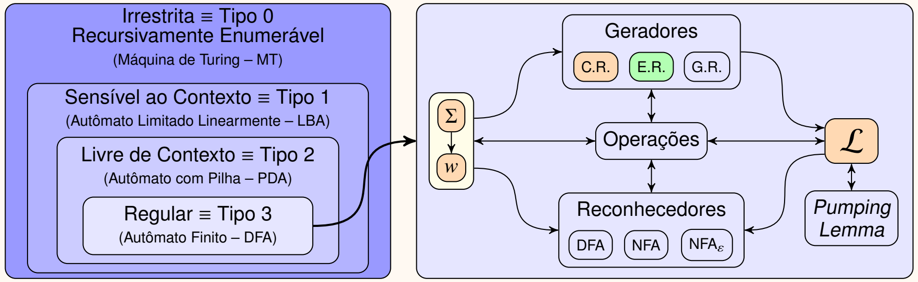
Linguagens Formais e Autômatos

Humberto Longo

Instituto de Informática
Universidade Federal de Goiás

Bacharelado em Ciência da Computação, 2021/1

Roteiro



Conjuntos e expressões regulares

- Expressões regulares são usadas para abreviar a descrição de conjuntos regulares:
  - O conjunto regular $\{a\}$ é representado por a.
  - As operações de união, fecho de Kleene e concatenação são designadas por $\cup$, * e justaposição, respectivamente.
  - Parênteses são usados para indicar a precedência dos operadores.

Expressões regulares

Definição 1.30

- Uma expressão regular sobre um alfabeto Σ é definida como:

Base: $\emptyset$, ε e a, para todo $a \in \Sigma$, são expressões regulares sobre Σ.

Recursão: Se u e v são expressões regulares sobre Σ, então $(u \cup v)$, (uv) e (u^*) também são expressões regulares sobre Σ.

Fecho: u é uma expressão regular sobre Σ se pode ser obtida, a partir das expressões regulares básicas, com a aplicação da recursão um número finito de vezes.

Linguagem de uma expressão regular

- Se u é uma expressão regular, então a linguagem gerada por u é denotada por $\mathcal{L}(u)$.

Expressões regulares

- Uma expressão regular define um padrão e uma cadeia pertence à linguagem gerada pela expressão se está de acordo com o padrão definido.
  - A concatenação especifica uma ordem relativa entre dois elementos.
  - O fecho de Kleene permite repetições.
  - A operação $\cup$ permite seleção.

Expressões regulares

- Precedência dos operadores regulares:
  1. Fecho de Kleene.
  2. Concatenação e união.
- Parênteses podem ser eliminados de expressões que são sequências de uniões ou concatenações.
  - União e concatenação são operações associativas.

Exemplo 1.31

- $(a \cup b)^* bb(a \cup b)^*$: expressão regular para o conjunto $\{a, b\}^* \circ \{bb\} \circ \{a, b\}^*$.
- $a(a \cup b)^* b(a \cup b)^* a$: expressão regular para o conjunto $\{a\} \circ \{a, b\}^* \circ \{b\} \circ \{a, b\}^* \circ \{a\}$.

Expressões regulares

Exemplo 1.32

- Expressões regulares especificam um padrão que as cadeias de uma determinada linguagem devem satisfazer.
- Em muitas situações são usadas notações diferentes, ou mesmo terminologias, mas o conceito básico permanece o mesmo:
  1. Nomes de usuários: $\wedge[a{-}z0{-}9\_{-}]\{3, 16\}\$$.
  2. Endereços de emails: $\wedge([a{-}z0{-}9\_\backslash.{-}]+)@([a{-}z0{-}9\_\backslash.{-}]+)\backslash.([a{-}z\backslash.]2, 6)\$$.

Expressões regulares

Definição 1.33

$$\begin{aligned} u^+ &= uu^*. \\ u^2 &= uu. \\ u^3 &= u^2u. \\ &\vdots \\ u^n &= u^{n-1}u. \end{aligned}$$

Linguagens de expressões regulares

Resumo

Expressão regular	Linguagem
$\emptyset$	$\mathcal{L}(\emptyset) = \{\} = \emptyset$
ε	$\mathcal{L}(\varepsilon) = \{\varepsilon\}$
$a,\ a \in \Sigma$	$\mathcal{L}(a) = \{a\}$
u	$\mathcal{L}(u)$
$u \cup v$	$\mathcal{L}(u \cup v) = \mathcal{L}(u) \cup \mathcal{L}(v)$
uv	$\mathcal{L}(uv) = \mathcal{L}(u)\circ\mathcal{L}(v)$
u^*	$\mathcal{L}(u^*) = (\mathcal{L}(u))^*$

Observações

- A notação $(u + v)$ também é usada para denotar a união $(u \cup v)$.
- Alguns parênteses podem ser eliminados. Em cada caso, a interpretação é feita de acordo com a ordem de precedência: fecho de Kleene, concatenação e, por último, união.
- $u^+ = uu^*$ e u^k, para k concatenações sucessivas, são notações válidas.

Linguagens de expressões regulares

Exemplo 1.34

Expressão regular	Linguagem
$a \cup b$	$\mathcal{L}(a \cup b) = \{a\} \cup \{b\} =$ $\{a, b\}$
a^*	$\mathcal{L}(a^*) = (\mathcal{L}(a))^* = (\{a\})^* =$ $\{\varepsilon, a, aa, aaa, aaaa, \dots\}$
$(a \cup b)(a \cup b)$	$\mathcal{L}((a \cup b)(a \cup b)) = \mathcal{L}((a \cup b))\circ\mathcal{L}((a \cup b)) = \{a, b\}\circ\{a, b\} =$ $\{aa, ab, ba, bb\}$
$a \cup (ab)^*$	$\mathcal{L}(a \cup (ab)^*) = \mathcal{L}(a) \cup \mathcal{L}((ab)^*) = \{a\} \cup \{\varepsilon, ab, abab, \dots\}$ $\{a, \varepsilon, ab, abab, ababab, \dots\}$

Linguagens de expressões regulares

Exemplo 1.35

- Seja o alfabeto $\Sigma = \{a, b\}$:

Expressão regular	Linguagem
$a \cup b$	$\{a, b\} = \Sigma$.
$(a \cup b)^*$	Todas as palavras sobre Σ (Σ^*).
$a(a \cup b)^*$	Todas as palavras que começam com a.
$b^*(a \cup \varepsilon)b^*$	Todas as palavras que contêm zero ou um a.
$a(\varepsilon \cup b)^*$	$\{a, ab, abb, abbb, \dots\}$.
$((a \cup b)(a \cup b))^*$	Todas as palavras de comprimento par.
$(a^*b)^*a^*$	Σ^*.

Expressões regulares

Exemplo 1.36

- O conjunto $\{bawab \mid w \in \{a, b\}^*\}$, sobre o alfabeto $\Sigma = \{a, b\}$ é regular:

	Conjunto	Expressão	Justificativa
1.	$\{a\}$	a	Base
2.	$\{b\}$	b	Base
3.	$\{a\}\circ\{b\} = \{ab\}$	ab	$1, 2$, Concatenação
4.	$\{a\} \cup \{b\} = \{a, b\}$	$a \cup b$	$1, 2$, União
5.	$\{b\}\circ\{a\} = \{ba\}$	ba	$2, 1$, Concatenação
6.	$\{a, b\}^*$	$(a \cup b)^*$	4, Fecho de Kleene
7.	$\{ba\}\circ\{a, b\}^*$	$ba(a \cup b)^*$	$5, 6$, Concatenação
8.	$\{ba\}\circ\{a, b\}^*\circ\{ab\}$	$ba(a \cup b)^*ab$	$7, 3$, Concatenação

Expressões regulares

Exemplo 1.37

- Alfabeto: $\Sigma = \{a, b\}$.
- Cadeias que contêm aa: $(a \cup b)^*aa(a \cup b)^*$.
- Cadeias que contêm bb: $(a \cup b)^*bb(a \cup b)^*$.
- Cadeias que contêm aa ou bb: $(a \cup b)^*aa(a \cup b)^* \cup (a \cup b)^*bb(a \cup b)^*$.

Exemplo 1.38

- Expressões regulares que definem o conjunto de cadeias, sobre o alfabeto $\Sigma = \{a, b\}$, que contêm pelo menos 2 b's:
  1. $a^*ba^*b(a \cup b)^*$.
  2. $(a \cup b)^*ba^*ba^*$.
  3. $(a \cup b)^*b(a \cup b)^*b(a \cup b)^*$.

Expressões regulares

Exemplo 1.39

- Expressão regular para o conjunto de cadeias, sobre o alfabeto $\Sigma = \{a, b\}$, que contêm exatamente 2 b's:
  - $a^*ba^*ba^*$.

Exemplo 1.40

- Expressão regular para o conjunto de cadeias, sobre o alfabeto $\Sigma = \{a, b\}$, que contêm número par de b's:
  - $a^*(a^*ba^*ba^*)^*$.
  - $a^*(ba^*ba^*)^*$.

Expressões regulares

Definição 1.41

- Duas expressões regulares r e s são equivalentes se $\mathcal{L}(r) = \mathcal{L}(s)$.

- Pode não existir uma única expressão regular para definir um conjunto.
- Duas expressões que representam o mesmo conjunto são ditas equivalentes.
- Expressões regulares podem ser manipuladas algebricamente para a construção de expressões equivalentes.

Expressões regulares

Definição 1.42 (Identidades Básicas)

$$
\begin{array}{rl}
1. & \emptyset u = u\emptyset = \emptyset \\
2. & \varepsilon u = u\varepsilon = u \\
3. & \emptyset^* = \varepsilon \\
4. & \varepsilon^* = \varepsilon \\
5. & u \cup v = v \cup u \\
6. & u \cup \emptyset = u \\
7. & u \cup u = u \\
8. & u^* = (u^*)^* \\
9. & u(v \cup w) = uv \cup uw \\
10. & (u \cup v)w = uw \cup vw \\
11. & (uv)^*u = u(vu)^* \\
12. & (u \cup v)^* = (u^* \cup v)^* \\
 & = u^*(u \cup v)^* = (u \cup vu^*)^* \\
 & = (u^*v^*)^* = u^*(vu^*)^* \\
 & = (u^*v)^*u^*
\end{array}
$$

Expressões regulares

Exemplo 1.43

- Expressão regular para o conjunto de cadeias, sobre o alfabeto $\Sigma = \{a, b\}$, que não contêm a subcadeia aa.
  - Uma cadeia pode conter um prefixo com qualquer número de b's;
  - Todo a deve ser seguido por pelo menos um b ou terminar a cadeia.

$$\begin{aligned} & b^*(ab^+)^* \cup b^*(ab^+)^*a \\ = \ & b^*(ab^+)^*(\varepsilon \cup a) \\ = \ & b^*(abb^*)^*(\varepsilon \cup a) \\ = \ & (b \cup ab)^*(\varepsilon \cup a) \end{aligned}$$

Expressões regulares

Exemplo 1.44

- Expressão regular para o conjunto de cadeias, sobre o alfabeto $\Sigma = \{a, b, c\}$, que contêm a subcadeia bc:
  - $(a \cup b \cup c)^*bc(a \cup b \cup c)^*$.

Exemplo 1.45

- Linguagem $\mathcal{L}$ definida por $c^*(b \cup ac^*)^*$:
  - c^* e ac^* garante que a's e b's podem ocorrer em qualquer ordem.
  - Quando ocorre um elemento de ac^*, este pode ser seguido por qualquer número de a's ou b's (em qualquer ordem).
  - $\mathcal{L}$ não contém cadeias com a subcadeia bc.

Exemplo 1.46

- Expressão regular para o conjunto de cadeias, sobre o alfabeto $\Sigma = \{a, b\}$, que não contêm a subcadeia abb:
  - $b^* \cup b^*a(a \cup ba)^*(b \cup \varepsilon) \equiv b^*(a^+b)^*a^*$.

Expressões regulares

Exemplo 1.47

Expressão regular	**Conjunto regular**
a^*ba^*	$\{w \in \{a,b\}^* \mid w$ contém apenas um símbolo $b\}$
$(a \cup b)^*b(a \cup b)^*$	$\{w \in \{a,b\}^* \mid w$ contém pelo menos um símbolo $b\}$
$a(a \cup b)^*a \cup b(a \cup b)^*b \cup a \cup b$	$\{w \in \{a,b\}^* \mid w$ começa e termina com o mesmo símbolo$\}$
$(a^*ba^*b)^*a^*$	$\{w \in \{a,b\}^* \mid$ número de ocorrências do símbolo b é par$\}$
$(a \cup b)^*aa(a \cup b)^*$	$\{w \in \{a,b\}^* \mid w$ contém pelo menos um par de a's consecutivos$\}$
$(a \cup b)^*b(a \cup b)(a \cup b)(a \cup b)$	$\{w \in \{a,b\}^* \mid w$ contém um b no quarto símbolo a partir do final$\}$
$(b^*abb^*)^*(a \cup \varepsilon) \cup b^*$	$\{w \in \{a,b\}^* \mid w$ não contém um par de a's consecutivos$\}$
$(a \cup b)^*(aa \cup ba \cup bb) \cup a \cup b \cup \varepsilon$	$\{w \in \{a,b\}^* \mid w$ não termina com $ab\}$

Expressões regulares

Exemplo 1.48

- Expressão regular para o conjunto de cadeias, sobre o alfabeto $\Sigma = \{a, b, c\}$, que contêm todos os símbolos pelo menos uma vez:

$$\begin{aligned} & (a \cup b \cup c)^*a(a \cup b \cup c)^*b(a \cup b \cup c)^*c(a \cup b \cup c)^* \cup \\ & (a \cup b \cup c)^*a(a \cup b \cup c)^*c(a \cup b \cup c)^*b(a \cup b \cup c)^* \cup \\ & (a \cup b \cup c)^*b(a \cup b \cup c)^*a(a \cup b \cup c)^*c(a \cup b \cup c)^* \cup \\ & (a \cup b \cup c)^*b(a \cup b \cup c)^*c(a \cup b \cup c)^*a(a \cup b \cup c)^* \cup \\ & (a \cup b \cup c)^*c(a \cup b \cup c)^*a(a \cup b \cup c)^*b(a \cup b \cup c)^* \cup \\ & (a \cup b \cup c)^*c(a \cup b \cup c)^*b(a \cup b \cup c)^*a(a \cup b \cup c)^* \end{aligned}$$

Expressões regulares

Teorema 1.49

- *Existem linguagens que não podem ser definidas por expressões regulares.*

Demonstração.

No decorrer do curso □

Algumas questões a serem respondidas

- Dada uma palavra w e uma expressão regular e, existe um algoritmo que decida se $w \in \mathcal{L}(e)$?
- Dadas expressões regulares e_1 e e_2, sobre o mesmo alfabeto, existe um algoritmo que decida se $\mathcal{L}(e_1) = \mathcal{L}(e_2)$?
- Todas as linguagens são regulares? Caso a resposta seja negativa, como provar que uma linguagem $\mathcal{L}$ não é regular?

Livros texto

R. P. Grimaldi
Discrete and Combinatorial Mathematics – An Applied Introduction.
Addison Wesley, 1994.

D. J. Velleman
How To Prove It – A Structured Approach.
Cambridge University Press, 1996.

J. E. Hopcroft; J. Ullman.
Introdução Ã Teoria de Autômatos, Linguagens e Computação.
Ed. Campus.

T. A. Sudkamp.
Languages and Machines – An Introduction to the Theory of Computer Science.
Addison Wesley Longman, Inc. 1998.

J. Carroll; D. Long.
Theory of Finite Automata – With an Introduction to Formal Languages.
Prentice-Hall, 1989.

M. Sipser.
Introduction to the Theory of Computation.
PWS Publishing Company, 1997.

H. R. Lewis; C. H. Papadimitriou
Elementos de Teoria da Computação.
Bookman, 2000.