Final Exam - Moed B
Total time for the exam: 3 hours!

Important: For **True / False** questions, copy the statement to your notebook and write clearly true or false. You should prove the statement if true, and provide counterexample otherwise.

1. **True or False (20 Points).**
   (a) If $X - Y - Z - W$ is a Markov chain, then $X - Y - Z$ and $Y - Z - W$ are Markov chains.
   (b) If $X - Y - Z$ and $Y - Z - W$ are Markov chains, then $X - Y - Z - W$ is a Markov chain.
   (c) If $P(x, y, z, w) = P(x)p(y|x)P(z, w)$, then $X - Y - (Z, W)$ is a Markov chain.
   (d) If $X \perp\!\!\!\perp Y$ and $Y \perp\!\!\!\perp Z$, then $X \perp\!\!\!\perp Z$.
   (e) If the conditional distribution $P(x|y, z)$ is a deterministic function of (x, y), then $X - Y - Z$ is a Markov chain.

2. **Markov with random index (10 points)** Let $A_1 - B_1 - C$ and $A_2 - B_2 - C$ be Markov chains, and let $i(C)$ be a binary (deterministic) function of C that emits 1 or 2.
   (a) $A_{i(C)} - B_{i(C)} - C$ holds if $P_{A_1,B_1}(a, b) = P_{A_2,B_2}(a, b)$ for all a, b in the alphabet.
   (b) $A_{i(C)} - B_{i(C)} - C$ holds if $P_{A_1,B_1}(a, b) \neq P_{A_2,B_2}(a, b)$ for some a, b in the alphabet.

3. **Erasure channel after discrete memoryless channel (20 Points):**
   Assume a discrete memoryless channel, $(\mathcal{X}, \mathcal{Y}, p(y|x))$ with capacity, C_1.
   The output of this channel serves as an input to an *erasure channel* with $|\mathcal{Y}|$ inputs and erasure probability ϵ.
   What is the capacity of the overall channel?

4. **Cross entropy (25 Points)**:
   Often in Machine learning, cross entropy is used to measure performance of a classifier model such as neural network. Cross entropy is defined for two PMFs P_X and Q_X as

$$H(P_X, Q_X) \stackrel{\triangle}{=} -\sum_{x\in\mathcal{X}} P_X(x) \log Q_X(x)$$

   In a shorter notation we write as

$$H(P, Q) \stackrel{\triangle}{=} -\sum_{x\in\mathcal{X}} P(x) \log Q(x)$$

   (a) Copy each of the following relations to your notebook and write **true** ir **false** and provide a proof/disproof.
      i. $0 \leq H(P, Q) \leq \log |\mathcal{X}|$ for all P, Q.
      ii. $\min_Q H(P, Q) = H(P, P)$ for all P.
      iii. $H(P, Q)$ is concave in the pair (P, Q).
      iv. $H(P, Q)$ is convex in the pair (P, Q).
   (b) Find an operation problem, such as in compression, communication (or even other fields) where the fundamental solution involve the cross entropy measure $H(P, Q)$. State the operational problem mathematically in less than half a page, and state the solution as a theorem. Provide a short proof to the theorem.

5. **Fast fading Gaussian channel (25 points)**:
   Consider a Gaussian channel given by $Y_i = G_iX_i + Z_i$, where $Z_i \stackrel{i.i.d}{\sim} \mathcal{N}(0, N)$ and $G_i \stackrel{i.i.d}{\sim} P_G(g)$.

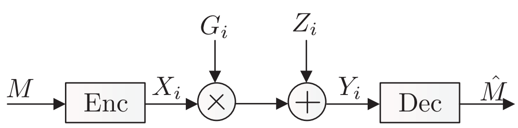

Figure 1: Fast fading Gaussian channel

The gains and noise are independent, i.e., $\{Z_i\} \perp\!\!\!\perp \{G_i\}$, and

$$P_G(g) = \begin{cases} 0.5 & \text{if } g = 1 \\ 0.5 & \text{if } g = 2 \end{cases}$$

(a) Assume that the states are known at the decoder only, and there is an input constraint P.

i. What is the capacity formula?

ii. Find the optimal inputs distribution in the formula you gave.

iii. Compute the capacity as a function of N and P.

(b) Now the states are known both to the encoder and decoder, and the input constraint is P.

i. What is the capacity formula?

ii. Compute the capacity as a function of N and P.
You can write your answer as an optimization problem.

(c) Assume

$$P_G(g) = \begin{cases} 0.5 & \text{if } g = 0 \\ 0.5 & \text{if } g = 1 \end{cases}.$$

Repeat 5b.

Good Luck!