Introduction.

information থাকলে তার probability থাকবে। যদি কোনো event এর probability 100% হয়, তাহলে তার মধ্যে কোনো information নাই। probability যত কমবে information তত বাড়বে।

Defination: Let a discrete random variable X with possible outcomes x_i, $i = 1, 2, \ldots, n$. The self information of the event $X = x_i$ is defined as

$$I(x_i) = \log\left(\frac{1}{P(x_i)}\right)$$

$$= -\log P(x_i).$$

So, high probability function convey less information than a low probability event.

If $P(x_i) = 1$, then $I(x_i) = 0$.

This is the mathematical model of information. And also the relationship between uncertainity and information.

The unit of $I(x_i)$ is determined by the base of logarithm.
when the base is 2, then bits
when the base is e, then nats.

Since $0 \leq P(x_i) \leq 1$, $I(x_i) \geq 0$; then self-information is non-negative.

Scanned by CamScanner

Example - 1.1 If $P(0) = P(1) = 0.5$,

Then the information ~~contain~~ content of each output from the source is

$$I(x_i) = -\log P(x_i)$$
$$= -\log_2 (0.5)$$
$$= 1 \text{ bit}$$

One bit is used to represent output.

* The self-information of an m-bit block is

$$I(x_i) = -\log_2 P(x_i)$$
$$= -\log_2 2^{-m} = m \text{ bits}$$

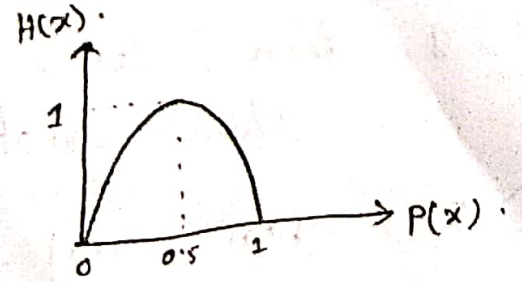
The Mutual information $I(x_i, y_j)$ between x_i and y_j is defined as

$$I(x_i, y_j) = \log \left(\frac{P(x_i | y_j)}{P(x_i)} \right)$$

In case of individual events, their probabilities are multiplied.

But information is added in logarithmic scale.

-যখন probability equal হয়, তখন entropy maximum হয়।



$$H(x) = -p\log_2(p) - (1-p)\log_2(1-p)$$
$$= -\sum_{i=0}^{1} P(x_i) \log P(x_i).$$

Scanned by CamScanner

Entropy: The entropy of a message is defined as the amount of a information to be transmitted to the random variable x.

একটা data or message এর মধ্যে কী পরিমান information আছে or determine করার জন্য entropy use হয়।

Why information theory?

1) Information theory is needed to determine the amount of information that can be transmitted through band-limited noisy channel.

2) How much error can be reduced and corrected.

3) Source coding is done to compress the data.

4) Channel coding is done to recover data by adding extra bit.

5) The amount of compression depends on the amount of information.

information বেশি হলে, compression কম হবে।
information কম হলে, compression বেশি হবে।

Scanned by CamScanner

Compression: Data compression means encoding information using fewer bits than original representation.

Compression is two types.

1) Lossy 2) Lossless.

Lossless when every single bit of data that was originally in the file remains after the file is uncompressed.

All of the information is completely restored from the compressed file.

In voice communication lossless compression is used.

Example: GIF, Huffman coding, arithmatic coding, LZW, Runlength coding.

Lossy:

Lossy compression reduces the redundanty information from the file.

When the file is uncompressed, only original information is still there.

Example: JPEG-2000, Wavelet based compression

Scanned by CamScanner

Compression Ratio :

~ it is the ratio between the uncompressed size and compressed size.

$$= \frac{\text{uncompressed sequence}}{\text{compressed sequence}}.$$

$$\therefore R = \frac{n_1}{n_2}$$

R is always greater than 1.

Redundancy :

→ It is the amount of wasted space used to transmit certain data.

→ It is the measure of how much unnecessary data can be removed.

→ It determines how much data can be compressed.

→ Redundancy $= (1 - 1/R)\%$

$$= \left(1 - \frac{n_2}{n_1}\right).$$

$$= \left(\frac{n_1 - n_2}{n_1}\right)\%$$

* The average self Information of a random variable X

$$H(X) = \sum_{i=1}^{n} P(x_i) \cdot I(x_i) = \sum_{i=1}^{n} P(x_i) \cdot \log P(x_i)$$

Scanned by CamScanner

Source Coding:

The primary motivation is the compression of data due to efficient representation of the symbols.

A code is a set of vectors called codewords.

When source symbols are not equally probable, a more efficient method is then Variable Length Coding. (VLC).

Run Length Coding:

→ Run length coding is a technique used to reduce the size of repeating string of characters.

X → It has high compression ratio.

→ RLE encodes a run of symbols into two bytes, a count and a symbol.

→ RLE compress any type of data regardless of its information content, but the content of data to be compressed affects the compression ratio.

→ RLE cannot acheive high compression ratio as compared to other compression method.

Scanned by CamScanner

→ It is easy to implement.

→ repeated signal যত বেশি, compression তত বেশি।

→ এটা secondary compression algorithm হিসেবে use হয়।

1) aabcaaabccc ccc ddd

$= (a,2), (b,1), (c,1), (a,3), (b,1), (c,5), (d,3)$.

orginally bit representation $= 16 \times 8 = 128$ bit.

after compression $= 14 \times 8 = 112$ bit.

2) S = 1111111111111111 00000000000000000000 1111

$= (15,1), (19,0), (4,1)$.

As maximum $= 19$. so 5 bit is needed to represent 19.

$\therefore (01111, 1), (10011, 0), (00100, 1)$.

$\therefore$ Compression ratio $= \frac{38}{18} = 2.11$.

→ RLE is highly suitable for FAX images.

→ RLE is also used for compression of images in the PCX format.

Scanned by CamScanner

Huffman coding:

→ variable length encoding algorithm.

→ dictionary based compression algorithm.

→ reconstruction এর জন্য same algorithm দরকার।

→ symbol গুলোকে probability অনুযায়ী descending order এ সাজাতে হবে। (বড় থেকে ছোট).

→ যার probability বেশি তাকে কম সংখ্যক bit দ্বারা representation করব।

→ The step of huffman coding Algorithm are given below:

i) Arrange the source symbols in a decreasing order of their probabilities.

ii) Take bottom two symbols and add ~~them~~ their probabilities. The branches are labeled as '0', '1'.

iii) Treate this sum of probabilities as new probability associated with new symbol.

Again two smallest probabilities are tied together and form a new probability.

Each time the total number of symbol is reduced by 1.

iv) Continue these until one probability is left.

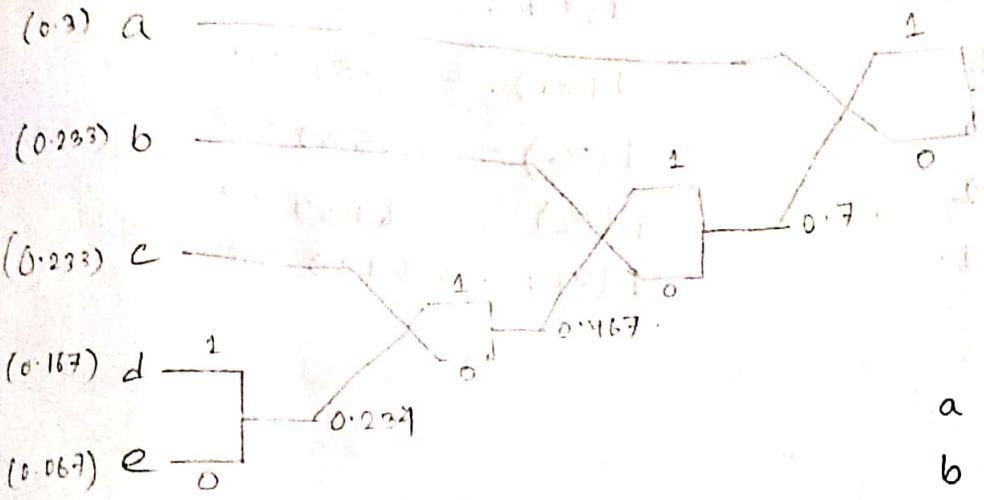
v) To findout prefix codeword, follow the branches from final node back to symbol.

Scanned by CamScanner

1. Given that.

$a = \frac{9}{30}$, $b = \frac{7}{30}$, $c = \frac{7}{30}$, $d = \frac{5}{30}$, $e = \frac{2}{30}$.

$a = 0.3$, $b = 0.233$, $c = 0.233$, $d = 0.167$, $e = 0.067$.



Codewords

a = 0

b = 10

c = 110

d = 1111

e = 1110

যদি একটা sequence দিয়ে বলে compress করতে,

b c c a.

10 110 110 0

এখন 101101100 কে decode করতে বললে,

10 110 110 0

b e e a.

যেহেতু এখানে codeword গুলোর length variable, তাই এটা variable length coding algorithm.

Scanned by CamScanner

Example - 1.13

$P(x_1) = 0.46$

$P(x_2) = 0.30$

$P(x_3) = 0.12$

$P(x_4) = 0.06$

$P(x_5) = 0.03$

$P(x_6) = 0.02$

$P(x_7) = 0.01$.

Self information. $I = -\log P(x)$.

$\therefore I(x_1) = 1.1203 \times P$

$I(x_2) = 1.737 \times P$

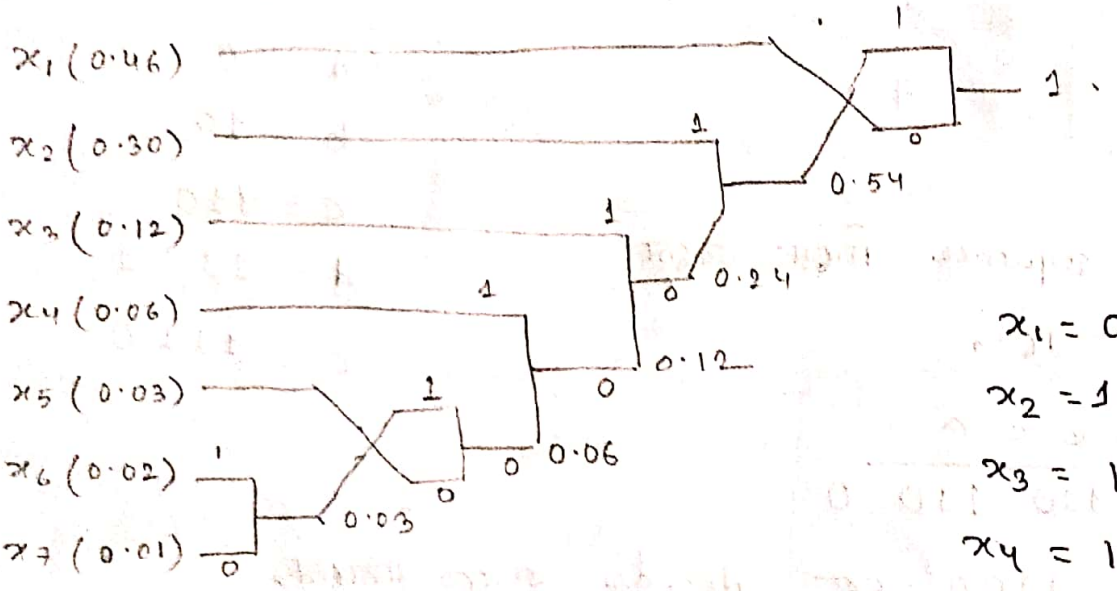
$I(x_3) = 3.0589 \times P$

$I(x_4) = 4.0589 \times P$

$I(x_5) = 5.0589 \times P$

$I(x_6) = 5.6439 \times P$

$I(x_7) = 6.6439 \times P$.



$x_1 = 0$

$x_2 = 11$

$x_3 = 101$

$x_4 = 1001$

$x_5 = 10000$

$x_6 = 100011$

$x_7 = 100010$

The entropy of the source is.

$$H(x) = -\sum_{k=1}^{7} P(x_k) \cdot \log_2 P(x_k).$$

$$= -(0.46 \times \log_2 0.46) - \{(0.30) \times \log_2 (0.30)\} - (0.03 \times \log_2 (0.03))$$
$$- (0.12 \times \log_2 0.12) - (0.06 \times \log_2 0.06)$$
$$- (0.02 \times \log_2 0.02) - (0.01 \times \log_2 (0.01)).$$

$$= 0.515 + 0.521 + 0.367 + 0.2435 + 0.1517 + 0.1128 + 0.066$$

$$= 1.9773 \text{ bits.}$$

Scanned by CamScanner

The average number of binary digit per symbol

n_k এর codeword এর length.

$$\bar{R} = \sum_{k=1}^{7} n_k P(x_k).$$

$$= 1(0.46) + 2(0.30) + 3(0.12) + 4(0.06) + 5(0.03) + 6(0.02) + 6(0.01)$$

$$= 1.99 \text{ bits}.$$

$\therefore$ efficiency of code, $\eta = \dfrac{1.9773}{1.99} = 0.9936.$

→ Huffmann coding isnot unique for a set of probabilities.

Arithmetic Coding:

→ one kind of variable length encoding.

→ haffman coding ও arithmatic coding এর difference:

haffmann coding converts a fixed no. of symbols into a variable length codeword.

Arithmatic coding convert a variable number of symbols into variable length codeword.

haffman → individual symbol individual bit দ্বারা represented হয়। Arithmatic → group of symbol কে একসাথে represent করা হয়।

Scanned by CamScanner

→ সূত্র: upper limit = lower limit + (difference × probability)

→ last output হবে lower limit টা।

→ ~~ex~~

1. Given that,

sequence is beda.

a = 0.4

b = 0.3

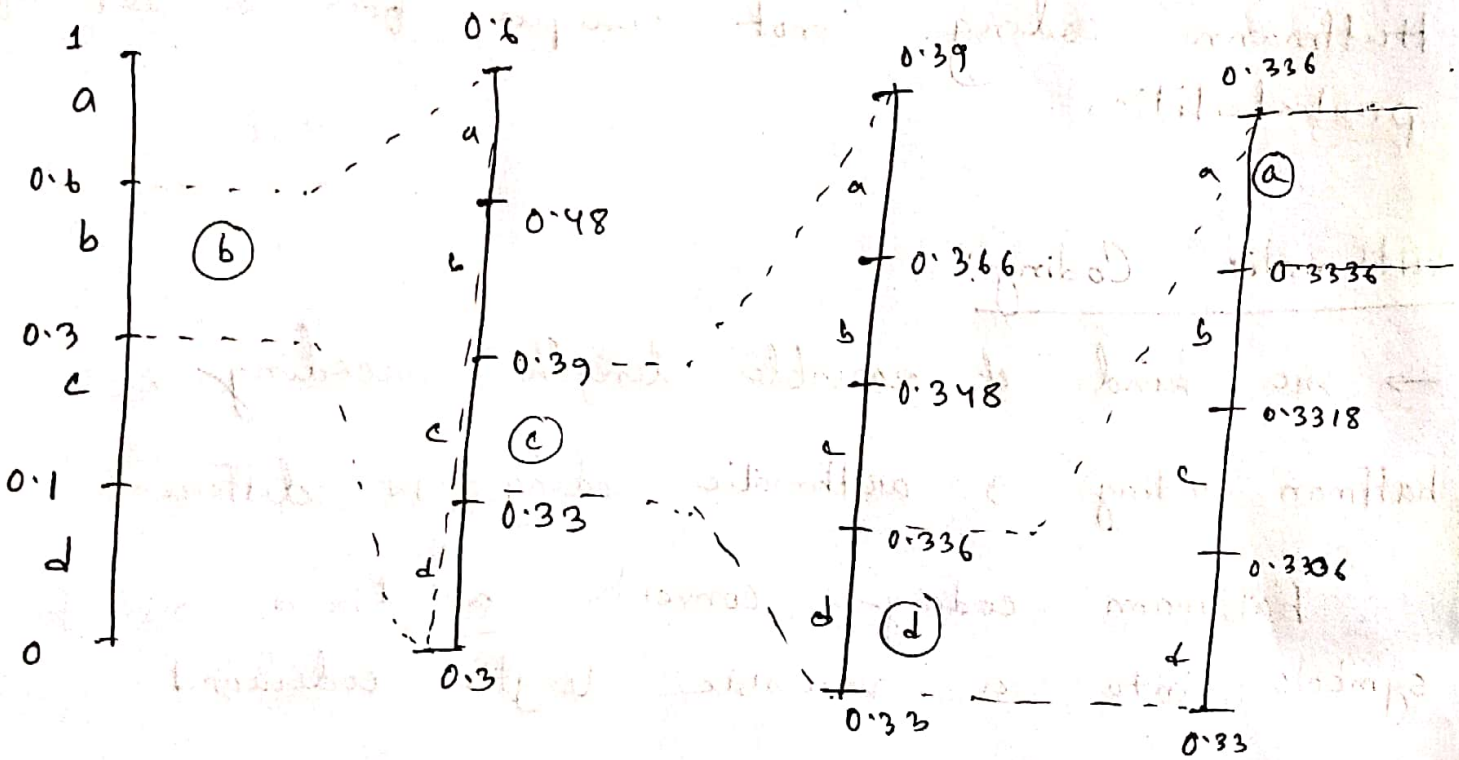
c = 0.2

d = 0.1.

descending order.

যখানে difference হল

(0.6 − 0.3) = 0.3



Ans: $(0.3336)_{10} = (.01010\ldots)$

$$\boxed{u = l + d \times P}$$

.3336
× 2
0.6672
× 2
1.3344
× 2
0.6688
× 2
1.3376
× 2
0.6752

a
c
b
d

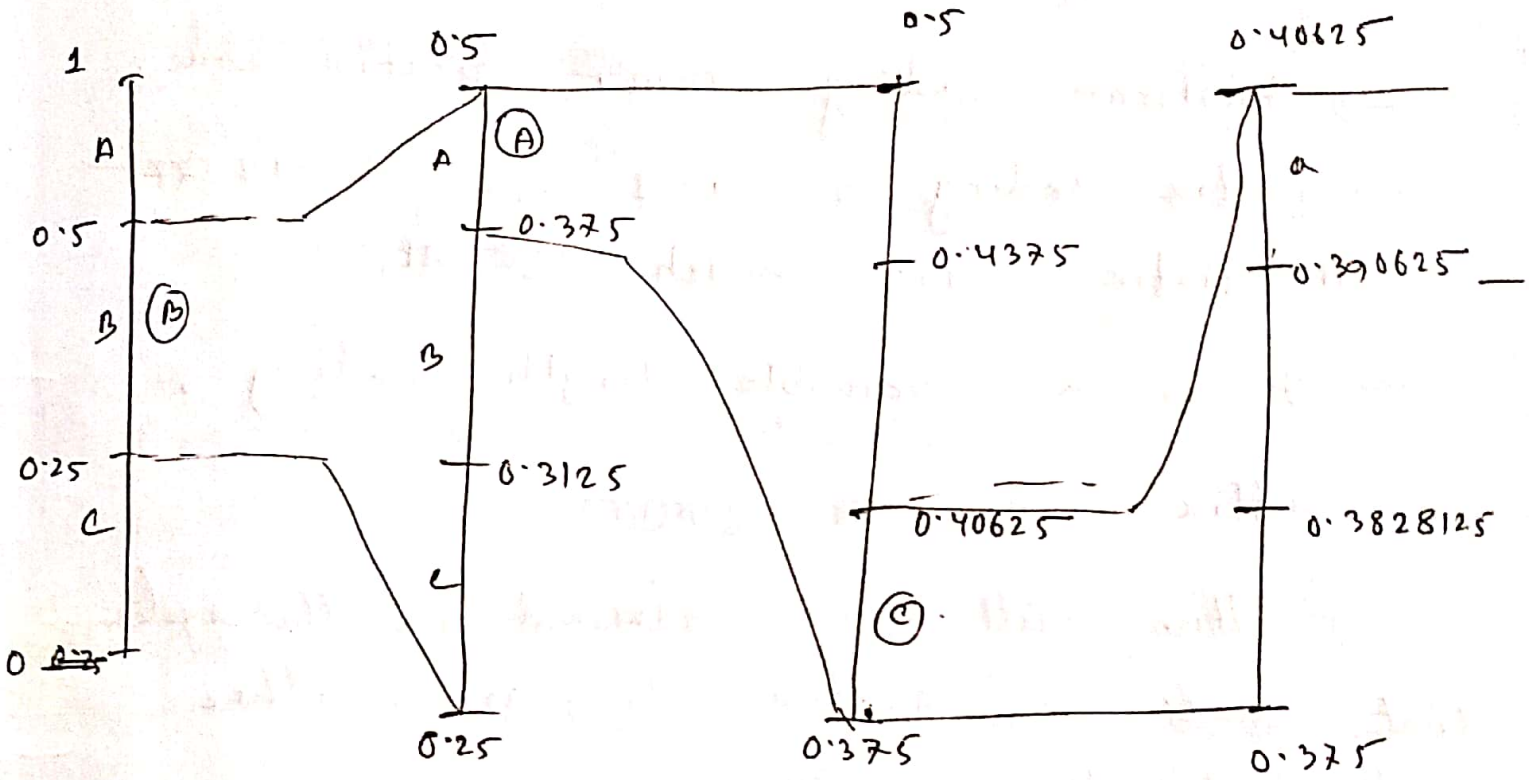
Scanned by CamScanner

symbol stream = BACA.

probability $P(A) = 0.5$

$P(B) = 0.25$

$P(C) = 0.25$.



Ans হবে 0.390625.

এখন binary convert করতে হবে।

ABC এর sequence A B C না হয়ে C B A হতে পারে।

Scanned by CamScanner

prefix code and Suffix Code:

There will no ~~col~~ codeword in the system that will be prefix to any other code of the system.

→ huffman coding হলো prefix code.

→ prefix coding এ a = 1 হলে, b = 12 হবে না। Prefix কখনও match হবে না।

→ It is a variable length coding.

→ Suffix code এর ক্ষেত্রে

There will no codeword in the system that will be suffix to any other code of the system.

~~দুইটা~~ codeword এর suffix কখনও same হবে না।

The Lempel-Ziv, Welch algorithm:

→ This algorithm doesnot need source statistics.

→ lossless compression.

→ It is prefix + variable length coding.

Scanned by CamScanner

Example: Code string: 1010110110101011

⇒ The first bit, a 1 has no ~~predd~~ predecessors. so, it has 'null prefix string'.

1, 010110110101011

1, 0, 10110110101011

1, 0, 10, 11, 01, 101, 010, 1011

AABABCDABBBACCCABCCDA

sequence টাকে uncommon part এ ভাঙবো,

position	1	2	3	4	5	6	7	8	9	10	11	12
Sequence	A	AB	ABC	D	ABB	B	AC	C	CA	BC	CD	A
Representation	0,1	1,2	2,3	0,4	2,2	0,2	1,3	0,3	8,1	6,3	8,4	1,1

representation = previous position + last symbol এর index.

Given, Index
A = 1
B = 2
C = 3
⋮
Z = 26.

এখানে problem হল কোনটা symbol এর index, কোনটা symbol এর position তা বুঝা যায় না, তাই Indexing করা হয়,

Scanned by CamScanner

Index	1	27	28	4	29	2	30	3	31	32	33	1
Sequence	A	AB	ABC	D	ABB	B	AC	C	CA	BC	CD	A
Output		1	27	28	4	29	2	30	3	31	32	33

output 1 ঘর shift করা হয়।
Then binary তে convert করা হয়।

Index

A = 1
B = 2
C = 3
⋮
Z = 26
AB = 27
ABC = 28
ABB = 29
AC = 30
CA = 31
BC = 32
CD = 33.

Normal Distribution:

A random variable x is said to be normal gaussian distributed with mean μ and variance σ^2 if its probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad [-\infty < x < \infty].$$

The normal or gausian distribution of x is usually represented by $X \sim N(\mu, \sigma^2)$ or also $X \sim N(x-\mu, \sigma^2)$ or $X \sim (\mu, \sigma)$.

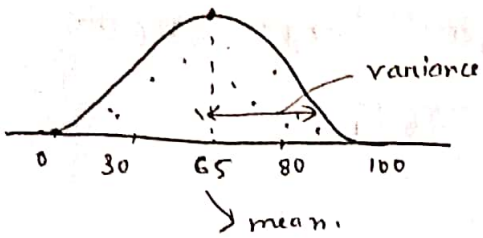
Properti

1) Ex
random

2) Va

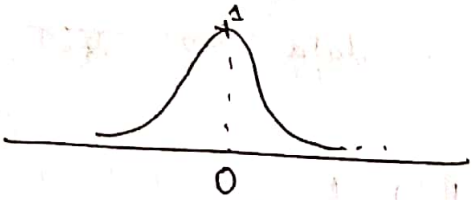
3) Th
a
aleou
for

Scanned by CamScanner



envelope এর peak হল average value.

normal distribution.



equivalent standard normal distribution.

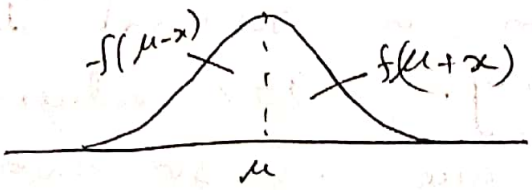
Properties of Gaussian Distribution:

1) Expected value, $E(x) = \mu$ for a normal random variable x. Again, $\mu = E(x) = \int_{-\infty}^{\infty} x f(x) . dx$

2) Variance $V(x) = \sigma^2 = E[(x-\mu)^2]$.

$$= \int_{-\infty}^{\infty} x^2 . f(x) . dx.$$

3) The probability density function f of a normal random variable x is symmetrical about the mean; normally $f(\mu - x) = f(\mu + x)$ for all real $x \in \mathbb{R}$.



Scanned by CamScanner

mean fixed রেখে তার রেখে deviation বাড়লে graph flat হবে।

Here, $\sigma_3^2 > \sigma_2^2 > \sigma_1^2$, data তত ছড়িয়ে যাবে।

Standard Normal Distribution:

A normal distribution with a mean of 0 and a standard deviation of 1 is called standard normal deviation.

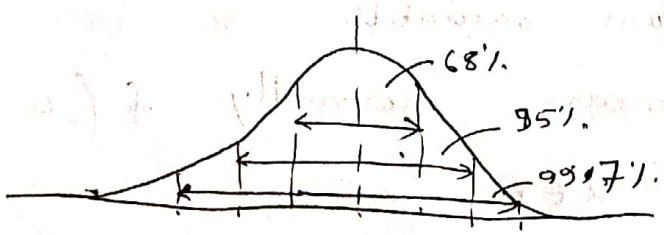
→ also known as gaussian distribution.

The cumulative probability distribution of the standard normal distribution is $P(Z \le z)$

random variable.

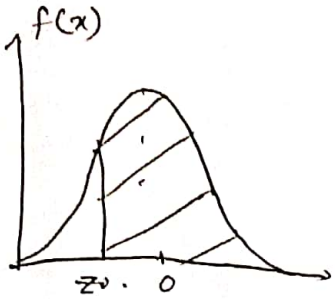
$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.$$

[3 sigma rule].

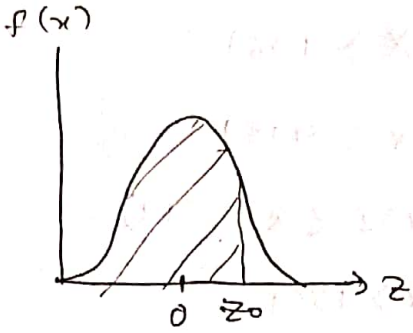


68% values drawn from a normal distribution are in one σ away. 95% values ... 2σ away. 99.7% values are ... 3σ away.

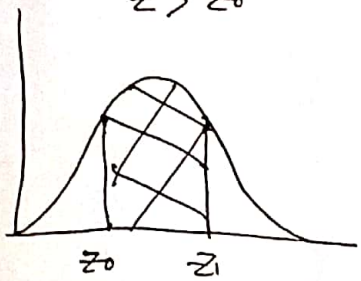
Scanned by CamScanner



$z > z_0$



$z < z_0$.



$P(z_0 \leq z < z_1) = P(z \leq z_1) - P(z \leq z_0)$.

$P(z > z_0) = P(z < z_0)$.

If X is a normal random variable with mean μ and standard deviation σ, then,

$$z = \frac{x-\mu}{\sigma}$$ is a standard normal random variable. This meant that to calculate $P(X \leq x)$ is the same as calculating $P\left(z \leq \frac{x-\mu}{\sigma}\right)$.

<u>Example:</u> The actual volume of soup in 500 ml jars follows a normal distribution with mean 500 ml and variance 16 ml. If X denotes the actual volume of soup, in a jar, what is

Scanned by CamScanner

i) $P(X > 496)$

ii) $P(X < 498)$

iii) $P(492 < x < 512)$.

iv) $P(X > 480)$.

Solution:

Here, $\mu = 500$ ml.

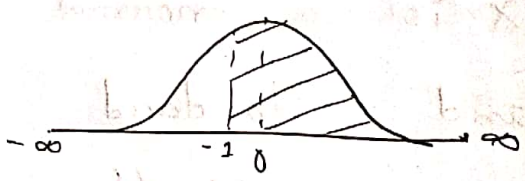
$\sigma^2 = 16$ ml.

i) We know,

$$z = \frac{x - \mu}{\sigma}.$$

$$= \frac{496 - 500}{\sqrt{16}} = -1.$$

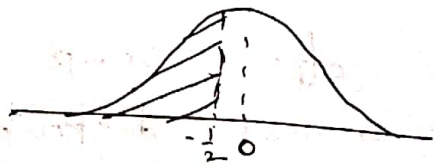
$\therefore P(z > -1)$.



This is the distribution of $P(z > -1)$

$P(X > 496)$.

ii) $$z = \frac{498 - 500}{\sqrt{16}} = -\frac{1}{2}.$$



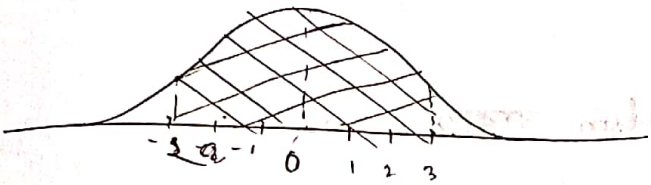
$P(z < -\frac{1}{2})$. $P(X < 498)$.

Scanned by CamScanner

iii) $P(492 < X < 512)$

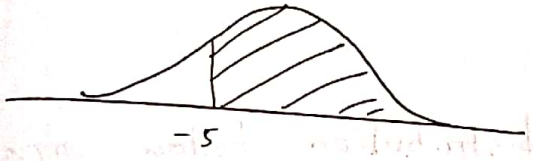
$$z_1 = \frac{492 - 500}{\sqrt{16}} = \frac{-8}{4} = -2.$$

$$z_2 = \frac{512 - 500}{\sqrt{16}} = \frac{12}{4} = 3.$$



iv) $P(X > 480)$.

$$z = \frac{480 - 500}{\sqrt{16}} = \frac{-20}{4} = -5.$$



$P(z > -5)$.

$e^{-z^2} = \sqrt{\pi}$

Now,

$$P(z > -1) = \int_{-1}^{\infty} f(z) \cdot dz.$$

$$= \int_{-1}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}z^2} \cdot dz.$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-1}^{\infty} e^{-\frac{1}{2}z^2} \cdot dz.$$

$$= \frac{1}{\sqrt{2\pi}} \left[4z e^{-\frac{1}{2}z^2} \right]_{-1}^{\infty}.$$

Scanned by CamScanner

$$= \frac{1}{\sqrt{2\pi}} \cdot \left[4\infty - e^{-\infty} - 4e^{-\frac{1}{2}} \right]$$

$$= \frac{1}{\sqrt{2\pi}} \left[0 - 0.606 \right]$$

$$= -\frac{.606}{\sqrt{2\pi}} .$$

output বাহির Integration করতে হবে.

$$P(z > z_0) = \int_{z_0}^{\infty} f(z) \cdot dz .$$

$$P(z < z_0) = \int_{-\infty}^{z_0} f(z) \cdot dz .$$

Noise gaussian distribution follow করে,

$$X \xrightarrow{\downarrow x_0} X + x_0 \qquad 1 - 0.01 \leftarrow 1\% \text{ error.}$$

$$X \xrightarrow{\downarrow x_1} X + x_1 \qquad 1 + 0.01$$

$$X \xrightarrow{\downarrow x_2} X + x_2 .$$

recently 10^{-9} bit error = standard rate.
1 gb data পাঠালে একটা bit error হয়।

There are two special sources.
1) memoryless source.
2) Markov source / source with memory,

Scanned by CamScanner

Memoryless Source:

Each simple independent of the previous ones

$$P(x_1, x_2, x_3 \cdots x_n) = P(x_1) \cdot P(x_2) \cdot \cdots P(x_n).$$

→ total probability = individual probability এর multiplication

Source with memory:

Each symbol depends on the previous ones

$$P(x_1, x_2, x_3 \cdots x_n) = P(x_1) \cdot P\left(\frac{x_2}{x_1}\right) \cdot P\left(\frac{x_3}{x_2}\right) \cdots P\left(\frac{x_n}{x_{n-1}}\right).$$

→ total probability = individual conditional probability এর multiplication.

Gaussian Channel:

→ Gaussian channel follow gaussian distribution.

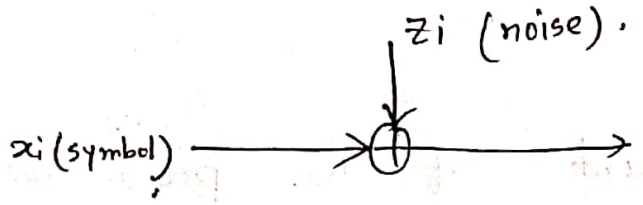
The noise Z_i with variance N, then,

$$Y_i = X_i + Z_i \longrightarrow Z_i \sim N(0, N).$$

If any codeword $(x_1, x_2, \cdots x_n)$ transmitted over the channel, then average power required,

$$\frac{1}{n} \sum_{i=1}^{n} x_i^2 \leq P.$$

Scanned by CamScanner



Here, z_i can be either additive or substract

- If ~~noise variance~~, $N = 0$, then the variance is zero, then the noise is zero.

  The receiver will receive the signal perfectly.

- Noise Power is not constant. Each has different power.

  The noise can be either removed or not. It depends on the power of codeword

  i) $z_i > x_i$ → noise power is greater than signal power. So, noise cannot be removed.

  ii) $x_i > z_i$ → noise power is less than signal power. So, noise cannot be removed.

Signal power বাড়ালে noise easily remove করা যাবে।

Scanned by CamScanner

But the power of signal x_i cannot be increased randomly. The average power is.

$$\frac{1}{n} \sum_{i=0}^{N} x_i^2 \leq P.$$

Here, P is power constant.

In gaussian distribution, mean $= 0$.

variance $= \infty$.

If x_i signal is repeated, then

$$\frac{z_1 + z_2 + z_3 \cdots\cdots + z_n}{n} = 0$$

because the mean $= 0$, So. the noise power will be zero. Noise will be removed if several repeated signals are send instead of one.

Error Probability of Gaussian Channel:

If 1 bit is send over the channel So, there are two possible values, either 1 or 0. with two power levels. $+\sqrt{P}$ or $-\sqrt{P}$. Both levels are equally likely.

Scanned by CamScanner

The optimum rule is to decide that $+\sqrt{P}$ was sent if $y>0$ and decide $-\sqrt{P}$ was sent if $y<0$.

The probability of error is

$$P_e = \frac{1}{2} P_r\left(y<0 \mid x=+\sqrt{P}\right) + \frac{1}{2} P_r\left(y>0 \mid x=-\sqrt{P}\right).$$

$$= \frac{1}{2}\left(z<-\sqrt{P} \mid x=+\sqrt{P}\right) + \frac{1}{2} P_r\left(z>\sqrt{P} \mid x=-\sqrt{P}\right).$$

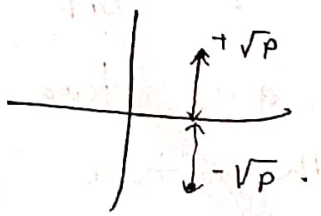
$$= P_r\left(z>\sqrt{P}\right).$$

$$= 1-\phi\left(\frac{\sqrt{P}}{N}\right).$$

where $\phi(x)$ is the cummulatative normal function.

$$\phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \, dt.$$

By following this theorem, continuous channel can be represented as discrete channel.



Scanned by CamScanner

Cover:

The information capacity of the gaussian channel with power constant P is.

$$\max_{P(x):\ E_x^2 \le P} I(X;Y).$$

Here, $I(X;Y) = h(y) - h(y/x)$

$= h(y) - h\left(\frac{x+z}{x}\right)$. $[Y = X+Z]$.

$= h(Y) - h(Z/X)$.

$= h(Y) - h(Z)$. [কারণ Z ; X এর উপর dependent না].

Z is independent of X, then entropy of $h(Z) = \frac{1}{2}\log 2\pi e N$.

Also, $E_y^2 = E(X+Z)^2$.

$= E_x^2 + 2E_x \cdot E_z + E_z^2$.

$$\boxed{\begin{aligned} E_z &= \frac{+1-1}{2} \\ &= 0 \\ E_z^2 &= (-1)^2 + (1)^2 \\ &= 2. \end{aligned}}$$

If $E_z = 0$, $E_y^2 = P+N$

entropy of y is bounded by $\frac{1}{2}\log 2\pi e(P+N)$.

Now, $I(X;Y) = h(y) - h(z)$

$\le \frac{1}{2}\log 2\pi e(P+N) - \frac{1}{2}\log 2\pi e N$.

$= \frac{1}{2}\log(1 + P/N)$.

$$C = \max_{E_x^2 < P} I(X;Y) = \frac{1}{2}\log(1 + P/N).$$

[where $X \sim N(0, P)$].

Scanned by CamScanner

where, P = signal power.

$N = E_z^2$ = noise power.

E_z হল noise voltage. এটা 0 হতে পারে।
E_z^2 হল noise power. এটা 0 হবে না।

Channel capacity depends on power. If $\frac{\text{power}}{\text{noise}}$ or SNR increases, then the capacity also increase.

$P = \infty$ হলে $C = \infty$

কিন্তু channel গুলো band limited.

Now, there are 2ω samples in signal.

$\therefore$ power per sample $= \frac{P_T}{2NT} = \frac{P}{2\omega}$.

$\therefore$ noise variance, $\frac{N_0}{2}$ per sample.

$$\left[\frac{T}{2\omega T} = \frac{N_0}{2}\right].$$

Then,

$$C = \frac{1}{2}\log\left(1 + \frac{P/2\omega}{N_0/2}\right).$$

Scanned by CamScanner

$$C = \frac{1}{2}\log\left(1+\frac{P}{N_0 W}\right) \text{ bits per sample.}$$

This is the maximum no. of bits that a symbol can integrate.

$\therefore$ As there are $2W$ samples, then.

$$2W \times \frac{1}{2}\log\left(1+\frac{P}{N_0 W}\right) = W\log\left(1+\frac{P}{N_0 W}\right).$$

$\rightarrow$ band limited gaussian channel.

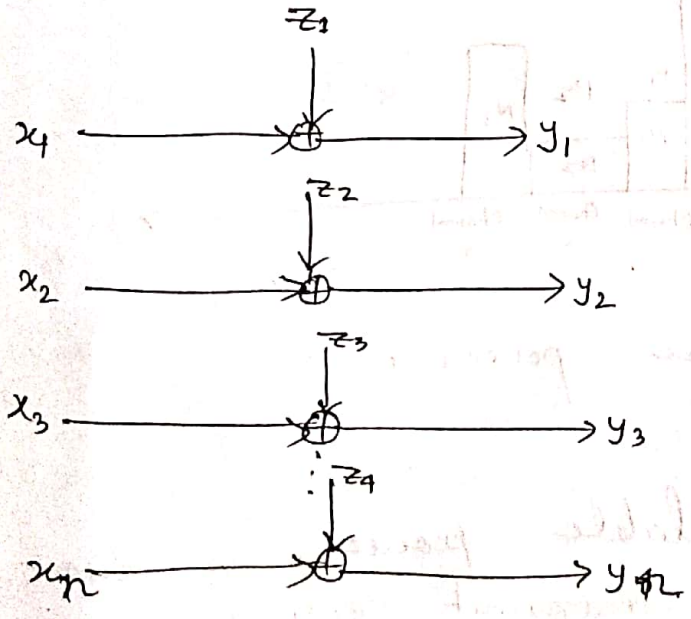
parallel Gaussian Channel:

$\rightarrow$ bits are send parallely.

Then,

$$\underset{f(x_1, x_2, x_3 \cdots x_k)\sum E x_i^2 \le P}{C = \max} I(x_1, x_2, x_k \cdots Y_1, Y_2 \cdots Y_k).$$

$$\le \sum_{i=0}^{n} \frac{1}{2}\log\left(1+\frac{P_i}{N_i}\right).$$



This is the parallel gaussian channel.

Scanned by CamScanner

[illegible] channel [illegible] [illegible] total power [illegible] [illegible] distribute [illegible]?

Here, $P_i = E[x_i^2]$ and $\sum P_i = P$.

Equality is achieved by $x_1, x_2, \ldots x_k^2 \sim N\left(0, \begin{bmatrix} P_1 & 0 & 0 & \cdots & 0 \\ 0 & P_2 & 0 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & P_k \end{bmatrix}\right)$

The problem of power allotment optimized by lagrange

$$J(P_1 \ldots P_k) = \sum \frac{1}{2} \log\left(1 + \frac{P_i}{N_i}\right) + \lambda\left(\sum P_i\right)$$

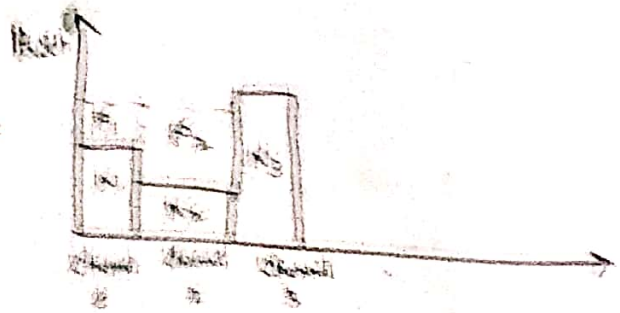
Differentiating with respect to P_i, we have $\frac{1}{2}\frac{1}{P_i + N_i} + \lambda = 0$ (1) and $P_i = -N_i$

Hence solution violation to verify the solution. The solution

$$P_i = (\nu - N_i)^+,$$

maximum capacity when ν is

$$\sum (\nu - N_i)^+ = P.$$



Here, V = total power

P_i = the power for each sample

= total power − noise power.

* The total available power is same [illegible] power transmit করা যাবে না।

→ In parallel channel, this power can be distributed.

→ In this figure, If there is an error in P_1, then the symbol cannot be extracted. In that case, the rest power of P_1 is distributed ~~among~~ between two other channels.

Water filling Process:

Here, channel three has more noise than the two others and the signal power is less. So, this signal power is distributed between two other channels. This is known as water-filling process.

$$(x)^* = \begin{cases} x ; & \text{if } x \geq 0 \\ 0 ; & \text{if } x \leq 0. \end{cases}$$

$y = h * x + z.$

$Y = HX + Z.$ → This is frequency representation

If this signal is transmitted for repeated times then $H = \frac{Y}{X}$ [where $Z = 0$; repeated signal এর জন্য].

Scanned by CamScanner

where, H = system's transfer function

[4] 1mV signal transmit করলে 0.001mV হয়ে যায় বা কমে যায়, এটা system response. এজন্য বিট trial bit পাঠানো যায় শুধু 3 বারের পাঠানো হয়।

Scanned by CamScanner

Information Theory.

Two fundamental things of communication theory.

1) entropy (H)
2) Channel capacity (c).

Information: Information is any entity or form that resolves uncertainity or provides the answer to a question of some kind.

purpose of Information theory:

A communication engineer has to fight with

→ The measure of source info.
→ The info capacity of the channel.
→ Coding.

If the rate of info from a source doesnot exceed the capacity of the channel, then there exist a coding scheme such that info can be transmitted over the communication channel with small amount of error despite the presence of noise. Information theory helps to deal all of these.

Scanned by CamScanner

Entropy: The entrophy is a measure of the average uncertainity in the random variable.

The entropy of a random variable X with a probability mass function $P(x)$ is

$$H(x) = -\sum P(x) \cdot \log_2 P(x).$$

Mutual Information for two random variabl X and Y is.

$$I(x,y) = H(x) - H(x|y).$$

$$= \sum_{x,y} P(x,y) \cdot \log \frac{P(x,y)}{P(x) \cdot P(y)}.$$

→ It measure the dependence between the two random variables.

Channel capacity:

→ The rate of communication of information over a channel.

Relative entropy $D(P||q)$ is a measure of the distance between two probability mass functions P and q.

$$D(P||q) = \sum_x P(x) \cdot \log \frac{P(x)}{q(x)}.$$

Scanned by CamScanner

Chapter - 2

Entropy: The entropy $H(x)$ of a discrete random variable X is defined by

$$H(x) = -\sum_{x \in X} P(x) \cdot \log P(x) = E_p \log \frac{1}{P(x)}$$

$$\therefore E_p = \sum_{x \in X} P(x)$$

→ measure of the uncertainity of a random variable.

→ যখনে log এর base 2 হলে unit bit হবে, যদি log এর base e হয়, তাহলে unit হবে nats. decimal system এর জন্য unit ~~to~~ decit.

→ entropy depends only on the probability not the actual value.

→ If the information is I, then.

$$I = \log \frac{1}{P(x)}$$

probability কম হয়, information বেশি হয়।

→ In case of independent events,

$$P = P_1 P_2 P_3 \longrightarrow \text{(multiplication)}.$$

$$I = \log \frac{1}{P} = \log \frac{1}{P_1 P_2 P_3}$$

$$= \log\left(\frac{1}{P_1}\right) + \log\left(\frac{1}{P_2}\right) + \log\left(\frac{1}{P_3}\right).$$

$$= I_1 + I_2 + I_3 \quad \text{(summation)}.$$

Scanned by CamScanner

$\therefore$ information, $I = -\sum_n \log P_n$.

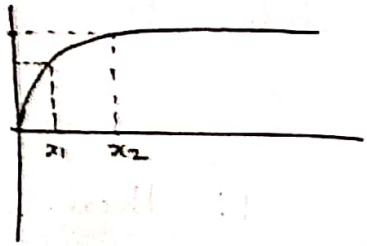
This is the measurement of information.

- entropy ↑ information ↑. probaleility ↓.

- In $H(x) = -\underbrace{\sum p(x)}_{\text{expectation}} \cdot \underbrace{\log p(x)}_{\text{information / marginal entropy}}$.

- probaleility 1 হলে information 0.
  probaleility 0 হলে information infinite.

- যার value বেশি, তার probaleility বেশি.



Here, $x_2 > x_1$

$\therefore P_2 > P_1$

- probaleility 3 information কখনও negative হয় না

- $H(x) \geq 0$. (property)

We know, range of probaleility is $0 \leq P(x) \leq 1$

So, $\log \frac{1}{P(x)} = -\log P(x)$.

$= -\log(1) = 0$. [for $P(x) = 1$]

for any value $P(x) \leq 1$.

$\log \frac{1}{P(x)} \geq 0$.

$\therefore H(x) \geq 0$.

Scanned by CamScanner

base conversion: $H_b(x) = (\log_b a) \cdot H_a(x)$. (property).

We know,

$$\log_b P = \log_b a \cdot \log_a P.$$

$$H_b(x) = (\log_b a) \cdot H_a(x).$$

[Prooved].

→ entropy can be changed from one base to another in this way.

If the QPSK modulated signal is

00 01 10 10 11 11 01 01 10 00 10 10 11 01 10

$P_{00} = \frac{2}{15}$ $P_{10} = \frac{6}{15}$

$P_{01} = \frac{4}{15}$ $P_{11} = \frac{3}{15}$.

So, the entropy is

$$H(x) = -\sum p(x) \cdot \log p(x).$$

$$= -P_{00} \cdot \log P_{00} - P_{01} \cdot \log P_{01} - P_{10} \cdot \log P_{10} - P_{11} \log P_{11}$$

$$= -\frac{2}{15} \log_2 \frac{2}{15} - \frac{4}{15} \log \frac{4}{15} - \frac{6}{15} \log \frac{6}{15} - \frac{3}{15} \log \frac{3}{15}.$$

$$= 1.899 \text{ bit}.$$

Entropy 1 এর কাছাকাছি বড় হলে তাতে information থাকবে।

Scanned by CamScanner

Another way,

$$H(x) = \frac{I_1 + I_2 + I_3 + I_4}{4}$$

$$I_1 = -\log \frac{2}{15}$$

$$I_2 = -\log \frac{4}{15}$$

$$I_3 = -\log \frac{6}{15}$$

$$I_4 = -\log \frac{3}{15}$$

$$= \frac{8.457637381}{4}$$

$$= 2.11 \text{ bits.}$$

Example:

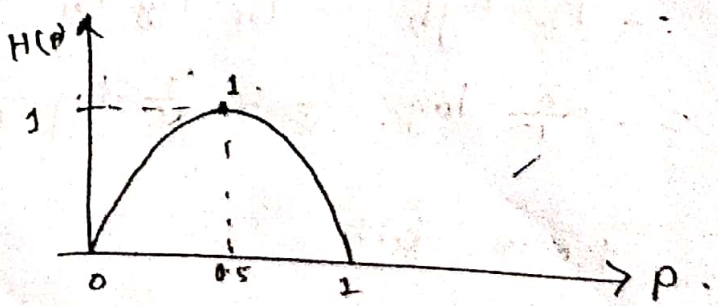
for BPSK signal.

$$x = \begin{cases} 1 & ; \text{ probability } = p \\ 0 & ; \text{ probability } = 1-p. \end{cases}$$

So, $H(x) = -\sum_{x=0}^{1} p(x) \log p(x)$,

or, $\boxed{H(p) = -p \log p - (1-p) \cdot \log (1-p).}$ defination of $H(p)$

when $p = \frac{1}{2}$, $H(x) = 1$.



→ entropy is concave function.

→ when $p = 0, 1$, there is no uncertainity. entropy $= 0$.

→ when $p = 0.5$, entropy is maximum
uncertainity is maximum.

Scanned by CamScanner

Example · 212

$$X = \begin{cases} a & \text{with probability } \frac{1}{2} \\ b & \text{with probability } \frac{1}{4} \\ c & \text{" " } \frac{1}{8} \\ d & \text{" " } \frac{1}{8} \end{cases}$$

Entropy · $H(x) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{4}\log_2\frac{1}{4} - \frac{2}{8}\log_2\frac{1}{8}$

$= \frac{7}{4}$ bits $= 1.75$ bits.

কখন সবচেয়ে বেশি info পাব?

probability 0 $\longrightarrow$ info ∞ $\rightarrow$ expectation = 0

probability 50% $\longrightarrow$ info max $\rightarrow$ expectation = max.

probability 100% $\longrightarrow$ Info : 0 $\rightarrow$ expectation = 0,

অর্থাৎ সবগুলোর probability যখন সমান হয়, তখন $H(x) = 2$ bit পাব. [max information].

$\rightarrow$ probability is cumulative (যোগফল 1)

$\rightarrow$ for same thing - $\frac{\Sigma x + \Sigma y}{n}$ for different things $\frac{\Sigma x + \Sigma y}{2n}$

$\rightarrow$ Entropy is the expectation of information.

$\rightarrow$ এই example এ সবগুলোর probability 1/4 হলে, entropy $H(x) = 2$ bit. অর্থাৎ information maximum হয়।

$\rightarrow$ 8PSK এর ক্ষেত্রে, প্রত্যেকের probability $\frac{1}{8}$ হলে, expectation অথবা information হবে 3 bit.

Scanned by CamScanner

Joint Entropy and Conditional Entropy:

Joint Entropy: The joint entropy $H(x,y)$ of a pair of discrete random variable (x,y) with joint distribution $P(x,y)$ is defined as

$$H(x,y) = -\sum_{x\in X}\sum_{y\in Y} P(x,y) \cdot \log P(x,y).$$

$$\text{or, } H(x,y) = -E \log P(x,y).$$

Conditional entropy:

The conditional entropy of a random variable given another, as the expected value of the entropies of the conditional distributions, averaged over conditioning random variable.

If $H(X|Y) \sim P(x,y)$, the conditional entropy $H(Y|X)$ is defined as

$$H(Y|X) = \sum_{x\in X} P(x) \cdot H(Y|X=x).$$

$$\cancel{= -\sum_{x\in X}\sum_{y\in Y} P(x,y)}$$

$$= -\sum_{x\in X} P(x) \cdot \sum_{y\in Y} P(y|x) \cdot \log P(y|x).$$

Scanned by CamScanner

$$= -\sum_{x \in X} \sum_{y \in Y} p(x,y) \cdot \log p(y|x) \quad | \text{ Bayes Rule } |$$

$$= -E_{p(x,y)} \log p(y|x)$$

Theorem 2.2.1 (Chain rule).

∴ The entropy of a pair of random variables is the entropy of one plus conditional entropy of the other. (Statement).

$$H(X,Y) = H(X) + H(Y|X).$$

(relation between joint entropy and conditional entropy).

<u>proof:</u>

$$H(X,Y) = -\sum_{x \in X} \sum_{y \in Y} p(x,y) \cdot \log p(x,y).$$

$$= -\sum_{x \in X} \sum_{y \in Y} p(x,y) \cdot \log p(x) \cdot p(y|x).$$

$$\frac{P(A,B)}{P(B)} = P(A|B) \quad = -\sum_{x \in X} \sum_{y \in Y} p(x,y) \cdot \log p(x) - \sum_{x \in X} \sum_{y \in Y} p(x,y) \log p(y|x)$$

$$= -\sum_{x \in X} p(x) \cdot \log p(x) - \sum_{x \in X} \sum_{y \in Y} p(x,y) \cdot \log p(y|x)$$

$$= H(X) + H(Y|X).$$

$$\therefore \underbrace{H(X,Y)}_{\text{joint entropy}} = \underbrace{H(X)}_{\text{marginal entropy}} + \underbrace{H(Y|X)}_{\text{Conditional entropy}}.$$

So, we can write,

log p(x|

Scanned by CamScanner

So, we can write :

$$\log p(X,Y) = \log p(X) + \log p(Y|X).$$

~~Coll~~ Corollary :

$$H(X,Y|Z) = H(X|Z) + H(Y|X,Z).$$

Example. 2.2.1 .

(X,Y) have following joint distribution .

$y \backslash x$	1	2	3	4
1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$
2	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$
3	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
4	$\frac{1}{4}$	0	0	0

যখন joint probability দেয়া থাকে তখন সবগুলোর যোগফল হবে marginal probability .

$$H(X) = H\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\right) \quad \text{(Colom যোগ)}$$

$$H(Y) = H\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right) \quad \text{(row যোগ)}$$

$$\therefore H(X) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{4}\log\frac{1}{4} - \frac{2}{8}\log\frac{1}{8}$$

$$= \frac{7}{4} \text{ bits.}$$

$$H(Y) = -4 \times \frac{1}{4}\log\left(\frac{1}{4}\right)$$

$$= 2 \text{ bits.}$$

Scanned by CamScanner

Also,

$$\frac{H(x,y)}{H(y)} = H(x$$

$$H(X|Y) = \sum_{i=1}^{4} P(Y=i)\, H(X|Y=i)$$

$$= \frac{1}{4} H\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\right) + \frac{1}{4} H\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{8}, \frac{1}{8}\right)$$

$$+ \frac{1}{4} H\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right) + \frac{1}{4} H(1, 0, 0, 0)$$

$$= \frac{1}{4} \times \frac{7}{4} + \frac{1}{4} \times \frac{7}{4} + \frac{1}{4} \times 2 + \frac{1}{4} \times 0.$$

$$= \frac{7}{16} + \frac{7}{16} + \frac{1}{2}$$

$$= \frac{7+7+8}{16} = \frac{22}{16} = \frac{11}{8} \text{ bits.}$$

$$H(Y|X) = \sum_{i=1}^{4} P(X=i) \cdot H(Y|X=i),$$

$$\frac{H(x,y)}{H(x)} = H($$

$$= \frac{1}{2} H\left(\frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{1}{2}\right) + \frac{1}{4} H\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}, 0\right) +$$

$$\frac{2}{8}\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}, 0\right)$$

Here,

$$H\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}, 0\right) = -\frac{2}{4} \log \frac{1}{4} - \frac{1}{2} \log \frac{1}{2}$$

$$= \frac{3}{2}$$

$$H\left(\frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{1}{2}\right) = -\frac{1}{4} \log \frac{1}{4} - \frac{2}{8} \log \frac{1}{8} - \frac{1}{2} \log \frac{1}{2}$$

$$= \frac{7}{4}.$$

Scanned by CamScanner

$$H(Y|X) = \frac{1}{2} \times \frac{7}{4} + \frac{1}{4} \times \frac{3}{2} + \frac{2}{8} \times \frac{3}{2}$$

$$= \frac{7}{8} + \frac{3}{8} + \frac{3}{8}$$

$$= \frac{6+7}{8} = \frac{13}{8} \text{ bits.}$$

$$\therefore H(X,Y) = H(X) + H(Y|X)$$

$$= \frac{7}{4} + \frac{13}{8} = \frac{27}{8} \text{ bits}$$

or, $H(X,Y) = H(Y) + H(X|Y)$

$$= 2 + \frac{11}{8} = \frac{27}{8} \text{ bits. (same).}$$

So,

$$H(Y|X) \neq H(X|Y).$$

$$H(X) - H(X|Y) = H(Y) - H(Y|X).$$

or. $H(X) + H(Y|X) = H(Y) + H(X|Y) = H(X,Y)$.

Bayes Rule:

$$P(x|y) = \frac{P(x,y)}{P(y)}$$

$$P(A|B) = \frac{P(A,B)}{P(B)}.$$

$$P(A,B) = P(B) \cdot P(A|B).$$

$$P(A,B) = P(A) \cdot P(B|A).$$

Scanned by CamScanner

$P(x|y=0)$, $y=0$ এর probability জানি, এর সাপেক্ষে x এর probability বের করব।

এখন $(x=0 \mid y=0) = 0.2$

$(x=1 \mid y=0) = 0.8$.

এর probability বেশি, receiver এর অনুযায়ী decis নিবে।

channel এর ক্ষেত্রে conditional entropy use করে bi alter করে কিন্তু info বাড়ানো যায়।

Relative Entropy and Mutual Information:

Relative entropy is a measure is a measure of the distance between two distri

The Relative entropy is the distance betwe two probability mass function $P(x)$ and $q(x)$ which is defined as

$$D(P\|q) = \sum_{x \in X} P(x) \cdot \log \frac{P(x)}{q(x)}.$$

$$= E_P \cdot \log \frac{P(X)}{q(X)}.$$

if $p=q$. then Relative entropy is zero

Mutual Information:

→ It is the measure of the amou of information that one random..

Scanned by CamScanner

variable contains about another random variable.

→ The mutual information $I(X;Y)$ is the relative entropy between the joint distribution and product distribution $P(x)$, $P(Y)$.

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} P(x,y) \cdot \log \frac{P(x,y)}{P(x) \cdot P(Y)}.$$

$$= D\left(P(x,Y) \,\|\, P(x) \cdot P(Y)\right).$$

$$= E_{P(x,y)} \cdot \log \frac{P(X,Y)}{P(X)\,P(Y)}.$$

The relative entropy $D(p\|q)$ is a measure of the inefficiency of assuming that the distribution is q when true distribution is p.

<u>Example - 2.3.1</u>

$$X = \{0, 1\}.$$

Let, $P(0) = 1-r$

$P(1) = r$

$q(0) = 1-s$

$q(1) = s$

$r = s$ হলে এর entropy or সবচেয়ে কমবে।

$$D(p\|q) = \sum p(x) \log \frac{p(x)}{q(x)}$$

$$\therefore D(p\|q) = (1-r) \cdot \log \frac{1-r}{1-s} + r \cdot \log \cancel{\log} \frac{r}{s}$$

$$D(q\|p) = (1-s) \cdot \log \frac{1-s}{1-r} + s \log \frac{s}{r}.$$

Scanned by CamScanner

if $r = s$, transmitted & received signal [illegible] are totally overlapped.

then $D(p\|q) = D(q\|p) = 0$.

If $r = \frac{1}{2}$, $s = \frac{1}{4}$.

$$D(p\|q) = \frac{1}{2}\log\frac{\frac{1}{2}}{\frac{3}{4}} + \frac{1}{2}\log\frac{\frac{1}{2}}{\frac{1}{4}}$$

$$= \frac{1}{2}\log\left(\frac{2}{3}\right) + \frac{1}{2}\log(2).$$

$$= 0.2075 \text{ bits.}$$

whereas,

$$D(q\|p) = \frac{3}{4}\log\frac{3/4}{1/2} + \frac{1}{4}\log\frac{1/4}{1/2}$$

$$= \frac{3}{4}\log\left(\frac{3}{2}\right) + \frac{1}{4}\log\left(\frac{1}{2}\right).$$

$$= 0.1887 \text{ bits.}$$

$\therefore D(p\|q) \neq D(q\|p)$.

[in general].

Relationship between Entropy and mutual Information:

mutual information.

$$I(X;Y) = \sum p(x,y)\log\frac{p(x,y)}{p(x)\cdot p(y)}$$

$$= \sum_{x,y} p(x,y)\cdot\log\frac{p(x|y)}{p(x)}.$$

Scanned by CamScanner

$$= -\sum_{x,y} p(x,y) \cdot \log p(x) + \sum_{x,y} p(x,y) \cdot \log p(x|y).$$

$$= -\sum_{x} p(x) \cdot \log p(x) + \sum p(x,y) \cdot \log p(x|y)$$

$$= -\sum_{x} p(x) \cdot \log p(x) - \left(-\sum_{x,y} p(x,y) \log p(x|y)\right).$$

$$I(X;Y) = H(X) - H(X|Y).$$ $H(x,y) = H(x) + H(y|x)$

Thus the mutual information $I(X;Y)$ is the reduction in the uncertainity of X due to the knowledge of Y.

We know,

$$H(X;Y) = H(X) + H(Y|X).$$

$$\therefore I(X;Y) = H(Y) - H(Y|X).$$

$$= H(Y) - H(X,Y) + H(X).$$

~~So, $I(X,X) = H(X) - H(X,X)$~~

$$= H(X) + H(Y) - H(X,Y).$$

Self information,

$$I(X,X) = H(X) - H(X|X)$$

$$= H(X).$$

Scanned by CamScanner

The mutual information of a random variable with itself is the entropy of the random variable. This is known as self-information.

So, Self information is like marginal entropy.

$$I(X;X) = H(X).$$

<u>Theorem: 2.4.1</u>

$$I(X;Y) = H(X) - H(X|Y).$$

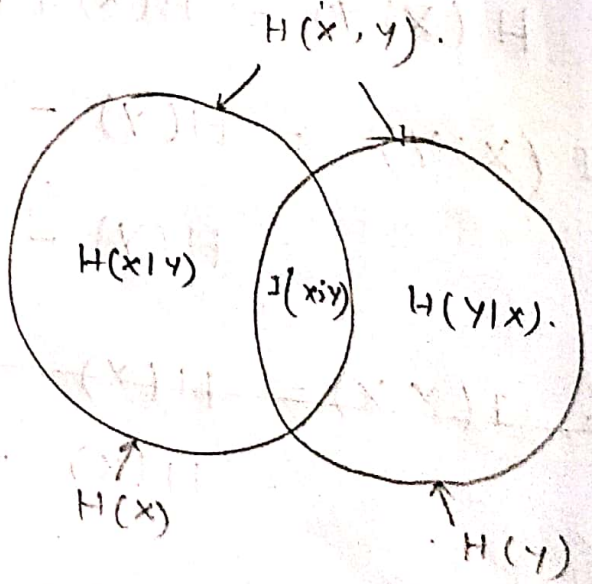
$$I(X;Y) = H(Y) - H(Y|X).$$

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

$$I(X;Y) = I(Y;X).$$

$$I(X;X) = H(X).$$

This is called venn dig. diagram.



In this digram, the relationship between $H(X)$, $H(Y)$, $H(X,Y)$, $H(X|Y)$, $H(Y|X)$ and $I(X;Y)$ is expressed.

relationship between entropy and mutual information.

~~I(X;Y) is the information~~ m

The mutual information $I(X;Y)$ corresponds to the intersection of the information in X with the information in Y.

Example (2.2.1) (continuous)

mutual information $I(X;Y) = H(X) - H(X|Y)$

$= H(Y) - H(Y|X)$.

$\frac{7}{4} - \frac{11}{8}$

$= \frac{14-11}{8} = \frac{3}{8}$.

$= 2 - \frac{13}{8}$

$= \frac{3}{8} = 0.375$ bits.

Summary

1) $H(X) \geq 0$

2) $H_b(X) = (\log_b a) H_a(X)$.

3) $H(X|Y) \leq H(X)$.

4) $H(P)$ is concave in P.

5) $H(X) = E_P \log \frac{1}{P(X)}$

6) $H(X,Y) = E_P \log \frac{1}{P(X,Y)}$.

7) $H(X|Y) = E_P \log \frac{1}{P(X|Y)}$

8) $I(X;Y) = E_P \log \frac{P(X,Y)}{P(X) \cdot P(Y)}$

9) $D(p\|q) = E_P \log \frac{P(X)}{q(X)}$.

Scanned by CamScanner

CC01

Chapter - 1.

Introduction

Shanon's Information theory.

The theory provides answers to two fundamental questions.

1) What is the irreducible complexity below which a signal cannot be compressed.

2) What is the ultimate transmission rate for reliable communication over a noisy channel?

Shannon's first theorem: (Source Coding theorem).

The theorem can be stated as follows: Given a discrete memoryless source of entropy H(s), the average code-word length L for any distortionless source coding is bounded as

$$L \geq H(s)$$

This theorem provides the mathematical tool for lossless data compression.

Scanned by CamScanner

⊞ The entropy of a source is a function of the probabilities of the source symbols. &

⊞ The entropy is the measure of the average information content per source symbol.

→ Source coding theorem is also known as "noiseless coding theorem".

Channel Coding theorem: (Shannon's 2nd theorem).

The channel coding theorem for a discrete memoryless channel is stated in two parts as follows:

a) Let a discrete memoryless source with an alphabet S have entropy H(S) and produce symbols once every T_s seconds. (প্রতি T_s সেকেন্ডে একটি symbol generate করে)

Let a discrete memoryless channel have capacity C and be used once every T_c seconds.

Then if $$\underbrace{\frac{H(S)}{T_s}}_{\text{(system rate)}} \leq \underbrace{\frac{C}{T_c}}_{\text{(channel rate)}}$$

There is a coding scheme for which the

Scanned by CamScanner

source output can be transmitted over the channel and reconstructed with an arbitary small probability of error.

(b) conversly if

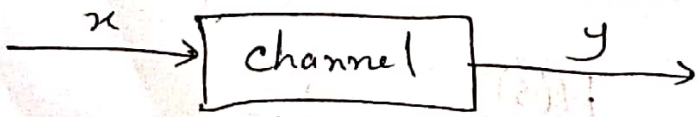
$$\frac{H(s)}{T_s} > \frac{C}{T_c}$$

It is not possible to transmit information over the channel and reconstruct it with an arbitarily small probability of error.

(channel rate হলে signal rate হবে হলে transmission possible না).

→ The theorem specifies that the ~~chap~~ channel capacity C as a fundamental limit on the rate at which transmission of reliable error-free message can take place over a discrete memoryless channel.

Mutual Information



Channel input x with entropy H.

$\bar{X} = \{x_0, x_1, \ldots x_{j-1}\}$.

Scanned by CamScanner

channel output y

$$\bar{Y} = \{y_0, y_1, \ldots, y_{k-1}\}.$$

How can the uncertainity about x after observing y?

The conditional entropy of X selected from $\bar{X}$, given that $Y = y_k$ is

$$H(\bar{X} \mid Y = y_k) = \sum_{j=0}^{J-1} P(x_j \mid y_k) \cdot \log_2 \left[\frac{1}{P(x_j \mid y_k)} \right].$$

The mean of entropy $H(\bar{X} \mid y_k)$ over the output $\bar{Y}$ is given by.

$$H(\bar{X} \mid \bar{Y}) = \sum_{k=0}^{k-1} H(\bar{X} \mid \bar{Y} = y_k) \cdot P(y_k).$$

$$= \sum_{k=0}^{k-1} \sum_{j=0}^{J-1} P(x_j, y_k) \cdot \log_2 \left[\frac{1}{P(x_j \mid y_k)} \right]$$

The conditional entropy represents the amount of uncertainity remaining the channel input after the channel output has been ~~offered~~ observed.

Mutual Info $= H(\bar{X}) - H(\bar{X} \mid \bar{Y})$.

It represents uncertainity about channel

Scanned by CamScanner

input that is resolved by observing the channel output.

$$\therefore I(\bar{X};\bar{Y}) = H(\bar{X}) - H(\bar{X}|\bar{Y}).$$

$$I(\bar{X};Y) = I(\bar{Y};X).$$

conditional entropy use করে আর mutual info বেশি, সেটাই decision.

$$\text{rate} = \frac{x}{X}$$

x ← any value, X ← total value.

$\bar{X}$ এর j component change হলে n component এর বেশে $x^J = c^n$ হয়,

$$\text{rate} = \frac{J}{n} \text{ হয়,}$$

Channel capacity:

The mutual information can be expressed

$$I(\bar{X};\bar{Y}) = \sum_{k=0}^{K-1}\sum_{j=0}^{J-1} P(x_j, y_k)\log_2\left[\frac{P(x_j|y_k)}{P(y_k)}\right].$$

$\{P(x_j)\}$ is obviously independent of channel. So. $I(\bar{X};\bar{Y}$ can be maximized with respect to $\{P(x_j)\}$.

So. channel capacity $C = \max_{\{P(x_j)\}} I(\bar{X};\bar{Y})$.

Scanned by CamScanner

channel capacity is measured in bit per channel.

→ Channel coding theorem is also known as noisy coding theorem.

→ maximum mutual info হল channel capacity. যখন x ও y স্বাধীন,

Information Capacity Theorem:

→ also known as Shannon. Hartley law or Shannon's 3rd theorem.

→ Statement:

The information capacity of continuous channel of bandwidth B Hz, with white gaussian noise of power spectral density $\frac{N_0}{2}$ and limited in Bandwidth to B is given by

$$C = B \log_2 \left(1 + \frac{P}{N_0 B}\right) \text{ bits/second.}$$

where P is the average transmitted power.

Scanned by CamScanner

This theorem implies that, for given average transmitted power P and channel bandwidth B, the information transmission rate = C bits per second. with small probability of error.

signal এর শুধু power বাড়ালেই capacity বাড়ে না, কিন্তু bandwidth বাড়ালে power তুলনায় better performance হয়।

Shanon limit:

for an ideal system that transmits bits at a rate $R_b = C$. (signal rate channel capacity সমান).

Thus,

$$P = E_b * R_b$$

$$= E_b * C$$

[E_b = power per bit]

~~Then~~

We know,

$$C = B\log_2\left(1 + \frac{P}{N_0 B}\right)$$

or, $$\frac{C}{B} = \log_2\left(1 + \frac{P}{N_0 B}\right)$$

or, $$\frac{C}{B} = \log_2\left(1 + \frac{E_b C}{N_0 B}\right)$$

or, $$\frac{C}{B} = \log_2\left(1 + \frac{E_b}{N_0} \cdot \frac{C}{B}\right).$$

Scanned by CamScanner

$\therefore \frac{E_b}{N_0}$

or, $2^{C/B} = 1 + \frac{E_b}{N_0} \cdot \frac{C}{B}$

or, $2^{C/B} - 1 = \frac{E_b}{N_0} \cdot \frac{C}{B}$

or, $\frac{E_b}{N_0} = \frac{2^{C/B} - 1}{\frac{C}{B}}$

For infinite bandwidth, the $\frac{E_b}{N_0}$

$$\left(\frac{E_b}{N_0}\right)_\infty = \lim_{B \to \infty} \left(\frac{E_b}{N_0}\right) = \ln 2$$

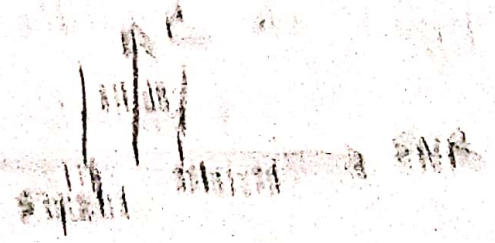
$$= 0.693$$

$$= -1.6 \text{ dB}$$

This value is called shannon limit

minimum -1.6 dB power supply for unlimited bandwidth ↳ reliable communication

signal ও power সমান হলে, power এর 70% হলে ও reliable communication হবে, when bandwidth unlimited.



Introduction to Error control coding:

Purpose of Error Control Coding:

1) In data communication, coding is used for controlling transmission errors induced by channel noise or other impairments so that error free communication can be acheived.

2) In data storage system, coding is used for controlling storage errors caused by storage medium defects so that error-free storage can be acheived.

Coding Principle:

Coding is acheived by adding properly designed redundant digits to each message.

↓

These redundant digits are used for detecting or correcting errors.

Types of Coding:

1) Block Coding
2) Convolutional Coding.

Scanned by CamScanner

Block Coding:

A message of k digits is mapped into a structured sequence of n digits, called a codeword.

If k = message length

n = code length

Code rate $= \frac{k}{n}$.

The mapping operation is called encoding.

→ This encoding is memoryless.

→ The collection of all codeword is called a 'block code'.

Convolutional Coding:

An information sequence is divided into blocks of k digit each.

↓

Each k digit is encoded into n-digit coded block

↓

The n digit coded block depends on k-digit message block and also previous message block

↓

encoder has memory of order m.

Scanned by CamScanner

→ The encoder has k inputs and n outputs.

→ The collection of all possible code sequence is called an convolutional code.

→ $1 \leq k \leq 8$; $\frac{k}{n}$ = code rate

$2 \leq n \leq 9$.

→ $n > k$; so, rate < 1.

Types of Errors and Channels

1) Random error (error bit by bit)
2) Burst error, (ত্রুটিগুলো একসাথে থাকে)

Type of channels:

Random Error channel:

1) Deep space channel
2) satellite channel
3) Line of sight transmission channel.

Burst error channel

1) radio link
2) terrestrial microwave link.
3) cable transmission link.

Decoding:

এটি হল received sequence এবং encoding rule এবং noise characteristics থেকে

উপর ভিত্তি করে decoder decision নেয় যে কী transmit করা হয়েছিল,

→ The decision making operation is called decoding. And the device which perform the decoding operation is called a decoder.

→ There are two types of decoding.

1) Hard decision decoding.
2) Soft decision decoding.

Hard decision decoding:

→ when binary coding is used, the modulator has only binary inputs.

If binary demodulator output quantization is used, the decoder has only binary inputs.

In this case, the decoder is said to make hard decision.

Decoding based on hard decision made by the modulator is called hard decision decoding.

Scanned by CamScanner

যদি binary (1, -1) transmit করি এবং যা channel
এর মধ্যে noise আছে (0.5, -0.5) হয়। তাহলে des.
এর মান 1 এবং -0.5 এর জন্য 0 হবে,

→ hard decision এ error বেশি।

Soft decision decoding:

If the output of demodulator consists of more than two quantization levels or left unquantized, then demodulator is said to make soft decision.

Decoding based on soft decision made by demodulator is called soft decision decoding.

(i) Hard decision decoding is much easier to implement than soft decision decoding.

(ii) Soft decision decoding offers much better performance.

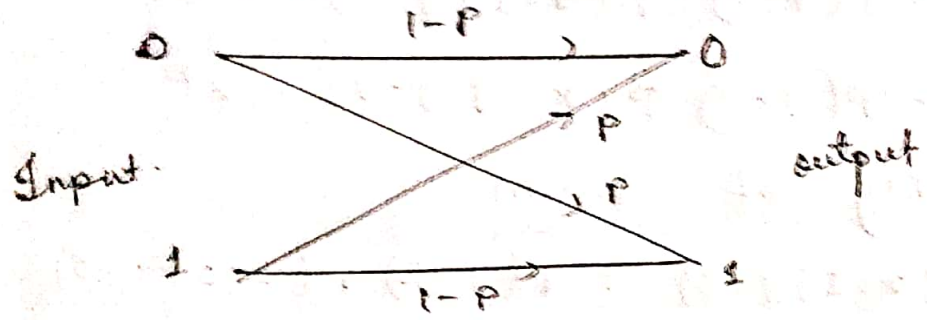
(iii) soft decision এর ক্ষেত্রে যার probability বেশি decision সেটাই হবে, এক্ষেত্রে time বেশি লাগবে।

Scanned by CamScanner

Some channel Model

Binary Symmetric Channel:

~~এটা~~ symmetric channel এতে 0 থেকে 1 হয়ে যাওয়ার probability = 1 থেকে 0 হয়ে যাওয়ার probability.



when BPSK modulation is used over the AWGN channel with optimum co-herent detection and binary output quantization, the bit error probability

$$P = Q\left(\sqrt{\frac{2E}{N_0}}\right).$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{y^2}{2}} . dy$

E: bit energy

$\frac{N_0}{2}$: power spectral density of AWGN.

→ এই model কে butterfly model ও বলে।

Scanned by CamScanner

$P(x=0 \mid y=0) = 1-p$

$P(x=1 \mid y=0) = 1-p$

$P(x=0 \mid y=1) = p$

$P(x=1 \mid y=1) = p$

probability of error

$$= \frac{p+p}{1-p+p+1-p+p}$$

$$= \frac{2p}{2} = \frac{2p}{2} = p.$$

$\therefore P(x=1, y=0) = P(y=0)\, P(x=1 \mid y=0) = \frac{1}{2}p.$

$\therefore P(x=0, y=1) = P(y=1)\, P(x=0 \mid y=1) = \frac{1}{2}p.$

$\therefore$ ~~total~~ probability of error

$= P(y=0)\, P(x=1 \mid y=0) + P(y=1)\, P(x=0 \mid y=1).$

$= p.$

Binary Input & ary Output discrete channel

Optimum Decoding:

$\bar{c}$ = codeword.

$\bar{m}$ = message.

$\bar{r}$ = output of demodulation.

$\hat{m}$ = estimated message from $\bar{r}$.

Scanned by CamScanner

→ An optimum decoding rule is that ~~minimizing~~ minimize the probability of a decoding error.

$\therefore P(\hat{c} \neq \bar{c} | \bar{r})$ is minimized.

or, $P(\hat{c} = \bar{c} | \bar{r})$ should be maximized.

$\hat{c}$ ও $\bar{c}$ match না হওয়ার probability কমাতে হবে
$\hat{c} = \bar{c}$ হওয়ার probability বাড়াতে হবে।

→ The decoding error is minimized by choosing $\hat{c}$ to be a codeword $\bar{c}$ that maximize

$$P(\bar{c} | \bar{r}) = \frac{P(\bar{r} | \bar{c}) P(\bar{c})}{P(\bar{r})}$$ → (MAP decoding এ যখন যদি P(c) এর estimation use হয়)

That is $\hat{c}$ is choosen to be most likely codeword.

decoding error minimize করার জন্য
$\hat{c}$ এমনভাবে choose করব যাতে $P(\bar{c}|\bar{r})$ maximize হয়

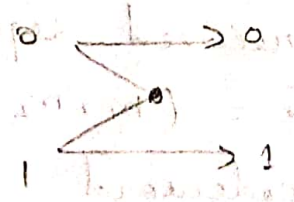
~~এবং decoding এর~~

$P(\hat{c} \neq \bar{c} | \bar{r})$ → minimum means square error.

$P(\hat{c} = \bar{c} | \bar{r})$ → maximum likelihood.

Binary Erase channel:

two input, 3 output.



0 → 0, 1 → 1 হলে receiver accept করে,

0, 1 ছাড়া কিছু হলে receiver retransmit করতে বলে,

Scanned by CamScanner

Maximum likelihood decoding:

All msg are equally likely.

↓

for every codeword $\bar{c}_j$, the conditional probability

$= P(\bar{r} \mid \bar{c}_j)$

↓

the codeword $\bar{c}_j$ with largest conditional probability is chosen as estimate of $\bar{c}$.

↓

This decoding rule is called maximum likelihood decoding.

MLD for BSC

$\bar{a} = (a_1, a_2, a_3, \ldots, a_n)$.

$\bar{b} = (b_1, b_2, b_3, \ldots, b_n)$.

Hamming distance btwn $\bar{a}$ and $\bar{b}$ $d_H(\bar{a}, \bar{b})$, is defined as no. of places where a and b differ.

If received sequence is $\bar{r}$,

$\bar{r} = (r_1, r_2, r_3, \ldots, r_n)$.

for codeword $\bar{c}_i$, the conditional probability $P(\bar{r} \mid \bar{c}_i)$ is

$P(\bar{r} \mid \bar{c}_i) = p^{d_H(r, c_i)} (1-p)^{n - d_H(r, c_i)}$

If and only if

$d_H(\bar{r}, \bar{c}_i) < d_H(\bar{r}, \bar{c}_j)$

then

$P(\bar{r} \mid \bar{c}_i) > P(\bar{r} \mid \bar{c}_j)$

Scanned by CamScanner

procedure of MLD:

1) compute $d_H(\bar{r}, \bar{c}_i)$ for all $\bar{c}_i$.

2) $\bar{c}_i$ is taken as transmitted codeword if $d_H(\bar{r}, \bar{c}_i) < d_H(\bar{r}, \bar{c}_j)$ for $j \neq i$.

3) Decode $\bar{c}_i$ into message $\bar{m}_i$.

So, received vector $\bar{r}$ is decoded into ~~closed~~ closest codeword.

এরে minimum distance decoding ও বলে।

Scanned by CamScanner

Chapter - 9

Code

- Linear
  - Systematic
    - Block Code
    - Convolutional Code
      - Recursive code
      - Non Recursive Code.
  - Non systematic
- Non linear.

properties

Linear Code ta সমষ্টিগত হৈ Code দুটাৰ যোগফলৰ পৰা নতুন Code হয়। এনে নিৰ্দিষ্ট কোডবোৰ Code, $C_0, C_1, C_2, \ldots, C_n$.

$C_0 + C_1 = C = C_2 / C_3 / C_4$ এনেকুৱা একাধিক মাত্ৰ মিলাব,

systematic code ত msg bits আৰম্ভত, extra bit থাকে হয়, সমান্য Parity check matrix.

block code ত block by block coding কৰা হয়, convolutional code infinite input ত্ৰে infinite output দিয়ে

recursive code এটা feedback code. যাতে complex but performance বেছি

non recursive code simple, but performance কম হয়।

Scanned by CamScanner

→ parity check code block code + systematic code.
কোন শুধু ১ position এর error detect করতে পারে কিন্তু correct করতে পারে না।

code rate, $R = \frac{k}{n} = \frac{k}{k+1}$.

→ channel code এর কাজ:

1) error detection 2) error correction.

Introduction to Linear Block code:

In block coding, the binary information sequence is segmented into message blocks of fixed length.

u = message block consists of k information digit.

$\therefore 2^k$ = distinct message.

each input message u is transformed into binary n-tuple v with $n > k$. The binary n-tuple v is referred to as the codeword of message u.

for 2^k message, 2^k codewords are possible. Set of 2^k codewords are called block code. প্রতি msg এর জন্য individual codeword.

Scanned by CamScanner

Definalion:

A block code of length n and 2^k codewords is called a linear code (n, k) if and only if its 2^k codewords ~~from~~ form a k-dimensional subspace of the vector space of all the n-tuples over the field GF.

Linear transformation of binary block code:

$$c = m \cdot G$$

$$= \sum_{i=0}^{k-1} m_i \cdot \bar{g}_i$$

$$= m_0 g_0 + m_1 g_1 + \cdots + m_{k-1} g_{k-1}$$

Here, c = codeword

G = generating matrix.

m = message to be encoded.

Binary block code is linear if and only if the sum of two code words is also a codeword.

Example: 3.1

$$G = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\underbrace{\qquad\qquad}_{\text{parity}} \quad \underbrace{\qquad\qquad}_{\text{identity}}$$

$u = (1101)$ is the message to be encoded.

Scanned by CamScanner

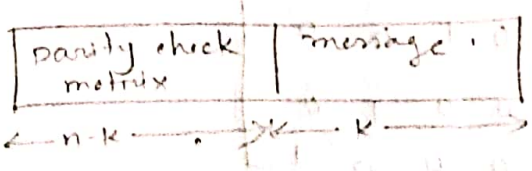
Here the parity matrix is $= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$.

codeword $c = u \cdot G$

$$= \begin{bmatrix} 1 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}.$$

Linear Systematic block code.



The codeword C is comprised of an information segment and a set of $n-K$ symbol that are linear combination of certain information symbol.

codeword:

$$(m_0, m_1, m_2, \ldots m_{k-1}) \longleftrightarrow (m_0, m_1, \ldots, m_{k-1}, \underbrace{c_k, c_{k+1}, \ldots c_{n-1}}_{\text{parity check matrix.}}).$$

An (n, K) linear systemetic code can be expressed as,

$G = [I_k \cdot P]$. where $I = k \times k$ matrix

$P = k \times (n-k)$ matrix.

Scanned by CamScanner

Example: (7, 4) hamming code over GF(2).

The encoding equation for this code is -

$C_0 = m_0$

$C_1 = m_1$

$C_2 = m_2$

$C_3 = m_3$

$C_4 = m_0 + m_1 + m_2$

$C_5 = m_1 + m_2 + m_3$

$C_6 = m_0 + m_1 + m_3$.

$\therefore$ So, $G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$

5
2
6
3

Example - 3.2

Given that,

$$G = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

$G = P \times I$.

$P^T . I$

$m = (u_0, u_1, u_2, u_3)$,

So, equation for codewords will be,

$C_0 = u_0 + u_2 + u_3$

$C_1 = u_0 + u_2 + u_1$

$C_2 = u_1 + u_2 + u_3$

$C_3 = u_0$.

$C_4 = u_1$

$C_5 = u_2$

$C_6 = u_3$.

If message $u = (1011)$ then codeword $= (1001011)$

Scanned by CamScanner

linear combination কে matrix দিয়ে represent করা হয় কেন?

matrix এর behaviour ও code এর behaviour,

matrix দিয়ে represent করলে প্রয়োজন parameter দ্বারা সহজে code থেকে data recover করতে সাহায্য করে matrix দেখে বুঝা যায়।

Parity Check Matrix:

An (n,k) linear code also be specified by an $(n-k) \times k$ matrix H.

Let. $\bar{c} = (c_0, c_1, \ldots c_{n-1})$.

$\bar{c}$ is a codeword if $\bar{c} \cdot H^T = 0$.

The H is called parity check matrix.

$$G = [I_k \; P]$$

$$H = [P^T, I_{n-k}]. \quad [P^T \text{ এর transpose of } P].$$

$$\therefore G \cdot H^T = 0.$$

Example:

Given that, A $(6,3)$ code is generated by.

$$G = \begin{array}{c} \\ \\ \\ \end{array}\begin{array}{c} \begin{array}{cccccc} n & & & k & & \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \\ \underbrace{\qquad\qquad}_{P} \end{array}$$

$P_0 = m_0 + m_1$

$P_1 = m_0 + m_2$

$P_2 = m_1 + m_2 + m_3$.

$$\therefore H = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Scanned by CamScanner

Now, $G \cdot H^T$

$$= \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 0.$$

$$\therefore G \cdot H^T = 0.$$

Hamming distance of linear block code and Error protection properties:

Distance between two n symbol vectors

$\bar{u} = (u_0, u_1, \ldots u_{n-1})$

$\bar{v} = (v_0, v_1, \ldots v_{n-1})$.

a) Euclidean distance

$$d_E(\bar{u}, \bar{v}) = \sqrt{\sum_{i=0}^{n-1} (u_i - v_i)^2}.$$ (Physical distance)

b) Hamming distance:

$$d_H(\bar{u}, \bar{v}) = \left| \{ i \mid u_i \neq v_i \; ; \; i = 0, 1, \ldots n-1 \} \right|.$$

→ no. of places where $\bar{u}$ and $\bar{v}$ differ.

(dissimilarity).

Scanned by CamScanner

Hamming weight and hamming distance of codewords.

(a) for a linear code C, the Hamming distance between any two codeword is

$$d_H(\bar{c}_1, \bar{c}_2) = wt(\bar{c}_1 - \bar{c}_2) = wt(\bar{c}_3).$$

$\bar{c}_3$ is the difference between $\bar{c}_1$ and $\bar{c}_2$.

→ $wt(\bar{c}_3)$ is the hamming weight of c_3.

→ Hamming weight means no. of non zero positions in a codeword $(\bar{c}_3)$.

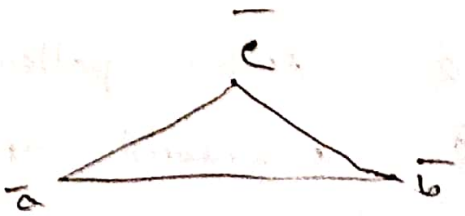
$$wt(\bar{c}) = \{ i \mid c_i \neq 0 \; ; \; i = 0, 1, \ldots n-1 \}.$$

[illegible] sequence [illegible] non zero position [illegible] hamming weight.

(b) Triangle inequality:

For codeword $\bar{a}, \bar{b}, \bar{c}$.

$$d_H(\bar{a}, \bar{c}) + d_H(\bar{c}, \bar{b}) \geq d_H(\bar{a}, \bar{b}).$$



c) $d_H(\bar{a}, \bar{b}) = wt(\bar{a} + \bar{b})$.

Minimum Distance of a block code:

Let c is a block code word.

The minimum distance of C

$$d_{min} \equiv \min \{ d(\bar{u}, \bar{v}) : \bar{v}, \bar{u} \in C, \bar{v} \neq \bar{u} \}.$$

Scanned by CamScanner

The minimum weight of C is

$$w_{min} = \min \{ w(\bar{v}) : \bar{v} \in C, \bar{v} \neq \bar{0} \}.$$

→ distance যত বাড়বে তত বেশি error detect correct করতে পারব।

Show that $w_{min} = d_{min}$.

$$d_{min} = \min \{ d(\bar{u}, \bar{v}) : \bar{u}, \bar{v} \in C, \bar{v} \neq \bar{u} \}$$

$$= \min \{ d(\bar{u} + \bar{v}) : \bar{v}, \bar{u} \in C, \bar{v} \neq \bar{u} \}.$$

$$= \min \{ w(x) : x \in C, \bar{x} \neq \bar{0} \}$$

$$= w_{min}.$$

minimum distance of linear code = minimum weight of its non-zero codewords

Maximum error correction capability:

$\bar{e}$ = error pattern

$\bar{c}$ = codeword transmission,

$\bar{r}$ = received word.

If d_{min} = minimum hamming distance,

then maximum error correction capability of the code, $t = \frac{d_{min} - 1}{2}$

garantee error detection capability of the code is $d_{min} - 1$.

Scanned by CamScanner

Hybrid modes of error channel:

One can correct t errors and still detect upto t_d errors, then

$$t + t_d < d_{min}.$$

$$2t + 1 \leq d_{min}.$$

Syndrome and error detection:

If codeword $\bar{c} = (c_0, c_1 \ldots c_{n-1})$ is transmitted.

receivedword $\bar{r} = (r_0, r_1 \ldots r_{n-1})$

So, the error, $\bar{e} = \bar{r} + \bar{c}$

$$= (r_0 + c_0,\ r_1 + c_1,\ \ldots,\ r_{n-1} + c_{n-1}).$$

$$= (e_0,\ e_1,\ \ldots,\ e_{n-1}).$$

total possible error $= 2^n$. (n received bit).

correctable error $= 2^{n-k}$. by a linear (n, k) block code.

syndrome, $\bar{s} = \bar{r} \cdot H^T$.

$$= (\bar{c} + \bar{e}) \cdot H^T.$$

$$= \bar{e} \cdot H^T.$$

$\bar{s} = (s_0, s_1 \ldots, s_{n-k-1})$.

$H = (n-k) \times n$ matrix.

if $\bar{s} = 0$, then $\bar{r}$ is codeword.

if $\bar{s} \neq 0$, then $\bar{r}$ is not a codeword and contains transmission errors.

Scanned by CamScanner

In this way, error can be detected.

* $\bar{s} = e \cdot H^T$

This ~~eq~~ equation gives a relationship between unknown error pattern and the syndrome.

→ $\bar{s}$ is called syndrome of error pattern. The elements of $\bar{s}$ are linear combination only of the errors, e_i.

<u>Example</u>

If parity check matrix of the (6,3) code is given.

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

the received word is $\bar{r} = (101001)$

$$\bar{s} = \bar{r} \cdot H^T = (101001) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

$$= (011).$$

$\therefore \bar{r}$ is not the codeword.

যদি $\bar{s} = (000)$ হত, তাহলে codeword হত।

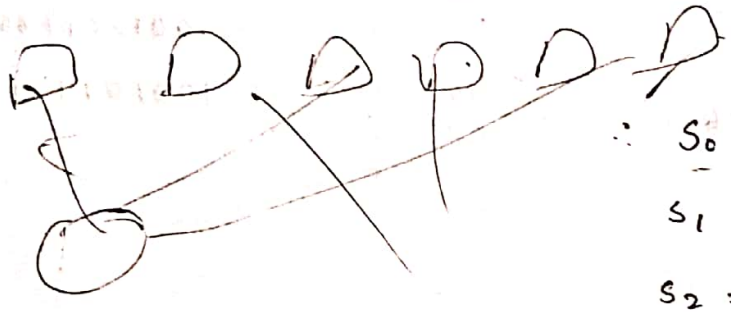
Scanned by CamScanner

Syndrome Circuit:

If $\bar{r} = (r_0, r_1, r_2, r_3\ r_4\ r_5)$ is a received word. and

$$H^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} ; \quad \bar{s} = (s_0, s_1, s_2) = r \cdot H^T$$

$$= \underbrace{(r_0\ r_1\ r_2\ r_3\ r_4\ r_5)}_{1\times 6} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}}_{6\times 3}$$



$\therefore\ s_0 = r_0 + r_4 + r_5$

$s_1 = r_1 + r_3 + r_5$

$s_2 = r_2 + r_3 + r_4$.

Example:

Given that,

$$H = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

$\bar{c} = (1001011)$.

$\bar{r} = (1001001)$.

So, syndrome of $\bar{r}$ is $s = (s_0\ s_1\ s_2)$

$= \bar{r} \cdot H^T$.

$$= (1001001) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} = (111).$$

Scanned by CamScanner

Let, $\bar{e} = (e_0\ e_1\ e_2\ e_3\ e_4\ e_5\ e_6)$ be the error pattern.

$\therefore \bar{s} = \bar{r} \cdot H^T$,

$= (111)$.

received এর সাথে error যোগ করলে true হয় তাহলে সা

$\therefore 1 = e_0 + e_3 + e_5 + e_6$

$1 = e_1 + e_3 + e_4 + e_5$

$1 = e_2 + e_4 + e_5 + e_6$,

$e' = r + e$

$= 1001001 + 0000010$

$= 1001011$

$\bar{e} = (\bar{r} + \bar{e})$

$= 1001001 + 1001011$

$= (0000010)$. This is the true error pattern.

Example - 3.4

Given that, $H^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$.

$\bar{r} = (r_0, r_1, r_2, r_3, r_4, r_5, r_6)$ is the received vector.

and syndrome, $\bar{s} = (s_0, s_1, s_2)$.

$= (r_0, r_1, r_2, r_3, r_4, r_5, r_6) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$

Scanned by CamScanner

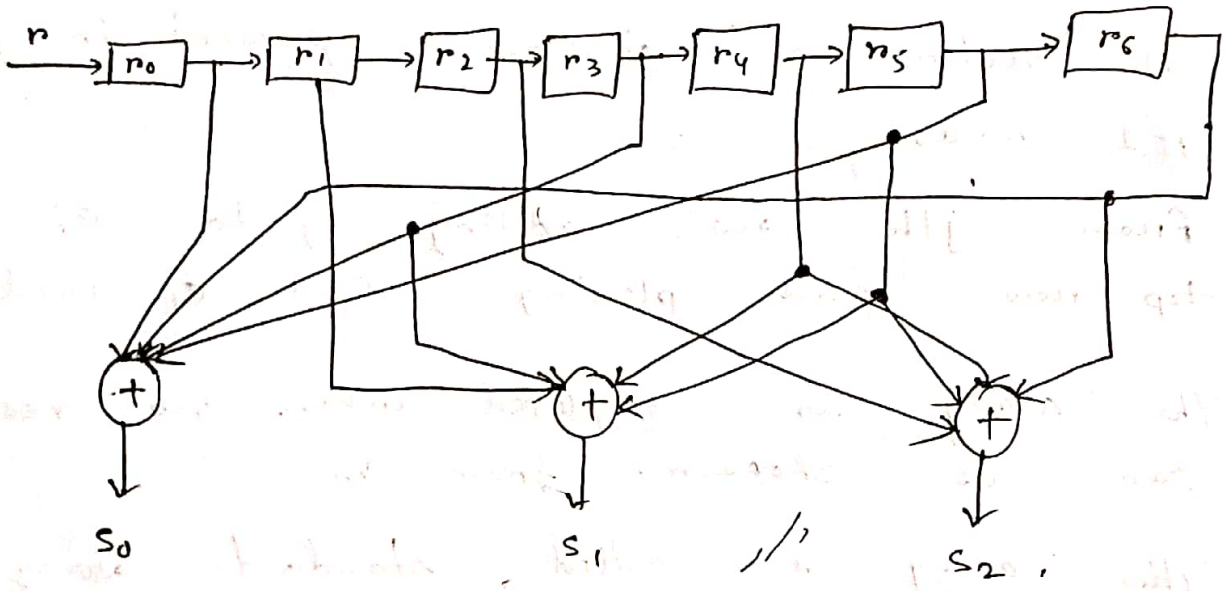
The syndrome digits are

$$s_0 = r_0 + r_3 + r_5 + r_6$$

$$s_1 = r_1 + r_3 + r_4 + r_5$$

$$s_2 = r_2 + r_4 + r_5 + r_6$$

So the syndrome circuit:



~~Syndrome circuit for linear n,k code:~~

Standard Array:

→ The standard array is a way of tabulating all of the q^n n tuples over GF(q).

→ The $q^{n-k} \times q^k$ rectangular array is used as decoding table.

$(6,3) = \{000, 111$

formulation of standard array:

for (n, k) linear binary code.

1) The 2^k codewords are arranged as the top row of the array with $\bar{c}_1 = 0$ as first element.

2) The $(j-1)$th row is formed.

3) A vector $\bar{e}_j$ is not present in previous $j-1$ rows.

4) From jth row, adding e_j to c_i in the top row and placing $\bar{e}_j + \bar{c}_i$ under $\bar{c}_i$.

5) The array is completed when no vector can be choosen from V_n.

6) This array is called standard array.

$\bar{c}_1 = 0$	$\bar{c}_2$	. . . .	$\bar{c}_i$	. . .	$\bar{c}_{2^k}$
$\bar{e}_2$	$\bar{e}_2 + \bar{c}_2$	. . .	$\bar{e}_2 + \bar{c}_i$		$\bar{e}_2 + \bar{c}_{2^k}$
⋮					
$\bar{e}_{2^{n-k}}$	$\bar{e}_{2^{n-k}} + \bar{c}_2$	. . .	$\bar{e}_{2^{n-k}} + \bar{c}_i$	. . .	$\bar{e}_{2^{n-k}} + \bar{c}_{2^k}$

→ each row is called a coset.

→ 2^{n-k} cosets are there.

→ first element of each coset is called coset leader.

→ binary vector appears once in the array.

→ ~~প্রতি~~ row ~~এর~~ syndrome same, different row

Scanned by CamScanner

लर syndrome [illegible]

Example: For a (6,3) linear block code.

$$G = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Syndrome Decoding using standard Array:

→ syndrome decoding can be done by using a table which consists of 2^{n-k} correctable error patterns and their corresponding syndrome.

$S_1 = 0$	$e_1 = 0$
$S_2 = e_2 \cdot H^T$	e_2
$S_3 = e_3 \cdot H^T$	e_3
⋮	
$S_{2^{n-k}} = e_{2^{n-k}} \cdot H^T$	$e_{2^{n-k}}$.

Scanned by CamScanner

Example: for (6, 3) code

$$G = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The parity check matrix

$s = e \cdot H^T$.

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Syndrome decoding table is constructed by using

$\bar{H}$ - (syndrome (সহ coset leader বের করা নিয়ম?

$\bar{s}$	$\bar{e}$
000	000000
110	000001
101	000010
011	000100
001	001000
010	010000
100	100000
111	000001
	~~001001~~

If the received word, $\bar{r} = (101001)$.

then $\bar{s} = \bar{r} \cdot H^T$

Scanned by CamScanner

$$= (101001)\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

$$= (011)$$

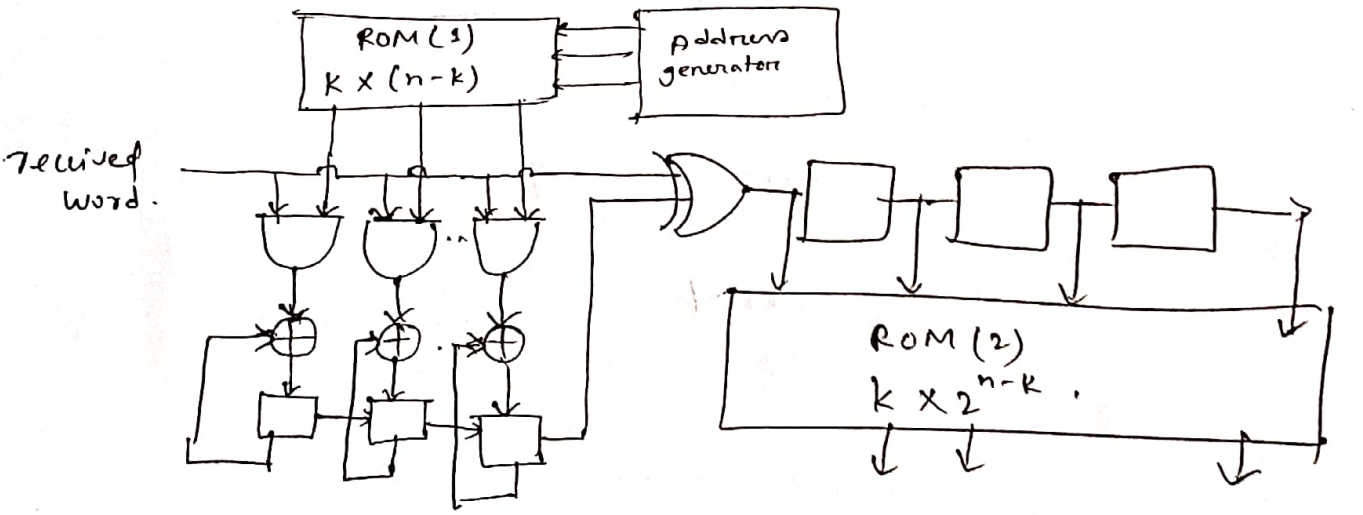
The corresponding error pattern $= (000100)$.

So correct codeword $\bar{c} = (101101)$. তিন নং bit alter হয়েছে।

$= (101001 + 000100)$

$= 101101$

Implementation of Syndrome Decoding:



ROM 2 ↳ standard array loaded থাকবে,
ROM 2 sequence আকারে চুকবে কোথায় error আছে এবং ঠিক করে output দেবে।

Scanned by CamScanner

Hamming code :

→ error detecting + error correcting code.

for positive integer $m \geq 3$,

block code length, $n = 2^m - 1$.

message length, $k = 2^m - 1 - m$.

minimum hamming distance, $d_{min} = 3$.

error correction capability $= \frac{3-1}{2} = 1$.

Golay code.

→ multiple error correcting code.

$r = (101001)$.

$S = r \cdot H^T$

$= (011)$ → 4

$C = 101101$

$e = r + c$

$= 101101 + 101001$

$= (000100)$

Scanned by CamScanner