LFA0001 – Linguagens Formais e Autômatos
Aula 06
Expressões Regulares

Karina Girardi Roggia
karina.roggia@udesc.br

Departamento de Ciência da Computação
Centro de Ciências Tecnológicas
Universidade do Estado de Santa Catarina

2016

Sumário

Expressão Regular

Formalismo denotacional (gerador)

Definida a partir de

- conjuntos básicos
- concatenação e união

Adequadas para a comunicação

- humano X humano
- humano X máquina

Expressão Regular

Definição

Uma *Expressão Regular* (ER) sobre o alfabeto Σ é indutivamente definida como segue:

Base

- $\varnothing$ é ER
  denota a linguagem vazia: $\varnothing$
- ε é ER
  denota a linguagem $\{\varepsilon\}$
- x é ER (para qualquer $x \in \Sigma$)
  denota a linguagem $\{x\}$

Expressão Regular

Passo se r e s são ER e denotam as linguagens R e S respectivamente

União $(r + s)$ é ER
denota a linguagem $R \cup S$

Concatenação (rs) é ER
denota a linguagem
$RS = \{uv \mid u \in R \text{ e } v \in S\}$

Concatenação Sucessiva (r^*) é ER
denota a linguagem R^*

Linguagem Gerada

Definição (Linguagem Gerada por ER)

Se r é uma ER, a correspondente linguagem denotada é chamada de **linguagem gerada por** r e escrita como $L(r)$ ou GERA(r).

Usualmente omite-se os parênteses em uma ER.
Precedência de operadores:

Concatenação Sucessiva $>$ Concatenação $>$ União

Exemplos

ER	Linguagem Gerada
aa	
ba^*	
$(a+b)^*$	
$(a+b)^*aa(a+b)^*$	
$a^*ba^*ba^*$	
$(a+b)^*(aa+bb)$	
$(a+\varepsilon)(b+ba)^*$	

Exemplos

ER	Linguagem Gerada
aa	somente a palavra aa
ba^*	
$(a+b)^*$	
$(a+b)^*aa(a+b)^*$	
$a^*ba^*ba^*$	
$(a+b)^*(aa+bb)$	
$(a+\varepsilon)(b+ba)^*$	

Exemplos

ER	Linguagem Gerada
aa	somente a palavra aa
ba^*	todas as palavras iniciadas por b, seguido por zero ou mais a
$(a+b)^*$	
$(a+b)^*aa(a+b)^*$	
$a^*ba^*ba^*$	
$(a+b)^*(aa+bb)$	
$(a+\varepsilon)(b+ba)^*$	

UDESC

Exemplos

ER	Linguagem Gerada
aa	somente a palavra aa
ba^*	todas as palavras iniciadas por b, seguido por zero ou mais a
$(a+b)^*$	todas as palavras sobre $\{a, b\}$
$(a+b)^*aa(a+b)^*$	
$a^*ba^*ba^*$	
$(a+b)^*(aa+bb)$	
$(a+\varepsilon)(b+ba)^*$	

Exemplos

ER	Linguagem Gerada
aa	somente a palavra aa
ba^*	todas as palavras iniciadas por b, seguido por zero ou mais a
$(a+b)^*$	todas as palavras sobre $\{a, b\}$
$(a+b)^*aa(a+b)^*$	todas as palavras contendo aa como subpalavra
$a^*ba^*ba^*$	
$(a+b)^*(aa+bb)$	
$(a+\varepsilon)(b+ba)^*$	

UDESC

Exemplos

ER	Linguagem Gerada
aa	somente a palavra aa
ba^*	todas as palavras iniciadas por b, seguido por zero ou mais a
$(a+b)^*$	todas as palavras sobre $\{a, b\}$
$(a+b)^*aa(a+b)^*$	todas as palavras contendo aa como subpalavra
$a^*ba^*ba^*$	todas as palavras contendo exatamente dois b
$(a+b)^*(aa+bb)$	
$(a+\varepsilon)(b+ba)^*$	

Exemplos

ER	Linguagem Gerada
aa	somente a palavra aa
ba^*	todas as palavras iniciadas por b, seguido por zero ou mais a
$(a+b)^*$	todas as palavras sobre $\{a, b\}$
$(a+b)^*aa(a+b)^*$	todas as palavras contendo aa como subpalavra
$a^*ba^*ba^*$	todas as palavras contendo exatamente dois b
$(a+b)^*(aa+bb)$	todas as palavras que terminam com aa ou bb
$(a+\varepsilon)(b+ba)^*$	

Exemplos

ER	Linguagem Gerada
aa	somente a palavra aa
ba^*	todas as palavras iniciadas por b, seguido por zero ou mais a
$(a+b)^*$	todas as palavras sobre $\{a, b\}$
$(a+b)^*aa(a+b)^*$	todas as palavras contendo aa como subpalavra
$a^*ba^*ba^*$	todas as palavras contendo exatamente dois b
$(a+b)^*(aa+bb)$	todas as palavras que terminam com aa ou bb
$(a+\varepsilon)(b+ba)^*$	todas as palavras que não possuem dois a consecutivos

Exemplo

Linguagem gerada pela ER $(a+b)^*(aa+bb)$

- a e b denotam $\{a\}$ e $\{b\}$, respectivamente
- $a+b$ denota $\{a\} \cup \{b\} = \{a, b\}$
- $(a+b)^*$ denota $\{a, b\}^*$
- aa e bb denotam $\{a\}\{a\} = \{aa\}$ e $\{b\}\{b\} = \{bb\}$, respectivamente
- $(aa+bb)$ denota $\{aa\} \cup \{bb\} = \{aa, bb\}$
- $(a+b)^*(aa+bb)$ denota $\{a, b\}^*\{aa, bb\}$

Portanto, GERA($(a+b)^*(aa+bb)$) é

$$\{aa, bb, aaa, abb, baa, bbb, aaaa, aabb, abaa, abbb, baaa, babb, bbaa, bbbb, \ldots\}$$

ER denota linguagens regulares

A classe das linguagens denotadas por *ER* é a classe das linguagens regulares.

Para provar a afirmação acima, é necessário:

- Dada uma ER r, construir um autômato finito (neste caso, iremos construir um AFε M_r) que reconheça $L(r)$
- Dado um autômato finito M, construir uma expressão regular r_M que denote $L(M)$

ER→AFε

Teorema: Se r é ER, então $L(r)$ é linguagem regular.
Prova: por indução no número de operadores de r.

Base: r é ER com zero operadores.

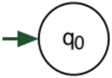
- $r = \varnothing$. Autômato:

ER→AFε

Teorema: Se r é ER, então $L(r)$ é linguagem regular.
Prova: por indução no número de operadores de r.

Base: r é ER com zero operadores.

- $r = \varnothing$. Autômato: $M_1 = \langle \varnothing, \{q_0\}, \delta_1, q_0, \varnothing \rangle$

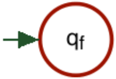


ER$\rightarrow$AFε

- $r = \varepsilon$. Autômato:

ER→AFε

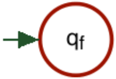
- $r = \varepsilon$. Autômato: $M_2 = \langle \varnothing, \{q_f\}, \delta_1, q_f, \{q_f\} \rangle$



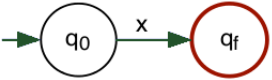
- $r = x$, sendo $x \in \Sigma$. Autômato:

ER→AFε

- $r = \varepsilon$. Autômato: $M_2 = \langle \emptyset, \{q_f\}, \delta_1, q_f, \{q_f\} \rangle$



- $r = x$, sendo $x \in \Sigma$. Autômato: $M_3 = \langle \{x\}, \{q_0, q_f\}, \delta_3, q_0, \{q_f\} \rangle$



ER→AFε

Hipótese: r é ER com até $n > 0$ operadores.
r terá um AFε R tal que $L(r) = L(R)$.

Passo: r é ER com $n + 1$ operadores.
Portanto r se enquadrará em um dos seguintes três casos:

- $r = r_1 + r_2$
- $r = r_1 r_2$
- $r = r_1^*$

onde r_1 e r_2 possuem no máximo n operadores
Pela hipótese de indução, existem

$$M_1 = \langle \Sigma_1, Q_1, \delta_1, q_{0_1}, \{q_{f_1}\} \rangle$$
$$M_2 = \langle \Sigma_2, Q_2, \delta_2, q_{0_2}, \{q_{f_2}\} \rangle$$

tais que

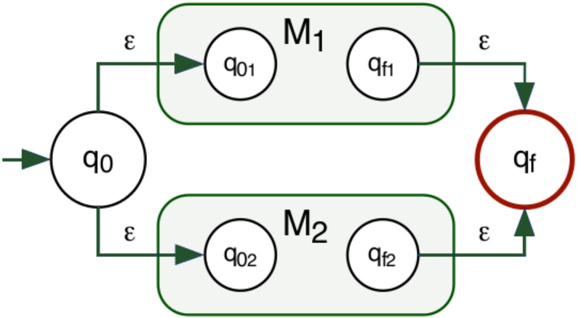
$$L(M_1) = L(r_1) \text{ e } L(M_2) = L(r_2)$$

ER→AFε

- $r = r_1 + r_2$. Autômato:

ER→AFε

- $r = r_1 + r_2$. Autômato:
  $M = \langle \Sigma_1 \cup \Sigma_2, Q_1 \cup Q_2 \cup \{q_0, q_f\}, \delta, q_0, \{q_f\} \rangle$

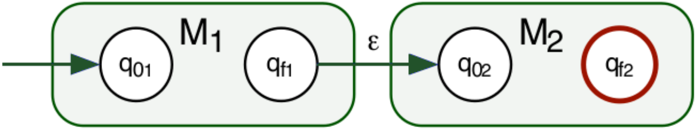


ER→AFε

- $r = r_1 r_2$. Autômato:

ER→AFε

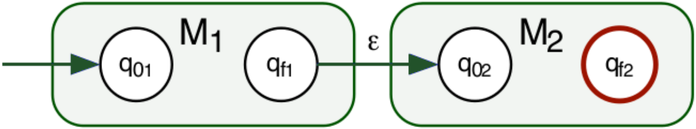
- $r = r_1 r_2$. Autômato: $M = \langle \Sigma_1 \cup \Sigma_2, Q_1 \cup Q_2, \delta, q_{0_1}, \{q_{f_2}\} \rangle$



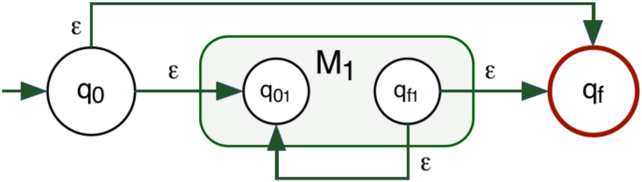
- $r = r^*$. Autômato:

ER→AFε

- $r = r_1 r_2$. Autômato: $M = \langle \Sigma_1 \cup \Sigma_2, Q_1 \cup Q_2, \delta, q_{0_1}, \{q_{f_2}\} \rangle$

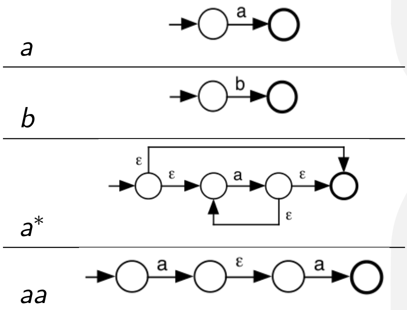


- $r = r^*$. Autômato: $M = \langle \Sigma_1, Q_1 \cup \{q_0, q_f\}, \delta, q_0, \{q_f\} \rangle$



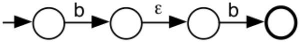
Exemplo

Construção do AFε de $a^*(aa + bb)$

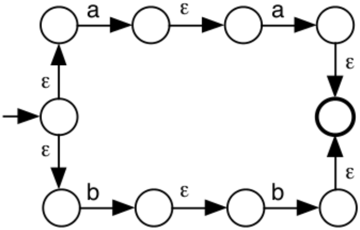


Exemplo

bb

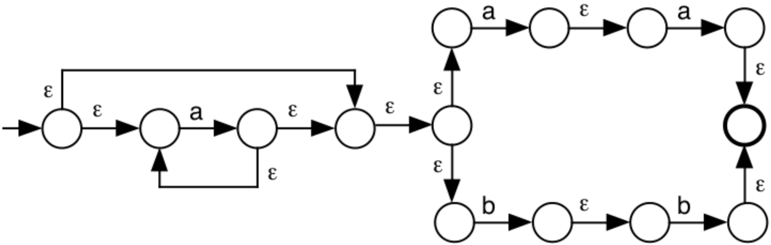


$(aa + bb)$



Exemplo

Autômato Final: $a^*(aa + bb)$



AFD→ER

Teorema: Se L é linguagem regular, então existe uma ER r tal que $L(r) = L$.

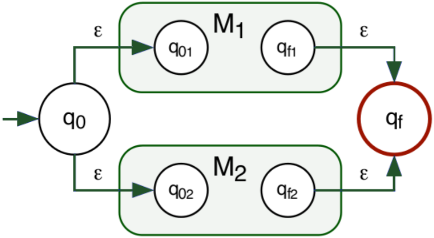
Prova: (Não será esmiuçada na disciplina)

Pode ser encontrada nas páginas 98 à 109 de

Hopcroft, J. E., Ullman, J. D. e Motwani, R. **Introdução à Teoria de Autômatos, Linguagens e Computação**. Tradução da segunda edição americana. Editora Campus, 2003.

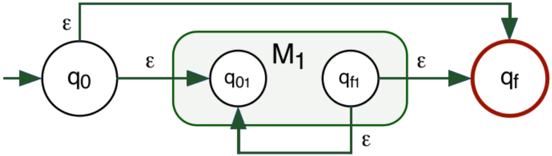
Exercícios

1. O que aconteceria caso não se inserissem os estados novos q_0 e q_f e, ao invés disso, unificásse-mos os estados iniciais ou finais de M_1 e M_2 no caso abaixo da construção de um AFε a partir de uma ER?



Exercícios

2. O que aconteceria caso não se introduzisse o estado q_f, mantendo q_{f_1} como o estado final e passando a transição vazia de q_0 para q_{f_1} de M_1 no caso abaixo da construção de um AFε a partir de uma ER?



Exercícios

3. Desenvolva expressões regulares que gerem as seguintes linguagens sobre $\{a, b\}$
   (a) $\{w \mid w$ tem no máximo um par de a como subpalavra e no máximo um par de b como subpalavra$\}$
   (b) $\{w \mid$ qualquer par de a antecede qualquer par de $b\}$
   (c) $\{w \mid w$não possui abacomo subpalavra$\}$