Objectivity, Information, and Maxwell's Demon

Steven Weinstein†‡

This paper examines some common measures of complexity, structure, and information, with an eye toward understanding the extent to which complexity or information-content may be regarded as objective properties of individual objects. A form of contextual objectivity is proposed which renders the measures objective, and which largely resolves the puzzle of Maxwell's Demon.

> *An* n *number of possible languages use the same vocabulary; in some of them, the symbol* library *allows the correct definition a* ubiquitous and lasting system of hexagonal galleries, *but* library *is* bread *or* pyramid *or anything else, and these seven words which define it have another value. You who read me, are you sure of understanding my language?*
>
> –*The Library of Babel* (Borges 1964, 57–58)

1. Introduction. In *The Library of Babel,* the Argentinian writer Jorge Luis Borges portrays a vast library consisting of filled with books having an identical format of four-hundred ten pages, forty lines per page, and eighty characters per line.

> The Library is total and . . . its shelves register all the possible combinations of the twenty-odd orthographical symbols (a number which, though extremely vast, is not infinite): Everything: the minutely detailed history of the future, the archangels' autobiographies, the faithful catalogs of the Library, thousands and thousands of false catalogs, the demonstration of the fallacy of those catalogs, the demonstration of the fallacy of the true catalog, the Gnostic gospel of Basilides, the commentary on that gospel, the commentary on the commen-

†To contact the author, write to Department of Philosophy, Dartmouth College, 6035 Thornton, Hanover, NH 03755; e-mail: steven.weinstein@dartmouth.edu.

‡Thanks to Jos Uffink and Janneke van Lith–van Dis for helpful discussions.

Philosophy of Science, 70 (December 2003) pp. 1245–1255. 0031-8248/2003/7005-0032$10.00
Copyright 2003 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/377404 Published online by Cambridge University Press

20523	21584	93712	83654	89761	90154	96345	37539	32556	74254
70603	97122	44978	78028	08943	13778	11080	34271	68276	85372
48410	94516	15427	75323	71685	70774	50342	33771	03678	42321
69788	41758	55004	30992	17402	63523	42328	87171	24751	15084
33884	83655	88345	69602	52606	57886	18034	03381	75796	35901
77480	28683	68324	66035	07223	14926	16128	13645	90370	31949
11057	98849	29499	21565	30786	83292	92392	37104	36899	49906
79368	43710	80365	88735	75275	21664	57965	19002	00301	12658
94385	01717	96191	50404	80166	93965	24688	27839	10812	31715
92127	42588	93307	80834	11317	26583	25769	98227	14884	58462
29148	68662	26872	72927	79021	51622	29521	33355	45701	45996
33782	93424	16530	96086	17329	74020	11501	46660	05583	22277
77653	55430	84644	00448	86828	58855	67451	95264	67386	82424
52611	60012	88620	72894	94716	22262	99813	69592	63464	33163
91857	47904	22209	78590	68615	52952	31441	41313	18550	72685

Figure 1. Excerpt from *A Million Random Digits* (Rand Corporation 1955, 17).

> tary on that gospel, the true story of your death, the translation of every book in all languages, the interpolations of every book in all books. (Borges 1964, 52)

The very idea of such a collection is mind-boggling. The library clearly contains great wisdom, but it is hidden amongst sophistry, lies, and above all, gibberish. Indeed, most of the volumes will contain vast stretches of meaningless prose; meaningless, anyway, in any given language. One might wonder what portion of the $\sim 2 \times 10^{1,800,000}$ books (assuming 25 orthographic symbols) are *truly random.*[1] Which begs the question, what does it mean to say that a sequence of letters is random?

Were Borges's library constructed from an alphabet of numerals, rather than letters, a typical page from a typical book might well look something like Figure 1 (above). This is an excerpt from *A Million Random Digits,* by The Rand Corporation (1955). Is there really a fact of the matter as to whether a given sequence is random? Is randomness objective? More generally, can we quantify the *complexity* of a string of symbols, or a physical object?

2. Shannon Information. The information theory of Claude Shannon (1948) provides a method for quantifying the amount of information produced by an information source. For a source emitting strings made up of an alphabet of n discrete symbols, the associated Shannon information is defined to be

$$H = -\sum_{i=1}^{n} p_i \log p_i \tag{1}$$

1. Note that there are "only" around 10^{79} atoms in the known universe!

https://doi.org/10.1086/377404 Published online by Cambridge University Press

where p_i is the probability of occurrence of the i'th symbol, and where the base of the logarithm is regarded as an arbitrary choice of units.[2] Sometimes called the entropy (because the expression for H is formally identical to the expression for the Gibbs entropy of statistical mechanics), this quantity is rightly said to represent the 'surprisal' of the source. That is, it tells us how much new information we obtain when we receive an additional character. Thus if our alphabet has two symbols, *X* and *Y,* and the source is such that *X* and *Y* are equally probable, we find that the information associated with the source is $H = -(\frac{1}{2}(\log_2 \frac{1}{2}) + \frac{1}{2}(\log_2 \frac{1}{2})) = 1$. On the other hand, if the probability of *X* is $\frac{3}{4}$ and the probability of *Y* is $\frac{1}{4}$, then we find that $H = -(\frac{3}{4}(\log_2 \frac{3}{4}) + \frac{1}{4}(\log_2 \frac{1}{4})) \approx 0.81$. The fact that the information content of the latter source is lower corresponds to the fact that there is less uncertainty in the value of the next symbol—it is more likely to be *X* than *Y.* The information is of course minimized if the probability is maximally concentrated on a single symbol; it is maximized if all symbols are equally likely. To the extent that one is inclined to regard a source with typical string *XXXXX* as 'more ordered' than a source in which a typical string has an equal number of *X*s and *Y*s, low information content can be said to correspond to a high degree of structure.

As noted, however, the Shannon information H quantifies the information associated with a *source,* rather than with any particular string emitted by that source. Thus there is no way to assign a distinctive information content to a particular string, such as *XXXXX.* Information content is a property of sources, or more abstractly, of statistical ensembles which represent the probabilistically weighted possible outputs of the source. (The same goes for the statistical-mechanical entropy of Gibbs.) Thus we can assign a Shannon information content to the Library of Babel, but we can't assign a distinctive information content to each of the individual books.

3. Algorithmic Complexity and Objectivity. The algorithmic complexity theory of Kolmogorov (1965), Solomonoff (1964), and Chaitin (1966) is an attempt to characterize the complexity of individual strings. In particular, it is claimed by some of its advocates and adherents to provide an objective characterization of randomness, where randomness is identified with maximal complexity. For example, Cover and Thomas (1991) enthuse, "This definition of complexity turns out to be universal, that is, computer independent, and is of fundamental importance" (3). And we find Dennett (1991) following suit, claiming that, "A pattern exists in some data—is real—if *there is* a description of the data that is more efficient than the bit map, whether or not anyone can concoct it" (34).

2. Here we are assuming that the probabilities are independent; that is, the probability of a given symbol does not depend on what symbol or symbols preceded it.

https://doi.org/10.1086/377404 Published online by Cambridge University Press

Dennett's use of "real" here, and his emphasis on "there is," suggests that he, too, has taken on board the idea that algorithmic complexity *simpliciter* is an objective property. We shall see that this is not the case. The idea behind algorithmic complexity is simple: the complexity of a string of symbols is the length of the shortest program on a universal Turing machine (essentially a general-purpose computer) that will generate the string. Thus a string of 10^6 ones will generally have a very low algorithmic complexity insofar as one can generate it by a program which says, in essence, "print 10^6 ones and then stop." On the other hand, a string $X_1X_2 \ldots X_n$ is regarded as random (maximally complex) if the only way to generate it is with a program of the form "print '$X_1X_2 \ldots X_n$' and stop." In such a case, the length of the program is (for long strings) roughly the length of the string itself.

But this is all a bit quick, for the minimum length of the program required to generate a string depends on *which* universal Turing machine is available. In fact, for any finite string at all, one can always design a machine that will generate the string with a program length of 1! This is of course done by effectively building the description of the string into the machine table (the hardware). Thus the complexity of a particular string is only unambiguously defined *with respect to* a particular machine.

Now it is often pointed out in discussions of this limitation that the complexity of a given string S on one machine U_1 can vary by no more than a constant from the complexity on another machine U_2. Formally,

$$C_{U_2}(S) \leq C_{U_1} + K(U_1, U_2) \tag{2}$$

where $K(U_1,U_2)$ is a constant that encodes the length of the shortest program that will emulate U_1 on U_2 (so that one may run the U_1 program on U_2). If this were an equality, rather than an inequality, we would be in business, because this would mean that both machines would rank all strings of a given length the same in terms of their algorithmic complexity, differing only by the overall constant K. Indeed, this is how the expression above is sometimes (mis)understood. For example, Caves (1990) writes, "Algorithmic information is defined up to an additive constant, which depends on the choice of universal computer (92)." Were the above inequality an *equality,* this would be correct. But it isn't, and it isn't.

Sometimes one sees a similar but more qualified claim along these lines:

> The constant [K] in the theorem may be very large. For example, [U_1] may be a large computer with a large number of functions built into the system. The computer [U_2] can be a simple microprocessor. The simulation program will contain the details of the implementation of all these functions, in fact, all the software available on the large computer. The crucial point is that the length of this simulation

https://doi.org/10.1086/377404 Published online by Cambridge University Press

> program is independent of the length of x, the string to be compressed. For sufficiently long x, the length of this simulation program can be neglected, and we can discuss Kolmogorov complexity without talking about the constants. (Cover and Thomas 1991, 148–149)

This is almost, but not quite, correct. Given two different machines, *most* suitably long strings will not differ significantly in their complexity, though this is only because most strings are maximally complex.[3] But this is not true of *all* suitably long strings, for one might have an extremely long string which has a complexity of 1 on machine U_1 and a complexity of $1+K$ on U_2. Furthermore, there will always be an infinite number of *other* machines which will assign an absolutely minimal complexity to *any* given string, no matter what the length.[4]

The purportedly objective characterization of the complexity/randomness of a single string of symbols given by algorithmic information theory turns out to be objective only insofar as one specifies a Turing machine. Algorithmic complexity *simpliciter* is no more an objective property of a string than velocity is an objective property of a physical object. But "objective" has a variety of meanings (Lloyd 1995) and so it behooves me to spell out what I have in mind.

When I call C_{U_1}, the algorithmic complexity relative to a particular machine U_1, *objective,* I mean that it is unambiguous. It is thus objective insofar as two reasonable people could not differ on its value. Without specifying the machine, the context, a Mac user and a PC user might quite reasonably differ! What starts out as an ambiguous property, whose value depends on a certain context (the particular machine) is rendered objective by explicitly including the context and thereby removing the ambiguity.

The concept of algorithmic complexity is similar to the concept of velocity. Two people may reasonably differ on the velocity of an object, depending on their state of motion, or on what they presumed to be the appropriate reference frame. Velocity is thus not an objective property of an object, insofar as its ambiguity (velocity relative to what?) allows reasonable people to differ over its value. But *relative* velocity, the velocity of an object relative to some specific frame of reference, is of course objective, because the ambiguity has been removed, the context specified.

4. Thermodynamic Entropy. The concept of entropy arose in the context of thermodynamics, which is the study of physical systems having temperatures (no surprise here!), and thus well-defined equilibrium

3. See theorem 7.2.4 in Cover and Thomas (1991).

4. To construct such a machine, simply modify an existing machine by adding an internal subroutine (a consecutively invoked sequence of internal states) which erases the tape and writes out the desired sequence if the initial string is a 1 followed by a blank.

https://doi.org/10.1086/377404 Published online by Cambridge University Press

states. The state of a system in thermodynamics is given by specifying the temperature T and one or more state variables $X, Y, \ldots$ which are understood to be stable in time, and thus "in equilibrium." These variables typically come in pairs, and satisfy one or more equations of state (hence the name 'state variable'). Thus for a so-called 'ideal gas,' we have the equation of state $PV = NRT$, where P and V are pressure and volume, N is the number of moles of gas, R is Boltzmann's constant, and T, of course, is the temperature. Neither N nor R are state variables (R is not even a variable), but P and V are, as they correspond to quantities which can be manipulated.

The change in entropy of a thermodynamic system is defined by $\Delta S := \int dQ/T$ (where the integral is taken over a reversible path in the space of state variables). The entropy itself is expressible (up to a constant) as a function of the state variables. For example the entropy S of N moles of an ideal gas is given by

$$S = N \ \log(VT^{\frac{3}{2}}) - N \ \log \ N. \tag{3}$$

Thus if one holds the volume of N moles fixed and raises the temperature, the entropy increases. Or, if one maintains the temperature while allowing the volume to increase, the entropy increases.

Changes in entropy are subject to the second law of thermodynamics. One version of this law reads,[5]

The entropy S of an isolated system cannot decrease.

"Isolated system" here means a system which receives no energy from the environment. According to the second law, the only way to *reduce* the entropy of a thermodynamic system is to do work on it (for example, by pushing on a piston) or to drain heat from it (cool it).

Now suppose the system is used to do work. For example, suppose our system is compressed air in a cylinder, which is allowed to push on a piston and thereby lift a weight. The work done corresponds to the increase in potential energy of the weight. It follows from the first and second laws that the work done W is related to the change in entropy ΔS by

$$W = T\Delta S. \tag{4}$$

Thus work done by an otherwise isolated system means an increase in entropy. Conversely, an isolated system whose entropy is constant can do no work.

5. See Uffink (2001) for an excellent discussion of the surprisingly wide variety of formulations of the Second Law.

https://doi.org/10.1086/377404 Published online by Cambridge University Press

While changes in entropy tell us something about how much work we can extract from a system, entropy itself is often said to characterize the amount of disorder in a system. Thus a system whose entropy is maximal is often understood to be maximally disordered or randomized. Having reflected on the objectivity and contextuality of algorithmic complexity, the question naturally arises as to whether these are features of entropy as well.

Here is a reason to think not. Since the possible changes in entropy quantify our ability to extract work from a system, there should be an objective fact of the matter as to the difference in entropy between two states of the system. In other words, entropy should be objective, in the sense of unambiguous, up to an additive constant. The idea here is that whether or not one can extract work from a system by, for example, removing a partition, must be an objective property of the system.

Consider a system consisting of a box filled with gas with a partition down the middle.[6] The temperature and pressure are identical on both sides of the partition, and the entropy should, as far as we know, be a function $S(T, V)$. Now remove the partition. Is there a change in the entropy? It would seem not, and the equation above for the entropy of an ideal gas is in line with this. Removing the partition makes no *real* difference, and in particular there is no loss of ability to do work when the two sides mix. There was never any such ability to start with, as the initial and final states are both states of the same, maximal entropy.

Now suppose someone else is watching, and that this someone else has special glasses. To this person, the initial state is one in which the gas on one side of the partition looks red, and that on the other side looks green. This person will use a different entropy function $S(T, V, C)$ which includes "color," since this person can clearly see that there are two different gasses. This person will certainly say that the entropy *increases* after the partition is removed.

At this point, it seems ambiguous as to whether the system has undergone a change in entropy following the removal of the partition. But one might think that the individual with more complete knowledge, the one aware of the red-green distinction, has the *true* entropy. Further evidence for this might be thought to come from the fact that someone who is aware of the distinction could use this information to do work by inserting a piston made of material which is transparent to the green gas and opaque to the red gas. Such a piston would be driven by the pressure of the red gas, and this could be harnessed to lift a weight (for example). So it would seem that the true entropy is the entropy which takes in all of the physical quantities governing the behavior of the gas.

6. In what follows I draw heavily on the exposition of Jaynes (1992).

https://doi.org/10.1086/377404 Published online by Cambridge University Press

However compelling this idea may be, it is incorrect. There is no "true" entropy. Rather, there are, as Grad (1961) and Jaynes (1992) have argued, many different entropies.

Consider N moles of some liquid—it might be characterized simply by its temperature, pressure, and volume. But suppose that the molecules of the liquid have an electric dipole moment, as is the case with nitrobenzene.[7] In this case, one can change the polarization M of the liquid by manipulating an electric field E (this is how a Kerr cell works). Whereas before we had one equation of state $f(T, P, V) = 0$, there will now be two equations of state $f(T, P, V, E) = 0$ and $g(T, M, E, P) = 0$. The additional state variables mean that entropy will now be a function $S(T, V, M)$, rather than $S(T, V)$, as before. Thus a sample of nitrobenzene which appears to be at maximum entropy with respect to V may or may not be at maximum entropy with respect to V and M. For instance, if there is no electric field, then the electrons will be unpolarized. If there is a (uniform) electric field, then they will tend to line up in the direction of the field. This configuration will have a lower entropy than the unpolarized liquid.

At this point, one might rightly point out that this is just the red-green example in a more realistic guise. One might further insist that *this* description, the one which invokes the polarization, is the true entropy. But the problem is that one can introduce further state variables as well, not only by discovering new sorts of properties, but by *manipulating the same properties in finer detail* (for example, by applying different electric fields to different regions of the sample). This process is limited only by the fact that, at extremely small scales, the system ceases to display equilibrium behavior. We do not expect the polarization of a single molecule to have a stable value, as that molecule is vibrating due to thermal fluctuations.

In conclusion, then, the entropy depends not only on the substance, but on the description of the substance, and there is no most-fundamental description. Which description is appropriate depends on which variables one is able to control, or interested in controlling. As is the case with algorithmic complexity and relative velocity, entropy can be contextualized, and entropy *relative to a description* can be considered an objective property of a physical system. But there is no most-fundamental description, just as there is no most-fundamental Turing machine or inertial reference frame.

5. Maxwell's Demon and the Objectivity of the Second Law. Let us now return to the second law. Given that entropy is only objective when contextualized, it would seem that the second law is in danger of losing

7. This example derives from Jaynes (1965).

https://doi.org/10.1086/377404 Published online by Cambridge University Press

its objectivity unless it is formulated in a way that incorporates this feature. This suggests the following reformulation:

> The entropy $S(T, X, Y, \ldots)$ of an isolated system cannot be reduced by manipulation of the state variables $(X, Y, \ldots)$.

This version of the second law (akin to that of Jaynes (1992)) explicitly acknowledges the contextual aspect of entropy, its dependence on the thermodynamic description, and furthermore gives this physical content by making the connection between the description and the possible physical manipulations of the system.

Perhaps the most important thing to note about this formulation of the second law is that it is a formulation which is *immune* to the threat of Maxwell's Demon (see Figure 2). Maxwell's demon is an imaginary creature who is able to take a box of gas at equilibrium and operate a trap door in a partition in the box, permitting (in one of the more popular versions) the fast molecules to go into one side and the slow ones to go into the other. This results in a decrease in entropy, in the sense that the pressure in one side goes up, and in fact one can use this entropy difference to do work. Coupled to a heat bath (an infinite source of heat to maintain the box of gas at constant temperature), this cycle can go on forever, creating what is called a 'perpetual motion machine of the second kind.'

Much ink has been spilled attempting to "exorcise" the demon (see Leff and Rex 1990), as it has seemed to many (though apparently *not* to

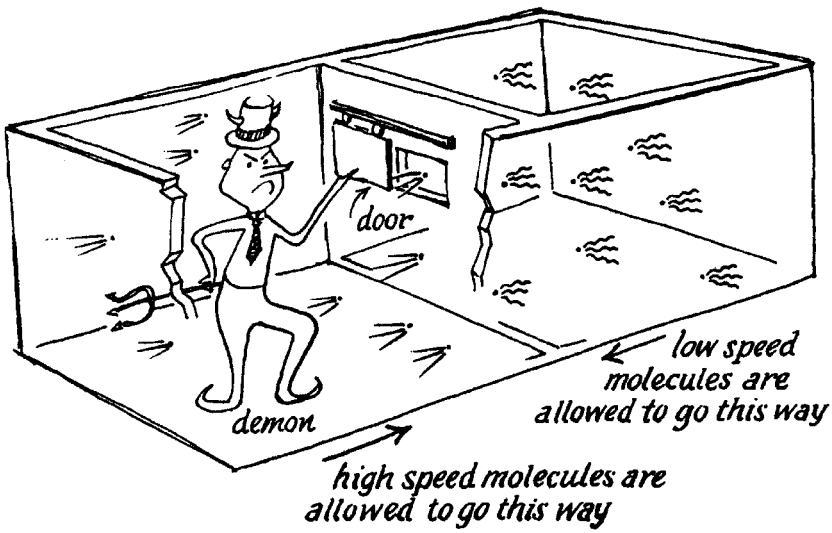


Figure 2. Maxwell's Demon, from Angrist and Hepler 1967; used by permission.

https://doi.org/10.1086/377404 Published online by Cambridge University Press

Maxwell) to be a threat to the universal validity of the second law.[8] Most exorcisms work on the assumption that the demon must do work in order to reduce the entropy, thereby ensuring no net gain of work, and thereby precluding a perpetual motion machine. These exorcisms are quite problematic (Earman and Norton 1998, 1999), but from the perspective of this paper, they are completely beside the point. If the second law is formulated properly, *the demon doesn't violate the second law.* The demon is operating on a scale at which the second law does not apply, for he is manipulating the system on a microscopic level. The second law, properly understood, is a law of *thermodynamics,* and it should come as no surprise that its domain extends only to thermodynamic variables. The demon interacts with the gas on a scale at which the properties it is manipulating are constantly fluctuating.

To those who resist this conclusion, I would simply point out that we are not surprised that manipulations described by quantum mechanics give rise to phenomena which could not be achieved by classical manipulations. In fact, this is the primary *appeal* of the field of quantum computation: manipulating systems on a sufficiently small scale may allow us to compute things in polynomial time that could only be computed in exponential time on a machine whose behavior is governed by classical mechanics (Nielsen and Chuang 2000). To be sure, there remains a compelling question at the end of the day, which is, "Can one build a perpetual motion machine of the second kind?" However, if one could, it would not violate a properly formulated second law.

6. Conclusion. We have taken a brief tour through some measures of complexity, structure, and information. I have shown how two important measures, algorithmic complexity and thermodynamic entropy, can be thought of as objective properties of, respectively, strings of symbols and thermodynamic systems, only insofar as they are regarded as contextual. In one case, we contextualize to a Turing machine, in the other case, to a description. Once this is taken into account, the apparent paradox of Maxwell's demon disappears.

REFERENCES

Angrist, Stanley W., and Loren G. Hepler (1967), *Order and Chaos.* New York: Basic Books.

Borges, Jorge Luis (1964), "The Library of Babel", in Donald A. Yates and James E. Irby (eds.), *Labyrinths: Selected Stories and Other Writings.* New York: New Directions, pp. 51–58.

8. The demon seems to have been proposed by Maxwell (1871) to make a point closely related to the point I am making in this paper. However, it was subsequently thought by many to pose a threat to the Second Law.

https://doi.org/10.1086/377404 Published online by Cambridge University Press

Caves, Carlton (1990), "Entropy and Information: How Much Information is Needed to Assign a Probability", in Wojciech Zurek (ed.), *Complexity, Entropy and the Physics of Information, SFI Studies in the Science of Complexity, Vol. VIII,* Reading, PA: Addison-Wesley, pp. 91–113.
Chaitin, Gregory (1966), "On the Length of Programs for Computing Binary Sequences", *Journal of the ACM* 13: 547–569.
Cover, Thomas, and Joy Thomas (1991), *Elements of Information Theory.* New York: John Wiley and Sons.
Dennett, Daniel (1991), "Real Patterns", *Journal of Philosophy* 88: 27–51.
Earman, John, and John Norton (1998), "Exorcist XIV: The Wrath of Maxwell's Demon, part I—from Maxwell to Szilard", *Studies in the History and Philosophy of Modern Physics* 29: 435–471.
——— (1999), "Exorcist XIV: The Wrath of Maxwell's Demon, part II—from Szilard to Landauer and Beyond", *Studies in the History and Philosophy of Modern Physics* 30: 1–40.
Grad, Harold (1961), "The Many Faces of Entropy", *Communications on Pure and Applied Mathematics* 14: 323–354.
Jaynes, Edwin T. (1965), *Thermodynamics,* unpublished, available at http://bayes.wustl.edu/etj/thermo.html.
——— (1992), "The Gibbs paradox", in *Maximum-Entropy and Bayesian Methods.* Dordrecht: Kluwer.
Kolmogorov, Andrei (1965), "Three Approaches to the Quantitative Definition of Information", *Problems of Information Transmission* 1: 4–7.
Leff, Harvey, and Andrew Rex (1990), *Maxwell's Demon: Entropy, Information, Computing.* Princeton, NJ: Princeton University Press.
Lloyd, Elisabeth (1995), "Objectivity and the Double Standard for Feminist Epistemologies", *Synthese* 194: 351–381.
Maxwell, James Clerk (1871), *Theory of Heat.* London: Longmans, Green, and Co.
Nielsen, Michael A., and Isaac L. Chuang (2000), *Quantum Computation and Quantum Information.* Cambridge: Cambridge University Press.
Rand Corporation (1955), *A Million Random Digits with 100,000 Normal Deviates.* Glencoe, IL: The Free Press.
Shannon, Claude (1948), "A Mathematical Theory of Communication", *Bell System Technical Journal* 27: 379–423, 623–656.
Solomonoff, Ray (1964), "A Formal Theory of Inductive Inference", *Information and Control* 7: 1–22, 224–254.
Uffink, Jos (2001), "Bluff Your Way in the Second Law of Thermodynamics", *Studies in the History and Philosophy of Modern Physics* 32: 305–394.

https://doi.org/10.1086/377404 Published online by Cambridge University Press