Departamento de Ciência de Computadores **FCUP**

Modelos de Computação (CC1004) 2014/15

1º Teste – 11.04.2015 **Uma resolução (v3)** *duração: 2h + 30m*

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

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

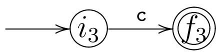
Resposta:
Uma expressão regular sobre Σ é qualquer sequência finita de símbolos de $\Sigma \cup \{\emptyset, \varepsilon, +, *,), (\}$ que se possa obter por aplicação das regras seguintes: ε, $\emptyset$, a, b e c são expressões regulares sobre Σ; se r e s são expressões regulares sobre Σ então $(r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

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

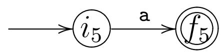
b) Determine o autómato finito que resulta da aplicação do método de Thompson à expressão regular r. Apresente **os passos relevantes** dessa construção.

Resposta:
Na continuação da resposta anterior, os AFNDs-ε para r_3, r_5, e r_6 são:

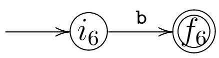
$A(r_3)$:



$A(r_5)$:

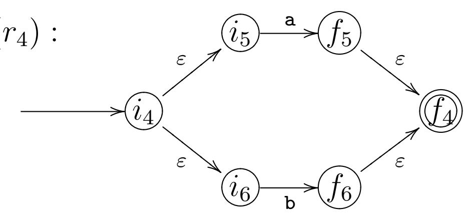


$A(r_6)$:

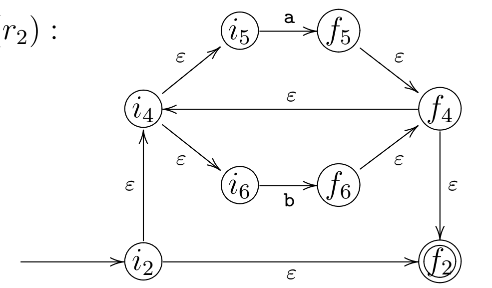


Os AFNDs-ε para $r_4 = (r_5 + r_6)$ e $r_2 = (r_4^\star)$ são:

$A(r_4)$:

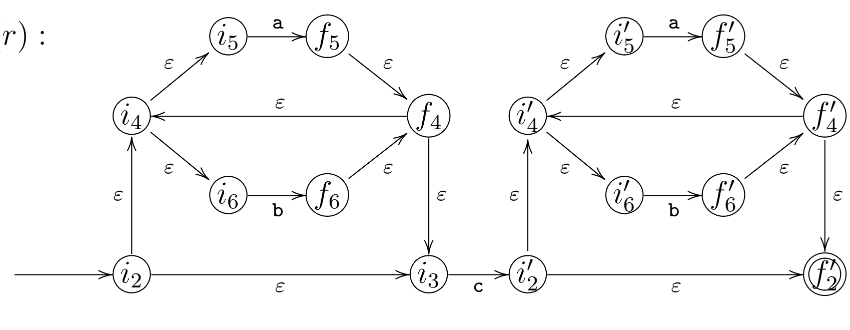


$A(r_2)$:



A partir dos AFNDs-ε $A(r_2)$ e $A(r_3)$ construimos o autómato $A(r_1)$ para $r_1 = (r_2 r_3)$, por identificação do estado inicial de $A(r_3)$ com o final de $A(r_2)$. Usando $A(r_1)$ e um clone de $A(r_2)$, construimos o autómato para $r = (r_1 r_2)$, por identificação do estado final de $A(r_1)$ com o inicial do clone 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.

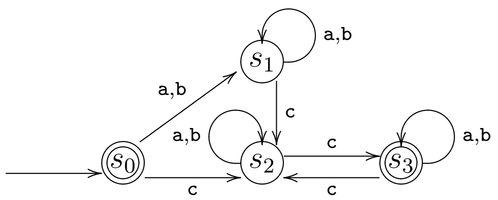
Resposta:
A expressão abreviada é $(\mathtt{a}+\mathtt{b})^\star\mathtt{c}(\mathtt{a}+\mathtt{b})^\star$.

A linguagem descrita pela expressão r é o conjunto das palavras de $\{\mathtt{a},\mathtt{b},\mathtt{c}\}^\star$ que têm exatamente um $\mathtt{c}$.

d) Descreva informalmente a linguagem descrita pela expressão regular $((rr)^\star)$. 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 $\mathtt{c}$'s e em que esse número não é zero, a menos que a palavra seja ε.

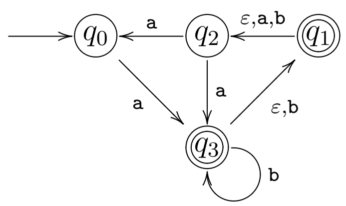
A linguagem é reconhecida pelo AFD representado à esquerda.



O que se sabe em cada estado:
- s_1: a palavra não tem $\mathtt{c}$'s mas tem algum $\mathtt{a}$ ou $\mathtt{b}$.
- s_2: a palavra tem número ímpar de $\mathtt{c}$'s.
- s_3: a palavra tem número par de $\mathtt{c}$'s e tem $\mathtt{c}$'s.
- s_0: a palavra é vazia.

O estado s_0 é estado final porque $\varepsilon \in \mathcal{L}((rr)^\star)$. Em s_0 muda para s_1 quando lê $\mathtt{a}$ ou $\mathtt{b}$ pois as palavras que não têm $\mathtt{c}$'s mas têm $\mathtt{a}$'s ou $\mathtt{b}$'s não pertencem a $\mathcal{L}((rr)^\star)$. Os estados s_2 e s_3 são necessários para controlar a paridade do número de $\mathtt{c}$'s e s_3 não é equivalente a s_0 pois em s_0 a palavra ainda não tem $\mathtt{c}$'s e em s_3 já tem (e o autómato deveria rejeitar palavras εz com $z \in \{\mathtt{a},\mathtt{b}\}^\star \setminus \{\varepsilon\}$ e aceitar $\mathtt{cc}z$).

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



a) Qual é o valor de $\delta(q_0,\mathtt{a})$, $\delta(q_2,\mathtt{b})$, $\delta(q_3,\varepsilon)$, e $\delta(q_3,\mathtt{b})$? 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, $\delta(q_0,\mathtt{a}) = \{q_3\}$, $\delta(q_2,\mathtt{b}) = \emptyset$, $\delta(q_3,\varepsilon) = \{q_1\}$, $\delta(q_3,\mathtt{b}) = \{q_1,q_3\}$.

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

Resposta:
$\hat{\delta}(\{q_0\}, \texttt{abb}) = \hat{\delta}(Fecho_\varepsilon(\{q_3\}), \texttt{bb}) = \hat{\delta}(\{q_3, q_1, q_2\}, \texttt{bb}) = \hat{\delta}(Fecho_\varepsilon(\{q_3, q_1, q_2\}), \texttt{b}) = \hat{\delta}(\{q_3, q_1, q_2\}, \texttt{b}) = \hat{\delta}(Fecho_\varepsilon(\{q_3, q_1, q_2\}), \varepsilon) = Fecho_\varepsilon(\{q_3, q_1, q_2\}) = \{q_3, q_1, q_2\}$.

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

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

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

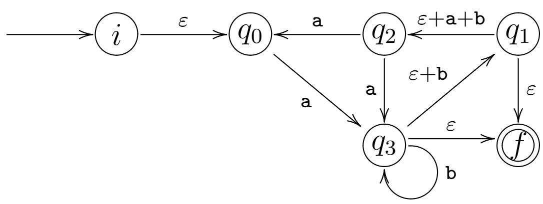
A palavra abb pertence a $\mathcal{L}(A)$ porque $\hat{\delta}(\{q_0\}, \texttt{abb}) = \{q_3, q_2, q_1\}$ contém estados finais (nomeadamente, os estados q_1 e q_3).

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.

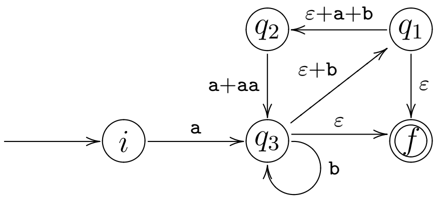
Resposta:

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.

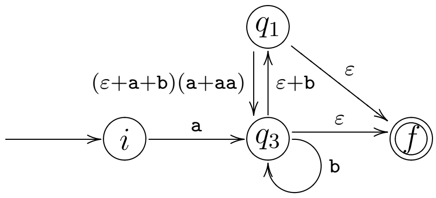
Começamos por substituir os estados finais por um único estado final f (do qual não saem transições). Inserimos um novo estado inicial i, para garantir que não chegam transições ao estado inicial. Substituimos as etiquetas dos ramos por expressões regulares.



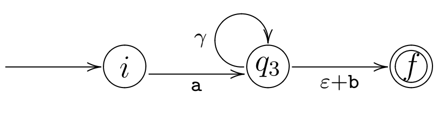
Eliminamos q_0, substituindo os percursos $q_2q_0q_3$ e iq_0q_3 por arcos (q_2, q_3) e (i, q_3), com etiquetas aa e $\varepsilon\texttt{a} \equiv \texttt{a}$, respetivamente. Como já existia um arco de q_2 para q_3, substituimos a sua expressão regular por $\texttt{a} + \texttt{aa}$.



Eliminamos q_2, substituindo o percurso $q_1q_2q_3$ por um arco (q_1, q_3) com etiqueta $(\varepsilon + \texttt{a} + \texttt{b})(\texttt{a} + \texttt{aa})$.



Eliminamos q_1, substituindo o percurso q_3q_1f e $q_3q_1q_3$ por arcos (q_3, f) e (q_3, q_3) com etiquetas $(\varepsilon + \texttt{b})\varepsilon \equiv \varepsilon + \texttt{b}$ e $(\varepsilon + \texttt{b})(\varepsilon + \texttt{a} + \texttt{b})(\texttt{a} + \texttt{aa})$.
No diagrama, $\gamma = \texttt{b} + (\varepsilon + \texttt{b})(\varepsilon + \texttt{a} + \texttt{b})(\texttt{a} + \texttt{aa})$.



(cont.)

Resposta 2d) cont.:
Finalmente eliminamos q_3, substituindo os percursos $iq_3q_3^\star f$ por um arco (i,f), etiquetado pela expressão $\mathtt{a}\gamma^\star(\varepsilon+\mathtt{b})$.

Como $\gamma^\star = (\mathtt{b}+(\varepsilon+\mathtt{b})(\varepsilon+\mathtt{a}+\mathtt{b})(\mathtt{a}+\mathtt{aa}))^\star \equiv (\mathtt{a}+\mathtt{b})^\star$, pois $\varepsilon\varepsilon\mathtt{a} = \mathtt{a} \in \mathcal{L}((\varepsilon+\mathtt{b})(\varepsilon+\mathtt{a}+\mathtt{b})(\mathtt{a}+\mathtt{aa}))$, a expressão obtida para $\mathcal{L}(A)$ pode ser simplificada assim:

$$\mathtt{a}\gamma^\star(\varepsilon+\mathtt{b}) \equiv \mathtt{a}(\mathtt{a}+\mathtt{b})^\star(\varepsilon+\mathtt{b}) \equiv \mathtt{a}(\mathtt{a}+\mathtt{b})^\star + \mathtt{a}(\mathtt{a}+\mathtt{b})^\star\mathtt{b} \equiv \mathtt{a}(\mathtt{a}+\mathtt{b})^\star.$$

Portanto, $\mathcal{L}(A) = \mathcal{L}(\mathtt{a}(\mathtt{a}+\mathtt{b})^\star)$.

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.

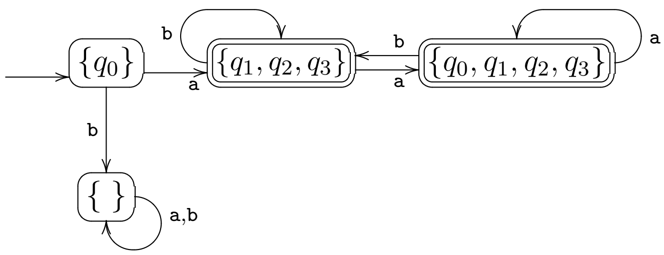
Resposta:
O AFD $A' = (2^S, \Sigma, \delta', Fecho_\varepsilon(q_0), F')$, com $F' = \{E \mid E \in 2^S \text{ e } E \cap F \neq \emptyset\}$ e

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

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.



Nota adicional: o AFD obtido reconhece $\mathcal{L}(\mathtt{a}(\mathtt{a}+\mathtt{b})^\star)$, como seria de esperar, por 2d).

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 $F = \{s_1,s_2\}$ 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_3	s_4
s_4	s_2	s_5
s_5	s_2	s_1

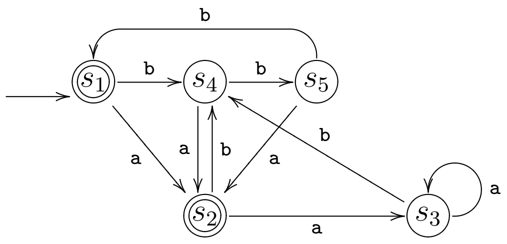
a) Desenhe o diagrama de transição de A e descreva informalmente L.

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 $\mathcal{L}(r_{13}^{(3)})$. Justifique sucintamente.

Resposta:

3a)



L é o conjunto das sequências finitas de a's e b's que terminam em a mas não em aa ou o número de b's depois do último a é múltiplo positivo de três. Se não têm a's, têm um número de b's que é multiplo de três, podendo ser zero.

3b)

Sendo R_L a relação do teorema de Myhill-Nerode que carateriza o AFD mínimo para L, temos:

$(\varepsilon, \mathtt{b}) \notin R_L$ pois $\varepsilon \in L$ e $\mathtt{b} \notin L$.

$(\varepsilon, \mathtt{a}) \notin R_L$ pois $\varepsilon\mathtt{a} \in L$ e $\mathtt{aa} \notin L$.

$(\mathtt{b}, \mathtt{a}) \notin R_L$ pois $\mathtt{a} \in L$ e $\mathtt{b} \notin L$.

$(\varepsilon, \mathtt{aa}) \notin R_L$ pois $\varepsilon \in L$ e $\mathtt{aa} \notin L$.

$(\mathtt{a}, \mathtt{aa}) \notin R_L$ pois $\mathtt{a} \in L$ e $\mathtt{aa} \notin L$.

$(\mathtt{b}, \mathtt{aa}) \notin R_L$ pois $\mathtt{ba} \in L$ e $\mathtt{aaa} \notin L$.

$(\varepsilon, \mathtt{bb}) \notin R_L$ pois $\varepsilon \in L$ e $\mathtt{bb} \notin L$.

$(\mathtt{a}, \mathtt{bb}) \notin R_L$ pois $\mathtt{aa} \notin L$ e $\mathtt{bba} \in L$.

$(\mathtt{b}, \mathtt{bb}) \notin R_L$ pois $\mathtt{bb} \notin L$ e $\mathtt{bbb} \in L$.

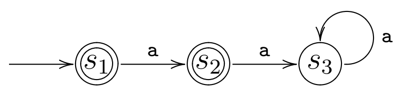
$(\mathtt{aa}, \mathtt{bb}) \notin R_L$ pois $\mathtt{aaa} \notin L$ e $\mathtt{bba} \in L$.

Então, o AFD mínimo para L tem pelo menos cinco estados ($[\varepsilon]$, $[\mathtt{a}]$, $[\mathtt{b}]$, $[\mathtt{aa}]$ e $[\mathtt{bb}]$), e como $L = \mathcal{L}(A)$ e A é um AFD com cinco estados, então A é o AFD mínimo para L.

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 $\mathcal{L}(r_{13}^{(3)})$ é descrita pela expressão regular $\mathtt{aaa}^\star$ pois apenas podemos considerar palavras que correspondem a percursos do estado s_1 para o estado s_3 que não passem em s_4 nem s_5, o que restringe a análise ao diagrama seguinte:



(Fim)