Homework 4: Introduction to Information Theory

University of Windsor
Department of Electrical and Computer Engineering
ELEC 4190 - Digital Communications

> **Instructions**
>
> Submissions should be through Brightspace. There is a 24-hour grace period after the due date without a penalty. Late submissions and email submissions will not be accepted.

> **Note**
>
> Course materials are made available to you for your own study and instructional purposes only and are not for redistribution, resale, or profit. Please note that students may not reuse or repost these course materials without the permission of the instructor. Parts of this document are adapted from the textbooks.

1. (4 points) A discrete memoryless source has an alphabet $\{x_1, x_2, x_3, x_4, x_5, x_6\}$ with corresponding probabilities $\{0.12, 0.08, 0.2, 0.35, 0.2, 0.05\}$.

- **(a)** Find the entropy of this source.
- **(b)** What is the minimum required average codeword length to represent this source for error-free reconstruction?
- **(c)** Compare the entropy with the entropy of a uniformly distributed source with the same alphabet.
- **(d)** Design a Huffman code for the source and compare the average length of the Huffman code with the entropy of the source.

2. (4 points) A discrete memoryless source has an alphabet $\{a_1, a_2, a_3\}$ with corresponding probabilities $\{0.2, 0.4, 0.4\}$.

- **(a)** Find the entropy of this source.
- **(b)** Design a Huffman code for the source. What is the average codeword length?
- **(c)** Design a Huffman code for the second extension of the source (take two letters at a time). What is the average codeword length? What is the average number of required binary letters per each source output letter?
- **(d)** Which is a more efficient coding scheme: the Huffman coding of the original source or the Huffman coding of the second extension of the source?

3. (4 points) Design a ternary Huffman code, using A, B, and C as letters, for a source with output alphabet probabilities given by $\{0.05, 0.2, 0.05, 0.16, 0.18, 0.23, 0.13\}$. Compare the average codeword length with the entropy of the source.

4. (4 points) Find the Lempel-Ziv source code for the source sequence below and recover the original sequence from the Lempel-Ziv source code. Use decimal numbers to represent the dictionary location.

```
BAAAABAAAABABABBBBBBBAABABAABABABBABABBAABBAABBABB
```

5. (4 points) Two binary random variables X and Y are distributed according to the joint distributions $P(X = 0, Y = 0) = P(X = 0, Y = 1) = P(X = 1, Y = 1) = 1/5$. Compute $H(X)$, $H(Y)$, $H(X, Y)$, $H(X|Y)$, $H(Y|X)$, and $I(X; Y)$.