

## Aula 2: Contextualização: alfabeto, palavra, linguagem formal e gramática.

Prof. Lucio A. Rocha

Engenharia de Computação  
Universidade Tecnológica Federal do Paraná, UTFPR  
Câmpus Apucarana, Brasil

1º semestre / 2023

# Sumário

1 Teoria da Computação

2 Conceitos Básicos

## Seção 1

# Teoria da Computação

# Teoria da Computação

- Lida com modelos matemáticos de computação.
- Herança da década de 50: teorias sobre processamento de linguagens naturais.
- Estudo de linguagens artificiais.
- Estudo de linguagens de programação de computadores.

# Teoria da Computação

- Aplicações:
  - Compiladores.
  - Modelagem de Redes Lógicas.
  - Modelagem de Circuitos Lógicos.
  - Modelagem de Sistemas Biológicos.
  - Linguagens Matemáticas.
  - Entre muitas outras.

# Teoria da Computação

- Aplicações:

- Análise Léxica.

- Reconhecer palavras do vocabulário do programa.
    - Ex.: Palavras-chave, operadores, constantes, símbolos

- Análise Sintática.

- Ex.: comandos simples (if-else, switch, for, etc.), blocos, estrutura do programa como um todo.

- Análise Semântica.

- Verificar relacionamentos entre partes do código no programa.
    - Ex.:

```
int metodo(int X);
```

```
...
```

```
int A,B;
```

```
A=metodo(B);
```

# Teoria da Computação

- *Linguagem Formal*: é a linguagem descrita com regras e composta de palavras de um alfabeto.
- A sintaxe da linguagem possui base lógica e matemática.
- Linguagem de Programação:
  - É uma linguagem formal.
  - É livre, i.e., sem qualquer significado associado.
  - Porém, exige interpretação/tradução do seu significado.

# Teoria da Computação

- Abordagem deste curso:
  - Análise de linguagens lineares abstratas.
  - Linguagens associadas a problemas na Computação.
- Classificação dos formalismos das linguagens:
  - Axiomático
  - Denotacional
  - Operacional



# Teoria da Computação

- Formalismo Axiomático:
  - Componentes da linguagem são associados a regras.
  - Regras permitem afirmar o que será verdadeiro após a execução de cada cláusula.
  - Formalismo Gerador: verifica se uma palavra é gerada por uma dada gramática.
  - Abordagem é sobre Gramáticas.

# Teoria da Computação

- Formalismo Denotacional:
  - Restrito às Expressões Regulares (ER).
  - Valor denotado por uma construção com ER.
  - Formalismo Gerador: verifica se uma palavra da linguagem é gerada.
  - Define um domínio com um conjunto de palavras admissíveis na linguagem.

# Teoria da Computação

- Formalismo Operacional:
  - Autômato (ou Máquina Abstrata).
    - Estados
    - Instruções primitivas
    - Definição de como cada instrução modifica em cada estado.
  - Máquina Abstrata
    - Suficientemente simples
    - Não gera dúvidas sobre a execução do código
    - Formalismo Reconhecedor: verifica se uma entrada é válida
  - Principais Máquinas Abstratas:
    - Autômato Finito
    - Autômato com Pilha
    - Máquina de Turing

## Seção 2

# Conceitos Básicos

# Conceitos Básicos

- *Linguagem*: é um conjunto de palavras formadas por símbolos.
- *Palavra*: é um conjunto de símbolos de um alfabeto. Palavras de uma linguagem são formadas com regras de produção.
- *Alfabeto*: é o conjunto de símbolos válidos na linguagem.
- Ex.:
  - Alfabeto da Língua Portuguesa:  $A = \{a, b, c, d, \dots, z\}$
  - Alfabeto de linguagem de máquina:  $B = \{0, 1\}$

# Conceitos Básicos

- Palavra:
  - Prefixo: qualquer sequência inicial de símbolos da palavra.
  - Sufixo: qualquer sequência final de símbolos da palavra.
  - Subpalavra: qualquer sequência de símbolos contíguos da palavra.

# Conceitos Básicos

- Exemplo:  $w=abcb$ ,  $\Sigma = \{a, b, c\}$ 
  - Prefixos:  $\varepsilon, a, ab, abc, abcb$  (e apenas estes)
  - Sufixos:  $\varepsilon, b, cb, bcb, abcb$  (e apenas estes)
  - Subpalavra: qualquer prefixo ou sufixo é uma subpalavra.

# Conceitos Básicos

- Em uma linguagem de programação como C:
  - Uma palavra é um programa.



# Conceitos Básicos

- Concatenação de palavras
  - Justaposição da primeira palavra com a segunda palavra.
  - Propriedades:
    - Elemento Neutro:  $\varepsilon w = w = w\varepsilon$
    - Associatividade:  $v(wt) = (vw)t$

# Conceitos Básicos

- Exemplos de Concatenação de palavras
  - Seja  $A = \{a, b\}$  um alfabeto. Para as palavras  $v = aa$  e  $w = bb$ 
    - $vw = aabb$
    - $v\varepsilon = v = aa$
    - $v^0 = \varepsilon$
    - $w^2 = bbbb$

# Conceitos Básicos

- Linguagem Formal

- Uma linguagem é um conjunto de palavras sobre um alfabeto.
- Uma linguagem formal  $L$  é aquela que, dado um alfabeto  $A$ :

$$L \subseteq A^*$$

# Conceitos Básicos

- *Palavra (sentença ou string)*: é um conjunto concatenado de símbolos de uma dada linguagem.
- A palavra vazia (nenhum símbolo) é indicada por  $\varepsilon$ .
- Se  $A$  é um alfabeto, então:
  - $A^*$  (clausura de  $A$ ): é o conjunto de todas as palavras formadas com símbolos de  $A$ .
  - $A^+ = A^* - \varepsilon$

# Conceitos Básicos

- Exemplo:
  - Seja  $B = \{0, 1\}$  o alfabeto para a linguagem:
    - $L = \{0^n 1^n \mid n \geq 0\}$
  - Quais das seguintes palavras são válidas nesta linguagem?
    - $\varepsilon$
    - 01
    - 0110
    - 0011

# Conceitos Básicos

- Gramática:
  - A gramática define as regras para a formação de palavras válidas para uma dada linguagem.
  - *Produções*: é o conjunto de regras de formação de palavras.
  - *Alfabeto*: é o conjunto de símbolos da linguagem E o conjunto de símbolos auxiliares:
    - **Símbolos terminais**: símbolos da linguagem.
    - **Símbolos não-terminais**: símbolos auxiliares. Indicam o ponto de partida para formação de palavras na linguagem.

# Conceitos Básicos

- Definição formal de Gramática:

$$G = (V_T, V_N, \mathbb{P}, S_i)$$

- $V_T$ : alfabeto de símbolos terminais.
- $V_N$ : alfabeto de símbolos não-terminais.
- $\mathbb{P}$ : conjunto de regras (produções), expressos na forma:
  - $\alpha \rightarrow \beta$ , onde  $\alpha \in V_N^+$ ,  $\beta \in V^*$
  - $S_i$ : símbolo sentencial, símbolo não-terminal inicial ou axioma:
    - É o símbolo de início da produção de palavras na linguagem.  $S_i \in V_N$

# Conceitos Básicos

- Produções:

$$E \rightarrow D$$

- Indica que o símbolo ou sequência de símbolos  $E$  (o lado esquerdo) pode ser substituído pelo símbolo ou sequência de símbolos  $D$  (o lado direito) na formação de uma palavra, a partir do símbolo sentencial.



# Conceitos Básicos

- Exemplo:

$$G_1 = (\{0, 1\}, \{Z\}, \{Z \rightarrow 0Z1, Z \rightarrow \varepsilon\}, Z)$$

- As duas produções são:
  - $Z \rightarrow 0Z1$ : onde há o símbolo  $Z$  é possível substituí-lo pela sequência  $0Z1$ .
  - $Z \rightarrow \varepsilon$ : onde há o símbolo  $Z$  é possível substituí-lo pela palavra vazia  $\varepsilon$  (ou seja, eliminá-lo).

# Conceitos Básicos

- Derivação:

- É a substituição do lado esquerdo de uma produção de uma gramática pelos símbolos do lado direito:

$$Z \Rightarrow 0Z1 \Rightarrow 01$$

- O resultado da derivação pode ser:
  - **Sentença:** Forma sentencial apenas com os símbolos terminais.
  - **Forma sentencial:** Sequência de símbolos, terminais ou não terminais, que pode ser derivada a partir do símbolo sentencial da gramática.

# Conceitos Básicos

- Derivação:

- Dadas duas formas sentenciais gama  $\gamma$  e delta  $\delta$ :

- $\gamma \Rightarrow \delta$ :  $\delta$  é imediatamente derivável de  $\gamma$  com uma única produção.
    - $\gamma \Rightarrow^+ \delta$ :  $\delta$  é derivável de  $\gamma$  pela aplicação de uma ou mais produções.

- Exemplo:

- 0Z1 é imediatamente derivável de Z
    - 01 é imediatamente derivável de 0Z1
    - 01 é derivável de Z

# Conceitos Básicos

- Derivação: Reconhecimento de sentenças válidas.
  - A sentença sigma  $\sigma$  faz parte da linguagem  $L$  se há sequências de derivações a partir do símbolo sentencial  $S$  que leve a  $\sigma$ :

$$S \Rightarrow^+ \sigma$$

- Exemplo: Seja a gramática:

$$G_1 = (\{0, 1\}, \{Z\}, \{Z \rightarrow 0Z1, Z \rightarrow \varepsilon\}, Z)$$

- Quais sentenças são válidas nesta gramática  $G_1$ ?
  - 01
  - 010
  - 0011
  - 0110
  - $\varepsilon$

# Conceitos Básicos

- Derivação: Reconhecimento de sentenças válidas.

$$G_1 = (\{0, 1\}, \{Z\}, \{Z \rightarrow 0Z1, Z \rightarrow \varepsilon\}, Z)$$

- $01 : Z \Rightarrow 0Z1 \Rightarrow 01$  (válida)
- $010 : Z \Rightarrow 0Z1 \Rightarrow ?$  (inválida)
- $0011 : Z \Rightarrow 0Z1 \Rightarrow 00Z11 \Rightarrow 0011$  (válida)
- $0110 : Z \Rightarrow 0Z1 \Rightarrow ?$  (inválida)
- $\varepsilon : Z \Rightarrow \varepsilon$  (válida)

# Conceitos Básicos

- Classificação de Gramáticas
  - Gramáticas podem ter produções com diferentes graus de complexidade em seu formato:
    - Quantos símbolos no lado esquerdo?
    - Como os símbolos que aparecem no lado esquerdo podem aparecer no lado direito (recursividade)?
  - Quanto menor for a restrição no formato das produções, maior será o poder de expressão da gramática
    - Podem representar gramáticas mais complexas.
    - Mais complexo o método de reconhecimento de sentenças.