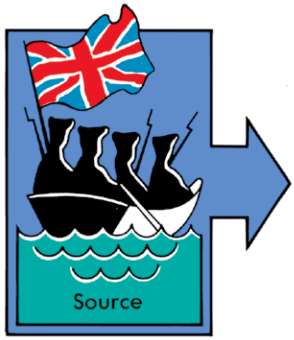
Information Theory

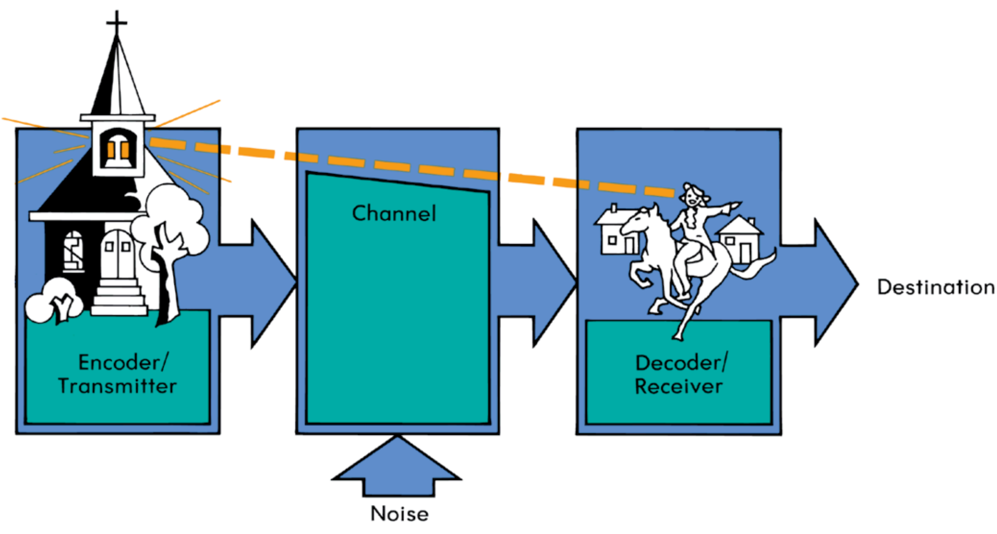
The era we are living in is sometimes called the age of information. But what is information, and how much of it is in any message? Let's look at two situations to determine their information content. Suppose you planned to play tennis with a friend at a nearby park but a heavy rain prevents you from leaving the house. Then the telephone rings and your friend tells you the game is off because it is raining. This message holds no information because you already know it is raining. Suppose, however, that you planned to play at a park by your friend's house several miles away, the sky was overcast, and the weatherman the night before had said that the chance of rain in the morning was high. Then your friend calls and says the game is off because it is raining there. This message contains information because prior to receiving it you were not certain whether it was raining at the park. Information, therefore, is anything that resolves uncertainty.

Information theory measures information. It also investigates the efficient use of information media. A system of information communication is composed of several components. First, the information source produces the message, or information, to be transmitted. Next, the encoder, or transmitter, processes the message and changes it to a signal suitable for the channel. (The channel is the medium over which the signal is sent. Random disturbances, or noise, on the channel cause the received signal to be somewhat different from the transmitted one. Noise imposes serious limitations on the amount of information that can be transmitted over a channel.) Finally, the decoder, or receiver, processes the received signal to recreate the original message.



Block Diagram
of a Communication System

The block diagram illustrates these components by using the key events of Paul Revere's famed midnight ride in 1775 to warn colonial patriots northwest of Boston of an impending British raid. The information source was the British army, and the source output was the information whether the British troops would attack the Massachusetts towns of Lexington and Concord by land or by sea. The encoder/transmitter was the sexton of the North Church who was to place one lantern in the belfry to indicate attack by land or two lanterns for attack by sea. The channel was the line of sight from the belfry to the place where Paul Revere, the decoder/receiver, was waiting. Anything that obstructed Revere's view, such as wind-tossed tree branches, created noise. The communication was only a selection between two alternatives—attack by land or attack by sea. As long as prior agreement existed between



A block diagram depicts the concept of a communication system in information theory.

encoder and decoder (the sexton and Revere) on the alternatives and their proper signals, this communication system could have sent a selection from any set of two alternatives. For example, were they infantry or artillery troops? A more complex system would have been needed to communicate the number of troops in the raid because of the larger number of possible alternatives.

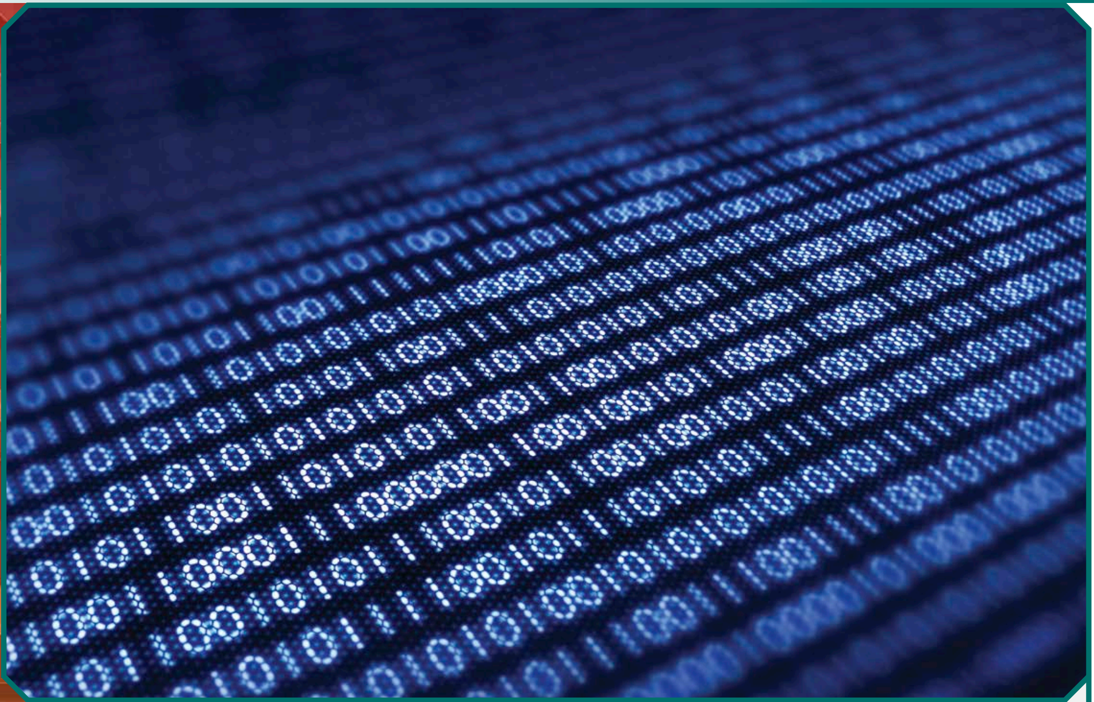
Binary Digits and Information Transmission

Teletype communication offers a good example of information transmission. In teletype communication the source is the person operating a teletypewriter, an instrument that transmits and receives messages via telephone cables or radio relay systems. The teletypewriter, in turn, is the encoder. The machine is connected by telephone wire (the channel) to a teleprinter (the decoder) in a distant city. The person sends a message by pressing down on the teletypewriter keys. Those keys are a set of alternatives. The pressing of each key triggers transmission of a sequence of binary digits that represent the letter on that key. A binary digit is either a 0 or a 1. From the teletypewriter

Information transmission can be represented by teletype communication. A person at the keyboard is the source; the machine is the transmitter, receiver, and encoder; and the telephone wire is the channel to the teleprinter, or the decoder.

pulsating information encoded in binary digits is sent over commercial telephone lines. At the other end, the sequence of binary digits is decoded back to the appropriate letters. Binary digits underlie the language of computers.

For information theory purposes, the length of the binary sequence needed for encoding individual letters is more important than how the teletypewriter actually encodes the information. Only two binary digits are needed to represent four alternatives—00, 01, 10, and 11. Similarly, three binary digits can represent



This photograph shows an example of binary code. A binary code is a code used in computers, based on a binary number system, and involves a method of calculating and representing information by using the numbers 0 and 1.

eight alternatives—000, 001, 010, 011, 100, 101, 110, and 111. As a general principle, each increase of one in the sequence length doubles the number of alternatives (2 binary digits—4 alternatives, 3—8, 4—16, and so on). In mathematical shorthand, we can say a sequence of *n* binary digits can represent 2^n alternatives.

Logarithms, especially base-two logarithms ($\log_2$), are used in the computation of information content. The base-two logarithm of a number is defined as the power to which two must be raised in order to get that number; hence, for any number n, $\log_2(2^n) = n$. As a result, if a teletypewriter has m keys on its keyboard, any individual key must activate a sequence of binary digits that is at least $\log_2 m$ long. For example, if a keyboard holds 32 keys, a sequence of five binary digits can represent any given key (00000 for A, 00001 for B, and so on) because the base-two logarithm of 32 is 5.00000.

Human communication is still another example of information transmission. A person with an idea is the source, and the idea is the source output. The idea is encoded into words and written. Another person decodes the words when reading them and tries to reconstruct the writer's idea. Each written word was selected from a set of alternatives—that is, all the words of the language. Information theory has not shed much light on biological communication because living communication systems are not well enough understood for the application of the mathematical concepts of the theory. However, the basic notion of information as a transmission of binary digits that represents a choice from a set of alternatives should prove helpful in future research into biological information systems.

Claude Shannon (1916–2001)

American engineer Claude Shannon's master's thesis, "A Symbolic Analysis of Relay and Switching Circuits" (1940), established the theoretical underpinnings of digital circuits. Digital circuits are fundamental to the operation of modern computers and telecommunications equipment, making Shannon's dissertation one of the most significant theses of the 20th century.

Shannon's 1948 paper "A Mathematical Theory of Communication" built on the foundations of other researchers at Bell Labs, such as Harry Nyquist and R.V.L. Hartley. Shannon's paper, however, went far beyond the earlier work. It established the basic results of information theory in such a complete form that his framework and terminology are still used.

Shannon separated the technical problem of delivering a message from the problem of understanding what a message means. He focused on two key concerns: determining the most efficient encoding of a message using a given alphabet in a noiseless environment, and understanding what additional steps need to be taken in the presence of noise.

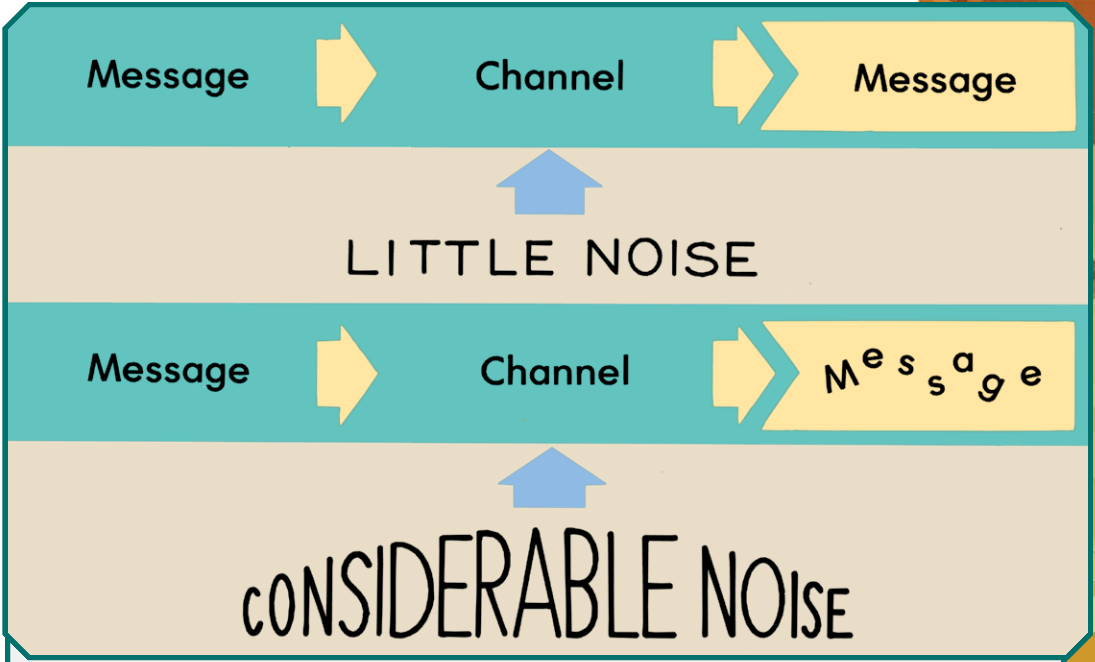
Shannon solved these problems successfully for a widely applicable model of a communications system that includes both discrete (digital) and continuous (analog) systems. His contributions to information theory were an immediate success with communications engineers and continue to prove useful. Shannon's work inspired many attempts to apply information theory in other areas, such as biology, linguistics, psychology, economics, and physics.

Renowned for his eclectic interests and capabilities—including such activities as juggling while riding a unicycle down the halls of Bell Labs—Shannon produced

many provocative and influential articles on information theory, cryptography, and chess-playing computers, as well as designing various mechanical devices.

Channels and Noise

Whether the channel of an information system is a wire, a radio wavelength, a laser beam, or some other medium, the output signal of the system is never quite the same as the



When great enough, random disturbances called noise can garble or distort a message being transmitted over a channel.

input signal because of noise intruding on the channel. Noise can be caused by heat, atmospheric disturbances, or cross talk, which is a form of interference from other channels. Static in a radio, snow on a television screen, and background hum in a telephone are familiar examples of noise. To send information over radio, television, and telephone channels, a modem system (digital data *mo*dulator and *dem*odulator) must be used. At one end of the channel the modulator converts the sequence of symbols making up the information into a signal suitable for transmission over the channel. At the other end the demodulator changes the received signal back into the original sequence of symbols. Noise, however, interferes by causing occasional errors at the receiving end of the channel.