This is a practice assignment and will not be worth any marks. You are still welcome to submit solutions if you want feedback on them.

1. Suppose that a text source generated from the alphabet $X = \{\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}, \mathrm{e}, \mathrm{f}, \mathrm{g}\}$ is encoded via the code $\psi : X \to W(\mathbb{Z}_2)$ defined as follows:

$$\phi(\mathrm{a}) = 11, \quad \phi(\mathrm{b}) = 01, \quad \phi(\mathrm{c}) = 0001, \quad \phi(\mathrm{d}) = 10, \quad \phi(\mathrm{e}) = 0011, \quad \phi(\mathrm{f}) = 0010, \quad \phi(\mathrm{g}) = 0000.$$

(a) Encode the word $faded$.

(b) Decode the codeword 00011100100011.

(c) Find a set of frequencies (probabilities) for the letters in the source that would produce this code as a Huffman code. Justify your answer by constructing the corresponding binary tree using Huffman's method.

(d) Find a set of frequencies for the letters such that the code is optimal, but not a Huffman code.

2. We want to compress a bitstring from a source which generates (independent) random bits with frequencies $p_0 = \frac{1}{4}$ and $p_1 = \frac{3}{4}$ by grouping it into substrings of length 3, and then constructing a Huffman code for these substrings (you may assume that the length of the whole bitstring is divisible by 3, so that the grouping into substrings actually works out without remainder).

(a) For each $x \in \mathbb{Z}_2^3$ compute the frequency p_x of x.

(b) Construct a Huffman code $\psi \colon \mathbb{Z}_2^3 \to W(\mathbb{Z})$ with respect to these frequencies.

(c) What is the average length of a codeword with respect to this Huffman code? How much space do we save compared to the original encoding?

3. Let $\psi : \mathbb{Z}_2^{14} \to W(\mathbb{Z}_2)$ be Fitingof's code.

(a) What are the maximal and minimal length of $\psi(\mathbf{x})$ for $\mathbf{x} \in \mathbb{Z}_2^{14}$?

(b) For which Hamming weights $w(\mathbf{x})$ is the encoded codeword $\psi(\mathbf{x})$ shorter than 14 bits?

(c) Encode the word $\mathbf{x} = 00100000001100$.

(d) A file which was compressed using the code ψ starts as follows:

$$00110000011000110\ldots$$

Determine the first 14 bits of the decoded file, in other words, decode the first 14-bit block.