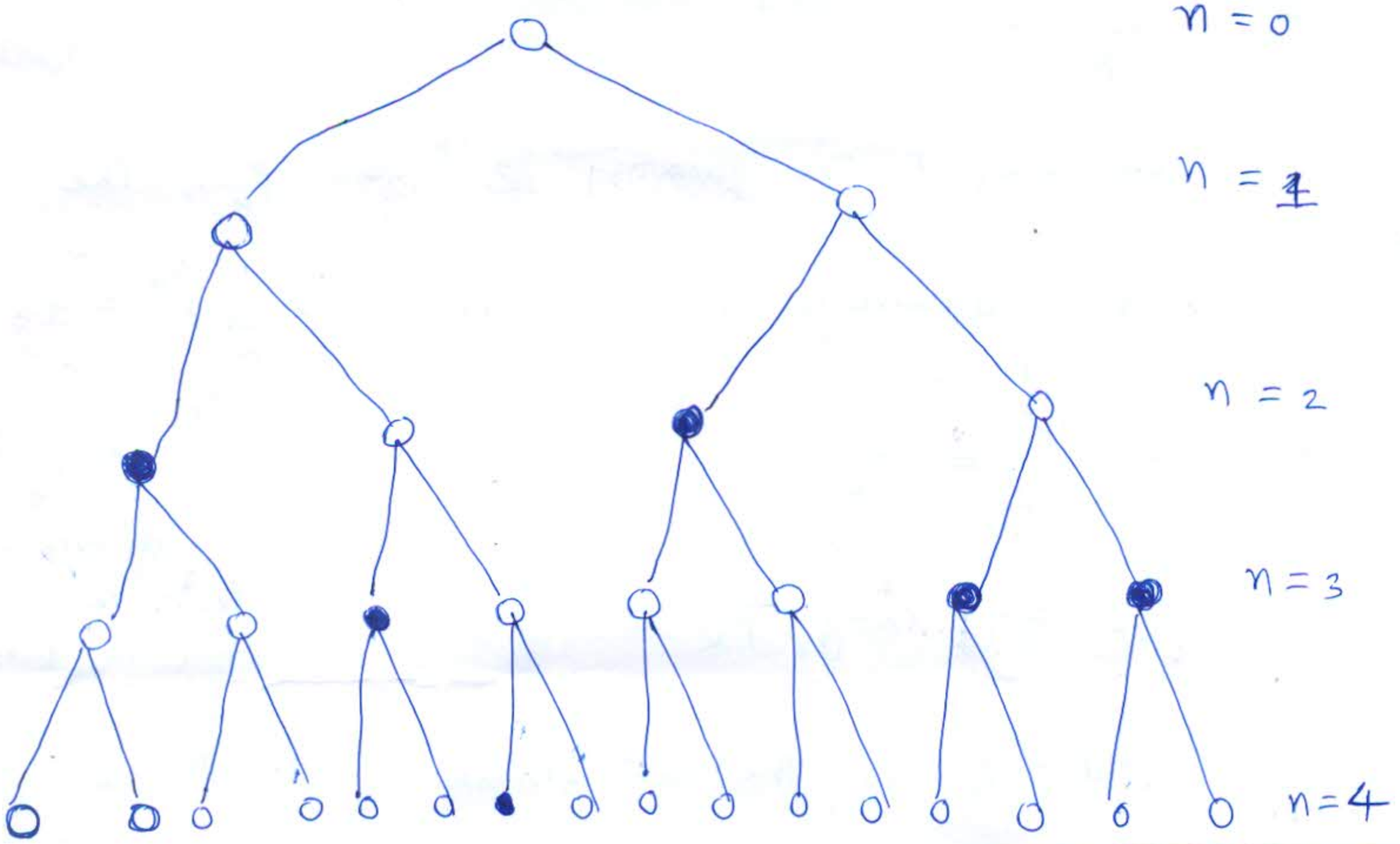
A Binary Tree Representation of Prefix-free Codes:



Kraft's Inequality:

→ Let the height of k^{th} prefix-free codeword be denoted as n_k.

→ Let number of leaf nodes that are precluded from consideration as codewords be denoted as A_k.

→ $\boxed{A_k = 2^{n - n_k}}$, where n is the height of the binary tree.

→ Now $\sum_k A_k \leq 2^n \Rightarrow \sum_k 2^{n-n_k} \leq 2^n$

→ This proves $\boxed{\text{Kraft's Inequality: } \sum_k 2^{-n_k} \leq 1}$

Kraft's Inequality : An Interpretation :

Let us consider a case when inequality is replaced by equal sign.

$$\sum_k 2^{-n_k} = 1 .$$

One can interpret the quantity 2^{-n_k} as p_k, the probability of occurrence of k^{th} symbol.

$\Rightarrow \sum_k 2^{-n_k} = 1$

$\Rightarrow \sum p_k = 1$

This is as expected. Probabilities should sum to 1!

Now, given that

$$2^{-n_k} = p_k$$

$$\Rightarrow \quad n_k = -\log_2 p_k$$

However $-\log_2 p_k$ is the self-information, of k^{th} symbol.

Thus, the length of the codeword ~~should~~ equals its self-information

ENTROPY & Asymptotic Equipartition Property or AEP

→ Consider a sequence of length N emitted by a binary iid ergodic source, with "0" probabity of p and "1" probability of $(1-p)$.

→ If N is sufficienty large, this sequence should have approximately Np "0"s and $N\times(1-p)$ "1"s.

→ Since the source is iid, probability of occurrence of each such string is $p^{Np} \times (1-p)^{N(1-p)} = P_{TYP}$

AEP and Entropy :

→ How many such strings are There?

→ Max # of strings : 2^N

→ Actual # of strings $m < 2^N$ if $p \neq 1/2$.

→ With N sufficiently large, it is almost guaranteed That all strings will have Np zeros and $N(1-p)$ ones.

→ Prob. of each such string is P_{TYP} .

→ $m \times P_{TYP} \approx 1$

$$\Rightarrow m = 1/P_{TYP} = p^{-Np} \times (1-p)^{-N(1-p)}$$

$$\Rightarrow \log_2 m = N\left(-p\log_2 p - (1-p)\log_2(1-p)\right)$$

$$= N \times H(x)$$

1	0	0	0	0
2	0	0	0	1
3	0	0	1	0
4	0	1	0	0
5	1	0	0	0
	⋮			
16	1	1	1	1

$m = 5 \qquad 2^N = 16$

$\lceil \log_2 m \rceil^{\approx 3}$ bits to encode this source..

If one takes length N binary strings, one is likely to see <u>ONLY</u> those strings for which # of zeros is $N \times p$ and # of ones is $N \times (1-p)$, where p is prob. of zero.

All these strings can be thought of as belonging to the "Typical set", and they all occur with equal prob. given as

$$\boxed{P_{Typ} = p^{N \times p} \cdot (1-p)^{N(1-p)} = \frac{1}{m}}$$

, where m is the size of the typical set.

$$\log_2 m = -\log_2 P_{TYP}$$

$$= N\,H(x)$$

$$\Rightarrow m = 2^{N\,H(x)}$$

i.e., Instead of 2^N strings, we are going to get $2^{N \times H(x)}$ strings.

i.e., Instead of N bits, we need $N \times H(x)$ bits to represent the outcome of the source.

Therefore, average rate R of the source is $H(x)$ bits. It cannot become smaller.