Worksheet:
Information: Coding and Redundancy

1. A five symbol source has the following probabilities and codes for each of the symbols:

symbol	a	e	i	o	u
probability	0.3	0.5	0.1	0.05	0.05
code	00	011	01	0101	0111

(a) Calculate the following:

    i. Redundancy of the source in terms of the symbol probabilities

    ii. Code capacity and Source efficiency:

(b) How could you make the source more efficient? Hint: change the code words for each symbol to what you think should produce a more efficient source.

2. Hamming (7,4) Block Coding: This is can detect and correct a single error unlike a single parity bit however Hamming (7,4) Block Coding introduces more redundancy:

p_1	p_2	I_1	p_3	I_2	I_3	I_4

Parity bits P_1 to P_3 are (xor) functions of the 4 data bits I_1 to I_4: $P_1 = I_2 \oplus I_3 \oplus I_4$, $P_2 = I_1 \oplus I_3 \oplus I_4$ and $P_3 = I_1 \oplus I_2 \oplus I_4$. For (7,4) Block coding what is the redundancy?

3. For Hamming (7,4) Block Coding, errors can be detected upon reception with check bits:

$$C_1 = P1 \oplus I_2 \oplus I_3 \oplus I_4, \quad C_2 = P2 \oplus I_1 \oplus I_3 \oplus I_4 \quad \text{and} \quad C_3 = P3 \oplus I_1 \oplus I_2 \oplus I_4.$$

Use an example data sequence of $(I_1, I_2, I_3, I_4) = (1, 1, 1, 1)$ for the following:

(a) What value are the check bits if there is no error?

(b) What value are the check bits if there is an error in I_1?

(c) What value are the check bits if there is an error in I_2?

(d) What value are the check bits if there is an error in I_3?

(e) What value are the check bits if there is an error in I_4?