

UNIVERSIDADE DE AVEIRO

DEPARTAMENTO DE ENGENHARIA DE ELECTROTELECOMUNICAÇÕES E ESTATÍSTICA

Informação e Codificação (2021/22)

Conjunto de problemas 1

1. O código $\{00, 11, 0101, 111, 1010, 100100, 0110\}$ é decodificável de forma única?
2. O código (ternário) $\{00, 012, 0110, 0112, 100, 201, 212, 22\}$ é decodificável de forma única?
3. Encontrar uma distribuição de probabilidade $\{p_1, p_2, p_3, p_4\}$ tal que existam dois códigos óptimos que atribuem comprimentos diferentes $\{l_1, l_2, l_3, l_4\}$ aos quatro símbolos.
4. Explique por que razão cada método de compressão irá expandir alguns dos seus inputs, em vez de os comprimir.
5. Considerar um código Golomb com parâmetro $m = 5$ (note que m não é uma potência de dois).
 - (a) De acordo com este código, dar, justificando, uma sequência de bits que representa os números inteiros $a = 12$ e $b = 13$ tão eficientemente quanto possível.
 - (b) Indicar a distribuição ótima de probabilidade para este código, ou seja, os valores de $P(n)$, $n \in \mathbb{N}_0$.
6. Considerar um sistema de codificação para números inteiros não negativos, em que um certo valor n é preenchido por n "0" bits seguido por um "1" bits. Para que distribuição de probabilidade, $P(n)$, $n \in \mathbb{N}_0$, este sistema de codificação é maximamente eficiente?
7. Considerar a seguinte tabela (incompleta) de correspondências entre símbolos σ_i , probabilidades p_i , e codewords:

σ_1	$(p_1 = 0,30)$	\rightarrow	00
σ_2	$(p_2 = 0,30)$	\rightarrow	?
σ_3	$(p_3 = 0,15)$	\rightarrow	110
σ_4	$(p_4 = 0,15)$	\rightarrow	?
σ_5	$(p_5 = 0,10)$	\rightarrow	111

- (a) Complete a tabela de modo a obter um código Huffman.
 - (b) Calcular a redundância do código, considerando a entropia da primeira ordem da fonte de informação.
8. Considere o seguinte quadro de correspondências entre símbolos σ_i , probabilidades p_i , e

codewords:

$$\begin{aligned}\sigma_1 (p_1 = 0.1) &\rightarrow 000 \\ \sigma_2 (p_2 = 0.2) &\rightarrow 01 \\ \sigma_3 (p_3 = 0.2) &\rightarrow 10 \\ \sigma_4 (p_4 = 0.5) &\rightarrow 1\end{aligned}$$

- (a) Assumindo que os símbolos ocorrem independentemente, calcular a entropia desta fonte de informação.
 - (b) Este código de comprimento variável não é construído correctamente. Porquê?
 - (c) Mostrar que não é possível construir um código sem prefixos com este conjunto de comprimentos de palavras de código.
 - (d) Propor um código de comprimento variável apropriado para a distribuição de probabilidade dada.
 - (e) Calcular a redundância (em relação à entropia da fonte) do código construído na pergunta anterior.
9. Explicar, brevemente, os princípios de um dos algoritmos de codificação baseados em dicionários, apontando as principais vantagens e desvantagens.
 10. Explicar, brevemente, o princípio de funcionamento de um codificador aritmético.
 11. Dê a sequência de código produzida por um codificador LZ78, se a sequência de entrada for: zzxyxyxxxxyzxxxzxxxzzxyxx
Além disso, fornecer o estado final do dicionário.
 12. Dê a sequência de código produzida por um codificador LZ77 se a sequência de entrada for: zzxyxyxxxxyzxxxzxxxzzxyxx
Considere que o buffer de entrada tem tamanho 4 e a janela do dicionário tem tamanho 12.
 13. Dê a sequência de código produzida por um codificador LZW se a sequência de entrada for: zzxyxyxxxxyzxxxzxxxzzxyxx
Considerar que o alfabeto de entrada é $\Sigma = \{x, y, z\}$.
 14. Tendo em consideração os princípios da codificação aritmética, e que $P(0) = P(1) = 0,5$, indique um valor no intervalo $[0, 1)$ que representa todas as mensagens que começam com a sequência binária "1011".
 15. Considere que um certo modelo de segunda ordem de contexto finito (ou seja, que utiliza os dois símbolos anteriormente ocorridos para condicionar a probabilidade de ocorrência do símbolo seguinte) já observou a seguinte sequência binária:

10010110111101010011000

Indicar uma estimativa para a probabilidade dada por este modelo de que o próximo símbolo será "1".