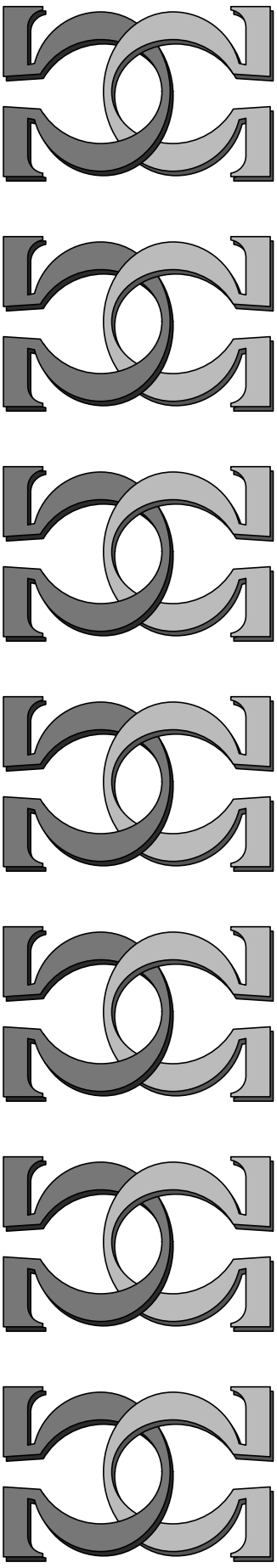

CDMTCS
Research
Report
Series

Who Is Afraid of Randomness?

C. S. Calude
University of Auckland, New Zealand

CDMTCS-143
September 2000

Centre for Discrete Mathematics and
Theoretical Computer Science

Who Is Afraid of Randomness?*

Cristian S. Calude
Department of Computer Science
University of Auckland
Private Bag 92019, Auckland
New Zealand
E-mail: cristian@cs.auckland.ac.nz

We finally understand. Things are the way they are because they can't be otherwise.
Ivar Ekeland

1 Introduction

Randomness–the mark of anxiety, the cause of disarray or misfortune, the cure for boring repetitiveness, is, like it or not, one of the most powerful driving forces of life. Is it bad? Is it good? The struggle with uncertainty and risk caused by natural disasters, market downturns or terrorism is balanced by the role played by randomness in generating diversity and innovation, in allowing complicated structures to emerge through the exploitation of serendipitous accidents.

To many minds any discussion about randomness is purely academic, just another mathematical or philosophical pedantry. False! Randomness could be a matter of life or death, as in the case of Sudden Infant Death Syndrome (SIDS), a merciless child-killer.

The present paper describes some difficulties regarding the mathematical modelling of randomness, contrasts silicon-computer generated pseudorandom bits with quantum-computer "random" bits, succinctly presents the algorithmic definition of randomness proposed by algorithmic information theory and the relations between randomness and (logical) incompleteness, briefly presents some applications of algorithmic randomness in physics and finishes by advocating "experimental mathematics", a quasi-empirical, more pragmatic manner view of mathematics.

2 Various Meanings

The term random has a variety of meanings. For example, in mathematics a 'random object' usually means a typical object, that is, one that does not seem to have any particular structure or properties that single it out among the other objects of its kind. In computer science 'random access memory' (RAM) refers to the place in a computer where the operating system, application programs, and data in current use are kept so that they can be quickly reached by the computer's processor.[1]

Collins English Dictionary[2] lists four meanings for random: 1) lacking any definite plan or prearranged order; haphazard, 2) having a value which cannot be deterministic but only described probabilistically; chosen without regard to any characteristics of the individual members of the population so that each has an equal chance of being selected, 3) bank (chiefly Brit.), 4) in a purposeless fashion; not

*Draft paper to be discussed at the Millennium Symposium "Defining the Science of Stochastics", Würzburg, Germany, October, 2000.

[1]The term suggests that any storage location can be accessed directly. Originally, the term distinguished regular core memory from offline memory, usually on magnetic tape in which an item of data could only be sequentially accessed. Perhaps it should have been called "nonsequential memory" because RAM access is hardly random. Note that other forms of storage such as the hard disk and CD-ROM are also accessed directly but the term random access is not used to these forms of storage.

[2]Third edition, Harper Collins, 1991.

following any prearranged order. Its origin is traced to the French *randon* (from *randir*, to gallop) and German *rinnan*, to run.

3 Is Randomness Simple?

> *I am convinced that the vast majority of my readers, and in fact the vast majority of scientists and even nonscientists, are convinced that they know what 'random' is. A toss of a coin is random; so is a mutation, and so is the emission of an alpha particle.... Simple, isn't it?*

wrote Kac in a famous paper [30]. Well, no! Kac knew very well that randomness could be called many things, but not simple, and in fact his essay shows that randomness is complicated, and it can be described in more than one way, even by mathematicians and scientists.

We will illustrate this with a simple, almost trivial, example. Marilyn vos Savant[3] is (in)famous for questioning Wiles' proof of Fermat's Last Theorem and Einstein's relativity theory in her controversial book [58]. Vos Savant runs every Sunday the column *Ask Marilyn* in the *Parade Magazine* (New York). In the 9 September 1990 column she described the following game: you are on a show and you are given the choice of three doors, behind which there are two goats and a flashy car. You choose one door, say door 3, and the host, who knows where the car is, opens *another door*, behind which there is a goat. She then gives you the choice of remaining with the first choice, i.e. door 3, or switching to another door (door 1 or door 2).[4] What should you do?

Vos Savant advises to change your mind, to switch doors, arguing that with the first choice you have only one-third chance of winning, but with the second choice you are doubling the odds to two-thirds. Her advice has generated instantly a prompt reaction of disagreement from the public![5] Even the famous mathematician P. Erdös is reported by Hoffman [28] (p. 253) to have said in the first instance: 'It should make no difference'.

The answer is rather simple as the following brute force enumeration of possibilities shows: if you choose to stick with door 3, then you get one-third chance of winning

Door 1	**Door 2**	**Door 3**	**Outcome**
car	goat	goat	loose
goat	car	goat	loose
goat	goat	car	win

Table 1. Stick with your first choice

but if you change your mind you get two-thirds:

Door 1	**Door 2**	**Door 3**	**Outcome**
car	goat	goat	win
goat	car	goat	win
goat	goat	car	loose

Table 2. Change your mind

A minor modification in the game reveals the conundrum: assume that after the first door was open a person no participating in the show, so not knowing what door was originally chosen, is asked to pick up one of the remaining two unopened doors. In this case chances that she makes a right guess are fifty-fifty, and the reason comes from her being disadvantaged of not knowing the original choice. When evidence clashes so violently with intuition people, even experts, are shaken.

The following Mathematica code simulates the game and computes the odds for both strategies:

[3] She thinks of herself as the person with the highest recorded IQ (an amazing 228 featured on page 26 of the 1989 edition of the "Guinness Book of World Records Hall of Fame").

[4] Actually, this was the Monty Hall dilemma faced by guests on the classic American TV game show "Let's Make a Deal".

[5] Many letters came from professional mathematicians and statisticians; see also the web site "Marilyn is Wrong!" at `http://www.wiskit.com/marilyn/`.

```
Myproced[ ] := (
      Montylist = {0, 0, 0};
      n = Random[Integer, {1, 3}];
      Montylist = Delete[Insert[Montylist, 1, n], 4];
      c = Random[Integer, {1, 3}]; k = k + Extract[Montylist, {c}];
      Montylist = Delete[Montylist, c];
      s = s + Extract[Montylist, {1}] + Extract[Montylist, {2}]);
Clear[Montylist, k, s, m, n];
k = 0; s = 0; m = 1;
Do[Myproced[ ], {m, 10000000}];
{k, s} // TableForm
```

Running it on various samples has produced the results in the following table, results which are consistent with the above discussion.

No of simulations/strategy	10	100	1,000	10,000	1,000,000	10,000,000
Keep the door	30	321	3,326	33,413	337,700	3,333,764
Change the door	70	679	6,6674	66,587	666,300	6,666,236

Table 3. Simulation results

4 More Difficulties

Suppose that one is watching a simple pendulum swing back and forth, recording 0 if it swings clockwise at a given instant and 1 if it swings counterclockwise. Suppose further that after some time the record looks as follows:

$$10101010101010101010101010101010.$$

At this point one would like to deduce a "theory" explaining the experiment. The "theory" should account for the data presently available and make "predictions" about future observations. How should one proceed? It is obvious that there are many "theories" that one could write-down for the given data, for example:

$$\begin{aligned}
&10101010101010101010101010101000000000000000000000000000000\ldots\\
&10101010101010101010101010101011111111111111111111111111111\ldots\\
&10101010101010101010101010101000100100100100100100100100100\ldots\\
&10101010101010101010101010101000011100011100011100011100011\ldots\\
&10101010101010101010101010101010111011100101101000110001111\ldots
\end{aligned}$$

Consistently with the requirements formulated above, each "theory" starts with the experimental data (which is finite) and continues with "predictions" about how the system future (which is potentially infinite). The results of the experiment have a simple pattern, always 01's, so probably the best prediction is that the system will continue to produce 01's for ever. Is there any rational, objective way of deciding among various possible "theories" that does not rely only upon intuition? Occam's razor states that the "best" theory is the "simplest" theory. Now the question becomes: What is a "simple theory"? For Solomonoff [52] the "simplest theory" is the one with the shortest length, i.e. the one printed by a shortest length computer program. First we have to produce the experimental data; then, we have to "guess" a continuation. For example, the following program will account for the first sequence:

```
PRINT  10101010101010101010101010101010, PRINT  0.
```

Can we do it better? Given the regularity of the record a considerable shorter program can be written to produce it, namely the program "`PRINT 01 16 times`". This program can be used to print out the first four "theories" above:

```
PRINT 10 16 times, PRINT 0
PRINT 10 16 times, PRINT 1
```

```
PRINT 10 16 times, PRINT 001
PRINT 10
PRINT 10 16 times, PRINT 000111
```

The program "`PRINT 01 16 times`" can be generalised to a "law" expressed by the program "`PRINT 01 X times`". Note that the length of printouts predicted by this program grows much more rapidly than the length of itself. Can one write a simple program to print the fifth and last "theory"? In this case the continuation of the experiment does not follow an obvious pattern ... or maybe there is no such pattern. To understand this "theory" we will follow Chaitin [18] who was interested in defining the "complexity" of finite binary strings. A typical question motivating this approach is: Are the first one million of digits of the binary expansion of the number π less complex than a string produced by flipping a fair coin one million times?[6] Chaitin defined the complexity of a finite binary string as the size of the smallest program which calculates it. If the string can be compressed into a very short program then one would conclude that the string has a pattern, that it follows a law, that it is simple; if the string cannot be compressed at all, then it is maximally complex or random. Let's test this idea on some examples. Suppose that someone claims to have tossed a fair coin 64 times and the result is:

$$x = 0101010101010101010101010101010101010101010101010101010101010101$$

Then, the experiment is repeated and the result is:

$$y = 1010101010010010101001011011000110001100111100110011111000011011$$

Almost everyone would be surprised or suspicious to see x, but the string y probably would be acceptable. Why? Here is typical flawed explanation: the probability of x is extraordinarily small, i.e. 2^{-64}, so it is unreasonable to believe that x has been actually produced by a real experiment. However, from a probabilistic point of view there is nothing special about x: all of the 2^{64} possible strings of length 64 have identical probabilities of appearance, i.e. 2^{-64}. The difference between x and y *is not* probabilistic, but *structural*: x is ordered, but there is no apparent pattern in y.[7] Laplace [31], pp.16-17, was, in a sense, aware of the above difficulty:

> *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.*

Instead of computing probabilities of specific strings let's instead discuss the "typicalness" of some strings with respect to some particular stochastic processes. For the process of flipping a fair coin, incompressible strings are typical, and highly compressible strings are atypical. And, because the number of highly compressible strings (of a given length) is *small*, the occurrence of such a string is *extraordinary*: our surprise at seeing x is explained.

Of course, the above explanation is informal, and a lot needs to be done to turn it into a rigorous theory. Before presenting some technical details let indulge ourselves in a simple counting analysis. A string of length n will be said to be c–incompressible if its compressed length is greater than or equal to $n - c$. For example, the 16–incompressible strings of length 64 are exactly the strings that can be compressed to a length of 48 or larger. Note that every $(n+1)$–incompressible string is n–incompressible, so every 5–incompressible string is 4–incompressible. Based on the fact that the number of strings of length n is 2^n, it turns out that at least half of all the strings of every length are 1–incompressible, at least $\frac{3}{4}$ are 2–incompressible, at least $\frac{7}{8}$ are 3–incompressible, so on. In general, at least $1 - \frac{1}{2^c}$ of all strings of length n are c–incompressible. For example, about 99.9% of all strings of length 64 cannot be compressed by more than 16% and about 99.99999998% of these strings cannot be compressed by more than 50%.

Note that a similar analysis can be done for the notion of *entropy* assigned by Shannon [49] to an ensemble of possible messages. In case all messages are equally probable, the entropy gives the number of bits needed to count all possibilities, expressing the fact that any message in the ensemble can be

[6] Record 1 for heads and 0 for tails.

[7] Of course, we may argue that the presence of the pattern 01 in x has no significance at all because a) the number of tosses is relatively small, b) finding patterns and meanings is just a human subjective predilection.

communicated using that number of bits. However, the entropy *says nothing about the number of bits needed to transmit any individual message in the ensemble.*

A source of these difficulties comes from the fact that books on probability theory do not even attempt to define randomness. In Beltrami [3] words:

> *The subject of probability begins by assuming that some mechanism of uncertainty is at work giving rise to what is called randomness, but it is not necessary to distinguish between chance that occurs because of some hidden order that may exist and chance that is the result of blind lawlessness. This mechanism, figuratively speaking, churns out a succession of events, each individually unpredictable, or it conspires to produce an unforeseeable outcome each time a large ensemble of possibilities is sampled.*

5 Two More Negative Results

In an extreme sense there is no such notions as "true randomness", "genuine randomness". Confining again to binary sequences, we cite a theorem by van der Waerden which indicates a nontrivial regularity shared by all sequences:

Theorem 1 *In every binary sequence at least one of the two symbols must occur in arithmetical progressions of every length.*

Note the nonconstructive nature of the proof of Theorem 1: there is no algorithm which will tell in a finite amount of time which alternative is true: 0 occurs in arithmetical progressions of every length or 1 occurs in arithmetical progressions of every length.

Even more disturbing, the answer to the question "how random is a coin toss?" seems to be: not too much. Indeed, following Ford [24] and Jaynes [29], let's look at the mechanics of the toss. The ellipsoid of inertia of a thin disc is an oblate spheroid of eccentricity $\sqrt{2}/2$. The displacement does not affect the symmetry of the spheroid, hence the polhodes circles remain concentric with the axis of the coin (cf. Routh [43]), so the character of the tumbling motion of the coin while in flight is exactly the same for a biased as an unbiased coin. There is however a subtle difference: for the biased coin the center of gravity (not the geometrical center) describes the parabolic trajectory. According to the law of conservation of angular momentum, the coin maintains a fixed direction in space, but not a fixed velocity (hence tumbling "looks" chaotic). The direction is determined by the twist you give the coin at launching, but doesn't depend on whether the coin is or isn't biased: the coin will show the same face when viewed from that direction.[8]

Based on this observations you can "cheat" with extremely good results at the usual coin-toss game: toss the coin with a twist so that the unit vector passing through the coin along its axis with its point on the "heads" side makes an acute angle with fixed direction maintained by the angular momentum, and catch it in a plan normal to that direction. On successive tosses you can let the magnitude of the angular momentum and unit vectors vary subject to the above constraints, the tumbling will appear chaotic.

The above scenario may seem too simple. Still, Jaynes [29], p. 1005, makes the following interesting remark:

> *While accepting this criticism, we cannot suppress the obvious comment: scanning the literature of probability theory, isn't it curious that so many mathematicians, usually far more careful than physicists to list all the qualifications needed to make a statement correct, should have failed to see the need for any qualifications here?*

What about tossing a quantum mechanical coin? Before answering this question let's pause a minute to look at pseudorandom bits.

[8] With the exception when the direction is perpendicular to the axis of the coin when no face will be visible.

6 Computer-Generated Pseudorandom Bits

All modern computers have "pseudorandom number generators" capable to produce "pseudorandom numbers". A variety of clever algorithms have been developed which generate sequences of numbers which pass many statistical tests used to distinguish "random" sequences from those containing some pattern or internal order. Any computer can only feign randomness; thinking otherwise is not only wrong, but as von Neumann said,

> *Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin.*

Scientists all over the world use millions and millions of computer-generated pseudorandom numbers with realising that although these numbers pass a good deal of statistical tests for randomness they are not genuinely random; more troublesome, very little is known about how nonrandom these numbers are and in which ways these nonrandomness might affect the results.[9]

> *I'm worried about the way people use probability theory models. ... It seems to me there is a lot of automated, nonthinking use of probability theory. People write down probabilistic models and often assume standard postulates of randomness. But frequently their assumptions are just crazy and their conclusions are meaningless.*

said P. Diaconis about 15 years ago (reported by Kolata [35]). Today things are not really better and using pseudorandom numbers as if they were random is still dangerous.

7 Quantum Random Bits

Randomness is at the very heart of quantum physics. When a physical state that is in a superposition of states is measured, then it collapses into one of its possible states in a completely unpredictable way–we can only evaluate the probability of obtaining various possible outcomes. An extreme view is to claim with Peres [39] that

> *in a strict sense quantum theory is a set of rules allowing the computation of probabilities for the outcomes of tests which follow specific preparations.*

According to Milburn [38], p.1, a fundamental quantum principle is

> *physical reality is irreducible random.*[10]

Quantum mechanics "seems"[11] capable to produce, with probability one, truly random strings. To describe a way to do it we need some elementary facts about qubits.

A classical bit (e.g. the position of gear teeth in Babbage's differential engine, a memory element or wire carrying a binary signal, in contemporary machines) is a system comprising many atoms. Typically, the system is described by one or more continuous parameters, for example, voltage. Such a parameter is used to separate the space into two well-defined regions chosen to represent 0 and 1. Manufacturing imperfections, local perturbations may affect, so signals are periodically restored toward these regions to prevent them from drifting away. An n-bit register of memory can exist in any of 2^n logical states, from $00\ldots0$ (n zeros) to $11\ldots1$ (n ones).

A quantum event in which we have two possible mutually exclusive outcomes is the elementary act of observation: all knowledge of the physical world is based upon such acts. An elementary act of observation is simultaneously like a coin-toss and not like a coin-toss. The information derived from an elementary act of observation is no more than a single bit, but *there is more on it than that.* To

[9] G. Marsaglia found simple randomness tests not passed by typical computer-generated pseudorandom numbers; today many algorithms produce pseudorandom numbers passing Marsaglia's tests, but it is not unlikely that other simple tests will be constructed that show up the new pseudorandom generators. Other examples have been discussed by Maddox [33].

[10] HotBits is an web resource (see `http://www.fourmilab.ch/hotbits/`) that claims to produce "genuine random numbers" generated by a process fundamentally governed by the inherent uncertainty in the quantum mechanical laws of nature. HotBits are generated by timing successive pairs of radioactive decays detected by a Geiger-Müller tube interfaced to a computer.

[11] We write "seems" because this is a postulate, not a fact deduced from the axioms of any model of quantum mechanics.

mark this difference Schumaker [48] has coined the name qubit. A quantum bit, qubit, is typically a microscopic system, such as an atom or nuclear spin or polarized photon. For example, the state of a spin-$\frac{1}{2}$ particle, when measured, is always found to be in one of two possible states, represented as

$$|+\frac{1}{2}\rangle \text{ (spin-up) or } |-\frac{1}{2}\rangle \text{ (spin-down).}$$

Due to quantization (see more in Calude and Păun [15]) one can use one spin state to represent 0, and the other spin state to represent 1. There is nothing special about spin systems–any 2-state quantum system can be equally used to represent 0 and 1. What is really special here is the existence of a continuum of intermediate states which are superpositions of 0s and 1s. Unlike the intermediate states of a classical bit (for example, any voltages between the "standard" representations of 0 and 1) which can be distinguished from 0 and 1, but do not exist from an informational point of view, quantum intermediate states cannot be reliably distinguished, even in principle, from the basis states, but do have an informational "existence".

An n-qubit system can exist in any superposition of the form

$$\Psi = \sum_{x=00\ldots0}^{11\ldots1} c_x|x\rangle, \tag{1}$$

where c_x are (complex) numbers such that $\sum_x |c_x|^2 = 1$. The exponential "explosion" represented by formula (1) distinguishes quantum systems from classical ones: in a classical system a state is described by a number of parameters growing only linearly with the size of the system, but quantum systems may not admit such a description (because quantum states may be "entangled").

Now consider the operator

$$R_\theta = \begin{pmatrix} cos\,\theta & -sin\,\theta \\ sin\,\theta & cos\,\theta \end{pmatrix},$$

which rotates a qubit $a|0\rangle + b|1\rangle$ through an angle θ. In particular, $R_{\frac{\pi}{4}}$ transforms that state $|0\rangle$ into an equally weighted superposition of 0 and 1:

$$\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle. \tag{2}$$

So, to make a quantum device to produce random bits one needs to place a 2-state quantum system in the $|0\rangle$ state, apply the operator $R_{\frac{\pi}{4}}$ to rotate the state into the superposition (2), and the observe the superposition. The act of observation produces the collapse into either $|0\rangle$ or $|1\rangle$, with equal chances. Consequently, one can use the quantum superposition and indeterminism to simulate, with probability one, a "fair" coin toss. Random digits produced with quantum random generators of the type described above are, with probability one, free of subtle correlations that haunt classical pseudorandom number generators. Of course, the problem of producing algorithmic random strings is still open. Indeed, let's assume that we have a classical silicon computer that simulates, using a high-quality pseudo-random generator, the quantum mechanics dynamics and quantum measurement of a 2-state quantum system. The simulated world will be statistically almost identical (up to some degree) with the "real" quantum system. However, all simulated bits will be, in the long run, highly compressible. How can we be sure that the "real" quantum system is not just a superpowerful pseudorandom generator?

8 Algorithmic Randomness

The discussion in the first section suggests two properties which should be possessed by any random sequence:

1. *a random sequence should be typical,* that is it should belong to any "reasonable" majority;
2. *a random sequence should be chaotic,* that is no simple law should be capable to produce the terms of the sequence.

To address typicalness let's isolate the set of all sequences having "all verifiable" properties that from the point of view of classical probability theory are satisfied with "probability one" with respect to the unbiased discrete probability. Let us denote by Σ the binary alphabet $\{0,1\}$ and by Σ^* the set of all binary strings. The unbiased discrete probability on Σ defined by the function $h(\{0\}) = h(\{1\}) = 2^{-1}$ induces the product measure μ on the set of all Borel subsets of the set of all binary sequences Σ^ω. If $x = x_1x_2\dots x_n$ is a string of length n, then the cylinder induced by x, $x\Sigma^\omega$, i.e. the set of all sequences starting with $x_1x_2\dots x_n$, will have the probability $\mu(x\Sigma^\omega) = 2^{-n}$. This number can be interpreted as "the probability that a sequence $\mathbf{y} = \mathbf{y_1y_2}\dots\mathbf{y_n}\dots$ has the first element $y_1 = x_1$, the second element $y_2 = x_2, \dots$, the nth element $y_n = x_n$". Independence means that the probability of an event of the form $y_i = x_i$ does not depend upon the probability of the event $y_j = x_j$. Note that every open set, i.e. a union of cylinders, is μ measurable. Finally, $S \subset \Sigma^\omega$ is a *null set* in case for every real $\varepsilon > 0$ there exists an open set which contains S and has measure less than ε. For instance, every enumerable subset of Σ^ω is a null set.

A property P of sequences is said to be *true almost everywhere (in the sense of μ)* in case the set of sequences not having the property P is a null set. The main example of such a property is the famous *Law of Large Numbers* discovered by Borel (a first version of which was known to Jakob Bernoulli around 1700):

> *For every sequence $x = x_1x_2\dots x_m\dots$ and natural number $n \geq 1$, the limit of $S_n(\mathbf{x})/\mathbf{n}$, when n tends to ∞, exists almost everywhere in the sense of μ and has the value $1/2$.*[12]

In other words, *there exists a null set $S \subset \Sigma^\omega$ such that for every $\mathbf{x}$ not in S, we have $S_n(\mathbf{x})/\mathbf{n} = \mathbf{1/2}$.*

It is clear that a sequence satisfying a property false almost everywhere with respect to μ is very "particular". Accordingly, it is tempting to say that

> *a sequence $\mathbf{x}$ is "random" if it satisfies every property true almost everywhere with respect to μ.*

Unfortunately, we may define, for every sequence $\mathbf{x}$, the property $P_{\mathbf{x}}$:

> *a sequence $\mathbf{y}$ satisfies $P_{\mathbf{x}}$ if and only if for every $n \geq 1$ there exists a natural $m \geq n$ such that $x_m \neq y_m$.*

Every $P_{\mathbf{x}}$ is an asymptotic property which is true almost everywhere with respect to μ and $\mathbf{x}$ does not have property $P_{\mathbf{x}}$. Accordingly, *no sequence can verify all properties true almost everywhere with respect to μ.* The above definition is vacuous!

Keeping in mind van der Waerden's result (see Theorem 1 and the comment following it) we are led to consider not *all* asymptotic properties true almost everywhere with respect to μ, but only a *countable set* of such properties. So, the important question becomes: *Which properties should be considered?* Clearly, the "larger" the chosen class of properties is, the "more random" will be the sequences satisfying those properties. A constructive selection seems to be suggested by both statistical practice and philosophical intuition. One such definition, suggested by Martin-Löf [36, 37], is based on randomness tests. We fix a standard recursive pairing function $\lambda k, y\ \langle k, y\rangle$ defined on $\mathbf{N} \times \Sigma^*$ with values in Σ^*; here $\mathbf{N}$ is the set of non-negative integers. For a set $A \subseteq \Sigma^*$ let $A_k = \{x \in \Sigma^* \mid \langle k, x\rangle \in A\}$. A *Martin-Löf test* is an computably enumerable (c.e.) set $A \subset \Sigma^*$ such that $\mu(A_i\Sigma^\omega) \leq 2^{-i}$, for all natural i. The set $\bigcap_{i\geq 0}(A_i\Sigma^\omega)$ is the set of all sequences which do not pass the randomness test A. With this apparatus we can say that a sequence $\mathbf{x}$ is *Martin-Löf random* if for every Martin-Löf test A, $\mathbf{x} \notin \bigcap_{\mathbf{i}\geq\mathbf{0}}(\mathbf{A_i\Sigma^\omega})$.

Martin-Löf [37] proved the existence of a *universal Martin-Löf test,* a test W with the property that for every Martin-Löf test A there is a constant c such that $A_n \subseteq W_{n+c}$, for all n. So, Martin-Löf 's definition can be rephrased as: *A sequence $\mathbf{x}$ is Martin-Löf random if and only if $\mathbf{x}$ passes a universal Martin-Löf test.* This result captures "typicality": for each Martin-Löf test A, the set $\bigcap_{i\geq 0}(A_i\Sigma^\omega)$ is constructively null, so

Theorem 2 *Constructively, with probability one (in the sense of μ), every sequence is Martin-Löf-random.*

[12] $S_n(\mathbf{x}) = \mathbf{x_1} + \mathbf{x_2} + \dots + \mathbf{x_n}$.

Hence, from the probabilistic point of view, the set of random sequences is *large*. However, from a topological point of view[13] the situation is completely different (cf. Calude and Chiţescu [12]) as Martin-Löf random sequences form a small set:

Theorem 3 *The set of Martin-Löf random sequences is constructively a first Baire category set.*

Solovay [53] proposed another measure-theoretic definition of random sequences aiming to capture typicality: a sequence $\mathbf{x}$ is *Solovay random* if for every c.e. set $A \subset \Sigma^*$ such that $\sum_{i\geq 1} \mu(A_i\Sigma^\omega) < \infty$, there exists a natural N such that for all $i > N$, $\mathbf{x} \notin A_i\Sigma^\omega$.

"Chaoticity" appears in the following two complexity-theoretic definitions (see Chaitin [17]): an infinite sequence $\mathbf{x}$ is *Schnorr random* if there is a constant c such that $H(\mathbf{x}(\mathbf{n})) > \mathbf{n} - \mathbf{c}$, for every integer $n > 0$, and, apparently the stronger definition, an infinite sequence $\mathbf{x}$ is *Chaitin random* if $\lim_{n\to\infty} H(\mathbf{x}(\mathbf{n})) - \mathbf{n} = \infty$.[14]

Finally, we present Hertling and Weihrauch topological approach to define randomness [27]. A *randomness space* is a triple (X, B, μ), where X is a topological space, $B : \mathbf{N} \to \mathbf{2^X}$ is a total numbering of a subbase of the topology of X, and μ is a measure defined on the σ-algebra generated by the topology of X.[15] Let (W_n) be a sequence of open subsets of X; a sequence (V_n) of open subsets of X is called *W–computable* if there is a c.e. set $A \subseteq \mathbf{N}$ such that $V_n = \bigcup_{\pi(n,i)\in A} W_i$ for all $n \in \mathbf{N}$.[16] Next we define $W'_i = W'(i) = \bigcap_{j\in D_{(1+i)}} W_j$, for all $i \in \mathbf{N}$, where $D : \mathbf{N} \to \{\mathbf{E} \mid \mathbf{E} \subseteq \mathbf{N} \text{ is finite}\}$ is the computable bijection defined by $D^{-1}(E) = \sum_{i\in E} 2^i$. Note that if B is a numbering of a subbase of a topology, then B' is a numbering of a base of the same topology. A *randomness test on X* is a B'–computable sequence (W_n) of open sets with $\mu(W_n) \leq 2^{-n}$, for all $n \in \mathbf{N}$. We say that an element $x \in X$ is called *Hertling–Weihrauch random* if $x \notin \bigcap_{n\in\mathbf{N}} W_n$, for every randomness test (W_n) on X.

Consider now the canonical topology on Σ^ω and the numbering B of a subbase (in fact a base) of the topology is given by $B_i = \{string_i\}\Sigma^\omega$. The general definition applies, so we get: *A sequence is Hertling–Weihrauch random if it is random in the space (Σ^ω, B, μ).*

All the above approaches lead to the same class of sequences:

Theorem 4 *Let $\mathbf{x} \in \mathbf{\Sigma^\omega}$. The following statements are equivalent:*

1. *The sequence $\mathbf{x}$ is Martin-Löf random.*
2. *The sequence $\mathbf{x}$ is Chaitin random.*
3. *The sequence $\mathbf{x}$ is Schnorr random.*
4. *The sequence $\mathbf{x}$ is Solovay random.*
5. *The sequence $\mathbf{x}$ is Hertling–Weihrauch random.*

In what follows we will simply call "algorithmically random" a sequence satisfying one of the above equivalent conditions. Theorem 4 motivates the following "randomness hypothesis" formulated in Calude [7]:

A sequence is "random" if it satisfies one of the equivalent conditions in Theorem 4.

[13] As mentioned before, Σ comes equipped with the discrete topology and Σ^ω is endowed with the product topology.

[14] $H(\mathbf{x}(\mathbf{n}))$ is the *program-size complexity* of the string $\mathbf{x}(\mathbf{n}) = \mathbf{x_1 x_2} \dots \mathbf{x_n} \in \mathbf{\Sigma^*}$, that is, the minimum length of a string which produces $\mathbf{x}(\mathbf{n})$ on a universal self-delimiting Turing machine.

[15] Recall that a subbase of a topology is a set β of open sets such that the sets $\bigcap_{W\in E} W$, for finite, nonempty sets $E \subseteq \beta$ form a basis of the topology.

[16] The function $\pi(n, i)$ is a computable bijection, for example, $\pi(n, i) = (n + i)(n + i + 1)/2 + i$.

Various arguments supporting this hypothesis, e.g. *algorithmically random sequences are Borel absolutely normal*[17] have been analyzed in the literature, e.g. Calude [6].[18] Here is recent argument due to Fouché [25]: *if $X \subseteq \Sigma^\omega$ is a measure one Σ_1^0 set, then it contains at least one algorithmically random sequence.* In particular, if X is Π_1^0 set which contains some algorithmically random sequence, then it has nonzero measure. So, if a Π_1^0 event is reflected in some algorithmically random sequence, then the event must be probabilistically significant.

The definition of algorithmically random sequences depends upon the underlying measure μ. The whole theory can be developed with respect to an arbitrary computable measure μ (see Martin-Löf [37]); however, the general case is haunted by paradoxes as those described by Calude and Chiţescu [11].

It is not difficult to destroy randomness. For example, start with an algorithmically random sequence $x_1x_2 \ldots x_n \ldots$ over the alphabet $\{0, 1\}$ and define a new sequence $y_1y_2 \ldots y_n \ldots$, over the alphabet $\{0, 1, 2\}$, by

$$y_1 = x_1, y_n = x_{n-1} + x_n,\ n \geq 2.$$

Then, the new sequence *is not algorithmically random*. The motivation is simple: the strings 02 and 20 (and, infinitely many more others) never appear, so the sequence has clear regularities (which can, actually, be detected by simple statistical randomness tests).

It is much more demanding to "generate" a truly random long string starting from an initial state with a simple description.[19] Note that the condition of simplicity of the initial state is *crucial*: starting from an algorithmically random string one can generate, in a pure algorithmic way, many other algorithmically random strings. For example, if $x_1x_2 \ldots x_{2n-1}x_{2n}$ is an algorithmically random binary string, then break the string into pairs and then code 00, 01, 10, 11 by a, b, c, d: the result is again an algorithmically random sequence. So, the problem is to start from an initial state which can be precisely controlled and has a low program-size complexity and produce measurements of unbounded program-size complexity out its natural dynamical evolution.

Finally let's note that *no algorithmically random sequence is computable*, so no pseudorandom sequence is algorithmically random. The uncomputability of the algorithmic definition of randomness makes it unsuited when it comes to generating a "practical" random sequence, or to check for "practical randomness" a real world sequence.[20] Algorithmically random sequences form an ideal class of sequences; in a sense, algorithmically random sequences are to pseudorandom sequences as real numbers are to rationals. A series of papers by Pincus, Singer, Kalman [41, 40] proposed a computable randomness test based on the so-called "approximate entropy", which has found many practical applications (see also Casti [16] and Beltrami [3]), but it is still far away from being understood from a theoretical point of view. The approximate entropy may be thought as a "computable approximation" of the complexity-theoretic definition of randomness. Indeed, using the algorithmic coding theorem (see Chaitin [19], Gács [26], Calude [6]), a sequence is algorithmically random if and only if the entropy of its first N terms comes ever closer to 1 as N grows. The approximate entropy computes the entropy of larger and larger blocks of digits with respect to an estimation of probabilities given by the Law of Large Numbers, that is, the probabilities are considered to be roughly equal to their relative frequencies.

9 Randomness and Incompleteness

In 1931 Kurt Gödel showed that if you assume a formal axiomatic system containing elementary arithmetic is consistent, then you can prove that it is incomplete. Turing [55] showed that no mechanical procedure, and therefore no formal axiomatic theory, can solve Turing's halting problem, the question of whether a given computer program will eventually halt. Turing's argument was based on computable real numbers. A real number is computable if there is a computer program or algorithm for calculating its digits one by one; of course, nearly all real numbers are not computable. Turing showed that if

[17] Every string appears in an algorithmically random sequence with the probability 2^{-n}, where n is the length of the string.

[18] Bailey and Crandall [1] discussed a hypothesis which implies the normality of many natural real numbers, e.g. π, e. A different approach was discussed in Pincus and Singer [41] and Pincus and Kalman [40].

[19] Note that human beings are not doing a better job in generating "random" bits as Shannon [50] has argued. Biases observed in people's preferences for popular lottery numbers are manifest. See also Bar-Hillel and Wagenaar [2].

[20] Such as a baby's heart-beat: a healthy heart beats in an irregular rhythm, as it responds to a multitude of stimuli from the brain, muscles, digestive organs, etc., and doctors believe that a symptom of SIDS is a strange tendency of the heart to descend into a deadly regular pattern beating. See also [42].

you could find a mechanical procedure to decide if a computer program will ever halt, then you could compute a real number that is not actually computable, which is impossible.

In 1975 Chaitin [17] has introduce his Ω number, the probability that an arbitrary computer program will eventually halt.[21]

The base two expansion of Ω is algorithmically random. The first 10,000 bits of Ω_U include a tremendous amount of mathematical knowledge. In Bennett's words [4]:

> *[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.*

In general, given the first n bits of Ω_U one can decide whether $U(x)$ halts or not on an arbitrary program x of length at most n. However, it is worth noting that even if we get, by some kind of miracle, the first 10,000 digits of Ω, the task of solving the problems whose answers are embodied in these bits is computable but unrealistically difficult: the time it takes to find all halting programs of length less than n from its n digits grows faster than any computable function of n.

Although the infinite amount of information contained in Ω's digits is algorithmically incompressible, it turns out that Ω is computably enumerable: it can be calculated by an infinite process during which one can never know how close one is to the final value.[22] In this way, the halting probability Ω shares two apparently irreconcilable properties: 'algorithmic randomness' and 'computable enumerability'. Recent results due Calude, Hertlinger, Khoussainov, Wand [13], Slaman [51] and Solovay [54] and Calude [8, 9] have increased our understanding of Ω and the depth and pervasiveness of its algorithmic randomness, reinforcing the limits placed on the power of mathematical reasoning by this theory.

Here are the facts. Chaitin [17] has proven that no formal mathematical theory can determine more than a finite number of digits of an Ω. One can explicitly compute a limit on the number of digits of Ω that a specific theory can determine, but apparently it is not possible to prove constructively the theorem. Solovay [54] has now constructed the 'worst ever' Ω for which no bit can be determined even with the help of the most powerful formal axiomatic system used by mathematicians, known as Zermelo-Fraenkel set theory (ZFC). Calude [9] has proven that that every computable enumerable random real is the halting probability of some universal self-delimiting Turing machine for which ZFC (if sound) cannot determine more than its initial block of 1 bits; Solovay construction leads to a real less than one-half, so its binary expansion starts with a zero, hence ZFC cannot determine any bit of it. Finally, Calude [9] has obtained the following constructive version of Chaitin's incompleteness theorem:

Theorem 5 *If ZFC is arithmetically sound and $s = s_1 s_2 \dots s_n$ is a binary string, then we can construct effectively a universal self-delimiting Turing machine U such that the following statements*

"The 0^{th} binary digit of the expansion of Ω_U is 0",

"The 1^{th} binary digit of the expansion of Ω_U is s_1",

[21] Technically, when run on a self-delimiting universal Turing machine; in fact, Ω depends upon the underlying self-delimiting universal Turing machine U, Ω_U, so we have a class of numbers not a number.

[22] A systematic run of all programs will produce better and better approximations, without being able to compute its digits exactly.

"The 2^{th} binary digit of the expansion of Ω_U is s_2",

$\vdots$

"The $(n+1)^{th}$ binary digit of the expansion of Ω_U is s_n",

are true but unprovable in ZFC.

10 Algorithmic Randomness and Physics

Algorithmic information theory, mainly through the algorithmic coding theorem has been successfully applied to a variety of physical problems (mainly in conjunction with Landauer's principle [32][23]): the Maxwell demon paradox (Bennett [5], Zurek [56]), the irreversibility in classical Hamiltonian chaotic systems (Schack and Caves [45]), the characterization of quantum chaos within the framework of statistical physics (Schack and Caves [46, 47]). The program-size complexity (algorithmic information) with respect to two different universal machines differs at most by an unknown, additive, computer-dependent, constant. This type of uncertainty is a serious issue of concern for a physical theory, so various attempts have been made to eliminate it (see, for example, Schack [44]). A sharper versions of the algorithmic coding theorem in which the uncertainty is reduced to a minimum or no assumption is made on the computability of the semi-distribution was recently obtained in Calude, Ishihara and Yamaguchi [14].

11 Experimental Mathematics

The message of algorithmic information theory is that algorithmic randomness is as fundamental and as pervasive in pure mathematics as it is in theoretical physics. This strongly supports "experimental mathematics", a quasi-empirical view of mathematics which sustains that although mathematics and physics are different, it is more a matter of degree than black and white (see Chaitin [21, 22], Calude and Chaitin [10]). Physicists are used to working with assumptions that explain a lot of data, but that can be contradicted by subsequent experiments. Not mathematicians!. Even after Gödel, Turing, Chaitin showed that Hilbert's dream didn't work, in practice most mathematicians carried on as before, in Hilbert's spirit. However, things are changing because of computers. It is easier to run a mathematical experiment on a computer, but you can't always find a proof to explain the results. So in order to cope with complexity and urgency, mathematicians are sometimes forced to proceed in a more pragmatic manner, like physicists. The results are not 100% sure, but the price is worth paying.

Acknowledgments

I am indebted to Elart von Collani, Monica Dumitrescu and Takeyudi Hida for useful discussions and criticism.

References

[1] D. H. Bailey, R. C. Crandall. On the Random Character of Fundamental Constant expansions, `http://www.perfsci.com`, May 2000.

[2] M. Bar-Hillel, W. Wagenaar. The perception of randomness, *Advances in Applied Mathematics* 12 (1991), 428–454.

[3] E. Beltrami. *What is Random? Chance and Order in Mathematics and Life*, Springer-Verlag, New York, 1999.

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

[5] C. H. Bennett. The thermodynamics of computation. *International Journal of Theoretical Physics* 21 (1982), 905–940.

[23] Which specifies the cost of energy dissipation for the erasure a bit of information.

[6] C. S. Calude. *Information and Randomness. An Algorithmic Perspective*, Springer-Verlag, Berlin, 1994.

[7] C. S. Calude. A glimpse into algorithmic information theory, in P. Blackburn, N. Braisby, L. Cavedon, A. Shimojima (eds.). *Logic, Language and Computation*, Volume 3, CSLI Series, Cambridge University Press, Cambridge, 2000, 65–81.

[8] C. S. Calude. A characterization of c.e. random reals, *Theoret. Comput. Sci.*, to appear.

[9] C. S. Calude, Chaitin Ω numbers, Solovay machines and incompleteness, *Theoret. Comput. Sci.*, accepted.

[10] C. S. Calude, G. J. Chaitin. Randomness everywhere, *Nature*, 400 22 July (1999), 319–320.

[11] C. Calude, I. Chiţescu. On a (too) general theory of random sequences, in M. G. Demetrescu, M. Iosifescu (eds.). *Studies in Probability Theory and Related Topics, Papers in Honour of Octav Onicescu on His 90th Birthday*, Nagard Publisher, 1983, 65–69.

[12] C. Calude, I. Chiţescu. Random sequences: some topological and measure-theoretical properties, *An. Univ. Bucureşti, Mat.-Inf.* 2 (1988), 27–32.

[13] C. S. Calude, P. Hertling, B. Khoussainov, and Y. Wang. Recursively enumerable reals and Chaitin Ω numbers, in: M. Morvan, C. Meinel, D. Krob (eds.), *Proceedings of the 15th Symposium on Theoretical Aspects of Computer Science (Paris)*, Springer–Verlag, Berlin, 1998, 596–606. Full paper to appear in *Theoret. Comput. Sci.*

[14] C. S. Calude, H. Ishihara, T. Yamaguchi. Minimal programs are almost optimal, *International Journal of Foundations of Computer Science*, accepted.

[15] C. S. Calude, G. Păun. *Computing with Cells and Atoms*, Taylor & Francis Publishers, London, 2000.

[16] J. L. Casti. Truly, madly, randomly, *New Scientist*, 23 Aug (1997), 32–35.

[17] G. J. Chaitin. A theory of program size formally identical to information theory, *J. Assoc. Comput. Mach.* 22 (1975), 329–340. (Reprinted in: [19], 113–128)

[18] G. J. Chaitin. On the length of programs for computing finite binary sequences, *J. Assoc. Comput. Mach.* 13(1966), 547–569. (Reprinted in: [19], 219-244)

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

[20] G. J. Chaitin. *Information-Theoretic Incompleteness*, World Scientific, Singapore, 1992.

[21] G. J. Chaitin. *The Limits of Mathematics*, Springer-Verlag, Singapore, 1997.

[22] G. J. Chaitin. *The Unknowable*, Springer-Verlag,

[23] I. Ekeland. *The Broken Dice and Other Mathematical Tales of Chance*, University of Chicago Press, Chicago, 1993.

[24] J. Ford. How random is a coin toss? *Physics Today* 36 (1983), 40–47.

[25] W. L. Fouché. Descriptive complexity and reflective properties of combinatorial configurations, *J. London. Math. Soc.* 54 (1996), 199-208.

[26] P. Gács. On the symmetry of algorithmic information, *Soviet Math. Dokl.* 15 (1974), 1477-1480; correction, Ibidem **15** (1974), 1480.

[27] P. Hertling, K. Weihrauch. Randomness spaces, in K. G. Larsen, S. Skyum, and G. Winskel (eds.). *Automata, Languages and Programming, Proceedings of the 25th International Colloquium, ICALP'98* (Aalborg, Denmark), Springer-Verlag, Berlin, 1998, 796–807.

[28] P. Hoffman. *The Man Who Loved Only Numbers*, Hyperion, New York, 1998.

[29] E. T. Jaynes. *Probability Theory: The Logic of Science*, Fragmentary Edition, march 1996.

[30] M. Kac. What is random? *American Scientist* 71 (1983), 405-406.

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

[32] R. Landauer. Irreversibility and heat generation in the computing process, *IBM J. Res. Develop.* 5 (1961), 183–191.

[33] J. Maddox. The poor quality of random numbers, *Nature* 372 (1994), 403.

[34] M. May. What is random? *American Scientist* 85, 3 (1997), 222.

[35] G. Kolata. What does it mean to be random? *Science* 7 (1986), 1068.

[36] P. Martin-Löf. *Algorithms and Random Sequences*, Erlangen University, Nürnberg, Erlangen, 1966.

[37] P. Martin-Löf. The definition of random sequences, *Inform. and Control* 9 (1966), 602–619.

[38] G. Milburn. *The Feynman Processor. An Introduction to Quantum Computation*, Allen & Unwin, St. Leonards, 1998.

[39] A. Peres. *Quantum Theory: Concepts and Methods*, Kluwer Academic Publishers, Dordrecht, 1993.

[40] S. Pincus, R. E. Kalman. Not all (possibly) "random" sequences are created equal, *Proc. Nat. Acad. Sci. USA* 94 (1997), 3513–3518.

[41] S. Pincus, B. H. Singer. Randomness and degrees of irregularity, *Proc. Nat. Acad. Sci. USA* 93 (1996), 2083–2088.

[42] I. Popescu. Calcul des intègrales multiples par la mètode de Monte Carlo, *Rev. Roumaine Math. Pures Appl.* 26 (1981), 19–29.

[43] E. J. Routh. *Stability of Motion*, Taylor & Francis, New York, 1975 (edited by A. T. Fuller).

[44] R. Schack. Algorithmic information and simplicity in statistical physics, *Int. J. Theor. Physics* 36 (1997), 209–226.

[45] R. Schack, C. M. Caves. Information and entropy in the baker's map, *Phys. Rev. Lett.* 69, 23 (1992), 3413–3416.

[46] R. Schack, C. M. Caves. Information-theoretic characterization of quantum chaos, *Phys. Rev., E* (3) (4) 53 (1996), 3257–3270.

[47] R. Schack, C. M. Caves. Chaos for Liouville probability densities, *Phys. Rev., E* (3) (4) 53 (1996), 3387–3401.

[48] B. Schumaker. Quantum coding, *Physical Review A*, 51, 4 (1995), 2738–2747.

[49] C. E. Shannon. A mathematical theory of communication, *Bell. System Technical Journal* 27 (1948), 379–423, 623–656.

[50] C. E. Shannon. Computers and automata, *Proceedings of the I. R. E.* 41 (1953), 1235–1241.

[51] T. A. Slaman. Randomness and recursive enumerability, *SIAM J. Comput.* (to appear).

[52] R. J. Solomonoff. A formal theory of inductive inference, Part 1 and Part 2, *Inform. and Control* 7(1964), 1–22 and 224–254.

[53] R. M. Solovay. *Draft of a paper (or series of papers) on Chaitin's work ... done for the most part during the period of Sept.–Dec. 1974*, unpublished manuscript, IBM Thomas J. Watson Research Center, Yorktown Heights, New York, May 1975, 215 pp.

[54] R. M. Solovay. A version of Ω for which ZFC can not predict a single bit, in C.S. Calude, G. Păun (eds.). *Finite Versus Infinite. Contributions to an Eternal Dilemma*, Springer-Verlag, London, 2000, 323-334.

[55] A. M. Turing. On computable numbers with an application to the Entscheidungsproblem, *Proc. Amer. Math. Soc.* 42 (1936-7), 230-265; a correction, ibid., 43 (1937), 544-546.

[56] W. H. Zurek. Algorithmic randomness, physical entropy, measurements, and the Demon of choice, in ed. J. G. Hey. *Feynman and Computation. Exploring the Limits of Computers*, Perseus Books, Reading, Massachusetts, 1999, pp. 393–410.

[57] R. von Mises. *Mathematical Theory of Probability and Statistics*, Edited and Complemented by Hilda Geiringer, Academic Press, New York, 1974.

[58] M. vos Savant. *The World's Most Famous Math Problem*, St. Martin's Press, New York, 1993.

[59] K. Weihrauch. *Computability*, Springer-Verlag, Berlin, 1987.