

Teoria de Informação

Ficha Teórico-Prática nº2

“Codificação”

Período de execução: 2 semanas

Objectivo: Pretende-se que o aluno adquira sensibilidade para as questões relacionadas com a teoria dos códigos.

Trabalho

1. Considere os códigos cujos conjuntos de palavras são os seguintes:

- a. $\{0, 10, 11\}$
- b. $\{00, 11\}$
- c. $\{00, 01, 10, 100\}$
- d. $\{00, 01, 10, 110\}$
- e. $\{0, 10, 110, 111\}$.

Para cada código, diga, justificando, se é instantâneo ou não e se pode ser um código ótimo para uma dada fonte.

2. Considere um alfabeto composto pelos símbolos $\{1,2,\dots,10\}$ com as probabilidades de ocorrência que se indicam na tabela.
- a. Determine o majorante do bit-rate garantido pela aplicação de uma codificação com base numa árvore de Huffman estática e com base num código aritmético em que em cada TAG sejam transmitidos 4 símbolos.
 - b. Determine a TAG a que corresponderá a sequência “1 2 3 10”.

Símbolo	Prob de ocorrência
1	1/4
2	1/4
3	1/8
4	1/16
5	1/16
6	1/32
7	1/32
8	1/32
9	1/32
10	1/8

3. Considere uma fonte de informação $S=\{1,2,3,4\}$, cujos símbolos se encontrem distribuídos de acordo com a seguinte distribuição estatística:

Símbolo	Probabilidade
1	0,6
2	0,2
3	0,15
4	0,05

Assumindo um código aritmético, qual deverá ser a TAG para codificar a seguinte sequência: “1112”?

4. Considere uma fonte de informação com o seguinte alfabeto: $\{1,2,3,4,5,6,7,8,9,10\}$. Admitindo que a sequência dos primeiros símbolos da referida fonte são “1,1,2,3,4,1,2,5” indique a sequência de bits transmitidos para cada um dos símbolos codificados. Para o efeito considere que o VLC é implementado com base no algoritmo de Huffman adaptativo. Em cada iteração, indique o estado da árvore de Huffman.
5. Considere uma fonte de informação composta pelos 26 caracteres do alfabeto anglo-saxónico. Considere que se pretende codificar a mensagem “bbsdgbsh” de forma adaptativa. Indique para cada símbolo da cadeia de caracteres (i) o código binário a transmitir e (ii) o estado da árvore sibilante.
6. Considere a seguinte cadeia de símbolos: “barrayar#bar#by#barrayar#bay”. Assuma que $C(a)=1$, $C(b)=2$, $C(\#)=3$, $C(r)=4$ e que $C(y)=5$
 - a. Codifique a sequência usando LZW.
 - b. Codifique a sequência usando LZ77. Para o efeito considere que a dimensão da janela é 30 e que o look-ahead buffer é de 15 símbolos. Observe que inicialmente o buffer de pesquisa é vazio.
7. Codifique a seguinte sequência usando LZ78: “barrayar#bar#by#barrayar#bay”. Para o efeito considere que $C(a)=1$, $C(b)=2$, $C(\#)=3$, $C(r)=4$ e que $C(y)=5$.
8. Assinale a(s) resposta(s) certa(s).
 - a) ☐ - A diferença entre um código Huffman e um código LZ77 é que o segundo é mais eficiente qualquer que seja o contexto.
☐ - O código aritmético e o código Fano-Elias têm o mesmo princípio.
☐ - A relação entre o LZW e o LZ77 é que o segundo não consegue capturar os padrões locais.
☐ - Nenhuma das anteriores está correcta
 - b) ☐ - O código $\{0,10,110\}$ é ótimo.
☐ - O código $\{0,01,011\}$ é instantâneo
☐ - O código $\{0,01,011\}$ é unicamente decodificável
☐ - Nenhuma das anteriores está correcta
- 9 - Considerando o alfabeto (abcdefghijklmnopqrstuvwxyz), com indexação 0, 1, 2, ..., 25, indique a sequência final de índices resultante da aplicação do algoritmo Move to Front à palavra “**pleeease**”

<input type="checkbox"/> 16, 13, 7, 1, 1, 4, 19, 3	<input type="checkbox"/> 15, 12, 6, 0, 0, 3, 18, 2
<input type="checkbox"/> 15, 12, 6, 0, 0, 3, 20, 2	<input type="checkbox"/> Nenhuma das anteriores
- 10 – Considere uma fonte de informação descrita pelo alfabeto $A=\{A, B, C, D, E, F, G, H, I, J, K\}$. Pretende-se codificar a fonte usando uma árvore de Huffman adaptativa.
 - a) O código pré-acordado para o símbolo **G** é:

<input type="checkbox"/> 0101	<input type="checkbox"/> 011
<input type="checkbox"/> 010	<input type="checkbox"/> 0011

- b) Assumindo um algoritmo PPM com contexto de ordem 1, indique a codificação dos 3 primeiros símbolos da sequência referida em a).

11 - Considere uma fonte de informação pertencente ao dicionário $S=\{1,2,3,4,5\}$.

Assumindo que os símbolos são todos equiprováveis, indique a sequência de bits resultante da codificação da sequência “1233554424” usando um código aritmético inteiro.

- a) Indique o bitstream resultante da codificação da seguinte mensagem: HHFGHF

☐ 10010010100011001

☐ 10010010100011101

☐ 10010001010111000101

☐ Nenhum dos anteriores

- b) Assumindo um algoritmo PPM com contexto de ordem 1, indique a codificação dos 3 primeiros símbolos da sequência referida em a).