Information and coding theory HW#1

Evgeny Marshakov

Problem 1

We know that

$$I(X;Z|Y) = H(X|Y) - H(X|Y,Z) = H(X,Y) - H(Y) - H(X,Y|Z) + H(Y|Z)$$

So we have that

$$I(X;Z|Y)+I(Y;Z) = H(X,Y)-H(Y)-H(X,Y|Z)+H(Y|Z)+H(Y)-H(Y|Z) = H(X,Y)-H(X,Y|Z) = I(X,Y;Z)$$

Problem 2

Using the chain rules for $H(X^n)$ and $H(X_{\bar{i}})$ we obtain

$$H(X^n) = \sum_{j=1}^{n} H(X_j|X_{j-1},\dots,X_1)$$

$$H(X_{\bar{i}}) = \sum_{j=1}^{n} H(X_j|X_{\bar{i}}^{j-1}) = \sum_{j=1}^{n} H(X_j|X_{j-1},\dots,X_{i+1},X_{i-1},\dots,X_1)$$

Obviously, we have the following inequalities

$$\begin{cases} H(X_j|X_{j-1},\dots,X_1) \le H(X_j|X_{\bar{i}}^{j-1}), & i<j \\ H(X_j|X_{j-1},\dots,X_1) = H(X_j|X_{\bar{i}}^{j-1}), & i>j \end{cases}$$

We see that there is no term corresponding to i in $H(X_{\bar{i}})$. So using these inequalities we can easily find that

$$\sum_{j\neq i} H(X_j|X_{j-1},\dots,X_1) \leq \sum_{j=1}^{n} H(X_j|X_{\bar{i}}^{j-1})$$

or

$$H(X^n) - H(X_i|X_{i-1},\dots,X_1) \leq H(X_{\bar{i}})$$

Summing over i we obtain

$$\sum_{i=1}^{n} [H(X^n) - H(X_i|X_{i-1},\dots,X_1)] \leq \sum_{i=1}^{n} H(X_{\bar{i}}) \Leftrightarrow (n-1)H(X^n) \leq \sum_{i=1}^{n} H(X_{\bar{i}})$$

Problem 3

1. Let us consider the following probabilities

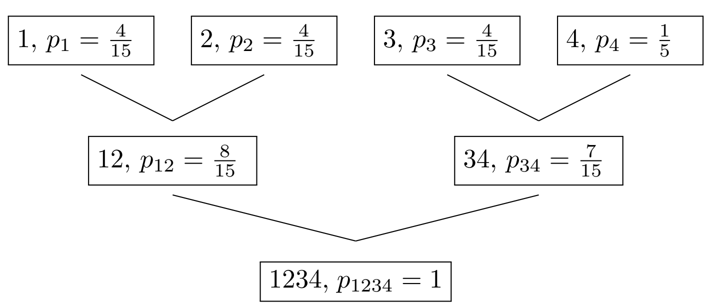
$$p_i = \mathbb{P}(X = i) = \frac{\#i \textit{ in the string}}{\#\textit{all elements in the string}}$$

   More precisely,

$$p_1 = p_2 = p_3 = \frac{4}{15},\ p_4 = \frac{1}{5}$$

2. $H(X) = -3 \cdot \frac{4}{15}\log_2\frac{4}{15} - \frac{1}{5}\log_2\frac{1}{5} \approx 1.98$ bits.

3.



So the Huffman code for these probabilities is the following

1	00
2	01
3	10
4	11

It can be easily seen that the average number of bits is 2.

4. Encoded file F

$$F \to 0000101010101111011101010101100000110110110010000100001110 01$$

The average number of bits in the file F is also 2.

Problem 4

1. We want to find the capacity of the Z channel. Let p_0 be the probability $\mathbb{P}(X=0)$. Set

$$q_0 = \mathbb{P}(Y=0) = p_0 + (1-p_0)\delta$$

$$q_1 = \mathbb{P}(Y=1) = (1-p_0)(1-\delta)$$

Then we have

$$H(Y) = -q_0 \log_2(q_0) - q_1 \log_2(q_1) = -[p_0+(1-p_0)\delta] \log_2[p_0+(1-p_0)\delta] - [(1-p_0)(1-\delta)] \log_2[(1-p_0)(1-\delta)]$$

$$H(Y|X) = p_0 H(Y|X=0) + (1-p_0) H(Y|X=1) = (1-p_0)(-\delta \log_2 \delta - (1-\delta)\log_2(1-\delta)) =: (1-p_0)\mathbf{H}(\delta)$$

Since $H(Y|X=0) = -1 \log_2 1 = 0$. So the the mutual information is the following

$$I(X;Y)(p_0) = H(Y) - H(Y|X) = -[p_0+(1-p_0)\delta] \log_2[p_0+(1-p_0)\delta] - [(1-p_0)(1-\delta)] \log_2[(1-p_0)(1-\delta)] -$$

$$-(1-p_0)\mathbf{H}(\delta) = \mathbf{H}\left((1-p_0)(1-\delta)\right) - (1-p_0)\mathbf{H}(\delta)$$

Here $\mathbf{H}(\delta) = -\delta \log_2 \delta - (1-\delta)\log_2(1-\delta.)$

After calculating the derivative and putting it to zero we obtain

$$\frac{dI(X;Y)}{dp_0} = -(1-\delta)\log_2\left[\frac{p_0 + (1-p_0)\delta}{(1-p_0)(1-\delta)}\right] + \mathbf{H}(\delta) = 0 \Leftrightarrow$$

$$\Leftrightarrow p_0 = 1 - \frac{1}{(1-\delta)(1+2^{H(\delta)/(1-\delta)})}$$

Substituting the result into the capacity formula we obtain

$$C(\delta) = \mathbf{H}\left(\frac{1}{1+2^{H(\delta)/(1-\delta)}}\right) - \frac{1}{(1-\delta)(1+2^{H(\delta)/(1-\delta)})}\mathbf{H}(\delta) = \log_2\left[1+2^{-\frac{\mathbf{H}(\delta)}{1-\delta}}\right]$$

2. Firstly, we use l'Hopital's rule

$$\lim_{\delta\to 1}\left(\frac{\frac{1}{1-\delta}}{1+2^{\mathbf{H}(\delta)/(1-\delta)}}\right) = \lim_{\delta\to 1}\frac{\left(\frac{1}{1-\delta}\right)'}{\left(1+2^{\mathbf{H}(\delta)/(1-\delta)}\right)'} = \lim_{\delta\to 1}\frac{1}{\ln 2}\frac{\frac{1}{(1-\delta)^2}}{2^{\mathbf{H}(\delta)/(1-\delta)}\cdot\left(\frac{\mathbf{H}(\delta)}{1-\delta}\right)'} = \tag{1}$$

$$= \frac{1}{\ln 2}\lim_{\delta\to 1}\frac{\frac{1}{(1-\delta)^2}}{2^{\mathbf{H}(\delta)/(1-\delta)}\cdot\left(-\frac{\log_2\delta}{(1-\delta)^2}\right)} = -\frac{1}{\ln 2}\lim_{\delta\to 1}\frac{1}{\log_2\delta\cdot 2^{-\frac{\delta\log_2\delta}{1-\delta}-\log_2(1-\delta)}} = \tag{2}$$

$$= -\frac{1}{\ln 2}\lim_{\delta\to 1}\frac{1}{\frac{\log_2\delta}{1-\delta}\cdot 2^{-\frac{\delta\log_2\delta}{1-\delta}}} \tag{3}$$

So we need to find the following limit

$$\lim_{\delta\to 1}\frac{\log_2\delta}{1-\delta} = -\log_2 e\cdot\lim_{x\to 0}\frac{1+x}{x} = -\log_2 e$$

because it is zamechatelniy limit. Substituting this result to (3) we obtain

$$-\frac{1}{\ln 2}\frac{1}{(-\log_2 e)\cdot e} = \frac{1}{e}$$

So we find that the capacity achieving distribution is the following

$$\mathbb{P}(X=0) = p_0 = 1 - \frac{1}{e}$$

$$\mathbb{P}(X=1) = 1 - p_0 = \frac{1}{e}$$

And the distribution is the following

$$C(1) = \lim_{\delta \to 1} \log_2 \left[1 + 2^{-\frac{\mathbf{H}(\delta)}{1-\delta}}\right] = 0$$

Problem 5

Let p_0 be the probability $\mathbb{P}(X = 0)$. Set

$$q_0 = \mathbb{P}(Y = 0) = (1 - \epsilon)(p_0(1 - p) + p(1 - p_0))$$

$$q_1 = \mathbb{P}(Y = 1) = (1 - \epsilon)(p_0 p + (1 - p_0)(1 - p))$$

$$q_* = \mathbb{P}(Y = \star) = \epsilon$$

It can be easily seen that this is symmetric channel so the capacity of this channel achieves with the uniform distribution of X, i.e. $p_0 = \frac{1}{2}$, hence

$$q_0 = q_1 = \frac{1}{2}(1 - \epsilon)$$

TO find the capacity we need to calculate $H(Y)$ and $H(Y|X)$.

$$H(Y) = -q_0 \log_2 q_0 - q_1 \log_2 q_1 - q_* \log_2 q_* = -(1 - \epsilon) \log_2 \left[\frac{1 - \epsilon}{2}\right] - \epsilon \log_2 \epsilon$$

$$H(Y|X) = -(1 - p)(1 - \epsilon) \log_2 [(1 - p)(1 - \epsilon)] - p(1 - \epsilon) \log_2(p(1 - \epsilon)) - \epsilon \log_2 \epsilon$$

Hence, the capacity is the following

$$C = (1 - \epsilon)(1 - \log_2(1 - \epsilon) + (1 - p) \log_2(1 - p) + (1 - p) \log_2(1 - \epsilon) + p \log_2(p) + p \log_2(1 - \epsilon)) =$$

$$(1 - \epsilon)(1 - H(p))$$

Problem 6

We have the optimization problem

$$\text{maximize} \quad \sum_{i=1}^{499} \frac{1}{2}\log_2\left(1+\frac{P_i}{100}\right)+\frac{1}{2}\log_2\left(1+\frac{P_{500}}{25}\right) \tag{4}$$

$$\text{subject to} \quad \sum_i P_i = 50 \tag{5}$$

We know that the solution if this problem has the following form

$$P_i = (\nu - N_i)^+$$

where ν is a constant that satisfies

$$\sum_{i=1}^{500}(\nu - N_i)^+ = 50 \tag{6}$$

In this case it is obvious that if $\nu \geq 100$ equality (6) doesn't hold. If $\nu < 100$ then the expression (6) has the form

$$(\nu - 25)^+ = 50$$

So $\nu = 75$ and we have the following solution

$$P_1 = \ldots = P_{499} = 0$$

$$P_{500} = 50$$

Hence, the capacity is the following

$$C = \frac{1}{2}\log_2\left(1+\frac{50}{25}\right) = \frac{1}{2}\log_2 3 \approx 0.79\text{bits}$$

Problem 7

Our capacity region is determined by the following inequalities

$$R_1 \leq I(X_1; Y|X_2) \tag{7}$$

$$R_2 \leq I(X_2; Y|X_1) \tag{8}$$

$$R_1 + R_2 \leq I(X_1, X_2; Y) \tag{9}$$

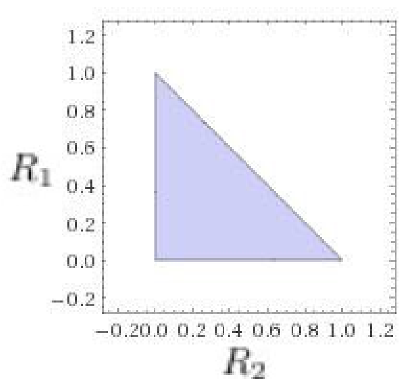
It can be easily seen that

$$I(X_1; Y|X_2) = H(X_1|X_2) - H(X_1|Y, X_2) \leq H(X_1) \leq 1$$

and, obviously, the equality holds when $\mathbb{P}(X_1 = 0) = \mathbb{P}(X_2 = 0) = \frac{1}{2}$. The same is true for (8). Also we can find that

$$R_1 + R_2 \leq I(X_1, X_2; Y) \leq H(Y) = 1$$

So we have the following plot



Problem 8

We have the following channel

X_1	X_2	Y
-1	-1	-2
1	-1	0
-1	1	0
1	1	2

It is obvious that we can consider symmetric case

$$\mathbb{P}(X_1 = 1) = \mathbb{P}(X_2 = 1) = p_1$$

Firstly, let us find $I(X_1; Y|X_2)$.

$$I(X_1; Y|X_2) \leq 1 \tag{10}$$

$$I(X_2; Y|X_1) \leq 1 \tag{11}$$

As in the previous exercise, and the equalities holds when $p_1 = \frac{1}{2}$. Now let us find $I(X_1, Y)$

$$I(X_1, Y) = H(X) - H(X|Y) = p_0 \log_2 \frac{1}{p_0} + (1 - p_0) \log_2 \left(\frac{1}{1 - p_0} \right) - \mathbb{P}(Y = -2)H(X_1|Y = -2) - \tag{12}$$

$$\mathbb{P}(Y = 2)H(X_1|Y = 2) - \mathbb{P}(Y = 0)H(X_1|Y = 0) = \tag{13}$$

$$p_0 \log_2 \frac{1}{p_0} + (1 - p_0) \log_2 \left(\frac{1}{1 - p_0} \right) - 4p_0(1 - p_0) \left(p_0(1 - p_0) \log_2 \frac{1}{p_0(1 - p_0)} \right) \tag{14}$$

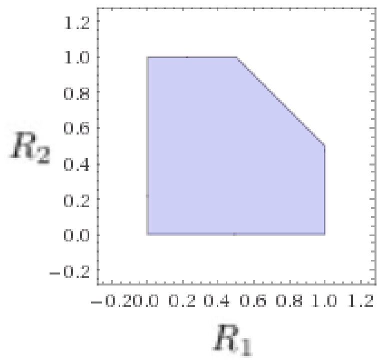
It can be easily calculated that

$$\frac{dI(X_1; Y)}{dp_0} = 0 \Leftrightarrow p_0 = \frac{1}{2}$$

And for $p_0 = \frac{1}{2}$

$$I(X_1; Y) \left[\frac{1}{2} \right] = \frac{1}{2}$$

We have the same for $I(X_2; Y)$. Finally, the $\max I(X_1, X_2, Y) = \frac{3}{2}$ and we have the following picture


R_2
R_1
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.20.0 0.2 0.4 0.6 0.8 1.0 1.2