1.

Kraft's ineq: $\sum_{i=1}^{|X|} r^{-l_i} \le 1$

$\sum r^{-l_i} = 4^{-1} \times 2 + 4^{-2} \times 7 = 0.9375 \le 1$ ✓

To construct such a code we should make sure that any code is not a prefix of any other codeword in the code. We have

{0. 1. 20. 21. 22. 30. 31. 32. 33} instantaneous code

{0.15 0.15 0.1 0.1 0.1 0.1 0.1 0.1 0.1} prob

Expected length:

$E(l) = \sum P(l_i) l_i = 1 \times 0.15 \times 2 + 2 \times 0.1 \times 7$

2.

To check a codebook we first verify Kraft's ineq:

if it satisfies → there exists an instantaneous code.

if a code is <u>instantaneous</u> (no codeword is a prefix of others)

↓

uniquely decodable (its extension must be non-singular)

↓

non-singular ($x_1 \neq x_2 \Rightarrow C(x_1) \neq C(x_2)$)

2.

a) {1, 01, 000, 001}

Kraft's ineq:

$$2^{-1} + 2^{-2} + 2^{-3} \times 2 \leq 1$$

inst. → uniquely decodable → non-singular

b) {0, 10, 000, 100}

nonsingular (no codewords equal) ? ✓

uniquely decodable (no ambiguity / different interpretation) ? ✗

instantaneous (decode without ref. to future?) ✗

c) {01, 01, 110, 100}

$$\sum r^{-l_i} = 2^{-2} \times 2 + 2^{-3} \times 2 < 1$$

two codewords equal ⇒ not nonsingular

↓

not u.d.

↓

not inst.

d) $\{0, 01, 011, 0111\}$

$2^{-1}+2^{-2}+2^{-3}+2^{-4}<1$

non-singular ✓

u.d. ✓ (only one possible string. C^+ is non-singular)

inst. ✗ (0 is prefix of the others)

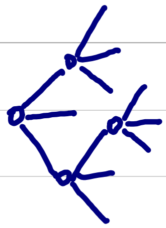
e) $\{10, 01, 0010, 0111\}$

non-singular ✓

u.d. ✓

inst. ✗ (01 is prefix of 0111)

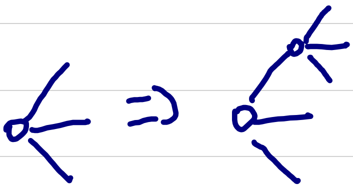
3.



$m(D-1)+1$

m: integer

D: number of leaves at each node



$\underline{-1+D}$

every time remove one leaf.

add D leaves

$+$

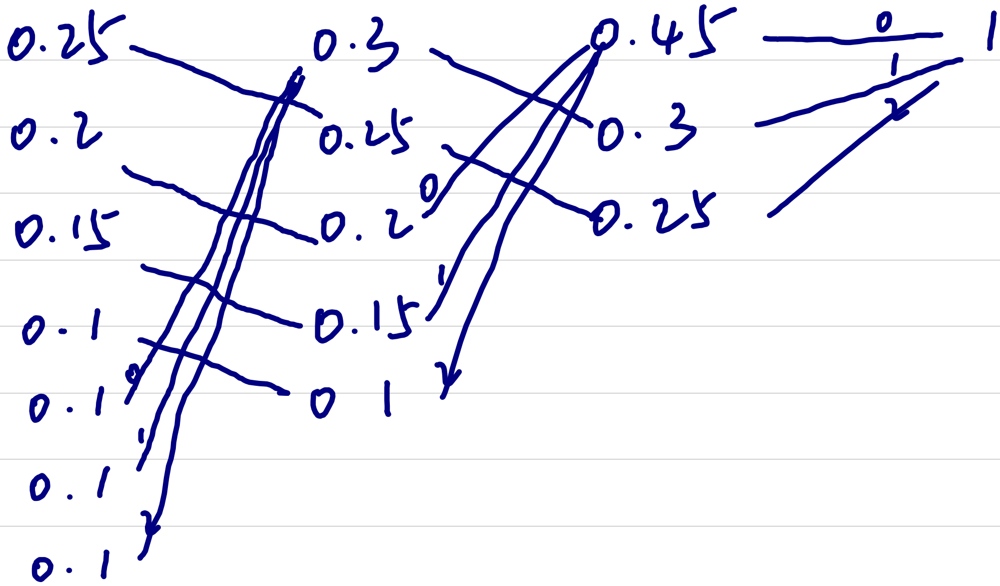
initial leaf $= 1$

$1+m(D-1)$

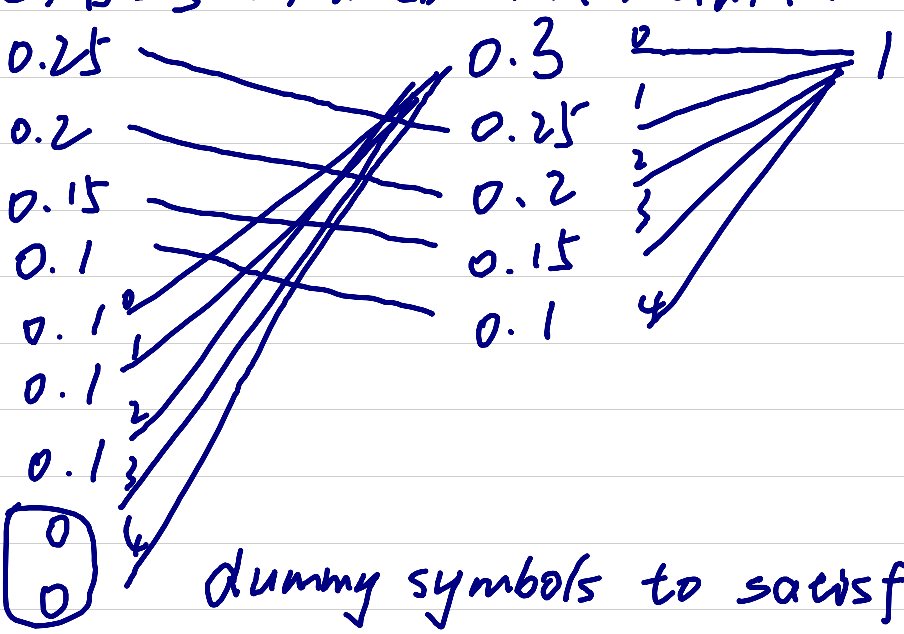
4.

a) $P = \{0.25, 0.2, 0.15, 0.1, 0.1, 0.1, 0.1\}$

$D = 3 \Rightarrow m(D-1)+1 = 2m+1$



c) $D = 5 \Rightarrow m(D-1)+1 = 4m+1$



dummy symbols to satisfy $4m+1$

5.

The entropy of a source (prob. vector) is equal to the expected value of codelength if it satisfies:

$$\sum l \cdot 2^{l} = H(P)$$

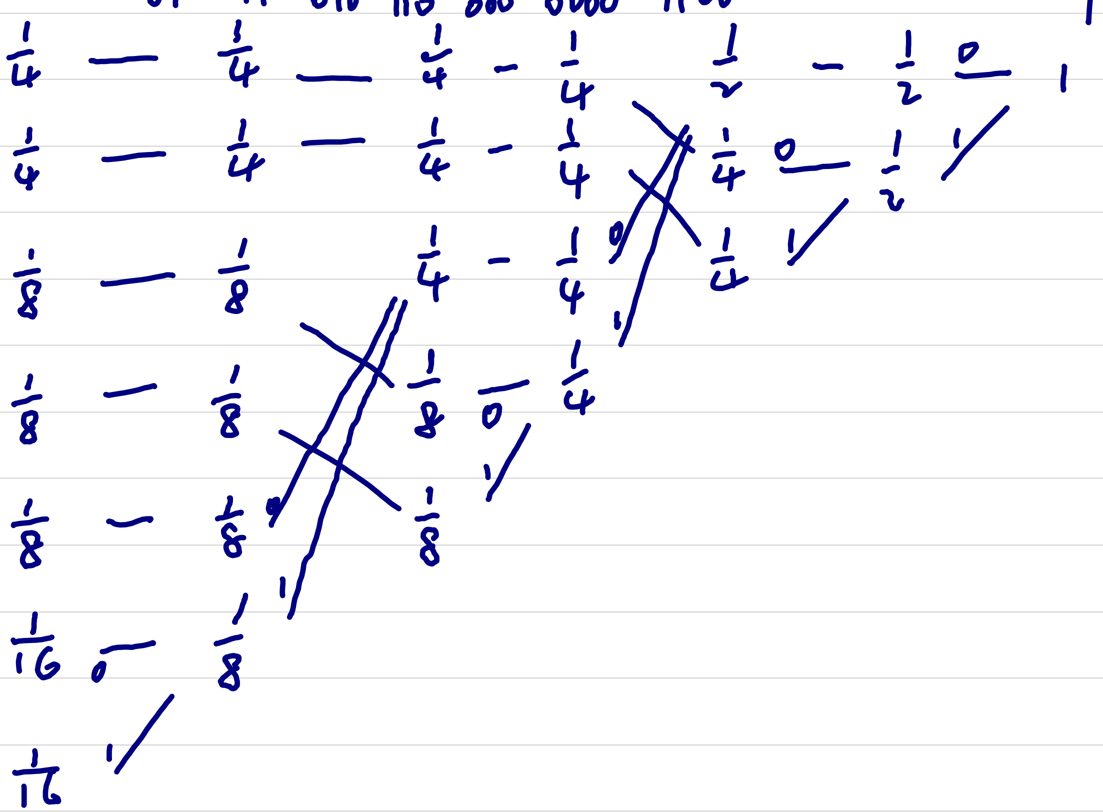
For this we need the probabilities equal 2^{-l} for each l

$\{2, 2, 3, 3, 3, 4, 4\}$ (binary $\rightarrow D = 2$)

$\downarrow$

$\{\frac{1}{4}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{16}, \frac{1}{16}\}$ → sums up to 1 to be a prob. vec.

01 11 010 110 000 0000 1100



6.

$$H(X_i \mid X_{i-1}, X_{i-2} \cdots X_{i-n}) = H(X_i \mid X_{i+1}, X_{i+2} \cdots X_{i+n})$$

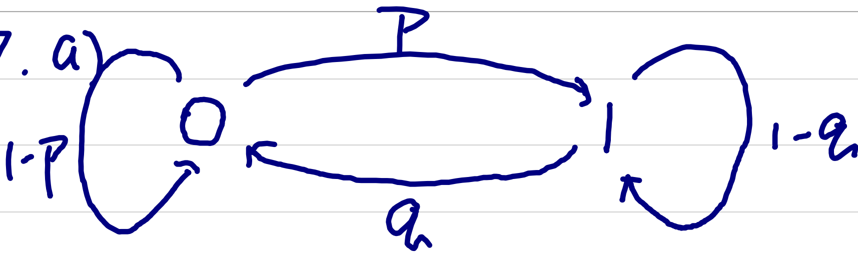
$$H(X_i, X_{i-1}, \cdots X_{i-n}) - H(X_{i-1}, X_{i-2} \cdots X_{i-n})$$

$$\overset{(a)}{=} H(X_{i+n}, X_{i+n-1}, \cdots X_i) - H(X_{i+n}, X_{i+n-1} \cdots X_{i+1})$$

$$= H(X_i \mid X_{i+1} \cdots X_{i+n})$$

(a) because of stationary we shift the seq.s n and n+1 samples.

7. a)



$$T = \begin{array}{cc} \begin{pmatrix} 1-P & P \\ q & 1-q \end{pmatrix} & \begin{array}{c} 0 \\ 1 \end{array} \\ \begin{array}{cc} \quad 0 \quad & \quad 1 \quad \end{array} & \end{array}$$

To find the stationary distribution we should have:

$$\begin{pmatrix} 1-P & P \\ q & 1-q \end{pmatrix} \begin{pmatrix} \alpha \\ 1-\alpha \end{pmatrix} = \begin{pmatrix} \alpha \\ 1-\alpha \end{pmatrix} \Rightarrow \alpha = \frac{q}{P+q}$$

b)

$$H(X) = \underbrace{H(X_n | X_{n-1})}_{}$$ $\rightarrow$ two-state Markov chain

in order to maximise $H(X)$ we have

$$H(X_n | X_{n-1}) \leq H(X_n) \leq 1$$

We can attain $H(X) = 1$ if $P = q = \frac{1}{2}$

* Another approach would have to differentiate w.r.t. q and P and equate with zero to solve it.

prob. state 0 is chosen

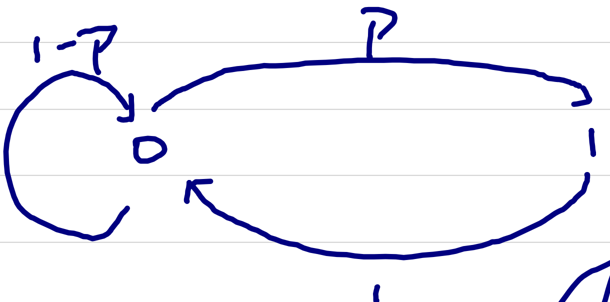
c)

(i) $H(X) = H(X_2 | X_1) = H(X_2 | X_1 = 0) P(X_1 = 0) + H(X_2 | X_1 = 1) P(X_1 = 1)$

$= \alpha H(p) + (1-\alpha) H(q_1)$

entropy of going to state $X_2 = 0$ or 1 when we observe state 1

$\stackrel{q_1 = 1}{=} \alpha H(p)$



$S_1(n) = S_0(n-1)$

$S_0(n) = S(n-1)$

once you see 1 the next r.v. is 0

the next r.v. can be either 0 or 1.

(ii)

$S_1(n) = S_0(n-1) = S(n-2)$

using this we have

$\underbrace{S(n)} = S_0(n) + S_1(n) = S(n-1) + S(n-2)$

the total number of 0/1 sequences

$f(1) = S(1) = 2$

$f(2) = S(2) = 3$

$f(3) = S(3) = 2+3 = 5$

$f(4) = S(4) = 8$

$f(5) = S(5) = 13$

Fibonacci

$\varphi = \frac{1+\sqrt{5}}{2}$ golden number

$f(n) = \frac{\varphi^n + (1-\varphi)^n}{\sqrt{5}}$

(iii) $\overset{\text{entropy rate}}{H(\mathcal{X})} = \frac{1}{n} H(\overset{(X_1, \dots X_n)}{X^n})$

The maximum is attained when X_n follows a uniform distribution over its space

The space of X^n is the sequence of 0/1 with length n which is S_n.

$\Rightarrow \max H(\mathcal{X}) \le \max \frac{H(X^n)}{n} = \frac{1}{n} \log |S_n| = \frac{1}{n} \log \frac{\varphi^n - (1-\varphi)^n}{\sqrt{5}}$

$n \to \infty$: $\max\limits_{n\to\infty} H(\mathcal{X}) = \frac{1}{n}(n \log \varphi - \frac{1}{2} \log 5) = \log \varphi = \log \frac{1+\sqrt{5}}{2}$

8. $X_i \in \{A. B\} \sim \{0.9\ 0.1\}$

$H(X) = 0.9 \log 0.9 + 0.1 \log 0.1$

$\{AA\ AB\ BA\ BB\}$ → use Huffman coding to design
0.81 0.09 0.09 0.01 the codebook

obtain the average length of the codebook by $\sum_{i=1}^{4} l_i P_i$

redun: $H(P) - E(L)$