

# Linguagens

# Linguagens

- Símbolos e Cadeias
- Linguagens
- Gramáticas

# Símbolos

- Os símbolos são representações gráficas, indivisíveis, empregadas na construção de cadeias
  - São exemplos de símbolos: a, abc, begin, if , 5, 1024, 2.017e4.
  - Não há uma definição formal para “símbolo”.
    - Entidade abstrata
    - Pode-se dizer que se trata de um conceito primitivo.

# Cadeia

- String, palavra ou cadeia
  - Seqüência finita de símbolos do alfabeto
- Comprimento
  - O comprimento de uma cadeia é um número natural que designa a quantidade de símbolos que a compõem. O comprimento de uma cadeia  $\alpha$  é denotado por  $|\alpha|$ .

Considerem-se as cadeias  $\alpha = 1$ ,  $\beta = 469$ ,  $\chi = bce60$  e  $\varphi = df$

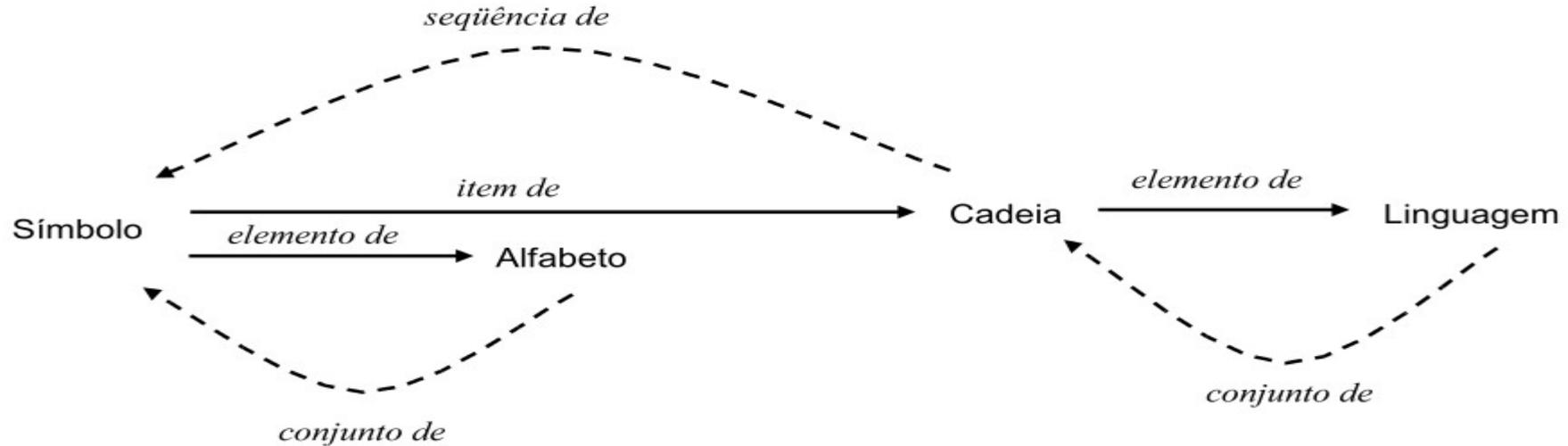
Então,  $|\alpha| = 1$ ,  $|\beta| = 3$ ,  $|\chi| = 5$  e  $|\varphi| = 2$ .

# Alfabeto

- Alfabeto
  - Conjunto de símbolos (finito e não-vazio)
  - $\Sigma = \{0,1\}$ , o alfabeto binário
  - $\Sigma = \{a,b,c,\dots,z\}$ , conjunto de letras minúsculas
  - Conjunto de caracteres ASCII
- $\Sigma^*$  é o conjunto de todas as cadeias possíveis de serem construídas sobre  $\Sigma$

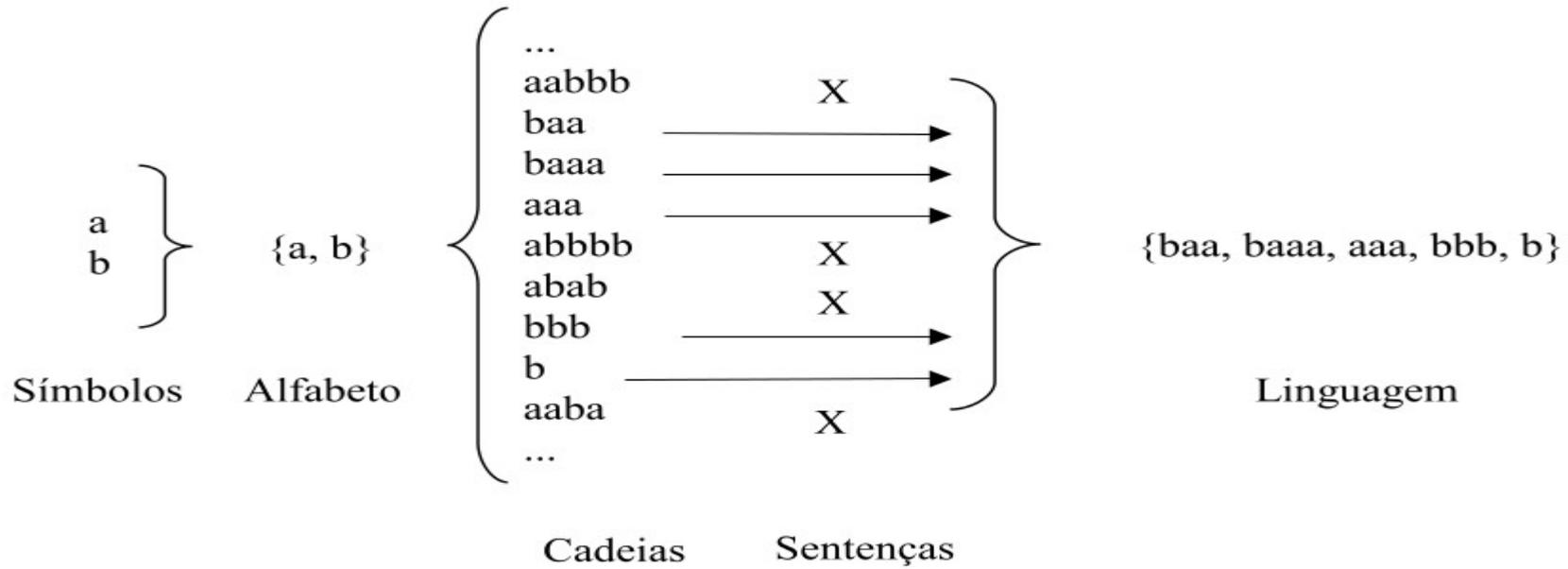
# Linguagem

- Uma linguagem formal é um conjunto, finito ou infinito, de cadeias de comprimento finito, formadas pela concatenação de elementos de um alfabeto finito e não-vazio.



# Linguagens

- Exemplo:



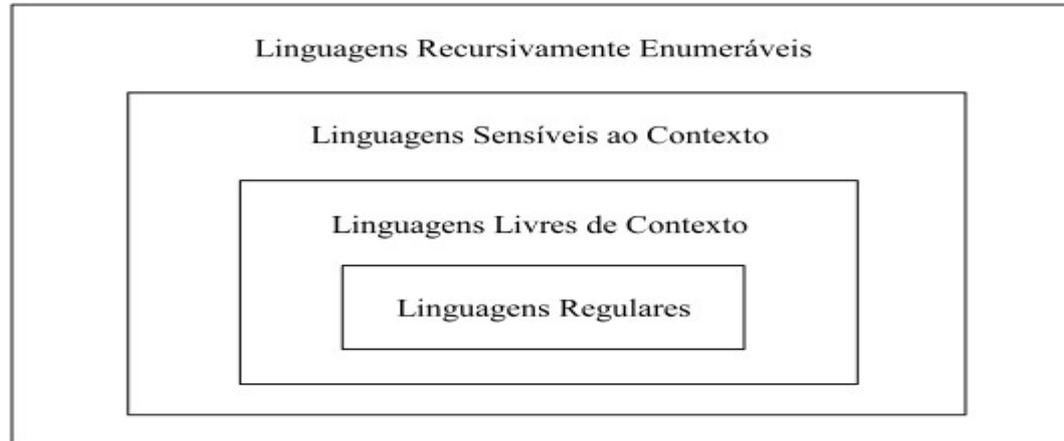
# Linguagens

- Linguagens formais de interesse, em sua maioria, contêm, se não uma quantidade infinita, ao menos um número finito, porém muito grande, de cadeias.
- Há um interesse muito grande em relação a métodos que permitam especificar linguagens, sejam elas finitas ou não, através de representações finitas.

# 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

# 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 (“U”) por um símbolo “+” ou “|”

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

# Expressões Regulares

- Exemplos

- $(ab \mid c^*) = ((ab) \mid c^*) = ((ab) \mid (c^*))$

- representa o conjunto  $\{ab, \varepsilon, c, cc, ccc, \dots\}$

- $a(b \mid c)^*$

- representa o conjunto  $\{a, ab, ac, abc, abb, acc, \dots\}$

- $(ab \mid c)^*$

- representa o conjunto  $\{\varepsilon, ab, c, abc, cab, abab, cc, \dots\}$

# 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.