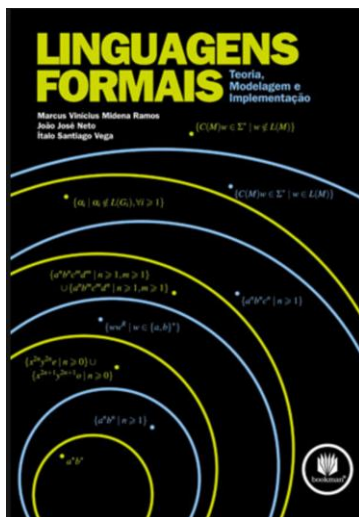




Nossos **objetivos** nesta aula são:

- conhecer os conceitos de alfabeto, palavra e linguagem
- praticar com construções e provas envolvendo estes conceitos



Para esta semana, usamos como referências as **Seções 2.1 (Símbolos e Cadeias)** e **2.2 (Linguagens)** do nosso livro da referência básica:

RAMOS, M.V.M., JOSÉ NETO, J., VEJA, I.S. Linguagens Formais: Teoria, Modelagem e Implementação. Porto Alegre: Bookman, 2009.

Não deixem de ler estas seções depois desta aula!

TEORIA: ALFABETOS, PALAVRAS E LINGUAGENS

- **Alfabeto (Σ):** conjunto finito não-vazio de símbolos (ou letras). **Ex:** $\Sigma=\{0,1\}$ (alfabeto binário)
 - **Notação:** $|\Sigma|$ indica a **cardinalidade** do alfabeto, ou seja, quantos elementos ele possui.
- **Palavra ou cadeia (ω) sobre um alfabeto Σ :** sequência finita (possivelmente vazia) de símbolos de um alfabeto Σ .
 - Para o alfabeto binário ($\Sigma=\{0,1\}$), temos como exemplos de palavras: $\omega=0$, $\omega=1$, $\omega=0101$, $\omega=110101010101110$
 - **Notação:**
 - ε denota a **palavra vazia**, isto é, a palavra que não tem símbolos
 - $|\omega|$ indica o tamanho da palavra, isto é, quantos símbolos ela tem
Ex: $|0110| = 4$ $|\varepsilon|=0$
 - Σ^* denota o conjunto de todas as palavras sobre o alfabeto Σ

- **Operações com palavras:**
 - **concatenação($\alpha\beta$):** junção dos símbolos de duas palavras α e β .
 - **alternação ($\alpha|\beta$):** denota a palavra α ou a palavra β
 - **prefixo:** uma palavra α é prefixo de β se e somente se $\beta=\alpha\gamma$, para alguma palavra γ .
 - **sufixo:** uma palavra α é sufixo de β se e somente se $\beta=\gamma\alpha$, para alguma palavra γ .
 - **reverso (α^R):** consiste em inverter a sequência de símbolos de α
 - **potência (α^n):** consiste em n concatenações da palavra. $\alpha^0=\varepsilon$, $\alpha^1=\alpha$, $\alpha^2=\alpha\alpha$, $\alpha^3=\alpha\alpha\alpha, \dots$ α^* denota 0 ou mais concatenações de α , α^+ denota 1 ou mais concatenações de α
- **Linguagem Formal (L) sobre um alfabeto Σ :** uma linguagem formal L sobre um alfabeto Σ é qualquer subconjunto de Σ^* ($L \subseteq \Sigma^*$).
 - Exs: $L=\emptyset$ (linguagem vazia), $L=\{\varepsilon\}$ (linguagem que contém a palavra vazia) e $L=\{0,1,001,110\}$ são linguagens sobre o alfabeto binário $\Sigma=\{0,1\}$.

EXERCÍCIO TUTORIADO

Considere-se o alfabeto binário $\Sigma=\{0,1\}$. Construa uma linguagem formal sobre o alfabeto Σ que denote todos os números inteiros pares sem sinal.

EXERCÍCIO COM DISCUSSÃO EM DUPLAS

Considere-se novamente o alfabeto binário $\Sigma=\{0,1\}$. Discuta com o seu colega como você construiria uma linguagem formal sobre o alfabeto Σ que denotasse todos os números inteiros que são potências de 2.

EXERCÍCIO TUTORIADO

Seja α uma palavra. Prove que $\alpha\varepsilon = \varepsilon\alpha = \alpha$.

EXERCÍCIO COM DISCUSSÃO EM DUPLAS

Considerando que α , β e γ sejam palavras, prove que $\alpha(\beta\gamma) = (\alpha\beta)\gamma$.

PROBLEMA DE APLICAÇÃO

Uma sequência de DNA é formada pelas bases A(Adenina), T(Timina), G (Guanina) e C(Citosina). Durante processos de análise genômica (um teste de paternidade, por exemplo) são geradas grandes sequências para classificação como, por exemplo, a mostrada abaixo:

```
GXGGAATTGCACATGGGCGCAGCTGCATCAGNNTTGCCGCGTGTGTGAAGAAGGCCTTATGG
TTGTAAAGCACTTTAATGGCGAGGAGGACGGCTACTTTAGATAATACCTAGAGATACGTGGA
CGTTACTCGCAGAATAAGCAGCCAGGCTAACTCTGTGCCAGCAGCCGCGGTAATACAGAGGG
TGCAAGCGTTAATCGGATTTACTGGGCGTAAAGCGCGCGTAGGCGGCTAATTAAGTCAAATG
TGAAATCCCCGAGCTTAACTTGGAATTGCATTCGATACTGGTTAGCTAGAGTGTGGGAGAG
GATGGTAGAAXTCCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGATGGCGA
AGGCAGCCATCTGGCCTAACACTGACGCTGAGGTGCGAAAGCATGGGGAGCAAACAGGATTA
GATACCCTGGTAGTCCATGCCGTAAACGATGTCTACTAGCCGTTGGGGCCTTTGAGGCCTTTA
GTGGCGCAGCTAACGCGATAAGTAGACCGCCTGGGGAGTACGGTCGCAAGACTAAAAC TCA
ATGAATTGACGGGGGGCCCGCACAAAGCGGTGGAGCATGTGGTTTAATTCGATGCAACGCGAAG
AACCTTACCTGGCCTTGACATACTAGAACTTTCCAGAGATGGATTTGGTGCXTC
```

Nestas sequências, quando ocorre um valor de X, temos um caso de falta de consenso na classificação da base como A,T,G ou C. Numa sequência de DNA, existem marcadores genômicos especiais como a sequência TATA, marcador importante para início do processo de transcrição e conhecido como promotor de iniciação do processo de transcrição.

- Qual alfabeto temos associado a este problema?
- Descreva, através de uma linguagem formal, as sequências genômicas que possuem o marcador TATA.

EXERCÍCIOS EXTRA-CLASSE

- Considerando que α , β e γ denotem palavras e que ε seja a palavra vazia, qual(is) alternativa(s) está(ão) correta(s)?
 - $\alpha(\beta\gamma)=(\alpha\beta)\gamma$
 - $\alpha\varepsilon=\varepsilon\alpha=\alpha$
 - $|\alpha\beta|=|\alpha|+|\beta|$
 - $|\alpha^R|=|\alpha|$
 - A palavra vazia é sufixo e prefixo de qualquer palavra.
 - Uma palavra α , $\alpha \neq \varepsilon$, é sufixo e prefixo de uma palavra β se e somente se $\alpha=\beta$.
- Considere-se o alfabeto $\Sigma=\{0,1\}$. Construa uma linguagem formal sobre o alfabeto Σ que denote todos os números inteiros ímpares sem sinal.
- Considere-se o alfabeto $\Sigma=\{a,b\}$. Construa uma linguagem formal sobre o alfabeto Σ que represente somente as palavras sobre Σ que tenham o segmento ab .
- O Código Morse é um sistema de comunicação via telegrafia, que utiliza sons curtos (ponto) e sons longos (traço) para formação das letras:

Letra	Código Internacional	Letra	Código Internacional
A	. -	N	- .
B	- . . .	O	---
C	- . - .	P	. - - .
D	- . .	Q	-- . -
E	.	R	. - .
F	. . - .	S	. . .
G	-- .	T	-
H	U	. . -
I	. .	V	. . . -
J	. - - -	W	. - -
K	- . -	X	- . . -
L	. - . .	Y	- . - -
M	--	Z	-- . .

- Qual o menor alfabeto para o Código Morse segundo a descrição acima ?

- b) Como seria codificada a mensagem de socorro internacional MAYDAY neste alfabeto ?