Q 7.1

a) $X \to Y \to \hat{Y}$

and using $\hat{Y}$ to recover the message results in a capacity

$$\hat{C} = \max_{p(x)} I(X;\hat{Y})$$

but $I(X;\hat{Y}) \le I(X;Y)$ by Data processing inequality.

$$\Rightarrow \hat{C} = \max_{p(x)} I(X;\hat{Y}) \le \max_{p(x)} I(X;Y) = C$$

b) We have equality iff $X \to \hat{Y} \to Y$ is a Markov chain.

7.2 Let $p_X(0) = \alpha$, $p_X(1) = 1-\alpha$

Let $a = 1$

$$C = \max_{p_X} I(X;Y) = \max_{\alpha} H(Y) - H(Y|X)$$

$$= \max_{\alpha} H(X+Z) - H(X+Z|X)$$

$$= \max_{\alpha} H(X+Z) - H(Z|X)$$

$$= \max_{\alpha} H(X+Z) - H(Z)$$

$$= \max_{\alpha} H(X+Z) - 1$$

Now, $p_Y(0) = \frac{1}{2}\alpha$

$p_Y(1) = \frac{1}{2}\alpha + \frac{1}{2}(1-\alpha) = \frac{1}{2}$

$p_Y(2) = \frac{1}{2}(1-\alpha)$

$$\Rightarrow H(Y) = -\frac{\alpha}{2}\log\frac{\alpha}{2} - \frac{1}{2}\log\frac{1}{2} - \frac{1-\alpha}{2}\log\frac{1-\alpha}{2}$$

$$= -\frac{\alpha}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} - \frac{1-\alpha}{2}\log\frac{1}{2}$$

$$+ \frac{1}{2}\left(-\alpha\log\alpha - (1-\alpha)\log(1-\alpha)\right)$$

$$= 1 + \frac{1}{2}h(\alpha)$$

$$\Rightarrow C = \max_{\alpha}\left(1 + \frac{1}{2}h(\alpha) - 1\right) = \frac{1}{2}$$

when $\alpha = \frac{1}{2}$.

- The calculation is essentially identical if $a=-1$, except that

$$P_Y(-1) = \tfrac{1}{2}\alpha$$

$$P_Y(0) = \tfrac{1}{2}$$

$$P_Y(1) = \tfrac{1}{2}(1-\alpha).$$

- If $a \neq \pm 1$, then

$$X=0,\quad Z=0 \Rightarrow Y=0$$

$$X=0,\quad Z=a \Rightarrow Y=a$$

$$X=1,\quad Z=0 \Rightarrow Y=1$$

$$X=1,\quad Z=a \Rightarrow Y=1+a$$

results in 4 different values of Y, and thus X can be recovered without error from Y. Then $X \to Y \to X$, and thus

$$I(X;Y) \geq I(X;X) = 1$$

which can be achieved.

7.3 $Y_\ell = X_\ell \oplus Z_\ell$, $\ell = 1, \dots, n$

If noise Z_ℓ is iid, then capacity is $1 - h(p) = C$ which is achieved with $X_1, \dots, X_n$ iid $\sim p_X$,

i.e., $p(x_1, \dots, x_n) = \prod_{\ell=1}^{n} p_X(x_\ell)$

Now

$$
\begin{aligned}
&\max_{p(x_1,\dots x_n)} I(X_1, \dots, X_n; Y_1, \dots, Y_n) \\
&\geq I(X_1, \dots, X_n; Y_1, \dots, Y_n) \quad \text{for } X_1^n \sim \text{iid } p_X \\
&= H(X_1, \dots, X_n) - H(X_1, \dots, X_n | Y_1, \dots, Y_n) \quad \text{for } X_1^n \text{ iid} \sim p_X \\
&= \sum_{\ell=1}^{n} H(X_\ell) - H(X_\ell | Y_1^n, X_1^{\ell-1}) \quad \text{for } X_1^n \text{ iid} \sim p_X \\
&\overset{(a)}{\geq} \sum_{\ell=1}^{n} H(X_\ell) - H(X_\ell | Y_\ell) \quad \text{for } X_1^n \text{ iid} \sim p_X \\
&= \sum_{\ell=1}^{n} I(X_\ell; Y_\ell) = nC
\end{aligned}
$$

where (a) follows since conditioning reduces entropy.

Note that if $Z_1, \dots, Z_n$ was memoryless then (a) would hold with equality.

7.4

a) $C = \max_{p_X} I(X;Y)$

$$= \max_{p_X} H(Y) - H(Y|X)$$

$$= \max_{p_X} H(X \oplus Z) - H(X \oplus Z|X)$$

$$= \max_{p_X} H(X \oplus Z) - H(Z|X)$$

$$= \max_{p_X} H(X \oplus Z) - H(Z)$$

$$\leq \log_2 11 - \log_2 3$$

with equality when p_X is uniform on $0, 1, \ldots, 10$ since then $Y = X \oplus Z$ is also.

7.5 Let $\bar{X} = (X_1, X_2)$
$\bar{Y} = (Y_1, Y_2)$

Then
$$C = \max_{p_{\bar{X}}} I(\bar{X}; \bar{Y})$$
$$= \max_{p_{X_1, X_2}} I(X_1 X_2; Y_1 Y_2)$$
$$= \max_{p_{X_1, X_2}} H(Y_1 Y_2) - H(Y_1 Y_2 | X_1 X_2)$$
$$\overset{(a)}{\leq} \max_{p_{X_1 X_2}} H(Y_1) + H(Y_2) - H(Y_1 Y_2 | X_1 X_2)$$

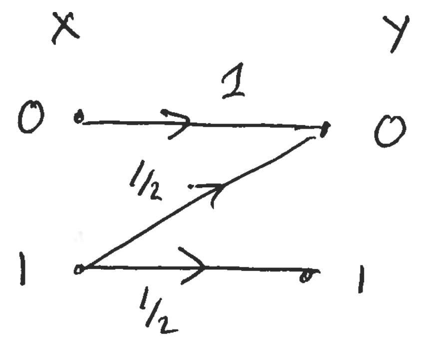
But
$$H(Y_1 Y_2 | X_1 X_2) = H(Y_1 | X_1 X_2) + H(Y_2 | X_1 X_2 Y_1)$$
$$= H(Y_1 | X_1) + H(Y_2 | X_2)$$

Since $P_{Y_1 Y_2 | X_1 X_2}(y_1 y_2 | x_1 x_2) = P_{Y_1|X_1}(y_1|x_1) P_{Y_2|X_2}(y_2|x_2)$

$$\Rightarrow C \leq \max_{p_{X_1 X_2}} H(Y_1) + H(Y_2) - H(Y_1|X_1) - H(Y_2|X_2)$$
$$= \max_{p_{X_1 X_2}} I(X_1; Y_1) + I(X_2; Y_2)$$
$$= \max_{p_{X_1}} I(X_1; Y_1) + \max_{p_{X_2}} I(X_2; Y_2)$$
$$= C_1 + C_2$$

and we have equality in (a) when X_1 & X_2 are independent

7.8



$$C = \max_{p_X} H(Y) - H(Y|X)$$

Let $p_X(0) = \alpha \qquad p_X(1) = 1-\alpha$

$$H(Y|X) = \sum_{x \in \{0,1\}} H(Y|X=x)\, p_X(x)$$

$$= H(Y|X=0)\, p_X(0) + H(Y|X=1)\, p_X(1)$$

$$= 0 \cdot \alpha + 1 \cdot (1-\alpha)$$

$$p_Y(0) = 1 \cdot \alpha + \tfrac{1}{2}(1-\alpha) = \tfrac{1}{2} + \tfrac{1}{2}\alpha$$

$$p_Y(1) = \tfrac{1}{2}(1-\alpha) = \tfrac{1}{2} - \tfrac{1}{2}\alpha$$

$$H(Y) = -\tfrac{1}{2}(1+\alpha)\log\frac{1+\alpha}{2} - \tfrac{1}{2}(1-\alpha)\log\frac{1-\alpha}{2}$$

$$= -\tfrac{1}{2}(1+\alpha)\log \tfrac{1}{2} - \tfrac{1}{2}(1-\alpha)\log\tfrac{1}{2}$$

$$+ \tfrac{1}{2}\left[(1+\alpha)\log(1+\alpha) - (1-\alpha)\log(1-\alpha)\right]$$

$$= 1 - \tfrac{1}{2}\left[(1+\alpha)\log(1+\alpha) + (1-\alpha)\log(1-\alpha)\right]$$

$$I(X;Y) = 1 - \frac{1}{2}\left[(1+\alpha)\log(1+\alpha) + (1-\alpha)\log(1-\alpha)\right] - (1-\alpha)$$

$$= \alpha - \frac{1}{2}\left[(1+\alpha)\log(1+\alpha) + (1-\alpha)\log(1-\alpha)\right]$$

$$\frac{d}{d\alpha} I(X;Y) = 1 - \frac{1}{2\ln 2}\left[\frac{1+\alpha}{1+\alpha} + \ln(1+\alpha) + \frac{1-\alpha}{1-\alpha}\cdot -1 - \ln(1-\alpha)\right]$$

$$= 1 - \frac{1}{2\ln 2}\left[\ln \frac{1+\alpha}{1-\alpha}\right]$$

$$= 1 - \frac{1}{2}\log_2\left(\frac{1+\alpha}{1-\alpha}\right) = 0$$

$$\log_2\left(\frac{1+\alpha}{1-\alpha}\right) = 2$$

$$\frac{1+\alpha}{1-\alpha} = 4$$

$$1+\alpha = 4 - 4\alpha$$

$$5\alpha = 3$$

$$\alpha = 3/5$$

$$DC = \frac{3}{5} - \frac{4}{5}\log_2\frac{8}{5} - \frac{1}{5}\log_2\frac{2}{5}$$

7.11 $I(X_1^n; Y_1^n)$

$$= H(Y_1^n) - H(Y_1^n | X_1^n)$$

$$= H(Y_1^n) - \sum_{l=1}^{n} H(Y_l | X_l)$$

$$\overset{(a)}{\leq} \sum_{l=1}^{n} H(Y_l) - H(Y_l | X_l)$$

$$= \sum_{l=1}^{n} I(X_l ; Y_l) \qquad (*)$$

where we have equality in (a) when $X_1, \ldots, X_n$ are independent.

Now, maximizing (*) yields

$$\max_{p(x_1^n)} \sum_{l=1}^{n} I(X_l ; Y_l) = \sum_{l=1}^{n} 1 - h(p_l)$$

where the maximizing input distribution is $p(x_1^n) = \prod_{l=1}^{n} p_X(x_l)$ and p_X is uniform on $\{0,1\}$.

But since $X_1, \ldots, X_n$ are indendent, we have equality in (a)

Thus, the upper bound $\sum_{l=1}^{n} [1 - h(p_l)]$ is achievable.

7.20 Note that $Y_1 \to X \to Y_2$ $(*)$

a) $I(X;Y_1,Y_2) = I(X;Y_1) + I(X;Y_2|Y_1)$

also

$I(Y_1;X) = I(Y_2;X)$ since $P_{Y_1|X} = P_{Y_2|X}$

$= I(Y_2;\cancel{X}\,Y_1)$ by $(*)$

$= I(Y_1;Y_2) + I(X;Y_2|Y_1)$

$\Rightarrow I(X;Y_2|Y_1) = I(X;Y_1) - I(Y_1;Y_2)$

$\Rightarrow I(X;Y_1,Y_2) = 2I(X;Y_1) - I(Y_1;Y_2)$

b) $C = \max\limits_{P_X} I(X;Y_1,Y_2)$

$= \max\limits_{P_X} 2I(X;Y_1) - I(Y_1;Y_2)$

$\leq 2 \max\limits_{P_X} I(X;Y_1)$

7.25

$$C = \max_{p_x} I(X;Y)$$

and $I(X;Y) \le I(X;V)$ by data processing inequality

$$\le H(V)$$

$$\le \log k$$

$$\Rightarrow C \le \max_{p_x} \log k = \log k$$

7.31 We need $H(V) < C$

Here $H(V) = h(\alpha) = -\alpha \log \alpha - (1-\alpha)\log(1-\alpha)$

$$C = \max_{p_x} I(X;Y)$$

$$= \max_{p_x} H(Y) - H(Y|X)$$

$$= \max_{p_x} H(Y) - h(p)$$

$$= 1 - h(p)$$

$\Rightarrow h(\alpha) \le 1 - h(p)$ or $h(\alpha) + h(p) < 1$.

7.34

a) $C = \max_{p_X} I(X;Y)$

$$I(X;Y) = H(Y) - H(Y|X)$$

$$H(Y|X) = \sum_{x=1}^{4} H(Y|X=x)\, p_X(x)$$

$$= \sum_{x=1}^{4} h(p)\, p_X(x)$$

$$= h(p)$$

$$\Rightarrow C = \max_{p_X} H(Y) - h(p)$$

$$\leq \log_2 4 - h(p) = 2 - h(p)$$

and we see that when p_X is uniform on $\{1,2,3,4\}$ that Y is uniform on $\{1,2,3,4\}$ so $H(Y) = 2$, and therefore $C = 2 - h(p)$.

7.34 b) $C = \max_{p_X} I(X;Y)$

$$I(X;Y) = H(Y) - H(Y|X)$$

$$H(Y|X) = \sum_{x=1}^{3} H(Y|X=x)\, p_X(x)$$

$$= h(p)\cdot p_X(1) + h(p)\cdot p_X(2) + 0\cdot p_X(3)$$

$$= h(p)\cdot[1 - p_X(3)]$$

where we used the fact that $p_X(1) + p_X(2) + p_X(3) = 1$

$$\Rightarrow p_X(2) + p_X(1) = 1 - p_X(3)$$

$$\Rightarrow I(X;Y) = H(Y) - h(p)[1 - p_X(3)]$$

Now, let $p_Y = (p_Y(1)\ \ p_Y(2)\ \ p_Y(3))$

be the distribution on Y when $p_X = (p_X(1)\ \ p_X(2)\ \ p_X(3))$

Then, when distribution on X is $\bar{p}_X = (p_X(2)\ \ p_X(1)\ \ p_X(3))$

distribution on Y is $\bar{p}_Y = (p_Y(2)\ \ p_Y(1)\ \ p_Y(3))$

$\Rightarrow \quad H(p_Y) = H(\bar{p}_Y)$

Also, if $\hat{p}_X = \left(\frac{p_Y(1)+p_Y(2)}{2} \quad \frac{p_Y(1)+p_Y(2)}{2} \quad p_Y(3) \right)$

$$= \frac{p_X + \bar{p}_X}{2}$$

then $\hat{p}_Y = \frac{1}{2}(p_Y + \bar{p}_Y) = \left(\frac{1-p_Y(3)}{2} \quad \frac{1-p_Y(3)}{2} \quad p_Y(3) \right)$

But, by convexity of entropy,

$$H\left(\tfrac{1}{2} p_Y + \tfrac{1}{2} \bar{p}_Y\right) \geq \tfrac{1}{2} H(p_Y) + \tfrac{1}{2} H(\bar{p}_Y)$$

$$= H(p_Y)$$

and since $p_Y(3) = p_X(3)$,

$$H(Y) \geq -\frac{1-p_X(3)}{2} \log \frac{1-p_X(3)}{2} - \frac{1-p_X(3)}{2} \log \frac{1-p_X(3)}{2}$$

$$- p_X(3) \log p_X(3)$$

with equality when $p_X(1) = p_X(2) = \frac{1-p_X(3)}{2}$.