Chapter 13
Claude Shannon

Key Topics

Boolean Algebra and Switching Circuits
Information Theory
Cryptography

13.1 Introduction

Claude Shannon was an American mathematician and engineer who made fundamental contributions to computing. He was born in Michigan in 1916, and his primary degree was in mathematics and electrical engineering at the University of Michigan in 1936. He was awarded a PhD in mathematics from the Massachusetts Institute of Technology (MIT) in 1940.

He was the first person[1] to see the applicability of Boolean algebra to simplify the design of circuits and telephone routing switches. He showed that Boole's symbolic logic developed in the nineteenth century provided the perfect mathematical model for switching theory and for the subsequent design of digital circuits and computers.

His influential master's thesis is a key milestone in computing, and it shows how to lay out circuits according to Boolean principles. It provides the theoretical foundation of switching circuits, and his insight of using the properties of electrical switches to do Boolean logic is the basic concept that underlies all electronic digital computers.

[1] Victor Shestakov at Moscow State University also proposed a theory of electric switches based on Boolean algebra around the same time as Shannon. However, his results were published in Russian in 1941, whereas Shannon's were published in 1937.

G. O'Regan, *A Brief History of Computing*, DOI 10.1007/978-1-4471-2359-0_13,

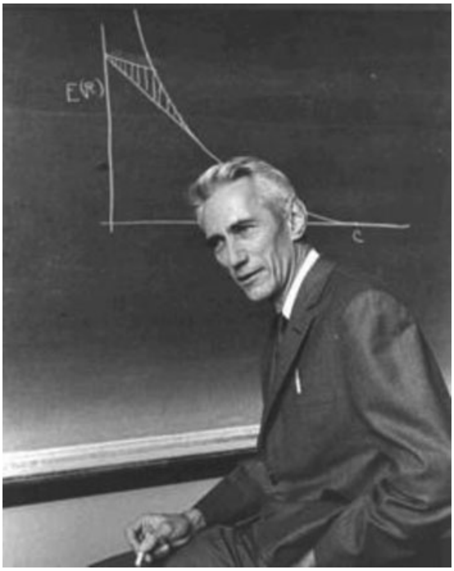


Fig. 13.1 Claude Shannon

Shannon realised that you could combine switches in circuits in such a manner as to carry out symbolic logic operations. The implications of this were enormous, as it allowed binary arithmetic and more complex mathematical operations to be performed by relay circuits. He showed the design of a circuit which could add binary numbers; he moved on to realising circuits which could make comparisons and thus is capable of performing a conditional statement. This was the birth of digital logic (Fig. 13.1).

He moved to the mathematics department at Bell Labs in the 1940s and commenced work that would lead to the foundation of modern information theory. This is concerned with defining a quantitative measure of information and using this to solve related problems. It includes the well-known 'channel capacity theorem' and is also concerned with the problem of the reliable transfer of messages from a source point to a destination point. The messages may be in any communications medium, for example, television, radio, telephone and computers. The fundamental problem is to reproduce at a destination point either exactly or approximately the message that has been sent from a source point. The problem is that information may be distorted by noise, leading to differences between the received message and that which was originally sent.

Shannon provided a mathematical definition and framework for information theory in '*A Mathematical Theory of Communication*' [Sha:48]. He proposed the idea of converting data (e.g. pictures, sounds or text) to binary digits, that is, binary bits of information. The information is then transmitted over the communication medium. Errors or noise may be introduced during the transmission, and the objective is to reduce and correct them. The received binary information is then converted back to the appropriate medium.

Shannon is considered the father of digital communication, and his theory was an immediate success with communications engineers with wide reaching impacts.

He also contributed to the field of cryptography in '*Communication Theory of Secrecy Systems*' [Sha:49].

He also made contributions to genetics and invented a chess playing computer program in 1948. He was a talented gadgeteer who built some early robot automata, game playing devices and problem-solving machines. He was able to juggle while riding a unicycle, and he designed machines to juggle and to ride unicycle like machines. He received many honours and awards and died in 2001.

13.2 Boolean Algebra and Switching Circuits

Vannevar Bush[2] was Shannon's supervisor at MIT, and Shannon's initial work was to improve Bush's mechanical computing device known as the differential analyser.

The differential analyzer was an analog computer made of gears, pulleys and rods. Its function was to evaluate and solve first-order differential equations. The machine was developed by Vannevar Bush and others in 1931 and funded by the Rockefeller Foundation. It weighed over a 100 tons, had 2,000 vacuum tubes, several thousand relays and automated punch-tape access units.

It had a complicated control circuit that was composed of 100 switches that could be automatically opened and closed by an electromagnet. Shannon's insight was the realisation that an electronic circuit is similar to Boolean algebra. He showed how Boolean algebra could be employed to optimise the design of systems of electromechanical relays used in Bush's analog computer, and that circuits with relays could solve Boolean algebra problems.

He showed in his master's thesis '*A Symbolic Analysis of Relay and Switching Circuits*' [Sha:37] that the binary digits (i.e. 0 and 1) can be represented by electrical switches. The convention employed in Shannon's thesis is the opposite to the convention used today. In his thesis, the digit 1 is represented by a switch that is turned off and the symbol 0 by a switch that is turned on. The *convention today is that the digit* 1 *is represented by a switch that is turned on, whereas the digit* 0 *is represented by a switch that is turned off*. We shall follow the convention in Shannon's thesis in this chapter for historical reasons.

He used Boolean algebra to show that complex operations could be performed automatically on these electrical circuits. The implications of true and false being denoted by the binary digits 1 and 0 were enormous since it allowed binary arithmetic and more complex mathematical operations to be performed by relay circuits. This provided electronics engineers with the mathematical tool they needed to design digital electronic circuits, and these techniques are the foundation of digital electronic design.

The design of circuits and telephone routing switches could be simplified with Boolean algebra. Shannon showed how to lay out circuitry according to Boolean

[2] Vannevar Bush was discussed in an earlier chapter.

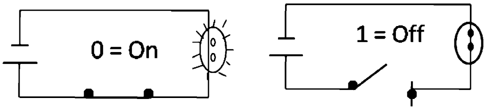


Fig. 13.2 Switching circuit representing Boolean logic

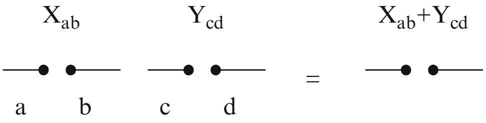


Fig. 13.3 Serial circuits

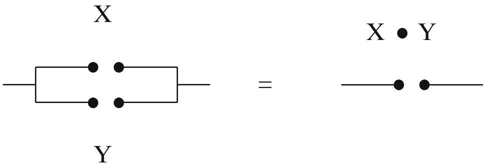


Fig. 13.4 Parallel circuits

principles, and his influential master's thesis became the foundation for the practical design of digital circuits. These circuits are fundamental to the operation of modern computers and telecommunication systems, and Shannon's master's thesis is considered a key milestone in the development of modern computers.

His insight of using the properties of electrical switches to do Boolean logic is the basic concept that underlies all electronic digital computers (Fig. 13.2).

A circuit may be represented by a set of equations with the terms in the equations representing the various switches and relays in the circuit. He developed a calculus for manipulating the equations, and this calculus is similar to Boole's propositional logic. The design of a circuit consists of algebraic equations, and the equations may be manipulated to yield the simplest circuit. The circuit may then be immediately drawn.

A switching circuit X_{ab} between two terminals *a* and *b* is either open (with $X_{ab} = 1$) or closed (with $X_{ab} = 0$). The symbol 0 is employed to denote a closed circuit, and the symbol 1 is used to denote an open circuit.

The expression $X_{ab} + Y_{cd}$ (usually written as X + Y) denotes the circuit formed when the circuit X_{ab} is joined serially to the circuit Y_{cd}. Similarly, the expression X_{ab}•Y_{cd} (written as XY) denotes the circuit formed when the circuit X_{ab} is placed in parallel to the circuit Y_{cd}. The operators + and • are, of course, different from the standard addition and multiplication operators (Figs. 13.3 and 13.4).

13.2.1 Properties of Circuits

It is evident from the above definitions that the laws found in Table 13.1 are true.

The circuit is either open or closed at a given moment of time, that is, $X = 0$ or $X = 1$. The algebraic properties found in Table 13.2 hold:

The negation of a circuit X is denoted by X′. X′ is open when X is closed and vice versa. Its properties are shown in Table 13.3.

Table 13.1 Properties of circuits

Property name	Property	Interpretation
Idempotent property	$0 \bullet 0 = 0$	A closed circuit in parallel with a closed circuit is a closed circuit
	$1 + 1 = 1$	An open circuit in series with an open circuit is an open circuit
Additive identity (0)	$1 + 0 = 0 + 1 = 1$	An open circuit in series with a closed circuit (in either order) is an open circuit
Multiplicative identity (1)	$1 \bullet 0 = 0 \bullet 1 = 0$	A closed circuit in parallel with an open circuit (in either order) is a closed circuit
Additive identity (0)	$0 + 0 = 0$	A closed circuit in series with a closed circuit is a closed circuit
Multiplicative identity (1)	$1 \bullet 1 = 1$	An open circuit in parallel with an open circuit is an open circuit

Table 13.2 Properties of circuits (ctd.)

Property name	Property
Commutative property	$x + y = y + x$
	$x \bullet y = y \bullet x$
Associative property	$x + (y + z) = (x + y) + z$
	$x \bullet (y \bullet z) = (x \bullet y) \bullet z$
Distributive property	$x \bullet (y + z) = (x \bullet y) + (x \bullet y)$
	$x + (y \bullet z) = (x + y) \bullet (x + z)$
Identity property	$x + 0 = x = 0 + x$
	$1 \bullet x = x \bullet 1 = x$
	$1 + x = x + 1 = 1$
	$0 \bullet x = x \bullet 0 = 0$

Table 13.3 Properties of circuits (ctd.)

Property name	Property
Negation	$X + X' = 1$
	$X \bullet X' = 0$
De Morgan's law	$(X')' = X$
	$(X + Y)' = X' \bullet Y'$
	$(X \bullet Y)' = X' + Y'$

Functions of variables may also be defined in the calculus in terms of the '+', '•' and negation operations. For example, the function $f(X,Y,Z) = XY' + X'Z' + XZ'$ is an example of a function in the calculus, and it describes a circuit.

13.2.2 Digital Circuits and Boolean Logic

The calculus for circuits is analogous to Boole's symbolic logic. The symbol X denotes a circuit in digital circuits and a proposition in Boolean logic;

Fig. 13.5 Sample circuit

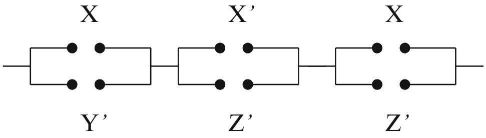
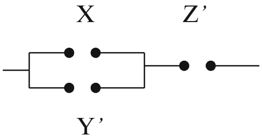


Fig. 13.6 Simplified circuit



0 denotes a closed circuit and the proposition false; 1 denotes an open circuit and the proposition true. The expression X + Y denotes serial connection of two circuits and denotes disjunction in Boolean logic; the expression XY denotes the parallel connection of two circuits and denotes logical conjunction in Boolean algebra; the expression X' denotes the complement of the circuit and the complement in Boolean algebra. De Morgan's laws hold for circuits and propositional logic.

Any expression formed with the additive, multiplicative and negation operators forms a circuit containing serial and parallel connections only. To find the simplest circuit with the least number of contacts, all that is required is to manipulate the mathematical expression into the form in which the fewest variables appear. For example, the circuit represented by $f(X,Y,Z) = XY' + X'Z' + XZ'$ is given by Fig. 13.5:

However, this circuit may be simplified by noting that (Fig. 13.6):

$$\begin{aligned} f(X,Y,Z) &= XY' + X'Z' + XZ' \\ &= XY' + (X' + X)Z' \\ &= XY' + 1Z' \\ &= XY' + Z'. \end{aligned}$$

13.2.3 Implementation of Boolean Logic

Digital circuits may be employed to implement Boolean algebra with the Boolean value 0 represented by a closed circuit and the Boolean value 1 represented by an open circuit. The logical conjunction and disjunction operations may be represented by combining circuits in parallel and series.

Complex Boolean value functions can be constructed by combining these digital circuits. Next, we discuss Shannon's contribution to information theory.

13.3 Information Theory

Early work on information theory was done by Nyquist and other engineers at Bell Labs in the 1920s. Nyquist investigated the speed of transmission of information over a telegraph wire [Nyq:24] and proposed a logarithmic rule ($W = k \log m$)[3] that set an upper limit on the amount of information that may be transmitted. This rule is a special case of Shannon's later logarithmic law.

There were several communication systems in use prior to Shannon's 1948 paper. These included the telegraph machine which dated from the 1830s, the telephone dating from the 1870s, the AM radio from the early 1900s and early television from the 1930s. These were all designed for different purposes and used various media. Each of these was a separate field with its own unique problems, tools and methodologies.

Shannon is recognised as the father of information theory due to his classic 1948 paper [Sha:48] which provided a unified theory for communication and a mathematical foundation for the field. The key problem in the field is the reliable transmission of a message from a source point over a communications channel to a destination point.[4] There may be noise in the channel that distorts the message, and the engineer wishes to ensure that the message received is that which has been sent. Information theory provides answers as to how rapidly or reliably a message may be sent from the source point to the destination point. The meanings of the messages are ignored as they are irrelevant from an engineering viewpoint. Shannon identified five key parts of an information system, namely (Fig. 13.7):

- Information source
- Transmitter
- Channel
- Receiver
- Destination

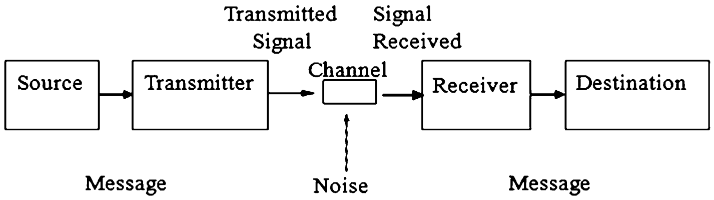


Fig. 13.7 Information theory

[3] W stands for the speed of transmission of information; m is the number of voltage levels to choose from at each step; and k is a constant.

[4] The system designer may also place a device called an encoder between the source and the channel and a device called a decoder between the output of the channel and the destination.

He derived formulae for the information rate of a source and for the capacity of a channel including noiseless and noisy cases. These were measured in bits per second, and he showed that for any information rate R less than the channel capacity C,[5] it is possible (by suitable encoding) to send information at rate R, with an error rate less than any preassigned positive ε, over that channel.

Shannon's theory of information is based on probability theory and statistics. One important concept is that of *entropy*[6] which measures the level of uncertainty in predicting the value of a random variable. For example, the toss of a fair coin has maximum entropy as there is no way to predict what will come next. A single toss of a fair coin has entropy of one bit.

The concept of entropy is used by Shannon as a measure of the average information content missing when the value of the random variable is not known. English text has fairly low entropy as it is reasonably predictable because the distribution of letters is far from uniform.

Shannon proposed two important theorems that establish the fundamental limits on communication. The first theorem deals with communication over a noiseless channel, and the second theorem deals with communication in a noisy environment.

The first theorem (Shannon's source coding theorem) essentially states that the transmission speed of information is based on its entropy or randomness. It is possible to code the information (based on the statistical characteristics of the information source) and to transmit it at the maximum rate that the channel allows. Shannon's proof showed that an encoding scheme exists, but it did not show how to construct one. This result was revolutionary as communication engineers at the time thought that the maximum transmission speed across a channel was related to other factors and not on the concept of information.

Shannon's *noisy-channel coding theorem* states that reliable communication is possible over noisy channels provided that the rate of communication is below a certain threshold called the 'channel capacity'. This result was revolutionary as it showed that a transmission speed arbitrarily close to the channel capacity can be achieved with an arbitrarily low error. The assumption at the time was that the error rate could only be reduced by reducing the noise level in the channel. Shannon showed that this assumption was not quite true, and he showed that a transmission speed arbitrarily close to the channel capacity can be achieved by using appropriate encoding and decoding systems.

Shannon's theory also showed how to design more efficient communication and storage systems.

[5] The channel capacity C is the limiting information rate (i.e. the least upper bound) that can be achieved with an arbitrarily small error probability. It is measured in bits per second.

[6] The concept of entropy is borrowed from the field of thermodynamics.

13.4 Cryptography

Shannon is considered the father of modern cryptography with his influential 1949 paper ‘Communication Theory of Secrecy Systems’ [Sha:49]. He established a theoretical basis for cryptography and for cryptanalysis and defined the basic mathematical structures that underlie secrecy systems.

A secrecy system is defined to be a transformation from the space of all messages to the space of all cryptograms. Each possible transformation corresponds to encryption with a particular key, and the transformations are reversible. The inverse transformation allows the original message to be obtained from the cryptogram provided that the key is known. The basic operation of a secrecy system is described in Fig. 13.8.

The first step is to select the key and to send it securely to the intended recipient. The choice of key determines the particular transformation to be used for the encryption of the message. The transformation converts the message into a cryptogram (i.e. the encrypted text). The cryptogram is then transmitted over a channel that is not necessarily secure to the receiver, and the recipient uses the key to apply the inverse transformation to decipher the cryptogram into the original message.

The enciphering of a message is a functional operation. Suppose M is a message, K is the key and E is the encrypted message, then:

$$E = f(M, K)$$

This is often written as a function of one variable $E = T_i\, M$ where the index i corresponds to the particular key being used. It is assumed that there are a finite number of keys $K_1, \ldots K_m$ and a corresponding set of transformations $T_1, T_2, \ldots, T_m$. Each key has a probability p_i of being chosen as the key. The encryption of a message M with key K_i is therefore given by:

$$E = T_i\, M$$

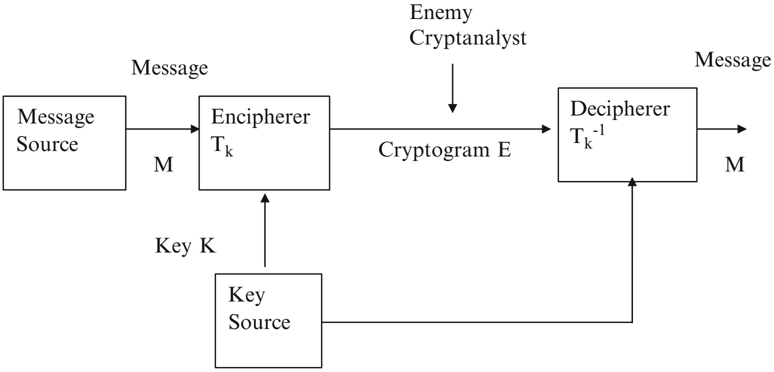


Fig. 13.8 Cryptography

It is then possible to retrieve the original message from the received encrypted message by:

$$M = T_i^{-1} E.$$

The channel may be intercepted by an enemy who will attempt to break the encrypted text. The enemy will examine the cryptogram and attempt to guess the key and the message from the cryptogram.

Shannon also showed that Vernam's cipher (also known as the *onetime pad*) is an unbreakable cipher. He showed that perfect secrecy of the cipher requires a secret random key with length greater than or equal to the number of information bits to be encrypted in the message. The secret key must be used once only, and the key needs to be sent securely to the receiver. This cipher was invented by Gilbert Vernam at Bell Labs.

13.5 Review Questions

1. Explain how Boolean algebra is applied to switching circuits.
2. Describe information theory and explain the importance of Shannon's fundamental theorems on communication.
3. Explain encryption and decryption in cryptology.
4. Describe the Vernam cipher.

13.6 Summary

Claude Shannon was an American mathematician and engineer who made fundamental contributions to computing. He was the first person to see the applicability of Boolean algebra to simplify the design of circuits and telephone routing switches. He showed how to lay out circuits according to Boolean principles, and his insight of using the properties of electrical switches to do Boolean logic is the basic concept that underlies all electronic digital computers.

He laid the foundation of modern information theory which is concerned with the problem of reliable transfer of messages from a source point to a destination point. This includes the transfer of messages over any communications medium, for example, television, radio, telephone and computers. His influential 1948 paper provided a mathematical foundation and framework for information theory that remains the standard today. He proposed two key theorems that establish the fundamental limits on communication in a noiseless and in a noisy channel.

He also made important contributions to the emerging field of cryptography, and his 1949 paper provided a theoretical foundation for cryptography and cryptanalysis.