Exercise 10.1

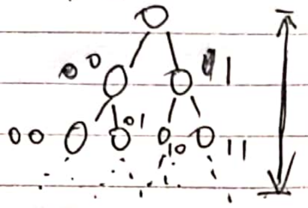
Following the statement of this binary huffman code of this sequence.

$$Pr(X_i) = \frac{1}{n}$$

Then we use Huffman algorithm,

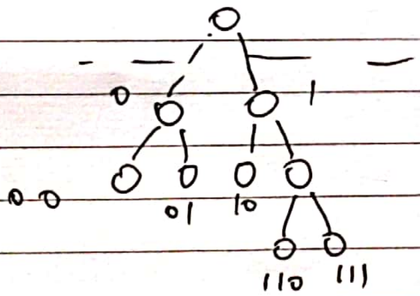
① when $n = 2^L$, we can know that this huffman tree structure can construct a perfect distribution

symbol | length



Therefore, every symbol has same length L and the average length is L

② When $n = 2^L + 1$, it means one node was added in the previous huffman tree, but for the existence of huffman tree, one node must be combined ~~the~~ with this new node.



For $(2^L - 1)$ symbol, their length are all L

For 2 symbol, their length are $L+1$

$$\text{Average length} = (2^L - 1) \cdot \frac{1}{2^L+1} \cdot L + 2 \cdot \frac{1}{2^L+1} \cdot (L+1)$$

$$= \frac{2^L \cdot L + L + 2}{2^L + 1}$$

$$= L + \frac{2}{2^L + 1}$$

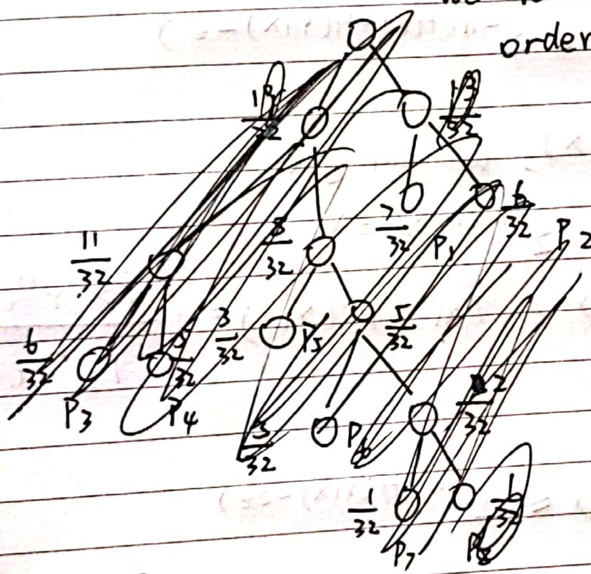
Exercise 10.3

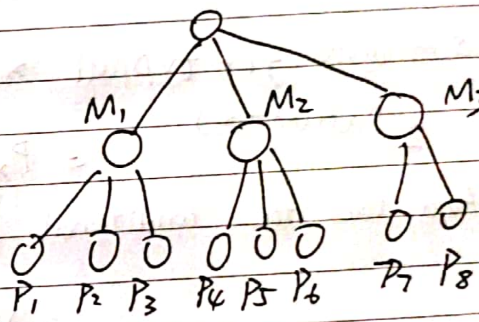
Date No

a) Without mixing the water, the minimum expected number of taste is 1 because you select the sugared water at the first time

In this order: $\left(\frac{7}{32}, \frac{6}{32}, \frac{6}{32}, \frac{5}{32}, \frac{3}{32}, \frac{3}{32}, \frac{1}{32}, \frac{1}{32}\right)$
$P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8$

b). Allowing mixing, ~~we use huffman code strategy.~~ we let ~~this order~~ 8 bottles of water in this order to be mixed





After tasting (M_1, M_2, M_3), we can output the sugar mixture, then tasting each part in this mixture

Expected number of ~~taste~~ taste : 5.9375
Maximal number of taste = 6.

In this strategy, $[P_1, P_2, P_3]$ mixture should be tasted first.

Exercise 10.4

a). ~~F~~ From the statement of joint typical sequence, we know

$$A_\varepsilon^{(n)}(XY) := \left\{ (x,y) \in X^n \times Y^n : \left| \frac{1}{n} \log \frac{1}{P_{X^nY^n}(x,y)} - H(XY) \right| \le \varepsilon \right\}$$

$$\Downarrow$$

$$H(X,Y) - \varepsilon \le \frac{1}{n} \log \frac{1}{P_{X^nY^n}(x,y)} \le H(X,Y) + \varepsilon, \text{ for all sequence } (x,y) \in A_\varepsilon^n$$

$$\Downarrow$$

$$H(X) + H(Y|X) - \varepsilon \le \frac{1}{n} \log \frac{1}{P_{X^nY^n}(x,y)} \le H(X) + H(Y|X) + \varepsilon$$

$$\Downarrow$$

$$2^{-n(H(X)+H(Y|X)+\varepsilon)} \le P_{X^nY^n}(x,y) \le 2^{-n(H(X)+H(Y|X)-\varepsilon)}$$

b). Similarily, for typical $x \in A_\varepsilon^{(n)}(X)$, we can get

$$2^{-n(H(X)+\varepsilon)} \le P_{X^n}(x) \le 2^{-n(H(X)-\varepsilon)}$$

And use the conditional probability, $P_{Y^n|X^n}(y|x) = \dfrac{P_{X^nY^n}(x,y)}{P_{X^n}(x)}$

$$\Downarrow$$

$$2^{-n(H(Y|X)+2\varepsilon)} \le P_{Y^n|X^n}(y|x) \le 2^{-n(H(Y|X)-2\varepsilon)}$$

~~c). Since every sequence in $A_\varepsilon^{(n)}(X,Y)$ has probability at least $2^{-n(H(X,Y)+\varepsilon)}$ by definition, there can be at most $2^{n(H(XY)+\varepsilon)}$ such sequence in the typical set, otherwise the total probability of~~

(c). suppose the new random variables $Z_i = \log \dfrac{1}{P_{X^nY^n}(x,y)} - H(XY)$

$$P[x^n, y^n \in A_\varepsilon^n(x,y)] = P\left[\left| \frac{1}{n} \log \frac{1}{P_{X^nY^n}(x,y)} - H(XY) \right| \le \varepsilon \right]$$

$$= P\left[\left| \frac{1}{n} \log \frac{1}{\prod_{i=1}^n P_{X^nY^n}(x_i, y_i)} - H(XY) \right| \le \varepsilon \right]$$

$$= P\left[\left|\frac{1}{n}\sum_{i=1}^{n}\left(\log\frac{1}{P_{XY}(x,y)} - H(XY)\right)\right| \le \varepsilon\right]$$

$$= P\left[\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \le \varepsilon\right]$$

$$= 1 - P\left[\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| > \varepsilon\right]$$

we can apply the weak law of large numbers (because Z_i are i.i.d and zero mean

$$\Downarrow$$

$$\lim_{n\to\infty} P\left[\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| > \varepsilon\right] = 0$$

$$\Downarrow$$

$$\lim_{n\to\infty} P\left[x^n, y^n \in A_\varepsilon^{(n)}(x,y)\right] = 1$$

$$\Downarrow$$

$$P_{x^n, y^n}\left(\{x, y\} \in \mathcal{X}^n \times \mathcal{Y}^n : x \in A_\varepsilon^{n}(x), (x,y) \in A_\varepsilon^{(n)}(XY)\}\right) \to 1$$