

TEORIA DA COMPUTAÇÃO I

Aula 04 – Linguagens e Expressões Regulares

Prof. Dr. Guilherme Pina Cardim

guilhermecardim@fai.com.br

13 de março de 2020

Linguagem (L)

Uma linguagem é constituída por um conjunto de **palavras** ou strings válidas.

Palavra (w)

Uma palavra é definida como uma sequência finita de **símbolos** pertencentes a um determinado **alfabeto**.

Alfabeto (Σ)

Conjunto finito de símbolos.

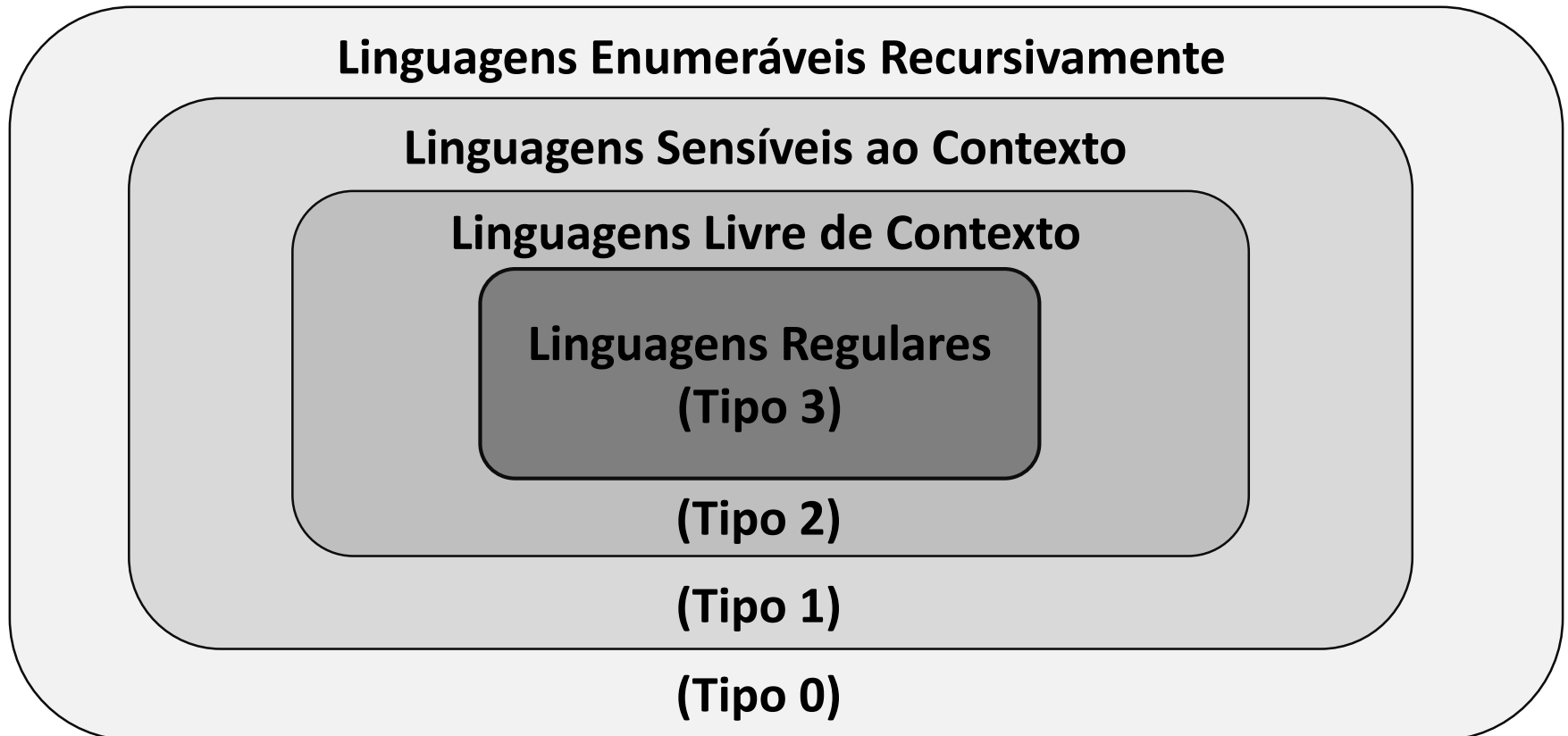
- Uma linguagem formal possui:
 - **Sintaxe** bem definida: dada uma sentença, é possível sempre saber se ela pertence ou não a uma linguagem;
 - **Semântica** precisa: de modo que não contenha sentenças sem significado ou ambíguas.

- Quando um programador projeta um novo sistema, é necessário estabelecer uma **linguagem formal** de comunicação com o usuário final;
- Se a linguagem de comunicação com o usuário é complexa indica que o sistema foi mal projetado.

- Na teoria da computação, um **problema** comum é **identificar** e **decidir** se determinada *string* / palavra é um elemento de uma linguagem:
- Se Σ é um alfabeto e L é uma linguagem sobre Σ , então:
 - Dado uma palavra w em Σ^* , definir se w pertence ou não em L ;
- Exemplo: Dada uma linguagem L , onde as palavras devam possuir quantidade igual de caracteres 0's e 1's:
 - $w = 01100 \rightarrow w$ **não** é um elemento de L ;
 - $w = 0110 \rightarrow w$ **é** um elemento de L .

- É uma coleção de palavras ou cadeias de símbolos de comprimento finito;
- Estas cadeias são denominadas sentenças da linguagem;
- Exemplos:
 - Na computação: Linguagem Pascal, C/C++, Java, etc.
 - $\{ab, bc\}$;
 - $\{ab^n, a^n b; n > 0\}$;

Hierarquia de Chomsky



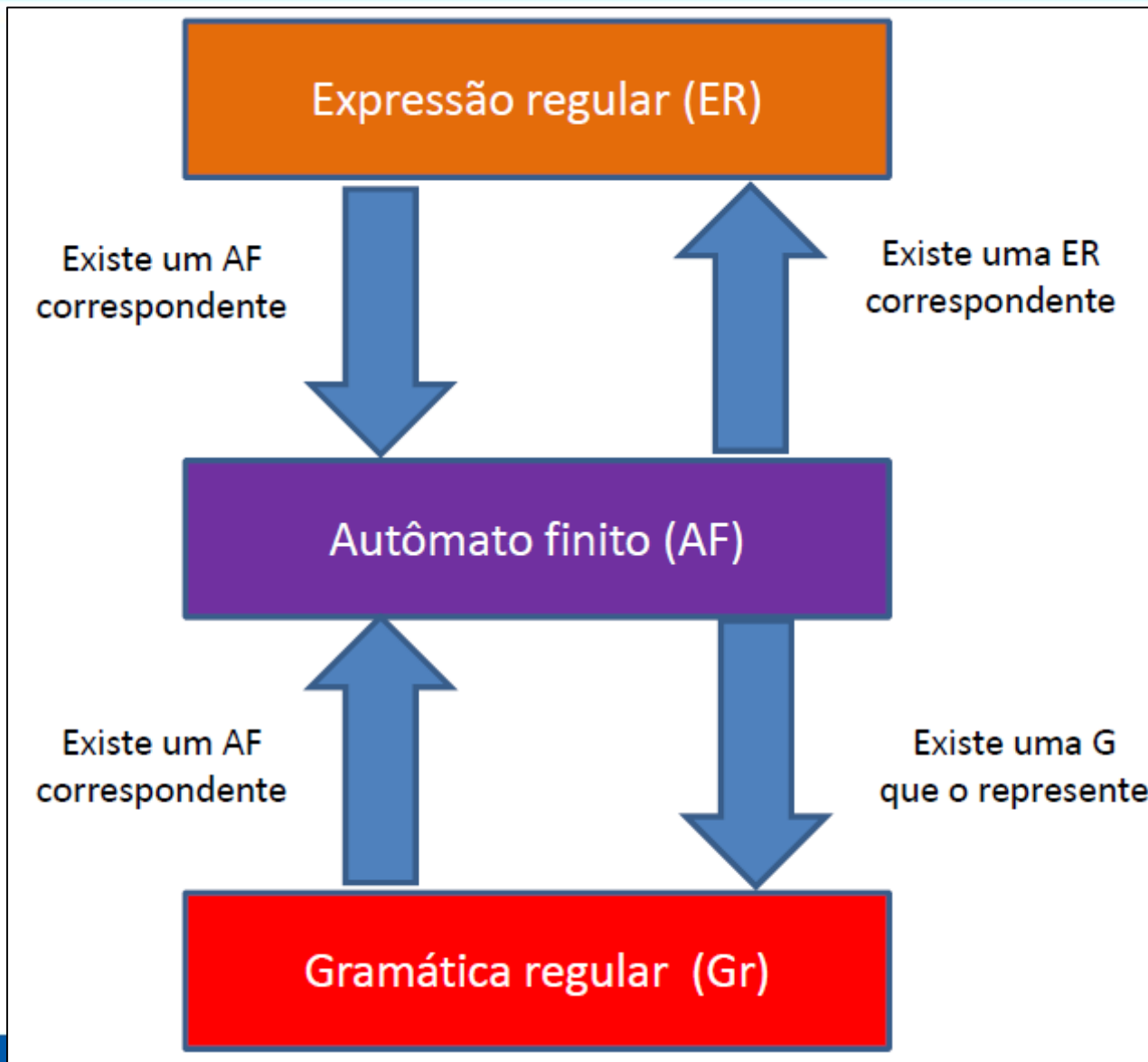
Equivalência entre LR

Linguagens Regulares
(Tipo 3)

{ Expressões Regulares
Autômatos Finitos
Gramáticas Regulares

- Uma linguagem regular (**LR**) pode ser representada de três formas:
 - Expressões Regulares (**ER**);
 - Autômatos Finitos (**AF**);
 - Gramáticas Regulares (**GR**).

Equivalência entre LR



Expressões Regulares



Padrões que descrevem um conjunto grande (até mesmo infinito) de subpalavras que são buscados dentro de uma palavra maior;
(MONTANHER, 2015)

Uma Expressão Regular é uma representação para que você encontre padrões em um texto;
(EIS, 2016)

Expressões Regulares

- Podem ser consideradas como mini-linguagens que permitem especificar as regras da construção de um conjunto de subpalavras válidas;
- Definida a partir de conjuntos básicos de operações de concatenação e união;
- Oferece um modo declarativo de expressar as cadeias de caracteres que devem ser aceitas dentro de uma linguagem.

- Desenvolvidas a partir de:
 - Símbolos do alfabeto: Σ
 - Operadores de:
 - **União:** conjunto de cadeias que está em um conjunto **ou** em outro conjunto – **representado por** \cup **ou** $+$ **ou** $|$
 - **Concatenação:** junção dos elementos – **representado por** \cdot **ou** sem ponto
 - **Fecho-estrela (fechamento):** cadeias que podem ser formadas por um número de repetições do elemento do conjunto (inclusive nenhuma vez) – **representado por** $*$
 - Parênteses: Utilizado para definir precedências.

Expressões Regulares



ER	Palavras Reconhecidas
aa	
ba*	
(a+b)	
(a+b)*	
(a+b)*aa(a+b)*	
a*ba*ba*	
(a+b)*(aa+bb)	

Expressões Regulares



ER	Palavras Reconhecidas
aa	Somente aa
ba*	
(a+b)	
(a+b)*	
(a+b)*aa(a+b)*	
a*ba*ba*	
(a+b)*(aa+bb)	

Expressões Regulares



ER	Palavras Reconhecidas
aa	Somente aa
ba*	Palavras iniciadas por b seguido por 0 ou mais ocorrências de a
(a+b)	
(a+b)*	
(a+b)*aa(a+b)*	
a*ba*ba*	
(a+b)*(aa+bb)	

Expressões Regulares



ER	Palavras Reconhecidas
aa	Somente aa
ba*	Palavras iniciadas por b seguido por 0 ou mais ocorrências de a
(a+b)	Palavras formadas por a ou b
(a+b)*	
(a+b)*aa(a+b)*	
a*ba*ba*	
(a+b)*(aa+bb)	

Expressões Regulares



ER	Palavras Reconhecidas
aa	Somente aa
ba*	Palavras iniciadas por b seguido por 0 ou mais ocorrências de a
(a+b)	Palavras formadas por a ou b
(a+b)*	Palavras formadas por qualquer quantidade de a ou b , inclusive nenhuma vez
(a+b)*aa(a+b)*	
a*ba*ba*	
(a+b)*(aa+bb)	

Expressões Regulares

ER	Palavras Reconhecidas
aa	Somente aa
ba*	Palavras iniciadas por b seguido por 0 ou mais ocorrências de a
(a+b)	Palavras formadas por a ou b
(a+b)*	Palavras formadas por qualquer quantidade de a ou b , inclusive nenhuma vez
(a+b)*aa(a+b)*	Contém aa como elemento
a*ba*ba*	
(a+b)*(aa+bb)	

Expressões Regulares



ER	Palavras Reconhecidas
aa	Somente aa
ba*	Palavras iniciadas por b seguido por 0 ou mais ocorrências de a
(a+b)	Palavras formadas por a ou b
(a+b)*	Palavras formadas por qualquer quantidade de a ou b , inclusive nenhuma vez
(a+b)*aa(a+b)*	Contém aa como elemento
a*ba*ba*	Palavras contendo exatamente dois b 's
(a+b)*(aa+bb)	

Expressões Regulares



ER	Palavras Reconhecidas
aa	Somente aa
ba^*	Palavras iniciadas por b seguido por 0 ou mais ocorrências de a
$(a+b)$	Palavras formadas por a ou b
$(a+b)^*$	Palavras formadas por qualquer quantidade de a ou b , inclusive nenhuma vez
$(a+b)^*aa(a+b)^*$	Contém aa como elemento
$a^*ba^*ba^*$	Palavras contendo exatamente dois b 's
$(a+b)^*(aa+bb)$	Palavras terminadas com aa ou bb

Expressões Regulares

- É possível encontrar diversos simuladores online para verificar uma expressão regular:
 - <https://pt.infobyip.com/regexexpressioncalculator.php>
 - https://tools.lymas.com.br/regexp_br.php#
 - <https://regex101.com/>

Expressão regular tester

Regex

Texto

Ação

Ponto alto ▼

Bandeiras

☒ global
 ☒ ignorar caso
 ☐ multilinha

Computar

Expressão

Texto 1

abb

✓

Texto 2

ababaab

✗

REGULAR EXPRESSION

/ (a+b)*aa

TEST STRING

aabaa

Expressões Regulares

- As ER são muitas vezes consideradas como uma “linguagem de programação”;
- São utilizadas:
 - Em aplicações de pesquisas em textos;
 - Por compiladores em traduções de algoritmos;
 - Para descrever componentes de softwares;
 - Em autômatos finitos para reconhecimento de expressões aceitas.

1. Forneça expressões regulares que denotem os seguintes conjuntos:
 - a) $\{w \in \{a, b\} \mid |w| \geq 3\}$
 - b) $\{w \in \{a, b\} \mid w \text{ começa com } a \text{ e termina com } b\}$
 - c) $\{w \in \{a, b\} \mid w \text{ contém } bb\}$
 - d) $\{w \in \{a, b\} \mid w \text{ contém pelo menos dois } a's\}$

2. Descreva em português e dê exemplos para as expressões regulares a seguir:

a) $0(0 \cup 1)^*1$

b) $0^*(0 \cup 1)1^*$

c) $(0 \cup 1)^*1(0 \cup 1)(0 \cup 1)$

d) $0(10 \cup 1)^*$

Material Referência

- DIVERIO, Tiarajú A. e MENEZES, Paulo B. **Teoria da computação: máquinas universais e computabilidade**. 3ed. Porto Alegre: Bookman, 2011.
- TANENBAUM, Andrew S. **Organização estruturada de computadores**. 5.ed. São Paulo: Pearson Prentice Hall, 2007.
- MONTANHER, Marcelo P. **Autômatos Finitos Determinísticos e Não-Determinísticos**. ICMC/USP, 2015.
- EIS, Diego. **O Básico Sobre Expressões Regulares**. Tableless, 2016.