

# COMPRESSIBILITY AND THE ALGORITHMIC THEORY OF LAWS

BILLY WHEELER

*Department of Philosophy (Zhuhai), Sun Yat-Sen University, CHINA*

`billy.wheeler@cantab.net`

---

**Abstract.** The algorithmic theory of laws claims that the laws of nature are the algorithms in the best possible compression of all empirical data. This position assumes that the universe is compressible and that data received from observing it is easily reproducible using a simple set of rules. However, there are three sources of evidence that suggest that the universe as a whole is incompressible. The first comes from the practice of science. The other two come from the nature of the universe itself: the presence of chaotic behavior and the nature of quantum systems also suggests that the universe is incompressible. This paper evaluates these sources and argues that none provides a convincing case to reject the algorithmic theory of laws.

**Keywords:** Laws of nature • the algorithmic theory of laws • the best system account • data compression • algorithmic information theory • incompressibility

---

RECEIVED: 15/03/2019

REVISED: 08/04/2019

ACCEPTED: 28/05/2019

## 1. The Algorithmic Theory of Laws

It can hardly be denied that one virtue of a scientific theory is its ability to present simple laws of nature that hold for as many objects and events as possible. But why should this be? Is there an intrinsic connection between simplicity in the formulation of our scientific laws and the laws of nature themselves? In other words, is simplicity just a desirable *epistemic* factor in scientific theory discovery and appraisal? Or does it reflect some deeper truth about the *metaphysical* nature of laws and their role in the universe?

One well-known attempt to connect simplicity in a fundamental way to the nature of laws is David Lewis's (1973, 1983, 1986 and 1994) "best system account" (BSA). According to Lewis, laws of nature are the axioms in the best possible deductive systematization of all empirical truths. Deductive systems can be very informative and they can be very simple. Lewis argued that these virtues tended to conflict and that the "best system" is one that best balances these virtues. These axioms exist as abstract objects and are mind-independent: there is a best-balanced system regardless of whether anybody has yet to discover it.



Lewis's BSA faces a number of formidable problems.<sup>1</sup> I won't rehearse them here. Instead, I want to draw attention to a different theory of laws, which like the BSA, connects lawhood to simplicity in a fundamental way. It is much less discussed than Lewis's theory, despite having the potential to overcome a number of its shortcomings. It is also, I believe, a much more faithful description of the role laws play in scientific practice and how they are arrived at. The account I am referring to is the "algorithmic theory of laws".<sup>2</sup> This theory differs significantly from Lewis's account: whereas for Lewis laws are *true statements*, according to the algorithmic theory—as the name suggests—laws are *algorithms*. They are not algorithms for designing programs to run on computers (although they can be used in this way), but algorithms for compactly and simply describing the world.

The basic idea behind the algorithmic theory has its origins in Ernst Mach's position on the role of theories and laws in science:

The communication of scientific knowledge always involves description, that is, a mimetic reproduction of facts in thought, the object of which is to replace and save the trouble of new experience. Again, to save the labor of instruction and of acquisition, concise, abridged description is sought. This is really all the natural laws are. Knowing the value of the acceleration of gravity, and Galileo's laws of descent, we possess simple and compendious directions for reproducing in thought all possible motions of falling bodies. (1894, p. 193)

With the laws of nature, we are able to describe a large amount of information about physical systems and their properties. Provided we have available some knowledge of initial conditions, we can "unpack" further information about a system by applying these laws. There is a very strong analogy here between this practice and data compression in computer science. A position similar to Mach's is given by Paul Davies, although this time Davies explicates the idea specifically in the language of computing:

The existence of regularities may be expressed by saying that the world is algorithmically compressible. Given some data set, the job of the scientist is to find a suitable compression, which expresses the causal linkages involved. For example, the positions of the planets in the solar system over some interval constitute a compressible data set, because Newton's laws may be used to link these positions at all times to the positions (and velocities) at some initial time. In this case Newton's laws supply the necessary algorithm to achieve the compression. (1995, p. 249)

To help explain Davies's idea, consider the following two strings of binary digits:

(A) 0101010101010101010101010101

(B) 011010010110101111010101000101

If we needed to store or transmit this information we could do so by simply making a note of each digit in the sequence. In this case it would be clear that we would require just as many resources to describe A as we would to describe B. However, we can describe A in a simpler way using the following rule:

(C) Print "01" 15 times

The information in C is of a different character to both A and B. It is a *command* or *rule* that tells you how to get to the information in A. If the length of C is shorter than A (and requires less resources to store and transmit), then we say that C is an *algorithmic compression* of A. According to the algorithmic theory, this is what the laws of nature are: they are algorithms in algorithmic compressions of nature. On this picture, nature is seen as a source of information. This information is encoded into basic elements of data and the job of the scientist is to find the shortest description of this data.

If all there is to the algorithmic theory is the claim that scientific laws compress empirical data, then this would not be enough to mark it out as a genuine alternative. It would be open, for example, for an advocate of the BSA, or an advocate of Armstrong's relations of nomic necessitation to claim that laws compress data whilst maintaining that—at the fundamental level—laws are axioms or second-order relations of nomic necessity. For the algorithmic theory, the fact laws can be used to compress empirical data goes beyond a mere observation about the practice of science: it is at the very core of what it is *to be* a law of nature.

The algorithmic theory relies heavily on results from a branch of mathematics called "algorithmic information theory" (AIT). This field was devised independently by Ray Solomonoff (1964) and Andrey Kolmogorov (1965) and later developed further by Gregory Chaitin (1987). AIT provides a measure of the simplicity (or conversely the "complexity") of a data structure in terms of the total amount of computational resources required to output the data on a universal Turing machine. One important result of AIT is that for any data structure *S*, such as a string of binary digits, there exists its *shortest possible description*. This is often called its "Kolmogorov Complexity":

The Kolmogorov Complexity relative to a given language is the length of the shortest program *p* plus input *y*, such that, when given input *y*, *p* computes (outputs) *x* and then halts. Thus:

$$K(x) = l(p) + l(y)$$

where  $l(p)$  denotes the length of input *p*, and  $l(y)$  denotes the length of program *y*. (Grünwald & Vitanyi 2008, p. 291)

AIT gives us confidence that for any set of data  $S$  we can say there exists, as an abstract object, a set of algorithms which constitute the best possible compression of  $S$ . According to the algorithmic theory, the laws of nature are just these algorithms when  $S$  is all the empirical data about the universe.

**Law of Nature:**  $X$  is a law of nature if, and only if,  $X$  is an algorithm in the best possible compression of nature's data.

If we like, we can take something similar to Lewis's "Humean Mosaic" as the basis from which all the data originates. The laws of nature then emerge from this basis as the algorithms in its shortest description (Wheeler, 2012, 2016 and 2017; Tomkow, 2013 and 2014). However, this view does not commit us to adopting Lewis's particular version of the atomic facts: we need not, for example, construe them as being about intrinsic natural properties attached to space-time points. The algorithmic theory is therefore not committed to Humeanism and is compatible with both Humean and non-Humean approaches to metaphysics. It is far more likely to be appealing to philosophers who prefer a sparse ontology, however, since it is able to explain the existence of laws without needing to appeal to irreducibly nomic objects such as necessity relations or capacities.

If the algorithmic theory of laws is correct, then the laws of nature exist because the world is compressible. This seems obvious: after all, nature is a regular place, and barring skeptical arguments about induction, the past is a good guide to the future. Despite this, there exist powerful arguments to suggest the opposite. Nature may appear regular, and indeed it may be regular in parts, but the empirical data we gather from it is not compressible. If this conclusion is correct, then it would pose both epistemological and metaphysical problems for the algorithmic theory of laws.

The argument for incompressibility comes chiefly from three different sources. Firstly, James McAllister (2003 & 2005) has been an active opponent of the algorithmic theory and has presented a number of arguments based largely around scientific practice which show that scientists do not, and cannot, compress empirical data with their laws. The second and third sources of the argument come from theories about nature itself. Chaos theory and quantum mechanics present pictures of nature that likewise seem to imply that empirical data will be incompressible.

If these arguments are correct, it would provide a strong reason to reject the algorithmic theory. In the rest of this paper I will consider these arguments and show why I think none of them gives us a good reason to reject the algorithmic theory of laws as it stands. Nonetheless, some of the observations about incompressibility are important in helping to further develop the algorithmic theory—especially when it comes to understanding why scientists choose to adopt laws which frequently *do fail* to compress their empirical data.

## 2. McAllister's Arguments

James McAllister has been a vocal opponent of the algorithmic theory of laws. According to him, there is no work for such laws because “empirical data sets are *algorithmically random*: they are characterized by maximal complexity, maximal entropy and zero redundancy, and are therefore maximally efficient carriers of information about the world” (2003, p. 634). To say that empirical data is algorithmically random is equivalent to saying that it cannot be compressed. As a result, either there are no laws of nature or the algorithmic theory is false.

McAllister's claim would strike many as counterintuitive, and he is well aware of this. But he claims the reason for this is that when most people reflect on the nature of empirical data they are almost always talking about *processed data*. This is data that has been deliberately modified to either remove the effects of perturbing causes, instrumental bias or measurement error (2003, p. 634). But if the laws of nature are compressors of information, then it must be what he calls *raw data* that is compressed. It is this raw data that McAllister argues is algorithmically random.

I can identify four distinct strands of argument in McAllister's writings for his claim that nature is incompressible:

- (i) *The Historical Argument*. All previous attempts in the history of science to compress empirical data have failed. Therefore, this constitutes good evidence that all such attempts will fail.
- (ii) *The Statistical Argument*. Raw empirical data pass statistical tests for “randomness”. Therefore, they are incompressible.
- (iii) *The Information Argument*. If data sets are compressible, then, according to AIT, they are low carriers of information. Since such data is very informative for scientists, so they cannot be compressed.
- (iv) *The Argument from Local Causes*. Scientists provide laws for isolated “phenomena”. In practice these phenomena are influenced by local variable causes which produce data that cannot be summarized by a simple rule.

In this section I will evaluate arguments (i)–(iii) and reserve discussion of argument (iv) until the next section. The reason for this is that argument (iv) is very similar to another argument based on the presence of “chaotic systems” in the world. It therefore makes sense to discuss these two arguments together.

### 2.1. The Historical Argument

At first glance it seems obvious that McAllister's claim is wrong. After all, we can simply produce empirical data to refute it. If we take a loaded coin and toss it a

thousand or so times, it will produce a string of data that is highly compressible. If it is loaded to always land heads, then it can be compressed to a simple algorithm:

(D) Print “H” 1000 times

Despite this, McAllister argues that most of the data that scientists are interested in—that is, data produced by “natural processes”—have frequently failed past attempts at compression:

The most revealing cases are laws formulated by scientists who take an empiricist or instrumentalist approach, showing minimal commitment to theoretical presuppositions and striving to fit a curve to a data set by all available mathematical techniques. Examples are Johannes Kepler’s first law of planetary motions of 1609, Johann Balmer’s law of the wavelengths of the emission lines of the hydrogen spectrum of 1884, and Max Planck’s law of blackbody radiation of 1900. If any scientific law can be expected to constitute a compression of a data set, these can. In fact, of course, none of these laws precisely generates either the empirical data available at the time of its formulation or data gathered subsequently. (2003, p. 638)

He gives as an example Balmer’s law for predicting the wavelengths of emission lines in Hydrogen:

$$\lambda = B \left( \frac{n^2}{n^2 - m^2} \right) = B \left( \frac{n^2}{n^2 - 2^2} \right)$$

Balmer used the data he had available to him to help formulate this law, but his data only agreed with it within one part in 40, 000. As McAllister remarks: “an impressive achievement, but it does not amount to an algorithmic compression of a data set” (2003, p. 638). Worse still, Balmer would later use this formula to correct previous data that had been gained using the method of spectroscopy. In this example then, although the data might have been used for inspiration for the law, it can hardly be said to be a compression of it. In addition to these examples, McAllister writes that he can “find no documented claim by a scientist that a specified law or theory constitutes a compression of a specified data set” (2003, p. 638)—further supporting his claim that the history of science provides little evidence in support of the algorithmic theory.

One explanation of the observations made by McAllister is that the algorithmic theory is wrong. But other explanations are possible. Firstly, an optimist might argue that the reason why the history of science has failed to find compressions is that the measuring equipment available to previous generations lacked the ability to provide accurate enough recordings. It could be argued, for example, that had Balmer’s equipment been accurate enough, he would have been able to generate data that was

compressed using his formula. Similar points might be made in defense against the claims about Kepler and Planck.

Another alternative explanation is that the reason why Kepler, Planck and Balmer did not find laws that compress is that they weren't actually looking for them. This need not necessarily count against the algorithmic theory as an account of the metaphysics of laws. As I will explain in more detail in section 3, the algorithmic theory recognizes two different kinds of laws in science: lossless laws (which provide complete compression without loss of fidelity) and lossy loss (which provide greater amounts of compression but with some loss of fidelity within the data). The reason for this is that the epistemology of laws, especially in science, is more complicated than other notions of lawhood have assumed. Taking its cue from Mach, the algorithmic theory recognizes competing virtues in scientific understanding. On the one hand, scientists aim for accuracy in prediction, but they also value simple and effective rules for making calculations. Lossy compressions provide a shortcut to balancing these aims. This does not refute the idea that there are perfectly lossless compression algorithms—ones that correspond to the laws of nature—but that sometimes the practical aims of science take precedence over the search for these.

However, the best argument against McAllister's historical argument is to actually find examples of scientific practice where lossless compression takes place. Such examples have previously been given by Twardy, Gardner and Dowe (2005). They argue that *minimum message length* (MML) compression is used in a range of scientific disciplines to compress empirical data:

Most of statistics, machine learning, econometrics, and data mining consists of finding patterns which do or could compress the data. For MML alone, dozens if not hundreds of papers apply it to empirical data in psychiatry, protein structure, archaeo-astronomy, spectral analysis, medicine, taxonomy, economics, lost person behavior, wildfire prediction, image segmentation and compression, and more [...] Each such paper is a specific counterexample to McAllister's conjecture. (2005, p. 400)

MML compression was developed by Chris Wallace and combines elements of both AIT and optimal coding (Wallace & Boulton, 1968). The principle behind it is that for any scattering of data points, it is always possible to reduce the amount of computational resources required to reproduce them. The compression has two steps. First, a curve of best fit  $y = xm$  is found that reduces the average distance between each data point, and second, the distance between the curve and each point is then recorded, typically as a whole number between  $-2$  and  $2$ . Overall, the amount of resources required to encode, store and transmit these two components is less than it would be to send each individual data point one by one.

McAllister has previously responded to the examples given by Twardy, Gardner

and Dowe. His main point of contention is whether or not MML counts as an example of *algorithmic* compression. One reason he gives is that unlike ordinary examples of algorithmic compression, MML compression requires the computer to store or transmit *just as many* data points as the original data set. For strings of finite length this might not seem such a worry; but problems emerge for the algorithmic theory if the universe—as a whole—contains an infinite amount of data:

Applying MML technique to an infinite data set yields an infinite string, which specifies the distance of each data point in the original set from the chosen pattern. MML technique therefore brings no message length reduction when it comes to communicating the information contained in an infinite data set. (2005, p. 405)

The algorithmic theory of laws, as a metaphysical theory, says that the laws of nature are the algorithms in the best possible compression of nature's data. But if nature's data is infinite in length, then clearly there are no benefits to compressing it using MML: the resulting string will also be infinite and as a result there is no reduction in message length.

McAllister's response only gets traction, of course, if it is true that the universe contains an infinite amount information. But it is not at all obvious that this is the case and there are many who reject the idea that the universe is infinite. For example, a recent study using NASA's Wilkinson Microwave Anisotropy Probe has detected that ripples in the background microwave radiation decrease in their distance. This supports a finite model of the universe as a bounded three-dimensional space (Luminet *et al.*, 2003). Clearly, the infinite nature of the cosmos has yet to be established and so McAllister's point against MML has little support.

McAllister also argues that MML does not count as *algorithmic* compression because of the way each element in the data set is repeated. Even though the amount of resources it uses is reduced because of the algorithm for  $y = xm$ , it counts instead as an example of what he calls "*enumerative compression*", a fundamentally distinct form of compression (2005, p. 406). Does this count against the algorithmic theory as a metaphysics of laws? I don't think so. Recall the definition of a law from section 1:

**Law of Nature:**  $X$  is a law of nature if, and only if,  $X$  is an algorithm in the best possible compression of nature's data.

Very few algorithms can reproduce data without the need of additional, unstructured singular data to "work with". This fact itself is reflected in scientific practice which uses fundamental constants and initial conditions. Nothing in the definition of law according to the algorithmic theory commits us to placing a limit on the length of this additional data. We might hope that it is small. That would make the laws easier to work with from a scientific point-of-view. But the algorithmic theory does



not commit us to this and so it is entirely feasible that the best compression just happens to be one with long strings of singular data. The law of nature is still identified with the algorithm in this best compression. As a result, if MML is the best method for compressing nature's data, then it will be a good method for finding the laws as well.

## 2.2. The Statistical Argument

McAllister's second argument for incompressibility comes from his claim that raw empirical data pass standard statistical tests of randomness. The three tests he discusses are (i) equal probabilities, (ii) no correlations and (iii) random walk. I will argue that empirical data only passes the random walk test of randomness. What is more, this is an unreliable indicator of compressibility because data sets that are known to be compressible often come out as "random" on this test.

According to the equal probabilities test, a string of data is random if, and only if, every symbol in that data set has an equal chance of occurring. A good example would be the data obtained from recording the outcome of tossing a fair coin. In each place in the sequence, every possible symbol in the base vocabulary ("heads" and "tails") has an equal chance of occurring.

However, as McAllister himself points out, raw empirical data are known to fail this test because of a curious phenomenon known as "Benford's Law". According to this law, for any physical measurement, the average probability of "1" occurring as the first significant digit is 0.301, whereas the average probability that "9" will occur is 0.046 (McAllister 2003, p. 639). This shows that for recordings of physical quantities, the chance of a value occurring does not have equal probability and therefore is not random. It is not known why measurements have this property. One idea (that McAllister endorses) is that this result is a side-effect of the standard units adopted by scientists. Presumably, if one accounted for this effect, say by using a different scale or set of units, then the data would appear more random.

Nevertheless, this response is not one that McAllister can adopt in defense of his claim. Recall that for McAllister, when he claims data is incompressible, he means raw empirical data that has not been "processed" to remove perturbing causes, instrumental bias or measurement error (2003, p. 634). Therefore, any data that has been "corrected" to account for standard units or measuring scales would count as processed and so would no longer be applicable. If processed data is allowed, then it is far easier to prove that scientific data is compressible. Such processed data is regularly compressed—we need only draw a line of best fit and note the equation. This will provide an effective compression algorithm of the processed data.

The second test of randomness is similar to the first in that it depends on the probability of each value in a string of data. Whereas the first takes the view that

each value or symbol has an equal chance of occurring, “no correlations” looks for the probability of a symbol occurring given its neighboring symbols. If no statistical correlation or probabilistic dependence is found, then the string is said to be random.

The problem with this proposal is that we may record empirical data in any order we like, and we regularly do so in a way that reflects known dependencies between numerical values. For example, suppose we use measuring apparatus to record values for the pressure ( $P$ ), volume ( $V$ ) and temperature ( $T$ ) of a sample of gas. After every measurement, we change the temperature by one degree, and measure again, and so on. We end up with a string of data like the following:

$$\langle P_1, V_1, T_1 \rangle, \langle P_2, V_2, T_2 \rangle, \langle P_3, V_3, T_3 \rangle, \dots \langle P_n, V_n, T_n \rangle$$

We know from experience and the Boyle–Charles gas law, that  $P$ ,  $V$  and  $T$  are highly correlated and that therefore there will be a strong statistical dependency. The value for  $P$  say is highly likely to fall within a certain narrow range, given values for  $V$  and  $T$ .

The fact that we cannot say what the value for  $P$  will be in advance to any degree of accuracy (given  $V$  and  $T$ ) is used by McAllister to reject this response. Although the value for  $P$  is likely to fall within a very narrow range of error given  $V$  and  $T$ , McAllister argues that this “narrowness” of error is a psychological illusion and merely reflects a fact about our measurement scales:

If the measured values of a parameter are expressed as 999, 1000, and 1001 units, it appears that there is a strong correlation among the measurement outcomes: they lie within 0.2 percent of one another. By shifting the zero point, however, we can express these values as  $-1$ ,  $0$  and  $+1$  units. The correlation between the measurement outcomes now appears weak. (2003, p. 640)

McAllister’s response here is therefore similar to the first. Likewise, where one places the “zero point” becomes a matter of choice. Arguably, this seems like another instance of the very data processing McAllister is against. But equally problematic for his position is the arbitrary nature of the zero point. If it is arbitrary whether or not the dependency is probabilistically significant or not, then it is arbitrary whether or not the data can be said to be random. This is not a very strong argument and therefore if anything, this test is neutral with respect to whether or not empirical data is random.

The final test McAllister considers is a “random walk”. The underlying principle of a random walk can be observed in data that has been gathered from the time-evolution of a system where at each stage there is more than one option available for the individual components. For example, a classic example of a random walk

is Brownian motion, where the position of each particle in the system over time is equivalent to making a random choice with each step. McAllister argues that many known empirical data sets (such as those from astrophysics, atmospheric temperature fluctuations, financial markets and linear polymer modelling) are indistinguishable from a random walk and therefore ought to be considered random (2003, p. 640).

What does the indistinguishability of empirical data from a random walk prove? As McAllister himself admits, this test is the weakest of the three (2003, p. 640). The main problem with this test is that it is known to apply to data that is widely accepted as algorithmically compressible. For example, the decimal expansion of pi has been shown to be indistinguishable from a random walk (Artacho 2012). Yet this data is easily compressible to a short computer program. As a result, random walk seems an unreliable test of randomness and cannot provide convincing evidence of the incompressibility of empirical data.

### 2.3. The Information Argument

According to AIT the amount of information contained within an object or string of symbols is equal to the length of the shortest program that can produce it (Grünwald & Vitanyi 2008, p. 289). If this is true and the algorithmic theory of laws is correct, then it would suggest that the universe as a whole contains “very little information”. McAllister argues that this is inconsistent with scientific practice because “scientists would gain no information by gathering additional empirical data about a phenomenon” (2003, p. 641). As scientists do perform experiments on the same phenomena, and gain new information from doing so, so the algorithmic theory of laws must be false.

There are a number of responses that can be given to McAllister’s third argument. Firstly, even if AIT and the algorithmic theory of laws are correct, this would not imply that scientists cannot gain new information by studying instances of the same empirical phenomena again. Every new occurrence or instant of a phenomenon provides new information. All the algorithmic theory says is that this information can be highly compressed if corresponding to that phenomenon there exists some compression algorithm. It does not say that the phenomenon is reducible to the algorithm or that repeat occurrences are uninformative. Secondly, human scientists are not infallible. Even if they have discovered what they take to be the best algorithm for compressing data related to some phenomenon, there is still the chance they might be wrong, or that the data is compressible using an even better algorithm. As such, the epistemology of the algorithmic theory of laws would never endorse the view that repeat experiments are uninformative. The Kolmogorov Complexity of a data set is not computable, as a consequence, there is always the possibility of improving on one’s existing best algorithms.

The main problem with McAllister's third argument is that it saddles the algorithmic theory of laws with a theory of information that it does not need to accept. Although the algorithmic theory of laws is a natural fit with AIT (and indeed was endorsed by many of its developers), the most essential aspect of AIT for the algorithmic theory is its definition of *complexity*. By providing a rigorous definition of the complexity of an object in terms of its length in some specified computing language, it allows for the comparison of different representations and gives us assurance that there exists the "best possible" description. And although it is true that some use this as an absolute measure of information, this is a different concept and need not be adopted by the algorithmic theory of laws to either define what a law of nature is metaphysically or to explain how we might come to have evidence in support of our knowledge of the laws.

It has been pointed out a number of times that our everyday concept of "information" is multifaceted and seems to defy simple characterization (see for example Adriaans & van Benthem, 2008; Floridi, 2010). There are a number of alternative definitions of information that attempt to capture these different aspects of what we mean by "being informed". The complexity of an object or message is one such aspect, but it isn't the only one. Two definitions of information attempt to capture another important aspect: that of being surprised by the outcome. Classic information theory, as given by Shannon and Weaver (1949), measures this aspect mathematically via the concept of "entropy". This measure, which is related to its namesake in thermodynamics, defines information as the amount of uncertainty reduced by receiving a message. The less probable a message (overall or relative to a previous message), the more informative it is to the recipient. A similar account is given by epistemic logic (van Benthem & Martinez, 2008), which measures information as "knowledge update". Luciano Floridi (2010) argues that any attempt to understand the nature of information in a purely syntactic or mathematical way is likely to ignore the different uses of information embedded in human mental and social life. Here, information must be interpreted semantically: in addition to being true, it must also be well-formed and cognitively significant to the recipient.

This shows that there are many different interpretations of information the advocate of the algorithmic theory can avail themselves of. They need not be tied to AIT on this front, and as such, need not be concerned about McAllister's third criticism.

### 3. Chaotic Systems

#### 3.1. The Argument from Local Variable Causes

McAllister's fourth and final argument is based on how phenomena types repeat in nature and how, with every time they repeat, they evolve differently to different rules.

If correct, this argument is the most serious one against the algorithmic theory. To appreciate the full force of McAllister's argument, we need to understand what he means by the difference between "phenomena" and "perturbations":

Each data point is the effect of a large number of causal factors. These can be divided, at least pragmatically, into two classes: *phenomena* and *perturbations*. A phenomenon is a feature of the world that constitutes the focus of a scientist's investigation. To the extent that the data set is causally determined by a phenomenon, it contains information about that phenomenon. Perturbations are all the causal factors distinct from the phenomenon that affect the data gathered. Typical perturbations in physics are dirt, friction, viscous drag, vibrations, temperature fluctuations, electrostatic charges, impurities in samples and variations in electric, magnetic and gravitational fields. Perturbations also include the influence of scientists' physiological and psychological dispositions on the recorded value. (2003, p. 637)

The idea that no physical system exists "in isolation" is nothing new. But McAllister's point cuts deeper than this. Phenomena are supposed to be the repeatable features of nature that scientists study: such as harmonic motion, biological organisms, chemical reactions, radioactive decay, etc. They are also the source of data for the formulation of laws. If the algorithmic theory of laws is correct, then these sources must provide evidence of compressibility. McAllister is arguing that there can be no algorithm that compresses *all* of these repeating tokens of a phenomenon-type, because with each occurrence, the system is subject to different "local" perturbing factors. In other words, although each system fails to exist in isolation, it fails to exist in isolation in different ways. As a result, there can be no algorithm that compresses the data for every system in the same way.

If correct, this is a powerful argument and shows that on the algorithmic view, there really are no "laws of nature" corresponding to all repeated phenomena in nature. At best, we can hope for a patchwork of rules and principles that might work in some circumstances, but none of these will ever amount to a compression.

In this section I will show that the existence of local variable causes is not so much a problem for the algorithmic theory as it might seem. I will argue that McAllister's observation about the nature of phenomena and how they repeat differently in nature is an expected outcome of *chaos theory*. What is more, such differences in how similar systems occur in nature can be emulated using larger systems, which are ultimately compressible to very simple computer programs. On the grand picture, the existence of local variable causes does not, therefore, give us reason to doubt the metaphysical picture of laws provided by the algorithmic theory. It does, however, have important epistemological consequences. We are not in a position to take a "God's eye view" of the universe. All we have are the individual occurrences of the phenomena and the effects of local variable causes on them. This does have an impact on what scientists

can expect to know. As I will argue, that provides an explanation for the practice of idealization in science and supports the algorithmic theory's postulation of lossy laws in scientific theories.

### 3.2. How Phenomena Emerge from Chaos

Chaos theory is a branch of mathematics that studies so-called “non-linear systems”. These are systems that are sensitive to initial conditions such that small differences at the beginning of two identical systems running in parallel results in larger differences after a small number of iterations. Scientists study chaotic behavior in a number of natural phenomena, including: weather patterns, celestial dynamics, population genetics, financial markets, environmental systems, etc. (Hall 1992). It only became practically possible to study chaotic systems, however, after the development of the digital computer in the 1940s and 1950s. One of the reasons is that chaotic systems often involve different parts with many variables that change over time. Before the creation of the electronic computer, it was very difficult to model the behavior of these systems and study their properties analytically.

One of the first and most famous to do so was Benoit Mandelbrot (1977). He utilized the processing speed and visual graphics of computers in order to produce striking two-dimensional and three-dimensional pictures of the behavior of chaotic systems. One well-known program is based on the so-called "Mandelbrot Set" and a picture of its output is given in Figure 1.

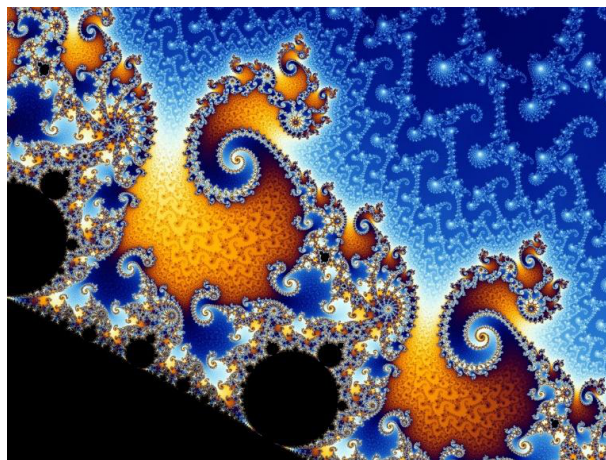


Figure 1: The Mandelbrot Set.

Many of the chaotic systems that Mandelbrot studied exhibited regularly occurring patterns, which he named “fractals”. A key feature of fractals is what Mandelbrot

termed “self-similarity” (1992, p. 123). These are shapes and patterns that occur time and again in the running of the program, and when viewed visually, can be found at every order of magnitude. Examples of self-similar shapes in the Mandelbrot set are “Sierpinski’s Triangle” and the “Koch Snowflake”. These recognizable patterns, that repeat again and again, look similar to McAllister’s “phenomena”. But there is one difference: unlike phenomena, these regularly occurring features are not subject to perturbations. As a result, they are perfectly orderly and compressible. In addition to “self-similar” patterns, there are also what Mandelbrot termed “approximately self-similar” fractals (1977, p. 34). Mandelbrot postulated that these are more likely to exist in nature and are the result of linear and non-linear systems working in tandem.

This idea, that phenomena that are approximately self-similar might arrive at the interplay between “randomness” on the one hand, and “order” on the other, has also been demonstrated in the work of Stephen Wolfram (2002). Wolfram has studied a different kind of computer model called “cellular automata” (CA), that involve a grid or matrix of cells that can be in either one of two states (“on” and “off”). The state or value of a cell at any given time is a function of its neighboring cells. CA can be designed to be in one, two, or  $n$  dimensions, can be given any set of initial conditions, can have any number of cells, and can be designed to follow any algorithm.

To illustrate, consider a simple one-dimensional CA following Wolfram’s “Rule-250”. This rule states that a cell is on if either its neighbors are on, or off if both its neighbors are off. The CA is evolved in a clockwise fashion, the rule being applied to the grid at every step. After a number of steps a pattern emerges from Rule-250 that looks similar to a chessboard (Fig. 2).

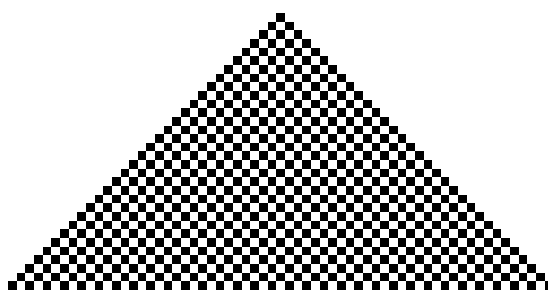


Figure 2: Wolfram’s Rule-250.

The system that emerges from Rule-250 is orderly and linear. It is not, therefore, an example of a chaotic system. In this respect it is very dissimilar to the kinds of patterns we see in the natural world. Nonetheless, rules can be designed for CA that produce patterns that are more complex. An example is Rule-30. Starting with a single cell in the “on” state, this rule gives the following command: “First, look at each cell and its right-hand neighbor. If both of these were [off] in the previous step,

then take the new color of the cell to be whatever the previous color of its left-hand neighbor was. Otherwise, take the new color to be the opposite of that” (Wolfram 2002, p. 27). The pattern that emerges from applying this rule after many iterations is given in Figure 3.

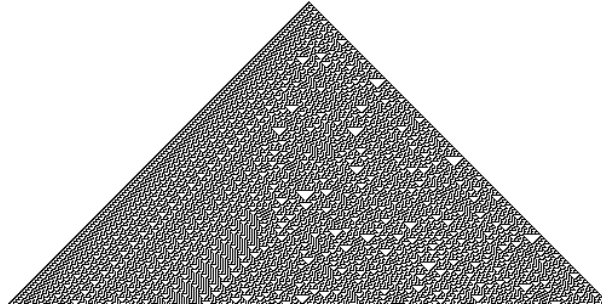


Figure 3: Wolfram's Rule-30.

As can be seen, the result of Rule-30 is very different from Rule-250. As Wolfram puts it: “it involves a mixture of order and randomness: localized structures are produced which on their own are fairly simple, but these structures move around and interact with each other in very complicated ways” (2002, p. 234). The localized structures that are visible in Rule-30 are very similar to McAllister's phenomena. They are repeatable tokens of a similar type, but their occurrence and their precise shape and size are affected by local interactions. It has been proven independently that Rule-30 satisfies the definition of a chaotic system (Cattaneo *et al.*, 2000) and therefore serves as a good model of chaotic systems in the natural world.

What this shows is that features in the world can emerge, which from a local point of view, produce data that cannot be easily compressed. However, taken as a whole, the emergence of the pattern results from the application of a very simple rule. The data that could be gathered from the entire system is highly compressible. Wolfram claims this was one of the remarkable results of his years of study into the behavior of CA:

Over and over again we see the same kind of thing: that even though the underlying rules for a system are simple, and even though the system is started from simple initial conditions, the behavior that emerges can be highly complex. And I will argue that it is this basic phenomenon that is ultimately responsible for most of the complexity that we see in nature. (2002, p. 28)

In order for the natural world to be compressible, it does not need to resemble a chess board. It can be chaotic and exhibit phenomena which are affected by local variable causes. Wolfram's CA have proven an effective means to model the occurrence of many phenomena in nature, and have been applied to help improve our



understanding of mollusk shells (Coombes, 2009), plant leaves (Peak *et al.*, 2003) and skin patterning (Packard, 2001), among others.

### 3.3. Idealization and Lossy Compression

The existence in nature of what McAllister calls phenomena subject to local variable causes does not, therefore, undermine the algorithmic theory as an account of the metaphysics of laws. As has been proven in studies in chaos theory, such behavior emerges quite naturally from very simple rules and initial starting conditions. If we took a God's eye view of the universe as a whole, from its beginning to its end, then it ought to be compressible to a very simple set of algorithms. The laws of nature would then be equivalent to these algorithms.

Although the existence of chaotic systems does not pose a challenge to the metaphysics behind the algorithmic theory, it might raise an epistemic problem. Clearly, scientists cannot take a God's-eye position. They cannot survey the totality of the universe from its beginning to its end. To continue with the CA model of nature, we are like individual cells that can only experience the self-similar patterns in our local neighborhood. But as we know, data gathered from observing these will in all likelihood not be compressible. They will depend on local perturbing factors, which if we attempt to include in an algorithmic description of the phenomenon, would result in little or no compression.

The algorithmic theory of laws has the resources to explain what is going on here. As I have argued elsewhere (Wheeler 2012, 2017 and 2018), it is prudent for an advocate of the algorithmic theory to follow computer science in acknowledging two kinds of data compression. The example of algorithmic compression given in section 1 is an example of "lossless compression": given the algorithm (and any other singular input data) it is possible to return the original string intact, without any information loss. Lossless compression is very important when data is being used for a purpose that cannot tolerate any errors. In image compression, for example, images for scientific purposes and medical diagnoses are compressed losslessly: any errors in the final product could lead to incorrect hypotheses or misdiagnosis. But when some errors are tolerable, a different kind of compression is possible—what is termed "lossy compression". Here significant gains in compression are possible if we are willing to sacrifice some of the information in the original string (Braddon-Mitchell, 2001).

We know that when scientists study individual phenomena in nature they frequently resort to idealization. Instead of describing a complete list of all the factors that can influence the values of a magnitude in a physical system, scientists will typically focus on only a small set. Consider the so-called "law of the pendulum":

$$T = 2\pi\sqrt{\frac{l}{g}}$$

This law describes the motion of pendulums and other bodies subject to simple harmonic motion. The law itself tells us how the system evolves over time with respect to the length of the pendulum and the strength of gravity. It is well known that this law cannot perfectly describe the evolution of an actual pendulum. Most pendulums are subject to friction, drag, and for larger swings, the time taken for a complete cycle is affected by its amplitude. But scientists are well aware of these other factors and when greater accuracy is called for, can compensate for them by using more complicated laws. The algorithmic theory of laws can explain this practice by claiming that what scientists are doing when they idealize is they *simplify* the data and formulate *lossy* compression algorithms. If a scientist was to use that algorithm for real-world data, they would not be able to use it to predict values with complete accuracy.

What the algorithmic theory of laws can do is provide a more nuanced explanation for why scientific laws do not perfectly describe the behavior of most observed physical systems. According to Lewis, scientists look for a “best balance” of strength and simplicity. But this explanation assumes that scientists work out *a priori* what the best information-to-compression tradeoff will be. There is no reason for thinking that scientists would take a 50 : 50 line in valuing informativeness and compression gains. Surely this is something that should be decided by practice, taking into consideration what scientists want to do with their laws (i.e., how much error can be accommodated) and how much computational resources they have at their disposal. The algorithmic theory perfectly accounts for this: when scientists want very simple descriptions, they idealize a lot and their algorithms will take on a more lossy character. When scientists want to focus on a particular occurrence of a phenomenon, and they want greater accuracy, then they can take into consideration data about the local perturbing factors.

On a day-to-day level, scientists therefore choose the level of compression they want for their laws. Perhaps for teaching and easy communication purposes highly compressed algorithms are used—but when more detail is needed, such as for accurate prediction, more complex algorithms are used in addition to more detailed singular data.

This leaves open the question of whether scientists can ever know what the laws of nature are (i.e., the algorithms in the best possible compression) based on observations of the world. Nothing in the existence of chaotic systems rules out that scientists might one day discover the final “theory of everything”. If they do, then, according to the algorithmic theory, this will involve hitting upon the algorithms in

the best compression. The existence of chaos is no more a barrier to preventing this for the algorithmic theory than it is for knowing the laws in any other regularity-based account.

## 4. Quantum Mechanics

### 4.1. Are Quantum Systems Algorithmically Random?

The behavior of objects at the quantum level provides an additional source of concern for the algorithmic theory. It is known that quantum processes behave in ways quite unfamiliar to everyday objects: wave–particle duality, entanglement and superposition are examples of this. It is also commonly believed that quantum events are “random” in the sense that at one point in time it is not possible to predict with certainty how a system will behave at a later time. This randomness is very different in kind from the sort of unpredictability that might emerge from scientists trying to predict chaotic systems. There the problem emerges because of the complexity of the phenomenon and its interactions. The randomness in quantum mechanics is believed by some to be far more fundamental.

The main reason for this randomness can be appreciated by looking at the most celebrated law of quantum mechanics—the Schrödinger Equation:

$$i\hbar \frac{\partial}{\partial t} \Psi(r, t) = -\frac{\hbar^2}{2m} \nabla^2 \Psi(r, t) + V(r) \Psi(r, t)$$

This equation describes the evolution of the “wave function” for a quantum system, such as the hydrogen atom. The wave function is a probability measure concerning the momentum and position of a particle at a given time. Does the Schrödinger Equation provide an algorithmic compression of empirical data? That is a difficult question to answer. On the one hand, the Schrödinger Equation only describes an “ideal quantum system”, so in this respect it does not provide a lossless compression. But even putting this to one side, there are reasons for thinking that no scientific law about behavior at the quantum level could be an effective compression.

Much of the reason for thinking that quantum systems are incompressible stems from the Copenhagen Interpretation. This view, which was held by Niels Bohr and Werner Heisenberg at the University of Copenhagen, makes three claims about the nature of quantum systems:

- (i) *Non-Realism*. Before an act of measurement there is no value for the position or momentum of a sub-atomic particle such as an electron. There is no fact of the matter as to “where the electron is” unless it is interacting with a measuring device.

- (ii) *Non-Determinism*. It is not possible to predict where a particle will be at the moment of measurement. All previous facts about the quantum state underdetermine its evolution after observation.
- (iii) *Non-Locality*. Two entangled particles can affect one another's state instantaneously regardless of their distance apart. The measurement of one particle will cause the entangled partner to assume a complimentary value.

The last of these claims, non-locality, is not much of a problem for the algorithmic theory. If the universe operates in a fashion that breaks the light speed barrier (such that one event can cause another to occur simultaneously in another part of the universe), then there could still be algorithms which efficiently compress empirical data. Only if we take a stronger metaphysical stance and assume the universe is something like a computer might this pose a problem. But the algorithmic theory as outlined here makes no assumptions about this.

The first and second claims are more of a problem for the algorithmic theory. Although by itself the first claim might not seem such a problem (if there is no “fact of the matter” about the electron's position, then there is no data to compress), this combined with non-determinism becomes an issue. It is effectively like saying that although the world contains a body of information, the information itself contains many gaps. As has been shown by Ulvi Yurtsever (2000) and Cristian Calude (2004), such gaps would make empirical data about the quantum world incompressible. This is even more worrying when we appreciate that what goes on at the quantum level could be important for what happens at the macro level. If there really are gaps in information at the quantum level, then this would suggest many parts of nature are equally incompressible.

This quantum behavior is no doubt puzzling for many theories of lawhood and it is made worse by the fact that there is no widely held consensus within the scientific community regarding how quantum mechanics should be interpreted. Some accept the Copenhagen Interpretation and as a result believe that nature is, at the fundamental level, non-deterministic and random. If that is the case, then the best the algorithmic theory can do is posit the existence of probabilistic laws for these kinds of systems. Such laws could still provide effective compressions, although they are unlikely ever to be lossless. In other words, they would be lossy laws. But they wouldn't be ideal laws. Ideal laws involve some degree of idealization with known idealization conditions. Probabilistic laws posit a brute probability. This might allow us to get more-or-less accurate predictions, but we could not posit conditions that explain when the law gets it wrong and *why* it gets it wrong. This is different, for example, from the law of the pendulum which does provide such conditions.

Another solution would be to abandon the Copenhagen Interpretation entirely. The most common way to do this is to adopt a “hidden-variable theory”. These al-

ternative interpretations deny the non-realism, non-determinism and non-locality claims by positing as yet undiscovered objects and processes that contain the information missing from the “gaps”. There are a number of different hidden-variable theories and if one of these turns out to be correct, it would provide an obvious way out for the algorithmic theory. As the scientific view has yet to be decided on this point, I do not think there are grounds to say that the existence of quantum processes refutes the algorithmic position. Before concluding, I want to draw attention to a recent hidden-variable theory, which like the work of Wolfram, uses CA to model nature.

#### 4.2. Gerard 't Hooft's CA Interpretation of Quantum Mechanics

Famously, Einstein proclaimed “God does not play dice”, reflecting his apparent disdain for the Copenhagen Interpretation of quantum mechanics. The Nobel-prize winning physicist Gerard 't Hooft agrees:

Einstein may have been right, when he objected to the conclusions drawn by Bohr and Heisenberg. It may well be that, at its most basic level, there is no randomness in Nature, no fundamentally statistical aspect to the laws of evolution. Everything, up to the minutest detail, is controlled by invariable laws. Every significant event in our universe takes place for a reason, it was caused by the action of physical law, not just by chance. (2016, p. 10)

One of the reasons 't Hooft thinks quantum mechanics appears so puzzling is its tendency to mix quantized processes (that come in discrete values) with continuous laws and equations. As he puts it: “quantum mechanics as we know it today, is the imperfect logic resulting from an incomplete discretization” (2016, p. 14). It is not that 't Hooft disagrees with the quantization of nature—he just thinks physicists haven't gone far enough. This is what attracted him to CA as a way of modelling quantum phenomena. Because CA are discrete systems and carry finite amounts of information, they are an ideal way to model the evolution of quantum states:

We have to select a very special set of mutually orthogonal states in Hilbert space that are endowed with the status of being ‘real’. This set consists of the states the universe can ‘really’ be in. At all times, the universe chooses one of these states to be in, with probability 1, while all others carry probability 0. We call these states ontological states, and they form a special basis for Hilbert space, the ontological basis. (2016, p. 8)

In the final analysis, 't Hooft believes that our current nomenclature of “particles”, “waves”, “spin” etc., will ultimately be shown to be non-referring. However, they are still useful templates for describing what comes down to something equivalent to the

cells in a CA—what he calls “beables” (2016, p. 69). These objects are very small, perhaps nothing larger than the Planck length, and are responsible for all the phenomena that we see in nature. Part of ’t Hooft’s ongoing project is to show how by simply following certain rules or programs, the time evolution of beables explains the behavior we experimentally determine for quantum processes. At the basic level, there is no Schrödinger Equation or Hilbert space, just classical algorithms determining the time-evolution of cells in a gigantic CA.

If ’t Hooft is correct, then his CA model of the quantum world can account for what we currently observe without needing to advocate non-realism or non-determinism. For every current quantum state, there is a fact of the matter concerning the current values of all of its physical properties. In turn, this allows us to predict with certainty how it will evolve. One problem for any hidden-variable theory is Bell’s Theorem:

In a theory in which parameters are added to quantum mechanics to determine the results of individual measurements, without changing the statistical predictions, there must be a mechanism whereby the setting of one measuring device can influence the reading of another instrument, however remote. Moreover, the signal involved must propagate instantaneously, so that such a theory could not be Lorentz invariant. (1964, p. 407)

Named after John Bell, this widely-accepted result shows that if an interpretation of quantum mechanics resorts to hidden variables that carry information, such as ’t Hooft’s CA, then one of two options must be true: either non-locality fails, or, human choice and behavior is fully determined (1993, p. 45–46). Interestingly, none of these results are a problem for the algorithmic theory of laws. As we have already seen, non-locality is not inconsistent with the algorithmic theory. Non-locality is a problem for mechanistic theories of laws that explain nomic processes by involving some kind of transference of quantity or force. But the algorithmic theory makes no such claims. All that is required is that patterns in the data emerge which ultimately makes it algorithmically compressible. This is perfectly at home with activities and processes in nature that can operate faster than the speed of light.

The second option, that human beings have no free choice—especially at the moment of measurement—might seem worrying, but philosophers have explored this possibility for many centuries, and it doesn’t require the truth of quantum mechanics to believe that the world is deterministic. The existence of compatibilist theories of free will show that even in a deterministic world, there are many who believe that free will is possible. If we take this option we need not, therefore, abandon our traditional institutions that depend upon free will such as morality and the law.

To sum up this section, the argument from quantum mechanics does not show that the natural world is incompressible. If the Copenhagen Interpretation is correct, then the algorithmic theory will have to adopt probabilistic laws as its basic algo-

rithms. This can still provide compression, although it is likely to be of a lossy kind. And if hidden-variable theories are correct, such as 't Hooft's CA interpretation, then it is reasonable to expect that the quantum world provides empirical data that is algorithmically compressible. The two main worries from Bell's theorem turn out not to be a problem for the algorithmic theory of laws, which by not requiring locality, is able to escape this worry surprisingly well.

## 5. Conclusion

This paper has considered three potential sources of incompressibility in the universe. If any of them turned out to be true, it would pose a serious challenge to the algorithmic theory of laws. As I have argued, none of these three sources provides a convincing case for incompressibility. As such, the algorithmic theory of laws is a viable alternative to other regularity-based accounts and ought to be given much greater attention by philosophers than it currently receives.

## References

- Adriaans, P. & van Benthem, J. 2008. Information is What Information Does. In: *The Philosophy of Information*, p. 7–31. Amsterdam: Elsevier.
- Armstrong, D. 1983. *What is a Law of Nature?* Cambridge: Cambridge University Press.
- Artacho, F. 2012. 100 billion step walk on the digits of pi. <https://gigapan.com/gigapans/106803>. Access: 04/06/2019.
- Barrow, J. 1991. *Theories of Everything*. Oxford: Oxford University Press.
- Bell, J. 1964. On the Einstein Podolsky Rosen Paradox. *Physics* 1: 195–200.
- Bird, A. 2008. The Epistemological Argument Against Lewis's Regularity View of Laws. *Philosophical Studies* 138: 73–89.
- Braddon-Mitchell, D. 2001. Lossy Laws. *Nous*, 35(2): 260–277.
- Calude, C. 2004. Algorithmic Randomness, Quantum Physics and Incompleteness. In: M. Margenstern (ed.), *Proceedings of the 4th International Conference on Machines Computations, and Universality*, p. 1–17. Berlin: Springer.
- Cattaneo, G.; Finelli, M. & Margara, L. 2001. Investigating Topological Chaos by Elementary Cellular Automata Dynamics. *Theoretical Computer Science* 244 (2000): 219–241.
- Chaitin, G. 1987. *Algorithmic Information Theory*. Cambridge: Cambridge University Press.
- Chaitin, G. 2005. *Meta Maths — The Quest for Omega*. London: Atlanta Books.
- Coombes, S. 2009. *The Geometry and Pigmentation of Seashells*. Nottingham: Department of Mathematical Sciences, University of Nottingham.
- Davies, P. 1995. Algorithmic Compressibility, Fundamental and Phenomenological Laws. In: F. Weinert (ed.), *Laws of Nature: Essays on the Philosophical, Scientific and Historical Dimensions*, p. 248–267. Berlin: Walter de Gruyter & Co.
- Feynman, R. 1963. *Lectures on Physics (Vol. 1)*. Reading: Addison-Wesley.
- Floridi, L. 2010. *Information: A Very Short Introduction*. Oxford: Oxford University Press.

- Gell-Mann, M. 1987. Simplicity and Complexity in the Description of Nature. Talk Delivered to *The Caltech Associates*, Pasadena 01/10/1987.
- Grünwald, P. & Vitanyi, P. 2008. Algorithmic Complexity. In: P. Adriaans & J. van Benthem (eds.), *Philosophy of Information*, p. 289–328. Amsterdam: Elsevier.
- Hall, N. 1992. Introduction. In: *The New Scientist Guide to Chaos*, p. 7–10. London: Penguin.
- Kolmogorov, A. 1965. Three Approaches to the Definition of the Quantity of Information. *Problems of Information Transmission*, 1(1): 1–7.
- Landauer, R. 1996. The Physical Nature of information. *Physics Letters A*, 217: 188–193.
- Lewis, D. 1973. *Counterfactuals*. Oxford: Blackwell.
- Lewis, D. 1983. New Work for a Theory of Universals. *Australasian Journal of Philosophy*, 61(4): 343–377.
- Lewis, D. 1986. *Philosophical Papers Volume II*. Oxford: Oxford University Press.
- Lewis, D. 1994. Humean Supervenience Debugged. *Mind* 103(412): 473–490.
- Lloyd, S. 2006. *Programming the Universe: A Quantum Computer Scientist takes on the Cosmos*. New York: Alfred Knopf.
- Luminet, J.; Weeks, J.; Riazuelo, A.; Lehouq, R. & Uzan, J. 2003. Dodecahedral space topology as an explanation for weak wide-angle temperature correlations in the cosmic microwave background. *Nature* 425: 593–595.
- Mach, E. 1894. *Popular Scientific Lectures*. Illinois: Open Court.
- Mandelbrot, B. 1977. *The Fractal Geometry of Nature*. New York: W. H. Freeman and Company.
- Mandelbrot, B. 1992. Fractals — a Geometry of Nature. In: N Hall (ed.), *The New Scientist Guide to Chaos*, p. 122–135. London: Penguin.
- McAllister, J. 2003. Algorithmic Randomness in Empirical Data. *Studies in the History and Philosophy of Science* 34 (3): 633–646.
- McAllister, J. 2005. Algorithmic Compression of Empirical Data: Reply to Twardy, Gardner and Dowe. *Studies in the History and Philosophy of Science* 36 (2): 403–410.
- Packard, A. 2001. A ‘Neural’ Net that Can Be Seen with the Naked Eye. In: W. Backhaus (ed.), *Neuronal Coding of Perceptual Systems*, p. 397–402. London: World Scientific.
- Peak, D.; West, J.; Messinger, S. & Mott, K. 2003. Evidence for complex, collective dynamics and emergent, distributed computation in plants. *Proceedings of the National Academy of Sciences of the United States of America* 101 (4): 918–922.
- Shannon, C, and W Weaver. 1949. *The Mathematical Theory of Information*. Chicago: University of Illinois Press.
- Solomonoff, R. 1964. A Formal Theory of Inductive Inference: Part I. *Information and Control* 7(1): 1–22.
- T Hooft, G. 2016. *The Cellular Automaton Interpretation of Quantum Mechanics*. Heidelberg: Springer.
- Tomkow, T. 2013. The Computational Theory of the Laws of Nature. <http://tomkow.typepad.com/tomkowcom/2013/09/the-computational-theory-of-natural-laws.html>. Access: 08/06/2019.
- Tomkow, T. 2014. Computation, Laws and Supervenience. <https://tomkow.typepad.com/tomkowcom/2014/02/computation-laws-and-supervenience.html>. 08/06/2019.
- Twardy, C.; Gardner, S. & Dowe, D. 2005. Empirical Data Sets are Algorithmically Compressible: Reply to McAllister? *Studies in the History and Philosophy of Science* 36 (2): 391–402.



- van Benthem, J. & Martinez, M. 2008. The Stories of Logic and Information. In: P. Adriaans & J. van Benthem (eds.) *The Philosophy of Information*, p. 255–288. Amsterdam: Elsevier.
- van Fraassen, B. 1980. *The Scientific Image*. Oxford: Clarendon Press.
- Wallace, C. & Boulton, D. 1968. An information Measure for Classification. *Computer Journal* 11 (2): 185–194.
- Wheeler, B. 2012. *The Metaphysics of Ideal Laws*. PhD Dissertation. University of Cambridge.
- Wheeler, B. 2016. Simplicity, Language-Dependency and the Best System Account of Laws. *Theoria: An International Journal for Theory, History and Foundations of Science* 31(2): 189-206.
- Wheeler, B. 2017. Humeanism and Exceptions in the Fundamental Laws of Physics. *Principia: An International Journal of Epistemology* 21(3): 317–337.
- Wheeler, B. 2018. *Idealization and the Laws of Nature*. Geneva: Springer.
- Wolfram, S. 2002. *A New Kind of Science*. Champaign: Wolfram Media.
- Yurtsever, U. 2000. Quantum Mechanics and Algorithmic Randomness. *Complexity* 6 (1): 27–34.

## Notes

<sup>1</sup>For critical discussion, see Armstrong (1983, p. 66–74), Bas van Fraassen (1989, p. 40–64) and Alexander Bird (2008).

<sup>2</sup>See Terence Tomkow (2013 and 2014) for a detailed and clear presentation of the algorithmic theory of laws from a Humean perspective. Versions of this theory have been advocated in one form or another by Richard Feynman (1963), Ray Solomonoff (1964), Andrey Kolmogorov (1965), Murray Gell-Man (1987), John Barrow (1991), Paul Davies (1995), Rolf Landauer (1996), David Braddon-Mitchell (2001), Gregory Chaitin (2005) and Seth Lloyd (2006). I have also developed and defended this view elsewhere (see Wheeler 2012, 2016, 2017 and 2018).