Solutions to Practice Final

1. **Huffman code**
   Give a Huffman encoding into an alphabet of size $D = 4$ of the following probability mass function:

$$\mathbf{p} = (\frac{8}{36}, \frac{7}{36}, \frac{6}{36}, \frac{5}{36}, \frac{4}{36}, \frac{3}{36}, \frac{2}{36}, \frac{1}{36})$$

Solution: Huffman code

(1)	8	8	15	36
(2)	7	7	8	
(3)	6	6	7	
(00)	5	5	6	
(01)	4	4		
(02)	3	3		
(030)	2	3		
(031)	1			
(dummy)	0			
(dummy)	0			

2. **Ternary Huffman word lengths**
   Which of the following sequences of word lengths *cannot* be the word lengths of a 3-ary Huffman code and which *can*?

   (a) $\mathbf{L} = (1, 1, 2, 2, 3, 3, 3)$
   (b) $\mathbf{L} = (1, 1, 2, 2, 3, 3)$
   (c) $\mathbf{L} = (1, 1, 2, 2, 3)$
   (d) $\mathbf{L} = (1, 2, 2, 2, 2, 2, 2)$
   (e) $\mathbf{L} = (1, 2, 2, 2, 2)$

Solution: Ternary Huffman word lengths
The simplest way is to draw a ternary tree for each case. Once we have

the tree, we can easily determine whether the set of codes derived from that tree can be Huffman codes (no redundancy) or not. Based on this approach, we can easily see that only codes in (a), (b), and (d) are valid ternary Huffman codes.

Alternatively, since the Huffman code tree should be complete including the dummies at the deepest level of the tree, we can obtain a simple condition for Huffman codewords as

$$\sum_i 3^{-l_i} = 1 - n3^{-L},$$

where L is the length of the longest codewords and n is the number of dummies, which is at most 1 for the ternary case.

In general, for a D-ary code with m codewords, $\sum 3^{-l_i} = 1 - nD^{-L}$, where L is the length of the longest codewords and $n = (D-m) \bmod (D-1)$ is the number of dummies.

3. **Cascade**
Consider the two discrete memoryless channels $(\mathcal{X}, p_1(y|x), \mathcal{Y})$ and $(\mathcal{Y}, p_2(z|y), \mathcal{Z})$. Let the channel transition matrices for the cascade channels in the previous problem be

$X\backslash Y$	1	e	0
1	0	1	0
e	0	1	0
0	0	0	1

$p_1(y|x)$

$Y\backslash Z$	1	e	0
1	1	0	0
e	0	1	0
0	0	1	0

$p_2(z|y)$

(a) What is the capacity C_1 of $p_1(y|x)$?

(b) What is the capacity C_2 of $p_2(z|y)$?

(c) We now cascade these channels. Thus $p_3(z|x) = \sum_y p_1(y|x)p_2(z|y)$. What is the capacity C_3 of $p_3(z|x)$?

(d) Now let us actively intervene between channels 1 and 2, rather than passively transmit y^n. What is the capacity of channel 1 followed by channel 2 if you are allowed to decode the output y^n of channel 1 and then reencode it as $\tilde{y}^n$ for transmission over channel 2? (Think $W \longrightarrow x^n(W) \longrightarrow y^n \longrightarrow \tilde{y}^n(y^n) \longrightarrow z^n \longrightarrow \hat{W}$.)

(e) What is the capacity of the cascade in part c) if the receiver can view *both* Y and Z?

Solution: Cascade

This problem is analogous to Question 3 of Homework Set #8 and exactly same reasoning applies here.

(a) Since $H(Y|X) = 0$ and Y can be only 0 or e, $C_1 = \max I(X;Y) = \max H(Y) = 1$, which is attained by any distribution $p(x)$ with $p(1) + p(e) = 1/2$ and $p(0) = 1/2$.

(b) Similarly, $C_2 = 1$.

(c) The transition matrix for the cascaded channel is

$X \backslash Z$	1	e	0
1	0	1	0
e	0	1	0
0	0	1	0

Since $Z = e$ for any input X, $0 \le C \le \max H(Z) = 0$ so that $C = 0$.

(d) As shown in Question 3 of Homework Set #8, the capacity is $\min(C_1, C_2) = 1$.

(e) Again as in Question 3 of Homework Set #8, the capacity is $C_1 = 1$.

4. **Noisy typewriter**

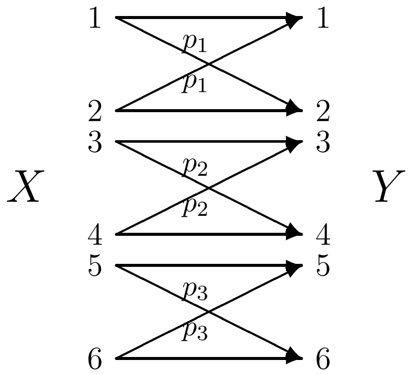
Find the capacity of th m-input channel in which $Y = X + Z \pmod{m}$, where $X \in \{0, 1, 2, \ldots, m-1\}$ and

$$Z = \begin{cases} 1, & \text{w.p. } \frac{3}{4} \\ 0, & \text{w.p. } \frac{1}{4} \end{cases}$$

Solution: Noisy typewriter
Since the channel transition matrix is symmetric, by Theorem 8.2.1 of the text, $C = \log m - H(1/4)$.

5. **Several BSC's**

   (a) What is the capacity of the 6-input, 6-output channel:



   (b) What is the capacity if $p_1 = p_2 = p_3 = 0$?

   (c) What is the capacity if $p_1 = p_2 = p_3 = \frac{1}{2}$?

Solution: Several BSC's
From Question 6 of Homework Set #8, the capacity C of the union of two discrete memoryless channels is $\log(2^{C_1} + 2^{C_2})$. Since this union channel itself is also a DMC, we can easily extend this result to the union of three channels to conclude that

$$2^C = 2^{\log(2^{C_1}+2^{C_2})} + 2^{C_3} = 2^{C_1} + 2^{C_2} + 2^{C_3},$$

or $C = \log(2^{C_1} + 2^{C_2} + 2^{C_3})$.

(a) $C = \log(2^{1-H(p_1)} + 2^{1-H(p_2)} + 2^{1-H(p_3)})$.

(b) $C = \log(2 + 2 + 2) = \log 6$.

(c) $C = \log(1 + 1 + 1) = \log 3$.