Probability, cryptology and meaning in Claude Shannon (1916-2001)'s works

Marie-José Durand-Richard

Probability, cryptology and meaning in Claude Shannon (1916-2001)'s works

DURAND-RICHARD Marie-José
Associated Researcher, Research Unity : SPHERE (UMR 7219 CNRS-Université Denis Diderot Paris)
Correspondence address:
Email : marie-jo.durand-richard@orange.fr

Abstract :
Between 1943 and 1949, the engineer and mathematician Claude Shannon (1916-2001) developped a new theory, founded on the measure of « information quantity ». Clearly, this quantifying design of information was supported by the theory of probability. But it was also fostered by cryptology, which Shannon also worked to innovate in the same period, for the needs of World War II and of its consequences. The first interest of my talk is to make explicit the relationship between informtion theory and cryptology in Shannon's inventive approach.
The reception of this Shannon's work was often accompanied by a specific claim, that is, by introducing prabability theory in this domain, whereby Shannon effectuated a radical break between « information » and « meaning ». Nevertheless, Shannon's texts insisted very systématically on the meaning of his work for the engineer. They rather introduced the idea that meaning depends on the observer's point of view, specifically, that of the engineer who manages the whole communication system. In that way, they ar linked with the 20th century theories of language, especially the pragmatic realm. I inted to examine the epistemological involvements of the reception modalities of this information theory.

Keywords :
Shannon, probability, information, cryptology, entropy, meaning

Introduction

I would like to discuss the sources and impacts of probability theory and cryptology in Shannon's investigation of communication processes as an exemplification of the conceptual transference in inventive thinking of this engineer who was very well-trained in mathematics. Claude Shannon (1916-2001) introduced probability theory in his definition of « amount of information », a theory that he had used prior in physical signal theory for the calculation of error, and – which is much less known – he had also used it in cryptology, where probability and statistics had long been used in cases of finite range. From the tenth century onwards, when algebra was initially being developed within a linguistic context, an approach now called « frequency analysis » was initiated in early Arab trials to decipher secret messages. Over the course of centuries up until the mechanization of cryptanalytical processes as developped in the XXth century, this approach was maintained in philological contexts. And when William Friedmann (1891-1969) introduced the index of coincidence in 1920, probability was still a tool in cryptanalysis. In his own work, Shannon invoked probability as a foundation for developing information as a whole mathematical theory, which he exhibited in a constructive way, so as to link it closely with the specific questions he needed to answer as an engineer at the Bell Telephone Laboratories. In this talk, I will:

• Firstly, explain how I approach Shannon's works from the perspective of an historian of mathematics, and I will discuss the importantce of understanding Shannon's proper location and perspective as it was laid down before and during World War II.

• Secundly, retrace the chronological growth of Shannon's effective works, from his training as an engineer to his papers on the mathematical analysis of the theory of communication and secret systems, as published and jointly conceived in 1948 and 1949 respectively. This will help to apprehend the mutually interactive fostering of these two fields.

• Thirdly, analyze, from this chronology of Shannon's works, ideas and papers, to show how secret systems and the theory of communication worked symbioltically to provide Shannon with the means of establishing theory of information and cryptology as a mathematical « system ».

• fourthly, examine the fate of « meaning » in this new system of accounting for the exchange of messages.

1. My work as the one of an historian of science and mathematics

History of science is often told in a retrospective manner, distilling from the past whatever appears to be most like the present, and wondering why its advancement did not come more quickly. Moreover, wherever mathematics is concerned, particularly in the history of cryptology, historians often forget to observe that mathematics is, itself, in progress. For this reason, the very cumbersome advancement of cryptology before the XIXth century is not only due to the fact that there were difficulties in the transmission of secret practices, but also that, from its earliest days, cryptology was more strongly linked to the analysis of language than to mathematics, just as as logic had long been associated with language as opposed to mathematics until the XIXth century.

Even by the XVIIth century, certain mathematical tools were missing, despite the fact that Marin Mersenne (1588-1648), then « la boîte aux lettres de l'Europe savante », and his close friend Pierre de Fermat (1601-1665), were so fond of analysing language as a combinatorial system, which led them to forge numbers that we now know of by the names Mersenne numbers and Fermat numbers. Today, cryptology is taught at universities as a mathematical subject. This is the case with the Master of Cryptology degree at University Paris 8 where I taught. And Shannon's work was a turning point in the radical transformation of this domain. Until the 1920s and 1930s, the theory of abstract algebraical structures was not really at disposal. When the French and Duchman Auguste Kerckhoffs (1835-1901) first referred to cryptography in terms of « cryptographic systems », and enunciated conditions which are used today to characterize them, there was nothing mathematically new being described. Kerckhoffs was priimarily motivated by the new conditions – the « desiderata » of military cryptology – in which cryptography was evolving with the telegraph: crytpgraphy was no longer concerned with private exchanges between isolated persons, but rather with the various levels of army commandment (Kerckhoffs, 1883, p. 12). When, in 1917, the engineer Gilbert Vernam (1890-1960), who was working on transmission security by teletype for American Telephon and Telegraph Company in Manhattan, invented – what is today known as – the « one-time-pad system », he did not write it with binary digits; he simply doubled the punched tape for the message, written in the Baudot code with sequences of five impulses, by another similar random punched tape for the key (Vernam, 1926). Shannon would later prove the system was unbreakable only after the process was translated into mathematical language using binary digits, that had to wait until the introduction of group theory in cryptology. Only in 1929 did the American Lester Sanders Hill (1890-1961) present his « Cryptography in an Algebraic Alphabet » (Hill, 1929), in which he explicitely considered this alphabet as a group with a law of combination *modulo n*, thereby introducing matrices in a ring and in a field.

Furthermore, it is important to keep in mind that the genius of individual scientists is not a sufficient explanation for why certain scientific achievements come about. Geniuses are often people who are able to generate new syntheses from ideas which are present in their conceptual environments. The history of mathematics overflow with instances in which several individuals thought of similarly new concepts at the same time but in different places. And we have to understand what is at stake when the special work of one of them prevails over the others: in our case, for instance, why did Shannon's 1937 algebraical analysis of relays and switching circuits prevail on that of Paul Ehrenfest (1880-1933) and Vladimir I. Shestakov in Russia, later in URSS, in 1910 and 1935, as well as A. Nakashima in Japan at the same period, as all actors worked on automatic telephone systems.

For my own part, I try to highlight the novelty from the past, avoiding such references as « anticipation » and « precursor ». My view focuses on the effective conditions in which new ideas were produced, at times in which the « images of knowledge » were not necessarily the same as our own. As Leo Corry wrote it about *Modern Algebra and the Rise of Mathematical Structures* (1996):

« the images of knowledge cover both cognitive and normative views of scientists concerning their own discipline » (Corry, 2004, p. 3).

Showing how « the images of science » change implies observing at small steps in history – steps that gives us more of a sense of continuous scientific advancement, with permanent rearrangements between several disciplines, including physics and mathematics since the XVIIth century onwards, as well as engineering since the Industrial Revolution.

I intend to show how Shannon's work drew on contemporary images of information emerging from his mathematical knowledge on probability theory, from his initial domains of research, to theory of communication and cryptology.

A common view of Shannon is that he was an original researcher, juggling in the corridors of the Bell Labs and inventing all sorts of devices, and always working as alone as possible on specific issues. What is important to me here rather is what he inherited from his training as an engineer at the University of Michigan in which he had been immersed during the Interwar period and the organization of research during World War II which helped him to develop new conceptions of information. These two periods created very new conditions for research and development. I shall present Shannon's work in a context in which science, industry and defence became more intricately intertwinded.

2. The chronology of Shannon's involvments as an engineer and a mathematician

Shannon's *Collected Papers* were edited in 1993 ; they expanded to almost one thousand pages. Of course, they involved many precious sources (Sloane & Wyner, 1993). But unfortunately, the *Collected Papers* are ordered according to their historical weight, so they are not faithful to the chronology of Shannon's achievements, and they do not really help to understand the route that his intellectual developments took throughout the 1930s-1940s.

2.1. Shannon's training and early works

Shannon's mathematical formation as an engineer in the 1920s was at least new, if not exceptional. The University of Michigan, where he studied, was precisely one of these state-funded universities that had begun to compete with the more venerable private universities of the colonial era (i. e. Ivy League schools) during the Interwar period, by introducing strong mathematical programs in their engineering curricula. The Massachusetts Institute of Technology, where Shannon would later work on the differential analyzer from 1936 onwards, as researcher assistant, had grown to become an exceptional center for mathematical research, both in applied mathematics with utilitarian objectives, and in pure mathematics, with Norbert Wiener (1894-1964) as a teacher from the 1920s onwards (Parshall & Rowe, 1994, p. 445).

Indeed, the differential analyzer was built at the MIT by Vannevar Bush (1890-1974) and his team between 1927 and 1931 (Bush, 1931; Crank, 1947). It was an analogue engine, intended to graph the curve corresponding to the solution of a differential equation, initial conditions being given. This machine was an essential tool in applied physics as well as in mathematical physics, since a general analytical theory of differential equations was missing. And its operation required the efforts of mathematicians, physicists and engineers. The main principle – an integrator system with a roller rolling on a disc turning on its axis – was involved in simple planimeters since the 19th century. Bush's main contribution was the torque amplifier, which was built to carry out an idea of William Thomson or Lord Kelvin (1824-1907). With the help of electricity in connecting together several integrators, the objective was to solve differential equations by successive approximations (Thomson, 1876).

Essential for the effective operation of the engine were the connections between the integrators, which necessarily involved numerous technical feedbacks, what Shannon named « hidden sneak circuits ». Shannon worked in simplifying the whole organisation of the engine, in order to avoid its numerous jams. His 1936 master's thesis, « A Symbolical Analysis of Relay and Switching Circuits », was defended in 1937. It circulated quickly among engineers (Shannon, 1938), even though it was only published by MIT in 1940, when it received the Alfred Noble award for American engineers. Historians of science have rightly considered that thesis to be an application of Boole's algebra of logic to simplify circuits, by translating them in terms of logical equations, in which the values of the variables were only 0 and 1, and by applying to those logical equations the research of normal forms. But in Shannon's paper, the symbol 0 was assigned to an closed circuit and the symbol 1 to an opened one, which is the dual manner by which we proceed nowadays.

What was crucially new in this master's thesis was that it provided for the first time a common language for both engineers and mathematicians, which will be of utmost importance when designing computers during the World War II. And Shannon's « Mathematical Theory of the Differential Analyzer » in 1941 helps us to remember that his work on this engine must not be considered to be marginal. Very broad hopes were placed on this type of engine, up to the Vannevar Bush's and Samuel Hawk Caldwell (1904-1960)'s project of the huge Rockefeller Differential Analyzer, which was intended to be « the centerpiece of MIT's Center of Analysis » (Owens, 1986, p. 63), and was announced, in spite of the Second World War, as « one of the great scientific instruments of modern times », a one-hundred-

tonne machine with 2000 vacuum tubes and 150 motors. It was, however, made obsolete by the electronic digital computer ENIAC (Electronic Numerator Integrator Analyzer and Computer) by the end of the war.

Thus, from his early works, Shannon showed a deep proficiency for expressing the technical and theoretical concerns in a common mathematical language. Both his previous training and the institutions in which he was working played a major role in shaping Shannon's way of thinking.

2.2. The commitment of Shannon's work in the war effort

The accuracy of the Differential Analyzer was a permanent cause of concern for this community of researchers and practitionners. Shannon was equally concerned with it, and he invoked probabilistic methods to handle it. By 1940, he relied upon these methods when he began to work with Hendrik W. Bode (1905-1982) and Ralph B. Blackman on the automated anti-aircraft gun M 9. This work took place in the Bell Telephon Laboratories from 1940 onwards, and it was invested in the industrial war efforts within the fire control section (D-2) of the NDRC (National Development and Research Committee), directed by the same Vannevar Bush (Segal, 2003, p. 87-106). MIT joined the project from 1941 onwards. By this time, Bell Labs were the hugest private research laboratory in communications in the United States, and even in the world, with 1400 researchers in the 1920s. Mathematics were highly cultivated there too, as Bell Labs were involved in research on automatic control systems and stability analysis.

The Director M 9 was the military version of the Director T-15 first prepared by Bode in the Bell Labs. It was developed for the automatic shooting down of ennemy aircraft. The calculation of the predicted coordinates of the target, in order to optimize the trajectory of the shooting, mobilised a large amount of statistical analysis and probability theory, as did the smoothing out of the data acquired from the target by radars, intended to lessen signal fluctuations and noise effects. The new algorithms introduced by Bode, so as the method of finite differences which he preferred to differentiation for the smoothing of data disturbances, and for calculating the velocity of the target, already reintroduced discrete processes besides continous Fourier analysis usually involved in the differential analyzer as an analogous device. So, the path was opened for Shannon to carry on with his twofold analysis of discrete and continuous processes of communication. The 1946 common Report of Bode, Blackman and Shannon, « Data Smoothing and Prediction in Fire-Control Systems », already viewed this problem as a special case of *transmission, manipulation and use of intelligence*, which referred both to « secret » and « information », and led to a paper by Bode and Shannon (1950). Shannon's later collective manuscript on « The Philosophy of Pulse Code Modulation » (1948) directly inherited this trend of research (Oliver, Pierce & Shannon, 1948).

As for Shannon's knowledge of abstract algebraical structures, he still used them in his 1940 mathematical thesis: « An Algebra for theoretical genetics » (Shannon, 1940 ; Smith, 1982), as well as in his analysis of electrical circuits of the M 9 (Shannon, 1942). If we look to the origins to his symbolical approach of circuits on the model of Boolean algebra, we can see that from the XIXth century, it paved the way towards mathematical operative analogies between several kinds of phenomena. George Boole (1815-1864) belonged to the network of English algebraists which, from Cambridge, developped this strictly Symbolical approach of algebra, looking for the logic of operations in different fields, especially complex numbers and differential operators. Babbage, De Morgan were also among that group. This symbolical approach of algebra was long to find its way towards the study of abstract structures, but, amongst the engineers, Oliver Heaviside (1850-1927) inherited of this way of thinking the resolution of differential equations. So, already at the beginning of World War II, the research of mathematical operative analogies between several communication systems was part of the engineering methodology.

Already in Shannon's 1937 memoir on the mathematical analysis of circuits, we find a move from the survey of the physical phenomena of propagation to that of the structure of the circuits, which could later on be applied to all kinds of networks. In his later papers, this move would come to concern the operative working of the transmitted « intelligence », and its nature for the engineer, which he would express often as: « The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point » (Shannon, 1948, p. 5).

2.3. Shannon's two main contributions during World War II

During Word War II, from 1943 to 1945, Shannon worked in cryptology for the Bell Labs inside the D-2 section of the NDRC. In september 1945, he gave a confidential report, entitled « A Mathematical Theory of Cyptography », which afterwards was declassified, and published in 1949 as « Communication Theory of Secrecy Systems » (Shannon, 1949). Meanwhile, in 1948, he published his famous « Mathematical Theory of Communication », which dealt with an issue that had been of special concern for him since 1939, as his letter to Vannevar Bush on the 16th of February demonstrate (Shannon, 1939). Yet in this letter, although we can already find a rough outline of his famous representation of a channel of communication, there is no trace of probabilities, and the entire text was concerned with continuous analysis.

So the claims made in the two main preceeding papers was stronly entwined, and I would like to illustrate in the fourth part of this paper, how the theory of probability fostered the analysis of the two domains, which were henceforth treated as systems in the mathematical sense.

•3• Probability theory as the core of Shannon's mathematical unifying discourse

As it is well known, Shannon used probability to define what he named « the amount of information » H as a measure of what have to be transmitted on a channel, and also as a mesure of the uncertainty of the occurrence of one message being choosen among a set of messages with given probabilities.

$$H = -\sum_{i=0}^{n} p_i \log_2 p_i \qquad (1)$$

Here, the p_i are the probabilities of occurrences of each possible message, and $\log_2 2$ is chosen as the unit of measure. This definition accompanied what became the classical schema of a communication channel, which Shannon proposed both in his « Mathematical Theory of Communication », and in his « Theory of Secrecy Systems ». With these similar diagrams, the sending of an enciphering message could be identified with an information source, and the whole vocabulary of communications – particularly that of « channel » – could invade the domain of cryptology, just as some concepts in cryptology could be transferred for analysing communication processes..

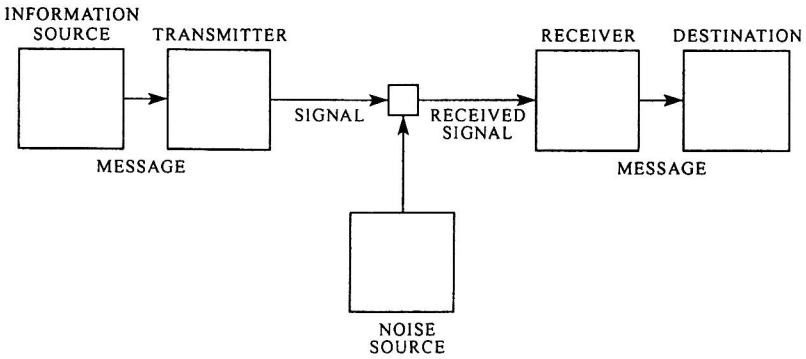

Fig. 1—Schematic diagram of a general communication system.

Fig. 1. Shannon's 1948 diagram of a general communication system (p. 7)

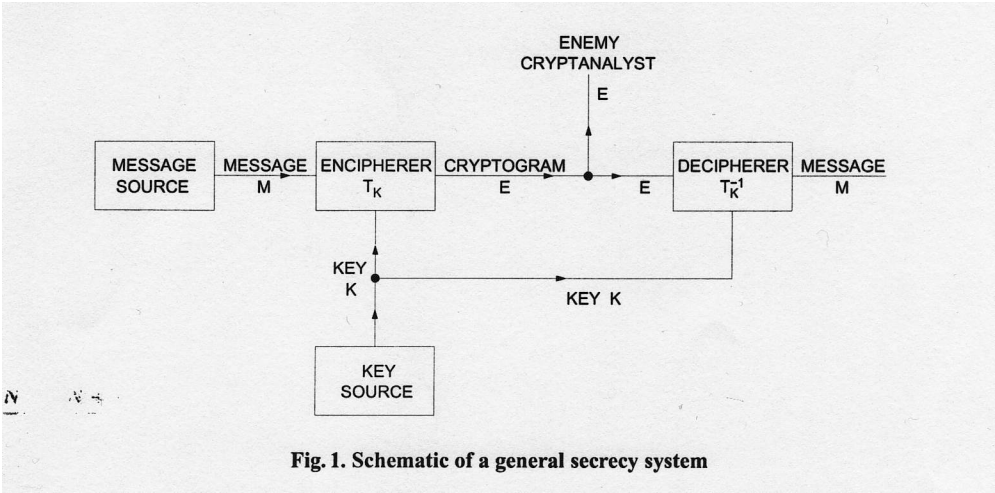

Fig. 1. Schematic of a general secrecy system

Fig. 2. Shannon's 1949 diagram of a general secrecy system (p. 89)

3.1. Shannon's new vocabulary and views

Before him, as Shannon remembered at once in 1948, two other engineers from the Bell Labs, Harry Nyquist (1889-1970) and Ralph V.L. Hartley (1888-1970), looked for such a definition: both of them introduced a logarithmic measure for what was still merely « intelligence », but they did not refer to probabilities. What was new in Shannon's paper was to apply the logarithmic measure for this selection of choices, what he justified from the linear variations of the essential parameters in communication systems.

Throughout these two main papers in 1948 and 1949, the author relied on the more recent mathematical theories. The vocabulary of the theory of sets, the theory of functions and the theory of systems run throughout the two papers. And probabilities were there referred to the modern theory of measure, with the major works of Andrei N. Kolmogoroff (1903-1987), J. L. Doob and Maurice Fréchet (1878-1973) (Shannon, 1949, p. 92 ; Shannon, 1948, p. 15), whose Shannon was also aware of in application to physics and astronomy (Shannon, 1948, p. 11; Changrasekhar, 1943). These concepts sustained the structuration of the two realms. General communication systems were immediately treated by their different states and transitions from one state to another. The enciphering process could be defined as the performance of a family of transformations T_i, where *i* designated the key, giving the cryptogram E by application on a message M:

$$E = T_i M \qquad (2)$$

And inversible transformations were implied in order to define enciphering and deciphering as reciprocal processes. In that way, when these transformations were endomorphisms – as Shannon used this specific vocabulary – secrecy systems were investigated as « linear associative algebras », with all their mathematical properties (Shannon, 1949, *Collected Papers,* p. 88-107). Multiplication and weighted addition no longer concerned these transformations, but rather the whole secrecy systems, defined as « sets of transformations of one space, the set of all possible messages, into another space, the set of all possible cryptograms » (Shannon, 1949, p. 657). Encoding and decoding in communication systems were considered similar processes (Shannon, 1948, p. 25).

Shannon was no longer analysing one signal corresponding to an exchange of messages from one person to another one. Thanks to his definition founded on probability theory, he could consider the whole system in order to ensure the quality of the entire exchange, whatever it could be:

« *The significant aspect is that the actual message is one selected from a set of possible messages. The system must be designated to operate for each possible selection, not just the one which will actually be chosen, since this is unknown at the time of the design* » (Shannon, 1948, p. 5).

Eventually, there are two fields where the practitionner does not necessarily know the « possible selection »: the

receiver of messages from a communication channel (telegraphy, telephone, teletype, radio) ; and the cryptanalyst. In both cases, the practitionner has to restore the hidden message which was previously encoded – or enciphered – and decoded – or deciphered – between emission, transmission and reception, and the engineer has to think, firstly, how these processes ought to be managed.

With the preceding formula, the definition of the amount of information as « choice and uncertainty » takes the same form as that of the entropy of a system in statistical mechanics (Shannon, 1948, p. 18 ; Tolman, 1938). It helps Shannon to consider as mathematically similar:

• the situation of the engineer, facing the « noise » which disturbs the quality of the transmission, which Shannon considered for the first time as a chance variable,

• and the one of the cryptanalyst, facing the cryptogram, and trying to recover the initial message hidden under the encoded one.

Moreover, Shannon could use all the properties of entropy elaborated by statistical mechanics to explore those of the « amount of information »: in his 1948 paper, he treated the entropy of a communication system, and that of an information source in this way; in the 1949 paper, he treated that of the set of messages and that of the set of keys. Entropy was also used to define the redundancy of a language, which helped to determine the maximum compression possible when the alphabet is encoded (Shannon, 1948, p. 24). In both cases, the statistical structure of language was at the core of the analysis, each message being viewed as a sequence of letters, each one chosen from a given set of letters with given probabilities. Moreover, in the case of natural languages, especially for communication systems, the set of these letters sequences was treated as a particular discrete process of Markov – namely, an ergodic process (Shannon, 1948, p. 15-18). And the whole 1948 paper abounded with examples and data from cryptography, mostly tables with letters and bigrams probabilities.

In both cases, the uncertainty of the situation was characterized in the same mathematical way, by a same quantity, called « equivocation ». It was defined as « the conditional entropy of the transmitted signal when the received signal is known » (Shannon, 1949, p. 113), and was introduced:

• in communication, « to measure the average ambiguity of the received signal » (Shannon, 1949, p. 33):

« *Roughly then, [the equivocation] is the amount of additional information that must be supplied per second at the receiving point to correct the receiver message* » (Shannon, 1948, p. 34)

• in cryptology, « to measure in a statistical way how uncertain is the ennemy of the original message after intercepting a cryptogram », and more precisely « how near the average cryptogram of N letters is to a unique solution » (Shannon, 1949, p. 87). Here too, the equivocation concerned the key as well as the message. For instance, for the key:

$$H_E(K) = H(M) + H(K) - H(E) \qquad (3)$$

where H(M) is the entropy of message, H(K) is the entropy of key – i. e. its *a priori* uncertainty –, H(E) is the *a priori* uncertainty of key, and H_E (K) is the equivocation of key (from the cryptogram E).

This introduction of conditional entropy provided Shannon with the opportunity to specify his parallel treatment of communication and secrecy systems:

« A similar situation arises in communication theory when a transmitted signal is perturbed by noise. It is necessary to set up a suitable measure of the uncertainty of what was actually transmitted knowing only the perturbed version given by the received signal. In *Mathematical Theory of Communication* it was shown that a natural mathematical measure of this uncertainty is the conditional entropy of the transmitted signal when the received signal is known. This conditional entropy was called, for convenience, the aquivocation.

From the point of view of the cryptanalyst, a secrecy system is almost identical with a noisy communication system. The message (transmitted signal) is operated on by a statistical element, the enciphering system, with its statistically chosen key. The result of this operation is the cryptogram (analogous to the perturbed signal) which is available for analysis » (Shannon, 1949, p. 113)

We can see there the operative analogy which led Shannon to consider the « noise » as a chance variable, and to view a secrecy system and a noisy communivation system as mathematically similar.

3.2. The constructive exhibition of secrecy and communication systems

In these two main papers of 1948 and 1949, Shannon developed his mathematical theory in a very constructive way. He systematically started from simple examples to lead the reader to more complex situations. And above all, he always specified first the requested operative needs for each new introduced notion, and could thus assess the mathematical properties as « natural ». So, Shannon caustiously isolated the mathematical conditions of the problem before giving his definitions, and he came back to the real situation after exhibiting the mathematical formulae. From this perspective, Shannon did not only worked as a mathematically minded engineer, he also enforced the methodology and vocabulary of Boole's symbolical approach of algebra, where experiment helped to « suggest » the « laws of combination » that were, in turn, followed by their meaning or « interpretation », as Shannon used Boole's vocabulary. Unlike for Boole, these interpretations were always existing in Shannon's developments, which were always carried by economical requirements of optimizing communication work. This was not the case in Boole's thought, for which mathematical laws of thought were always existent in the mind, even if no interpretation could be found in the external world.

For instance, in 1948, Shannon firstly introduced his definition of the « amount of information » after he gave the requested conditions for such a « measure »: to address parameters important for the engineer, to be linear and mathematically more suitable from the choice of the bit as the unit of information measure. In 1949, the definition of a secret system as a « family of unique reversible transformations » was preceded by the characterisation of a cryptographic system, as a suitably idealised situation for mathematical expression, i. e. separated from physical phenomena. Each time Shannon introduced a new mathematical expression, he indicated how it was suitable to the engineer's intuitive approach of the situation. Each time he introduced a new mathematical concept for information or crytpological theory, he gave many examples and diagrams to illustrate and visualise his subject.

In this constructive way of thinking, we can follow in several ways the manner in which cryptology was closely linked with this constructive introduction of probability in Shannon's work. I shall give there characteristic examples:

• in his 1948 paper, Shannon first characterized a discrete stochatic process by means of what he called « artificial languages » consisting of five letters, and of « approximation of languages », built on probability of letters, more and more close to natural English language. The author gave examples of such approximations from order 0 to order 5, built on the probabilities of letters, digrams, trigrams, and words, in the English language. All these probabilities came from recent crytpological books, referred to in the notes (Shannon, 1948, p. 13 ; Dewey, 1923 ; Pratt, 1939).

• in his 1949 paper, Shannon characterised « the generalized problem of cryptanalysis » as « the calculation of *a posteriori* probabilities ». He carefully distinguished between:
- the *a priori* probability, which comes from the statistical structure of lanugage, considered as an ergodic process, and which is known, for this reason, by the sender of the message as well as by the ennemy or the cryptoanalyst,
- and the *a posteriori* probability, which constitutes the cryptanalyst knowledge of the message and the key after interception. As he symetrically worked from the point of view of the « ennemy » or of the cryptanalyst, he claimed that: « Knowledge » is thus identified with a set of propositions having associated probabilities » (Shannon, 1949, p. 85).

The *a priori* probability and the *a posteriori* probability are linked by Bayes' theorem:

$$P_E(M) = \frac{P(M).P_M(E)}{P(E)} \qquad (4)$$

in which P(M) is the *a priori* probability of message M, P(E) is the probability of obtaining cryptogram E from any cause, $P_M(E)$ is the conditional probability of cryptogram E if message M is chosen, and $P_E(M)$ the *a posteriori* probability of message M if cryptogram E is intercepted » (Shannon, 1949, p. 108).

Through this analysis, the paper goes on to use Bayes' theorem between these two kinds of probabilities in order to specify several species of secret systems, which had to help the optimization of cryptanalysis:

• « pure secrecy », where all keys lead to the same *a posteriori* probabilities (Shannon, 1949, p. 101-106). Shannon there used the properties of a group and showed that such a system can be subdivised in closed subsystems.

• « perfect secret system », when the *a posteriori* probabilities are equal to the *a priori* probabilities, and so, when « intercepting the message has given no information to the cryptanalyst » (Shannon, 1949, p. 107-111). In this case, the entropies of message and key are equal, and the equivocation of key $H_E(K)$ equals the *a priori* uncertainty of key H(K). Such a system is particularly useful for « correspondence between the highest levels of command » (Shannon, 1949, p. 111). It required an infinite amount of key, and can be exemplified by Vernam cipher.

• « ideal system », for which the equivocations of key and message never vanished when the number of intercepted letters increases ; and « strongly ideal system », for which the equivocation of key remains constant at the entropy of key. Such a system prevents fhe enemy to find a unique solution from an intercepted cryptogram. It could fairly replace a perfect secret system, for which the making and transmission of keys was a delicate affair.

3.3. How probability theory supported the optimization of communication problems

Even if Shannon treated cryptology and information theory as a mathematician, he still constantly worked as an engineer, whose main concern it was to optimize the quality of communication or the work of cryptanalysis. It is the reason why, if he followed the way Kerkchoffs characterized a cryptographic system from the birth of telegraph by assuming that the ennemy could know the system itself, he rather referred to Von Neumann and Morgernstern's *Theory of Games* (1947 ; Shannon, 1948, p. 91 & p. 132), and he explicitly identified the cryptanalysis problem with a « zero-sum game ». Then he could use Morgenstern's min-max technics to « maximise the minimum amount of work the enemy must do to break [the cipher] » (Shannon, 1949, p. 132)

The whole of his two papers was organised by a systematic investigation of the better means – the more economic ones in time and work – to treat communication problems: either to render more difficult the attacks on enciphered messages, or to investigate the efficiency of encoding systems, so that to ensure the best transmission in a communication channel according to its capacity. And the whole project was realised by means of probability theory. In both cases, Shannon had to determine the different ways to obtain, from the ambiguous received message, the uniqueness of the possible solutions for the original message.

The major issue for the engineer and the cryptanalyst was to determine the original message, and for the designer of the secrecy system to make this work as difficult as possible. In the pursuit of that goal, Shannon investigated the variations of the equivocation from the variation of the number N of intercepted letters of the message. He named the courve « equivocation characteristics of secrecy systems » and showed that it approached zero when N tends to infinity. Thus, he could define « how much intercepted material is required to obtain a solution to a secrecy system » (Shannon, 1949, p. 88) which is for N the « unicity distance ». It approximates H(K)/D, where D is the redundancy of the language, from which it was possible to try to find a unique solution for the original message from the cryptogram.

So, as we can see, cryptology was strongly involved in Shannon's design of communication theory, first as an heuristic way to view and to present his mathematical theory of communication. With this heuristic function, cryptology played a major role in inducing Shannon to assess that « the discrete case forms a foundation for the continuous [one] » (Shannon, 1949, p. 8). The whole structure and methodology of the paper was built accordingly to this assumption. Shannon did not give an axiomatical foundation, but this assumption was more than just a pedagogical one. It was rather a mathematical methodology for solving problems, from the known to the unknown, treating first the simple cases, and building on them solutions form more complicated situations. In fact, the outset of the 1948 paper was built upon cryptological knowledge of English language, such as the probabilies of letters, bigrams, and so on. However, the continuous case would be treated only in 1949, in his book with William Weaver (1894-1978), entitled *The Mathematical Theory of Communication* (Shannon & Weaver, 1949).

4. The destiny of « meaning » in Shannon's work

Concerning this mutual fostering of both communication and cryptological problems in Shannon's work, I would like to conclude on the meaning issue.

I would insist on the fact that Shannon dit not at all eliminate the issue of meaning in his work, as it is commonly assessed from his 1949 book, where Weaver tried to reintroduce meaning inside the communication process, and afterwards with the latest developments of information theory (Segal, 2003, p. 143-538).

On the contrary, meaning was always an essential issue for Shannon, for instance, as we can see when he expressed his theorems, both in mathematical language and in engineering language. always specifying what mathematical formulae he was implying. Moreover, when Shannon treated of cryptanalysis from Bayes' theorem, he explicitly tackled the difficult epistemological questions connected with subjective probabilities, distinguishing between the logical validity of probability, and its application to physical situations (Shannon, 1949, p. 92). Eventually, Shannon did not say meaning does not matter in information theory, he specified, rather, that « the semantic aspects of communication are irrelevant to the *engineering problem* » (Shannon, 1948, p. 5 ; emphasis mine).

But the issue of meaning was greatly renewed by Shannon's work, as he showed that meaning depends on the location of the enunciation: meaning for Shannon as an engineer was not at all the same as meaning for « Alice and Bob », as his cryptological followers would often referred to, i. e. from inside the system, when the engineer and the organizer of the communication channel were outside.

Nowadays, Shannon's theory of communication is invading the whole field of communication, and seem to do so as if the issue of meaning was not a significant concern. But eventually, it is fully significant in the following ranges:

• firstly, it is crucial for those who organize and control communication systems, as it was for Shannon and the military requirements he adressed in his work during the Second World War;

• secundly, it makes significant that the meaning issue depends of the enunciation location: meaning is not the same for the practitionners inside the system, and for those who build and control it from outside the system. Along with contemporary linguistics theories, Shannon's work contributed to make clear that, for each situation, one has to keep in mind the perspective from which people are speaking in order to generate an appropriate manner of treating what they say.

In mathematics too, especially in applied mathematics, the issue of meaning remained an open one. Shannon wrote in a note of his 1949 paper:

« The word « ennemy » stemming from military applications, is commonly used in cryptographic work to denote anyone who may intercept a cryptogram » (Shannon, 1948, note p. 5).

As cryptology to day is expanding in very broad areas of public life, we have to be cautious about the possible implication of this way of thinking on our perception of the unknown.

Conclusion

Shannon was educated and professionalized both into engineering and mathematics. He first worked on devices where these two domains were closely interwined. In each of his two major contributions on communication theory and on secrecy systems, he used mathematics, and especially modern probability theory, to express in the same way the operative working of the two processes, in an inventive turn of mind in which they mutually enriched one another. I think it is important to underline this point, as the history of cryptology has been somewhat ignored until these past few years (Kahn, 1978). Both the mathematisation of the engineering training, and the closeness of disciplines during World War II, stood as conditions of possibility of such an explicitation.

From this step, mathematics occurred as a unifying discourse, giving a common language for both engineers and academics. As such, this unifying discourse sustained a closer vicinity for the whole scientific achievement of the period, including that of the digital computer. It marked the mathematization of cryptology, which was a radical turning point in its history. With cryptological analysis crossing the quantizing of continuous signal, it made proeminent of the discrete foundation of information theory, as well as its probabilistic approach, where the operative structure and running of the whole system became more important than the specific meaning of each message. This new foundation characterized Shannon's own view of knowledge. It was also a determining step in the theory of probability, which conversely tended to be founded back on information theory in the following next

decades.

After World War II, information theory also invaded whole sections of knowledge theory, which soon became « cognitive sciences ». But we have to keep in mind that, from the perspective of human history, language could not be reduced to a quantitative probabilistic approach, and that, paradoxally, security in peace time requires more complex means of proceeding than security of communication systems in wartime.

Ackowledgments

I am particularly grateful to the organizers of the International Conference togive me the honor of delivering a plenary lecture for this meeting. I would like to thank my colleagues in Paris 8 University, who gave me the occasion to become interested in cryptology, my research team SPHERE who permanently sustain my work, and the association M2REAL for its commitment in investigating the relationships between mathematics, engineering and society, that I estimate as essential to any understanding of our present world. I am also deeply indebted to Josipa Petrunic (doctor in the history of science, Edinburgh University, post-dpctor in University College London) who regularly works to improve my English writings.

A first short investigation of Shannon's work on cryptology was recently delivered at the 2009 Cryptologic History Symposium, organized by the Center for Cryptologic History of the NSA Museum, on the 15th-16th of october 2009 at the John Hopkins University Applied Physics Laboratory, on the topic: « Global Perspectives on Cryptologic History ». It was entiled: « How probability Transofrmed Cryptanalysis in Shannon's investigation on Secrecy Systems ».

References

Bode, Hendrik W., & Shannon, C. E., 1950, « A Simplified Derivation of Linear Least Square Smoothing andPrediction Theory », Decimal Classification: R 150. Reprinted in Shannon's *Collected Papers,* pp. 628-656.

Bush, V., 1931, « The Differential Analyzer. A New Machine for Solving Differential Equations », *Journal of the Franklin Institute,* vol. 212, pp. 447-488.

Corry, Leo, 2004, *Modern Algebra and the Rise of Mathematical Structures,* BaseBoston-Berlin, Birkhäuser Verlag, 2d ed.

Crank, J., *The differential Analyser,* Jondon & New York, Longamns, Green & Co, 1947.

Dewey, G., 1923, *Relative Frequency of English Speech Sounds,* Harvard Un. Press.

Doob, J. L., 1941, « Probability as Measure », *Annals of Mahtmematical Statistics,* vol. 12, pp. 206-214.

Durand-Richard, Marie-José, 2000, « Logic versus algebra : English debates and Boole's mediation », *Anthology on Boole,* (ed.) James Gasser, Kluwer Academic Publishers, Synthese Library, pp. 139-166.

Fréchet, Maurice, 1938, *Méthode des fonctions arbitraires. Théorie des événements en chaîne dans le cas d'un nombre fini d'états possibles*, Paris, Gauthier-Villars.

Hill, Lester S., 1929, « Cryptography in an Algebraic Alphabet », *American Mathematical Monthly*, vol 36, pp. 306-312.

Hill, Lester S., 1931, « Concerning certain Linear Transformation Apparatus of Crytpgraphy », *American Mathematical Monthly*, vol 38, pp. 135-154.

Kahn, David, 1978, *The codebreakers, the Story of Secret Writing.* New York. McMillan Pub.

Kerckhoffs, Auguste, 1883, « La cryptographie militaire », *Journal des sciences militaires,* vol. IX, p. 5-38, & vol. X, févrieir 1883, p. 161-191.

Kolmogorov, Andrei N., 1933, « Grunbegriffe der Wahrscheinlichkeits rechnung », *Ergebniss der Mathematic,* vol. 2, n° 3.

Oliver, M., Pierce, J. M., Shannon, Claude E., 1948, « The Philosophy of Pulse Code Modulation », Decimal classification R 148, Reprinted in *Collected Papers,* pp. 151-159.

Owens, L., 1986, « Vannevar Bush and the Differential Analyzer: the text and context of an Early Computer », *Technology and Culture*, n° 27, 63-95.

Parshall, Karen H., & Rowe, David E., 1994, *The Emergence of the American Mathematical Community, 1876-1900: J.J. Sylvester, F. Klein, and E.H. Moore*, Ed. the American Mathematical Society & the Ld Mathcal Soc., vol. 8. History of Maths, « Changes on the American Scene after 1876 ».

Pratt, Fletcher, 1939, *Secret and Urgent,* Blue Ribbon Books.

Segal, Jérôme, 2003, *Le zéro et le un, histoire de la notion scientifique d'information au 20 !me siècle,* Paris, Syllepse.

Shannon, Claude E., 1938, « A Symbolical Analysis of Relay and Switching Circuits », *Transactions of the American Institute of Electrical Engineering,* vol. 5, pp. 38-40. Reprinted in Shannon *Collected Papers,* pp. 471-495.

Shannon, Claude E., 1939, A letter from Shannon to Vannevar Bush, dated: 16th february 1939. Reprinted in Shannon's *Collected Papers,* pp. 455-456.

Shannon, Claude E., 1940, « An algebra for theoretical genetics », Ph. D. Doctoral mathematical Disseration, Reprinted in *Collected Papers,* pp. 891-920.

Shannon, Claude E., 1942, « The Theory and Design of Linear Differential Equation Machines », Reprinted in *Collected Papers,* pp. 514-570.

Shannon, Claude E., 1948, « A Mathematical Theory of Communication », *The Bell System Technical Journal,* vol. 27, pp. 379-423 et 623-656, july-october 1948. Reprinted in Shannon's *Collected Papers,* pp. 5-82.

Shannon, Claude E., 1946-49, « Communication Theory of Secrecy Systems », *The Bell System Technical Journal,* vol. 28, pp. 656-711. Reprinted in Shannon'*Collected Papers,* pp. 83-143.

Shannon, Claude E., 1993, *Collected Papers,* (eds) Sloane, N. J. A., & Wyner, Aaron D., 1993, *Claude Elwood Shannon Collected Papers,* New York, John Wiley's & Sons. Sponsored by the Institute of Electrical Electronics Engineers Information Theory Society.

Shannon, C.E., & Weaver, W., 1949, *The Mathematical Theory of Communication,* Urbana, Illinois, University of Illinois Press. French translation by Dahan, G., 1952.

(eds) Sloane, N. J. A., & Wyner, Aaron D., 1993, *Claude Elwood Shannon Collected Papers,* New York, John Wiley's & Sons. Sponsored by the Institute of Electrical Electronics Engineers Information Theory Society.

Smith, G.C., 1982, *The Boole-De Morgan Correspondance, 1842-1864,* Oxford, Clarendon Press.

Thomson, William, 1876, « Mechanical Integration of Linear Differential Equations of the Second Order with Variable Coefficients », *Proceedings of the Royal Society,* 24, pp. 269-271. Reprinted in Thomson, William & Tait, Peter Guthrie, 1879, *Treatise on Natural Philosophy,* vol 1, part 1, Cambridge, Cambridge University Press, 2nd ed., pp. 497-499.

Tolman, R. C., 1938, *Principles of Statistical Mechanics,* Oxford Clarnedon Press.

Vernam, Gilbert S., 1926, « Cipher Printing Telegraph Systems For Secret Wire and Radio Telegraphic Communications », *Journal of the American Institue of Electrical Engineers,* Vol 55, p. 109-115.

Von Neumann, J. & Morgenstern, 1947, *The Theory of Games,* Princeton.