# An Algorithmic Approach to Information and Meaning*

## A Formal Framework for a Philosophical Discussion

Hector Zenil
hectorz@labores.eu
Institut d'Histoire et de Philosophie des Sciences et des Techniques (Paris 1/ENS Ulm/CNRS)
http://www.algorithmicnature.org/zenil

**Abstract**

While it is legitimate to study ideas and concepts related to information in their broadest sense, that formal approaches properly belong in specific contexts is a fact that is too often ignored. That their use outside these contexts amounts to misuse or imprecise use cannot and should not be overlooked. This paper presents a framework based on algorithmic information theory for discussing concepts of relevance to information in philosophical contexts. Special attention will be paid to the intersection of syntactic and semantic variants and connections.

**Keywords:** information theory, meaning, algorithmic randomness and complexity, logical depth, philosophy of information.

*Presented at the *Interdisciplinary Workshop: Ontological, Epistemological and Methodological Aspects of Computer Science*, Philosophy of Simulation (SimTech Cluster of Excellence), Institute of Philosophy, at the Faculty of Informatics, University of Stuttgart, Germany, July 7, 2011.

## 1 Introduction

It is not unusual to come across surveys and volumes devoted to information (in the larger sense) in which the mathematical discussion does not venture beyond the state of the field as Shannon [36] left it some 60 years ago, who together with Wiener [41], first recognized that communication is a statistical problem, yet made clear that little had to do with information content (the semantical side of information). Recent breakthroughs in the development of information theory in its algorithmic form—both theoretical and empirical developments possessing applications in diverse domains (e.g. [28, 29, 30, 44])—are often overlooked in the semantical study of information, and it is philosophy and logic (e.g. epistemic temporal logic) what has been, one would say, forced to account for what is said to be the semantic formalism of information. As examples one may cite the work of [19, 20, 37]. Such partial and ill-considered accounts of a field can be dangerous when they become the exclusive focus of attention of semantic information. Even in the best of cases, algorithmic information theory is not given due weight. Cursorily treated, its basic definitions are inaccurately rendered, as, to cite one instance, the definition of Bennett's [3] logical depth in (p. 25 [35])—the definition provided being incomplete and therefore incorrect. This is unacceptable in a book on information, and is a reflection of the author's being ill-equipped to discuss the connection of the theory to nature.

In Floridi's works, to cite another instance [20], the only reference to algorithmic information theory as a formal context for the discussion of information content and meaning is a negative one–appearing in van Benthem's contribution (p. 171 [20]). It reads:

> To me, the idea that one can measure information flow one-dimensionally in terms of a number of bits, or some other measure, seems patently absurd...

I think this position is misguided. When Descartes transformed the notion of space into an infinite set of ordered numbers (coordinates), he did not deprive the discussion and study of space of any interest, but on the contrary advanced and expanded the philosophical discussion to encompass concepts such as dimension and curvature, which wouldn't be seriously possible otherwise in the light of the development of Descartes. Perhaps this answers the following question that Benthem poses (immediately after the above comment p. 171 [20]):

> But in reality, this quantitative approach is spectacularly more successful, often much more so than anything produced in my world of logic and semantics. Why?

On the other hand, accepting a formal framework such as algorithmic complexity for information content does not mean that the philosophical discussion of information will be reduced to the discussion of the numbers involved, just as it did not in the case of the philosophy of geometry after Descartes.

The foundational thesis upon which the state of information theory rests today, thanks to Shannon's seminal contribution, is that information can be reduced to a sequence of symbols, with the bit being the most basic unit, since any sequence of symbols can be translated into a binary sequence, thereby preserving the original content (as it can be translated back and forth from the original to the binary and vice versa). Despite the possibility of legitimate discussions of information on the basis of different foundational hypotheses, in its syntactic variant, information theory can be considered in large part achieved by Shannon's theory of communication. Epistemological discussions are, however, impossible to conceive of in the absence of a notion of semantics. There is prolific work from the side of logic to capture the concept of meaning in a broader and formal sense. Too few or nothing has, however, been done to explain meaning with pure computational models as a natural extension of Shannon's work on information and the later developments by Turing merging information and computation and, in its current state, epitomized by the theory of algorithmic information theory.

Semantics is concerned with content. Both the syntactic and semantic components of information theory are concerned with order, the former particularly with the number of symbols and their combinations, while the latter is intimately related to structure. The context provided by the theory of algorithmic information to discuss the concept of information is the theory of computation, in which the description of a message is interpreted in terms of a program. The following sections are an overview of the different formal directions in which information has developed in the last decades. They leave plenty of room for fruitful philosophical discussion, discussion focusing on information per se as well as on its connections to aspects of physical reality.

# 2 Information, computation and communication

Among the several contributions made by Alan Turing on the basis of his concept of computational universality is the unification of the concepts of data and program. Turing machines are extremely basic abstract symbol-manipulating devices, which despite their simplicity, can be adapted to simulate the logic of any computer that could possibly be constructed. While one can think of a Turing machine input as data, and a Turing machine rule table as its program, each of them being separate entities, they are in fact interchangeable as a consequence of universal computation, as shown by Turing himself, since for any input $x$ for a Turing machine $M$, one can construct $M'$ with empty input such that $M$ and $M'$ accept the same language, with $M'$ a (universal) Turing machine accepting an encoding of $M$ as input and emulating it for an input x for $M$ in $M'$. In other words, one can always embed data as part of the rule table another machine. The identification of something as data or a program is, therefore, merely a customary convention and not a fundamental distinction.

On the other hand, Shannon's conception of information inherits the pitfalls of probability. Which is to say that one cannot talk about the information content of individual strings. However, misinterpretations have dogged Shannon's information measure from the inception, especially around the use of the term *entropy*, as Shannon himself acknowledged. The problem has been that Shannon's entropy is taken to be a measure of order (or disorder), as if it were a complexity measure (and in analogy to physical entropy in classical thermodynamics). Measures based in probability theory inherit the limitation that they can only study distributions but not individual objects. This is the case, for example, with both Shannon's entropy and Fisher's notion of information. Shannon acknowledges that his theory is a theory of communication and transmission and not one of information.

That Shannon's measure is computable and easily calculable in practice may account for its frequent and unreasonable application as a complexity measure. The fact that algorithmic complexity is not computable, however, doesn't mean that one cannot approximate it—and get a much better result when it comes to the measurement of an object's complexity than a trivial upper bound offered by Shannon's measure.

# 3 Information content and algorithmic meaning

But Shannon's notion of information makes it clear that information content is subjective (Shannon himself):

> Frequently the messages have meaning: that is they are referred to or correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one selected from a set of possible messages. [36].

*Subjective* doesn't mean, however, that one cannot define information content formally, only that one should include the plausible interpretation in the definition, a point we will explore in the next section.

Shannon's contribution is seminal in that he defined the bit as the basic unit of information, as do our best current theories of information complexity. Shannon's information theory approaches information syntactically as a physical phenomenon: whether and how much information (rather than what information) is conveyed. The basic idea is that if a message contains some redundancy, this redundancy can be removed to make the channel more efficient. The elimination of redundancy is one among many possible ways to compress a message. Just as it is in algorithmic complexity.

But Shannon's doesn't help to define meaning. Think of a number like $\pi$ which is believed to be normal (that is, that its digits are equally distributed), and therefore has little or no redundancy. $\pi$, however, can be greatly compressed using any of the known briefly describable formulae and processes generating its digits. Some redundancy, in the way of patterns, is usually expected of something that is meaningful, but $\pi$ has no repeating pattern (because is an irrational number) yet it seems highly meaningful as the relationship between any circumference and its diameter. It is generally accepted that meaning is imparted by the observer. While meaning is semantic by definition, I will argue that meaning can be treated formally and syntactically.

## 3.1 Defining lack of intrinsic meaning

As an attempt to define lack of meaning think of a single bit, a single bit does not carry any information, and so it cannot but be meaningless if there is no recipient to interpret it as something that does not lack a message. The Shannon entropy of a single bit is 0 because one cannot implement a communication channel of 1 bit only, 1 and 0 having the same meaning both for Shannon and for algorithmic complexity if isolated, it cannot have any information content. In other words, it is intrinsically meaningless because there is no context (one cannot interpret a single bit if is not preceded or followed by anything else). A string of $n$ identical bits (either 1s or 0s) is also intrinsically meaningless (one is forced to make an arbitrary external interpretation to give it some meaning), because even if it carries a message it cannot be intrinsically rich (neither by the standards of algorithmic complexity nor by those of Shannon's entropy), because it cannot carry much information.

At the other extreme, a random string may or may not be taken to be meaningful depending on the measure. What one can say with certainty is that something lying in between the two extremes would definitely represent what we may consider to be *meaningful*, the two extremes in question being: *no information* (trivial) or *complete nonsense* (random). Algorithmic complexity associates randomness with the highest level of complexity, but Bennett's logical depth [3] (also based on algorithmic complexity) is able to distinguish between something that looks organized and something that looks random or trivial by introducing time (a parameter that seems unavoidable in reality, which makes it reasonable to associate this measure with *physical complexity*).

In order for the information conveyed to have any semantical value, it must in some manner add to the knowledge of the receiver. I claim that logical depth is the measure to be resorted to when it comes to mapping meaning onto information content in the real world. Logical depth is defined as the execution time required to generate a string by a near-incompressible program, i.e. one not produced by a significantly shorter program. Logically deep objects contain internal evidence of having been the result of a long computation and satisfy a slow-growth law (by definition).

## 3.2 Meaning is logically deep

The main point made by Shannon when formulating his measure in the context of communication is that in practice a message with no redundancy is more likely to carry information if one assumes one is transmitting more than just random bits. If something is random-looking, then it will usually be considered meaningless. To say that something is meaningful usually implies that one can somehow arrive at a conclusion based on it. Information has meaning only if it has a context, a story behind it. Connecting meaning to the concept of logical depth has the advantage of taking into account the context of a message, and therefore of potentially accounting for the plausible recipient's interpretation. As is known, the problem with meaning is that it is highly dependent on the recipient and its interpretation. Even if incompressible, a single bit has low algorithmic complexity when evaluated through algorithmic probability [17]. However, a single bit can still trigger a long calculation if a computing machine is designed to perform a lot of work when provided with the single bit. But taking into account both the message (in this case the single bit) and the computation associated with it (the interpretation) one can measure the organized complexity of the resulting outcome from a combination of its program-size complexity and the computing time. If the result is not trivial, nor algorithmically random, one can say that the message is meaningful. A meaningful message (short or long) contains a long computational history when taken together with the associated computation, otherwise it has little or no meaning. Hence the pertinence of the introduction of logical depth.

One might think that the approach may not be robust enough if a Turing machine performs a lot of computation when provided with a random input, in which case something that would be taken as meaningful may actually be just a random computation triggered by a random meaningless message. There are, however, two acceptable answers to this objection. The probability of a machine to undertake a long computation by chance is very low. Among the machines that halt, most machines will halt after a few steps [8]. This happens for most strings, meaning that most messages are meaningless if both the message and the computation do not somehow *resonate* to each other, something close to what one intuitively may think for a meaningful message, for example, among human beings. Algorithmic probability guarantees the almost non-existence of *Rube Goldberg machines* (a toy machine that does a lot of stuff for achieving a trivial task). In other words, strings tend to be generated by short program implementations.

So it is algorithmic probability that provides the robustness of this algorithmic approach to meaning. On the one hand, the meaning of a message makes only sense in the context of the recipient, a message that has meaning for someone may have not for someone else. This is what happens when some machines react to a *meaningful input*. In other words, the study of meaning is dependable of the recipient subject to the way in which a particular machine interprets the message. On the other hand, algorithmic probability guarantees that most machines will halt almost immediately with no computational history at the same time that most machines, for what we have defined as a *meaningful input* will perform a computation resulting in a structured output. In other words, there will be some correspondence between a meaningful input, a computation and a structured output.

A more down-to-earth example is a winning number in a lottery. The number by itself may be meaningless for a recipient, but if two parties had shared information on how to interpret it, the information shared beforehand becomes part of the computational history and as such not unrelated to the subsequent message. The only way to interpret a number as being the winning number of a lottery is to have a story, not just a story that relates the number to a process, but one that narrates the process itself. Since winning a prize is no longer a matter of apparent chance but has to do with the release of information (both the number and the interpretation of the number) it is therefore not the number alone that represents the content and meaning of the message (the number), but the story around it.

There are also messages that contain the story in themselves. If instead of a given number one substitutes the interpretation of such a number, the message can be considered meaningful in isolation. But both cases have the same logical depth, as they have the same output and computing time and are the result of the same history (even if in the first case such a history may be rendered in two separate steps) and origin.

In a move that parallels the mistaken use and overuse of Shannon's measure as a measure of complexity, the notion of complexity is frequently associated, in the field of complex systems, with the number of interacting elements or the number of layers of a system. Researchers who make such an association should continue using Shannon's entropy since it quantifies the distribution of elements, but they should also be aware that they are not measuring the complexity of a system or object, but rather its diversity, which may be a different thing altogether (despite being grossly related).

Of course this sketch of a possible algorithmic approach may or may not solve all problems related to meaning, but it is worth trying. The contribution is, however, to provide a formal computational framework to discuss these matters.

## 3.3 Finite randomness is in the beholder eye

As has been shown by Stephen Wolfram [42], it is not always the case that the greater the number of elements the greater the complexity, nor is it the case that a greater number of layers or interactions make for greater complexity, for the simplest computing systems are capable of the greatest apparent complexity. Despite what reservations one may legitimately entertain about it, it is clear that Wolfram's work has been misunderstood by a large sector of the complex systems community, as is epitomized by the misuse and misapplication of basic concepts of information and complexity in contexts that are properly the domain of probability.

If it is true that the complexity manifest in the systems studied by Wolfram is not identical to the complexity studied by algorithmic information theory, this does not mean that they are essentially incompatible. As an example, there is Wolfram's elementary cellular automaton, Rule 30. While the evolution of Rule 30 from the simplest possible initial configuration (a single black cell) looks random, the generating code has very low algorithmic complexity, given that it can be encoded in, at most, 8 bits (plus the program interpreting those 8 bits as a cellular automaton rule). Hence it is highly compressible.

Algorithmic randomness, however, does not guarantee that a string of finite length cannot be algorithmically compressed. Nonetheless, any string is guaranteed to occur as a substring (with equal probability) in any algorithmically random infinite sequence. But this has to do with the semantic value of algorithmic information theory, given that a finite string has meaning only in a particular context, as a substring of a larger, potentially complex sequence. Therefore, one can declare a string to be random-looking only as long as it does not appear as a substring embedded in another finite or infinite string. One can, however, declare a string non-random if the length of a shorter program (measured in bits) is significantly shorter than the string itself. Hence, Wolfram deterministic randomness is of epistemological nature, compatible with the fact that algorithmic randomness can only be guaranteed for infinite sequences given than any finite sequence can only be declared random-looking (as far as no short program producing it is known).

At the other extreme, Chaitin's $\Omega$ number [10] has the greatest possible meaning because it encodes all possible messages in the form of answers to all possible questions encoded by Turing machines. That Chaitin's $\Omega$ is in practice inaccessible seems a necessary characteristic, given the algorithmic definition, and accords with our intuitive notion of meaning. In other words, the meaning of all, or all possible meanings, is unattainable in this algorithmic approach, as it is ultimately uncomputable. Just as one would intuitively expect as a main feature of *meaning* in its broadest (and philosophical) sense.

# 4 A philosophical agenda

The previous discussion sketches a possible agenda for a philosophy of information in the context of the current state of the theory of algorithmic information. Focusing more on the core of the theory itself, there are several directions that a deeper exploration of the foundations of algorithmic information might take.

There is, for example, Levin's contribution to algorithmic complexity in the form of the eponymous semi-measure, motivated by a desire to fundamentally amend Kolmogorov's plain definition of complexity in light of the realization that information should follow a law of non-growth conservation. This is an apparently different motivation from the one behind Gregory Chaitin's definition of algorithmic complexity in its prefix-free version.

There are also laws of symmetry and mutual information discovered by Gács [22], and Li and Vitányi [28], for example, which remain to be explored, fully understoood and philosophically dissected.

Furthermore there are the subtle but important differences involved in capturing organization in information through the use of algorithmic randomness versus doing so using Bennett's logical depth [3], a matter brought to our attention in, for example [16]. The motivation behind Bennett's formulation of his concept of logical depth was to capture the notion of complexity taking into account the history of an object. It had the important consequence of classifying intuitively shallow physical objects as objects deprived of meaning.

There is also the question of the dependence of the definitions on the context in which meaning is evaluated (the choice of universal Turing machine), up to an additive constant, which has recently been addressed in [14, 17], who propose reasonable choices leading to reasonable evaluations (e.g. the comparison of different computational formalisms producing strongly correlated measures as shown in [15]). In other words, the problem of a stable framework for the measurement of information content. Also the connections of the theory of algorithmic information to the cognitive sciences [23] and to the physical world [15].

Van Benthem in [20] highlights an issue of great philosophical interest when he expresses a desire to understand the *unreasonable effectiveness*—a phrase he claims is borrowed—of quantitative information theories.

Paradoxically, my concern would be with the *unreasonable ineffectiveness* of qualitative information theories, notably algorithmic complexity, given that it is the *unreasonable effectiveness* of quantitative information theories, notably Shannon's notion of information entropy, that has mistakenly led researchers to use it in frameworks in which a true (and universal) measure of complexity is needed (Shannon's measure is, in the best of cases, a gross upper bound of algorithmic complexity [24], and almost anything can be an upper bound).

On the other hand, it is the ineffectiveness of algorithmic complexity that imbues information content with its deepest character, given that its full characterization cannot effectively be achieved even if it can be precisely defined. Van Benthem opens up, hence, a rich vein, this discussion being potentially fruitful and of great interest, even if it has been largely ignored so far.

I think the provisional formulations of the laws of information, together with their underlying motivations should be a central part of a discussion, if not the main focus of the semantics approaches to information as closer based on current mathematical developments.

It may be objected that the study of information aligned with the so-called semantic wing should not be reduced to the algorithmic, sometimes considered syntactic digital view of information. That, however, would be rather an odd objection, given that most texts on the information start with Shannon's information theory, without however taking the next natural step and engaging the current state of information as exemplified by algorithmic information theory. So either the philosophy of information ought to take a completely different path from Shannon's, which inevitably leads to the current state of algorithmic information and prompts deeper exploration of it, or else it should steer clear of the algorithmic side as being separate and strange to it[1]. In other words, I don't find it consistent to cover Shannon's work while leaving out all further developments of the field by, among others, Kolmogorov, Chaitin, Solomonoff, Levin, Bennett, Gács and Landauer. As I have pointed out in the previous section, there is a legitimate agenda concerning what some may call the syntactic mechanistic branch of the study of Information, which paradoxically, I think is the most interesting and fruitful part of the semantic investigation, that mainstream Philosophy of Information has traditionally steered clear of for the most part. It must be acknowledged, however, that some philosophical papers do try to more closely follow the current mathematical developments, engaging them in a philosophical discussion. Examples are Parrochia [32] and McAllister [31] (even if I do not agree with McAllister's conclusions), to mention just two.

No complete account of what information might be can be considered

[1] Pieter Adriaans has presented similar arguments [1] in relation to the current mainstream trend of semantics approaches to information.

complete without taking into account the interpretations of quantum information. One issue is the one partially raised by Wheeler, although perhaps at a different scale, that is whether an observer is necessary for information to exist and the meaning of an observation. There does not exist a universally accepted interpretation of quantum mechanics, although the "Copenhagen Interpretation" is considered the mainstream. Discussions about the meaning of quantum mechanics and its implications do not, however, lead to a consensus. It is, however, beyond the scope of this paper to further discuss the quantum approach other than for pointing out its pertinence in an encompassing discussion.

## 4.1 The basics to agree upon

One can agree upon fundamental developments from the theory of information and the theory of algorithmic information that can serve as the basis of a mathematical framework for a philosophical discussion.

- The basic unit is the bit and information is subjective (Shannon [36])
- Shallowness is meaningless (Kolmogorov [25])
- Randomness implies the impossibility of information extraction (Chaitin [10])
- Randomness is structurally meaningless (Bennett [3])
- Meaningful information can be transformed into energy (Landauer [26], Bennett [3])
- There are strong connections between logical and thermodynamic (ir)reversibility (Landauer [26], Bennett [5], Fredkin [21], Toffoli [21])
- Information follows fundamental laws: symmetry, non-growth, mutual information and (ir)reversibility (Gács [22], Zvonkin [46], Levin [27], Bennett [5], Landauer [26])
- Physics and information are related (Wheeler [40], Feynman [18], Bennett [5], Landauer [26], Fredkin [21])

The current state of the theory and the framework for a discussion of algorithmic information to be agreed upon:

- Shannon's information measure cannot capture content, organization or meaning as largerly accepted, but neither complexity.
- Shannon's measure is only incidentally connected to random complexity, and not at all to structured complexity.
- Shannon's measure is only a gross upper bound of random complexity.
- Information follows rules, e.g. conservation, non-additivity, symmetry.
- Algorithmic complexity is a universal framework for information content.
- Bennett's logical depth is a fruitful framework for an acceptable definition of meaning.
- Information (and mostly its nature) has played a major role in quantum mechanics and is assuming foundational status in modern physics as it did in classical physics, notably in thermodynamics.

# 5 Concluding remarks

A common language and formal framework to agree upon seems to be necessary to reach maturity for a fruitful and rich discussion of information, information content and meaning from the perspective of algorithmic information theory.

I've claimed that algorithmic information is suitable for defining individual information content and for providing a characterization of the concept of meaning in terms of, for example, logical depth. This rather syntactic characterization reduced to a number (size of a program, decompression time or algorithmic probability of output production) does not mean that a discussion of algorithmic information would be deprived of legitimate philosophical interest.

I have briefly drawn attention to and discussed some of the questions germane to a philosophy of algorithmic information. If our mapping of information and information content is well understood, it will be clear why we can claim that meaning is context (recipient) dependent in a rather objective, eventually formalizable way. Levin's universal distribution, taken together with Bennett's concept of logical depth, can constitute an appropriate informational framework within which to discuss these concepts. That meaning can be fully formalized doesn't mean, however, that it loses its subjectivity with respect to the recipient. On the contrary, such a subjective dimension can also be captured by the formal framework and can constitute a point of departure for an organized philosophical discussion.

# References


[1] P. Adriaans, A Critical Analysis of Floridis Theory of Semantic Information, in *Knowlegde, Technology and Policy*, 23:41–56, 2010.

[2] J.D. Bekenstein, Information in the Holographic Universe, *Scientific American*, Volume 289, Number 2, p. 61, 2003.

[3] C.H. Bennett, Logical Depth and Physical Complexity in Rolf Herken (ed), *The Universal Turing Machine–a Half-Century Survey*, Oxford University Press 227-257, 1988.

[4] C.H. Bennett, How to define complexity in physics and why. In W.H. Zurek (ed.), *Complexity, entropy and the physics of information*, Addison-Wesley, SFI studies in the sciences of complexity, p 137–148, 1990.

[5] C.H. Bennett, Demons, Engines and the Second Law, *Scientific American*, No. 257, 108–116, 1987.

[6] T. Bolognesi, Reality Is Ultimately Digital, and Its Program Is Still Undebugged in *FQXi Essay Contest - Is Reality Digital or Analog?*, 2011.

[7] A. Cabello, J.J. Joosten, Hidden Variables Simulating Quantum Contextuality Increasingly Violate the Holevo Bound. in C.S. Calude et al. (eds.) *Unconventional Computation 2011*, Lecture Notes in Computer Science 6714, Springer 2011.

[8] C.S. Calude and M.A. Stay, Most Programs Stop Quickly or Never Halt, arXiv:cs/0610153v4 [cs.IT], 2006.

[9] C.S. Calude, M.J. Dinneen, M. Dumitrescu, K. Svozil, Experimental Evidence of Quantum Randomness Incomputability, arXiv:1004.1521v2 [quant-ph], 2010.

[10] G.J. Chaitin A Theory of Program Size Formally Identical to Information Theory, *J. Assoc. Comput. Mach.* 22, 329-340, 1975.

[11] R. Cilibrasi and P. Vitányi, Clustering by compression, *IEEE Trans. on Information Theory*, 51(4), 2005.

[12] M. Daumer et al. *The Message of the Quantum?*, AIP Conference Proceedings 844, 129-132, 2006.

[13] D. Deutsch, *The Fabric of Reality: The Science of Parallel Universes and Its Implications.* Penguin (Non-Classics), 1998.

[14] J.P. Delahaye and H. Zenil, On the Kolmogorov-Chaitin complexity for short sequences, in Cristian Calude (eds), *Complexity and Randomness: From Leibniz to Chaitin*, World Scientific, 2007.

[15] J.-P. Delahaye and H. Zenil, Towards a stable definition of Kolmogorov-Chaitin complexity, arXiv:0804.3459v3 [cs.IT], 2008.

[16] J.-P. Delahaye, *Complexité aléatoire et complexité organisée*, Editions Quae, 2009.

[17] J.-P. Delahaye and H. Zenil, Numerical Evaluation of the Complexity of Short Strings: A Glance Into the Innermost Structure of Algorithmic Randomness, arXiv:1101.4795v2 [cs.IT], 2011.

[18] R. P. Feynman, *Feynman Lectures on Computation*, in J.G. Hey and W. Allen (eds.), Addison-Wesley, 1996.

[19] L. Floridi, *The Blackwell Guide to the Philosophy of Computing and Information (Blackwell Philosophy Guides)*, Wiley-Blackwell, 2003.

[20] L. Floridi (ed.) *Philosophy of Computing and Information: 5 Questions*, Automatic Press / VIP, 2008.

[21] E. Fredkin and T. Toffoli, Conservative logic, *International Journal of Theoretical Physics*, 21:219–253, 1982.

[22] P. Gács. On the symmetry of algorithmic information, *Soviet Mathematics Doklady*, 15:1477–1480, 1974.

[23] N. Gauvrit, H. Zenil, J.-P. Delahaye, Assessing Cognitive Randomness: A Kolmogorov Complexity Approach, arXiv:1106.3059v1 [cs.CC], 2011.

[24] P. Grünwald and P. Vitányi, Shannon Information and Kolmogorov Complexity, *Computing Research Repository - CORR*, 2004.

[25] A. N. Kolmogorov, Three approaches to the quantitative definition of information *Problems of Information and Transmission*, 1(1):1–7, 1965.

[26] R. Landauer, Irreversibility and heat generation in the computing process, *IBM Journal of Research and Development*, vol. 5, pp. 183–191, 1961.

[27] L. Levin, Laws of information conservation (non-growth) and aspects of the foundation of probability theory, *Problems of Information Transmission*, 10(3):206–210, 1974.

[28] M. Li, P. Vitányi, *An Introduction to Kolmogorov Complexity and Its Applications*, Springer, 3rd. ed., 2008.

[29] M. Li, J. Badger, X. Chen, S. Kwong, P. Kearney and H. Zhang, An Information-Based Sequence Distance and Its Application to Whole Mitochondrial Genome Phylogeny, *Bioinformatics*, 17(2):149154, 2001.

[30] M. Li, X. Chen, X. Li, B. Ma, P. Vitányi, The similarity metric, in *Proc. of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 863–872, 2003.

[31] J. McAllister, "Algorithmic randomness in empirical data." *Studies in History and Philosophy of Science Part A*, 34 (3):633-646, 2003.

[32] D. Parrochia, L'idée d'une information-limite, *Philosophie du langage et informatique*, Hermès, 1996.

[33] H. Leff and A.F. Rex (eds.), *Maxwell's Demon 2: Entropy, Classical and Quantum Information, Computing*, Taylor & Francis, 2002.

[34] S. Lloyd, *Programming the Universe: A Quantum Computer Scientist Takes On the Cosmos,* Alfred A. Knopf, 2006.

[35] J.G. Roederer, *Information and Its Role in Nature*, Springer, 2010.

[36] C. Shannon, A Mathematical Theory of Communication, *Bell Systems Technical Journal* 27: 279–423, 623–656, 1948.

[37] P. Godfrey-Smith and K. Sterelny, "Biological Information", The Stanford Encyclopedia of Philosophy (Fall 2008 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/fall2008/entries/information-biological.

[38] R.J. Solomonoff. A formal theory of inductive inference: Parts 1 and 2, *Information and Control*, 7:1–22 and 224–254, 1964.

[39] K. Svozil, *Randomness & Undecidability in Physics*, World Scientific, 1994.

[40] J.A. Wheeler, Information, physics, quantum: The search for links in W. Zurek (ed.) *Complexity, Entropy, and the Physics of Information*, Addison-Wesley, 1990.

[41] N. Wiener, *Cybernetics*, John Wiley and Sons, Inc., 1948.

[42] S. Wolfram, *A New Kind of Science*, Wolfram Media, 2002.

[43] H. Zenil and J.P. Delahaye, On the Algorithmic Nature of the World, in Gordana Dodig-Crnkovic and Mark Burgin (eds), *Information and Computation*, World Scientific, 2010.

[44] H. Zenil, J.-P. Delahaye and C. Gaucherel, Image Information Content Characterization and Classification by Physical Complexity, arXiv:1006.0051v4 [cs.CC], forthcoming in the journal of *Complexity*.

[45] W.H. Zurek (ed.), *Complexity, entropy and the physics of information*, Addison-Wesley, SFI studies in the sciences of complexity, p 137–148, 1990.

[46] A. K. Zvonkin, L. A. Levin, The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms, *jour Uspekhi Mat. Nauk*, vol 25, 6(156), pp. 85–127, 1970.