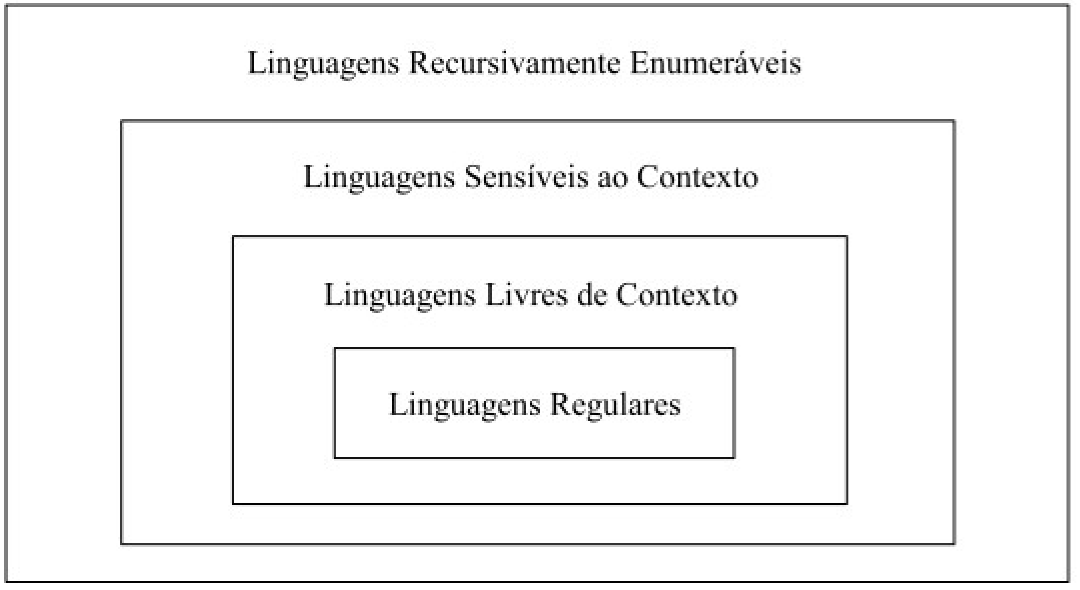
Linguagens Regulares

Introdução

- Hierarquia de Chomsky



Introdução

- Conjuntos e expressões regulares são notações alternativas utilizadas para representar essa classe de linguagens
- A classe mais restrita dentro da Hierarquia de Chomsky

Definição

- Conjuntos regulares sobre um alfabeto finito Σ são linguagens definidas recursivamente da seguinte forma:

  - $\emptyset$ é um conjunto regular sobre Σ;
  - $\{ \sigma \}$, $\forall \sigma \in \Sigma$, é um conjunto regular sobre Σ.

Definição

- Se X e Y são conjuntos regulares sobre Σ, então também são conjuntos regulares sobre Σ:

  - $X \cup Y$;
  - $X \cdot Y$, também denotado XY;
  - X^*

Exemplo 1

- Seja $L = \{0^m 1^n \mid m \geq 0, n \geq 0\}$ sobre $\Sigma = \{0, 1\}$
  - concatenação de um número arbitrário de símbolos “0” (incluindo nenhum) com um número também arbitrário de símbolos “1” (incluindo nenhum)
  - $L = \{ \varepsilon , 0, 1, 00, 01, 11, \ldots\}$

- Linguagem é regular?

Exemplo 1

- Considerem-se as linguagens sobre o mesmo Σ
  - $L1 = \{0\}$
  - $L2 = \{1\}$
  - $L3 = \{0^i \mid i \geq 0\}$
  - $L4 = \{1^i \mid i \geq 0\}$
  - $L5 = \{0^p 1^q \mid p \geq 0, q \geq 0\}$

Exemplo 1

- L1 e L2 são conjuntos regulares sobre Σ, por definição
- L3 e L4 são obtidos a partir de L1 e L2
  - $L3 = L1^*$
  - $L4 = L2^*$
- L5 = L pode ser expresso pela concatenação dos conjuntos L3 e L4

Exemplo 2

- A linguagem ℕ formada pelos números naturais decimais é um conjunto regular sobre o alfabeto dos algarismos arábicos
- Seja D = {0, 1, 2, 3, 4, 5, 6, 7, 8, 9}
- então

$$\mathbb{N} = DD^*$$

Expressões regulares

- Visa obter maior concisão e facilidade de manipulação
  - Desenvolvido por Kleene na década 1950
  - eliminação do uso dos símbolos “{” e “}”, bem como a substituição do símbolo de união (“∪”) por um símbolo “+” ou “|”

Precedência	**Operador**	**Representação**
Mais alta	Fechamento	x^*
Intermediária	Concatenação	$x \cdot y$ ou xy
Mais baixa	União	$x \mid y$ ou $x + y$

Expressões Regulares

- Exemplos
  - $(ab \mid c^{*}) = ((ab) \mid c^{*}) = ((ab) \mid (c^{*}))$
    - representa o conjunto {ab, ε , c, cc, ccc...}
  - $a(b \mid c)^{*}$
    - representa o conjunto {a, ab, ac, abc, abb, acc, …}
  - $(ab \mid c)^{*}$
    - representa o conjunto { ε , ab, c, abc, cab, abab, cc, ...}

Expressões Regulares

- Abreviação
  - Uma abreviação muito comum consiste na substituição da expressão regular xx^* por x^+

- $L = \{0^m 1^n \mid m \geq 0, n \geq 0\}$ pode ser reescrita como $((0)^* (1)^*)$, ou simplesmente, 0^*1^*
- Para $m \geq 0$ e $n \geq 1$, a expressão seria 0^*11^*
- $0^*11^* = 0^*1^*1 = 0^*1^+$

Exercícios

- Obter expressões regulares que representam as linguagens cujas sentenças estão descritas a seguir

Exercícios

1. Começam com aa;
2. Não começam com aa;
3. Terminam com bbb;
4. Não terminam com bbb;
5. Contém a subcadeia aabbb;
6. Possuem comprimento maior ou igual a 3;
7. Possuem comprimento menor ou igual a 3;
8. Possuem comprimento par;
9. Possuem comprimento ímpar;
10. Possuem quantidade par de símbolos a;
11. Possuem quantidade ímpar de símbolos b.