Shannon Coding

Information Theory Exercises

1. We perform Shannon coding on an information source with $H(S) = 20\text{b}$.
   a. What are the possible values for the efficiency of the code?
   b. What are the possible values for the redundancy of the code?
   c. What is the minimum number of messages the source may possibly have?

2. A discrete memoryless source has the following distribution:

$$S : \begin{pmatrix} s_1 & s_2 & s_3 & s_4 & s_5 \\ 0.05 & 0.4 & 0.1 & 0.25 & 0.2 \end{pmatrix}$$

   a. Encode the source with Shannon, Shannon-Fano coding and Huffman coding and compute the average length in every case.
   b. Find the efficiency and redundancy of the Huffman code
   c. Compute the probabilities of the symbols 0 and 1, for the Huffman code

3. For the following source, perform Huffman coding and obtain three different codes with same average length, but different individual codeword length.

$$S : \begin{pmatrix} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \\ 0.05 & 0.05 & 0.15 & 0.25 & 0.25 & 0.25 \end{pmatrix}$$

   a. Compute the average length in all three cases and show it is the same

4. A discrete memoryless source has the following distribution:

$$S : \begin{pmatrix} s_1 & s_2 & s_3 \\ 0.1 & 0.7 & 0.2 \end{pmatrix}$$

   a. Find the average code length obtained with Huffman coding on the original source and on its second order extension.

   b. Encode the sequence $s_7 s_7 s_3 s_7 s_7 s_7 s_1 s_3 s_7 s_7$ with both codes.

5. A discrete memoryless source has the following distribution

$$S : \begin{pmatrix} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \\ 0.4 & 0.3 & 0.2 & 0.04 & 0.03 & 0.02 & 0.009 & 0.001 \end{pmatrix}$$

Find the Huffman code for a code with 4 symbols, x_1, x_2, x_3 and x_4, and encode the sequence

$$s_1 s_7 s_8 s_3 s_3 s_1$$