

## Sistemas Multimédia

2022/2023

### Guião 08

#### I. Codificação de Informação

1. Carregue para o *workspace* do MATLAB a variável **Mensagem** que se encontra no Moodle. Esta variável contém uma sequência de caracteres (*string*) que compõem uma determinada mensagem. Desenvolva a função no MATLAB:

$\text{Simbolos} = \text{Alfabeto1}(\text{Texto})$

que devolve, na variável *Simbolos*, a lista de caracteres que são usados para compor o texto passado na *string* *Texto*, e teste essa função identificando a lista de caracteres considerados no texto de *Mensagem*.

2. Desenvolva, agora, a função:

$[\text{Simbolos}, \text{Frequencia}] = \text{Alfabeto2}(\text{Texto})$

que adicione à função anterior a capacidade de devolver, também, a frequência com que cada um dos símbolos identificados aparece no texto da variável passada como argumento, e teste-a com o texto de *Mensagem*.

3. Defina um **código binário** para os símbolos identificados nas alíneas anteriores, sendo que todos os códigos têm o mesmo número de bits (devendo esse número ser o menor possível). Quantos bits seriam necessários para representar o texto contido em *Mensagem*? Caracterize essa codificação quanto à sua ambiguidade e instantaneidade.

4. Considere, agora, um outro código binário onde: i) os símbolos são ordenados por ordem decrescente da sua frequência de ocorrência; ii) o símbolo de índice  $k$  é composto por  $k$  bits todos iguais a 1, exceto o último que é sempre 0. Desenvolva a função

$\text{NumBits} = \text{NumeroBits}(\text{Texto})$

que, podendo fazer usos das funções desenvolvidas anteriormente, determina o número de bits necessário para representar a mensagem de texto contida no argumento *Texto*. Teste-a com o texto contido em *Mensagem*. Caracterize essa codificação quanto à sua ambiguidade e instantaneidade. Compare esta codificação com a da alínea anterior quanto à **eficiência**.

5. Construa a **árvore binária** correspondente a cada um dos códigos anteriores.

6. Desenvolva a função:

$H = \text{Entropia}(\text{Texto})$

que calcula a entropia (em bits por símbolo) associada ao texto contido na variável *Texto* (tendo por base a estatística dada por esse mesmo texto). Esta função poderá fazer uso de funções desenvolvidas anteriormente. Calcule a entropia do texto contido em *Mensagem*.