End Sem Examination
Information and Communication (Spring 2023)
Time : 3 hours, Total Marks: 100

Prasad Krishnan

Instructions:

- Pick questions that seem to be easily answerable first, rather than go linearly through the question paper. Reasons for all steps should be given, in general.
- Calculators are allowed.
- This is a closed book, traditional, exam.
- Malpractice will directly result in 0 and further academic action will be initiated.
- You can request (not demand) for an additional hint from the course instructor (not from anyone else), who should be around. The discretion of providing (or not providing) the hint for a particular question will be left to the instructor. If the instructor is absent, no hint will be provided. *No debate or discussion will be there during the time of evaluation or post-evaluation regarding these hints.*

Questions:

1. (10 marks) 720 720 720

   (a) (2 marks) An video is a stream of images or frames, at rate 30 frames per second. Each frame has resolution 720p, which basically means it has 720×1080 pixels in each frame, where each pixel can take values in the set $\{0, 1, \ldots, 254, 255\}$ (we assume grayscale images; imagine 0 indicates pitch black and 255 is pure white, values in between denoting grays of different degree). This video signal is to be transmitted via a point-to-point link (or binary channel). What would be the bit-rate (bits per second) requirement of this link, if there is no compression algorithm used?

   (b) (4 marks) Suppose it turns out that each pixel in each image of the video takes the 256 values with probability distribution as follows $\{\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{1012}, \cdots, \frac{1}{1012}\}$ (that is, three of the 256 values occur with probability $\frac{1}{4}$ each, and the remaining 253 values occur with probability $\frac{1}{1012}$ each). Now, what would be the bit rate required on the link, if we used an optimal source code for each pixel? (Hint: Use some property of the expected code length of optimal code)

   (c) (3+1 marks) If each frame in the video is just an all-white frame or an all-black frame or all-equally-gray frame with probabilities $\{0.2, 0.6, 0.2\}$ respectively, what would be the bit rate required on the link, assuming an optimal source code for the entire frame? If this channel has capacity 1.2 bits per second, reason if it can support this required bit rate.

2. (12 marks) Consider the Hamming code with length 7 and generator matrix $[I_4 \;\; P]$ (concatenation of identity matrix of size 4, and matrix P given as $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$).

   (a) (3 marks) Find the codeword generated by Hamming code corresponding to message vector $(0, 0, 1, 1)$. (Remember that additions are XOR).

   (b) (5 marks) Suppose when transmitting the above codeword, the sixth bit is flipped. Show that the decoding procedure (which uses the parity check matrix H) succeeds in correcting this bit flip and finding the correct codeword. $(H^T$

   (c) (4 marks) Suppose there are two flips (assume some two coordinates of the codeword are flipped) when the above codeword is transmitted. Show that the decoding procedure using the H matrix fails to correctly decode the transmitted codeword.

3. (8+8=16 marks) Consider a **binary symmetric channel** with with bit-flip probability $p < 0.5$.

   (a) (8 marks) Derive the capacity of this channel.

   (b) (8 marks) Derive an expression for the probability of error, when using a repetition code of length n on this channel, with the decoding rule being the maximum aposteriori probability rule. Assume that transmitted codeword is having a uniform distribution over the codewords of the repetition code.

4. (20 marks) Consider a source X with probability distribution on six letters being

$$0.3, 0.05, 0.05, 0.1, 0.2, 0.3.$$

   (a) ($5+6 = 11$ marks) Design two source codes; one using the Shannon-Fano technique (i.e., a prefix-free code using the specific length choices as taught in class, which satisfy Kraft inequality), and the second using the Huffman algorithm. You have to clearly specify the codewords along with encoding function in each case.

   (b) (5 marks) Compare the expected lengths of the two source codes, and reason out which code is better.

   (c) (2+2 =4 marks) Calculate the entropy of X. Compare this with the two expected length from part (b) and give comments.

5. (8 marks) Consider a Gaussian channel (AWGN channel) with noise distribution being $\mathcal{N}(0, \frac{N_0}{2})$. Assume that we are choosing to transmit on this channel a random variable X, which takes values in $\{A, -A\}$ (A is some positive number) with uniform probability. It is given that the decoder is doing MAP (Max. Aposteriori Probability) detection (i.e., given that output random variable Y takes value y, the decoder finds the most probable input value for X). Show an expression for the probability of error for this decoder. (Hint: Find the conditional distributions of Y when the input is A. Based on this, using Bayesian rule, you can understand for what values of Y you will fail to detect correctly under the MAP detection rule, given that the transmission was A. Similarly you can do for the other possible input. These will help you to find the probabilities of error given each input. Take their expectation to get the (average) probability of error).

6. (2+4+2 = 8 marks) If p_X and p_Y are the probability distributions of two random variables X and Y, prove or disprove: $q(x,y) = p_X(x)p_Y(y), \forall x, y$ is a valid joint distribution. Now, using what you verified, and the fact that $log(\cdot)$ is a concave function, show that the mutual information between X and Y is non-negative. Show that mutual information between X and Y is also upper bounded by the minimum of the entropies of X and Y.

7. (19 marks) Given $x(t)$ has the Fourier transform $X(f)$, express the Fourier transform of the signals listed below in terms of $X(f)$. Let f_0 be a large positive number.

- (3 marks) $x_1(t) = x(3t - 6)$
- (3 marks) $x_2(t) = x(t)e^{j2\pi f_0 t}$

$x(3(t-2))$

Now answer the following questions.

(a) (4 marks) Assuming some shape for $X(f)$, draw the spectrum of the sampled version of $x_1(t)$, when the sampling rate is $1/T_s$, where $1/T_s$ is slightly above Nyquist rate. Mark important values carefully on the axes.

(b) (3 marks) Remark with reasons on which of the signals $x_1(t)$ and $x_2(t)$ is closest to the idea of (analog) modulation presented in the class.

(c) (6 marks) If the bandwidth of $x(t)$ is W, what can you say about the bandwidth of $x_1(t)$ and $x_2(t)$? (Hint: you may want to remember your answer for part (b))

8. (7 marks) Assume that X and Y are random variables with joint distribution $p_{X,Y}$. The definition of the expectation of a function $g(X,Y)$ of the two random variables X and Y, is given by

$$\mathbb{E}(g(X,Y)) \triangleq \sum_{x,y} p_{X,Y}(x,y)g(x,y).$$

(a) (5 marks) Using the given statements, show that $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$ (Hint: You have to use definitions of $\mathbb{E}(X)$ and $\mathbb{E}(Y)$).

(b) (2 marks) Write a generic theorem which captures the similar result for any number of random variables, including proper assumptions and definitions for the quantities similar to what is given in this question. You don't need to prove this, just give the statement as clearly as the question is given.

$x(2(t-2))$

$x(t-2) \rightarrow e^{j2\pi f 2} \cdot X(f)$

$x(3(t-2)) \rightarrow e^{j2\pi f 2} \cdot \frac{1}{3} X(\ldots$

$x(3t) \rightarrow \frac{1}{3} X(\frac{f}{3})$

$x(3t-6) \rightarrow e^{j2\pi f}$