

Entropia e dicionários

1. Abra um romance e conte: o número de caracteres por linha (não esquecer o espaço), o número de linhas por página. Quantos caracteres tem um livro? Num CD de $700MB$ ¹ quantos livros de tamanho semelhante pode armazenar?
2. Um armazém tem 45 835 artigos. Pretende construir um código com C bits para indexar os artigos. Qual é o valor de C ?
3. As notas dos alunos são atribuídas na escala $\{A, B, C, D\}$. Calcule a entropia assumindo que
 - (a) $P(A) = P(B) = P(C) = P(D) = \frac{1}{4}$
 - (b) $P(A) = \frac{1}{4}, P(B) = \frac{1}{2}, P(C) = P(D) = \frac{1}{8}$
 - (c) Se tiver 100 alunos, diga quantos bits precisa para codificar as respectivas notas.
4. Uma mensagem contém os símbolos seguintes $\{*, +, i, j, :, p, q\}$. As probabilidades de cada símbolo são $\{0.25, 0.15, 0.07, 0.03, 0.16, 0.3, 0.04\}$
 - (a) Quantos bits são necessários para representar cada símbolo?
 - (b) Calcule a entropia da mensagem.
 - (c) Desenhe um codificador Huffman
 - (d) Utilizando o código de Huffman quantos bits precisa para codificar a mensagem?
5. Numa imagem de 256×256 os pixéis estão representados com 8 bits. Após fazer um histograma dos valores dos pixéis verifica que 50% dos pixéis têm o valor 100, 25% tem o valor 10 e os restantes têm o valor 220. Como é que pode representar eficientemente a informação?
 - Quantos bits precisa para representar a imagem
 - Qual é o rácio de compressão

¹Mega Bytes (1 byte= 8 bits).

6. Considere que os pixels da primeira linha de uma imagem têm os valores seguintes:

$$\{68, 68, 68, 68, 68, 68, 10, 40, 50, 20, 20, 20, 6, 6, 6, 6, 6\}$$

Qual é o resultado da codificação com RLE.

7. Considere que inicia uma codificação LZW com o dicionário seguinte: $\{P, A, T, I\}$

- qual é o dicionário final para a mensagem: "PATATIPATATA".
- Que sequência (os índices do dicionário) é que obtém para a mensagem.
- Mostre que no decodificador só precisa de saber o dicionário inicial (a tabela) para decodificar a sequência.