Algorithmic Complexity

Appendix to
The Prophets of the Thinking Machines:
Artificial General Intelligene and Transhumanism
History, Theory and Pioneers
Past, Present and Future
by Todor Arnaudov

Edition: 20.7.2025[1]
http://twenkid.com/agi
https://github.com/twenkid/sigi-2025
#complexity #сложност

©Authors: The authors of all cited and reviewed works, such as:
Hector Zenil, Felipe Abrahao, Andrey Kolmogorov, Gregory Chaitin, Ray Solomonoff, Leonid Levin; Charles Benett, David Deutsch, Nicolas Gauvrit, Jesper Tegnér, Leong Ting Lui, Santiago Hernández-Orozco, Jean-Paul Delahaye, Cristian S. Calude, Gordana Dodig-Crnkovic, Oliver Korten Noam Nisan, Avi Wigderson, Martin Davis, Alexander Shen, P. Martin-Löf, Marvin Minsky, Korb and Dorin, Luciano Floridi et al.,; various encyclopaedic sources such as Wikipedia etc.
… & Todor Arnaudov – author of *The Prophets of the Thinking Machines* and editor, summarizer, introductory texts, notes and author of the contextual reviews and mini-sub-articles.

*** The Sacred Computer**: Thinking Machines, Creativity and Human Development
http://github.com/twenkid   http://twenkid.com   https://artificial-mind.blogspot.com

[1] 32 pages, added morediscussions, fixed a few references, mentioned more authors on the first page; editing and more line-spaceing. First edition: 18.7.2025, 27 pages, 19.7: 30 p.

You're invited to join or help the "Sacred Computer": in the R&D and as authors of future or refined works, including appendices to The Prophets.
Visit the year-long virtual conference Thinking Machines 2025/Self-Improved General Intelligence 2025: https://github.com/twenkid/sigi-2025

Why this appendix was compiled?

The algorithmic complexity and probability – AC, AP – are important concepts in AGI and the theories of the Universe as Computer as they are about general prediction and compression for any kind of data, based on a program running on some model computer – Turing machine, cellular automaton etc.

One specific reason for compiling this work *now* was that I recently discovered a paper from the circle of Hector Zenil about the SuperARC[2] test, and then a recent strategic program paper about the *Simulation Intelligence*[3]. I decided to refresh and deepen my knowledge by doing a literature review, using Hector's works as a starting point[4].

This work intends to be food for thought for readers who also want to improve and widen their knowledge. Other goals are to share my comments and additions to some of the reviewed ideas from the papers and in the context of *The Prophets…* – to shed light on the yet little known early 2000s works from TUM, in relation to Artificial General Intelligence, whose insights keep being rediscovered and restated in academical publications 10 or 20-some years later.

A broad conclusion which I made for myself after this new review of the field was that the Turing machines in particular, in their original form, and the original Kolmogorov-Chaitin complexity, computed with these particular machines, are perhaps not the best way to study computation for *Thinking Machines*, because the *Turing* machines are too simple, "mindless" and stiff. Blind enumeration of "all random programs", generating flat strings of bits or symbols, without another meaning, mapping, relationsan, explicit structure, and just calculation of the K(x)[5] that way, also doesn't seem too interesting or enlightening per se for me, beyond computing some measurements of some complexity for producing some ranking,

[2] See below. This paper has technical and extensive continuations of generalizations done also in TUM in the early 2000s
[3] Simulation Intelligence: Towards a New Generation of Scientific Methods Alexander Lavin, David Krakauer, Hector Zenil et al., 2021/2022, https://arxiv.org/abs/2112.03235
[4] Part of the reviews here are included also in the bigger appendix **#Listove** which has broader and interdisciplinary scope and currently is bilingual - Bulgarian and English.
[5] This work could be criticized for the lack of (many) formulas with mathematical notation – see the cited papers for this part. The goal here are broad generalizations and new insights, not calculations. In additions the formulas are usually very simple and "not stacking", unlike a complex computer program with many explicit functions, classes, data structures. See: 1. Ultimate AI, Free Energy Principle and Predictive Coding vs Todor and CogAlg - Discussion in Montreal.AI forum and Artificial Mind, 7.12.2018 … and "Is Mortal Computation…",2025 about the *"Sigma-Product-Log-Probability mathematical formula fetishism"*. https://artificial-mind.blogspot.com/2018/12/ultimate-ai-free-energy-principle-and.html

which sometimes is useful for deciding; however the other material *reasons* and *connections*, *logic* etc. for computing and constructing the complexities are not transparent or present in flat enumerating approaches on disconnected linear data as just strings of bits; *the interpretation* matters. If only the K(x) or AP(x) are computed, the actual meaning and structure are derived by or mapped to, taken from other structures, sources and analysis[6]. A thinking machine also has *Will*.

See Zrim, Vursherod, Kazborod, InR etc. in the future work: "*Creating Thinking Machines*" and the articles in "The Prophets…", citing Arthur Schopenhauer's thought "*Where Calculation Begins, Comprehension Ceases*"[7].

Complexity and The Universe Computer from the "Sacred Computer"

* **World's first university course in Artificial General Intelligence,** Plovdiv 2010, 2011,
https://artificial-mind.blogspot.com/2010/04/universal-artificial-intelligence.html
https://research.twenkid.com/agi/2010/
https://research.twenkid.com/agi/2010/Complexity_Probability_Chaos_10-2010_MTR.pdf
* **Lecture slides (Bulgarian):** Complexity and Theory of Information. Probability Theory – statistical (empirical) probability. Chaos Theory. Systems Theory. Emergent functions and behavior. Universe as a computer – digital physics. Algorithmic Probability. Kolmogorov's Complexity and Minimum Message Length. Occam's Razor. (slides originally in Bulgarian)
* **Definition of Machine Intelligence**, **Legg and Hutter, 2007** – Slides for the course: Bulgarian: "Универсална интелигентност: Дефиниция на агента на Маркус Хутер" по работата: … Universal Intelligence:"
https://research.twenkid.com/agi/2010/Intelligence_by_Marcus_Hutter_Agent_14_5_2010.pdf
*** Slides in English "Universal Intelligence According to Marcus Hutter's Definition"**
https://research.twenkid.com/agi/2010/en/Machine_Intelligence_Hutter_Legg_Eng_MTR_Twenkid_Research.pdf

*** See TUM = "Theory of Universe and Mind",** 2001-2004, Todor Arnaudov,
"The Universe Computer" etc. (Bulgarian and some translations)
Notes: Tosh = Todor

https://eim.twenkid.com/old/razum/index.htm - Originals in Bulgarian
Translations in English of some parts:
https://research.twenkid.com/agi/2010/en/Todor_Arnaudov_Theory_of_Universe_and_Mind_3.pdf

---

[6] At least this is my insight for now.
[7] * Where calculations begin, comprehension ceases* - on understanding and superintelligent AGI. Todor's comment in "Strong Artificial Intelligence" at Google+
https://artificial-mind.blogspot.com/2014/08/where-calculations-begin-comprehension.html
***** On Understanding and Calculation, Quantitative and Qualitative Reasoning: Where Calculation begins Comprehension Ceases: Part II
https://artificial-mind.blogspot.com/2023/11/on-understanding-and-calculation.html

https://research.twenkid.com/agi/2010/en/Todor_Arnaudov_Theory_of_Universe_and_Mind_4.pdf
https://research.twenkid.com/agi/2010/en/Todor_Arnaudov_Theory_of_Hierarchical_Universal_Simulators_of_universes_Eng_MTR_3.pdf
* Semantic analysis of a sentence. Reflections about the meaning of the meaning and the Artificial Intelligence … 2004 (Bulgarian)/2010 (English) (Анализ на смисъла на изречение въз основа на базата знания на действаща мислеща машина. Мисли за смисъла и изкуствената мисъл, 2004) https://artificial-mind.blogspot.com/2010/01/semantic-analysis-of-sentence.html

*** Introduction, general information and comparisons of TUM, 2023-2024:**

https://github.com/Twenkid/Theory-of-Universe-and-Mind
Detailed review of other schools of thought and theories (such as Free Energy Principle/Active Inference, Michael Levin's et al. generalizations etc.) are given in the main volume of *"The Prophets of the Thinking Machines"*. A lot is spread in the old posts in the blog "Artificial Mind".

*** Is Mortal Computation Required for the Creation of Universal Thinking Machine**, 2025 (currently in Bulgarian) – challenging measurements or the requirement of efficiency etc.
*** Universe and Mind 6**, ~2023+, to be published on SIGI-2025 (English)
*** The main volume of "The Prophets of the Thinking Machines …"**, 2025 .
* **Ultimate AI, Free Energy Principle and Predictive Coding vs Todor and CogAlg - Discussion in Montreal.AI forum and Artificial Mind,** 7.12.2018 …
https://artificial-mind.blogspot.com/2018/12/ultimate-ai-free-energy-principle-and.html

Overview of Basic Concepts about Turing Machines and Computation

See: Computable_function * https://en.wikipedia.org/wiki/Turing_machine

A **tape**, divided into **cells**; a head; ***state register***, a **finite table of instructions**: given the state and the symbol that is read from the tape, what to do: erase a symbol and replace it with a given one, move the head left or right. One of the states is **HALT** – the machine stops. For **one-sided** TM a shift from the 0-th position to the first out, where the head couldn't go, is also used for a HALT state for reducing the number of states. The head of a **two-sided** TM can move in both directions.

A Post–Turing machine[8]:a **binary** alphabet (1,0), *infinite sequence of storage locations, instructions for bi-directional movement among the storage locations and alteration of their contents one at a time.*

Basic TM types and parameters: one-sided or two-sided tape = one or two directions of expansion; numbers of symbols in the alphabet; single tape or multi-tape; states; halting state; number of steps; multi-tape have many heads, while a **multi-track** TM has many tapes, but one head writes to all at once. The **halting problem** – predicting whether a random program or/and input will end ("halts"), i.e. the TM will enter into the halt state, after following the procedure fed to the TM, or the machine will work indefinitely like in an endless loop. This problem is unsolvable for the general case for a random program and input. **An extension of Turing machine** defined by Turing himself in 1939, called **Oracle machine,** provides non-computable answers as a black-box, e.g. whether a program halts. The term "oracle" as a source of information about the correctness of the solution is used in modern fields of computer science, for example in **program synthesis**[9]. The halting problem is resolved in practice and in the computers by limiting the conditions, adding time limits and halting by a time out, watchdogs to interrupt the machine; in computers – with interrupts and preemtive operating systems etc.

Some other types of abstract machines are: **Counter machines | Pointer machines (graphs) | Register machines & Random access machines,** where RA-machines are an extension of the Register machines, they both have memory … **Hypercomputation** are methods that can compute TM-uncomputable problems. See Model_of_computation: Finite-state machines, Post machines (Post–Turing machines and tag machines), Pushdown automata, Decision tree model, External memory models, Functional models = Abstract rewriting systems, Combinatory logic, General recursive functions. Lambda calculus, Concurrent models = Actor model, Cellular automaton, Interaction nets, Kahn process

[8] https://en.wikipedia.org/wiki/Post%E2%80%93Turing_machine

[9] Oracle-Guided Component-Based Program Synthesis, Susmit Jha et al, 2010, https://people.eecs.berkeley.edu/~sseshia/pubdir/synth-icse10.pdf
https://cmu-program-analysis.github.io/2021/lecture-slides/20-oracle-guided.pdf
See also the literature and notes on **Program Synthesis** in the appendix "A survey of various papers …" **#lotsofpapers** file: lazar-…
https://www.cs.cmu.edu/~aldrich/courses/17-355-18sp/notes/notes13-synthesis.pdf
https://www.microsoft.com/en-us/research/wp-content/uploads/2017/10/program_synthesis_now.pdf
https://synthesis.to/presentations/introduction_to_program_synthesis.pdf

networks, Logic gates and digital circuits, Petri nets, Process calculus, Synchronous Data Flow. These are all ways to represent and construct models of systems that travel through a state space given initial configuration and possible input data.

See: **Schönhage's storage modification machine (SMM)**, a type of pointer machine that modifies a graph; "linking automaton" and Kolmogorov–Uspenskii (KU-machine). Pointer_machine

See also: Lambda_calculus General_recursive_function (**μ-recursive:** a finite tuples of natural numbers, returning a single natural number). Types of functions: constant, successor (x+1), projection ... **Total recursive function** is one defined for **all** possible inputs or that can be solved by a **total Turing machine,** called also a **Decider** – it halts for every input (in programming, the function returns). Decider_(Turing_machine) Total functional programming – restricted to the programs which are provably terminating. Total_functional_programming

Computation is connected to **Enumeration.**

Algorithmic and Non-Algorithmic Problems

The algorithmic or computable problems are ones having well defined procedures of explicit instructions, which lead to solution in all cases (or to the "halt state" in a TM). Other types of problems don't have a guarantee and are e.g. **heuristic** or empirical – where guesses are made; **optimization** problems of searching for the solution, which may be **stochastic** and also may not guarantee that the result that is found is the globally optimal one; **physical** (applied and interacting with the "real world"); **ill-defined** (subjective, vague, lacking clear criteria for correctness; some call them "AI complete", but this is confused); non-computable (undecideable). Some algorithmic problems are intractable/NP-hard – their solution is computationally hard to find (O(3) etc.), but the verification of a proposed solution for correctness is easy; one such problem is the "travelling salesman" for more than a few cities, nodes in a graph.

*** Algorithmic Information Theory and Kolmogorov Complexity**, Alexander Shen, 19.11.2007 – Lecture notes of an inroductory course on Kolmogorov complexity. https://www.researchgate.net/publication/228600376_Algorithmic_information_theory_and_Kolmogorov_complexity

Selected papers, summaries and notes in Algorithmic Complexity, Randomness etc.

* **THE COMPLEXITY OF FINITE OBJECTS AND THE DEVELOPMENT OF THE CONCEPTS OF INFORMATION AND RANDOMNESS BY MEANS OF THE THEORY OF ALGORITHMS**, A.K. Zvonkin and L.A. Levin, 1970
One of the foundational works, including: … etc. See also the work of G.Chaitin.
* A.N. Kolmogorov, **Three approaches to the concept of the "amount of information"** , Problemy Peredachi Informatsii 1 : 1(1965), 3-11. MR32# 2273.
http://alexander.shen.free.fr/library/Kolmogorov65_Three-Approaches-to-Information.pdf
"p.1. There are two common approaches to the quantitative definition of "information": combinatorial and probabilistic. .. A universal coding method that permits the transmission of any sufficiently long message in an alphabet of s letters with no more than nh binary digits is not necessarily excessively complex; in particular, it is not essential to begin by determining the frequencies pr for the entire message. In order to make this clear, it is sufficient to note that by splitting the message S into m segments S1, S2,..., Sm, we obtain the inequality" .. p.2 various methods of "guessing continuations." .. It is more difficult to estimate the combinatorial entropy of texts subject to definite, more elaborate constraints. It would, for example, be of interest to estimate the entropy of Russian texts that could be regarded as sufficiently accurate (in terms of content) translations of a given foreign-language text. …
* A.N. Kolmogorov, Towards a logical foundation of information theory and probability theory, Problemy Peredachi Informatsii 5 : 3(1969), 3-7.
* A.N. Kolmogorov, Logical basis for information theory and probability theory, IEEE Trans. Information theory m (1968), 662-664.
* R.J. Solomonoff, **A formal theory of inductive inference,** Information and Control 1 (1964), 1-22. [Part I] … *PSG - phrase structure grammars .. p.3 "The "solutions" that are proposed involve Bayes' Theorem. A priori probabilities are assigned to strings of symbols by examining the manner in which these strings might be produced by a universal Turing machine. Strings with short and/or numerous "descriptions" (a "description" of a string being an input to the machine that yields that string as output) are assigned high a priori probabilities. Strings with long, and/or few descriptions are assigned small a priori probabilities"… p.5 Occam's razor; the principle of indifference (="indifferently equal a priori" likelihoods of the inputs), and the inversion of Huffman codes. P.8. "In general, any regularity in a corpus may be utilized to write a shorter description of that corpus. Remaining regularities in the descriptions can, in turn, be used to write even shorter descriptions, etc." 4.1…, a Bernoulli sequence. .. any "regularities" (i.e., deviations of the relative frequencies of various symbols from the average), result in shorter and/or more numerous descriptions. P.19 "probability evaluation method" (PEM) and normalized PEM. P.22 Acknowledgments: **Many of the basic ideas on induction that have been presented are the outgrowth of numerous discussions over many years with Marvin Minsky.***

https://www.sciencedirect.com/science/article/pii/S0019995864902232?via%3Dihub
See the remark of Minsky from a 2010 panel discussion, mentioned in the paper reviewed below: *Algorithmic Probability-Guided Machine Learning on Non-Differentiable Spaces.*

* R.J. Solomonoff, **A Formal Theory of Inductive Inference.** Part II, Reprinted from INFORMATION AND CONTROL. Volume 7, No. 2, June 1964, pp. 224–254. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=bca98303fcc3ef7d8b482ad23712f6e26c2140cb p.21: *"The problem of finding a grammar that can best be used to describe a set of strings (i.e., that results in a description having highest possible a-priori probability)".. promiscuous grammar – all sequences of symbols drawn from the alphabet are acceptable; "ad hoc grammar – only the given sequences are acceptable sentences, and no others.p.6. "the relative probability of the symbol A following the sequence α, rather than the symbol B following α, will be approximated by the total apriori probabilities of all sequences of length m + n + 1 that start out with the sequence αA, divided by the corresponding expression for sequences that start out with αB"*

* L. A. Levin, **Universal Sequential Search Problems**, 1973
https://lance.fortnow.com/papers/files/Levin%20Universal.pdf

* L.A. Levin. Laws of information conservation (nongrowth) and aspects of the foundation of probability theory. Problemy Peredachi Informatsii, 10(3):30–35, 1974.
http://alexander.shen.free.fr/library/Levin74_LawsOfInformationConservation.pdf
randomness deficiency, ..

* Leonid A. Levin, One way functions and pseudorandom generators, 1987, Combinatorica
The one way functions are *easy to compute, but hard to invert on a non-negligible fraction of instances.* https://dl.acm.org/doi/pdf/10.1145/22145.22185
* Leonid A. Levin, On the notion of random sequence, 1972
https://www.cs.bu.edu/fac/lnd/pdf/ran72.pdf
"P. Martin-Lof ..introduced a separate criterion (test) of randomness" … sequences of natural numbers, finite (corteges); seq. x,y; one is a beginning of the other – coordinated seq; computable operator … semicomputable measure R(x) - a priori probability of the sequence x. A sequence a is random with respect to the distribution P …

* P. Martin-Löf, **The definition of random sequences**, Information and Control 9 (1966), 602-619. MR 36 #6228.

* Compression is Comprehension: The Unreasonable Effectiveness of Digital Computation in the Natural World, Hector Zenil

https://arxiv.org/pdf/1904.10258 9 Jun 2021
Tosh: Many conclusions and reasoning in this work match TUM 2001-2004. "Algorithmic" in TUM is also "computer" (The Universe Computer), "predictable", tendency towards higher precision, range, resolution etc. of the predictions. E.g.: **H.Zenil, 2021:** *"Science moves empirical data from apparent randomness to algorithmic non-randomness .. Figure 9:*

Science moves apparently random data from observation to model/theory. Indeed, scientific models tend to cover previously random data together with other previously unexplained phenomena. … We have made great progress at taming apparent randomness since ancient times … *in some sense, science can be illustrated as the practice of moving natural phenomena from what we as observers perceive as random, towards nonrandom phenomena. .. The fact that these models compress more observations from natural phenomena in a reduced number of models is an indication that, indeed, compression is comprehension. "evidence .. also suggests that what we think is noise is often a signal whose source is unknown or irrelevant to a system of interest. … Time has ~~tell~~ told again and again that every time there is two seemingly disparate phenomena, they are often sides of a common underlying duality or symmetry … Computer programs can be looked at from a certain vantage point as laws of physics."* (…) *my team and I aim to tackle a fundamental challenge in science: that of developing tools for causal discovery. This in order to unveil the design principles and generating mechanisms of arbitrary dynamic systems. .. interventional calculus based upon the theory of computability and algorithmic probability .. p.17-19.. In other words, we gain much more explanatory and even predictive power with an increasingly compact number of theories over time. This is not only a clear indication that our world and universe is removed from randomness … Indeed, that the universe can be explained with theories and models of apparently ever-decreasing length means that the data they describe is formally of low algorithmic randomness, and that finding ever smaller models among all possible computable models is algorithmically very likely, which in turn, by the coding theorem, means that the world is algorithmically simple.*

*** A Review of Methods for Estimating Algorithmic Complexity: Options, Challenges, and New Directions** † , Hector Zenil, 5.2020 https://www.mdpi.com/1099-4300/22/6/612 Alternatives of statistical lossless compression (LZW etc.) – CTM, BDM. … RLE .. 1.1 *Like* [the entropy] *H, algorithms such as LZW cannot characterise algorithmic randomness because they implement a version or function of Shannon entropy (entropy rate), not only because they are not universal in the Turing sense, but also because they are designed only to characterise statistical redundancy (repetitions). However, accounting for non-statistical regularities ought to be crucial in the theory of algorithmic information, since these regularities represent the chief advantage of using algorithmic complexity. … The incomputability of C and AP [algorithmic probability] … The original formulation of algorithmic probability is of fundamental interest to science because it can address some of its most pressing challenges, such as inference, inverse problems, model generation, and causation, which happen to be the topics of interest in our research programme (beyond simple estimations of algorithmic complexity represented by a single real-valued number). .. resource-bounded Kolmogorov complexity .. Minimum Message Length – separating data from the model; and Minimum Description Length (MML, MMD) .. algorithmic complexity is one of those concepts … different formulations based on* ***uncompressibility***, ***unpredictability*** *and* ***atypicality***, ***converge**** … ***Statistical methods capture redundancy in the form of string repetitions, with no connection to state variables, unlike the generation of a set of algorithmic/mechanistic generative mechanisms from CTM, which can be regarded as***

computable models of the data. CTM+BDM combines the best of the statistical and algorithmic approaches. Notice that algorithmic also implies statistical because any statistical feature is also (trivially) algorithmic, though the reverse is not true. 5. Conclusion: .. *Future efforts to measure algorithmic complexity should focus less on obtaining single real-value numbers at any cost, as when using statistical lossless compression, and instead work to imbue with a deeper purpose applications meant to help with fundamental questions in science.*

Todor, 19.7.2025: Reaching to uncompressibility and unpredictability is an indication of a point of convergence or reaching to a limit; this is a signal to wrap up or finish current operation, increase the range and span (or reduce it), change the direction, complete a given phase of search or expand or shrink etc. In brief: **reorient.** In Zrim: дъно или връх (bottom & peak/top), depending on the search direction.

* Vitányi, P.M.B**. How incomputable is Kolmogorov complexity?** Entropy 2020, https://www.mdpi.com/1099-4300/22/4/408

* Schnorr, C.P. A unified approach to the definition of a random sequence. Math. Syst. Theory 1971, 5, 246–258.

*** Calculating Kolmogorov complexity from the output frequency distributions of small Turing machines**, F Soler-Toscano, H Zenil, JP Delahaye, N Gauvrit - PloS one, 2014 - https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0096223&type=printable

* The Information-theoretic and Algorithmic Approach to Human, Animal and Artificial Cognition, Nicolas Gauvrit, Hector Zenil, Jesper Tegnér, 2015/1/17, Book, Representation and reality : Humans, animals and machines, Springer https://arxiv.org/pdf/1501.04242 … *passing the Turing test is a trivial computational problem … the computational nature of the human mind more than it does on the challenge of artificial intelligence … the need to consider the question of resources and therefore of computational complexity. This means that the mind harness mechanisms to compress large amounts of information in efficient ways* … 1.1. *The Turing test is trivial, ergo the mind is algorithmic .. we will show that the human mind behaves more like a machine .. Block decomposition method BDM, ACSS (short strings). .. Coding theorem method … cognition & complexity … Our results in [Gauvrit et al. 2014c], suggest that the mind is intrinsically wired to believe that the world is algorithmic in nature, that what happens in it is likely the output of a random computer program and not of a process producing uniform classical randomness. .. p 12.: In many ways, animal behavior (including, notably, human behavior) suggests that the brain acts as a data compression device. .. Cognitive and behavioral science deals for the most part with small series, barely exceeding a few tens of values. For such short sequences, estimating the algorithmic complexity is a challenge.*

Tosh: Compare the compression conclusion with TUM. Also, thus you should or could work with sufficiently uncompressed or normalized representations in order to find “juice” that can be compressed.

* **Complexity Measurement Based on Information Theory and Kolmogorov Complexity,** Leong Ting Lui, Germán Terrazas, Hector Zenil, Cameron Alexander, N. Krasnogor, 2015 https://www.researchgate.net/publication/271533265_Complexity_Measurement_Based_on_Information_Theory_and_Kolmogorov_Complexity
Shannon - entropy, Kolmogorov – Alg.Complexity, Zurek - physical entropy = Shannon's+Kolmogorov's: H(d) + K(d). Gell-Mann & Lloyd: effective complexity. Korb and Dorin – **Controllability complexity** - *not just the complexity of the event itself, but also the complexity of the system that generated it.* Minimum Message Length. Introducing **EC, emergent complexity:** *amount of additional complexity a system adds to the initial states measured* – a merging of Shannon information and Kolmogorov complexity. … cellular automata: the output states evolve into: periodic states, random patterns; both ordered and random patterns. … porphyrin molecule self-assembly. Porphyrins: *planar molecules with fourfold symmetry and a chemical structure comprising four structural units that can be synthesized with substituent functional groups… diverse self-assembly complexity together with a high degreeof reversibility and highly dynamic pattern formation;* porphyrin-tile , isofunctionalized and heterofunctionalized; position, motion, and rotation; Monte Carlo family of methods; porphyrin-tile kinetic Monte Carlo (kMC) system; *a list of the possible transitions of the system (deposition, motion, and rotation of a porphyrin molecule) and their corresponding rates*.. 5.1 Complexity Analysis … both the outputs of the cellular automata and *the porphyrin simulation results are converted intoPNG images and optimally compressed using pngcrush* … into a new measure of **emergent complexity:** EC; diversity of the messages.. *an index of emergent, or excess, complexity(EC) of output states in relation to their corresponding input states, that is, the amount of addi-tional complexity a system adds to the initial states measured.* **p.3** *"Korb and Dorin [11] argued that when analyzing the emergence of complex events one should consider not just the complexity of the event itself, but also the complexity of the system that generated it. "*

* Korb, K. B., & Dorin, A. (2011). **Evolution unbound: Releasing the arrow of complexity. Biology and Philosophy**, 26(3), 317–338
– **Compare to** TUM, 2001-2004, starting from:
* "Man and Thinking Machine: Analysis of the Possibility a Thinking Machine to be Created and some Disadvantages of Humans and Organic Matter", T.Arnaudov, 2001, The Sacred Computer – the **complexity** of a cell vs computers, machines, human artefacts. Machines or computers are actually **more complex** than a cell, because they require the Universe first to create societies of thinking beings – humans – millions and billions of them, which together to work for thousands of years, to produce intermediate instruments, earlier versions of tools and machines and other humans, and finally, by orchestrating many advanced machines, to invent, develop and generate all the required science, electronics, engineering and the organization between these billions of humans with their "*more complex*" cells and brains, which individually as designs were created by the Unvierse billions or hundreds of millions of years earlier and starting from a supposedly much "less complex" state. That reasoning is related to Charles Benett's Logical Depth (see below) - I didn't know the concept at the time (discovered on ~3.7.2025).

* Zurek, W. (1989). **Algorithmic randomness and the physical entropy**. Physical Review A, 40,4731–4751.
* Zurek, W. (1990). **Algorithmic information content, Church-Turing thesis, physical entropy, and Maxwell's demon**. In E. Jen (Ed.), Lectures in Complex Systems, SFI Studies in the Sciences of Complexity,Vol. II (pp. 49–65).
* Korb, K. B., & Dorin, A. (2011). **Evolution unbound: Releasing the arrow of complexity.** Biology and Philosophy, 26(3), 317–338
* Wallace, C. S. (2005). **Statistical and inductive inference by minimum message length.** Springer-Verlag.
* Lopez-Ruiz, R., Sañudo, J., Romera, E., & Calbet, X. (2011). **Statistical complexity and Fisher-Shannoninformation: Applications**. In K. D. Sen (Ed.), Statistical complexity (pp. 65–127) 12.1.2012

Algorithmic Probability-Guided Machine Learning on Non-Differentiable Spaces, Santiago Hernández-Orozco, Hector Zenil, Jürgen Riedel, Adam Uccello, Narsis A Kiani, Jesper Tegnér, 2021/1/25, Frontiers in artificial intelligence https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2020.567356/full robustness and fitness; algorithmic probability & thermodynamics of computation .. *algorithmic search to navigate .. discrete non-differentiable spaces* ... identification of generative rules from data observations; .. more resilient against pixel attacks compared to neural networks; AP loss f . .. 3 Categorical Algorithmic Probability Classification … an algorithmic information model must output an object that minimizes the algorithmic distance to all the members of the class, so that we can classify them. .. " However, in one of his latest interviews, if not the last one, Marvin Minsky suggested that the most important direction for AI was actually the study and introduction of Algorithmic Probability (Minsky, 2014)."* – actually the cited one is from 2010, but published on Youtube in 2014. See a later one from 2013:
* Minsky, M. (2014). **The Limits of Understanding.** Panel discussion (New York, NY: World Science Festival, ~1:30 h Minsky: https://youtu.be/DfY-DRsE86s?t=5389 also with G.Chaitin et al. *The most important discovery after Goedel* … Algorithmic Probability … *How to make predictions, given a collection of experiences…* A bit earlier M.Minsky explains that in the early days of computing they tried out the first thousands random Turing machine and they discovered that most of them didn't do intereseting things, except one which was counting, but that implies that arithmetics is one of the simple operations, which is discovered early even in so simple enumerative process.
* **Marvin Minsky on AI: The Turing Test is a Joke!,** Socrates / Podcasts, https://www.singularityweblog.com/marvin-minsky/ , **Posted on:** July 12, 2013 / **Last Modified:** June 10, 2022
Todor: The search in the field of **Program Synthesis** is in non-differentiable spaces, in comparison to the "differentiable" space of functions where smooth, continuous, gradients are computed as in the ANNs; see the literature, reviewed in the appendices **#lotsofpaper** and

#**listove** → **#bongard** etc.
* See also H.Zenil, ***A Review of Methods for Estimating Algorithmic Complexity: Options, Challenges, and New Directions***.
→ * 25. M. Mitchell. What Is Meant By "Biological Computation"? ACM Communications Ubiquity. vol 2011, number, Pages 1–7, February, 2011
* H. Zenil and J-P. Delahaye. **On the Algorithmic Nature of the World**. In G. Dodig-Crnkovic and M. Burgin (eds), Information and Computation. World Scientific Publishing Company, 2010. [see below]
* H. Zenil. Information **Theory and Computational Thermodynamics: Lessons for Biology from Physic**s. Information. In revision. [see below]

*** On the possible computational power of the human mind**, Hector Zenil, Francisco Hernandez-Quiroz, 2007, https://arxiv.org/pdf/cs.NE/0605065 - ANN, SNN/PNN (spiking or pulsed NN) … Analog Recurrent NN (ARNN), … p.10: [Proposed reasons] Why minds are not computers (…) *(a) There are special operations that occur in the brain which are not Turing computable, a claim made most famously by Penrose . (b) The mind could be a machine but with access to a certain oracle (from an external source or from a previously coded internal source).* p.19. The computational power of the mind could be determined if some important properties of the physical universe and the mind could be first determined such as: * discrete or continuous * can it hold non-computable numbers and operations * does the brain inherits those properties; * does the mind takes advantage of them, *which depends on the precise relationship between the mind and the physical operation of the brain.*

* **Empirical Encounters with Computational Irreducibility and Unpredictability,** Hector Zenil, Fernando Soler-Toscano, Joost J. Joosten, 4.2011/6.2011 https://arxiv.org/abs/1104.3421 … decidability, intractability, nonlinearity; halting problem; Irreducibility measures. *Blum's speed-up theorem asserts that there exist problems for which finding the optimal algorithm cannot be achieved ..* ***David Deutsch(1998):*** *"The theory of computation has traditionally been studied almost entirely in the abstract, as a topic in pure mathematics. This is to miss the point of it. Computers are physical objects, and computations are physical processes. What computers can or cannot compute is determined by the laws of physics alone, and not by pure mathematics."*

Elementary cellular automata; impossible to find shortcuts for computing the final state and all steps have to be passed in a **complete detailed simulation**. .. *most non-trivial systems would require a simulation with roughly the same number of steps as that of the simulated one.*

Tosh: Other shortcuts for predicting faster: I see at least one other **shortcut**, especially for "systems"*: **if the differences at that finer or highest level of resolution is actually *irrelevant*.** A more advanced causality-control unit may understand that it doesn't need to generate all pixels and all frames, it needs only a summary, borders, a rough classification of some properties etc. and the target effect would be the same, with given selected criteria, as if it computed at the finest resolution. The mind doesn't have to remember the details either, because it knows that they will be sufficiently stable and they can be sampled when needed and the Universe Computer will take care of the intermediate computations. The noise in the simplest sense, which is often said to be "incompressible", "unpredictable", "random" in either algorithmic complexity and normal statistical sense, for example is actually extremely compressible down to its distribution which is simpler than the one of sufficiently "realistic" and broad "real data"; if a noise pattern covers a particular region of an input space as a layer over another signal, as a mask which covers the other data, then the location and shape of that region has also to be encoded. However the exact intensities and pixel values of the noisy area shouldn't matter if the pattern is really noise or considered as such, because if the finest details of the signal considered "noise" also matter and have to be remembered exactly, then this "noise" should be relabeled to *signal*, it is not noise for the evaluator-observer anymore. See also TUM.UM2, 2002, the three types of prediction in the Unvierse Computer, a letter from Todor to Angel, 27.8.2002:((…) *In fact,* ***if we run a program on any computer and the program doesn't need "side" information which should be entered during the computation of the solution – that is, we are given a "closed universe", where WE KNOW EVERYTHING on which the future depends, then the final outcome – the answer of the problem = the future is unique and it is clear right in the moment when the program is initiated. It is clear to someone, who: (1) can use a more advanced computer, which would compute the answer before the first computer (2) someone who doesn't have a faster computer, but has found a simpler solution thus he can get the solution faster than the first, or (3) who knows the answer beforehand, because he has already solved the same problem… We always predict using the second method. Even I can predict the future using the first one – I program the same algorithm on my old Pravetz-8M and on my current, also old :), computer. I start the program on the Pravetz, then I start it on the current one – it outputs the result before the Pravetz and now I know THE FUTURE of the computer – I know what exactly will be printed on its screen!*** [Bold is from the original text]

** Advanced here is "faster", but it could be also one using more powerful instruction sets or representations; in the given frame it would convert to time. [Note 6.7.2025]

* Not specifically Turing machines, which are a limiting abstraction; they can emulate each others and Turing machines, but the *emulator* should also be provided or developed. IMO it is more productive to think and operate with more powerful and developing abstractions. See works to be published in the future: **"Creating Thinking Machines" (Създаване на мислещи машини)**

* N. Israeli and N. Goldenfeld, Computational Irreducibility and the Predictability of Complex Physical Systems, Phys. Rev. Lett. 92, 2004.

* Blum, Manuel (1967). "A Machine-Independent Theory of the Complexity of Recursive

Functions" (PDF). Journal of the ACM. 14 (2): 322–336. doi:10.1145/321386.321395. S2CID 15710280.
* D. Deutsch. **The Fabric of Reality: The Science of Parallel Universes and Its Implications**. Penguin (Non-Classics), 1998 [first published 8.1997] See Ch. 5: Virtual Reality, p. 71 - p.87 … universal image generator, accuracy, … *If an image generator is playing a recording taken from life, its accuracy may be defined as the closeness of the rendered images to the ones that a person in the original situation would have perceived.* (or to intended one in case of cartoons etc.) *If the rendering is so close as to be indistinguishable by the user from what is intended, then we can call it perfectly accurate. (So a rendering that is perfectly accurate for one user may contain inaccuracies that are perceptible to a user with sharper senses, or with additional senses.)* … The VR should also respond in an intended way to all possible user's actions; etc.
https://ia801208.us.archive.org/24/items/TheFabricOfReality/The_Fabric_of_Reality.pdf
See also: * **Stanislav Lem**, "**Summa Technologiae", 1964**, "phantomatics" (фантоматика);
* See also T.A., TUM, 2001-2004, "Letters between…"/The Universe Computer 2, 2002; The Truth, 2002, **"The Matrix in the Matrix is a matrix in the matrix…",** 2003; Universe and Mind 3, 2003; Universe and Mind 4, 2004. … Deutsch's discussion about accuracy is related to TUM's resolution of perception and control (causality) (РСВУ), "true/real causality control" at the highest possible resolution or the lowest level from the POV of a given causality-control unit. The coverage of all possible actions is general for all virutal universe simulators: "what can be done" (?ВМдсП) in Zrim.

* **Life as Thermodynamic Evidence of Algorithmic Structure in Natural Environments**, Hector Zenil 1,*,Carlos Gershenson 2,James A. R. Marshall 1 and David A. Rosenblueth 2, Entropy, 11.2012, https://www.mdpi.com/1099-4300/14/11/2173?source=techstories.org https://doi.org/10.3390/e14112173 – *We contribute to the discussion of the algorithmic structure of natural environments and provide statistical and computational arguments for the intuitive claim that living systems **would not be able to survive in completely unpredictable environments**, even if adaptable and equipped with storage and learning capabilities by natural selection (brain memory or DNA). ..*
2.1. What can be learned about the relationship between the information content of a stochastic environment and its degree of predictability and structure versus randomness from the way in which organisms gather information from it in order to survive and reproduce? .. requisite variety … the foraging strategies, how the environments are scanned: Brownian motion, exhaustive search … *2.3. Markov chains: An organism's representation of the world, however, is always limited, not only because it has access to limited resources, but also because **organisms can only process a finite amount** of relevant information in order to make a decision. 3. Computation and Life … Computers may be thought of as engines transforming energy into information (the information may already be there, but without energy one is unable to extract it.* 4.1. Simulation of Increasingly Predictable Environments .. 4.3. Organisms Survive (only) in Predictable Environments* .. Representational systems – prediction; Reactive systems – adaptation; *computationally, either system requires the same*

environmental predictability in order to survive, independent of the particular strategy or mechanism used.

Todor, 3.7.2025: The requirement of predictability for the development of minds is explained/claimed in TUM, since 2001-2004 (and by that theory a mind (intelligence: razum (Разум)) is present at all scales at different degrees and different manifestations, where the basic/main principles for maximizing prediction and match, causality-control etc. are preserved at all scales). Life is a way for implementing nested, hierarchical, higher level causality-control units: agents, having higher degrees, levels of cognition, intelligence, "computation" capabilities.

These capacities would be hindered not only in "*completely* unpredictable" environment, but also in unpredictable *beyond* some "borderline" and also if the unpredictable events lead to destruction – if they are, or they start at low levels of the Universe computer, the collapse after an "error" would be quick, i.e. if the chemical and physical forces fluctuated in a way that couldn't be encoded or amortized sufficiently within a cell, the cells would just die or wouldn't form at all. Such catastrophe happens at once when the cells are "attacked" by "aggressive" chemicals which cause too quick and powerful chemical reactions, e.g. acids; similar sequences unfold in case of exposure to higher temperatures than the tolerable, too low temperatures etc.

Both the possibilities in the environment: predictability – in theory – and the *practical* ***capability*** of the agents to predict in *that environment* are required. Both would shape what "ecological niches" could and will be constructed and could exist and then persist, what kind of organisms could develop and how they could navigate with what amount of "complexity" and "reservoirs" for added one, built-in locally in their bodies or delegated to the environment (see "ecological psychology", "enactivism", "emobiment" etc., for example the notes in the section for Schools of thoughts in the early parts of The Prophets..) etc.

*** 5.7.2025: Energy is already information:** where it is located, what is its amount, its form of storage (potential, kinetic, electric, magnetic, chemical – and all specific details about that). Computers and all other processes or systems convert different forms of energy = different forms of information. In TUM: for example, food is also information and data for an organism; this type of information however has to be encoded in a specific format in order to be "downloadable". The process of consumming food and extracting its content is "decoding" or "downloading", similar to decompression or decrypting, and spreading the "packets" through the "cellular network" for each cell, mitochondria and eventually all molecules to "read" and process by the "computers" at all scales in the system, down to atoms and electrons and possibly the quarks etc. Similarly to a compressed file which is decompressed and then its bits can be imagined or represented as traveling through the electric vessels of the processor and RAM and affecting the logical gates and individual transistors, which process or store them. The Universe is a Computer, therefore everything is information and machine code in our informationalist-computationalist view. On informationalism see also Gordana Dodig-Crnkovic, "*Where do new ideas come from* (…)", 2007 cited below. See also Free Energy Principle, Active Inference by Karl Friston et al.

*** Algorithmic probability of large datasets and the simplicity bubble problem in machine learning,** Felipe S Abrahao, Hector Zenil, Fabio Porto, Klaus Wehmuth, 22.12.2021, 4.2023 (extended version) https://arxiv.org/abs/2112.12275 … *data generating source; externally generated events; partial computable function; infinite computably enumerable set; "**too much information** tends to behave like **very little information**" phenomenon and the bias toward simplicity .. worst-case scenarios in which overfitting is expected to always occur once the datasets get sufficiently large, even if the datasets are generated following the universal distributiontwo proofs … if a locally optimal model found by an arbitrarily chosen machine learning algorithm using a sufficiently large dataset has an **algorithmic complexity arbitrarily smaller than that of the available data**, the algorithmic probability of **the deceiving dataset still dominates** the algorithmic probability of any other randomly generated dataset of size greater than or equal to the size of the deceiving dataset. .. a **deceiving dataset** is one for which the learning algorithm (in addition to any formal theory and performance measure one may choose to embed into this algorithm) generates an optimal model that completely satisfies the assumed performance criteria to be considered a global one, but that in fact is a locally optimal model... **As available data becomes sufficiently large**, its underlying generating process (or mechanism) likely becomes more complex than the learning algorithm. … the data generating process may "trap" the learning algorithm into a "simplicity bubble" in which any solution found from the available data alone is necessarily not complex enough to match the complexity of the actual data generating process. Only by adding a larger amount of (unpredictable) fresh data, can the learning algorithm hypothetically be capable of "bursting" this bubble p.*5 .. p.8 ***the machine learning process:*** *.. a well-defined computable sequence (or combination) of total computable functions (and, in some cases, calls to stochastically random number generators) that finally returns a model at the end of this process. …* p.35 **How to avoid the simplicity bubble?** *..the complexity of the phenomena about which the learning algorithms can make predictions cannot be greater than (up to a small constant) the complexity of the learning algorithms themselves. … … Thus, **the algorithms for sampling and retrieving the training set, test set and validation set, calculating the loss/cost/error functions, applying gradient descent, crossvalidation, regularisation, dropouts, early stopping, bias-variance trade-off, hyperparameter tunning, etc are already embedded into P**. … **Complexity caging principle/hypothesis**. … p.32-33: given a bias toward low algorithmic complexity (or high algorithmic probability), no computable learning process can in general give reliable predictions for any randomly generated computably constructible collection of data. That is, there is always an **optimization limit** such as with the no free lunch theorems.*

Tosh: Bold: mine. Note that in static ML, with stationary distribution, the actual genesis of the data and how it was actually generated, is missing, and in the general case the genesis is from limited points of view, so the reconstructed model is "wrong" in that regard even if it reconstructs correctly the data in the artificial setting and dataset. A learner may be optimal, but in its fantasy world of segmentations and features, as shown with adversarial attacks etc.
* The complexity of the learning algorithm vs data: compare to Ashby's Requisite Variety.
* In TUM: the more advanced a mind and the bigger the data/memory of an intelligent system (a mind, agent, causality-control unit), the more the new input can be interpreted as "*non-procedural knowledge*" and more varying meanings and possible intepretations can be deduced or generated by the mind, because the system has higher complexity, operates at higher levels of abstractions etc. In such circumstances the interpretation of the meaning is more dependent on the knowledge and the state of the intelligent system, rather than the literal content of the actual input as the ratio between the complexity of the two grows^. Similarly with the behavior: the bandwidth of the effective control signals from the "consciousness" to, say, the hands of a human, can be represented as a dozen or a few dozens of bits, while the description of the actual motions in the "real" Universe should include the trajectories and states of all particles which the Universe can discriminate at its own maximum resolution of causality and control and this task requires an astronomical amount of bits of "special type" in the machine language of the Universe. Converesely, the entity that actually controls, causes the future, of the physical representation is not "the consciousness" or "the brain", but a system that includes the whole body and the Universe, or the latter two can become entangled with the consciousness. This phenomenon is exemplified in TUM also with the generative fairy tales about genies which fulfill wishes and that is repeated with the prompts in modern AI: the prompt can be miniscule in comparison to the parameters and the information in the generative model. Another related phenomenon is the controllability, the "alignment" – the model builders and users want to control with a high precision a system of billions and trillions parameters with complex, multi-step and deep relations between them and astronomical amount of calculations, by using a few bits which they can type in a few seconds. That is possible, but then the actual complexity of the model will also collapse or the "masters" will be fooled that the "genie" is following their orders "exactly" (enough), because their own resolution of causality-control is sufficiently low and they cannot discriminate the difference and details or/and don't care about them.
^ Except if the input is specifically directed towards *precise interventions*, exact instructions etc. or it controls such facilities which cause such changes.

See also *** Multiclass classification utilising an estimated algorithmic probability prior,** Kamaludin Dingle,1, 2, ∗ Pau Batlle,1 and Houman Owhadi1, 12.2022, https://arxiv.org/pdf/2212.07426 p.3 Algorithmic complexity and probability, etc.: *"Thus, high-complexity outputs have exponentially low probability, and simple outputs must have high probability."*

Todor, 10.7.2025: This is the "default", however it is true only in a general case and overall, possibly, for some whole "universe" which is being analyzed where this probabiilty was actually measured to hold; also it is true in a simplified case, if counting prefixes, e.g. a longer string of bits is constructed from subsequences of shorter strings, and the shorter strings are *parts* of the whole, thus they have a lower complexity and length, compared to the whole string. However for example the structural complexity of the typical molecules inside a living organism is *higher* than the one in the surrounding air or water – if we don't compute it with "prefixes", the organic molecules are built by atoms, whose individual "complexity" is lower; however at their smaller scale the atoms are "diferent" and *uncomparable* to molecules, because in that configuration they **lack the structure**, which molecules have and which causes dramatic effects in the interactions between the particles. The structure is related also to the scale; the bigger molecules can act and interact as a system and a whole on many locations with many radicals like a parallel computer with more cores or threads and different atoms and radicals may have vastly incompatible behavior and cant't be interchanged. The molecules in the space of the Universe where we can access them, for now and for our "readouts" of their states are not just "strings of bits"* and their abstract representation with a code book of letters etc. is not really describing the process of their initial genesis and assembly and how difficult and rare, up to impossible, the molecular synthesis is in a "general" case, outside of a living organism, without special calatytic compounds, temperature, a sequence of precise RNA-DNA operations etc.

In brief**, inside an organism, it is *more probable* to find *algorithmically more complex* "strings of atoms", than simpler ones,** if measuring it with a general or average scale, which includes in the calculation the "empty spaces" and the matter "at the lower scales of organization".

* If the atoms and molecules, their configurations, location and energy could possibly be recorded as bits by out rechnology, they will need also the bits for the surrounding environments and generative processes, and in order to turn again into atoms and molecules, a "materializer" interpreter or compiler is needed, a "replicator" which assembles or even produces the elementary particles.

* See also the **Assembly theory**, reviewed in the Appendices of The Prophets of the Thinking Machines #listove; H.Zenil et al. criticize it for being only a form of compression algorithm.

* Another example is if you are located inside a **cryptography table** or a compressed file – all strings there have a higher complexity, compared to "the average" or a particular uncompressed text file on the computer. However, if you *know* that you are *already inside* an encrypted or compressed environment, then you should *expect* that it will be "complex", so encountering a low complexity string *now* will be *more unexpected*, thus of *lower probability.* When the *context* changes, the distributions and expectations change and a more granular model should recognize it and switch accordingly. A "solution" or an *excuse* to this contextual difference is using "random Turing machines" or whatever machines and large scale or "infinity", thus they are uninformed or the local variations become irrelevant etc., then they would be correct when processing all contexts etc. but this seems like a limitation for simplification, which "feels" related to the "simiplicity bubble", mentioned in the previous paper.

Note also that with the advent of computers, higher compression ratios are achieved, thus in its "dormant" state as a file, the pieces of data have a higher internal complexity than the *local* views of the decompressed data, which is finally printed, rendered, played etc. Executable files are often obfuscated in order to make their reverse engineering harder. In addition, encryption of increasing difficulty is more often applied on any data: on the Internet, locally and for the whole disk drives, in the GPUs for "texture compression", image, video etc. and during training, the weights of the artificial neural networks also turn into a "noise-like" blob of bits with a high complexity, as they are building a compressed representation. One reason for this is that they are "distributed" representation and the local values per "neuron" or a weight are influenced by many other "neurons" or parameters, which may be located a "long distance" away in the serialized representation; furthermore, the representation itself is, say, floating point numbers, which are an encoding themselves etc.

Therefore on a modern computer it is more probable to find "higher complexity" files and bit sequences, than lower, given a particular scale or size of the span of data, and not "cheating" with prefix-excuses (shorter strings are simpler than the longer ones).

Another trend is that with the development of the systems, they operate with longer elemental sequences, their chunking, tokenizing procedures aggregate or build bigger structures with "one" operation or "per …" some metrics, and at a particular level of causality-control, complexity of the lower level components doesn't matter. The "assembler" or "processor" doesn't care about the complexity of the **content bits** of the elements, but only of the **structural bits** – recall the example with the molecules.

(…)

* See also CTM & BDM in the other papers (Coding Theorem Method, …)
See also the classical TUM for more general informal discussion on complexity and randomness.

* **Non-Random Data Encodes Its Geometric and Topological Dimensions**, Hector Zenil, Felipe S. Abrahao, Luan Ozelim, https://arxiv.org/pdf/2405.07803 20 Dec 2024
– Signal deconvolution; preferrably non-compressed input, finding its dimensions by **pertrubations** (transforms) .. *In this case in which the message is encoded in an uncompressed form, or equivalently in which the message is compressible (therefore non-random), one would have a large standard deviation in the range of information/complexity values for the different ways of decomposing, parsing or perturbing the signal. Algorithmic Information Dynamics … we estimate the algorithmic probability changes for a message under distinct partitions that result from perturbations .. perturbation analysis .. of compressed data; bitwise perturbations ..* [Reviewed: 5.7.2025]
* H. Zenil, N. A. Kiani, A. A. Zea, and J. Tegner, "Causal deconvolution by algorithmic generative models,", Nature Machine Intelligence, vol. 1, no. 1, pp. 58–66, jan 2019.

* Levin, L. A. (1974). Laws of information conservation (nongrowth) and aspects of the foundation of probability theory. Probl. Peredachi Inf. 10 (3), 30–35.
* Minsky, M. (2014). Panel discussion on the limits of understanding. (New York, NY: World Science Festival). **The Limits of Understanding,** World Science Festival, 1,35 млн.

Абонати, 1 663 852 показвания 15.12.2014 г. #briangreene #science #consensus **Original Program Date: June 4, 2010**

*** Natural scene statistics mediate the perception of image complexity**, Nicolas Gauvrit, Fernando Soler-Toscano, and Hector Zenil July 28th, 2014 https://arxiv.org/pdf/1509.03970 See Fig. 2, page 6, a sample of 100 binary patterns. Compare to the tokens in Visual LM.

* Lui, Leong Ting, Germán Terrazas, Hector Zenil, Cameron Alexander, & Natalio Krasnogor. 2015. **Complexity measurement based on information theory and Kolmogorov complexity**. Artificial Life, uncorrected proof. – various types of entropy and complexity … "*Mandelbrot set can be seen as a simple mathematical algorithm having a complex phase diagram*"

Tosh: if the complexity account includes the memory cells (their count, relations, possibilities for changing and possible values etc.), the relations between them and the actual performed calculations, the simple formula and representation is not so simple anymore. That's one of the tricks for appearing simpler, takint into account a compressed abstract representation in one hand, but then comparing to "exploded" specific representation. How much "bits" or instructions and calculations are needed in order to describe the computer hardware, the program for it, the human who discovered and devised the formula and the display technology, needed to render the Mandelbrot set? See also Benett's *Logical depth* below.

*** Some Computational Aspects of Essential Properties of Evolution and Life,** Hector Zenil and James A.R. Marshall**,** Behavioural and Evolutionary Theory Lab**,** Department of Computer Science, The University of Sheffield, UK, 2.6.2012 https://arxiv.org/pdf/1206.0375

* Hector Zenil, **An Algorithmic Approach to Information and Meaning,** 30.9.2011, https://arxiv.org/pdf/1109.3887 *van Benthem's contribution (p. 171*). 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...* communication, diversity, complexity, data+program+interpretation .. *Turing on the basis of his concept of computational universality is the unification of the concepts of data and program .. C(s) – algorithmic complexity; Conditional C(s|x); logical depth; message interpretation – a program for a TM; halt or not halt …*
* L. Floridi (ed.) Philosophy of Computing and Information: 5 Questions, Automatic Press / VIP, 2008

*** Image Characterization and Classification by Physical Complexity,** Hector Zenil, Jean-Paul Delahaye† and C´edric Gaucherel, 3.7.2011, https://arxiv.org/pdf/1006.0051 (…) *2.2* ***Bennett's logical depth*** .. *combining the notions of algorithmic information content* ***and time.*** *.. the time that an unfolding process takes to reproduce the string from its shortest description. The longer it takes, the more complex. Hence complex objects are those which can be seen as "containing internal evidence of a* ***nontrivial causal history**** *". Unlike algorithmic complexity, which assigns a high complexity to both random and highly organized objects,* *placing them at the same level, logical depth assigns a low complexity to both random and trivial objects … a measure of organized complexity …* ***string's depth at significance d****… the* ***number of steps T(p) in the computation*** *… the length of the shortest program for s, (therefore K(s)). … * p.17 Figure 12: Compressed path: Going from all-white to all-black by randomly inserting black lines …* Compression tests and ranking with different classes of images (e.g. noise with different densities, fractals, faces, processor die shots, mandala, writing, city; outputs from cellular automata, tiles, simple patterns of all black or white, overlapping white and black rectangles etc. – p.20 (..) *plain algorithmic complexity does not help to distinguish between complexity associated with randomness and complexity associated with highly structured objects*; random images are not ***physically*** **complex**. .. Decompression time ranking; apparent physical complexity. *Unsupervised method for evaluating the information content of an image by* ***physical complexity****…*

Tosh: However the degree of triviality of the nontrivial causal history has to be defined and measured as well and it is not the same everywhere. For example a completely white piece of paper of an area of the land, given some sufficiently low variation, may seem "simple" or "trivial" *now* or it can be encoded as *"a white piece of paper"* or a program with a loop filling the memory with "255" or "1.0"; however it took the Universe Computer **billions of years** until producing the first sheet of white paper from its initial configuration, as long as current science has extrapolated the past.

Straight lines and linearity are considered among the simple mathematical dependencies, easier for mathematical abstract models and compuation, but they are hard for actual physical systems, where many forces apply and all of them are not linear, but quadratic, exponential, stochastic with various natural distributions etc. The apparent simplicity in the abstract representations is grounded by the complexity of the computer systems, the mind which have to compensate for the physical non-linearities with other reflected non-linearities, or as the artificial neural network do with their non-linearities and many iterations of adjustments until they find sufficiently correct settings for all parameters.

Noise that is considered "easy" or lacking structure also can be difficult to produce in some contexts, for example in computers with limited memory, especially with some desirable properties or a "real" one, if there is not a true source of randomness; the generation process may require a sufficiently big look-up table with enough samples from real noise – if it is not sufficiently big or smoothed, it may produce repetitive patterns and artifacts which may reduce the appearance as random or "noisyness" of the output. [5.7.2025;+19.7]

* 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. https://www.academia.edu/116589715/Logical_Depth_and_Physical_Complexity
Logical Depth: *the time required by a standard universal Turing machine to generate it from an input that is algorithmically random (i.e. Martin-L¨of random) .. obeying a slow-growth law: deep objects cannot be quickly produced from shallow ones by any deterministic process, nor with much probability by a probabilistic process, but can be produced slowly.*

* Charles H. Bennett, **Complexity in the Universe,** IBM Research Division, T. J. Watson

Research Center, Yorktown Heights, NY 10598. USA, https://cqi.inf.usi.ch/qic/94_Bennett.pdf
"Logical Depth, the plausible number of computational steps in an object's causal history, is the complexity measure we chiefly recommend. A logically deep object, in other words, is one containing internal evidence of having resulted from a long computation, or from a dynamical process requiring a long time for a computer to simulate …"
Todor: How many are *plausible* and for whom?

* **Human behavioral complexity peaks at age 25,** Nicolas Gauvrit, Hector Zenil , Fernando Soler-Toscano, Jean-Paul Delahaye, Peter Brugger**,** April 13, 2017
https://doi.org/10.1371/journal.pcbi.1005408 – *Random Item Generation tasks (RIG) are commonly used to assess high cognitive abilities such as inhibition or sustained attention. They also draw upon our approximate sense of complexity. A detrimental effect of aging on pseudo-random productions has been demonstrated for some tasks, but little is as yet known about the developmental curve of cognitive complexity over the lifespan.*

Tosh, 2.7.2025**:** Sequence prediction and generation are indeed general metrics for intelligence, creativity, originality, however I'd put them in other more complex and diverse experimens, such as making decisions and their qualities in more "real" or interactive environments; I'd measure the diversity (rareness, given some metric, e.g. a comparison: first, "qualified model creative explorers" will explore the space in multiple variations and will search, enumerate and record the possibilities; then the participants which would be subjects in the experiment will be asked to pass through similar conditins, environments, games etc. and their decisions, creative choices etc. will be collected and compared to each other and to the first "expert" or "model creative person" or a "model of complex behavior"; the time for figuring out some "creative decisions" or solutions will be measured and compared (more creative ones are supposed to figure out and construct more elaborate configurations faster; less creative may also reach to that solution, but after more steps, more time etc.).

Another direction – the human generated pseudo random numbers can be used for modulating the generation of "realistic" or "natural-looking" data, as this is done with Perlin noise and in Pixel Shader Art such as in Shadertoy. The "naturalness" can be measured using signal-to-noise ratio, image-based metrics (now possibly with vision-text models, CLIP, matching of features' embeddings etc.). One kind of such data is composing texts given directions, keywords, topics, parts of sentences, beginnings, constraints, size limits, images etc. – "prompts".

In the experiments, described in the paper, the subjects are asked to generate too simple sequences of "random" variables such as 12 heads and tails of coin tosses or similar series of dice throws, which they believe "would appear random for somebody else"; they are asked to guess what card would appear after random shuffle etc. (Table 1).

To me this kind of guessing is more like "cheating", "manipulative", deception and guessing what the others believe. This is not (*own*) *behavioral* complexity. **Anything could appear random or ordered to unknown person.** I disagree with the term "*behavioral complexity*" for what they have measured. See also my critics of the interpretation of the "Marshmallow test". An extended version in "The Prophets…"

*** Delayed Gratification is an ill-defined concept**, T.Arnaudov, 2018, https://artificial-mind.blogspot.com/2018/06/delayed-gratification-is-ill-defined.html

* Zenil, H. et al. **A Decomposition Method for Global Evaluation of Shannon Entropy and Local Estimations of Algorithmic Complexity**. Entropy 20, 605 (2018). https://www.mdpi.com/1099-4300/20/8/605 Hector Zenil, Santiago Hernández-Orozco, Narsis A. Kiani, Fernando Soler-Toscano, Antonio Rueda-Toicen – **introduces BDM** which extends the Coding Theorem Method (CTM). **CTM**: *"unlike common implementations of lossless compression algorithms, the main motivation of CTM is to find algorithmic features in data rather than just statistical regularities that are beyond the range of application of Shannon entropy and popular lossless compression algorithms"* blocks; sliding window size; object boundaries; approximations of Kolmogorov's Algorithmic complexity; short strings, long strings; … lossless compression; Shannon entropy's limitations – statistical. CTM: … *compressing very short strings, for which no implementation of lossless compression gives reasonable results.*

Tosh: Compare **Computational irreducibility, Wolfram.** Compression or redution is finite. The length of the code for compression or "it's complexity"* becomes longer or higher than just copying the data directly. *The representation may be more … complex, multi-faceted, hierarchical, graph-based etc. than just a string of bits of symbols from a single alphabet etc.; the complexity can be evaluated via a more … complex method etc. Apply BDM for program synthesis.

See also the works of **Hector Zenil** etc.:
https://scholar.google.com/scholar?q=Hector+Zenil
https://scholar.google.com/citations?user=P6z3U-wAAAAJ&hl=en&oi=sra

* **A Review of Methods for Estimating Algorithmic Complexity: Options, Challenges, and New Directions**, Hector Zenil, *Entropy* **2020**, *22*(6), 612; https://doi.org/10.3390/e22060612, https://www.mdpi.com/1099-4300/22/6/612 … See fig.1, a statistical approach: Run-length encoding RLE overn an unknown process vs reverse engineering the generative process; CTM, alternative to lossless compression for short string; LZW – not efficient for < 1000 bits etc. Lempel–Ziv–Welch, dictionary; Borel normality

*** Causal deconvolution by algorithmic generative models,** Hector Zenil, Narsis A. Kiani, Allan A. Zea & Jesper Tegnér **,** Nature Machine Intelligence volume 1, pages58–66 (2019) https://www.researchgate.net/publication/330203174_Causal_deconvolution_by_algorithmic_generative_models Algorithmic similarity – not only statistical; hierarchical decomposition, causal clustering, inverse design problems, simulation-Based Inference, intractable inference; efficiency and inference quality; "forced deconvolution";

*** A Computable Universe: Understanding and Exploring Nature as Computation**, Editor Hector Zenil, 2012, Singapore: World Scientific Publishing Company/Imperial College Press (dozens of authors of articles, including Fredkin, Schmidhuber, Zuse, Chaitin, Wolfram etc. a

section with open discussion)
https://books.google.bg/books?hl=en&lr=&id=SGG6CgAAQBAJ&oi=fnd&pg=PR7&dq=Hector+Zenil&ots=3ksMjCrQ3l&sig=NToU0OqmtChlcGmbD5WJ6aDPac0&redir_esc=y#v=onepage&q=Hector%20Zenil&f=false

* **The Future of Fundamental Science Led by Generative Closed-Loop Artificial Intelligence**. Hector Zenil, Jesper Tegnér, Felipe S. Abrahão et al. (20) 7.2023/29.8.2023 https://arxiv.org/abs/2307.07522

* **On the Algorithmic Nature of the World,** Hector Zenil, Jean-Paul Delahaye, 2009/6/19/ 11.8.2010 Book, Information and Computation, 477-496 – testing the hypothesis of the Universe as a computer, data generated by simple algorithmic rules, rather than "truly complicated" or random ones or normal distribution etc. "1.*6. Conclusions: … the information in the world might be the result of processes resembling processes carried out by computing machines. …[the] general physical processes are dominated by algorithmic simple rules. … processes involved in the replication and transmission of the DNA have been found[Li (1999)] to be concatenation, union, reverse, complement, annealing and melting, all they very simple in nature. The same kind of simple rules may be the responsible of the rest of empirical data in spite of looking complicated or random* (…) *As opposed to simple rules one may think that nature might be performing processes represented by complicated mathematical functions, such as partial differential equations or all kind of sophisticated functions and possible algorithms."*

Todor: Compare with Theory of Universe and Mind, "The Universe Computer" (Вселената сметач), T.Arnaudov, **2001-2004** which precedes this work. Regarding representations with partial differential equations or whatever functions of whatever combined complexity – they are **not "complicated or sophisticated"** by default or per se *in their finer grained building blocks*, as long as they can also be decomposed into and solved by "simple" building blocks and instructions. This is done with electronic computers since their creation.
* A review of methods for estimating algorithmic complexity: Options, challenges, and new directions, Hector Zenil, 2020/5/30, Entropy https://www.mdpi.com/1099-4300/22/6/612
* Chapter 1 **On the Kolmogorov-Chaitin Complexity for short sequences**, 17.12.2010, Ch.1: https://arxiv.org/pdf/0704.1043v5

* H. Zenil. **Information Theory and Computational Thermodynamics: Lessons for Biology from Physics**. Information. *"Biology and Computational Universality: Information, in living beings, is maintained one-dimensionally through a double-stranded polymer called DNA"*1* … DNA – natural selection *2 … Charles Benett, RNA polymerase – *"truly chemical Turing machine";* true randomness is not required as quantum mechanics suggets; the complexity of the world can be explained by computation and informational worldview, the information may explain some quantum phenomena, and not the quantum mechanics the computation, the structures in the world and their algorithmic unfolding, so they *put computation at the lowest level underlying physical reality.* … ***Conclusion****: "the information in the world is the result of processes resembling computer programs rather than of dynamics characteristic of a more random, or analog, world"* *

Tosh, *2.7.2025***:** Compare the computationalism and informationalism with the Theory of Universe and Mind, "The Universe Computer" etc. However: 1) The one dimensionality of DNA is questionable, this is not a purely mathematical structure and DNA is meaningless without its environment, immediate cell, cells, tissue and organism to interpret it and to convert it to something else. The forces of nature together, the "physical laws", the way all kinds of particles interact, in particular the protein molecules, and the overall configuration with the environment contains information about the organism, not only the DNA sequence as a linear string of "symbols", this is not abstract representation in a mind which is thinking about the content of computer memory* **2)** Natural selection as in Darwinism is too simple way to explain the selection of DNA, hiding all actual processes and replacing them with a wrong tautology ("survival of the fittest", as cited, which in fact was intended to mean the one who **reproduces the most**, which is wrong in many ways: not only the "fittest", the most populous survive, but all who reproduce **enough;** furthermore "fittest" is often not clarified and is suggested as having particular "adaptations"; the latter can be implied (if you can't breath under water you won't survive in that "ecological niche").

What are the **reasons** for particular configurations being better suited, what actual lower level processes shape them, how exactly the transofmations happen etc. – that's more interesting to me. The process of "testing" whether an organism is "fit" (does it reproduce, or reproduce better than competition) is "brute forcing", let it "compute", "process", without understanding and explaining. Yet it also has particular details, forces, laws, dependencies that can be derived or observed and are more precise guide and predictor of whether an individual or species will survive.

* Note the contrast H.Zenil makes about computation: *"a **process** rather than a random **event**".* That suggests that event is regarded as momentary or discrete.

* C.H. Bennett. The Thermodynamics of Computation–A Review. International Journal of Theoretical Physics. vol. 21, no. 12, pp. 905–940, 1982
* R. Landauer, Irreversibility and Heat Generation in the Computing Process, https://worrydream.com/refs/Landauer_1961_-_Irreversibility_and_Heat_Generation_in_the_Computing_Process.pdf
…
* G.J. Chaitin**. Metaphysics, Metamathematics and Metabiology**. in H. Zenil (ed.), **Randomness Through Computation**. pp. 93–103, World Scientific, 2011.
Cited in H.Zenil, "Information Theory and Computational Thermodynamics…":
"DNA is essentially a programming language that computes the organism and its functioning; hence the relevance of the theory of computation for biology"

Tosh: Wasn't that supposed to be generally interpreted as such and adopted decades before 2011?, also it is not only the DNA, but the rules of interaction between the proteins and all molecules, see the note in the frame above.

* 7 G.J. Chaitin. **Life as evolving software**. in H. Zenil (ed.), A Computable Universe. World Scientific, forthcoming 2012.

*** Some Computational Aspects of Essential Properties of Evolution and Life**, Hector Zenil and James A.R. Marshall, Behavioural and Evolutionary Theory Lab, Department of Computer Science, The University of Sheffield, UK, 2012
– Compare with examples in TUM, 2001-2004. E.g. in TUM the need to "connect to the electrical grid" after certain limited amount of independent/autonomous work. Within another point of view in the explanation in the work of mine: The need to connect to a wider spatio-temporal range, because the amount of energy – resources for doing processing independently or "more independently" – that can be stored in some lower-scale spatio-temporal range or causality-control unit is limited; a mobile computer etc. – or a more mobile one, easier to be moved, lighter – sooner or later needs to connect to the power grid, and the power grid may be connected with the electric grid of the whole world, the Sun from solar power, thus to communicate directly with the minimal delay of 8:20 min with system of the scale of the Solar system etc. In another wording, the device, agent, causality-control unit has to pick another source of *causal energy* or causal power, bu connecting, becoming dependent, influenced by another broader causality control unit that is spanning a bigger range, span etc., either the electric power grid, solar power, get connected to another power bank/battery pack - another agent or a robot has to change the battery of the laptop etc.

* On Randomness, regularity …

* **What is Nature-like Computation? A Behavioural Approach and a Notion of Programmability**∗ Hector Zenil, 22.11.2012
*… Gregory Chaitin .. his strong belief that mathematicians should transcend the millenary theorem-proof paradigm in favor of a quasiempirical method based on current and unprecedented access to computational resources … The Kolmogorov-Chaitin complexity (or algorithmic complexity) of a string s is defined as the length of its shortest description p on a universal Turing machine U, formally $K(s) = \min\{l(p) : U(p) = s\}$. The major drawback of K, as measure, is its uncomputability. So in practical applications it must always be approximated by compression algorithms. A string is uncompressible if its shorter description is the original string itself**. If a string is uncompressible it is said that the string is random since no patterns were found*.** . Floridi's .. two types of information … instructional information..* Approaches to information: behavioral, syntactic, semantic… interpreter – *converting inf. Into a set of instructions .. Deutsch's: computers are physical objects, and computations are physical processes governed by the laws of physics. .. Wolfram's Principle of Computational Equivalence …* sensitivity to external stimuli … Cellular automata .. *Reversibility, 0-computers and conservation laws*

Todor: I don't agree with their/Kolmogorov-Chaitin's definition of randomness. Possibly and as it is explained, it is a terminology choice, but it doesn't feel the right term. The classical TUM claims that there's no true *objective* randomness, there is order in everything and the decision of the presence of randomness is decided by the evaluator-observer. These strings are rather ***basic***, ***elements*** than "random". In Zrim: букваче, дъно (bukvache, duhno).
As of complexity – see the notes in the frame about the expected average complexity of the molecules in a living organism and the sequences of bits in a cryptography table or a sufficiently trained neural network or already highly compressed file.

* Floridi, L. **Is Information Meaningful Data?** Philosophy and Phenomenological Research, 70(2): 351–370, 2005
* Floridi, L. The Method of Levels of Abstraction Minds and Machines, 18(3): 303–329, 2008.
* Wolfram, S. A New Kind of Science, Wolfram Media, 2002

* Levin, L.A. **Laws of information conservation (nongrowth) and aspects of the foundation of probability theory**. Probl. Pereda. Inf. 1974, 10, 30–35.
* Delahaye, J.-P. & Zenil, H. **Numerical evaluation of the complexity of short strings: A glance into the innermost structure of algorithmic randomness.** Applied Mathematics and Computation 219, 63–77 (2012). https://arxiv.org/abs/1101.4795 CTM …

*** Image information content characterization and classification by physical complexity**, Hector Zenil, Jean-Paul Delahaye, Cedric Gaucherel, 6.2010/7.2011 https://arxiv.org/abs/1006.0051 p.6: See the 100 sample blocks. Compare to the image tokens etc.[10]

* **The World is Either Algorithmic or Mostly Random**, Third Prize Winning Essay, 2011 FQXi Contest Is Reality Digital or Analog?, Hector Zenil, IHPST, Universit´e de Paris 1 – Panth´eon-Sorbonne .. *the notion that the universe is digital, not as a claim about what the universe is made of but rather about the way it unfolds*. … Digital universe, informational universe, randomness or order, algorithmic complexity, symmetry breaking, bit-string universe, prefix-free Turing machine, Busy Beaver game, halting problem
H.Zenil: *"..general physical processes are dominated by simple algorithmic rules, the same rules that digital computers are capable of carrying out. Our approach suggests that the information in the world is the result of processes resembling computer programs rather than of dynamics characteristic of a more random, or analog, world".*

[10] See also TUM, "Man and the Thinking Machine…", 2001 the basic elements of the hierarchical, also in the lecture slides about the TUM from the AGI course.

Todor: These definitions suggest that “algorithmic” is supposed to mean **“simple”** rules and “analog” means “random or complex”, which is not true. The **basic** building blocks always can be *simpler* than some bigger ones in a universe which is *constructive* and can build, or an evaluator-observer can recognize in it, hierarchical systems, segmentations, segments, borders etc. The definition seems to oppose “algorithmic” to “analog”, however the analog also can have “algorithmic” component (sequential etc.) and discreteness and the analog nature are a matter of degree, for example as precision of the Analog-to-Digital-Converter etc. and their estimation and classification are defined by and depend on the evaluator-observer and her resolution of causality-control and perception, specific technique, means or preferences of sampling etc. See other notes regarding that in the *Theory of Universe and Mind*, 2001-2004 and later, and notes in the main volume of *The Prophets of the Thinking Machines*, related to the work of Luciano Floridi. Bulgarian title of the section: **Вселената сметач, Информационната вселена, разделителна способност на възприятието и управлението** (The Universe computer, Informational Universe, Resolution of perception and causality-control), commenting:
* Against Digital Ontology, Luciano Floridi, 2009, 519 Views, 39 Pages, 2009, Synthese
https://www.academia.edu/327006/Against_Digital_Ontology
…
See Fig.1. *“From noise to highly organized structures.”* Regarding the evolution of the Universe. However, **were the initial or earlier states random**? In the interpretation of Theory of Universe and Mind they were not random. The current configuration is already implanted. Could random be defined in an initial state, to what it can be compared if there is no past? If that “random” state leads to order, then was it random? If any or many other “random” initial states would also lead to similar order, then are they also “random” or the process “random”, or it's rather **convergent** and some of the specifics of the configuration are equivallent and interconnected like the quantum entagled particles, and carry little own intrinsic complexity.

What has seemed as white noise wasn’t noise in its “deeper” meaning – just the resolution of perception and causation of the evaluator was or is not high enough to decode that representation with sufficient speed, precision etc., e.g. to unfold the code and simulate it faster than the Universe and to see that this “noise” eventually converts into the structure, which would be later classified as displaying order, organization etc. “Order”, regularity, patterns, matches, organization etc. are all according and within the cognitive limitations of the evaluator. “Regular” is what is “simple” enough for the observer (or similar enough to his filter/”allowed complexity”, allowed number of parts, segments, breadth etc.) is recognized as “organized”.

In p.3, section “1.1 Complexity from randomness” there is an example of code in C which computes the first 2400 digits of the number π / Pi -
H.Zenil.: “Physical laws, like computer programs, make things happen” Laws, calculations, prediction the orbits faster than real time. .. *“The existence of human-made digital computers in the universe is an obvious demonstration that the universe is capable of performing digital computation.”*

Todor: TUM/TOUM, 2001-2004: causality-control units, at higher levels – humans – are higher forms of physical laws and prediction as the main operation. Computers display the basic principles of operation of the Universe and the development towards their ubiquitiousness, universality is also a suggestion that they are connected to these core principles. (See e.g. Universe and Mind 2 (a.k.a. “The Universe Computer”, “Conception about the Universal Predetermination II”, “Letters between the 18-years old Todor Arnaudov and the philosopher Angel Grancharov” etc.)

Another recent related work in that line of research:

*** SuperARC: An Agnostic Test for Narrow, General, and Super Intelligence Based On the Principles of Causal Recursive Compression and Algorithmic Probability**, Alberto Hernández-Espinosa 1 , Luan Ozelim1,2, Felipe S. Abrahão1,2,3,4, and Hector Zenil, 79 p. ∗ 22.4.2025 https://arxiv.org/pdf/2503.16743 Recursive Compression-Decompression … Abstraction and Reasoning Corpus (ARC) – F.Chollet.

Test for LLMs; CTM; BDM … ***Design of experiments:*** … *1.* ***Low Complexity:*** *Sequences of digits or integers whose pattern is easily recognisable by a person and highly compressible; low CTM/BDM values. 2. Medium C.: Sequences of digits integers generated recursively with longer formulas than those in the simpler set; intermediate CTM/BDM. 3. High C.: Random-looking sequences of digits or integers; high CTM/BDM ...* ***1. Sort similarity:*** *This measures how many elements in the target sequence were predicted correctly, with their order being considered.* ***2. General similarity:*** *This measures the correctness of predicted elements, without considering their order.* ***3. Levenshtein****: This measures the Levenshtein distance between the expected and predicted sequences after converting them to strings*. (...) *for random sequences, which are considered highly complex, all models performed similarly, showing limited predictive* power. (..) *Recursive compression and optimal prediction go hand in hand .. Most of the [frontier] models demonstrate poor accuracy in replicating and predicting even simple and recursively generated sequences beyond clearly memorisation results from the training distribution (such as sequence labelling)*

Concl.: ***compression is proportional to prediction and vice versa .. if a system can better predict it can better compress, and if it can better compress, then it can better predict*.***

Tosh: This was predicted in TUM etc. since the early 2000s.

…

Randomness & Complexity, from Leibniz to Chaitin, Cristian S. Calude. 2007 – a book about Algorithmic Information Theory, Algorithmic Complexity, Computation, Turing Machine, Computability, Gregory Chaitin etc.,

https://www.academia.edu/1006593/WHERE_DO_NEW_IDEAS_COME_FROM_HOW_DO_THEY_EMERGE_EPISTEMOLOGY_AS_COMPUTATION_INFORMATION_PROCES

SING_
* Preface about G.Chaitin
*** On Random and Hard-to-Describe Numbers**, Charles H. Bennett, Ch.1., p.3 ..
Unprovability of Randomness, …Omega, Ω= the halting probability of a universal computer whose program is generated randomly, by tossing a fair coin whenever the computer requests another bit of input (Defined by G.Chaitin)
* Chaitin, G. J. "A Theory of Program Size Formally Identical to Information Theory." *J. Assoc. Comput. Mach.* **22**, 329-340, 1975.
* https://mathworld.wolfram.com/ChaitinsConstant.html

*** Where Do New Ideas Come From? How Do They Emerge? Epistemology as Computation (Information Processing)**, Gordana Dodig-Crnkovic, Chapter 15, M¨alardalen University,V¨aster˚ as, Sweden, p.263 – Computational universe, info-computationalism, pancomputationalism, … *"the question posed by Chaitin (2006) about the origin of creativity and novelty in a computational universe. …For even if the universe as a whole would be a totally deterministic mechanism, for humans to know its functioning and predict its behavior would take infinite time, as Chaitin already demonstrated that there are incompressible truths. In short****, in order to be able to predict the Universe-computer we would need the Universe-computer itself to compute its next state."***

Todor: Compare to TUM with exactly the same conclusion using the term "The Universe Computer" as well (Вселената сметач), regarding the highest resolution of causality-control; starting e.g. since 2002, **"Letters between the 18-years old Todor Arnaudov and the philosopher Angel Grancharov**", which has many titles, such as "The Universe Computer", "Conception about the Universal Predetermination Part II" etc.

…
*** What is a Computation?,** Martin Davis, Department of Computer Science, Courant Institute of Mathematical Sciences, New York University, USA: **Chapter 5,** p.89 – on Turing machine and Turing-Post programs (Emil Post): tape, 1-bit, STOP, Go To … halting problem and it's unsolvability
* Chapter 8 **Chaitin's Graph Coloring Algorithm,** James Goodman, Computer Science Department, University of Auckland, Auckland, New Zealand;
* G.J. Chaitin, M.A. Auslander, A.K. Chandra, J. Cocke, M.E. Hopkins, & P.W. Markstein. **Register allocation via coloring,** Computer Languages, Vol. , 6, pp. 47–57, 1981. https://pages.cs.wisc.edu/~fischer/cs701.f06/graph-coloring.pdf

*** Chapter 10, Ω in Number Theory**, Toby Ord and Tien D. Kieu, Faculty of Philosophy, Centre for Atom Optics and Ultrafast Spectroscopy, Swinburne University of Technology
[Omega is] [Ω is] *both recursively enumerable and algorithmically random.* … - it has a simple definition and can be computed "in the limit from below", yet only a finite number of digits can be determined with certainty, and the rest can't be predicted better than random. .. blank tape problem ..

* **Algorithmic Randomness and Complexity**, Rod Downey, 2011, slides, https://www.academia.edu/93559511/Algorithmic_Randomness_and_Complexity

* Robson, J. M. (1992). Deterministic simulation of a single-tape Turing machine by a random-access machine in sub-linear time. Information and Computation, 99(1), 109–121. https://doi.org/10.1016/0890-5401(92)90026-C

* **Derandomization from Time-Space Tradeoffs, Oliver Korten** # Columbia University, New York, NY, USA https://drops.dagstuhl.de/storage/00lipics/lipics-vol234-ccc2022/LIPIcs.CCC.2022.37/LIPIcs.CCC.2022.37.pdf - **pseudorandomness** and circuit complexity; incompressible strings, *i.e. finite objects which lack a specific type of structure or simplicity*; circuit minimization; RAM and 1-tape computation models; time-space tradeoffs; **non-constructive proof:** *an argument which demonstrates the existence of an object satisfying some special property, but which fails to indicate a particular example of such an object. Proofs which guarantee the existence of "pseudorandom objects." Key examples include the existence of truth tables of high circuit complexity and of pseudorandom generators capable of derandomizing algorithms. The task of making these proofs constructive is often referred to as an "explicit construction problem; the pigeonhole principle (Dirichlet's box)*

* **Hardness vs. Randomness,** Noam Nisan, Avi Wigderson, 1994 https://www.math.ias.edu/~avi/PUBLICATIONS/MYPAPERS/NOAM/HARDNESS/final.pdf *a pseudorandom generator is an "easy to compute" function which converts a "few" random bits to "many" pseudorandom bits that "look random" to any "small" circuit: "easy to compute" = polynomial time; "few" = n; ε (for some 0<ε<1); "many" = n; "look random" = subpolynomial difference in acceptance probability; and "small" = any polynomial.*
.. a long string that looks random to any algorithm from a complexity class C (eg P, NC, PSPACE,...) using an arbitrary function that is hard for C…

* **Metaphysics, Metamathematics and Metabiology,** February 2011, DOI: 10.1142/9789814327756_0006, Gregory Chaitin https://www.researchgate.net/publication/242674901_Metaphysics_Metamathematics_and_Metabiology