It is also possible to define ***"time-averaged autocorrelation function"*** as

$$\langle R_{XX}(\tau)\rangle = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{+T/2} X(\tau)X(t+\tau)d\tau \qquad \text{...... (0.75)}$$

and ***"time-averaged mean"*** as

$$\langle \mu_X\rangle = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{+T/2} X(t)dt$$

ERGODICITY : A random process is said to be ***"ergodic"*** in the most general form if ***"time averages"*** equal ***"ensemble averages"***.

In particular form, ergodicity is defined in the following way:

A random process is ***"ergodic in the mean"*** if

$E\{\langle\mu_x\rangle\} = E[X(t)]$ and Var $\{\langle\mu_x\rangle\} \to 0$ as $T\to\infty$ and a random process is ***"ergodic in the autocorrelation function"*** if $E[\langle R_{XX}(\tau)\rangle] = R_{XX}(\tau)$ and the variance Var $[\langle R_{XX}(\tau)\rangle] \to 0$ as $T\to\infty$.

POWER SPECTRAL DENSITY OF STATIONARY RANDOM PROCESS

If the autocorrelation function $R_{XX}(\tau)$ of a stationary random process is such that

$$\int_{-\infty}^{+\infty} |R_{XX}(\tau)|\, d\tau < \infty,$$

then its Fourier Transform $G_X(f)$ is given by

$$G_X(f) = \int_{-\infty}^{+\infty} R_{XX}(\tau)\exp(-j2\pi f\tau)\, d\tau \qquad \text{...... (0.76)}$$

is called as ***"power density spectrum"*** or the ***"power spectral density function"*** of X(t).

❏ ❏ ❏ ❏

UNIT 1

INFORMATION THEORY

1.1 INTRODUCTION

The block diagram of an information system can be drawn as shown in figure 1.1. The meaning of the word "*information*" in information theory is ***"message"*** or ***"intelligence"***. This message may be an electrical message such as voltage, current or power or speech message or picture message such as fascimile or television or music message. A source which produces these messages is called "*information source*".

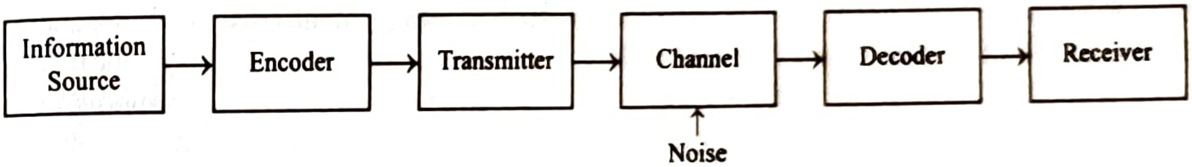

Fig. 1.1 : Block diagram of an Information System

Information sources can be classified into two categories: analog information sources and discrete information sources. Analog information sources, such as a microphone actuated by speech, or a TV camera scanning a scene, emit one or more continuous amplitude electrical signals with respect to time. The output of discrete information sources such as a teletype or the numerical output of a computer consists of a sequence of discrete symbols or letters. An analog information source can be transformed into a discrete information source through the process of sampling and quantizing.

Discrete information sources are characterized by (a) source alphabet (b) symbol rate (c) source alphabet probabilities and (d) probabilistic dependance of symbols in a sequence.

Example of source alphabet (discrete information source) → a teletype having 26 letters of the English Alphabet plus several special characters such as full stop, comma etc. along with numerals.

The ***Symbol rate*** refers to the rate at which the teletype produces characters. Ex: If the teletype operates at the speed of 10 characters/sec, then the symbol rate is said to be 10 symbols/sec.

If the teletype is producing messages in English language, then some letters appear more frequently than others. For example, the letter E appears more often than the letter Z and if a word starts with Q, the next letter will be U and so on. These structural properties of symbol sequences can be characterized by ***probability of occurrence of the individual symbols*** and also by the conditional properties of occurrence of symbols (i.e., probabilistic dependence).

In the block diagram of the information system shown in figure 1.1, let us assume that the information source is a discrete source emitting discrete message symbols $s_1, s_2, \ldots, s_q$ with probabilities of occurrence given by $p_1, p_2, \ldots, p_q$ respectively. The sum of all these probabilities must be equal to 1, since, if at all any symbol is emitted by the source, then it has to be one of $s_1, s_2, \ldots, s_q$ and not any other symbol.

$$\therefore p_1 + p_2 + \ldots + p_q = 1$$

$$\text{or} \quad \sum_{i=1}^{q} p_i = 1 \qquad \ldots\ldots (1.1)$$

SOURCE ENCODER : Let the input to the source encoder be a string of source symbols from the source Alphabet $S = \{s_1, s_2, \ldots, s_n\}$ occurring at a rate of "r_s" symbols/sec.

The source encoder converts the symbol sequence into a binary sequence of 0's and 1's by assigning code-words to the symbols in the input sequence. Binary coding is preferred because of its high efficiency of transmission and also the case with which they can be transmitted over the channel [other types of coding such as ternary, quarternary coding etc.are discussed in unit 3]. The simplest way of coding is to assign a fixed length binary code-word to each symbol in the input sequence. But, fixed-length coding of individual symbols in a source output is efficient only if the symbols occur with equal probabilities in a statistical independent sequence. In most practical situations, the symbols occur with unequal probabilities. The source encoder, then assigns variable length code-words to these symbols. The important parameters of a source encoder namely block size, length of code-words, average data rate and the encoder efficiency are discussed in detail in unit 3.

TRANSMITTER : The transmitter couples the input message signal to the channel. While it may sometimes be possible to couple the input transducer directly to the channel, it is often necessary to process and modify the input signal for efficient transmission over the channel. Signal processing operations performed by the transmitter include amplification, filtering and modulation. The most important of these operations is modulation - a process designed to match the properties of the transmitted signal to the channel through the use of a carrier wave.

CHANNEL : A communication channel provides the electrical connection between the source and the destination. The channel may be a pair of wires (2-line transmission system) or a telephone cable or free space over which the information bearing signal is radiated. Due to physical limitations, communication channels have only finite bandwidth and the information bearing signal suffers amplitude and phase distortion as it travels over the channel. In addition to the distortion, the signal power also decreases due to attenuation of the channel.

Furthermore, the signal is corrupted by unwanted unpredictable electrical signals referred as ***noise***. While some of the degrading effects of the channel can be removed or compensated for, the effects of noise cannot be completely removed. The main objective of a communication system design is to suppress the ill effects of noise as much as possible. When binary symbols are transmitted over the channel, the effect of noise is to convert some of the 0's into 1's and some of the 1's to 0's. The signals are then said to be corrupted by noise.

DECODER AND RECEIVER : The source decoder converts the binary output of the channel decoder into a symbol sequence. The decoder for a fixed-length code-words is quite simple, but the decoder for a system using variable-length code-words will be very complex. Therefore, the function of the decoder is to convert the corrupted signals into a symbol sequence and the function of the receiver is to identify the symbol sequence and match it with the correct sequence.

In 1948, C.E. SHANNON, known as ***"Father of Information Theory"***, published a treatise on the mathematical theory of communication in which he established basic theoretical bounds for the performances of communication systems. Shannon's theory is based on probabilistic models for information sources and communication channels. In the forthcoming sections, we present some of the important aspects of Shannon's work.

1.2 MEASURE OF INFORMATION

In order to know and compare the ***"information content"*** of various messages produced by an information source, a measure is necessary to quantitatively know that information content. For this, let us consider an information source producing independent sequence of symbols from source alphabet $S = \{s_1, s_2, ..., s_q\}$ with probabilities $P = \{p_1, p_2, ..., p_q\}$ respectively.

Let S_K be a symbol chosen for transmission at any instant of time with a probability equal to p_K. Then the ***"Amount of Information"*** or ***"Self-Information"*** of message S_K (provided it is correctly identified by the receiver) is given by

$$I_K = \log \frac{1}{p_K} \qquad ...... (1.2)$$

If the base of the logarithm is 2, then the units are called ***"BITS"***, which is the short form of ***"Binary Units"***. If the base is "10", the units are ***"HARTLEYS"*** or ***"DECITS"***. If the base is "e", the units are ***"NATS"*** and if the base, in general, is "r", the units are called ***"r-ary units"***.

The most widely used unit of information is ***"BITS"*** where the base of the logarithm is 2. ***Throughout this book, log to the base 2 is simply written as log and the units can be assumed to be bits, unless or otherwise specified.***

Just to get the concept of the "amount of information", let us consider the following example:

Suppose that you are planning a trip to Chennai, Tamil Nadu state from Bangalore in peak winter time (may be Chrismas time). To know the weather at Chennai, you telephone the weather bureau or browse through internet and receiv_ one of the following forecasts:

1. Sun will rise in East on the day of trip,
2. Weather will be dry, sultry and sunny for most of the day,
3. It will be a very cold day,
4. There will be snowfall on that particular day in Chennai.

The 'amount of information' received is obviously different for these messages.

The 1st message contains absolutely ***no information*** at all, since everybody know that Sun will rise in East on everyday morning in Chennai. The probability of Sun rising in East = $p_K = 1$ and from equation (1.2) is clear that $I_K = 0$, irrespective of the base of the logarithm.

The 2nd message gives some information, since except during monsoon, it is almost true.

The 3rd message which forecasts a very cold day contains more information since it is not an event that occurs often in Chennai even during winter time.

The 4th message which forecasts snowfall conveys a very very large amount of information since the occurrence of snowfall in Chennai is almost an impossible event!! The probability of snowfall in Chennai is very very small so that the information content as confirmed from equation (1.2) is very very large.

From the above example, we can conclude that the message associated with an event least likely to occur contains most information. This statement can be confirmed with another numerical example:

Example 1.1 : The binary symbols '0' and '1' are transmitted with probabilities 1/4 and 3/4 respectively. Find the corresponding self-informations.

Solution

$$\text{Self-information in a '0'} = I_0 = \log \frac{1}{p_0} = \log 4 = 2 \text{ bits.}$$

$$\text{Self Information in a "1"} = I_1 = \log \frac{1}{p_1} = \log_2 \frac{4}{3}$$

$$\therefore \quad I_1 = \frac{\log_{10} \frac{4}{3}}{\log_{10} 2} = \frac{ln \frac{4}{3}}{ln 2} = 0.415 \text{ bits}$$

Thus, it can be observed that more information is carried by a less likely message.

Logarithmic expression is chosen for measuring information because of the following reasons:

1. The information content or self-information of any message cannot be negative. Each message must contain certain amount of information.
2. The lowest possible self-information is "zero" which occurs for a sure event since P (sure event) = 1.

   The English phrase ***"Zero-Level Talk"*** might have been derived from this concept. Certain class of people, especially in public meetings, keep on talking without conveying much information to the general public. In such cases, the information conveyed is equal to zero. Thus, the lowest possible amount of information is equal to zero.
3. More information is carried by a less likely message. This statement is clear from example 1.1.

4. When independent symbols are transmitted, the total self-information must be equal to the sum of individual self-informations.

To prove this statement, let us suppose that two independent symbols S_K and S_l are transmitted with probabilities p_K and p_l respectively. Then the total self information is given by,

$$I_{Kl} = \log \frac{1}{P(S_K \text{ and } S_l)} = \log \frac{1}{P(S_K \cap S_l)}$$

$$= \log \frac{1}{P(S_K).P(S_l)} \text{ by using eqn (1.25)}$$

$$= \log \frac{1}{p_K \cdot p_l} = \log \frac{1}{p_K} + \log \frac{1}{p_l} = I_K + I_l$$

$$\therefore \quad I_{Kl} = I_K + I_l \qquad \text{...... (1.3)}$$

Thus, the total self-information is equal to the sum of individual self-informations.

As an example of independent events, suppose that you read two news items in the newspaper: (1) Earth quake rocks Gujarat state and (2) One of the ***"Spice girls"*** gets married to her boy friend. It is reasonable to assume that the two events mentioned in the news are independent and that the total information received from the two messages is same as the sum of the informations contained in each of the two news items.

Zero-Memory Source:- It represents a model of a discrete information source emitting a sequence of symbols from a fixed finite source alphabet $S = \{s_1, s_2, ..., s_q\}$. Successive symbols are selected according to some fixed probability law and are statistically independent of one another. This means that there is no connection between any two symbols and that the source has no memory. Such type of sources are called ***"memoryless"*** or ***"zero-memory"*** sources.

1.3 AVERAGE INFORMATION CONTENT (ENTROPY) OF SYMBOLS IN LONG INDEPENDENT SEQUENCES

Let us consider a zero-memory source producing independent sequences of symbols. While the receiver of these sequences may interpret the entire message as a single unit, communication systems often have to deal with individual symbols. Let us consider the source alphabet $S = \{s_1, s_2, ..., s_q\}$ with probabilities $P = \{p_1, p_2, ..., p_q\}$ respectively.

Let us consider a long independent sequence of length L symbols. This long sequence then contains

p_1L number of messages of type s_1,

p_2L number of messages of type s_2,

$\vdots \quad \vdots \quad \vdots \quad \vdots$

and $p_q L$ number of messages of type s_q.

From equation (1.2), the self-information of $s_1 = \log \frac{1}{p_1}$ bits.

$\therefore$ $p_1 L$ number of messages of type s_1 contain $p_1 L \log \frac{1}{p_1}$ bits of information,

$p_2 L$ number of messages of type s_2 contain $p_2 L \log \frac{1}{p_2}$ bits of information,

$\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots$

$p_q L$ number of messages of type s_q contain $p_q L \log \frac{1}{p_q}$ bits of information.

$\therefore$ The total self-information content of all these message symbols is given by,

$$I_{total} = p_1 L \log \frac{1}{p_1} + p_2 L \log \frac{1}{p_2} + ... + p_q L \log \frac{1}{p_q} \text{ bits}$$

$$= L \sum_{i=1}^{q} p_i \log \frac{1}{p_i} \text{ bits.}$$

$$\therefore \text{ Average self-information} = \frac{I_{total}}{L}$$

$$= \sum_{i=1}^{q} p_i \log \frac{1}{p_i} \text{ bits/message symbol.}$$

Average self-information is also called ***"ENTROPY"*** of source S denoted by H(S).

$$\therefore \quad \mathbf{H(S)} = \sum_{i=1}^{q} \mathbf{p_i \log \frac{1}{p_i}} \text{ bits/message symbol} \qquad ...... (1.4)$$

Note that the definition of H(S) given in equation (1.4) is based on time-averaging. This definition is valid for ensemble averages provided the source is ergodic. (Refer section 0.17).

The source entropy given by equation (1.4) is similar to the expression for entropy in statistical mechanics. The source entropy can be interpreted as follows. On the average, we can expect to get H(S) bits of information per symbol in long messages from the information source eventhough we cannot say in advance what symbol sequences will occur in these messages.

Thus H(S) represents the ***"average uncertainty"*** or the ***'average amount of surprise"*** of the source. The following illustrations helps us to understand the meaning of entropy clearly.

Illustration I :

Let us consider a binary source with source alphabet $S = \{s_1, s_2\}$ with probabilities

$$P = \left\{\frac{1}{256}, \frac{255}{256}\right\}$$

$$\text{Then, entropy } H(S) = \sum_{i=1}^{2} p_i \log \frac{1}{p_i}$$

$$= \frac{1}{256}\log 256 + \frac{255}{256}\log\frac{256}{255}$$

$$= \mathbf{0.037\ bits/message\ symbol}$$

$\therefore$ The average uncertainty is very very small and is relatively very easy to guess whether s_1 or s_2 will occur.

Illustration II :

$$\text{Let } S' = \{s_3, s_4\} \text{ with } P' = \left\{\frac{7}{16}, \frac{9}{16}\right\}$$

$$\text{Then, entropy } H(S') = \frac{7}{16}\log\frac{16}{7} + \frac{9}{16}\log\frac{16}{9}$$

$$= \mathbf{0.989\ bits/message\ symbol}$$

In this case, it is hard to guess whether s_3 or s_4 is transmitted.

Illustration III :

$$\text{Let } S'' = \{s_5, s_6\} \text{ with } P'' = \left\{\frac{1}{2}, \frac{1}{2}\right\}$$

$$\text{Then, entropy } H(S'') = \frac{1}{2}\log 2 + \frac{1}{2}\log 2 = 1 \text{ bit/message symbol}$$

In this case, the uncertainty is maximum for a binary source and becomes impossible to guess which symbol is transmitted.

These illustrations clearly indicate the significance and dependence of entropy on probabilities of messages.

INFORMATION RATE : Let us suppose that the symbols are emitted by the source at a fixed time rate "r_s" symbols/sec. The ***"average source information rate R_s"*** in bits/sec is defined as the product of the average information content per symbol and the message symbol rate r_s.

$$\therefore \quad R_s = r_s H(S) \text{ bits/sec or BPS} \qquad \text{...... (1.5)}$$

Example 1.2 : Consider a source $S = \{s_1, s_2, s_3\}$ with $P = \left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right\}$. Find (a) self information of each message, (b) entropy of source S.

Solution

(a) Self-information of $s_1 = I_1 = \log\frac{1}{p_1} = \log 2 = 1$ bits

Self-information of $s_2 = I_2 = \log\frac{1}{p_2} = \log 4 = 2$ bits

Self-information of $s_3 = I_3 = \log\frac{1}{p_3} = \log 4 = 2$ bits

(b) From equation (1.4) entropy is given by

$$H(S) = \sum_{i=1}^{q} p_i \log\frac{1}{p_i} = \sum_{i=1}^{q} p_i I_i \quad \text{using equation (1.2)}$$

In this example, $q = 3$

$$\therefore\ H(S) = \sum_{i=1}^{q=3} p_i I_i = p_1 I_1 + p_2 I_2 + p_3 I_3$$

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

$$\therefore\ \mathbf{H(S) = 1.5\ bits/message\ symbol}$$

Example 1.3 : The collector voltage of a certain circuit is to lie between –5 and –12 volts. The voltage can take on only these values –5, –6, –7, –9, –11, –12 volts with respective probabilities $\frac{1}{6}, \frac{1}{3}, \frac{1}{12}, \frac{1}{12}, \frac{1}{6}, \frac{1}{6}$. This voltage is recorded with a pen recorder. Determine the average self-information associated with the record in terms of bits/level.

Solution

From equation (1.4), for 6 levels

$$H(S) = \sum_{i=1}^{6} p_i \log\frac{1}{p_i} \text{ bits/level}$$

$$= \frac{1}{6}\log 6 + \frac{1}{3}\log 3 + \frac{1}{12}\log 12 + \frac{1}{12}\log 12 + \frac{1}{6}\log 6 + \frac{1}{6}\log 6$$

$$\therefore\ \mathbf{H(S) = 2.418\ bits/level}$$

Example 1.4 : Find relationship between Hartleys, nats and bits.

Solution

From definition of Hartleys, nats and bits, we have

$$I = \log_{10}\frac{1}{p} \text{ Hartleys} \quad \text{...... (1.6)}$$

$$I = \log_e\frac{1}{p} \text{ nats} \quad \text{...... (1.7)}$$

$$I = \log_2\frac{1}{p} \text{ bits} \quad \text{...... (1.8)}$$

From equation (1.6) $1 \text{ Hartley} = \frac{I}{\log_{10} 1/p} = \frac{\log_e 1/p \text{ nats}}{\log_{10} 1/p}$ using equation (1.7)

$$= \frac{-\log_e p}{-\log_{10} p} \text{ nats}$$

$$= \frac{\log_p 10}{\log_p e} \text{ nats} \quad \left[\because \log_a b = \frac{1}{\log_b a}\right]$$

$$\therefore \quad 1 \text{ Hartley} = \log_e 10 \text{ nats}$$

$$\text{or } \mathbf{1\ Hartley = 2.303\ nats} \quad \text{...... (1.9)}$$

$$\text{Similarly, } 1 \text{ Hartley} = \log_2 10 \text{ bits}$$

$$= \frac{1}{\log_{10} 2} \text{ bits}$$

$$\text{or } \mathbf{1\ Hartley = 3.32\ bits} \quad \text{...... (1.10)}$$

$$\text{and } \quad 1 \text{ nat} = \log_2 e \text{ bits}$$

$$\therefore \quad 1 \text{ nat} = \frac{1}{\log_e 2} = \frac{1}{\ln 2} \text{ bits}$$

$$\text{or } \mathbf{1\ nat = 1.443\ bits} \quad \text{...... (1.11)}$$

Example 1.5 : A discrete source emits one of six symbols once every m-sec. The symbol probabilities are $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}$ and $\frac{1}{32}$ respectively. Find the source entropy and information rate.

Solution

Source entropy H(S) is given by equation (1.4) as

$$H(S) = \sum_{i=1}^{6} p_i \log \frac{1}{p_i}$$

$$= \frac{1}{2}\log 2 + \frac{1}{4}\log 4 + \frac{1}{8}\log 8 + \frac{1}{16}\log 16 + \frac{1}{32}\log 32 + \frac{1}{32}\log 32$$

$$= \frac{1}{2} + \frac{1}{4} \times 2 + \frac{1}{8} \times 3 + \frac{1}{16} \times 4 + \frac{1}{32} \times 5 + \frac{1}{32} \times 5$$

$\therefore$ **H(S) = 1.9375 bits/message-symbol.**

Information Rate R_s is given by equation (1.5) as

$$R_s = r_s H(S)$$

In the problem given r_s = 1 message symbol/m-sec

= 10^3 message symbols/sec

$\therefore$ R_s = (10^3 message symbols/sec) (1.9375 bits/message symbol)

$\therefore$ **R_s = 1937.5 bits/sec.**

Example 1.6 : The output of an information source consists of 150 symbols, 32 of which occur with a probability of 1/64 and the remaining 118 occur with a probability of 1/236. The source emits 2000 symbols/sec. Assuming that the symbols are chosen independently, find the average information rate of this source.

Solution

The entropy is given by

$$H(S) = \sum_{i=1}^{150} p_i \log \frac{1}{p_i}$$

$$= \sum_{i=1}^{32} p_i \log \frac{1}{p_i} + \sum_{i=33}^{150} p_i \log \frac{1}{p_i}$$

$$= \left[\frac{1}{64}\log 64\right] \times 32 + \left[\frac{1}{236}\log 236\right] \times 118$$

H(S) = 6.9413 bits/message symbol

Given r_s = 2000 message symbols/sec.

$\therefore$ Average Information rate = $R_s = r_s H(S)$

= 2000 × 6.9413

R_s = 13,882.6 bits/sec

Example 1.7 : A code is composed of dots and dashes. Assuming that a dash is 3 times as long as a dot and has one-third the probability of occurrence. Calculate

(i) the information in a dot and a dash

(ii) the entropy of dot-dash code

(iii) the average rate of information if a dot lasts for 10 m-sec and this time is allowed between symbols.

Solution

Since only dots and dashes are present, we must have

$$p_{dot} + p_{dash} = 1$$

$$\text{Given} \quad p_{dash} = \frac{1}{3} p_{dot}$$

$$\text{Substituting, } p_{dot} + \frac{1}{3} p_{dot} = 1$$

$$\therefore \quad p_{dot} = \frac{3}{4}$$

$$\therefore \quad p_{dash} = \frac{1}{4}$$

(i) Information in a dot $= I_{dot} = \log \dfrac{1}{p_{dot}} = \log \dfrac{4}{3} = \mathbf{0.415\ bits}$

Information in a dash $= I_{dash} = \log \dfrac{1}{p_{dash}} = \log 4 = \mathbf{2\ bits}$

(ii) The entropy of dot dash code is

$$H(S) = p_{dot} \log \frac{1}{p_{dot}} + p_{dash} \log \frac{1}{p_{dash}}$$

$$= \frac{3}{4} \log \frac{4}{3} + \frac{1}{4} \log 4$$

$$\mathbf{H(S) = 0.8113\ bits/message\text{-}symbol}$$

(iii) Since $p_{dot} = \dfrac{3}{4}$ and $p_{dash} = \dfrac{1}{4}$, for every 4 symbols sent on an average, there will be 3 dots and 1 dash. With a dot lasting for 10 m-sec, a dash for 30 m-sec and with 10 m-sec gap in between two successive symbols, a total of 100 m-sec is required to transmit 4 symbols.

$$\therefore \text{ Symbol rate} = r_s = 4 \text{ symbols/100 m-sec.}$$
$$= 40 \text{ symbols/sec}$$

30 10 10 10 10 10 10 10

=100ms

$$\therefore \quad \text{Information rate} = R_s = r_s H(S)$$
$$= (40 \text{ symbols/sec}) (0.8113 \text{ bits/symbol})$$

$$\therefore \quad \mathbf{R_s = 32.452\ bits/sec}$$

Example 1.8 : A card is drawn from a deck.

(i) You are told it is a spade. How much information did you receive?

(ii) How much information did you receive if you are told that the card drawn is an ace?

(iii) If you are told that the card drawn is an ace of spades, how much information did you receive?

(iv) Is the information obtained in (iii) the sum of informations obtained in (i) and (ii)?

Solution

(i) Since there are 13 'spade' cards in a deck of 52 cards.

$$P_{spade} = \frac{13}{52} = 0.25$$

$$\therefore \text{ Self information } I_{spade} = \log\frac{1}{P_{spade}} = \log 4 = \mathbf{2\ bits}$$

(ii) Since there are 4 "Aces" in a deck of 52 cards,

$$P_{ace} = \frac{4}{52} = \frac{1}{13}$$

$$\therefore \text{ Self information } I_{ace} = \log\frac{1}{P_{ace}} = \log 13 = \mathbf{3.7\ bits}$$

(iii) Since there is only 'one' ace-of-spades in a deck of 52 cards.

$$P_{ace\text{-}of\text{-}spades} = \frac{1}{52}$$

$$\therefore \text{ Self-information } I_{ace\text{-}of\text{-}spades} = \log\frac{1}{P_{ace-of-spades}} = \log 52$$

$$= \mathbf{5.7\ bits}$$

(iv) Yes. The information obtained in (iii) is the sum of information obtained in (i) and (ii), since drawing a 'spades' card and drawing an 'ace' card are two independent events and total self-information must be equal to individual self-informations as shown in equation (1.3).

Example 1.9 : A source emits one of 4 possible symbols x_0 to x_3 during each signalling interval. The symbols occur with probabilities as given in table 1.1:

Symbol	Probability
x_0	$p_0 = 0.4$
x_1	$p_1 = 0.3$
x_2	$p_2 = 0.2$
x_3	$p_3 = 0.1$

Table 1.1 : Source probabilities of example 1.9

Find the amount of information gained by observing the source emitting each of these symbols and also the entropy of the source.

Solution

From equation (1.2), the self-information I_K is given by

$$I_K = \log_2 \frac{1}{p_K} \text{ bits}$$

when $K = 0,\ I_0 = \log_2 \frac{1}{p_0} = \log_2 \frac{1}{0.4} = 1.322 \text{ bits}$

$K = 1,\ I_1 = \log_2 \frac{1}{p_1} = \log_2 \frac{1}{0.3} = 1.737 \text{ bits}$

$K = 2,\ I_2 = \log_2 \frac{1}{p_2} = \log_2 \frac{1}{0.2} = 2.322 \text{ bits}$

$K = 3,\ I_3 = \log_2 \frac{1}{p_3} = \log_2 \frac{1}{0.1} = 3.322 \text{ bits}$

From equation (1.4), the entropy of the source is given by

$$H(X) = \sum_{K=0}^{3} p_K \log \frac{1}{p_K} \text{ bits / message symbol.}$$

$$= \sum_{K=0}^{3} p_K I_K$$

$$= p_0 I_0 + p_1 I_1 + p_2 I_2 + p_3 I_3$$

$$= (0.4)(1.322) + (0.3)(1.737) + (0.2)(2.322) + (0.1)(3.322)$$

$$\therefore\ \mathbf{H(X) = 1.8465 \text{ bits/message symbol.}}$$

Example 1.10 : A source emits one of four symbols S_0, S_1, S_2 and S_3 with probabilites $\frac{1}{3}$, $\frac{1}{6}$, $\frac{1}{4}$ and $\frac{1}{4}$ respectively. The successive symbols emitted by the source are statistically independent. Calculate the entropy of the source.

Solution

From equation (1.4), the entropy of the source is given by

$$H(S) = \sum_{i=0}^{3} p_i \log \frac{1}{p_i}$$

$$= p_0 \log \frac{1}{p_0} + p_1 \log \frac{1}{p_1} + p_2 \log \frac{1}{p_2} + p_3 \log \frac{1}{p_3}$$

$$= \frac{1}{3}\log 3 + \frac{1}{6}\log 6 + \frac{1}{4}\log 4 + \frac{1}{4}\log 4$$

$$= 0.5282 + 0.43082 + 0.5 + 0.5$$

$$\mathbf{H(S) = 1.95914\ bits/message\ symbol}$$

Example 1.11 : Let X represent the outcome of a single roll of a fair die. What is the entropy of X?

Solution

Since a die is having six faces, we note that the probability of getting any number is $\frac{1}{6}$.

$$\therefore\ p_x = \frac{1}{6}$$

$$\therefore\ \text{Entropy of X} = H(X) = p_x \log \frac{1}{p_x} = \frac{1}{6}\log 6 = 0.431 \text{ bits}$$

Example 1.12 : A fair coin is tossed repeatedly. Let

A = {event of getting 3 heads out of 5 trials}

B = {event of getting 5 heads out of 8 trials}

Which event conveys more information? Support your answers by numerical computation of respective amounts of information. *[V sem EC/TE, July/Aug 2003, Q.1 (b)]*

Solution

Let us first find out the probabilities of events A and B by using the concept of binomial distribution discussed in section 0.10 in page 19.

Let us define the random variable

X = number of heads

= binomial random variable (since tossing a coin results in only 2 outcomes → head or tail)

n = number of trials

p = probability of getting head = 0.5

q = 1 − p = 0.5

From equation (0.47), we have

$$P(X = x) = {}_nC_x\, p^x q^{n-x}$$

$$\therefore\quad P(A) = P(X = 3) = {}_5C_3\,(0.5)^3\,(0.5)^{5-3} = 0.3125$$

And $$P(B) = P(X = 5) = {}_8C_5\,(0.5)^5\,(0.5)^{8-5} = 0.21875$$

$$\therefore\quad \text{Self-information of event A} = I_A = \log_2 \frac{1}{P(A)} = \log_2 \frac{1}{0.3125} = 1.678 \text{ bits}$$

And Self-information of event B = $I_B = \log_2 \frac{1}{P(B)} = \log_2 \frac{1}{0.21875} = 2.193$ bits

$\therefore$ **Event-B conveys more amount of information.**

Example 1.13 : Find the entropy of a source in nats/symbol of a source that emits one out of four symbols A, B, C and D in a statistically independent sequence with probabilities $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}$ and $\frac{1}{8}$. ***[VI sem EC/TE Jan/Feb. 2005, Q 1 (a)]***

Solution

From equation (1.4), the entropy of the source is given by

$$H(S) = \sum_{i=A}^{D} p_i \log \frac{1}{p_i}$$

$$= p_A \log \frac{1}{p_A} + p_B \log \frac{1}{p_B} + p_C \log \frac{1}{p_C} + p_D \log \frac{1}{p_D}$$

$$= \frac{1}{2} \log 2 + \frac{1}{4} \log 4 + \frac{1}{8} \log 8 + \frac{1}{8} \log 8$$

$$= \frac{1}{2} + \frac{1}{4} \times 2 + \frac{1}{8} \times 3 + \frac{1}{8} \times 3$$

$$= 1.75 \text{ bits/symbol.}$$

But from equation (1.11)

$$1 \text{ nat} = 1.443 \text{ bits}$$

or

$$1 \text{ bit} = \frac{1}{1.443} \text{ nat} = 0.693 \text{ nats}$$

$\therefore$ Entropy of source $H(S) = (1.75)\ (0.693$ nats/symbol)

$\therefore$ **H(S) = 1.213 nats/symbol**

Example 1.14 : A binary source is emitting an independent sequence of '0's and '1's with probabilities p and 1 – p respectively. Plot the entropy of the source versus p. ***[VI sem EC/TE, Aug. 2001, Q.1 (a)]***

Solution

From equation (1.4), the entropy of the binary source is given by

$$H(S) = \sum_{i=1}^{2} p_i \log \frac{1}{p_i}$$

$$= p_1 \log \frac{1}{p_1} + p_2 \log \frac{1}{p_2}$$

$$\therefore \quad \mathbf{H(S) = p \log \frac{1}{p} + (1-p) \log \frac{1}{(1-p)}}$$

Let $p = 0.1$, $\quad H(S) = 0.1 \log \frac{1}{0.1} + 0.9 \log \frac{1}{0.9} = 0.469$ bits/symbol.

Let $p = 0.2$, $\quad H(S) = 0.2 \log \frac{1}{0.2} + 0.8 \log \frac{1}{0.8} = 0.722$ bits/symbol.

Let $p = 0.3$, $\quad H(S) = 0.3 \log \frac{1}{0.3} + 0.7 \log \frac{1}{0.7} = 0.881$ bits/symbol.

Let $p = 0.4$, $\quad H(S) = 0.4 \log \frac{1}{0.4} + 0.6 \log \frac{1}{0.6} = 0.971$ bits/symbol.

Let $p = 0.5$, $\quad H(S) = 0.5 \log \frac{1}{0.5} + 0.5 \log \frac{1}{0.5} = 1$ bits/symbol.

Let $p = 0.6$, $\quad H(S) = 0.6 \log \frac{1}{0.6} + 0.4 \log \frac{1}{0.4} = 0.971$ bits/symbol.

Let $p = 0.7$, $\quad H(S) = 0.7 \log \frac{1}{0.7} + 0.3 \log \frac{1}{0.3} = 0.881$ bits/symbol.

Let $p = 0.8$, $\quad H(S) = 0.8 \log \frac{1}{0.8} + 0.2 \log \frac{1}{0.2} = 0.722$ bits/symbol.

Let $p = 0.9$, $\quad H(S) = 0.9 \log \frac{1}{0.9} + 0.1 \log \frac{1}{0.1} = 0.469$ bits/symbol.

For $p = 0$ and $p = 1$, $H(S) = 0$.

The above calculated values can be tabulated as shown in table 1.2.

p	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
H(S) bits/m-sym	0	0.469	0.722	0.881	0.971	1	0.971	0.881	0.722	0.469	0

Table 1.2 : Entropy values for various probabilities

The entropy $\dot{H}(S)$ can now be plotted as a function of p as shown in figure 1.2.

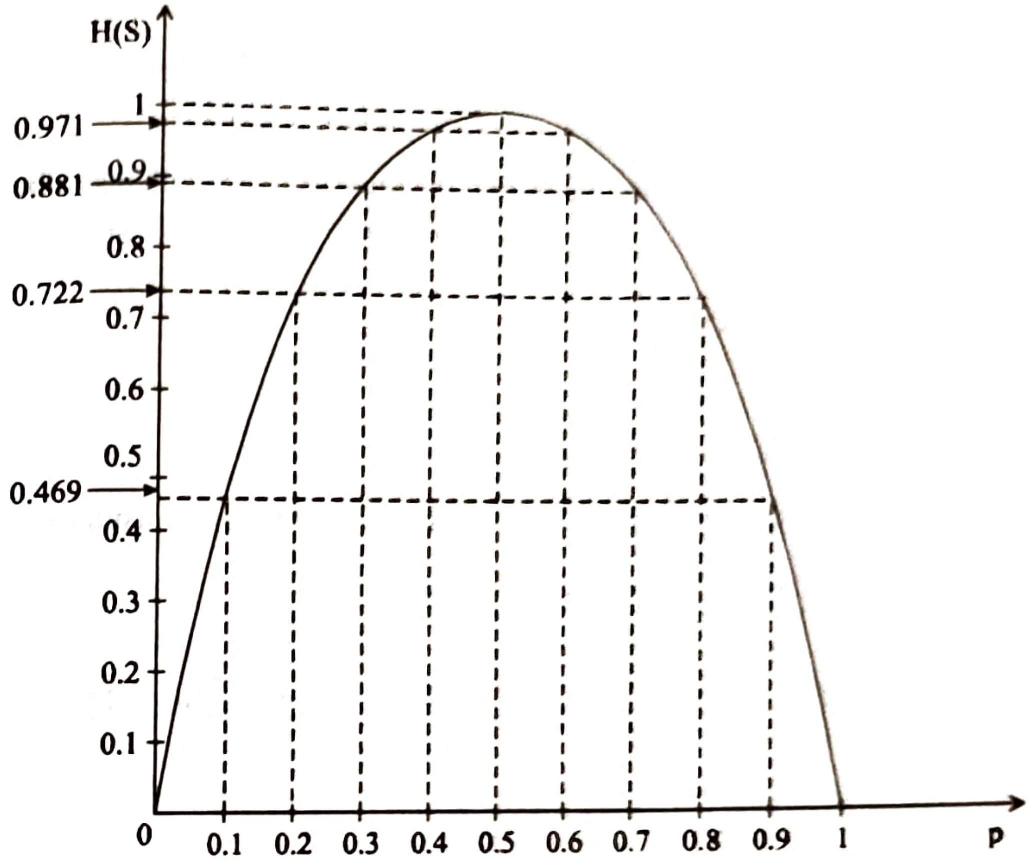

Fig. 1.2 : Plot of H(S) versus p of example 1.14

PROPERTIES OF ENTROPY

The entropy function is given by equation (1.4) for a source alphabet $S = \{s_1, s_2, \ldots, s_q\}$ with $P = \{p_1, p_2, \ldots, p_q\}$ where q = number of source symbols, as

$$H(S) = \sum_{i=1}^{q} p_i \log \frac{1}{p_i} = \sum_{i=1}^{q} p_i I_i \text{ bits/message-symbol}$$

Many interesting properties can be observed as listed below:

1. The entropy function is continuous for every independent variable p_K in the interval (0, 1). i.e., if p_K varies continuously from 0 to 1, so does the entropy function.
   [Note: Entropy function vanishes at both $p_K = 0$ and $p_K = 1$].
2. The entropy function is a symmetrical function of its arguments.
   i.e., $H[p_K, (1 - p_K)] = H[(1 - p_K), p_K]$ for all K = 1,2,....., q ...... (1.12)
   i.e., the value of H(S) remains the same irrespective of the locations of the probabilities. i.e., as long as the probabilities are same, it does not matter in which order they are arranged. Thus the sources S_A, S_B and S_C with probabilities.

$$\left.\begin{aligned} P_A &= \{p_1, p_2, p_3\} \\ P_B &= \{p_2, p_3, p_1\} \\ P_C &= \{p_3, p_1, p_2\} \end{aligned}\right] \begin{aligned} &\text{such that } \sum_{i=1}^{3} p_i = 1 \text{ will all have the same entropy.} \\ &\text{i.e., } H(S_A) = H(S_B) = H(S_C) \end{aligned}$$

3. **EXTREMAL PROPERTY :** Let us consider the same source S with q symbols $S = \{s_1, s_2, ..., s_q\}$ with probabilities $P = \{p_1, p_2, ..., p_q\}$. The entropy of S is given by equation (1.4) as

$$H(S) = \sum_{i=1}^{q} p_i \log \frac{1}{p_i}$$

$$\text{and } \sum_{i=1}^{q} p_i = 1 \text{ from equation (1.1)}$$

Let us now prove that the entropy H(S) has an ***"upper bound"*** by considering a quantity [log q – H(S)].

[It is obvious that the ***"lower bound"*** for H(S) is zero].

$$\therefore \quad \log q - H(S) = \left[\sum_{i=1}^{q} p_i\right] \log q - \sum_{i=1}^{q} p_i \log \frac{1}{p_i} \quad \left[\text{Since } \sum_{i=1}^{q} p_i = 1\right]$$

$$= \sum_{i=1}^{q} p_i \left[\log q - \log \frac{1}{p_i}\right]$$

$$= \sum_{i=1}^{q} p_i \log q p_i$$

$$= \sum_{i=1}^{q} p_i \frac{\log_e q p_i}{\log_e 2}$$

$$\therefore \quad \log q - H(S) = \log_2 e \sum_{i=1}^{q} p_i \, ln \, q p_i \qquad ...... (1.13)$$

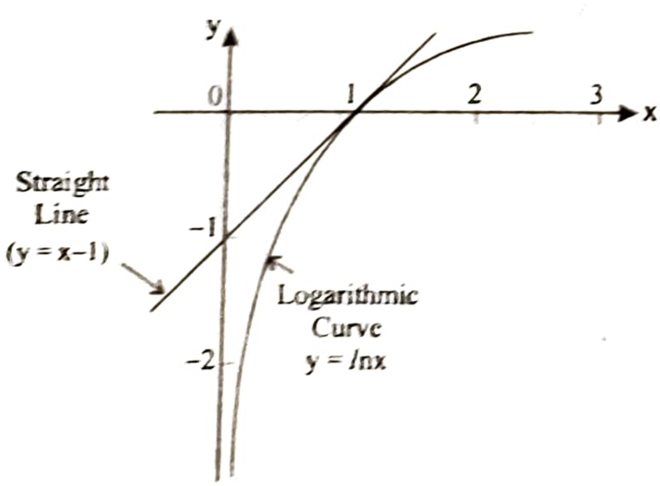

Fig. 1.3 : Illustrating graph of straight line y = x – 1 and logarithmic curve y = lnx

From the graph of figure 1.3 it is evident that the straight line $y = x - 1$ always lies above the logarithmic curve $y = ln\, x$ except at $x = 1$. Thus the straight line forms the "tangent" to the curve at $x = 1$.

$$\therefore \quad ln\, x \le x - 1$$

Multiplying by -1, we get

$$-ln\, x \ge 1 - x$$

$$\text{or} \quad ln\frac{1}{x} \ge 1 - x \qquad \text{...... (1.14)}$$

$$\text{If } x = \frac{1}{qp_i}, \text{ then } ln\, qp_i \ge 1 - \frac{1}{qp_i} \qquad \text{...... (1.15)}$$

Multiplying equation (1.15) both sides by p_i and then taking summation for all $i = 1,2,...,q$, we get

$$\sum_{i=1}^{q} p_i\, ln\, qp_i \ge \sum_{i=1}^{q} p_i \left[1 - \frac{1}{qp_i}\right]$$

Multiplying both sides by $\log_2 e$, we get

$$\log_2 e \sum_{i=1}^{q} p_i\, ln\, qp_i \ge \log_2 e \left[\sum_{i=1}^{q} p_i - \sum_{i=1}^{q} \frac{1}{q}\right] \qquad \text{...... (1.16)}$$

From equation (1.13), L.H.S. of equation (1.16) is $\log q - H(S)$ and R.H.S. of equation (1.16) is always 0.

$$\therefore \quad \log q - H(S) \ge 0$$

$$\text{or} \quad \mathbf{H(S) \le \log_2 q} \qquad \text{...... (1.17)}$$

The equality sign holds good when $p_i - \frac{1}{q} = 0$ for all values of i.

$$\text{i.e.,} \quad p_i = \frac{1}{q} \text{ for all } i = 1, 2, ..., q \qquad \text{...... (1.18)}$$

When the condition of equation (1.18) is satisfied, then the entropy becomes maximum given by

$$\mathbf{H(S)_{max} = \log_2 q \text{ bits/message-symbol}} \qquad \text{...... (1.19)}$$

i.e., the entropy attains a ***"maximum value"*** when all the source symbols become ***"equiprobable"***.

Illustration I :

From example (1.2) for $S = \{s_1, s_2, s_3\}$ and $P = \left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right\}$, we had

$$H(S) = 1.5 \text{ bits/message-symbol}$$

$$\therefore \quad H(S)_{max} = \log_2 3 \text{ since } q = 3$$

$$\therefore \quad H(S)_{max} = 1.585 \text{ bits/message-symbol and}$$

this maximum value occurs when $p_i = \frac{1}{q}$ for all i = 1, 2, 3

i.e., $p_1 = p_2 = p_3 = \frac{1}{3}$ i.e., when all message symbols become equiprobable.

Illustration II :

For $S = \{s_1, s_2, s_3, s_4\}$ and $P = \left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\right\}$, the entropy is given by

$$H(S) = \sum_{i=1}^{4} p_i \log \frac{1}{p_i}$$

$$= \frac{1}{2}\log 2 + \frac{1}{4}\log 4 + \frac{1}{8}\log 8 + \frac{1}{8}\log 8$$

$$= 1.75 \text{ bits/message-symbol}$$

From equation (1.19),

$$H(S)_{max} = \log_2 q$$

$$= \log_2 4$$

$$\therefore \quad H(S)_{max} = 2 \text{ bits/message-symbol}$$

which occurs when $p_1 = p_2 = p_3 = p_4 = \frac{1}{4}$ (i.e., equiprobable)

4. **PROPERTY OF ADDITIVITY :** Suppose that we split the symbol s_q into 'n' sub-symbols such that $s_q = s_{q_1}, s_{q_2}, \ldots\ldots, s_{q_n}$. Occuring with probabilities $p_{q_1}, p_{q_2}, \ldots\ldots, p_{q_n}$ such that

$$p_q = p_{q_1} + p_{q_2} + \ldots\ldots, + p_{q_n} = \sum_{j=1}^{q} p_{q_j} \qquad \ldots\ldots (1.20)$$

Then, the splitted symbol entropy is

$$H'(S) = H(p_1, p_2, \ldots\ldots, p_{q-1}, p_{q_1}, p_{q_2}, \ldots\ldots, p_{q_n})$$

$$= \sum_{i=1}^{q-1} p_i \log \frac{1}{p_i} + \sum_{j=1}^{n} p_{q_j} \log \frac{1}{p_{q_j}}$$

$$H'(S) = \sum_{i=1}^{q} p_i \log \frac{1}{p_i} - p_q \log \frac{1}{p_q} + \sum_{j=1}^{n} p_{q_j} \log \frac{1}{p_{q_j}}$$

$$= \sum_{i=1}^{q} p_i \log\frac{1}{p_i} - \sum_{i=1}^{n} p_{q_j} \log\frac{1}{p_q} + \sum_{j=1}^{n} p_{q_j} \log\frac{1}{p_{q_j}} \quad \text{using eqn. (1.20)}$$

$$= \sum_{i=1}^{q} p_i \log\frac{1}{p_i} + \sum_{j=1}^{n} p_{q_j}\left[\log\frac{1}{p_{q_j}} - \log\frac{1}{p_q}\right]$$

$$= \sum_{i=1}^{q} p_i \log\frac{1}{p_i} + p_q \sum_{j=1}^{n} \frac{p_{q_j}}{p_q}\left[\log\frac{p_q}{p_{q_j}}\right]$$

$\therefore \quad H'(S) = H(S) + \text{a positive quantity since } p_{q_j} \le p_q$ for all j

$\therefore \quad H'(S) \ge H(S)$ ...... (1.21)

i.e., partitioning of symbols into sub-symbols cannot decrease the entropy.

5. The "**SOURCE EFFICIENCY**" denoted by η_s is given by

$$\eta_s = \frac{H(S)}{H(S)_{max}} \qquad \text{...... (1.22)}$$

and the "**SOURCE REDUNDANCY**" denoted by R_{η_s} is given by

$$R_{\eta s} = 1 - \eta_s \qquad \text{...... (1.23)}$$

Note : η_s and R_{η_s} can also be expressed as percentage.

Example 1.15 : Verify the rule of additivity for the following source

$$S = \{s_1, s_2, s_3, s_4\} \text{ with } P = \left\{\frac{1}{2}, \frac{1}{3}, \frac{1}{12}, \frac{1}{12}\right\} = \{p_1, p_2, p_3, p_4\} \text{ (say)}$$

Solution

$$H(S) = \frac{1}{2}\log 2 + \frac{1}{3}\log 3 + 2 \times \frac{1}{12}\log 12$$

$\therefore \quad$ **H(S) = 1.6258 bits/message-symbol**

From the property of additivity, we have

$$H'(S) = p_1 \log\frac{1}{p_1} + (1-p_1)\log\frac{1}{(1-p_1)} +$$

$$(1-p_1)\left\{\frac{p_2}{(1-p_1)}\log\frac{(1-p_1)}{p_2} + \frac{p_3}{(1-p_1)}\log\frac{(1-p_1)}{p_3} + \frac{p_4}{(1-p_1)}\log\frac{(1-p_1)}{p_4}\right\}$$

$$\therefore \quad H'(S) = \frac{1}{2}\log 2 + \frac{1}{2}\log 2 + \frac{1}{2}\left\{\frac{\left(\frac{1}{3}\right)}{\left(\frac{1}{2}\right)}\log\frac{\left(\frac{1}{2}\right)}{\left(\frac{1}{3}\right)} + \frac{\left(\frac{1}{12}\right)}{\left(\frac{1}{2}\right)}\log\frac{\left(\frac{1}{2}\right)}{\left(\frac{1}{12}\right)} + \frac{\left(\frac{1}{12}\right)}{\left(\frac{1}{2}\right)}\log\frac{\left(\frac{1}{2}\right)}{\left(\frac{1}{12}\right)}\right\}$$

$\therefore \quad$ **H'(S) = 1.6258 bits/message-symbol which is same as H(S)**

Example 1.16 : A certain data source has 8 symbols that are produced in blocks of four at a rate of 500 blocks/sec. The first symbol in each block is always the same (presumably for synchronization). The remaining three are filled by any of the 8 symbols with equal probability. What is the entropy rate of this source?

Solution

A block of 4 symbols can be illustrated as shown in figure 1.4.

Given q = 8 symbols

r_s = 500 blocks/sec.

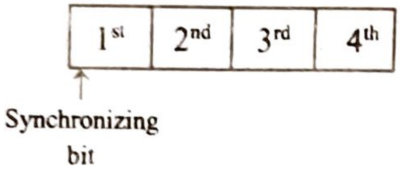

Fig. 1.4 : Illustrating example (1.10)

The symbols are produced in blocks of four, the total entropy H_T is

$$H_T = H_1 + H_2 + H_3 + H_4 \qquad \text{...... (1.24)}$$

where H_i is the entropy of the i^{th} bit.

Since 1^{st} bit is always same, no information is carried by that sure event and hence

$$\mathbf{H_1 = 0}$$

Given that the 2^{nd} position is occupied by any of the 8 symbols with equal probability

$$\therefore \quad H_2 = H(S)_{max} = \log_2 q \text{ from equation (1.19)}$$

$$\therefore \quad H_2 = \log_2 8 = 3 \text{ bits/m-sym}$$

Since 3^{rd} and 4^{th} positions are also occupied by any of the 8 equiprobable symbols, we have

$$H_3 = H_4 = 3 \text{ bits/message symbol}$$

Substituting these values of individual entropies in equation (1.44), we get

$$H_T = 0 + 3 + 3 + 3 = 9 \text{ bits/block}$$

$$\therefore \quad \text{Entropy rate} = \text{Information rate} = R_s = r_s H_T \,.$$

$$\therefore \quad R_s = (500 \text{ blocks/sec})\,(9 \text{ bits/block})$$

$$\therefore \quad \mathbf{R_s = 4500 \text{ bits/sec.}}$$

Example 1.17 : A black and white TV picture consists of 525 lines of picture information. Assume that each line consists of 525 picture elements (pixels) and that each element can have 256 brightness levels. Picture are repeated at the rate of 30 frames/sec. Calculate the average rate of information conveyed by a TV set to a viewer.

Solution

Figure 1.5 shows a TV frame consisting of 525 horizontal lines with each line containing 525 pixels.

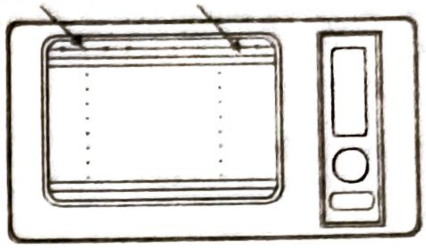

Fig. 1.5 : Illustrating a TV frame

Total number of pixels in one frame = 525 × 525 = 2,75,625 pixels. It is given that each pixel can have 256 different brightness levels from fully black to fully white. Then the total number of different frames possible

$$= (256)^{2,75,625} \text{ frames.}$$

Let us assume that all these frames occur with equal probability, the net maximum information (average) content per frame is

$$I = H(S)_{max} = \log_2 q = \log_2 (256)^{2,75,625} = 275625 \log_2 256$$

$$= 22.05 \times 10^5 \text{ bits/frame}$$

It is given that $r_s = 30$ frames/sec

∴ The average rate of information (maximum) is given by

$$R_s = r_s I$$

$$= (30 \text{ frames/sec})(22.05 \times 10^5 \text{ bits/frames})$$

$$\therefore \quad \mathbf{R_s = 66.15 \times 10^5 \text{ bits/sec}}$$

Example 1.18 : Suppose that S_1 and S_2 are two zero-memory sources with probabilities $p_1, p_2, ..., p_n$ for source S_1 and $q_1, q_2, ..., q_n$ for source S_2. Show that the entropy of source S_1

$$H(S_1) \leq \sum_{K=1}^{n} p_K \log \frac{1}{q_K}$$

Solution

Given S_1 and S_2 are zero-memory information sources

$$\therefore \quad H(S_1) = \sum_{K=1}^{n} p_K \log \frac{1}{p_K} \qquad \text{...... (1.25)}$$

$$\text{and} \quad \sum_{K=1}^{n} p_K = 1 \qquad \text{...... (1.26)}$$

$$\text{Similarly} \quad H(S_2) = \sum_{K=1}^{n} q_K \log \frac{1}{q_K} \qquad \text{...... (1.27)}$$

$$\text{and} \quad \sum_{K=1}^{n} q_K = 1 \qquad \text{...... (1.28)}$$

Consider $H(S_1) - \sum_{K=1}^{n} p_K \log \frac{1}{q_K} = \sum_{K=1}^{n} p_K \log \frac{1}{p_K} - \sum_{K=1}^{n} p_K \log \frac{1}{q_K}$

$$= \sum_{K=1}^{n} p_K \left[\log \frac{1}{p_K} - \log \frac{1}{q_K} \right]$$

$$= \sum_{K=1}^{n} p_K \log \frac{q_K}{p_K} = \sum_{K=1}^{n} p_K \frac{ln\left(\frac{q_K}{p_K}\right)}{ln 2}$$

$$= \log_2 e \sum_{K=1}^{n} p_K \, ln\left(\frac{q_K}{p_K}\right) \quad \text{...... (1.29)}$$

From equation (1.14)

$$ln \frac{1}{x} \geq 1 - x$$

$$-lnx \geq 1 - x$$

Removing -1 on both sides, $lnx \leq x - 1$

With $x = \frac{q_K}{p_K}$, we have,

$$ln\left(\frac{q_K}{p_K}\right) \leq \left[\frac{q_K}{p_K} - 1\right]$$

Multiplying by p_K, taking summation for all K = 1, 2, ..., n and then multiplying by $\log_2 e$, on both sides, we get

$$\log_2 e \sum_{K=1}^{n} p_K \, ln\left(\frac{q_K}{p_K}\right) \leq \log_2 e \sum_{K=1}^{n} p_K \left[\frac{q_K}{p_K} - 1\right] \quad \text{...... (1.30)}$$

Using equation (1.29) in (1.30), we get

$$H(S_1) - \sum_{K=1}^{n} p_K \log \frac{1}{q_K} \leq \log_2 e \left[\sum_{K=1}^{n} q_K - \sum_{K=1}^{n} p_K \right]$$

$\leq \log_2 e \, [1 - 1]$ using equations (1.26) and (1.28)

≤ 0

$$\therefore \quad H(S_1) \leq \sum_{K=1}^{n} p_K \log \frac{1}{q_K} \rightarrow \text{proved.}$$

Example 1.19 : Find the information content of a message that consists of a digital word 9 digits long in which each digit may take on one of five possible levels. The probability of sending any of the five levels is assumed to be equal, and the level in any digit does not depend on the values taken by previous digits.

Solution

Given length of digital word = 9 digits

Number of levels = 5 equiprobable.

The digital word of 9 digits can be illustrated as shown in table 1.3.

1^{st} Digit	2^{nd} Digit	....	9^{th} Digit
↑ Entropy H_1	↑ Entropy H_2	....	↑ Entropy H_9

Table 1.3 : Table of entropy for example 1.19

The total entropy of 9 digit word is given by

$$H_T = H_1 + H_2 + \ldots\ldots + H_9$$

where H_i is the entropy of the i^{th} digit, $i = 1, 2, \ldots., 9$. Since each digit can take on one of five equiprobable levels, the entropy of each digit is given by equation (1.19) as

$$H = H(S)_{max} = \log_2 q = \log_2 5 = 2.322 \text{ bits/level}$$

$$\therefore \quad H_1 = H_2 = \ldots\ldots = H_9 = 2.322 \text{ bits/level.}$$

$\therefore$ Total information content $H_T = (9)(1.322)$

$$\therefore \quad \mathbf{H_T = 20.898 \text{ bits/level.}}$$

Example 1.20 : A pair of dice are tossed simultaneously. The outcome of the first dice is recorded as x_1 and that of second dice as x_2. Two events are defined as follows:

3.a

$$A = \{(x_1, x_2) \text{ such that } x_1 + x_2 \le 7\}$$

$$B = \{(x_1, x_2) \text{ such that } x_1 > x_2\}$$

Which event conveys more information? Support your answer by numerical computation.

Solution

When a pair of dice of tossed simultaneously, then the sample space S consists of 36 combinations of (x_1, x_2) given by

$$S = \begin{bmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{bmatrix}$$

Each pair in the sample space above can occur with a probability $= \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{1}{36}$ and hence all the pairs are equiprobable.

The event A contains the pairs given by

$$A = \{(1, 1) (1, 2) (1, 3) (1, 4) (1, 5) (1, 6) (2, 1) (2, 2) (2, 3) (2, 4) (2, 5) (3, 1) (3, 2) (3, 3) (3, 4) (4, 1) (4, 2) (4, 3) (5, 1) (5, 2) (6, 1)\}$$

And the event B contains the pairs

$$B = \{(2, 1) (3, 1) (3, 2) (4, 1) (4, 2) (4, 3) (5, 1) (5, 2) (5, 3) (5, 4) (6, 1) (6, 2) (6, 3) (6, 4) (6, 5)\}$$

Since all the pairs are equiprobable.

$$P(A) = \frac{21}{36} = \frac{7}{12} \quad \text{and} \quad P(B) = \frac{15}{36} = \frac{5}{12}$$

$\therefore$ Self-information of event A $= I_A = \log \frac{1}{P(A)} = \log \frac{12}{7} = 0.7776$ bits

$\therefore$ Self-information of event B $= I_B = \log \frac{1}{P(B)} = \log \frac{12}{5} = 1.263$ bits

$\therefore$ $\mathbf{I_B > I_A}$

The event B conveys more information than event A.

Example 1.21 : A discrete message source "S" emits two independent symbols X and Y with probabilities 0.55 and 0.45 respectively. Calculate the efficiency of the source and its redundancy.

Solution

Given $p_x = P(X) = 0.55, \; p_y = P(Y) = 0.45.$

$$\therefore \quad \text{Entropy } H(S) = \sum_{i=x}^{y} p_i \log \frac{1}{p_i}$$

$$= p_x \log \frac{1}{p_x} + p_y \log \frac{1}{p_y}$$

$$= 0.55 \log \frac{1}{0.55} + 0.45 \log \frac{1}{0.45}$$

$$= 0.9928 \text{ bits/message symbol.}$$

From equation (1.19), the maximum entropy is given by

$$H(S)_{max} = \log_2 q = \log_2 2 = 1 \text{ bit/message symbol.}$$

From equation (1.22), the source efficiency is given by

$$\eta_s = \frac{H(S)}{H(S)_{max}} = \frac{0.9928}{1}$$

$$\therefore \quad \eta_s = \mathbf{99.28\%}$$

From equation (1.23), the source redundancy is given by

$$R_{\eta s} = 1 - \eta_s = 1 - 0.9928 = 0.0072$$

$$\therefore \quad \mathbf{R_{\eta s} = 0.72\%}$$

Example 1.22 : Shortly before a horse-race, a book-maker believes that several horses entered in the race have the following probability of winning:

Horse →	A	B	C	D	E
P (winning) →	0.04	0.42	0.31	0.12	0.11

He, then receives a message that owing to a minor injury, one of the horses is not participating in the race. Explain, how would you assess from an information theory point of view, the information value of this message.

(i) if the horse in question is known

(ii) if it is not known.

Solution

(i) The information value of the message when the horse in question is **known** is nothing but the **self-information**.

If the horse 'A' is not running, then the self-information of the message

$$I_A = \log \frac{1}{P(A)} = \log \frac{1}{0.04} = 4.644 \text{ bits}$$

If the horse 'B' is not running, then

$$I_B = \log \frac{1}{P(B)} = \log \frac{1}{0.42} = 1.252 \text{ bits}$$

If the horse 'C' is not running, then

$$I_C = \log \frac{1}{P(C)} = \log \frac{1}{0.31} = 1.69 \text{ bits}$$

If the horse 'D' is not running, then

$$I_D = \log \frac{1}{P(D)} = \log \frac{1}{0.12} = 3.059 \text{ bits}$$

And if the horse 'E' is not running, then

$$I_E = \log \frac{1}{P(E)} = \log \frac{1}{0.11} = 3.184 \text{ bits.}$$

(ii) The information value of the message when the horse in question is **not known** is nothing but the **average information** given by

$$H(S) = \sum_{i=A}^{E} p(i) \log \frac{1}{p(i)}$$

$$= P(A) \log \frac{1}{P(A)} + P(B) \log \frac{1}{P(B)} + P(C) \log \frac{1}{P(C)}$$

$$+ P(D) \log \frac{1}{P(D)} + P(E) \log \frac{1}{P(E)}$$

$$= 0.04 \log \frac{1}{0.04} + 0.42 \log \frac{1}{0.42} + 0.31 \log \frac{1}{0.31}$$

$$+ 0.12 \log \frac{1}{0.12} + 0.11 \log \frac{1}{0.11}$$

$$\therefore \quad H(S) = (0.04)(4.644) + (0.42)(1.252) + (0.31)(1.69) + (0.12)(3.059) + (0.11)(3.184)$$

$$\therefore \quad \mathbf{H(S) = 1.953 \text{ bits/horse}}$$

Example 1.23 : In a facsimile transmission of picture, there are about 2.25×10^6 pixels/frame. For a good reproduction 12 brightness levels are necessary. Assume all these levels are equally likely to occur. Find the rate of information if one picture is to be transmitted every 3 minutes. What is the source efficiency of this facsimile transmitter?

[VI sem EC/TE, Jan/Feb, 2003, Q.1 (b)]

Solution

$$\text{Total number of pixels in one frame} = 2.25 \times 10^6$$

$$\text{Number of brightness levels} = 12$$

$$\therefore \quad \text{Total number of different frames possible} = (12)^{2.25 \times 10^6}$$

Since all the levels are equally likely to occur, the net maximum information content per frame is

$$I = H(S)_{max} = \log_2 q = \log_2 (12)^{2.25 \times 10^6}$$

$$= 2.25 \times 10^6 \log_2 12$$

$$\therefore \quad I = 8.066 \times 10^6 \text{ bits/picture}$$

Given that one picture is transmitted in 3 minutes.

Therefore, the rate of transmission is given by

$$r_s = \frac{1}{3 \text{ minutes}} = \frac{1}{3 \times 60} \text{ picture/sec}$$

$\therefore$ The average rate of information is given by

$$R_s = r_s I = \frac{1}{3 \times 60} \times 8.066 \times 10^6$$

$$\therefore \quad \mathbf{R_s = 44812 \text{ bits/sec}}$$

Since the information transmitted is maximum (as all levels are equiprobable) the source efficiency

$$\eta_s = \frac{H(S)}{H(S)_{max}} = \frac{H(S)_{max}}{H(S)_{max}} = 1 = 100\%$$

Example 1.24 : A discrete memoryless source produces four symbols with probabilities p_1, p_2, p_3 and p_4. Using partial differentiation method show that the entropy of the source is maximum when all the four symbols occur with equal probability. Compute the value of the maximum entropy. ***[VI sem EC/TE, July/August 2003, Q.1 (c)]***

Solution

Let us choose an arbitrary probability p_m which depends on p_n where n takes values 1, 2, ... (m – 1). Let H represent the entropy function which depends on p_m and p_n. The derivative of entropy function must vanish in order that the entropy function H attains a maximum value.

i.e., $$\frac{dH}{dp_n} = 0$$

The LHS of the above expression can be expanded using the property of partial differentiation as

$$\frac{dH}{dp_n} = \sum_{j=1}^{m}\left[\frac{\partial H}{\partial p_j}\right]\left[\frac{\partial p_j}{\partial p_n}\right]$$

The entropy function H may be written as

$$H = H(p_n, p_m) = p_n \log_2 \frac{1}{p_n} + p_m \log_2 \frac{1}{p_m}$$

$$= p_n \frac{\log_e \frac{1}{p_n}}{\log_e 2} + p_m \frac{\log_e \frac{1}{p_m}}{\log_e 2}$$

$$= \log_2 e\left[p_n \ln \frac{1}{p_n} + p_m \ln \frac{1}{p_m}\right]$$

$$H = -\log_2 e\,[p_n \ln p_n + p_m \ln p_m]$$

where $p_m = 1 - p_n$ with n = 1, 2 and 3 for four symbols.

Differentiating the above expression with respect to p_n we get,

$$\frac{dH}{dp_n} = -\log_2 e\left[\frac{d}{dp_n}\{p_n \ln p_n\} + \frac{d}{dp_n}\{p_m \ln p_m\}\right]$$

$$= -\log_2 e\left[\frac{d}{dp_n}\{p_n \ln p_n\} + \frac{\partial}{\partial p_m}\{p_m \ln p_m\}\frac{\partial p_m}{\partial p_n}\right]$$

Substituting $p_m = 1 - p_n$, we get

$$\frac{dH}{dp_n} = -\log_2 e\left[\frac{d}{dp_n}\{p_n\, ln\, p_n\} + \frac{\partial}{\partial p_m}\{p_m\, ln\, p_m\}\frac{\partial(1-p_n)}{\partial p_n}\right]$$

$$= -\log_2 e\left[\left\{p_n \times \frac{1}{p_n} + ln\, p_n\right\} + \left\{p_m \times \frac{1}{p_m} + ln\, p_m\right\}(0-1)\right]$$

$$= -\log_2 e\,[1 + ln\, p_n - 1 - ln\, p_m]$$

$$= -\log_2 e\,[ln\, p_n - ln\, p_m]$$

$$= \log_2 e\,[ln\, p_m - ln\, p_n]$$

$$= \log_2 e\, ln\left(\frac{p_m}{p_n}\right)$$

$$= \log_2\left(\frac{p_m}{p_n}\right) \text{ from properties of logarithms}$$

For H to become maximum, we must have

$$\frac{dH}{dp_n} = 0$$

$$\therefore \quad \log_2\left(\frac{p_m}{p_n}\right) = 0$$

or $$\frac{p_m}{p_n} = 1$$

or $p_n = p_m$ for all n = 1, 2 and 3 and m = 4.

Since $p_m = p_4$ is only an arbitrary probability, we can conclude that,

For $n = 1$, $p_1 = p_4$ for $m = 4$

For $n = 2$, $p_2 = p_4$

For $n = 3$, $p_3 = p_4$

Combining, we get $p_1 = p_2 = p_3 = p_4$ in order that the entropy H becomes a maximum which is given by

$$H_{max} = \sum_{j=1}^{4} p_j \log\frac{1}{p_j}$$

$$= p_1 \log\frac{1}{p_1} + p_2 \log\frac{1}{p_2} + p_3 \log\frac{1}{p_3} + p_4 \log\frac{1}{p_4}$$

$$= 4\, p_1 \log\frac{1}{p_1}$$

But $\quad p_1 + p_2 + p_3 + p_4 = 1$

i.e., $\quad p_1 + p_1 + p_1 + p_1 = 1$

$$\therefore \quad 4p_1 = 1$$

$$\therefore \quad p_1 = \frac{1}{4}$$

$$\therefore \quad H_{max} = 4 \times \frac{1}{4} \log \frac{1}{(1/4)} = \log 4$$

$$\therefore \quad H_{max} = 2 \text{ bits/message symbol}$$

EXTENSION OF ZERO-MEMORY SOURCE : Extension of zero-memory sources becomes a necessity in some of the coding situations (will be discussed later). To understand the concept of source extension, let us consider a binary source S emitting symbols s_1 and s_2 with probabilities p_1 and p_2 respectively such that

$$p_1 + p_2 = 1 \quad \text{...... (1.31)}$$

Then the **"2ND EXTENSION"** of this binary source will have [(number of basic source symbols)$^{\text{Extension}}$] $2^2 = 4$ number of symbols given by

s_1s_1 occuring with probabilities $p_1p_1 = p_1^2$

s_1s_2 occuring with probabilities $p_1p_2 = p_1p_2$

s_2s_1 occuring with probabilities $p_2p_1 = p_1p^2$

and s_2s_2 occuring with probabilities $p_2p_2 = p_2^2.$

The sum of all probabilities of the 2nd extended source is

$$p_1^2 + 2p_1p_2 + p_2^2 = (p_1 + p_2)^2 = 1$$

The entropy of the 2nd extended source can be calculated as follows:

We have $S = \{S_1, S_2\}$ with $P = \{p_1, p_2\}$

$$\therefore \quad H(S) = \text{entropy of the basic binary source}$$

$$= \sum_{i=1}^{2} p_i \log \frac{1}{p_i}$$

$$\therefore \quad H(S) = p_1 \log \frac{1}{p_1} + p_2 \log \frac{1}{p_2} \quad \text{...... (1.32)}$$

The entropy of the 2nd extended source is given by

$$H(S^2) = \sum_{j=1}^{4} p_j \log \frac{1}{p_j}$$

$$= p_1^2 \log \frac{1}{p_1^2} + p_1p_2 \log \frac{1}{p_1p_2} + p_1p_2 \log \frac{1}{p_1p_2} + p_2^2 \log \frac{1}{p_2^2}$$

$$= 2\,p_1^2 \log\frac{1}{p_1} + 2\,p_1p_2 \log\frac{1}{p_1p_2} + 2\,p_2^2 \log\frac{1}{p_2}$$

$$= 2\,p_1^2 \log\frac{1}{p_1} + 2\,p_1p_2 \log\frac{1}{p_1} + 2\,p_1p_2 \log\frac{1}{p_2} + 2\,p_2^2 \log\frac{1}{p_2}$$

$$= 2\,p_1\,(p_1 + p_2) \log\frac{1}{p_1} + 2\,p_2\,(p_1 + p_2) \log\frac{1}{p_2}$$

$$= 2\left[p_1 \log\frac{1}{p_1} + p_2 \log\frac{1}{p_2}\right] \text{ since } p_1 + p_2 = 1$$

$$= 2\,H(S) \text{ using equation (1.32)}$$

$$\therefore\quad \mathbf{H(S^2) = 2\,H(S)} \qquad \text{...... (1.33)}$$

Similarly, the 3rd extension of the basic binary source will have $2^3 = 8$ number of symbols given by

$s_1s_1s_1$ occurring with a probability $p_1p_1p_1 = p_1^3$
$s_1s_1s_2$ occurring with a probability $p_1^2p_2$
$s_1s_2s_1$ occurring with a probability $p_1^2p_2$
$s_1s_2s_2$ occurring with a probability $p_1p_2^2$
$s_2s_1s_1$ occurring with a probability $p_1^2p_2$
$s_2s_1s_2$ occurring with a probability $p_1p_2^2$
$s_2s_2s_2$ occurring with a probability p_2^3

The sum of all probabilities of the 3rd extended binary source is

$p_1^3 + 3\,p_1^2p_2 + 3\,p_1p_2^2 + p_2^3 = (p_1 + p_2)^3 = 1$

The entropy of the 3rd extended source can similarly be shown to be

$$H(S^3) = 3\,H(S) \qquad \text{...... (1.34)}$$

Generalizing, the nth extension of the basic binary source will have 2^n symbols and the entropy of the nth extended source is given by

$$H(S^n) = n\,H(S) \qquad \text{...... (1.35)}$$

The general proof of equation (1.35) is given below.

Let us consider the source $S = \{s_1, s_2, \ldots, s_q\}$ with $P = \{p_1, p_2, \ldots, p_q\}$. The nth extension of this source, S^n, will have q^n number of symbols, namely $(\sigma_1, \sigma_2, \ldots\ldots, \sigma_{q^n})$. The probabilities of each of these σ_i [i varying from 1 to q^n] is given by

$$P(\sigma_i) = P(S_{i_1}).\,P(S_{i_2})\,.\,P(S_{i_3}) \ldots\ldots P(S_{i_n}),$$

$$\therefore \quad \sum_{S^n} P(\sigma_i) = \sum_{i=1}^{q^n} P(S_{i_1}).P(S_{i_2}).P(S_{i_3})......P(S_{i_n})$$

$$= \sum_{i_1=1}^{q} P(S_{i_1}).\sum_{i_2=1}^{q} P(S_{i_2}).\sum_{i_3=1}^{q} P(S_{i_3})......\sum_{i_n=1}^{q} P(S_{i_n})$$

$$= 1\,.\,1\,.\,1\,......\,1$$

$$\therefore \quad \sum_{S^n} P(\sigma_i) = 1 \qquad ......(1.36)$$

The entropy of the n^{th} extended source is

$$H(S^n) = \sum_{S^n} P(\sigma_i) \log \frac{1}{P(\sigma_i)}$$

$$= \sum_{S^n} P(\sigma_i) \log \frac{1}{P(S_{i_1}).P(S_{i_2}).....P(S_{i_n})}$$

$$= \sum_{S^n} P(\sigma_i) \log \frac{1}{P(S_{i_1})} + \sum_{S^n} P(\sigma_i) \log \frac{1}{P(S_{i_2})} +....+ \sum_{S^n} P(\sigma_i) \log \frac{1}{P(S_{i_n})} \qquad .....(1.37)$$

Consider

$$\sum_{S^n} P(\sigma_i) \log \frac{1}{P(S_{i_1})} = \sum_{S^n} P(S_{i_1}).P(S_{i_2}).P(S_{i_3})......P(S_{i_n}) \log \frac{1}{P(S_{i_1})}$$

$$= \sum_{i_1=1}^{q} P(S_{i_1}) \log \frac{1}{P(S_{i_1})} \cdot \sum_{i_2=1}^{q} P(S_{i_2}).\sum_{i_3=1}^{q} P(S_{i_3}).....\sum_{i_n=1}^{q} P(S_{i_n})$$

$$= H(S).1.1......1$$

$$\left.\begin{aligned}
&\therefore \sum_{S^n} P(\sigma_i) \log \frac{1}{P(S_{i_1})} = H(S) \\
&\text{Similarly } \sum_{S^n} P(\sigma_i) \log \frac{1}{P(S_{i_2})} = H(S) \\
&\qquad\quad \vdots \qquad\qquad\qquad \vdots \\
&\text{and} \quad \sum_{S^n} P(\sigma_i) \log \frac{1}{P(S_{i_n})} = H(S)
\end{aligned}\right\} \qquad ......(1.38)$$

Using equation (1.38) in (1.37), we get

$$H(S^n) = \underbrace{H(S) + H(S) +.......+ H(S)}_{\text{n terms}}$$

$$\therefore \quad \mathbf{H(S^n) = n\,H(S)} \text{ — proved}$$

Example 1.25 : A zero memory source has a source alphabet $S = \{s_1, s_2, s_3\}$ with $P = \left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right\}$. Find the entropy of this source. Also determine the entropy of its 2nd extension and verify that $H(S^2) = 2\,H(S)$.

Solution

For the basic source with 3 symbols,

$$H(S) = \sum_{i=1}^{3} p_i \log \frac{1}{p_i}$$

$$= \frac{1}{2}\log 2 + \frac{1}{4}\log 4 + \frac{1}{4}\log 4$$

$$= 1.5 \text{ bits/message-symbol}$$

The second extension of the basic source with 3 symbols will have $3^2 = 9$ symbols which can be listed along with their probability of occurrence as given below:

The symbol $S_1 S_1$ occurs with probability $= \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \left(\frac{1}{4}\right)$

S_1S_2 occurs with probability $= \left(\frac{1}{2}\right)\left(\frac{1}{4}\right) = \left(\frac{1}{8}\right)$

S_1S_3 occurs with probability $= \left(\frac{1}{2}\right)\left(\frac{1}{4}\right) = \left(\frac{1}{8}\right)$

S_2S_1 occurs with probability $= \left(\frac{1}{4}\right)\left(\frac{1}{2}\right) = \left(\frac{1}{8}\right)$

S_2S_2 occurs with probability $= \left(\frac{1}{4}\right)\left(\frac{1}{4}\right) = \left(\frac{1}{16}\right)$

S_2S_3 occurs with probability $= \left(\frac{1}{4}\right)\left(\frac{1}{4}\right) = \left(\frac{1}{16}\right)$

S_3S_1 occurs with probability $= \left(\frac{1}{4}\right)\left(\frac{1}{2}\right) = \left(\frac{1}{8}\right)$

S_3S_2 occurs with probability $= \left(\frac{1}{4}\right)\left(\frac{1}{4}\right) = \left(\frac{1}{16}\right)$

and S_3S_3 occurs with probability $= \left(\frac{1}{4}\right)\left(\frac{1}{4}\right) = \left(\frac{1}{16}\right)$

[*Note:* The sum of all probabilities of the 2nd extended source symbols must be equal to 1].

The entropy of the 2nd extended source is given by

$$H(S^2) = \sum_{j=1}^{9} p_j \log \frac{1}{p_j}$$

$$= \frac{1}{4}\log 4 + 4 \times \left(\frac{1}{8}\right)\log 8 + 4 \times \left(\frac{1}{16}\right)\log 16$$

$$= 3 \text{ bits/message symbol}$$

$$= 2 \times 1.5 \text{ bits/message symbol}$$

$$\therefore\ \mathbf{H(S^2) = 2\,H(S)} \quad \text{Proved}$$

Example 1.26 : An analog signal is band limited to 500 Hz and is sampled at "Nyquist Rate". The samples are quantized into 4 levels. The quantization levels are assumed to be independent and occur with probabilities $p_1 = p_4 = \frac{1}{8}, p_2 = p_3 = \frac{3}{8}$. Find the information rate of the source.

Solution

$$H(S) = \sum_{i=1}^{4} p_i \log \frac{1}{p_i} = \left(\frac{1}{8}\log 8\right)(2) + \left(\frac{3}{8}\log\frac{8}{3}\right)(2)$$

$$= 1.8113 \text{ bits/level or symbol.}$$

Since the signal is sampled at Nyquist rate, the symbol rate r_s is given by

$$r_s = 2B = 2 \times 500 = 1000 \text{ symbols/sec}$$

$$\therefore \text{ Information rate } R_s = r_s H(S)$$

$$= (1000)(1.81) = 1810 \text{ bits/sec}$$

Example 1.27 : Consider a discrete memoryless source with source alphabet $S = \{S_0, S_1, S_2\}$ with source statistics $\{0.7, 0.15, 0.15\}$.

(a) Calculate the entropy of the source

(b) Calculate the entropy of the second order extension of the source.

Solution : (a) From equation (1.4), the entropy of the source is given by

$$H(S) = \sum_{i=0}^{3} p_i \log \frac{1}{p_i}$$

$$\therefore \quad H(S) = 0.7 \log \frac{1}{0.7} + 0.15 \log \frac{1}{0.15} + 0.15 \log \frac{1}{0.15}$$

$$= 0.3602 + 0.41054 + 0.41054$$

$$\therefore \quad \mathbf{H(S) = 1.18128 \text{ bits/symbol.}}$$

(b) The entropy of the second-order extension of the source is given by equation (1.33) as

$$H(S^2) = 2\,H(S)$$

$$= (2)(1.18128)$$

$$\therefore \quad \mathbf{H(S^2) = 2.36256 \text{ bits/symbol.}}$$

Example 1.28 : A source emits one of the four probable messages M_1, M_2, M_3 and M_4 with probabilities of $\frac{7}{16}, \frac{5}{16}, \frac{1}{8}$ and $\frac{1}{8}$ respectively. Find the entropy of the source. List all the elements for the second extension of this source. Hence show $H(S^2) = 2\,H(S)$.

Solution

For the basic source $S = \{M_1, M_2, M_3, M_4\}$, the entropy is given by equation (1.4) as

$$H(S) = \sum_{i=1}^{4} P(M_i) \log \frac{1}{P(M_i)}$$

$$= \frac{7}{16}\log\frac{16}{7} + \frac{5}{16}\log\frac{16}{5} + \frac{1}{8}\log 8 + \frac{1}{8}\log 8$$

$= 1.7962$ bits/message symbol.

The second extension S^2 of the basic source will have $4^2 = 16$ symbols which can be listed along with their respective probabilities of occurrence as shown in table 1.4.

Symbol	Probability	Symbol	Probability
M_1M_1	49/256	M_3M_1	7/128
M_1M_2	35/256	M_3M_2	5/128
M_1M_3	7/128	M_3M_3	1/64
M_1M_4	7/128	M_3M_4	1/64
M_2M_1	35/256	M_4M_1	7/128
M_2M_2	25/256	M_4M_2	5/128
M_2M_3	5/128	M_4M_3	1/64
M_2M_4	5/128	M_4M_4	1/64

Table 1.4 : Source symbols and probabilities for 2nd extended source of example 1.28

The entropy of the second extended source is given by

$$H(S^2) = \sum_{j=1}^{16} p_j \log\frac{1}{p_j}$$

$$= \frac{49}{256}\log\frac{256}{49} + \left(\frac{35}{256}\log\frac{256}{35}\right)(2) + \frac{25}{256}\log\frac{256}{25}$$

$$+ \left(\frac{7}{128}\log\frac{128}{7}\right)(4) + \left(\frac{5}{128}\log\frac{128}{5}\right)(4) + \left(\frac{1}{64}\log 64\right)(4)$$

$= 3.5924$ bits/message symbol

$= 2 \times 1.7962$ bits/message symbol

$\therefore \quad \mathbf{H(S^2) = 2\,H(S)} \rightarrow$ Proved

1.4 AVERAGE INFORMATION CONTENT OF SYMBOLS IN LONG DEPENDENT SEQUENCES

For zero-memory sources which emit symbols in statistically independent sequences, the entropy or average information per message symbol and the source information rate can be found out using equation (1.4) and (1.5) respectively. In zero-memory sources, occurrence of one symbol does not alter the probability of occurrences of an other symbols. However, nearly all practical sources emit sequences of symbols that are statistically dependent. For example, in English language, the letter E occurs more frequently than the letter Z or Q; occurrence of letter Q implies that the following letter will most probably the letter U; the occurrence of a consonant implies that the following letter will most probably a vowel and so on. Thus the amount of information coming out from these sources will be less than that obtained from zero-memory sources. In the next section, let us develop a statistical model for sources emitting symbols in dependent sequences and calculate the corresponding entropy and information rate.

1.5 MARKOFF STATISTICAL MODEL FOR INFORMATION SOURCES

In the case of zero-memory sources, these is no intersymbol influence and in a long message, the symbols occur independently according to their $P(x_i)$. In real-life sources, there is intersymbol influence present such that the occurrence of x_i in the zeroth position s_0 of the message depends on the previous 'q' symbols $\{s_1, s_2, \ldots, s_q\}$. Such a source is known as q^{th} order ***"Markoff source"*** or ***"Markov source"*** and these sources are generally specified by a set of conditional probabilities $P(x_i/s_1, s_2, \ldots, s_q)$ where x_i is the symbol in the s_o position and each S has an m-symbol alphabet $\{x_1, x_2, \ldots, x_m\}$. Since $P(x_i)$ now depends on the earlier 'q' symbols, the transitional probabilities may be shown in a state diagram with m^q possible states previous to s_o and the behaviour of the Markov source may be predicted from the state diagram.

Example 1.29 : For the first order Markov source shown in figure 1.6, draw the tree diagram representing the states at the end of second symbol interval and find the corresponding probabilities. Assume P(1) = P(2) = P(3) = 1/3.

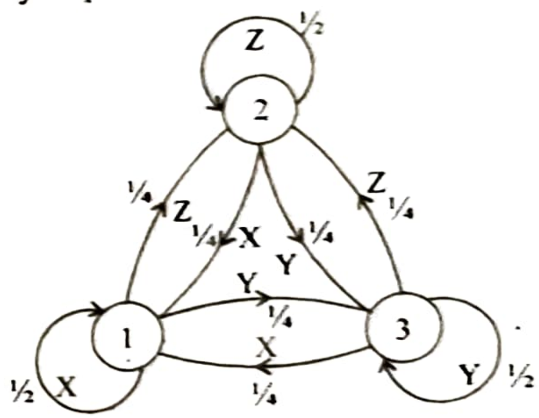

Fig. 1.6 : Markov source of example 1.29

Solution

The tree diagram representing the initial state, first state at the end of first symbol interval and the second state at the end of second symbol interval is shown in figure 1.7.

From the tree diagram, the symbol XX can be generated by either one of the following transitions 1→1→1 or 2→1→1 or 3→1→1. Thus the probability of the source emitting the 2 symbol sequence XX is given by,

$$P(XX) = P[(1\to1\to1) \text{ or } (2\to1\to1) \text{ or } (3\to1\to1)] \qquad \ldots\ldots (1.39)$$

Since the three transition paths are disjoint, we have

$$P(XX) = P[(1\to1\to1)] + P[(2\to1\to1)] + P[(3\to1\to1)]$$

$$= \left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{2}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right) = \tfrac{1}{6}$$

$$\text{Similarly } P(XZ) = \left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{4}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) = \tfrac{1}{12}$$

$$P(XY) = \left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{4}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) = \tfrac{1}{12}$$

$$P(ZX) = \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{4}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) = \tfrac{1}{12}$$

$$P(ZZ) = \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{2}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right) = \tfrac{1}{6}$$

$$P(ZY) = \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{4}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) = \tfrac{1}{12}$$

$$= \frac{7}{16}\log\frac{16}{7} + \frac{5}{16}\log\frac{16}{5} + \frac{1}{8}\log 8 + \frac{1}{8}\log 8$$

$= 1.7962$ bits/message symbol.

The second extension S^2 of the basic source will have $4^2 = 16$ symbols which can be listed along with their respective probabilities of occurrence as shown in table 1.4.

Symbol	Probability	Symbol	Probability
M_1M_1	49/256	M_3M_1	7/128
M_1M_2	35/256	M_3M_2	5/128
M_1M_3	7/128	M_3M_3	1/64
M_1M_4	7/128	M_3M_4	1/64
M_2M_1	35/256	M_4M_1	7/128
M_2M_2	25/256	M_4M_2	5/128
M_2M_3	5/128	M_4M_3	1/64
M_2M_4	5/128	M_4M_4	1/64

Table 1.4 : Source symbols and probabilities for 2nd extended source of example 1.28

The entropy of the second extended source is given by

$$H(S^2) = \sum_{j=1}^{16} p_j \log\frac{1}{p_j}$$

$$= \frac{49}{256}\log\frac{256}{49} + \left(\frac{35}{256}\log\frac{256}{35}\right)(2) + \frac{25}{256}\log\frac{256}{25}$$

$$+\left(\frac{7}{128}\log\frac{128}{7}\right)(4) + \left(\frac{5}{128}\log\frac{128}{5}\right)(4) + \left(\frac{1}{64}\log 64\right)(4)$$

$= 3.5924$ bits/message symbol

$= 2 \times 1.7962$ bits/message symbol

$\therefore$ **$H(S^2) = 2H(S)$** $\rightarrow$ Proved

1.4 AVERAGE INFORMATION CONTENT OF SYMBOLS IN LONG DEPENDENT SEQUENCES

For zero-memory sources which emit symbols in statistically independent sequences, the entropy or average information per message symbol and the source information rate can be found out using equation (1.4) and (1.5) respectively. In zero-memory sources, occurrence of one symbol does not alter the probability of occurrences of an other symbols. However, nearly all practical sources emit sequences of symbols that are statistically dependent. For example, in English language, the letter E occurs more frequently than the letter Z or Q; occurrence of letter Q implies that the following letter will most probably the letter U; the occurrence of a consonant implies that the following letter will most probably a vowel and so on. Thus the amount of information coming out from these sources will be less than that obtained from zero-memory sources. In the next section, let us develop a statistical model for sources emitting symbols in dependent sequences and calculate the corresponding entropy and information rate.

1.5 MARKOFF STATISTICAL MODEL FOR INFORMATION SOURCES

In the case of zero-memory sources, these is no intersymbol influence and in a long message, the symbols occur independently according to their $P(x_i)$. In real-life sources, there is intersymbol influence present such that the occurrence of x_i in the zeroth position s_0 of the message depends on the previous 'q' symbols $\{s_1, s_2, \ldots, s_q\}$. Such a source is known as q^{th} order ***"Markoff source"*** or ***"Markov source"*** and these sources are generally specified by a set of conditional probabilities $P(x_i/s_1, s_2, \ldots, s_q)$ where x_i is the symbol in the s_o position and each S has an m-symbol alphabet $\{x_1, x_2, \ldots, x_m\}$. Since $P(x_i)$ now depends on the earlier 'q' symbols, the transitional probabilities may be shown in a state diagram with m^q possible states previous to s_o and the behaviour of the Markov source may be predicted from the state diagram.

Example 1.29 : For the first order Markov source shown in figure 1.6, draw the tree diagram representing the states at the end of second symbol interval and find the corresponding probabilities. Assume $P(1) = P(2) = P(3) = 1/3$.

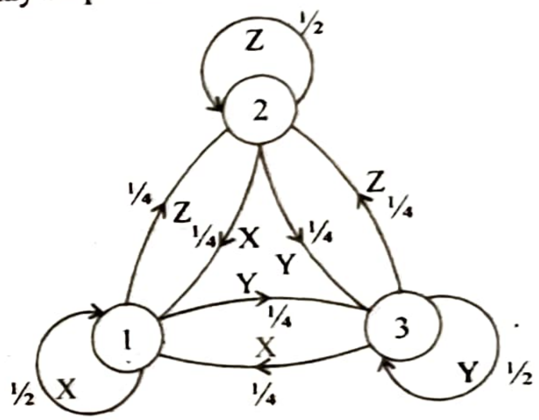

Fig. 1.6 : Markov source of example 1.29

Solution

The tree diagram representing the initial state, first state at the end of first symbol interval and the second state at the end of second symbol interval is shown in figure 1.7.

From the tree diagram, the symbol XX can be generated by either one of the following transitions $1\to1\to1$ or $2\to1\to1$ or $3\to1\to1$. Thus the probability of the source emitting the 2 symbol sequence XX is given by,

$$P(XX) = P[(1\to1\to1) \text{ or } (2\to1\to1) \text{ or } (3\to1\to1)] \qquad \ldots\ldots (1.39)$$

Since the three transition paths are disjoint, we have

$$P(XX) = P[(1\to1\to1)] + P[(2\to1\to1)] + P[(3\to1\to1)]$$

$$= \left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{2}\right)+\left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right)+\left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right) = \tfrac{1}{6}$$

$$\text{Similarly } P(XZ) = \left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{4}\right)+\left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right)+\left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) = \tfrac{1}{12}$$

$$P(XY) = \left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{4}\right)+\left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right)+\left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) = \tfrac{1}{12}$$

$$P(ZX) = \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right)+\left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{4}\right)+\left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) = \tfrac{1}{12}$$

$$P(ZZ) = \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right)+\left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{2}\right)+\left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right) = \tfrac{1}{6}$$

$$P(ZY) = \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right)+\left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{4}\right)+\left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) = \tfrac{1}{12}$$

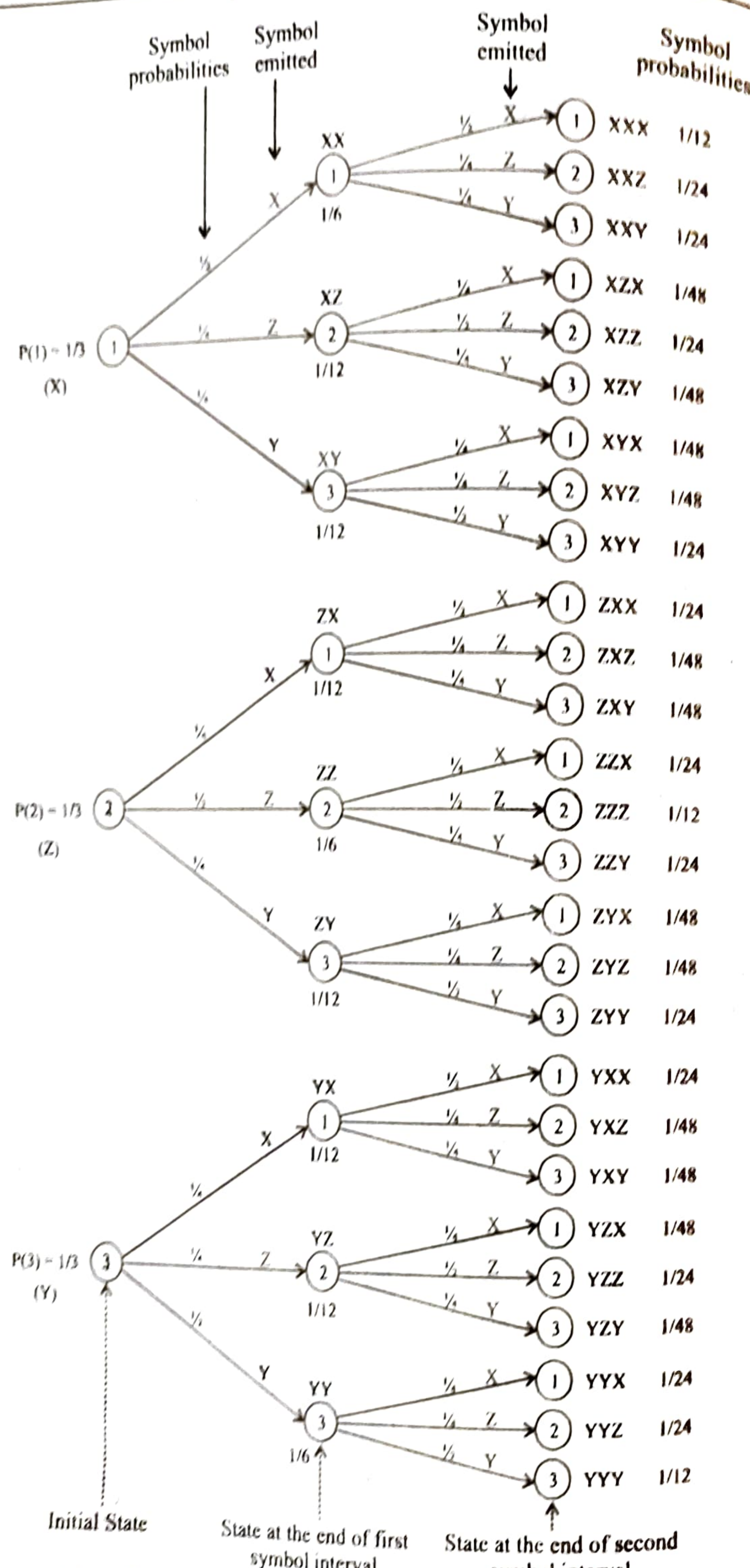

Fig. 1.7 : Tree diagram for the Markov source of figure 1.6

$$P(YX) = \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{4}\right) = \tfrac{1}{12}$$

$$P(YZ) = \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{4}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{4}\right) = \tfrac{1}{12}$$

$$P(YY) = \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right) + \left(\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{2}\right) = \tfrac{1}{6}$$

1.6 ENTROPY AND INFORMATION RATE OF MARKOFF SOURCES

The entropy of the source is defined as the weighted average of the entropy of the symbols emitted from each state, where the entropy of state 'i', denoted by H_i, is defined as the average information content of the symbols emitted from the i^{th} state.

$$\therefore \quad H_i = \sum_{j=1}^{n} p_{ij} \log \frac{1}{p_{ij}} \text{ bits/message-symbol} \qquad \text{...... (1.40)}$$

The entropy of the source is then the average of the entropy of each state. i.e.,

$$\left.\begin{aligned} H &= \sum_{i=1}^{n} p_i H_i \\ &= \sum_{i=1}^{n} p_i \left[\sum_{j=1}^{n} p_{ij} \log \frac{1}{p_{ij}}\right] \text{ bits/message symbol} \end{aligned}\right\} \qquad \text{...... (1.41)}$$

Where p_i is the probability that the source is in state i.

The average information rate R_s for the source is defined as,

$$R_s = r_s H \text{ bits/sec} \qquad \text{...... (1.42)}$$

Where r_s is the number of state transitions per second or the symbol rate of the source.

Theorem : If $P(m_i)$ is the probability of a sequence m_i of N symbols from the source and if

$$G_N = \frac{1}{N}\sum_i P(m_i)\log\frac{1}{P(m_i)} = \frac{1}{N} H(\bar{S}^N) \qquad \text{...... (1.43)}$$

where H(S) is entropy of adjoint source.

[The "adjoint source $\bar{S}$" is defined as a zero memory source that has the same source alphabet and same first order probabilities as the initial state. Then the 2nd extension of adjoint source $\bar{S}^2$ will have the source alphabet and second order probabilities at the start of the 2nd symbol interval (or at the end of first symbol interval) and so on].

where the sum is over all sequences m_i containing N symbols, then G_N is a monotonically decreasing function of N and

$$\lim_{N\to\infty} G_N = H \text{ bits/sec} \qquad \text{...... (1.44)}$$

The above definitions of entropy and information rate and the theorem can be used to solve many problems as shown below:

Example 1.30 : For the Markov source of figure 1.6 (refer previous example), find (i) the entropy of each state, (ii) the entropy of the source, (iii) G_1, G_2, G_3 and then show that $G_1 > G_2 > G_3 > H$.

Solution

The entropy of each state is given by equation (1.40) as,

$$H_i = \sum_{j=1}^{n} p_{ij} \log \frac{1}{p_{ij}}$$

(i)

$$H_1 = p_{11} \log \frac{1}{p_{11}} + p_{12} \log \frac{1}{p_{12}} + p_{13} \log \frac{1}{p_{13}}$$

$$= \frac{1}{2}\log 2 + \frac{1}{4}\log 4 + \frac{1}{4}\log 4$$

$$\therefore \mathbf{H_1 = 1.5 \text{ bits\textbackslash message-symbol}}$$

Similarly, $H_2 = p_{21} \log \frac{1}{p_{21}} + p_{22} \log \frac{1}{p_{22}} + p_{23} \log \frac{1}{p_{23}}$

$$= \frac{1}{4}\log 4 + \frac{1}{2}\log 2 + \frac{1}{4}\log 4$$

$$\therefore \mathbf{H_2 = 1.5 \text{ bits/message-symbol}}$$

and $H_3 = p_{31} \log \frac{1}{p_{31}} + p_{32} \log \frac{1}{p_{32}} + p_{33} \log \frac{1}{p_{33}}$

$$= \frac{1}{4}\log 4 + \frac{1}{4}\log 4 + \frac{1}{2}\log 2$$

$$\therefore \mathbf{H_3 = 1.5 \text{ bits/message-symbol}}$$

(ii) The entropy of the source is given by equation (1.41) as

$$H_i = \sum_{i=1}^{n} P_i H_i$$

$$= P_1 H_1 + P_2 H_2 + P_3 H_3$$

We have, $P_1 = P(1) = \frac{1}{3} = P_2 = P(2) = P_3 = P(3)$

$$\therefore H = \left(\frac{1}{3}\right)(1.5) + \left(\frac{1}{3}\right)(1.5) + \left(\frac{1}{3}\right)(1.5)$$

$$\therefore \mathbf{H = 1.5 \text{ bits/message-symbol}}$$

(iii) From the equation (1.43)

$$G_N = \frac{1}{N} \sum_i P(m_i) \log \frac{1}{P(m_i)}$$

For N = 1, the source is a zero-memory source $\bar{S}$ emitting symbols with equal probability.

$$\therefore G_1 = H(\bar{S}) = \frac{1}{3}\log 3 + \frac{1}{3}\log 3 + \frac{1}{3}\log 3$$

$$\therefore \mathbf{G_1 = H(\bar{S}) = 1.585 \text{ bits/message-symbol}}$$

$\therefore$ We can conclude that $G_1 > H$.

To find G_2, we require the second order symbol probabilities which have been calculated in the previous example as

$$P(XX) = P(YY) = P(ZZ) = \frac{1}{6}$$

$$P(XY) = P(XZ) = P(YX) = P(YZ) = P(ZX) = P(XZ) = \frac{1}{12}$$

$\therefore$ The entropy of the states at the end of second symbol interval (i.e., the entropy of the second order symbols) is

$$H(\bar{S}^2) = 3 \times \frac{1}{6} \log 6 + 6 \times \frac{1}{12} \log 12$$

$$\mathbf{H(\bar{S}^2) = 3.085 \text{ bits/message-symbol}}$$

$\therefore$ From the equation (1.43) for N = 2,

$$G_2 = \frac{1}{2} \sum_{i=1}^{n^2} p(2_i) \log \frac{1}{p(2_i)} = \frac{1}{2} H(\bar{S}^2) = \frac{1}{2}(3.085)$$

$$\therefore \mathbf{G_2 = 1.5425 \text{ bits/message-symbol}}$$

The third order symbol probabilities are listed in the tree diagram of figure 1.7 from which the entropy of the states at the end of third symbol interval is given by

$$H(\bar{S}^3) = \left(\frac{1}{12} \log 12\right)(3) + \left(\frac{1}{24} \log 24\right)(12) + \left(\frac{1}{48} \log 48\right)(12)$$

$$= \mathbf{4.585 \text{ bits/message-symbol}}$$

$$\therefore G_3 = \frac{H\left(\bar{S}^3\right)}{3} = \frac{4.585}{3}$$

$$\therefore \mathbf{G_3 = 1.528 \text{ bits/message symbol}}$$

Thus we observe that $\mathbf{G_1 > G_2 > G_3 > H}$.

Example 1.31 : For the first order Markov source with a source alphabet S = {A, B, C} shown in figure 1.8, (i) compute the probabilities of states (ii) find H(S) and H(S^2).

- 10M

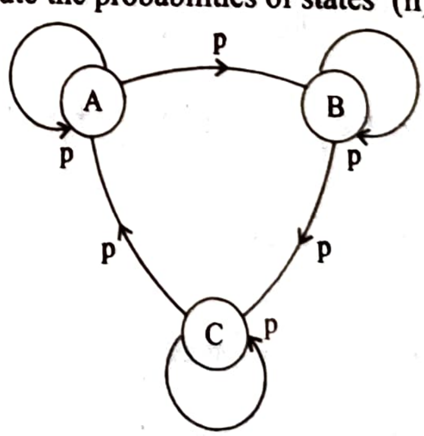

Fig. 1.8 : Markov source of example 1.31

Solution

(i) From the given state diagram, the state equations can be written as

$$P(A) = p\,P(A) + p\,P(C) \quad \text{...... (1.45)}$$

$$P(B) = p\,P(B) + p\,P(A) \quad \text{...... (1.46)}$$

$$P(C) = p\,P(C) + p\,P(B) \quad \text{...... (1.47)}$$

Adding all the three equations, we get

$$P(A) + P(B) + P(C) = 2p\,[P(A) + P(B) + P(C)]$$

$$\therefore\ 2p = 1 \Rightarrow \mathbf{p = \frac{1}{2}}$$

Using this value of p in equation (1.45), we get

$$P(A) = \frac{1}{2}P(A) + \frac{1}{2}P(C)$$

$$\therefore\ P(A) = P(C) \quad \text{...... (1.48)}$$

Again using $p = \frac{1}{2}$ in equation (1.46), we get

$$P(B) = \frac{1}{2}P(B) + \frac{1}{2}P(A)$$

$$\therefore\ P(B) = P(A) \quad \text{...... (1.49)}$$

Combining equations (1.48) and (1.49), we get

$$P(A) = P(B) = P(C)$$

But $P(A)+P(B)+P(C) = 1$

$$\therefore\ \mathbf{P(A) = P(B) = P(C) = \frac{1}{3}}$$

(ii) From equation (1.40)

$$H_i = \sum_{j=1}^{n} p_{ij}\log\frac{1}{p_{ij}}$$

$$\therefore \text{ For 1st state, } H_A = p_{AA}\log\frac{1}{p_{AA}} + p_{AB}\log\frac{1}{p_{AB}} + p_{AC}\log\frac{1}{p_{AC}}$$

$$= p\log\frac{1}{p} + p\log\frac{1}{p} + 0$$

$$= 2p\log\frac{1}{p} = 2\times\frac{1}{2}\log 2$$

$$\therefore\ \mathbf{H_A = 1\ bit/message\text{-}symbol}$$

$$\text{Similarly } H_B = p_{BA} \log \frac{1}{p_{BA}} + p_{BB} \log \frac{1}{p_{BB}} + p_{BC} \log \frac{1}{p_{BC}}$$

$$= 0 + p \log \frac{1}{p} + p \log \frac{1}{p}$$

$$= 2p \log \frac{1}{p}$$

$$\therefore \mathbf{H_B = 1 \text{ bit/message-symbol}}$$

$$\text{and } H_C = p_{CA} \log \frac{1}{p_{CA}} + p_{CB} \log \frac{1}{p_{CB}} + p_{CC} \log \frac{1}{p_{CC}}$$

$$= p \log \frac{1}{p} + 0 + p \log \frac{1}{p}$$

$$= 2p \log \frac{1}{p}$$

$$\therefore \mathbf{H_C = 1 \text{ bit/message-symbol}}$$

From equation (1.41)

$$H(S) = H = \sum_{i=1}^{n} P_i H_i$$

$$= P(A)H_A + P(B)H_B + P(C)H_C$$

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

$$\therefore \mathbf{H(S) = 1 \text{ bit/message-symbol}}$$

$$\text{and } H(S^2) = 2H(S)$$

$$\therefore \mathbf{H(S) = 2 \text{ bits/message-symbol}}$$

Example 1.32 : Consider the state diagram of the Markov source of figure 1.9.

(i) Compute the state probabilities

(ii) Find entropy of each state

(iii) Find the entropy of the source.

Solution

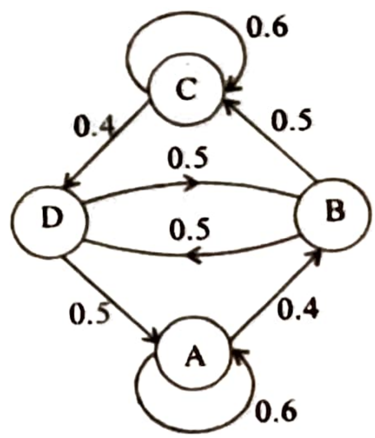

Fig. 1.9 : Markov source of example 1.32

(i) From the state diagram, the state equations can be immediately written as,

$$P(A) = 0.6\, P(A) + 0.5\, P(D) \qquad \text{...... (1.50)}$$

$$P(B) = 0.4\, P(A) + 0.5\, P(D)$$

$$P(C) = 0.6\, P(C) + 0.5\, P(B) \quad \text{......(1.51)}$$

$$P(D) = 0.4\, P(C) + 0.5\, P(B) \quad \text{......(1.52)}$$

$$\text{......(1.53)}$$

From equation (1.50),

$$0.4\, P(A) = 0.5\, P(D)$$

$$\therefore\ P(A) = 1.25\, P(D) \quad \text{......(1.54)}$$

Using equation (1.54) in (1.51), we get

$$P(B) = (0.4)(1.25)\, P(D) + (0.5)\, P(D)$$

$$\therefore\ \mathbf{P(B) = P(D)} \quad \text{......(1.55)}$$

From equation (1.52)

$$0.4\, P(C) = 0.5\, P(B)$$

$$= 0.5\, P(D) \text{ from equation } (1.55)$$

$$\therefore\ P(C) = 1.25\, P(D) \quad \text{......(1.56)}$$

But $P(A) + P(B) + P(C) + P(D) = 1$ ......(1.57)

Substituting equations (1.54) to (1.56) in (1.57), we get

$1.25\, P(D) + P(D) + 1.25\, P(D) + P(D) = 1$

$$4.5\, P(D) = 1$$

$$\therefore\ \mathbf{P(D) = \frac{2}{9}}$$

$$\therefore\ P(A) = 1.25\, P(D) = (1.25)\left(\frac{2}{9}\right) \qquad \therefore\ \mathbf{P(A) = \frac{5}{18}}$$

$$\therefore\ P(B) = P(D) \qquad \therefore\ \mathbf{P(B) = \frac{2}{9}}$$

$$\therefore\ P(C) = 1.25\, P(D) \qquad \therefore\ \mathbf{P(C) = \frac{5}{18}}$$

Therefore the probabilities are $P(A) = P(C) = \frac{5}{18}$, $P(B) = P(D) = \frac{2}{9}$.

(ii) From equation (1.40), the entropy of each state is

$$H_i = \sum_{i=1}^{n} p_{ij} \log \frac{1}{p_{ij}}$$

$$\therefore\ H_A = p_{AA} \log \frac{1}{p_{AA}} + p_{AB} \log \frac{1}{p_{AB}} + p_{AC} \log \frac{1}{p_{AC}} + p_{AD} \log \frac{1}{p_{AD}}$$

$$= 0.6 \log \frac{1}{0.6} + 0.4 \log \frac{1}{0.4} + 0 + 0$$

$$\therefore\ \mathbf{H_A = 0.971 \text{ bits/state or message-symbol}}$$

$$\text{Similarly } H_B = p_{BA} \log \frac{1}{p_{BA}} + p_{BB} \log \frac{1}{p_{BB}} + p_{BC} \log \frac{1}{p_{BC}} + p_{BD} \log \frac{1}{p_{BD}}$$

$$= 0 + 0 + 0.5 \log \frac{1}{0.5} + 0.5 \log \frac{1}{0.5}$$

$$\therefore \; H_B = 1 \text{ bit/state or message-symbol}$$

$$\text{and } H_C = p_{CA} \log \frac{1}{p_{CA}} + p_{CB} \log \frac{1}{p_{CB}} + p_{CC} \log \frac{1}{p_{CC}} + p_{CD} \log \frac{1}{p_{CD}}$$

$$= 0 + 0 + 0.6 \log \frac{1}{0.6} + 0.4 \log \frac{1}{0.4}$$

$$\therefore \; H_C = 0.971 \text{ bits/state or message-symbol}$$

$$\text{and } H_D = p_{DA} \log \frac{1}{p_{DA}} + p_{DB} \log \frac{1}{p_{DB}} + p_{DC} \log \frac{1}{p_{DC}} + p_{DD} \log \frac{1}{p_{DD}}$$

$$= 0.5 \log \frac{1}{0.5} + 0.5 \log \frac{1}{0.5} + 0 + 0$$

$$\therefore \; H_D = 1 \text{ bits/state or message-symbol}$$

(iii) The entropy of the source $S = \{A, B, C, D\}$ is calculated using equation (1.41) as

$$H(S) = H = \sum_{i=1}^{n} P_i H_i$$

$$= P(A)H_A + P(B)H_B + P(C)H_C + P(B)H_D$$

$$= \left(\frac{5}{18}\right)(0.971) + \left(\frac{2}{9}\right)(1) + \left(\frac{5}{18}\right)(0.971) + \left(\frac{2}{9}\right)(1)$$

$$\therefore \; H = 0.9839 \text{ bits/binary digits}$$

Example 1.33 : You are asked to design an information system, which gives the information, every year, for about 200 students passing out with B.E. Electronics and Communication degree from a certain university. The students can get into one of three fields as given below:

(i) go abroad for higher studies → A

(ii) join MBA or Civil Services → B

(iii) join Industries in India → C

Based on the data given below, construct the model for the source and find source entropy.

(a) On the average 100 students are going abroad.

(b) Out of 100 going abroad this year, 50 were reported going abroad next year, while 25 each went to MBA and Civil Services or joined Industries in India.

(c) Out of 100 remaining in India this year, 50 continued to do so, while 50 went abroad next year.

(d) Those joining MBA and Civil Services or Industry could not swap the two fields next year.

Solution

Based on the data (a) to (d) given above, the state diagram can be constructed as shown in figure 1.10 below:

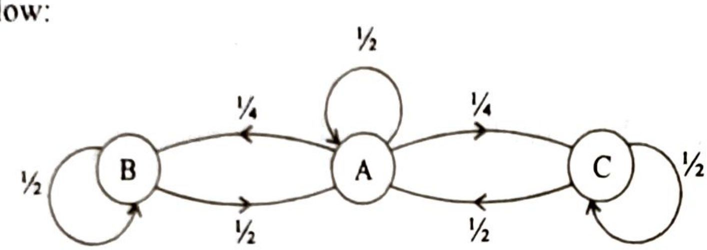

Fig. 1.10 : Markoff source for example 1.33

The state probabilities can be calculated from the state diagram as below:

For state A $\rightarrow P(A) = \frac{1}{2}P(A) + \frac{1}{2}P(B) + \frac{1}{2}P(C)$ ...... (1.58)

For state B $\rightarrow P(B) = \frac{1}{2}P(B) + \frac{1}{4}P(A)$ ...... (1.59)

For state C $\rightarrow P(C) = \frac{1}{2}P(C) + \frac{1}{4}P(A)$ ...... (1.60)

From equation (1.58),

$$P(A) = \frac{1}{2}[P(A) + P(B) + P(C)] = \frac{1}{2}(1) \quad \therefore \quad \mathbf{P(A) = \frac{1}{2}}$$

From equation (1.59),

$$P(B) = \frac{1}{2}P(B) + \frac{1}{4}\left(\frac{1}{2}\right) \quad \therefore \quad \mathbf{P(B) = \frac{1}{4}}$$

From equation (1.60),

$$P(C) = \frac{1}{2}P(C) + \frac{1}{4}\left(\frac{1}{2}\right) \quad \therefore \quad \mathbf{P(C) = \frac{1}{4}}$$

From equation (1.40); the entropy of each state is

$$H_A = p_{AA}\log\frac{1}{p_{AA}} + p_{AB}\log\frac{1}{p_{AB}} + p_{AC}\log\frac{1}{p_{AC}}$$

$$= \frac{1}{2}\log 2 + \frac{1}{4}\log 4 + \frac{1}{4}\log 4$$

$$= 1.5 \text{ bits/message-symbol or state}$$

$$\text{Similarly } H_B = p_{BA} \log \frac{1}{p_{BA}} + p_{BB} \log \frac{1}{p_{BB}} + p_{BC} \log \frac{1}{p_{BC}}$$

$$= \frac{1}{2} \log 2 + \frac{1}{2} \log 2 + 0$$

$$= 1 \text{ bit/message-symbol}$$

$$\text{and } H_C = p_{CA} \log \frac{1}{p_{CA}} + p_{CB} \log \frac{1}{p_{CB}} + p_{CC} \log \frac{1}{p_{CC}}$$

$$= \frac{1}{2} \log 2 + 0 + \frac{1}{2} \log 2$$

$$= 1 \text{ bit/message-symbol}$$

From equation (1.41), the entropy of the source is calculated as,

$$H = \sum_{i=1}^{n} P_i H_i$$

$$= P(A)H_A + P(B)H_B + P(C)H_C$$

$$= \left(\frac{1}{2}\right)(1.5) + \left(\frac{1}{4}\right)(1) + \left(\frac{1}{4}\right)(1)$$

$$\therefore \; H = 1.25 \text{ bits/message-symbol}$$

Example 1.34 : The state diagram of a Markoff source is shown in figure 1.11.

10 (i) Find the entropy H of the source.

(ii) Find G_1, G_2 and G_3 and verify that $G_1 > G_2 > G_3 > H$.

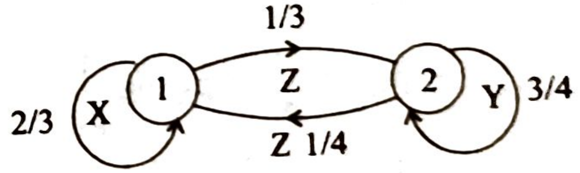

Fig. 1.11 : Markoff source of example 1.34

Solution

(i) From the given state diagram of figure 1.11, the state equations can be written as

$$P(1) = \frac{2}{3} P(1) + \frac{1}{4} P(2) \qquad \text{.... (1.61)}$$

$$P(2) = \frac{3}{4} P(2) + \frac{1}{3} P(1) \qquad \text{.... (1.62)}$$

From equation (1.61), $\frac{1}{3} P(1) = \frac{1}{4} P(2)$

$$\therefore \qquad P(1) = \frac{3}{4} P(2) \qquad \text{.... (1.63)}$$

We have $$P(1) + P(2) = 1$$

$$\therefore \qquad \frac{3}{4} P(2) + P(2) = 1$$

$$\therefore \qquad P(2) = \frac{4}{7}$$

From equation (1.63), $$P(1) = \frac{3}{7}$$

From equation (1.40), the entropy of each state is given by

$$H_i = \sum_{j=1}^{n} p_{ij} \log \frac{1}{p_{ij}}$$

For 1^{st} state, $$H_1 = \sum_{j=1}^{2} p_{1j} \log \frac{1}{p_{1j}}$$

$$= p_{11} \log \frac{1}{p_{11}} + p_{12} \log \frac{1}{p_{12}}$$

$$= \frac{2}{3} \log \frac{3}{2} + \frac{1}{3} \log 3 = 0.9183 \text{ bits/message symbol.}$$

For 2^{nd} state, $$H_2 = \sum_{j=1}^{2} p_{2j} \log \frac{1}{p_{2j}}$$

$$= p_{21} \log \frac{1}{p_{21}} + p_{22} \log \frac{1}{p_{22}}$$

$$= \frac{1}{4} \log 4 + \frac{3}{4} \log \frac{4}{3} = 0.8113 \text{ bits/message symbol}$$

$\therefore$ The entropy of the source is given by equation (1.41) as

$$H = \sum_{i=1}^{n} p_i H_i = \sum_{i=1}^{n} P(i) H_i = P(1) H_1 + P(2) H_2$$

$$= \left(\frac{3}{7}\right)(0.9183) + \left(\frac{4}{7}\right)(0.8113)$$

$$\therefore \qquad \mathbf{H = 0.8572 \text{ bits/message symbol}}$$

(ii) The tree diagram representing the initial state, first state at the end of first symbol interval, the second state at the end of second symbol interval and the third state at the end of third symbol interval is shown in figure 1.12.

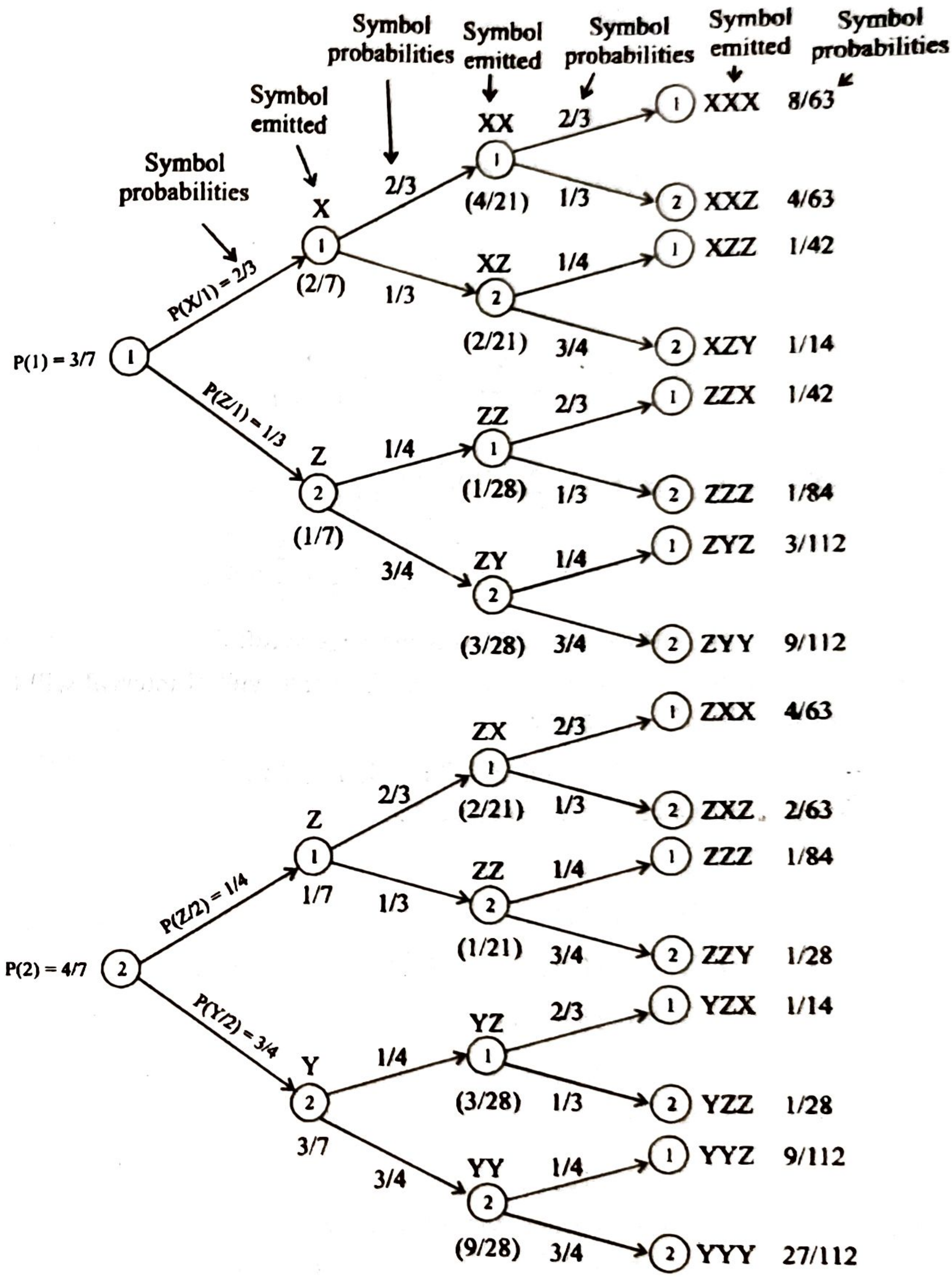

Fig. 1.12 : Tree diagram showing first three states for the Markoff source of fig. 1.11.

The symbols produced by the source will have an alphabet given by $\bar{S} = \{X, Y, Z\}$ and the corresponding probabilities at the end of first symbol interval is calculated as

$$P(X) = \frac{3}{7} \times \frac{2}{3} = \frac{2}{7}$$

$$P(Y) = \frac{4}{7} \times \frac{3}{4} = \frac{3}{7}$$

$$P(Z) = \frac{3}{7} \times \frac{1}{3} + \frac{4}{7} \times \frac{1}{4} = \frac{2}{7}$$

$$\therefore \quad P(\bar{S}) = \left\{\frac{2}{7}, \frac{3}{7}, \frac{2}{7}\right\}$$

$\therefore$ From equation (1.43), $\quad G_N = \frac{1}{N} \sum_{i=1}^{3} P(m_i) \log \frac{1}{P(m_i)}$

For N = 1, wehave, $\quad G_N = H(\bar{S}) = \sum_{i=1}^{3} P(m_i) \log \frac{1}{P(m_i)}$

$$\therefore \quad G_1 = H(\bar{S}) = \frac{2}{7} \log \frac{7}{2} + \frac{3}{7} \log \frac{7}{3} + \frac{2}{7} \log \frac{7}{2}$$

$$\therefore \quad \mathbf{G_1 = H(\bar{S}) = 1.557 \text{ bits/message-symbol}}$$

The symbols produced by the source at the end of second symbol interval will have an alphabet given by (note that ZZ is repeated).

$$\bar{S}^2 = \{XX, XZ, ZZ, ZY, ZX, YZ, YY\}$$

The corresponding probabilities are calculated as

$$P(XX) = \left(\tfrac{3}{7}\right)\left(\tfrac{2}{3}\right)\left(\tfrac{2}{3}\right) = \tfrac{4}{21}$$

$$P(XZ) = \left(\tfrac{3}{7}\right)\left(\tfrac{2}{3}\right)\left(\tfrac{1}{3}\right) = \tfrac{2}{21}$$

$$P(ZZ) = \left(\tfrac{3}{7}\right)\left(\tfrac{1}{3}\right)\left(\tfrac{1}{4}\right) + \left(\tfrac{4}{7}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{1}{3}\right) = \tfrac{1}{28} + \tfrac{1}{21} = \tfrac{1}{12}$$

$$P(ZY) = \left(\tfrac{3}{7}\right)\left(\tfrac{1}{3}\right)\left(\tfrac{3}{4}\right) = \tfrac{3}{28}$$

$$P(ZX) = \left(\tfrac{4}{7}\right)\left(\tfrac{1}{4}\right)\left(\tfrac{2}{3}\right) = \tfrac{2}{21}$$

$$P(YZ) = \left(\tfrac{4}{7}\right)\left(\tfrac{3}{4}\right)\left(\tfrac{1}{4}\right) = \tfrac{3}{28}$$

$$P(YY) = \left(\tfrac{4}{7}\right)\left(\tfrac{3}{4}\right)\left(\tfrac{3}{4}\right) = \tfrac{9}{28}$$

$\therefore$ The probabilities of second order symbols are given by

$$P(\bar{S}^2) = \left\{\frac{4}{21}, \frac{2}{21}, \frac{1}{12}, \frac{3}{28}, \frac{2}{21}, \frac{3}{28}, \frac{9}{28}\right\}$$

The entropy of the source at the end of second symbol interval is given by

$$H(S^2) = \frac{4}{21}\log\frac{21}{4} + \left(\frac{2}{21}\log\frac{21}{2}\right)(2) + \frac{1}{12}\log 12$$

$$+\left(\frac{3}{28}\log\frac{28}{3}\right)(2) + \frac{9}{28}\log\frac{28}{9}$$

$\therefore$ $H(S^2) = 2.6174$ bits/message symbol.

$\therefore$ From equation (1.43) for N = 2, we have

$$G_2 = \frac{H(S^2)}{2} = \frac{2.6174}{2}$$

$\therefore$ G_2 **= 1.3087 bits/message symbol**

The symbols produced by the source at the end of third symbol interval will have an alphabet given by (note that ZZZ is repeated).

$$\bar{S}^3 = \{XXX, XXZ, XZZ, XZY, ZZX, ZZZ, ZYZ, ZYY, ZXX, ZXZ, ZZY, YZX, YZZ, YYZ, YYY\}$$

The corresponding probabilities are calculated in a similar way as

$$P(\bar{S}^3) = \left\{\frac{8}{63}, \frac{4}{63}, \frac{1}{42}, \frac{1}{14}, \frac{1}{42}, \frac{1}{42}, \frac{3}{112}, \frac{9}{112}, \frac{4}{63}, \frac{2}{63}, \frac{1}{28}, \frac{1}{14}, \frac{1}{28}, \frac{9}{112}, \frac{27}{112}\right\}$$

The entropy of the source at the end of third symbol interval is given by

$$H(\bar{S}^3) = \frac{8}{63}\log\frac{63}{8} + \left(\frac{4}{63}\log\frac{63}{4}\right)(2) + \left(\frac{1}{42}\log 42\right)(3)$$

$$+\frac{1}{14}\log 14 + \frac{3}{112}\log\frac{112}{3} + \left(\frac{9}{112}\log\frac{112}{9}\right)(2)$$

$$+\frac{2}{63}\log\frac{63}{2} + \left(\frac{1}{28}\log 28\right)(2) + \frac{1}{14}\log 14 + \frac{27}{112}\log\frac{112}{27}$$

$$= 3.533 \text{ bits/message symbol.}$$

From equation (1.43) for N = 3, we have

$$G_3 = \frac{H(S^3)}{3} = \frac{3.533}{3}$$

$\therefore$ G_3 **= 1.178 bits/message symbol**

It can be easily observed that

$$G_1 > G_2 > G_3 > H$$

Example 1.35 : For the first order Markoff model shown in figure 1.13, find the state probabilities, entropy of each state and the entropy of the source.

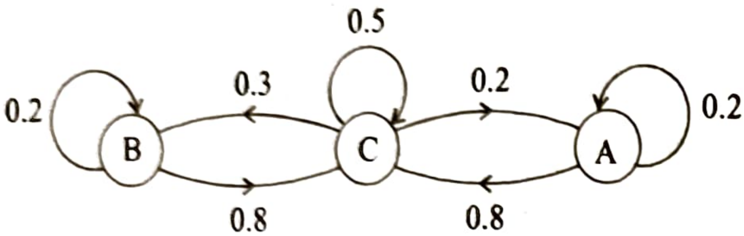

Fig. 1.13 : Markoff model for example 1.35

Solution

From the Markoff model of figure 1.13, the state equations are given by

for state A → $P(A) = 0.2\,P(A) + 0.2\,P(C)$ .... (1.64)

for state B → $P(B) = 0.2\,P(B) + 0.3\,P(C)$ .... (1.65)

for state C → $P(C) = 0.8\,P(A) + 0.8\,P(B) + 0.5\,P(C)$ .... (1.66)

From equation (1.64),

$$0.8\,P(A) = 0.2\,P(C)$$

$$\therefore \quad P(A) = \frac{1}{4}\,P(C) \qquad \text{.... (1.67)}$$

From equation (1.65)

$$0.8\,P(B) = 0.3\,P(C)$$

$$\therefore \quad P(B) = \frac{3}{8}\,P(C) \qquad \text{.... (1.68)}$$

We have

$$P(A) + P(B) + P(C) = 1$$

$$\therefore \quad \frac{1}{4}\,P(C) + \frac{3}{8}\,P(C) + P(C) = 1$$

$$\therefore \quad P(C)\left[\frac{2+3+8}{8}\right] = 1$$

$$\therefore \quad \mathbf{P(C) = \frac{8}{13}}$$

From equation (1.67)

$$\therefore \quad P(A) = \frac{1}{4}\,P(C) = \left(\frac{1}{4}\right)\left(\frac{8}{13}\right) = \frac{2}{13}$$

$$\therefore \quad P(A) = \frac{2}{13}$$

From equation (1.68), $P(B) = \left(\frac{3}{8}\right) P(C) = \left(\frac{3}{8}\right)\left(\frac{8}{13}\right)$

$$\therefore \quad P(B) = \frac{3}{13}$$

R.H.S. of equation (1.66) is

$$0.8\ P(A) + 0.8\ P(B) + 0.5\ P(C) = (0.8)\left(\frac{2}{13}\right) + (0.8)\left(\frac{3}{13}\right) + (0.5)\left(\frac{8}{13}\right)$$

$$= \frac{8}{13} = P(C) = \text{L.H.S of equation (1.66)} \rightarrow \text{Verified}$$

$\therefore$ The state probabilities are given by

$$P(A) = \frac{2}{13},\ P(B) = \frac{3}{13} \text{ and } P(C) = \frac{8}{13}$$

From equation (1.40), the entropy of each state is given by

$$H_i = \sum_{j=1}^{n} p_{ij} \log \frac{1}{p_{ij}}$$

For state A,

$$H_A = \sum_{j=A}^{C} p_{Aj} \log \frac{1}{p_{Aj}}$$

$$= p_{AA} \log \frac{1}{p_{AA}} + p_{AB} \log \frac{1}{p_{AB}} + p_{AC} \log \frac{1}{p_{AC}}$$

$$= 0.2 \log \frac{1}{0.2} + 0 + 0.8 \log \frac{1}{0.8}$$

$$\therefore \quad H_A = 0.722 \text{ bits/message symbol}$$

For state B,

$$H_B = \sum_{j=A}^{C} p_{Bj} \log \frac{1}{p_{Bj}}$$

$$= p_{BA} \log \frac{1}{p_{BA}} + p_{BB} \log \frac{1}{p_{BB}} + p_{BC} \log \frac{1}{p_{BC}}$$

$$= 0 + 0.2 \log \frac{1}{0.2} + 0.8 \log \frac{1}{0.8}$$

$$\therefore \quad H_B = 0.722 \text{ bits/message symbol}$$

For state C,

$$H_C = \sum_{j=A}^{C} p_{Cj} \log \frac{1}{p_{Cj}}$$

$$= p_{CA} \log \frac{1}{p_{CA}} + p_{CB} \log \frac{1}{p_{CB}} + p_{CC} \log \frac{1}{p_{CC}}$$

$$= 0.2 \log \frac{1}{0.2} + 0.3 \log \frac{1}{0.3} + 0.5 \log \frac{1}{0.5}$$

$\therefore$ H_C = **1.485 bits/message symbol**

From equation (1.41), the entropy of the source is given by

$$H = \sum_{i=A}^{C} P(i)\, H_i$$

$$= P(A)\, H_A + P(B)\, H_B + P(C)\, H_C$$

$$= \left(\frac{2}{13}\right)(0.722) + \left(\frac{3}{13}\right)(0.722) + \left(\frac{8}{13}\right)(1.485)$$

$\therefore$ H = **1.192 bits/message symbol**

Example 1.36 : For the first order Markoff source shown in figure 1.14,

(i) Find the stationary distribution

(ii) Find the entropy of each state and hence the entropy of the source.

(iii) Find the entropy of the adjoint source and verify that $H(\bar{S}) < H(S)$.

[VI sem EC/TE, July/August 2003, Q.3 (b)]

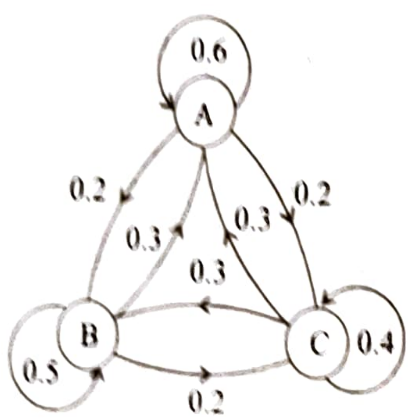

Fig. 1.14 : Markoff model for example 1.36

Solution

(i) From the state diagram of figure 1.14, the state equations are given by

for state A	$\rightarrow$	$P(A) = 0.6\, P(A) + 0.3\, P(B) + 0.3\, P(C)$	.... (1.69)
for state B	$\rightarrow$	$P(B) = 0.2\, P(A) + 0.5\, P(B) + 0.3\, P(C)$	.... (1.70)
for state C	$\rightarrow$	$P(C) = 0.2\, P(A) + 0.2\, P(B) + 0.4\, P(C)$	.... (1.71)

Substracting equation (1.71) from (1.70), we get

$$P(B) - P(C) = 0.3\ P(B) - 0.1\ P(C)$$

$$0.7\ P(B) = 0.9\ P(C)$$

$\therefore$

$$P(B) = \frac{9}{7} P(C) \qquad \text{.... (1.72)}$$

$\therefore$

Multiplying equation (1.69) by 2, we get

$$2\ P(A) = 1.2\ P(A) + 0.6\ P(B) + 0.6\ P(C) \qquad \text{.... (1.73)}$$

Multiplying equation (1.71) by 3, we get

$$3\ P(C) = 0.6\ P(A) + 0.6\ P(B) + 1.2\ P(C) \qquad \text{.... (1.74)}$$

Subtracting equation (1.74) from (1.74), we get

$$2\ P(A) - 3\ P(C) = 0.6\ P(A) - 0.6\ P(C)$$

$$1.4\ P(A) = 1.4\ P(C)$$

$\therefore$

$$P(A) = \frac{12}{7} P(C) \qquad \text{.... (1.75)}$$

$\therefore$

We have, $P(A) + P(B) + P(C) = 1$ .... (1.76)

Using equations (1.72) and (1.75) in (1.76) we get

$$\frac{12}{7} P(C) + \frac{9}{7} P(C) + P(C) = 1$$

$$\therefore \quad P(C)\left[\frac{12}{7} + \frac{9}{7} + 1\right] = 1$$

$$\therefore \quad P(C) = \frac{1}{4}$$

From equation (1.75), $P(A) = \frac{12}{7} \times \frac{1}{4}$

$$\therefore \quad P(A) = \frac{3}{7}$$

From equation (1.72), $P(B) = \frac{9}{7} P(C) = \frac{9}{7} \times \frac{1}{4}$

$$\therefore \quad P(B) = \frac{9}{28}$$

$\therefore$ The stationary distribution which is nothing but the state probabilities are given by

$$P(A) = \frac{3}{7}, \quad P(B) = \frac{9}{28} \quad \text{and} \quad P(C) = \frac{1}{4}$$

(ii) The entropy of each state is given by equation (1.40) as

$$H_i = \sum_{j=A}^{C} p_{ij} \log \frac{1}{p_{ij}}$$

For state-A,

$$H_A = \sum_{j=A}^{C} p_{Aj} \log \frac{1}{p_{Aj}}$$

$\therefore$

$$H_A = p_{AA} \log \frac{1}{p_{AA}} + p_{AB} \log \frac{1}{p_{AB}} + p_{AC} \log \frac{1}{p_{AC}}$$

$$= 0.6 \log \frac{1}{0.6} + 0.2 \log \frac{1}{0.2} + 0.2 \log \frac{1}{0.2}$$

$$= 1.371 \text{ bits/message symbol}$$

For state-B,

$$H_B = \sum_{j=A}^{C} p_{Bj} \log \frac{1}{p_{Bj}}$$

$$= p_{BA} \log \frac{1}{p_{BA}} + p_{BB} \log \frac{1}{p_{BB}} + p_{BC} \log \frac{1}{p_{BC}}$$

$$= 0.3 \log \frac{1}{0.3} + 0.5 \log \frac{1}{0.5} + 0.2 \log \frac{1}{0.2}$$

$$= 1.485 \text{ bits/message symbol.}$$

For state-C,

$$H_C = \sum_{j=A}^{C} p_{Cj} \log \frac{1}{p_{Cj}}$$

$$= p_{CA} \log \frac{1}{p_{CA}} + p_{CB} \log \frac{1}{p_{CB}} + p_{CC} \log \frac{1}{p_{CC}}$$

$$= 0.3 \log \frac{1}{0.3} + 0.3 \log \frac{1}{0.3} + 0.4 \log \frac{1}{0.4}$$

$$= 1.571 \text{ bits/message symbol}$$

From equation (1.41), the entropy of source is given by

$$H(S) = H = \sum_{i=A}^{C} P(i)\, H_i$$

$$= P(A)\, H_A + P(B)\, H_B + P(C)\, H_C$$

$$= \left(\frac{3}{7}\right)(1.371) + \left(\frac{9}{28}\right)(1.485) + \left(\frac{1}{4}\right)(1.571)$$

$\therefore$

$$\mathbf{H(S) = 1.458 \text{ bits/message symbol}}$$

(iii) From equation (1.43) for N = 1,

$$G_1 = H(\bar{S}) = \sum_{i=1}^{3} P(m_i) \log \frac{1}{P(m_i)}$$

$$= P(A) \log \frac{1}{P(A)} + P(B) \log \frac{1}{P(B)} + P(C) \log \frac{1}{P(C)}$$

$$= \frac{3}{7} \log \frac{7}{3} + \frac{9}{28} \log \frac{28}{9} + \frac{1}{4} \log 4$$

$\therefore$ $H(\bar{S})$ = 1.55 bits/message symbol

But $H(S)$ = 1.458 bits/message symbol

$\therefore$ **$H(S) < H(\bar{S})$ → Proved**

MISCELLANEOUS EXAMPLES

Example 1.37 : An analog signal band limited to 6 KHZ is sampled at thrice the Nyquist rate and then quantized into 11 levels $Q_1, Q_2, \ldots, Q_{11}$. Of these, three levels occur with probability of $\frac{1}{6}$ each, four levels with probability of $\frac{1}{12}$ each and the remaining four levels with probability $\frac{1}{24}$ each. Find the rate of information associated with the analog signal.

Solution : Given sampling rate r_s = 3 × Nyquist rate

= 3 × 2B

= 3 × 2 × 6K = 36 K-samples/sec.

The entropy of the analog signal is given by

$$H(S) = \sum_{i=1}^{11} p_i \log \frac{1}{p_i} \text{ bits/m-sym or sample}$$

$$= 3 \times \frac{1}{6} \log 6 + 4 \times \frac{1}{12} \log 12 + 4 \times \frac{1}{24} \log 24$$

= 3.252 bits/sample

$\therefore$ Rate of information = (r_s) [H(S)]

$\therefore$ R_1 = (36 K-samples/sec) (3.252 bits/sample)

$\therefore$ **R_1 = 117.072 KBPS**

Example 1.38 : A fair coin is tossed repeatedly.

Let X = {event of getting 5 heads out of 8 trials}

Y = {event of getting 7 heads out of 11 trials}

Which event conveys more information? Support your answer by numerical computation of respective amounts of information.

Solution : Let,

$$Z = \text{number of heads occurring}$$
$$= \text{binomial R.V. with parameters}$$
$$n = \text{number of trials}$$
$$p = \text{probability of getting head} = 0.5$$
$$q = 1 - p = 1 - 0.5 = 0.5$$

From equation (1.47), we have,

$$P(Z = z) = {}_nC_z p^z q^{n-z}$$

$$\therefore \quad P(X) = P(X = 5) = {}_nC_x p^x (1-p)^{n-x}$$
$$= {}_8C_5 (0.5)^5 (1-0.5)^{8-5}$$
$$= 0.21875$$

$$P(Y) = P(Y = 7) = {}_nC_y p^y (1-p)^{n-y}$$
$$= {}_{11}C_7 (0.5)^7 (1-0.5)^{11-7}$$
$$= 0.16113$$

$$\therefore \quad \text{Self-information of X} = I(X) = \log \frac{1}{P(X)} = \log \frac{1}{0.21875}$$

$$\therefore \quad I(X) = 2.193 \text{ bits}$$

$$\text{Self-information of Y} = I(Y) = \log \frac{1}{P(Y)} = \log \frac{1}{0.16113}$$

$$\therefore \quad I(Y) = 2.634 \text{ bits}$$

$\therefore$ Y carries more information.

Example 1.39 **:** A pair of dice are tossed simultaneously. The outcome of first dice is recorded as x_1 and that of second dice as x_2. Three events are defined as below:

$$A = \{(x_1, x_2) \text{ such that } (x_1 + x_2) \text{ is divisable exactly by } 4\}$$
$$B = \{(x_1, x_2) \text{ such that } 6 \le (x_1 + x_2) \le 8\}$$
$$C = \{(x_1, x_2) \text{ such that } x_1 x_2 \text{ is divisable exactly by } 3\}$$

Which events conveys maximum information? Support your answer by numerical computation.

Solution : From example 1.20, the sample space S has thirty six combinations given by

$$S = \begin{bmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{bmatrix}$$

The event-A contains the pairs given by

$$A = \{(1, 3), (2, 2), (2, 6), (3, 1), (3, 5), (4, 4), (5, 3), (6, 2), (6, 6)\}$$

$$\therefore \quad P(A) = \frac{9}{36} = 0.25$$

The event-B contains

$$B = \{(1, 5), (1, 6), (2, 4), (2, 5), (2, 6), (3, 3), (3, 4), (3, 5), (4, 2), (4, 3), (4, 4), (5, 1), (5, 2), (5, 3), (6, 1), (6, 2)\}$$

$$\therefore \quad P(B) = \frac{16}{36} = 0.444$$

The event C contains

$$C = \{(1, 3), (1, 6), (2, 3), (2, 6), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6), (4, 3), (4, 6), (5, 3), (5, 6), (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)\}$$

$$\therefore \quad P(C) = \frac{20}{36} = 0.556$$

$$\therefore \quad \text{Self-information of A} = I_A = \log\frac{1}{0.25} = 2 \text{ bits}$$

$$\text{Self-information of B} = I_B = \log\frac{1}{0.444} = 1.17 \text{ bits}$$

$$\text{Self-information of C} = I_C = \log\frac{1}{0.555} = 0.848 \text{ bits}$$

$\therefore$ The event-A carries the maximum information.

Example 1.40 : A voice signal in a PCM system is sampled at 2.5 times the Nyquist rate and is quantized into 16 levels with the following probabilities:

$$p_1 = p_2 = p_3 = p_4 = 0.08$$
$$p_5 = p_6 = p_7 = p_8 = 0.065$$
$$p_9 = p_{10} = p_{11} = p_{12} = 0.055$$
$$p_{13} = p_{14} = p_{15} = p_{16} = 0.05$$

Calculate the entropy and information rate of the PCM signal if the bandwidth of signal is 3.5 KHz.

Solution :

$$\text{Entropy} = H(S) = \sum_{i=1}^{16} p_i \log\frac{1}{p_i}$$

$$H(S) = 4 \times 0.08 \log\frac{1}{0.08} + 4 \times 0.065 \log\frac{1}{0.065} + 4 \times 0.055 \log\frac{1}{0.055} + 4 \times 0.05 \log\frac{1}{0.05}$$

$$\therefore \quad H(S) = 3.9763 \text{ bits/level or sample.}$$

Given sampling rate $= (2.5)$ (Nyquist rate)

$$= (2.5)(2B)$$
$$= (2.5)(2)(3.5\ K)$$
$$= 17.5 \text{ K-samples/sec.}$$

$\therefore$ Information rate $R_I = r_s H(S)$

$$= (17.5 \text{ K-samples/sec}) (3.9763 \text{ bits/sample})$$

$\therefore$ $\mathbf{R_I = 69.585\ KBPS}$

Example 1.41 : A discrete source emits the messages m_1, m_2, m_3 and m_4 with probabilities $\frac{7}{15}, \frac{4}{15}, \frac{3}{15}$ and $\frac{1}{15}$ respectively. Calculate the information content of each message and the average information contents per message. Show that the entropy of the source increases when the symbol m_4 is partitioned into three sub-symbols m_{41}, m_{42} and m_{43} with respective probabilities $\frac{1}{24}, \frac{1}{60}$ and $\frac{1}{120}$.

Solution :

Information content of each message :

$$\text{Self-information in } m_1 = I_{m_1} = \log \frac{1}{(7/15)} = 1.0995 \text{ bits}$$

$$\text{Self-information in } m_2 = I_{m_2} = \log \frac{1}{(4/15)} = 1.9069 \text{ bits}$$

$$\text{Self-information in } m_3 = I_{m_3} = \log \frac{1}{(3/15)} = 2.322 \text{ bits}$$

$$\text{Self-information in } m_4 = I_{m_4} = \log \frac{1}{(1/15)} = 3.9069 \text{ bits}$$

$$\text{Average information} = H(S) = \sum_{i=1}^{4} p_{m_i} \log \frac{1}{p_{m_i}}$$

$$\therefore \quad H(S) = \frac{7}{15} \log \frac{15}{7} + \frac{4}{15} \log \frac{15}{4} + \frac{3}{15} \log \frac{15}{3} + \frac{1}{15} \log 15$$

$$\therefore \quad \mathbf{H(S) = 1.7465\ bits/m\text{-}symbol}$$

When m_4 is partitioned into three sub-symbols, then the source S' will have six symbols given by m_1, m_2, m_3, m_{41}, m_{42} and m_{43} with respective probabilities of $\frac{7}{15}, \frac{4}{15}, \frac{3}{15}, \frac{1}{24}, \frac{1}{60}$ and $\frac{1}{120}$. The new entropy of source S' is given by

$$H(S') = \sum_{j=1}^{6} p_j \log \frac{1}{p_j}$$

$$= \frac{7}{15}\log\frac{15}{7} + \frac{4}{15}\log\frac{15}{4} + \frac{3}{15}\log\frac{15}{3}$$

$$+ \frac{1}{24}\log 24 + \frac{1}{60}\log 60 + \frac{1}{120}\log 120$$

$\therefore$ $$H(S') = 1.8331 \text{ bits/m-symbol}$$

$\therefore$ $$H(S') > H(S) \quad \rightarrow \text{ proved}$$

Example 1.42 : Consider a source with alphabets m_1 and m_2 with respective probabilities of $\frac{5}{6}$ and $\frac{1}{6}$. Determine the entropy of source S and the entropy of its third extension. Hence show that $H(S^3) = 3H(S)$.

Solution : Entropy of source S is given by

$$H(S) = \sum_{i=1}^{2} p_{m_i} \log \frac{1}{p_{m_i}}$$

$$= \frac{5}{6}\log\frac{6}{5} + \frac{1}{6}\log 6$$

$\therefore$ $$H(S) = 0.65 \text{ bits/message-symbol}$$

3rd Extension :

Symbols(σ_i)	$m_1m_1m_1$	$m_1m_1m_2$	$m_1m_2m_1$	$m_2m_1m_1$	$m_1m_2m_2$	$m_2m_1m_2$	$m_2m_2m_1$	$m_2m_2m_2$
$P(\sigma_i)$	$\frac{125}{216}$	$\frac{25}{216}$	$\frac{25}{216}$	$\frac{25}{216}$	$\frac{5}{216}$	$\frac{5}{216}$	$\frac{5}{216}$	$\frac{1}{216}$

$\therefore$ Entropy of 3rd extended source is given by

$$H(S^3) = \sum_{i=1}^{8} P(\sigma_i) \log \frac{1}{P(\sigma_i)}$$

$$= \frac{125}{216}\log\frac{216}{125} + 3\times\frac{25}{216}\log\frac{216}{25}$$

$$+ 3\times\frac{5}{216}\log\frac{216}{5} + \frac{1}{216}\log 216$$

$\therefore$ $$H(S^3) = 1.95 \text{ bits/m-symbol}$$

$$= 3 \times 0.65 \text{ bits/message-symbol}$$

$\therefore$ $$H(S^3) = 3\,H(S) \quad \rightarrow \text{ proved}$$

Example 1.43 : The state diagram of a Markov source is shown in fig. 1.15(i). Find entropy H of the source (ii) Find G_1 and G_2 and hence show that $G_1 > G_2 > H$.

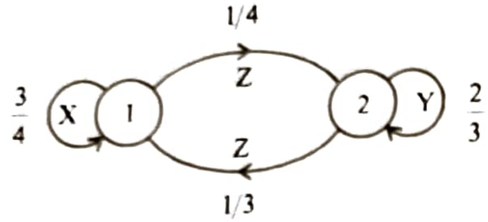

Fig. 1.15 : Markov Source for Example 1.43

Solution : The state equations of the given Markov source are given by

for state 1 : $$P(1) = \frac{3}{4}P(1) + \frac{1}{3}P(2) \quad \text{..... (1.77)}$$

for state 2 : $$P(2) = \frac{1}{4}P(1) + \frac{2}{3}P(2) \quad \text{..... (1.78)}$$

And : $$P(1) + P(2) = 1 \quad \text{..... (1.79)}$$

From equation (1.77), $$\frac{1}{4}P(1) = \frac{1}{3}P(2)$$

$$\therefore \quad P(2) = \frac{3}{4}P(1) \quad \text{..... (1.80)}$$

Using equation (1.80) in (1.79)

$$P(1) + \frac{3}{4}P(1) = 1$$

$$\therefore \quad \frac{7}{4}P(1) = 1$$

$$\therefore \quad \mathbf{P(1) = \frac{4}{7}}$$

$$\therefore \quad P(2) = \frac{3}{4}P(1) = \frac{3}{4} \times \frac{4}{7} \quad \therefore \quad \mathbf{P(2) = \frac{3}{7}}$$

(i) From equation (1.40), the entropy of the 1st state is given by

$$H_1 = \sum_{j=1}^{2} p_{1j} \log \frac{1}{p_{1j}}$$

$$= p_{11} \log \frac{1}{p_{11}} + p_{12} \log \frac{1}{p_{12}}$$

$$= \frac{3}{4} \log \frac{4}{3} + \frac{1}{4} \log 4$$

$$\therefore \quad \mathbf{H_1 = 0.8113 \text{ bits/m-sym}}$$

And
$$H_2 = \sum_{j=1}^{2} p_{2j} \log \frac{1}{p_{2j}}$$
$$= p_{21} \log \frac{1}{p_{21}} + p_{22} \log \frac{1}{p_{22}}$$
$$= \frac{1}{3} \log 3 + \frac{2}{3} \log \frac{3}{2} \quad \therefore \ \mathbf{H_2 = 0.9183 \text{ bits/message-symbol}}$$

From equation (1.41), the entropy H of the source is given by
$$H = \sum_{i=1}^{2} P(i)\, H_i$$
$$= P(1)\, H_1 + P(2)\, H_2$$
$$= \left(\frac{4}{7}\right)(0.8113) + \left(\frac{3}{7}\right)(0.9183) \quad \therefore \mathbf{H = 0.8572 \text{ bits/message-symbol}}$$

(ii) Fig. 1.16 shows the tree corresponding to initial state and the state at the beginning of the second symbol interval.

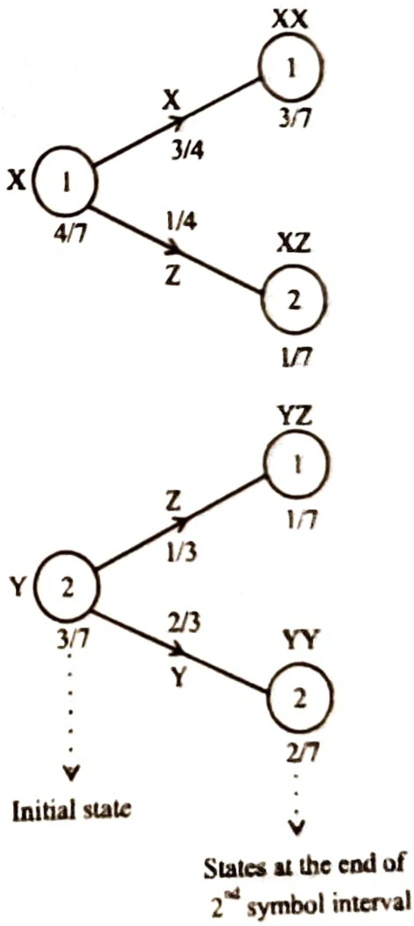

Fig. 1.16 : Tree diagram for Markov source of example 1.43

From equation (1.43) with N = 1 (i.e., at the start of the 1st symbol interval equal to initial state)

$$G_1 = \frac{1}{1} H(\bar{S}^1)$$

$$= H(\bar{S})$$

$$= \frac{4}{7} \log \frac{7}{4} + \frac{3}{7} \log \frac{7}{3}$$

$$\therefore \quad \mathbf{G_1 = 0.9852 \text{ bits/message-symbol}}$$

Comparing H and G_1 values, we conclude that

$$\mathbf{G_1 > H}$$

Again from equation (1.43), with N = 2 (i.e., at the start of the 2nd symbol interval), we have,

$$G_2 = \frac{1}{2} H(\bar{S}^2)$$

$$= \frac{1}{2}\left[\frac{3}{7} \log \frac{7}{3} + \frac{1}{7} \log 7 + \frac{1}{7} \log 7 + \frac{2}{7} \log \frac{7}{2}\right]$$

$$\therefore \quad \mathbf{G_2 = 0.9212 \text{ bits/message-symbol}}$$

Again comparing values of H, G_1 and G_2, we conclude that,

$$\mathbf{G_2 > G_1 > H}$$

1.7 REVIEW QUESTIONS

1. With a neat diagram explain the block diagram of an information system.
2. Define (a) symbol rate (b) self-information (c) zero-memory source (d) average self-information (entropy) (e) information rate.
3. Obtain an expression for entropy of a zero-memory information source emitting independent sequence of symbols.
4. Discuss the various properties of entropy.
5. Discuss extremal property of entropy with example.
6. Discuss additive property of entropy with example.
7. What do you understand by extension of zero-memory information source.
8. Derive an expression for the entropy of n^{th} extension of a zero-memory source.
9. What is MARKOFF'S source? Explain with an example.
10. Obtain an expression for entropy of Markoff's source.
11. What is adjoint source and how its entropy is related to entropy of Markoff source?

1.8 PROBLEMS

1. If quadruple is the unit of information associated with the use of base 4, determine its relation with a nat.

2. A discrete source emits one of four possible message s_1, s_2, s_3 and s_4 with probabilities $\frac{5}{18}, \frac{1}{6}, \frac{1}{3}$ and $\frac{2}{9}$ respectively. Calculate the information content of each message and the average information content per message.
3. The output of an information source consists of 128 symbols, 16 of which occur with a probability of 1/32 each and the remaining 112 occur with a probability of 1/224 each. The source emits 1000 symbols/sec. Assuming that the symbols are chosen independently, find the average information rate of this source.
4. A source emits an independent sequence of symbols from an alphabet consisting of five symbols A, B, C, D and E with symbol probabilities $\frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{3}{16}, \frac{5}{16}$ respectively. Find the entropy of the source.
5. A binary source $S = \{s_1, s_2\}$ has $P = \{p, (1-p)\}$. Plot the entropy of the source H(S) versus p, as p varies from 0 to 1.
6. A message source produces two independent symbols A and B with probabilities $P(A) = 0.3$ and $P(B) = 0.7$. Calculate the efficiency of the source and its redundancy. What is the significance of redundancy?
7. An analog signal bandlimited to 4 KHz, is sampled at twice the Nyquist rate and then quantized into 8 levels Q_1, Q_2, ..., Q_8. Of these, 2 levels occur with a probability of 1/4 each, 2 with probability of 1/8 each and the remaining four with a probability of 1/16 each respectively. Find the information rate associated with the analog signal.
8. A fair coin is tossed repeatedly. Let

   X = {event of getting 4 heads out of 7 trials}

   Y = {event of getting 5 heads out of 9 trials}

   Which event conveys more information? Support your answer by numerical computation of respective amounts of information.
9. Find the entropy of a source in Hartley/symbol of a source that emits one out of four symbols A, B, C and D in a statistically independent sequence with probabilities $\frac{5}{13}, \frac{4}{13}, \frac{3}{13}$ and $\frac{1}{13}$.
10. Shortly before a horse-race, a book-maker believes that several horses entered for the race have the following probability of winning:

A	B	C	D	E	F	G	H
0.3	0.25	0.15	0.10	0.05	0.05	0.05	0.05

   He, then receives a message that owing to an accident one of the horses will not run. Explain how you would assess, from an information theory point of view, the information value of this message, (i) if the horse in question is known and (ii) if it is not known.
11. In a facsimile transmission of picture, there are about 3×10^6 pixels in one frame. For a good reproduction, 16 brightness levels are found to be necessary. Assuming all these levels to occur equally likely, determine the rate of information transmission if 1 picture frame is to be transmitted every 2 minutes.

12. Conventional telegraphy uses two symbols, the dot and the dash. Assuming that the dash is twice as long as the dot and half as probable, find the average symbol rate and the entropy rate.
13. A man is informed that when a pair of dice were rolled, the result was seven. How much information is there in the message?
14. Suppose that a large field is divided into 64 squares. In the dark night, a cow has entered in this field and it is equally likely to be in any of the squares. This cow is located by a member of searching party who sends back an information giving the location of the cow as 43rd square. Calculate the amount of information obtained in the reception of the message.
15. Find the information content of a message that consists of a digital word 10 digits long in which each digit may take on one of six possible levels. The probability of sending any of the six levels is assumed to be equal, and the level in any digit does not depend on the values taken by the previous digits.
16. A pair of dice are tossed simultaneously. The outcome of the first dice is recorded as "x_1" and that of second dice as "x_2". Three events are defined as follows:

    $X = \{(x_1, x_2)$ such that $(x_1 + x_2)$ is exactly divisable by $3\}$

    $Y = \{(x_1, x_2)$ such that $(x_1 + x_2)$ is an odd number$\}$

    $Z = \{(x_1, x_2)$ such that $5 \le (x_1 + x_2) \le 7\}$

    Which event conveys more information? Support your answer by numerical computation.
17. A discrete message source S emits three independent symbols X, Y and Z with probabilities 0.35, 0.37 and 0.28 respectively. Calculate the efficiency of the source and its redundancy.
18. Consider a source with alphabets x_1, x_2, x_3, x_4 with respective probabilities $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}$. Determine the entropy of the source H(S) and show that for the 2nd extension of the source $H(S^2) = 2H(S)$ by listing all the 2nd extension symbols and their respective probabilities.
19. Consider a source with alphabets x_1 and x_2 with respective probabilities $\frac{1}{4}$ and $\frac{3}{4}$. Determine the entropy of the source H(S) and show that for the third extension of the source $H(S^3) = 3H(S)$ by listing the symbols of third extended source along with their probabilities.
20. The probabilities of occurrence of various letters of the English alphabet are given below:

A	0.081	J	0.001	S	0.066
B	0.016	K	0.005	T	0.096
C	0.032	L	0.040	U	0.031
D	0.037	M	0.022	V	0.009
E	0.124	N	0.072	W	0.020
F	0.023	O	0.079	X	0.002
G	0.016	P	0.023	Y	0.019
H	0.051	Q	0.002	Z	0.001
I	0.072	R	0.060		

(a) Which letter conveys the maximum amount of information?
(b) Which letter conveys the minimum amount of information?
(c) What is the entropy of English text if you can assume that letters are chosen independently to form words and sentences (which is not true indeed!).

21. A source emits one of the four probable messages m_1, m_2, m_3 and m_4 with probabilities $\frac{4}{11}, \frac{3}{11}, \frac{2}{11}$ and $\frac{2}{11}$ respectively. Find the entropy of the source. List all the elements for the second extension of this source. Hence show that $H(S^2) = 2\,H(S)$.

22. Consider an information source modelled by a Markov process whose state diagram is shown in figure 1.17. Find the source entropy H and the average information content per symbol in messages containing one, two and three symbols. i.e., find G_1, G_2 and G_3.

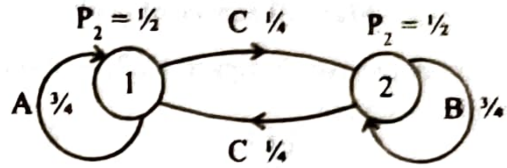

Fig. P-22 : Markov source for problem 22

23. For the first order Markov source shown in figure 1.18, calculate the state probabilities, the entropy of the states and the entropy of the source.

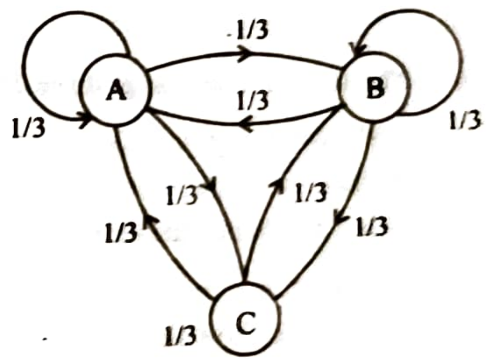

Fig. P-23 : First order Markov source of problem 23

24. For the first order Markov source shown in figure 1.19, repeat problem 23.

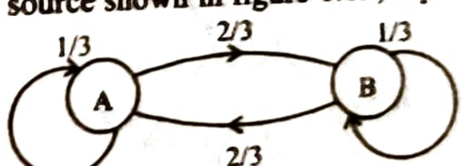

Fig. P-24 : First order Markov source of problem 24

25. The state diagram of a second-order Markov source is given in figure 1.20. Calculate the state probabilities, the entropy of the states and the entropy of the source.

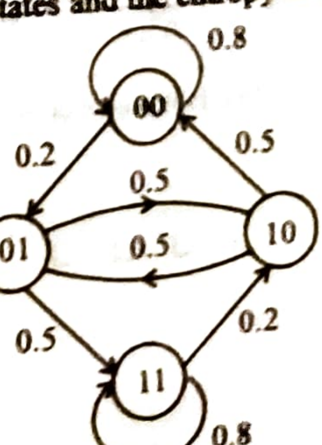

Fig. P-25 : Markov Source of problem 25