Introduction to Information and Coding Theory July 24, 2016
(Prof. Permuter Haim, Mr. Oron Sabag and Mr. Iddo Gattengo)

Final Exam - Moed A
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) **Cascaded BSCs (21 Points)**: Given is a cascade of k identical and independent binary symmetric channels, each with crossover probability α.
   a) In the case where no encoding or decoding is allowed at the intermediate terminals, what is the capacity of this cascaded channel as a function of k, α.
   b) Now, assume that encoding and decoding is allowed at the intermediate points, what is the capacity as a function of k, α.
   c) What is the capacity of each of the above settings in the case where the number of cascaded channels, k, goes to infinity?
2) **True or False on conditional independence probabilities (10 Points)**:
   Given are three discrete random variables X, Y, Z.
   a) **True**/**False**: If $X \perp\!\!\!\perp Y$ then $X \perp\!\!\!\perp Y|Z$.
   b) **True**/**False**: If $X \perp\!\!\!\perp Y|Z$ then $X \perp\!\!\!\perp Y$.
3) **Disjoint sets on discrete random variable (28 Points)**:
   Let X_0 and X_1 be discrete random on the alphabets $\mathcal{X}_0 = \{1, ..., m\}$ and $\mathcal{X}_1 = \{m+1, ..., n\}$, respectively. Let θ be a binary random variable with $P(\theta = 1) = p$, for some p. Let
   $$X = \begin{cases} X_0 & \text{if } \theta = 0 \\ X_1 & \text{if } \theta = 1. \end{cases}$$
   a) Find $H(X)$ in terms of $H(X_0)$, $H(X_1)$, and p.
   b) Prove the inequality: $2^{H(X)} \leq 2^{H(X_0)} + 2^{H(X_1)}$.
   c) Find a sufficient and necessary condition for equality to hold in (b). The condition should be stated using $H(X_0)$, $H(X_1)$, and p only.
   d) Using (b), prove that $H(X) \leq \log |\mathcal{X}|$.
4) **True or False on the concatenation order (10 Points)**: Given are channel A and channel B both have binary inputs and binary outputs. The channels are concatenated so the output of the channel A is the input to channel B and the capacity of this channel is denoted by $C_{A\to B}$. The definition of $C_{B\to A}$ is similar but channel B comes first.
   a) **True**/**False**: If channels A and B are binary symmetric channels, then $C_{A\to B} = C_{B\to A}$.
   b) **True**/**False**: For arbitrary binary channels, the order of the concatenation has no effect on the capacity.
5) **Huffman Code (31 Points)** : Let X^n be a an i.i.d. source that is distributed according to p_X:

x	0	1	2	3
$p_X(x)$	0.5	0.25	0.125	0.125

   a) What is the optimal lossless compression rate R^* for the source sequence? (4 points)
   b) Build a binary Huffman code for the source X. (4 points)
   c) What is the expected length of the resulting compressed sequence. (4 points)
   d) What is the expected number of zeros in the resulting compressed sequence. (5 points)
   e) Let $\tilde{X}^n$ be an another source distributed i.i.d. according to $p_{\tilde{X}}$.

$\tilde{x}$	0	1	2	3
$p_{\tilde{X}}(\tilde{x})$	0.3	0.4	0.1	0.2

   What is the expected length of compressing the source $\tilde{X}$ using the code constructed in (b). (4 points)
   f) Answer (d) for the code constructed in (b) and the source $\tilde{X}^n$. (5 points)
   g) Is the relative portion of zeros (the quantity in (d) divided by the quantity in (c)) after compressing the source X^n and the source $\tilde{X}^n$ different? For both sources, explain why there is or there is not a difference. (5 points)

Good Luck!