

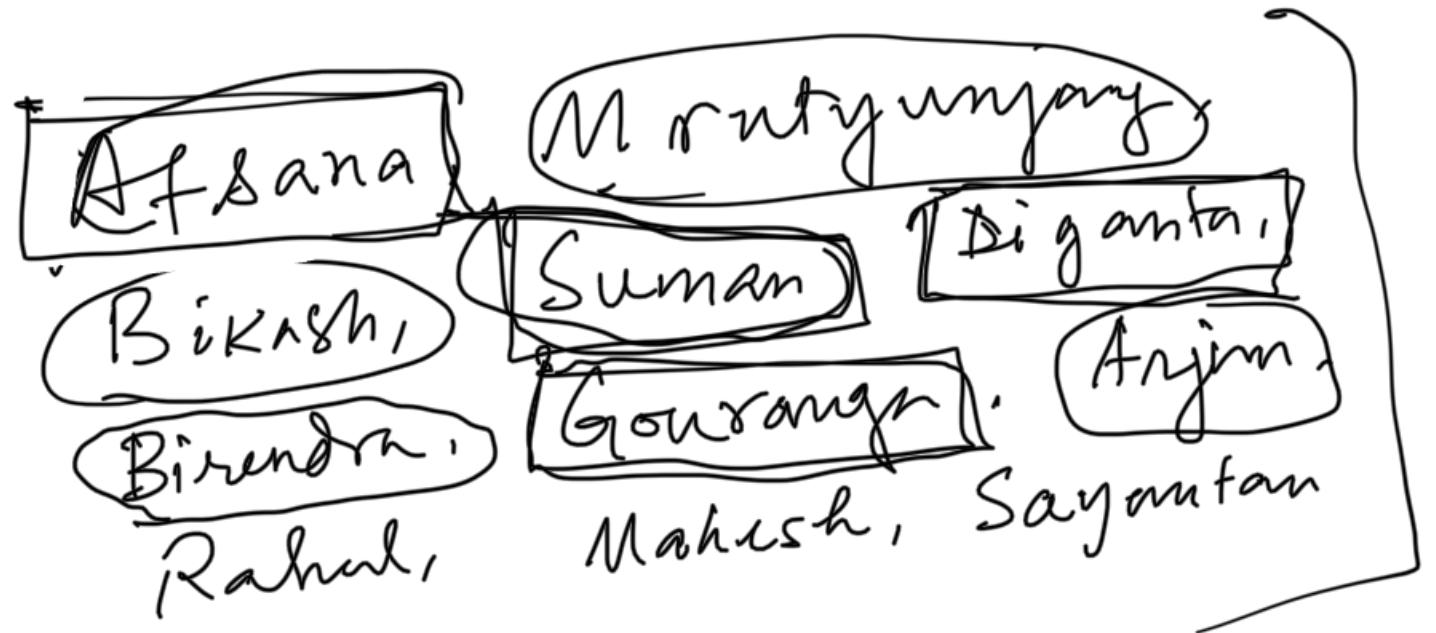
# Information and Coding theory.

Lee-7.

Planning - Viva Exam - on Feb 14th.  
Monday.

10% weightage.

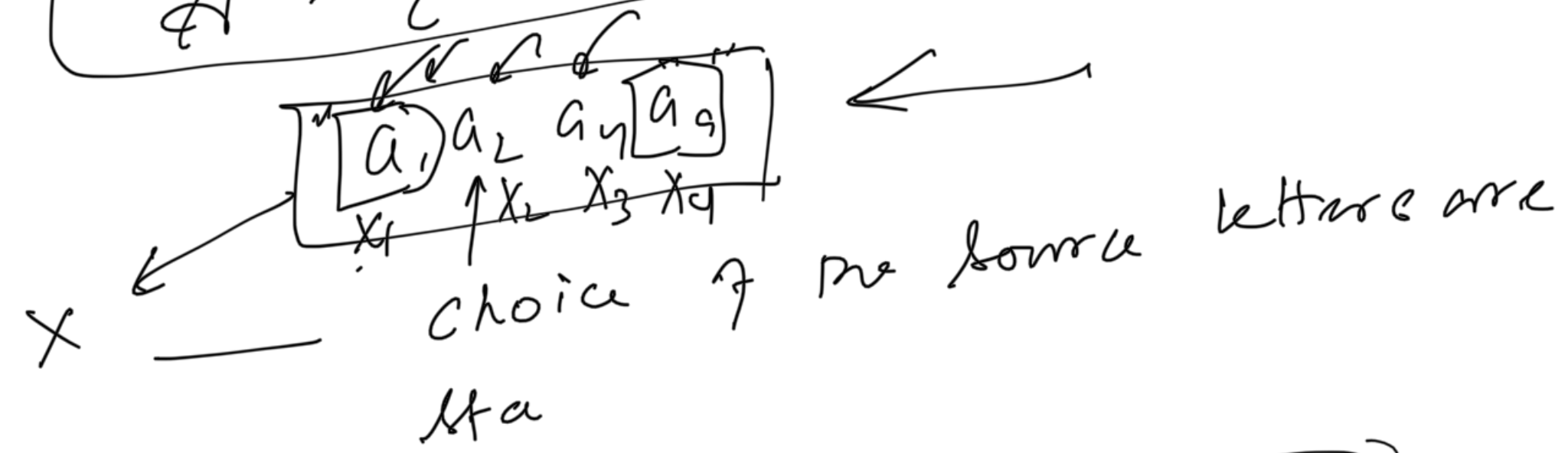
Notes -





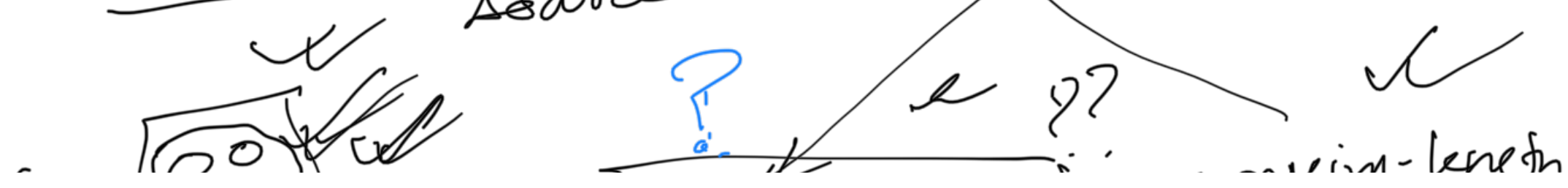
Source ———→ memoryless  
                            → with memory.

$$A = \{a_1, a_2, \dots, a_k\} \quad p(a_i)$$



HT HT

Recall: there is a codeword for each source letter.





fixed-length

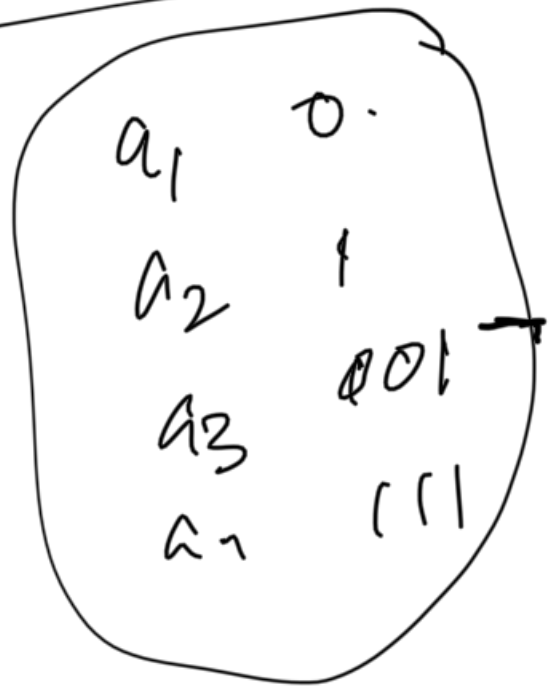
varying

10 bits

0010011



channel



Suppose we have a memoryless channel source with fixed-length code.

thank

✓

Let us assume  
Source sequence:  $\bar{x} = (x_1, x_2, \dots, x_{|L|})$   
 is a sequence of  $L$  letters generated  
 from the source.

$$x_i \in A = \{a_1, \dots, a_{|K|}\}$$

Q. How many seq<sup>s</sup>  $\bar{x}$  can be generated

Ans.  $K^L$  ✓

Codeword / code - alphabet

$\{0, 1\}$  ✓

Let us assume that the code  
 alphabet contains  $|D|$  letters. ✓  
 (so  $D=2$  in case of binary code)

number of codeword for

Suppose that the <sup>original</sup> source letter is  $N$

Ans.

$$D^N$$

different seqs of  
code words are possible.

Then

Observation:

$$D^N \geq K^2$$

$$\overline{x} = \begin{pmatrix} \vdots \\ \vdots \\ \vdots \end{pmatrix}$$

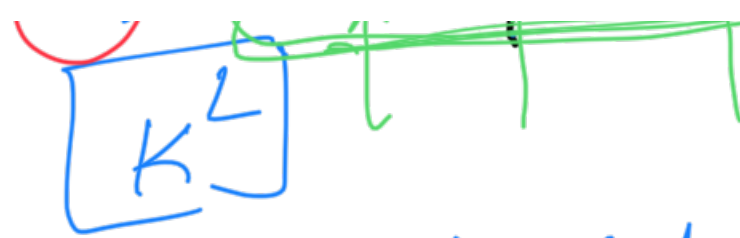
$$\Rightarrow N \log D \geq 2 \log K$$

$$\Rightarrow \boxed{\frac{N}{2}} \geq \frac{\log K}{\log D}$$

Source letters :  $\{a_1, a_2, \dots, a_K\} = A$

$\overline{x}_L = (x_1, x_2, \dots, x_L), [x_i \in A]$

Source Seq:



Code letters : # of search code letter is  $D$ .  
(for instance  $D=2$ )  
 $\{0, 1\}$   $\leftarrow$   $\{0, 1, 3\}$   
 $\{0, 1\}$   $\leftarrow$   $\{0, 1, 3\}$

$$a_1 = 00$$

$$a_1 = 130 \leftarrow$$

Fixed-length code:

$N$   $\rightarrow$  the length of  
code word for each source  
letter.  $\uparrow$



$$\frac{N}{L} \geq$$

$$\frac{\log K}{\log D}$$

and corresponding

i.e. for each code word  
 a source seq<sup>n</sup> ~~look~~ at letters  
 $\frac{\log K}{\log D}$  code letters are required for  
 each source letter.

If  $\bar{x} = (x_1, x_2, \dots, x_L) \in X_1 \times X_2 \times \dots \times X_L$  is a source seq<sup>n</sup>.

$\hookrightarrow \underbrace{X_1 X_2 \dots X_L}_{L\text{-times}}$

$\Pr(\bar{x}_L)$  ✓

$= \Pr((x_1, x_2, \dots, x_L) = \bar{x}) \underbrace{(x_1, x_2, \dots, x_L)}_{L\text{-tuple}}$

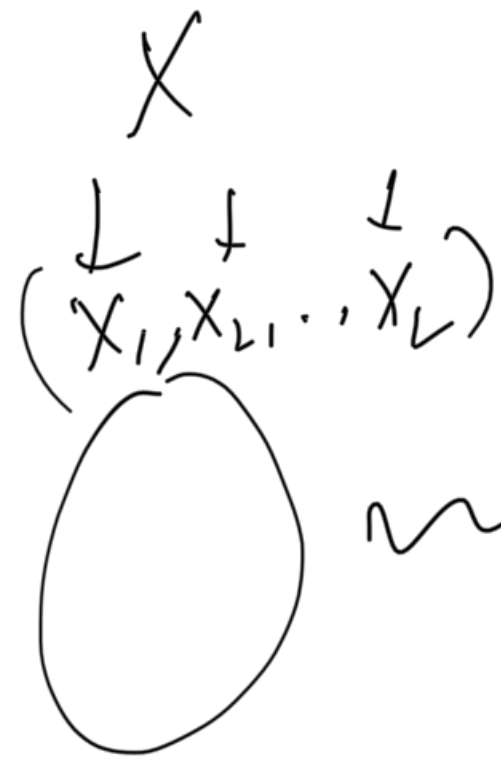
$= \Pr(\underbrace{x_1 = x_1}_{\text{...}}, \underbrace{x_2 = x_2}_{\text{...}}, \dots, \underbrace{x_L = x_L}_{\text{...}})$

~~$X \rightarrow p(x)$~~

$$) = P_X(\underbrace{X=x_1, \dots, X=x_L}_{\text{sequence}})$$

$$= P(X=x_1) P(X=x_2) \dots P(X=x_L)$$

$$= \prod_{l=1}^L P(x_l)$$



$(X, Y) \leftarrow \text{joint } (X, Y)?$

$f(x, y) = P[X=x, Y=y]$

$= P[X=x] P[Y=y]$

Q. what is the self-information of such a source sequence?

$$I(\bar{x}_L) = -\log \prod_{l=1}^L P(x_l)$$

$$= - \sum_{l=1}^L \log p(x_l)$$

$$= \sum_{l=1}^L - \log p(x_l)$$

$$= \sum_{l=1}^L I(x_l)$$

Recall:

convergence of random variables  
 $\xrightarrow{\text{convergence}} X$  as  $n \rightarrow \infty$

Let  $S$  be sample space.  $X_1, X_2, \dots$

①

In probability:

if for any  $\varepsilon > 0$

$$P(|X_n - X| > \varepsilon) \rightarrow 0$$

$$\Rightarrow P(X - \varepsilon < X_n < X + \varepsilon) \rightarrow 0$$

$$P(|X_n(\omega) - X(\omega)| > \varepsilon) \rightarrow 0$$

$$X_n: S \rightarrow \mathbb{R}$$

$$X: S \rightarrow \mathbb{R}$$

$$X_n(\omega) \in L^2$$

WES,  $X_n(\omega) \in L^2$

(2)

In mean square:

$$E (X_n - X)^2 \rightarrow 0$$

(Asymptotic equipartition property)??  
A.E.P

Thm  
If.

$X_1, X_2, \dots$  are i.i.d random variables with pmf  $[p(x)]$

then  $-\frac{1}{n} \log p(X_1, X_2, \dots, X_n) \rightarrow H(X)$   
in probability

$$X_i \sim X$$

Ben weak law of large numbers.

Proof.

WLLN

$X_1, \dots, X_n, \dots, X_{1,2,3}$

$x_1, \dots, x_n$

Sample mean

$x_1 + \dots + x_n$

$n$

$E(X)$

$g(X_i)$   
 $\log p(X_i)$

$-\frac{1}{n} \log p(X_1, X_2, \dots, X_n)$

$= -\frac{1}{n} \sum_i \log p(X_i)$

WLLN

$= -E(\log p(X))$   
 $= H(X)$

$\log p(X_1) + \log p(X_2) + \dots + \log p(X_n)$

$$\begin{aligned} & \rightarrow E(\log p(x)) \\ & \sum p(x) \log p(x) = H(X) \end{aligned}$$

$$\frac{I(x_1)}{L} \rightarrow H(X)$$

Goal!

deciding the value of  $N$ ?

length of ~~each~~ word of each

Source letters.

$\bar{x}_L \equiv (x_1, \dots, x_L) \rightarrow$  source sequence.

$K^L$   $\rightarrow$  # of source letters  
 $D^N$   $\rightarrow$  # of code letters.



Typical set:

typical set

The  $A^{(L)}_\epsilon$

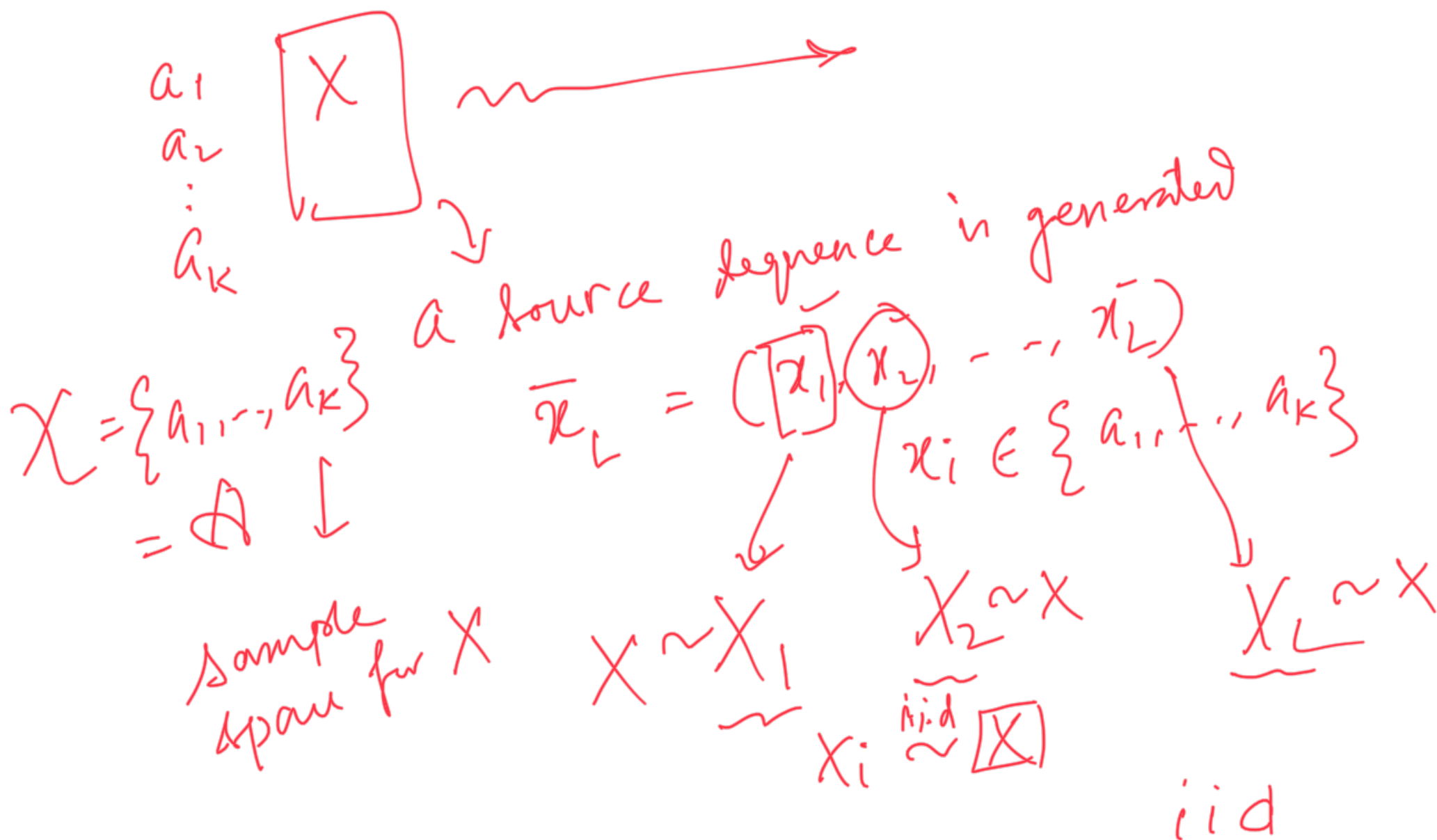
is the set of  $\bar{x}_L = (x_1, x_2, \dots, x_L)$  such that

$$2^{-L(H(x) + \epsilon)} \leq p(x_1, x_2, \dots, x_L) \leq 2^{-L(H(x) - \epsilon)}$$

$$P(\bar{x}_L)$$

$$\frac{L - L_H(X)}{2}$$

Lec-8



AEP:

$$X_1, X_2, \dots \rightarrow H(X)$$

$$-\frac{1}{L} \log p(X_1, X_2, \dots, X_L) \rightarrow H(X)$$

$$-\frac{1}{L} \log p(x_1, x_2, \dots, x_L) \rightarrow H(X)$$

↓  
convergence in probability

$$-\log p(x_1, x_2, \dots, x_L)$$

$$= -\log p(\bar{x}_L)$$

$$= I(\bar{x}_L)$$

$$\frac{I(\bar{x}_L)}{L} \approx H(X) \checkmark$$

$$\Rightarrow -\log Pr(\bar{x}_L) \approx L H(X)$$

$$-\log Pr(\bar{x}_L) \approx L H(X) \checkmark$$

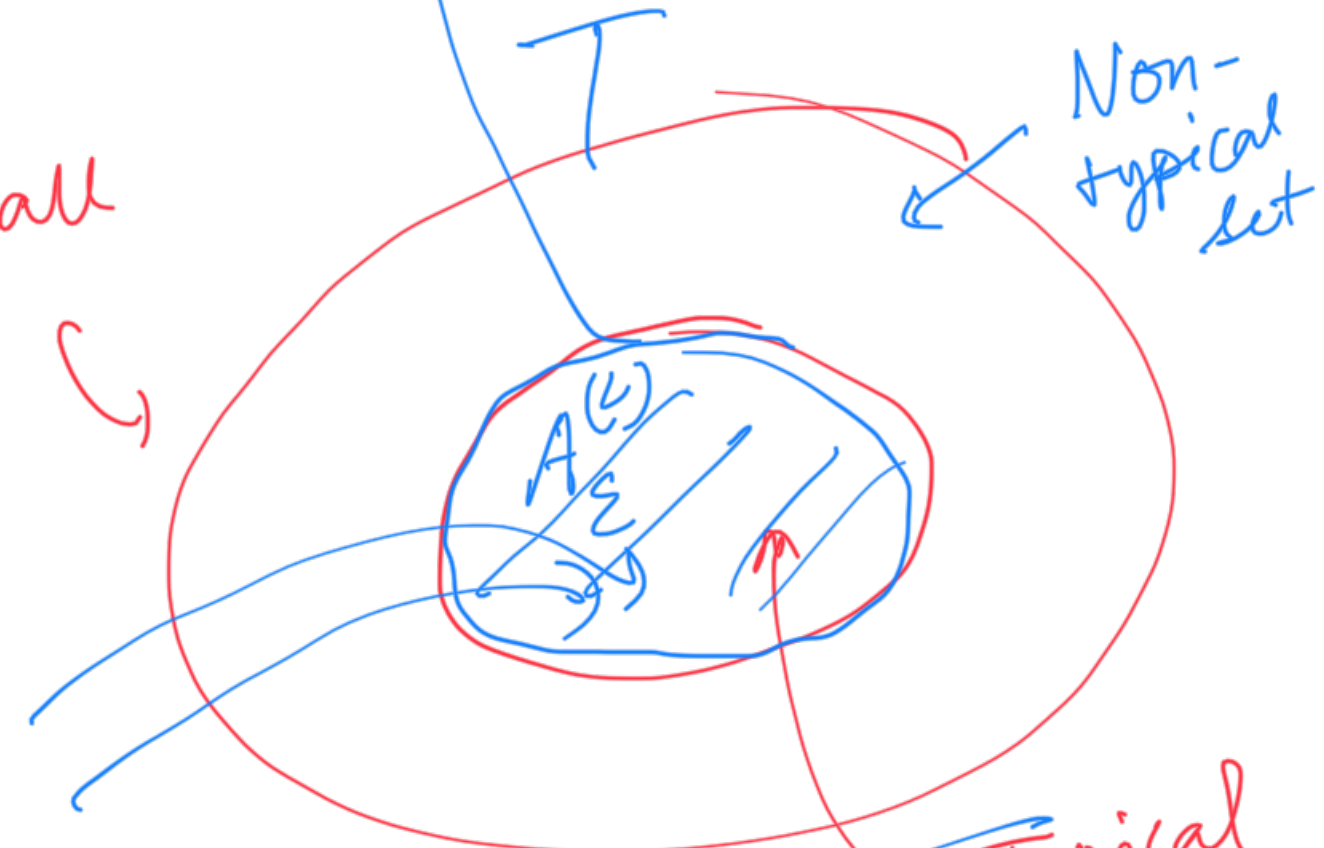
$\Rightarrow$

$P_r(\bar{x}_L) \leq 2^{-L(H(X) \pm \epsilon)}$

$-L(H(X) \pm \epsilon)$

$2$

Set of all  $\bar{x}_L$



$2^{-L(H(X) \pm \epsilon)} \leq P_r(\bar{x}_L) \leq 2^{-L(H(X) - \epsilon)}$

Typical set

non-typical set

Thm.

Properties of the

① if  $\bar{x}_L \in A_\varepsilon^{(L)}$  then

$$H(x) - \varepsilon \leq \frac{\log p(x_1, x_2, \dots, x_L)}{L} \leq H(x) + \varepsilon$$

②

$$P_r \{ A_\varepsilon^{(L)} \} > 1 - \varepsilon$$

when  $L$  is large.

Pf.

$$-\frac{1}{L} \log p(x_1, \dots, x_L) \rightarrow H(x)$$

$$\text{i.e. } P_r \left\{ \left| -\frac{\log p(x_1, \dots, x_L)}{L} - H(x) \right| < \varepsilon \right\} > 1 - \delta$$

Then setting  $\delta = \epsilon$

(2)

$$|A_\epsilon^{(L)}| \leq 2$$

$$L(H(x) + \epsilon)$$

the # of ~~elem~~ source seq.  
in the typical set.

Pf.



$$|A_\epsilon^{(L)}| =$$

$$\sum_{\bar{x}_L \in A_\epsilon^{(L)}} P(\bar{x}_L)$$

$$= \sum_{\bar{x}_L \in A_\epsilon^{(L)}} P(\bar{x}_L) + \sum_{\bar{x}_L \in (A_\epsilon^{(L)})^c} P(\bar{x}_L)$$



$$\sum_{\bar{x}_L \in A_\epsilon^{(L)}} P(\bar{x}_L)$$

$$\sum_{\bar{x}_1} p(\bar{x}_1) = 1$$

$$\geq \sum_{\bar{x}_2 \in A_\varepsilon^{(L)}} 2^{-L(H(x) + \varepsilon)}$$

$$= |A_\varepsilon^{(L)}| 2^{-L(H(x) + \varepsilon)}$$

$$\Rightarrow |A_\varepsilon^{(L)}| \leq 2^{L(H(x) + \varepsilon)}$$

---


$$\left\{ \boxed{100}, \boxed{200}, \frac{1}{11}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots \right\}$$

$$\begin{array}{ccccccc}
 & & x_{n-1} & x_n & x_1 & \frac{1}{n} & \dots
 \end{array}$$

$\left( -\frac{1}{2}, 0, \frac{1}{2} \right) \rightarrow \bigcirc$

$\left( -\frac{1}{2}, \frac{1}{2} \right) \quad \left( 0 - \varepsilon, 0 + \varepsilon \right)$

$\varepsilon = \frac{1}{2}$

$x_n \rightarrow l \quad \exists \quad K \quad \text{A.t.}$

if  $\varepsilon > 0$

$|x_n - l| < \varepsilon, \quad n \geq K$

i.e.  $x_n \in (l - \varepsilon, l + \varepsilon), \quad n \geq K$

④

$|A_\varepsilon^{(L)}| \geq (1 - \varepsilon) 2^{L(A(x) \varepsilon)}$

when  $L$  is large

Pf.  
=

H.W.



Code: the set of codewords.

Nonsingular Code:

if ~~every~~ element of  
 $x \neq x' \Rightarrow$

$x, x' \in X$

$C(x) \neq C(x')$

\* code  $C$ :

## Extension $C^n$ of a

if we take a sequence of source letters  $(x_1, x_2, \dots, x_L)$  then

$$C^*(x_1, x_2, \dots, x_L) = C(x_1) C(x_2) \dots C(x_L)$$

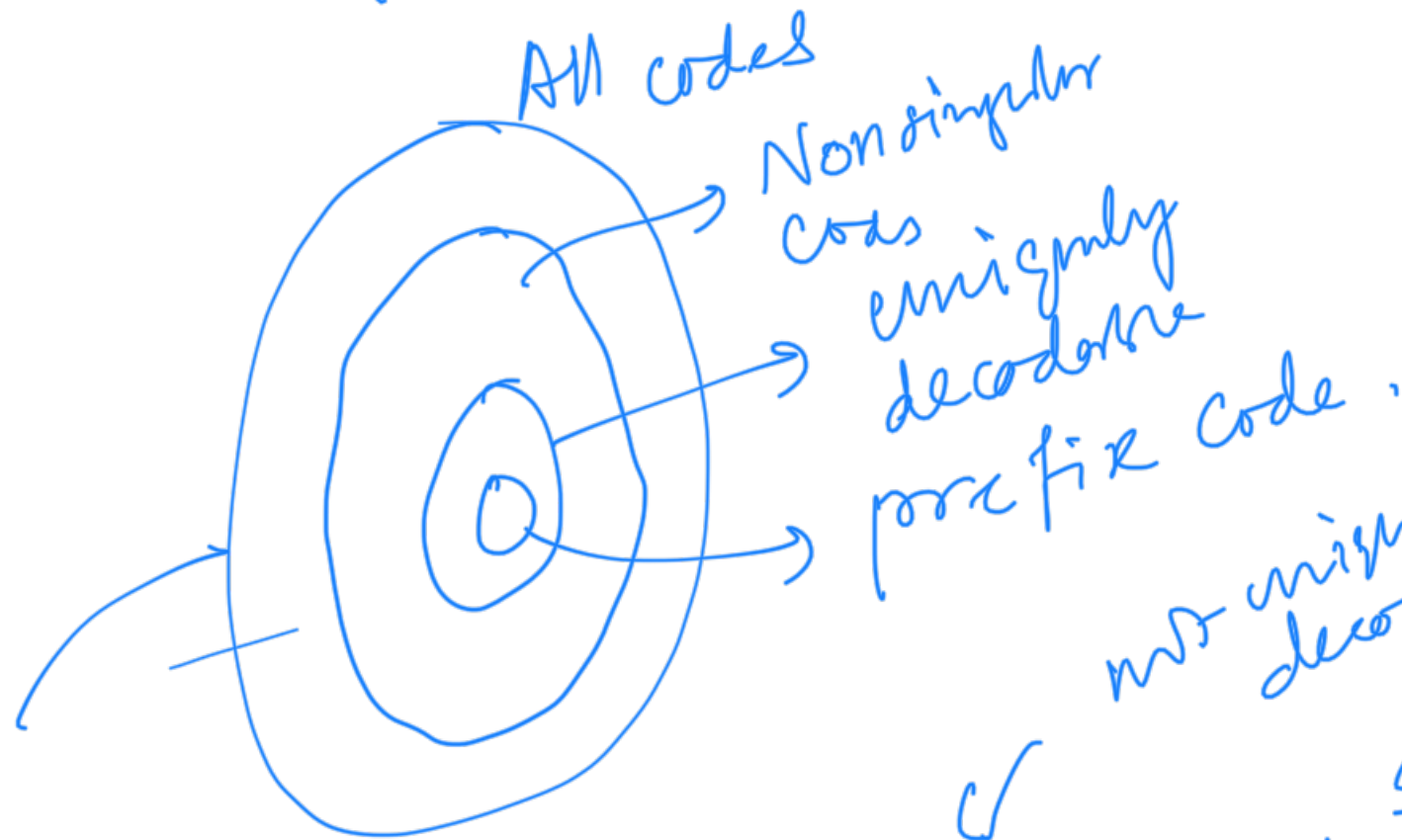
Exp.  $a_1 = 00$   $C^*(a_1, a_2)$   
 $a_2 = 01$   $= C(a_1) C(a_2)$   
 $= 0001$

## Uniquely decodable Code:

A code is called uniquely decodable if  $C^*$  is non-singular.

# Prefix Code

A code is said to be prefix code if no codeword is a prefix of any other codeword.



H.W.

not unique decoder

uniquely decodable but not prefix

1 0

0

|       |   |     |            |
|-------|---|-----|------------|
| $a_1$ | 0 | 010 | 00         |
| $a_2$ | 0 | 01  | 11         |
| $a_3$ | 0 |     | <u>110</u> |
| $a_4$ | 0 | 10  | 001        |

$\boxed{010}$   $\begin{cases} \rightarrow a_2 \\ \rightarrow a_1, a_4 \end{cases}$

Is it prefix.

$I(x; y)$

$\boxed{I(X; Y)}$   $z$

$\lim_{n \rightarrow \infty} \frac{p(n, y)}{p(n) p(y)}$

$\lim_{n \rightarrow \infty} \frac{p(n|y)}{p(n)}$

$p(y)$

|       |             |       |
|-------|-------------|-------|
|       | $y_1$       | $y_2$ |
| $x_1$ | $\boxed{1}$ | 0     |
| $x_2$ | 0           | 1     |

$$\begin{bmatrix} p(y_1|x_1) & p(y_2|x_1) \\ p(y_1|x_2) & p(y_2|x_2) \end{bmatrix}$$

