Tribute to Shannon

Gérard Battail
École nationale supérieure des télécommunications (retired),
46, rue Barrault, 75634 Paris Cedex 13
Electronic mail : gbattail@club-internet.fr

Introduction

Claude Shannon died on 24 February 2001 of the after-effects of Alzheimer disease. With him, one of the greatest scientific minds of the century, and even of all times, disappears. His work exercised a deep influence, although often misknown, in the communication techniques hence in the world where we are living, as well as in the thoughts of the XX$^{\text{e}}$ century. I shall try, after having evoked the carrier and work of Shannon, to show in what his approach was extraordinarily innovative and also, which is more risky, to bring out promises for the future that this work contains. The use I make of the first person should be understood as intended to claim a deliberate subjectivity. I do not indeed pretend to evoke all facets of Shannon's genius but only those which my experience and my reflection enabled me, I hope, to grasp. Beyond the anecdotes and picturesque details I chose mainly to evoke the creator of *information theory*.

Shannon's papers were collected by N.J.A. Sloane and A.D. Wyner [1]. For convenience, I shall cite Shannon's works by reference to this collection. My main source of historic information, except for the few biographic data contained in [1], is the excellent and monumental doctoral thesis of Jérôme Ségal [2].

His life

Claude Elwood Shannon was born on 30 April 1916 in Petoskey, Michigan, the United States. His father, businessman and for a time Judge of Probate, was a descendant of early New Jersey settlers; his mother, a daughter of German immigrants, was a language teacher and Principal of Gaylord High School in Gaylord, Michigan, where he spent all his childhood. He then admires Edison, a distant cousin of the family, and exhibits ingenuity in tinkering and inventions in mechanics, electricity and radioelectricity. He leaves Gaylord High School in 1932 and enters Michigan University at Ann Arbor. He obtains in 1936 the degree of Bachelor of Science in Electrical Engineering and Bachelor of Science in Mathematics. He then becomes a research assistant in the Department of Electrical Engineering at the Massachusetts Institute of Technology (MIT) near Boston, a part-time position which enables him to continue studying. His master's thesis is devoted to the application of Boolean algebra to relay and switching circuits. It is published in 1938, meets a very great success and is awarded in 1940 the Alfred D. Nobel prize, an award given each year in the United States to an engineer less than 30 (do not confuse ...).

In 1938, he leaves the Department of Electrical Engineering for the Department of Mathematics, at instigation of the vice-chairman of MIT Vannevar Bush (who will become during and after the war a consultant to President Franklin Roosevelt). Bush was an engineer of visionary imagination who invented machines predating the computer but failed by the technology of the time. He was just named as chairman of the Carnegie Institution in Washington, a branch of which was studying genetics (and eugenics, which will be discredited only after the war has revealed the monstrous usage made of it). With Shannon's memoir, the design of switching circuits passed from the status of an art to that of a science, thanks to a mathematical formulation of the problem, and Bush hoped that a similar approach by the same Shannon would be fecund to genetics. Back to MIT after a stay at the genetics laboratory of the Carnegie Institution at Cold Spring Harbor, Shannon wrote, under the supervision of the algebraist Frank L. Hitchcock, his thesis entitled 'An algebra for theoretical genetics'. Incidentally, Shannon's work was examined by Barbara Burks, a psychologist expert in the 'genetics of geniuses' member of the American Eugenics Society. Her diagnosis was devoid of ambiguity: the young Shannon is a genius she compares, in a letter to Bush in 1939, with Pascal re-inventing Euclid's geometry at the age of 12 [2].

Shannon obtains his Ph.D. degree in the Spring of 1940. He spends the summer of the same year at the Bell Telephone Laboratories (Bell Labs) in successfully applying the method of his 1938 memoir to simplifying switching circuits (an important stake in the design of telephonic exchanges). After he worked during the academic year 1940–1941 at the Institute for Advanced Studies in Princeton, under the supervision of Hermann Weyl, he comes back to the Bell Labs in 1941, called to integrate a research team (the main members of which were H.S. Black and H.W. Bode) working on anti-aircraft defence systems: a pressing problem in this war time. The works of this team eventually resulted in perfecting and manufacturing the fire control system M6 which enabled England to limit the damage due to the German missiles V1 and V2, and helped the Allies to get the mastering of the airs, a decisive step towards their victory. The war context is the reason why Shannon worked also as a consultant in cryptography to the National Defense Research Committee (NDRC), created even before the United States entered the war and chaired by Vannevar Bush. For this reason, he had the opportunity to meet several times Alan Turing. It seems that cryptography has been for Shannon a source of inspiration but also mainly a mask, honourable in war time, for the studies he already undertook on communication theory and information: they did not contribute to the war effort and their possible usefulness could not be justified but *a posteriori* [2].

Bell Labs were a very fecund assembly of researchers and engineers in all domains of physics and mathematics. Information theory (and many other works by Shannon, the main ones to be found in [1]) are not the least production of Bell Labs. The invention of the transistor is another one, miraculously complementary to information theory, to which it provided, as well as to computer technique, implementation means which badly lacked in 1948. Shannon remained 15 years at Bell Labs which he left only to get teaching at the MIT.

Reliable witnesses met Shannon in the corridors of Bell Labs riding a unicycle and juggling (with three balls, he said). Beyond the anecdote, this attests the immoderate taste for playing which was characteristic of Shannon's personality, his interest to precarious equilibria and, of course, a nonconformism he dared display. Maybe it was a paradoxical means to protect himself from inquisitive people: Shannon did not open up easily and lived retired. For instance he used to get rid of the journalists who tried to interview him by letting them visit his collection of 'toys'.

Shannon indeed loved play, all plays. Gambling, chess, music (he played clarinet and collected instruments of all kind) and, maybe still more, the sophisticated toys he constructed himself. His deep interest in roulette made him undesirable in casinos. Should we consider financial investments as gambling? Shannon was successful here to the point of making a fortune, which enabled him no longer to financially depend on the Bell Labs. He was an excellent chess player (during a trip in USSR, in 1965, he brilliantly resisted the world champion Mikhail Botvinnik, just missing the draw), which naturally led him to get interested in chess playing machines. His 1950 paper "Programming a computer for playing chess" [1, pp. 637–656] made him a pionneer in this field.

This theoretician of genius was also a great handyman, constructing himself play machines, very diverse gadgets having only in common their almost surrealistic gratuitousness. Here are a few: computing machine entirely operating in Roman numerals, mind-reading machine, cybernetic turtles and mice learning to direct themselves in a maze, cycles with eccentric wheels, cycling and juggling robots, ... The 'ultimate machine' alone deserves to be described: it is a coffin-shaped box with a switch on one face. If turned on, an angry buzz rings out, the lid slowly rises, a hand emerges from beneath and turns off the switch, thus ending what may hardly be called the machines's activity!

Let us go back to Shannon's career. Invited professor at MIT in 1956, he became there a permanent teacher in 1959, supervising doctoral dissertations of researchers many of them made a brilliant carrier in information theory and coding. He formally remained at MIT until 1978, but with a progressively reduced activity. He then retired in a large house near a lake at Winchester (Massachusetts), where he could devote himself to his favourite pastimes. His last papers in the field of information theory and coding were published in 1967, cosigned by R.G. Gallager et E.R. Berlekamp [1, pp. 385–423, 424–454]. One of his last papers published under his sole name, in 1959, "Probability of error for optimal codes in a Gaussian channel", besides being outstanding, maybe contains a key as regards Shannon's behaviour with respect to information theory. It contains several times a word which is unusual in scientific literature: 'tedious'. Shannon actually had to make lengthy and non-fascinating calculations so as to obtain bounds on error probabilities (with the help of his wife Betty he explicitly acknowledges), in contrast with the exaltation which he obviously felt with the discoveries of the begin-

nings. There is no doubt that the fear of boredom was a major motivation of this passionate lover of plays.

I feel nevertheless an impression of mystery as regards the behaviour of this man in front of the continuation of the researches stemming from his own work, which reminds Moses gazing at the Promised Land without entering it. His famous theorem of channel coding stated the existence of codes making the error probability arbitrarily small provided the source entropy is less than the channel capacity, a result the proof of which relied on an extraordinary process which left no hope of an actual implementation: random coding, probably inspired by cryptology and maybe also by Darwin. Why did not Shannon contribute to the search for explicitly defined coding means, as opposed to random, although efficient according to the criteria of information theory? His coauthors of the aforementioned papers, Gallager and Berlekamp, have both been eminent actors of this search which was to look like the Holy Grail quest (and still continues nowadays, although it enjoyed a decisive progress with the invention of turbo codes in 1993 [3, 4]; truly, the way which led to this invention was sinuous and far from the initial directions of this research). Was it the awareness that many efforts had still to be made, that this work could only be a collective, slow and rather boring task? Maybe, too, the reluctance of Shannon to any utilitarian finality, as illustrated by the gratuitousness of the 'toys' he constructed?

His work: information theory

I shall restrict myself to *information theory*, generally considered as Shannon's main contribution. His contributions to other fields are however by no means ignorable, but I do not feel competent to deal with them. Moreover, I do not wish to depart from what I believe the essentials.

It happens that a science originates from a founding text so obscure that it needs the work of many exegetes before it is eventually understood, and then years and efforts in order to exploit the ideas it contains only in germ. Far from being so, Shannon's seminal text ("A mathematical theory of communication", [1, pp. 5–83], issued in July and October 1948 in the prestigious journal of Bell Labs, the *Bell System Technical Journal*) looks like a popularization work. Not only did he discuss with great clarity the premises of this new science, but he developed it so coherently and so completely that he left little to find to his successors. The great ease of his discussion was not devoid of casualness in the mathematical treatment. This way of introducing a new science was disliked by some keepers of the mathematical orthodoxy, especially in the United States. Greater mathematicians, but in the Soviet Union, became enthousiastic of this emerging science despite the cold war. Thus, Khintchin undertook more rigourously proving Shannon's theorems [5]. Kolmogorov, who made no mystery of his admiration for Shannon, later introduced using the concept of 'algorithmic complexity' a variant of information theory which was more complementary than competing with Shannon's [6, 7]: instead of letting the measure of information depend on probability distributions assumed to be known, Kolmogorov introduces information as a basic concept, freeing it from that of probability the philosophical bases of which are rather weak. Its practical usefulness, however, is restricted since the quantities of Kolmogorov's theory can not be computed, at variance with Shannon's information.

The best account I can give of information theory consists of the analysis of its founding text. In the very *introduction* of "A mathematical theory of communication", Shannon discards semantics from the field of his discussion, only considering as relevant the fact that the message to be communicated in just an element chosen in a certain set and that the communication system must work regardless of the chosen message. The problem of communication is thus basically of statistical nature, the information brought by a particular message being in fact measured by the number of messages among which it is chosen. Explicitly referring to Hartley, he shows the interest of using a logarithmic measure, according to the practical, intuitive and mathematical viewpoints. The choice of the logarithmic base determines the unit used for measuring information. If the base is 2, he names this unit the *bit*, an acronym for *binary digit*. Among other possible bases, he considers also 10, directly consistent with decimal numbers, and Euler's constant e, which is more convenient when integrations and derivations have to be performed. As a model of the communication process, he introduces the famous scheme (or paradigm)

$$\textit{source} \text{ — } \textit{channel} \text{ — } \textit{destination}$$

the 'channel' being actually split into the transmitter which generates a signal, a medium subject to noise (which he names 'channel' in a restricted sense) and a receiver. He then distinguishes three kinds of communication systems: the *discrete* ones where both the message and signal are strings of discrete symbols or signals (a typical example of

which is Morse telegraphy); the *continuous* ones where both the message and the signal are dealt with as continuous functions (as in radio and television); and *mixed* ones where both discrete and continuous variables appear (as in pulse code modulation, PCM, transmission of speech).

The *first part* is naturally devoted to discrete noiseless systems, the simplest case. He first considers the noiseless discrete channel the capacity of which he defines in information unit per time unit. It depends on the duration of the symbols and the constraints which determine their succession. He then considers discrete sources of which he gives many examples: 'natural' languages; discrete sources deriving from continuous ones by quantization; discrete sources mathematically defined by stochastic processes which specify the symbol choices and their possible mutual dependence, especially in the form of Markov chains where the probability of a choice only depends on the previous choices through the present state of the system. He shows how the models of discrete source he just introduced can provide a series of statistical approximations to English with an increased fidelity according to whether one takes into account the frequency of letters, digrams, trigrams, ..., or of the frequency of words, couples, word triplets .... He gives examples which look like what will become later the exercises of Oulipo[1]. He graphically represents the Markov chains by transition diagrams which he then assumes to be ergodic (he briefly explains this term). In order to evaluate the average information rate of such a source, he states the axioms which should be satisfied by an information measure in terms of the probabilities which describe it and thus obtains the *entropy* function $H = -\sum p_i \log p_i$. He discusses its main properties and defines redundancy as the difference between the maximum possible entropy and its actual value. He notices that the possibility of cross-word puzzles depends on redundancy of the language. It is the more difficult to construct a grid, the stronger the constraints hence the higher the language redundancy. He also considers coding operations aimed at minimizing the average message length. He states that the lower bound on this minimum length is proportional to the source entropy, which is the fundamental theorem of channel coding, and gives some examples of optimum source coding.

The *second part* deals with noisy discrete systems, i.e., where the channel input/output transition probabilities differ from 0 and 1. He then introduces the quantity compatible with the entropy definition which measures the average information quantity that the output variable provides as regards the input variable (it is called mutual information due to its symmetry), and he defines the channel *capacity* as the maximum of this quantity with respect to all information sources which can be connected to its input. He sets out the fundamental theorem of channel coding, which paradoxically states that errorless communication of a message is possible if a proper code is employed, provided only that the source entropy is less that the channel capacity. He sketches the proof of this theorem based on the extraordinary idea of random coding. Since he cannot exhibit a particular code with good enough error probability, he considers a *probabilistic ensemble of codes*, computes the average error probability for this ensemble and shows that it can be made arbitrarily small by increasing the word length provided the above condition is satisfied. In the considered ensemble of codes, thus, there exists at least a code which is at least as good as the average. Still better, the error probability thus obtained with a peculiar code is almost surely (asymptotically as the word length approaches infinity) close to the average so it vanishes. From this point of view, one can thus say that 'all codes are good'. Shannon comments these results, recognizes that random coding cannot be actually implemented, insists on the role of redundancy in protecting against noise, gives examples of noisy channels and computes their capacity. Finally, he gives an example of coding which turns out to be the (7,4) Hamming code, still unpublished when Shannon's paper brings out.

The *third part* considers the information measure for sets of functions depending on random parameters, especially the set of band-limited functions. Shannon defines entropies simply derived from those of the discrete case by replacing sums by integrals and probabilities by probability density functions (what is now called differential entropies, a term that Shannon does not use). He states the properties of these entropies homologous of the discrete ones, but notices that these quantities now depend on the coordinate system, at variance with the discrete case entropies.

In the *fourth part*, he computes the capacity of a continuous channel. The mutual information is then expressed as a difference of entropies as defined in the third part. At variance which each of the terms in the difference, it remains unchanged

[1] Oulipo (*OUvroir de LIttérature POtentielle*) was a group of writers interested in mathematical games and combinatorics. It was founded by the mathematician François le Lionnais, and its most famous members were Raymond Queneau and Georges Perec.

with respect to a change of coordinates, so the definition of the channel capacity does not change. He considers the case of additive noise, and especially that of Gaussian and white noise which is the usual model of thermal noise, when the average received power is given. Applying his definitions, he gets the capacity per time unit of this channel, C, probably the most famous formula of information theory (if not the best understood), namely:

$$C = W \log \frac{P+N}{N},$$

where W is the bandwidth, P the average received signal power and N that of the additive noise. He mentions that this formula was independently found by other researchers, especially Norbert Wiener and W.G. Tuller. He also considers other cases where he can only give lower and upper bounds of the capacity.

The *fifth and last part* considers the extension of what precedes to a continuous source. The information rate cannot then be defined without introducing a fidelity criterion, two messages close enough for this criterion being considered as equivalent.

I gave a long summary of these papers, especially as regards the introduction and the first two ones, which contain the main innovations and the meaning of which is not made obscure by mathematical difficulties, in order to show the extreme richness of their content. Another paper by Shannon, maybe less famous, remarkably complements the preceding ones: "Communication in the presence of noise" published in 1949 [1, pp. 160–172]. Starting from the sampling theorem (often wrongly ascribed to Shannon, although he refers to previous works) which states that signals the spectrum of which is limited to a frequency band of width W can be exactly recovered from the values (or samples) they assume at time intervals separated by $1/2W$, he introduces a geometric representation of signals and of additive white Gaussian noise as vectors in a high-dimensional Euclidean space. Some unclear properties of analog modulation systems found an obvious explanation in this representation, which also was much later used for the design of systems combining modulation and coding. Shannon uses it to sketch a direct proof of the capacity of the additive white Gaussian noise channel, i.e., to prove that the information rate must be less than the above expression of this channel capacity so that no errors occur. Shannon also discusses his very smart solution, referred to as water-filling, to the problem of communication in the presence of non-white noise, i.e., where the noise spectral density is not a constant in the signal band.

Shannon's influence

Engineers and scientists interested in information theory formed under the banner of the Institute of Radio Engineers (IRE, later become after a merger with another society, IEEE) a working group which started publishing a journal, in 1953, the *IRE* (then *IEEE*) *Transactions on Information Theory*. From a few hundreds of pages for each of the first years, its volume did not cease to increase up to more than 2,000 pages yearly now. This quantitative increase has nevertheless coincided with a narrowing of the field which was covered. The issues of the first years were indeed much more eclectic than they are today. Problems of signal theory, automatics or psycho-plysics found a place in it, while these topics are now relevant to other journals. A reason of this trend is a reaction against a fashion effect, ephemeral by definition, I shall more lengthily deal with. The problem was to avoid that works really useful to information theory be diluted in the flood of papers about more or less relevant applications, and the policy of restricting the scope prevailed, after debates between the members of the working group to which Shannon participated, as we shall see it.

I already mentioned the reluctance initially expressed by certain mathematicians with respect to Shannon's work. On the contrary, it was enthusiastically welcome by many researchers of other fields: genetics, neurology, psycho-physics, psychology, economy, linguistics, sociology ... It was unfortunately an irrational fad, the intensity of which was often matched by the lack of understanding. The vocabulary of information theory then had often a decorative role and papers like "Information theory, photosynthesis and religion" proliferated (this emblematic title, from an editorial of the *IRE Trans. on Information Theory* in which Peter Elias mocked at this fad [8], is of course invented but it is hardly caricatured). Even the indeniable influence that Shannon's work had on first rate linguists and philosophers, like Roman Jakobson or Claude Lévi-Strauss, remained limited to its more superficial aspects. The deepest and most innovative ideas of information theory, especially the possibility of errorless transmission despite the channel perturbations and the extraordinary method of random coding to prove it seem to have escaped any comment by established philosophers.

Shannon himself reacted against the fashion he unwillingly initiated. He thus wrote an editorial in the same journal in March 1956 [1, p. 462], which I shall more lengthily comment when dealing with the

future of information theory. Despite its brevity, it seems indeed to me that it opens up a program much of which remains to be performed.

One said once that the fecundity of a work is measured by the number of misunderstandings it gave rise. In this case, Shannon's work is immense! The obituaries just published in the French newspapers eloquently witness it, if I may write so. They are few, short and, far from helping to know his work, show the misunderstandings it suffers. This also shows how the importance of Shannon's work was underestimated: rarely the futility of the media was so obvious.

Certain of these misunderstandings have as sole origin an erroneous reading. Thus, a myth that nothing (or almost nothing) justifies sees in Shannon the exalter of the binary: the core of his theory would be the possibility of transmitting a message, regardless of its nature, by the means of binary symbols or signals. I wrote 'almost'. Shannon could not imagine how journalists would dress up his work and he imprudently proposed to name 'bit', an acronym for 'binary digit', the unit of information quantity which results of choosing 2 as logarithmic base (he also contemplated other bases, as we have seen it). A digit and a unit are objects of different nature and defining 'bit' as the acronym for *binary unit* could maybe have avoided the misunderstanding. It turns out that 'bit' is usually employed for binary digit in technical jargon, even when it bears no information or an information quantity less than the binary unit. Everybody aware of information theory knows that and distinguishes with no risk of error the two meanings of the word 'bit' (personally, I use the word 'shannon' for the binary unit, which avoids any ambiguity). Hasty and inexpert readers unfortunately fell in the trap which was unwillingly set. I do not know if Shannon has been angry or, more probably, amused at that.

Other misunderstandings have a much deeper origin which it is important to analyse. The role of semantics is a point of major divergence between certain of the authors who tried to apply information theory outside its original domain, and the now unanimous opinion of engineers. The wide development of information theory in the mathematical and technical fields amply justified the exclusion of semantics which is a premise in Shannon's theory. This exclusion actually appears as a *methodological necessity* which enables distinguishing the information from both the message which bears it and the meaning ascribed to it. On the contrary, many people coming from different horizons, especially biologists, have felt when reading Shannon the exclusion of semantics as a congenital defect to be repaired. This misunderstanding is not recent. Under the title "The mathematical theory of communication" (the definite article substituted for the indefinite one cancelled the modesty of the original title), the two 1948 papers were reprinted as a book as early as 1949 [9]. A lengthy postface by the biologist Warren Weaver, then administrator of the Rockefeller foundation, has been appended to them. Shannon claims that he discards semantics in his very introduction, in a few sentences, arguing that the semantic content of a message has no incidence on how the messenger works. On the contrary, most of the comments by Weaver deplore the exclusion of semantics and suggest remedies for it. Rather strangely, the two authors of the book thus express irreconcilable points of view. Time did not attenuate this misunderstanding. I shall try later to analyse the reasons for it at the same time I shall outline perspectives for the future.

The future of Shannon theory

I believe that Shannon theory has a great future outside the technical domain, as applied to nature sciences. I shall in the following restrict myself to uphold this opinion as regards biology. Many engineers do not share this opinion, which moreover contradicts the one which currently prevails among biologists.

Answering his (few) interviewers, Shannon claimed his atheism. He saw no fundamental difference between the machines and the living things, including men. His tinkering was maybe intended to imitate nature (widely anticipating on François Jacob!), in a pathetic way which made intuitively perceptible the distance between the technical means available in the middle of the XX$^{\text{e}}$ century and that of nature, generally endowed by evolution of an extreme refinement. He thus could not be hostile in principle to the idea of applying information theory to biology.

The short editorial of the *IRE Transactions on Information Theory* I mentioned above is entitled 'The bandwagon'. Shannon has mocked there at speculations which referred to his work, calling for patience and modesty. Asserting as a personal opinion the rightfulness of applying his theory to sciences of nature, he suggests however that this approach will be fruitful only after information theory will be firmly enough established in its domain of origin. One cannot but admire how lucid is this editorial. Most of the speculations Shannon denounced

fell indeed into a well deserved oblivion, whereas information theory has confirmed its validity and its fecundity in the mathematical and technical domains; at the same time, the attempts to apply information theory to other sciences became more and more unfrequent. One may deplore this withdrawal into an ivory tower, but hope that the reflection acquired in the technical domain will eventually enable applying it to the sciences of nature freed from the naïvety and vague approximations of the first attempts.

The fairies who leaned over Shannon's cradle (I think of Barbara Burks and Vannevar Bush) as well as Norbert Wiener who had a less direct but inescapable influence on him were fervent advocates of interdisciplinarity. If one defines information theory as the science of symbol strings (with Shannon and Kolmogorov) then it can obviously be applied to biology: Crick and Watson identified in 1953 the DNA molecule as the bearer of the hereditary information, made of a string of quaternary symbols. That attempts aimed at applying information theory to biology failed until now is not a reason to give up (I would like to say: on the contrary). The initial fad having passed and the misunderstanding I mentioned as regards the role of semantics being stronger and stronger, the biologists turned away from information theory. After its too discreet triumph in the technical domain, I think it is now mature for eventually fecund applications to biology and perhaps physics. A mandatory condition for the success of this ambitious plan is a clear awareness of the origin of the misunderstandings which as yet hindered it.

The comments by Weaver gave me the first example of a misunderstanding between information theorists and biologists which got worse with time. Weaver worries about the congenital inability of information theory to take semantics into account. Shannon accepts it, on the contrary. All the subsequent development of information theory has shown he was right, since the exclusion of semantics never appeared as a drawback or a brake. Information plays with respect to semantics the role of a *container*, and should not be confused with its symbolic supports, i.e., messages, and still not with the physical supports of the messages. Much more than mastering the mathematical difficulties of some of its chapters (but the discrete finite case, the most important one, does not suffer such difficulties), it is the understanding of the status of information as an intermediate which is the key of its fruitful application. Information is indeed abstracted from the set of supports and messages which can bear it, but it is also the bearer of a meaning which is completely independent of it and not amenable to a quantitative measure. The difficulty of information theory is thus not so much intrinsic than conceptual, insofar as it is the epistemological status of the main quantity it deals with which is far from obvious. At the turning point between the abstract and the concrete, information revealed itself as an unexpected intermediate. This status is now well perceived by the engineers who have learned by experience that 'it works', but not at all by the upholders of other disciplines, especially physicists and biologists.

This reflection shows how deeply innovative Shannon theory is. With the discovery of a measurable quantity as fundamental as hidden, it is a new world that it opened to science.

References

[1] Claude Elwood Shannon (N.J.A. Sloane et A.D. Wyner, eds), *Collected papers*, IEEE Press, 1993.

[2] Jérôme Ségal, *Théorie de l'information : sciences, techniques et société de la seconde guerre mondiale à l'aube du XXI^{e} siècle*, thèse de Doctorat, Faculté d'Histoire de l'Université Lyon II, defended in December 1998.

[3] C. Berrou, A. Glavieux and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding : turbo-codes," *Proc. of ICC'93*, Geneva, Switzerland, pp. 1064–1070, 23–26 May 1993.

[4] C. Berrou and A. Glavieux, "Near optimum error-correcting coding and decoding: turbo-codes", *IEEE Trans. Com.,* Vol. 44, No. 10, pp. 1261–1271, Oct. 1996.

[5] A.I. Khinchin, "Mathematical foundations of information theory", Dover, 1957.

[6] A.N. Kolmogorov, "Three approaches to the quantitative definition of information", *Problems of Information Transmission*, Vol. 1, pp. 4–7, 1965.

[7] A.N. Kolmogorov, "Logical basis for information theory and probability theory", *IEEE Trans. on Inf. Th.,* Vol. IT-14, No. 5, pp. 662–664, Sep. 1968.

[8] P. Elias, "Two famous papers", *IRE Trans. on Information Theory*, Vol. 4, No. 3, p. 99, Sep. 1958.

[9] C.E. Shannon and W. Weaver, “The mathematical theory of communication”, University of Illinois Press, 1949.