FEUP Universidade do Porto
Faculdade de Engenharia

Departamento de Engenharia Electrotécnica e de Computadores

Exercícios de

Teoria da Informação

Sílvio A. Abrantes

2000

1. Teoria de Shannon e códigos de fonte

1.1. Uma fonte produz letras estatisticamente independentes e equiprováveis, extraídas de um alfabeto (a_1, a_2) à velocidade de 1 letra em cada 3 segundos. Essas letras são transmitidas através de um canal binário simétrico, usado uma vez em cada segundo codificando a letra de fonte a_1 na palavra de código **000** e codificando a_2 na palavra de código **111**. Se, na saída e no intervalo de 3 segundos correspondente, qualquer das sequências **000**, **001**, **010** e **100** for recebida, a_1 é descodificado; senão, a_2 é descodificado. Seja $\varepsilon < 1/2$ a probabilidade de erro do canal.

a) Para cada sequência de 3 dígitos possível, recebida no intervalo correspondente a uma dada letra da fonte, determine a probabilidade de a_1 ter sido produzido pela fonte dada essa sequência recebida.

b) Usando a alínea a) mostre que a regra de descodificação acima descrita minimiza a probabilidade de uma decisão incorrecta.

c) Determine a probabilidade de uma decisão incorrecta *(usar a alínea a) não é a melhor maneira!)* .

1.2. Numa população feminina **X**, consistindo em 1/4 de loiras, 1/2 de morenas e 1/4 de ruivas, as loiras chegam sempre a horas nos seus compromissos, as ruivas chegam sempre atrasadas e cada morena lança sempre uma moeda ao ar para decidir se há-de chegar atrasada ou não.

a) Que quantidade de informação é fornecida pela afirmação "x, um membro de **X**, chegou a horas" relativamente a cada uma das seguintes proposições:

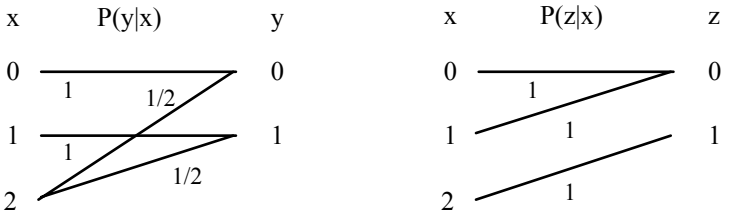
(1) x é loira,

(2) x é morena,

(3) x é ruiva?

b) Que quantidade de informação é fornecida pela afirmação "x, um membro de **X**, chegou a horas três vezes seguidas" relativamente à proposição "x é morena"?

1.3. Uma fonte **X** produz letras de um alfabeto de três símbolos com as probabilidades $P_X(0)=1/4$, $P_X(1)=1/4$ e $P_X(2)=1/2$. Cada letra *x* da fonte é transmitida directa e simultaneamente através de dois canais com saídas *y* e *z* e com as probabilidades de transição indicadas.

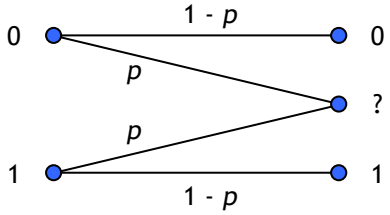


(Note que estes dois canais poderiam ser considerados como um único canal com saída yz).

Calcule H(X), H(Y), H(Z), H(YZ), I(X,Y) e I(X,Z). Interprete as expressões da informação mútua.

1.4. Uma fonte gera os símbolos 0 e 1, cada um com a duração de 2 e 4 segundos, respectivamente, sendo o número de 1's metade do número de 0's. Determine a taxa de informação desta fonte.

1.5. Na figura está representado um canal discreto designado por *binary erasure channel* (ou BEC).



Probabilidade de transmitir um "0": p_0

Probabilidade de transmitir um "1": $p_1 = 1 - p_0$

a) Determine a informação mútua média do canal quando $p_O = 1/4$ e $p = 0{,}1$.

b) Represente graficamente a capacidade do canal em função de p.

c) Determine a capacidade do canal, se $p = 0{,}25$.

1.6. Imagine uma máquina de escrever com 26 teclas.

a) Se ao carregar numa tecla a letra correspondente é escrita, determine a capacidade C em bits/letra.

b) A máquina está-se a avariar: ao carregar numa tecla a letra correspondente, ou a seguinte, é escrita (isto é, A $\rightarrow$ A ou B, …, Z $\rightarrow$ Z ou A. Quanto vale a nova capacidade?

1.7. Um canal tem à entrada e à saída símbolos do conjunto {0, 1, 2, 3, 4}. As probabilidades de transição são da forma

$$p(y|x) = \begin{cases} 1/2 & \text{se } y = x \pm 1 (\text{mod } 5) \\ 0 & \text{outros valores} \end{cases}$$

Determine a capacidade do canal.

1.8. Um conjunto de oito palavras equiprováveis é codificado no seguinte conjunto de oito palavras de código, para transmissão através de um canal binário simétrico com probabilidade de transição p = 0,2:

$x_1 = 0000$	$x_5 = 1001$
$x_2 = 0011$	$x_6 = 1010$
$x_3 = 0101$	$x_7 = 1100$
$x_4 = 0110$	$x_8 = 1111$

Se a sequência de dígitos y = 0000 for recebida à saída do canal, determine a quantidade de informação que o primeiro dígito recebido fornece acerca de x_1.

1.9. Encontraram-se as seguintes probabilidades num canal discreto binário: P(1|0) = 1/4; P(0|1) = 1/2. Se P(0) = 2/5, determine a informação mútua média do sistema.

1.10. Uma série de cinco jogos entre duas equipas termina logo que uma delas ganhe três vezes. Seja X a variável aleatória que representa o resultado dos jogos entre as equipas A e B; exemplos de valores possíveis de X são AAA, $BABAB$ e $BBAAA$. Seja Y o número de jogos jogados (Y=3, 4 ou 5).

a) Admitindo que as duas equipas tem igual nível competitivo e que os jogos são independentes, calcule $H(X)$, $H(Y)$, $H(Y|X)$ e $H(X|Y)$.

b) Seja Z a equipa vencedora. Determine $H(X|Z)$ e compare com $H(X)$. Determine ainda $H(Z|X)$.

1.11. Considere o seguinte método de construção de palavras de código de fonte binárias para um conjunto de mensagens A com probabilidades de símbolos $P(a_i)$: seja $P(a_k) \leq P(a_j)$ para $k > j \geq 1$ e defina-se

$$Q_i = \sum_{k=1}^{i-1} P(a_k) \quad \text{para } i > 1; Q_1 = 0$$

A palavra de código atribuida à mensagem a_i é formada determinando a expansão "decimal" de $Q_i < 1$ no sistema binário (isto é, $1/2 \rightarrow 1000\ldots$, $1/4 \rightarrow 01000\ldots$, $5/8 \rightarrow 101000\ldots$) e depois truncando esta expansão aos primeiros n_i dígitos, em que n_i é o inteiro igual ou imediatamente superior a $I(a_i)$ bits.

a) Construa as palavras binárias de código para o conjunto de oito mensagens que ocorrem com as probabilidades 1/4, 1/4, 1/8, 1/8, 1/16, 1/16, 1/16 e 1/16.

b) Prove que o método acima descrito origina em todos os casos um conjunto de palavras de código satisfazendo a condição de prefixação e cujo comprimento médio $\overline{N}$ satisfaz o teorema da codificação de fonte $H(A) \leq \overline{N} < H(A) + 1$.

1.12. Verifique se existe um código binário com palavras de código de comprimentos 1, 2, 3, 3 e 4 que satisfaça a condição de prefixação.

1.13. Uma fonte tem um alfabeto de quatro letras. Em baixo apresentam-se as probabilidades das letras e dois conjuntos possíveis de palavras de código binárias:

Letras	Probabilidades	Código I	Código II
a_1	0,4	1	1
a_2	0,3	01	10
a_3	0,2	001	100
a_4	0,1	000	1000

Para cada código responda às seguintes questões:

a) O código satisfaz a condição de prefixação?

b) O código é unicamente descodificável?

c) Suponha que a primeira letra da palavra de código é "1"**.** Qual é a informação mútua que esta ocorrência fornece acerca do acontecimento "a letra da fonte é a_1"?

d) Qual é a informação mútua média que a especificação da primeira letra da palavra de código fornece relativamente à letra da fonte?

1.14. Considere uma fonte com K=3 e probabilidades dos símbolos 0,6, 0,3 e 0,1.

a) Obtenha o código de Shannon-Fano e calcule a sua eficiência.

b) Repita a alínea a) com um código de extensão 2 (agrupando os símbolos em blocos de dois).

1.15. Uma fonte binária tem símbolos com probabilidades 0,8 e 0,2.

a) Agrupe os símbolos em blocos de dois, obtenha o correspondente código de Shannon-Fano de extensão 2 e calcule a sua eficiência.

b) Repita a alínea a) para o código de extensão 3.

1.16. Suponha que uma fonte discreta produz as cinco letras E, R, T, C e O com as probabilidades de ocorrência 0.5, 0.09, 0.15, 0.01 e 0.25, respectivamente.

a) Determine a entropia da fonte.

b) Determine a sequência original de letras que deu origem à sequência codificada 111110111011100111111010. O código usado foi o de Shannon-Fano, com E $\rightarrow$ '0'.

c) Determine, para o código referido, o número médio de bits por cada letra da fonte.

1.17. Considere duas fontes discretas sem memória. A fonte 1 tem um alfabeto de 6 letras com as probabilidades 0,3, 0,2, 0,15, 0,15, 0,12 e 0,08 e a fonte 2 tem um alfabeto de 7 letras com as probabilidades 0,3, 0,27, 0,13, 0,12, 0,08, 0,05 e 0,05. Para cada fonte:

a) Construa um código de Huffman binário. Determine o número médio de letras de código por letra de fonte.

b) Construa um código de Huffman ternário. Determine o número médio de letras de código por letra de fonte.

1.18. Determine um conjunto óptimo de comprimentos de palavras binárias $n_1, n_2, \ldots$ de um código instantâneo se as probabilidades dos símbolos da fonte forem dadas pelos seguintes conjuntos:

a) $p = \left(\frac{10}{41}, \frac{9}{41}, \frac{8}{41}, \frac{7}{41}, \frac{7}{41}\right)$.

b) $p = \left(\frac{9}{10}, \frac{9}{10}\times\frac{1}{10}, \frac{9}{10}\times\left(\frac{1}{10}\right)^2, \frac{9}{10}\times\left(\frac{1}{10}\right)^3, \ldots\right)$

1.19. Dão-lhe 6 garrafas de vinho, A, B, …, F, e dizem-lhe que numa delas o vinho está estragado (sabe mal). É-lhe dito ainda que a probabilidade de cada garrafa estar estragada é $(p_A, p_B, p_C, p_D, p_E, p_F) = \left(\frac{2}{23}, \frac{4}{23}, \frac{6}{23}, \frac{1}{23}, \frac{8}{23}, \frac{2}{23}\right)$. Pedem-lhe para encontrar o vinho estragado através de provas gustativas.

Vamos supor que experimenta uma garrafa de cada vez. Escolha a ordem de prova de forma a minimizar o número esperado de provas requeridas para escolher o mau vinho. É claro que se as primeiras 5 garrafas passarem o teste já não precisa de provar a sexta! Sendo assim:

a) Quantas garrafas espera provar?

b) Qual é a garrafa que deve experimentar primeiro?

Já concluiu que o método anterior não é o melhor e agora vai mudar de táctica: primeiro mistura alguns dos vinhos num copo e prova a mistura, depois prossegue misturando e provando até que o mau vinho tenha sido encontrado.

c) Qual é o número esperado mínimo de provas que tem de fazer desta vez?

d) Qual é a mistura que deve provar primeiro?

1.20. Uma fonte gera letras de um alfabeto α={A, E, I, O, U} com as probabilidades P(A)=P(I)=0,2, P(E)=0,4 e P(O)=P(U)=0,1.

a) Codifique as letras com um código de Huffman binário e determine o número médio de bits usado para cada letra.

b) Determine a variância do comprimento das palavras de código.

c) Repita a alínea a) procurando obter um código de Huffman de variância mínima.

d) Determine a variância do novo código.

1.21. Considere os dois códigos de Huffman da tabela seguinte:

Símbolo	Probabilidade	Código 1	Código 2
x_1	0,2	01	10
x_2	0,4	1	00
x_3	0,2	000	11
x_4	0,1	0010	010
x_5	0,1	0011	011

Poderá verificar que o número médio de bits/símbolo, $\overline{N}$, é igual em ambos os códigos e que a variância do código 2 é a menor das duas.

a) A sequência $x_2x_1x_3x_2x_1x_2x_4$ foi codificada com o código 1 e enviada através de um canal, que provocou um erro no primeiro bit da sequência binária (em vez de se receber um "1" recebeu-se um "0", ou vice-versa). Quantos caracteres errados ocorrem antes do primeiro correctamente descodificado?

b) Repita o mesmo para o código 2.

c) Repita as alíneas anteriores mas supondo agora que é o terceiro bit recebido que está errado.

1.22. O alfabeto de uma fonte é constituído por símbolos a_i que ocorrem com probabilidades

$$p(a_1)=0,4 \qquad p(a_2)=0,2 \qquad p(a_3)=0,03$$
$$p(a_4)=0,05 \qquad p(a_5)=0,02 \qquad p(a_6)=0,3$$

Pretende-se codificar os símbolos gerados pela fonte através de um código de Huffman ternário.

a) Determine as palavras de código usando um agrupamento prévio de símbolos. Determine $\overline{N}$ e a eficiência da codificação. Determine ainda a *redundância da codificação*, definida como a diferença entre $\overline{N}$ e o seu valor mínimo possível.

b) Repita a alínea a) mas sem recorrer a um agupamento prévio de símbolos. Compare os resultados com os anteriores.

1.23. Um alfabeto de entrada (por exemplo, um teclado de um processador de texto) consiste em 100 caracteres.

a) Se as teclas forem codificadas através de um código de comprimento fixo, determine o número requerido de bits para a codificação de cada tecla.

b) Suponhamos que 10 das teclas são equiprováveis e que cada uma ocorre com probabilidade 0,05. Suponhamos também que as restantes 90 teclas são batidas com igual probabilidade. Determine o número médio de bits requerido para codificar este alfabeto usando um código de Huffman.

1.24. Uma palavra foi codificada usando o código de Huffman, tendo-se obtido a sequência binária

1 0 1 1 1 0 1 1 0 1 0 1 1 1 0 0 1 1 1 0 1 0 0

O alfabeto original era constituido pelas letras A, B, C, D, E, I, L, R e T e a letra I foi codificada como "00". Supondo que estas letras ocorriam com as probabilidades

P(A) = 0,26	P(D) = 0,01	P(L) = 0,01
P(B) = 0,09	P(E) = 0,07	P(R) = 0,23
P(C) = 0,08	P(I) = 0,22	P(T) = 0,03

qual terá sido a palavra codificada?

1.25. Uma fonte ternária apresenta as seguintes probabilidades de ocorrência de símbolos: $p(a_1)=0,8$, $p(a_2)=0,02$ e $p(a_3)=0,18$. Codifique a sequência $a_1a_3a_2a_1$ usando codificação aritmética.

1.26. Dadas as probabilidades p(A)=0,2, p(B)=0,3 e p(C)=0,5, determine um valor real, usando codificação aritmética, que represente a sequência AACBCA.

1.27. Dadas as probabilidades $P(A) = 0,37$, $P(B) = 0,38$ e $P(C) = 0,25$, use codificação aritmética para determinar a menor sequência binária correspondente à sequência ABACABB.

1.28. Uma mensagem de seis símbolos $\{a_1, a_2, a_3\}$ é representada pelo número real 0,927430. Sabe-se que na fonte a probabilidade de ocorrência dos símbolos é $p(a_1) = 0,6$, $p(a_2) = 0,3$ e $p(a_3) = 0,1$. Descodifique a mensagem.

1.29. Uma fonte discreta possui um alfabeto de 10 símbolos $X_1, X_2, \ldots, X_{10}$ que ocorrem com as seguintes probabilidades:

{1/50, 2/50, 3/50, 4/50, 5/50, 5/50, 6/50, 7/50, 8/50, 9/50}

a) Codifique os símbolos da fonte com um código de Huffman ternário.
b) Determine o comprimento médio das palavras do código.
c) Imagine que em vez de um codificador de Huffman se dispunha de um codificador aritmético binário. Quantos bits seriam necessários para representar a sequência de símbolos $X_3X_2X_{10}X_2X_6$?

1.30. Com um alfabeto de 47 caracteres uma fonte gerou a mensagem (de onde as aspas não fazem parte)

"MUITO BEM, SÓ QUE QUEM VIU NÃO DIZ QUE VIU, DIZ QUE OUVIU ALGUÉM QUE VIU."

Esta mensagem vai ser codificada com um codificador LZ77 com uma janela de observação de 60 caracteres dos quais 10 pertencem ao "look-ahead buffer".

a) Quantos bits são necessários para representar cada palavra de código?
b) Suponha que o início da mensagem já foi codificado, de tal modo que no corpo da janela já se encontram 34 caracteres. Indique a sequência de apontadores que se obtém à saída do codificador a partir desse momento.

1.31. A sequência de apontadores (7, 3, C) (5, 2, C) (4, 2, B) (3, 5, B) (3, 4, C) apresenta-se à entrada de um descodificador LZ77 quando no corpo da sua janela se encontra a sequência descodificada AAAABABCCA. Descodifique a mensagem restante.

1.32. Considere a mensagem binária seguinte:

ABAABBABABAAAABABBABBABBAABABABABBBABAABAABABABABAB

Esta mensagem vai ser codificada usando o código LZ78 com um dicionário que inicialmente contém as entradas A e B (nas posições 1 e 2, respectivamente) e cujo tamanho máximo é 30.

a) Seccione a mensagem.
b) Construa uma tabela com as primeiras dez entradas do dicionário.
c) Quantos bits necessita para codificar toda a sequência?

1.33. Deseja-se codificar a mensagem AAAABABCCAABACAACCABCABCABCABCC com o código LZ78.

a) Seccione a mensagem.
b) Construa o dicionário de codificação partindo de um dicionário inicial contendo as letras A, B e C (por esta ordem).
c) Obtenha a sequência codificada.
d) Quantos bits são precisos por cada palavra de código de saída, admitindo que o dicionário tem capacidade para albergar 16 entradas?

1.34. A sequência binária 000101000100001010001100001001100010 representa uma dada mensagem codificada em LZ78. A fonte de mensagens é ternária (produz as letras A, B e C) e o dicionário de codificação, com um tamanho de 16 caracteres, foi desenvolvido a partir de um dicionário inicial contendo apenas aquelas letras, e por aquela ordem. Qual é a mensagem original?

1.35. Imagine que um veículo móvel equipado com uma câmara de TV a preto e branco foi proposto para explorar a superfície de Marte. As imagens de TV serão digitalizadas para serem transmitidas para Terra. A largura de banda B é tal que $B/R > 10$, em que R é o ritmo de transmissão. Deseja-se estimar o tempo requerido para transmitir uma imagem, dadas as seguintes especificações:

- Imagem digitalizada: n_p = 400x300 pixels (cada um com 16 níveis possíveis de luminância)
- Ligação Marte-Terra: microondas, com frequência de portadora $f_c = 2GHz$, distância l =3.10^8 km.
- Emissor do veículo: $S_T = 20W$; Antena do veículo: 1 m de diâmetro.
- Antena da estação terrena: 30 m de diâmetro; temperatura de ruído do receptor: T_N = 58K.

Considere as seguintes expressões:

- Sinal no receptor: $S = \frac{g_T g_R}{L} S_T$, em que g_T e g_R são os ganhos das antenas e L é a perda em espaço livre
- Ganho de uma antena parabólica: $g = \frac{4\pi A_e f^2}{c^2}$ A_e ≈ área da antena

- Perda em espaço livre: $L = \left(\frac{4\pi f l}{c} \right)^2$
- Densidade espectral de potência do ruído: $N_0 = 4.10^{-21} \frac{T_N}{T_0}$, $T_O = 300K$

2. Códigos detectores e correctores de erros: ARQ , códigos de blocos e códigos cíclicos

2.1. Num determinado sistema ARQ verifica-se que $t_d = 0{,}2$ms e que as mensagens são geradas à cadência de $r = 72$ kbits/s. A probabilidade de erro de transição do canal (binário simétrico) é $p = 10^{-3}$ e para detectar os erros usa-se um código de paridade dupla, com $k = 8$ e $n = 10$. Sabendo que as limitações do canal impõem que o ritmo de transmissão seja $r_b \leq 120$ kbits/s, mostre, recorrendo a R'_c (taxa de transferência), que o método "Go-Back-N" seria aceitável, ao contrário do método "Stop-and-Wait".

2.2. Considere o código BCH (2047, 2014), que corrige até 3 erros por palavra de código, associado a um sistema ARQ. Este código satisfaz

$$P_{end} \leq [1-(1-p)^k]2^{-(n-k)}$$

em que p representa a probabilidade de erro de transição do canal e P_{end} é a probabilidade de erro não detectado do código. Como um padrão de erros não detectável pode ocorrer na transmissão inicial de uma palavra ou em qualquer retransmissão, a probabilidade P(E) (probabilidade de uma palavra recebida ser aceite pelo sistema ARQ e um erro de descodificação ser cometido) vem dada por

$$P(E) = P_{end} + p_R P_{end} + p_R^2 P_{end} + \ldots = P_{end}(1 + p_R + p_R^2 + \ldots) =$$
$$= P_{end} \frac{1}{1-p_R} = \frac{P_{end}}{P(0,n) + P_{end}} \qquad p_R - \text{probabilidade de retransmissão}$$

Suponha então que o canal é de tal modo ruidoso que $p = 10^{-3}$. Verifique que, apesar disso e apesar de cada palavra conter poucos bits de paridade, $P(E) \leq 8.10^{-10}$, isto é, o sistema ARQ associado a este código BCH é muito fiável.

2.3. Num sistema de comunicações ARQ o número de palavras de código que podem ser transmitidas durante o "round-trip delay" da comunicação é de 4 e a probabilidade de erro é de 10^{-3}. Determine a taxa de transferência ("throughput") dos três métodos ARQ estudados, se se usar um código de Hamming (7, 4).

2.4. Considere um código de blocos sistemático (6,3) gerado pela submatriz **P**

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

Escreva as equações dos bits de paridade e construa uma tabela com as palavras de código e respectivos pesos, mostrando que $d_{min} = 3$.

2.5. Queremos obter códigos de blocos de comprimento $n = 255$ com capacidades de correcção $t = 1$, 2 ou 3. De acordo com o limite de Hamming, quantos bits de paridade, no mínimo, devemos usar para cada valor de t?

2.6. As equações de paridade de um código binário (8,4) são

$$c_0 = x_1 + x_2 + x_3$$
$$c_1 = x_0 + x_2 + x_3$$
$$c_2 = x_0 + x_1 + x_3$$
$$c_3 = x_0 + x_1 + x_2$$

em que x_0, x_1, x_2 e x_3 representam os bits da mensagem.

a) Determine as matrizes geradora e de verificação de paridade deste código.

b) Mostre analiticamente (isto é, sem determinar todas as palavras de código) que a distância mínima deste código é 4.

2.7. Considere um código de blocos linear (127,92) capaz de correcções de erros triplos usado num canal com uma probabilidade de erro de 10^{-4}.

a) Qual é a probabilidade de erro na mensagem, para um bloco não codificado de 92 bits?

b) Qual é a probabilidade de erro na mensagem, quando se usa o código de blocos (127,92)?

2.8. Projecte um código de blocos linear sistemático (4,2).

a) Determine as palavras de código, e escolhendo-as com o objectivo de maximizar d_{min}.

b) Determine a matriz geradora do código.

c) Calcule a matriz de verificação de paridade.

d) Coloque os dezasseis conjuntos de 4 bits numa matriz padrão.

e) Quantos (e quais) padrões de erro consegue corrigir e detectar?

f) Construa uma tabela de síndromes para os padrões de erros corrigíveis.

2.9. Considere um código de blocos linear com cada palavra de código definida por

$\mathbf{X} = x_1 + x_2 + x_4 + x_5, x_1 + x_3 + x_4 + x_5, x_1 + x_2 + x_3 + x_5, x_2 + x_3 + x_4 + x_5, x_1, x_2, x_3, x_4, x_5$

a) Determine a matriz geradora.

b) Determine a matriz de verificação de paridade.

c) Determine n, k e d_{min}.

2.10. Um código de blocos linear (15,11) pode ser definido pela seguinte matriz de paridade:

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}^T$$

a) Indique a matriz de verificação de paridade deste código.

b) Faça a lista dos "coset leaders" da matriz padrão. Este código é um código perfeito? Justifique.

c) Um vector recebido é **V** = [0 1 1 1 1 1 0 0 1 0 1 1 0 1 1]. Calcule a síndrome. Supondo que um único bit está errado, determine a palavra de código correcta.

2.11. Um código de blocos caracterizado pela submatriz $P = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ é usado num canal com probabilidade de erro $p = 10^{-4}$. Determine:

a) Determine a probabilidade de eventuais erros não serem detectados.

b) Faça uma lista de "coset leaders".

c) Calcule a probabilidade de correcção errada.

2.12. A matriz geradora de um código de blocos é

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

a) Codifique a sequência 0011010101.

b) Descodifique a sequência 1001011000010101.

c) Determine o número de bits corrigíveis por palavra recebida, t. O código é perfeito?

d) Estime a distância mínima do código.

2.13. Um código (7,3) é um código perfeito? E um código (7,4)? E um código (15,11)? Justifique as respostas.

2.14. Exprima a distribuição de pesos de um código de Hamming aumentado em função da distribuição de pesos do código de Hamming de onde ele foi obtido.

2.15. Considere um código sistemático (7,3) gerado por $g(p) = p^4+p^3+p^2+1$. Determine o polinómio de paridade $C(p)$ e a palavra de código **Y** correspondentes à mensagem **X** = (101). Determine também $Q(p)$, o quociente da divisão de $Y(p)$ por $g(p)$. Tome então $\mathbf{Z}=\mathbf{Y}_1$ (**Y** deslocado de uma casa para a esquerda) e confirme que a síndrome é nula.

2.16. Determine a matriz geradora de um código de Hamming (7,4) gerado por $g(p) = p^3 + p + 1$.

2.17. Um código cíclico (15,5) tem o polinómio gerador $g(p) = p^{10} + p^8 + p^5 + p^4 + p^2 + p + 1$.

a) Desenhe o diagrama de um codificador para este código.

b) Determine o polinómio de código (na forma sistemática) para a mensagem $X(p) = 1 + p^2 + p^4$.

c) O polinómio $V(p) = p^{14} + p^8 + p^6 + p^4 + 1$ é um polinómio de código deste sistema? Justifique.

2.18. A palavra 111100101110 é codificada usando o código de Golay de polinómio gerador $g(p) = p^{11} + p^9 + p^7 + p^6 + p^5 + p + 1$. Determine a palavra de código correspondente.

2.19. Mostre que para qualquer código binário linear (n,k) com distância mínima maior ou igual a **2t+1** o número de dígitos de verificação de paridade satisfaz a seguinte desigualdade (chamada *limite de Hamming*):

$$n - k \geq \log_2 \left[1 + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{t} \right]$$

(Como vê, o limite de Hamming indica um valor máximo para a capacidade de correcção de erros, *t*, do código).

2.20. Um código cíclico com n=15 é gerado pelo polinómio $g(p) = p^8 + p^7 + p^6 + p^4 + 1$.

a) Determine o polinómio de verificação de paridade h(p).

b) Qual é a capacidade de correcção de erros aleatórios, t, do código? Acha que este código é perfeito?

c) Qual é o vector síndrome S(p) correspondente ao polinómio recebido $r(p) = p^{10} + p^9 + p^8 + p^7 + p^5 + p$? Confirme que a síndrome calculada corresponde ao padrão de erro [0 0 0 0 1 0 0 0 0 0 0 0 0 0 0].

d) Determine a matriz geradora, na forma sistemática, de um código cíclico (7,3) gerado por $g_1(p) = p^4 + p^2 + p + 1$.

2.21. O polinómio gerador de um código cíclico é $g(p) = (p^{15} + 1)/(p^2 + p + 1)$.

a) Construa uma tabela com as palavras de código.

b) Determine a distância mínima e o valor de *t*.

c) Determine o polinómio de verificação de paridade, h(p).

2.22. Determine os valores (n,k) do código cíclico gerado pelo polinómio $p^{10} + p^8 + p^5 + p^4 + p^2 + p + 1$.

2.23. O comprimento das palavras de um código cíclico binário gerado pelo polinómio $g(p) = p^5 + p^4 + p^2 + 1$ é 15.

a) Calcule o polinómio de verificação de paridade deste código.

b) Qual é o tamanho das mensagens e quantas palavras de código existem?

c) Calcule as matrizes geradora e de verificação de paridade do código.

d) Calcule os polinómios de código correspondentes aos polinómios de mensagem seguintes, supondo que o código é sistemático: $x_1(p) = p^2$; $x_2(p) = p^7 + p^3 + p$.

e) Determine a síndrome correspondente a cada um dos seguintes polinómios recebidos: $z_1(p) = p^{10}$; $z_2(p) = p^8 + p^6 + p + 1$.

f) Desenhe um circuito codificador sistemático para este código.

g) Desenhe um circuito de cálculo de síndromes para este código.

2.24. O polinómio $p^4 + p + 1$ é o polinómio gerador de um código de Hamming (15,11).

a) Determine a matriz geradora do código.

b) Desenhe um circuito codificador.

2.25. O polinómio de paridade de um código de comprimento máximo ("maximum length shift register code") é $p^5 + p^3 + 1$.

a) Determine o tamanho de cada palavra de código.

b) Determine o polinómio gerador.

c) Determine a síndrome correspondente ao polinómio $p^6 + p^2 + p$.

d) Este código é perfeito? Porquê?

2.26. A palavra binária 10001100 foi recebida num descodificador cíclico. Estime a sequência de cinco bits que lhe terá dado origem.

2.27. Um sistema de comunicações usa um código RS (255, 223).

a) A transmissão é afectada por uma interferência que provoca *bursts* de 1000 erros. Verifique se o descodificador corrige todos estes erros.

b) Analise com clareza a mesma situação quando se usa um sistema de entrelaçamento e desentrelaçamento de blocos com profundidade de entrelaçamento de 10 (número de colunas da matriz).

c) Se a interferência provocar 1456 erros em bits consecutivos, qual é a distância entre eventuais *bursts* de erros consecutivos à saída do desentrelaçador, e qual é o seu tamanho, em bits?

2.28. Os bits de paridade de um determinado código de blocos linear são determinados através do sistema de equações (do Exercício 2.6)

$$\begin{aligned} c_0 &= x_1 + x_2 + x_3 \\ c_1 &= x_0 + x_2 + x_3 \\ c_2 &= x_0 + x_1 + x_3 \\ c_3 &= x_0 + x_1 + x_2 \end{aligned}$$

a) O código é perfeito?

b) Determine as equações de cálculo da síndrome de uma palavra genérica $Z = [z_0 z_1 \ldots z_7]$ e esboce o respectivo circuito combinatório.

c) Recebeu-se a palavra [10111111]. Determine a síndrome e o bloco de quatro bits de informação que terá sido codificado.

3. Códigos detectores e correctores de erros: códigos convolucionais

3.1. Desenhe o diagrama de estados, a árvore do código e a treliça do código convolucional de taxa 1/3 e comprimento de restrição (*constraint length*) 3 gerado pelos polinómios

$$g_1(x) = x + x^2$$

$$g_2(x) = 1 + x$$

$$g_3(x) = 1 + x + x^2$$

3.2. Um código convolucional com taxa 1/3 e *constraint length* igual a 3 tem os polinómios geradores $g_1(x) = x^2 + x + 1$, $g_2(x) = x^2 + x + 1$ e $g_3(x) = x^2 + 1$. Determine:

a) a função de transferência T(D) do código.

b) a distância livre.

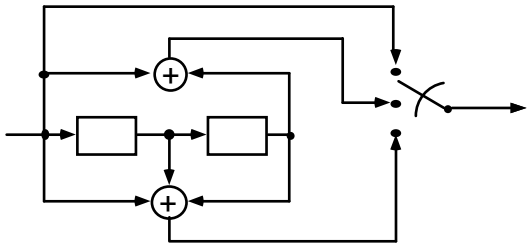
3.3. A função de transferência de um código convolucional é

$$\frac{L^3 D^4 (1 + L - LD^2)}{1 - (L + L^2 + L^3)D^2 + L^3 D^4}$$

a) Dos percursos que na treliça saem do estado nulo e a ele regressam quantos têm peso 6 e são compostos por 6 ramos?

b) Qual é a distância livre do código?

3.4. Considere o seguinte codificador convolucional:



a) Desenhe o diagrama de estados adequado à determinação da função de transferência T(D, L, N) do codificador.

b) Considere todos os percursos da treliça que começam num estado inicial "nulo" e a ele regressam. Quantos percursos existem com peso 5, 6, 7 e 8? Cada percurso é composto por quantos ramos?

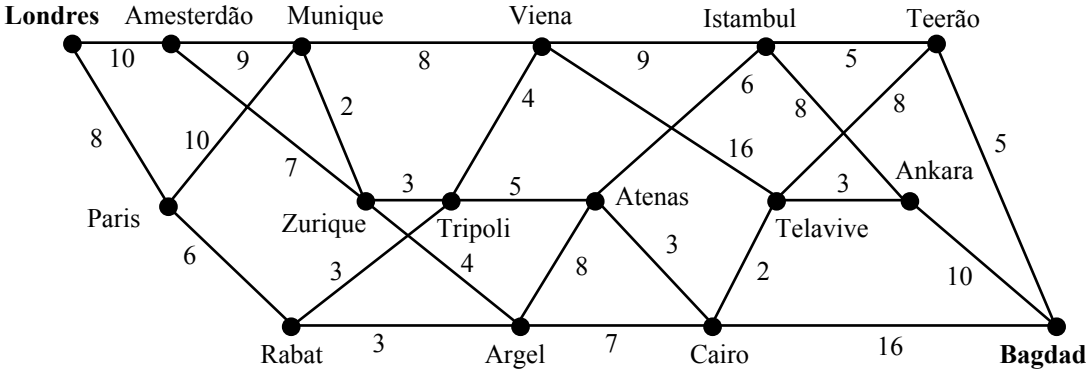
3.5. Construa a treliça para um código (2,1,2) com $x'_j = m_{j-1} + m_j$ e $x''_j = m_{j-2} + m_{j-1}$. Aplique depois o algoritmo de Viterbi para determinar a mensagem original e a sequência codificada estimada quando a sequência recebida é Z = 10 11 01 01 10 01 10 11 00. Se dois percursos chegarem a um dado nó com igual métrica acumulada escolha o percurso de cima.

3.6. Considere o codificador convolucional descrito pelos polinómios de ligação $p_1(x) = x$ e $p_2(x) = 1 + x + x^2$ e com k=1.

a) Desenhe o diagrama de estados do codificador.

b) Uma sequência binária codificada foi enviada através de um canal ruidoso e à entrada de um descodificador de Viterbi foi recebida a sequência **001011100111**. Determine os primeiros bits da sequência binária enviada pelo codificador. (Para bom entendedor: escolha o percurso de cima em caso de empate).

3.7. (No tempo da Guerra do Golfo) Imagine que era um jornalista de televisão que, estando em Londres, tinha a máxima urgência em chegar a Bagdad. Só dispondo do "mapa" apresentado, no qual estão indicadas as horas de viagem inter-cidades, indique qual o percurso que escolheria e quanto tempo demoraria até pôr as primeiras imagens no ar (se o deixassem, claro!).



3.8. Considere um codificador convolucional (2,1,2) caracterizado pelos polinómios $g_1(x) = 1$ e $g_2(x) = x + x^2$. O canal de comunicação BSC introduz ruído, com probabilidade de erro de transição $p = 10^{-3}$ e os primeiros bits da sequência recebida são 10 11 00 01 11 10 01 01 10 … . Suponha que se usa um descodificador sequencial com $\Delta = 2$. Trace o percurso de descodificação, convenientemente anotado, e indique qual a mensagem original (não-codificada) estimada.

TEORIA DA INFORMAÇÃO

Resultados dos exercícios

1.1 a) $P(a_1|000)=\frac{(1-\varepsilon)^3}{(1-\varepsilon)^3+\varepsilon^3}$; $P(a_1|100)=P(a_1|010)=1-\varepsilon$; $P(a_1|011)=\varepsilon$

c) $3\varepsilon^2-2\varepsilon^3$

1.2 a) I(x,y) = I (loira, a horas) = 1 bit; I(morena, a horas) = 0 bits; I(ruiva, a horas) = -∞

b) -1,32 bits

1.3 H(X) = 1,5 bits/símbolo; símbolo; H(Y) = 1 bit/símbolo; H(Z) = 1bit/símbolo;
H(YZ) = 2 bits/símbolo; I(X;Y) = 0,5 bit/símbolo; I(X;Z) = 1 bit/símbolo;

1.4 0,344 bit/s

1.5 a) 0,73 bits/símbolo

c) 0,75 bits/símbolo

1.6 a) log26.

b) log13.

1.7 $\log 5 \log_2 5 - 1 = \log_2 2,5 = 1,32$ bits/símbolo.

1.8 0,68 bits

1.9 H(Y) = 0,97 bits/símbolo; H(Y|X) = 0,92 bits/símbolo; I(X, Y) = 0,05 bits/símbolo.

1.10 a) $H(X) = 33/8 = 4,125$ bits; $H(Y) = 1,561$ bits; $H(Y|X) = 0$ bits;
$H(X|Y) = H(X) + H(Y|X) - H(Y) = 2,564$ bits.

b) $H(X|Z) = 3,125$ bits; $H(Z|X) = 0$ bits.

1.11 a) $a_1 \rightarrow 00$; $a_2 \rightarrow 01$; $a_3 \rightarrow 100$; $a_4 \rightarrow 101$; $a_5 \rightarrow 1100$; $a_6 \rightarrow 1101$; $a_7 \rightarrow 1110$; $a_8 \rightarrow 1111$

1.12 Não existe.

1.13 c) Código I: $I(a_1;y_1) = 1,32$ bits; Código II: $I(a_1;y_1) = 0$ bits

d) Código I: 0,971 bits/símbolo

1.14 a) 92,5%

b) 96,3%

1.15 a) 92,6%

b) 99,2%

1.16 a) 1,791 bits/símbolo

b) "CORRECTO"

1.17 a) Código binário: fonte 1 – $\overline{N} = 2,5$ binits/símbolo da fonte; fonte 2 – $\overline{N} = 2,53$ binits/símbolo da fonte.

b) Código ternário: fonte 1 – $\overline{N} = 1,7$ dígitos ternários/símbolo de fonte; fonte 2 – $\overline{N} = 1,61$ dígitos ternários/símbolo de fonte.

1.18 a) (2, 2, 2, 3, 3).

b) (1, 2, 3,…).

1.19 a) 2,39 garrafas

b) Deve-se experimentar a garrafa com probabilidade 8/23.

c) 2,35 garrafas (usar o código de Huffman).

d) Deve-se experimentar a mistura das duas primeiras garrafas.

1.20 a) $\overline{N} = 2,2$ bits / símbolo .

b) V(X)=1,36.

c) V(X)=0,16.

1.21 a) Três caracteres errados.

b) Um carácter errado.

c) Três e um, respectivamente.

1.22 a) $\overline{N} = 1,35$ dígitos ternários/símbolo; Eficiência = 93,3%; Redundância = 0,091

b) $\overline{N} = 1,70$ dígitos ternários/símbolo; Eficiência = 74%; Redundância = 0,441

1.23 a) 7 binits/tecla;

b) $\overline{\mathrm{N}}$ = 5,967 binits/tecla; H(X) = 5,907 bits/tecla; Eficiência = 99%

1.24 "ACERTEI"

1.25 Intervalo final: [0,7712 0,773504[

1.26 Intervalo final: [0,027 0,0276[

1.27 Intervalo final: [0,1796862; 0,180398[; sequência binária: 0010111

1.28 $a_3 a_1 a_1 a_2 a_1 a_2$

1.29 a)

$X_1 \rightarrow 221$ $X_6 \rightarrow 11$

$X_2 \rightarrow 220$ $X_7 \rightarrow 10$

$X_3 \rightarrow 21$ $X_8 \rightarrow 02$

$X_4 \rightarrow 20$ $X_9 \rightarrow 01$

$X_5 \rightarrow 12$ $X_{10} \rightarrow 00$

b) 2,06 símbolos ternários/símbolo de fonte.

c) 20 bits, no máximo.

1.30 a) 16 bits.

b) (21, 5, V) (16, 2, ,) (13, 9, O) (4, 1, V) (31, 3, A) (0, 0, L) (0, 0, G) (5, 1, É) (42, 2, Q) (30, 6, .)

1.31. AAAABABCCA **ABAC AAC CAB CABCAB CABCC**

1.32. a) AB / AA / BB / ABA / BA / AAA / BAB / BABB / ABB / AAB / ABAB / ABBB / ABAA / BAA / BABA / BABAB

b)

Nº entrada	Entrada	Representação
1	A	-
2	B	-
3	AB	1B
4	AA	1A
5	BB	2B
6	ABA	3A
7	BA	2A
8	AAA	4A
9	BAB	7B
10	BABB	9B
:	:	:

c) 96 bits.

1.33. a) AA / AAB / AB / CC / AABA / CA / AC / CAB / CABC / ABC / ABCC

b)

Nº entrada	Entrada	Representação
1	A	-
2	B	-
3	C	-
4	AA	1A
5	AAB	4B
6	AB	1B
7	CC	3C
8	AABA	5A
9	CA	3A
10	AC	1C
11	CAB	9B
12	CABC	11C
:	:	:

c) 1A 4B 1B 3C 5A 3A 1C 9B 11C 6B 13C

d) 6 bits.

1.34. ABAABCCABBBBC

1.35 ≥ 48 segundos

2.1 GBN: $R'_c \leq 0{,}772$; SW: $R'_c \leq 0{,}172$. Terá de ser $R'_c \geq 0{,}6$ logo o método "stop-and-wait" não serve.

2.3 $R'_{SW} = 0{,}199 \times \frac{4}{7}$, $R'_{GBN} = 0{,}972 \times \frac{4}{7}$, $R'_{SR} = 0{,}993 \times \frac{4}{7}$.

2.4 $c_1 = x_1 \oplus x_2$; $c_2 = x_1 \oplus x_3$; $c_3 = x_2 \oplus x_3$; $d_{min} = 3$

2.5 t=1: $n-k=8$; t=2: $n-k=15$; t=3: $n-k=22$.

2.6 a) $G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$ $H = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T$

2.7 a) $9{,}2.10^{-3}$

b) $1{,}03.10^{-9}$

2.8 e) Corrige 3 padrões de 1 erro e detecta 12 padrões de erro.

2.9 c) n = 9, k = 5, $d_{min} = 3$.

2.10 b) É um código perfeito, com $t = 1$; c) S=[0 1 0 1], $\hat{V}$ = [0 0 1 1 1 1 0 0 1 0 1 1 0 1 1]

2.11 a) 4.10^{-12}

c) $1{,}4.10^{-7}$

2.12 a) [00110100] [10101100]

b) [00010] [00011]

c) t=0. O código não é perfeito.

d) 2.

2.13 Não. Sim. Sim.

2.14 Hamming: $\{A_1, A_2, \ldots, A_n\} = \{1, 0, 0, A_3, A_4, \ldots, A_n\}$

Hamming aumentado: $\{1, 0, 0, 0, A_3 + A_4, 0, A_5 + A_6, 0, \ldots\}$

2.15 $Q(p) = p^2+p+1$; $C(p) = p+1$; Y = [1 0 1 0 0 1 1]

2.16 $R_1(p) = p^2+1$; $R_2(p) = p^2+p+1$; $R_3(p) = p^2+p$; $R_4(p) = p+1$

2.17 b) $p^{14}+p^{12}+p^{10}+p^9+p^6+p^2+p+1$.

c) Não.

2.18 Y = 1 1 1 1 0 0 1 0 1 1 1 0 1 0 1 1 1 1 1 1 0 1 0

2.20 a) $h(p) = p^7 + p^6 + p^4 + 1$;

b) t = 2. Não é perfeito;

c) $S(p) = p^7 + p^6 + p^5 + p^2 + p$;

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

2.21 b) t = 4

c) $h(p) = p^2 + p + 1$

2.22 (15, 5)

2.23 a) $h(p) = p^{10} + p^9 + p^8 + p^6 + p^5 + p^2 + 1$.

b) k=10; 1024 palavras de código.

c) $P = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}^T$; $G = [I_{10} | P]$; $H = \begin{bmatrix} P \\ \overline{I_5} \end{bmatrix}$

d) $y_1(p) = p^7 + p^3 + p + 1$; $y_2(p) = p^{12} + p^8 + p^6 + p^4 + p^2 + 1$.

e) $S_1(p) = p^2 + p + 1$; $S_2(p) = p^3 + p$

2.25 a) n=31.

b) $g(p) = p^{26} + p^{24} + p^{22} + p^{21} + p^{20} + p^{18} + p^{17} + p^{13} + p^{12} + p^{11} + p^{10} + p^9 + p^6 + p^5 + p^3 + 1$

c) $S(p) = p^6 + p^2 + p$

d) Não é perfeito.

2.26 $\hat{X} = [10001]$.

2.27 a) t = 16 símbolos. Não consegue.

b) Permite a correcção.

c) Distância mínima: 1895 bits; Tamanho máximo: 146 bits.

2.28 a) Não é perfeito;

b) Equações de cálculo da síndrome:

$$s_0 = (z_1 + z_2 + z_3) + z_4$$
$$s_1 = (z_0 + z_2 + z_3) + z_5$$
$$s_2 = (z_0 + z_1 + z_3) + z_6$$
$$s_3 = (z_0 + z_1 + z_2) + z_7$$

c) $\mathbf{S} = [1 \quad 0 \quad 1 \quad 1]$; bloco de quatro bits: [1111].

3.2 a) $T(D) = \dfrac{2D^8 - D^{10}}{1 - 3D^2 + D^4} = 2D^8 + 5D^{10} + 13D^{12} + \ldots$;

b) $d_f = 8$.

3.3 a) 2; b) 4.

3.4 b) $T(D,L,N) = L^3ND^6 + L^4N^2D^8 + L^5N^2D^8 + L^5N^3D^{10} +$ ….Peso 5: 0 percursos; peso 6: 1 percurso com 3 ramos; peso 7: 0 percursos; peso 8: 2 percursos, um com 4 ramos e outro com 5 ramos.

3.5 Sequência original estimada: 1 0 0 0 1 1 0…;
Sequência codificada original estimada: 10 11 01 00 10 01 10 ...

3.6 b) 01 10 11 10

3.7 Londres — Paris — Rabat — Tripoli — Atenas — Istambul — Teerão — Bagdad.

3.8 1 1 0 0 0 1 0 0 1.