

Aspectos Teóricos da Computação

Prof. Rodrigo Martins

rodrigo.martins@francomontoro.com.br



Cronograma da Aula



- ◆ Expressões Regulares e Linguagens

Linguagens Regulares

- ◆ Todos os formalismos reconhecedores foram vistos.
 - Autômatos Finitos Determinísticos
 - Autômatos Finitos Não Determinísticos
 - Autômatos Finitos com Movimentos ε
- ◆ Agora veremos um formalismo denotacional ou gerador
 - Expressões Regulares

Expressões Regulares - Definição

- ◆ As expressões regulares podem definir exatamente as mesmas linguagens que as diversas formas de autômatos descrevem: as linguagens regulares.
- ◆ As expressões regulares oferecem algo que os autômatos não oferecem: um modo declarativo de expressar os strings que queremos aceitar.
- ◆ As expressões regulares servem como a linguagem de entrada para muitos sistemas que processam strings.

Expressões Regulares - Definição

- ◆ Não há mais a preocupação em aceitar ou rejeitar uma palavra.
 - Localizar cadeias de um texto
 - Para criar analisadores léxicos, que são componentes fundamentais dos compiladores
 - Validação de campo

Expressões Regulares - Definição

Assim como uma expressão aritmética representa um número natural:

$$(10 + 5) \times 7$$

Uma **expressão regular** representa uma linguagem:

$$(0 + 1).0^*$$

Os Operadores de Expressões Regulares

As expressões regulares denotam linguagens.

Exemplo 1:

A expressão regular $01^* + 10^*$ denota a linguagem que consiste em todos os strings que são um único 0 seguidos por qualquer número de 1's ou um único 1 seguido por qualquer número de 0's.

Os Operadores de Expressões Regulares

Exemplo 2:

- ◆ Linguagem regular: o conjunto de cadeias de 0's e 1's tais que comece com qualquer quantidade de 1's (inclusive nenhum), seguidos necessariamente de um 0 e outra sequencia com qualquer quantidade de 1's
 - Essa linguagem aparentemente complexa pode ser escrita em forma de expressão regular facilmente:

1^*01^*

Expressões Regulares - Exemplo

Na expressão $(0 + 1) \cdot 0^*$:

- 0 representa o conjunto $\{0\}$
- 1 representa o conjunto $\{1\}$
- $(0 + 1)$ representa o conjunto $\{0\} \cup \{1\} = \{0, 1\}$
- 0^* representa $\{0\}^*$

Então $(0 + 1) \cdot 0^*$ representa a linguagem:

$\{uv: u \in \{0, 1\} \text{ e } v = 0^n, n \geq 0\}$

As Três operações sobre Linguagens

União

A união de duas linguagens L e M , denotadas por $L \cup M$, é o conjunto de strings que estão em L ou M , ou em ambas.

Operador $+$ é a ideia de **ou**

Por exemplo, se $L = \{001, 10, 111\}$ e

$M = \{\varepsilon, 001\}$, então

$L \cup M = \{\varepsilon, 10, 001, 111\}$

As Três operações sobre Linguagens (cont.)

Concatenação

A concatenação de linguagens L e M é o conjunto de strings que podem ser formados tornando-se qualquer string em L e concatenando-se esse string com qualquer string em M .

Operador $.$ é a ideia de **e**

Por exemplo, se $L = \{001, 110\}$ e $M = \{\varepsilon, 11, 110\}$, então $L.M$ (com um ponto) ou LM (sem ponto), onde $LM = \{001, 00111, 001110, 110, 11011, 110110\}$

As Três operações sobre Linguagens (cont.)

Por exemplo, se $L = \{001, 110\}$ e $M = \{\varepsilon, 11, 110\}$, então $L.M$ (com um ponto) ou LM (sem ponto), onde $LM = \{001, 00111, 001110, 110, 11011, 110110\}$

Atenção: LM é diferente de ML

As Três operações sobre Linguagens (cont.)

Exemplo:

Quais palavras a expressão **$(a+b)c$** representa?

Na expressão dada, temos uma união de **a** e **b** , que representa $\{a,b\}$. Em seguida, concatenada a **$(a+b)$** temos a expressão **c** que representa $\{c\}$.

O resultado da concatenação $\{a,b\}.\{c\}$ dá a linguagem $\{ac, bc\}$, que é a resposta esperada.

As Três operações sobre Linguagens (cont.)

Fechamento de Kleene ou Estrela

O fechamento (ou estrela, ou fechamento de Kleene) de uma linguagem L é denotado L^* e representa o conjunto dos strings que podem ser formados tomando-se qualquer número de strings de L , possivelmente com repetições (isto é, o mesmo string pode ser selecionado mais de uma vez) e concatenando-se todos eles.

Por exemplo se $L = \{0, 1\}$ então L^* representa todos os strings de 0's e 1's. Se $L = \{0, 11\}$, então L^* consiste nos strings de 0's e 1's tais que os símbolos 1 formam pares, por exemplo, 011, 11110 e ε , mas não 01011 ou 101.

Exemplo

- ◆ Operadores (seja r uma expressão)
 - concatenação sucessiva: r^*
 - Dá uma ideia de zero ou mais repetições de r
 - Denota $L = L_r^*$
 $= \{\text{palavras formadas pela concatenação de zero ou mais palavras de } L_r\}$
 - Exemplo
 - \mathbf{a}^* denota $L = \{\varepsilon, a, aa, aaa, aaaa, \dots\}$
 - \mathbf{ab}^* denota $L = \{a, ab, abb, abbb, \dots\}$

Exemplo

- ◆ Todas as palavras sobre $T = \{a, b\}$

$$(a + b)^*$$

- ◆ Palavras que terminam com aa ou bb

$$(a+b)^*(aa+bb)$$

Exemplo

O conjunto de todas as cadeias de 0's e 1's com exatamente três símbolos

- $(0+1)(0+1)(0+1)$

O conjunto de cadeias de 0's e 1's contendo pelo menos um símbolo 0

- $0(0+1)^* + (0+1)^*0 + (0+1)^*0(0+1)^*$

Forneça uma descrição em português da expressão:
 $(0+1)^*101(0+1)^*$

- o conjunto de todas as cadeias de zeros e uns que contém 101 como subcadeia



Autômatos Finitos e Expressões Regulares

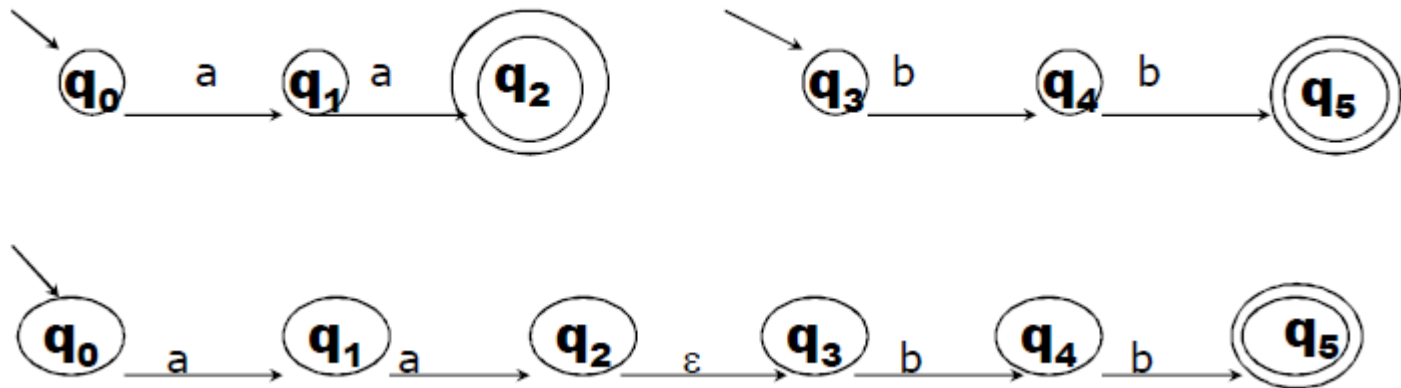


Provaremos a equivalência entre ER e AF da seguinte forma:

Como obter um AF a partir de uma ER

Autômatos Finitos e Expressões Regulares

Seja $A = \{aa\}$, $B = \{bb\}$



Aplicações

- ◆ Especificar endereços de e-mail válidos
- ◆ Procura (avançada) por arquivos
- ◆ Para especificar Linguagens de Programação
 - Especificar identificadores
 - Especificar números inteiros
 - Especificar números decimais.

Exercícios

- 1) Qual a expressão regular dada o conjunto de cadeias de 0's e 1's tais que comece com qualquer quantidade de 1's (inclusive nenhum), seguidos necessariamente de um 1 e outra sequencia com qualquer quantidade de 0's
- 2) Quais palavras a expressão **(a+c)b** representa?
- 3) Quais as três operações sobre linguagens, explique cada uma delas.
- 4) Represente todas as palavras sobre $T = \{c, d\}$
- 5) Represente todas as palavras sobre $T = \{c, d\}$ que terminam com cc ou dd

Exercícios

- 6) Represente o conjunto de todas as cadeias de 0's e 1's com exatamente cinco símbolos
- 7) Represente o conjunto de todas as cadeias de 0's e 1's contendo o símbolo 0 no final
- 8) Forneça uma descrição em português da expressão:
 $(0)^*11(0+1)^*$

Referências desta aula

- ◆ HOPCROFT, John E.; MOTWANI, Rajeev; ULLMAN, Jeffrey D. Introdução a teoria de autômatos, linguagens e computação. Rio de Janeiro: Campus, 2002.

FIM

Obrigado

Rodrigo