C. Calude, A. Salomaa. Algorithmically coding the universe, in G. Rozenberg, A. Salomaa (eds.). *Developments in Language Theory*, World Scientific, Singapore, 1994, 472-492.

Algorithmically Coding the Universe[1]

Cristian Calude[2] and **Arto Salomaa**[3]

Abstract

All science is founded on the assumption that the physical universe is ordered. Our aim is to challenge this hypothesis using arguments from the algorithmic information theory.

1 Introduction

Algorithmic information theory opens new vistas that extend far beyond the traditional boundaries of mathematics and computer science. How can we describe the seemingly random processes in nature and reconcile them with the supposed order? How much can a given piece of information be compressed? These are matters of fundamental scientific importance that will be discussed below, mainly from an informal or semi-formal point of view.

The *descriptional complexity* of a sequence of bits, finite or infinite, is the length of the shortest sequence of bits defining the originally given sequence. A given sequence being *random* means, roughly, that its descriptional complexity equals its length. In other words the simplest way to define the sequence is to write it down. This seems to be the case for the sequence 0110 0001 1011 1111, whereas the sequence $(01)^{1000}$ can surely be defined in much less than 2000 bits. These observations should be contrasted with the fact that all sequences of the

[1]This work has been supported by Grant A18/XXXXX/62090/3414012 of Auckland University and the Project 11281 of the Academy of Finland.

[2]Computer Science Department, The University of Auckland, Private Bag 92109, Auckland, New Zealand; Email: c_calude@cs.auckland.ac.nz.

[3]Academy of Finland and Department of Mathematics, University of Turku, 20500 Turku, Finland; Email: ASALOMAA@sara.cc.utu.fi.

same length are equally likely when viewed, for instance, as the result of coin tosses.

The sequence $(01)^{1000}$ can be described as the pattern 01 repeating 1000 times. There may be other ways to compress information than detected patterns. There is no pattern visible in tables of trigonometric functions. Even tables of modest size give rise to a rather long sequence of bits if everything is expressed as a single sequence. However, a much more compact way to convey the same information is to provide instructions for calculating the tables from underlying trigonometric formulas. Such a description is brief and, moreover, can be used to generate tables of any size.

Our considerations lead to Chaitin's mysterious number Ω, "the secret number", "the number of wisdom", "the number that can be known of but not known through human reason". Indeed, if we know some reasonably long prefix of Ω, say the first 10000 bits, then we are able to decide of formal systems F and well-formed formulas α, both of a reasonable size, whether α is provable, refutable or independent in F, Rozenberg and Salomaa [38].

A brief outline of the contents of the paper follows. Very little previous knowledge is required on part of the reader; Salomaa [39] may be consulted if need arises. A proof for the undecidability of the halting problem, based on algorithmic information theory, is given in Section 2. We believe that the proof is very illuminating also for classroom use. The secret number Ω is introduced in Section 3 in connection with the coding of the halting problem. Comparisons along these lines are made in Section 4, taking into account also problems of nameability. The concluding Section 5 and 6 deal with modeling reality, and the Hypothesis of Order as opposed to the Hypothesis of Randomness. The Final Remarks concern future work.

2 The Halting Problem Revisited

Suppose we have an enumerable set of mathematical "problems"

$$Problem = Problem_1, Problem_2, \ldots, Problem_n, \ldots$$

each of which has a yes/no answer. Assume further that the function $n \mapsto Problem_n$ is computable (i.e. the conditions in $Problem_n$ are written out in an effective manner, as a function of n). To such a sequence $Problem$ we may associate a set

$$Y(Problem) = \{i \in \mathbf{N} | Problem_i \text{ has "yes" for answer}\}.$$

The sequence $Problem$ is *algorithmically decidable* if the set $Y(Problem)$ is recursive; otherwise, it is *algorithmically undecidable.* It is possible to distinguish the case when $Y(Problem)$ is recursively enumerable (but not recursive); in fact, a whole hierarchy of degrees of undecidability can be constructed, but this lies outside the aim of this paper.

The most important undecidable problem is the halting problem. In this case (P_i) is an effective enumeration of all computer programs (without loss of generality we may assume that P_i accepts as inputs only natural numbers and eventually produces a natural number, if ever; the input is considered a part of the program itself). The set associated to the halting problem is

$$Y(HP) = \{i \in \mathbf{N} | P_i \text{ "eventually halts"}\}.$$

At the first glance the halting problem might seem decidable, a possible reason being that a program that eventually halts can be demonstrated by simply running the program! More, we know many examples of programs that can be proved to halt or not, even without running them (e.g. the program computing four naturals $x, y, z, n > 3$ such that

$$x^n + y^n = z^n$$

will never halt by virtue of the recent proof of Fermat's Great Theorem by Andrew Wiles; see Ribet [37]).

The main difficulty connected with the halting problem – as well as with many other decision problems – is not the possibility to solve particular instances, but *in solving the problem in general.* And this general case is undecidable, as Turing proved.

There are many different proofs of the undecidability of the halting problem, most of them relying on diagonalization (see, for instance, Salomaa [39], Calude [9]). We follow here a different path, just computing – following Chaitin – the quantity of information required by a computation. Assume that there exists a **halting program** deciding if an arbitrary program eventually halts, i.e. computing the characteristic function of the set $Y(HP)$. Construct the following program (which essentially makes use of the **halting program**):

- read a natural N;
- generate all programs up to N bits in size;
- use the **halting program** to check for each generated program whether it halts;
- simulate the running of the above generated programs;
- output the double of the biggest value outputed by these programs.

First, notice that the above program eventually halts for every natural N. How long is the above program ? It is about $\log_2 N$ bits. Indeed, the program consists of the input data N (which requires about $\log_2 N$ bits) and a constant part. Globally, the program has $\log_2 N + O(1)$ bits. *There is a big difference between the size – in bits – of our program and the size of the output produced by this program.* Indeed, for large enough N, our program will belong to the set of programs having less than N bits (because $\log_2 N + \mathrm{O}(1) < N$). Accordingly, the program will be generated by itself – at some stage of the computation. In this case we have got a contradiction since our program will output a natural number two times bigger than the output produced by itself!

So, the halting problem is undecidable. The reason is in the difference in size between coding resources (which are about $\log_2 N$ bits) and programs to be coded (of size less than N).

The above – seemingly esoteric – problem has many interesting consequences, one of them being Gödel's Incompleteness Theorem. Indeed, assume that one has a formal axiomatic system from which all arithmetical truths follow, in a consistent way (i.e. all and only all arithmetical true statements can be proven within the system). Then we can decide if an arbitrary program P_i eventually halts:

> *Run through all possible proofs until either one reaches a proof that P_i halts or one finds a proof that P_i does not halt*

(these statements can be expressed as simple arithmetical formulae, via a suitable coding). But this is impossible, as the halting problem is undecidable! So, there exists a true statement of the form "P_i does not halt" which cannot be derived within the axiomatic system. We have reached the "heart" of Gödel's argument, and even, we have got an extra bonus. In Gödel's terms, incompleteness is a relative result, i.e. it essentially depends upon the fixed axiomatic system. In the above argumentation an *absolute* claim has been proved, namely the **lack of a decision procedure for the truth/falsity of statements expressible within the formal axiomatic system**. Finally, let us recognize the fact that Gödel's Incompleteness Theorem is merely an assertion about a coding impossibility:

> *In a consistent axiomatic system for arithmetic there are not enough resources to code the true statements.*

In fact a more quantitative remark can be proven (see Chaitin [12], Calude [11]):

> *An N-bit set of axioms cannot yield a theorem that asserts that a specific object is of complexity substantially greater than N.*

3 Coding the Halting Problem

How much information is contained in the halting problem? The easiest way to measure it is to switch from the deterministic point of view to the probabilistic one. Put all programs in a bag and think of P_i as a binary string. For technical reasons (see, for instance, Calude [11], Rozenberg and Salomaa [38]) we may assume that the resulting set of strings is prefix-free, i.e. no string in the set is a proper prefix of another one. In other words, every program is self-delimiting: Its total length (say, in bits) is given by the program itself. Real programming languages are self-delimiting as they provide constructs for beginning and ending a program.

Fix a length N and proceed generating a program at random by tossing a fair coin N times (for each bit of the program). What is the chance that the program generated in this way will eventually halt? It is easy to see that the probability to get a particular program of length N is 2^{-N} and any program P_i that halts contributes

$$2^{-(length\ of\ P_i)}$$

to the halting probability. We have got Chaitin's formula

$$\Omega = \sum_{\{i\in \mathbf{N} \mid P_i \text{ halts}\}} 2^{-(length\ of\ P_i)}.$$

An alternative, but equivalent, way to think of the above experiment is to consider a universal program (Turing machine) which, instead of being given a specific program at the beginning of the computation, is fed with a "random string" of bits. Whenever the universal program requests another input bit we just toss a fair coin and input 1 or 0, according to whether the coin comes up heads or tails. Finally we ask the question: When the above procedure has begun, what is the probability that the universal program will eventually halt? Again, the answer is Ω.

It is important to realize that Ω depends upon the fixed enumeration P_i (or, equivalently, the universal program). So, we really have a class of probabilities Ω, not an absolute constant Ω, like π or e. But **all** these probabilities share a number of interesting properties, some of which are going to be discussed here.

First, $0 \leq \Omega \leq 1$, like all probabilities.[4] More precisely, $0 < \Omega < 1$. How big is Ω ? As a direct consequence of the "random experiment" it is very likely that the universal program will be instructed to do something impossible. This means that most probably the universal program stops almost immediately or enters a loop with a few instructions. This shows that actually Ω is quite close to 1, i.e. the binary expansion of Ω starts with a long string of 1's. However, this is not a "rule" for Ω: In the long run the digits of Ω become patternless,

[4]The fact that Ω is indeed a probability in the formal sense is not important for our discussion; see Calude [11] for more mathematical details.

unpredictable, immune to any conceivable gambling scheme. We shall return later with more concrete facts on this property. Now we discuss the way Ω encodes the solutions of the halting problem.

The number Ω is particularly important since it encodes the halting problem in a very compact way. In Bennett's words (see [7]):

> *[Ω] embodies an enormous amount of wisdom in a very small space ... inasmuch as its first few thousands digits, which could be written on a small piece of paper, contain the answers to more mathematical questions than could be written down in the entire universe.*

> *Throughout history mystics and philosophers have sought a compact key to universal wisdom, a finite formula or text which, when known and understood, would provide the answer to every question. The use of the Bible, the Koran and the I Ching for divination and the tradition of the secret books of Hermes Trismegistus, and the medieval Jewish Cabala exemplify this belief or hope. Such sources of universal wisdom are traditionally protected from casual use by being hard to find, hard to understand when found, and dangerous to use, tending to answer more questions and deeper ones than the searcher wishes to ask. The esoteric book is, like God, simple yet undescribable. It is omniscient, and transforms all who know it ... Omega is in many senses a cabalistic number. It can be known of, but not known, through human reason. To know it in detail, one would have to accept its uncomputable digit sequence on faith, like words of a sacred text.*

A finitely refutable statement is equivalent to the assertion that some program – searching systematically for some nonexistent object – never halts. Goldbach's Conjecture (*every even number greater than* 2 *is the sum of two primes,*[5] or Riemann Hypothesis (*the function*

$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \cdots$$

has all complex roots on the axis $x = \frac{1}{2}$) belong to this class. Solutions for both conjectures, as well as for many other similar problems, are contained in the first few thousand of bits of Ω. Indeed, due to the following inequalities

$$\Omega_n < \Omega < \Omega_n + 2^{-n}, n = 1, 2, \ldots$$

(here Ω_n is the number consisting of the first n bits of Ω) one can solve the halting problem for all n-bit programs as follows:

[5] As in the following examples: $6 = 3 + 3$, $8 = 3 + 5$, $10 = 3 + 7 = 5 + 5$, $12 = 5 + 7, \ldots$

Start a systematic (dovetailing) search through all programs that eventually halt until **enough** *halting programs have been found to overpass* Ω_n.

Notice that we will never get *all* these programs, but if we have enough patience (and computational resources) we finally get enough programs P_{i_1}, $P_{i_2}, P_{i_3}, \ldots, P_{i_k}$ of lengths $l_{i_1}, l_{i_2}, l_{i_3}, \ldots, l_{i_k}$, such that

$$\sum_{j=1}^{k} 2^{-l_{i_j}} > \Omega_n.$$

In the above list there are programs longer than n bits, as well as some shorter ones. It really doesn't matter; the main thing is that the list $P_{i_1}, P_{i_2}, P_{i_3}, \ldots, P_{i_k}$ contains **all** programs shorter than n bits (otherwise, their contribution to Ω would be larger than $\Omega_n + 2^{-n}$, a contradiction). If n is large enough, then among the halting programs

$$P_{i_1}, P_{i_2}, P_{i_3}, \ldots, P_{i_k}$$

we will find programs deciding almost all finitely refutable conjectures (i.e. all conjectures which can be expressed by reasonable long strings).

How large is the class of finitely refutable problems? This seems to be a difficult question. We will confine ourselves to some nontrivial examples. The first such example that comes to mind pertains the independent statements.

A statement expressible within a formal axiomatic system is *independent* of the system if neither the statement, nor the negation of it can be proven (within the system). The Parallel Postulate (*through a given point there is exactly one line parallel to a given line*), the Continuum Hypothesis (*there is no cardinal number strictly in between aleph -null, the cardinality of the the set of natural numbers, and aleph-one, the cardinality of the set of reals*) or (a slight variation of) Ramsey Theorem (asserting that *if a partition of a "big" finite set contains only a few classes, then at least one of these classes is "big enough"*) are probably the best known examples of independent statements (from Euclidean axioms, Zermelo-Fraenkel system or Peano arithmetic, respectively). Fix now an axiomatic system S and a statement s expressible in S. Construct the program $P(s)$ that searches systematically among the proofs of S for a proof or refutation of s. Then, s is independent with respect to S if only if $P(s)$ never halts.

In case the system S is recursively axiomatizable, i.e. the set of theorems in S is recursively enumerable, and S is sound with respect to the set of propositions of the form

"the nth bit of Ω is a 0"

"the nth bit of Ω is a 1"

(such a statement is in S only if it is true), then S can enable us to determine the positions and values of at most finitely many scattered bits of Ω. So, if N

is large enough, then the statement

$$s = \text{“the } N\text{th bit of } \Omega \text{ is a 0”}$$

is true, but unprovable in S. Neither provable is the negation of s, so s is independent of S. We can effectively construct the program $P(s)$, as above, which – *we know* – will never halt. If $P(s)$ is the Mth program, then the statement

$$\text{“the } M\text{th bit of } \Omega \text{ is a 0”}$$

is true and $M > N$. But the above statement is itself independent of S and the procedure can be iterated. In this way we generate an infinity of positions all of which are 0 in Ω, which is not possible. The conclusion is that *there is no way to effectively compute a bound for the finite set of provable positions and values of* Ω *(within a given recursively axiomatisable sound theory).*

4 More about the Inner Structure of the Halting Problem

Of course, Ω isn't the only possibility to encode the halting problem. Consider, for instance, the number

$$K = \sum_{i \geq 1} 2^{h(P_i)},$$

where $h(P_i)$ is 0 or 1 depending whether P_i eventually halts or not. Both K and Ω contain "globally" the same amount of information; they are both noncomputable. However, the number K_n consisting of the first n bits of K contains about $\mathrm{O}(\log n)$ bits of information,[6] in contrast with Ω_n that contains approximately n bits of information. So, "locally" K and Ω are fairly distinct. The basic difference between them lies in the degree of "organisation" of the information. In K the information is ordered, well structured, predictable. Knowing an infinity of halting programs (which can be obtained even by virtue of trivial reasons) makes us wiser in revealing the bits of K. Betting on these bits one can win consistently! Nothing similar is possible for Ω. We may know only a tiny finite set of bits of Ω, and even a bound for the number of such bits

[6] This is because K_n is completely specified by the number of indices $i \in \{1, 2, \ldots, n\}$ such that the ith Turing machine M_i halts on input i. Once the $\mathrm{O}(\log n)$-bit number is *known*, a direct simulation of the Turing machines M_j with $j = 1, 2, \ldots, n$ on all inputs $1, 2, \ldots, n$ produces *all* bits of K_n.

has to remain unknown for ever! Ω is chaotic, random, immune to any gambling scheme. The information in K is dilute; in Ω it is very compact.

We need to go deeper into the inner structure of Ω in order to reveal its chaotic character. On several occasions we have said that a certain object is a notation for, or denotes, another element. A convenient general term for this relantionship is *naming.* A successfully chosen name may be of a tremendous value as it may be seen as a bridge between the formal knowledge and common sense. To what extent are we free to choose names? A drastic limitation comes from the well known Berry's paradox.

Consider the number

one million, one hundred one thousand, one hundred twenty one.

This number appears to be

the first number not nameable in under ten words.

However, the above expression has only **nine** words!

It follows that the property of nameability is inherently ambiguous and, consequently, too powerful to be freely used. The list of similar properties is indeed very long; another famous example refers to the classification of numbers in *interesting* versus *dull.* There can be no dull numbers: If they were, the first such a number would be *interesting* on account of its dullness.

Naming is intimately linked to meaning. A famous model called the "meaning-text" model (due to I. A. Melchuk) states that to a large extent linguistics is the theory of two-way algorithmic translation between two infinite sets, "texts" and "meanings". "Texts" belong to a *natural language*, whereas "meanings" usually are constructed in terms of an *artificial language* having semantic emphasis. This translation process is active, in the sense that "texts" to be translated pass through a hierarchy of intermediate steps to reach a "meaning" and neither the set of meanings, nor the translation algorithms are *a priori* given, they are synthesized during the translation process; see Calude [10]

To be more precise we shall consider here only the process of naming by compression of information. A way to do this is to detect patterns. For instance, consider the following two binary strings of length 32:

$$10011001100110011001100110011001,$$

$$01101000100110101101100110100101.$$

The first string follows an obvious pattern: 1001 is written 8 times. No such pattern is visible in the second one, that actually was generated by tossing a coin. Tossing a fair coin 32 times can "theoretically" produce each of the 2^{32} binary strings of length 32. Classical Probability Theory assures us that there is no preference among the strings of the same length. Still, it is hard to believe that the first string was actually produced by tossing a fair coin! Laplace [28], pp.16-17 was, in a sense, aware of this paradox, as it may be clear from the following phrase:

> *In the game of heads and tails, if head comes up a hundred times in a row then this appears to us extraordinary, because after dividing the nearly infinite number of combinations that can arise in a hundred throws into regular sequences, or those in which we observe a rule that is easy to grasp, and into irregular sequences, the latter are incomparably more numerous.*

There are many other ways to compress information than by detecting patterns. There is no visible pattern in a long table of trigonometric functions. A much more compact way to convey the information in such a table is to provide instructions for calculating the table, e.g. using Euler's equation $e^{ix} = \cos x + i \sin x$. Such a description is not only compact, but it can be used to generate arbitrarily long trigonometric tables.

The above method fails to be adequate for empirical data. For instance, consider the collection of the results of the gold medal winners in the Olympic Games since 1896 (see Rozenberg and Salomaa [38]). For such an information the amount of compression is practically null, especially if the attention is restricted to the least significant digits. More, since the tendency is of (slow) improvement, the most significant digits have a kind of regularity which even makes predictions possible. Empirical data give rise to binary strings which have to be "explained" and new ones have to be predicted. This can be done by theories. A crude model of a theory is just a computer program which reproduces the empirical observations. Usually, there exist infinitely many such programs, but the interesting ones are clearly the *minimal* programs. These minimal programs can be used to measure the amount of compression on the initial data: Just compare the size of the program to the size of the input data. **Random** observations are characterized by the lack of compression: The most concise way to represent them is just to list them.

The above discussion suggests that a string w is random if the shortest program describing w is roughly of the same length as w. An infinite sequence (like the digits of Ω) is random if all the prefixes of the sequence are random strings. For words randomness is a relative property; it indicates how close is the length of the shortest program generating the word to the maximal value (computed for strings having the same length). For infinite sequences there is a sharp distinction between random and nonradom sequences.

5 Ω as a Model of the Universe

The idea that the Universe is a living organism is very old. Aristotle thought that the entire Universe “resembles a gigantic organism, and it is directed towards some final cosmic goal”.[7]

But, *What is life ? When must life arise and evolve?* or, maybe better, *How likely is life to appear and evolve? How common is life in the Universe*?

The evolution of life on Earth is seen as a deterministic affair, but a somewhat creative element is introduced through *random* variations and natural selection. Essentially, there are two views as regards the origins of life. The first one claims that the precise physical processes leading to the first living organism are *exceedingly improbable*, and life is in a way intimately linked to planet Earth (the events preceding the appearance of the first living organism would be very unlikely to have been repeated elsewhere). The second one puts no sharp division between living and non-living organisms. So, the origin of life is only one step, maybe a major one, along the long path of the progressive complexification and organisation of matter. See more in Barrow [1], Davies [17], Davies and Gribbin [18], Hawking [22].

To be able to analyse these views we need some coherent concept of *life*! Do we have it? It is not difficult to recognize life when we see it, but it looks tremendously difficult to set up a list of distinct features shared in common by all and only all living organisms. The ability to reproduce, the response to external stimuli, and growth are among the most frequent cited properties. But, unfortunately, none of these properties “defines” life. Just consider an example: the *virus* does not satisfy any of the above criteria of life though the viral diseases clearly imply biological activity.

Along the line of reasoning of Miller and Urey primeval soup and Darwin evolution, it appears that the spontaneous generation of life from simple inanimate chemicals occurs far more easily than its deep complexity would suggest. In other words, *life appears to be a rather common feature in the Universe*! Von Neumann [44] wished to isolate the mathematical essence of life as it evolves from physics and biochemistry. He succeeded to make the first step, showing that the exact reproduction of universal Turing machines is possible in a particular deterministic model Universe. These results are consistent with Holland's model of Universe [23, 24] and his claims concerning a possibility “to demonstrate that self-replicating systems can emerge from unorganized initial states”.[8]

Following this path of thought it may be possible to formulate a way to

[7]Teleology is the idea that physical processes can be determined by, or drawn towards, an *a priori* determined end-state.

[8]His mathematical arguments appear to be seriously flowed (see the critical discussion in Haji [21]).

differentiate between dead and living matter: by the *degree of organisation*.[9] According to Chaitin [13] an organism is a highly interdependent region, one for which the complexity of the whole is much less than the sum of complexities of its parts. Life means *unity*. Dead versus living can be summarised as the whole versus the sum of its parts. Charles Bennett's thesis is that a structure is deep *if it is superficially random but subtly redundant, in other words, if almost all its algorithmic probability is contributed by slow-running programs.* To model this idea Bennett has introduced the notion of "logical depth": a string logical depth reflects the amount of computational work required to expose its "buried redundancy" ([5]; see also Juedes, Lathrop, Lutz [27]):

> *... the value of a message is the amount of mathematical or other work plausibly done by its originator, which its receiver is saved from having to repeat.*

For John Wheeler the Universe is a gigantic information-processing system in which the output is as yet undetermined. He coined the slogan: *It from bit*! That is, *it* – every force, particle, etc. – is ultimately present through *bits* of information. And Wheeler is not unique on this path. Ed Fredkin and Tom Toffoli emphatically say **yes**. The Universe is a gigantic cellular automaton. No doubt! The only problem is that somebody else is using it.

> *All we have to do is "hitch a ride" on his huge ongoing computation, and try to discover which parts of it happen to go near where we want*

says Toffoli [43].

In view of the above discussion we may consider Ω as a crude model for the Universe. The first objection, a strong one, pertains to the *discrete* nature of the model, which contrasts with current models allowing the space to go down to infinitesimal distances.[10] Time and space are considered continuous! Let's follow Paul Davies argumentation which starts with the observation that the continuity of time and space are just *working hypotheses* and the current state of art in physics cannot assure us that indeed this is the case: Time might be discrete, at some extremely small scale of size.[11] The situation could be very similar to the well-known example on movies: the film appears to the human eye to be *continuous*, but in reality it is **discrete**.[12] The reason of this

[9] Recall that Ω and K encode the same amount of information, but differ in the structuralization of this information!

[10] See also Holland [23, 24].

[11] For instance, experiments in physics cannot measure intervals of time less than 10^{-26} seconds.

[12] It advances one frame at a time.

apparent paradox comes from the fact that human eyes cannot resolve the short time intervals between frames. For a detailed discussion we quote Mellor [31], Prigogine [36].

Bennett [4] raised the following important question: *Is self-organisation an asymptotical qualitative phenomenon (like phase transitions)?* More precisely, are there physically reasonable models in which complexity – appropriately defined – increases without bound in time and space? The answer is "yes" for the case of the model Ω. There is a sense in which Ω and the underlying Chaitin complexity H reflect a well-defined qualitative property of infinite systems.

Adopting this model we apparently contradict a large part of the knowledge provided by Science, i.e. the successful application of mathematics to make predictions expressed by means of the laws of physics.[13]

Where do the physical laws come from? Why do they operate universally and unfailingly?

Nobody seems to have reasonable answers to these questions. The most we can do is to explain that the Hypothesis of Order is supported by our daily observations: the rhythm of day and night, the pattern of planetary motion, the regular ticking of clocks. However, there is a limit to this perceived order: the vagaries of weather, the devastation of earthquakes or the fall of meteorites are perceived as fortuitous.

How are we to reconcile these seemingly random processes with the supposed order? There are, at least, two ways. The most common one starts by observing that even if the individual chance events may give the impression of lawlessness, disorderly processes may still have deep (statistical) regularities. This is the case for most interpretations of quantum mechanics.

It is not too hard to notice some limits of this kind of explanation. It is common sense to say that "casino managers put as much faith in the laws of chance as engineers put in the laws of physics". We may ask: How can the same physical process obey two contradictory laws, the laws of chance and the laws of physics? As an example consider the *spin of a roulette wheel.*

There is a second, "symmetric" approach, which is mainly suggested by algorithmic information theory. As our direct information refers to finite experiments, it is not out of question to discover *local rules*, functioning on large, but finite scales, even if the global behaviour of the process is truly random.[14] But, to perceive this *global randomness* we have to have access to **infinity**, which is not physically possible!

It is important to notice that, consistently with our common experience, facing global randomness does not imply the impossibility of making predictions.

[13] It appears that everything is based on the assumption that the *physical Universe is ordered and rational.*

[14] Recall that in a random sequence every string – of any length – appears infinitely many times. So, in such a random sequence the first billion of digits may be exactly the first digits of the expansion of π!

Space scientists can pinpoint and predict planetary locations and velocities "well enough" to plan missions years in advance. Astronomers can predict solar or lunar eclipses centuries before their occurrences. We have to be aware that all these results – as superb as may be – are only true within a certain degree of precision. Of course, in the process of solving equations, say of motion, small errors accumulate, making the predictions less reliable as the time gets longer. We face the limits of our methods!

Why are our tools so imperfect?

The reason may be found in the fact that a random sequence cannot be "computed", it is only possible to approximate it very crudely. Algorithmic information theory gives researchers an appreciation of how little complexity in a system is needed to produce extremely complicated phenomena and how difficult is it to describe the Universe. This statement is consistent with the main conclusions of Svozil [42]:

> **Chaos in physics corresponds to randomness in mathematics.**
>
> **Randomness in physics may correspond to uncomputability in mathematics.**

Nowdays the lack of computability doesn't look strange at all due to the striking results obtained by Pour-El and Richards [35] (for an ample discussion see Penrose's book [33]) for the wave equation. They have proven that even though solutions of the wave equation behave *deterministically*, in the most common sense, there exist *computable initial data*[15] with the strange property that for a later computable time the determined value of the field is *non-computable.* Thus, we get a certain possibility that the *equations – of a possible field theory – give rise to a non-computable evolution.* In the same spirit, da Costa and Doria [16] have proven that the problem whether a given Hamiltonian can be integrated by quadratures is undecidable; their approach led to an Incompleteness Theorem for Hamiltonian mechanics.

Perhaps the most important relation between randomness and the Universe is provided by the quantum mechanics. Let us examine it very briefly. This theory pertains to events involving atoms and particles smaller than atoms, events such as collisions or the emission of radiation. In all these situations the theory is able to tell what will **probably** happen and not what will **certainly** happen.

The classical idea of causality (i.e. the idea that the present state is the effect of an anterior state and cause of the state which is to follow) implies that in order to predict the future we must know the present, with enough precision.[16] Not so, in quantum mechanics! For quantum events this is impossible in view of

[15] More precisely, the initial condition is C^1 (i.e. continuous, with continuous derivative), but not twice differentiable.

[16] In company with Laplace: ... *a thing cannot occur without a cause which produces it.*

Heisenberg's Uncertainty Principle. According to this principle it is impossible to measure both the *position* and the *momentum* of a particle accurately at the same time. Worse than this, there exists an *absolute limit* on the product of these inaccuracies expressed by the formula

$$\Delta p.\Delta q \geq h,$$

where q, p refer, respectively, to the position and momentum and $\Delta p, \Delta q$ to the corresponding inaccuracies. In other words, the more accurately the position q is measured, the less accurately can the momentum p be determined, and *vice versa.* The measurement with an infinity of precision is ruled out: If the position were measured to infinite precision, then the momentum would become completely uncertain and if the momentum is measured exactly, then the particle's location is uncertain.

To get some concrete feeling let us assume that the position of an electron is measured within the accuracy of 10^{-9} m; then the momentum would become so uncertain that one could not expect that, one second later, the electron would be closer than 100 kilometres away (see Penrose [33], p. 248). Borel [8] proved that if a mass of one gram is displaced through a distance of one centimetre on a star at the distance of Sirius it would influence the magnitude of gravitation on the Earth by a factor of only 10^{-100}. More recently, it has been proven that the presence/absence of an electron at a distance of 10^{10} light years would affect the gravitational force at the Earth by an amount that could change the angles of molecular trajectories by as much as one radian after about 56 collisions.

Einstein was very upset about this situation! His opposition to the probabilistic aspect of quantum mechanics is very well-known:

> *Quantum mechanics is very impressive. But an inner voice tells me that it is not yet the real thing. The theory produces a good deal but hardly brings us closer to the secret of the Old One. I am at all events convinced that* He *does not play dice.*[17]

It is important to note that Einstein was not questioning the use of probabilities in quantum theory (as a measure of temporary ignorance or error), but the implication that the individual microscopic events are themselves indeterminate, unpredictable, random.

Quantum randomness is precisely the kind of randomness usually considered in probability theory. It is a "global" randomness, in the sense that it addresses processes (e.g. measuring the diagonal polarization of a horizontally-polarized photon) and not individuals (it does not allow one to call a particular measurement random). Algorithmic information theory succeeds in formalizing the notion of an individual random sequence using a self-delimiting universal computer. However, we have to pay a price: If a more powerful computer is used –

[17] From his reply to one of Niels Bohr's letters in 1926, quoted from Penrose [33], p. 280.

for instance, a computer supplied with an oracle for the halting problem – then the definition changes. Moreover, there is no hope of obtaining a "completely invariant" definition of random sequences because of Berry's paradox. In Bennett's words [6]: *The only escape is to conclude that the notion of definability or nameability cannot be completely formalized, while retaining its usual meaning.*

6 The Hypothesis of Randomness

Following the discussion in the preceding section we would like to suggest replacing the Hypothesis of Order by its opposite, the **Hypothesis of Randomness: The Universe is random.**

First let us note that the ancient Greeks and Romans would not have objected to the idea that the Universe is essentially governed by chance – in fact they made their Gods play dice quite literally, by throwing dice in their temples, to see the will of Gods; the Emperor Claudius even wrote a book on the art of winning at dice.[18]

Poincaré may have suspected and even understood the chaotic nature of our living Universe. More than 85 years ago he wrote:

> *If we knew exactly the laws of nature and the situation of the universe at the initial moment, we could predict exactly the situation of that universe at a succeeding moment. But even if it were the case that the natural law no longer had any secret for us, we could still only know the initial situation* approximately. *If that enabled us to predict the succeeding situation with the* same approximation, *that is all we require, that [it] is governed by the laws. But it is not always so; it may happen that small differences in the initial conditions produce very great ones in the final phenomena. A small error in the former will produce an enormous error in the latter. Prediction becomes impossible, and we have the fortuitous phenomenon.*

Of course, one may discuss this hypothesis and appreciate its value (if any) by its fruitfulness. We may observe, following Davies [17], that "random" events in the Universe "may not be random at all". As we previously noticed, randomness is not algorithmically testable: A sequence of quantum mechanical measurements appears random, but we cannot *prove* this! Most of those discussions

[18] However, from the point of view of Christianity, playing dice with God was definitely a pagan practice – it violates the first commandment. St. Augustine is reported to say that nothing happens by chance, because everything is controlled by the will of God.

have been focussed on quantum indeterminism, which in the light of algorithmic information theory is a severe limitation. Randomness is omnipresent in the Universe, and by no means is it a mark of the microscopic Universe!

Gödel [20] discusses the essence of *time*. Under the influence of Einstein – during their stay at the Institute of Advanced Study in Princeton – Gödel produced some new solutions for Einstein's gravitational field equations. His main conclusion is that the lapse of time might be unreal and illusory.[19] In his own words:

> *It seems that one obtains an unequivocal proof for the view of those philosophers who, like Parmenides and Kant, and the modern idealists, deny the objectivity of change and consider change as an illusion or an appearance due to our special perception.*

His model describes a rotating Universe giving rise to space-time trajectories that loop back upon themselves. *Time is not a straight linear sequence of events* – as it is commonly suggested by the arrow – *but a curving line.* There is no absolute space; matter has inertia only relative to other matter in the Universe.

> *By making a round trip on a rocket ship in a sufficiently wide curve, it is possible in these worlds to travel into any region of the past, present, and future, and back again.*

Loschmidt's Paradox (particles obeying time reversible equations of motion can exhibit time irreversible behaviour) is particularly relevant in this context. The solutions suggested rely on probabilistic arguments (the low probability of the initial conditions gives rise to the time reversed behaviour; Prigogine [36], Holian, Hoover and Posh [25]) or on "time asymmetric physical laws" (Penrose [32]). Sulis [40] made an interesting suggestion: the paradox might be related to the failure in distinguishing between local and global properties of physical systems.[20] The microscopic equations are local equations of motion whereas the arrow time is a global aspect of the Universe.

Many other fundamental questions like:

- Is the existence of God an axiom or a theorem?
- Is God omnipotent?
- Is God rational? Do the laws of physics contradict the laws of chance?

[19]Karl Svozil pointed out in [41] that "Gödel himself looked into celestial data for support of his solutions to the Einstein equations; physicists today tend to believe that the matter distribution of the universe rules out these solutions, but one never knows..."

[20]Chaitin's Ω number reflects this distinction.

can be discussed from this new point of view. We hope to pursue this analysis in another paper.

Finally, let us go back to the largely based conviction that the future is determined by the present, and therefore a careful study of the present allows us to unveil the future. As it is clear, we do not subscribe to the first part of the statement, but we claim that our working hypothesis is consistent with the second part of it.

7 Final Remarks

The present paper can be viewed as an introductory one. We hope to return to a more detailed analysis of some of the questions mentioned in the preceding section. A really dramatic consequence of the properties of the number Ω is that we can exhibit a specific (exponential) Diophantine equation $P(i, X_1, \ldots, X_m) = 0$ having, for each fixed value of the parameter i, infinitely many solutions in $X_1, \ldots, X_m$ if and only if the ith bit of Ω equals 1. Any formal theory can answer this question for finitely many values of i only. No matter how many additional answers we learn, for instance, by experimental methods or just flipping a coin, this won't help us in any way as regards the remaining infinitely many values of i. As regards these values, mathematical reasoning is helpless, and a mathematician is not better off than a gambler flipping a coin. This holds in spite of the fact that we are dealing with basic arithmetic.

Adopting the Randomness Hypothesis may have some methodological implications as well. For instance, it gives some evidence toward the new tendency in the philosophy of mathematics called **experimental mathematics**;[21] see more in the recent intriguing discussions in Chaitin [15],[22] Jaffe and Quinn [26].[23]

---

[21] There is a journal entitled *Journal of Experimental Mathematics.*

[22] "Perhaps number theory should be pursued more openly in the spirit of experimental science! To prove more, one must sometimes assume more." [12], p. 160. In contrast, Levin [29] expresses a pessimistic view: "Our thesis contradicts the conviction of some mathematicians that the truth of any true proposition can be established in the course of the development of science by means of informal methods (it is impossible to do so by formal methods, due to Gödel's theorem."

[23] One distinguishes between "theoretical mathematics" (referring to the speculative and intuitive work) and "rigorous mathematics" (the proof-oriented phase) in an attempt to build a framework assuring a positive role for speculation and experiment.

References

[1] J. Barrow. *Pi in the Sky*, Clarendon Press, Oxford, 1992.

[2] J. Barrow, F. J. Tipler. *The Anthropic Cosmological Principle*, Oxford University Press, Oxford, 1986.

[3] C. H. Bennett. The thermodynamics of computation — a review, *Internat. J. Theoret. Physics* 21(1982), 905-940.

[4] C. H. Bennett. Logical depth and physical complexity, in R. Herken (ed.). *The Universal Turing Machine. A Half-Century Survey*, Oxford University Press, Oxford, 1988, 227-258.

[5] C. H. Bennett. Dissipation, information, computational complexity and the definition of organization, in D. Pines (ed.). *Emerging Syntheses in Science*, Addison-Wesley, Boston, 1987, 297-313.

[6] C. H. Bennett. E-mail to C. Calude, April 25, 1993.

[7] C. H. Bennett, M. Gardner. The random number omega bids fair to hold the mysteries of the universe, *Scientific American* 241(1979), 20-34.

[8] É. Borel. *Le hasard*, Alcan, Paris, 1928.

[9] C. Calude. *Theories of Computational Complexity*, North-Holland, Amsterdam, New York, Oxford, Tokyo, 1988.

[10] C. Calude. Meanings and texts: An algorithmic metaphor, in M.Balat, J.Deledalle-Rhodes (eds.). *Signs of Humanity*, Mouton de Gruyter, 1992, 95-97.

[11] C. Calude. *Information and Randomness — An Algorithmic Perspective*, Springer-Verlag. [in press]

[12] G. J. Chaitin. *Algorithmic Information Theory*, Cambridge University Press, Cambridge,1987. (third printing 1990)

[13] G. J. Chaitin. *Information, Randomness and Incompleteness, Papers on Algorithmic Information Theory*, World Scientific, Singapore, New Jersey, Hong Kong, 1987.(2nd ed., 1990)

[14] G. J. Chaitin. *Information-Theoretic Incompleteness*, World Scientific, Singapore, New Jersey, Hong Kong, 1992.

[15] G. J. Chaitin. Randomness in arithmetic and the decline and fall of reductionism in pure mathematics, *EATCS Bull.* 50(1993), 314-328.

[16] N. C. A. da Costa, F. A. Doria. Undecidability and incompleteness in classical mechanics, *Internat. J. Theoret. Physics* 30(1991), 1041-1073.

[17] P. Davies. *The Mind of God, Science and the Search for Ultimate Meaning*, Penguin Books, London, 1992.

[18] P. Davies, J. Gribbin. *The Matter Myth. Beyond Chaos and Complexity*, Penguin Books, London, 1992.

[19] S. Feferman, J. Dawson, Jr., S. C. Kleene, G. H. Moore, R. M. Solovay, J. van Heijenoort (eds.). *Kurt Gödel Collected Works*, Volume II, Oxford University Press, New York, Oxford, 1990.

[20] K. Gödel. An example of a new type of cosmological solutions of Einstein's field equations of gravitation, *Reviews of Modern Physics* 21(1949), 447-450. (Reprinted in [19], p. 190-198.)

[21] N. A. Haji. *Spontaneous Emergence of Self-Replicating Systems*, Master of Science Thesis, University of Western Ontario, London, Canada, 1989.

[22] S. W. Hawking. *A Brief History of Time, From the Big Bang to Black Holes*, Bantam Press, London, New York, Auckland, 1988.

[23] J. H. Holland. Studies of the spontaneous emergence of self-replicating systems using cellular automata and formal grammars, in A. Lindenmeyer and G. Rozenberg (eds.). *Automata, Languages and Development*, North-Holland, Amsterdam, 1976, 385-404.

[24] J. H. Holland. *Adaptation in Natural and Artificial Systems. An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, The University of Michigan Press, Ann Arbor, 1975.

[25] B. Holian, W. G. Hoover, H. A. Posch. *Phys. Rev. Lett.* 59(1987), 10-13.

[26] A. Jaffe, F. Quinn. "Theoretical mathematics": toward a cultural synthesis of mathematics and theoretical physics, *Bull. Amer. Math. Soc.* 29(1993), 1-13.

[27] D. W. Juedes, J. I. Lathrop, J. H. Lutz. Computational depth and reducibility, *Theoret. Comput. Sci.* [in press]

[28] P. S. Laplace. *A Philosophical Essay on Probability Theories*, Dover, New York, 1951.

[29] L. A. Levin. Randomness conservation inequalities: information and independence in mathematical theories, *Problems Inform. Transmission* 10(1974), 206-210.

[30] M. Li, P. M. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*, Springer-Verlag, 1993.

[31] D. H. Mellor. *Real Time*, Cambridge University Press, Cambridge, 1985.

[32] R. Penrose. Singularities and time-asymmetry, in S. W. Hawking and W. Israel (eds.). *General Relativity: An Einstein Centenary Survey*, Cambridge University Press, Cambridge, 1979, 581-638.

[33] R. Penrose. *The Emperor's New Mind. Concerning Computers, Minds, and the Laws of Physics*, Oxford University Press, Oxford, New York, Melbourne, 1989.

[34] R. Penrose. Précis of *The Emperor's New Mind. Concerning Computers, Minds, and the Laws of Physics* (together with responses by critics and a reply by the author) *Behavioural and Brain Sciences* 13(1990), 643-705.

[35] M. Pour-El, I.Richards. *Computability in Analysis and Physics*, Springer-Verlag, Berlin, Heidelberg, New York, 1989.

[36] I. Prigogine. *From Being to Becoming*, G. H. Freeman, San Francisco, 1980.

[37] K. A. Ribet. Wiles proves Taniyama's Conjecture: Fermat's Last Theorem follows, *Notices Amer. Math. Soc.* 40 (1993), 575-576.

[38] G. Rozenberg, A. Salomaa. *Cornerstones of Undecidability*, Prentice Hall. [in press]

[39] A. Salomaa. *Computation and Automata*, Cambridge University Press, Cambridge, 1985.

[40] W. H. Sulis. *Order Automata*, Ph. D. Thesis, The University of Western Ontario, London, Canada, 1989.

[41] K. Svozil. E-mail to C.Calude, June 14, 1993.

[42] K. Svozil. *Randomness & Undecidability in Physics*, World Scientific, Singapore, New Jersey, Hong Kong, 1993. [in press]

[43] T. Toffoli. Physics and computation, *Internat. J. Theoret. Physics* 21(1982), 165-175.

[44] J. von Neumann. *Theory of Self-Reproducing Automata*, Edited and Complemented by A.W.Burks, University of Illinois Press, Urbana, 1966.