

TEORIA DA COMPUTAÇÃO

---

UNIDADE 2: LINGUAGENS LIVRES-DO-CONTEXTO

AULA 2: INTRODUÇÃO

PROFESSOR: LUCAS CAMBUIM

# Contexto do que vamos começar a estudar

Até agora nós vimos dois métodos diferentes, embora equivalentes, de descrever linguagens:

## Autômatos finitos e Expressões regulares

Mostramos que algumas linguagens simples ,  $\{0^n 1^n \mid n \geq 0\}$  não podem ser descritas por estes métodos.

# Contexto do que vamos começar a estudar

As linguagens também podem ser definidas formalmente por meio de gramáticas livres-do-contexto, que é um método poderoso de descrever formalmente uma linguagem.

# Gramáticas Livres de Contexto

- Usos – especificação de linguagens formais , compiladores, parsers, ...
- A coleção de linguagens associadas com gramáticas livres de contexto é chamado de **linguagens livres-do-contexto**
  - Incluem as linguagens regulares
  - e muitas linguagens adicionais

# Gramáticas Livres de Contexto

Seja a gramática livre-de-contexto  $G_1$

$$A \rightarrow 0A1$$

$$A \rightarrow B$$

$$B \rightarrow \#$$

# Gramáticas Livres de Contexto

- Uma gramática consiste de uma coleção de ***regras de substituição***, também chamadas ***produções***.

# Gramáticas Livres de Contexto

- Uma gramática consiste de uma coleção de **regras de substituição**, também chamadas **produções**.
- Regra:  $\langle \text{Variável} \rangle \rightarrow \langle \text{cadeia} \rangle$

# Gramáticas Livres de Contexto

- Uma gramática consiste de uma coleção de **regras de substituição**, também chamadas **produções**.
- Regra:  $\langle \text{Variável} \rangle \rightarrow \langle \text{cadeia} \rangle$
- A cadeia consiste de variáveis e outros símbolos chamados de **terminais**.

# Gramáticas Livres de Contexto

- Uma gramática consiste de uma coleção de **regras de substituição**, também chamadas **produções**.
- Regra:  $\langle \text{Variável} \rangle \rightarrow \langle \text{cadeia} \rangle$
- A cadeia consiste de variáveis e outros símbolos chamados de **terminais**.
- **Variáveis** são representados por letras maiúsculas.

# Gramáticas Livres de Contexto

- Uma gramática consiste de uma coleção de **regras de substituição**, também chamadas **produções**.
- Regra:  $\langle \text{Variável} \rangle \rightarrow \langle \text{cadeia} \rangle$
- A cadeia consiste de variáveis e outros símbolos chamados de **terminais**.
- **Variáveis** são representados por letras maiúsculas.
- Os terminais são análogos ao alfabeto de entrada

# Gramáticas Livres de Contexto

- Uma gramática consiste de uma coleção de **regras de substituição**, também chamadas **produções**.
- Regra:  $\langle \text{Variável} \rangle \rightarrow \langle \text{cadeia} \rangle$
- A cadeia consiste de variáveis e outros símbolos chamados de **terminais**.
- **Variáveis** são representados por letras maiúsculas.
- Os terminais são análogos ao alfabeto de entrada
- Os terminais são representados por letras minúsculas, números, ou símbolos especiais.

# Gramáticas Livres de Contexto

- Uma gramática consiste de uma coleção de **regras de substituição**, também chamadas **produções**.
- Regra:  $\langle \text{Variável} \rangle \rightarrow \langle \text{cadeia} \rangle$
- A cadeia consiste de variáveis e outros símbolos chamados de **terminais**.
- **Variáveis** são representados por letras maiúsculas.
- Os terminais são análogos ao alfabeto de entrada
- Os terminais são representados por letras minúsculas, números, ou símbolos especiais.
- A **variável inicial** ocorre no lado esquerdo da 1a regra

# Gramáticas Livres de Contexto

- Para gerar uma cadeia de uma linguagem:
- 1. Escreva a variável inicial. Ela é a variável no lado esquerdo da primeira regra, a menos que especificado em contrario.

# Gramáticas Livres de Contexto

- Para gerar uma cadeia de uma linguagem:
  - 1. Escreva a variável inicial. Ela é a variável no lado esquerdo da primeira regra, a menos que especificado em contrario.
  - 2. Encontre uma variável que está escrita em uma regra que começa com essa variável. Substitua a variável escrita pelo lado direito dessa regra.

# Gramáticas Livres de Contexto

- Para gerar uma cadeia de uma linguagem:
  - **1.** Escreva a variável inicial. Ela é a variável no lado esquerdo da primeira regra, a menos que especificado em contrario.
  - **2.** Encontre uma variável que está escrita em uma regra que começa com essa variável. Substitua a variável escrita pelo lado direito dessa regra.
  - **3.** Repita o passo 2 ate que não reste nenhuma variável.

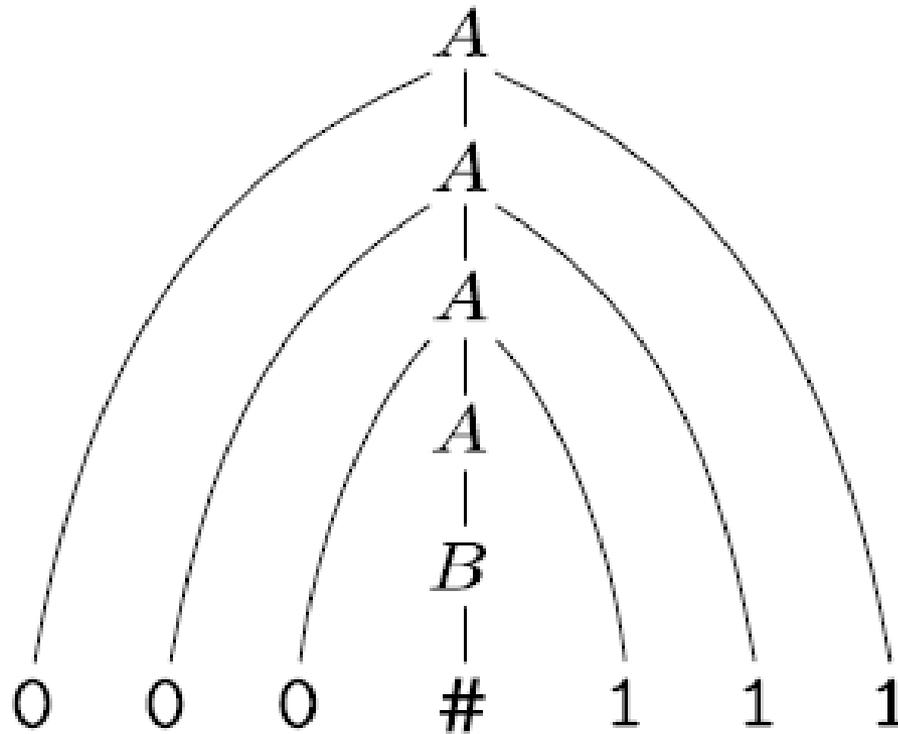
# Gramáticas Livres de Contexto

$$A \rightarrow 0A1$$
$$A \rightarrow B$$
$$B \rightarrow \#$$

- A gramática  $G_1$  gera a cadeia  $000\#111$ .
- A sequência de substituições para obter uma cadeia é denominada uma **derivação**.
- $A \rightarrow 0A1 \rightarrow 00A11 \rightarrow 000A111 \rightarrow 000B111 \rightarrow 000\#111$

# Gramáticas Livres de Contexto

Árvore sintática



# Denominações

- O conjunto de todas as cadeias geradas dessa maneira constitui a ***linguagem da gramática***.
- Escrevemos  $L(G_1)$  para a linguagem da gramática  $G_1$
- Mesma variável do lado esquerdo:

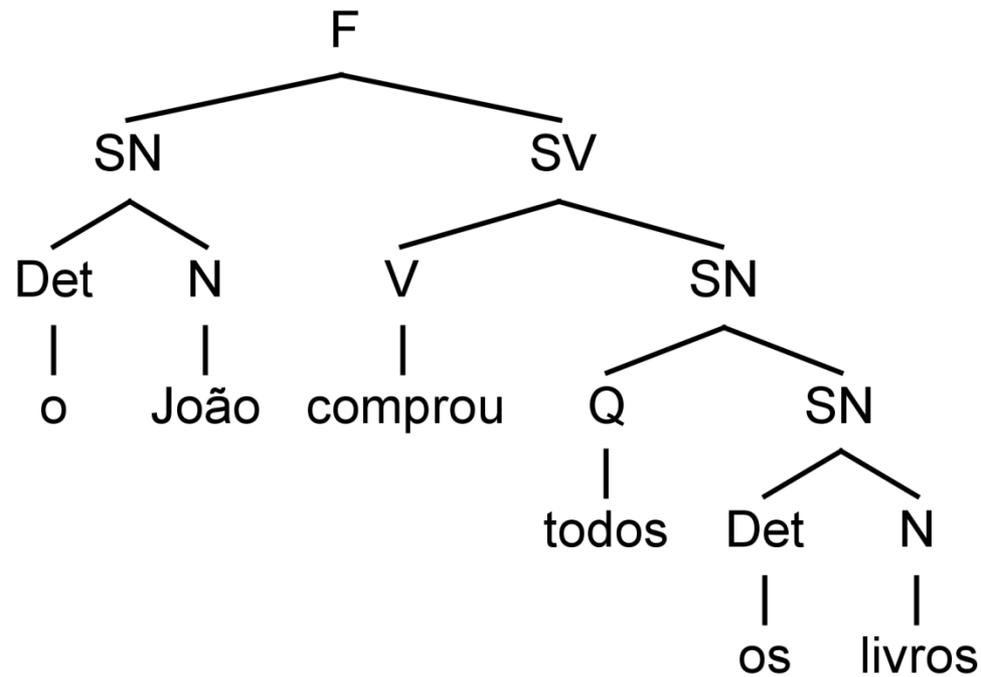
$$A \rightarrow 0A1$$

$$A \rightarrow B$$

Podem ser escritos como:  $A \rightarrow 0A1 \mid B$

# Gramáticas Livres de Contexto

## Árvore sintática



# Gramáticas Livres de Contexto

- $\langle \text{SENTENCE} \rangle \rightarrow \langle \text{NOUN-PHRASE} \rangle \langle \text{VERB-PHRASE} \rangle$
- $\langle \text{NOUN-PHRASE} \rangle \rightarrow \langle \text{CMPLX-NOUN} \rangle \mid \langle \text{CMPLX-NOUN} \rangle \langle \text{PREP-PHRASE} \rangle$
- $\langle \text{VERB-PHRASE} \rangle \rightarrow \langle \text{CMPLX-VERB} \rangle \mid \langle \text{CMPLX-VERB} \rangle \langle \text{PREP-PHRASE} \rangle$
- $\langle \text{PREP-PHRASE} \rangle \rightarrow \langle \text{PREP} \rangle \langle \text{CMPLX-NOUN} \rangle$
- $\langle \text{CMPLX-NOUN} \rangle \rightarrow \langle \text{ARTICLE} \rangle \langle \text{NOUN} \rangle$
- $\langle \text{CMPLX-VERB} \rangle \rightarrow \langle \text{VERB} \rangle \mid \langle \text{VERB} \rangle \langle \text{NOUN-PHRASE} \rangle$
- $\langle \text{ARTICLE} \rangle \rightarrow a \mid the$
- $\langle \text{NOUN} \rangle \rightarrow boy \mid girl \mid flower$
- $\langle \text{VERB} \rangle \rightarrow touches \mid likes \mid sees$
- $\langle \text{PREP} \rangle \rightarrow with$

A gramática tem 10 variáveis

# Gramáticas Livres de Contexto

- $\langle \text{SENTENCE} \rangle \rightarrow \langle \text{NOUN-PHRASE} \rangle \langle \text{VERB-PHRASE} \rangle$
- $\langle \text{NOUN-PHRASE} \rangle \rightarrow \langle \text{CMPLX-NOUN} \rangle \mid \langle \text{CMPLX-NOUN} \rangle \langle \text{PREP-PHRASE} \rangle$
- $\langle \text{VERB-PHRASE} \rangle \rightarrow \langle \text{CMPLX-VERB} \rangle \mid \langle \text{CMPLX-VERB} \rangle \langle \text{PREP-PHRASE} \rangle$
- $\langle \text{PREP-PHRASE} \rangle \rightarrow \langle \text{PREP} \rangle \langle \text{CMPLX-NOUN} \rangle$
- $\langle \text{CMPLX-NOUN} \rangle \rightarrow \langle \text{ARTICLE} \rangle \langle \text{NOUN} \rangle$
- $\langle \text{CMPLX-VERB} \rangle \rightarrow \langle \text{VERB} \rangle \mid \langle \text{VERB} \rangle \langle \text{NOUN-PHRASE} \rangle$
- $\langle \text{ARTICLE} \rangle \rightarrow a \mid the$
- $\langle \text{NOUN} \rangle \rightarrow boy \mid girl \mid flower$
- $\langle \text{VERB} \rangle \rightarrow touches \mid likes \mid sees$
- $\langle \text{PREP} \rangle \rightarrow with$

# Gramáticas Livres de Contexto

A gramática tem:

- 10 variáveis (os termos gramaticais em maiúsculas escritos dentro de colchetes);
- 27 terminais (o alfabeto inglês-padrão mais um caractere de espaço em branco)
- 18 regras

# Gramáticas Livres de Contexto

- G2 produz:
- a boy sees
- the boy sees a flower
- a girl with a flower likes the boy

# Gramáticas Livres de Contexto

a boy sees

<SENTENCE> ⇒ <NOUN-PHRASE><VERB-PHRASE>  
⇒ <CMPLX-NOUN><VERB-PHRASE>  
⇒ <ARTICLE><NOUN><VERB-PHRASE>  
⇒ a <NOUN><VERB-PHRASE>  
⇒ a boy <VERB-PHRASE>  
⇒ a boy <CMPLX-VERB>  
⇒ a boy <VERB>  
⇒ a boy sees

# Gramáticas Livres de Contexto

- Uma **gramática livre-do-contexto** é uma 4-upla  $(V, \Sigma, R, S)$ :
  - 1.  $V$  é um conjunto finito das **variáveis**,
  - 2.  $\Sigma$  é um conjunto finito, disjunto de  $V$ , dos **terminais**,
  - 3.  $R$  é um conjunto finito de **regras**, com cada regra sendo uma variável e uma cadeia de variáveis e terminais, e
  - 4.  $S \in V$  é a variável inicial.

## Exemplo 2.3

Considere a gramática  $G_3 = (\{S\}, \{a, b\}, R, S)$ . O conjunto de regras,  $R$ , é:

$$S \rightarrow aSb \mid SS \mid \epsilon$$

## Exemplo 2.3

Considere a gramática  $G_3 = (\{S\}, \{a, b\}, R, S)$ . O conjunto de regras,  $R$ , é:

$$S \rightarrow aSb \mid SS \mid \epsilon$$

Cadeias que podem ser geradas: abab, aaabbb e aababb.

## Exemplo 2.3

Se pensar em a como '(' e b como ')':

$$S \rightarrow (S) \mid SS \mid \epsilon$$

$L(G_3)$  é a linguagem de todas as cadeias de parênteses apropriadamente aninhadas.

## Exemplo 2.4

Considere a gramática  $G_4 = (V, \Sigma, R, \langle \text{EXPR} \rangle)$ .

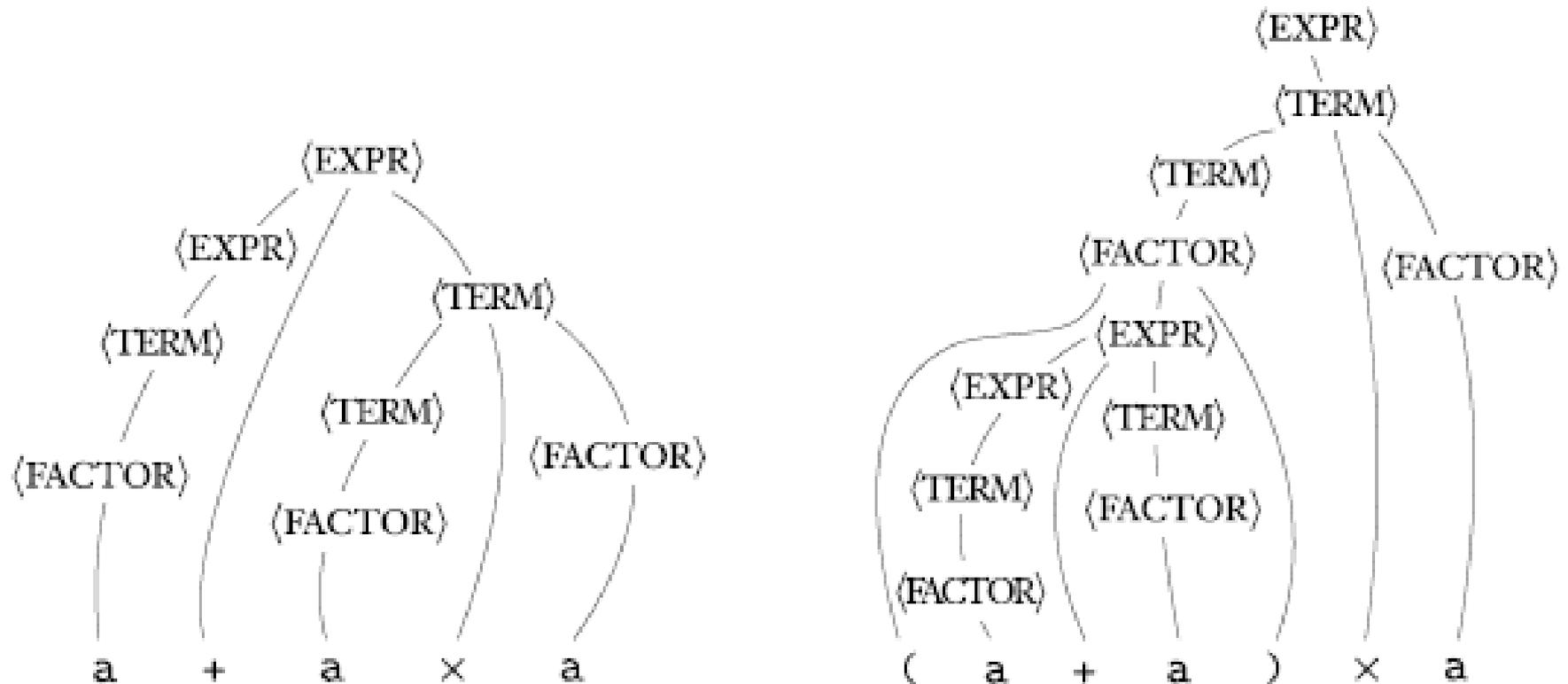
$V$  é  $\{\langle \text{EXPR} \rangle, \langle \text{TERM} \rangle, \langle \text{FACTOR} \rangle\}$  e  $\Sigma$  é  $\{a, +, x, (, )\}$ .

As regras são:

$$\langle \text{EXPR} \rangle \Rightarrow \langle \text{EXPR} \rangle + \langle \text{TERM} \rangle \mid \langle \text{TERM} \rangle$$
$$\langle \text{TERM} \rangle \Rightarrow \langle \text{TERM} \rangle x \langle \text{FACTOR} \rangle \mid \langle \text{FACTOR} \rangle$$
$$\langle \text{FACTOR} \rangle \Rightarrow (\langle \text{EXPR} \rangle) \mid a$$

As duas cadeias  $a+axa$  e  $(a+a)xa$  podem ser geradas com a gramática  $G_4$ . Mostre as árvores sintáticas para gerar essas cadeias.

# Exemplo 2.4



**FIGURA 2.5**

Árvores sintáticas para as cadeias  $a+a*a$  e  $(a+a)*a$

# Gramáticas Livres de Contexto

- Muitas linguagens livre de contexto são uniões de linguages livres de contexto mais simples;
- Se você precisa construir uma gramática para uma linguagem assim, você pode construir gramáticas para cada uma das partes;
- Depois pode uní-las com uma nova regra:
- $S \rightarrow S_1 \mid S_2 \mid \dots \mid S_n$ , onde cada variável  $S_i$  são variáveis iniciais das gramáticas construídas a parte;

# Exemplo

- Construção da Gramática para a seguinte linguagem  $\{0^n 1^n \mid n \geq 0\} \cup \{1^n 0^n \mid n \geq 0\}$ ;

# Exemplo

- Construção da Gramática para a seguinte linguagem  $\{0^n 1^n \mid n \geq 0\} \cup \{1^n 0^n \mid n \geq 0\}$ ;
- Primeiro construímos a regra para  $\{0^n 1^n \mid n \geq 0\}$ ;
- $S_1 \rightarrow 0S_11 \mid \varepsilon$

# Exemplo

- Construção da Gramática para a seguinte linguagem  $\{0^n 1^n \mid n \geq 0\} \cup \{1^n 0^n \mid n \geq 0\}$ ;
- Primeiro construímos a regra para  $\{0^n 1^n \mid n \geq 0\}$ ;
- $S_1 \rightarrow 0S_11 \mid \varepsilon$
- Em seguida construímos a regra para  $\{1^n 0^n \mid n \geq 0\}$
- $S_2 \rightarrow 1S_20 \mid \varepsilon$

# Exemplo

- E por fim uma regra em que  $S \rightarrow S_1 \mid S_2$
- $S \rightarrow S_1 \mid S_2$
- $S_1 \rightarrow 0S_11 \mid \varepsilon$
- $S_2 \rightarrow 1S_20 \mid \varepsilon$

## Gramáticas Livres de Contexto a partir de AFDs

- Construir uma Gramática Livre de Contexto para uma linguagem regular é simples, caso você primeiro construa um AFD para esta linguagem;

## Gramáticas Livres de Contexto a partir de AFDs

- Para converter um AFD em uma Gramática equivalente:

## Gramáticas Livres de Contexto a partir de AFDs

- Para converter um AFD em uma Gramática equivalente:
- Faça uma variável  $R_i$  para cada estado  $q_i$  do AFD;

## Gramáticas Livres de Contexto a partir de AFDs

- Para converter um AFD em uma Gramática equivalente:
- Faça uma variável  $R_i$  para cada estado  $q_i$  do AFD;
- Adicione a regra  $R_i \rightarrow aR_j$ , se  $\delta(q_i, a) = q_j$ ;

## Gramáticas Livres de Contexto a partir de AFDs

- Para converter um AFD em uma Gramática equivalente:
- Faça uma variável  $R_i$  para cada estado  $q_i$  do AFD;
- Adicione a regra  $R_i \rightarrow aR_j$ , se  $\delta(q_i, a) = q_j$ ;
- Adicione a regra  $R_i \rightarrow \varepsilon$ , se  $q_i$  é um estado de aceitação;

# Mais exercícios

Desenvolva Gramáticas Livre de Contexto (GLC) que geram as seguintes linguagens:

a)  $L = \{a,b\}^*$

b)  $L = \{a^i b^{2i} \mid i \geq 1\}$

# Mais exercícios

Considere a GLC

$$G=(V,T,P,S)$$

$$V=\{S,A,B\}$$

$$T=\{0,1\}$$

$$P=\{S \rightarrow A1B, A \rightarrow 0A \mid \varepsilon, B \rightarrow 0B \mid 1B \mid \varepsilon\}$$

Qual a linguagem gerada? Essa linguagem é regular ou livre de contexto? Justifique sua resposta.