Ubiquity Symposium

What is Computation?

What is Information?
Beyond the Jungle of Information Theories
by Paolo Rocchi
IBM and Luiss University

Editor's Introduction

This fourteenth piece is inspired by a question left over from the Ubiquity Symposium *entitled* What is Computation?

Peter J. Denning
Editor

Ubiquity Symposium

What is Computation?

What is Information?
Beyond the Jungle of Information Theories
by Paolo Rocchi
IBM and Luiss University

Computing saw the light as a branch of mathematics in the '40s, and progressively revealed ever new aspects [gol97]. Nowadays even laymen have become aware of the broad assortment of functions achieved by systems, and the prismatic nature of computing challenges thinkers who explore the various topics that substantiate computer science [mul98].

What Might Information Be?

The dualism analog/digital reveals one of the most intriguing facets of computing that draws experts' attention since years. The contrasting properties of analog and digital computing have been debated in various contexts. Actually computation and information are themes strictly related as "computations process information" [hub90]; and even the odd nature of analog and digital symbols attract authors. Computer experts and humanists attempt to discern the essential differences emerging between the two technologies.

A circle of commentators agrees that analog information appears somewhat close to Nature, the analog is defined by its always having a relation to things. Blachowicz [bla98] and others claim that the analog communication has a necessary relation to what it represents. Instead discretization produces representations "less real" or "less authentic" than analog forms. This interpretation sounds somewhat reasonable and one could be inclined to sanction the analog/digital paradigms on the basis of the natural/artificial criterion in a definitive way. But this rigid conclusion raises some doubts.

Digital devices handle discrete signals rather alien to natural forms found in the world but this is not an absolute feature: there are digital solutions very close to Nature. Take the digital sound that consists of a finite series of ones and zeros. A converter transmutes a continuous wave into a stream of discrete signals and ensures the realism of the digital sequence. The sampling

theorem regulates the conversion between continuous and numeric signals; and sanctions the symmetry extant between the two representations. As a practical case, digital signals generated at 96 kHz with the resolution of 24 bits make a high-definition format which represents sounds much more accurately than a hi-fi vinyl record.

Analog describes any fluctuating or evolving event, while digital uses pulses that are separated values to represent events. The vast majority of technical writers are inclined to take '*analog information'* as synonymous with a continuously variable value and '*digital information'* as synonymous with discrete values [rao09]. The criterion continuous/discrete has many fathers. This idea seems to be self-evident and infiltrated the scientific community since decades. But some cases bring this criterion into question. Take the Frequency-Shift Keying (FSK), a modulation pattern in which digital information is transmitted through frequency changes of a carrier wave. The signal consists of high-frequency wave-blocks and slow-frequency wave-blocks. All the wave-blocks have the same duration and the final outcome is a continuous wave that creates a bit stream. One may wonder: Are the FSK waves to be classified as digital since they deliver numbers? Or otherwise are they analog due to the continuous change of values?

These and other queries direct us to the definition of information. The appropriate notion of information could be a significant key to decipher the analog and the digital technologies; however this definition proved to be a troublesome issue.

We have learned that ancient Greek philosophers were the pioneers who investigated the domain of information in the western hemisphere. But men of letters were involved with information in a broad sense and explored topics close to their cultural stances. They argued over logic, language and related matters such as grammar, interpretation, philology, lexis and semantics. Psychologists, sociologists and others added their comments on human communications over the centuries.

The radio engineer R.V.L. Hartley was the first to develop a theory of information that was not centered on humans. In his 1928 paper, he introduced the word '*information'* into the engineering environment through an original method of calculus. "Transmission of Information" by Hartley circulated in small circles of experts but did not enjoy wide discussion. Two decades later the paper "A Mathematical Theory of Communication" by C. Shannon attracted vivid attention of researchers from many different environments. Information became an argument of study that raises ample debates in the scientific community so far as now. Through a bibliographical inquiry we have inventoried over twenty-five information theories proposed during the last six decades [roc10a]. About twenty per cent of theories attempt to perfect and extend Shannon's paper. The remaining works propose alternative views on information. The high number of contributions could appear as a sign of vitality but

sometimes the clashing opinions of writers surprise the reader. The intellectual domain looks to be rather confusing and disappointing to an observer.

Carnap's view on information revolves around on semantics; in contrast Shannon deliberately ignores the aspects of semantics. Kolmogorov reasons at the pure technical level, whereas Maturana and Varela aim at unifying the view of the mind with the external world. Norbert Wiener rejects the idea that information is physical, and Tom Stonier sees information as much a part of the physical universe as energy and matter. Whilst to Shannon information is inversely proportional to probability, to Wiener it is directly proportional to probability: the one is simply the negative of the other.

Theorists even do not agree on the nature of the issue; one group sees information as a quantity to measure—for instance Shannon, Fisher and Klir—whereas others —see Floridi—are convinced of the manifold constitution of information. The former are prone to attack the problem using analytical methods; the latter reject any analytical approach and claim that pure philosophy can enlighten the argument. There are odd positions even inside the circle of Shannon followers. Marschak and Miller deem the master's frame insufficient and attempt to complete and to enrich it. Others—such as Petitjean and Leydesdorff—debate philosophical subject matters though their master intended only to calculate the capacity of channels and not to develop philosophical arguments.

In sum authors have not come into accord on a common definition. They still quarrel over the method of study, the theoretical assumptions, the application ranges, and the scopes of their inquiries. Also the terminology seems to obstruct communication amongst experts, as W. Neville Holmes notes in [hol01]. Broad divergences prevent authors from having constructive dialogues. They often deny any validity to their opponents and rarely accept confrontation. Lack of cooperation and discussion averts researchers from progressing toward the abstract and shared interpretation of information.

Two Popular Ideas

Despite the substantial dissension on the appropriate definition of information, the sorry state of theoretical investigations did not deter us to go ahead and to investigate what is analog and digital.

This effort paid the dividends and we made an interesting discovery.

The vast majority of experts share two basic ideas on information despite their otherwise diverging conceptualizations. Almost all practitioners and scholars are inclined to believe that (a) *a piece of information has physical origin* and (b) *stands for something*. They agree that a sign has a body and this body symbolizes something else that can either be material or abstract. Because this pair of ideas emerged in linguistics we shall call the body of information as *signifier*, and the represented entity as *signified* [cow77]. Sometimes commentators popularize this pair of notions using the terms *form* and *content*.

We have verified in literature how dozens of disciplines employ the material properties of signifiers to great advantage. We have found examples in medicine, neurology, physics, zoology, linguistic and jurisprudence. Electronic engineers devised several signifiers and may be considered the inventors of the most advanced forms of communication.

The pair signifier/signified has infiltrated numerous fields of research and is usually gained through intuition. Shannon, for example, recognizes that signals conveyed in a channel are physical quantities and carry a message, represented by a code that utilizes the signals of the medium. Shannon reminds that in the Morse Code, a short electrical impulse - a dot - symbolizes the alphabet letter E; three dashes and a dot make the letter Q, and so forth. Physical quantities become symbols for something else. Shannon holds that semantic aspects of communication are irrelevant to his theory, but provides an account of what a signifier and a signified are.

Authors accept (a) and (b) on the basis of their direct experiences. Engineers normally design and build systems such as radio equipment and computers, without having to make any special assumptions because the signifier and the signified are easy ideas; it is enough to use these popular notions in intuitive terms. The broad acceptance of (a) and (b) understates the success of electronic appliances that have been established on the basis of progressive, pragmatic discovery of effective, charming signifiers.

The ideas of signifier and signified perceived through intuition are sufficient to make optimal equipment, however naïf ideas are not enough to clarify a tricky topic such as the analog/digital argument. The formal definitions for (a) and (b) appear necessary to introduce the discussion.

Toward Formalization

Writers from various fields [ges97] [hir92] [pel90] acknowledge that a sign comes into existence when an observer can distinguish sign's physical body. A sign disappears as soon as the observer cannot perceive the sign's material base. For example you can talk to a friend

provided the sound of your voice – namely the signifier – is discernible. When the words are feeble or the noise covers them, then your friend cannot discern your voice and the message vanishes. Music, texts, pictures, and images must all be sharp or otherwise they vanish and convey no information. The lack of distinction results in the disappearance of a simple signifier and even of compound signifiers. We represent this special feature as the *property of sharpness*:

(1) E NOT=R E^*

That is to say, the entity E is a signifier if E differs from any adjacent entity E^* with respect to the observer R. The verb "differs" is represented by the symbol 'NOT=' ('not equal'). If E does not contrast in a way, E does not exist as a signifier in the physical reality.

The property expressed by (b) connects E to NE, namely the signifier E symbolizes the object or concept NE. There is no tangible, physical connection between E and NE, and we formalize this as the *semantic property*:

(2) $E \longrightarrow NE$

Both the properties (1) and (2) are not absolute concepts; they are relative to an observer. Philosophers make a lot of discussion on the intervention of R. They argue on the intellectual capabilities of humans who perceive the objects and recognize their meanings. People learn how to identify an item and realize it from past experiences; they relate sense observations to their own culture and assign the meaning to a sign in accordance to individual cognition. Thinkers develop a lot of comments instead engineers focus on more straightforward notions. Technicians basically have concern for the physical detection (1) and the linear association with a meaning (2). The observer R may be a living being, a machine, or an electric component in engineering. It is not necessary that the observer be a human brain; it could be a mechanical or electrical device with no self awareness at all. Thus, we restrict our attention to objective representations and to objective perception processes and we shall disregard philosophical arguments.

Digital Paradigm versus Analog Paradigm

A method to separate the digital from the analog. Almost any physical object is a potential signifier in accord with the mathematical rule (1). Any body $E,$ which contrasts with an adjacent body E^* respect to the observer, can be classified as a signifier. Thus all the elements of the universe are potential signifiers as long they are perceived trough a sensory organ or a tool. A chair is a signifier, and even the floor, a flower, and a mountain.

Physical entities that have not been built up for the specific purpose of informing—such as the mountain—will be called *spontaneous* or *natural signifiers*. Physical entities that have been built to designate information will be called *artificial signifiers*. They are related to the natural in that engineers often use spontaneous signifiers to build up an appliance. For example, a heap of sand falls down from a higher level to the ground by the action of gravity. This natural effect has been imported to create sand-glasses. The sand that flows from the upper glass bottle into the lower bottle is the signifier that marks the passing time. The needle that moves along the scale of an instrument is another example of E that is analogue of a finger pointing at something. Signifiers brought in from the environment are usually named *analog signals* as they are similar to natural entities. Electrical signals that replicate sounds, music etc. are perhaps the most popular analog signals in ICT (information and communication technology).

Analog experts are attracted by spontaneous signifiers while sharpness and semantics seem to play a secondary role for them. The properties (1) and (2) are not a worry for analog designers and signals often exhibit various defects. Usually E is the continuous function of time $E(t)$, and has a certain margin of error $|\Delta E(t)|>0$. Frequently $E(t)$ denotes the signified $NE(t)$ in a generic way, namely the signifier itself induces the significance NE. Consider the case of a signal whose voltage signifies a number. Suppose $E(tk)=4.0$ volt at the instant tk has $\Delta E(tk)=\pm0.2$ volt and represents the number four. This implies that any signal within the range 3.8 volt and 4.2 volt will be interpreted as the signifier for the number four due to the fuzziness of $E(tk)$. It is evident how the physical and semantic performances of $E(t)$ are below the expected standard.

One could say that digital experts take the opposite direction. They want to rise above the uncertainties of Nature by using separate signifiers. Hardware designers place the physical values at a certain distance in order to make improper discernment extremely improbable. Discrete signifiers are distinct even under the worst conditions and do not resemble any spontaneous item. This is a noteworthy virtue as no physical restriction compels or influences the meanings of discrete items. Bits can symbolize sounds, pictures, texts etc. Digital experts do not explicitly use expressions (1) or (2), but behave as if they observed them scrupulously. The bits E and $E*$ are perfect pieces of information according to the sharpness property and to the semantic property above introduced.

We are inclined to conclude that digital experts rigorously apply the principles of sharpness and semantics whereas analog experts pay less attention to these principles. Thus we put forward a way to distinguish the digital paradigm from the analog paradigm: the former conforms to principles (1) and (2) through precise rules; the latter adheres to these principles in generic and imprecise terms.

Turing's machine - Electronics teaches us that a circuit's response can deviate slightly from the theoretical response because various *error factors* distort the signals. These inaccuracies include offset and gain errors, integral errors and differential non-linearity errors. A distortion is generated by independent phenomena such as the temperature of the circuit, the quality of semiconductors, and the system noise. The error factors always appear at the output and should be classified as unavoidable, normal electrical effects. These factors should not be confused with random failures - e.g. short circuits - producing catastrophic effects.

Analog programmers use panels, cables, and jacks to connect the various circuits necessary for a computation. They also have to control and to regulate the response of each circuit. For example, a programmer verifies the output voltage using an oscilloscope and adds an inductor or a resistor to the system to mitigate the effects of the input offset error. Analog computers are equipped with potentiometers, rheostats, and other manual tools for tuning up the electrical signals and for correcting errors. See 16 potentiometer knobs which serve the computing circuits placed beneath (Figure 1).
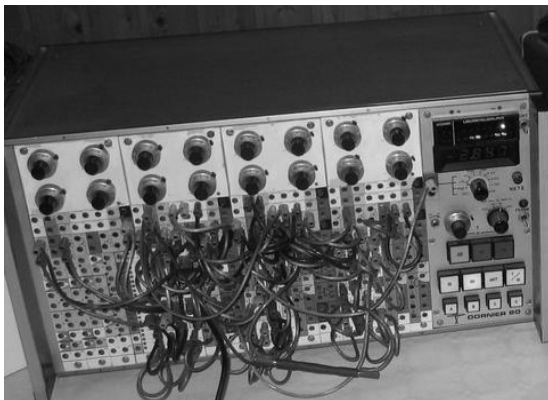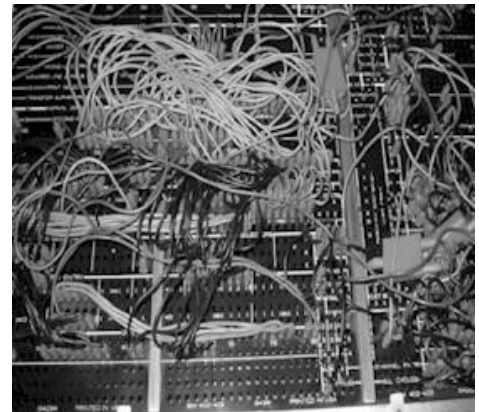

Figure 1.- Dornier DO-80 analog computer

Figure 2. Wired plug board for an IBM 402

Error factors affect also digital circuits, but the distance amongst the signals neutralizes those negative effects. Discrete values very rarely overlap and are always distinct.

The earliest digital experts operated with tabulating machines that were set up in a manner similar to analog computers, notably experts connected *n* circuits using cables and jacks as analog programmers did (Figure 2). However a great difference emerged: digital programmers were not concerned with the control of electrical signals; they simply joined the various parts

8

with plugs and jacks, and were sure of the correct execution of each function. Actually error factors do not distort the digital signals as they comply with (1) in any situation.

Digital programmers had not to control the quality of signals and used circuits as prefabricated blocks, thus over the years the tabulating machines evolved into computers. Programmers found better to join operations by means of written instructions than by cables. The Turing machine guided by a sequence of instructions gained success and fame all over the world.

The software program—exclusive to digital systems—is a pattern that triggers the circuits necessary to execute the algorithm. A programmer does not regulate digital circuits. The instructions run automatically without any manual support due to the high quality of bits that conform to properties (1) and (2). On the other hand the low quality of analog signifiers confronts them with three heavy consequences.

A] One cannot assign the meaning to an analog signal at will. The use of continuous signals forecloses symbol-oriented processing such as the manipulation of words, texts and pictures; *analog computers usually process only numerical information.*

When the value E forces the attribution of the signified *the distortion of E results in misinterpreting the message NE*. E.g. every value of the AM (Amplitude Modulation) radio waves signifies a precise sound and even small deformation interferes with listening

B] A wired plug board is sufficient to prepare a tabulating machine. Digital experts do not have to worry about electrical error factors thanks to the quality of bits that conform to (1) and (2). The high degree of confidence allowed the market to evolve from pluggable calculators toward programmable electronic computers. Programming advanced from hardware set-up to software coding. Instead the approximate application of (1) and (2) forbade symmetrical progress in the analog area. *A symbolic program cannot control the analog computer* because this system needs to be put into proper electrical conditions after manual control.

C] Electronic analog computers developed in the 1970s ran faster than digital systems solving the same differential equations. Although the higher performance of analog equipment was very attractive, the adjustment and programming of those systems turned out to be very demanding. For example, a software expert could set up a FORTRAN program to solve differential equations in a *matter of hours*; the corresponding analog expert would require *several days* to set up the analog computer. The cost of programming makes analog systems uneconomical compared to digital.

In sum analog computing has been left behind by digital computing since analog computers have limited semantic functions and fuzzy signals.

Conclusive Remarks

The present paper suggests a tentative way to discern the analog from the digital computations and in addition means to put forward a method of research revolving around the concepts of signifier and signified [roc10b].

The most famous for engineers is Shannon's theory, which soars as a solid reference for communication systems and computers. Other frames such as *semantic theory* of information by Carnap and Bar Hillel focus on linguistics and interpretations of words and sentences. Still others, such as the *autopoietic theory* of Maturana and Varela, offer an intriguing view on evolutive living beings. Unfortunately these various schools have not attempted much cooperation; the result is a jungle of information theories.

Today, most practitioners and theorists use the concepts of signifier and signified as they understand them at the intuitive level, prompted by a natural tendency. The ideas (a) and (b) have not been translated into mathematical language so far. We believe that when they are captured in mathematical models, the principles of sharpness and semantics will advance our understanding of computing and its applications in many domains**.**

If we imagine the principles of computing are written in a huge book placed in an elevated shelf of the library, we can conclude that the signifier and the signified are the stairs that take us to that book.

About the Author

Paolo Rocchi (paolorocchi@luiss.it) is a Docent Emeritus of IBM. At the present time he is an adjoined member of the CeRSI Scientific Board and serves the Luiss University of Rome as a professor.

Acknowledgements

Bibliography

[bla98] Blachowicz J. *Of Two Minds: the Nature of Inquiry*. SUNY Press, (1998).

[cow77] Coward R., Ellis J. *Language and Materialism: Development in Semiology and the Theory of the Subject.* Routledge & Kegan Paul Books, (1977).

[ges97] Gescheider G.A. *Psychophysics: The Fundamentals.* Lawrence Erlbaum Associates Publishers (1997).

[gol97] Goldweber M., Impagliazzo J., Bogoiavlenski I., Clear T., Davies G., Flack H., Mers J.P., Rasala R. Historical Perspectives on the Computing Curriculum. *ACM SIGCUE Outlook*, 25(4), (1997).

[hir92] Hirst R.J. *Problems of Perception.* Prometheus (1992).

[hol01] Holmes W.N. The Great Term Robbery. *IEEE Computer*, 34(5), (2001).

[hub90] Huber G.P. A Theory of the Effects of Advanced Information Technologies on Organizational Design, Intelligence, and Decision Making. *Academy of Management Review,* 15(1), (1990).

[mul98] Mulder F., Van Weert T. Towards Informatics as a Discipline: Search for Identity in Mulder F., Van Weert T.(eds) *Informatics in Higher Education: Views on Informatics and Non-Informatics Curricula*. Chapman & Hall, (1998).

[pel90] Peli E. Contrast in Complex Images *J. of the Optic Soc. of America A*, 7(10), (1990).

[rao09] Rao Yarlagadda R.K. *Analog and Digital Signals and Systems.* Springer (2009).

[roc10a] Rocchi P. Notes on the Essential System to Acquire Information in '*Quantum Information and Entanglement'* Special issue of *Advances in Mathematical Physics*, vol. 2010, id 480421, (2010).

[roc10b] Rocchi P. *Logic of Analog and Digital Machines.* Nova Science Publishers (2010).

11