Preprints.org

Article

Not peer-reviewed version

Assembly Theory of Binary Messages (How to Assemble a Black Hole and Use it to Assemble New Binary Information?)

Szymon Łukaszyk * and Wawrzyniec Bieniawski

Posted Date: 14 March 2024

doi: 10.20944/preprints202401.1113.v5

Keywords: assembly theory; emergent dimensionality; holographic principle; black holes; complexity measures; P versus NP problem; Gödel's incompleteness theorems; halting problem; elegant program; Fibonacci sequence; Collatz conjecture; addition chains; quantum orthogonalization interval theorems; second law of infodynamics; mathematical physics; binputation

Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Assembly Theory of Binary Messages (How to Assemble a Black Hole and Use It to Assemble New Binary Information?)

Szymon Łukaszyk * and Wawrzyniec Bieniawski

Łukaszyk Patent Attorneys, ul. Głowackiego 8, 40-052 Katowice, Poland

* Correspondence: szymon@patent.pl

Abstract: Using assembly theory, we investigate the assembly pathways of fixed-length binary strings formed by joining the individual bits present in the assembly pool and the strings that entered the pool as a result of previous joining operations. We show that the string assembly index is bounded from below and above, and we conjecture about the form of the upper bound. We show that a string having the smallest assembly index can be assembled by an elegant trivial program of length equal to this index, if the length of the string is expressible as a product of Fibonacci numbers. We conjecture that there is no binary program that has a length shorter than the length of the string having the largest assembly index that could assemble this string. We conjecture that a black hole surface is defined by a balanced distinct string that satisfies the upper bound of a distinct string assembly index. The results confirm that four Planck areas provide a minimum information capacity that provides a minimum thermodynamic (Bekenstein-Hawking) entropy. Knowing that the problem of determining the assembly index is at least NP-complete, we conjecture that this problem is NP-complete, while the problem of creating the string so that it would have a predetermined maximum assembly index is NP-hard. Therefore, once the new information is assembled by a dissipative structure or by a human, increasing the information entropy according to the 2$^{\text{nd}}$ law of infodynamics, it is subject to the 2$^{\text{nd}}$ law of thermodynamics, and nature seeks to optimize its assembly pathway.

Keywords: assembly theory; emergent dimensionality; holographic principle; black holes; complexity measures; P versus NP problem; Gödel's incompleteness theorems; halting problem; elegant program; Fibonacci sequence; Collatz conjecture; addition chains; quantum orthogonalization interval theorems; second law of infodynamics; mathematical physics; binputation

---

1. Introduction

Assembly Theory (AT) [1–7] provides a distinctive complexity measure, superior to established complexity measures used in information theory, such as Shannon entropy or Kolmogorov complexity [1,5]. AT does not alter the fundamental laws of physics [6]. Instead, it redefines *objects* on which these laws operate. In AT, *objects* are not considered as sets of point *particles* (as in most physics), but instead are defined by the histories of their formation (assembly pathways) as an intrinsic property, where, in general, there are multiple assembly pathways to create a given *object*.

AT explains and quantifies selection and evolution, capturing the amount of memory necessary to produce a given *object* [6]. This is because the more complex a given *object* is, the less likely an identical copy can be observed without the selection of some information-driven mechanism that generated that *object*. Formalizing assembly pathways as sequences of joining operations, AT begins with basic units (such as chemical bonds) and concludes with a final *object*. This conceptual shift captures evidence of selection in *objects* [1,2,6].

The Assembly Index, which represents the length of the shortest assembly pathway leading to an *object*, facilitates the quantification of the minimum memory required for its construction. In general, it increases with the *object's* size, but decreases with symmetry, so large *objects* with repeating substructures may have lower complexity than smaller *objects* with greater heterogeneity [1]. The copy number specifies the number of copies of an *object*, essential for assessing its structural complexity.

© 2024 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

AT has been experimentally confirmed in the case of molecules and probed directly experimentally with high accuracy with spectroscopy techniques, including mass spectroscopy, IR, and NMR spectroscopy [6]. It is a versatile concept with applications in various domains. Beyond its application in the field of biology and chemistry [7], its adaptability to different data structures, such as text, graphs, groups, music notations, image files, compression algorithms, etc., showcases its potential in diverse fields [2].

In this study, we investigate the assembly pathways of binary strings by joining individual bits present in the assembly pool [6] and strings that entered the pool as a result of previous joining operations.

In particular, we investigate the assembly of black-body objects (BBs: black holes (BHs), white dwarfs, and neutron stars) considered binary strings [8–10]. It is known [2,8–19] that information in the universe evolves toward increased structural complexity, decreasing information entropy.

We use emphasis for *object* as this term, understood as a collection of *matter*, is a misnomer, as it neglects the (quantum) nonlocality [20]. Nonlocality is independent of the entanglement among *particles* [21], as well as the quantum contextuality [22], and increases as the number of *particles* [23] grows [24,25]. Furthermore, the ugly duckling theorem [26,27] asserts that every two *objects* we perceive are equally similar (or equally dissimilar).

This study extends the findings of previous research [8–10,23] within the framework of AT and emergent dimensionality [8–10,15,17,18,20,23,28]. However, our results generally apply to information theory. Therefore, we put the BB-related content in frames like this one. The reader not interested in BBs may skip the text in these frames and the additional results presented in Section 7.

The paper is structured as follows. Section 2 introduces the basic definitions used in the paper. Section 3 shows that the assembly index of binary strings is bounded from below and provides the form of this bound. Section 6 introduces the concept of binary assembling program and shows that the trivial assembling program assembles binary strings that have a minimum assembly index. Section 4 shows that the assembly index of binary strings is bounded from above and conjectures about the exact form of this bound. Section 7 discusses additional results of this research in the context of BBs. Section 8 concludes the findings of this study.

2. Preliminaries

For K subunits of an *object* O the assembly index a_O of AT is bounded [1] from below by

$$\min(a_O) = \log_2(K), \tag{1}$$

and from above by

$$\max(a_O) = K - 1, \tag{2}$$

The lower bound (1) represents the fact that the simplest way to increase the size of a subunit in a pathway is to take the largest subunit so far and join it to itself [1] and in the case of the upper bound (2), subunits must be distinct so that they could not be reused from the pool, decreasing the index.

Here we consider binary strings $C_k^{(N)}$ containing bits $\{0,1\}$, which are our basis AT *objects* [2], with N_0 zeros and N_1 ones[1], having a fixed length $N = N_0 + N_1$.

[1] N_1 is called binary Hamming weight or bit summation of a string $C_k^{(N)}$.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Where the bit value can be either 0 or 1, we write $* = \{0, 1\}$ with $*$ being the same within the string $C_k^{(N)}$. If we allow for the 2nd possibility that can be the same as or different from $*$, we write $\star = \{0, 1\}$. Thus, $C_k^{(2)} = [*\star]$, for example, is a placeholder for all four 2-bit strings.

We consider strings to be *messages* transmitted through a communication channel between a source and a receiver, similarly to the Claude Shannon approach [29] used in the derivation of binary information entropy

$$H\left(C_k^{(N)}\right) = -p_0 \log_2(p_0) - p_1 \log_2(p_1), \tag{3}$$

where

$$p_0 = \frac{N_0}{N} \quad \text{and} \quad p_1 = \frac{N_1}{N}, \tag{4}$$

are the ratios of occurrences of zeros and ones within the string $C_k^{(N)}$ and the unit of entropy (3) is bit.

Here, we consider the process of formation of binary strings within the AT framework.

Definition 1. *A string assembly index $a^{(N)}$ is the smallest number of steps s required to assemble a binary string $C_k^{(N)}$ of length N by joining two distinct bits contained in the initial assembly pool $P = \{0, 1\}$ and strings joined in previous steps that are added to the assembly pool. Therefore, the assembly index $a^{(N)}(C_k)$ is a function of the string $C_k^{(N)}$.*

For example, the 8-bit string

$$C_k^{(8)} = [01010101] \tag{5}$$

can be assembled in at most seven steps:

1. join 0 with 1 to form $C_k^{(2)} = [01]$, adding [01] to $P = \{0, 1, 01\}$,
2. join $C_k^{(2)} = [01]$ with 0 to form $C_k^{(3)} = [010]$, adding [010] to $P = \{0, 1, 01, 010\}$,
3. ...
7. join $C_k^{(7)} = [0101010]$ with 1 to form $C_k^{(8)} = [01010101]$

(i.e. not using the pool), six, five, or four steps:

1. join 0 with 1 to form $C_k^{(2)} = [01]$, adding [01] to P,
2. join $C_k^{(2)} = [01]$ with [01] taken from P to form $C_k^{(4)} = [0101]$, adding [0101] to P,
3. join $C_k^{(4)} = [0101]$ with [01] taken from P to form $C_k^{(6)} = [010101]$, adding [010101] to P,
4. join $C_k^{(6)} = [010101]$ with [01] taken from P to form $C_k^{(8)} = [01010101]$,

or at least three steps:

1. join 0 with 1 to form $C_k^{(2)} = [01]$, adding [01] to P,
2. join $C_k^{(2)} = [01]$ with [01] taken from P to form $C_k^{(4)} = [0101]$, adding [0101] to P,
3. join $C_k^{(4)} = [0101]$ with [0101] taken from P to form $C_k^{(8)} = [01010101]$,

while the 8-bit string

$$C_l^{(8)} = [00010111] \tag{6}$$

can be assembled in at least six steps:

1. join 0 with 1 to form $C_l^{(2)} = [01]$, adding [01] to P,
2. join $C_l^{(2)} = [01]$ with [01] taken from P to form $C_l^{(4)} = [0101]$, adding [0101] to P,
3. join 0 with 0 adding [00] to P,
4. join $C_l^{(4)} = [0101]$ with [00] taken from P to form $C_l^{(6)} = [000101]$, adding [000101] to P,
5. join $C_l^{(6)} = [000101]$ with 1 to form $C_l^{(7)} = [0001011]$, adding [0001011] to P,
6. join $C_l^{(7)} = [0001011]$ with 1 to form $C_l^{(8)} = [00010111]$,

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

as only the doublet [01] can be reused from the pool. Therefore, strings (5) and (6) have respective assembly indices $a^{(8)}(C_k) = 3$ and $a^{(8)}(C_l) = 6$ that represent the lengths of their shortest assembly pathways, which in turn ensures that the assembly pool P is a distinct set.

Tables 1 and A6–A13 (Appendix D) show the distributions of the assembly indices among 2^N strings for $4 \leq N \leq 12$ taking into account the number of ones N_1. The sums of each column form Pascal's triangle read by rows (OEIS sequence A007318).

Theorem 1. *A string having length $N = 4$ is the shortest string having more than one string assembly index 1.*

Proof. The proof is trivial. For $N = 1$ the assembly index $a^{(1)}(C) = 0$, as all basis *objects* have a pathway assembly index of 0 [2] (they are not *assembled*). $N = 2$ provides four available strings with $a^{(2)}(C) = 1$. $N = 3$ provides eight available strings with $a^{(3)}(C) = 2$. Only $N = 4$ provides 16 strings that include four stings with $a^{(4)}(C) = 2$ and twelve strings with $a^{(4)}(C) = 3$ including $|B^{(4)}| = 6$ balanced strings, as shown in Tables 1 and 2.

For example, to assemble the string $B_1 = [0101]$ we need to assemble the string [01] and reuse it. Therefore, $a^{(N)}(C_k) = N - 1$ for $0 < N < 4, \forall k$ and $\min_k\left(\{a^{(N)}(C_k)\}\right) < N - 1$ for $N \geq 4$, where $\{a^{(N)}(C_k)\}$ denotes a set of different assembly indices. □

Table 1. Distribution of the assembly indices for $N = 4$.

		N_1				
$a^{(4)}(C)$	$\lvert a^{(4)}(C)\rvert$	0	1	2	3	4
2	4	1		2		1
3	12		4	4	4	
	16	1	4	$\lvert B^{(4)}\rvert = 6$	4	1

Table 2. $|B^{(4)}| = 6$ balanced strings $B_k^{(4)}$.

k	$B_k^{(4)}$				$a^{(4)}(B_k)$
1	(0	1)	(0	1)	2
2	(1	0)	(1	0)	2
3	0	1	1	0	3
4	1	1	0	0	3
5	1	0	0	1	3
6	0	0	1	1	3

Interestingly, Theorem 1 strengthens the meaning of $N_{\mathrm{BH}} = 4$ as the minimum information capacity that provides a minimum thermodynamic (BH) entropy [30–32].

There is no *disorder* or *uncertainty* in an *object* that can be assembled in the same number of steps $s \leq 2$.

Definition 2. *A string $B_k^{(N)}$ is a balanced string if it has the same number of bits, where $N_1 = N_0 - 1$ or $N_0 = N_1 - 1$ if N is odd.*

Without loss of generality, we assume that if N is odd, $N_1 < N_0$ (e.g., for $N = 5$, $N_1 = 2$, and $N_0 = 3$). However, our results are equivalently applicable if we assume the opposite (i.e. a larger number of ones for an odd N).

Considered as messages [29], only balanced and even length strings $B_k^{(N)}$ have binary entropies $H(B_k^{(N)}) = \{0, 1\}$, i.e. natural numbers. Conversely, only non-balanced and/or odd length strings $C_k^{(N)}$ have binary entropies $0 < H(C_k^{(N)}) < 1$.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

The number $|B^{(N)}|$ of balanced strings among all 2^N strings is[2]

$$|B^{(N)}| = \binom{N}{\lfloor N/2 \rfloor} = \binom{N}{\lceil N/2 \rceil} \approx \sqrt{\frac{2}{\pi N}} 2^N. \quad (7)$$

This is OEIS A001405 sequence, the maximal number of subsets of an N-set such that no one contains another, as asserted by Sperner's theorem, and approximated using Stirling's approximation for large N.

BBs emit Hawking black-body radiation having a continuous spectrum that depends only on one factor, the BB temperature $|T_{\text{BB}}| = T_{\text{P}}/(2\pi d_{\text{BB}})$ corresponding to the BB diameter $D_{\text{BB}} := d_{\text{BB}}\ell_{\text{P}}, d_{\text{BB}} \in \mathbb{R}$, where ℓ_{P} and T_{P} is the Planck length and temperature [8].

Triangulated BB surfaces contain a balanced number of Planck area triangles, each having binary potential $\delta\varphi_k = -c^2 \cdot \{0, 1\}$, where c denotes speed of light in vacuum, as has been shown [8,10], based on the Bekenstein-Hawking entropy [30–32] $S_{\text{BB}} = k_{\text{B}} N_{\text{BB}}/4$. Here k_{B} is the Boltzmann constant and $N_{\text{BB}} := \pi D_{\text{BB}}^2/\ell_{\text{P}}^2 = \pi d_{\text{BB}}^2$ is the information capacity of the BB surface, i.e., the $\lfloor N_{\text{BB}} \rfloor \in \mathbb{N}$ Planck triangles corresponding to bits of information [8–10,31,33,34], and the fractional part triangle(s) having the area $\{N_{\text{BB}}\}\ell_{\text{P}}^2 = (N_{\text{BB}} - \lfloor N_{\text{BB}} \rfloor)\ell_{\text{P}}^2$ too small to carry a single bit of information [8,9].

Therefore, a balanced string B_k represents a BB surface comprising $N_1 = \lfloor N_{\text{BB}} \rfloor/2$ active Planck triangles (APTs) with binary potential equal to $-c^2$ [9].

Definition 3. *A string $D_k^{(N)}$ is a distinct string if a ring formed with this string by joining its beginning with its end is unique among the rings formed from the other distinct strings $D_l^{(N)}, l \neq k$.*

There are at least two and at most N forms of a distinct string $D_k^{(N)}$ that differ in the position of the starting bit. For example for $|B^{(4)}| = 6$ balanced strings, shown in Table 2, two augmented strings with $a^{(4)} = 2$ correspond to each other if we change the starting bit

$$\begin{aligned} &[\ldots 1 \mid 0101 \mid 0101 \mid 01 \ldots] = \\ &[\ldots 10 \mid 1010 \mid 1010 \mid 1 \ldots]. \end{aligned} \quad (8)$$

Similarly, four augmented strings with $a^{(4)} = 3$ correspond to each other

$$\begin{aligned} &[\ldots \mid 0110 \mid 0110 \mid 011 \ldots] = \\ &[\ldots 0 \mid 1100 \mid 1100 \mid 11 \ldots] = \\ &[\ldots 01 \mid 1001 \mid 1001 \mid 1 \ldots] = \\ &[\ldots 011 \mid 0011 \mid 0011 \mid \ldots], \end{aligned} \quad (9)$$

after a change in the position of the starting bit. Thus, there are only two distinct strings for $N = 4$

The number of distinct strings $|D^{(N)}|$ among all 2^N strings is given by the OEIS sequence A000031. In general (for $N \geq 3$), the number $|D^{(N)}|$ of distinct strings is much lower than the number $|B^{(N)}|$ of balanced strings.

[2] "$\lfloor x \rfloor$" is the floor function that yields the greatest integer less than or equal to x and "$\lceil x \rceil$" is the ceiling function that yields the least integer greater than or equal to x.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

As asserted by the no-hair theorem [35], BH is characterized only by three parameters: mass, electric charge, and angular momentum.

However, BHs are fundamentally uncharged and non-rotating, since the parameters of any conceivable BH, that is, charged (Reissner-Nordström), rotating (Kerr) and charged rotating (Kerr-Newman), can be arbitrarily altered, provided that the area of a BH surface does not decrease [36] using Penrose processes [37,38] to extract electrostatic and/or rotational energy of a BH [39].

Thus, a BH is defined by a single real number, and no Planck triangle is distinct on a BH surface. We can define neither a beginning nor an end of a balanced distinct string $E_k^{(N_{\mathrm{BH}})}$ that represents a given BH.

By neglecting the notion of the beginning and end of a string, we focus on its length and content. In Yoda's language,

"complete, no matter where it begins. A message is".

The numbers of the balanced $|B_k|$, distinct $|D_k|$, and balanced distinct[3] $|E_k|$ strings are shown in Table 3 and Figure 1. The formula for $|E_k|$ remains to be researched.

Table 3. String length N, number of all strings 2^N, number of balanced strings B_N, number of distinct strings D_N, and number of balanced distinct strings E_N.

N	2^N	$\|B_N\|$	$\|D_N\|$	$\|E_N\|$	$\|B_N\|/\|E_N\|$
1	2	1	2	1	1
2	4	2	3	1	2
3	8	3	4	1	3
4	16	6	6	2	3
5	32	10	8	2	5
6	64	20	14	4	5
7	128	35	20	5	7
8	256	70	36	10	7
9	512	126	60	14	9
10	1024	252	108	26	9.6923...
11	2048	462	188	42	11
12	4096	924	352	80	11.55
13	8192	1716	632	132	13
14	16384	3432	1182	246	13.9512...
15	32768	6435	2192	429	15

[3] $|E_k|$ is close to OEIS A000014 up to the eleventh term.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

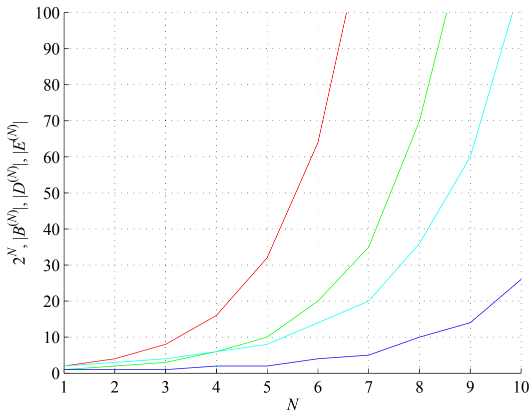

Figure 1. Numbers of all 2^N strings (red), balanced strings $|B^{(N)}|$ (green), distinct strings $|D^{(N)}|$ (cyan), and balanced distinct strings $|E^{(N)}|$ (blue) as a function of the string length N.

We note that, in general, the starting bit is relevant for the assembly index. Thus, different forms of a distinct string may have different assembly indices. For example, for $N = 7$ balanced strings B_{34} and B_{35}, shown in Table A16 have $a^{(7)} = 6$. However, these strings are not distinct, since they correspond to each other and to the balanced strings B_{13}, B_{18}, B_{20}, B_{28}, and B_{30} with $a^{(7)} = 5$. They all have the same triplet of adjoining ones.

Definition 4. *The assembly index of a distinct string $D_k^{(N)}$ is the smallest assembly index among all forms of this string.*

Thus, if different forms of a distinct string have different assembly indices, we assign the smallest assembly index to this string. In other words, we assume that the smallest number of steps

$$a^{(N)}(D_k) = \min_l \left(\{a^{(N)}(D_k)_l\} \right), \quad (D_k)_l \in D_k, \tag{10}$$

where $(D_k)_l$ denotes a particular form of a distinct string D_k, is the string assembly index of this distinct string.

> If an *object* that can be represented by a distinct string (a BB in particular) can be assembled in fewer steps, this procedure will be preferred by nature.

The distribution of the assembly indices of the balanced distinct strings E_k is shown in Table 4.

Table 4. Distribution of assembly indices among balanced distinct strings $E^{(N)}$ for $4 \leq N \leq 11$.

N	$\lvert E^{(N)} \rvert$	$a = 2$	$a = 3$	$a = 4$	$= 5$	$a = 6$	$a = 7$	$a^{(N)} = 8$
4	2	1	1					
5	2		1	1				
6	4		1	2	1			
7	5			2	3			
8	10		1	1	6	2		
9	14			1	4	7	2	
10	26			1	6	9	10	
11	42				2	14	20	6

3. Minimum Assembly Index

In the following, we derive the tight lower bound of the set of different string assembly indices 1.

Theorem 2 (Tight lower bound on the string assembly index)**.** *The smallest string assembly index $a^{(N)}(C_{min})$ as a function of N corresponds to the shortest addition chain for N (OEIS sequence A003313).*

Proof. Strings $C_{\min}$ for which $a^{(N)}(C_{\min}) = \min_k\left(\{a^{(N)}(C_k)\}\right)$, $\forall k = \{1, 2, \ldots, 2^N\}$ can be formed in subsequent steps s by joining the longest string assembled so far with itself until $N = 2^s$ is reached [1]. Therefore, if $N = 2^s$, then $\min_k\left(\{a_{(2^s)}(C_k)\}\right) = s = \log_2(N)$. Only four strings

$$\begin{aligned} C^{(2^s)}_{\min_1} &= [00\ldots], \\ C^{(2^s)}_{\min_2} &= [11\ldots], \\ C^{(2^s)}_{\min_3} &= [0101\ldots], \text{and} \\ C^{(2^s)}_{\min_4} &= [1010\ldots] \end{aligned} \tag{11}$$

have such an assembly index in this case.

An addition chain for $N \in \mathbb{N}$ having the minimum length $s \in \mathbb{N}$ (commonly denoted as $l(N)$) is defined as a sequence $1 = b_0 < b_1 < \cdots < b_s = N$ of integers such that for each $j \geq 1$, $b_j = b_k + b_l$ for $l \leq k < j$. The first step in creating an addition chain for N is always $b_1 = 1 + 1 = 2$ and this corresponds to assembling a doublet $[**]$ from the initial assembly pool P. Thus, the lower bound for s of the addition chain for N, $s \geq \log_2(N)$ is attained for $N = 2^s$. In the assembly case, this bound is attained by the strings (11). The second step in creating an addition chain can be either $b_2 = 1 + 1 = 2$ or $b_2 = 1 + 2 = 3$.

Thus, finding the shortest addition chain for N corresponds to finding an assembly index of a string containing bits and/or doublets and/or triplets generated by these doublets for $N \neq 2^s$ since due to Theorem 1 only they provide the same assembly indices $\{0, 1, 2\}$. Such strings correspond to linear molecules made of carbons [4] (Supplementary Materials, S3.2). □

Calculating both the minimum length of the addition chain for N and finding the assembly index of a string of length N have been shown to be at least as hard as NP-complete [4,40].

Table 5. The lower bound on the binary string assembly index (OEIS A003313).

N	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
$a^{(N)}_{\min}$	0	1	2	2	3	3	4	3	4	4	5	4	5	5	5	4	5	5	6

4. Maximum Assembly Index

In the following, we conjecture the form of the upper bound of the set of different string assembly indices 1.

In general, of all strings C_k having a given assembly index, shown in Tables 1 and A6–A13 (Appendix D), most are those having $N_1 = \lfloor N/2 \rfloor$. The only exceptions we found are $N = 8$ for $a^{(8)} = 4$ ($4 < 8$) and for $a^{(8)} = 6$ ($24 < 26$), $N = 10$ for $a^{(10)} = 4$ ($2 < 5$) and for $a^{(10)} = 5$ ($32 < 33$), and $N = 12$ for $a^{(12)} = 4$ ($2 < 3$).

Introducing the definition 2 of a balanced string allows us to reduce the search space of possible strings with maximal assembly indices to balanced strings only. With the exception of $N = 8$, of all strings $C^{(N)}_k$ having a maximum assembly index, most are balanced. We can further restrict the search space to distinct strings 3.

If a string $C_{\min}$ for which $a^{(N)}(C_{\min}) = \min_k\left(\{a^{(N)}(C_k)\}\right)$ is constructed from repeating patterns, then a string $C_{\max}$ for which $a^{(N)}(C_{\max}) = \max_k\left(\{a^{(N)}(C_k)\}\right)$ must be the most patternless. The string assembly index must be bounded from above and $a^{(N)}(C_{\max})$ must be a monotonically nondecreasing function of N that can increase at most by one between N and $N + 1$.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Identifying the shortest pathway is known to be computationally challenging [3]. This problem has been proven to be at least as hard as NP-complete [4]. However, certain heuristic rules apply in our binary case. For example,

- for $N = 7$ we cannot avoid two doublets (e.g. $2 \times [00]$) within a distinct string $E_{28}^{(7)} = [0011100]$ and thus $a^{(7)}(C_{\max}) = 5 < 6$,
- for $N = 8$ we cannot avoid two pairs of doublets (e.g. $2 \times [00]$ and $2 \times [11]$) within a distinct string $E_{7}^{(8)} = [00001111]$ and thus $a^{(8)}(C_{\max}) = 5 < 6$,
- for $N = 12$ we cannot avoid three pairs of doublets (e.g. $2 \times [00]$, $2 \times [10]$, and $2 \times [11]$) within a distinct string $E_{k}^{(12)} = [111000101100]$ and thus $a^{(12)}(C_{\max}) = 8 < 9$,
- for $N = 14$ we cannot avoid two pairs of doublets and one doublet three times (e.g. $2 \times [00]$, $2 \times [11]$, and $3 \times [01]$, and thus $a^{(14)}(C_{\max}) = 9 < 10$,
- etc.

Table 6 shows the exemplary balanced strings $B_{\max}$ having maximal assembly indices that we assembled (cf. also Apendix B). To determine the assembly index $a^{(18)} = 11$ of the string

$$E_{k}^{(18)} = [1(001)(11)(110)(110)(00)(001)0], \tag{12}$$

we look for the longest patterns that appear at least twice within the string, and we look for the largest number of these patterns. Here, we find that each of the two triplets [001] and [110] appear twice in $E_{k}^{(18)}$ and are based on the doublets [00] and [11] also appearing in $E_{k}^{(18)}$. Thus, we start with the assembly pool $\{0, 1, [00], [001], [11], [110]\}$ made in four steps and join the elements of the pool in the following seven steps to arrive at $a^{(18)}(E_k) = 11$. On the other hand, another form of this balanced distinct string

$$E_{l}^{(18)} = [(01)(11)(110)(110)00(001)(01)0], \tag{13}$$

has $a^{(18)}(E_l) = 12$.

Table 6. Exemplary balanced strings $B_{\max}^{(N)}$ having a maximum assembly index. Conjectured ($a_{\text{conj}}^{(N)}$) form of the maximum assembly index and its factual values for distinct ($a_{\text{dst}}^{(N)}$) and non-distinct ($a_{\text{ndst}}^{(N)}$) strings (red if below the conjectured value, green - if above).

N	$B_{\max}^{(N)}$																				$a_{\text{conj}}^{(N)}$	$a_{\text{dst}}^{(N)}$	$a_{\text{ndst}}^{(N)}$
1	0																				0	0	0
2	1	0																			1	1	1
3	0	0	1																		2	2	2
4	0	0	1	1																	3	3	3
5	0	0	0	1	1																4	4	4
6	0	0	0	1	1	1															5	5	5
7	0	0	1	1	1	0	0														5	5	6
8	0	0	0	1	0	1	1	1													6	6	6
9	0	0	0	0	1	1	1	0	1												7	7	7
10	0	0	0	0	1	1	1	1	0	1											7	7	8
11	0	0	0	0	0	1	0	1	1	1	1										8	8	8
12	1	1	1	0	0	0	1	0	1	1	0	0									9	8	8
13	0	0	0	0	0	0	1	0	1	1	1	1	1								9	9	9
14	0	0	0	0	0	1	0	1	0	1	1	1	1	1							9	9	9
15	0	0	0	0	0	1	0	1	0	1	1	1	1	1	0						10	10	10
16	1	0	0	0	0	0	0	1	0	1	0	1	1	1	1	1					11	10	10
17	0	0	0	0	0	0	1	0	1	0	1	1	1	1	1	1	0				11	11	11
18	1	0	0	1	1	1	1	1	0	1	1	0	0	0	0	0	1	0			11	11	12
19	1	0	0	0	0	1	0	1	0	1	0	0	1	1	1	1	1	0	1		12	11	12
20	1	0	1	0	0	1	1	1	1	1	0	1	1	0	0	0	0	0	1	0	13	12	13

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Conjecture 1 (Tight upper bound on a string assembly index)**.** *With exceptions for small N the largest string assembly index $a^{(N)}(C_{max})$ of a binary string as a function of N is given by a sequence formed by $\{+1, +1, k \times 0, +1, +1, k \times 0\}$ strings for $k \in \mathbb{N}_0$, where $+1$ denotes increasing $a^{(N)}(C_{max})$ by one, and* 0 *denotes maintaining it at the same level, and $a^{(0)} = -1$.*

However, at this moment, we cannot state whether this conjecture applies to distinct or non-distinct strings. The assembly indices for $N < 3$ are the same for a given N, whereas the assembly indices for $4 \leq N \leq 10$ were discussed above and are calculated in Appendix D for balanced and balanced distinct strings.

The conjectured sequence is shown in Figures 2 and 3 starting with $a^{(0)} = -1$ (we note in passing that $n = -1$ is a dimension of the void, the empty set $\varnothing$, or (-1)-simplex). Subsequent terms are given by $\{0, 1, 2, 3, 4, 5, 5, 6, 7, 7, 8, 9, 9, 9, 10, \dots\}$, which is periodic for $N = k(k+3)$ and flattens at $a^{(N)}(D_{\max}) = 4k - 3$, and $a^{(N)}(D_{\max}) = 4k - 1$, $k \in \mathbb{N}$, $k > 1$.

This sequence can be generated using the following procedure

```
step=1;                     % step flag
run =1;                     % run flag
flat=0;                     % flat counter

Nk = 0;
aub= -1;                    % the upper bound
while Nk < N
    if step < 3
        Nk = Nk+1;          % next Nk
        aub= aub + 1;       % increment the bound
    else                    % step==3
        for k=1:flat
            if flat > 0
                Nk = Nk+1;  % next Nk
            end
        end
        run = run+1;        % increment run flag
        if run > 2
            run  = 1;       % reset run flag
            flat = flat+1;  % increment flat counter
        end
    end
    step = step+1;          % increment step flag
    if step > 3
        step=1;             % reset step flag
    end
end
```

We note the similarity of this bound to the Aufbau rule[4], the Janet sequence (OEIS A167268) and the monotonically non-decreasing Shannon entropy of chemical elements, including observable ones [23]. Perhaps the exceptions in the sequence 1 vanish as N increases.

5. Strings between the bounds

The bounds 2 and 1 are shown in Tables 5 and 6 and illustrated in Figures 2 and 3. No binary string can be assembled in a smaller number of steps than given by a lower bound of Theorem 2. On the other hand, some strings cannot be assembled in a smaller number of steps than given by an upper bound (which for large N, as we suppose, has the form presented in Conjecture 1).

The Hamlet tragedy contains approximately 130,000 letters. Assigning five bits per letter (32 possibilities), the Hamlet tragedy can be encoded in a string having $N_{\text{Hamlet}} = 650000$ bits (81.25 kB)

---

[4] Only about twenty chemical elements (with only two nondoubleton sets of consecutive ones) violate the Aufbau rule.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

yielding the total number of possible strings $2^{N_{\text{Hamlet}}} \approx 1 \times 10^{195312}$ (including $|B_{N_{\text{Hamlet}}}| \approx 1 \times 10^{195309}$), and their assembly indices are bounded by

$$27 \leq a^{(N_{\text{Hamlet}})}(C_k) \ll 3217 \tag{14}$$

The lower bound (14) can be estimated using Theorem 2 (though for large N it can be at least as hard as NP-complete [4,40]). The upper bound (14) can be estimated by finding the smallest k that satisfies $k(k+3) \geq N_{\text{Hamlet}}$ and using the relation $a^{(N)}(C_{\max}) = 4\lceil k \rceil - 1$ of Conjecture 1.

We assume that the assembly index of the string encoding the actual Hamlet tragedy is close to the upper bound. Even if the probability of random typing of the Hamlet tragedy is unfathomably small, when constrained to the bounds of the physical universe [5], as asserted by the infinite monkey theorem, this tragedy was once created by William Shakespeare.

SARS-CoV-2 genome sequence contains 29903 nucleobases $\{A, C, G, T\}$. Assigning two bits per base it can be encoded in a string of $N_{\text{SARS-CoV-2}} = 59806$ bits having the assembly index bounded by

$$20 \leq a^{(N_{\text{SARS-CoV-2}})}(C_k) \ll 971. \tag{15}$$

The supermassive BH Sagittarius A* has an estimated mass $M_{\text{BH}} \approx 8.26 \times 10^{36}$ kg corresponding to the Schwarzschild diameter $D_{\text{BH}} \approx 2.45 \times 10^{10}$ m and the information capacity $N_{\text{Sagittarius A}^*} \approx 7.24 \times 10^{90}$ [8]. In this case, its assembly index is bounded by

$$332 \leq a^{(N_{\text{Sagittarius A}^*})} \ll 1.0763 \times 10^{46}. \tag{16}$$

However, we conjecture that

Conjecture 2. *A BB surface is defined by a balanced distinct string that satisfies the upper bound of a distinct string assembly index.*

To be the most patternless [8], a balanced BB surface must minimize not only Shannon entropy and Kolmogorov complexity (the latter is uncomputable), but also maximize its assembly index. A BB cannot be assembled in a suboptimal way, since black-body radiation is informationless.

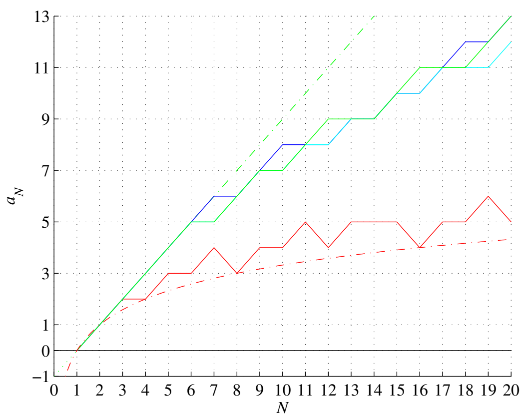

Figure 2. Lower bound on the binary string assembly index 2 (red) and $\log_2(N)$ (red, dash-dot), conjectured upper bound on the binary string assembly index 1 (green), factual values of the string assembly index (blue) and the distinct string assembly index (cyan) and $N-1$ (green, dash-dot), for the string length $0 \leq N \leq 20$.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

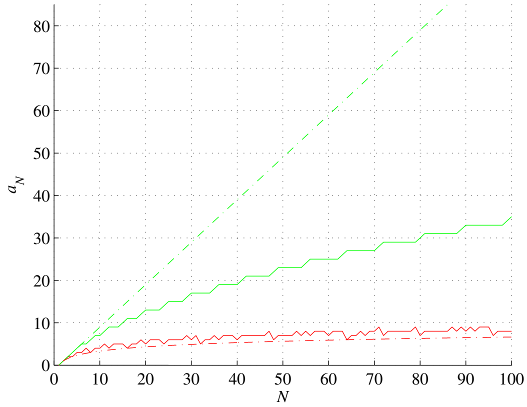

Figure 3. Lower bound on the binary string assembly index (red) and $\log_2(N)$ (red, dash-dot), conjectured upper bound on the string assembly index (green) and $N-1$ (green, dash-dot), for the binary string length $0 \le N \le 100$.

6. Binputation

Definition 5. *The binary assembling program Q_B is a binary string of length s_Q that acts on the assembly pool P and outputs the assembled strings, adding them to the pool.*

Definition 6. *The trivial assembling program Q is a binary assembling program 5 with particular bits denoting*

0 take the last two elements from P, join them with each other, and output; and
1 take the last element from P, join it with itself, and output.

As the assembly pool P is a distinct set to which strings are added in subsequent assembly steps, only these two commands are applicable to the initial assembly pool $P = \{0, 1\}$ containing only two bits.

Theorem 3. *A string $C_{min}^{(N)}$ can be assembled by an elegant trivial program of length $s_Q = a^{(N)}(C_{min})$ iff N is expressible as a product of Fibonacci numbers (OEIS A065108).*

Proof. An elegant program is the shortest program that produces a given output [41,42].

The 1st bit of the program Q is irrelevant as $Q = 0$ assembles $C_{\mathrm{min}_3}^{(2)} = [01]$ and $Q = 1$ assembles $C_{\mathrm{min}_1}^{(2)} = [11]$ and we can denote it by $*$. Then the programs $Q = *1\ldots1$ assemble the 2^{s_Q}-bit strings $C_{\mathrm{min}_{1,3}}^{(2^s)} = [*1*1\ldots]$ having the assembly index $a_{\mathrm{min}}^{(2^{s_Q})} = s_Q$, while strings $C_{\mathrm{min}_{2,4}}^{2^{s_Q}}$ with the smallest assembly index $a_{\mathrm{min}}^{(2^{s_Q})} = s_Q$ can be assembled with the same two programs starting with the pool $P = \{1, 0\}$.

The remaining $2^{s_Q} - 2$ programs will assemble some of the shorter strings with the assembly index $a_{\mathrm{min}}^{(N)} = s_Q$. In general, all programs Q assemble strings having lengths expressible as a product of Fibonacci numbers (OEIS A065108) as shown in Table A2 (Appendix C), wherein out of 2^{s_Q} programs (cf. Tables A5 and A2):

- 2^{s_Q-1} programs $Q = *1*\ldots$ assemble even length strings $C = [*1*1\ldots]$ having natural binary entropies (3) $H(C) = \{0, 1\}$, including strings $C_{\mathrm{min}_{1,3}}^{(2^{s_Q})}$ (11),
- 2^{s_Q-2} programs $Q = *01*\ldots$ assemble $[1*11*1\ldots]$ strings having lengths divisible by three and entropies $H(C) \approx \{0, 0.9183\}$,
- 2^{s_Q-3} programs $Q = *001*\ldots$ assemble $[*11*1*11*1\ldots]$ strings having lengths divisible by five and entropies $H(C) \approx \{0, 0.9710\}$,
- $\ldots$,
- the program $Q = *0\ldots1$ joins two shortest strings joined in a previous step into a string of length being twice the Fibonacci sequence (OEIS A055389).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

- the program $Q = *0\dots0$ assembles the shortest string that has length belonging to the set of Fibonacci numbers.

Thus, for $* = 0$, binary assembling programs Q assemble subsequent $2^{s_Q-1} = 2^{s_Q-2} + 2^{s_Q-3} + \dots + 2^0 + 1$ Fibonacci words (the fixed point of the morphism $0 \to 01$, $1 \to 011$, OEIS sequence A096270, another version of the Fibonacci word, OEIS A005614) having entropies (3) with ratios (4)

$$p_{0,m} = \frac{F_m}{F_{m+2}} \quad \text{and} \quad p_{1,m} = \frac{F_{m+1}}{F_{m+2}}, \tag{17}$$

where $m = \{1, 2, \dots s_Q\}$, and F is the Fibonacci sequence starting from 1 that rapidly converges to

$$\lim_{s_Q \to \infty} H_m \approx 0.9594187, \tag{18}$$

the binary entropy of the Fibonacci word limit.

For $s_Q \geq 4$, some of the programs are no longer elegant. Furthermore, for $s_Q \geq 4$, $Q = 000011\dots$ assembles a string

$$C^{(2^{s_Q-1})}_{\text{non-min}} = [10101101\dots] \tag{19}$$

with an assembly index $a^{(2^{s_Q-1})} = s_Q$ which is not the minimum for this length of the string. For example, 4-bit program $Q = *000$ assembles the string $C^{(8)} = [1 * 1 * 11 * 1]$, but if $* = 1$ this string can be assembled by a shorter, 3-bit program $Q = *11$, and if $* = 0$ this string does not have the minimal assembly index $a^{(8)}(C_{\text{min}}) = 3$ but $a^{(8)}(C_{\text{non-min}}) = 4$.

For $s_Q = \{4, 7\}$ and $s_Q \geq 10$ and for the shortest string assembled by the program Q the program Q is not elegant for $* = 1$ and the shortest string assembled by the program is not C_{min} for $* = 0$. However, the length s_Q of any program Q is not shorter than the assembly index of the string that this program assembles. □

The trivial assembly programs Q and the strings they assemble are listed in Tables 7 and A3–A5 (Appendix C) for one version of the assembly pool and for $1 \leq s_Q \leq 6$. If a binary string $C^{(N)}$ were to code four DNA nucleobases, (for example as, A= 00, G= 01, C= 10, and T= 11) then we note that the nucleobase encoded by 00 (or 11 for $P = \{1, 0\}$) would not be present in the strings generated by trivial assembly programs Q defined by 6. We note in passing that the Fibonacci sequence can be expressed through the golden ratio, which corresponds to the smallest Pythagorean triple $\{-3, 4, 5\}$ [28,43].

Table 7. 3-bit elegant programs assembling strings with $a^{(N)} = 3$.

Q	$C(s=1)$	$C(s=2)$	$C(s_Q=3)$	N
$*00$	$*1$	$1*1$	$*11*1$	5
$*01$	$*1$	$1*1$	$1*11*1$	6
$*10$	$*1$	$*1*1$	$*1*1*1$	6
$*11$	$*1$	$*1*1$	$*1*1*1*1$	8

Perhaps the minimum assembly index 2 and Theorem 3 are related to the Collatz conjecture, as the lengths of the strings (11) $N = 2^{2k}$ correspond to the numbers to which the Collatz conjecture converges, from $N = (4^k - 1)/3$, $k \in \mathbb{N}$ (OEIS A002450).

Theorem 3 is related to Gödel's incompleteness theorems and the halting problem. N cases of the halting problem correspond only to $s_Q = \log_2(N)$, not to N bits of information [44]. Therefore, s_Q-bit elegant programs assemble all four strings $C^{(2^{s_Q})}_{\text{min}}$ with $a^{(2^{s_Q})} = s_Q$ (with two versions of the assembly pool). Furthermore, we could consider all strings assembled by the s_Q-bit assembly program as corresponding to provable theorems. Any formal axiomatic system (such as our trivial program; Defnition 6) only enables proving provable theorems. Thus, proving a theorem equals halting. There is a more fundamental path to incompleteness that involves complexity, rather than self-reference [45].

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Theorem 3 would be violated if we defined the command "1" e.g. as "take the last element from the assembly pool, join it with itself, join with what you have already assembled (say at "the right"), and output". Then the 2-bit program "11" would produce [111111] with the assembly index $a^{(6)} = 3$. However, such a one-step command would violate the axioms of assembly theory since it would perform two assembly steps in one program step. An elegant program to output the gigabyte binary string of all zeros would take a few bits of code and would have a low Kolmogorov complexity [46]. However, such a string would be *outputted*, not *assembled*. Furthermore, the length of such a program that outputs the string $[1 \dots 1]$ would be shorter than the length of the program that outputs the string $[01 \dots]$, while in AT, the lengths of these programs must be the same if the strings have the same assembly indices. Theorem 3 is about *binputation*[5] of binary strings.

Conjecture 3. *There is no binary assembling program 5 that has a length shorter than the length of the string having the largest assembly index that could assemble this string.*

Partial Proof. When assembling the string having the largest assembly index (the largest complexity), we cannot rely solely on the last or two last strings in the assembly pool. Thus, we need to index the strings in the pool. However, we cannot predict in advance how many strings there will be in the assembly pool. Thus, we do not know how many bits will be needed to encode the indices.

Furthermore, no program P (for a Turing machine) shorter than an elegant program Q can find Q [45]. If it could, it could also generate Q's output. But if P is shorter than Q, then Q would not be elegant. Contradiction.
□

Conjecture 4. *The problem of determining the assembly index of a given binary string $C_k^{(N)}$ and the problem of assembling a non-trivial string so that it would have a minimum assembly index (Theorem 2) for N are NP-complete. The problem of assembling the string so that it would have a maximum assembly index for N is NP-hard.*

Partial Proof. We found it much easier to determine an assembly index of a given binary string $C_k^{(N)}$ than to assemble a string so that it would have a maximum assembly index. Similarly, a trivial string with a minimum assembly index for N can have the form $C_{\text{min}}^{(N)} = [* \dots]$ or the form of a Fibonacci word generated by the trivial program 6. □

A proof of conjecture 4 would also be the proof of the following conjecture.

Conjecture 5. $P \neq NP$

Partial Proof. Every computable problem and every computable solution can be encoded as a finite binary string. Here, determining whether the assembly index of a given string has its known maximal value corresponds to checking the solution to a problem for correctness, whereas assembling such a string corresponds to solving the problem. □

Thus, AT would solve the P versus NP problem in theoretical computer science. There is ample pragmatic justification for adding P $\neq$ NP as a new axiom [44].

7. Additional Results

The [perceivable] universe is not big enough to contain the future; it is deterministic going back in time and non-deterministic going forward in time [47]. But we know [2,8–19] that it has evolved to the present since the Big Bang.

---

[5] As an analog to chemputation, where assembly theory is applied in digital chemistry.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Perceivable information about any *object* can be encoded by a binary string [26,27]. This does not imply that a binary string defines an *object*. Information that defines a chemical compound, a virus, a computer program, etc. can be encoded by a binary string. However, a dissipative structure [12] such as a living biological cell (or its conglomerate such as a human, for example) cannot be represented by a binary string (even if its genome can). This information can only be perceived (so this is not an *object defining* information). Each of us is given to ourselves as a mystery [48]. Therefore, since one bit is the smallest amount and the quantum of information, the lower bound and the upper bound of the string assembly index define the allowed region of the assembly indices for binary strings.

The bounds 2, 1, (1), and (2) on the assembly index are shown also in Figure 4 (adopted from [1] and modified). According to the authors of [1], the "green portion of the figure is illustrative of the location in the complexity space where life might reasonably be found. Regions below can be thought of as being potentially naturally occurring, and regions above being so complex that even living systems might have been unlikely to create them. This is because they represent structures with limited internal structure and symmetries, which would require vast amounts of effort to faithfully reproduce." [1].

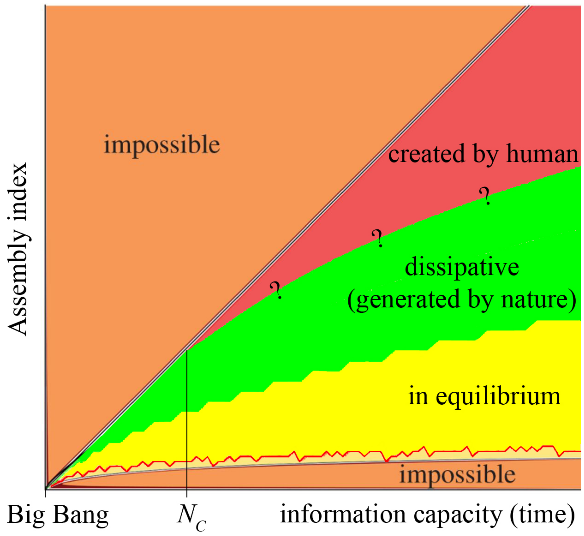

Figure 4. An illustrative graph of complexity against information capacity: orange regions are impossible, as they are above or below the assembly bounds; yellow region contains structures optimally assembled (in equilibrium); green region contains dissipative structures; and red region is the region of human creativity (figure not to scale).

We disagree with this statement. It is obvious that a binary string itself is neither dissipative nor creative. It is its assembly process that can be dissipative or creative. Evolution is about assembling new information and optimizing it until it reaches its assembly index.

That is why, we found determining the assembly index of a given binary string $C_k^{(N)}$ is easier than creating a string with a maximum assembly index for this length of the string (Conjecture 4). Once the new information is assembled (by a dissipative structure operating far from thermodynamic equilibrium, or created by humans) increasing the information entropy according to the 2$^{\text{nd}}$ law of infodynamics [16], it enters the realm of the 2$^{\text{nd}}$ law of thermodynamics, and nature seeks how to optimize its assembly pathway decreasing information entropy. And only humans are gifted with creativity. Any creation is required to be shaped by the unique personality of the creator to such an extent that it is statistically one-time in nature [49]; it is an imprint of the author’s personality.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

The total entropy of the universe S is constant and is the sum of the information entropy S_{info} and the physical entropy S_{phys}. Therefore, over time [19]

$$\frac{dS_{info}}{dt} + \frac{dS_{phys}}{dt} = 0. \tag{20}$$

The time corresponds to an increasing information capacity. Bit by bit: $dt = (N+1) - N = 1$

At first, the newly assembled information corresponds to the discovery by groping [11]. However, its assembly pathway does not attain its most economical or efficient form all at once. For a certain period of time, its evolution gropes about within itself. The try-out follows the try-out, not being finally adopted. Then finally perfection comes within sight, and from that moment the rhythm of change slows down [11]. The new information, having reached the limit of its potentialities, enters the phase of conquest. Stronger now than its less perfected neighbours, the new information multiplies and consolidates. When the assembly index is reached, new information attains its equilibrium (not necessarily a BH equilibrium) and its evolution terminates. It becomes stable.

There is a certain minimum amount of information N_C required to establish a creation, as shown in Figure 4. Sixteen possibilities provided by the minimum of thermodynamic entropy [30–32] bifurcate the assembly pathways (cf. Theorem 1) but none of these possibilities can be considered a *creation*. However, the boundary between the green region of dissipative structures [12] and the red region of human creativity remains to be discovered.

> "Thanks to its characteristic additive power, living matter (unlike the matter of the physicists) finds itself 'ballasted' with complications and instability. It falls, or rather rises, towards forms that are more and more improbable. Without orthogenesis life would only have spread; with it there is an ascent of life that is invincible." [11]

BB having the energy given by mass-energy equivalence

$$\begin{aligned} E_{\text{BB}} &= \frac{k}{2} M_{\text{BB}} c^2 = \frac{k}{2} m_{\text{BB}} E_{\text{P}} = \frac{d_{\text{BB}}}{4} E_{\text{P}} = \frac{1}{4}\sqrt{\frac{N_{\text{BB}}}{\pi}} E_{\text{P}}, \\ 2 \leq k &\leq k_{\max} = \frac{2\alpha^2}{\sqrt{\alpha^4 - \alpha_2^4}} \approx 6.7933 \end{aligned} \tag{21}$$

where $M_{\text{BB}} := m_{\text{BB}} m_{\text{P}}$, $m_{\text{BB}} \in \mathbb{R}$ denote the BB mass, and E_{P}, m_{P} denote the Planck energy and mass, $\alpha \approx 1/137.036$ is the fine-structure constant and $\alpha_2 \approx -1/140.178$ is the 2$^{\text{nd}}$ fine-structure constant related to α by $(\alpha + \alpha_2)/(\alpha\alpha_2) = -\pi$, and k is the BB size-to-mass ratio (STM) [10] ($k = 2$ if BB is BH).

It was shown [9] based on the Mandelstam-Tamm [50], Margolus–Levitin [51], and Levitin-Toffoli [52] theorems on the quantum orthogonalization interval that BBs generate (or rather *assemble*) a pattern forming nonequilibrium shell (VS) through the solid-angle correspondence, as shown in Figure 5. The BB entropic work

$$\begin{aligned} W_{\text{BB}} &= T_{\text{BB}} S_{\text{BB}} = T_{\text{BB}} \frac{1}{4} k_{\text{B}} N_{\text{BB}} = T_{\text{BB}} \frac{1}{4} k_{\text{B}} \pi d_{\text{BB}}^2 \\ &= \frac{E_{\text{P}} d_{\text{BB}}}{4k} \left(1 \pm i\sqrt{\frac{k^2}{4} - 1}\right), \end{aligned} \tag{22}$$

is the work done by all APTs of a BB. It is the product of the BB entropy [30–32] and the general, complex BB temperature

$$T_{\text{BB}} = \frac{T_{\text{P}}}{k\pi d_{\text{BB}}} \left(1 \pm i\sqrt{\frac{k^2}{4} - 1}\right), \tag{23}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

which in modulus and for a BH ($k = 2$) reduces [10] to Hawking temperature

$$T_{\mathrm{BH}} = \frac{\hbar c^3}{8\pi G M_{\mathrm{BH}} k_{\mathrm{B}}} = \frac{T_{\mathrm{P}}}{2\pi d_{\mathrm{BH}}}, \tag{24}$$

where $\hbar = h/(2\pi)$ is the reduced Planck constant, G is the gravitational constant, and T_{P} is the Planck temperature. In particular [10]

$$T_{\mathrm{BB}}(k_{\max}) = \frac{T_{\mathrm{P}}}{2\pi d_{\mathrm{BB}} \alpha^2} \left(\sqrt{\alpha^4 - \alpha_2^4} \pm i\alpha_2^2 \right), \tag{25}$$

$$T_{\mathrm{BB}}(k_{\mathrm{eq}}) = \frac{T_{\mathrm{P}}}{2\pi d_{\mathrm{BB}}} \frac{\alpha^2 \pm i\alpha_2^2}{\sqrt{\alpha^4 + \alpha_2^4}}, \tag{26}$$

where

$$k_{\mathrm{eq}} = \frac{2}{\alpha^2} \sqrt{\alpha^4 + \alpha_2^4} \approx 2.7665. \tag{27}$$

is the energy equilibrium STM.

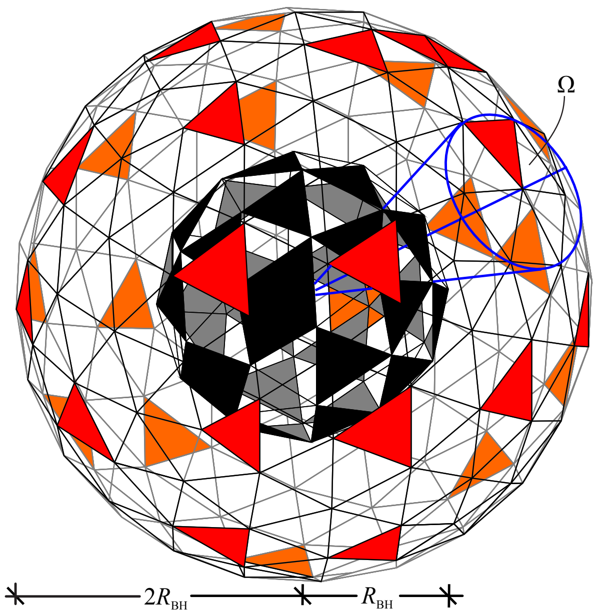

Figure 5. A black body object as a generator of an entropy variation shell (VS) through the solid angle Ω correspondence.

A VS has the information capacity bounded by

$$\begin{aligned} &N_{\mathrm{BB}} \leq N_{\mathrm{VS}} \leq 4N_{\mathrm{BB}}, \\ &N_{\mathrm{VS}} := lN_{\mathrm{BB}}, \quad 1 \leq l \leq 4, \end{aligned} \tag{28}$$

where l is a VS defining factor. The number of APTs is bounded by

$$\left\lfloor \frac{1}{4} N_{\mathrm{BB}} \right\rfloor \leq N_1 \leq \left\lfloor \frac{1}{2} N_{\mathrm{BB}} \right\rfloor, \tag{29}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

as shown in Figure 6, and thus its binary potential $\delta\varphi_{\mathrm{VS}} = -N_1 c^2/N_{VS}$ [8,9] is bounded by

$$-\frac{1}{2}c^2 \leq \delta\varphi_{\mathrm{VS}} < \begin{cases} \left(\frac{1}{N_{\mathrm{BB}}} - \frac{1}{4}\right)c^2 > 0, & N_{\mathrm{BB}} < 4 \\ \left(\frac{1}{4N_{\mathrm{BB}}} - \frac{1}{16}\right)c^2 < 0, & N_{\mathrm{BB}} > 4 \end{cases}. \tag{30}$$

and the theoretical probability $p_1 := N_1/N_{\mathrm{VS}}$ for a triangle on a VS to be an active Planck triangle is also bounded [9] by

$$\frac{1}{16} - \frac{1}{4N_{\mathrm{BB}}} < p_1 \leq \frac{1}{2}. \tag{31}$$

On the other hand, the entropy variation [8,53] $\delta S/k_{\mathrm{B}} = -c\delta\varphi$ so that for $N_{\mathrm{BB}} < 4$ the lower bound (31) is negative and the upper bound (30) is positive ($N_1 \leq 1$ in this range). The Planck triangle of VS is located *somewhere* on the VS surface defined by a solid angle

$$\Omega = \frac{\ell_{\mathrm{P}}^2}{R_{\mathrm{BB}}^2} = \frac{4\pi\ell_{\mathrm{P}}^2}{4\pi R_{\mathrm{BB}}^2} = \frac{4\pi}{N_{\mathrm{BB}}}, \tag{32}$$

that corresponds to the BB Planck triangle.

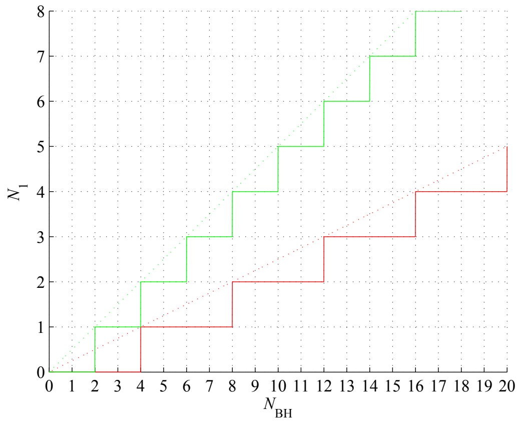

Figure 6. Lower (red) and upper (green) bound on the number of APTs N_1 on a VS as a function of the information capacity of the generating BB [9].

The BB information capacity is dictated by its diameter and the BB energy (21) as a function of its diameter is the same for all BBs (it is independent on k). However, the BB mass and density

$$\rho_{\mathrm{BB}} = \frac{M_{\mathrm{BB}}}{V_{\mathrm{BB}}} = \frac{3}{kN_{\mathrm{BB}}}\rho_{\mathrm{P}}, \tag{33}$$

are not.

Based on the orbiting condition $V_O^2 \leq V_R^2 \leq V_E^2$, where $V_O = \sqrt{GM_{\mathrm{C}}/R_{\mathrm{avg}}}$ is the orbital, and $V_E = \sqrt{2GM_{\mathrm{C}}/R_{\mathrm{avg}}}$ the escape speed of an orbiting *object*, R_{avg} is the average distance from the center of the central *object* to the center of the orbiting *object*, and M_{C} is the mass of the central *object*, the bounds

$$N_{\mathrm{BB}} \leq 4v_R^4 N_{\mathrm{VS}} \leq 4N_{\mathrm{BB}}, \tag{34}$$

containing the velocity term $V_R = v_R c$, $v_R \in \{\mathbb{R}, \mathbb{I}\}$ were also derived [9]. Plugging N_{VS} from the bounds (28) into the bounds (34) we arrive at

$$\frac{1}{4l} \leq v_R^4 \leq \frac{1}{l}, \tag{35}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

which is satisfied by real and imaginary (but not complex) velocities (for example, for $l = 1$ by $-1 \leq v_R \leq -1/\sqrt{2}$, $1/\sqrt{2} \leq v_R \leq 1$, $-i \leq v_R \leq -i/\sqrt{2}$, and $i/\sqrt{2} \leq v_R \leq i$). Taking the square root of the bounds (35), using $v_{LL}^2 + v_{RR}^2 = 1$, $v_R \in \{\mathbb{R}, \mathbb{I}\}$ [9], and squaring again, we arrive at

$$\frac{l - 2\sqrt{l} + 1}{l} \leq v_L^4 \leq \frac{4l - 4\sqrt{l} + 1}{4l}. \tag{36}$$

The bounds (35) and (36), shown in Figure 7, meet at $v = 1/\sqrt{2}$, where de Broglie and Compton wavelengths of mass M are the same

$$\lambda_{\mathrm{dB}} = \frac{h}{p} = h\frac{\sqrt{1 - \frac{V^2}{c^2}}}{MV} = \lambda_{\mathrm{C}} = \frac{h}{Mc} \Leftrightarrow \frac{V}{c} = \frac{1}{\sqrt{2}}, \tag{37}$$

where p is the relativistic momentum. The same is the ratio of orbital to escape speed: $\frac{V_O}{V_E} = \frac{1}{\sqrt{2}}$.

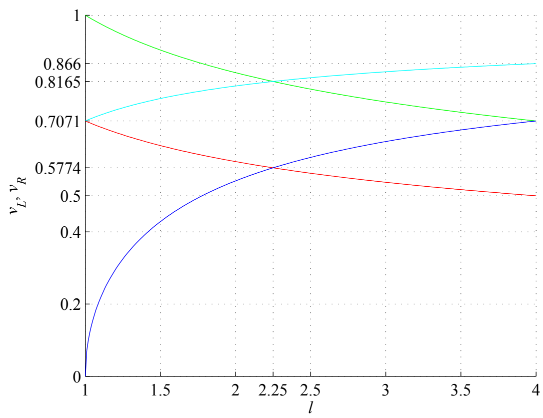

Figure 7. Lower (red) and upper (green) bounds on v_R and lower (blue) and upper (cyan) bounds on v_L as a function of l defining VS. Characteristic velocities are $\{0, \sqrt{1/4}, \sqrt{1/3}, \sqrt{2/4}, \sqrt{2/3}, \sqrt{3/4}, 1\}$, $v_L, v_R \in \mathbb{R}_+$.

Furthermore, the bounds (35) and (36) do not overlap only for $l = \{1, 4\}$. Therefore, $1 < l < 4$ defines the dissipativity or the assembly range. Furthermore, the intersection of the bounds (35) and (36) is the common region for both velocities. If v_L is within this region, then v_R is as well. We note that the average orbital velocity of each orbiting *object* only slightly exceeds its orbital speed V_O. This implies that the average VS defining factor $l_{\mathrm{avg}} \gtrsim 1$ in (28) for a VS orbiting *object* (cf. Appendix A).

BBs define a perfect thermodynamic equilibrium, and the bounds (28) and (29) show that nature uses optimally assembled information (cf. Conjecture 2) to assemble new information. Figure 8 shows the bounds on the string assembly indices and Figure 9 shows the BB temperature (24), energy (21), and entropic work (22) for $0 \leq N_{\mathrm{BB}} \leq 5$. $k_{\mathrm{B}}|T_{\mathrm{BB}}|/E_{\mathrm{BB}} = 2/N_{\mathrm{BB}}$ is a rational number for natural N_{BB}. Furthermore, $\log_2(N_{\mathrm{BB}}) > N_{\mathrm{BB}} - 1$ for $N_{\mathrm{BB}} \in \mathbb{R}$ and $1 < N_{\mathrm{BB}} < 2$.

Let us examine this process starting from the Big Bang during the Planck epoch and shortly thereafter, and for continuous $N_{\mathrm{BB}} \in \mathbb{R}$ (i.e., including fractional Planck triangle(s)).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

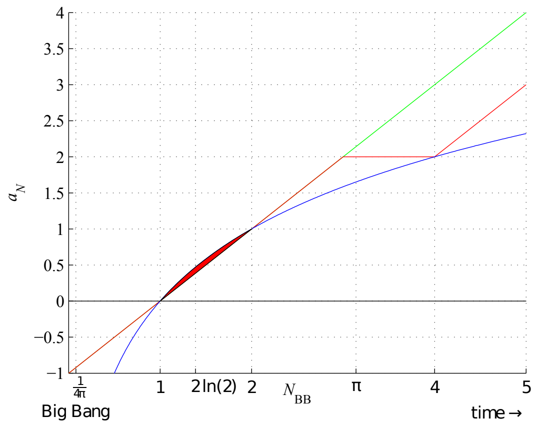

Figure 8. Lower (red) and upper (green) bounds on the binary string assembly index of length N_{BB} and $\log_2(N_{BB})$ (blue), for $0 \leq N_{BB} \leq 5$.

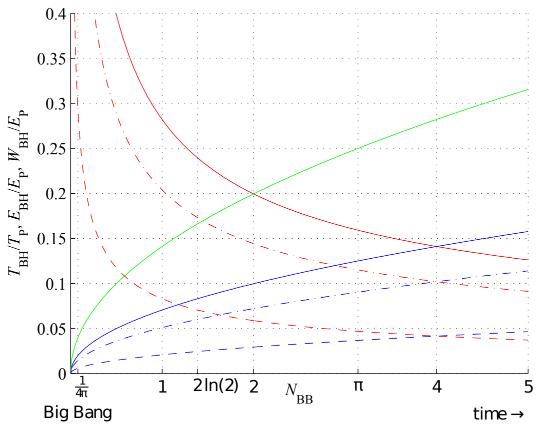

Figure 9. Black body object energy E_{BB} (green); temperature T_{BH} (red), $\text{Re}[T_{BB}(k_{eq})]$ (red, dash-dot), $\text{Re}[T_{BB}(k_{max})]$ (red, dash); and work W_{BH} (blue), $\text{Re}[W_{BB}(k_{eq})]$ (blue, dash-dot), $\text{Re}[W_{BB}(k_{max})]$ (blue, dash),as a function of its information capacity N_{BB} in terms of Planck units, for $0 \leq N_{BB} \leq 5$.

$N_{BB} = 0$

There is nothing to talk about. It is a mystery.
$0 < N_{BB} < 1$

The Big Bang has occurred, forming the 1st BB. At $N_{BB}(k_{max}) = (\alpha^4 - \alpha_2^4)/(4\pi\alpha^4) \approx 0.0069$ the BB temperature (24) and subsequently at $N_{BH} = 1/(4\pi) \approx 0.0796$ the BH temperature (24) become equal to the Planck temperature, but any BB in this range is still too small to carry a single bit of information and cannot be triangulated. However, independent BBs merge [9,10] summing their entropies and increasing the information capacity.
$N_{BB} = 1$

The first bit (a degree of freedom [9]) becomes available, and the BH energy reaches the limit of the equipartition theorem for one bit ($E_{BH} = \frac{1}{2}k_B T_{BH}$). APTs on BBs begin to fluctuate. However, the bounds (29) make them unable to generate any APTs on a VS ($N_1 = 0$).
$1 < N_{BB} < 2$

This is the only range in which the lower AT bound (1) is greater than the upper AT bound (2).

The BH temperature (24) exceeds its energy (21) ($\frac{1}{2}k_B T_{BH} < E_{BH} < k_B T_{BH}$) [9]. At $N_{BH} = 2\ln(2)$ the BH energy (21) is equal to the Landauer limit $E_{BH} = k_B T_{BH} \ln(2) \approx 1.3863$ [54]. Shortly thereafter, at $N_{BH} = 1.5$, the BH density reaches the level of the Planck density For a BB [10] Still $N_1 = 0$. Merging BBs expand fractional Planck triangle(s) to form the 2nd bit.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

$N_{\text{BB}} = 2$

The first nonvanishing $N_1 = 1$ becomes available on a VS generated by a BB. The BH temperature (24) is equal to its energy (21) ($k_{\text{B}} T_{\text{BH}} = E_{\text{BH}} = E_{\text{P}}/(2\sqrt{2\pi})$).

$2 < N_{\text{BB}} < 3$

At $N_{\text{BB}} = 4\ln(2)$ the BH entropic work (22) is equal to the Landauer limit ($k_{\text{B}} T_{\text{BH}} = W_{\text{BH}} = E_{\text{P}}/(4\sqrt{\pi})$). At $N_{\text{BB}} > 2.4507$ the density of the least dense BB ($k_{\max} \approx 6.7933$) decreases below the modulus of its temperature. $N_1 = \{0, 1\}$.

$N_{\text{BB}} = 3$

$3 < N_{\text{BB}} < 4$

With $N_{\text{BB}} > 3$ BBs can finally be triangulated. Yet, containing only one APT ($N_1 = \{0, 1\}$), they are not ergodic [9].

At $N_{\text{BH}} > \pi$ the BH surface gravity $g_{\text{BH}} = 1/d_{\text{BH}}$ decreases below the Planck acceleration and the tangential acceleration [8,9] becomes real ($a_L \in \mathbb{R}$).

$N_{\text{BB}} = 4$

The BB assembly index bifurcates, minimal thermodynamic entropy [31] is reached, and the relation (29) provides the second bit on a VS ($N_1 = 2$). At this moment BB can be assembled in a different number of steps and nature seeks to minimize this number following the dynamics induced by the relation (20). The BH temperature (24) is equal to its entropic work (22) ($k_{\text{B}} T_{\text{BH}} = W_{\text{BH}}$).

$4 < N_{\text{BB}} < 6$

The BH temperature (24) finally decreases below the entropic work (22) limit and $N_1 \geq 2$.

$N_{\text{BB}} = 6$

A BB reaches the upper bound on distinct assembly index.

$6 < N_{\text{BB}} < 7$

The imaginary Planck time appears at the BH surface [8] heralding the end of the Planck epoch. After crossing this threshold, the VSs begin to operate with $1 \leq N_1 \leq 3$ on $2\pi < N_{\text{VS}} \leq 8\pi$, and the first dissipative structures can be assembled.

Nature enters a directed exploration phase ($\alpha < 1$) and selectivity emerges, limiting the discovery of new objects [6].

$N_{\text{BB}} = 7$

A BB reaches the upper bound on nondistinct assembly index.

...

$N_{\text{BB}} > 12$

At $N_{\text{BB}} = 4\pi$ a first precise diameter relation can be established between the vertices of the BB surface. Furthermore, for $N_{\text{BB}} = 4\pi$, the solid angle (32) equals one steradian.

...

$N_{\text{BB}} > N_C$

The onset of human creativity.

8. Conclusions

The results reported here can be applied in the fields of cryptography, data compression methods, stream ciphers, approximation algorithms [55], reinforcement learning algorithms [56], information-theoretically secure algorithms, etc. Another possible application of the results of this study could be molecular physics and crystallography.

Overall, the results reported here support the AT, the Bekenstein's minimum of thermodynamic entropy [30–32], the holographic principle [33], entropic gravity [34], emergent dimensionality [8–10,15,17,18,20,23,28], the second law of infodynamics [16,19], and invite further research.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Author Contributions: WB: Conjecture concerning the diversification of strings in Theorem 1; partitioning conjecture for $N \geq 16$ resulting in the flattening of the string assembly index upper bound curve; observation that the string assembly index upper bound curve should BE monotonically non-decreasing; linear interpolation of VS defining factors; prior-art search; numerous clarity corrections and improvements; SŁ: The remaining part of the study.

Data Availability Statement: The public repository for the code written in MATLAB computational environment is given under the link https://github.com/szluk/Evolution_of_Information (accessed on November 23, 2023).

Acknowledgments: The authors thank Piotr Masierak for his creative remarks on the definition of a distinct string 3, research on the general strategy for determining the string assembly indices, and creating the $E_k^{(18)}$ string (*Peter's rock*), Andrzej Tomski for a clarity correction, and Mariola Bala for noting that "this is logical". SŁ thanks his wife Magdalena Bartocha for her unwavering support and motivation and his partner and friend, Renata Sobajda, for her prayers.

Abbreviations

The following abbreviations are used in this manuscript:

AT	assembly theory;
BH	black hole;
BB	black-body object (BH, white dwarf, neutron star);
VS	nonequilibrium shell;
APT	active Planck triangle;
N	length of a binary string;
N_0	number of 0's in the binary string;
N_1	number of 1's in the binary string (number of APTs);
$C_k^{(N)}$	binary string of length N;
$B_k^{(N)}$	balanced string of length N;
$D_k^{(N)}$	distinct string of length N;
$E_k^{(N)}$	balanced distinct string of length N;
$\lvert C^{(N)}\rvert$	number of binary strings of length N (2^N);
$\lvert B^{(N)}\rvert$	number of balanced strings of length N (OEIS A001405);
$\lvert D^{(N)}\rvert$	number of distinct strings (OEIS A000031);
$\lvert E^{(N)}\rvert$	number of balanced distinct strings;
$a^{(N)}$	assembly index of a string of length N;
P	assembly pool;
s	assembly step;
Q	binary assembling program;
s_Q	length of the binary assembling program;
F	Fibonacci sequence.

Appendix A. Orbital Velocities and the VS Defining Factor l

Table A1 shows the orbital speed V_O and escape speed V_E of some celestial objects, their minimal $V_{\min}$ and maximal $V_{\max}$ velocities[6]. The former ones lie below the orbital speed limits. The average VS defining factor $l_{\text{avg}} = (l_{\max} - l_{\min})/2$, where $l_{\min/\max} = 3(V_{\min/\max} - V_O)/(V_E - V_O) + 1$ were determined by linear interpolation.

[6] Based on https://sci.esa.int/web/solar-system.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Table A1. Exemplary orbital speeds and velocities, and the average VS defining factor l_{avg}.

Object	V_O [km/s]	V_{min} [km/s]	V_{max} [km/s]	V_E [km/s]	l_{avg}
Mercury	47.88	38.86	58.98	67.71	1.158
Venus	35.02	34.79	35.26	49.53	1.000
Earth	29.79	29.29	30.29	42.13	1.000
Mars	24.13	21.97	26.50	34.13	1.030
Jupiter	13.06	12.44	13.72	18.47	1.011
Saturn	9.62	9.09	10.18	13.61	1.009
Uranus	6.8	6.49	7.11	9.61	1.000
Neptune	5.43	5.37	5.50	7.68	1.001
Pluto	4.74	3.71	6.10	6.70	1.247
The Moon	1.02	0.96	1.08	14.40	1.011

Appendix B. Exemplary Strings with Maximal Assembly Indices

For the exemplary balanced distinct strings E_{max}, shown in Table 6:

- all forms of $E_k^{(4)} = [0011]$ have $a^{(4)} = 3$,
- all forms of $E_6^{(5)} = [00011]$ have $a^{(5)} = 4$,
- all forms of $E_{16}^{(6)} = [000111]$ have $a^{(6)} = 5$,
- the form $E_{28}^{(7)} = [0011100]$ has $a^{(7)} = 5$ but the form $E_{34}^{(7)} = [0001110]$ has $a^{(7)} = 6$,
- all forms of $E_{45}^{(8)} = [00010111]$ have $a^{(8)} = 6$,
- all forms of $E_{13}^{(9)} = [000011101]$ have $a^{(9)} = 7$,
- the form $E_{22}^{(10)} = [0000111101]$ has $a^{(10)} = 7$ but the form $E_l^{(10)} = [0111101000]$ has $a^{(10)} = 8$,
- all forms of $E_7^{(11)} = [00000101111]$ have $a^{(11)} = 8$,
- all forms of $E_9^{(12)} = [111000101100]$ have $a^{(12)} = 8$,
- all forms of $E_8^{(13)} = [0000001011111]$ have $a^{(13)} = 9$,
- all forms of $E_k^{(14)} = [00000101011111]$ have $a^{(14)} = 9$,
- all forms of $E_k^{(15)} = [000001010111110]$ have $a^{(15)} = 10$,
- all forms of $E_k^{(16)} = [1000000101011111]$ have $a^{(16)} = 10$,
- all forms of $E_k^{(17)} = [00000010101111110]$ have $a^{(17)} = 11$,
- all forms of $E_k^{(18)} = [000000101010111111]$ have $a^{(18)} = 11$,
- some forms of $E_k^{(19)} = [1000010101001111101]$ have $a^{(19)} = 12$,
- some forms of $E_k^{(20)} = [10100111110110000010]$ have $a^{(20)} = 13$.

Appendix C. Trivial Assembling Programs

Table A2 shows the lengths of the strings assembled by the trivial assembling program introduced in Section 6. The table is divided into sections corresponding to sets of assembled strings having the same form but different lengths. For example, thirty two 7-bit programs in the bottom section assemble strings $C = [*1*1\ldots]$.

Table A2. Lengths of the strings assembled by Qs (OEIS A065108).

$a^{(N)}$	0	1	2	3	3,4	4,5	5,6	6,7
s_Q		1	2	3	4	5	6	7
1	1	2	3	5	8	13	21	34
		│	│	│	│	│	⋆	42
		│	│	│	│	1	26	39
		│	│	│	│		⋆	52
		│	│	│	1	16	24	40
		│	│	│		│	⋆	48
		│	│	│		1	32	48
		│	│	│			⋆	64
		│	│	1	10	15	25	40
		│	│		│	│	⋆	50
		│	│		│	⋆	30	45
		│	│		│		⋆	60
		│	│		⋆	20	30	50
		│	│			│	⋆	60
		│	│			⋆	40	60
		│	│				⋆	80
		│	1	6	9	15	24	39
		│		│	│	│	⋆	48
		│		│	│	⋆	30	45
		│		│	│		⋆	60
		│		│	⋆	18	27	45
		│		│		│	⋆	54
		│		│		⋆	36	54
		│		│			⋆	72
		│		⋆	12	18	30	48
		│			│	│	⋆	60
		│			│	⋆	36	54
		│			│		⋆	72
		│			⋆	24	36	60
		│				│	⋆	72
		│				⋆	48	72
		│					⋆	96
		1	4	6	10	16	26	42
			│	│	│	│	⋆	52
			│	│	│	⋆	32	48
			│	│	│		⋆	64
			│	│	⋆	20	30	50
			│	│		│	⋆	60
			│	│		⋆	40	60
			│	│			⋆	80
			│	⋆	12	18	30	48
			│		│	│	⋆	60
			│		│	⋆	36	54
			│		│		⋆	72
			│		⋆	24	36	60
			│			│	⋆	72
			│			⋆	48	72
			│				⋆	96
			⋆	8	12	20	32	52
				│	│	│	⋆	64
				│	│	⋆	40	60
				│	│		⋆	80
				│	⋆	24	36	60
				│		│	⋆	72
				│		⋆	48	72
				│			⋆	96
				⋆	16	24	40	64
					│	│	⋆	80
					│	⋆	48	72
					│		⋆	96
					⋆	32	48	80
						│	⋆	96
						⋆	64	96
							⋆	128

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Table A3. 4-bit programs assembling strings with $a^{(N)} = \{3, 4\}$.

Q	$C(s = 3)$	$C(s_Q = 4)$	N
$*000$	$*11*1$	$1*1*11*1$	$8^{‡}$
$*001$	$*11*1$	$*11*1\ldots$	10
$*010$	$1*1\ldots$	$1*1\ldots$	9
$*011$	$1*1\ldots$	$1*1\ldots$	12
$*100$	$*1\ldots$	$*1\ldots$	10
$*101$	$*1\ldots$	$*1\ldots$	12
$*110$	$*1\ldots$	$*1\ldots$	12
$*111$	$*1\ldots$	$*1\ldots$	16

‡. This program is not elegant for $* = 1$ and the assembled string is not C_{min} for $* = 0$.

Table A4. 5-bit programs assembling strings with $a^{(N)} = \{4, 5\}$.

Q	$C(s = 4)$	$C(s_Q = 5)$	N
$*0000$	$1*1*11*1$	$*11*11*1*11*1$	13
$*0001$	$1*1*11*1$	$1*1*11*1\ldots$	$16^{‡}$
$*0010$	$*11*1\ldots$	$*11*1\ldots$	15
$*0011$	$*11*1\ldots$	$*11*1\ldots$	20
$*0100$	$1*1\ldots$	$1*1\ldots$	15
$*0101$	$1*1\ldots$	$1*1\ldots$	18
$*0110$	$1*1\ldots$	$1*1\ldots$	18
$*0111$	$1*1\ldots$	$1*1\ldots$	24
$*1000$	$*1\ldots$	$*1\ldots$	$16^{†}$
$*1001$	$*1\ldots$	$*1\ldots$	20
$*1010$	$*1\ldots$	$*1\ldots$	18
$*1011$	$*1\ldots$	$*1\ldots$	24
$*1100$	$*1\ldots$	$*1\ldots$	20
$*1101$	$*1\ldots$	$*1\ldots$	24
$*1110$	$*1\ldots$	$*1\ldots$	24
$*1111$	$*1\ldots$	$*1\ldots$	32

†. This program is not elegant (the same string can be assembled using the shorter 4-bit program $*111$). ‡. This program is not elegant for $* = 1$ and the assembled string is not C_{min} for $* = 0$.

Table A5. 6-bit programs assembling strings with $a^{(N)} = \{5, 6\}$.

Q	$C(s_Q = 6)$	N
$*00000$	$1*1*11*1*11*11*1*11*1$	21
$*00001$	$*11*11*1*11*1\ldots$	26
$*0001\star$	$1*1*11*1\ldots$	$24^{†}, 32^{‡}$
$*001\star\star$	$*11*1\ldots$	$25, 30, 40$
$*01\star\star\star$	$1*1\ldots$	$24^{†}, \ldots, 48$
$*1\star\star\star\star$	$*1\ldots$	$26, 32^{†}, \ldots, 64$

†. This program is not elegant. ‡. This program is not elegant for $* = 1$ and the assembled string is not C_{min} for $* = 0$.

Appendix D. Binary Strings and Their Assembly Indices

Table A2 show the lengths of the strings assembled by programs Fs having the minimal assembly indices. Tables A6-A13 show distributions of the assembly indices for $5 \leq N \leq 12$. Tables A14-A18 show balanced strings $B^{(N)}$ and their assembly indices for $5 \leq N \leq 8$. Tables A19-A24 show the balanced distinct strings $E^{(N)}$ and their assembly indices for $5 \leq N \leq 10$. Tables A25-A27 show selected balanced distinct strings $E^{(N)}$ and their assembly indices for $11 \leq N \leq 13$.

Table A6. Distribution of the assembly indices for $N = 5$.

		N_1					
$a^{(5)}(C)$	$\lvert a^{(5}(C)\rvert$	0	1	2	3	4	5
3	18	1	3	5	5	3	1
4	14		2	5	5	2	
	32	1	5	10	10	5	1

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Table A7. Distribution of the assembly indices for $N = 6$.

$a^{(6)}(C)$	$\lvert a^{(6)}(C)\rvert$	N_1: 0	1	2	3	4	5	6
3	10	1		3	2	3		1
4	44		6	10	12	10	6	
5	10			2	6	2		
	64	1	6	15	20	15	6	1

Table A8. Distribution of the assembly indices for $N = 7$.

$a^{(7)}(C)$	$\lvert a^{(7)}(C)\rvert$	N_1: 0	1	2	3	4	5	6	7
4	50	1	5	7	12	12	7	5	1
5	74		2	14	21	21	14	2	
6	4				2	2			
	128	1	7	21	35	35	21	7	1

Table A9. Distribution of the assembly indices for $N = 8$.

$a^{(8)}(C)$	$\lvert a^{(8)}(C)\rvert$	N_1: 0	1	2	3	4	5	6	7	8
3	4	1				2				1
4	38			9	8	4	8	9		
5	132		8	17	22	40	22	17	8	
6	82			2	26	24	26	2		
	256	1	8	28	56	70	56	28	8	1

Table A10. Distribution of the assembly indices for $N = 9$.

$a^{(9)}(C)$	$\lvert a^{(9)}(C)\rvert$	N_1: 0	1	2	3	4	5	6	7	8	9
4	24	1	3		3	5	5	3		3	1
5	184		4	17	35	36	36	35	17	4	
6	248		2	19	42	61	61	42	19	2	
7	56				4	24	24	4			
	512	1	9	36	84	126	126	84	36	9	1

Table A11. Distribution of the assembly indices for $N = 10$.

$a^{(10)}(C)$	$\lvert a^{(10)}(C)\rvert$	N_1: 0	1	2	3	4	5	6	7	8	9	10
4	20	1		3		5	2	5		3		1
5	198		8	22	20	33	32	33	20	22	8	
6	502		2	18	68	108	110	108	68	18	2	
7	288			2	32	62	96	62	32	2		
8	16					2	12	2				
	1024	1	10	45	120	210	252	210	120	45	10	

Table A12. Distribution of the assembly indices for $N = 11$.

$a^{(11)}(C)$	$\lvert a^{(11)}(C)\rvert$	N_1: 0	1	2	3	4	5	6	7	8	9	10	11
5	184	1	7	14	23	18	29	29	18	23	14	7	1
6	686		4	32	69	104	134	134	104	69	32	4	
7	970			9	69	178	229	229	178	69	9		
8	208				4	30	70	70	30	4			
	2048	1	11	55	165	330	462	462	330	165	55	11	

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Table A13. Distribution of the assembly indices for $N = 12$.

		N_1												
$a^{(12)}$	$\lvert a^{(12)}\rvert$	0	1	2	3	4	5	6	7	8	9	10	11	12
4	10	1				3		2		3				1
5	94			13	4	10	12	16	12	10	4	13		
6	1034		12	42	94	141	130	196	130	141	94	42	12	
7	1688			11	106	196	354	354	354	196	106	11		
8	1180				16	143	282	298	282	143	16			
9	90					2	14	58	14	2				
	4096	1	12	66	220	495	792	924	792	495	220	66	12	1

Table A14. $|B^{(5)}| = 10$ balanced strings.

k	$B_k^{(5)}$					$a^{(5)}(B_k)$
1	0	(0	1)	(0	1)	3
2	(0	1)	0	(0	1)	3
3	(0	1)	(0	1)	0	3
4	(1	0)	0	(1	0)	3
5	(1	0)	(1	0)	0	3
6	0	0	0	1	1	4
7	0	0	1	1	0	4
8	0	1	1	0	0	4
9	1	0	0	0	1	4
10	1	1	0	0	0	4

Table A15. $|B^{(6)}| = 20$ balanced strings.

k	$B_k^{(6)}$						$a^{(6)}(B_k)$
1	(0	1)	(0	1)	(0	1)	3
2	(1	0)	(1	0)	(1	0)	3
3	0	(0	1)	(0	1)	1	4
4	0	(0	1)	1	(0	1)	4
5	(0	1)	0	(0	1)	1	4
6	(0	1)	(0	1)	1	0	4
7	(0	1)	1	0	(0	1)	4
8	(0	1)	1	(0	1)	0	4
9	(1	0)	0	(1	0)	1	4
10	(1	0)	0	1	(1	0)	4
11	(1	0)	(1	0)	0	1	4
12	(1	0)	1	(1	0)	0	4
13	1	(1	0)	0	(1	0)	4
14	1	(1	0)	(1	0)	0	4
15	0	0	1	1	1	0	5
16	0	0	0	1	1	1	5
17	0	1	1	1	0	0	5
18	1	0	0	0	1	1	5
19	1	1	0	0	0	1	5
20	1	1	1	0	0	0	5

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Table A16. $|B^{(7)}| = 35$ balanced strings.

k	$B_k^{(7)}$							$a^{(7)}(B_k)$
1	0	(0	1)	(0	1)	(0	1)	4
2	(0	1)	(0	1)	(0	1)	0	4
3	(1	0)	(1	0)	(1	0)	0	4
4	(0	1)	(0	1)	0	(0	1)	4
5	(1	0)	(1	0)	0	(1	0)	4
6	(0	1)	0	(0	1)	(0	1)	4
7	(1	0)	0	(1	0)	(1	0)	4
8	(1	0	0)	(1	0	0)	1	4
9	(1	0	0)	1	(1	0	0)	4
10	1	(1	0	0)	(1	0	0)	4
11	(0	0	1)	1	(0	0	1)	4
12	(0	0	1)	(0	0	1)	1	4
13	1	(0	0)	(0	0)	1	1	5
14	1	0	0	(0	1)	(0	1)	5
15	(1	0)	0	0	1	(1	0)	5
16	(1	0)	(1	0)	0	0	1	5
17	(1	0)	1	(1	0)	0	0	5
18	1	1	(0	0)	(0	0)	1	5
19	1	(1	0)	(1	0)	0	0	5
20	1	1	1	(0	0)	(0	0)	5
21	(0	1)	(0	1)	1	0	0	5
22	(0	1)	1	0	0	(0	1)	5
23	(0	1)	1	0	(0	1)	0	5
24	(0	1)	1	(0	1)	0	0	5
25	(0	1)	0	(0	1)	1	0	5
26	0	(0	1)	(0	1)	1	0	5
27	0	(0	1)	1	(0	1)	0	5
28	(0	0)	1	1	1	(0	0)	5
29	(0	1)	0	0	(0	1)	1	5
30	(0	0)	(0	0)	1	1	1	5
31	0	0	(0	1)	(0	1)	1	5
32	0	0	(0	1)	1	(0	1)	5
33	1	(1	0)	0	0	(1	0)	5
34	0	0	0	1	1	1	0	6
35	0	1	1	1	0	0	0	6

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Table A17. $|B^{(8)}| = 70$ balanced strings (1^{st} part).

k	$B_k^{(8)}$								$a^{(8)}(B_k)$
1	((0	1)	(0	1))	((0	1)	(0	1))	3
2	((1	0)	(1	0))	((1	0)	(1	0))	3
3	((0	0)	(1	1))	((0	0)	(1	1))	4
4	((0	1)	(1	0))	((0	1)	(1	0))	4
5	((1	0)	(0	1))	((1	0)	(0	1))	4
6	((1	1)	(0	0))	((1	1)	(0	0))	4
7	(0	0)	(0	0)	(1	1)	(1	1)	5
8	(0	0	1)	(0	0	1)	1	1	5
9	0	(0	1)	(0	1)	(0	1)	1	5
10	0	(0	1)	(0	1)	1	(0	1)	5
11	0	(0	1)	1	(0	1)	(0	1)	5
12	(0	0	1)	1	1	(0	0	1)	5
13	(0	0)	(1	1)	(1	1)	(0	0)	5
14	(0	1)	0	(0	1)	(0	1)	1	5
15	(0	1)	0	(0	1)	1	(0	1)	5
16	(0	1)	(0	1)	0	(0	1)	1	5
17	(0	1)	(0	1)	(0	1)	1	0	5
18	(0	1)	(0	1)	1	0	(0	1)	5
19	(0	1)	(0	1)	1	(0	1)	0	5
20	(0	1	1)	0	0	(0	1	1)	5
21	(0	1)	1	0	(0	1)	(0	1)	5
22	(0	1)	1	(0	1)	0	(0	1)	5
23	(0	1)	1	(0	1)	(0	1)	0	5
24	(0	1	1)	(0	1	1)	0	0	5
25	(1	0	0)	(1	0	0)	1	1	5
26	1	0	(0	1)	(0	1)	(0	1)	5
27	(1	0)	0	(1	0)	1	(1	0)	5
28	(1	0)	0	1	(1	0)	(1	0)	5
29	(1	0	0)	1	1	(1	0	0)	5
30	(1	0	1)	0	0	(1	0	1)	5
31	(1	0)	(1	0)	0	1	(1	0)	5
32	(1	0)	(1	0)	(1	0)	0	1	5
33	(1	0)	(1	0)	1	(1	0)	0	5
34	(1	0)	1	(1	0)	0	(1	0)	5
35	(1	0)	1	(1	0)	(1	0)	0	5
36	(1	1)	(0	0)	(0	0)	(1	1)	5
37	(1	1	0)	0	0	(1	1	0)	5
38	1	1	(0	0	1)	(0	0	1)	5
39	1	1	0	0	1	0	1	0	5
40	1	(1	0)	(1	0)	0	(1	0)	5
41	(1	1	0)	(1	1	0)	0	0	5
42	1	1	0	1	0	1	0	0	5
43	1	1	(1	0	0)	(1	0	0)	5
44	(1	1)	(1	1)	(0	0)	(0	0)	5
45	0	0	(0	1	1)	(0	1	1)	5
46	0	(0	1	1)	(0	1	1)	0	5

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Table A18. $|B^{(8)}| = 70$ balanced strings (2[nd] part).

k	$B_k^{(8)}$								$a^{(8)}(B_k)$
47	0	0	(0	1)	(0	1)	1	1	6
48	0	0	(0	1)	1	1	(0	1)	6
49	0	0	0	(1	1)	(1	1)	0	6
50	0	(0	1)	(0	1)	1	1	0	6
51	0	0	1	1	(1	0)	(1	0)	6
52	(0	1)	0	0	(0	1)	1	1	6
53	(0	1)	0	(0	1)	1	1	0	6
54	(0	1)	(0	1)	1	1	0	0	6
55	(0	1)	1	1	0	0	(0	1)	6
56	(0	1)	1	1	0	(0	1)	0	6
57	(0	1)	1	1	(0	1)	0	0	6
58	0	(1	1)	(1	1)	0	0	0	6
59	1	(0	0)	(0	0)	1	1	1	6
60	(1	0)	0	0	(1	0)	1	1	6
61	1	0	0	(0	1)	1	(0	1)	6
62	(1	0)	0	0	1	1	(1	0)	6
63	(1	0)	(1	0)	0	0	1	1	6
64	(1	0)	1	(1	0)	0	0	1	6
65	(1	0)	1	1	(1	0)	0	0	6
66	1	(1	0)	0	0	(1	0)	1	6
67	1	1	(0	1)	0	0	(0	1)	6
68	1	1	1	(0	0)	(0	0)	1	6
69	1	1	(1	0)	0	0	(1	0)	6
70	1	1	(1	0)	(1	0)	0	0	6

Table A19. $|E^{(5)}| = 2$ balanced distinct strings.

k	$E_k^{(5)}$					$a^{(5)}(E_k)$
1	0	(0	1)	(0	1)	3
6	0	0	0	1	1	4

Table A20. $|E^{(6)}| = 4$ balanced distinct strings.

k	$E_k^{(6)}$						$a^{(6)}(E_k)$
1	(0	1)	(0	1)	(0	1)	3
3	0	(0	1)	(0	1)	1	4
4	0	(0	1)	1	(0	1)	4
16	0	0	0	1	1	1	5

Table A21. $|E^{(7)}| = 5$ balanced distinct strings.

k	$E_k^{(7)}$							$a^{(7)}(E_k)$
1	0	(0	1)	(0	1)	(0	1)	4
12	(0	0	1)	(0	0	1)	1	4
30	(0	0)	(0	0)	1	1	1	5
31	0	0	(0	1)	(0	1)	1	5
32	0	0	(0	1)	1	(0	1)	5

Table A22. $|E^{(8)}| = 10$ balanced distinct strings.

k	$E_k^{(8)}$								$a^{(8)}(E_k)$
1	((0	1)	(0	1))	((0	1)	(0	1))	3
3	(0	0)	(1	1)	(0	0)	(1	1)	4
7	(0	0)	(0	0)	(1	1)	(1	1)	5
8	0	(0	1)	0	(0	1)	1	1	5
9	0	(0	1)	(0	1)	(0	1)	1	5
10	0	(0	1)	(0	1)	1	(0	1)	5
11	0	(0	1)	1	(0	1)	(0	1)	5
46	0	0	(0	1	1)	(0	1	1)	5
45	0	0	(0	1)	(0	1)	1	1	6
47	0	0	(0	1)	1	1	(0	1)	6

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Table A23. Selected balanced distinct strings $|E^{(9)}| = 14$.

k	$E_k^{(9)}$									$a^{(9)}(E_k)$
1	0	((0	1)	(0	1))	((0	1)	(0	1))	4
2	0	((0	0)	(1	1))	((0	0)	(1	1))	5
3	(0	(0	1))	(0	1)	(0	0	1)	1	5
4	(0	(0	1))	(0	0	1)	1	(0	1)	5
5	(0	(0	1))	(0	0	1)	(0	1)	1	5
6	0	(0	0	1)	1	1	(0	0	1)	6
7	0	0	(0	1)	1	(0	1)	(0	1)	6
8	0	0	(0	1)	(0	1)	1	(0	1)	6
9	0	0	(0	1)	(0	1)	(0	1)	1	6
10	0	(0	0	1)	(0	0	1)	1	1	6
11	(0	0)	(0	0)	(1	1)	0	(1	1)	6
12	0	(0	0)	(0	0)	(1	1)	(1	1)	6
13	(0	0)	(0	0)	1	1	1	0	1	7
14	(0	0)	(0	0)	1	0	1	1	1	7

Table A24. $|E^{(10)}| = 26$ balanced distinct strings.

k	$E_k^{(10)}$										$a^{(10)}(E_k)$
1	((0	1)	(0	1))	((0	1)	(0	1))	(0	1)	4
2	0	((0	1)	(0	1))	((0	1)	(0	1))	1	5
3	(0	1)	(1	(0	1)	0)	(1	(0	1)	0)	5
4	(0	(0	1)	1)	(0	0	1	1)	(0	1)	5
5	0	((0	1)	0	1)	1	(0	1	0	1)	5
6	0	((1	0)	1	0)	1	(1	0	1	0)	5
7	(0	1)	((0	1)	1	0)	(0	1	1	0)	5
8	(0	(0	1))	(0	1)	(0	0	1)	1	1	6
9	(0	(0	1))	(0	0	1)	1	1	(0	1)	6
10	(0	(0	1))	(0	0	1)	1	(0	1)	1	6
11	(0	(0	1))	(0	0	1)	(0	1)	1	1	6
14	0	(0	0	1	1)	1	(0	0	1	1)	6
15	0	0	((0	1)	1)	(0	1	1)	(0	1)	6
16	0	0	((0	1)	1)	(0	1)	(0	1	1)	6
17	0	(0	0	1	1)	(0	0	1	1)	1	6
19	0	0	(0	1)	((0	1)	1)	(0	1	1)	6
12	(0	0)	0	(1	1)	(1	1)	(0	0)	1	7
13	0	0	(0	1)	1	1	(0	1)	(0	1)	7
18	0	0	(0	1)	(0	1)	1	1	(0	1)	7
20	0	0	(0	1)	(0	1)	(0	1)	1	1	7
21	(0	0)	0	1	(0	0)	(1	1)	(1	1)	7
22	(0	0)	(0	0)	(1	1)	(1	1)	0	1	7
23	(0	0)	(0	0)	(1	1)	1	0	(1	1)	7
24	(0	0)	(0	0)	(1	1)	0	(1	1)	1	7
25	(0	0)	(0	0)	1	0	(1	1)	(1	1)	7
26	(0	0)	(0	0)	0	1	(1	1)	(1	1)	7

Table A25. Selected balanced distinct strings $E^{(11)}$.

k	$E_k^{(11)}$											$a^{(11)}(E_k)$
1	0	(0	1)	((0	1))	(0	1))	(0	1	0	1)	5
2	(0	(0	1)	(0	1))	(0	0	1	0	1)	1	5
3	(0	0)	((0	0)	1	1)	(1	0	0	1	1)	6
4	(0	(0	1))	(0	1)	(0	1)	(0	0	1)	1	6
5	(0	0)	(0	0)	(0	0)	(1	1)	(1	1)	1	7
6	(0	0)	(1	1	0)	1	(0	0)	(1	1	0)	7
7	(0	0)	(0	0)	(0	1)	(0	1)	1	1	1	8

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

Table A26. Selected balanced distinct strings $E^{(12)}$.

k	$E_k^{(12)}$												$a^{(12)}(E_k)$
1	((0	1)	(0	1))	(0	1	0	1)	(0	1	0	1)	4
2	(0	(0	1)	1	(0	1))	(0	0	1	1	0	1)	5
3	((0	1)	1	(0	(0	1)))	((0	1)	1	(0	0	1))	5
4	(0	(0	1)	1)	(0	0	1	1)	(0	1)	(0	1)	6
5	((0	1)	0	(0	1))	(0	1	0	0	1)	1	1	6
6	(0	0	1)	(0	0	1)	(0	0	1)	1	1	1	7
7	(0	0)	(0	0)	(0	0)	(1	1)	(1	1)	(1	1)	7
8	(0	0)	(0	0)	(1	1)	(1	1)	1	(0	0)	1	8
9	(0	0)	(1	0)	(1	1)	(0	0)	(1	1)	(1	0)	8
10	(1	1)	(1	1)	(0	1)	(0	1)	(0	0)	(0	0)	8
11	(1	1)	(1	1)	(0	0)	(0	0)	(1	0)	(1	0)	8

Table A27. Selected balanced distinct strings $E^{(13)}$.

k	$E_k^{(13)}$													$a^{(13)}(E_k)$
1	0	((0	1)	(0	1))	(0	1	0	1)	(0	1	0	1)	5
2	0	((1	0)	0	1	(1	0))	(1	0	0	1	1	0)	6
3	(0	((0	1)	(0	1))	(0	0	1	0	1)	(0	1)	1	6
4	0	(0	0)	((0	0)	(1	1))	(0	0	1	1)	(1	1)	7
5	(0	0)	((0	0)	(1	1))	(0	0	1	1)	0	(1	1)	7
6	(0	0)	(0	0)	(0	0)	0	(1	1)	(1	1)	(1	1)	8
7	(0	0	(0	1))	(0	0	0	1)	(0	1)	1	1	1	8
8	(0	0)	(0	0)	(0	0)	1	0	(1	1)	(1	1)	1	9

References

1. Marshall, S.M.; Murray, A.R.G.; Cronin, L. A probabilistic framework for identifying biosignatures using Pathway Complexity. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **2017**, *375*, 20160342. doi:10.1098/rsta.2016.0342.
2. Murray, A.; Marshall, S.; Cronin, L. Defining Pathway Assembly and Exploring its Applications, 2018. arXiv:1804.06972 [cs, math].
3. Marshall, S.M.; Mathis, C.; Carrick, E.; Keenan, G.; Cooper, G.J.T.; Graham, H.; Craven, M.; Gromski, P.S.; Moore, D.G.; Walker, S.I.; Cronin, L. Identifying molecules as biosignatures with assembly theory and mass spectrometry. *Nature Communications* **2021**, *12*, 3033. doi:10.1038/s41467-021-23258-x.
4. Liu, Y.; Mathis, C.; Bajczyk, M.D.; Marshall, S.M.; Wilbraham, L.; Cronin, L. Exploring and mapping chemical space with molecular assembly trees. *Science Advances* **2021**, *7*, eabj2465. doi:10.1126/sciadv.abj2465.
5. Marshall, S.M.; Moore, D.G.; Murray, A.R.G.; Walker, S.I.; Cronin, L. Formalising the Pathways to Life Using Assembly Spaces. *Entropy* **2022**, *24*, 884. doi:10.3390/e24070884.
6. Sharma, A.; Czégel, D.; Lachmann, M.; Kempes, C.P.; Walker, S.I.; Cronin, L. Assembly theory explains and quantifies selection and evolution. *Nature* **2023**, *622*, 321–328. doi:10.1038/s41586-023-06600-9.
7. Jirasek, M.; Sharma, A.; Bame, J.R.; Mehr, S.H.M.; Bell, N.; Marshall, S.M.; Mathis, C.; Macleod, A.; Cooper, G.J.T.; Swart, M.; Mollfulleda, R.; Cronin, L. Determining Molecular Complexity using Assembly Theory and Spectroscopy, 2023. arXiv:2302.13753 [physics, q-bio].
8. Łukaszyk, S., Black Hole Horizons as Patternless Binary Messages and Markers of Dimensionality. In *Future Relativity, Gravitation, Cosmology*; Nova Science Publishers, 2023; chapter 15, pp. 317–374. doi:10.52305/RLIT5885.
9. Łukaszyk, S. Life as the Explanation of the Measurement Problem. *Journal of Physics: Conference Series* **2024**, *2701*, 012124. doi:10.1088/1742-6596/2701/1/012124.
10. Łukaszyk, S. The Imaginary Universe. preprint, PHYSICAL SCIENCES, 2023.
11. de Chardin, P.T. *The Phenomenon of Man*; Harper, New York, 1959.
12. Prigogine, I.; Stengers, I. *Order out of Chaos: Man's New Dialogue with Nature*; Bantam Books, 1984.
13. Melamede, R. Dissipative Structures and the Origins of Life. Unifying Themes in Complex Systems IV; Minai, A.A.; Bar-Yam, Y., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2008; pp. 80–87.
14. Vedral, V. *Decoding Reality: The Universe as Quantum Information*; Oxford University Press, 2010. doi:https://doi.org/10.1093/osc
15. Łukaszyk, S. Four Cubes, 2020. doi:10.48550/ARXIV.2007.03782.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

16. Vopson, M.M.; Lepadatu, S. Second law of information dynamics. *AIP Advances* **2022**, *12*, 075310. doi:10.1063/5.0100358.
17. Łukaszyk, S. Novel Recurrence Relations for Volumes and Surfaces of n-Balls, Regular n-Simplices, and n-Orthoplices in Real Dimensions. *Mathematics* **2022**, *10*. doi:10.3390/math10132212.
18. Łukaszyk, S.; Tomski, A. Omnidimensional Convex Polytopes. *Symmetry* **2023**, *15*. doi:10.3390/sym15030755.
19. Vopson, M.M. The second law of infodynamics and its implications for the simulated universe hypothesis. *AIP Advances* **2023**, *13*, 105308. doi:10.1063/5.0173278.
20. Łukaszyk, S. A No-go Theorem for Superposed Actions (Making Schrödinger's Cat Quantum Nonlocal). In *New Frontiers in Physical Science Research Vol. 3*; Purenovic, D.J., Ed.; Book Publisher International (a part of SCIENCEDOMAIN International), 2022; pp. 137–151. doi:10.9734/bpi/nfpsr/v3/17106D.
21. Qian, K.; Wang, K.; Chen, L.; Hou, Z.; Krenn, M.; Zhu, S.; Ma, X.s. Multiphoton non-local quantum interference controlled by an undetected photon. *Nature Communications* **2023**, *14*, 1480. doi:10.1038/s41467-023-37228-y.
22. Xue, P.; Xiao, L.; Ruffolo, G.; Mazzari, A.; Temistocles, T.; Cunha, M.T.; Rabelo, R. Synchronous Observation of Bell Nonlocality and State-Dependent Contextuality. *Physical Review Letters* **2023**, *130*, 040201. doi:10.1103/PhysRevLett.130.040201.
23. Łukaszyk, S. Shannon Entropy of Chemical Elements. *European Journal of Applied Sciences* **2024**, *11*, 443–458. doi:10.14738/aivp.116.16194.
24. Tran, D.M.; Nguyen, V.D.; Ho, L.B.; Nguyen, H.Q. Increased success probability in Hardy's nonlocality: Theory and demonstration. *Phys. Rev. A* **2023**, *107*, 042210. doi:10.1103/PhysRevA.107.042210.
25. Colciaghi, P.; Li, Y.; Treutlein, P.; Zibold, T. Einstein-Podolsky-Rosen Experiment with Two Bose-Einstein Condensates. *Phys. Rev. X* **2023**, *13*, 021031. doi:10.1103/PhysRevX.13.021031.
26. Watanabe, S. *Knowing and Guessing: A Quantitative Study of Inference and Information*; Wiley, 1969.
27. Watanabe, S. Epistemological Relativity. *Annals of the Japan Association for Philosophy of Science* **1986**, *7*, 1–14. doi:10.4288/jafpos1956.7.1.
28. Łukaszyk, S. Metallic Ratios and Angles of a Real Argument. *IPI Letters* **2024**, pp. 26–33. doi:10.59973/ipil.55.
29. Shannon, C.E. A Mathematical Theory of Communication. *Bell System Technical Journal* **1948**, *27*, 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x.
30. Bekenstein, J.D. Black holes and the second law. *Lettere Al Nuovo Cimento Series 2* **1972**, *4*, 737–740. doi:10.1007/BF02757029.
31. Bekenstein, J.D. Black Holes and Entropy. *Phys. Rev. D* **1973**, *7*, 2333–2346. doi:10.1103/PhysRevD.7.2333.
32. Hawking, S.W. Particle creation by black holes. *Communications In Mathematical Physics* **1975**, *43*, 199–220. doi:10.1007/BF02345020.
33. Hooft, G.t. Dimensional Reduction in Quantum Gravity, 1993. doi:10.48550/ARXIV.GR-QC/9310026.
34. Verlinde, E. On the origin of gravity and the laws of Newton. *Journal of High Energy Physics* **2011**, *2011*, 29. doi:10.1007/JHEP04(2011)029.
35. Misner, C.W.; Thorne, K.S.; Wheeler, J.A. *Gravitation*; W. H. Freeman: San Francisco, 1973.
36. Gould, A. Classical derivation of black-hole entropy. *Physical Review D* **1987**, *35*, 449–454. doi:10.1103/PhysRevD.35.449.
37. Penrose, R.; Floyd, R.M. Extraction of Rotational Energy from a Black Hole. *Nature Physical Science* **1971**, *229*, 177–179. doi:10.1038/physci229177a0.
38. Christodoulou, D.; Ruffini, R. Reversible Transformations of a Charged Black Hole. *Physical Review D* **1971**, *4*, 3552–3555. doi:10.1103/PhysRevD.4.3552.
39. Stuchlík, Z.; Kološ, M.; Tursunov, A. Penrose Process: Its Variants and Astrophysical Applications. *Universe* **2021**, *7*, 416. doi:10.3390/universe7110416.
40. Downey, P.; Leong, B.; Sethi, R. Computing Sequences with Addition Chains. *SIAM Journal on Computing* **1981**, *10*, 638–646, [https://doi.org/10.1137/0210047]. doi:10.1137/0210047.
41. Chaitin, G.J. Randomness and Mathematical Proof. *Scientific American* **1975**, *232*, 47–52. doi:10.1038/scientificamerican0575-47.
42. Chaitin, G.J. *The unknowable*; Springer series in discrete mathematics and theoretical computer science, Springer: Singapore ; New York, 1999.
43. Rajput, C. Metallic Ratios in Primitive Pythagorean Triples : Metallic Means embedded in Pythagorean Triangles and other Right Triangles. *JOURNAL OF ADVANCES IN MATHEMATICS* **2021**, *20*, 312–344. doi:10.24297/jam.v20i.9088.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 March 2024 doi:10.20944/preprints202401.1113.v5

44. Chaitin, G. *From Philosophy to Program Size: Key Ideas and Methods : Lecture Notes on Algorithmic Information Theory from the 8th Estonian Winter School in Computer Science, EWSCS'03 : [2-7 March]*; Institute of Cybernetics at Tallinn Technical University, 2003.
45. Chaitin, G. Doing Mathematics Differently. *Inference: International Review of Science* **2016**, *2*. doi:10.37282/991819.16.5.
46. Kolmogorov, A. On tables of random numbers. *Theoretical Computer Science* **1998**, *207*, 387–395. doi:10.1016/S0304-3975(98)00075-9.
47. Cronin, Leroy. Lee Cronin: Controversial Nature Paper on Evolution of Life and Universe | Lex Fridman Podcast #404, 2023. Accessed: 2023-12-18.
48. Zatorski, W. *W życiu chodzi o życie*, wydanie i ed.; Tyniec Wydawnictwo Benedyktynów: Kraków, 2020. OCLC: 1243002797.
49. Barta, J.; Markiewicz, R., Eds. *Prawo autorskie: przepisy, orzecznictwo, umowy miedzynarodowe*, wyd. 4., rozsz. i zaktualizowane ed.; Dom Wydawniczy ABC: Warszawa, 2002.
50. Mandelstam, L.; Tamm, I. The Uncertainty Relation Between Energy and Time in Non-relativistic Quantum Mechanics. *J. Phys. (USSR)* **1945**, *9*, 249—-254.
51. Margolus, N.; Levitin, L.B. The maximum speed of dynamical evolution. *Physica D: Nonlinear Phenomena* **1998**, *120*, 188–195. doi:10.1016/S0167-2789(98)00054-2.
52. Levitin, L.B.; Toffoli, T. Fundamental Limit on the Rate of Quantum Dynamics: The Unified Bound Is Tight. *Physical Review Letters* **2009**, *103*, 160502. doi:10.1103/PhysRevLett.103.160502.
53. Hossenfelder, S. Comments on and Comments on Comments on Verlinde's paper "On the Origin of Gravity and the Laws of Newton", 2010. arXiv:1003.1015 [gr-qc].
54. Landauer, R. Irreversibility and Heat Generation in the Computing Process. *IBM Journal of Research and Development* **1961**, *5*, 183–191. doi:10.1147/rd.53.0183.
55. Łukaszyk, S. A new concept of probability metric and its applications in approximation of scattered data sets. *Computational Mechanics* **2004**, *33*, 299–304. doi:https://doi.org/10.1007/s00466-003-0532-2.
56. Castro, P.S.; Kastner, T.; Panangaden, P.; Rowland, M. MICo: Improved representations via sampling-based state similarity for Markov decision processes. *Advances in Neural Information Processing Systems* **2021**, *34*, 30113–30126. doi:10.48550/ARXIV.2106.08229.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.