



Algorithmic Information Dynamics

Item Type	Article
Authors	Zenil, Hector;Kiani, Narsis;Abrahão, Felipe;Tegner, Jesper
Citation	Zenil, H., Kiani, N., Abrahão, F., & Tegnér, J. (2020). Algorithmic Information Dynamics. Scholarpedia, 15(7), 53143. doi:10.4249/scholarpedia.53143
Eprint version	Post-print
DOI	10.4249/scholarpedia.53143
Publisher	Scholarpedia
Journal	Scholarpedia
Rights	"Algorithmic Information Dynamics" by Hector Zenil, Narsis A. Kiani, Felipe S Abrahão and Jesper N. Tegnér is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. Permissions beyond the scope of this license are described in the Terms of Use.
Download date	2024-12-27 06:04:27
Item License	http://creativecommons.org/licenses/by-nc-sa/3.0/deed.en_US
Link to Item	http://hdl.handle.net/10754/666314

Algorithmic Information Dynamics

Hector Zenil et al. (2020), Scholarpedia, 15(7):53143.

doi:10.4249/scholarpedia.53143

revision #194701 [link to/cite this article]

- **Dr. Hector Zenil**, Oxford Immune Algorithmics and Algorithmic Dynamics Lab, Oxford and Stockholm, United Kingdom and Sweden
- **Dr. Narsis A. Kiani**, Algorithmic Dynamics Lab, Unit of Computational Medicine, Center for Molecular Medicine, Karolinska Institutet, Stockholm, Sweden
- **Dr. Felipe S Abrahão**, National Laboratory for Scientific Computing (LNCC), Petrópolis, Rio de Janeiro, Brazil
- **Prof. Jesper N. Tegnér**, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

Algorithmic Information Dynamics (AID) is an algorithmic probabilistic framework for causal discovery and causal analysis. It enables a numerical solution to inverse problems based or motivated on principles of algorithmic probability. AID studies dynamical systems in software space where all possible computable models can be found or approximated under the assumption that discrete longitudinal data such as particle orbits in state and phase space can approximate continuous systems by Turing-computable means. AID combines perturbation analysis and algorithmic information theory to guide a search for sets of models compatible with observations and to precompute and exploit those models as testable generative mechanisms and causal first principles underlying data and systems. AID is an alternative or a complement to other approaches and methods of experimental inference, such as statistical machine learning and classical information theory.

AID connects with and across other parallel fields of active research such as logical inference, causal reasoning, and neuro-symbolic computation. AID studies how candidate discrete computable equations as generating mechanisms are affected by changes in observed phenomena over time as a result of a system evolving (e.g. under the influence of noise) or being externally perturbed.

AID is related to other areas such as computational mechanics and program synthesis. However, unlike methods such as Bayesian networks, AID does not rely on graphical models or the (often inaccessible) empirical estimation of mass probability distributions. AID encompasses the foundations and methods that make the area of algorithmic information and algorithmic complexity more relevant to scientific discovery and causal analysis.

Contents

- 1 Foundations and Motivation
- 2 The Coding Theorem Method (CTM) and Causal Discovery
- 3 The Block Decomposition Method (BDM)
- 4 Algorithmic Intervention Analysis
- 5 Information Difference and Algorithmic Information Calculus
 - 5.1 Connection to Dynamical Systems
- 6 Applications to Data Science
 - 6.1 Dimensionality Reduction and Minimal Loss of Information
- 7 Relation to other concepts
- 8 Challenges and limitations
- 9 References
- 10 Further reading
- 11 Recommended reading
- 12 External links

Foundations and Motivation

Algorithmic Probability

Algorithmic probability in its modern formulation was introduced by Solomonoff and Levin, and by Kolmogorov and Chaitin through their work on algorithmic complexity (Chaitin, 1987; Calude, 2002; Downey & Hirschfeldt, 2010; Li & Vitányi, 2019). Algorithmic Probability considers the probability of a (discrete) object being produced by a computable mechanistic set of rules capable of Turing universality. It imposes a mapping distribution called the universal distribution between input and output. Formally, a computable process that produces a string s is a program p that when executed on a universal Turing machine U produces the string s as output and halts.

As p is itself a binary string, we can define the discrete universal a priori probability $m(s)$ as the probability that the output of an arbitrary binary input of a universal prefix-free Turing machine U is s when this input is provided with fair coin flips on the input tape. Formally,

$$m(s) := \sum_{p : U(p)=s} 2^{-\ell(p)},$$

where the sum is over all halting programs p for which U outputs the string s . As U is a prefix-free universal Turing machine, the set of valid programs forms a prefix-free set (or self-delimited programming language), and thus the sum is bounded, given Kraft's inequality (i.e., it defines a probability semi-measure).

The methods underpinning AID can be described as a combination of Bayes' theorem and computability theory, where the default agnostic prior distribution is the universal distribution instead of some other agnostic distribution such as the uniform distribution. In this way, any updates to the specific distribution of the problem proceed according to algorithmic probability.

Introduced by Ray Solomonoff, algorithmic probability, or the theory of universal inductive inference as he also called it, is a theory of prediction based on observations that assumes that each individual observed object is generated by arbitrary computable processes. An example is the prediction of the next digit in the sequence $s = 1234567\dots$. According to algorithmic probability, the next digit would be 8; if the simplest model (i.e., the shortest self-delimiting program) able to generate that sequence is the successor function $x_0 = 1; x_i := x_{i-1} + 1$, such a model would generate 8 as a predictor of the next digit given the data. The only assumption is that the prior probability follows a computable probability distribution even if such a distribution is unknown because as proven by Levin, all computable distributions converge in what is known as the universal distribution. As one of the pillars of algorithmic information theory (AIT), the universal distribution that is an offshoot of the algorithmic coding theorem conjoins algorithmic complexity, universal a priori probability, and a universal/maximal computably enumerable discrete semi-measure into a single precise and ubiquitous mathematical formalization of a necessary "bias toward simplicity" for spaces of computably generated objects. This is a result that has been called 'miraculous' in the scientific literature and been lauded by the late Marvin Minsky as the most important scientific theory of relevance to AI (Zenil, 2020; Zenil et al., 2019).

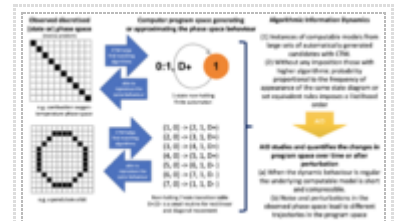


Figure 1: The basic principles of Algorithmic Information Dynamics consist in finding computable models or sequences of (short) computable models able to explain a piece of data or an observation and study the effect that interventions have on such models in the software space. Fully deterministic systems subject to no noise or no external influence will find an invariant set of candidate models with description lengths varying by only a logarithmic or sublogarithmic term; any deviation will suggest noise or external influence.

Algorithmic probability captures (albeit without actually setting out to do so) longstanding principles on which science has been founded: the principle (also known as Occam's razor) that the simplest explanation (in this case the most algorithmically probable which turns out to be the shortest algorithmic description) is most likely the correct one; and the principle of multiple explanations (Epicurus), which mandates the retention of all explanations consistent with the data; and Bayes's Rule, requiring the transformation of the a priori distribution into a posterior distribution according to the evidence, to keep all hypotheses consistent with the data (Zenil, 2020).

Numerical Methods

AID was conceived and introduced in the early 2010s and is currently an active area of research, but the methods enabling AID were introduced in the mid-2000s with the first publication of a calculation and systematic study of output probability distributions of different types of models of computation appeared in (Delahaye & Zenil, 2007). While it was not known--no experimental verification being then available--whether the concept of algorithmic probability would empirically capture the intuition of Occam's razor other than as established in the theory, Zenil and colleagues studied the behavior of these probability distributions in an exhaustive fashion (Delahaye & Zenil, 2012; Soler-Toscano et al., 2014). Their results suggested that the output distributions were more compatible than theoretically expected, and furthermore, met both the theoretical and empirical expectations entailed in Occam's razor, as the elements assigned the highest probability were also found to be the most algorithmically simple according to various complementary order parameters, including computable ones, which converge in value and thus provide further confirmation, while the least frequent elements were also more random by all measures (Zenil, 2020; Zenil, 2020a).

The numerical application of AID beyond its theoretical formulation relies strongly upon numerical methods designed to encompass and expand classical information theory to characterize randomness, using measures other than popular compression algorithms, such as Lempel-Ziv, widely used to produce purported approximations to algorithmic complexity (Zenil, 2020). These methods on which AID relies upon (but is also independent of) are the so-called coding theorem (CTM) and block decomposition (BDM) methods that have as their main features (1) that can go beyond entropic and statistical compression approaches by providing the means to explore and find computable models allowing causal discovery (and as opposed to e.g. obfuscated compressed files with no state correspondence to a model or evolving dynamical system), and (2) are sensitive enough to allow (small) perturbation causal analysis (Zenil et al., 2019; Zenil et al., 2018; Zenil et al., 2019b).

The Coding Theorem Method (CTM) and Causal Discovery

Most attempts to approximate algorithmic complexity have proceeded by way of popular lossless compression algorithms such as LZ or LZW. However, if we take the example of the sequence $s = 1234567\dots$, an algorithm such as LZ or LZW will fail at compressing s despite its obvious compressed form ($x_0 = 1; x_i := x_{i-1} + 1$). This is because algorithms such as LZ and LZW are entropy estimators and they build a dictionary of most frequent words to assign them shorter codes. For example, if s is a computable Borel normal number (Calude, 2002), no contiguous subsequence is approximately represented more frequently than any other of the same length, and the resulting compressed file tends to be of about the same size as s itself (modulo change of data type transformation, which is only a transliteration from, e.g., ASCII to binary) (Zenil et al., 2017a). Because of these limitations of popular lossless compression algorithms, and for other reasons, Zenil et al. introduced a method based on the so-called algorithmic coding theorem. Formulated by L.A. Levin (Levin, 1974), the algorithmic coding theorem in the context of algorithmic probability establishes equality between universal a priori probability $m(s)$ and algorithmic complexity $K(s)$. Formally:

$$m(s) = 2^{-K(s)} + c$$

or equivalently,

$$-\log m(s) = K(s) + c,$$

where c is a constant.

Based on this fundamental theorem, and under the assumption of optimality of the reference Turing machine, the coding theorem Method (CTM) is an alternative to statistical compression algorithms such as Lempel-Ziv, widely used to approximate algorithmic complexity. CTM does not rely upon statistical methods such as those based on the dictionaries on which popular lossless compression algorithms such as LZW are based (being designed to find statistical regularities and thus more closely related to classical information theory than to algorithmic complexity) (Delahaye & Zenil, 2012; Soler-Toscano et al., 2014a; Soler-Toscano et al., 2013; Zenil et al., 2015). The aim of CTM is to embrace Turing universality and explore the space of computer programs able to capture properties beyond statistical patterns (as a consequence of which has been overlooked in statistical approaches). Then, the method pumps it into the generative models that support a natural or artificial phenomenon, following the usual scientific *modus operandi*--whereas in statistical and probabilistic approaches the part of the probabilistic content of an object or state that belongs to the model itself and the part that belongs to the explanation of the possible explanatory models have traditionally been conflated.

For example, when modeling the outcome of throwing a dice, what a statistical model ends up quantifying is an uncertainty extrinsic to the process itself, the degree of uncertainty of an observer. Thus, it quantifies a property of the observer and not the process undergone by the dice itself. This is because the process of throwing the dice is deterministic according to the laws of classical mechanics, which are known to govern the dice trajectory (though quantum fluctuations may be believed to be probabilistic, at short distances they wouldn't have any effect). The probabilistic content of the model describing the dice outcome is thus extrinsic to the actual process. What an algorithmic-probability-like approach like CTM would do in principle--and this is its major difference--is to produce a set of deterministic models describing the dice trajectory and outcome without recourse to probability. Consequentially, it is in the distribution of possible computable models explaining the dice that a probability emerges; it is not assumed by the individual models themselves. In turn, not only can these models be tested beyond their outcome predictions as mechanistic (step-by-step) descriptions of the process, but they also offer a non-black-box approach where each model can be followed step-by-step, with model states corresponding to constructive (e.g. physical) states, as opposed to random variables with no state-to-state correspondence between the model and phenomenological data.

Unlike popular lossless compression algorithms that are guaranteed not to be able to characterize objects such as s , CTM can in principle do so, because when running all possible computer programs up to the size of s in bits there is a non-zero probability that CTM will find such a program if it exists. Indeed, this is guaranteed, as the worst case is when the program in question has the form of *print[s]* and s is algorithmically random (i.e., s is incompressible or with an algorithmic complexity value equal to the length of s , up to a constant). This holds because of the ergodicity of the algorithmic complexity approximation by the CTM over the software space, if enough computational resources are expended, which in turn is enabled by the lower semi-computability of the universal a priori probability (Zenil et al., 2018; Delahaye & Zenil, 2012).

The Block Decomposition Method (BDM)

One way to see BDM is as a weighted version of Shannon's entropy that introduces algorithmic randomness into the original formulation of classical information, a version that is able to help tell apart statistical randomness and algorithmic randomness (Zenil et al., 2019; Zenil et al., 2018; Zenil, 2020). To illustrate this let's again consider the

sequence $s = 1234567\dots$. While the digits of s are distributed with maximal entropy because it can be seen as the digital expansion of the so-called Champernowne constant, proven to be Borel normal (because it does not show any statistical patterns in the sense of overrepresentation of any digit subsequence, no subsequence of any fixed length is repeated more times than any other subsequence of the same length), when considering the set of all possible sequences, some of which do show statistical patterns, the appearance of each digit in s would be characterized as unexpected, where we may otherwise have been inclined to say that the most salient feature of s was, in reality, its non-random nature, as it can be generated by a simple deterministic/computable generative process (Zenil et al., 2017a).

This is a key methodological distinction because (1) the sequence would be characterized as maximally 'disordered' by statistical approaches such as Shannon Entropy, unless upon its application one already had the information that s had been generated by a deterministic process, which renders the use of the measure redundant (in science this is the rule rather than the exception when observing data: one investigates the nature of an object when the nature of that object is as yet unknown); and (2) it is of high empirical value in application to science.

For example, if we set out to quantify human memory, an experimental subject would not need to learn s digit by digit in order to generate it, illustrating the algorithmic nature of human cognitive processes, that go beyond statistical patterns. The sequence s may look very special, but in fact, most sequences are of this type. In addition to not having any statistical regularity, most sequences--if we consider the set of all possible sequences--do not even have any short description (low algorithmic randomness). That is, most sequences are algorithmically random. Moreover, the difference between those that do not have a short algorithmic description versus those that appear statistically random but have a short description is divergent. We only chose s because it was the most obvious for purposes of illustration (another example would be the digits of a mathematical constant such as π). Indeed, CTM and BDM have found many applications in psychometrics and cognition due to this advantage (see Applications of AID).

What BDM does is to extend the power of CTM to quantify algorithmic randomness by implementing a divide-and-conquer approach whereby the data is decomposed into pieces small enough that an exhaustive computational search finds the set of all computable models able to generate each piece of data (Zenil et al., 2019; Zenil et al., 2018; Zenil, 2020). The sequence of small generators then supports the larger piece of data from which insight is gained as to their algorithmic properties. Small pieces of data supported by even smaller programs than those pieces are called causal patches in the context of AID, as they are the result of a causal relationship between the set of these computable models and the observed data. On the other hand, models of about the same length as the data are not causally explained by a shorter generating mechanism and are therefore considered (algorithmically) random and not causally supported (Zenil, 2020a).

A sequence of computer programs smaller than their matched patches constitutes a sufficient test for non-randomness and is therefore informative of its causal content, as its components can be explained by underlying small computable models shorter than the original whole data itself. This way, the stronger the coarse-graining of the original data (i.e., the weaker the decomposition) the closer the value resulting from the BDM is to the theoretical optimal value that corresponds to the whole data's algorithmic complexity. Hence, the more fine-grained (i.e., the stronger the decomposition), the more the BDM value can diverge from the whole data's algorithmic complexity. This is because the BDM value and the algorithmic complexity are proven to converge (up to a constant that only depends on the error of the CTM relative to the algorithmic complexity) as the data partition size tends to the whole data size (Zenil et al., 2018).

Algorithmic Intervention Analysis

Unlike other approaches, such as Pearl's do-calculus, algorithmic information dynamics can help in the initial process of causal discovery and is able to perform fully unsupervised hypothesis generation from causal computable models. It also provides the tools to explore the algorithmic effects that perturbations to systems and data may have on their underlying computable models, effectively also providing a framework for causal analysis similar to the do-calculus, but without recourse to traditional probability distributions. AID can substitute for or complement other approaches to causal analysis. For example, the methods for causal analysis developed by Judea Pearl et al. assume the existence of an educated causal model but cannot provide the means for primary causal discovery.

One can formulate a cause-effect question in the language of the do-calculus as a probability between a random variable and an intervention $P(L|do(D))$, and one can also do so in the context of AID. In one of the typical examples used by Pearl himself, if L represents the human lifespan and $do(D)$ the use of some drug D , the probability P of D having an effect on L is quantified by classical probability to calculate P . What AID does is to substitute AP for P , the algorithmic probability that D exerts an effect on L with the chief advantage that not only does one obtain a probability for the dependency (and direction) between D and L but also a set of candidate models explaining such a connection, with no classical probability involved in the inference or description of each individual model. AP would then impose a natural non-uniform algorithmic probability distribution over the space of candidate models connecting D and L based on estimations of the universal distribution, which in effect introduces a simplicity bias that favors shorter models as opposed to algorithmically random ones, all of which have already been found to explain the causal connection between D and L . AID thus removes the need for classical probability distributions, and more importantly produces a set of generative models no longer derived from traditional statistics (e.g. regression, correlation), which even the do-calculus uses to determine the probability of a dependency, thereby falling back on the methods the do-calculus set out to circumvent.

While perturbation analysis is a major improvement in the area of causal discovery, AID complements it, offering a path that leaves classical probability behind and takes the next step towards independence from methods that confound correlation and causation. Another key difference between the original do-calculus and the algorithmic probability calculus is that the do-calculus makes a distinction between $P(L|do(D))$ and $P(L|D)$, but in AID both hypotheses $AP(L|D)$ and $AP(D|L)$ can be tested independently, as they have different meanings under algorithmic probability. $AP(L|D)$ means that L is the generative mechanism of D and $AP(D|L)$ means that D is the generative mechanism of L . Each of them would trigger a different explorative process under AID. The former would look for the set of smaller to larger computable models denoted by L that can explain D , while the latter would look for all the generative models $\{D\}$ that generate L . Intuitively, this means that, for example, an attempt to explain how a barometer falling (X) could be the cause for a storm would likely not yield many shorter generative mechanisms to causally explain the occurrence of the storm (Y), whereas an attempt to do the reverse would, indicating that barometric fall is effect rather than cause. Clearly, regardless of the result, $AP(X|Y)$ and $AP(Y|X)$ are not equivalent.

In the language of AID, conforming more to a typical notation in graph and set theory, the first example would be written as $AP(L \setminus D)$ or $C(L \setminus D)$, where C is the algorithmic complexity of L with intervention D (in the context of AID this is usually a deletion to a model that includes D), and the second example as $AP(X \setminus Y)$ or $C(X \setminus Y)$. Alternatively, $AP(L \setminus \{D\})$ would be L with a set of interventions $\{D\}$. Multivariate causal analysis can be achieved by conditional versions of algorithmic complexity and is currently an area of active research.

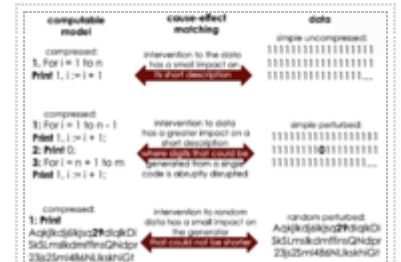


Figure 2: Causal intervention matching with AID: Central to AID is the study of the effect of interventions at the data level to the set of computable candidate models, their changes in program space being an indication both of the nature of the original data and the nature of the perturbation.

In Pearl's account of causal analysis, the so-called causal ladder leading up to human-grade reasoning consists of 3 levels, with the first being that of pattern observation covered by traditional statistics and long based on regression and correlation, the second being interventions such as the one his do-calculus suggests and that AID allows, and the 3rd level is that of counterfactuals or the power to imagine what would happen if conditions were different. AID also provides the most fundamental step in the causal analysis, which is causal discovery, within a single framework and without the need of resorting to making recourse to different approaches for different purposes (discovery vs analysis). While the do-calculus comes with no initial guidelines for how to come up with a first testable model, AID can generate a set of computable models ranked by likelihood and offers a path towards the full replacement of regression and correlation in the description of a model, taking the statistical nature out of the causal description. In addition, AID can also potentially cover all three levels of causality in Pearl's ladder, and provides the methodological framework to address each without the need of classical probability, regression, or correlation.

Both interventions and counterfactuals in AID are covered by the underlying simulation of all possible computable hypotheses (hence all 'if' questions) up to the length of the original observation, and methods such as CTM and BDM allow the estimation of processes at all these 3 levels by decomposing the original problem into smaller problems, with the additional advantage that AID ranks the counterfactuals naturally by likely algorithmic probability based on the universal distribution as a prior (Zenil et al., 2019).

Information Difference and Algorithmic Information Calculus

With its ability to quantify the departure of models away from or toward algorithmic randomness by perturbing a piece of data or a system, AID enables the investigation of which elements of the data contribute the most to the information necessary for the underlying causal computable model (the positive information elements) and which elements would trigger an increase in the system's algorithmic complexity (the negative information elements) (Zenil et al., 2019; Zenil et al., 2019b). This way, the "algorithmic randomness control" performed by the algorithmic intervention analysis puts forward new general methods for studying perturbation effects on non-linear systems, beyond those derived from classical control theory, and without making strong a priori assumptions of linearity. For example, AID has shown not only how to reconstruct the space-time evolution of Elementary Cellular Automata by gathering disordered states, but also correlated, as expected, positive information genes in E-Coli with homeostatic processes, and more negative information genes with specialization processes in the cellular network (Zenil et al., 2019).

Let $|S|$ denote the whole size of the object S , and not only the number of constitutive elements. Let M denote the total number of constitutive elements of S (Therefore M may be smaller than $|S|$, but $|S| = O(f_c(N))$, where f_c is a total computable function). For example, in the case of networks, $|S|$ grows as the number of all possible edges of a graph (in $O(N^2)$) and M is the total number of vertices.

Let $F \neq \emptyset$ be a subset of the set of elements of the object S to be deleted. In order to grasp such a formal measure of the algorithmic-informational contribution of each element, one may look into the difference in algorithmic complexity between the original data S and the destructively perturbed data $S \setminus F$. We define the *information difference* (Zenil et al., 2019; Zenil et al., 2020b) between S and $S \setminus F$ as

$$I(S, F) = K(S) - K(S \setminus F).$$

Note that the same holds analogously for constructive perturbation, i.e., the insertion of elements. When $F = \{e\}$ with $I(S, F) = I(S, e)$ and this difference assumes an absolute value

$$|I(S, e)| = |K(S) - K(S \setminus e)| \leq \log(N) + O(1).$$

we say e is a *neutral information element*, since the algorithmic information contribution of the insertion of e back into S (or its deletion from S) would necessarily be of the same order as a computable reversible procedure. Furthermore, we say e is a *positive information element*, if it is not neutral and $I(S, e) > 0$; and a *negative information subset* of elements, if it is not neutral and $I(S, e) < 0$. Thus, negative content is immediately associated with information gain from perturbations, and positive content with information loss.

In the context of causal analysis, when considering a stochastically random bit-flip being performed on a string pattern, such as 0000000... repeated n times, such a perturbation will have a significant effect with respect to the length of the original shortest program p consisting of a loop printing n number of 0s, because the new program p' would need to take into account the 1 inserted at a stochastically random position $x \leq n$. If, however, the string is algorithmically random, the effect of any stochastically random single-bit perturbation will have no effect (up to a logarithmic term of the string's algorithmic complexity, which is already maximal), because every bit is already necessary information for the string's own description. Not only for strings, but also for graphs and other multidimensional data structures, (stochastically) random single perturbations have similar worst-case effects in both simple and algorithmically random objects amounting to up to about $\log(|S|)$. Thus, in the case of objects for which $|S| \leq N^C$, where the constant $C \geq 2$ that does not depend on N (e.g. if the object is a graph), the information difference impact of moving S toward or away from algorithmic randomness may be:

1. of the same order (i.e., $I(S, e) = \Theta(\log(N))$) as the object's algorithmic complexity $K(S)$, if, e.g., S is computable and therefore logarithmically compressible in the form $K(S) \leq \log(N) + O(\log(\log(N))) + O(1)$;
2. or strongly asymptotically dominated by the object's algorithmic complexity $K(S) \geq |S| - O(1)$ as $I(S, e) = O(\log(|S|)) = o(|S|)$, if, e.g., S is algorithmically random and therefore incompressible (up to a constant);

In other words, at the asymptotic limit when $N \rightarrow \infty$, the ratio of worst-case information difference per bit of algorithmic complexity

$$\frac{I(S, e)}{K(S)}$$

tends to 0 for algorithmically random objects and to a constant $\epsilon > 0$ for computable objects. The more the object's algorithmic complexity diverges from the maximal value (i.e., the larger the algorithmic randomness deficiency), the larger the relative impact of the worst-case on the object's algorithmic complexity (i.e., on its optimal lossless compressibility). Thus, within the context of computably generated objects, this phenomenon suggests some kind of "robustness" in the face of stochastically random perturbations relative to objects' intrinsic algorithmic randomness. Indeed, as one of the main results of AID, these impacts were studied and demonstrated to be related to a thermodynamic-like behavior in stochastic-randomly moving an object toward or away from algorithmic randomness. In addition, this, in turn, led to the existence of optimum points of (re)programmability between both extremes of algorithmic complexity (Zenil et al., 2019b).

Connection to Dynamical Systems

While some theoretical results have connected algorithmic complexity and dynamical systems, not many applications have done so. One of the main results in AID is that if a chain of causally connected processes is unaffected and generated from the same generative mechanism, then its algorithmic complexity remains constant up to a (double) logarithmic term accounting for the time step when the system is to be reproduced up to a certain observable time (Zenil et al., 2019). Anything departing from such a quantity indicates that the process has been subject to an external perturbation or interaction which can be pinpointed by AID. Numerical experiments quantifying the change in the number of attractors in a dynamical system (and therefore the average depth or

shallowness of these attractors) demonstrates that, on the one hand, removing elements that reduce the algorithmic complexity of a dynamical system (with respect to the average length of the models explaining its original description) systematically reduces the number of attractors (Zenil et al., 2019). On the other hand, when removing elements that increase the dynamical system's algorithmic complexity (thus making it more algorithmically random), the number of attractors increases. Current open topics of AID research include multivariable modeling, e.g., describing multiple particles individually rather than as a system.

Applications to Data Science

Dimensionality Reduction and Minimal Loss of Information

Besides causation, control, and interventional analysis, AID methods and algorithmic information difference can also be applied to study lossy compression, so that the information loss is minimized or quantified. One can study the rate at which information is lost, recovered or reconstructed in reduced complex artificial and real networks, while retaining the typical features of biological, social, and engineering networks, such as scale-free edge distribution and the small-world property (Zenil et al., 2016; Zenil et al., 2020b).

For example, the ability to preserve the information content of a network using previous dimensionality reduction methods based on motif analysis, graph sparsification, and graph spectra has been evaluated (Zenil et al., 2016). Graph spectra have been found to be the most irregular, losing most of the information content of the original network, and network motif analysis has been shown to be the method that better preserves the original information content, which is consistent with the notion that local regularities can preserve global algorithmic information.

On the other hand, AID methods can be applied directly in order to contribute to areas of feature selection and dimensionality reduction. In the case of graphs, for example, new methods for network summarization based on successive local perturbations in the form of single edge deletions that minimize loss of algorithmic information have been introduced (Zenil et al., 2020b). These methods constitute an interesting computation-time-feasible approach to designing theoretically optimal lossy compression techniques based on principles and estimations of theoretical optimal lossless compression. It has been shown that graph-theoretic indices (ranging from degree distribution and clustering coefficient to edge betweenness and degree and eigenvector centralities) measured on synthetic and real-world networks were preserved, achieving equal or significantly better results than other data reduction methods and some of the leading network sparsification methods.

The same methods have been proposed to equip machine intelligence with an inference engine based on AID, in order to help AI algorithms build computable hypotheses from data that can then be tested against observation (Hernández-Orozco et al., 2019). This would complement other approaches such as statistical machine learning that fail at tasks such as inference and abstraction.

Relation to other concepts

There is an independent and somewhat parallel line of research by G. Chaitin concentrating on incompleteness and epistemology (Chaitin, 1975; Chaitin, 1987; Wuppuluri & Doria (ed), 2020). Chaitin's metabiology (Chaitin, 2012; Chaitin et al., 2014; Chaitin & Chaitin, 2018) also proposed to study open-ended biological evolution in software space, a concept that AID has taken inspiration from and has contributed to (Hernández-Orozco et al., 2018; Hernández-Orozco et al., 2018a).

Resource-bounded complexity

AID, as based on CTM, is a resource-bounded algorithmic complexity measure, as it takes informed runtime cutoffs (usually very short, to deal with short strings) to estimate candidate upper bounds of algorithmic complexity under assumptions of optimality and for the reference universal enumeration chosen.

Minimum Message and Minimal Description Length methods

While AID, MML, and MDL are ultimately related and are based on the principles of algorithmic probability, these minimum length approaches depart from AID (and AIT) in fundamental ways. MDL is a model selection principle rather than a model generating method. As such, AID incorporates the principle of MDL and represents a generalization, with MDL being a particular case based on learning algorithms using the statistical notion of information rather than the more general--and powerful--notion of algorithmic information used by AID that requires Turing-completeness. In practice, however, MDL and AID can complement each other because AID is more difficult to estimate while MDL provides some statistical shortcuts, and this is similar to and mostly already exemplified in the Block Decomposition Method (BDM), which combines both classical and algorithmic information theories.

In fact, AID is agnostic as regards the underlying method, even if it currently relies completely on the use of CTM and BDM, as described above. So, for example, when AID confines itself to statistical models not produced by CTM, it would collapse into methods such as MDL or MML, but any methods that introduce attempts to go beyond the statistical would approximate the spirit (if not the measure, under the assumption of optimality) of AID as exemplified in CTM and BDM.

Computational Mechanics

AID can be considered a special case or generalization of the area of computational mechanics, depending on one's perspective. They differ in that (1) Crutchfield's formulation involves the introduction of stochastic processes into the models themselves (rather than on the probability distribution of the models), while with AID there is no probabilistic content at the core of the candidate generative model, only at the level of the universal distribution (the distribution governing all computer programs), and (2) AID by way of CTM and BDM provides the means to implement other computational-mechanical approaches, including potentially Crutchfield's, as the means to mine the space of computable finite (deterministic or stochastic) automata.

Measures of sophistication

Because AID has the potential property (under assumptions of optimality) to distinguish simple from random and assign them similar values, AID can be considered a measure of sophistication similar to Bennett's logical depth. Moreover, a measure of (re)programmability is a measure of sophistication by design, and is based on AID. Some thermodynamic-like properties associated with reprogramming systems have been found and reported in the literature.

Granger causality and transfer entropy

This family of methods is statistical in nature and used mostly for causal analysis, not causal discovery, especially because they are unable to deal with models in phase-space corresponding to possible physical state models. Some relationships between variables, causal or temporal, can be derived as they can be from intervention analyses similar to Pearl's do-calculus, but they don't produce candidate mechanistic models in the first place, and the causal analysis falls back into the purely classical probabilistic framework which, though partially circumvented, must be resorted to again when it comes to testing the associative nature of 2 or more variables.

Challenges and limitations

The computational resources needed to compute CTM, at the core of AID, limit the size of its application, but BDM extends and combines CTM with other computable measures, including Shannon entropy itself. One of the most active areas of AID research is to find ways to make the algorithmic measures more relevant and to find possible shortcuts towards the computation of the computable region in the uncomputable space that is relevant for applications in science.

References

- Calude, Cristian (2002). "Information and randomness: an algorithmic perspective (<https://dl.acm.org/doi/book/10.5555/1965511>)" (2nd ed., rev. and extended ed.). Berlin: Springer. ISBN 3-540-43466-6. OCLC 50022712 (<http://worldcat.org/oclc/50022712>) . Archived from the original (<https://dl.acm.org/doi/book/10.5555/1965511>) on 2002. <https://dl.acm.org/doi/book/10.5555/1965511>.
- Chaitin, Gregory J. (1966). On the Length of Programs for Computing Finite Binary Sequences (<https://dl.acm.org/doi/10.1145/321356.321363>) (in en). *Journal of the ACM* 13 (4): 547–569. doi:10.1145/321356.321363 (<http://dx.doi.org/10.1145%2F321356.321363>) . ISSN 0004-5411 (<http://worldcat.org/issn/0004-5411>) . Archived from the original (<https://dl.acm.org/doi/10.1145/321356.321363>) on 1966. <https://dl.acm.org/doi/10.1145/321356.321363>.
- Chaitin, Gregory J. (1975). A Theory of Program Size Formally Identical to Information Theory (<https://dl.acm.org/doi/10.1145/321892.321894>) (in en). *Journal of the ACM* 22 (3): 329–340. doi:10.1145/321892.321894 (<http://dx.doi.org/10.1145%2F321892.321894>) . ISSN 0004-5411 (<http://worldcat.org/issn/0004-5411>) . Archived from the original (<https://dl.acm.org/doi/10.1145/321892.321894>) on 1975. <https://dl.acm.org/doi/10.1145/321892.321894>.
- Chaitin, Gregory J. (1987). "Algorithmic Information Theory (<https://doi.org/10.1017/CBO9780511608858>)" (in en) (1 ed.). Cambridge University Press. doi:10.1017/cbo9780511608858 (<http://dx.doi.org/10.1017%2Fcb09780511608858>) . ISBN 978-0-521-34306-0. Archived from the original (<https://doi.org/10.1017/CBO9780511608858>) on 1987. <https://doi.org/10.1017/CBO9780511608858>.
- Chaitin, Gregory (2012). "Life as Evolving Software (http://doi.org/10.1142/9789814374309_0015)" . in Zenil, Hector (in en). *A Computable Universe*. World Scientific Publishing. pp. 277–302. doi:10.1142/9789814374309_0015 (http://dx.doi.org/10.1142%2F9789814374309_0015) . ISBN 978-981-4374-29-3. Archived from the original (http://doi.org/10.1142/9789814374309_0015) on 2012. http://doi.org/10.1142/9789814374309_0015. Retrieved on 2020-10-03.
- Chaitin, Gregory; Chaitin, Virginia M.F.G. and Abrahão, Felipe S. (2014). Metabiología: los orígenes de la creatividad biológica (<https://www.investigacionyciencia.es/revistas/investigacion-y-ciencia/la-era-de-los-macrodatos-591/metabiologia-los-orgenes-de-la-creatividad-biolgica-11696>) (in es). *Investigación y Ciencia*. Archived from the original (<https://www.investigacionyciencia.es/revistas/investigacion-y-ciencia/la-era-de-los-macrodatos-591/metabiologia-los-orgenes-de-la-creatividad-biolgica-11696>) on 2014. <https://www.investigacionyciencia.es/revistas/investigacion-y-ciencia/la-era-de-los-macrodatos-591/metabiologia-los-orgenes-de-la-creatividad-biolgica-11696>. Retrieved on 2020-10-03.
- Chaitin, Virginia M. F. G. and Chaitin, Gregory J. (2018). "A Philosophical Perspective on a Metatheory of Biological Evolution (http://doi.org/10.1007/978-3-319-72478-2_29)" . in Wuppuluri, Shyam and Doria, Francisco Antonio. *The Map and the Territory*. Cham: Springer International Publishing. pp. 513–532. doi:10.1007/978-3-319-72478-2_29 (http://dx.doi.org/10.1007%2F978-3-319-72478-2_29) . ISBN 978-3-319-72477-5. Archived from the original (http://doi.org/10.1007/978-3-319-72478-2_29) on 2018. http://doi.org/10.1007/978-3-319-72478-2_29. Retrieved on 2020-10-06.

- Delahaye, Jean-Paul and Zenil, Hector (2007). "On the Kolmogorov-Chaitin Complexity for Short Sequences (http://doi.org/10.1142/9789812770837_0006) ". in Calude (in en). *Randomness and Complexity, From Leibniz to Chaitin*. World Scientific Publishing. pp. 123–129. doi:10.1142/9789812770837_0006 (http://dx.doi.org/10.1142%2F9789812770837_0006) . ISBN 978-981-277-082-0. Archived from the original (http://doi.org/10.1142/9789812770837_0006) on 2007. http://doi.org/10.1142/9789812770837_0006.
- Delahaye, Jean-Paul and Zenil, Hector (2012). Numerical evaluation of algorithmic complexity for short strings: A glance into the innermost structure of randomness (<https://linkinghub.elsevier.com/retrieve/pii/S0096300311012367>) (in en). *Applied Mathematics and Computation* 219 (1): 63–77. doi:10.1016/j.amc.2011.10.006 (<http://dx.doi.org/10.1016%2Fj.amc.2011.10.006>) . Archived from the original (<https://linkinghub.elsevier.com/retrieve/pii/S0096300311012367>) on 2012. <https://linkinghub.elsevier.com/retrieve/pii/S0096300311012367>.
- Downey, Rodney G. and Hirschfeldt, Denis R. (2010). "Algorithmic Randomness and Complexity (<http://doi.org/10.1007/978-0-387-68441-3>) ". Theory and Applications of Computability. New York, NY: Springer New York. doi:10.1007/978-0-387-68441-3 (<http://dx.doi.org/10.1007%2F978-0-387-68441-3>) . ISBN 978-0-387-95567-4. Archived from the original (<http://doi.org/10.1007/978-0-387-68441-3>) on 2010. <http://doi.org/10.1007/978-0-387-68441-3>.
- Gàcs, P. (1974). On the symmetry of algorithmic information. *Soviet Mathematics Doklady*, 15:1477--1480.
- Gauvrit, Nicolas; Soler-Toscano, Fernando and Zenil, Hector (2014). Natural scene statistics mediate the perception of image complexity (<http://www.tandfonline.com/doi/abs/10.1080/13506285.2014.950365>) (in en). *Visual Cognition* 22 (8): 1084–1091. doi:10.1080/13506285.2014.950365 (<http://dx.doi.org/10.1080%2F13506285.2014.950365>) . ISSN 1350-6285 (<http://worldcat.org/issn/1350-6285>) . Archived from the original (<http://www.tandfonline.com/doi/abs/10.1080/13506285.2014.950365>) on 2014. <http://www.tandfonline.com/doi/abs/10.1080/13506285.2014.950365>.
- Gauvrit, Nicolas; Zenil, Hector; Delahaye, Jean-Paul and Soler-Toscano, Fernando (2014). Algorithmic complexity for short binary strings applied to psychology: a primer (<https://doi.org/10.3758/s13428-013-0416-0>) (in en). *Behavior Research Methods* 46 (3): 732–744. doi:10.3758/s13428-013-0416-0 (<http://dx.doi.org/10.3758%2Fs13428-013-0416-0>) . ISSN 1554-3528 (<http://worldcat.org/issn/1554-3528>) . Archived from the original (<https://doi.org/10.3758/s13428-013-0416-0>) on 2014. <https://doi.org/10.3758/s13428-013-0416-0>.
- Gauvrit, Nicolas; Singmann, Henrik; Soler-Toscano, Fernando and Zenil, Hector (2016). Algorithmic complexity for psychology: a user-friendly implementation of the coding theorem method (<https://doi.org/10.3758/s13428-015-0574-3>) (in en). *Behavior Research Methods* 48 (1): 314–329. doi:10.3758/s13428-015-0574-3 (<http://dx.doi.org/10.3758%2Fs13428-015-0574-3>) . ISSN 1554-3528 (<http://worldcat.org/issn/1554-3528>) . Archived from the original (<https://doi.org/10.3758/s13428-015-0574-3>) on 2015. <https://doi.org/10.3758/s13428-015-0574-3>.
- Hernández-Orozco, Santiago; Hernández-Quiroz, Francisco and Zenil, Hector (2018). Undecidability and Irreducibility Conditions for Open-Ended Evolution and Emergence (https://www.mitpressjournals.org/doi/abs/10.1162/ARTL_a_00254) (in en). *Artificial Life* 24 (1): 56–70. doi:10.1162/ARTL_a_00254 (http://dx.doi.org/10.1162%2FARTL_a_00254) . ISSN 1064-5462 (<http://worldcat.org/issn/1064-5462>) . Archived from the original (https://www.mitpressjournals.org/doi/abs/10.1162/ARTL_a_00254) on 2018. https://www.mitpressjournals.org/doi/abs/10.1162/ARTL_a_00254.
- Hernández-Orozco, Santiago; Kiani, Narsis A. and Zenil, Hector (2018a). Algorithmically probable mutations reproduce aspects of evolution, such as convergence rate, genetic memory and modularity (<https://royalsocietypublishing.org/doi/10.1098/rsos.180399>) (in en). *Royal Society Open Science* 5 (8):

180399. doi:10.1098/rsos.180399 (<http://dx.doi.org/10.1098%2Frso.180399>) . ISSN 2054-5703 (<http://worldcat.org/issn/2054-5703>) . PMID: 30225028 (<http://www.ncbi.nlm.nih.gov/pubmed/30225028>) . PMC: PMC6124114 (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=PMC6124114>) . Archived from the original (<https://royalsocietypublishing.org/doi/10.1098/rsos.180399>) on 2018. <https://royalsocietypublishing.org/doi/10.1098/rsos.180399>.
- Hernández-Orozco, Santiago; Zenil, Hector; Riedel, Jürgen; Uccello, Adam; Kiani, Narsis A. and Tegnér, Jesper (2019). Algorithmic Probability-guided Supervised Machine Learning on Non-differentiable Spaces (<http://arxiv.org/abs/1910.02758>) . *arXiv:1910.02758 [cs, stat]*. Archived from the original (<http://arxiv.org/abs/1910.02758>) on 2019. <http://arxiv.org/abs/1910.02758>.
 - Hutter, Marcus (2007). On universal prediction and Bayesian confirmation (<https://linkinghub.elsevier.com/retrieve/pii/S0304397507004586>) (in en). *Theoretical Computer Science* 384 (1): 33–48. doi:10.1016/j.tcs.2007.05.016 (<http://dx.doi.org/10.1016%2Fj.tcs.2007.05.016>) . Archived from the original (<https://linkinghub.elsevier.com/retrieve/pii/S0304397507004586>) on 2007. <https://linkinghub.elsevier.com/retrieve/pii/S0304397507004586>.
 - Kolmogorov, A. N. (1968). Three approaches to the quantitative definition of information (<http://www.tandfonline.com/doi/abs/10.1080/00207166808803030>) (in en). *International Journal of Computer Mathematics* 2 (1-4): 157–168. doi:10.1080/00207166808803030 (<http://dx.doi.org/10.1080%2F00207166808803030>) . ISSN 0020-7160 (<http://worldcat.org/issn/0020-7160>) . Archived from the original (<http://www.tandfonline.com/doi/abs/10.1080/00207166808803030>) on 1968. <http://www.tandfonline.com/doi/abs/10.1080/00207166808803030>.
 - Kolmogorov, A N (1983). Combinatorial foundations of information theory and the calculus of probabilities (<http://mr.crossref.org/iPage?doi=10.1070%2FRM1983v038n04ABEH004203>) . *Russian Mathematical Surveys* 38 (4): 29–40. doi:10.1070/rm1983v038n04abeh004203 (<http://dx.doi.org/10.1070%2FRm1983v038n04abeh004203>) . ISSN 0036-0279 (<http://worldcat.org/issn/0036-0279>) . Archived from the original (<http://mr.crossref.org/iPage?doi=10.1070%2FRM1983v038n04ABEH004203>) on 1983. <http://mr.crossref.org/iPage?doi=10.1070%2FRM1983v038n04ABEH004203>.
 - Levin, Leonid A. (1974). Laws of Information Conservation (Nongrowth) and Aspects of the Foundation of Probability Theory (<http://mi.mathnet.ru/eng/ppi1039>) . *Problems Inform. Transmission* 10 (3): 206–210. Archived from the original (<http://mi.mathnet.ru/eng/ppi1039>) on 1974. <http://mi.mathnet.ru/eng/ppi1039>. Retrieved on 2020-10-03.
 - Li, Ming and Vitányi, Paul (2019). "An Introduction to Kolmogorov Complexity and Its Applications (<http://doi.org/10.1007/978-3-030-11298-1>) " (in en). Texts in Computer Science. Cham: Springer International Publishing. doi:10.1007/978-3-030-11298-1 (<http://dx.doi.org/10.1007%2F978-3-030-11298-1>) . ISBN 978-3-030-11297-4. Archived from the original (<http://doi.org/10.1007/978-3-030-11298-1>) on 2019. <http://doi.org/10.1007/978-3-030-11298-1>.
 - Soler-Toscano, Fernando; Zenil, Hector; Delahaye, Jean-Paul and Gauvrit, Nicolas (2013). Correspondence and Independence of Numerical Evaluations of Algorithmic Information Measures (<https://www.medra.org/servlet/aliasResolver?alias=iospress&doi=10.3233/COM-13019>) . *Computability* 2 (2): 125–140. doi:10.3233/COM-13019 (<http://dx.doi.org/10.3233%2FCOM-13019>) . Archived from the original (<https://www.medra.org/servlet/aliasResolver?alias=iospress&doi=10.3233/COM-13019>) on 2013. <https://www.medra.org/servlet/aliasResolver?alias=iospress&doi=10.3233/COM-13019>.
 - Soler-Toscano, Fernando; Zenil, Hector; Delahaye, Jean-Paul and Gauvrit, Nicolas (2014). Dehmer, Matthias. ed. Calculating Kolmogorov Complexity from the Output Frequency Distributions of Small Turing Machines

- (<https://dx.plos.org/10.1371/journal.pone.0096223>) (in en). *PLoS ONE* 9 (5): e96223. doi:10.1371/journal.pone.0096223 (<http://dx.doi.org/10.1371%2Fjournal.pone.0096223>) . ISSN 1932-6203 (<http://worldcat.org/issn/1932-6203>) . PMID: 24809449 (<http://www.ncbi.nlm.nih.gov/pubmed/24809449>) . PMC: PMC4014489 (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=PMC4014489>) . Archived from the original (<https://dx.plos.org/10.1371/journal.pone.0096223>) on 2014. <https://dx.plos.org/10.1371/journal.pone.0096223>.
- Soler-Toscano, Fernando; Zenil, Hector; Delahaye, Jean-Paul and Gauvrit, Nicolas (2014a). Dehmer, Matthias. ed. Calculating Kolmogorov Complexity from the Output Frequency Distributions of Small Turing Machines (<https://dx.plos.org/10.1371/journal.pone.0096223>) (in en). *PLoS ONE* 9 (5): e96223. doi:10.1371/journal.pone.0096223 (<http://dx.doi.org/10.1371%2Fjournal.pone.0096223>) . ISSN 1932-6203 (<http://worldcat.org/issn/1932-6203>) . PMID: 24809449 (<http://www.ncbi.nlm.nih.gov/pubmed/24809449>) . PMC: PMC4014489 (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=PMC4014489>) . Archived from the original (<https://dx.plos.org/10.1371/journal.pone.0096223>) on 2014. <https://dx.plos.org/10.1371/journal.pone.0096223>.
 - Soler-Toscano, Fernando and Zenil, Hector (2017). A Computable Measure of Algorithmic Probability by Finite Approximations with an Application to Integer Sequences (<https://www.hindawi.com/journals/complexity/2017/7208216/>) (in en). *Complexity* 2017: 1–10. doi:10.1155/2017/7208216 (<http://dx.doi.org/10.1155%2F2017%2F7208216>) . ISSN 1076-2787 (<http://worldcat.org/issn/1076-2787>) . Archived from the original (<https://www.hindawi.com/journals/complexity/2017/7208216/>) on 2017. <https://www.hindawi.com/journals/complexity/2017/7208216/>.
 - Solomonoff, R.J. (1964). A formal theory of inductive inference. Part I (<https://linkinghub.elsevier.com/retrieve/pii/S0019995864902232>) (in en). *Information and Control* 7 (1): 1–22. doi:10.1016/S0019-9958(64)90223-2 (<http://dx.doi.org/10.1016%2FS0019-9958%2864%2990223-2>) . Archived from the original (<https://linkinghub.elsevier.com/retrieve/pii/S0019995864902232>) on 1964. <https://linkinghub.elsevier.com/retrieve/pii/S0019995864902232>.
 - Solomonoff, R.J. (1964a). A formal theory of inductive inference. Part II (<https://linkinghub.elsevier.com/retrieve/pii/S0019995864901317>) (in en). *Information and Control* 7 (2): 224–254. doi:10.1016/S0019-9958(64)90131-7 (<http://dx.doi.org/10.1016%2FS0019-9958%2864%2990131-7>) . Archived from the original (<https://linkinghub.elsevier.com/retrieve/pii/S0019995864901317>) on 1964. <https://linkinghub.elsevier.com/retrieve/pii/S0019995864901317>.
 - Terrazas, Germán; Zenil, Hector and Krasnogor, Natalio (2013). Exploring programmable self-assembly in non-DNA based molecular computing (<https://doi.org/10.1007/s11047-013-9397-2>) (in en). *Natural Computing* 12 (4): 499–515. doi:10.1007/s11047-013-9397-2 (<http://dx.doi.org/10.1007%2FS11047-013-9397-2>) . ISSN 1572-9796 (<http://worldcat.org/issn/1572-9796>) . Archived from the original (<https://doi.org/10.1007/s11047-013-9397-2>) on 2013. <https://doi.org/10.1007/s11047-013-9397-2>.
 - Zenil, Hector; Ball, Gordon and Tegnér, Jesper (2013). Testing Biological Models for Non-linear Sensitivity with a Programmability Test (<https://www.mitpressjournals.org/doi/abs/10.1162/978-0-262-31709-2-ch188>) . *Advances in Artificial Life, ECAL 2013* (MIT Press): 1222–1223. doi:10.7551/978-0-262-31709-2-ch188 (<http://dx.doi.org/10.7551%2F978-0-262-31709-2-ch188>) . ISBN 978-0-262-31709-2. Archived from the original (<https://www.mitpressjournals.org/doi/abs/10.1162/978-0-262-31709-2-ch188>) on 2013. <https://www.mitpressjournals.org/doi/abs/10.1162/978-0-262-31709-2-ch188>.

- Zenil, H. (2014). Programmability: A Turing Test Approach to Computation. In L. De Mol and G. Primiero (eds.), *Turing in Context*, Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten (Belgian Academy of Sciences and Arts), Contactforum.
- Zenil, Hector; Soler-Toscano, Fernando; Dingle, Kamaludin and Louis, Ard A. (2014a). [Correlation of automorphism group size and topological properties with program-size complexity evaluations of graphs and complex networks](https://linkinghub.elsevier.com/retrieve/pii/S0378437114001691) (<https://linkinghub.elsevier.com/retrieve/pii/S0378437114001691>) (in en). *Physica A: Statistical Mechanics and its Applications* 404: 341–358. doi:10.1016/j.physa.2014.02.060 (<http://dx.doi.org/10.1016%2Fj.physa.2014.02.060>) . Archived from the original (<https://linkinghub.elsevier.com/retrieve/pii/S0378437114001691>) on 2014. <https://linkinghub.elsevier.com/retrieve/pii/S0378437114001691>.
- Zenil, Hector; Soler-Toscano, Fernando; Delahaye, Jean-Paul and Gauvrit, Nicolas (2015). [Two-dimensional Kolmogorov complexity and an empirical validation of the Coding theorem method by compressibility](https://peerj.com/articles/cs-23) (<https://peerj.com/articles/cs-23>) (in en). *PeerJ Computer Science* 1: e23. doi:10.7717/peerj-cs.23 (<http://dx.doi.org/10.7717%2Fpeerj-cs.23>) . ISSN 2376-5992 (<http://worldcat.org/issn/2376-5992>) . Archived from the original (<https://peerj.com/articles/cs-23>) on 2015. <https://peerj.com/articles/cs-23>.
- Zenil, Hector (2015a). [Algorithmicity and programmability in natural computing with the Game of Life as in silico case study](http://www.tandfonline.com/doi/abs/10.1080/0952813X.2014.940686) (in en). *Journal of Experimental & Theoretical Artificial Intelligence* 27 (1): 109–121. doi:10.1080/0952813X.2014.940686 (<http://dx.doi.org/10.1080%2F0952813X.2014.940686>) . ISSN 0952-813X (<http://worldcat.org/issn/0952-813X>) . Archived from the original (<http://www.tandfonline.com/doi/abs/10.1080/0952813X.2014.940686>) on 2015. <http://www.tandfonline.com/doi/abs/10.1080/0952813X.2014.940686>.
- Zenil, Hector; Kiani, Narsis A. and Tegnér, Jesper (2015b). Ortuño, Francisco and Rojas, Ignacio. eds. [Numerical Investigation of Graph Spectra and Information Interpretability of Eigenvalues](https://link.springer.com/chapter/10.1007%2F978-3-319-16480-9_39) (https://link.springer.com/chapter/10.1007%2F978-3-319-16480-9_39) (in en). *Bioinformatics and Biomedical Engineering*. Lecture Notes in Computer Science (Cham: Springer International Publishing): 395–405. doi:10.1007/978-3-319-16480-9_39 (http://dx.doi.org/10.1007%2F978-3-319-16480-9_39) . ISBN 978-3-319-16480-9. Archived from the original (https://link.springer.com/chapter/10.1007%2F978-3-319-16480-9_39) on 2015. https://link.springer.com/chapter/10.1007%2F978-3-319-16480-9_39.
- Zenil, Hector; Kiani, Narsis A. and Tegnér, Jesper (2016). [Quantifying loss of information in network-based dimensionality reduction techniques](https://academic.oup.com/comnet/article-lookup/doi/10.1093/comnet/cnv025) (<https://academic.oup.com/comnet/article-lookup/doi/10.1093/comnet/cnv025>) (in en). *Journal of Complex Networks* 4 (3): 342–362. doi:10.1093/comnet/cnv025 (<http://dx.doi.org/10.1093%2Fcomnet%2Fcnv025>) . ISSN 2051-1310 (<http://worldcat.org/issn/2051-1310>) . Archived from the original (<https://academic.oup.com/comnet/article-lookup/doi/10.1093/comnet/cnv025>) on 2016. <https://academic.oup.com/comnet/article-lookup/doi/10.1093/comnet/cnv025>.
- Zenil, Hector; Kiani, Narsis A. and Tegnér, Jesper (2016a). [Methods of information theory and algorithmic complexity for network biology](https://linkinghub.elsevier.com/retrieve/pii/S1084952116300118) (<https://linkinghub.elsevier.com/retrieve/pii/S1084952116300118>) (in en). *Seminars in Cell & Developmental Biology* 51: 32–43. doi:10.1016/j.semcdb.2016.01.011 (<http://dx.doi.org/10.1016%2Fj.semcdb.2016.01.011>) . Archived from the original (<https://linkinghub.elsevier.com/retrieve/pii/S1084952116300118>) on 2016. <https://linkinghub.elsevier.com/retrieve/pii/S1084952116300118>.
- Zenil, H. (2017). Approximations to Algorithmic Probability. In Robert A. Meyers (ed), 2nd. Edition of the Springer Encyclopedia of Complexity and Systems Science.

- Zenil, Hector; Kiani, Narsis A. and Tegnér, Jesper (2017a). [Low-algorithmic-complexity entropy-deceiving graphs](http://link.aps.org/doi/10.1103/PhysRevE.96.012308) (<http://link.aps.org/doi/10.1103/PhysRevE.96.012308>) (in en). *Physical Review E* 96 (1): 012308. doi:10.1103/PhysRevE.96.012308 (<http://dx.doi.org/10.1103%2FPhysRevE.96.012308>) . ISSN 2470-0045 (<http://worldcat.org/issn/2470-0045>) . Archived from the original (<http://link.aps.org/doi/10.1103/PhysRevE.96.012308>) on 2017. <http://link.aps.org/doi/10.1103/PhysRevE.96.012308>.
- Zenil, Hector; Hernández-Orozco, Santiago; Kiani, Narsis; Soler-Toscano, Fernando; Rueda-Toicen, Antonio and Tegnér, Jesper (2018). [A Decomposition Method for Global Evaluation of Shannon Entropy and Local Estimations of Algorithmic Complexity](http://www.mdpi.com/1099-4300/20/8/605) (<http://www.mdpi.com/1099-4300/20/8/605>) (in en). *Entropy* 20 (8): 605. doi:10.3390/e20080605 (<http://dx.doi.org/10.3390%2Fe20080605>) . ISSN 1099-4300 (<http://worldcat.org/issn/1099-4300>) . Archived from the original (<http://www.mdpi.com/1099-4300/20/8/605>) on 2018. <http://www.mdpi.com/1099-4300/20/8/605>.
- Zenil, Hector; Kiani, Narsis A.; Marabita, Francesco; Deng, Yue; Elias, Szabolcs; Schmidt, Angelika; Ball, Gordon and Tegnér, Jesper (2019). [An Algorithmic Information Calculus for Causal Discovery and Reprogramming Systems](https://linkinghub.elsevier.com/retrieve/pii/S2589004219302706) (<https://linkinghub.elsevier.com/retrieve/pii/S2589004219302706>) (in en). *iScience* 19: 1160–1172. doi:10.1016/j.isci.2019.07.043 (<http://dx.doi.org/10.1016%2Fj.isci.2019.07.043>) . PMID: 31541920 (<http://www.ncbi.nlm.nih.gov/pubmed/31541920>) . PMC: PMC6831824 (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=PMC6831824>) . Archived from the original (<https://linkinghub.elsevier.com/retrieve/pii/S2589004219302706>) on 2019. <https://linkinghub.elsevier.com/retrieve/pii/S2589004219302706>.
- Zenil, Hector; Kiani, Narsis A.; Zea, Allan A. and Tegnér, Jesper (2019a). [Causal deconvolution by algorithmic generative models](http://www.nature.com/articles/s42256-018-0005-0) (<http://www.nature.com/articles/s42256-018-0005-0>) (in en). *Nature Machine Intelligence* 1 (1): 58–66. doi:10.1038/s42256-018-0005-0 (<http://dx.doi.org/10.1038%2Fs42256-018-0005-0>) . ISSN 2522-5839 (<http://worldcat.org/issn/2522-5839>) . Archived from the original (<http://www.nature.com/articles/s42256-018-0005-0>) on 2019. <http://www.nature.com/articles/s42256-018-0005-0>.
- Zenil, Hector; Kiani, Narsis A. and Tegnér, Jesper (2019b). [The Thermodynamics of Network Coding, and an Algorithmic Refinement of the Principle of Maximum Entropy](https://www.mdpi.com/1099-4300/21/6/560) (<https://www.mdpi.com/1099-4300/21/6/560>) (in en). *Entropy* 21 (6): 560. doi:10.3390/e21060560 (<http://dx.doi.org/10.3390%2Fe21060560>) . ISSN 1099-4300 (<http://worldcat.org/issn/1099-4300>) . Archived from the original (<https://www.mdpi.com/1099-4300/21/6/560>) on 2019. <https://www.mdpi.com/1099-4300/21/6/560>.
- Zenil, Hector; Badillo, Liliana; Hernández-Orozco, Santiago and Hernández-Quiroz, Francisco (2019c). [Coding-theorem like behaviour and emergence of the universal distribution from resource-bounded algorithmic probability](https://www.tandfonline.com/doi/full/10.1080/17445760.2018.1448932) (<https://www.tandfonline.com/doi/full/10.1080/17445760.2018.1448932>) (in en). *International Journal of Parallel, Emergent and Distributed Systems* 34 (2): 161–180. doi:10.1080/17445760.2018.1448932 (<http://dx.doi.org/10.1080%2F17445760.2018.1448932>) . ISSN 1744-5760 (<http://worldcat.org/issn/1744-5760>) . Archived from the original (<https://www.tandfonline.com/doi/full/10.1080/17445760.2018.1448932>) on 2019. <https://www.tandfonline.com/doi/full/10.1080/17445760.2018.1448932>.
- Zenil, Hector; Kiani, Narsis A. and Tegnér, Jesper (2019d). "Algorithmic Information Dynamics of Emergent, Persistent, and Colliding Particles in the Game of Life * (<https://www.taylorfrancis.com/books/e/9781315167084/chapters/10.1201/9781315167084-17>) ". in Adamatzky, Andrew; Akl, Selim and Sirakoulis, Georgios Ch. (in en). *From Parallel to Emergent Computing* (1 ed.). Boca Raton, Florida : CRC Press, [2019] | Produced in celebration of the 25th anniversary of the International Journal of Parallel, Emergent, and Distributed Systems.: CRC Press. pp. 367–384. doi:10.1201/9781315167084-17 (<http://dx.doi.org/10.1201%2F9781315167084-17>) . ISBN 978-1-315-16708-4.

Archived from [the original](#)

(<https://www.taylorfrancis.com/books/e/9781315167084/chapters/10.1201/9781315167084-17>) on 2019.

<https://www.taylorfrancis.com/books/e/9781315167084/chapters/10.1201/9781315167084-17>. Retrieved on 2020-10-07.

- Zenil, Hector and Minary, Peter (2019e). Training-free measures based on algorithmic probability identify high nucleosome occupancy in (<https://www.ncbi.nlm.nih.gov/pubmed/31511887>) DNA sequences. *Nucleic Acids Research* 47 (20): e129. doi:10.1093/nar/gkz750 (<http://dx.doi.org/10.1093%2Fnar%2Fgkz750>) . ISSN 1362-4962 (<http://worldcat.org/issn/1362-4962>) . PMID: 31511887 (<http://www.ncbi.nlm.nih.gov/pubmed/31511887>) . PMC: 6846163 (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=6846163>) . Archived from the original (<https://www.ncbi.nlm.nih.gov/pubmed/31511887>) on 2019. <https://www.ncbi.nlm.nih.gov/pubmed/31511887>.
- Zenil, Hector (2020). A Review of Methods for Estimating Algorithmic Complexity: Options, Challenges, and New Directions (<https://www.mdpi.com/1099-4300/22/6/612>) (in en). *Entropy* 22 (6): 612. doi:10.3390/e22060612 (<http://dx.doi.org/10.3390%2Fe22060612>) . ISSN 1099-4300 (<http://worldcat.org/issn/1099-4300>) . Archived from the original (<https://www.mdpi.com/1099-4300/22/6/612>) on 2020. <https://www.mdpi.com/1099-4300/22/6/612>.
- Zenil, Hector (2020a). "Compression is Comprehension and the Unreasonable Effectiveness of Digital Computation in the Natural World (http://doi.org/10.1142/9789811200076_0011) " (in en). *Unravelling Complexity*. World Scientific Publishing. pp. 201–238. doi:10.1142/9789811200076_0011 (http://dx.doi.org/10.1142%2F9789811200076_0011) . ISBN 978-981-12-0006-9. Archived from the original (http://doi.org/10.1142/9789811200076_0011) on 2020. http://doi.org/10.1142/9789811200076_0011. Retrieved on 2020-10-03.
- Zenil, Hector; Kiani, Narsis A.; Abrahão, Felipe S.; Rueda-Toicen, Antonio; Zea, Allan A. and Tegnér, Jesper (2020b). Minimal Algorithmic Information Loss Methods for Dimension Reduction, Feature Selection and Network Sparsification (<http://arxiv.org/abs/1802.05843>) . *arXiv:1802.05843 [physics]*. Archived from the original (<http://arxiv.org/abs/1802.05843>) on 2020. <http://arxiv.org/abs/1802.05843>.

Further reading

- A Brief Introduction to Algorithmic Information Dynamics by H. Zenil at the conference AUTOMATA 2020 (<https://youtu.be/HXM3BUXsY4g>)
- Santa Fe Institute Course on Algorithmic Information Dynamics, 2018 (<http://web.archive.org/web/20180626020129/https://www.complexityexplorer.org/courses/63-algorithmic-information-dynamics-a-computational-approach-to-causality-and-living-systems-from-networks-to-cells>)
- Gallager, Robert G.. "Information theory and reliable communication (<https://link.springer.com/book/10.1007/978-3-7091-2945-6>) ". Wein. ISBN 978-3-7091-2945-6. OCLC 883323349 (<http://worldcat.org/oclc/883323349>) . Archived from the original (<https://link.springer.com/book/10.1007/978-3-7091-2945-6>) on 1972. <https://link.springer.com/book/10.1007/978-3-7091-2945-6>.
- Wuppuluri, Shyam, ed (2020). "Unravelling Complexity: The Life and Work of Gregory Chaitin (<https://www.worldscientific.com/worldscibooks/10.1142/11270>) " (in en). World Scientific Publishing. doi:10.1142/11270 (<http://dx.doi.org/10.1142%2F11270>) . ISBN 978-981-12-0006-9. Archived from the original

(<https://www.worldscientific.com/worldscibooks/10.1142/11270>) on 2020.

<https://www.worldscientific.com/worldscibooks/10.1142/11270>.

- Zvonkin, A K and Levin, L A (1970). The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms (<http://mr.crossref.org/iPage?doi=10.1070%2FRM1970v025n06ABEH001269>) . *Russian Mathematical Surveys* 25 (6): 83–124. doi:10.1070/rm1970v025n06abeh001269 (<http://dx.doi.org/10.1070%2FRM1970v025n06abeh001269>) . ISSN 0036-0279 (<http://worldcat.org/issn/0036-0279>) . Archived from the original (<http://mr.crossref.org/iPage?doi=10.1070%2FRM1970v025n06ABEH001269>) on 1970. <http://mr.crossref.org/iPage?doi=10.1070%2FRM1970v025n06ABEH001269>.

Internal references

- Marcus Hutter (2007) Algorithmic information theory. Scholarpedia, 2(3):2519.
- Marcus Hutter, Shane Legg, Paul M.B. Vitanyi (2007) Algorithmic probability. Scholarpedia, 2(8):2572.
- Rodney G. Downey and Jan Reimann (2007) Algorithmic randomness. Scholarpedia, 2(10):2574.
- Paul M.B. Vitanyi (2007) Andrey Nikolaevich Kolmogorov. Scholarpedia, 2(2):2798.
- Ming Li and Paul M.B. Vitanyi (2007) Applications of algorithmic information theory. Scholarpedia, 2(5):2658.
- Olaf Sporns (2007) Complexity. Scholarpedia, 2(10):1623.
- Tomasz Downarowicz (2007) Entropy. Scholarpedia, 2(11):3901.
- Matteo Gagliolo (2007) Universal search. Scholarpedia, 2(11):2575.

Recommended reading

For an introduction to Kolmogorov complexity, its history, and for further references, see:

- An Introduction to Kolmogorov Complexity and Its Applications (<http://www.springer.com/chl/home?SGWID=2-102-22-1519244-o&changeHeader=true&referer=www.springeronline.com&SHORTCUT=www.springer.com/o-387-94868-6>)
- Methods and Applications of Algorithmic Complexity, Beyond Lossless Compression Algorithms, forthcoming (2020), Springer Verlag
- Algorithmic Information Dynamics, forthcoming (2020), Cambridge University Press

External links

- Algorithmic Information Dynamics course at the Sta Fe Institute platform Complexity Explorer (<https://youtu.be/2K5w28wZmzo>)
- Video produced by Nature to explain an application of AID to machine intelligence (<https://www.youtube.com/watch?v=rkmz7DAA-t8>)
- The works of Gregory Chaitin (<https://uba.academia.edu/GregoryChaitin>)
- Homepages of
 - Algorithmic Dynamics Lab (<https://www.algorithmicdynamics.net/>)
 - Algorithmic Nature Group (<https://algorithmicnature.org/>)
 - Living Systems Lab (<https://livingsystems.kaust.edu.sa>)
 - Jesper Tegner (https://scholar.google.se/citations?hl=en&user=_DUppAgAAAAJ&view_op=list_works&sortby=pubdate)

- [Hector Zenil \(https://www.hectorzenil.net/\)](https://www.hectorzenil.net/)

See also

Algorithmic "Kolmogorov" Complexity, Algorithmic "Solomonoff" Probability, Universal "Levin" Search, Algorithmic "Martin-Loef" Randomness, Recursion Theory, Applications of AIT.

Sponsored by: Eugene M. Izhikevich, Editor-in-Chief of Scholarpedia, the peer-reviewed open-access encyclopedia

Reviewed by (http://www.scholarpedia.org/w/index.php?title=Algorithmic_Information_Dynamics&oldid=194039) :

Prof. Gregory J Chaitin, University of Buenos Aires, Buenos Aires, Federal District, Argentina

Reviewed by (http://www.scholarpedia.org/w/index.php?title=Algorithmic_Information_Dynamics&oldid=0) :

Eugene M. Izhikevich, Editor-in-Chief of Scholarpedia, the peer-reviewed open-access encyclopedia

Accepted on: 2020-07-20 17:07:26 GMT (http://www.scholarpedia.org/w/index.php?title=Algorithmic_Information_Dynamics&oldid=194061)

[title=Algorithmic_Information_Dynamics&oldid=194061](http://www.scholarpedia.org/w/index.php?title=Algorithmic_Information_Dynamics&oldid=194061))

Categories: Pages containing cite templates with deprecated parameters | Algorithmic Information Theory | Artificial Intelligence | Computational Intelligence | Computer Science

This page was last modified on 8 October 2020, at 17:57.



This page has been accessed 197,315 times.

"Algorithmic Information Dynamics" by Hector Zenil, Narsis A. Kiani, Felipe S Abrahão and Jesper N. Tegnér is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. Permissions beyond the scope of this license are described in the Terms of Use