Introduction to Information and Coding theory 10 July 2012

Dr. Permuter Haim, Mr. Ziv Foldfeld and Mr. Lior Dikstein

Final Exam (Moed A)

1) **True or False**
Copy each relation to your notebook and write **true** or **false**. Then, if it's true, prove it. If it is false give a counterexample or prove that the opposite is true.
   a) Let X, Y, Z be three random variables. Then $H(X,Y,Z|g(X)) = H(X,Y,Z) - I(X;g(X))$, for any function g.
   b) Let X, Y be jointly Gaussian random variables and let $\hat{X}^{\text{lin}}$ be the MMSE (minimum mean square error) linear estimator of X given Y.

$$h(X|Y) \overset{(*)}{=} h(X - \hat{X}^{\text{lin}}|Y) \overset{(**)}{=} h(X - \hat{X}^{\text{lin}})$$

      i) Is $(*)$ true/false?
      ii) Is $(**)$ true/false?
   c) Which of the following sequences of code-lengths are valid binary Huffman codes?
      i) 1,2,3,4,4
      ii) 2,2,2,3,3
      iii) 2,2,3,3,3
   d) Let $f(x)$ be a convex function. Is the function $(a+bt)f(\frac{x}{a+bt})$ convex for all $a, b \neq 0$?

2) **Channel with state**
A discrete memoryless (DM) state dependent channel with state space $\mathcal{S}$ is defined by an input alphabet $\mathcal{X}$, an output alphabet $\mathcal{Y}$ and a set of channel transition matrices $\{p(y|x,s)\}_{s\in\mathcal{S}}$. Namely, for each $s \in \mathcal{S}$ the transmitter sees a different channel. The capacity of such a channel where the state is know causally to both encoder and decoder is given by:

$$C = \max_{p(x|s)} I(X;Y|S). \tag{1}$$

Let $|\mathcal{S}| = 3$ and the the three different channel (one for each state $s \in \mathcal{S}$) are as illustrated in Fig.1

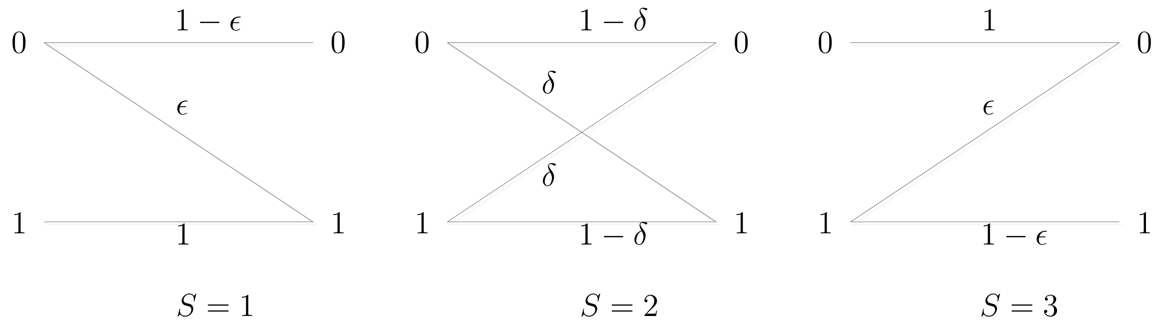

Fig. 1. The three state dependent channel.

The state process is i.i.d. according to the distribution $p(s)$.
   a) Find an expression for the capacity of the S-channel (the channel the transmitter sees given $S = 1$) as a function of ϵ.
   b) Find an expression for the capacity of the BSC (the channel the transmitter sees given $S = 2$) as a function of δ.
   c) Find an expression for the capacity of the Z-channel (the channel the transmitter sees given $S = 3$) as a function of ϵ.

d) Find an expression for the capacity of the DM state dependent channel (using formula (1)) for $p(s) = [\frac{1}{2} \;\; \frac{1}{3} \;\; \frac{1}{6}]$ as a function of ϵ and δ.

e) Let us define a conditional probability matrix $P_{X|Y}$ for two random variables X and Y with $|\mathcal{X}| = \{0, 1\}$ and $|\mathcal{Y}| = \{1, 2, 3\}$, by:

$$\left[P_{X|Y}\right]_{i=1,j=1}^{3,2} = p(x = j - 1|y = i). \tag{2}$$

What is the input conditional probability matrix $P_{X|S}$ that achieves the capacity you have found in (d)?

3) **Entropy of a Markov process**

Consider a binary Markov chain with the following transition matrix:

$$P = \begin{pmatrix} 1-a & a \\ 1 & 0 \end{pmatrix}.$$

The matrix entry i, j is the probability $\Pr(x_n = i-1|x_{n-1} = j-1)$. For example: $\Pr(x_n = 0|x_{n-1} = 1) = [P]_{1,2} = a$.

a) A stationary probability vector is a vector, μ, such that $\mu P = \mu$. This means that if we have an initial state drawn according to the stationary distribution, the next state, which is determined using the transition matrix, is distributed the same. Find the stationary vector for the given P.

b) Assume that the Markov chain defined by P has an initial state drawn according to the stationary distribution found earlier. The entropy rate of a stochastic process $\{X_i\}$ is defined by

$$H(\mathcal{X}) = \lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \ldots, X_n).$$

Find the entropy rate of the Markov chain defined by P.

4) **Network coding for disk array**

Assume we have a source of information which we want to save for back-up on three different hard-disks of the same size such that given any two hard-disks we can reconstruct the whole source without an error. For instance, we can simply copy the source to each hard disk, but in such a case we need disks of the size of the sources, hence if the sources are 12Gbit we need 3 hard-disks of 12Gbit.

a) What is the minimum size of hard-disks that is needed in order to be able to reconstruct the data from any two hard-disks, assuming the source size is 12Gbit?

b) Draw a network-coding problem and show how this can be achieved. (hint: the source of the network coding is the source of data, there should be three (relay) nodes that are the hard-disks and three destination nodes where each destination has access to two hard-disks).

c) Now consider the case where there are 4 hard-disks used for back up and we should be able to reconstruct the data without loss from every two of them. Draw the network-coding problem and solve it.

Good Luck!