



LFA0001 – Linguagens Formais e Autômatos

Aula 02

Linguagens e Gramáticas

Karina Girardi Roggia
karina.roggia@udesc.br

Departamento de Ciência da Computação
Centro de Ciências Tecnológicas
Universidade do Estado de Santa Catarina

2016



Sumário

Introdução

Alfabeto

Palavra

Linguagem Formal

Gramática





Linguagens e Gramáticas

Linguagem: Dicionário Aurélio

o uso da palavra articulada ou escrita como meio de expressão e comunicação entre pessoas

Como utilizar o conceito para linguagens de programação?

- definição matemática formal

Para isto, é necessário definir:

- alfabeto
- palavra (ou cadeia)



Alfabeto

Símbolo ou Caractere

- entidade abstrata básica, não definida formalmente
- base para as demais definições a serem usadas
- *exemplos*: letras e dígitos

Definição (Alfabeto)

Um alfabeto Σ é um conjunto **finito** de símbolos ou caracteres.

Portanto

- conjunto infinito **não** é alfabeto
- \emptyset é um alfabeto



Alfabeto

São alfabetos

- $\{a, b, c\}$
- \emptyset
- $\{\text{☺}, \text{©}, \text{♥}, \text{★}\}$
- $\{0, 1\}$

Não são alfabetos

- \mathbb{Z}
- $\{0, 1, 2, 3, 4, \dots\}$
- $\{a, aa, aaa, aaaa, \dots\}$
- $[0, 1]$



Exemplo: Alfabeto para uma Linguagem de Programação

Conjunto de todos os símbolos usados nos programas

- letras
- dígitos
- caracteres especiais como $=, +, *, \dots$
- espaço ("branco")

Atenção: cada palavra reservada é considerada **um símbolo**



Palavra

Definição (Palavra)

Uma palavra sobre um alfabeto Σ é uma sequência finita de símbolos justapostos.

Palavra Vazia (ou cadeia vazia)

- é a cadeia sem símbolos
- palavra de tamanho zero
- representada pelo símbolo ϵ



Prefixo, Sufixo, Subpalavra

Prefixo (sufixo)

- qualquer sequência inicial (final) de símbolos da palavra

Subpalavra

- qualquer sequência de símbolos contíguos da palavra

Exemplo: seja *abcb* uma palavra sobre o alfabeto $\{a, b, c\}$

- ϵ , *a*, *ab*, *abc*, *abcb* são os prefixos
- ϵ , *b*, *cb*, *bcb*, *abcb* são os sufixos
- qualquer prefixo ou sufixo, e as palavras *bc* e *c*, são subpalavras



Concatenação

Definição (Concatenação)

A concatenação de duas palavras é uma operação que associa a cada par de palavras uma outra, formada pela justaposição da primeira com a segunda.

Usualmente denotada pelo operador \cdot , ou omitido.

Propriedades:

- Elemento neutro: $\varepsilon \cdot w = w = w\varepsilon$
- Associativa: $v(wt) = (vw)t = vwt$



Concatenação

Exemplo:

Seja $\Sigma = \{a, b\}$ um alfabeto e as palavras $v = baaaa$ e $w = bb$

- $vw = baaaabb$
- $v\varepsilon = v = baaaa$

Definição (Concatenação Sucessiva)

A concatenação sucessiva de uma palavra (com ela própria) w^n é definida indutivamente como se segue, a partir da operação de concatenação:

- $w^0 = \varepsilon$
- $w^n = w \cdot w^{n-1}$ para $n > 0$



Conjunto de Todas as Palavras

Se Σ é um alfabeto

- Σ^* é o conjunto de **todas** as palavras sobre Σ
- $\Sigma^+ = \Sigma^* - \{\varepsilon\}$

Definição (Conj. Todas as Palavras)

Seja Σ um alfabeto. Então Σ^* é indutivamente definido como se segue:

- $\varepsilon \in \Sigma^*$
- para qualquer $x \in \Sigma$, tem-se que $x \in \Sigma^*$
- para quaisquer $u, v \in \Sigma^*$, tem-se que $uv \in \Sigma^*$



Palavras

Portanto, uma palavra sobre um alfabeto Σ é qualquer elemento de Σ^*

Exemplo: Conjunto de todas as palavras. Seja $\Sigma = \{a, b\}$, então:

- $\Sigma^+ = \{a, b, aa, ab, ba, bb, aaa, \dots\}$
- $\Sigma^* = \{\epsilon, a, b, aa, ab, ba, bb, aaa, \dots\}$



Tamanho de uma palavra

O comprimento (ou tamanho) de uma palavra w , representado por $|w|$, é o número de símbolos pertencentes ao alfabeto que compõem a palavra.

Definição (Tamanho de uma Palavra)

O tamanho de uma palavra sobre o alfabeto Σ é uma função $|\cdot| : \Sigma^* \rightarrow \mathbb{N}$ indutivamente definida como segue:

- $|\varepsilon| = 0$
- $|x| = 1$ para todo $x \in \Sigma$
- $|xv| = 1 + |v|$ para todo $x \in \Sigma$ e $v \in \Sigma^*$



Linguagem Formal

Definição (Linguagem Formal)

Uma linguagem L sobre um alfabeto Σ é um conjunto qualquer de palavras sobre Σ , ou seja,

$$L \subseteq \Sigma^*$$

Exemplo: Linguagem Formal

\emptyset e $\{\varepsilon\}$ são linguagens sobre *qualquer* alfabeto

$$\emptyset \neq \{\varepsilon\}$$

Σ^* e Σ^+ são linguagens sobre um Σ qualquer

O conjunto de palíndromos sobre $\Sigma = \{a, b\}$ é uma linguagem

$\varepsilon, a, b, aa, bb, aaa, aba, bab, bbb, aaaa, \dots$



Mais Exemplos

Conjunto de todas as Linguagens sobre um Alfabeto

- É o conjunto das partes de Σ^*
- 2^{Σ^*}

Linguagem de Programação

- conjunto de todos os programas (palavras) da linguagem



Gramática

Especificação formal **finita** de linguagens (eventualmente) infinitas.

- Conjunto finito de regras
- Aplicações sucessivas geram palavras
- A linguagem é definida pelo conjunto de todas as palavras geradas por uma gramática



Gramática

Definição (Gramática)

Uma gramática é uma quádrupla ordenada

$$G = \langle V, T, P, S \rangle$$

onde:

- V é um conjunto *finito* de símbolos, chamados **variáveis**
- T é um conjunto *finito* de símbolos, chamados **terminais**, sendo que $V \cap T = \emptyset$
- $P : (V \cup T)^+ \rightarrow (V \cup T)^*$ é uma relação *finita* onde cada par da relação é denominado **regra de produção**
- $S \in V$ é a **variável inicial** da gramática



Gramática

Uma regra de produção (α, β) é representada por

$$\alpha \rightarrow \beta$$

Abreviação para $\alpha \rightarrow \beta_1, \alpha \rightarrow \beta_2, \dots, \alpha \rightarrow \beta_n$

$$\alpha \rightarrow \beta_1 \mid \beta_2 \mid \dots \mid \beta_n$$



Derivação

É a aplicação de uma regra de produção

Sucessivas derivações:

- fecho transitivo da relação de derivação
- permite derivar palavras da linguagem



Derivação

Definição (Derivação)

Seja uma gramática $G = \langle V, T, P, S \rangle$. Uma derivação é um par da relação $\Rightarrow: (V \cup T)^+ \rightarrow (V \cup T)^*$ representado por

$$\alpha \Rightarrow \beta$$

e tal relação é indutivamente definida como segue:

- para toda produção da forma $S \rightarrow \beta$, tem-se

$$S \Rightarrow \beta$$

- para todo par $\eta \Rightarrow \rho\alpha\sigma$ da relação de derivação, se $\alpha \rightarrow \beta \in P$, tem-se

$$\eta \Rightarrow \rho\beta\sigma$$



Derivação

Derivação: substituição de uma subpalavra de acordo com uma regra de produção.

Sucessivos passos de derivação

- \Rightarrow^* : zero ou mais passos sucessivos de derivação
- \Rightarrow^+ : um ou mais passos sucessivos de derivação
- \Rightarrow^n : exatos n passos de derivação



Linguagem Gerada

Definição (Linguagem Gerada)

Seja $G = \langle V, T, P, S \rangle$ uma gramática. A linguagem gerada por G , denotada por $L(G)$ é o conjunto

$$L(G) = \{w \in T^* | S \Rightarrow^+ w\}$$

$L(G)$ são todas as palavras compostas **somente de símbolos terminais** deriváveis a partir do símbolo inicial da gramática.



Exemplo: Gramática para Números Naturais

$$G = \langle V, T, P, N \rangle$$

- $V = \{N, D\}$
- $T = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$
- $P = \{N \rightarrow D, N \rightarrow DN, D \rightarrow 0 \mid 1 \mid \dots \mid 9\}$

Gera sintaticamente o conjunto dos números naturais

- **distingue** os zeros à esquerda
- $123 \neq 0123$
- Desafio: definir uma gramática que não produza números com zeros à esquerda



Exemplo: Gramática para Números Naturais

Derivação da palavra 243

$N \Rightarrow$	$N \rightarrow DN$
$DN \Rightarrow$	$D \rightarrow 2$
$2N \Rightarrow$	$N \rightarrow DN$
$2DN \Rightarrow$	$D \rightarrow 4$
$24N \Rightarrow$	$N \rightarrow D$
$24D \Rightarrow$	$D \rightarrow 3$
243	

Portanto

- $S \Rightarrow^* 243$
- $S \Rightarrow^+ 243$
- $S \Rightarrow^6 243$



Exemplo: Gramática para Palavra Duplicada

$$G = \langle \{S, X, Y, A, B, F\}, \{a, b\}, P, S \rangle$$

onde:

$$P = \{ \begin{array}{l} S \rightarrow XY \\ X \rightarrow XaA \mid XbB \mid F \\ Aa \rightarrow aA, Ab \rightarrow bA, AY \rightarrow Ya \\ Ba \rightarrow aB, Bb \rightarrow bB, BY \rightarrow Yb \\ Fa \rightarrow aF, Fb \rightarrow bF, FY \rightarrow \varepsilon \end{array} \}$$

gera a linguagem

$$\{ww \mid w \text{ é palavra de } \{a, b\}^*\}$$



Exemplo: Gramática para Palavra Duplicada

Derivação da palavra *baba*

$S \Rightarrow$	$S \rightarrow XY$
$XY \Rightarrow$	$X \rightarrow XaA$
$XaAY \Rightarrow$	$AY \rightarrow Ya$
$XaYa \Rightarrow$	$X \rightarrow XbB$
$XbBaYa \Rightarrow$	$Ba \rightarrow aB$
$XbaBYa \Rightarrow$	$BY \rightarrow Yb$
$XbaYba \Rightarrow$	$X \rightarrow F$
$FbaYba \Rightarrow$	$Fb \rightarrow bF$
$bFaYba \Rightarrow$	$Fa \rightarrow aF$
$baFYba \Rightarrow$	$FY \rightarrow \varepsilon$
<i>baba</i>	

Existe mais alguma derivação de *baba*?



Gramáticas Equivalentes

Definição (Gramáticas Equivalentes)

Duas gramáticas G_1 e G_2 são equivalentes se e somente se

$$L(G_1) = L(G_2)$$

Notações a serem usadas:

- A, B, C, \dots, X, Y, Z para símbolos **variáveis**
- a, b, c, \dots, s, t para símbolos **terminais**
- u, v, w, x, y, z para **palavras** de símbolos **terminais**
- α, β, \dots para **palavras** compostas por **variáveis e/ou terminais**