ECE 634/CSE 646 InT: Final Examination

Instructor: Manuj Mukherjee

Total: 30 points

Instruction: Answer all questions from the INFANT level, and then answer either three questions from the ADULT level, or one one question from the ADULT level and one question from the SENSEI level. SENSEI level questions will be used to distinguish between A and A+ grades.

In the paper $h(\cdot)$ is used for both the binary entropy and the differential entropy. It should be clear from the context what a specific instance of $h(\cdot)$ means since argument to binary entropy is a number in $[0, 1]$, whereas the argument to differential entropy is a random variable.

I. INFANT LEVEL

1) Find the differential entropy of an $\exp(\lambda)$ random variable.

   In case you have forgotten, the pdf of an $\exp(\lambda)$ random variable, where $\lambda > 0$, is given by $\mathrm{f}(x) = \lambda e^{-\lambda x}$, for $x \geq 0$.

   [5 points]

2) Consider the channel shown in Figure 1, where $\mathrm{P}_{Y|X}(i|i) = 1 - \epsilon$ for all $i = 0, 1, 2$. Find the capacity of this channel.

   [5 points]

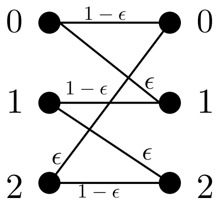

Fig. 1: Figure for Prob. 2

3) As shown in Figure 2 we have fed the output Y of a BSC(ϵ) to an arbitrary channel to obtain Z. If $X \sim$ Be$(1/2)$, can you write down a strictly positive lower bound on $I(X; Z)$?

   [5 points]

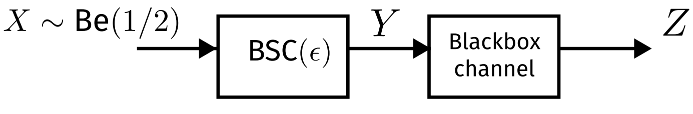

Fig. 2: Figure for Prob. 3

II. Adult level

4) *['Combined' DMCs]* Consider two DMCs $(\mathcal{X}_0, \mathcal{Y}_0, \mathrm{P}^{(0)}_{Y|X})$ and $(\mathcal{X}_1, \mathcal{Y}_1, \mathrm{P}^{(1)}_{Y|X})$, with respective capacities C_0 and C_1, and consider a DMC obtained by combining them, i.e., the new DMC has input alphabet $\mathcal{X}_0 \cup \mathcal{X}_1$, output alphabet $\mathcal{Y}_0 \cup \mathcal{Y}_1$,[1] and the transition probability $\mathrm{P}_{Y|X}$ being given by

$$\mathrm{P}_{Y|X}(y|x) = \begin{cases} 0, & \text{if } y \in \mathcal{Y}_i, x \in \mathcal{X}_{\bar{i}}, i = 0,1 \\ \mathrm{P}^{(i)}_{Y|X}(y|x), & \text{if } x \in \mathcal{X}_i, y \in \mathcal{Y}_i, i = 0,1. \end{cases}$$

a) Let X, Y be the random variables corresponding to the input and output of the parallel channel. Let $Z \sim \mathrm{Be}(p)$, and let $X \in \mathcal{X}_i$ if $Z = i$, and the Markov chain $Z - X - Y$ holds. Argue that Z can be completely determined by either X or Y, i.e., $H(Z|Y) = H(Z|X) = 0$, and hence show that $I(X;Y) = h(p) + I(X;Y|Z)$.

[**Hint**: Expand $I(X;YZ)$ in two-ways using chain rule.]

b) Let C be the capacity of the parallel DMC. Use part a) to argue that $C = \max_{p\in[0,1]} \max_{\mathrm{P}_{X|Z=1}} \max_{\mathrm{P}_{X|Z=0}} \{h(p) + I(X;Y|Z)\}$, and hence show that $C = \max_{p\in[0,1]}\{h(p) + pC_1 + (1-p)C_0\}$.

c) Use part c) to show that $2^C = 2^{C_0} + 2^{C_1}$.

[**Hint**: Use Captain Haddock's best mate, Calculus. Also, you can argue for the maximization without using the second derivative test through properties of concave functions.]

[1.5+1+2.5= 5 points]

5) *[Information theoretic security]* A symmetric key cryptosystem consists of two parties with a shared key K, which is independent of the message M. The encrypting party computes a function $C = f(M,K)$, where C is called the ciphertext, and sends it to a decrypting party, which uses a decryption function g to get back the exact message M as $M = g(K,C)$. This system is said to be information theoretically secure if $M \perp C$.

a) Show that $H(K) \geq H(M)$ is necessary for information theoretic security.[2]

[**Hint**: Expand $I(C;MK)$ in two different ways, and use the various properties of the cryptosystem such as $M \perp K$, $M \perp C$, $H(M|K,C) = 0$, and so on.]

b) Let $K, M \sim$ i.i.d Be$(1/2)$, and let $C = K \oplus M$. Show that this is an information theoretic secured symmetric key cryptosystem, and identify the decryption function g.[3]

[4+1 = 5 points]

[1] Here you should assume that $\mathcal{X}_0, \mathcal{X}_1$ (also, $\mathcal{Y}_0, \mathcal{Y}_1$) are mutually exclusive. Suppose not, we simply relabel the elements to ensure this. For example, if both channels were BSCs, simply relabel $\mathcal{X}_1 = \{0', 1'\}$, and $\mathcal{Y}_1 = \{0', 1'\}$, and hence $\mathcal{X}_0 \cup \mathcal{X}_1 = \mathcal{Y}_0 \cup \mathcal{Y}_1 = \{0, 1, 0', 1'\}$.

[2] This is a classical result by Shannon from 1949.

[3] This scheme is called Vernam's one-time pad.

6) Let $X \sim$ Poi(λ), and let $Y|_{X=i} \sim N(i, 2^{2i})$. Find $h(Y|X)$.

   [**Hint**: In case you have forgotten, if $X \sim$ Poi(λ), where $\lambda > 0$, then $\mathrm{P}_X(i) = \frac{e^{-\lambda}\lambda^i}{i!}$, for all $i \in \mathbb{Z}$.]

[5 points]

7) Let X_1, X_2 be i.i.d. uniform bits and let $X_3 = X_1$ AND X_2. Let $g : \{0,1\} \to \{0,1\}^2$ be any function used to estimate X_1, X_2 from X_3. That is, on observing X_3, we declare $g(X_3)$ to be an estimate of X_1, X_2. Let $\mathrm{P}_e \triangleq \mathrm{P}((X_1, X_2) \neq g(X_3))$ be the probability of error with the estimating function g. Show that $\mathrm{P}_e \geq \frac{3}{8}\log 3 - \frac{1}{2}$.

   [**Hint**: Try to think about how we prove converses of channel coding theorems, and recall how lower bounds on the probability of error are obtained.]

[5 points]

III. Sensei level

8) We will prove that for $k \leq \frac{n}{2}$, $\sum_{i=0}^{k}\binom{n}{i} \leq 2^{nh(\frac{k}{n})}$. Assume that $X_1, \ldots, X_n$ represent the coordinates of a string drawn uniformly at random from all n-bit strings of weight at most k.

   a) Note that by symmetry $X_1, \ldots, X_n$ are identically distributed, but not independent. Use this property and the fact that X^n has weight at most k to show that $\mathbf{E}[X_i] \leq \frac{k}{n}$.

   b) Show that $h(p) \geq p$ for any $p \in [0, \frac{1}{2}]$.

   c) Using a) and b), show that $H(X_1, \ldots, X_n) \leq nh(\frac{k}{n})$.

   d) Using c) and the definition of X^n, complete the proof.

   e) In which of the steps did you use the fact that $k \leq \frac{n}{2}$?

[3+1.5+4+1+0.5 = 10 points]

9) *[Leftover Hash Lemma]* Let $k \geq 1$, let $\mathcal{U}$ be a finite set, and let $\mathcal{F}$ be a family of functions $\phi : \mathcal{U} \to [k]$, where $[k] \triangleq \{1, 2, \ldots, k\}$. The family $\mathcal{F}$ is called a universal hash family (UHF) if for any $u \neq u'$, the family satisfies

$$\frac{1}{|\mathcal{F}|}\sum_{\phi \in \mathcal{F}} \mathbf{1}\{\phi(u) = \phi(u')\} \leq \frac{1}{k}.$$

   In this question, we shall prove the leftover hash lemma, which states that any universal hash family $\mathcal{F}$ satisfies

$$\frac{1}{|\mathcal{F}|}\sum_{\phi \in \mathcal{F}} \mathrm{TV}(\mathrm{P}_{\phi(U)}, \mathrm{unif}\{[k]\}) \leq \frac{1}{2}\sqrt{k2^{-H_{\min}(\mathrm{P}_U)}},$$

   where $H_{\min}(\mathrm{P}_U) \triangleq \min_{u \in \mathcal{U}} \log \frac{1}{\mathrm{P}_U(u)}$.

   a) Show that for any $\phi : \mathcal{U} \to [k]$, we have $\mathrm{TV}(\mathrm{P}_{\phi(U)}, \mathrm{unif}\{[k]\}) \leq \frac{1}{2}\sqrt{k\sum_{i=1}^{k}(\sum_{u:\phi(u)=i}\mathrm{P}_U(u) - \frac{1}{k})^2}$.

   b) Show that $\sum_u \mathrm{P}_U(u)^2 \leq 2^{-H_{\min}(\mathrm{P}_U)}$.

      [**Hint**: You will need to use Jensen's inequality.]

c) Use part b) to show that $\sum_{i=1}^{k}(\sum_{u:\phi(u)=i} \mathrm{P}_U(u) - \frac{1}{k})^2 \leq 2^{-H_{\min}(\mathrm{P}_U)} + \sum_{u,u':u\neq u'} \mathrm{P}_U(u)\mathrm{P}_U(u')\mathbf{1}\{\phi(u) = \phi(u')\} - \frac{1}{k}$.

d) Use parts a) and c) to complete the proof.

[**Hint**: You will first need to use Jensen's inequality, and somewhere down the line the definition of UHF.]

[2+2+3+3 = 10 points]