Name ________________________ ID3 _ _ _

ECS 452: Digital Communication Systems	**2020/2**

Additional Examples 1

Lecturer: Prapun Suksompong, Ph.D.

Source Coding
CH2

Problem 1. These codes cannot be Huffman codes. Why?

(a) {00, 01, 10, 110} $p_1\ p_2\ p_3\ p_4$

(b) {01, 10}

(c) {0, 01}

(a) The code {00,01,10,110} can be shortened to {00,01,10,11}. The prefix-free property is preserved; so the shortened code is still UD. Regardless of the probabilty associated with each source symbol (as long as the last symbol has non-zero probability), the shortened code will have smaller expected length. Because Huffman code is optimal, the original code can't be Huffman.

(b) The code {01,10} can be shortened to {0,1}. The prefix-free property is preserved; so the shortened code is still UD. Regardless of the probabilty associated with each source symbol, the shortened code will have smaller expected length. Because Huffman code is optimal, the original code can't be Huffman.

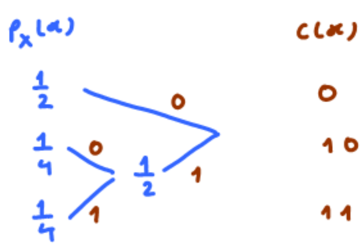
(c) The code {0,01} is not prefix-free. Any Huffman code must be prefix-free. Therefore, it cannot be a Huffman code.

Alternatively, one can also use the same reasoning in part (a) and (b): The code {0,01) can be shortened to {0,1}. The new code is prefix-free and hence still UD. Regardless of the probabilty associated with each source symbol, the shortened code will have smaller expected length. Because Huffman code is optimal, the original code can't be Huffman.

Problem 2. Construct a random variable X (by specifying its pmf) whose corresponding Huffman code is $\{0, 10, 11\}$.

{0,10,11} is a Huffman code for

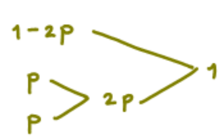
$$p_X(x) = \begin{cases} 1/2, & x = 1 \\ 1/4, & x = 2,3 \\ 0, & \text{otherwise.} \end{cases}$$



Note that the answer for this question is not unique.
You may check that

$$p_X(x) = \begin{cases} 1-2p, & x = 1 \\ p, & x = 2,3 \\ 0, & \text{otherwise} \end{cases}$$

works when



$1-2p \geq p$

$\Updownarrow$

$p \leq \frac{1}{3}$

This condition is here to guarantee that x = 2 and x = 3 are grouped frist.

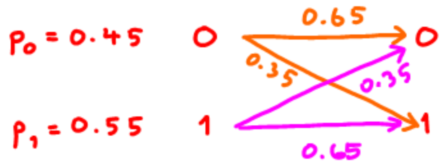
Note also that when $p = \frac{1}{3}$, we have uniform pmf.

Problem 3. Consider a BSC whose crossover probability for each bit is $p = 0.35$. Suppose $P[X = 0] = 0.45$. $\equiv p_0$

$1-p = 0.65$

$p_1 = 1 - p_0 = 1 - 0.45 = 0.55$

(a) Draw the channel diagram.



Method 2: Directly read the transition (conditional) probabilities from the channel diagram.

(b) Find the channel matrix **Q**.

Method 1

$$Q = \begin{array}{c|cc} x\backslash y & 0 & 1 \\ \hline 0 & 1-p & p \\ 1 & p & 1-p \end{array} = \begin{array}{c|cc} x\backslash y & 0 & 1 \\ \hline 0 & 0.65 & 0.35 \\ 1 & 0.35 & 0.65 \end{array}$$

(c) Find the joint pmf matrix **P**.

the row vector $\underline{p}$ of input probabilities: $\underline{p} = \begin{bmatrix} P[X=0] & P[X=1] \end{bmatrix} = [p_0 \quad p_1] = \begin{bmatrix} 0.45 & 0.55 \end{bmatrix}$

To get the matrix P, we simply scale each row of matrix Q by the corresponding p(x)

$$Q = \begin{array}{c|cc} x\backslash y & 0 & 1 \\ \hline 0 & 0.65 & 0.35 \\ 1 & 0.35 & 0.65 \end{array} \begin{array}{c} \xrightarrow{\times 0.45} \\ \xrightarrow[\times 0.55]{} \end{array} \begin{array}{c|cc} x\backslash y & 0 & 1 \\ \hline 0 & 0.2925 & 0.1575 \\ 1 & 0.1925 & 0.3575 \end{array} = P$$

(d) Find the row vector $\underline{q}$ which contains the pmf of the channel output Y.

$\downarrow \Sigma \qquad \downarrow \Sigma$

Method 1: $\underline{q} = \begin{bmatrix} 0.485 & 0.515 \end{bmatrix}$

Method 2: $\underline{q} = \underline{p}Q = \begin{bmatrix} 0.45 & 0.55 \end{bmatrix} \begin{bmatrix} 0.65 & 0.35 \\ 0.35 & 0.65 \end{bmatrix} = \begin{bmatrix} 0.485 & 0.515 \end{bmatrix}$

(e) Analyze the performance of all four reasonable detectors for this binary channel. Complete the table below:

$\hat{x}(y)$	$\hat{x}(0)$	$\hat{x}(1)$	$P(\mathcal{C})$	$P(\mathcal{E}) = 1 - P(\mathcal{C})$
y	0	1	$0.65(p_0+p_1) = 0.65$	0.35
$1-y$	1	0	$0.35(p_0+p_1) = 0.35$	0.65
1	1	1	$(0.35+0.65)p_1 = p_1 = 0.55$	0.45
0	0	0	$(0.65+0.35)p_0 = p_0 = 0.45$	0.55

$$Q = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

$$= \begin{array}{c|cc} x\backslash y & 0 & 1 \\ \hline 0 & 0.65 & 0.35 \\ 1 & 0.35 & 0.65 \end{array} \begin{array}{c} \xrightarrow{\times p_0} \\ \xrightarrow[\times p_1]{} \end{array} \begin{array}{c|cc} x\backslash y & 0 & 1 \\ \hline 0 & \boxed{0.65 p_0} & \boxed{0.35 p_0} \\ 1 & 0.35 p_1 & 0.65 p_1 \end{array} = P$$

The value of $\hat{x}(y)$ tell us which row to select in the column y of the P matrix.

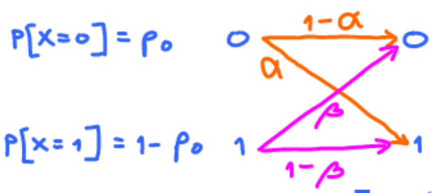
$Q(0|0) = 0.65$, $Q(1|1) = 0.45$

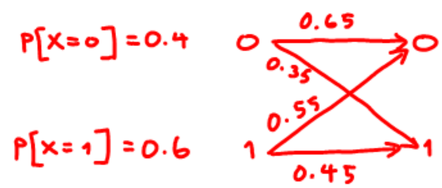
Problem 4. Consider a BAC whose $Q(1|0) = 0.35$ and $Q(0|1) = 0.55$. Suppose $P[X = 0] = 0.4$. $= p_0 \Rightarrow p_1 = 1 - 0.4 = 0.6$

In some of these parts, we will also try to derive the answer in a general form.

(a) Draw the channel diagram.

So, we will start with $Q(1|0) = \alpha$ and $Q(0|1) = \beta$.





To get the matrix P, we simply scale each row of matrix Q by the corresponding p(x).

(b) Find the joint pmf matrix $\mathbf{P}$.

First, we find the Q matrix,

$$Q = \begin{array}{c|cc} x \backslash y & 0 & 1 \\ \hline 0 & 1-\alpha & \alpha \\ 1 & \beta & 1-\beta \end{array} \quad \begin{array}{c} \xrightarrow{\times p_0} \\ \xrightarrow{\times 1-p_0} \end{array} \quad \begin{array}{c|cc} x \backslash y & 0 & 1 \\ \hline 0 & p_0(1-\alpha) & p_0\alpha \\ 1 & (1-p_0)\beta & (1-\beta)(1-p_0) \end{array} = P$$

For this question,

$$Q = \begin{array}{c|cc} x \backslash y & 0 & 1 \\ \hline 0 & 0.65 & 0.35 \\ 1 & 0.55 & 0.45 \end{array} \quad \begin{array}{c} \xrightarrow{\times 0.4} \\ \xrightarrow{\times 0.6} \end{array} \quad \begin{array}{c|cc} x \backslash y & 0 & 1 \\ \hline 0 & 0.26 & 0.14 \\ 1 & 0.33 & 0.27 \end{array} = P$$

(c) Find the row vector $\underline{\mathbf{q}}$ which contains the pmf of the channel output Y.

Method 2: $\downarrow \Sigma$ $\downarrow \Sigma$ $[0.59 \quad 0.41]$

Method 1: Recall that $\underline{q} = \underline{p}Q$.

For this question, $\underline{p} = [p_0 \quad 1-p_0] = [0.4 \quad 0.6]$.

Therefore, $\underline{q} = [0.4 \quad 0.6]\begin{bmatrix} 0.26 & 0.14 \\ 0.33 & 0.27 \end{bmatrix} = [0.59 \quad 0.41]$

(d) Analyze the performance of all four reasonable detectors for this binary channel. Complete the table below:

$\hat{x}(y)$	$\hat{x}(0)$	$\hat{x}(1)$	$P(\mathcal{C})$	$P(\mathcal{E})$ = $1 - P(\mathcal{C})$
y	0	1	0.26+0.27 = 0.53	0.47
$1-y$	1	0	0.33+0.14 = 0.47	0.53
1	1	1	0.33+0.27 = 0.60	0.40
0	0	0	0.26+0.14 = 0.40	0.60

The value of $\hat{x}(y)$ tells us which row to select in the column y of the P matrix.

$$Q = \begin{array}{c|cc} x \backslash y & 0 & 1 \\ \hline 0 & 0.65 & 0.35 \\ 1 & 0.55 & 0.45 \end{array} \quad \begin{array}{c} \xrightarrow{\times 0.4} \\ \xrightarrow{\times 0.6} \end{array} \quad \begin{array}{c|cc} x \backslash y & 0 & 1 \\ \hline 0 & 0.26 & 0.14 \\ 1 & 0.33 & 0.27 \end{array}$$

The convention for our class is that these numbers are ordered in the same way that they are specified in the support.

Problem 5. Consider a DMC whose $\mathcal{X} = \{1, 2, 3\}$, $\mathcal{Y} = \{1, 2, 3\}$, and
$$\mathbf{Q} = \begin{array}{c} x \backslash y \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.2 & 0.6 \end{bmatrix} \end{array}.$$

Suppose the input probability vector is
$$\begin{array}{c} \quad\ \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \underline{\mathbf{p}} = [0.2, 0.4, 0.4]. \end{array}$$

(a) Find the joint pmf matrix **P**.

We can get the P matrix by scaling each row of the Q matrix using the corresponding input probability p(x).

$$Q = \begin{array}{c} x \backslash y \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.2 & 0.6 \end{bmatrix} \end{array} \begin{array}{c} \\ \xrightarrow{\times 0.2} \\ \xrightarrow{\times 0.4} \\ \xrightarrow{\times 0.4} \end{array} \begin{array}{c} x \backslash y \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0.10 & 0.04 & 0.06 \\ 0.12 & 0.16 & 0.12 \\ 0.08 & 0.08 & 0.24 \end{bmatrix} \end{array} = P$$

(b) Find the row vector $\underline{\mathbf{q}}$ which contains the pmf of the channel output Y.

Method 1: $\Sigma \downarrow$ each column: $[0.3 \quad 0.28 \quad 0.42]$

Method 2:
$$q = \underline{p}P = [0.2 \quad 0.4 \quad 0.4] \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.2 & 0.6 \end{bmatrix} = \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ [0.3 \quad 0.28 \quad 0.42] \end{array}$$

(c) Find the following probabilities:

(i) $P[X = 1] = 0.2$

(ii) $P[Y = 2] = 0.28$

(iii) $P[X = 1, Y = 2] = 0.04$

(iv) $P[Y = 2|X = 1] = 0.2$

(v) $P[X = 1|Y = 2] \overset{P(A|B)}{=} \dfrac{P(A \cap B)}{P(B)} = \dfrac{P[X=1, Y=2]}{P[Y=2]} = \dfrac{0.04}{0.28} = \dfrac{4}{28} = \dfrac{1}{7}$

(vi) Find the error probability of the naive decoder. Naive Decoder : $\hat{x}_{Naive}(y) = y$

$$Q = \begin{array}{c|ccc} x \backslash y & 1 & 2 & 3 \\ \hline 1 & 0.5 & 0.2 & 0.3 \\ 2 & 0.3 & 0.4 & 0.3 \\ 3 & 0.2 & 0.2 & 0.6 \end{array} \quad \begin{array}{l} \xrightarrow{\times 0.2} \\ \xrightarrow{\times 0.4} \\ \xrightarrow{\times 0.4} \end{array} \quad \begin{array}{c|ccc} x \backslash y & 1 & 2 & 3 \\ \hline 1 & 0.10 & 0.04 & 0.06 \\ 2 & 0.12 & 0.16 & 0.12 \\ 3 & 0.08 & 0.08 & 0.24 \end{array} = P$$

For naive decoder, look at each column of P and select (circle) the element whose corresponding x value is the same as y in that column.

$P(\mathcal{C}) = 0.10 + 0.16 + 0.24 = 0.5$

$P(\mathcal{E}) = 1 - P(\mathcal{C}) = 1 - 0.5 = 0.5$

(vii) Find the error probability of the (DIY) decoder $\hat{x}(y) = 4 - y$. $\Rightarrow$ Decoder table :

y	$\hat{x}(y)$
1	3
2	2
3	1

$$Q = \begin{array}{c|ccc} x \backslash y & 1 & 2 & 3 \\ \hline 1 & 0.5 & 0.2 & 0.3 \\ 2 & 0.3 & 0.4 & 0.3 \\ 3 & 0.2 & 0.2 & 0.6 \end{array} \quad \begin{array}{l} \xrightarrow{\times 0.2} \\ \xrightarrow{\times 0.4} \\ \xrightarrow{\times 0.4} \end{array} \quad \begin{array}{c|ccc} x \backslash y & 1 & 2 & 3 \\ \hline 1 & 0.10 & 0.04 & 0.06 \\ 2 & 0.12 & 0.16 & 0.12 \\ 3 & 0.08 & 0.08 & 0.24 \end{array} = P$$

For DIY decoder, look at each column of P and select the element whose corresponding x value is the same as $\hat{x}(y)$ in the decoder table.

$P(\mathcal{C}) = 0.08 + 0.16 + 0.06 = 0.30$

$P(\mathcal{E}) = 1 - 0.30 = 0.70$

Problem 6. ***Optimal code lengths that require one bit above entropy***: The source coding theorem says that the Huffman code for a random variable X has an expected length strictly less than $H(X) + 1$. Give an example of a random variable for which the expected length of the Huffman code (without any source extension) is very close to $H(X) + 1$.

We want to come up with some simple example. Therefore, we shall start by considering the Bernoulli RV X which has only two possible values:

Consider $X \sim \text{Bernoulli}(p_1)$: $p(x) = \begin{cases} p_1, & x = 1, \\ 1 - p_1, & x = 0, \\ 0, & \text{otherwise.} \end{cases}$

Because there are only two possible values, the pairing in the Huffman coding process must be between these two values:

x	$p(x)$	$c(x)$	$\ell(x)$
0	$1-p_1$	0	1
1	p_1	1	1

Therefore, Huffman coding (without extension) always needs 1 bit per symbol.

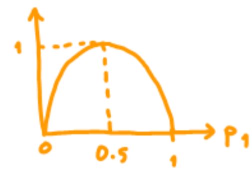
Now, we want the $\underbrace{\text{expected length}}_{=1}$ to be $\approx H(X) + 1$

So, we must have $H(X) \approx 0$.

By definition, $H(X) = -\sum_x p(x) \log_2 p(x) = -(1-p_1)\log(1-p_1) - p_1 \log p_1$

You may recall that in Section 2.4, this expression is called the binary entropy function. The plot of this function is shown in the lecture notes; it is sketches below:

Note that to get this function to be ≈ 0, we need to consider $p_1 \approx 0$ or $p_1 \approx 1$.



Also note that we don't want to have $p_1 = 0$ or $p_1 = 1$ because they would make our RV X degenerated (deterministic). For degenerated RV, we don't have to waste any bit to convey its value. (The value is already pre-determined.) For example, if $p_1 = 1$, we know that P[X = 1] = 1 and hence X = 1 all the time. There is no uncertainty. Anyone can guess the value of X with 100% accuracy by simply guessing the value 1 every time. Alternatively, we may think of the situation here as sending the empty string (ε) to the receiver.

So, $\mathbb{E}[\ell(X)] = \mathbb{E}[0] = 0$.

Because $H(X)$ is also 0, we have $\mathbb{E}[\ell(X)] = H(X)$ and not $H(X)+1$.

More formally we can take $\lim_{p_1 \to 0}$ or $\lim_{p_1 \to 1}$ on the function and show that $H(X) \to 0$.

$$\left(\lim_{x \to 0^+} x \ln x = \lim_{x \to 0^+} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x \to 0^+} (-x) = 0 \right)$$

L'Hôpital's rule

So, if p_1 is close to 1 or 0, the entropy will be almost 0.

But we still need $\mathbb{E}[\ell(X)] = 1$ bit to send X.

So, the expected length will be very close to $H(X) + 1$.

Problem 7. A channel encoder maps blocks of two bits to five-bit (channel) codewords. The four possible codewords are 00000, 01000, 10001, and 11111. A codeword is transmitted over the BSC with crossover probability $p = 0.1$. What is the minimum (Hamming) distance d_{min} among the codewords?

$\underline{x}$

$d_{min} = 1$

$\underline{x}$	00000	01000	10001	11111
00000		1	2	5
01000			3	4
10001				3
11111				

The given expression for the joint pmf can be expressed using the joint pmf matrix as

$$P = \begin{array}{c|cc} x \backslash y & 2 & 4 \\ \hline 1 & 3c & 5c \\ 3 & 5c & 7c \end{array}$$

Problem 8. Consider random variables X and Y whose joint pmf is given by

$$p_{X,Y}(x,y) = \begin{cases} c(x+y), & x \in \{1,3\} \text{ and } y \in \{2,4\}, \\ 0, & \text{otherwise.} \end{cases}$$

Evaluate the following quantities.

(a) $c = 1/20$

The sum of all elements in the P matrix should be 1:

$$\left.\begin{aligned} &3c + 5c + 5c + 7c = 1 \\ &\Rightarrow 20c = 1 \\ &\Rightarrow c = \frac{1}{20}. \end{aligned}\right\}$$

(b) $H(X,Y) = H(\underbrace{[\tfrac{3}{20}\ \tfrac{1}{4}\ \tfrac{1}{4}\ \tfrac{7}{20}]}_{\text{all elements in the P matrix}}) = -\frac{3}{20}\log_2\frac{3}{20} - \frac{2}{4}\log_2\frac{1}{4} - \frac{7}{20}\log_2\frac{7}{20}$

≈ 1.9406 bits.

$$\begin{array}{c|cc|l} x \backslash y & 2 & 4 & p(x) \\ \hline 1 & 3c & 5c & \xrightarrow{\Sigma} 8c = \frac{8}{20} = \frac{2}{5} \\ 3 & 5c & 7c & \xrightarrow{\Sigma} 12c = \frac{12}{20} = \frac{3}{5} \\ \hline & \Sigma\downarrow & \Sigma\downarrow & \\ & 8c & 12c & \\ & \| & \| & \\ q(y) & 2/5 & 3/5 & \end{array}$$

(c) $H(X) = H([\tfrac{2}{5}\ \tfrac{3}{5}]) = -\frac{2}{5}\log_2\frac{2}{5} - \frac{3}{5}\log_2\frac{3}{5}$

≈ 0.9710

same

(d) $H(Y) = H([\tfrac{2}{5}\ \tfrac{3}{5}]) \approx 0.9710$

(e) $H(X|Y) = H(X,Y) - H(Y) \approx 0.9697$

(f) $H(Y|X) = H(X,Y) - H(X) \approx 0.9697$

(g) $I(X;Y) = H(X) + H(Y) - H(X,Y) \approx 0.0013$