

LECTURE 1: INTRODUCTION TO COMMUNICATION ENGINEERING

ECEN 310

Communication Engineering



VICTORIA UNIVERSITY OF
WELLINGTON
TE HERENGA WAKA

Lecture Outline



We will look at:

- How communication systems work

What is Communication Engineering (CE)?

- It is related to the production and transmission of information, through waveforms.
- The waveform at the receiver (user) is unknown until after it is received.
- Otherwise, no information would be transmitted and there is no need for CE.
- More information is communicated to the receiver when the user is “more surprised” by the message that was transmitted.
- That is, transmission of information implies that the messages are not known ahead of time (a priori).

What is Communication Engineering (CE)?

- Noise limits our ability to communicate.
- Communication systems are designed to transmit information bearing waveforms to the receiver. There are many possibilities for selecting waveforms to represent the information.
- For example, how does one select a waveform to represent the letter A in a typed message?
- Waveform selection depends on many factors. Some of these are bandwidth (frequency span) and center frequency of the waveform, waveform power or energy, the effect of noise on corrupting the information carried by the waveform, and the cost of generating the waveform at the transmitter and detecting the information at the receiver.

About this course (10 lectures)

- Lecture 1 introduces some key concepts, such as the definition of information, and provides a method for evaluating the information capacity of a communication system.
- Lecture 2 covers the basic techniques related to signals and systems.
- Amplitude modulation will be studied in Lecture 3, and Angle modulation in Chapters 4, followed by A2D in Lecture 5.
- Lectures 6 to 10 we will look at Bandpass systems and digital modulation techniques.

Essence of CE

In summary, communication systems are designed to transmit information.

Communication system designers have four main concerns:

1. Selection of the information-bearing waveform
2. Bandwidth and power of the waveform
3. Effect of system noise on the received information
4. Cost of the system.

Historical Perspective

- The telephone was developed late in the 19th century,
- The first transatlantic telephone cable was not completed until 1954. Previous to that date, transatlantic calls were handled via shortwave radio.
- Similarly, although the British began television broadcasting in 1936, transatlantic television relay was not possible until 1962, when the Telstar I satellite was placed into orbit.
- Digital transmission systems—embodied by telegraph systems—were developed in the 1850s before analog systems—the telephone—in the 20th century.
- Now, digital transmission is again becoming the preferred technique.

Digital and Analog Sources and Systems

DEFINITION. A digital information source produces a finite set of possible messages.

A telephone touchtone pad is a good example of a digital source. There is a finite number of characters (messages) that can be emitted by this source.

DEFINITION. An analog information source produces messages that are defined on a continuum.

A microphone is a good example of an analog source. The output voltage describes the information in the sound, and it is distributed over a continuous range of values.

Digital and Analog Sources and Systems

- ❑ DEFINITION. A digital communication system transfers information from a digital source to the intended receiver (also called the sink).
- ❑ DEFINITION. An analog communication system transfers information from an analog source to the sink.

Strictly speaking, a digital waveform is defined as a function of time that can have only a discrete set of amplitude values. If the digital waveform is a binary waveform, only two values are allowed. An analog waveform is a function of time that has a continuous range of values.

- An electronic digital communication system usually has voltage and current waveforms that have digital values; however, it may have analog waveforms. For example, the information from a binary source may be transmitted to the receiver by using a sine wave of 1,000 Hz to represent a binary 1 and a sine wave of 500 Hz to represent a binary 0. Here the digital source information is transmitted to the receiver by the use of analog waveforms, but the system is still called a digital communication system. From this viewpoint, we see that a digital communication engineer needs to know how to analyze analog circuits as well as digital circuits.

Why use Digital?

Digital communication has a number of advantages:

- Relatively inexpensive digital circuits may be used.
- Privacy is preserved by using data encryption.
- Greater dynamic range (the difference between the largest and smallest values) is possible.
- Data from voice, video, and data sources may be merged and transmitted over a common digital transmission system.
- In long-distance systems, noise does not accumulate from repeater to repeater.
- Errors in detected data may be small, even when there is a large amount of noise on the received signal.
- Errors may often be corrected by the use of coding.

The downside of Digital

Digital communication also has disadvantages:

- Generally, more bandwidth is required than that for analog systems.
- Synchronization is required.

The advantages of digital communication systems usually outweigh their disadvantages.

Consequently, digital systems are becoming dominant.

Deterministic and Random Waveforms

In communication systems, we are concerned with two broad classes of waveforms: deterministic and random (or stochastic).

□ DEFINITION. A deterministic waveform can be modeled as a completely specified function of time.

For example, if $w(t) = A \cos(\omega_0 t + \phi_0)$

describes a waveform, where A , ω_0 , and ϕ_0 are known constants, this waveform is said to be deterministic because, for any value of t , the value $w(t)$ can be evaluated. If any of the constants are unknown, then the value of $w(t)$ cannot be calculated, and consequently, $w(t)$ is not deterministic.

Deterministic and Random Waveforms

DEFINITION. A random waveform (or stochastic waveform) cannot be completely specified as a function of time and must be modeled probabilistically.

For example, in a digital communication system, we might send information corresponding to any one of the letters of the English alphabet.

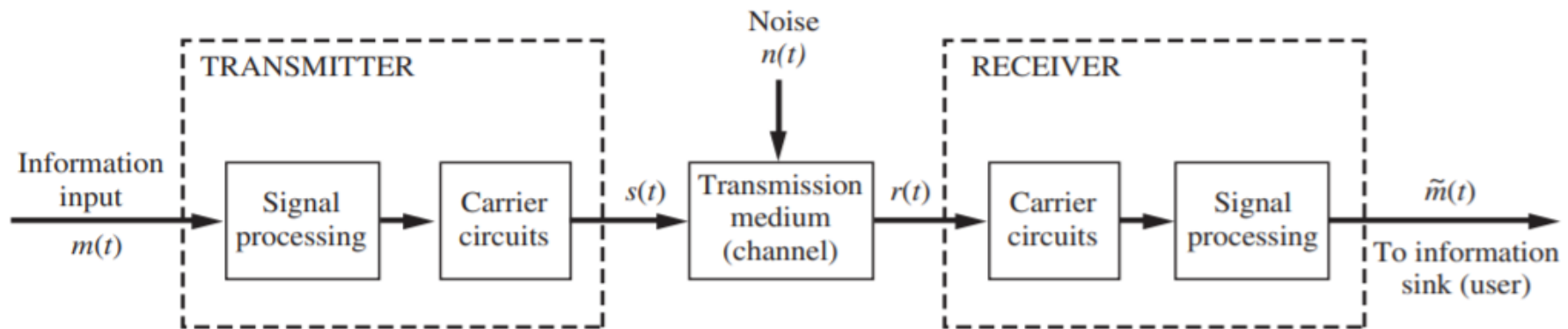
Each letter might be represented by a deterministic waveform, but when we examine the waveform that is emitted from the source, we find that it is a random waveform because we do not know exactly which characters will be transmitted.

Consequently, we really need to design the communication system by using a random signal waveform. Noise would also be described by a random waveform.

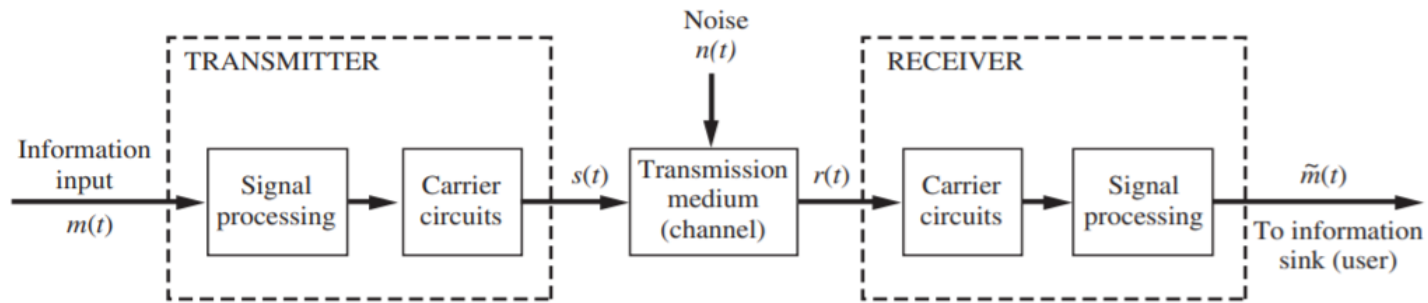
This requires the use of probability and statistical concepts (covered in Lectures 6 and 7) that make the design and analysis procedure more complicated.

However, if we represent the signal waveform by a “typical” deterministic waveform, we can obtain most, but not all, of the results we are seeking. That is the approach taken in the first five lectures of this book.

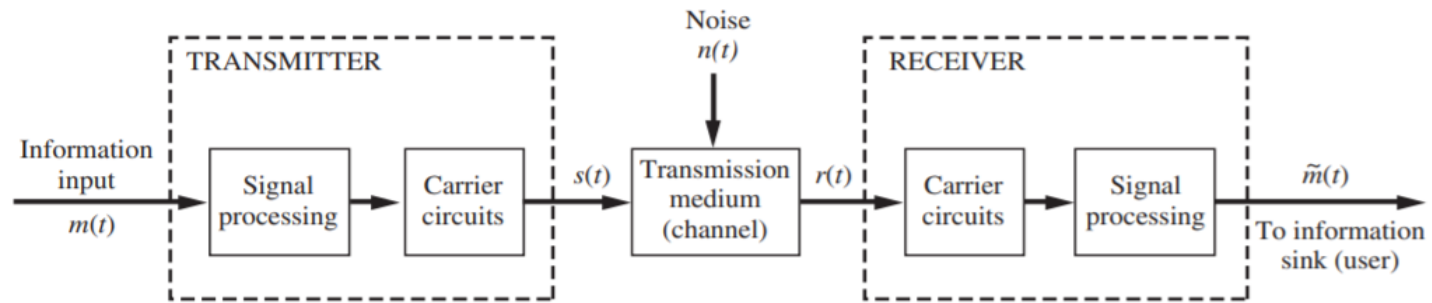
Block diagram of a communication system



- 3 main parts: transmitter, channel and receiver.
- The message from the source is represented by the information input waveform $m(t)$. The message delivered by the receiver is denoted by $\tilde{m}(t)$.
- The $[\sim]$ indicates that the message received may not be the same as that transmitted. That is, the message at the sink may be corrupted by noise in the channel, or there may be other impairments in the system, such as undesired filtering or undesired nonlinearities.
- The message information may be in analog or digital form, depending on the particular system, and it may represent audio, video, or some other type of information. In multiplexed systems, there may be multiple input and output message sources and sinks.
- The spectra (or frequencies) of $m(t)$ and $\tilde{m}(t)$ are concentrated about $f = 0$; consequently, they are said to be baseband signals.



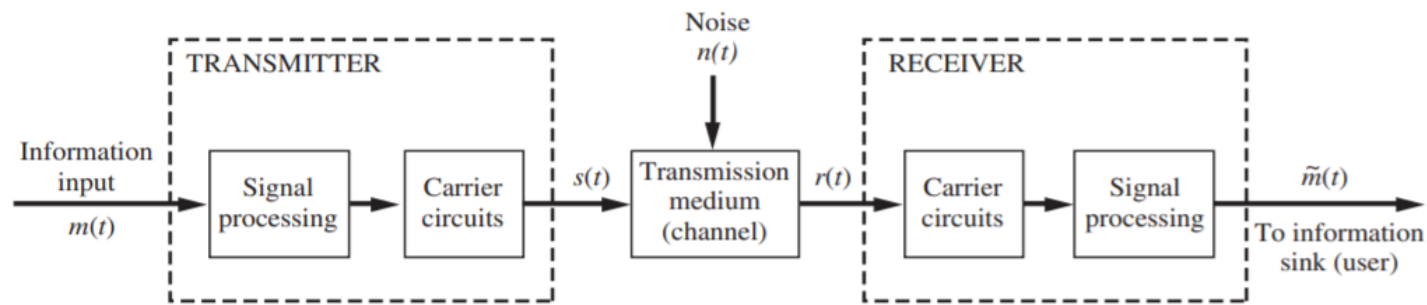
- The signal-processing block at the transmitter conditions the source for more efficient transmission. For example, in an analog system, the signal processor may be an analog low-pass filter that is used to restrict the bandwidth of $m(t)$.
- In a hybrid system, the signal processor may be an analog-to-digital converter (ADC), which produces a “digital word” that represents samples of the analog input signal (as described in Lecture 3 in the section on pulse code modulation).
- In this case, the ADC in the signal processor is providing source coding of the input signal.
- In addition, the signal processor may add parity bits to the digital word to provide channel coding so that error detection and correction can be used by the signal processor in the receiver to reduce or eliminate bit errors that are caused by noise in the channel.
- The signal at the output of the transmitter signal processor is a baseband signal, because it has frequencies concentrated near $f=0$.



The transmitter carrier circuit converts the processed baseband signal into a frequency band that is appropriate for the transmission medium of the channel.

For example, if the channel consists of a fiber-optic cable, the carrier circuits convert the baseband input (i.e., frequencies near $f=0$) to light frequencies, and the transmitted signal, $s(t)$, is light.

If the channel propagates baseband signals, no carrier circuits are needed, and $s(t)$ can be the output of the processing circuit at the transmitter.

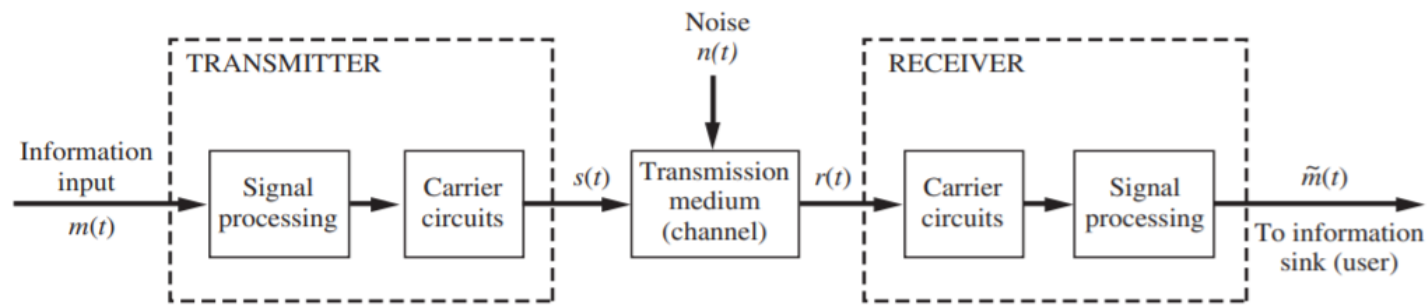


Carrier circuits are needed when the transmission channel is located in a band of frequencies around $f_c \gg 0$. (The subscript c denotes “carrier” frequency.)

In this case, $s(t)$ is said to be a bandpass, because it is designed to have frequencies located in a band about f_c .

For example, an amplitude-modulated (AM) broadcasting station with an assigned frequency of 850 kHz has a carrier frequency of $f_c = 850 \text{ kHz}$.

The mapping of the baseband input information waveform $m(t)$ into the bandpass signal $s(t)$ is called modulation. [$m(t)$ is the audio signal in AM broadcasting.]



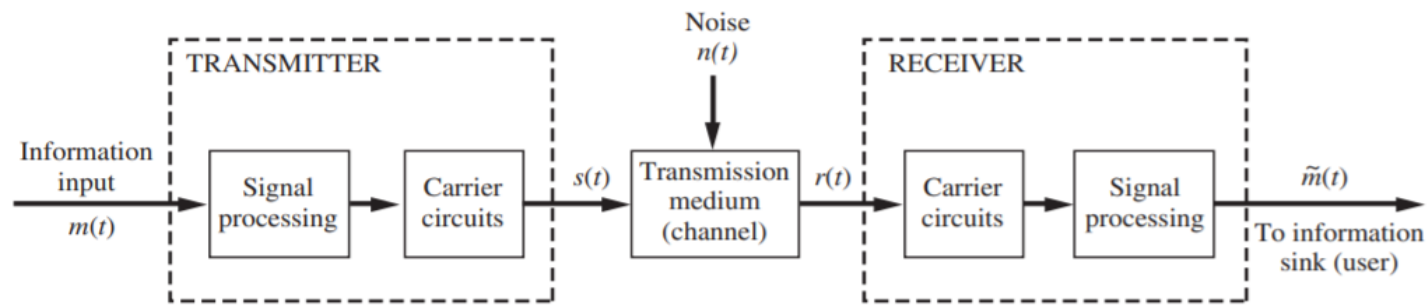
In Lecture 4, it will be shown that any bandpass signal has the form $s(t) = R(t) \cos [\omega_c t + \theta(t)]$ where $\omega_c = 2\pi f_c$.

If $R(t) = 1$ and $\theta(t) = 0$, $s(t)$ would be a pure sinusoid of frequency $f = f_c$ with zero bandwidth.

In the modulation process provided by the carrier circuits, the baseband input waveform $m(t)$ causes $R(t)$ or $\theta(t)$ or both to change as a function of $m(t)$.

These fluctuations in $R(t)$ and $\theta(t)$ cause $s(t)$ to have a nonzero bandwidth that depends on the characteristics of $m(t)$ and on the mapping functions used to generate $R(t)$ and $\theta(t)$.

In Lecture 5, practical examples of both digital and analog bandpass signaling are presented.



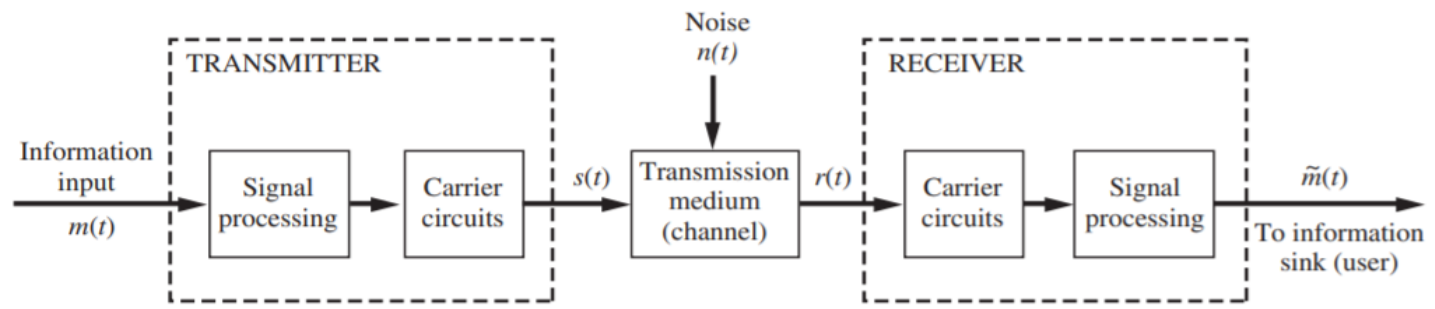
Channels may be classified into two categories: wire and wireless.

Some examples of wire channels are twisted-pair telephone lines, coaxial cables, waveguides, and fiber-optic cables.

Some typical wireless channels are air, vacuum, and seawater.

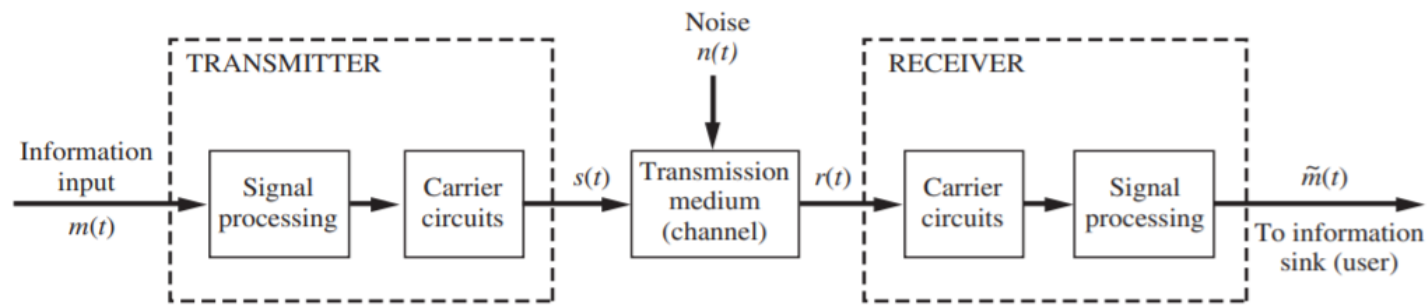
Note that the general principles of digital and analog modulation apply to all types of channels, although channel characteristics may impose constraints that favor a particular type of signaling.

In general, the channel medium attenuates the signal so that the noise of the channel or the noise introduced by an imperfect receiver causes the delivered information to be deteriorated from that of the source.



The channel noise may arise from natural electrical disturbances (e.g., lightning) or from artificial sources, such as high-voltage transmission lines, ignition systems of cars, or switching circuits of a nearby digital computer.

The channel may contain active amplifying devices, such as repeaters in telephone systems or satellite transponders in space communication systems.

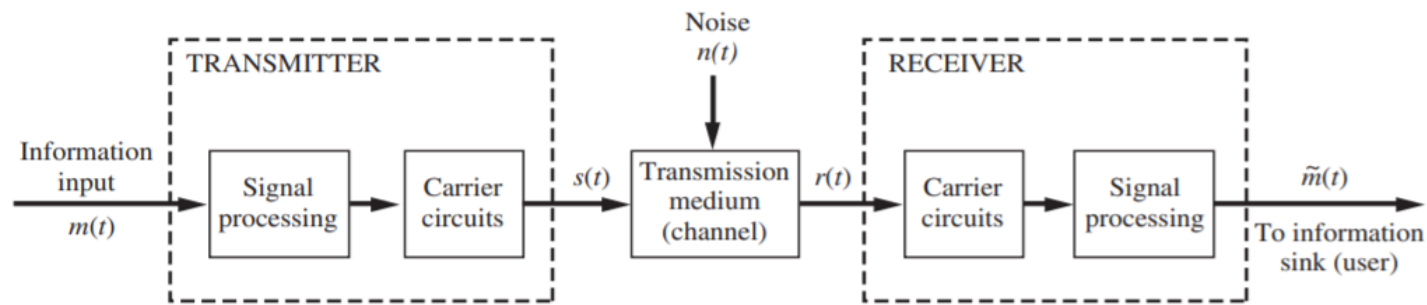


The channel may also provide multiple paths between its input and output. This can be caused by the signal bouncing off of multiple reflectors.

This multipath can be approximately described by two parameters—delay spread and Doppler spread. Delay spread is caused by multiple paths with varying lengths, which will cause a short transmitted pulse to be spread over time at the channel output because of the combination of received pulses with different delays from the multiple paths.

Different motions of the various multipath reflectors cause the received pulses to have different Doppler frequency shifts so that there is a spread of Doppler frequency shifts on the components of the combined received signal.

If the multipath reflectors move around slowly and, moreover, appear and disappear, the received signal will fade due to the individual received signals cancelling each other (when the composite received signal fades out).



The receiver takes the corrupted signal at the channel output and converts it to a baseband signal that can be handled by the receiver baseband processor.

The baseband processor “cleans up” this signal and delivers an estimate of the source information to the communication system output.

The goal is to design communication systems that transmit information to the receiver with as little deterioration as possible while satisfying design constraints, of allowable transmitted energy, allowable signal bandwidth, and cost.

In digital systems, the measure of deterioration is usually taken to be the probability of bit error (P_e)—also called the bit error rate (BER)—of the delivered data .

In analog systems, the performance measure is usually taken to be the signal-to-noise ratio at the receiver output.

Information Measure

What exactly is information, and how do we measure it? We know qualitatively that it is related to the surprise that is experienced when we receive the message. For example, the message “The ocean has been destroyed by a nuclear explosion” contains more information than the message “It is raining today.”

DEFINITION. The information sent from a digital source when the j th message is transmitted is given by

$$I_j = \log_2 \left(\frac{1}{P_j} \right) \text{ bits}$$

where P_j is the probability of transmitting the j th message.

Information Measure

From this definition, we see that messages that are less likely to occur (smaller value for P_j) provide more information (larger value of I_j).

We also observe that the information measure depends on only the likelihood of sending the message and does not depend on possible interpretation of the content as to whether or not it makes sense.

In general, the information content will vary from message to message because the P_j 's will not be equal. Consequently, we need an average information measure for the source, considering all the possible messages we can send.

DEFINITION. The average information measure of a digital source is

$$H = \sum_{j=1}^m P_j I_j = \sum_{j=1}^m P_j \log_2 \left(\frac{1}{P_j} \right) \text{ bits}$$

where m is the number of possible different source messages and P_j is the probability of sending the j th message (m is finite because a digital source is assumed). The average information is called **entropy**.

Example 1-3

Example 1-3 EVALUATION OF INFORMATION AND ENTROPY

Find the information content of a message that consists of a digital word 12 digits long in which each digit may take on one of four possible levels. The probability of sending any of the four levels is assumed to be equal, and the level in any digit does not depend on the values taken on by previous digits.

In a string of 12 symbols (digits), where each symbol consists of one of four levels, there are $4 \cdot 4 \dots 4 = 4^{12}$ different combinations (words) that can be obtained. Because each level is equally likely, all the different words are equally likely. Thus,

$$p_j = \frac{1}{4^{12}} = \left(\frac{1}{4}\right)^{12}$$

or

$$I_j = \log_2 \left(\frac{1}{\left(\frac{1}{4}\right)^{12}} \right) = 12 \log_2(4) = 24 \text{ bits}$$

Example 1-3

In this example, we see that the information content in every one of the possible messages equals 24 bits. Thus, the average information H is 24 bits.

Suppose that only two levels (binary) had been allowed for each digit and that all the words were equally likely. Then the information would be for the binary words, and the average information would be $I_j = 12$ bits.

Here, all the 12-bit words gave 12 bits of information, because the words were equally likely.

If they had not been equally likely, some of the 12-bit words would contain more than 12 bits of information and some would contain less, and the average information would have been less than 12 bits.

For example, if half of the 12-bit words (2,048 of the possible 4,096) have probability of occurrence of $P_i = 10^{-5}$ for each of these words (with a corresponding $I_j = 16.61$ bits) and the other half have $P_i = 4.78 \times 10^{-4}$ (for a corresponding $I_j = 11.03$ bits), then the average information is $H = 11.14$ bits.

Rate of Information

The rate of information is also important.

DEFINITION. The source rate is given by

$$R = \frac{H}{T} \text{ bits/s}$$

where H is evaluated by using $H = \sum_{j=1}^m P_j I_j = \sum_{j=1}^m P_j \log_2 \left(\frac{1}{P_j} \right) \text{ bits}$

and T is the time required to send a message.

The definitions given apply to digital sources.

Channel Capacity

Many criteria can be used to measure the effectiveness of a communication system to see if it is ideal or perfect.

For digital systems, the optimum system might be defined as the system that minimizes the probability of bit error at the system output subject to constraints on transmitted energy and channel bandwidth.

Thus, bit error and signal bandwidth are of prime importance and are covered in subsequent lectures.

This raises the following question: Is it possible to invent a system with no bit error at the output even when we have noise introduced into the channel?

Channel Capacity

This question was answered by Claude Shannon in 1948–1949. The answer is yes, under certain assumptions.

Shannon showed that (for the case of signal plus white Gaussian noise), a channel capacity C (bits/s) could be calculated such that if the rate of information R (bits/s) was less than C , the probability of error would approach zero. The equation for C is

$$C = B \log_2 \left(1 + \frac{S}{N} \right)$$

where B is the channel bandwidth in hertz (Hz) and S/N is the signal-to-noise power ratio (watts/watts, not dB) at the input to the digital receiver.

Shannon does not tell us how to build this system, but he proves that it is theoretically possible to have such a system. Thus, Shannon gives us a theoretical performance bound that we can strive to achieve with practical communication systems. Systems that approach this bound usually incorporate error-correction coding.

Channel Capacity

In analog systems, the optimum system might be defined as the one that achieves the largest signal-to-noise ratio at the receiver output, subject to design constraints such as channel bandwidth and transmitted power.

Here, the evaluation of the output signal-to-noise ratio is of prime importance.

We might ask the question, Is it possible to design a system with infinite signal-to-noise ratio at the output when noise is introduced by the channel?

The answer is no.

Other important theorems

Other fundamental limits for digital signaling were discovered by Nyquist in 1924 and Hartley in 1928.

Nyquist showed that if a pulse represents one bit of data, noninterfering pulses could be sent over a channel no faster than $2B$ pulses/s, where B is the channel bandwidth in hertz.

This is now known as the dimensionality theorem and is discussed in Lecture 2.

Hartley generalized Nyquist's result for the case of multilevel pulse signaling, as discussed in Lectures 3 and 5.

Channel Coding

- Also known as error correction coding.
- Channel Coding can be used to reduce bit errors, hence helping to approach Shannon's capacity.
- However, this is not covered in this course.
- If you are interested, please do a self-study.