

Linguagens Formais e Autômatos

Prof. Me. Stéfano Schwenck Borges Vale Vita

Material adaptado UEM-PR:

Prof. Yandre Maldonado e Gomes da Costa



Introdução

- Problema: definir um conjunto de cadeias de símbolos;
- Exemplo: conjunto M dos números binários que têm 2 dígitos
 - $M=\{00,01,10,11\}$



Introdução

- Representação clara:
 - humanos x computador
- Representação Formal → Computador;
- Um objetivo de LFA é estudar uma maneira **precisa e formal** de descrever seqüências de símbolos pertencentes à um determinado conjunto;



Introdução

- Em especial conjuntos que **não podem ser trivialmente enumerados**;
- Os estudos iniciais foram em torno de Linguagens Naturais (LN);
- Algumas características de LN introduziram dificuldades no tratamento computacional das mesmas:
 - LN é extensa, complexa, não tem sintaxe rígida e semântica bem determinada (rica em ambigüidade);



Introdução

- Entretanto, os estudos iniciais apresentaram resultados significativos na descrição de linguagens computacionais;
- Linguagens Computacionais são muito mais simples, são linguagens criadas artificialmente de forma que possam ser tratadas computacionalmente de forma satisfatória;



Introdução

- As maneiras sistemáticas de descrever uma linguagem de programação são:
 - um método que permite construir programas sintaticamente corretos - geração (Gramática);
 - um método que permite verificar se um programa escrito está sintaticamente correto - reconhecimento (Autômatos);



Conceitos Básicos

- Alfabeto:
 - Conjunto de finito de símbolos;
 - Normalmente descrito por Σ ;
 - Exemplos:
 - $\Sigma=\{a, b\}$
 - $\Sigma=\{1, 2, 3\}$
 - $\Sigma=\{00, 11\}$



Conceitos Básicos

- Símbolo ou letra:
 - é todo elemento pertencente à um alfabeto;
 - a é um símbolo de Σ sse $a \in \Sigma$;
- Exemplo: dado o alfabeto $\Sigma = \{0, 1, 23\}$
 - 0 é um símbolo de Σ ;
 - 1 é um símbolo de Σ ;
 - 23 é um símbolo de Σ ;



Conceitos Básicos

- Essas duas primeiras definições são bastante livres. Embora os símbolos também possam ser chamados de letras, eles não precisam ter necessariamente um único caractere. E além disso, os símbolos de um alfabeto não precisam todos ter o mesmo número de caracteres. A única restrição é que o tamanho do símbolo seja finito.



Conceitos Básicos

- Cadeia ou palavra:
 - É uma concatenação de símbolos de um mesmo alfabeto;
 - É uma seqüência finita de símbolos do alfabeto justapostos;
 - Assim, dado um alfabeto Σ e uma seqüência de símbolos $x = a_1 a_2 a_3 \dots a_n$, x é uma cadeia sobre Σ sse $a_i \in \Sigma$ para $i = 1, 2, \dots, n$



Conceitos Básicos

○ Comprimento de Cadeia ou Tamanho de Palavra:

- É o número de símbolos que compõem uma dada cadeia (ou palavra).
- O comprimento de uma cadeia x é denotado por $|x|$
- Então, a cadeia $x=a_1a_2a_3...a_n$, tem seu comprimento $|x| = n$
- Cadeia nula ou palavra vazia: é um caso especial, ela é denotada por λ (ou ε) e tem tamanho igual a zero.



Conceitos Básicos

- Exercício: dado o alfabeto $\Sigma=\{a, b, c, de\}$, verifique se as cadeias a seguir são formadas sobre este alfabeto, e se for, verifique qual o comprimento das mesmas:
 - $x = ababac$
 - $y = abdec$
 - $z = abedc$
 - $w = abdceaba$
 - $s = d$
 - $t = a$



Conceitos Básicos

- Exponenciação de Alfabetos: Σ^k é o conjunto de todas as cadeias com tamanho k , formadas sobre o alfabeto Σ .
 - Exemplo: considere $\Sigma = \{0, 1\}$
 - $\Sigma^0 = \{\lambda\}$
 - $\Sigma^1 = \{0, 1\} = \Sigma$
 - $\Sigma^2 = \{00, 01, 10, 11\}$
 - ...
- **Exercício:** encontre Σ^3 para o exemplo anterior.



Conceitos Básicos

- Fechamento de um Alfabeto: Seja Σ um alfabeto, então o fechamento de Σ , descrito por Σ^* é definido como

$$\Sigma^* = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \dots \cup \Sigma^n \cup \dots$$

- Σ^* é o conjunto de todas as cadeias possíveis de se formar sobre o alfabeto Σ .
- Fechamento positivo $\Sigma^+ = \Sigma^* - \{\lambda\}$



Conceitos Básicos

- Concatenação de cadeias: dado o alfabeto Σ e as cadeias $x, y \in \Sigma^*$, a concatenação de x e y , indicada por xy , produz uma cadeia formada pelos símbolos de x seguidos pelos símbolos de y .
- Se $x = a_1a_2\dots a_n \in \Sigma^*$ e $y = b_1b_2\dots b_m \in \Sigma^*$, então $xy = a_1a_2\dots a_nb_1b_2\dots b_m$



Conceitos Básicos

- Exemplos:

$$\Sigma = \{a, b\}$$

$$x = abaa, y = ba, z = \lambda$$

$$xy = abaaba$$

$$yx = baabaa$$

$$yz = ba = zy = y$$

- A cadeia nula (λ) é o elemento neutro da concatenação.



Conceitos Básicos

- Concatenação sucessiva:
concatenação de uma palavra com ela mesma;
 - Representada através de um expoente: w^n
 - Onde w é uma palavra e n indica o número de concatenações sucessivas;

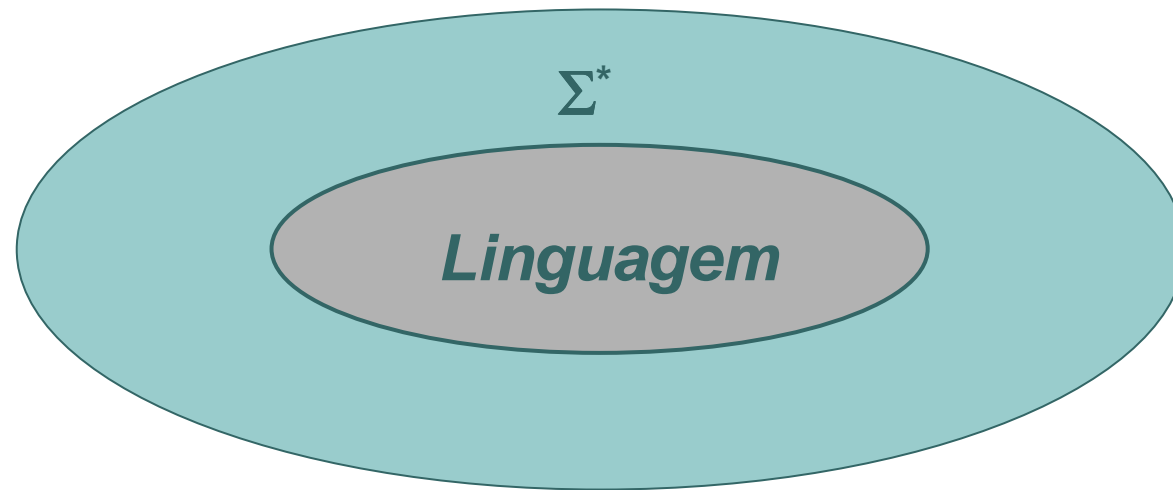


Conceitos Básicos

- Dado um alfabeto Σ e $x, y \in \Sigma^*$, diz-se que:
 - x é um **prefixo** de y sse $\exists w \in \Sigma^*$ tal que $y = xw$;
 - x é um **sufixo** de y sse $\exists w \in \Sigma^*$ tal que $y = wx$;
 - x é um **subpalavra** de y sse $\exists w, u \in \Sigma^*$ tal que $y = wxu$;

● ● ● | Conceitos Básicos

- Linguagem: dado o alfabeto Σ , o conjunto de palavras L é uma linguagem sobre Σ , sse $L \subset \Sigma^*$.





Conceitos Básicos

- Operações sobre linguagens;
- Considere L_1 e L_2 linguagens definidas sobre Σ :
 - União: $L_1 \cup L_2 = \{x \mid x \in L_1 \vee x \in L_2\}$
 - Intersecção: $L_1 \cap L_2 = \{x \mid x \in L_1 \wedge x \in L_2\}$
 - Diferença: $L_1 - L_2 = \{x \mid x \in L_1 \wedge x \notin L_2\}$
 - Concatenação: $L_1.L_2 = \{x \mid x = yz, y \in L_1 \wedge z \in L_2\}$
 - Complemento: $\overline{L_1} = \{x \mid x \in \Sigma^* \wedge x \notin L_1\}$



Conceitos Básicos

- Exemplos de operações:
 - Sejam L_1 e L_2 definidas sobre $\{0,1\}$:
 - $L_1 = \{0, 11\}$
 - $L_2 = \{0, 1, 00\}$
 - $L_1 \cup L_2 = \{0, 1, 00, 11\}$
 - $L_1 \cap L_2 = \{0\}$
 - $L_1 - L_2 = \{11\}$
 - $L_1 \cdot L_2 = \{00, 01, 000, 110, 111, 1100\}$



Conceitos Básicos

- Comparando as definições:
 - **Linguagem Natural:**
 - Uma palavra em português equivale à um símbolo;
 - Uma sentença da língua portuguesa é uma cadeia composta por vários símbolos;
 - **Linguagem Computacional:**
 - Cada programa escrito numa linguagem computacional corresponde a uma cadeia de símbolos que podem ser:
 - identificadores;
 - palavras reservadas;
 - símbolos especiais e operadores;
 - constantes numéricas.

● ● ● | Conceitos Básicos

- Uma linguagem computacional como linguagem formal:

└ Alfabeto da linguagem Pascal

{program, var, integer, real, char, begin, end, if, then, else, for,..., ;, “,”, :, :=, ., ...}



O código fonte de um programa corresponde à uma cadeia formada a partir de símbolos do alfabeto.

```
Program Teste;  
Var  
  i: integer;  
Begin  
  i:=0;  
End.
```



LINGUAGEM
Conjunto de todas as cadeias
descritas a partir do alfabeto que
respeitam um conjunto de regras
sintáticas.

● ● ● | Conceitos Básicos

○ Hierarquia de Chomsky

