CS6351 September 1, 2022
Youssef

Homework 1
Due Date: September 22, 2022

Problem 1: (20 points)

Consider the alphabet $\{a, b, c, d, e, f\}$ with the following probabilities: $P[a] = \frac{12}{32}, P[b] = \frac{10}{32}, P[c] = \frac{3}{32}, P[d] = \frac{4}{32}, P[e] = \frac{1}{32}, P[f] = \frac{2}{32}$.

a. Compute the entropy of this channel.

b. Build the Huffman tree for this alphabet, and derive the Huffman codes of the 5 symbols of this alphabet. (Always make the child with the smaller probability the left child. In case of equal probabilities, make the child of smaller alphabetical order to be the left child.)

c. What is the Huffman bitrate?

d. Code the following string: *aaaaabbbdddbdebaaaaccbaf*. What is the bitrate for this input?

Problem 2: (20 points)

Let $p = \frac{61}{64}$ and $q = \frac{3}{64}$. Consider the following 1st order Markov source with alphabet P[0|0]=P[1|1]= p and P[0|1]=P[1|0]= q. Assume that P[1st bit = 0] = P[1st bit = 1] =$\frac{1}{2}$.

a. Use the arithmetic coding algorithm (presented in class) to code the following binary sequence from the source in question: 000111. Indicate what the interval is after the processing of each bit. What is the bitrate?

b. Compute P[000], P[001], P[010], P[011], P[100], P[101], P[110], and P[111]. Afterwards, perform block-Huffman coding on 000111 where each block is 3 bits long. Make sure to show the block-Huffman tree and the block-Huffman Codewords. What is the bitrate for block-Huffman coding of 000111? Compare it to the bitrate in part (a)?

Problem 3: (20 points)

Denote by a^n the string of n a's. Let $x=0^{15}1^{11}0^{10}1^{13}0^{14}1^{9}$

a. Apply run-length encoding on x. Allocate 4 bits to represent each length.

b. Apply Golomb coding of order m on x, where m=nearest power of 2 of $\frac{p\,Ln(2)}{1-p}$, where p is the probability of the more probable symbol.

c. Apply differential Golomb coding on x. Choose the m that is most appropriate.

d. Compute the bitrates of the three techniques for this particular input x, and indicate the best.

Problem 4: (20 points)

Let x be the same as in problem 3, and let y= *aaaabbbbabbbabaabaabbaaa*

a. Code *x* using LZ. What is the bitrate? Compare it to the best bitrate in Problem 3.

b. Code *y* using LZ. What is the bitrate?

Problem 5: (20 points)

In this problem, you will generalize Arithmetic Coding to ternary sources. Consider input strings where the alphabet is $\{a, b, c\}$, and assume that the input follows a Markov model of order 1 where $P[a|a], P[a|b], P[a|c], P[b|a], P[b|b], P[b|c], P[c|a], P[c|b],$ and $P[c|c]$ are given to you, and that $P[a] = P[b] = P[c] = \frac{1}{3}$. Give an Arithmetic Coding algorithm that codes an input string $x_1 x_2 \ldots x_n$ where each x_i is a, b or c, and the input obeys the Markov model presented above. Note: You need not worry about implementation issues such as undeflow.