Teoria da Informação (2004-2005) Exame de 1ª chamada (29-6-05)

# Resolução

**1.** Matriz de transição do canal:

$$\begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 1/2 & 0 & 0 & 1/2 \end{bmatrix}$$

a) Determinação das probabilidades de $Y$:

$$P(Y=0)=P(y_0)=\frac{1}{2}P(x_0)+\frac{1}{2}P(x_3)=\frac{1}{8}+\frac{1}{6}=\frac{7}{24}$$

$$P(Y=1)=P(y_1)=\frac{1}{2}P(x_0)+\frac{1}{2}P(x_1)=\frac{1}{8}+\frac{1}{12}=\frac{5}{24}$$

$$P(Y=2)=P(y_2)=\frac{1}{2}P(x_1)+\frac{1}{2}P(x_2)=\frac{1}{12}+\frac{1}{8}=\frac{5}{24}$$

$$P(Y=3)=P(y_3)=1-\left[P(y_0)+P(y_1)+P(y_2)\right]=\frac{7}{24}$$

Entropia de $Y$: $H(Y)=H\left(\frac{7}{24},\frac{5}{24},\frac{5}{24},\frac{7}{24}\right)=1{,}980$ bits/símbolo.

b) De $I(X;Y)=H(Y)-H(Y\mid X)=H(X)-H(X\mid Y)$ tiramos a equivocação $H(X|Y)$:

$$H(X\mid Y)=H(X)-H(Y)+H(Y\mid X).$$

Ora $H(Y\mid X)=\sum_{i=0}^{3}P(x_i)H(Y\mid x_i)=\sum_{i=0}^{3}P(x_i)\underbrace{H(1/2,1/2,0,0)}_{1}=1$

e $H(X)=H\left(\frac{1}{4},\frac{1}{6},\frac{1}{4},\frac{1}{3}\right)=1{,}959$. Logo,

$$H(X\mid Y)=H(X)-H(Y)+H(Y\mid X)=1{,}959-1{,}980+1=0{,}979 \text{ bits/símbolo.}$$

c) $C=\max I(X;Y)=\max\left[H(Y)-H(Y\mid X)\right]$.

Viu-se que $H(Y|X)=1$. Portanto, $C=\max\left[H(Y)\right]-H(Y\mid X)=\max\left[H(Y)\right]-1$. O máximo de $H(Y)$ é $\log_2 4=2$ bits/símbolo e atinge-se se os símbolos $y_i$ forem equiprováveis. Mas só o são para certas distribuições de probabilidades de $X$. Atendendo à simetria do canal uma delas (mas não a única) tem as probabilidades iguais (símbolos $x_i$ equiprováveis): (1/4, 1/4, 1/4, 1/4). A capacidade deste canal quaternário é, pois, $C = 2 - 1 = 1$ bit/símbolo.

Nota: a distribuição optimizada de probabilidades $(p_1, p_2, p_3, p_4)$ de $X$ com a qual $P(y_i) = 1/4$ e se atinge a capacidade é $p(1/2, 0, 1/2, 0) + (1-p)(0, 1/2, 0, 1/2)$, com $p \in [0,1]$. A distribuição equiprovável de $X$, (1/4, 1/4, 1/4, 1/4), atinge-se com $p = 1/2$.

**2.** a) O número ISBN 0-87805-748-**A** é calculado de modo que

$$1 \times 0 + 2 \times 8 + 3 \times 7 + 4 \times 8 + 5 \times 0 + 6 \times 5 + 7 \times 7 + 8 \times 4 + 9 \times 8 + 10 \times A = 252 + 10A$$

seja um múltiplo de 11, isto é, que $(252 + 10A) \bmod 11 = \underbrace{252 \bmod 11}_{10} + 10A \bmod 11 = 0 \bmod 11$,

obrigando a que $10A$ mod 11 = 1. Não será difícil concluir que $A = 10$. Sempre que o dígito de controlo é 10 ele é substituído pelo carácter X. Portanto o número ISBN completo é 0-87805-748-**X**.

Nota: este ISBN pertence ao livro “The Color Curtain”, de Richard Wright, Banner Books, University Press of Mississippi/Jackson, 1956.

b) Os dois números têm os algarismos 2 e 9 trocados.

Verificação de 0-306-81296-7:

$$1 \times 0 + 2 \times 3 + 3 \times 0 + 4 \times 6 + 5 \times 8 + 6 \times 1 + 7 \times 2 + 8 \times 9 + 9 \times 6 + 10 \times 7 = 286.$$

286 é múltiplo de 11, logo, o número ISBN dado é válido.

Verificação de 0-306-81926-7:

$$1 \times 0 + 2 \times 3 + 3 \times 0 + 4 \times 6 + 5 \times 8 + 6 \times 1 + 7 \times 9 + 8 \times 2 + 9 \times 6 + 10 \times 7 = 279.$$

279 **não é** múltiplo de 11, logo, este número **não é** válido.

**3.** O código de comprimento máximo (15, 4) tem, de acordo com o enunciado, um polinómio enumerador de pesos $A(z) = 1 + (2^k - 1)z^{2^{k-1}} = 1 + 15z^8$. Logo, $A_0 = 1$, $A_8 = 15$ e os restantes $A_i$ são nulos.

a) Como todas as palavras de código excepto a nula têm peso 8 concluímos que a palavra de código dada, 11001_ _ _0010100, terá de ser completada com três bits 1, ou seja, 11001**111**0010100.

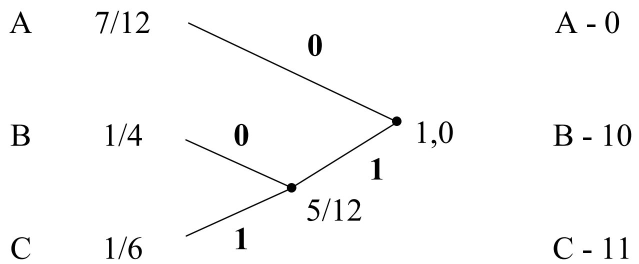
b) Do que se disse atrás conclui-se que $d_{min} = 8$.

c) $P_{end} = A_8 p^8 (1-p)^7 = 1{,}4.10^{-15}$ (com $A_8 = 15$ e $p = 10^{-2}$).

**4.** Sequência a codificar: ABCA. Probabilidades: $P(A)$=7/12, $P(B)$ = 1/4 e $P(C)$ = 1/6.

a) Entropia: $H(7/12, 1/4, 1/6) = 1{,}3844$ bits/símbolo

b) Codificação de Huffman:

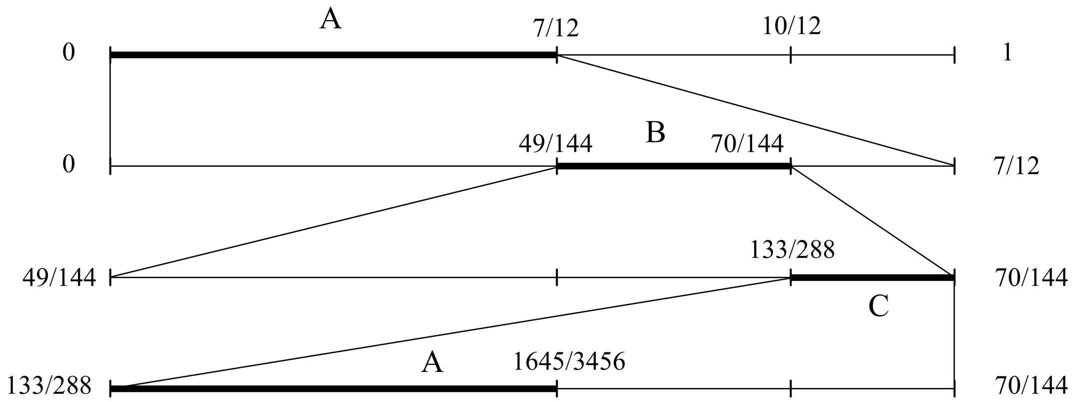



Sequência codificada: 0 10 11 0

$$\bar{N} = \frac{7}{12}\times 1 + \frac{1}{4}\times 2 + \frac{1}{6}\times 2 = \frac{17}{12} = 1{,}4167 \text{ binits/símbolo}$$

Eficiência da codificação: $\frac{H}{\bar{N}} = \frac{1{,}3844}{1{,}4167} = 97{,}7\%$ (se H = 1,4 $\Rightarrow \frac{H}{\bar{N}} = \frac{1{,}4}{1{,}4167} = 98{,}8\%$ )

c) Determinação da fracção diádica em codificação aritmética. Primeiro é preciso encontrar o intervalo semi-aberto de codificação.




Intervalo final: $[\alpha;\beta[ \ = \ [\frac{133}{288};\ \frac{1645}{3456}[ \ = \ [0{,}4618;\ 0{,}4760[ \quad \Rightarrow \quad L = 0{,}0142$

Fracção diádica: $\alpha \le \frac{x}{2^t} < \beta$, com $t = \lceil -\log_2 L \rceil = \lceil 6{,}14 \rceil = 7$.

$$0{,}4618 \le \frac{x}{2^7} < 0{,}4760 \quad \Rightarrow \quad 59{,}11 \le x < 60{,}93 \quad \Rightarrow \quad x = 60$$

A fracção diádica com menor denominador é, portanto, $\frac{x}{2^t} = \frac{60}{128} = \frac{15}{32}$

d) A representação binária de 15/32 é, com $\log_2 32 = 5$ bits, **01111**.

Como se fez? Representando o numerador, 15, com cinco bits. Outra maneira seria multiplicar 15/32 sucessivamente por 2: se o resultado for inferior a 1 toma-se um 0, se for superior ou igual a 1 toma-se um 1. Nesta última situação subtrai-se 1 ao resultado e retoma-se a multiplicação por 2, prosseguindo até termos cinco bits.

**5.** Código (8, 4) de matriz geradora

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

a) Matriz de verificação de paridade, **H**:

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

b) A distância mínima do código não será superior a 4 pois a matriz **G**, que contém palavras do código, tem linhas de peso 4. Olhando para a matriz **H** vemos que não tem linhas nulas nem iguais (logo, $d_{\min} \geq 3$). Observando melhor vemos que se somarmos as linhas 1, 3 e 6, por exemplo, o resultado é zero. Concluímos, portanto, que $d_{min} = 3$ (número mínimo de linhas de **H** de soma nula).

c) As equações de paridade podem obter-se das colunas da submatriz **P** ou das colunas da matriz **H**:

$$\begin{aligned} x_1 + x_2 + x_3 + x_5 &= 0 \\ x_1 + x_2 + x_4 + x_6 &= 0 \\ x_1 + x_3 + x_4 + x_7 &= 0 \\ x_1 + x_2 + x_3 + x_4 + x_8 &= 0 \end{aligned}$$

Substituindo $x_i$ pelo bit $y_i$ da sequência $\mathbf{y} = y_1y_2\ldots y_8 = 11010001$ obtemos

$$\begin{aligned} 1+1+0+0 &= 0 \\ 1+1+1+0 &= 1 \\ 1+0+1+0 &= 0 \\ 1+1+0+1+1 &= 0 \end{aligned}$$

Portanto a síndrome pedida é **S** = 0100. Este resultado poderá ser confirmado multiplicando **y** pela matriz **H**.

d) A matriz geradora do código encurtado obtém-se da original retirando-lhe a 3ª linha e a 3ª coluna:

$$\tilde{\mathbf{G}} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

A distância mínima de um código encurtado não é inferior à distância mínima do código original (3, neste caso (ver alínea b)). Sendo a matriz $\tilde{\mathbf{H}}$ igual a

$$\tilde{\mathbf{H}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

concluímos que $d_{min} = 3$ no código encurtado pois existem três linhas de **H** de soma nula (por exemplo as linhas 1, 3 e 4).

e) O polinómio gerador satisfaz $g(p)\,\text{mod}(p^3+p^2+p+1) = p+1$, por hipótese. Sabemos que $g(p)$ é um factor de $p^n + 1$ de grau $n - k$. Neste caso $n = 7$ e o grau é 4. Consultando a tabela de factores de $p^n + 1$ vemos que $p^n + 1$ = 3.13.15 = 011.001011.001101, isto é,

$$p^7+1 = \underbrace{(p+1)}_{y_1}\underbrace{\left(p^3+p+1\right)}_{y_2}\underbrace{\left(p^3+p^2+1\right)}_{y_3}.$$

Temos duas hipóteses de polinómio, $y_1y_2$ ou $y_1y_3$, mas só uma satisfaz a condição dada. Vamos ver qual:

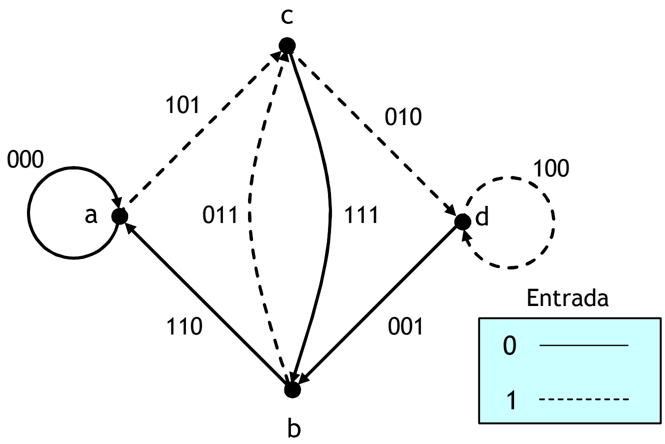
$$g_1(p) = p^4+p^3+p^2+1 \qquad \Rightarrow \qquad g_1(p)\,\text{mod}(p^3+p^2+p+1) = p+1 \quad \text{(Serve)}$$

$$g_2(p) = p^4+p^2+p+1 \qquad \Rightarrow \qquad g_2(p)\,\text{mod}(p^3+p^2+p+1) = p^2+p \quad \text{(Não serve)}$$

Resposta: $g_1(p) = p^4+p^3+p^2+1$.

**6.** Matriz geradora: $\mathbf{G} = \begin{bmatrix} 7 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 111 & 011 & 110 \end{bmatrix} \Rightarrow G(D) = \begin{bmatrix} 1+D+D^2 & D+D^2 & 1+D \end{bmatrix}$.

a) O codificador tem dois andares (maior grau dos polinómios geradores) e, portanto, dois bits de estado pelo que tem 4 estados. O seu diagrama de estados é o seguinte:




b) Condição necessária e suficiente para que o codificador não seja catastrófico: $m.d.c.\left[g_1(D), g_2(D), g_3(D)\right] = D^m$, $m \geq 0$. O máximo divisor comum é:

$$m.d.c.\left(1+D+D^2, D+D^2, 1+D\right) = m.d.c.\left(1+D+D^2, \underbrace{m.d.c.\left(D+D^2, 1+D\right)}_{1+D}\right) = 1$$

(pois $1+D+D^2$ e $1+D$ não são factorizáveis). Concluímos que o codificador não é catastrófico.

c) Codificador RSC equivalente: $\tilde{G}(D) = \begin{bmatrix} 1 & \dfrac{D+D^2}{1+D+D^2} & \dfrac{1+D}{1+D+D^2} \end{bmatrix}$