

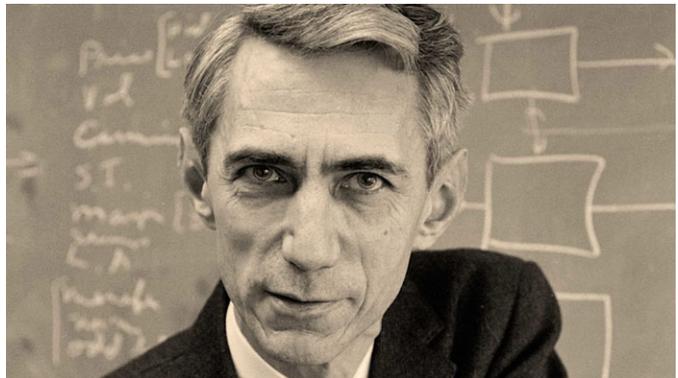
THE UNICYCLING GENIUS WHO INVENTED INFORMATION THEORY: A REVIEW OF *A MIND AT PLAY*, BY JIMMY SONI AND ROB GOODMAN

Laurence B. Siegel
November 28, 2017

Where do technological innovations come from? We have two mental images. One is of a lone genius working in a laboratory or garage, misunderstood until, at long last, the world appreciates her contribution. The other is of a team of busy bees, experts working at a corporation or government agency, the Manhattan Project being the best-known example.

The life of the inventor, mathematician and engineer Claude Shannon merges the two stereotypes. Temperamentally a loner and very much a genius, Shannon was never misunderstood — at an early age he was a protégé of the MIT School of Engineering dean Vannevar Bush — and he became part of the legendary research team at Bell Laboratories. While no one person invented the computer, Shannon's discovery of the parallelism between the zeroes and ones of binary, or Boolean, logic and the on-off status of electronic circuits was the concept that made electronic computers possible. And, because Shannon was an engineer as well as a theoretician, he *built* computers, something that the better-known John von Neumann, Norbert Wiener, and Alan Turing never did.

In *A Mind at Play: How Claude Shannon Invented the Information Age*, Jimmy Soni, best known as the former editor of the *Huffington Post*, and Rob Goodman, a graduate student and political speechwriter, chronicle Shannon's life and scientific achievements. Their style blends traditional biographical information with a considerable amount of scientific content. The book is readable and solidly written, but falls a little short of captivating.



Claude Shannon (photo by Alfred Eisenstadt)

But Shannon's life and personality are so rich a tale that they shine through the mild blandness of the authors' presentation.

EARLY YEARS AND ENGINEERING TOMFOOLERY

A country boy from Michigan who wanted to be an electrical engineer like his distant cousin Thomas Edison, Shannon showed an obvious gift for engineering as a youth: "his creations included a makeshift elevator, a backyard trolley, and a telegraph system that sent coded messages along a barbed-wire fence." He was awkward, looking like he was "always on the verge of being mugged or hit by a bus." Shannon was so self-evidently brilliant that his flight instructor at first declined to teach him, fearing he would crash the plane and die, depriving the world of a first-rate mind.

Shannon could also be sublimely silly. He maintained a fleet of unicycles – including, writes *The New Yorker's* Siobhan Roberts, “one without pedals, one with a square tire, and a particularly confounding unicycle built for two.”¹ He built a robot whose only ability was to turn itself off using a mechanical hand, and invented a flame-throwing trumpet and a rocket-propelled Frisbee. In the authors’ clever phrase, Shannon “worked with levity and played with gravity” — thus the emphasis, in their title, on a mind that was not at work but at play. Despite some distressing moments, it must have been fun to be Shannon, never feeling as though he had done a day of work in his extraordinarily productive life.

THE MAGICAL PARALLELISM OF BITS AND CIRCUITS

When Shannon was 22, he wrote a master’s thesis that would define his career.² “Following a discussion of complex telephone switching circuits with Amos Joel, famed Bell Laboratories expert in the topic,” writes *IEEE Spectrum's* John Horgan,³

Shannon showed how an algebra invented by the British mathematician George Boole in the mid-1800s — which deals with such concepts as “if X or Y happens but not Z , then Q results” — could represent the workings of switches and relays in electronic circuits.

The implications of the paper by the 22-year-old student were profound: Circuit designs could be tested mathematically, before they were built, rather than through tedious trial and error.

Even more important, Shannon demonstrated that the correspondence between Boolean algebraic expressions and electronic circuits is exact, so that if you wanted to construct a machine that could perform operations involving Boolean logic, you could build it out of electronic circuits. That is what a computer is. Consequently, Shannon’s paper has been described as the most influential master’s thesis ever written.

Almost everyone who uses computers now has some sense of the relationship to Boolean algebra, but Shannon discovered it. Boolean logic and electronic circuits had developed along different paths. The 19th century computing pioneers Charles Babbage, who designed — but did not quite build — a mechanical computer called an analytical engine, and Lady Ada Lovelace (Lord Byron’s daughter), who wrote the first algorithmic program for Babbage’s proposed machine, followed the Boolean path. Electronic circuitry emerged from the telephone industry’s need to manage a blizzard

¹ Roberts, Siobhan. 2016. “Claude Shannon, the Father of the Information Age, Turns 1100100.” *The New Yorker* (April 30), <https://www.newyorker.com/tech/elements/claude-shannon-the-father-of-the-information-age-turns-1100100>.

² The thesis, “A Symbolic Analysis of Relay and Switching Circuits,” *Transactions, American Institute of Electrical Engineers*, vol. 57 (1938), is at <http://www.ccapitalia.net/descarga/docs/1938-shannon-analysis-relay-switching-circuits.pdf>.

³ Horgan, John. 2016. “Claude Shannon: Tinkerer, Prankster, and Father of Information Theory.” *IEEE Spectrum* (April 27), <https://spectrum.ieee.org/tech-history/cyberspace/claude-shannon-tinkerer-prankster-and-father-of-information-theory>. Later, John Horgan became a popular writer on science issues and a contributor to *Scientific American*.

of intersecting requests for the system to connect phone calls. The two paths collided in Shannon's brain and in his lab, and we now take "logic circuits" for granted.

GENETICS AS INFORMATION SCIENCE

Just when Shannon seemed poised to become one of the world's leading electronic engineers and theoreticians, his interests took a 90-degree turn: to population genetics, the topic of his Ph.D. studies at Cold Spring Harbor Laboratory on Long Island, New York. Genetics is, of course, the branch of biology that studies how *information* is transmitted reproductively from one organism to another, and the information is digital (conveyed, we now know, by the locations of the four nitrogenous bases — adenine, thymine, guanine, and cytosine — in a DNA molecule). Thus there is a close connection to Shannon's other work.

This change in direction was spurred by Shannon's mentor and thesis advisor, Vannevar Bush, who remarked, "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."⁴ Of Shannon's contribution, James F. Crow writes, in *Genetics*,⁵

The main purpose of [Shannon's] thesis was to develop a genetic algebra. Shannon's formalism was original and quite different from any previous work. The idea was to predict the genetic makeup in future generations of a population starting with arbitrary frequencies.

While this work was original and creative, it remained unpublished and therefore had little influence. Crow concludes,

Shannon went to work at the Bell Labs immediately after receiving his [Ph.D.] degree. There he found a stimulating environment with outstanding engineers, physicists, and mathematicians interested in communication. This got him started on a new career, and genetics was dropped.

THE THEORY OF INFORMATION *INFORMATION IS STOCHASTIC*

So, what is information theory, Shannon's central achievement? Perhaps the best one-sentence explanation is that it is the science that treats information as *stochastic* rather than *deterministic*.

Whether we know it or not, we usually think of information as deterministic: "five troop ships," the classic textbook example of a brief message, contains 14 letters and two spaces — that is, three "words" — in a very particular order. There is no mistaking any of the letters or words for anything other than what they are. And when we transmit the message to another person, through the air in a room, through a wire, or wirelessly using satellites, we expect the message to be received exactly as it was sent.

⁴ <http://math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf>

⁵ Crow, James F. 2001. "Shannon's Brief Foray into Genetics." *Genetics*, vol. 159, no. 3 (November 1), pp. 915-917, <http://www.genetics.org/content/159/3/915>.

But, it turns out, the message is received exactly as it was sent because of the intervention of engineers using information theory. These engineers, following principles first established by Shannon, start by recognizing that some information will inevitably be lost in transmission, whether across a room (as in the children's game of "telephone") or in an electronic communications system. There is just no way around information loss. If you push signals, even simple discrete ones like the dots and dashes of Morse code, through a communications system *at or near the capacity* of the system, some of them will come out wrong. That's what "capacity" means.

So the engineers pack extra information into the message to make up for what they expect will be lost. They make the message *redundant*. MIT's Larry Hardesty explains:⁶

In a noisy channel, the only way to approach zero error is to add some redundancy to a transmission. For instance, if you were trying to transmit a message with only three bits, like 001, you could send it three times: 001001001. If an error crept in, and the receiver received 001011001 instead, she could be reasonably sure that the correct string was 001.

Information theory, then, says that when some of a redundant message is lost, the heart of it is not lost. This is because, as Soni and Goodman tell it, "information is stochastic. It is neither fully unpredictable nor fully determined. It unspools in roughly guessable ways... Whenever we communicate, rules everywhere restrict our freedom to choose the next letter and the next pineapple." The authors, in an excess of cuteness, note that any recipient of the message will figure out that "pineapple" is a transmission error because it makes no sense.

The word that belongs in place of "pineapple" almost has to be "word"; if it's not, we instinctively feel that a rule has been broken. By eliciting the rules of language from large volumes of text, a linguist — or a computer — can take advantage of the partial predictability of language to detect and correct transmission errors.

CAPACITY, BANDWIDTH, AND ERROR-FREE MESSAGE TRANSMISSION

Recall that I earlier referred to the capacity of a system. But how do you measure the capacity of a message-bearing system? This was one of the central concerns of Shannon's information theory. Hardesty continues:

Shannon...showed that any communications channel — a telephone line, a radio band, a fiber-optic cable — could be characterized by two factors: bandwidth and noise. Bandwidth is the range of electronic, optical, or electromagnetic frequencies that can be used to transmit a signal; noise is anything that can disturb that signal.

Given a channel with particular bandwidth and noise characteristics, Shannon showed how to calculate the maximum rate at which data can be sent over it with zero error. He called that rate the channel capacity, but today, it's just as often called the Shannon limit.

⁶ Hardesty, Larry. 2010. "Explained: The Shannon limit." MIT News Office (January 19), <http://news.mit.edu/2010/explained-shannon-0115>.

Recall that Hardesty described adding extra digits to a binary message to help identify and correct mistakes. He explains,

Any such method of adding extra information to a message so that errors can be corrected is referred to as an error-correcting code. The noisier the channel, the more [of this code] you need.... As codes get longer, however, the transmission rate goes down: you need more bits to send the same fundamental message. So the ideal code would minimize the number of extra bits while maximizing the chance of correcting error.

Shannon wrote the rules for solving this optimization problem in any setting. "He was able to prove," Hardesty writes, "that for any communications channel, there must be an error-correcting code that enables transmissions to approach the Shannon limit." Shannon did not, however, "explain how to construct such a code [that worked with certainty]. Instead, it relied on probabilities" — information being stochastic, or subject to a considerable, but not unlimited, amount of randomness. Hardesty concludes,

Say you want to send a single four-bit message over a noisy channel. There are 16 possible four-bit messages. Shannon's proof would assign each of them its own randomly selected code — basically, its own serial number.

Consider the case in which the channel is noisy enough that a four-bit message requires an eight-bit code. The receiver, like the sender, would have a codebook that correlates the 16 possible four-bit messages with 16 eight-bit codes. Since there are 256 possible sequences of eight bits, there are at least 240 that don't appear in the codebook. If the receiver receives one of those 240 sequences, she knows that an error has crept into the data. But of the 16 permitted codes, there's likely to be only one that best fits the received sequence — that differs, say, by only a digit.

Shannon showed that, statistically, if you consider all possible assignments of random codes to messages, there must be at least one that approaches the Shannon limit. The longer the code, the closer you can get: eight-bit codes for four-bit messages wouldn't actually get you very close, but two-thousand-bit codes for thousand-bit messages could.

Thus Shannon outlined the theoretical basis for the methods now used for transmitting information through the phone system, on the Internet, and everywhere else that message senders and recipients rely on getting the message right with very high probability. You pack in extra information, but *not too much* of it — resources are expensive, and you don't get paid for wasting them. The whole design is quite an achievement, economic as well as technological, and it is Shannon's.

IS THE UNIVERSE A COMPUTER?

While Soni and Goodman explain these principles with some skill, at times their reach exceeds their grasp. At the end of the knowledge-packed chapter 16 of *A Mind at Play*, the authors engage in one of their many expositions that remind us they are not scientists but journalists: "[We] reimagine the universe in the image of our tools. We made clocks, and found the world to be clockwork; steam engines, and found the

world to be a machine processing heat.” They conclude by saying that, having invented information networks, we found the world to be one of those, too.

But we didn’t. We built clocks to imitate the way the Earth rotates beneath the Sun. We built steam engines to do work, but we intuit, correctly, that Earth is not a machine built to do work; it just *is*. And no one seriously believes that the Earth or the universe is a computer. Art imitates nature and not the other way around. The authors would be better off sticking to biography, rather than injecting their homespun philosophy of science in a place where it does not help educate.

WORLD WAR II: CODE TALKERS

Like many scientists, Shannon contributed to the war effort on the home front, working on problems of importance to the military. Among the most important problems were encoding messages and breaking the codes used by the enemy, an obvious application of the emerging science of information.

Shannon’s approach to codebreaking was closely related to his insight into accurate message transmission, the former constituting a kind of mirror image of the latter. If the letters in a message arrived truly randomly, redundancy would be zero and codebreaking would be impossible. But, Soni and Goodman write, “our messages are less, much less, than fully uncertain.” Given part of a message, the next part is somewhat predictable — as we saw with “pineapple.” The authors continue, “[Shannon’s] work on information and his work on codes grew from a single source: the unexamined statistical nature of messages.”

What makes code-cracking possible is redundancy, which “means that more symbols are transmitted in a message than are actually needed to bear the information.”⁷ We do this in ordinary messages that are not intended to be stuffed into a wire: “When we write English,” Shannon said, “half of what we write is determined by the structure of the language and half is chosen freely”;⁸ he later revised his estimate of the structural part to 80%. If we see the same patterns over and over, and with known frequencies — U after Q, lots of pairs of E’s, S’s, and T’s and almost no pairs of H’s and Y’s, “of” often followed by “a” or “the” — then we can crack code. And a machine designed to detect subtle statistical properties can help us greatly.

Cracking code, then, resembles the information transmission problem discussed earlier in that both problems involve separating information from noise by understanding their statistical properties. The mirror-image part is that, in one case, the transmitter is trying to maximize the clarity of communication while in the other case he is trying to minimize it — to conceal rather than to reveal.

Cryptography, unlike some other aspects of Shannon’s career, is straightforward enough that a skilled popular science writer can teach a great deal about it, and Soni and Goodman do this successfully. Overall, the book is fairly generous in its detailed,

⁷ Soni and Goodman (p. 151), quoting David Kahn, a historian of cryptography.

⁸ Soni and Goodman, p. 152.

elucidating explanations of scientific concepts, and the section on codebreaking is one of the best parts.

ANECDOTES AND ADVENTURES

That's most of the science in *A Mind at Play*. The rest of the book is a compendium of sometimes brief, sometimes lengthy anecdotes about Shannon's career, personality, colleagues, and adventures. He interacted with many of the other great minds of his time: Norbert Wiener, who thought Shannon's work was following a wrong track (actually both men were pursuing fruitful, but quite different, paths); Marvin Minsky, who suggested the goofy mechanical-hand robot; and even Einstein, who asked him where a cup of tea could be found.

We learn that Shannon built an Erector Set robot that, like Shannon himself, could juggle. He channeled his passion for chess into a design for a chess-playing computer that would influence the builders of Deep Blue a half-century later. He befriended the finance proto-quant Ed Thorp, who introduced him to the joys of gambling and the stock market. Shannon would have made a fine finance quant himself! But the world is probably better off that he chose information technology.

WHY IS CLAUDE SHANNON NOT BETTER KNOWN?

Claude Shannon is a legend among information scientists, computer engineers, and historians of technology. But, despite his immense contribution to the Information Age, he is not well known to the public, even to those who pay attention to technology and science.

Why not? Rob Goodman, in a *Forbes* article, wrote,⁹

Because that's how he wanted it. Shannon certainly earned comparisons to the likes of Turing and Einstein during his lifetime...and when Shannon made a surprise appearance at an information theory conference in 1985, the conference chairman reflected, 'It was as if Newton had showed up at a physics conference.'

But...Shannon consciously stepped away from fame. After the publication of his landmark information theory paper in 1948, he did experience a brief period of notoriety... Yet, at the height of that brief fame, when his information theory had become the buzz-phrase to explain everything from geology to politics to music, Shannon published a four-paragraph article kindly urging the rest of the world to...focus on research...

In other words, Shannon was, at heart, a working engineer, and he was uncomfortable making the leap to professional pontificator, public intellectual, or scientific oracle... Those options simply didn't interest him: he preferred to spend his time tinkering in his two-story workshop, inventing new gadgets...and studying the mathematics of juggling.

⁹ <https://www.forbes.com/sites/quora/2017/07/20/claude-shannon-was-a-genius-on-par-with-einstein-and-turing-why-isnt-he-as-famous/amp/>

Goodman also speculates that Shannon wasn't a tragic enough figure to fit our taste in heroes. He did not have a cruel upbringing like Weiner's, nor was he persecuted by his government like Turing. Shannon had a decent life (marred, at its end, by Alzheimer's). On the second try, he had a satisfying marriage. He succeeded at almost everything he tried.

"The 'trouble' with that," Goodman writes, 'is that it doesn't necessarily lend itself to a tidy narrative of 'genius overcoming the odds'...What we take from Shannon's story is a reminder that creatively fruitful lives can also be joyful ones."

DEPARTURE

Claude Shannon died in 2001. Rob Goodman writes,

[T]here's no better memorial to Shannon than the one he planned himself: later in life, but still in a lucid moment, he sketched out a memorial parade for himself featuring a jazz combo, unicycling pallbearers, juggling machines, a "chess float" atop which a human grandmaster squared off against a computer in a live match, a phalanx of joggers, and a 417-instrument marching band.

The procession never took place, of course. But it tells us a great deal about the person who planned it.



CLAUDE SHANNON AT A HINGE OF HISTORY

Technology is the driving force of our economy, and investors and their advisors would be wise to learn as much about it and its history as they can. Jimmy Soni and Rob Goodman's *A Mind at Play*, while covering material that is somewhat distant both in time and in topic from today's immediate concerns, opens a window into a crucial period in the creation of the Information Age in which we now live. It is skillfully, although not brilliantly, written, and is a good read.

Thomas Cahill, the historian, likes to refer to "hinges of history," events that, on benefit of reflection, made all the difference in bringing about some important aspect of our world. Located at the end of physics' golden age and the beginning of both the Information Age and the Age of Biology, Claude Shannon's remarkable career is one of those hinges.

And his joyous and quirky life is fun to read about.

Laurence B. Siegel is the Gary P. Brinson Director of Research at the CFA Institute Research Foundation and an independent consultant. He may be reached at lbsiegel@uchicago.edu.