Homework no. 5

1. **Conditions for unique decodability.** Prove that a code C is uniquely decodable if (and only if) the extension
$$C^k(x_1, x_2, \ldots, x_k) = C(x_1)C(x_2)\cdots C(x_k)$$
is a one-to-one mapping from $\mathcal{X}^k$ to D^* for every $k \geq 1$. (The only if part is obvious.)

2. **Unused code sequences.** Let C be a variable length code that satisfies the Kraft inequality with equality but does *not* satisfy the prefix condition.

   (a) Prove that some finite sequence of code alphabet symbols is not the prefix of any sequence of codewords.

   (b) (Optional) Prove or disprove: C has infinite decoding delay.

3. **Huffman Codes and Kraft's inequality** Consider a Huffman code with binary code alphabet. Show that it satisfies Kraft's inequality with equality.

4. **Arithmetic Coding For a 3-Symbol Source.** Consider a data source with a ternary alphabet, say A, B, and C. Consider the non-adaptive Arithmetic Coding scheme. Assume $\mathbb{P}(A) = 0.3$, $\mathbb{P}(B) = 0.5$, and $\mathbb{P}(C) = 0.2$.

   (a) Let the source sequence be $BABCB$. What is the length of the encoded sequence?

   (b) Find the encoded sequence.

   (c) Show me how you would decode the encoded sequence.

   Note This problem has multiple solutions, since there are different implementations of Arithmetic Coding. For example, if we used integer arithmetic coding with few bits of precision, we could have truncated the intervals. Alternately, if we allowed for bit carry over, the result might have been slightly different.