Programming Assignment

You can use any programming language you prefer (MATLAB, Python, C/C++, or Julia, for example). Write down your code as clearly as possible and add suitable comments. Turn in the hard copy of answers to each problem and submit your code to `ece154ucsd@gmail.com` with the exact subject `ECE 154C (HW2)`.

1. Write a program for a function `daryHuffman(pmf, d)` that takes a probability vector `pmf` and the size of alphabet d as inputs, and outputs a d-ary Huffman code. For example,

   ```
   ["0","1","20","21","22"] = daryHuffman([0.6,0.2,0.1,0.05,0.05],3)
   ```

   (a) Run the program for `pmf` $= \{0.6, 0.3, 0.1\}$ with $d = 3$ and compute the average length of the resulting Huffman code.

   (b) Consider block lengths $n = 2, 3, 5$. For each block length, find the ternary Huffman code and the average length per symbol.

2. Consider a binary source $\{A, B\}$, and suppose we have a probability vector `src_pmf` on this set. Write a program for a function `Tunstall(src_pmf, num_phrases)` that takes the probability vector `src_pmf` and the desired number of phrases `num_phrases` as inputs and outputs the resulting phrases of Tunstall coding and the probability vector of it. For example,

   ```
   ["B", "AB", "AAA", "AAB"], [0.1,0.09,0.729,0.081]
                = Tunstall([0.9, 0.1], 4),
   ```

   and

   ```
   ["B", "AB", "AAB", "AAAA", "AAAB"], [0.2,0.16,0.128,0.4096,0.1024]
                = Tunstall([0.8, 0.2], 5).
   ```

   (a) For `src_pmf` $= \{0.9, 0.1\}$, find the length of codewords per average number of symbols for `num_phrases` $= 4, 8, 16, \ldots, 2^{10} = 1024$.

   (b) For `src_pmf` $= \{p, 1-p\}$, plot the length of codewords per average number of symbols for `num_phrases` $= 1024$ as a function of p. The plot should be evaluated at 100 values of p.

   (c) Compare the result in part (b) with the Huffman code rate in Q4 of Homework Set #1.