

TEORIA DA COMPUTAÇÃO

AULA 7: AUTÔMATOS E LINGUAGENS (CAP 1)

EXPRESSÕES REGULARES

PROFESSOR: LUCAS CAMBUIM

Expressões Regulares

Na aritmética, podemos usar as operações $+$ e $\times$ para montar expressões tais como $(5 + 3) \times 4$.

Similarmente, podemos usar as operações regulares para montar expressões descrevendo linguagens, que são chamadas ***expressões regulares***.

Por exemplo: $(0 \cup 1)0^*$

O valor de uma expressão aritmética é um número. No exemplo é 32.

O valor de uma expressão regular é uma **linguagem**.

No exemplo: a linguagem das cadeias binárias que começam com 1 ou 0 seguido por um número qualquer de 0s.

Expressões Regulares

Como obtemos o valor de $(0 \cup 1)0^*$?

Primeiro, os símbolos 0 e 1 são abreviações para os conjuntos $\{0\}$ e $\{1\}$.

Dessa forma, $(0 \cup 1)$ significa $(\{0\} \cup \{1\})$.

O valor dessa parte é a linguagem $\{0,1\}$.

Segundo, como o símbolo $\circ$ em álgebra, o símbolo da concatenação frequentemente está implícito nas expressões regulares.

Por conseguinte, $(0 \cup 1)0^*$ é, na realidade, uma abreviação de $(0 \cup 1) \circ 0^*$.

A concatenação junta as cadeias das duas partes para obter o valor da expressão inteira.

Expressões Regulares

Outro exemplo: $(0 \cup 1)^*$

Ela começa com a linguagem $(0 \cup 1)$ e aplica a operação $*$.

O valor dessa expressão é a linguagem consistindo de todas as possíveis cadeias de 0s e 1s.

Se $\Sigma = \{0,1\}$, podemos escrever Σ como abreviação para a expressão regular $(0 \cup 1)$.

Expressões Regulares

Genericamente, se Σ for um alfabeto qualquer, a expressão regular Σ descreve a linguagem consistindo de todas as cadeias de comprimento 1 sobre esse alfabeto.

Σ^* descreve a linguagem consistindo de todas as cadeias sobre esse alfabeto.

Expressões Regulares

Qual é a linguagem Σ^*1 ?

Expressões Regulares

Qual é a linguagem Σ^*1 ?

É a linguagem que contem todas as cadeias que terminam em 1.

Expressões Regulares

E a linguagem $(0 \Sigma^*) \cup (\Sigma^* 1)$?

Expressões Regulares

E a linguagem $(0 \Sigma^*) \cup (\Sigma^* 1)$?

Consiste de todas as cadeias que começam com 0 ou terminam com 1.

Precedência dos operadores

A operação estrela é feita primeiro, seguida por concatenação, e finalmente união, a menos que parênteses sejam usados para mudar a ordem usual.

Expressões Regulares: Definição Formal

Dizemos que R é uma *expressão regular* se R for

1. a para algum a no alfabeto Σ ,
2. ε ,
3. $\emptyset$,
4. $(R_1 \cup R_2)$, onde R_1 e R_2 são expressões regulares,
5. $(R_1 \circ R_2)$, onde R_1 e R_2 são expressões regulares, ou
6. (R_1^*) , onde R_1 é uma expressão regular.
7. Nos itens 1 e 2, as expressões regulares a e ε representam as linguagens $\{a\}$ e $\{\varepsilon\}$, respectivamente.

Expressões Regulares: Notação

$$R^+ = RR^*$$

$$R^* = R^+ \cup \{\varepsilon\}$$

R^k : concatenação de k's Rs.

Quando queremos distinguir entre uma expressão regular R e a linguagem que ela descreve, escrevemos $L(R)$ como sendo a linguagem de R .

Expressões Regulares: Exemplos onde $\Sigma=\{0,1\}$

0^*10^*

$\{w \mid w \text{ contem um único } 1\}.$

$\Sigma^*1\Sigma^*$

$\{w \mid w \text{ tem pelo menos um símbolo } 1\}.$

$1^*(01^+)^*$

$\{w \mid \text{todo } 0 \text{ em } w \text{ é seguido por pelo menos um } 1\}.$

$(\Sigma\Sigma)^*$

$\{w \mid w \text{ é uma cadeia de comprimento par}\}.$

$01 \cup 10$

$\{01, 10\}$

Expressões Regulares: Exemplos onde $\Sigma=\{0,1\}$

$$0\Sigma^*0 \cup 1\Sigma^*1 \cup 0 \cup 1$$

$\{w \mid w \text{ começa e termina com o mesmo símbolo}\}.$

$$(0 \cup \varepsilon)1^*$$

$$01^* \cup 1^*$$

$$(0 \cup \varepsilon)(1 \cup \varepsilon)$$

$$\{\varepsilon, 0, 1, 01\}.$$

$$1^* \emptyset$$

$$\emptyset$$

$$\emptyset^*$$

$$\{\varepsilon\}$$

Expressões Regulares: Identidades

$$R \cup \emptyset$$

R

$$R \circ \varepsilon$$

R

$$R \cup \{\varepsilon\}$$

pode ser diferente de R

$$R \circ \emptyset$$

É sempre $\emptyset$

Expressões Regulares

Expressões regulares são ferramentas úteis no desenho de compiladores para linguagens de programação.

Objetos elementares, *tokens*, tais como os nomes de variáveis e constantes podem ser descritos com expressões regulares.

A partir da sintaxe dos tokens, sistemas automáticos podem gerar analisadores léxicos, a parte do compilador que inicialmente processa o programa de entrada.

Expressões Regulares: Exemplo

Uma constante numérica que pode incluir uma parte fracionária e/ou um sinal pode ser descrita como um membro da linguagem:

$$(+ \cup - \cup \varepsilon) (D^+ \cup D^+ . D^* \cup D^* . D^+)$$

Onde $D = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

Cadeias geradas: 72, 3.14159, +7. e -.01 .