})^\star))$ e $r_2 = \mathtt{a}$, sendo $r_1 = (r_3 r_4)$, com $r_3 = \mathtt{a}$ e $r_4 = ((\mathtt{b}+\mathtt{c})^\star)$, e $r_4 = (r_5^\star)$, para $r_5 = (\mathtt{b}+\mathtt{c})$, e $r_5 = (r_6 + r_7)$, com $r_6 = \mathtt{b}$ e $r_7 = \mathtt{c}$.

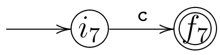
b) Determine o autómato finito que resulta da aplicação do método de Thompson à expressão regular r. Apresente **os passos relevantes** dessa construção.

Resposta:
Na continuação da resposta anterior, os AFNDs-ε para r_6, r_7, r_3 e r_2 são:

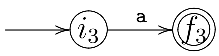
$A(r_6)$:



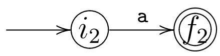
$A(r_7)$:



$A(r_3)$:

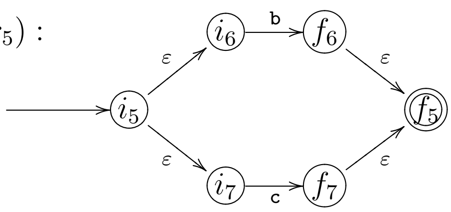


$A(r_2)$:

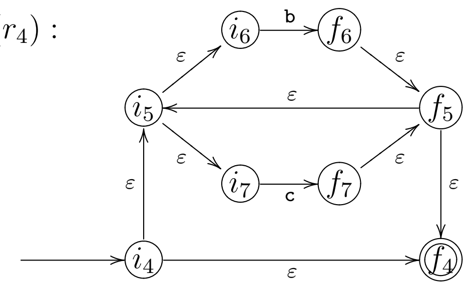


Os AFNDs-ε para $r_5 = (r_6 + r_7)$ e $r_4 = (r_5^\star)$ são:

$A(r_5)$:

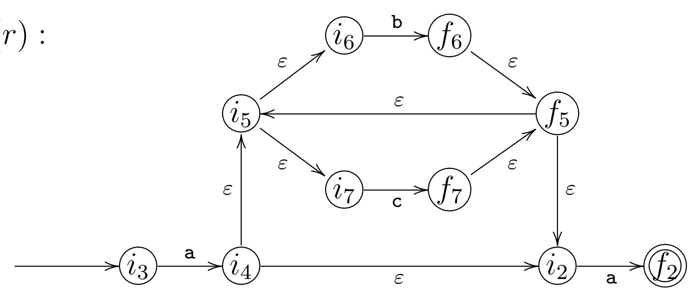


$A(r_4)$:



Os autómatos para $r_1 = (r_3 r_4)$ e $r = (r_1 r_2)$ obtém-se de $A(r_3)$ e $A(r_4)$ e, depois, de $A(r_1)$ e $A(r_2)$, por identificação do estado inicial de $A(r_4)$ com o final de $A(r_3)$, e do estado final de $A(r_1)$ (que é o final de $A(r_4)$) com o inicial de $A(r_2)$.

$A(r)$:



c) Apresente a expressão r na forma *abreviada*, retirando parentesis desnecessários, e descreva informalmente a linguagem de $\Sigma^\star$ que é caraterizada pela expressão regular r.

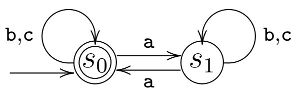
Resposta:
A expressão abreviada é $\mathtt{a}(\mathtt{b}+\mathtt{c})^\star\mathtt{a}$.

A linguagem descrita pela expressão r é o conjunto das palavras de $\{\mathtt{a},\mathtt{b},\mathtt{c}\}^\star$ que têm exatamente dois a's, e começam e terminam em a.

d) Descreva informalmente a linguagem descrita pela expressão regular $((r+(\mathtt{b}+\mathtt{c}))^\star)$. Partindo dessa descrição, determine um AFD que reconheça tal linguagem. Justifique sucintamente a correção da resposta, descrevendo o que memoriza cada estado (e explicando a necessidade das mudanças de estado).

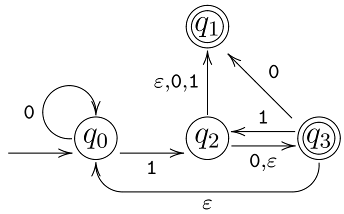
Resposta:
A linguagem descrita pela expressão indicada é o conjunto das palavras de $\{\mathtt{a},\mathtt{b},\mathtt{c}\}^\star$ que têm número par de a's.

A linguagem é reconhecida pelo AFD seguinte:



As palavras que levam este AFD do estado inicial s_0 ao estado s_1 são as que têm número ímpar de a's. As palavras que têm número par de a's levam o autómato do estado s_0 ao estado s_0. Os dois estados são necessários porque apenas as palavras que têm número par de a's pertencem à linguagem.

2. Seja $A = (S, \Sigma, \delta, q_0, F)$ o autómato finito não determinístico com transições por ε representado pelo diagrama seguinte, com alfabeto $\Sigma = \{0, 1\}$.



a) Qual é o valor de $\delta(q_2, 0)$, $\delta(q_0, \varepsilon)$, $\delta(q_3, 0)$, e $\delta(q_1, \varepsilon)$? Justifique sucintamente.

Resposta:
A função de transição δ do AFND-ε $A = (S, \Sigma, \delta, q_0, F)$ é uma função de $S \times (\Sigma \cup \{\varepsilon\})$ em 2^S. O diagrama de transição tem um arco de s para s' com etiqueta α se e só se $s' \in \delta(s, \alpha)$. Se houver várias transições de um estado para outro (ou para si próprio), essas transições representam-se por um único arco, com os símbolos correspondentes separados por vírgulas.

Assim, $\delta(q_2, 0) = \{q_1, q_3\}$, $\delta(q_0, \varepsilon) = \emptyset$, $\delta(q_3, 0) = \{q_1\}$, e $\delta(q_1, \varepsilon) = \emptyset$.

b) Determine $\hat{\delta}(\{q_0\}, 100)$. Apresente os cálculos intermédios.

Resposta:
$\hat{\hat{\delta}}(\{q_0\}, 100) = \hat{\hat{\delta}}(Fecho_\varepsilon(\{q_2\}), 00) = \hat{\hat{\delta}}(\{q_2, q_3, q_1, q_0\}, 00) = \hat{\hat{\delta}}(Fecho_\varepsilon(\{q_3, q_1, q_0\}), 0) = \hat{\hat{\delta}}(\{q_3, q_1, q_0\}, 0) = \hat{\hat{\delta}}(Fecho_\varepsilon(\{q_1, q_0\}), \varepsilon) = Fecho_\varepsilon(\{q_1, q_0\}) = \{q_1, q_0\}$.

(a definição de $\hat{\hat{\delta}}$ que estamos a seguir está em **2e)**)

c) Que interpretação tem $\hat{\delta}(\{q_0\}, \texttt{100})$? É verdade ou é falso que $\texttt{100} \in \mathcal{L}(A)$? Justifique.

Resposta:

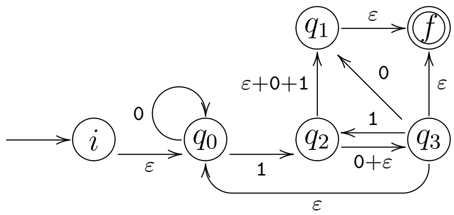
$\hat{\delta}(\{q_0\}, \texttt{100})$ representa o conjunto de estados em que o autómato se pode encontrar se consumir a palavra $\texttt{100}$ partindo do estado q_0 (estado inicial do autómato).

A palavra $\texttt{100}$ pertence a $\mathcal{L}(A)$ porque $\hat{\delta}(\{q_0\}, \texttt{100}) = \{q_0, q_1\}$ contém um estado final (o estado q_1).

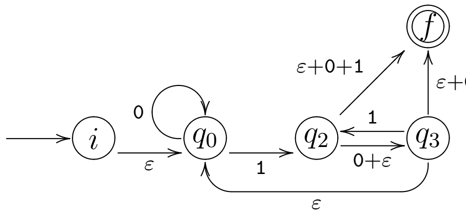
d) Por aplicação do método de eliminação de estados, determine uma expressão regular que descreva a linguagem que A reconhece. Deverá apresentar os passos intermédios da aplicação do algoritmo. Pode apresentar expressões abreviadas, usando as propriedades e precedência das operações para retirar parentesis desnecessários. Sempre que for óbvio, simplifique as expressões obtidas em cada passo.

Resposta:

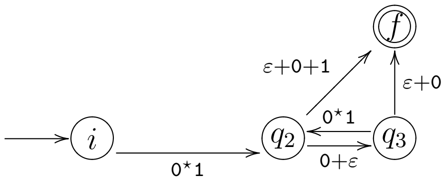
Todos os estados do autómato são acessíveis do estado inicial e permitem aceder a algum estado final. Assim, não há estados que se possam eliminar por serem inutéis.
Começamos por substituir os estados finais por um único estado final f (do qual não saem transições). Inserimos um novo estado inicial i, para garantir que não chegam transições ao estado inicial. Substituimos as etiquetas dos ramos por expressões regulares.



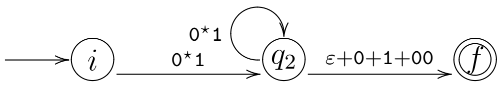
Eliminamos q_1, substituindo os percursos q_3q_1f e q_2q_1f por arcos (q_3, f) e (q_2, f), com etiquetas $\texttt{0}\varepsilon$ e $(\varepsilon + \texttt{0} + \texttt{1})\varepsilon$, respetivamente (ou seja, $\texttt{0}$ e $\varepsilon + \texttt{0} + \texttt{1}$). Como já existia um arco de q_3 para f, substituimos a sua expressão regular por $\varepsilon + \texttt{0}$.



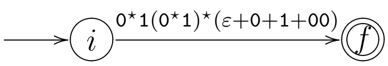
Eliminamos q_0, substituindo os percursos $q_3q_0q_0^\star q_2$ e $iq_0q_0^\star q_2$ por arcos (q_3, q_2) e (i, q_2) com etiquetas $\varepsilon\texttt{0}^\star\texttt{1}$, ou seja, $\texttt{0}^\star\texttt{1}$. A expressão de (q_3, q_2) seria substituída por $\texttt{1} + \texttt{0}^\star\texttt{1}$, que é equivalente a $\texttt{0}^\star\texttt{1}$.



Eliminamos q_3, substituindo $q_2q_3q_2$ e q_2q_3f por arcos (q_2, q_2) e (q_2, f) com etiquetas $(\texttt{0} + \varepsilon)\texttt{0}^\star\texttt{1}$ e $(\texttt{0} + \varepsilon)(\varepsilon + \texttt{0})$. A expressão de (q_2, f) seria substituída por $\varepsilon + \texttt{0} + \texttt{1} + (\texttt{0} + \varepsilon)(\varepsilon + \texttt{0})$ que é equivalente a $\varepsilon + \texttt{0} + \texttt{1} + \texttt{00}$. A expressão $(\texttt{0} + \varepsilon)\texttt{0}^\star\texttt{1}$ também se pode simplificar como $\texttt{0}^\star\texttt{1}$.



Eliminando q_2, por substituição de percursos $iq_2q_2^\star f$ por um arco (i, f), obtém-se



concluindo-se que $\mathcal{L}(A)$ é descrita pela expressão regular (abreviada):

$$\texttt{0}^\star\texttt{1}(\texttt{0}^\star\texttt{1})^\star(\varepsilon + \texttt{0} + \texttt{1} + \texttt{00})$$

e) Por aplicação do método de conversão descrito nas aulas para obter um AFD equivalente a um dado AFND-ε, determine o diagrama de transição de um AFD equivalente ao autómato A. Explique.

Resposta:

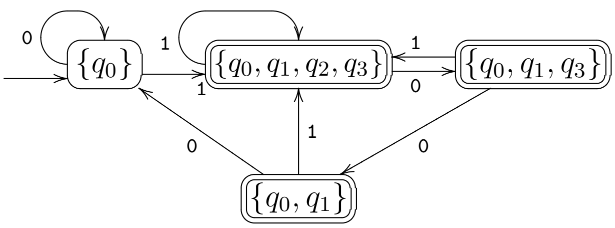
O AFD $A' = (2^S, \Sigma, \delta', Fecho_\varepsilon(q_0), F')$, com $F' = \{E \mid E \in 2^S \text{ e } E \cap F \neq \emptyset\}$ e

$$\delta'(E, a) = Fecho_\varepsilon \left(\bigcup_{s \in Fecho_\varepsilon(E)} \delta(s, a) \right) = \hat{\delta}(E, a)$$

para todo $E \in 2^S$ e $a \in \Sigma$, é equivalente ao AFND-ε $A = (S, \Sigma, \delta, q_0, F)$.

Por definição, $Fecho_\varepsilon(s) = \{s\} \cup \{s' \mid \text{existe um percurso de } s \text{ para } s' \text{ formado por transições-}\varepsilon\}$ e $Fecho_\varepsilon(E) = \bigcup_{s \in E} Fecho_\varepsilon(s)$.

Como o número de estados deste AFD genérico é exponencial no número de estados do AFND-ε dado, vamos tentar obter um AFD com menos estados, criando apenas os estados que são acessíveis do seu estado inicial $Fecho_\varepsilon(q_0) = \{q_0\}$. Obtém-se o seguinte AFD.



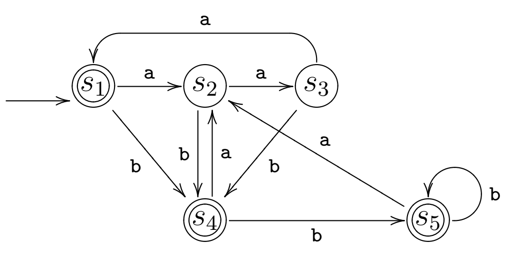
3. Seja L a linguagem de alfabeto $\Sigma = \{\mathtt{a}, \mathtt{b}\}$ que é aceite pelo AFD $A = (\{s_1, s_2, s_3, s_4, s_5\}, \Sigma, \delta, s_1, F)$, com $F = \{s_1, s_4, s_5\}$ e δ dada pela tabela representada à esquerda.

	a	b
s_1	s_2	s_4
s_2	s_3	s_4
s_3	s_1	s_4
s_4	s_2	s_5
s_5	s_2	s_5

a) Desenhe o diagrama de transição de A e descreva informalmente L.
b) Diga, justificando, se o AFD dado é o AFD mínimo para L.
c) Assuma que, para aplicação do método de Kleene a A, se designa o estado s_i apenas pelo símbolo i, para $i = 1, 2, 3, 4, 5$. Indique uma expressão regular (abreviada) que descreva a linguagem $\mathcal{L}(r_{11}^{(3)})$. Justifique sucintamente.

Resposta:

3a)



L é o conjunto das sequências finitas de a's e b's que depois do último b têm um número de a's que é múltiplo de três, possivelmente zero (se não têm b's, têm um número de a's que é multiplo de três, podendo ser zero).

3b)

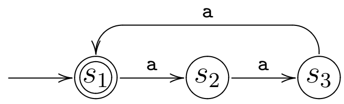
Segundo a definição da linguagem, tem-se: $\mathtt{b}z \in L$ se e só se $z \in L$ e $\mathtt{bb}z \in L$ se e só se $z \in L$. Isto implica que as palavras b e bb são equivalentes segundo R_L (relação do teorema de Myhill-Nerode que carateriza o AFD mínimo para L), pelo que, sendo δ' a função de transição do AFD mínimo, teriamos $\delta'([\mathtt{b}], \mathtt{b}) = [\mathtt{bb}] = [\mathtt{b}]$. Assim, os dois estados s_4 e s_5 dão origem a um único estado no AFD mínimo. Portanto, o AFD dado não é o AFD mínimo para L.

Resposta (cont):

3b) Resposta alternativa:
O AFD dado não é mínimo pois teriamos um AFD equivalente se retirassemos s_5 e colocassemos um lacete com $\mathtt{b}$ em s_4. Os estados s_4 e s_5 não são distinguíveis pois, para quaisquer x e y tais que $\hat{\delta}(s_1, x) = s_4$ e $\hat{\delta}(s_1, y) = s_5$, tem-se $xz \in \mathcal{L}(A)$ se e só se $yz \in \mathcal{L}(A)$, para todo $z \in \Sigma^\star$. De facto, se $z = \varepsilon$, a condição verifica-se pois s_4 e s_5 são ambos estados finais. Se $z = \mathtt{b}w$, para algum $w \in \Sigma^\star$, então $\hat{\delta}(s_1, xz) = \hat{\delta}(s_5, w) = \hat{\delta}(s_1, yz)$ e, se $z = \mathtt{a}w$ então $\hat{\delta}(s_1, xz) = \hat{\delta}(s_2, w) = \hat{\delta}(s_1, yz)$. Logo, se $z \neq \varepsilon$, as palavras xz e yz são equivalentes para A e, portanto, $xz \in \mathcal{L}(A)$ sse $yz \in \mathcal{L}(A)$.

3c)
No método de Kleene, a expressão regular $r_{ij}^{(k)}$ descreve a linguagem das palavras que levam o autómato do estado i ao estado j, podendo passar por estados *intermédios* numerados até k (inclusivé).

Assim, a $\mathcal{L}(r_{11}^{(3)})$ é descrita pela expressão regular $(\mathtt{aaa})^\star$ pois além de ε apenas podiamos considerar palavras que correspondem a percursos do estado s_1 para o estado s_1 que não passam em s_4 nem s_5, o que restringe a análise ao diagrama seguinte:



(Fim)