Claude Shannon

Mathematical Theory of Claude Shannon

A study of the style and context of his work up to the genesis of information theory.

by

Eugene Chiu, Jocelyn Lin, Brok Mcferron,
Noshirwan Petigara, Satwiksai Seshasai

6.933J / STS.420J The Structure of Engineering Revolutions

Table of Contents

Acknowledgements

We would like to thank professors Robert Fano, Bob Gallagher, Chris Kaiser, Charles Vest and Hartley Rogers for their academic and personal insight into Claude Shannon's work and the fields of switching theory, genetics and information theory. Shannon's advisees Trenchard More, William Sutherland and Henry Ernst provided us with a unique perspective on working with Dr. Shannon. Betty Shannon graciously discussed her husband's personality and work. MIT staff Be Blackburn and Laura Mersky helped facilitate our research. The staff of the MIT Archives and Library of Congress were instrumental in locating important documents and correspondences between Shannon and his colleagues. Professor David Mindell and Teaching Assistant Chen-Pang Yeang went beyond their roles as instructors and provided us with insight into their own research on this topic.

Finally, we would also like to thank Peter Elias, who before passing away on Dec. 7, 2001 allowed us to search through his extensive archive of books and documents related to Claude Shannon and his fields. This paper is dedicated to his memory and the memory of Claude Shannon, who passed away on February 24, 2001.

Introduction

When Trenchard More first met Claude Elwood Shannon, he was taking his oral exam for his Ph. D at the Massachusetts Institute of Technology[1]. The written exam was especially difficult at the time, recalled More, so doing well on the oral portion of the exam was vital. Shannon had agreed to be on More's committee because More was a TA under Sam Caldwell, an advisor for Shannon's own Master's thesis. The questions Shannon asked were drastically different from the rest and concentrated on the mathematical ideas behind the topics being discussed. Shannon was "after the ways of thought," said More. He cared more about how More was thinking and whether he understood the fundamental mathematical concepts behind his research. Shannon felt that someone who really understood ideas could recreate them before your eyes, related More. Fortunately, despite having stumbled through the technical details in the written and oral exam, More had developed a solid understanding of the mathematical concepts and passed the oral exams by successfully answering Shannon's questions. He remembered another student meeting a different fate with Shannon - despite his perfect GPA, the student could not answer Shannon's questions, and it was revealed that he was simply memorizing concepts. The focus Shannon displayed in his questioning of More was representative of the guiding vision which drove much of his work.

The most popular and revolutionary pieces of Shannon's work came very early in his life. Many experts including MIT professor and colleague Robert Fano suggested that his two most important scientific contributions were his Master's thesis and his revolutionary 1948 paper on communication theory[2]. His Master's thesis, developing a method for using Boolean logic to represent circuits, and his 1948 paper on communication theory each helped to define a field and revolutionized the way in

---

[1] Interview with Trenchard More, 2001.
[2] Interview with Robert Fano, 2001.

which we viewed our world. At the heart of these two feats and much of his other work was the idea that mathematical concepts could be used to bring structure and understanding to almost anything.

This paper is an attempt to understand the context of Shannon's early work and examine the elements that contributed to his research and the subsequent explosion of study in his fields. Our goal is to not only understand the research, but also the man who gave birth to it. By exploring the common threads between the many diverse areas of Shannon's research, the fundamental ideas that drove Shannon become evident.

Much of the existing research on Shannon focuses on the work that had the greatest impact on the scientific community. As the creator of the field of information theory, he has been widely hailed as the "father of the digital age."[3] His 1948 paper detailed a method for representing information and forever revolutionized the way in which information was conceptualized. This paper begins by attempting to understand the context in which Shannon produced this work on information theory in the late 1940s.

After understanding Shannon's role in this field, we return to the earlier work and examine a few of his major accomplishments before his landmark paper. Although his style was very independent and his results were very unique, the areas in which he worked were very heavily influenced by external factors. Each of the areas of Shannon's work we consider - switching, genetics, information theory and cryptography - were motivated by the environment which Shannon was in at the time, and the specific interests of those advising him. He rarely displayed a sincere interest in promoting the

---

[3] Waldrop, 2001.

application and understanding of his work. Instead, he relied on the external factors to drive the infusion of his work into the appropriate field.

This paper pays careful attention to Shannon's Ph.D. thesis in genetics, which provides an interesting example of the role of external influence in Shannon's work. The genetics thesis did not receive nearly as much attention as his other work, but his contributions in the thesis are similar in scope and style to his work in switching theory and information theory. Again, the domain was provided by his environment, but his role was quite independent and dealt with using mathematical theory to represent the system. As with his other work, Shannon did little to promote the widespread awareness or acceptance of his results. However, unlike his work in switching and information, his genetics work was not embraced by the community, and thus did not have as great an impact as the other pieces.

Examining Shannon's work style and interests further confirmed this notion of an individual focused solely on the abstraction of a problem to its simplest form in any given field. As a student and colleague, Shannon was described as very shy and independent, yet incredibly bright. As he moved on to practice research professionally, both in industry and academia, his wife, advisees and colleagues all described a man who avoided collaboration, focused on the mathematical theory, and moved from subject to subject once being satisfied with having conquered the theoretical underpinnings of his topic of study. The professor who helped Trenchard More pass his Ph.D. exams by emphasizing the underlying mathematical theory employed this devotion in almost every aspect of his life.

The essence of Shannon's contributions was his style of work - his ability to take a problem and apply mathematical theory to revolutionize the way in which the field was viewed. The impact of his work was brought about by societal influences. The following pages will expose this reality and demonstrate the power such a style has to change the world.

Methodology

Our journey into the work of Shannon began just there - with an exploration of his major works in switching, genetics, information theory and cryptography. After examining his work specifically, we collected other primary source material of the time to get a sense of the context in which he was working. We examined works that Shannon cited, other related works of the time, and correspondences related to Shannon. Textbooks and commencement exercises of the time were consulted to obtain a sense of the state of his fields. Secondary historical sources provided a broader framework for our research, and provided many links between the various pieces we studied. Our choices in secondary source material were also driven by the fields of study we explored, rather than studies of Claude Shannon himself. But to return the focus back to Shannon, it was vital to speak to those who had known him and worked with him. His wife, advisees, colleagues and friends provided invaluable insight into the style and interests of Shannon, and helped us understand many of the unique characteristics of his life.

Information Theory

Information Theory before Shannon

To understand the contributions, motivations and methodology of Claude Shannon, it is important to examine the state of communication engineering before the advent of Shannon's 1948 paper, "A Mathematical Theory of Communication".

Before 1948, communication was strictly an engineering discipline, with little scientific theory to back it up. In fact, one might even go as far as to liken communication engineering of the time to a black art rather than the hard science it is today. Still, by the 1940's there were a large number of communication systems that were in use. Verdu, in his paper, "50 Years of Shannon Theory", lists some of the major ones as:

- Telegraph (from the 1830's)
- Telephone (1870's)
- Wireless Telegraph (1890's)
- AM Radio (1900's)
- Single-Sideband Modulation (1920's)
- Television (1930's)
- Teletype (1930's)
- Frequency Modulation(1930's)
- PCM (1930's)
- Vocoder (1930's)
- Spread Spectrum (1940's)

Reprinted from http://fohnix.metronet.com/~nmcewen/tel_off-page.html

1921 Newspaper Advertisement for RCA. Reprinted from http://fohnix.metronet.com/~nmcewen/tel_off-page.html

As is evident from this list, the systems of the time were diverse in not only the media used to deliver the message, but also in the methods used to transfer messages from one point to another. Separate fields emerged to deal with the problems associated with each medium, each with their own set of tools and methodologies. For example, it would have been inconceivable to an engineer that one would be able to send video over a phone line, as is commonplace today with the advent of the modem. The reason for this skepticism would not have been because no technology existed for sending moving pictures, or that telephone technology was not advanced. However, engineers treated both of these as separate entities and did not see the connection in the transmission of 'information'– a concept that would cross the boundaries of these disparate fields and bind them together.

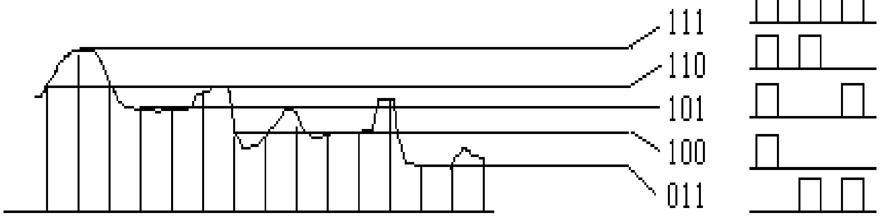
Although, there was no unifying theory to bring these fields together, some components that would prove to be key elements to a scientific theory of communication could already be seen in some of these systems. However, as the disciplines were thought of as separate entities. The first of these elements is the Morse code. Morse code had been in use since the early days of the telegraph. The Morse code is significant because it is a coding system that takes into account the frequency of symbols in order to transmit efficiently. Although, it was not envisioned as such, the Morse code, models the information source (which in the case of telegraph is the English language) probabilistically in order to maximize the speed of transmission of a set of symbols. This model of an information source is complementary with Shannon's concepts of an information source.

−−• •• •••− •   ••− •••   •− −•   •−

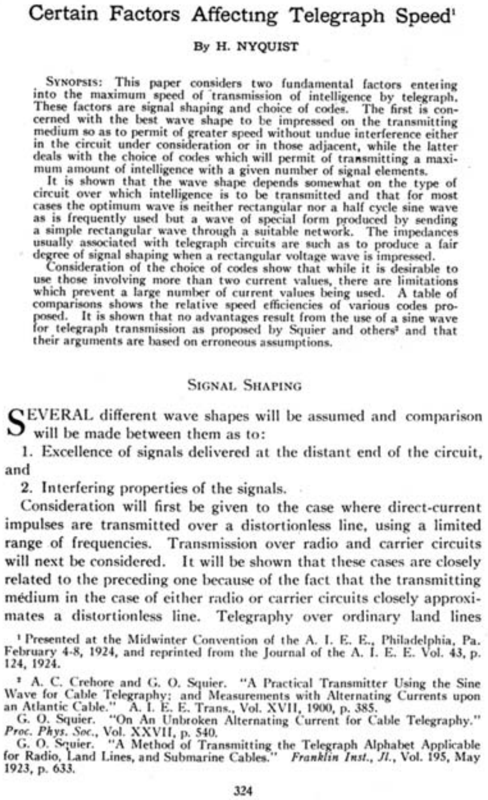
An example of morse code

The second component of importance was pulse-code modulation (PCM). PCM was a 'digital' system, although it transmitted along analog continuous-time signals. This was significant because Shannon explains his theory in terms of discrete systems rather than analog systems (the analog realm is a special case of the discrete system

Diagramatic representation of PCM. Reprinted from http://www.oneoffcd.com/info/historycd.cfm



As was the method of the time, such components had wended their way into different systems, without being grouped under an overarching theory. More significant however, was the work done by engineers at Bell Labs in the 1920's. These works captured some early efforts to come to grips with the concept of information as being distinct from the semantics of the message.



Certain Factors Affecting Telegraph Speed[1]

By H. NYQUIST

SYNOPSIS: This paper considers two fundamental factors entering into the maximum speed of transmission of intelligence by telegraph. These factors are signal shaping and choice of codes. The first is concerned with the best wave shape to be impressed on the transmitting medium so as to permit of greater speed without undue interference either in the circuit under consideration or in those adjacent, while the latter deals with the choice of codes which will permit of transmitting a maximum amount of intelligence with a given number of signal elements.

It is shown that the wave shape depends somewhat on the type of circuit over which intelligence is to be transmitted and that for most cases the optimum wave is neither rectangular nor a half cycle sine wave as is frequently used but a wave of special form produced by sending a simple rectangular wave through a suitable network. The impedances usually associated with telegraph circuits are such as to produce a fair degree of signal shaping when a rectangular voltage wave is impressed.

Consideration of the choice of codes show that while it is desirable to use those involving more than two current values, there are limitations which prevent a large number of current values being used. A table of comparisons shows the relative speed efficiencies of various codes proposed. It is shown that no advantages result from the use of a sine wave for telegraph transmission as proposed by Squier and others[2] and that their arguments are based on erroneous assumptions.

SIGNAL SHAPING

SEVERAL different wave shapes will be assumed and comparison will be made between them as to:

1. Excellence of signals delivered at the distant end of the circuit, and

2. Interfering properties of the signals.

Consideration will first be given to the case where direct-current impulses are transmitted over a distortionless line, using a limited range of frequencies. Transmission over radio and carrier circuits will next be considered. It will be shown that these cases are closely related to the preceding one because of the fact that the transmitting medium in the case of either radio or carrier circuits closely approximates a distortionless line. Telegraphy over ordinary land lines

[1] Presented at the Midwinter Convention of the A. I. E. E., Philadelphia, Pa. February 4-8, 1924, and reprinted from the Journal of the A. I. E. E. Vol. 43, p. 124, 1924.

[2] A. C. Crehore and G. O. Squier. "A Practical Transmitter Using the Sine Wave for Cable Telegraphy; and Measurements with Alternating Currents upon an Atlantic Cable." A. I. E. E. Trans., Vol. XVII, 1900, p. 385.
G. O. Squier. "On An Unbroken Alternating Current for Cable Telegraphy." *Proc. Phys. Soc.*, Vol. XXVII, p. 540.
G. O. Squier. "A Method of Transmitting the Telegraph Alphabet Applicable for Radio, Land Lines, and Submarine Cables." *Franklin Inst., Jl.*, Vol. 195, May 1923, p. 633.

324

First page of Nyquist''s "Certain Factors Affecting Telegraph Speed" .From Bell Labs Technical Journal.

The first of these engineers was Harry Nyquist, who is famous for his sampling theorems. Nyquist's 1924 paper, "*Certain Factors Affecting Telegraph Speed*"[4], is very focused on the engineering aspects of telegraph. However, there is a theoretical section entitled *Theoretical Possibilities of Using Codes with Different Numbers of Current Values*, in which he

---

[4] H. Nyquist, Certain Factors Affecting Telegraph Speed, Bell Labs Technical Journal 1924. pg. 324

produces two interesting results[5]. In this section, Nyquist discussed the "speed of transmission of intelligence."[6] This concept of intelligence is similar to Shannon's concept of information. Nyquist was beginning to understand the need to abstract away the actual content of the signal from the information carried within the message. While most of the paper considers the engineering aspects of a communication systems, it documents a logarithmic rule that governs the maximum amount of 'intelligence' that can be sent across a telegraph wire.

$\mathbf{W = K \log} \boldsymbol{m}$, where W is the speed of transmission of intelligence, m is the number of current values, and K is a constant.

Although this law was not general enough to describe all communication systems, it remains a special case of Shannon's later logarithmic law. In addition, he explained that the content of a message could be encoded with an "optimum code" such that it would "permit transmitting a maximum amount of intelligence with a given number of signal elements."[7] This was the first instance of an analysis of the theoretical bounds for ideal transmission codes.

In 1928, R.V.L. Hartley, another Bell Labs engineer, expanded on Nyquist's work by formalizing and generalizing some of the ideas contained in his paper[8]. Hartley was aware that Nyquist's concept of "intelligence" was still plagued by psychological aspects. Hartley wanted to develop a theory that was sufficiently general to encompass all of the major transmission media of the time. Hartley clearly stated the difference between the meaning and information in a particular message:

---

[5] H. Nyquist, Certain Factors Affecting Telegraph Speed, Bell Labs Technical Journal 1924, pg 332

[6] Ibid pg 332

[7] ibid pg 334

[8] R.V.L. Hartley, Transmission of Information, BellLabs Technical Journal, 1928. Pg 535

> Hence in estimating the capacity of the physical system to transmit information we should ignore the question of interpretation, make each selection perfectly arbitrary, and base our results on the possibility of the receiver's distinguishing the result of selecting any one signal from that of selecting any other. By this means the psychological factors and their variations are eliminated and it becomes possible to set up a definite quantitative measure of information based on physical considerations alone[9]

He stressed that the capacity of a system to transmit any sequence of symbols depended solely on distinguishing at the receiving end between the results of various selections at the sending end and not on the meaning of the sequence.

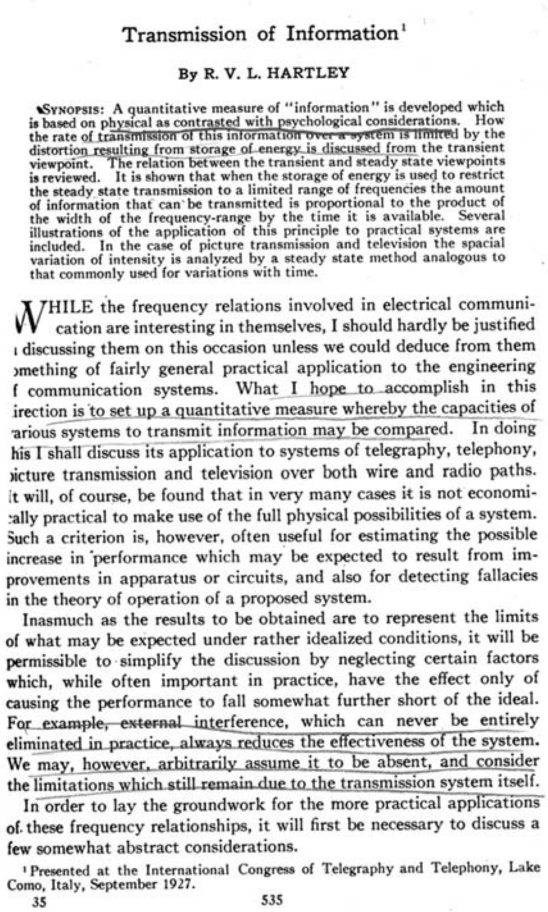
Hartley also generalized Nyquist's logarithm law for the amount of information transmitted.

The form he gives for the equation is

$$\mathbf{H = \log S^n}$$

Where ***S*** in the number of possible symbols, and

n is the number of symbols in a transmission.



First page of Hartkey's "Transmission of Information.From Bell Labs Technical Journal. 1928

Transmission of Information[1]

By R. V. L. HARTLEY

SYNOPSIS: A quantitative measure of "information" is developed which is based on physical as contrasted with psychological considerations. How the rate of transmission of this information over a system is limited by the distortion resulting from storage of energy is discussed from the transient viewpoint. The relation between the transient and steady state viewpoints is reviewed. It is shown that when the storage of energy is used to restrict the steady state transmission to a limited range of frequencies the amount of information that can be transmitted is proportional to the product of the width of the frequency-range by the time it is available. Several illustrations of the application of this principle to practical systems are included. In the case of picture transmission and television the spacial variation of intensity is analyzed by a steady state method analogous to that commonly used for variations with time.

WHILE the frequency relations involved in electrical communication are interesting in themselves, I should hardly be justified discussing them on this occasion unless we could deduce from them mething of fairly general practical application to the engineering f communication systems. What I hope to accomplish in this irection is to set up a quantitative measure whereby the capacities of arious systems to transmit information may be compared. In doing his I shall discuss its application to systems of telegraphy, telephony, icture transmission and television over both wire and radio paths. t will, of course, be found that in very many cases it is not economially practical to make use of the full physical possibilities of a system. Such a criterion is, however, often useful for estimating the possible increase in performance which may be expected to result from improvements in apparatus or circuits, and also for detecting fallacies in the theory of operation of a proposed system.

Inasmuch as the results to be obtained are to represent the limits of what may be expected under rather idealized conditions, it will be permissible to simplify the discussion by neglecting certain factors which, while often important in practice, have the effect only of causing the performance to fall somewhat further short of the ideal. For example, external interference, which can never be entirely eliminated in practice, always reduces the effectiveness of the system. We may, however, arbitrarily assume it to be absent, and consider the limitations which still remain due to the transmission system itself.

In order to lay the groundwork for the more practical applications of these frequency relationships, it will first be necessary to discuss a few somewhat abstract considerations.

[1] Presented at the International Congress of Telegraphy and Telephony, Lake Como, Italy, September 1927.

35 535

---

[9] R.V.L. Hartley, Transmission of Information, BellLabs Technical Journal, 1928, pg 536

Shannon cited both these papers in his *Mathematical Theory of Communication.* In addition, in a 1984 interview, Shannon said, "I had already read Hartley's paper, and that it had been an important influence on my life."[10]

When looking at information theory as proposed by Shannon, in the broadest sense, we can divide it into two parts. The first of these parts is the conceptualization of information and the modeling of information sources. The second part of the theory encompasses the sending of information across the channel – what are the limits on the amount of information that can be sent and what is the effect of noise on this communication.

Aspray, in his paper, "Information: A Survey"[11] presents some interesting theories on the evolution of the concept of information and communication theory. He examines the roots of information science in nineteenth- and early twentieth century mathematical logic, physics, psychology, and electrical engineering and then focuses on how Warren McCulloh, Walter Pitts, Claude Shannon, Alan Turing, John von Neumann, and Norbert Weiner combined these diverse studies into a coherent discipline. Aspray writes that 5 areas in particular led to the scientific conceptualization of information:

1.James Clerk Maxwell, Ludwig Bolzmann, and Leo Szilar's work in thermodynamics and statistical mechanics, especially on the concept of entropy.

2. The research work in communication and control that arose from the development of telegraphy, radio, and television.

---

[10] Robert Price, A Conversation with Claude Shannon, IEEE Communications Magazine, May 1984, pg 123

3. Work starting in the nineteenth century on the physiology of the nervous system, and work in the 20th century on homeostasis and internal regulation of living organism.

4. The development of functionalist and behaviorist theories of the mind in psychology, leading to a view of the brain as a processor of information and to a demand for the experimental verification of theories of the mind through observation of external behavior.

5. The development of recursive function theory in mathematical logic as a formal, mathematical characterization of the human computational process.

Aspray identifies Claude Shannon, Norbert Wiener, Warren McCulloh, Walter Pitts, Alan Turing, and John Von Neumann as the leaders of a movement that took place during and after the war to unify the theories of information characterization and processing that arose from diverse roots.

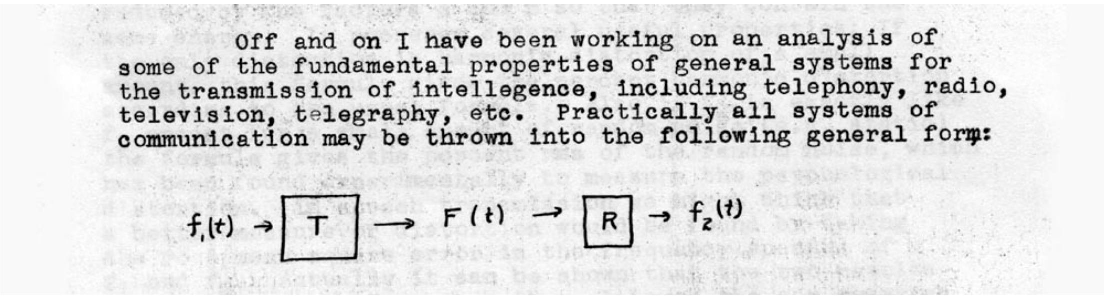
Although these scientists came from different disciplines, the single most important factor that unified them was their background in mathematical logic. The reason why this background was so important is because mathematical logic is the study of the laws of thought in abstraction. One might hypothesize that this ability to take a process and abstract it to a high level enabled Shannon to formalize the process as viewing information as distinct from semantics.

---

[11] William Aspray, Information: A Survey, Annals of the History of Computing, Volume 7, 1985. Pg. 117

What was Missing

While we have taken a brief look at how some of the components of information theory came into being, by the 1940's there was still no unifying theory for communication.



Off and on I have been working on an analysis of some of the fundamental properties of general systems for the transmission of intellegence, including telephony, radio, television, telegraphy, etc. Practically all systems of communication may be thrown into the following general form:

$$f_1(t) \rightarrow \boxed{T} \longrightarrow F(t) \rightarrow \boxed{R} \rightarrow f_2(t)$$

Excerpt of a letter from Shannon to Bush. Feb. 16, 1939. From Library of Congress

Shannon showed Bush a diagram of a generalized communication system. We can clearly see that Shannon had already started abstracting away layers in the communication hierarchy to get at the root of the problem behind communication. He does not yet refer to the transmissions as information, but rather as "intelligence", as was done by Nyquist. Interestingly, Shannon was still caught in the analog realm and did not discuss discrete systems. In addition, he discussed the effect of "distortion" on the transmitted signal. We see the beginnings of Shannon Theory in his discussion on fundamental limits and bandwidth-delay products.

While we have seen the early beginnings of Shannon Theory, in order to study his methodologies, it is useful to first examine the finished product before we delve into an examination of various bodies of work that Shannon created between these times.

1948 Mathematical theory of communication

In his 1948 paper, Shannon listed the parts of a communication system as:

1. Information source
2. Transmitter
3. Channel
4. Receiver
5. Destination

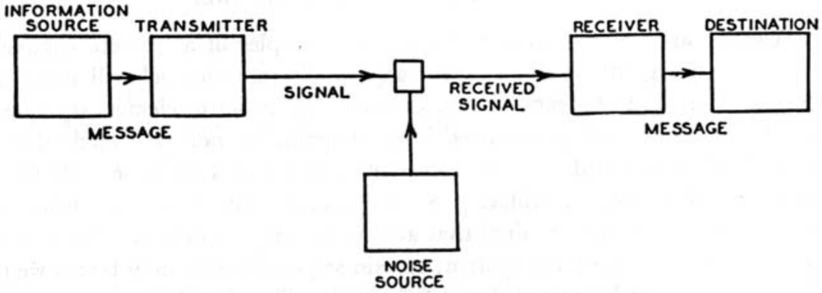


Fig. 1—Schematic diagram of a general communication system.

From "A Mathematical Theory of Communication

Shannon grouped communication systems into 3 types – discrete, continuous, mixed, stating that the discrete case is the foundation for the other two and "has applications not only in communication theory, but also in the theory of computing machines, the design of telephone exchanges and other fields."

Shannon stated that information is defined in the simplest cases to be measured by the logarithm of the number of available choices of symbols. As the base of the logarithm, Shannon chooses the number two, thus making the smallest unit of information a bit.

Building on Nyquist and Hartley's work, right in the beginning of his paper, Shannon clearly distinguished between the information and semantics of a message by stating, "The semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one selected from a set of possible messages"

After establishing this, Shannon introduced the idea of an information source being probabilistic by posing the following question to the reader - If an information source is producing messages by successively selecting symbols from a finite set, and the probability of a symbol appearing is dependent on the previous choice, then what is the amount of information associated with this source?

Shannon explained the answer to this question by describing information in terms of entropy. If a communication source is not characterized by a large degree of randomness or choice, then the information (or entropy) is low i.e.
1-(actual entropy/maximum entropy) = redundancy.

Shannon understood that to derive the fundamental theorems for a communication system, he would need to provide a precise definition of information, such that it was a physical parameter that could be quantified. He did this by providing his entropic definition of information. Once Shannon had established this concept of information, he was able to work within this framework to discover his two fundamental theories of communication.

The first theorem deals with communication over a noiseless channel.

> Let a source have entropy H(bits per symbol) and a channel have a capacity C (bits per transmit at the average rate C/H - ε symbols per second over the channel where ε is arbitrarily small. It is not possible to transmit at an average rate greater than C/H.

The main idea behind this theorem is that the amount of information that is possible to transmit is based on its entropy or randomness. Therefore based on the statistical characteristic of the information source, it is possible to code the information so that it is possible to transmit it at the maximum rate that the channel allows. This was revolutionary as communication engineers previously thought that the maximum signal that could be transported across a medium was related to various factors such as frequency, not on the concept of information.

Shannon's second theorem deals with communication in a noisy environment. Shannon states:

> Let a discrete channel have the capacity C and a discrete source the entropy per second H. If H<=C there exists a coding system such that the output of the source can be transmitted over the channel with an arbitrary small frequency of errors. If H > C it is possible to encode the source so that the equivocation is less than H – C + ε, where ε is arbitrarily small. There is no method of encoding which gives an equivocation less than H –C.

The idea that Shannon is conveying is that no matter what the noise, there is an encoding scheme that allows you to transmit the information error-free over the channel (so long as H < C). Again, this idea was revolutionary as it was believed that after a certain level of noise, it would be impossible to transmit the desired signal.

The Shannon Methodology

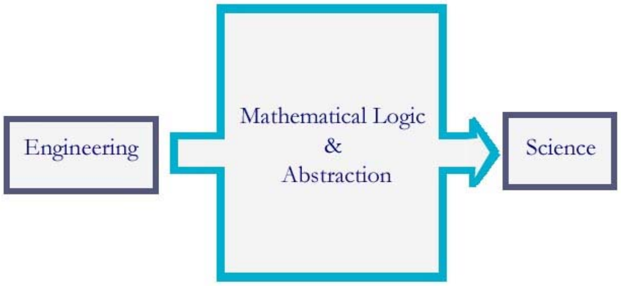
We have seen how in 1939, Shannon was still far from a complete theory of communication. However, 9 years later, he published *A Mathematical Theory of Communication*, in which he gave a complete but simple description of a generalized communication system. His architecting of the parts of a communication system, his modeling of information as entropy, and his theorizing on the limits of communication both with and without noise were leaps beyond contemporary thinking in communication engineering.

In the following sections, we will see a thorough analysis of the content and context of the various topics that Shannon researched, from which we can make a detailed analysis of the methodology and forces that shaped Shannon's work. However, from what we have seen in this section, we can come to certain conclusions.

Shannon as the architect

As one reads *A Mathematical Theory of Communication*, one notices Shannon's framing of the problem, which is evident from his famous communication diagram. In this diagram, which is the cornerstone of his theory of communication, Shannon abstracted away most of the complexities of communication systems. In fact, each part of the system is just a simple box. He then deals with each of these components separately. This concept of abstraction and simplification is one that we will see as key to Shannon's work style. Shannon himself talks about this in a talk that he gave in 1953:

> The first one that I might speak of is the of simplification. Suppose that you are given a problem to solve, I don't care what kind of problem – a machine to design, or a physical theory to develop, or a mathematical theorem to prove or something of that kind – probably a very powerful approach to this is to attempt to eliminate everything from the problem except the essentials; that is, cut it down to size. Almost every problem that you come across is befuddled with all kinds of extraneous data of one sort or another; and if you can being this problem down into the main issues, you can see more clearly what you're trying to do and perhaps find a solution. Now in so doing you may have stripped away the problem you're after. You may have simplified it to a point that it doesn't even resemble the problem that you started with; but very often if you can solve this simple problem, you can add refinements to the solution of this until get back to the solution of the one you started with.[12]



Once Shannon had abstracted his problem to a high enough level, he would then apply the mathematics that he was so familiar with from his days at Michigan – Boolean Logic in the case of his Master's thesis on switching circuits, algebra in the case of his work on genetics, and probability

---

[12] speech by Claude Shannon – Creative Thinking , Claude Elwood Shannon: Miscellaneous Writings

in the case of communication. By doing this, Shannon could get at the scientific theories that are the basis for the field.

While it would certainly be simplistic to credit all of Shannon's tremendous discoveries solely to his methodology, we feel that it is very important as it counters current thoughts on engineering and science. Shannon's methodology is depicted pictorially above, in which we see the engineering problem being tackled at a high level in order to get at the science behind it. However, in contrast, conventional wisdom dictates that engineering is the application of science to a particular problem.

In the following sections, while we examine the context and content of Shannon's other works, we will keep this unique methodology in mind to see how it has affected Shannon's work.

In addition to this common ground, all of these scientists had wartime exposure that heightened their awareness of the importance of information theory. Weiner developed cybernetics based on his work with Vannevar Bush, with Y. W. Lee on wave filters, and his collaborations with Julian Bigelow on fire control for antiaircraft artillery. Similarly, Aspray writes that a lot of Shannon's theory of communication was spawned by the tremendous advances in communication engineering during the development of the radar and partly from the need for secure communications during the war.

Switching Theory

Background and Contemporary Work

Written in 1938, *A Symbolic Analysis of Relay and Switching Circuits* did not present a slew of research which had never been seen before. Rather, it was the insight of linking Boolean algebra to circuits which made it so novel. Looking at the types of Master's theses which were written in the MIT Electrical Engineering Department at the time can give a good idea of what else was going on. The hot topics seemed to deal more with energy, power, and motors, no doubt because major advancements were needed to match the growing need for transportation. 1936 thesis titles *Influence of Harmonic Fields Upon the Induction Motor Torque Characteristic* and *Some Technical and Economic Problems of Energy Storage in Electric Power Systems* reflect this movement.[13]

In the years 1937-1940, there were several theses pertaining to circuits, all with the same title *Equivalent Circuits for the Representation of Large Metropolitan Area in Transient Stability Studies*.[14] This showed that researchers were looking a lot at how equivalent circuits could be used to represent different kinds of networks, which Shannon delved into slightly in his thesis. In this time span there were no theses in the mathematics department that covered Boolean Algebra topics.

The connection between Boolean algebra and circuits was actually first recognized as early as 1886 by an American philosopher and logician Charles Pierce, but he never did much to further the research.[15] In 1910, Paul Ehrenfest of St. Petersburg University was the first to propose the idea of applying logic in a technical fashion; he had in mind the telephone commutation station as the best object for this purpose. Soviet engineer Gersevanov actually implemented mathematical logic on a

---

[13] Graduation Exercises. 1936

[14] Graduation Exercises. 1940

[15] Povarov, 2001.

broad scale in his design for a hydraulic power plant in 1923. Then in 1936, a psychologist from Chicago named Benjamin Burack constructed a device which implemented electric light bulbs for displaying the logical relationships between a collection of switches. However he did not publicize this work until 1949[16]. Burack's logic device was constructed two years before Shannon's master's thesis was completed.

The work that Shannon did in his master's thesis was closely paralleled by two contemporary researchers in the former Soviet Union and Japan, two areas of the world where switching theory and logic design were catching on. The Russian physicist Vladimir Shestakov published some articles and did his dissertation thesis on the subject. He made his first reports in 1935 after studying telegraphy and telephone systems and reading *Algebra of Logic.* Also in 1935, Nakashima published an article called *A Realization Theory for Relay Circuits.* Switching theory and logic design developed in Eastern Europe and the former Soviet Union because of automatic control problems in system theory. In Japan, there was ongoing development in applying Boolean logic to technical systems because of interest in the design of relay network systems.[17]

Building Blocks to Shannon's Master's Thesis

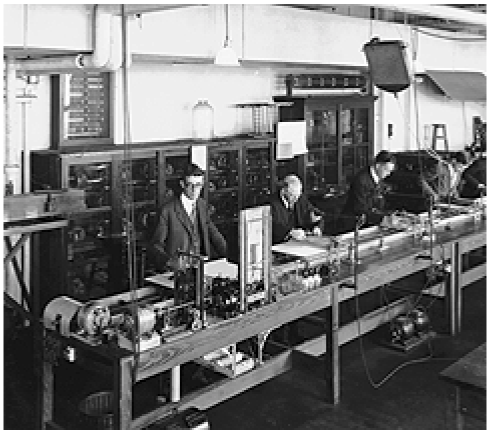
Shannon started pursuing his master's degree at MIT in 1936 as a research assistant in Vannevar Bush's lab. He had just graduated from the University of Michigan with dual degrees in Mathematics and Electrical Engineering, which would lay the groundwork for Shannon's research for years to come. Reportedly Shannon took up the position in Bush's lab after seeing it advertised on a postcard nailed to a campus bulletin board.[18] At the time Bush was MIT's vice president and

---

[16] Povarov, 2001.
[17] Ibid.
[18] Interview with Charles Vest, 2001.

dean of engineering. Bush gave Shannon the job of taking care of the Differential Analyzer, which was an analog computer made of gears, pulleys, and rods whose function it was to evaluate and solve differential equations. Shannon would help visiting scientists set up their problems on the analyzer by rearranging mechanical linkages in between the rods so that their movements would match up with the corresponding mathematical equations.

Before Shannon's arrival, Vannevar Bush and his colleagues had labored on the differential analyzer project for about ten years. The final machine funded by the Rockefeller Foundation weighed almost a hundred tons and was comprised of two thousand vacuum tubes, several thousand relays, a hundred and fifty motors, and automated punched-tape access units. During the second World War the Rockefeller Differential Analyzer was probably the most important computer in operation in the United States,[19] and it provided the spark for Shannon's groundbreaking master's thesis *A Symbolic Analysis of Relay and Switching Circuits.*


Intergraph. Reprinted from Mindell's "MIT Differential Analyzer" webpage

The Differential Analyzer arose from work on the Intergraph. Bush, along with F.G. Kear, H.L. Hazen, H.R. Stewart, and F.D. Gage, developed it in 1927, a machine which could be mechanically set to solve sets of first-order differential equations.[20]

It was hailed on the front page of the New York Times: *'Thinking Machine' Does Higher Mathematics; Solves Equations That Take Humans Months.* Using both electrical and mechanical devices, the integraph suffered from imprecision and unneeded

---

[19] Owens. 1984

complexity. Getting inspiration from the mathematical elegance and mechanical simplicity of a mechanical disc-integrator, Bush decided to build a new mechanical machine.[21]

The Differential Analyzer developed in 1931 was an improvement on the integraph and could handle sixth-order differential equations.

It consisted of a long table-like, supporting framework which was crossed by connectable shafts. One side contained an array of drawing boards while the other side had six disc-integrators. The shafts would control the pens so that they could trace curves on the drawing boards. The operator could also manually cause a pen to follow a given curve, thus imparting a desired rotation to particular shafts. Through association of terms of an equation with the shaft rotations combined with employing an assortment of gearings, the machine could be used for all the basic mathematical operations in addition to integration.[22]

It was the Rockefeller Differential Analyzer's extremely complicated control circuit, composed of approximately one hundred switches that could be automatically opened and closed by an electromagnet, which led to Shannon's momentous discovery.

Rockefeller Differential Analyzer. Reprinted from Mindell's "MIT Differential Analyzer" webpage

At the time that Shannon was writing his thesis, the Rockefeller Differential Analyzer was still under

---

[20] "Symposium Vannevar Bush at MIT" website, 2001.

[21] Owens. 1984

[22] Owens. 1984

construction. Dr. Charles Vest recalled Shannon describing how one night it just dawned on him that the circuits he was building so closely resembled the Boolean logic he had studied as a student at Michigan.[23] He then realized that you could combine switches in circuits in such a manner as to carry out symbolic logic operations. It was a discovery which linked two well known fields, but the link had up to then not been fully explored or widely publicized.

The implications of true and false being denoted by the numerical digits one and zero were enormous, since then it would be possible for the relays could perform operations of binary arithmetic. In his own words, Shannon wrote, "It is possible to perform complex mathematical operations by means of relay circuits."[24] From initially showing the design of a circuit which could add binary numbers, he moved on to realizing circuits which could make comparisons and thus be capable of doing courses of action such as "if the number X equals the number Y, then do operation A." A machine could now have the capability of making decisions, which spawned a whole new era of computers and artificial intelligence.[25] Thus digital logic was born.

A Symbolic Analysis of Relay and Switching Circuits

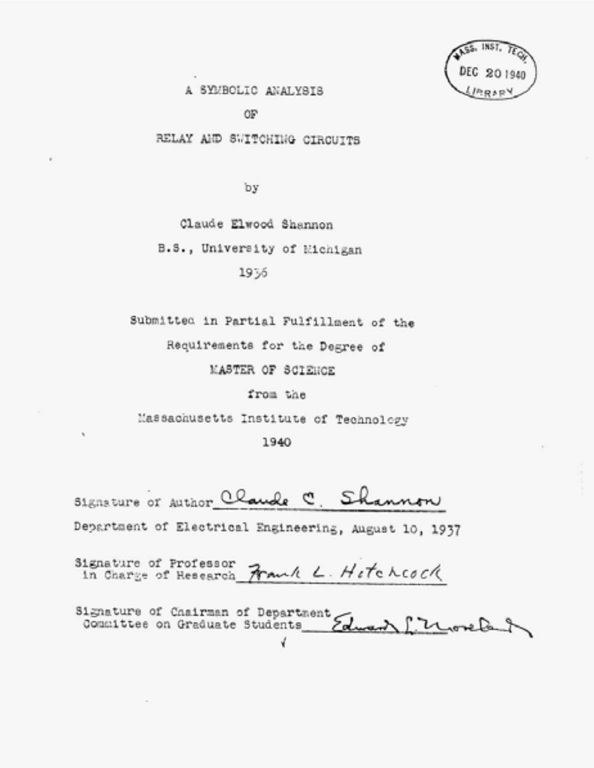
Shannon's master thesis presented a method of attack for finding the simplest circuit representation of a network synthesis problem with certain characteristics. Given a synthesis problem the first step was to express the desired characteristics as a system of equations, with the terms of the equations being the various relays or switches in the circuit. The next step was to manipulate the equations into the simplest form which had the least occurrence of terms. A calculus was developed for this purpose using simple algebraic algorithms. This calculus was almost exactly analogous to the

---

[23] Interview with Charles Vest, 2001.

[24] Waldrop. 2001

Calculus of Propositions, an algebra of logic originated by George Boole. In the final step the circuit representing the synthesis problem could then be immediately drawn from the manipulated equations. The thesis started out by introducing the postulates and theorems necessary to construct the calculus for simplifying terms, then it went into an analysis of circuit properties with multiple examples of solving real problems with a circuit representation.



A SYMBOLIC ANALYSIS
OF
RELAY AND SWITCHING CIRCUITS

by

Claude Elwood Shannon
B.S., University of Michigan
1936

Submitted in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
from the
Massachusetts Institute of Technology
1940

Signature of Author Claude E. Shannon
Department of Electrical Engineering, August 10, 1937

Signature of Professor
in Charge of Research Frank L. Hitchcock

Signature of Chairman of Department
Committee on Graduate Students Edward L. Moreland

Cover page from Shannon's 1938 Master's thesis. Reprinted from the MIT Libraries Thesis Archives

In the treatment of circuits, a terminal could either be open (infinite impedance) or closed (zero impedance). If a circuit between two terminals *a* and *b*, denoted X_{ab}, was closed than it was represented by the symbol 0 (zero). The symbol 1(unity) was used to represent the hindrance of an open circuit. Given these basic definitions, particular postulates were put forth. These postulates were used to develop the theorems used in connection with circuits containing only series and parallel connections. Negation was a new operation which when applied to a hindrance X was written X' and was defined as a value which was opposite of X.

---

[25] Waldrop. 2001

Postulates

a. 0*0 = 0	A closed circuit in parallel with a closed circuit is a closed circuit.
b. 1+1 = 1	An open circuit in series with an open circuit is an open circuit.
a. 1+0 = 0+1 = 1	An open circuit in series with a closed circuit in either order(i.e., whether the open circuit is to the right or left of the closed circuit) is an open circuit
b. 0*1 = 1*0 = 0	A closed circuit in parallel with an open circuit in either order is a closed circuit
a. 0+0 = 0	A closed circuit in series with a closed circuit is a closed circuit.
b. 1*1 = 1	An open circuit in parallel with an open circuit is an open circuit.
At any give time either X=0 or X=1	

Theorems

X+Y = Y+X
XY = YX
X+(Y+Z) = (X+Y)+Z
X(YZ) = (XY)Z
X(Y+Z) = XY+XZ
X+YZ = (X+Y)(X+Z)
1*X = X
0+X = X
1+X = 1
0*X = 0

Negation Theorems

X+X' = 1
XX' = 0
0' = 1
1' = 0
(X')' = X

The basis was now set to demonstrate the equivalence of the above calculus with the Calculus of Propositions. The symbols of Boolean algebra allowed a variable to be represented by the values 0 and 1. The link that Shannon made between Boolean logic and symbolic relay analysis could be summed up as follows:

Analogue Between the Calculus of Propositions and the Symbolic Relay Analysis

Symbol	Interpretation in Relay Circuits	Interpretation in the Calculus of Propositions
X	The circuit X	The proposition X
0	The circuit is closed	The proposition is false
1	The circuit is open	The proposition is true
X+Y	The series connection of circuits X and Y	The proposition which is true if either X or Y is true
XY	The parallel connection of circuits X and Y	The proposition which is true if both X and Y are true
X'	The circuit which is open when X is closed and closed when X is open	The contradictory of proposition X
=	The circuits open and close simultaneously	Each proposition implies the other

Using De Morgan's theorem, many theorems useful for simplifying expressions could be derived. To simplify a circuit the equations which describe it should be manipulated into the form in which the least number of elements occur, each element being a hindrance. These methods were very useful for circuits containing only series and parallel connections.

An Example of a Synthesis Problem

Shannon applied these concepts to an example of the design of an electric combination lock. This showed the method of finding a circuit representation of a sequential system. There were several characteristics of the lock:

- Five pushbutton switches are available on the front of the lock, labeled *a, b, c, d, e,* and these buttons are to be pressed in the order: *c, b, a and c simultaneously, d.*
- A lock will unlock in the case that the correct order is pressed, but if there is an incorrect button pressed an alarm *U* will operate.
- A switch *g* must be pressed to relock the system.
- A switch *h* must be pressed to release the alarm once it has started.
- Let *w, x, y,* and *z* correspond to the sequential relays for the sequential system

Given these conditions the system of equations could be derived:

$w = cw + z' + U',$

$x = bx + w + z' + U',$

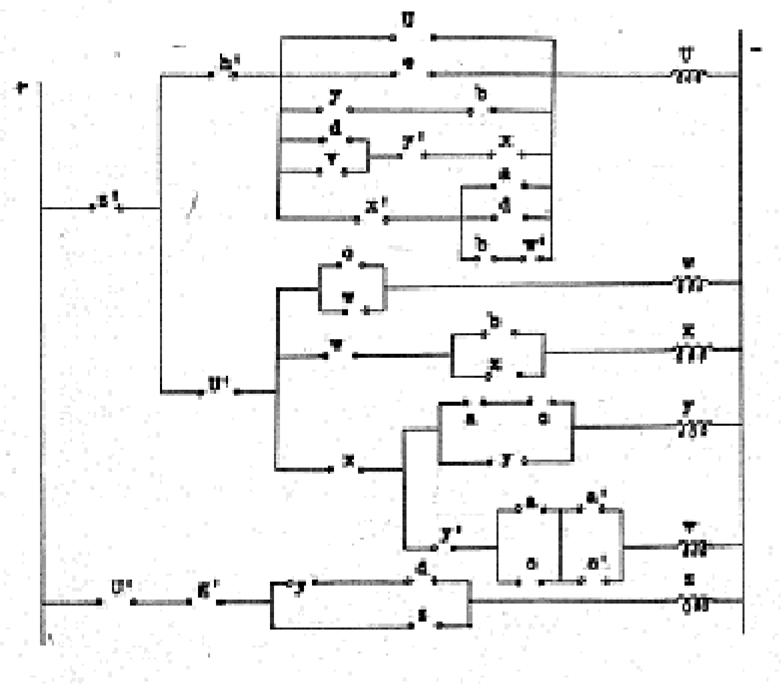
$y = (a + c)y + x + z' + U',$

$z = z(d + y) + g' + U',$

$v = x + ac + a'c' + z' + U',$

$U = e(w' + abd)(w + x' + ad)[x + y' + dv(t - s)][y + bv(t - s)]U + h' + z'.$

Using the theorems derived earlier in the thesis the system could be simplified and the circuit drawn up as:



Electric Combination lock circuit

The thesis also discussed multi-terminal and non-series-parallel circuits which include star-mesh and delta-wye transformations. In addition to the electric combination lock it presented more examples of systems represented by simplified circuits like the base two electric adder and the selective circuit.

From these examples, one can see the significance of relating Boolean logic to circuits. This insight was the most important part of the thesis. In his 1958 textbook on switching theory, Shannon's thesis advisor Sam Caldwell described the thesis as "an opportunity to supplement skill with methods based on science, and thus to increase the productivity of circuit designers."[26] This statement epitomizes Shannon's main manner of thinking, an effort to introduce mathematical structure to every area he researched. His application of mathematical thinking to the area of circuit design produced what H.H. Goldstine called: "one of the most important master's theses ever written... a landmark in that it changed circuit design from an art to a science."[27]

Popular Recognition

Although Shannon's discovery was groundbreaking and led to a whole new era of digital logic and computers, the connection between Boolean algebra and circuits had been recognized and applied in several parts of the world, as early as 1886. Why were there so many instances of researchers using the idea of Boolean logic for technical applications but never pursuing it very far or not publishing it? Since the idea had been stumbled upon many times before Shannon, it obviously was not a difficult concept to grasp, unlike Shannon's later work in information theory.

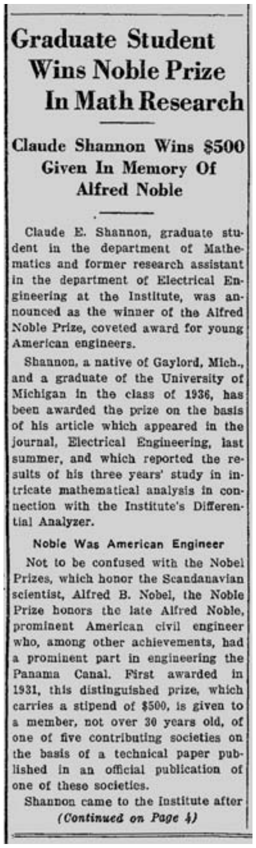
Shannon's theories became widely publicized and read because it was the right time for technology to take that step. This was in part due to the development of machines which started to take on "thinking" functions and which would later evolve to make decisions. The social and economic pressures came from worldwide changes which demanded something like digital logic to lead machines into a new age. During the previous fifty years when other smart minds in many countries

---

[26] Caldwell. 1958. pg. vii

came up with similar theories, there simply were no opportunities to take advantage of the discovery. Rather, the work would remain an interest of the researcher. Shannon's thesis was the basis for switching theory and logic design which would be applied to a myriad of problems arising in switching in railroad systems, information transmission and error correction, and automatic telephony.

Another element in the recognition of his work was his receipt of the Alfred Noble prize, an engineering award given to a young author of a technical paper of exceptional merit. It was an honor he was not expecting at all as shown in his letter to Bush "You may have heard that I received the Alfred Noble prize … for my paper on switching circuits. In fact I have a sneaking suspicion that you have not only heard about it but had something to do with my getting it …. if so, thanks a lot. I was so surprized[sic] and pleased to receive the letter announcing the award that I nearly fainted!"[28] Apparently Bush had submitted Shannon's master thesis to the Alfred Noble committee for consideration without consulting Shannon. This episode clearly emphasized how Bush was influential in the recognition of Shannon's work. It also depicted how Shannon was uninterested in pursuing fame – he focused most of his interest in the research which was abstracted away from the engineering aspect. While set to work on the Differential Analyzer, it was the of hundreds of switches on the control circuit made him think about applying Boolean logic to circuits. It was social factors like Vannevar Bush and the technological pressures of the time which made Shannon's research famous and allowed it to make great contributions to the digital age.



Graduate Student Wins Noble Prize In Math Research

Claude Shannon Wins $500 Given In Memory Of Alfred Noble

Claude E. Shannon, graduate student in the department of Mathematics and former research assistant in the department of Electrical Engineering at the Institute, was announced as the winner of the Alfred Noble Prize, coveted award for young American engineers.

Shannon, a native of Gaylord, Mich., and a graduate of the University of Michigan in the class of 1936, has been awarded the prize on the basis of his article which appeared in the journal, Electrical Engineering, last summer, and which reported the results of his three years' study in intricate mathematical analysis in connection with the Institute's Differential Analyzer.

Noble Was American Engineer

Not to be confused with the Nobel Prizes, which honor the Scandanavian scientist, Alfred B. Nobel, the Noble Prize honors the late Alfred Noble, prominent American civil engineer who, among other achievements, had a prominent part in engineering the Panama Canal. First awarded in 1931, this distinguished prize, which carries a stipend of $500, is given to a member, not over 30 years old, of one of five contributing societies on the basis of a technical paper published in an official publication of one of these societies.

Shannon came to the Institute after

(Continued on Page 4)

The Tech,
Feb. 6. 1940

---

[27] AT&T research biography

After the thesis on Boolean logic and switching, Vannevar Bush was determined to nurture Shannon's talent. He knew that Shannon was a "very shy and retiring sort of individual, exceedingly modest, and one who would readily be thrown off the track"[29] and tried to create the best circumstances for him. In 1938, with Bush's help, Shannon entered the Mathematics department at MIT as a doctoral student. The two also began a correspondence, exchanging ideas on possible research topics and directions. This was how Shannon began his work in genetics.

---

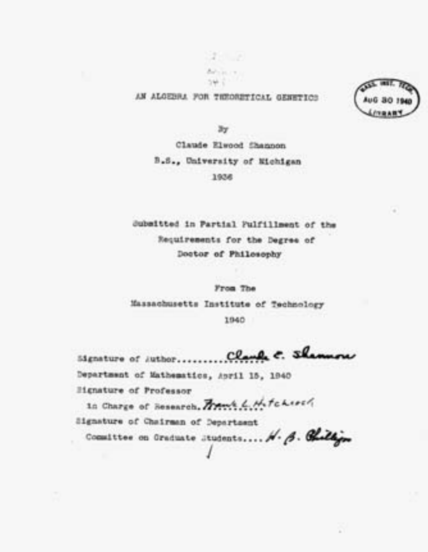
[28] Correspondence from Shannon to Bush. December 13, 1939.

[29] Correspondence from Bush to Wilson. December 15, 1938.

Genetics

An Algebra for Theoretical Genetics, the beginnings

Vannevar Bush asked Claude Shannon to see whether he could find something interesting to work on in the field of genetics. He remarked to a colleague, "It occurred to me that, just as a special algebra had worked well in his hands on the theory of relays, another special algebra might conceivably handle some of the aspects of Mendelian heredity."[30] However, Shannon had at best only a vague knowledge of genetics. According to Bush, "At the time that I suggested that he try his queer algebra on this subject, he did not even know what the words meant ..."[31]. Despite his unfamiliarity with the subject, he soon produced a manuscript with some ideas of an algebraic notation. The idea was so novel that Bush arranged for Shannon to work under Barbara Burks, a researcher at the Eugenics Record Office in Cold Spring Harbor in the summer of 1939. The research was then refined into a doctoral thesis submitted in the spring of 1940.



AN ALGEBRA FOR THEORETICAL GENETICS

By
Claude Elwood Shannon
B.S., University of Michigan
1936

Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

From The
Massachusetts Institute of Technology
1940

Signature of Author..........Claude E. Shannon
Department of Mathematics, April 15, 1940
Signature of Professor
in Charge of Research..........Frank L. Hitchcock
Signature of Chairman of Department
Committee on Graduate Students....H. B. Phillips

Cover page from Shannon's 1940 doctoral thesis.
Reprinted from the MIT Libraries Thesis Archives

Why did Bush suggest the field of genetics? In 1939, he had just become the president of the Carnegie Institution of Washington, which had the Cold Spring Harbor Genetics Laboratory as one of its research centers. As a result, he knew a significant amount about what was going on in the field of genetics and saw that this would be a good time for Shannon to try his "queer algebra", as he put it. To understand further the societal influences on this suggestion, the next few sections discuss events in genetics leading up to 1940.

---

[30] Correspondence from Bush to Wilson. December 15, 1938.
[31] Correspondence from Bush to Burks. January 5, 1938.

History of population genetics

Prior work in genetics had a strong tie to mathematics from the very beginning. Gregor Mendel's famous experiments with pea plants correlated traits throughout different generations with particular proportions, arriving at the concept of a gene. However, this research went unknown for a few decades until in the 1900's, when the results were rediscovered by others. Researchers then raced to correlate common mathematical knowledge with the possibilities of genetic combinations. Thus began the field of genetics. The applications to living beings soon spread from plants to animals, then to human beings. This study of inheritable human traits was called eugenics.

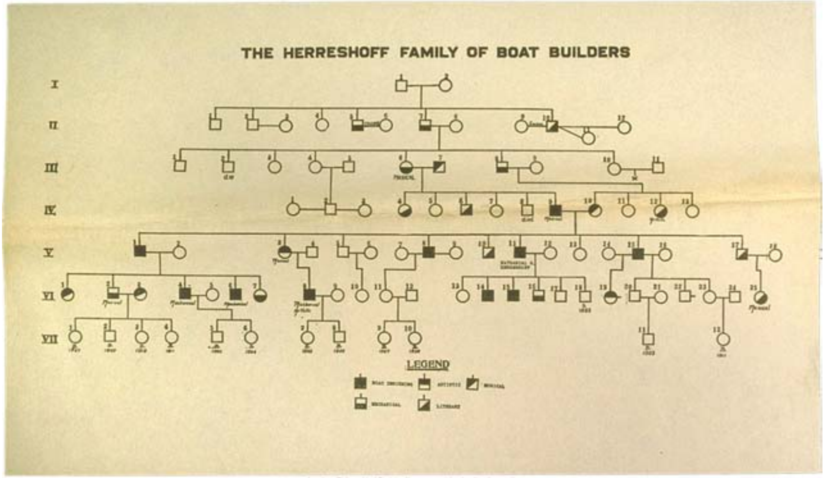
Eugenics

In 1904, at the same time that studies began in earnest of mathematical applications to genetics, the Eugenics Record Office was founded at Cold Spring Harbor Laboratories by Charles Davenport. This became the center of all eugenics study in the United States. Eugenicists were interested in finding out how diseases such as hemophilia and Parkinson's disease were passed from generation to generation and how traits such as eye color were also inherited through a Mendelian model.[32] This interest soon spread to research on much less definable traits, such as the "love of the sea", musicianship, feeblemindedness, and intelligence.

Eugenics tree logo. Reprinted from the Image Archive on the American Eugenics Movement.

---

[32] Image Archive on the American Eugenics Movement, website.



American Philosophical Society. Noncommercial, educational use only.

Boat Builder Pedigree. Reprinted from the Image Archive on the American Eugenics Movement. From studying this pedigree, one may surmise that boatbuilding is a sex-linked trait, expressing itself in males -- much of the research in eugenics was conducted in a similar manner.

Unfortunately, eugenics research methodology was highly suspect to inaccuracies and personal prejudice – anybody that was from another country or of a lower economic status was often assumed to have undesirable traits. In spite of the drawbacks, or perhaps because of them, the field soon caught the attention of the public.

Eugenics gained popularity through the 1920's and 30's. There was much public support for the preservation and propagation of good traits, while prevention of the bad traits was condoned. In the United States, Britain, and Germany, this was particularly evident in the movement to discourage the propagation of the unfit traits. Sterilization soon became advocated for those who were not of good eugenic stock and sometimes forced upon patients of mental institutions.[33]

---

[33] Image Archive on the American Eugenics Movement, website..

Genetics in late 1930's

By the mid 1930's, eugenics had split off from genetics as its own field due to vast differences in their methods of research. Geneticists were focused on developing more mathematical methods, using statistics to study their experiments. Hogben's 1933 publication (referenced by Shannon's thesis) moved away from the mere applications of algebra into matrix representation. This paper described a 3x2 matrix notation for a population as well as its possible operations. The mathematic structure was then applied to sex-linked and autosomal genetic combinations, as well as in-breeding studies.[34] From these applications, Hogben wrote a table of operations to describe recurring expressions in in-breeding.

$$P_n = \begin{vmatrix} \frac{1}{16}(16l + 8m + p) & \frac{1}{8}(4m + p) & \frac{1}{16}p \\ \frac{1}{8}(4m + p) & \frac{1}{4}(2m + 8n + p + 2q) & \frac{1}{8}(4q + p) \\ \frac{1}{16}p & \frac{1}{8}(4q + p) & \frac{1}{16}(p + 8q + 16t) \end{vmatrix}$$

The matrix notation for a population. Reprinted from Hogben's "A Matrix Notation for Mendelian Populations."

At the same time, Haldane, Fisher, and Wright were leading a new emphasis on scientific and mathematical treatments. These three laid down the foundation of a more precise mathematical treatment of populations and genes to create the "classical period of population genetics."[35] Continuing the technical work from before, they made contributions to the studies of inbreeding, mutation, and selection. Haldane did not use any new notation as Shannon did, but instead, evaluated the frequencies of alleles and derived approximate solutions from the formulae they created. Sewall Wright began using statistics to derive relationships between different genetic

---

[34] Hogben. 1933

[35] Interview with Kaiser. 2001. Dr. Kaiser is an Associate Professor of Biology at MIT.

factors. The papers written by the trio had a much more direct application in biology, drawing from experimental data.[36]

$$\bar{q} = \int_0^1 q\phi(q)dq = \tfrac{1}{3}$$

$$\sigma_q^2 = \int_0^1 (q - \bar{q})^2\phi(q)dq = \frac{2}{9(6N + 1)} \cdot$$

Calculations in finding the mean and variance of allele distributions.

Reprinted from Wright's "Self-Sterility Alleles."

It was these same geneticists who began to oppose the mainline eugenics. People such as Hogben, who valued scientific integrity, began pointing out flaws in eugenic reasoning. Haldane, who would also contribute much to the mathematics of genetics, was a vocal critic, remarking that "Da Vinci ... would have been sterilized in some American states because of certain abnormalities."[37]

Genetics in early 1940's

By the end of the 1930s, mainstream eugenics was no longer popular. The questionable methods of many eugenicists had been brought to light. Vannevar Bush, upon becoming the president of Carnegie Institutions of Washington, closed the Eugenics Record Office in 1940. Finally, the death blow to the field came when it was revealed that the Nazis used it to justify their genocide.

Vannevar Bush, Claude Shannon, and genetics

It was these recent events in the deterioration of mainstream eugenics and the rise of mathematical applications that likely inspired Vannevar Bush to suggest genetics to Shannon. As mentioned before, with only the most basic knowledge of the field, Shannon submitted a manuscript of algebraic notations for approval.

---

[36] Kimura. 1968

Bush was so impressed by “the directness with which he obtained some rather general relationships”[38] that he asked the advice of colleagues on the importance of the material. Among the people he sought out was Barbara Burks of the Eugenics Record Office at Cold Spring Harbor. Dr. Burks had done previous research on the study of intelligence and so was fascinated by such a novel approach to the problem of genetics and the opportunity to read the manuscript, commenting that “surely Shannon is gifted -- perhaps to a very high degree.”[39] She noted some problems in the paper -- much of the material had already been discovered, but it was generally agreed that the paper ought to be published since the methodology was unique. These comments were relayed to Shannon by way of Bush, who wrote that "The results that you obtain are most them old but the manner of arriving at them is apparently new". Bush concluded the letter with the advice to look into genetics and a strong endorsement to immediate publication.[40] Although working on several other projects concurrently within the Mathematics department (among them, a paper “on the mathematics of switching circuits”, the problem of “distortion” in communications, and his main project, “the machine for performing symbolic mathematical operations”), he decided to revise and later with Bush's help, to publish the material.[41]

The association with Dr. Burks turned into a summer of research for Shannon at the Cold Spring Harbor genetics laboratory. There, he refined his ideas being "able to generalize the method to take

---

[37] Ibid. pg. 147
[38] Kimura. 1968.
[39] Correspondence from Burks to Bush. January 10, 1938.
[40] Correspondence from Bush to Shannon. January 27, 1939.
[41] Correspondence from Shannon to Bush. February 16, 1939.

care of practically all important cases, including linked factors and multiple alleles."[42] The result of this work was his doctoral thesis, titled "An Algebra for Theoretical Genetics".

Shannon's Ph.D.

His thesis summarized his new method of depicting genetic information with mathematical structures. Specifically, he used a Greek letter notation with subscripts to represent Mendelian populations and described various operations that could be performed on them. With this basic framework in place, he proceeded to prove known theories about genetics with mathematical operations. As with his master's thesis, he produced this work in virtual isolation, referring only to two other works, those of Robbins and Hogben, which dealt with simpler mathematical applications to genetics. Dr. Burks had sent him references to Haldane's work through Bush, but Shannon deemed it irrelevant to his studies. In the doctoral paper, he mentioned that "Inasmuch as no work has been done previously along the specific algebraic lines indicated in this thesis, our references must be of a fairly general nature."[43]

Shannon's ideas flowed along similar lines to the current research in population genetics. He was using mathematics to study how different allele combinations propagated through several generations of breeding. However, the manner in which he did it was new. An interesting analysis of his thesis by modern geneticists pointed out Theorem 12 as being a unique idea, which was not discovered independently for five to ten years[44]. This was an extension of Theorem 11, which Shannon himself felt was the most important piece. "It is the first time that a general expression has been found for the offspring of a population after n generations, under a random mating system

---

[42] Correspondence from Shannon to Bush. December 18, 1939.

[43] The remainder of his references were to basic genetics texts.

[44] Sloane and Wyner. 1993. Questions were put to Professors Crow and Nagylaki, "experts in population genetics".

considering two linked linked [sic] factors." Theorem 11 covered the case for two alleles while Theorem 12 projected results for three.[45] To Shannon, this was trivial mathematically, while to geneticists, this extension was original.

Algebraic details

In "An Algebra for Theoretical Genetics", the symbol

$$\lambda^{hi}_{jk}$$

Basic notation for a population.

represented a population with two gene loci, the pairs of hj and ik. Addition of two separate populations was equivalent to simply combining them, while cross multiplication was defined as the child population of the two parents. These definitions formed the basis of his algebraic system.

$$\begin{aligned}
\nu^{hij}_{klm} &= \lambda^{hij}_{klm} \times \mu^{hij}_{klm} \\
&= \tfrac{1}{2}\left[p_{00}\lambda^{hij}_{\cdot\cdot\cdot} + p_{01}\lambda^{hi\cdot}_{\cdot\cdot j} + p_{10}\lambda^{h\cdot\cdot}_{\cdot ij} + p_{11}\lambda^{h\cdot j}_{\cdot i\cdot}\right] \\
&\cdot \left[p_{00}\mu^{klm}_{\cdot\cdot\cdot} + p_{01}\mu^{kl\cdot}_{\cdot\cdot m} + p_{10}\mu^{k\cdot\cdot}_{\cdot lm} + p_{11}\mu^{k\cdot m}_{\cdot l\cdot}\right] \\
&+ \tfrac{1}{2}\left[p_{00}\lambda^{klm}_{\cdot\cdot\cdot} + p_{01}\lambda^{kl\cdot}_{\cdot\cdot m} + p_{10}\lambda^{k\cdot\cdot}_{\cdot lm} + p_{11}\lambda^{k\cdot m}_{\cdot l\cdot}\right] \\
&\cdot \left[p_{00}\mu^{hij}_{\cdot\cdot\cdot} + p_{01}\mu^{hi\cdot}_{\cdot\cdot j} + p_{10}\mu^{h\cdot\cdot}_{\cdot ij} + p_{11}\mu^{h\cdot j}_{\cdot i\cdot}\right]
\end{aligned} \tag{11}$$

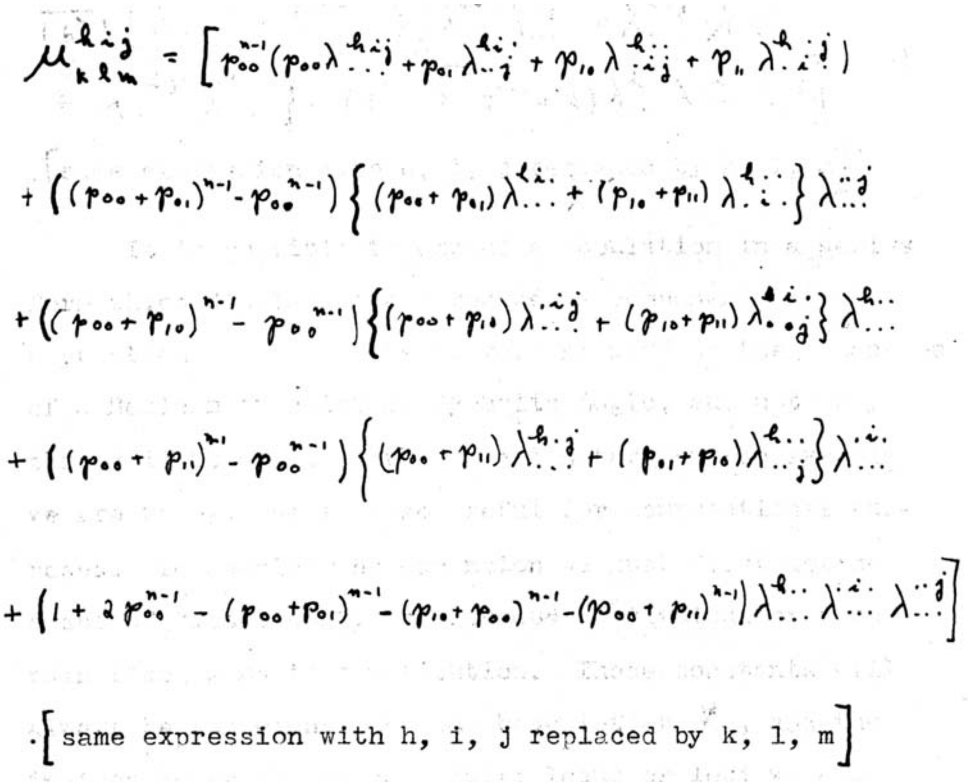
Equation 11 in Shannon's Ph.D.

He used it in Theorem 12 to describe the results of finding the genetic makeup of an nth generation of inbreeding a parent population with three gene loci. This theorem was a result of an equation he had derived earlier in the paper which gave the result of crossing two populations with three gene

---

[45] Correspondence from Shannon to Bush. March 8, 1940.

loci. Shannon then extended it to find the expression as n approached infinity. By using mathematical induction methods, showing that the theorem was true for the first generation, then extending it to the n and the (n+1) generations, he proved his theorem:



$$\mu^{hij}_{klm} = \Big[p_{00}^{n-1}\left(p_{00}\lambda^{hij}_{\cdot\cdot\cdot} + p_{01}\lambda^{hi\cdot}_{\cdot\cdot j} + p_{10}\lambda^{h\cdot\cdot}_{\cdot ij} + p_{11}\lambda^{h\cdot j}_{\cdot i\cdot}\right)$$

$$+ \left((p_{00}+p_{01})^{n-1} - p_{00}^{n-1}\right)\left\{(p_{00}+p_{01})\lambda^{hi\cdot}_{\cdot\cdot\cdot} + (p_{10}+p_{11})\lambda^{h\cdot\cdot}_{\cdot i\cdot}\right\}\lambda^{\cdot\cdot j}_{\cdot\cdot\cdot}$$

$$+ \left((p_{00}+p_{10})^{n-1} - p_{00}^{n-1}\right)\left\{(p_{00}+p_{10})\lambda^{\cdot ij}_{\cdot\cdot\cdot} + (p_{10}+p_{11})\lambda^{\cdot i\cdot}_{\cdot\cdot j}\right\}\lambda^{h\cdot\cdot}_{\cdot\cdot\cdot}$$

$$+ \left((p_{00}+p_{11})^{n-1} - p_{00}^{n-1}\right)\left\{(p_{00}+p_{11})\lambda^{h\cdot j}_{\cdot\cdot\cdot} + (p_{01}+p_{10})\lambda^{h\cdot\cdot}_{\cdot\cdot j}\right\}\lambda^{\cdot i\cdot}_{\cdot\cdot\cdot}$$

$$+ \left(1 + 2p_{00}^{n-1} - (p_{00}+p_{01})^{n-1} - (p_{10}+p_{00})^{n-1} - (p_{00}+p_{11})^{n-1}\right)\lambda^{h\cdot\cdot}_{\cdot\cdot\cdot}\,\lambda^{\cdot i\cdot}_{\cdot\cdot\cdot}\,\lambda^{\cdot\cdot j}_{\cdot\cdot\cdot}\Big]$$

$$\cdot\Big[\text{same expression with h, i, j replaced by k, l, m}\Big]$$

Theorem 12 in the Ph.D. paper. The above three images are from Shannon's Ph.D. thesis, stored in the Institute Archives at MIT.

In the expanding of his population formula into a series form, he noted an interesting parallel to his previous work, saying "This series is very similar to the expansion of a Boolean function in Symbolic Logic, and not only throws light on the mathematical nature of the symbols we are using, but is also useful for computational purposes."[46]

---

[46] Correspondence from Shannon to Bush. March 8, 1940..

Analysis and Comparison

Although Shannon did derive a new theorem, the real innovation in his work was the unique notation. Instead of conducting research like the geneticists, who would do experiments then derive appropriate formulae, Shannon abstracted the common knowledge to achieve a basic set of representation. Here, he showed again his interest in simplifying and reducing real-life phenomena to discrete mathematical terms. He had evolved the research until it was "generalized from the first ideas I had on the subject" to include other cases, such as linked genes and multiple alleles.[47] In addition, Shannon drew from his prior mathematical techniques in Boolean switching to arrive at his thesis. With this genetics paper, he expanded his applications into different areas, which would later affect his other work. Roch, in his evaluation of the thesis, argued that it also let him deepen his knowledge of statistics and algebra, which would again aid in his future research.

A Dead End

Although everyone who knew of the thesis regarded it as original and exciting, it never made any contribution to the field of genetics and remained unknown. Many factors contributed to this result.

First, the paper itself was never published in a journal. This fact was rather puzzling since Vannevar Bush and Barbara Burks were adamant in their insistence on publishing. Bush wrote to Shannon after the first draft of the manuscript, "This, I feel strongly, should be polished a bit and then published. I will be very glad indeed, if you wish me to do so, to transmit it for publication for you with my indorsement [sic], and to recommend some place where it might properly appear" and further admonished "I do not believe you ought to wait to examine further problems before making

---

[47] Correspondence from Shannon to Bush. March 8, 1940.

a publication."[48] Bush even contacted the mathematics department from his new position at Carnegie Institutions "to be sure that he has careful guidance just at the present time."[49] Burks had also suggested the periodical *Genetics* as a good place to begin such an endeavor.[50] Before he began work at Cold Spring Harbor, Shannon had every intention of publishing.

Upon the completion of the doctoral paper, Bush sent a copy to several colleagues. Again, the response was very positive. Reed thought that the paper was "a very good one" and that the notation with its operations gave "a very suitable mechanism for dealing with fairly complex hereditary problems."[51] Dunn, a statistician in government, also praised the paper highly.[52]

Despite all of the encouragement, the publication did not occur. Somewhere between working on his Ph.D. and his acceptance to the National Research Council fellowship, Shannon lost interest in further pursuing the matter and neglected the paper. The reasons behind this were both personal and professional. At the time, Shannon had just married his first wife, and was very busy settling into his new home. Between the relationship and the rush to complete his Ph.D. requirements, he most likely had no time to arrange for publication. After the finish of his doctoral degree, Shannon had indicated a desire to work on something new for his fellowship, rather than continuing old research. In the mathematics department, he had been working on several other interesting problems concurrently. This included his ideas on lens focusing and on "distortion" in communication, which he would keep working on for the next few years to arrive at information

---

[48] Correspondence from Bush to Shannon. January 27, 1939.
[49] Correspondence from Bush to Burks. January 27, 1939.
[50] Correspondence from Burks to Bush. January 30, 1939.
[51] Correspondence from Reed to Dunn. April 9, 1940. Reed was part of the Department of Biochemistry at Johns Hopkins.
[52] Correspondence from Dunn to Bush. April 19, 1940. Dunn was an M.D. and the Chief Statistician of Vital Statistics in the census bureau. He had been so impressed by the paper that he had passed a copy to Reed, who better understood mathematics as it applied to genetics.

theory. All of the described events combined to distract him from a finished work and neglect its publication.

Even had Shannon published the paper, it was highly unlikely that the work would be taken up by other researchers. Burks had warned Bush that such an original piece of work required the continued attention of its “creator” since "few scientists are ever able to apply creatively a new and unconventional method furnished by some one else - at least of their own generation."[53] Bush passed this message on to Shannon, "I doubt very much whether your publication will result in further work by others using your method, for there are very few individuals in this general field who would be likely to do so." The environment in genetics was drastically different from the other working places Shannon found himself in. When Shannon worked on the Differential Analyzer, his theories had much relevance to electrical engineering and the design of circuits. He was also surrounded by people who understood the technical aspects as well as the implications of his ideas. The same could be said for his work during the war, and at Bell Labs. In contrast, many geneticists lacked enough appreciation of mathematics without experimentation to be able to propel his work forward, and mathematicians were not interested in the problem of population genetics. His tendency to work in solitude and his noted shyness and modesty only exacerbated the matter.

Finally, other events in the world most likely served to terminate this study in genetics. The Eugenics Record Office (which Dr. Burks, his advisor, belonged to) was in the process of being shut down by Vannevar Bush. That probably introduced much chaos and distracted those who could have encouraged and helped Shannon. This time was also the beginning of World War II, which

---

[53] Burks to Bush, January 20, 1939.

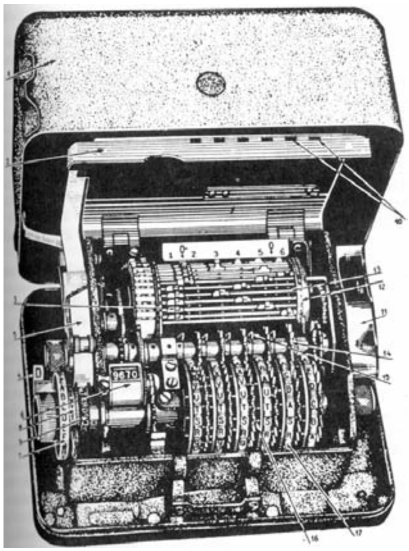
would show the world the depths which eugenics could reach, and which would call Shannon into service as a mathematician in cryptography.

Cryptography

Relation to Information Theory

After his fellowship at the Institute for Advanced Study in Princeton, he went to Bell Labs where the pioneering switching and circuit research was being conducted. The nation's focus had shifted from purely academic efforts to defense research because of the second World War. Like many other researchers in academia, Shannon was recruited to Bell Labs to conduct research that was largely government defense sponsored.

While working at Bell Labs he studied and contributed to fire control processes and data smoothing. The targets of this war were faster then ever before so firing weapons like the M-9 anti aircraft gun needed to more accurately calculate the trajectory of a speeding plane. The advances in technology for the war put an even greater emphasis on communication, and the work that Claude Shannon contributed formally in cryptography at Bell Labs more closely resembles his information theory work than did his fire control and data smoothing research. The war required an unparalleled amount of communication to coordinate multifaceted attacks, and it was vital that those communications were kept secret and that enemy communications were cracked. From World War I to World War II, the number of cryptanalysts grew from 400 to 16,000. Machines like the M-209 were cracking codes at an average rate of 1 for every 8 hours. Key patterns were drawn up by hundreds of privates sitting in Arlington Hall.



The M-209 broke a number of encryption systems at a rate of 1 every 8 hours. The inventor and producer of this device Bruce Hagelin became the first millionaire from the field of cryptography. Reprinted from Kahn's The Codebreakers.

These machines were merely physical extensions of the current cryptography theory. Developing these theories further Claude Shannon said,

> Bell Labs were working on secrecy systems. I'd work on communication systems and I was appointed to some of the committees studying cryptanalytic techniques. The work on both the mathematical theory of communications and the cryptography went forward concurrently from about 1941. I worked on both of them together and I had some of the ideas while working on the other. I wouldn't say that one came before the other--they were so close together you couldn't seperate them.[54]

Information theory and cryptology were very linked and the cryptography work is very much presented in terms of the information theory work. Much of the work was completed for these two documents by 1944 although they were refined until they were published in 1948 and 1949.

A large sector of Claude Shannon's research involved the idea of redundancy in language. An example of this in the English language is that the letter "q" is always followed by a "u" and therefor the "u" is redundant because it contains no additional information. Also, the word "the" is often redundant information and telegrams completely omit the word.

Shannon showed that if all letters in English were pronounceable or usable in any combination then all letter permutations of four letters would be sufficient for 456,976 words (about the number of words in an unabridged dictionary). He realized, however, that if such a language were used it would make error checking difficult because any letter combination would be an acceptable word. Only

---

[54] Kahn, 1967.

context clues within the paragraph or sentence could reveal a word written in error. One or two letters incorrect would create a perfectly usable word.

Shannon realized that redundancy is the major grounds for cryptanalysis. He wrote, "In...the majority of ciphers...it is only the existence of redundancy in the original messages that makes a solution possible." This meant that a message with less redundancy had a cryptogram that was more difficult to crack. Following this theory Shannon insisted on operation with text that had gone through a transducer to remove all redundancies. The most vital part in this process was to eliminate all vowels that did not cause multiple definitions when rebuilding the message. The remains should then be reduced as much as further possible before encryption.

In Kahn's book, *The Codebreakers*, he stated, "Shannon has managed to quantify the amount of material needed to achieve a unique and unambiguous solution when the plaintext has a known degree of redundancy. He calls the number of letters the 'unicity distance'...One of the most interesting uses of the unicity-point formula is in determining the validity of an alleged solution to a cryptogram, especially one of the questionable solutions, such as those claimed to be hidden in the Shakespearean plays to prove that Francis Bacon wrote them. "

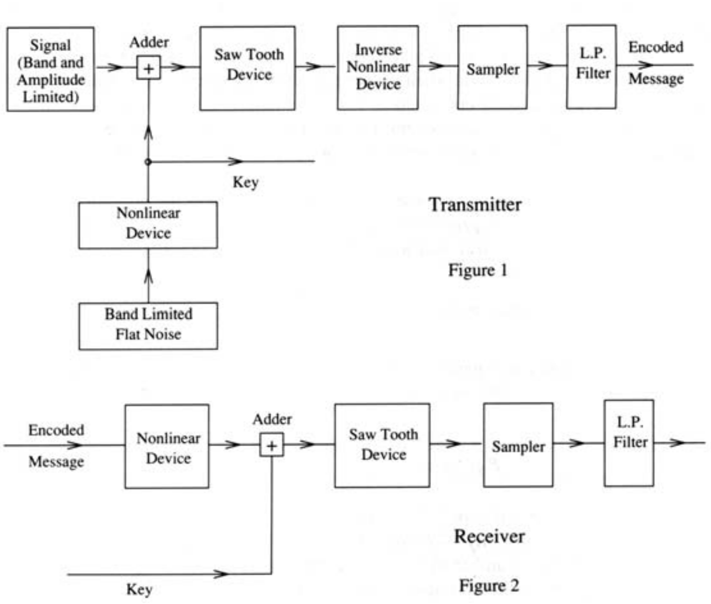
It is also important to note that the Shannon Test had come later in the war. By then, a large number of poor encryption and decryption systems had already been put into place which meant messages were not as safe as they could have been and poorly decrypted messages had been produced. Shannon realized that when someone is trying to decrypt a message, they become engulfed in trying to see a pattern and often will eliminate or alter things so that they can begin to see the message that they are hoping to find. His tests would have eliminated a number of these falsely decoded systems.

Shannon demonstrated how to make cryptograms more difficult to decipher as well as how much encrypted text is needed to decipher it to the original message. Shannon drew the analogy that, " a secrecy sytem is almost identical with a noisy communication system." Although this was a good analogy, Shannon was not limited to this model noting," The chief difference in the two cases...are: first, that the operation of the enciphering transformation is generally of a more complex nature than the perturbing noise in a channel; and, second, the key for a secrecy system is usually chosen from a finite set of possibilities while the noise in a channel is more often continually introduced, in an effect chosen from an infinite set."

Vernam System

Shannon did wartime sponsored cryptography work at Bell Laboratories, and he himself noted that cryptography was a well respected and formal way for him to introduce a number of his information theory concepts. His mathematical understanding of the topic of cryptology allowed him to explain cryptographic systems. The widely used Vernam system was a mathematical abstraction presented by Shannon in a Bell Lab memo. He was able to show why the system was so effective and actually why it was a "perfect secrecy" system.

Perfect secrecy meant that if the enemy has some known knowledge of the intercepted message, it does not help him in any way. The perfect secrecy of the Vernam system is proved by probability arguments and an analogous secrecy system for continuous functions. If one knew that a message was originally English they might know that the letter 'e' is the most common; but when the message is encoded using a perfect secrecy system then that information does not help them to crack the message



From an Analogue of the Vernam System for Continuous Time Series

The reason for this is that a perfect secrecy system does not make any symbol in the encoded message more redundant than any other. Shannon demonstrated his proof by using a probablilistic analysis and stated in his paper, "This theorem is a simple consequence of Bayes' theorem in inverse probablilites, which states that the a posteriori probability of a "cause" A when a "result" B is given by

$$P_B(A) = \frac{P(A)P_A(B)}{P(B)}$$

Bayes' Theorem

where P(A) is the a priori probablility of A, $P_A(B)$ is the probablility of B if A is known to have occurred, and P(B) is the probability of B from any cause. In our case A is any particular uncoded message, of the same number of symbols as the intercepted message." This mathematical analysis of a physical system was a recurrent theme in Shannon's work.

Link to Information Theory

The major contribution of Shannon's cryptographic efforts were refined in his "Communication Theory of Secrecy Systems." The paper introduced a model of secrecy systems stating, "A secrecy system is defined abstractly as a set of transformation of one space(the set of all possible message) into a second set(the set of all possible cryptograms)." Most of the work of this document was conducted at the same time as the work for information theory. Similarities can be drawn between the graphs, illustrations, and mathematical expressions found in documents he wrote.

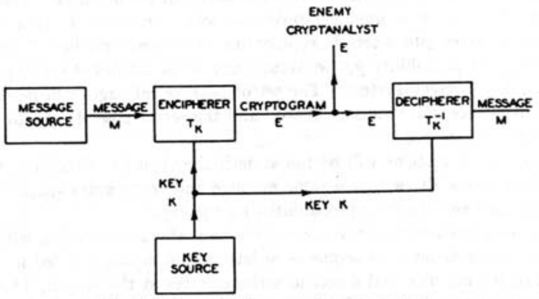


Fig. 1—Schematic of a general secrecy system.

Figure 1 from The Mathematical Theory of Communication

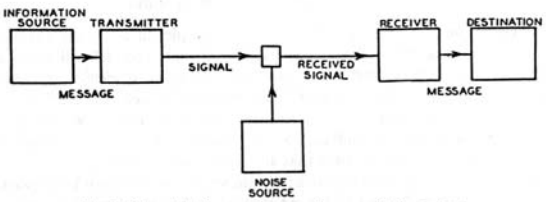
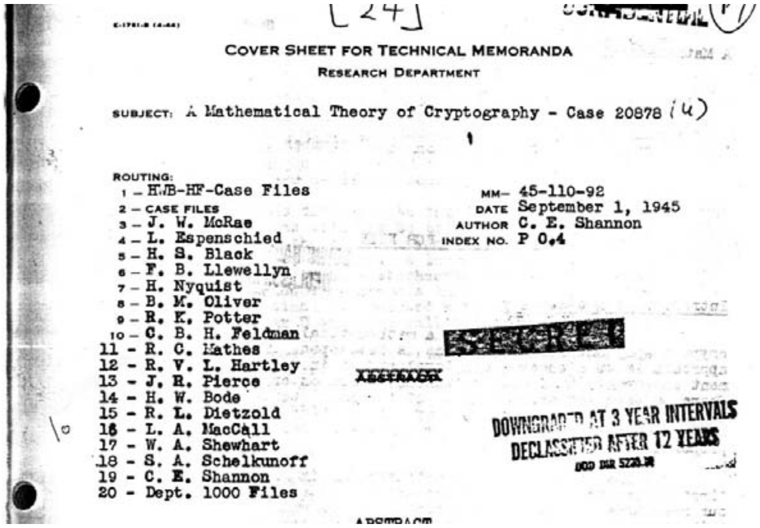


Fig. 1—Schematic diagram of a general communication system.

Figure 1 from Communication Theory of Secrecy Systems

Both diagrams follow the same methodology for generalizing the real world semantic components of the problem. The network diagrams and mathematics can then be applied to each of these boxes in the box and arrow diagram. Once Shannon had worked out the math to the abstract representations of each of these engineering problems, he would reconstruct or explain the system with his new mathematical insights.

One can see that these subjects were inherently linked both in topic and in the analysis patterns used by Dr. Shannon. Both styles of diagrams were not uniquely Shannon's as other Bell Labs memos relating to communication networks contained similar diagrams. Many of the people working with Shannon in cryptography were also involved with information theory. On Shannon's original release of a confidential report "A Mathematical Theory of Cryptography" in 1945, the list of readers includes Black, Nyquist, Bode, and Hartley who are all known as the communications pioneers and this graphic can be seen below.



[24]

[redacted]

COVER SHEET FOR TECHNICAL MEMORANDA
RESEARCH DEPARTMENT

SUBJECT: A Mathematical Theory of Cryptography - Case 20878 (4)

ROUTING:
1 – HWB-HF-Case Files
2 – CASE FILES
3 – J. W. McRae
4 – L. Espenschied
5 – H. S. Black
6 – F. B. Llewellyn
7 – H. Nyquist
8 – B. M. Oliver
9 – R. K. Potter
10 – C. B. H. Feldman
11 - R. C. Mathes
12 - R. V. L. Hartley
13 - J. R. Pierce
14 - H. W. Bode
15 - R. L. Dietzold
16 - L. A. MacColl
17 - W. A. Shewhart
18 - S. A. Schelkunoff
19 - C. E. Shannon
20 - Dept. 1000 Files

MM– 45-110-92
DATE September 1, 1945
AUTHOR C. E. Shannon
INDEX NO. P 0.4

[redacted]

XXXXXXXX

DOWNGRADED AT 3 YEAR INTERVALS
DECLASSIFIED AFTER 12 YEARS
DOD DIR 5200.10

ABSTRACT

Notice Shannon's famous fellow researches that were the readers for his major work in the field of cryptography. The list includes Black, Bode, Hartley, and Nyquist.

Shannon's Style

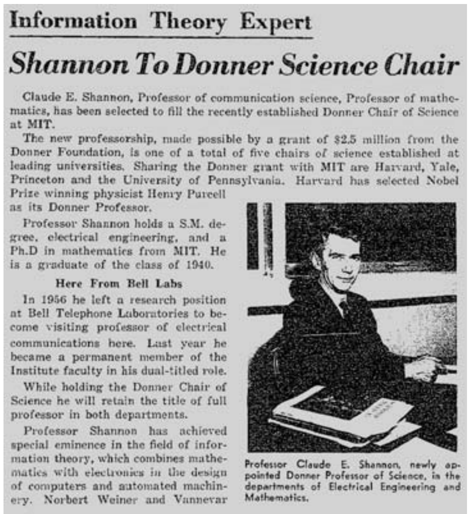
The manner in which Shannon worked and the way in which he interacted with others provide insight into the type of man he was and the motivations behind his work. As a collaborator and advisor, he chose to work alone as much as possible, as this was the best way to focus his thinking. Many who worked with him described having limited but incredibly penetrating interactions which often changed their perspective. This closely resembled his impact on the fields in which he worked, as he went from field to field producing very minimal, but often revolutionary work. His changing interests were also revealed through his non-professional activities.

In his work, Shannon was solely interested in examining these fundamental conceptual ideas, and this was revealed in the way he treated students and colleagues who were looking for his advice. He rarely sought interaction with others and even less frequently sought application or recognition for his work, once again demonstrating his focus on theory. The best way to study this part of Shannon – a step back from his work and towards the actual person he was – is to talk to the people who knew him. Their perspectives did not really relate to a specific section of his work, but instead relate to the broader themes this paper has developed regarding Shannon's work and the nature of engineering revolutions.

Collaboration

"He was just a loner and liked to work alone," said Betty Shannon. "He wouldn't go out of his way to collaborate with other people."[55] "He was not someone who would listen to other people about what to work on," said Fano.[56] His work habits were "not exemplary - he slept when he felt like sleeping," and often spent hours at the kitchen table just pondering ideas, remembers his wife.

Shannon was so focused on the fundamental theories he was considering that he would ask his wife to type his results as he would dictate. As a graduate in mathematics, Mrs. Shannon was also helpful at times with working out the simpler mathematics involved in his work. More also noted that Shannon tended to stay home a lot and receive many visitors, especially Dr. Marvin Minsky. [57] Moving to the house really changed Shannon's life, said More.



Information Theory Expert

Shannon To Donner Science Chair

Claude E. Shannon, Professor of communication science, Professor of mathematics, has been selected to fill the recently established Donner Chair of Science at MIT.

The new professorship, made possible by a grant of $2.5 million from the Donner Foundation, is one of a total of five chairs of science established at leading universities. Sharing the Donner grant with MIT are Harvard, Yale, Princeton and the University of Pennsylvania. Harvard has selected Nobel Prize winning physicist Henry Purcell as its Donner Professor.

Professor Shannon holds a S.M. degree, electrical engineering, and a Ph.D in mathematics from MIT. He is a graduate of the class of 1940.

Here From Bell Labs

In 1956 he left a research position at Bell Telephone Laboratories to become visiting professor of electrical communications here. Last year he became a permanent member of the Institute faculty in his dual-titled role.

While holding the Donner Chair of Science he will retain the title of full professor in both departments.

Professor Shannon has achieved special eminence in the field of information theory, which combines mathematics with electronics in the design of computers and automated machinery. Norbert Weiner and Vannevar

Professor Claude E. Shannon, newly appointed Donner Professor of Science, in the departments of Electrical Engineering and Mathematics.

The Tech, October 7, 1958 – As a professor at MIT, Shannon held a dual-appointment in communication science and mathematics.

When Shannon came to MIT as a professor, his reputation was so great and he was so much in demand that his whole environment changed. "He was really lionized," said More. "He was going to be the luminary that led the electrical engineering department into the future of information theory." Shannon found himself engaged in an environment which was not conducive to careful research thinking. "To create new ideas, you really need to limit the amount of information that comes in," said More. "There are times when you act as a vacuum cleaner and suck everything in, other times you shut everything out and just think." Before he arrived at MIT as a professor, Shannon had embraced the latter approach, but as More states, "it would take a person of almost inhuman discipline" to have continued this in his environment at MIT. This could suggest why Shannon’s most ground-breaking work came so early in his career.

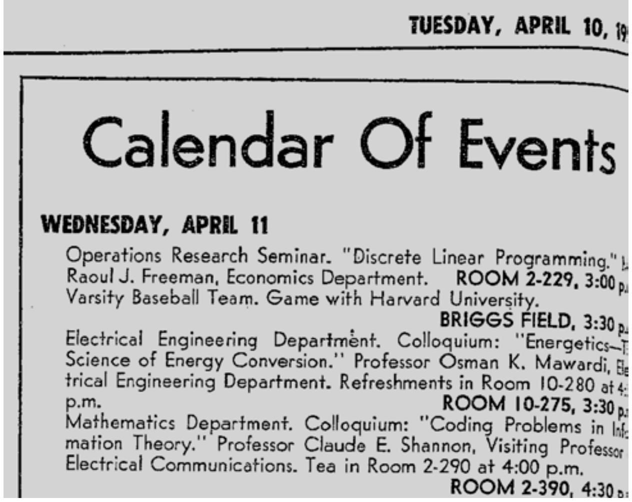
---

[55] Interview with Betty Shannon

[56] Interview with Bob Fano

[57] Interview with Trenchard More

Advising


TUESDAY, APRIL 10,

Calendar Of Events

WEDNESDAY, APRIL 11

Operations Research Seminar. "Discrete Linear Programming."
Raoul J. Freeman, Economics Department. **ROOM 2-229, 3:00**
Varsity Baseball Team. Game with Harvard University.
BRIGGS FIELD, 3:30
Electrical Engineering Department. Colloquium: "Energetics—
Science of Energy Conversion." Professor Osman K. Mawardi,
trical Engineering Department. Refreshments in Room 10-280 at
p.m. **ROOM 10-275, 3:30**
Mathematics Department. Colloquium: "Coding Problems in
mation Theory." Professor Claude E. Shannon, Visiting Professor
Electrical Communications. Tea in Room 2-290 at 4:00 p.m.
ROOM 2-390, 4:30

The Tech, April 10, 1956 – His many seminars and colloquia on information theory were the principle manner in which students were exposed to Shannon, as he did not teach many courses or advise many students. The colloquia motivated one of his few advisees, Henry Ernst, to pursue working with Shannon.

As a professor at MIT, Shannon took on very few students as advisees. His principle contact with students occurred through seminars he would give on information theory. "His talks/lectures were perfect, but he didn't like to lecture. He didn't teach a course regularly at MIT," said Fano.[58] Shannon rarely sought out students, as a couple accounts from his advisees suggest. William Sutherland was one of Shannon's Ph.D. thesis advisees. Sutherland said he came to be Shannon's advisee partially because he knew him from his visit to the Michigan high school and his older brother Ivan was already one of Shannon's advisees.[59] More and Shannon's common interest in mathematical logic led to Shannon serving as an advisor to More's Master's thesis. More was exploring the role of mathematical logic, natural deduction and prepositional calculi in the foundation of computer science, but Shannon ended his role as an advisor to More after the master's thesis, deciding this was not his interest.

Shannon's advising style was consistent with his behavior in other areas. Sutherland said that in order to obtain feedback on his work, he would have to visit Shannon at his Winchester home. While he never had trouble getting in touch with Shannon when he desired, Henry Ernst described their interaction as an "arms-length relationship."[60] According to Ernst, both Shannon and Dr. Minsky were not involved in the specifics of Ernst's research and development of the mechanical

---

[58] Interview with Bob Fano
[59] Interview with William Sutherland
[60] Interview with Henry Ernst

hand, beyond providing the initial inspiration of bringing computers closer to humans. Unlike his relationship with William Sutherland, who described having to travel to Shannon's Wincester home to discuss his research, Shannon never had Ernst travel to his home and was at MIT at least every other day during the years he was there. Despite these differences, both advisee's relationships with Shannon reaffirm his characterization as an independent and remote individual.

It is important to note the distinction between being physically out of touch with the world and lacking the ability to interact with the outside world. The lack of interaction described by the students who worked with Shannon is complemented by stories of tremendous interest in improving the human experience. Shannon was deeply concerned with the fundamental ideas which form our understanding of the world and must have believed that exploration and development of these ideas was only possible through intense individual thought. However, Shannon seemingly did not dismiss the notion of communicating his ideas with the outside world, as he was always described as accessible by his advisees. Ernst even remembers an instance where Shannon spontaneously stepped into an interview with PBS when Ernst was unable to take part. Without preparation, Shannon was able to perform the interview and explain their research to a broad audience. One of the common threads throughout his work and his work style remained his interest in developing mathematical theories relevant to the human experience.

The "great gadgeteer"

Shannon was known for taking up many interests which were completely unrelated to his professional work. Studying Shannon's hobbies and extracurricular interests reveals further insight into his peculiar style and motivation. Betty Shannon described her husband as having a "split personality," because he had an interest in understanding mechanical things and how they operated, but also the very theoretical foundations of science and engineering.[61]

According to Vest, Shannon became interested in engineering as a child when his father would give him erector sets to play with and learn.[62] The erector sets would increase in complexity year after year. His interest in building and gadgetry never waned. Mrs. Shannon remembered fondly the many mechanical interests Shannon pursued at their home in Wincester, MA. "Well he just puttered - he built several machines of course," she said. One machine she particularly remembered was a chair lift which carried people from a back porch along a long run to the lake behind the house. The family also took a trip across the country in a Volkswagon van converted into a camper in 1957, visiting many of the country's national parks. He was always looking for "a challenge of some sort," explained his wife. "We had a very informal house - if there was something that interested us, we did it."

Shannon had a great sense of humor and he loved building things, remembers Fano.[63] A prime example of the combination of these things is seen in "The Ultimate Machine" - when a button was pushed, it would open and a mechanical hand would slowly go towards the button and press it again to close the box.

Shannon was known as a talented juggler, and Vest remembers visiting his house in Wincester, MA and observing the juggling machines Shannon had created. These machines demonstrated Shannon's mindset and passion for integrating mathematical theory into every aspect of the world.

---

[61] Interview with Betty Shannon
[62] Interview with Charles Vest
[63] Interview with Bob Fano

Vest noted that interestingly, over 50% of the members of the American Jugglers Association are mathematicians.

Ernst remembers Shannon spending a considerable amount of time struggling with the theory behind the stock market, concluding that arbitrage was the only true way to make money. Even in his work with the stock market, Shannon worked independently, recalled Ernst.[64]

Sutherland said that upon arrival to Dr. Shannon's residence that he would often find Shannon engrossed in one of his many hobbies and specifically recalled interrupting oboe practice on more than one occasion.[65] Dr. Sutherland had met Dr. Shannon when he spoke at his Michigan high school. He said Dr. Shannon had demonstrated his mouse and maze project to them. In the same vein, Dr. Sutherland had noted Claude Shannon to be a "great gadgeteer" and mentioned Shannon's well known inventions at his Winchester home. Besides having an intense passion for juggling, Sutherland also mentioned Shannon's interest in running; Shannon had even thought of designing a harness to return energy to the legs for running.

Although no technical link could be drawn between his research and his outside hobbies, "the connection was his sense of humor," said Shannon's wife. Shannon was widely known for his sense of humor, which was another example of his ability to look at the world in unique and interesting ways. In fact, the intense interest and creativity Shannon demonstrated in his projects around the house closely resembled the style and diversity of his early groundbreaking research. In both cases, he sought to take fields or environments which were seemingly understood and take them to an entirely new level by introducing a revolutionary concept.

---

[64] Interview with Henry Ernst

Application of his work

Shannon himself said that did not really care whether his research was useful or not. In a 1984 magazine interview, he said, "I am very seldom interested in applications. I am more interested in the elegance of a problem. Is it a good problem, an interesting problem?"[66] He simply worked on what interested him at the time. His wife also remembered this defining element of his style. "Once he was done with something, he was done with it," said Mrs. Shannon.

In the same interview, Shannon was asked about the specifics of anti-jamming and worst possible noise. He replied:

> Bob, I think you impute a little more practical purpose to my thinking that actually exists. My mind wanders around, and I conceive of different things day and night. Like a science-fiction writer, I'm thinking 'What if it were like this', or 'Is there and interesting problem of this type' and I'm not caring whether someone is working on it or not. It's usually just that I like to solve a problem, and I work on these all the time.[67]

"Oddly enough, I don't think he even realized what it turned into," said Betty Shannon about her husband's concept of the impact of his research. "That was the farthest thing from his mind - mostly he didn't pay attention to the outside world."[68] Shannon was not too familiar with the explosion of personal computers and the birth of the digital age, said his wife. The family had computers in the house but by then Shannon had fallen victim to Alzheimer's disease and his interest in technology

---

[65] Interview with William Sutherland

[66] Price, Robert. Interview with IEEE Communication Journal, 1984.

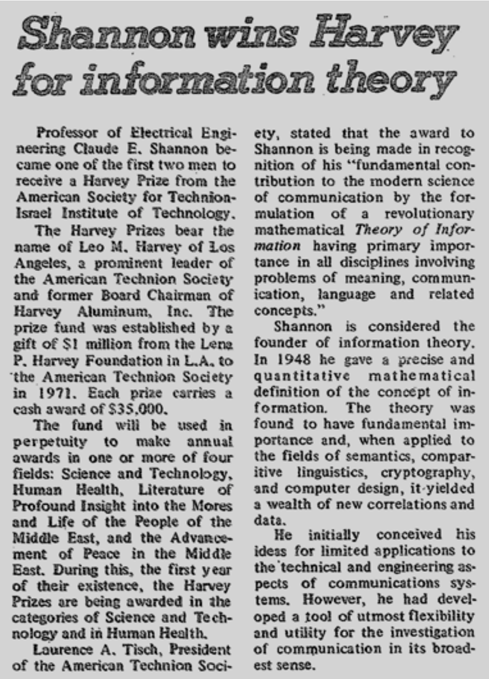
[67] Ibid.

[68] Interview with Betty Shannon

was gone by then. He was home until 1993. If he had been aware of the current state of technology, "he would have been absolutely astounded with computers," said Mrs. Shannon.

Awards and Recognition

Shannon was never particularly interested in the awards or recognition he received. The popularity he had achieved by the time he returned to MIT did not seem to be something he had sought or relished. "He was a very modest guy," said his wife. "He got a lot of awards but they never went to his head and he never talked about them." Mrs. Shannon remains certain the recognition meant something to Shannon, but he simply was not the type of person to make a point of it. Towards the end of his life, "[his eminence] was beginning to dawn on the family a little, but I didn't forsee statues of Claude all over the place," said Mrs. Shannon. "We just didn't think of those terms - *he was just dad in our house*."[69]



Shannon wins Harvey for information theory

Professor of Electrical Engineering Claude E. Shannon became one of the first two men to receive a Harvey Prize from the American Society for Technion-Israel Institute of Technology.

The Harvey Prizes bear the name of Leo M. Harvey of Los Angeles, a prominent leader of the American Technion Society and former Board Chairman of Harvey Aluminum, Inc. The prize fund was established by a gift of $1 million from the Lena P. Harvey Foundation in L.A. to the American Technion Society in 1971. Each prize carries a cash award of $35,000.

The fund will be used in perpetuity to make annual awards in one or more of four fields: Science and Technology, Human Health, Literature of Profound Insight into the Mores and Life of the People of the Middle East, and the Advancement of Peace in the Middle East. During this, the first year of their existence, the Harvey Prizes are being awarded in the categories of Science and Technology and in Human Health.

Laurence A. Tisch, President of the American Technion Society, stated that the award to Shannon is being made in recognition of his "fundamental contribution to the modern science of communication by the formulation of a revolutionary mathematical *Theory of Information* having primary importance in all disciplines involving problems of meaning, communication, language and related concepts."

Shannon is considered the founder of information theory. In 1948 he gave a precise and quantitative mathematical definition of the concept of information. The theory was found to have fundamental importance and, when applied to the fields of semantics, comparitive linguistics, cryptography, and computer design, it yielded a wealth of new correlations and data.

He initially conceived his ideas for limited applications to the technical and engineering aspects of communications systems. However, he had developed a tool of utmost flexibility and utility for the investigation of communication in its broadest sense.

The Tech, October 3, 1972 – The impact of Shannon's landmark 1948 paper continued to be recognized as he won the first Harvey Prize for developing a theory which had "primary importance in all disciplines involving problems of meaning, communication, language and related concepts."

Shannon's low-profile and unwillingness to seek the popular limelight continued throughout his career. In fact, Dr. Vest first met Shannon at an alumni event at the University of Michigan, when the dean of engineering saw a man wearing a nametag reading "Claude Shannon" and wanted to know if it was "*the* Claude Shannon." He sent then-associate dean of engineering Vest to explore,

---

[69] Interview with Betty Shannon

and after some conversation it was determined that this was in fact, "*the* Claude Shannon."[70] Given the extraordinary impact Shannon's research had on the world, his popular familiarity was staggeringly low. This was quite representative of his work style, as he was incredibly devoted to pursuing his mathematical theories, but cared little about the popular aspects of promoting his ideas, gaining worldwide recognition, and pursuing the practical application of his work.

---

[70] Interview with Charles M. Vest

Conclusion

Studying the work of Claude Shannon not only provides an insight into the man himself, but also into the very nature of engineering revolutions. The impact of Shannon's work in certain areas – switching theory and information theory – is widely reported on and acknowledged. However, a closer examination of the context of each of these works, as well as work which did not have the same impact on its field or receive the same attention, shows that there was a much deeper quality to Shannon's work. As his wife described, "When he was working on a theory, he was thinking of things that were beautiful mathematically."[71] Shannon's dual background in mathematics and engineering allowed him to look at the world in a unique manner, searching for the conceptual foundation of every issue he faced. He demonstrated the power of reversing the commonly viewed relationship of engineering as simply an applied science, and instead began with engineering problems and searched for the conceptual scientific core. The essence of Shannon's contribution to the scientific community and to the world was truly his style of work and devotion to abstraction.

Shannon's early work also demonstrates that there are many elements necessary for an engineering revolution. While many view his work as revolutionary, and the catalyst for the creation of multiple fields, there are many factors that contributed to its impact. First, the level of technology in the field must be ready to progress. The reason Shannon's work had such an impact in certain fields and not in others was partly due to the previous work done in each field and the need, ability and interest to move forward at the time of his work. Shannon also showed that while innovative thought is at the core of an engineering revolution, it does not always have to be specifically related to the field, or produced by one with a career of expertise in the field. Shannon provided the fundamental innovative jump by representing circuits and information in new ways, but the environment at the time was what led it to revolutionize the field. The vision for a revolution is the product of many innovators, with the strong example in this case being Vanneuvar Bush. Bush was able to recognize the essence of Shannon's abilities and guide it in a way that would have profound impact on the world. Another key element in an engineering revolution is the willingness of leaders such as Bush to accept change, and members of the field being ready to promote change. Shannon's work was promoted both socially and technologically through the efforts of many of his colleagues. In fact, if Shannon had lived in another time, it is very possible that he would have been hailed as the father of another age.

---

[71] Interview with Betty Shannon

References

Some sources here were originally in German, and translated by the Google online language tools at <http://www.google.com/language_tools>. Edits were made by the authors with the help of the online dictionary at <http://dict.leo.org>.

Correspondences

Burks, Barbara. Correspondences to Vannevar Bush. January 1939. Vannevar Bush Collection, Manuscript Division, Library of Congress.

Bush, Vannevar.
  - Correspondences to Barbara Burks. January 1939. Vannevar Bush Collection, Manuscript Division, Library of Congress.
  - Correspondences to Claude Shannon, December 1938 through July 1940. Library of Congress V. Bush Collection.
  - Correspondences to and from Warren Weaver, January 33 to March 10, 1934. AC233, folder 16 in MIT Archives.
  - Correspondences to E.B.Wilson. June 10, 1932. Bush Manuscript Collection, MIT Institute Archives. AC233, folder 16. December 15, 1938. Vannevar Bush Collection, Manuscript Collection, Library of Congress.

Dunn, Halbert L. Correspondence to Vannevar Bush. April 19, 1940. Vannevar Bush Collection, Manuscript Division, Library of Congress.

Reed, Lowell J. Correspondence to Halbert Dunn. April 9, 1940. Vannevar Bush Collection, Manuscript Division, Library of Congress.

Shannon, Claude E.
  - Correspondences to Vannevar Bush. December 1938 through July 1940. Vannevar Bush Collection, Manuscript Division, Library of Congress.
  - Correspondence to Norbert Wiener. October 13, 1948. MIT Institute Archives. MC22, box 6:folder 85.

Wiener, Norbert. Correspondence to Claude E. Shannon. Undated. MIT Institute Archives. MC22, box 6:folder 85.

Interviews

Ernst, Henry. Interview on November 27, 2001.

Fano, Robert Interview on November 1, 2001.

Gallagher, Robert. Interview on November 20, 2001.

Kaiser, Chris. Interview on November 29, 2001.

More, Trenchard. Interview on November 28, 2001.

Sutherland, William. Interview on November 27, 2001.

Shannon, Elizabeth (Betty). Interview on November 30, 2001.

Vest, Charles M. Interview on November 29, 2001.

Publications

Aspray, William. "Information: A Survey". *Annals of the History of Computing*, Volume 7, 1985.

Astola, Jaakko, Radomir S. Stankovic, Tsutomu Sasao. “Publications in the First Twenty Years of Switching Theory and Logic Design.” Viewed in 2001 at <http://sigwww.cs.tut.fi/TICSP/Publications/report14.htm>.

"AT&T Research Claude Shannon biography." Viewed in 2001 at <http://www.research.att.com/~njas/doc/ces5.html>.

Blackman R.B.; Bode H.W., Shannon C.E. "Data Smoothing and Prediction in Fire-Control Systems." Uncollected Papers of Claude Shannon.

Brumbaugh, Robert. "The Most Mysterious Manuscript". Carbondale: Southern Illinois University Press, 1978.

"Calendar of Events". *The Tech: Newspaper of the Undergraduates of the Massachusetts Institute of Technology.* Vol 76:33, April 10, 1956.

"Graduate Student Wins Noble Prize in Math Research". *The Tech: Newspaper of the Undergraduates of the Massachusetts Institute of Technology.* Vol 60:1, February 6, 1940.

Hagemeyer, Friedrich W. *The Emergence of Information Theory in Communications Engineering.* 1979. Viewed in 2001 at <http://www.radiobremen.de/online/hagemey/hag.htm>.

Hartley, R.V.L. "Transmission of Information". *Bell System Technical Journal*, 1928.

Hogben, L. "A Matrix Notation for Mendelian Populations." *Proceedings of the Royal Society of Edinburgh.* Vol 53, 1933. Pgs 7 - 25.

Image Archive on the American Eugenics Movement. Viewed in 2001 at <http://vector.cshl.org/eugenics/>.

Jennings, H.S. "Numerical Results of Diverse Systems of Breeding". *Genetics.* Vol 1, 1916. Princeton: Princeton University Press. Pg 53.

Kay, Lily. *Who Wrote the Book of Life?* Stanford: Stanford University Press, 2000.

Kevles, Daniel J. *In the Name of Eugenics.* Cambridge: Harvard University Press, 1985, 1995.

Khan, David. *The Codebreakers.* New York: Macmillan, 1967.

Kimura, Motoo. "Haldane's Contributions to the Mathematical Theories of Evolution and Population Genetics". *Haldane and Modern Biology.* Ed. Dronamraju, K.R. Baltimore: Johns Hopkins Press, 1968. Pgs 133 - 140.

Lewand, Robert,"Cryptological Mathematics". Mathematical Society of America, 2000.

Massachusetts Institute of Technology, Graduation Exercises, Class of 1936. 1936.

Massachusetts Institute of Technology, Graduation Exercises, Class of 1937. 1937.

Massachusetts Institute of Technology, Graduation Exercises, Class of 1940. 1940.

"Shannon to Donner Science Chair". *The Tech: Newspaper of the Undergraduates of the Massachusetts Institute of Technology.* Vol 78:33, October 7, 1958.

"Shannon Wins Harvey for Information Theory". *The Tech: Newspaper of the Undergraduates of the Massachusetts Institute of Technology.* Vol 92:36, October 3, 1972.

"Symposium Vannevar Bush at MIT." Viewed in 2001 at <http://www.histech.rwth-aachen.de/www/quellen/bush/photos.htm>.

Waldrop, M. Mitchell. "Claude Shannon: Reluctant Father of the Digital Age". Viewed in 2001 at <http://www.technologyreview.com/magazine/jul01/waldrop.asp>.

Mindell, David.
- *Datum for its Own Annihilation.* MIT Ph.D. thesis for STS, 1996. MIT Institute Archives.
- "MIT Differential Analyzer." Viewed in 2001 at <http://web.mit.edu/mindell/www/analyzer.htm>

Nyquist, H. "Certain Factors Affecting Telegraph Speed". *Bell System Technical Journal* ,1924.

Owens, Larry. *The Text and Context of an Early Computer: Vannevar Bush and the Differential Analyzer.* Princeton, 1984.

Pearson, Karl. "A Mathematical Contribution to the Theory of Evolution, on a generalized theory of alternative inheritance, with special reference to Mendel's Laws". *Philosophical Transactions of the Royal Society of London.* Vol 203 A, September 1904. London: Harrison & Sons. Pgs 53 - 86.

Povarov, Gelliy N. "The First Russian Logic Machines." Viewed in 2001 at <http://www.taswegian.com/TwoHeaded/PartI5.htm>.

Price, Robert. "A Conversation with Claude Shannon". *IEEE Communications Magazine*, May 1984.

Robbins, R.B. "Some Applications of Mathematics to Breeding Problems." *Genetics.* Vol 2, 1917. Princeton: Princeton University Press. Pgs 489 - 504.

Roch, Axel. "Evaluation of Claude Shannon's 'An Algebra of Theoretical Genetics'". Viewed in 2001 at <http://waste.informatik.hu-berlin.de/MTG/archiv/1_roch.htm>.

Shannon, Claude E.

- *An Algebra for Theoretical Genetics*. MIT Ph.D. thesis, Department of Mathematics, 1940. MIT Institute Archives.
- "Analogue of the Vernam System for Continuous Time Series." Memorandum Bell Laboratories, 1943. reprinted in *Claude Elwood Shannon: Collected Papers*. New York: IEEE Press, 1993.
- *"Communication Theory of Secrecy Systems".* Bell System Technical Journal*, July 1948, p. 379; Oct. 1948, p623. Reprinted in* Claude Elwood Shannon: Collected Papers. *New York: IEEE Press, 1993.*
- "Creative Thinking". *Claude Elwood Shannon: Miscellaneous Writings*. Mathematical Sciences Research Center, ATT. 1993.
- "A Mathematical Theory of Communication". *Bell System Technical Journal*, July and October 1948. Reprinted in *Claude Elwood Shannon: Collected Papers*. New York: IEEE Press, 1993.
- *A Symbolic Analysis of Relay and Switching Circuits*. MIT Department of Electrical Engineering M.S. 1938. MIT Institute Archives.

Shannon, Claude E. and Weaver, Warren. *The Mathematical Theory of Communication*. Urbana: The University of Illinois Press. 1949.

Sloane, N.J.A. and Wyner, Aaron D. "Notes to Part C, An Algebra for Theoretical Genetics". *Claude Elwood Shannon: Collected Papers*. New York: IEEE Press, 1993. Pgs 921 - 922.

Sturtevant, A.H. and Beadle G.W. *An Introduction to Genetics*. New York: Dover Publications, Inc., 1962, 1939. Pgs 358 - 365.

Wright, Sewall. "The Distribution of Self-Sterility Alleles in Populations". *Genetics*. Vol 5, 1939. Pgs 538-541.