

Tecnologias Multimédia

Aula 6

Manuela. Pereira
mpereira@di.ubi.pt

April 14, 2020

1. A tabela abaixo mostra um código para a fonte discreta S, sem memória, com alfabeto de cinco símbolos e respectivas probabilidades:

símbolo	s_0	s_1	s_2	s_3	s_4
probabilidade	0.4	0.15	0.15	0.15	0.15
código	1	10	010	001	000

- (a) Estabeleça um código de Huffman para esta fonte.
 - (b) Estabeleça um código de Shannon-Fano para esta fonte.
 - (c) Mostre qual destes códigos é o mais eficiente. Se a fonte gerasse 1000 símbolos qual a diferença de débito entre as duas diferentes codificações?
 - (d) Compare o melhor dos seus códigos (Huffman e Shannon-Fano) com o código da tabela. Comente a eficiência e a possibilidade de utilização de cada um dos códigos (o seu melhor e o da tabela).
2. Uma fonte com alfabeto pertencente ao conjunto de caracteres { 'a', 'b', 'c'}. gerou a seguinte sequência: *aaaaaaaaabc*
 - (a) Codifique a sequência acima com o algoritmo Huffman.
 - (b) Comparando a codificação de Huffman com a de Shannon, qual seria mais eficiente a codificar esta mensagem? Justifique.
 3. Seja

$$X = \begin{cases} 1 & \text{com probabilidade } p, \\ 0 & \text{com probabilidade } 1 - p. \end{cases}$$

Calcule a entropia de X . Indique para que valores de p a entropia é máxima e para que valores de p a entropia é mínima.

4. Uma dada fonte transmite $r = 3000$ símbolos por segundo de um conjunto de quatro símbolos, com as probabilidades que se seguem:

x_i	$p(x_i)$
A	1/3
B	1/3
C	1/6
D	1/6

- (a) Qual o menor débito (em bits) de informação possível (por segundo, considerando codificação por símbolo).
- (b) Considere as codificações estatísticas por símbolo abordadas na disciplina. Com qual consegue uma codificação que permita um débito de informação mais próximo do menor débito possível? Qual o débito conseguido?
5. Uma fonte gera símbolos do alfabeto $\{A, B, C, D, E, F\}$, com probabilidades $\{0.3, 0.2, 0.15, 0.15, 0.15, 0.05\}$. Qual dos seguintes códigos binários é óptimo para esta fonte?
- ☐ $\{00, 01, 10, 1100, 1101, 1110\}$
- ☐ $\{10, 11, 011, 010, 001, 000\}$
- ☐ $\{10, 01, 011, 010, 001, 000\}$
- ☐ $\{00, 01, 10, 110, 1110, 1111\}$
- ☐ Nenhuma das anteriores.
6. Um ficheiro codificado com Huffman canónico continha a seguinte informação:

4 DIMRTV

3 .AESU

101111011000010101010001100011101110010011100001000011011001010000011110

- (a) Descodifique a mensagem anterior considerando que os ficheiros codificados com o codificador em causa apresentam a seguinte estrutura:
- N linhas começadas por um número. Nestas linhas, aparece um número inteiro seguido de um espaço e de uma lista de caracteres. O número inteiro corresponde ao número de bits usados para codificar os caracteres que aparecem na lista.
 - uma linha vazia
 - o resultado da codificação da mensagem original
- (b) Calcule a largura média, a entropia e a taxa de compressão.

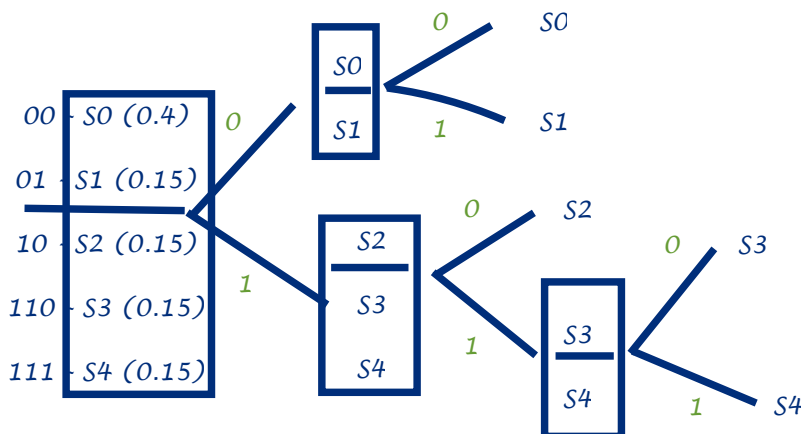
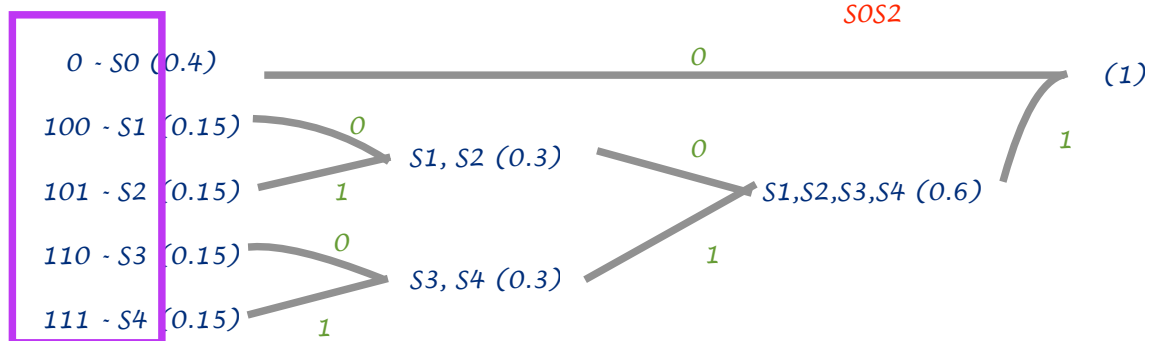
Resolução - pg1

$H = -\sum_i p_i * \log_2(p_i)$, sendo p_i a probabilidade associada ao símbolo i .

$L = \sum_i p_i * l_i$, sendo l_i o tamanho de código associado ao símbolo i .

símbolo	s_0	s_1	s_2	s_3	s_4
probabilidade	0.4	0.15	0.15	0.15	0.15
código	1	10	010	001	000

S1S1
1010
1010
S0S2



$$L_H = 0.4 * 1 + 0.15 * 3 * 4 = 2.2 \quad M_{1000} = 2.2 * 1000 = 2200 \text{ bits}$$

$$L_S = 2 * 0.7 + 3 * 0.3 = 2.3 \quad M_{1000} = 2.3 * 1000 = 2300 \text{ bits}$$

$$H = -0.4 * \log_2(0.4) - 0.15 * \log_2(0.15) * 4 = 2.171;$$

Resolução - pg2

aaaaaaaaabc

$$\begin{aligned} H_2 &= -10/12 * \log_2(10/12) - 2 * 1/12 * \log_2(1/12) = 0.8167; \\ L_2 &= 1 * 10/12 + 2 * 2/12 = 1.1667; \end{aligned}$$

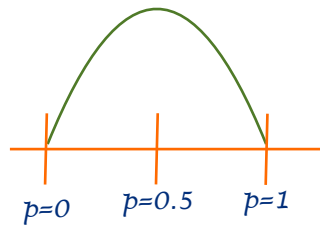
Código de Huffman: { 0, 10, 11 }

Código de Shannon: { 0, 10, 11 }

Resolução - pg3

$$X = \begin{cases} 1 & \text{com probabilidade } p, \\ 0 & \text{com probabilidade } 1 - p. \end{cases}$$

```
p=0.01:0.01:1;  
H = -p.*log2(p) - (1-p).*log2(1-p);  
plot(p,H)
```



Resolução - pg4

x_i	$p(x_i)$
A	1/3
B	1/3
C	1/6
D	1/6

$$H_4 = -1/3 * \log_2(1/3) * 2 - 1/6 * \log_2(1/6) * 2 = 1.9183;$$

$$L = 2$$

*Para 3000 símbolos, 2*3000 bits, ou seja 6000 bits.*

Resolução - pg5

Uma fonte gera símbolos do alfabeto $\{A, B, C, D, E, F\}$, com probabilidades $\{0.3, 0.2, 0.15, 0.15, 0.15, 0.05\}$. Qual dos seguintes códigos binários é ótimo para esta fonte?

- ☐ $\{00, 01, 10, 1100, 1101, 1110\}$
- ☒ $\{10, 11, 011, 010, 001, 000\}$
- ☐ $\{10, 01, 011, 010, 001, 000\}$ ✗
- ☐ $\{00, 01, 10, 110, 1110, 1111\}$
- ☐ Nenhuma das anteriores.

Resolução - pg6