

Information Is Four Quantities, and They Need Different Things

— A Distribution, a State, a Family, or Nothing at All: The One That Needs Nothing Supplied Is the One That
Cannot Be Computed — [Paper 252]

Yuuki YAMAGISHI (山岸 勇輝)

Yamagishi Complexity Torus Lab, Japan

September 2026 (Second Edition)

— *From thinking to syncing.* —

Abstract

One word, information, names at least four different quantities. This paper calls the roll and writes down what each one cannot be defined without. No new mathematical theorem and no new law is claimed. First, Shannon needs a distribution. Fitting three distributions to the same 200000 symbols gives H of **1.253874709**, **1.254310170** and **2.000000000** bits: not one symbol was changed (Section 2). Second, von Neumann needs a state. The pure state $|+\rangle$ and the completely mixed state both return 0.5 and 0.5 when measured along z , yet S is **0.000000000** and **1.000000000** bits: the outcomes do not fix it (Section 3). Third, Fisher needs a family, not a distribution. Embedding the same standard normal in $N(\theta, 1)$ gives **1** and in $N(0, \sigma^2)$ gives **2** (numerically **1.000050** and **1.998993**) (Section 4). Fourth, Kolmogorov needs nothing, and cannot be computed. The 40000 decimal digits of π come out of a self-contained program of **257** bytes, while the best compressor reaches 17271 bytes, **67** times as long (Section 5). Fifth, this is the core. Two quantities called information return opposite verdicts on one and the same object. For the digits of π , Shannon gives **3.321837** bits per digit, all but the maximum $\log_2 10 = 3.321928$, while Kolmogorov gives at most **0.051400**, a factor of **65**, one near the ceiling and one near the floor (Section 6). Sixth, the alignment is exact. A distribution, a state, a family: the three that demand something can be computed, and the one that demands nothing cannot (Section 7).

Scope of this paper (scope note) : No new mathematical theorem and no new law is claimed: Shannon entropy, von Neumann entropy, Fisher information, Kolmogorov complexity and its uncomputability (Chaitin) are all standard. No theory of algorithmic information is built: Section 5 exhibits one measured upper bound and nothing more. The ratio of compressor performance is not used as a lower bound on K : a compressor failing to shorten a string is no evidence that the string is incompressible (Paper 223, Section 7). It is not claimed that the digits of π are normal. That is an open problem, and the chi-squared of Section 6 says only that this test does not see a departure from uniformity. It is not claimed that four is all of them: Renyi, Tsallis, mutual information, the Fisher information matrix and minimum description length are further names. It is not judged which is the right information. What can be written is what each one demands. Mutual information and conditional entropy are not treated. That is the province of Paper 162. No quantum information theory is built: Section 3 uses two 2×2 density matrices. The alignment of Section 7 is not claimed as a law. It is an observation on four points, not a theorem. The relation to earlier papers. Paper 223 showed that a sequence can pass every test and still be short to write, and stated the uncomputability of K . This paper does not repeat that demonstration; it cites 223 and adds only where that uncomputability sits in the roll call. Paper 133 showed the uniqueness of the Fisher metric is a classical statement and Paper 137 that the second derivative of KL is a metric: Section 4 takes neither, only the single point that the number is undetermined until a family is supplied. Paper 162 showed the whole can be known less than the part: Section 3 uses one state and goes no further. Paper 88 warned against double counting two entropies: that concerns physical quantities and this a roll call of a word; the axes differ. Paper 248 showed the demon's cost was placed in five places and Paper 164 gave probability four addresses. This paper performs the same kind of roll call on information. Paper 163 showed that existence and construction are different sentences: the asymmetry of Section 7 is a relative. What is added is ordering the four by what cannot be defined without it, obtaining 1 and 2 from one and the same standard normal, obtaining $S = 0$ and $S = 1$ from one and the same set of outcomes, and showing with numbers that two informations return opposite verdicts on a single object.

1. Notation and standing

The setting: the quantities called information are not one. **Four are treated here:** Shannon entropy H , von Neumann entropy S , Fisher information I , and Kolmogorov complexity K .

The question: what separates them? **Not the unit:** all but I are written in bits. **Not the size:** on one object they can be large and small at once (Section 6).

The separator of this paper is what each cannot be defined without. Handed the object alone and asked to compute each quantity, what would one have to say is missing?

It is not judged which is the right information. Each of the four answers correctly once given what it demands.

2. Shannon: not defined without a distribution

Definition: for a probability distribution p , $H(p) = -\sum_i p_i \log_2 p_i$. The argument is a distribution, not data.

Check 2.1: 200000 symbols over four letters are generated, and with that one sequence held fixed, three distributions are fitted.

distribution fitted	H (bits)
the true one, (0.7, 0.2, 0.06, 0.04)	1.253874709
the empirical one, from the sequence	1.254310170
assumed uniform, (0.25, 0.25, 0.25, 0.25)	2.000000000

Observation 2.2

Not one symbol changed; what changed is the assumed distribution. What H measures is not the sequence but the distribution fitted to it.

Observation 2.3

The question what is the entropy of this file has, as it stands, no answer. An answer appears only after a distribution has been chosen, and using the empirical one is itself a choice.

3. Von Neumann: not defined without a state, and outcomes are not enough

Definition: for a density matrix ρ , $S(\rho) = -\text{tr} \rho \log_2 \rho$. The argument is a state, not the statistics of a measurement.

Check 3.1: take the pure state $|+\rangle$, that is $\rho = |+\rangle\langle+|$, and the completely mixed $\rho = I/2$. Measured along z , both give up and down equally.

state	outcomes along z	S (bits)
pure $ +\rangle$	0.5 / 0.5	0.000000000
completely mixed $I/2$	0.5 / 0.5	1.000000000

Observation 3.2

The outcomes are the same and S is 0 against 1. Given only the z statistics, S cannot be computed.

Observation 3.3

There is a way to tell them apart: measure along x . The pure state answers + with certainty; the mixed one answers evenly. The point is that the z statistics are not enough as a way of being given the object. What S demands is the state itself, not a distribution in one basis.

Observation 3.4

This is also the difference from Shannon. Applying H to the z statistics (0.5, 0.5) gives one bit for both: Shannon cannot separate them. Two quantities return different answers to the same question.

4. Fisher: not defined without a family, and a distribution is not one

Definition: for a family $f(x; \theta)$ indexed by a parameter, $I(\theta) = E[(\partial_\theta \log f)^2]$. The argument is the family, not one distribution inside it.

Check 4.1: fix one distribution, the standard normal $N(0, 1)$, and embed it in two families.

family	point	I (analytic)	I (numerical, $n = 4 \times 10^6$)
$N(\theta, 1)$	$\theta = 0$	1	1.000050
$N(0, \sigma^2)$	$\sigma = 1$	2	1.998993

Observation 4.2

The distribution is one and the same. Only the parametrisation differs, and that alone makes I equal to 1 or to 2.

Observation 4.3

So the phrase the Fisher information of this distribution carries no meaning as it stands. What carries meaning is the Fisher information of this family at this point.

Observation 4.4

Paper 133 showed that the uniqueness of the Fisher metric is a classical statement, and Paper 137 that the second derivative of KL is a metric. This paper enters neither. It takes only the single point that before any question of a metric, the number is undetermined until a family is given.

5. Kolmogorov: nothing need be supplied, and it cannot be computed

Definition: for a finite string x , $K(x)$ is the length of the shortest program that outputs x . No distribution, no state, no family. It is fixed by the string alone (the choice of description language costs an additive constant).

Check 5.1: take the 40000 decimal digits of π . A self-contained digit generator was written, using no library (an unbounded spigot, Gibbons 2006), and its length measured: **257** bytes. Its first 3000 digits were produced and checked to agree exactly.

Check 5.2: those digits, and pseudorandom digits of the same length, are put through three general purpose compressors.

compressor	π (bytes)	random (bytes)	difference	as a fraction
gzip	19883	19924	+41	+0.21%
bz2	17271	17271	0	0.00%
lzma	17596	17564	-32	-0.18%

Observation 5.3

For bz2 the difference is exactly 0 bytes. None of the three separates the digits of π from random digits.

Observation 5.4

Yet the true article is 257 bytes. A naive packing of 40000 decimal digits floors at 16610 bytes and the best compressor reaches 17271: 67 times the program.

Observation 5.5

Paper 223 already wrote this asymmetry in general form. Exhibiting one short program gives an upper bound; a lower bound does not come that way: K is uncomputable (Chaitin). What this paper adds is not that fact but where it sits in the roll call (Section 7).

6. The core: opposite verdicts on one and the same object

The setting: fix one object, the 40000 decimal digits of π , and apply both H and K to it.

Check 6.1: count the digits.

digit	0	1	2	3	4	5	6	7	8	9
count	3989	4060	3892	3972	4014	4040	4026	3977	4032	3998

The chi-squared is 5.0495 on nine degrees of freedom. This test does not see a departure from uniformity.

Check 6.2: place both quantities in the same unit, bits per digit.

quantity	value (bits per digit)	against its own range
Shannon H , empirical	3.321837	all but the maximum $\log_2 10 = 3.321928$
Kolmogorov K , measured upper bound	≤ 0.051400	257 bytes $\times 8/40000$, near the floor

Observation 6.3

The gap is a factor of 65. On one string Shannon says all but maximal and Kolmogorov says all but minimal.

Observation 6.4

Neither is wrong. H is a quantity about the distribution fitted, and the digit distribution really is close to uniform. K is a quantity about the object itself, and the string really is short to write. They measure different things, so opposite answers are no contradiction.

Observation 6.5

Hence the sentence this has high information carries no content until it says which. Of the digits of π it is both true and false.

Observation 6.6

It is not claimed that the digits of π are normal. That is an open problem. Check 6.1 says only that this test sees no departure from uniformity, and that H is near its maximum is itself a statement inside the limits of the test.

7. The roll call: what each cannot be defined without

quantity	cannot be defined without	a quantity about	computable
Shannon H	a distribution	the distribution fitted	yes
von Neumann S	a state (density matrix)	the state	yes
Fisher I	a parametrised family	a position within the family	yes
Kolmogorov K	nothing at all	the object itself	no

Observation 7.1

The last two columns line up exactly. The three that demand something can be computed, and the one that demands nothing cannot.

Observation 7.2

A reason can be given. Having something to supply means the quantity is a function of something finite that was handed over: of a distribution, a state, a family, and so it can be computed. K has nothing to supply because K ranges over all programs and takes a minimum, and that ranging is what halts the computation. Demanding nothing and being uncomputable are one and the same fact.

Observation 7.3

Four points are not a law. This is an observation on an alignment; it does not prove that a quantity demanding nothing must be uncomputable. What can be said is that across four the alignment did not break.

Observation 7.4

Paper 163 showed that a good code exists and here is a good code are different sentences. The asymmetry of Section 5, an upper bound but no lower bound, is a relative. As proving existence differs from exhibiting a construction, exhibiting an upper bound differs from proving a lower one, and for K the latter is closed in principle.

8. Where the roll call bites

Observation 8.1

Reading any claim that contains the word information, one thing is checked. What has been supplied. A distribution means Shannon, a state means von Neumann, a family means Fisher, and nothing at all means Kolmogorov. If a quantity is used that demands what was not supplied, there is a gap there.

Observation 8.2

The dangerous move is fitting the empirical distribution and then speaking about the object itself. Section 6 is the worked example: the empirical distribution of the digits of π is near uniform and H is near its maximum. Continuing from there to this string is random leaves the jurisdiction of H .

Observation 8.3

Paper 248 showed that the demon's cost was placed in five places. Of those five, Szilard and Brillouin put the cost on measurement and Landauer and Bennett on erasure: both are within Shannon's jurisdiction, both speaking with a distribution in hand. This roll call shows that the whole of that debate took place inside one of the four.

9. Roll call and closing

1. Shannon needs a distribution (Check 2.1): **1.253874709 / 1.254310170 / 2.000000000** on one sequence.
2. Von Neumann needs a state (Check 3.1): **$S = 0$ and $S = 1$** on the same z outcomes.
3. Fisher needs a family (Check 4.1): **$I = 1$ and $I = 2$** on the same $N(0, 1)$.
4. Kolmogorov needs nothing (Check 5.1): **a program of 257 bytes**.
5. And cannot be computed (Observation 5.5): **an upper bound yes, a lower bound no**.
6. Opposite verdicts on one object (Check 6.2): **3.321837** against **≤ 0.051400** , a factor of 65.
7. The alignment is exact (Observation 7.1): **the three that demand something can be computed; the one that demands nothing cannot**.

Closing: all four are called information, and none of them knows what the others measure. Applied to one object, one says all but maximal and another says all but minimal. Both are right: their jurisdictions differ. One thing separates them, what each cannot be defined without. And the one with nothing to supply is the one that can be defined and cannot be computed. Demanding nothing, and never yielding an answer, were the two faces of a single property.

Revision Record

Second edition (September 2026): no claim, theorem or check has been altered from the first edition. In the English references, the lines citing earlier papers of this corpus still carried their Japanese titles; **they now carry the English titles**.

Nothing of substance has changed, and nothing has been withdrawn.

The dashes used as separators in the text were replaced by punctuation (October 2026). No claim, number or reference was changed.

References

1. Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27, 379-423 and 623-656.
2. Cover, T. M., Thomas, J. A. (2006). *Elements of Information Theory*, 2nd edition. Hoboken: Wiley.
3. Kolmogorov, A. N. (1965). Three Approaches to the Quantitative Definition of Information. *Problems of Information Transmission*, 1(1), 1-7.
4. Chaitin, G. J. (1975). A Theory of Program Size Formally Identical to Information Theory. *Journal of the ACM*, 22(3), 329-340.
5. Li, M., Vitanyi, P. (2019). *An Introduction to Kolmogorov Complexity and Its Applications*, 4th edition. Cham: Springer.
6. Nielsen, M. A., Chuang, I. L. (2010). *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.
7. Amari, S., Nagaoka, H. (2000). *Methods of Information Geometry*. Providence: American Mathematical Society.
8. Gibbons, J. (2006). Unbounded Spigot Algorithms for the Digits of Pi. *American Mathematical Monthly*, 113(4).
9. Yamagishi, Y. (2026). *Two Entropies* [Paper 88]. Zenodo. 10.5281/zenodo.21747809
10. Yamagishi, Y. (2026). *The Fisher Metric Is Unique Only on the Classical Side* [Paper 133]. Zenodo. 10.5281/zenodo.22080550
11. Yamagishi, Y. (2026). *The KL Divergence Is Not a Distance, but Its Second Derivative Is a Metric* [Paper 137]. Zenodo. 10.5281/zenodo.22081369
12. Yamagishi, Y. (2026). *The Whole Can Be Better Known Than Its Parts* [Paper 162]. Zenodo. 10.5281/zenodo.22092568
13. Yamagishi, Y. (2026). *“A Good Code Exists” and “Here Is a Good Code” Are Different Sentences* [Paper 163]. Zenodo. 10.5281/zenodo.22092677
14. Yamagishi, Y. (2026). *Probability Has Four Addresses* [Paper 164]. Zenodo. 10.5281/zenodo.22092815
15. Yamagishi, Y. (2026). *There Are Four Independences, and They Say Nothing to One Another* [Paper 220]. Zenodo. 10.5281/zenodo.22119838
16. Yamagishi, Y. (2026). *It Passes Every Test, and the Points Lie on Fifteen Planes* [Paper 223]. Zenodo. 10.5281/zenodo.22120728
17. Yamagishi, Y. (2026). *Whose Cost Is the Demon’s? Five People Put It in Five Places* [Paper 248]. Zenodo. 10.5281/zenodo.22160169

— *From thinking to syncing.* —

Yuuki YAMAGISHI (山岸 勇輝)

© 2026 Yamagishi Complexity Torus Lab, Japan · Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)