CIE6020: Selected Topics in Information Theory

Final Examination
SSE, CUHK(SZ)

May 9, 2018

Name: ______________________ Student ID: ______________________

Answer all the questions in the Answer Book.
No questions in this page!

1. (12 points) We wish to encode a Bernoulli(α) process for transmission over a binary symmetric channel with crossover probability p. For an integer $n > 0$, the encoding function $f : \{0,1\}^n \to \{0,1\}^n$ maps n bits of the source to n bits to be transmitted through the channel.

   (a) (5 points) Find conditions on α and p so that the probability of error can be made to go to zero as n tends to infinity.

   Solution: By the source-channel separation theorem, $H(\alpha) < 1 - H(p)$ is a necessary and sufficient condition such that the source can be reliabily transfered through the channel.

   (b) (7 points) Give an example of f and the corresponding decoding algorithm such that the probability of error can be made to go to zero as n tends to infinity. (You only need to outline the justification.)

   Solution: Let T_p^n be the typical set of Bernoulli(p). We can define the encoding function f as follows: First find uniformly at random a subset $\mathcal{C}$ of $T_{1/2}^n$ with cardinality $|T_\alpha^n|$, which is feasible since $|T_\alpha^n| < |T_{1/2}^n|$. f maps each $\mathbf{x} \in T_\alpha^n$ is mapped to a unique sequence in $\mathcal{C}$. The sequnces not in T_α^n are count as decoding error, which contribultes a vanishing probablity. As the rate $\mathcal{C}$ is less than the capacity of the BSC, we know the error probability of decoding error of the sequences in T_α^n tends to zero as $n \to \infty$.

2. (21 points) Consider a linear code with parity check matrix

$$H = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

   (a) (3 points) Give a generator matrix of the code.

   Solution:

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

   (b) (3 points) What is the minimum distance of the code? What are the error correction and erasure correction capabilities of the code?

   Solution: $d_{\min} = 3$, 1-error correction and 2-erasure correction.

   (c) (3 points) Draw the Tanner graph of the code using the parity check matrix H.

   (d) (5 points) Use the code for a binary erasure channel. Suppose the outputs of the channel are $(0, ?, ?, 1, 0, ?, 0)$, where ? represents an erasure. When using the message passing decoder for erasure channels, what is the output codeword? Please illustrate the message passing procedure.

Solution: $(0, 1, 0, 1, 0, 1, 0)$.

(e) (2 points) Use the code for a binary symmetric channel and the outputs are $(1, 0, 1, 1, 1, 0, 1)$. What is output codeword of the minimum distance decoder?

Solution: $(1, 0, 1, 0, 1, 0, 1)$.

(f) (5 points) Consider the use of Gallager algorithm A for decoding the above outputs, where a variable node sends to check node c

- value b if all the incoming messages from check nodes other than c are the same value b, or
- its received value otherwise (or at round 0).

What is value sent by a check node to its neighboring variable node in Gallager algorithm A? Please illustrate the decoding procedure and give the output of the algorithm for one round of message passing.

Solution: For a neighboring variable node v, a check node sends the sum of the values it gets from all neighboring variable nodes other than v.

3. (12 points) A low-density generator matrix (LDGM) code has a sparse generator matrix. (Consider LT code as an example.) This question is about binary (n, k) LDGM code ensemble. The generator matrix of this code can be described by a Tanner graph with k variable nodes and n check nodes. The variable nodes degree distribution is $L = (L_1, L_2, \dots, L_{l_{\max}})$ and the check node degree distribution is $R = (R_1, R_2, \dots, R_{r_{\max}})$. In other words, the fraction of the variable (check) nodes with degree i is L_i (R_i). The random connection betwee the variable nodes and the check nodes is defined similar to that of the LDPC code ensemble.

(a) (3 points) Give an equality constraint of L, R, n and k. What is the design rate of the code in terms of L and R?

Solution: $n \sum iR_i = k \sum iL_i$.

(b) (3 points) Suppose that we use the code in a binary erasure channel with erasure probability ϵ. Give an upper bound on the achievable rate of the code, which should be less than 1.

Solution: The achievable rate is upper bounded by the capapcity of the erasure channel $1 - \epsilon$.

(c) (3 points) Describe a message passing decoding algorithm of the code for erasure channels.

(d) (3 points) Suppose that $L_4 = 1$ and $R_3 = 1$. Draw a typical subgraph related to a variable node for two rounds of message passing when n is large .

4. (10 points) This problem is about AND-OR tree analysis. An AND-OR tree of $l+1$ levels is formed as follows. The nodes of the tree are in levels labelled from 0 to l. The root is at level l, the children of the root is at level $l-1$, so and so forth. The nodes at levels $0, 2, 4, \dots$ are OR nodes and the nodes at levels $1, 3, \dots$ are AND nodes. Associated with each node is a binary random variable. The random variables in the same level are independent. The value of an OR node is one if and only if one of its descendent AND nodes has value one. The value of an AND node is one if and only if all of its descendent OR nodes have value one.

(a) (5 points) Fig. 1 is a deterministic AND-OR tree of three levels with the random variables associated with each node labelled. According to the definition, we have

$$\begin{aligned}X_1 &= Y_1^1 \vee Y_1^2,\\ Y_1^1 &= X_0^1 \wedge X_0^2 \wedge X_0^3,\\ Y_1^2 &= X_0^4 \wedge X_0^5.\end{aligned}$$

Suppose that $\Pr\{X_0^i = 1\} = p$, $i = 1, \dots, 5$. Calculate $\Pr\{X_1 = 1\}$.

Solution: First, $\Pr\{Y_1^1 = 1\} = \Pr\{X_0^1 = 1, X_0^2 = 1, X_0^3 = 1\} = p^3$ and similarly $\Pr\{Y_1^2 = 1\} = p^2$. Hence $\Pr\{X_1 = 1\} = \Pr\{Y_1^1 = 1 \cup Y_1^2 = 1\} = 1 - \Pr\{Y_1^1 = 0, Y_1^2 = 0\} = 1 - (1-p^3)(1-p^2)$.

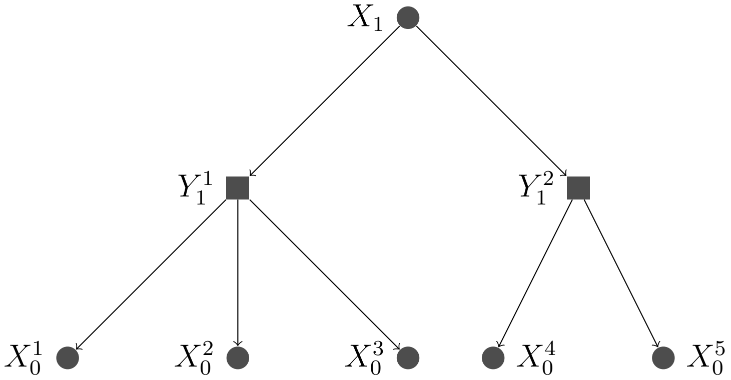

Figure 1: The example of an And-Or tree.

(b) (5 points) Consider a random AND-OR tree of three levels, where the root (OR) node X_1 has i children with probability λ_i and each AND node has i children with probability ρ_i. Suppose that each node in level 0 has value 1 with probability p. Give the formula of $\Pr\{X_1 = 1\}$ in terms of p, $\{\rho_i\}$ and $\{\lambda_i\}$.

Solution: $\Pr\{X_1 = 1\} = 1 - \sum_i \lambda_i (1 - \sum_j \rho_j p^j)^i = 1 - \lambda(1 - \rho(p))$.

5. (12 points) Consider random variable $X \in \{0, 1\}$ with the distribution Bernoulli(1/2), and let the distortion measure be given by

$$d(x, \hat{x}) = \begin{cases} 0 & x = \hat{x}, \\ \infty & \hat{x} = x \oplus 1, \\ 1 & \hat{x} = e, \end{cases}$$

where $\oplus$ denotes exclusive or, and e represents a symbol that is not in $\{0, 1\}$. Let $R(D)$ be the rate distortion function for this source.

   (a) (4 points) For a given transition matrix $Q(\hat{x}|x)$ such that $E[d(X, \hat{X})]$ is finite, calculate $E[d(X, \hat{X})]$.

   (b) (2 points) What is $R(0)$?

   (c) (2 points) Find the minimum value of D such that $R(D) = 0$.

   (d) (4 points) Calculate $R(D)$.

Solution:

(a) We see that $E[d(X, \hat{X})] < \infty$ iff $Q(1|0) = Q(0|1) = 0$. So $E[d(X, \hat{X})] = \frac{1}{2}[Q(e|0) + Q(e|1)] = p_{\hat{X}}(e)$.

(b) To have the zero distortion, $Q(0|0) = Q(0|0) = 1$. Hence $I(X, \hat{X}) = 1$, i.e., $R(0) = 1$.

(c) $R(1) = 0$ using $Q(e|0) = Q(e|1) = 1$.

(d) Consider $0 < D < 1$ and $Q(\hat{x}|x)$ such that $E[d(X, \hat{X})] \leq D$, i.e., $Q(1|0) = Q(0|1) = 0$ and $p_{\hat{X}}(e) \leq D$. Hence $I(X; \hat{X}) = H(X) - H(X|\hat{X}) = 1 - p_{\hat{X}}(e) \geq 1 - D$. So $R(D) = 1 - D$.

6. (20 points) (Channel polarization) For a binary input channel W, recall that

$$Z(W) = \sum_y \sqrt{W(y|0)W(y|1)}.$$

Consider the two transformations $^-$ and $^+$ of W defined as follows:

$$W^-(y_1, y_2|x_1) = \sum_{x_2 \in \{0,1\}} \frac{1}{2} W(y_1|x_1 \oplus x_2) W(y_2|x_2)$$

and

$$W^+(y_1, y_2, x_1|x_2) = \frac{1}{2} W(y_1|x_1 \oplus x_2) W(y_2|x_2).$$

Note that both W^- and W^+ are binary input channels. Suppose that the input distributions of these binary input channels are *uniform*.

(a) (5 points) Show that $I(W^-) + I(W^+) = 2I(W)$.

Solution: $I(W^-)+I(W^+) = I(X_1;Y_1,Y_2)+I(X_2;X_1,Y_1,Y_2) = I(X_1;Y_1,Y_2)+I(X_2;Y_1,Y_2|X_1) = I(X_1,X_2;Y_1,Y_2) = I(X_1 \oplus X_2, X_2; Y_1, Y_2) = I(X_1 \oplus X_2; Y_1) + I(X_2;Y_2) = 2I(W)$.

(b) (5 points) Consider that W is the binary erasure channel. Show that $Z(W^+) = Z(W)^2$ and $Z(W^-) = 2Z(W) - Z(W)^2$.

Solution: For the erasure channel with erasure probability ρ, $Z(W) = \rho$. As, $Z(W^-) = 2\rho - \rho^2$ and $Z(W^+) = \rho^2$, the question is solved.

(Some notations to be used in (c) and (d).) For integer $n \geq 0$ and integer $0 \leq i \leq 2^n - 1$, define

$$W_{n+1}^{(2i)} = (W_n^{(i)})^-,$$

and

$$W_{n+1}^{(2i+1)} = (W_n^{(i)})^+,$$

with $W_0^{(0)} = W$. Let $Y(W) = Z(W)(1 - Z(W))$.

(c) (5 points) Show that for the binary erasure channel

$$Y(W_{n+1}^{(2i)}) = Y(W_n^{(i)})(1 - Z(W_n^{(i)}))(2 - Z(W_n^{(i)})),$$

and

$$Y(W_{n+1}^{(2i+1)}) = Y(W_n^{(i)})Z(W_n^i)(1 + Z(W_n^i)).$$

(d) (5 points) Show that for the binary erasure channel

$$\frac{1}{2^n}\sum_{i=0}^{2^n-1}\sqrt{Y(W_n^{(i)})} \leq \frac{1}{2}\left(\frac{\sqrt{3}}{2}\right)^n.$$

(Hint: apply $\sqrt{z(1+z)} + \sqrt{(1-z)(2-z)} \leq \sqrt{3}$.)

7. (13 points) Let $A(n,d)$ be the maximum number of codewords in any binary code of length n and minimum distance d. Let $\mathcal{C}$ be an (n, M, d) code with $M = A(n,d)$ and $d = 2t+1$. Let S_i be the set of vectors at distance i from $\mathcal{C}$, i.e.,

$$S_i = \{u \in \{0,1\}^n, d(u,c) \geq i \ \forall c \in \mathcal{C} \text{ and } d(u,v) = i \text{ for certain } v \in \mathcal{C}\}.$$

(a) (5 points) Show that $S_0 \cup S_1 \cup \cdots \cup S_{d-1} = \{0,1\}^n$.

(b) (4 points) Find the cardinality of S_i for $i = 0, 1, \ldots, t$.

(c) (4 points) Find an upper bound on $A(n,d)$.

(d) (5 points (bonus)) Can you estimate S_{t+1}?

——— No more questions! ———