31. Concept of Information and Entropy

The information content of a message must be established to determine if the message can be transmitted over a channel with a *channel capacity* of C_C.

In the context of information theory, signals which change unpredictably with time give useful information. Repetitive signals contain no information. The greater the uncertainty of a signal occurring, the more the information carried. For simplicity, we shall assume discrete signals.

Consider the voltage-time diagram of Figure 31.1 (a).

Figure 31.1 (a) Voltage-time diagram, (b) message 1, and (c) message 2.

Suppose we take a duration of T seconds and divide it into m intervals of τ seconds each; then

$$m = \frac{T}{\tau}. \qquad (31.1)$$

In each interval, there exists n possible signal levels. Figure 31.1 (b) and (c) show two possible discrete signal messages.

Given n possible signal levels per interval of τ seconds, the total number of signal combinations in T seconds is $n^{T/\tau}$. Clearly, the presence of noise can reduce our capacity to distinguish distinct signals.

How much information can be transmitted in a message lasting T seconds? Intuitively, we can assume that the information associated with a message is proportional to T. By taking the logarithm of $n^{T/\tau}$, the amount of information of the message in T seconds can now be made proportional to $\frac{T}{\tau}\log n$. If base e is used, the unit of information is in *nats*. If base-2 is used, the unit of information is in *bits*. In practice, we prefer base-2. A combination of 1s and 0s can then be used to represent a n-level signal rather than sending n discrete values in the same interval.

Let us assume that the amount of information in a message lasting T seconds is

$$\frac{T}{\tau}\log_2 n \qquad \text{bits in } T \text{ seconds} \qquad (31.2)$$

Thus the amount of information transmitted per second is

$$\frac{1}{\tau}\log_2 n \qquad \text{bits/second} \qquad (31.3)$$

The information content of a message thus relates not only to the relative frequency of occurrence but also to the total number of signal combinations.

Sketch of Proof.

If n possible events are specified to be the n possible signal levels at any interval, then the probability of occurrence for equally likely events is $p = 1/n$.

The information carried by the appearance of any one event in one interval is

$$\log_2 n = -\log_2 p \qquad \text{bits/interval}$$

and the total information in a message lasting m intervals is

$$-m\log_2 p \qquad \text{bits in } m \text{ intervals}$$

Since $m = \frac{T}{\tau}$, the total information in a message lasting T seconds is

$$-\frac{T}{\tau}\log_2 p = \frac{T}{\tau}\log_2 n \qquad \text{bits in } T \text{ seconds}$$

and the total information carried per second is

$$\frac{1}{\tau}\log_2 n \qquad \text{bits/second} \qquad \square$$

Entropy

Consider the case of transmission of n possible signal levels in an interval of τ seconds.

$$\sum_{i=1}^{n} p_i = 1$$

If the n possible events are specified to be the n possible signal levels at any interval, and event i with p_i occurs a_i times in a_{tot} samples, then each time the event i occurs, we get $-\log_2 p_i$ bits of information per interval. Thus, the average information content of a_{tot} appearances of all events per interval is

$$H_{av} = \frac{1}{a_{tot}} \sum_{i=1}^{n} a_i (-\log_2 p_i)$$
$$= - \sum_{i=1}^{n} p_i \log_2 p_i \qquad \text{bits/interval} \qquad (31.4)$$

The quality H_{av} is called the *entropy*.

The amount of information in a message carried in m intervals is

$$H = - mH_{av}$$
$$= - m \sum_{i=1}^{n} p_i \log_2 p_i \qquad \text{bits in } m \text{ intervals} \qquad (31.5)$$

Since $m = \frac{T}{\tau}$, the amount of information in a message carried in T seconds is

$$H = -\frac{T}{\tau} \sum_{i=1}^{n} p_i \log_2 p_i \qquad \text{bits in } T \text{ seconds} \qquad (31.6)$$

The average amount of information carried in each second is

$$H_{av} = - \frac{1}{\tau} \sum_{i=1}^{n} p_i \log_2 p_i \qquad \text{bits/second} \qquad (31.7)$$

A channel capable of transmitting this information should have an average capacity

$$C_c \geq H_{av}. \qquad (31.8)$$

C_c is called the *channel capacity*. It is defined as the *maximum* amount of information per second, usually in bit/s, that can be sent over a channel.

Example 31.1

If $n = 2$, $p_1 = p$ and $p_2 = q$, we have

$$H = -m(p \log_2 p + q \log_2 q) \qquad \text{bits in } m \text{ intervals}$$

and

$$H_{av} = -(p \log_2 p + q \log_2 q) \qquad \text{bits/interval}$$

When $p = 0.5$, we get

$$H_{av} = -(0.5 \log_2 0.5 + 0.5 \log_2 0.5)$$
$$= 1 \qquad \text{bit/interval}$$

Figure 31.2 A plot of H_{av} against p.

Example 31.2

If $p_1 = p_2 = ... = p_i = 1/n$ for $1 \le i \le n$, we have

$$H = -m \sum_{i=1}^{n} \frac{1}{n} \log_2 \frac{1}{n}$$
$$= m \log_2 n \qquad \text{bits in } m \text{ intervals}$$

and

$$H_{av} = -\sum_{j=1}^{n} \frac{1}{n} \log_2 \frac{1}{n}$$
$$= \log_2 n \qquad \text{bits/interval}$$

Reference

[1] M. Schwartz, Information Transmission, Modulation, and Noise, 4/e, McGraw-Hill, 1990.

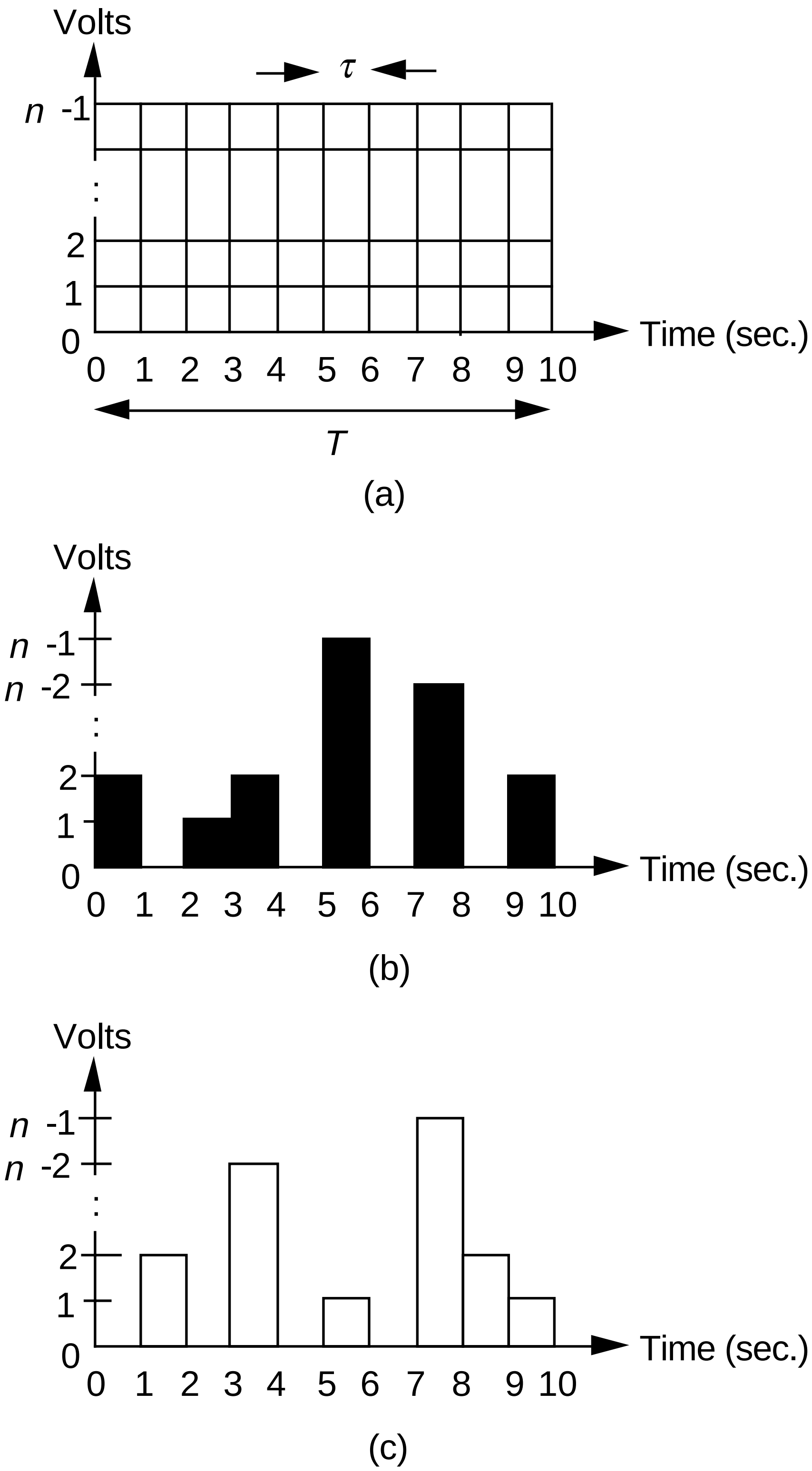

Figure 31.1 (a) Voltage-time diagram, (b) message 1, and (c) message 2.

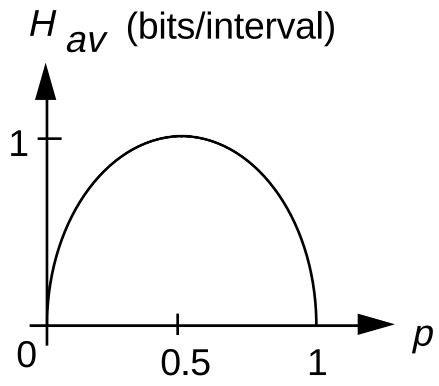

Figure 31.2 A plot of H_{av} against p.