UvA-DARE (Digital Academic Repository)

Algorithmic information theory

van Lambalgen, M.

DOI

[10.2307/2274821](https://doi.org/10.2307/2274821)

Publication date

1989

Published in

Journal of Symbolic Logic

[Link to publication](#)

Citation for published version (APA):

van Lambalgen, M. (1989). Algorithmic information theory. *Journal of Symbolic Logic*, 54(4), 1389-1400. <https://doi.org/10.2307/2274821>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Algorithmic Information Theory

Michiel van Lambalgen

Journal of Symbolic Logic, Volume 54, Issue 4 (Dec., 1989), 1389-1400.

Stable URL:

<http://links.jstor.org/sici?sici=0022-4812%28198912%2954%3A4%3C1389%3AAIT%3E2.0.CO%3B2-H>

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Journal of Symbolic Logic is published by Association for Symbolic Logic. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/asl.html>.

Journal of Symbolic Logic

©1989 Association for Symbolic Logic

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2002 JSTOR

ALGORITHMIC INFORMATION THEORY

MICHIEL VAN LAMBALGEN

Abstract. We present a critical discussion of the claim (most forcefully propounded by Chaitin) that algorithmic information theory sheds new light on Gödel's first incompleteness theorem.

§0. Introduction. The basic concept of algorithmic information theory, as presented by Chaitin, is that of the *information content* of an individual object, which is a measure of the difficulty of specifying, describing or calculating that object. Its aim is the attempt to apply information-theoretic and probabilistic ideas to recursive function theory and metamathematics. In particular, it purports to be able to *refine* Gödel's first incompleteness theorem by introducing *degrees* of incompleteness of formal systems and relate these to the "algorithmic information content" of those theories (see [1]–[6]).

Chaitin's views enjoy a certain popularity, witness for instance the inclusion of two of his papers in the recent anthology *New directions in the philosophy of mathematics* [25]. Thomas Tymoczko, its editor, writes in his introductory remarks:

In these essays, Chaitin blends standard mathematical logic with a computer oriented information theory to obtain a new version of Gödel's fundamental results. This investigation, he argues, clarifies the significance of Gödel's incompleteness theorem, showing that it is not an isolated paradox, but a natural consequence of the constraints imposed by information theory. . . . Actually, the moral is more general: in so far as we precisely and consistently specify the methods of reasoning permitted, we determine an upper bound to the complexity of our results. This upper bound is the information theoretic limit imposed by Gödel's theorem [25, 287–289].

In fact, in the physical literature connected with chaotic dynamics, the moral is taken to be still more general; the interested reader may wish to consult Ford [27] for a representative example.

Received January 25, 1988; revised July 25, 1988.

Acknowledgment. I am indebted to Dick de Jongh and Albert Visser for illuminating comments, and to the referee for essential improvements. The author is a Huygens fellow of the Netherlands Organization for Scientific Research (NWO).

©1989, Association for Symbolic Logic
0022-4812/89/5404-0013/\$02.20

The purpose of this note is to throw some cold water on these developments, by pointing out that Chaitin's mathematics does not support his philosophical conclusions. We study two applications of algorithmic information theory: in §1 we examine the proposed explanation of the first incompleteness theorem as being in fact a theorem on information content [1], [3], [4], and in §2 we discuss the alleged discovery of "randomness in mathematics" [5], [6].

§1. Information-theoretic limitations on formal systems. G. J. Chaitin's claim, that algorithmic information theory provides the proper setting to discuss Gödel's first incompleteness theorem, consists of two parts:

(a) It is easy to show that to each sound formal system S containing, say, arithmetic, one can associate a constant $c_S \in \mathbb{N}$ such that S cannot prove true statements of the form " $K(w) > c_S$ ", where w is a finite binary string and $K(w)$ is the *Kolmogorov complexity* of w , the definition of which will be recalled below. Chaitin now suggests that there exists a precise quantitative relationship between c_S and the "information content" or "entropy" of the formal system S , where the latter quantity is to be measured by the Kolmogorov complexity of the axioms (*qua* finite binary string):

Here is our incompleteness theorem for formal axiomatic theories whose arithmetical consequences are true. The set-up is as follows: the axioms are a finite string, the rules of inference are an algorithm for enumerating the theorems given the axioms and we fix the rules of inference and vary the axioms. Within such a formal system a specific string cannot be proven to be of entropy [= complexity] greater than the entropy of the axioms of the theory. Conversely, there are formal theories whose axioms have entropy $n + O(1)$ in which it is possible to establish all true propositions of the form " $K(\text{specific string}) > n$ " [3, 336].

(b) Chaitin's ultimate aims are even more ambitious:

Gödel's original proof constructed a paradoxical assertion that is true but not provable within the usual formalizations of number theory. In contrast I would like to measure the power of a set of axioms and rules of inference. I would like to be able to say that if one has ten pounds of axioms and a twenty-pound theorem, then the theorem cannot be derived from the axioms. [...] To be more specific, I will apply the viewpoint of thermodynamics and statistical mechanics to Gödel's theorem, and will use such concepts as probability, randomness, entropy, and information to study the incompleteness phenomenon and to attempt to evaluate how widespread it is [4, 942].

Although the theorem alluded to in the first quotation is valid (cf. 1.2.6 below), the emphasis on "the *entropy* of the axioms of the theory" is slightly misleading. In fact, it will be shown by a simple application of a reflection theorem due to Kreisel and Levy that there exists an infinite collection C of essentially different arithmetical theories S for which there is no interesting relationship between the information content of S and the *minimal* c_S such that for all w , $S \not\vdash K(w) > c_S$.

More precisely, all the S in C have the same minimal c_S such that for all w , $S \nvdash K(w) > c_S$, whereas by the properties of Kolmogorov complexity, entropy is not even bounded on C . It follows that the information content of a theory (still defined as the Kolmogorov complexity of its axioms) is only marginally relevant to the incompleteness phenomenon mentioned in (a), and hence the expectation expressed in (b) seems somewhat premature. Moreover, it will be seen that only by confusing object language and metalanguage can (a) be said to support (b). The overall conclusion is that Chaitin's constructions do not, as yet, provide an explanation of the first incompleteness theorem.

1.1. Kolmogorov complexity. We briefly review the properties of Kolmogorov complexity that we shall need in the sequel. Although Chaitin [2]–[6] uses a slightly different concept of complexity (involving prefixalgorithms), the essentials can be illustrated using the simpler concept.

1.1.1. DEFINITION. Let $A: 2^{<\omega} \rightarrow 2^{<\omega}$ be a partial recursive function with Gödelnumber $\ulcorner A \urcorner$. If w is a finite binary string, then $|w|$ denotes the length of w . The *complexity* $K_A(w)$ of w with respect to A is defined to be

$$K_A(w) = \begin{cases} \infty & \text{if there is no } p \text{ such that } A(p) = w, \\ |p| & \text{if } p \text{ is a shortest input such that } A(p) = w. \end{cases}$$

A universal machine U is said to be *asymptotically optimal* if it is specified by the requirement that on inputs of the form $q = 0^{\ulcorner A \urcorner} 1p$ (i.e. a sequence of $\ulcorner A \urcorner$ zeros followed by a one, followed by a string p), U simulates the action of A on p . Fix a Gödelnumbering and an asymptotically universal machine U and put $K(w) := K_U(w)$. K is called the *Kolmogorov-complexity* of w (Kolmogorov [12]–[16]). Inputs will also be called *programs*.

Observe that, identifying a natural number with its binary representation, it makes sense to speak of the complexity of a natural number. Clearly, we have

1.1.2. LEMMA. (a) For any partial recursive $A: 2^{<\omega} \rightarrow 2^{<\omega}$ and for all w , $K(w) \leq K_A(w) + \ulcorner A \urcorner + 1$.

(b) For some constant c and for all w , $K(w) \leq |w| + c$.

A simple counting argument will show that for each m , there exist infinitely many sequences w with $K(w) > m$. In the sequel, the expression “ $\#A$ ” stands for the cardinality of the (finite) set A and 2^n denotes the set of binary strings of length n .

1.1.3. LEMMA. (a) $\#\{w \in 2^n \mid K(w) \leq n - m\} \leq 2^{n-m+1} - 1$.

(b) $\#\{w \in 2^n \mid K(w) > n - m\} > 2^n \cdot (1 - 2^{-m+1})$.

(c) Fix m ; then for all sufficiently large n there is $w \in 2^n$ such that $K(w) > n - m$.

PROOF. (a) The number of programs on U of length $\leq n - m$ is $\leq 2^{n-m+1} - 1$. Hence (b) at least $2^n - 2^{n-m+1} = 2^n \cdot (1 - 2^{-m+1})$ sequences in 2^n satisfy $K(w) > n - m$. (c) is an immediate consequence of (b). $\square$

Note the extreme simplicity of the argument: it can be formalized in any formal system based on classical logic which is capable of handling finite sets of integers. This is to be contrasted with the fact, proved in 1.2, that the set $\{w \mid K(w) > g(|w|)\}$, where g is a total recursive increasing function, though infinite, contains no infinite recursively enumerable subsets.

1.2. Information-theoretic incompleteness. In this section we prove that for each sound formal system S there exists a constant c_S such that for all finite binary strings

$w, S \nVdash K(w) > c_S$. For reasons that will become clear in the sequel, we give two proofs, both different from Chaitin's. We use Rogers' notation for partial recursive functions and recursively enumerable sets [21]: ϕ_n denotes the partial recursive function from $\mathbf{N}$ to $\mathbf{N}$ with Gödelnumber n and W_e denotes the r.e. subset of $\mathbf{N}$ with Gödelnumber e . As usual, we shall assume that sets such as $2^{<\omega}$ or $2^{<\omega} \times \omega$ etc. are coded into the natural numbers.

1.2.1. LEMMA. $\{\langle w, m \rangle \in 2^{<\omega} \times \omega \mid K(w) \leq m\}$ is recursively enumerable.

PROOF. If U is the universal machine defined in 1.1.1, we have, using the definition of K ,

$$\begin{aligned} & \{\langle w, m \rangle \in 2^{<\omega} \times \omega \mid K(w) \leq m\} \\ &= \{\langle w, m \rangle \in 2^{<\omega} \times \omega \mid \exists p (U(p) = w \ \& \ |p| \leq m)\}; \end{aligned}$$

the condition on the right hand side is Σ_1 . □

Hence $\{\langle w, m \rangle \in 2^{<\omega} \times \omega \mid K(w) > m\}$ is Π_1 ; but it also satisfies a stronger property.

1.2.2. DEFINITION. (a) A set A is *immune* if it is infinite but contains no infinite recursively enumerable subset.

(b) A set A is *effectively immune* if, for some total recursive function $g: \omega \rightarrow \omega$, $W_e \subseteq A$ implies $\#W_e \leq g(e)$.

(c) A set B is (effectively) *simple* if B is r.e. and B^c is (effectively) immune.

1.2.3. THEOREM. *There exists a constant d such that any r.e. subset W_e of $\{\langle w, m \rangle \in 2^{<\omega} \times \omega \mid K(w) > m\}$ is bounded in the second coordinate by $K(e) + d$.*

PROOF. Although the result is stated for K only, it holds for a wide variety of complexity measures. To bring this out, we give an abstract proof. Let U be the universal algorithm and define a partial recursive function f as follows. f operates on inputs of the form $0^n 1 q$. Given this input, f first calculates $U(q)$; if and when it has found $e = U(q)$, it generates W_e until it has found a pair $\langle w, m \rangle \in W_e$ such that $m > |q| + n + 1$; it then outputs w . Now suppose $W_e \subseteq \{\langle w, m \rangle \in 2^{<\omega} \times \omega \mid K(w) > m\}$. Apply the recursion theorem to get an n such that for all q , $\phi_n(q) \cong f(0^n 1 q)$. (That is, the left hand side is defined iff the right hand side is, and when defined the two sides are equal.) Let q_0 be such that $e = U(q_0)$; we claim that $\phi_n(q_0)$ is undefined. For suppose that $\phi_n(q_0) = w$. Then on the one hand, by construction,

$$(1) \quad K(w) > m > |q_0| + n + 1;$$

on the other hand, by Lemma 1.1.2,

$$(2) \quad K(w) \leq K_{\phi_n}(w) + n + 1 \leq |q_0| + n + 1.$$

Hence $\phi_n(q_0)$ is undefined. It follows that $K(e) + n + 1$ is an upper bound for the second coordinate of W_e . □

We first apply the theorem to obtain some recursion-theoretic information on K .

1.2.4. COROLLARY. *Let $g: \omega \rightarrow \omega$ be total recursive and suppose that $\lim_{n \rightarrow \infty} g(n) = \infty$ and $g(n) \leq n - m$ for some m . Then $\{w \mid K(w) > g(|w|)\}$ is immune. In addition, if $\lim_{n \rightarrow \infty} g(n) = \infty$ recursively, then $\{w \mid K(w) > g(|w|)\}$ is effectively immune.*

PROOF. Let $W_e \subseteq \{w \mid K(w) > g(|w|)\}$. Put $V_e := \{\langle w, g(|w|) \rangle \mid w \in W_e\}$; then for some total recursive f , $V_e = W_{f(e)}$. Since $W_{f(e)} \subseteq \{\langle w, m \rangle \mid K(w) > m\}$, $W_{f(e)}$ is bounded in the second coordinate, e.g. by $\log_2 f(e)$. But then, if $\lim_{n \rightarrow \infty} g(n) = \infty$,

W_e must be finite and if $\lim_{n \rightarrow \infty} g(n) = \infty$ recursively, we can choose effectively $n_0(e)$ such that for $n \geq n_0(e)$, $\log_2 f(e) \leq g(n)$. In the latter case we therefore have $\#W_e \leq 2^{n_0(e)+1}$. $\square$

It follows from the corollary that the r.e. relation $\{\langle w, m \rangle \mid K(w) \leq m\}$ is not recursive and likewise that the function $K: 2^{<\omega} \rightarrow \omega$ is not recursive.

1.2.5. EXAMPLE. The set of *random strings* $\{w \mid K(w) > |w| - m\}$ is effectively immune. By a theorem of Martin (see Soare [24, 87]) it follows that $\{w \mid K(w) \leq |w| - m\}$ is a *complete recursively enumerable set*.

We now formulate the first half of Chaitin's incompleteness theorem. Recall that for any natural number m all except finitely many w satisfy $K(w) > m$. We proved this in 1.1 using only elementary properties of finite sets; the proof can be formalized in any (classical) theory which contains a modicum of arithmetic. Nevertheless, as the following theorem shows, it is well nigh impossible to verify that some *specific string* has high complexity.

1.2.6. THEOREM. *Let S be a sound formal system, identified with its r.e. set of theorems. Let p be an r.e. index for S . Then for some constant d , independent of S , and for all w , $S \vdash K(w) > K(p) + d$.*

PROOF. Delete from S all theorems not of the form " $K(w) > m$ " and call the resulting sound formal system S' . This operation is primitive recursive. S' may be identified with an r.e. subset of $\{\langle w, m \rangle \in 2^{<\omega} \times \omega \mid K(w) > m\}$ with Gödelnumber p' . By Theorem 1.2.3, S' is bounded in the second coordinate by $K(p') + d'$, for some constant d' not depending on p' . But for some c , $K(p') \leq K(p) + c$. $\square$

1.2.7. REMARK. Since the operation that brings us from S to S' is primitive recursive, it is of course true that for some constant d , S proves no statements of the form $K(w) > K(p) + d$, where p is an r.e. index for S . But the above proof already gives a hint that $K(p)$, the information content of the formal system S , might not be relevant to the determination of the *minimal* c such that for all w , $S \vdash K(w) > c$. Let us call the minimal such c , which depends on S , the *characteristic constant* of the formal system S , henceforth denoted c_S . Then the above proof leaves open the possibility that such widely differing theories as PA or ZF still have the same characteristic constant. We shall see below that such situations do indeed occur.

We now sketch a completely different way of determining a constant c such that (for instance) PA proves no statement of the form " $K(w) > c$ ". This observation is due to Albert Visser and Dick de Jongh.

1.2.8. THEOREM. *There exists a constant c such that PA proves no statement of the form " $K(w) > c$ ".*

PROOF. Fix a listing of the proofs of PA and define a partial recursive function ϕ_e as follows: $\phi_e(m) = n$ iff n is the k in the first proof in PA of a statement of the form " $\phi_m(m) \neq k$ ". We claim that $\phi_e(e)$ is undefined. For if $\phi_e(e) = n$, then $\text{PA} \vdash \phi_e(e) \neq n$, a contradiction. It follows that $\text{PA} + \{\phi_e(e) = n\}$ is consistent for each n , for if not, then $\text{PA} \vdash \phi_e(e) \neq n$, hence $\phi_e(e)$ would be defined.

Since for all n , $\text{PA} + \{\phi_e(e) = n\}$ is consistent, $\text{PA} + \{K(n) \leq 2e + 1\}$ is also consistent for all n ; hence PA proves no statement of the form " $K(n) > 2e + 1$ ". $\square$

The argument of course applies to other formal systems. The above argument shows convincingly that there is no a priori reason to expect that c_S and the information content of S (still defined as the Kolmogorov complexity of its axioms) are related in some interesting way.

1.2.9. REMARK. In the quotation from Chaitin given in the Introduction, there is also a weak converse to Theorem 1.2.6: “There are formal theories whose axioms have entropy $n + O(1)$ in which it is possible to establish all true propositions of the form ‘ $K(\text{specific string}) > n$.’” For our purposes it is not necessary to prove this result here. We remark only that a full converse to Theorem 2.6 would read: “All formal theories whose axioms have entropy $n + O(1)$ establish all true propositions of the form ‘ $K(\text{specific string}) > n$.’” But the arguments below show that *this* statement is false.

1.3. Discussion.

1.3.1. Although Theorem 1.2.6 was hailed as a “dramatic extension of Gödel’s theorem” (Davis [8]), we should not forget that there is a big difference between the two results. Gödel’s first incompleteness theorem is an *explicit* construction of an undecidable (hence true) Π_1 formula: the fixed point lemma [23, 827] associates with any formal system S in a primitive recursive way a formula ψ_S which says of itself “I am unprovable in S ”. But neither Theorem 1.2.6 nor Theorem 1.2.8 provides an explicit construction. Consider for instance Theorem 1.2.6. First, its proof shows that the information content of a formal system S is not a recursive function of S . Second, suppose we take some recursive upper bound $f(S)$ for the information content of S ; then it is *still* not possible to determine recursively a word $w(S)$ such that $K(w(S)) > f(S) \geq c_S$. If this were so, we could define an infinite r.e. sequence of formal systems S_n and words $w(S_n)$ such that $K(w(S_n)) > f(S_n)$ and $f(S_n) \rightarrow \infty$ follows: $S_0 = \text{PA}$, $S_1 = S_0 \cup \{K(w(S_0)) > f(S_0)\}$, etc. By construction, $c_{S_n} < c_{S_{n+1}}$, hence also $\lim_{n \rightarrow \infty} f(S_n) = \infty$. But Corollary 1.2.4 implies that we can construct only finitely many $w(S_n)$. Hence it is impossible to determine effectively, given a formal system S , a word $w(S)$ such that $K(w(S)) > c(S)$. In *this* sense, Theorem 1.2.6 is a weak form, rather than an extension, of the first incompleteness theorem. An analogous argument applies to Theorem 1.2.8.

1.3.2. We now show that c_S is in general not roughly equal to the information content of the formal system S , where by “roughly equal” we mean that the difference between the two quantities is bounded.

By Theorem 11 of Kreisel and Levy [17, 121], the arithmetical fragment of ZF is not finitely axiomatizable over PA. Theorem 1.2.6 assigns finite constants c_{PA} and c_{ZF} such that no statement “ $K(w) > c_{\text{PA}}$ ” (“ $K(w) > c_{\text{ZF}}$ ”) is provable in PA (ZF). (Note that we do not even know whether $c_{\text{ZF}} > c_{\text{PA}}$!) It follows that an infinity of ever stronger number theories S_n , which lie in between PA and (the arithmetical fragment of) ZF, must have the *same* characteristic constant c and must prove the *same* set of statements of the form “ $K(w) > m$ ”. Since K is unbounded on axioms for the S_n , the characteristic constants of the S_n are not roughly equal to the information contents of these axioms. In other words, although the information content of the axioms indeed provides the “outer limits of the possible” (cf. Schwartz’ introduction to [6]), the true limits have to be sought elsewhere.

1.3.3. Furthermore, there is nothing in Theorem 1.2.6 which supports Chaitin’s claim (stated as (b) in the Introduction) that the undecidability of a formula can be explained as the result of an excess of information content. Observe that we said nothing about the *information content* of the *formulas* “ $K(w) > c_S$ ” (whatever that may mean); all that mattered was that the undecidable formula *asserts* that some

specific string contains too much information, which is something entirely different.

Once one has seen this confusion between object language and metalanguage, however, it is easy to construct by diagonalization true statements which are undecidable in PA more or less because they contain too much information, if we stipulate that the information content of a statement is the complexity of its Gödelnumber. The construction of such a statement utilizes the fixed point lemma:

1.3.4. LEMMA [23, 827]. *Let ϕ be an arithmetical formula in one free variable. Then, for infinitely many ψ , $PA \vdash (\psi \leftrightarrow \phi(\ulcorner \psi \urcorner))$.*

We use the fixed point lemma to define a sentence ψ which says roughly “I contain too much information for PA”. Put $k_0 := \max\{k \mid K(k) \leq c_{PA}\}$. Choose (noneffectively!) ψ such that $\ulcorner \psi \urcorner > k_0$ and $PA \vdash (\psi \leftrightarrow K(\ulcorner \psi \urcorner) > c_{PA})$. Then $PA \nvdash \psi$, since otherwise $PA \vdash K(\ulcorner \psi \urcorner) > c_{PA}$, which is impossible by Theorem 1.2.6; but ψ is true, for if $\neg \psi$ were true then $K(\ulcorner \psi \urcorner) \leq c_{PA}$, which implies $\ulcorner \psi \urcorner \leq k_0$. Since PA is sound, $PA \nvdash \neg \psi$. Hence ψ is true but undecidable in PA. The construction is trivial, however, since we essentially use the fact that there exist fixed points of “ $K(\ulcorner \psi \urcorner) > c_{PA}$ ” with arbitrarily large Gödelnumber (hence this argument could be used to show that $K(\ulcorner \phi \urcorner)$ is *not* a good measure of the information content in ϕ). Furthermore the fixed point does not *quite* express “I contain too much information for PA”, since there is no effective determination of c_{PA} .

The considerations above do not completely rule out the possibility that some kind of information concept is useful in studying incompleteness. They do show, however, that the *complexity* of the axioms is not a good measure of information. Furthermore, if the information is an *integer-valued* function and obeys something like Theorem 1.2.6, then we must accept the consequence that a theory S_1 may be stronger than S_2 , while having the same information content as S_2 . It is difficult to imagine a concept of information which allows this possibility. Even if the information concept turns out to be useless for the study of formal systems, it may be worthwhile to investigate what *other* properties of formal systems are relevant for the values of their characteristic constants. This investigation, however, is seriously hampered by the extreme scarcity of concrete examples: as noted above, we do not know whether $c_{PA} < c_{ZF}$, and, worse, we even have no idea how to establish results of this sort.

1.4. Digression: complexity and intuitionistic logic. In this subsection we briefly indicate why the proof of Proposition 1.1.3(c) essentially uses classical logic. This is a consequence of the existence property of Heyting Arithmetic (HA): if HA proves $\exists n A(n)$, then for some numeral n , HA also proves $A(n)$. As a consequence, if HA would prove $\exists w (K(w) > c_{PA})$, it would also prove, for some w , $K(w) > c_{PA}$, which is impossible. The argument implies that Chaitin’s “information-theoretic incompleteness theorem” does not apply to intuitionistic theories. It also follows that occasional references [26, 21] to the constructivity of those lower bound arguments in computational complexity which involve random strings have to be taken with a grain of salt.

§2. Incompleteness theorems for random reals. In [5] and [6] Chaitin attempts to give a kind of “ultimate form” of the first incompleteness theorem by showing that mathematics is not only beset by incompleteness, but by randomness as well. To

explain his idea, we have to say a few words on one of his fundamental constructions, first introduced in [2], the *halting probability* Ω of a universal Turing machine. This concept can only be defined for *prefixalgorithms*, i.e. algorithms A which satisfy “ $\{p \mid A(p) \text{ halts}\}$ is prefixfree” (a set of binary words is prefixfree if no element is an initial segment of some other element). Let U be a universal prefixalgorithm; then $\Omega := \sum \{2^{-|p|} \mid U(p) \text{ halts}\}$ is well-defined. He then constructs an exponential Diophantine equation $Q = Q'$ with the following property:

... $Q = Q'$ has finitely or infinitely many solutions depending on whether the n^{th} bit of the base two expansion of the halting probability Ω is a zero or a one. $Q = Q'$ is quite a remarkable equation, as it shows that there is a kind of uncertainty principle even in pure mathematics, in fact, even in the theory of whole numbers. [...] It may be said that the truth or falsity of the assertion that there are infinitely many solutions is indistinguishable from the result of independent tosses of a fair coin [6, 131].

The last assertion is motivated by the fact that Ω satisfies the definition of randomness given by Martin-Löf [20]. The philosophical importance of the construction is emphasized in the conclusion of [6]:

...we see that proving whether particular exponential Diophantine equations have finitely or infinitely many solutions, is absolutely intractable. Such questions escape the power of mathematical reasoning. This is a region in which mathematical truth has no discernible structure or pattern and appears to be completely random. These questions are beyond the power of human reasoning. Mathematics cannot deal with them. Nonlinear dynamics and quantum mechanics have shown that there is randomness in nature. I believe that we have demonstrated in this book that randomness is already present in pure mathematics, in fact, even in rather elementary branches of number theory [6, 160].

To counter these grand claims, we first give a very short proof of Chaitin's result using general nonsense only (thereby making the need for an explicit description of the 900,000-character 17,000-variable universal exponential Diophantine equation, deemed to be necessary in [6], slightly less pressing). While this alone would lead one to suspect that the result is not deep, it is also argued that Chaitin's association of randomness in nature with the mathematical phenomenon loosely described above is rather unfortunate.

2.1. Martin-Löf randomness and Δ_2 -definability. We quickly review those properties of Martin-Löf's definition of randomness that we shall need for our purposes. A full exposition can be found in [20], [22] or [18].

2.1.1. DEFINITION. The probability measure μ on 2^ω is called *computable* if there exists a recursive function $g: 2^{<\omega} \times \omega \rightarrow \mathbf{Q}$ such that for all $w, k: |\mu[w] - g(w, k)| < 2^{-k}$, where $[w]$ is the set of all x having the same initial segment w .

2.1.2. DEFINITION (MARTIN-LÖF [20]). Let μ be a computable measure. $N \subseteq 2^\omega$ is a *recursive sequential test* with respect to μ if N can be written as a Π_2 set $\bigcap_n O_n$, where $O_n \in \Sigma_1$, the function $n \rightarrow O_n$ is recursive, $O_{n+1} \subseteq O_n$ and $\mu O_n \leq 2^{-n}$.

2.1.3. DEFINITION (MARTIN-LÖF [20]). Let μ be a computable measure. $x \in 2^\omega$ is *random* with respect to μ (denoted $x \in R(\mu)$) if for all recursive sequential tests N with respect to μ , $x \notin N$.

Obviously $\mu R(\mu) = 1$; and it can be shown that sequences random in this sense (for measures of the form $(1 - p, p)^\omega$) satisfy the usual probabilistic laws such as the strong law of large numbers, the law of the iterated logarithm, the invariance of limiting relative frequency under application of recursive place selections, etc. The following property is slightly surprising:

2.1.4. LEMMA (MARTIN-LÖF [20]). Let μ be a computable measure.

(a) *The collection of recursive sequential tests with respect to μ is recursively enumerable.*

(b) *There exists a universal recursive sequential test with respect to μ , i.e. a test U such that for all recursive sequential tests N with respect to μ , $N \subseteq U$.*

A curious consequence of the preceding lemma is that $R(\mu)$ has elements which are rather simple. Although it does not contain recursive sequences if μ is non-atomic, $R(\mu)$ does contain Δ_2 -definable sequences. This is a consequence of the following fact.

2.1.5. BASIS THEOREM (SOARE [24, 109]). Any nonempty Π_1 subset of 2^ω has a Δ_2 -definable element.

PROOF (sketch). A Π_1 subset of 2^ω can be viewed as the set of infinite paths through a recursive binary tree T . Call $w \in T$ *admissible* if $\forall n > |w| \exists v \in 2^n (w \subseteq v \text{ \& } v \in T)$. (By König's lemma, w is admissible iff there is an infinite branch of T through w .) The set of admissible words is Π_1 . Since the subset is nonempty, T has an infinite branch. The *leftmost* infinite branch can be constructed recursively in the set of admissible words, which is Π_1 ; hence this branch must itself be Δ_2 . $\square$

The following observation is originally due to Schnorr [22, 56]; we give an abstract proof.

2.1.6. LEMMA. Let μ be a nonatomic computable measure. Then $R(\mu)$ contains Δ_2 -definable sequences.

PROOF. By 2.1.4, $R(\mu)$ is a Σ_2 set of measure 1. Pick a Π_1 set $A \subseteq R(\mu)$ such that $\mu A > 0$ and apply the basis theorem. $\square$

2.1.7. REMARK. Chaitin's Ω is also Δ_2 -definable.

To get Chaitin's result [6, 159, Theorem D], we now proceed as follows. Let $\lambda = (\frac{1}{2}, \frac{1}{2})^\omega$, $x \in \Delta_2 \cap R(\lambda)$. A fortiori, x is Π_2 -definable. Since every r.e. relation is Diophantine [9], there exists a polynomial P such that

$$x_n = 1 \quad \text{iff} \quad \forall \vec{m} \exists \vec{k} P(\vec{m}, \vec{k}, n) = 0.$$

Hence one could say with Chaitin that whether or not the equation $P(\vec{m}, \vec{k}, n) = 0$ has a solution for all $\vec{m}$ "jumps around in a completely unpredictable manner as n varies". As we have seen, a trivial combination of a few known theorems yields Chaitin's result. Whether this situation is most adequately described by such phrases as "a dramatic version of Gödel's theorem" [6, 131] or "the strongest possible version of Gödel's incompleteness theorem" [6, ix], is perhaps a matter of taste. (An application of the ideas of algorithmic information theory in this context would presumably yield the conclusion that the information content of a theorem has to be

measured by its simplest proof.) In any case, we believe that the analogy with nonlinear dynamics and quantum mechanics is rather farfetched. The randomness of a quantum mechanical phenomenon such as electron spin does not disappear as soon as we decide to adopt a different definition of randomness, whereas in cases such as the above we seem to observe an artifact of the particular definition chosen. Of course other (arithmetical) definitions of random sequences would yield new examples of “randomness in mathematics” constructed with the help of a suitable basis theorem; but one looks in vain for examples which are stable under a change of definition. Hence the *robustness* which is present in the physical examples that Chaitin mentions is lacking here.

2.2. Digression. As a technical aside, we show here that Δ_2 -definable sequences, even when random, exhibit rather atypical complexity behavior, thus lending some support to the intuitive conviction that such simple sequences cannot be “truly” random. As a bonus, we obtain a new proof of a celebrated result of Martin-Löf on complexity oscillations.

The first theorem gives the typical complexity behavior of infinite binary sequences. If $x \in 2^\omega$, then $x(n)$ is the initial segment of x of length n , and x_n the n th coordinate.

2.2.1. THEOREM (MARTIN-LÖF). $\lambda\{x \in 2^\omega \mid \exists m \forall k \exists n \geq k \ K1(x(n)) > n - m\} = 1$, where $\lambda = (\frac{1}{2}, \frac{1}{2})^\omega$.

We now show that Δ_2 -definable sequences behave rather atypically in this respect.

2.2.2. THEOREM. If $x \in 2^\omega$ is Δ_2 -definable, then $\lim_{n \rightarrow \infty} (n - K2(x(n))) = \infty$.

PROOF. We use the modulus lemma (Soare [24, 56]): a sequence x in 2^ω is Δ_2 iff for some sequence $(\xi_k)_{k \in \mathbb{N}}$, where $\xi_k \in 2^\omega$ and $\{\langle k, n \rangle \mid (\xi_k)_n = 1\}$ is recursive: for all n , $x_n = \lim_{k \rightarrow \infty} (\xi^k)_n$. (In words: Δ_2 sequences x can be produced by Turing machines if the machine is allowed to correct itself a finite number of times per x_n .)

We first need a convention on coding pairs of strings. Given finite binary strings w and v we can code the pair $\langle w, v \rangle$ into a single string $\overline{wv}$ as follows: double the bits in w , write a 1 and then add v . Define an algorithm A as follows. A operates on inputs of the form $\overline{iq}$, where i is a natural number written in binary, and q an arbitrary binary string. On being presented with $\overline{iq}$, A first determines i and $|q|$; it then outputs

$$A(\overline{iq}) = \xi^{i+|q|}(i)q.$$

For $|q|$ large enough, $A(\overline{iq})$ is of the form

$$A(\overline{iq}) = x(i)w, \quad \text{for some } w.$$

Fix i ; then for $n > i$ large enough we have $K(x(n)) \leq K_A(x(n)) + c \leq (n - i) + 2\log i + c$. Hence

$$\forall i \exists n_0(i) \forall n \geq n_0(i) (n - K(x(n)) \geq i - 2\log i - c).$$

Because the right-hand side is unbounded, $\lim_{n \rightarrow \infty} (n - K(x(n))) = \infty$. $\square$

2.2.3. THEOREM (MARTIN-LÖF [19]). $\forall x \forall m \exists n \ K3(x(n)) \leq n - m$.

PROOF. Suppose that for some m , $\{x \mid \forall n \ K3(x(n)) > n - m\} \neq \emptyset$. By the basis theorem, this nonempty Π_1 set must contain a Δ_2 -definable element, in contradiction with the preceding theorem. $\square$

Martin-Löf's original proof of 2.2.3 essentially rests upon the following idea. Let

$f: \omega \rightarrow 2^{<\omega}$ be a recursive enumeration of $2^{<\omega}$. Then for any $x \in 2^\omega$, there are infinitely many initial segments $x(n+i) = x(n)w$, where $|w| = i$, such that $f(i) = x(n)$. It is perhaps of interest that 2.2.3 is also true for an entirely different reason.

2.3. Final remarks. In conclusion, we note that “randomness in mathematics” has been investigated at least since 1916, when Weyl proved that for irrational α , the sequence $(n\alpha)_{n \in \omega}$ is equidistributed mod 1. The general problem is to prove that certain simply defined sequences are random, where the concept of randomness in this context is specified by statistical criteria (such as equidistribution), without recourse to recursion theory. Examples of simply defined sequences which are strongly suspected to be random are

- the sequence whose n th term is 1 (0) if $(\frac{3}{2})^n \bmod 1$ is larger than (less than) $\frac{1}{2}$;
- Kolakoski’s 2–1 sequence [10];

but no one has been able to give a proof so far (see, e.g., Choquet [7] or Dekking [10]).

Observe that these sequences are recursive, hence nonrandom according to the usual definitions. This does not, however, diminish the interest of the above problems. There appear to be two ways to reconcile the theoretical definition of randomness with the practicing mathematician’s interest in this concept: either employing (variants of) generalized Kolmogorov complexity as in Ko [11], or else questioning the use of recursion theory in defining randomness altogether, as in [18].

REFERENCES

- [1] G. J. CHAITIN, *Information theoretic limitations on formal systems*, *Journal of the Association for Computing Machinery*, vol. 21 (1974), pp. 403–424.
- [2] ———, *A theory of program size formally identical to information theory*, *Journal of the Association for Computing Machinery*, vol. 2² (1975), pp. 329–340.
- [3] ———, *Algorithmic information theory*, *IBM Journal of Research and Development*, vol. 21 (1977), pp. 350–359, 496.
- [4] ———, *Gödel’s theorem and information*, *International Journal of Theoretical Physics*, vol. 21 (1982), pp. 941–954.
- [5] ———, *Incompleteness theorems for random reals*, *Advances in Applied Mathematics*, vol. 8 (1987), pp. 119–146.
- [6] ———, *Algorithmic information theory*, Cambridge Tracts in Computer Science, vol. 1, Cambridge University Press, Cambridge, 1987.
- [7] G. CHOQUET, *Repartition des nombres $k(3/2)^n \bmod 1$; mesures et ensembles associés*, *Comptes Rendus des Séances de l’Académie des Sciences, Séries A–B*, vol. 290 (1980), pp. A575–A580.
- [8] M. DAVIS, *What is a computation?* *Mathematics today* (L. A. Steen, editor), Springer-Verlag, Berlin, 1978, pp. 241–267.
- [9] M. DAVIS, Y. MATIJASEVIČ and J. ROBINSON, *Hilbert’s tenth problem. Diophantine equations: positive aspects of a negative solution*, *Mathematical developments arising from Hilbert problems* (F. E. Browder, editor), American Mathematical Society, Providence, Rhode Island, 1976, pp. 323–378.
- [10] F. M. DEKKING, *Regularity and irregularity of sequences generated by automata*, *Séminaire de théorie des nombres, 1979–1980*, Université de Bordeaux I, Talence, 1980, Exposé 9.
- [11] KEN-I KO, *On the notion of infinite pseudorandom sequences*, *Theoretical Computer Science*, vol. 48 (1986), pp. 9–33.
- [12] A. N. KOLMOGOROV, *On tables of random numbers*, *Sankhyā Series A*, vol. 25 (1963), pp. 369–376.
- [13] ———, *Three approaches to the definition of the concept of “amount of information”*, *Selected translations in mathematical statistics and probability*, vol. 7, American Mathematical Society, Providence, Rhode Island, 1968, pp. 293–302.

- [14] ———, *The logical basis for information theory and probability theory*, *IEEE Transactions on Information Theory*, vol. IT-14 (1968), pp. 662–664.
- [15] ———, *Combinatorial basis for information theory and probability theory*, *Russian Mathematical Surveys*, vol. 38 (1983), no. 4, pp. 29–40.
- [16] ———, *On logical foundations of probability theory*, *Probability theory and mathematical statistics (proceedings of the fourth USSR-Japan symposium, Tbilisi, 1982)*; K. Itô and J. V. Prokhorov, editors), *Lecture Notes in Mathematics*, vol. 1021, Springer-Verlag, Berlin, 1984, pp. 1–5.
- [17] G. KREISEL and A. LEVY, *Reflection principles and their use for establishing the complexity of axiomatic systems*, *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, vol. 14 (1968), pp. 97–142.
- [18] M. VAN LAMBALGEN, *Random sequences*, Ph.D. thesis, Department of Mathematics, University of Amsterdam, Amsterdam, 1987.
- [19] P. MARTIN-LÖF, a) *Algorithmen und zufällige Folgen*, lecture notes, University of Erlangen, Erlangen, 1966.
b) *Complexity oscillations in infinite binary sequences*, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 19 (1971), pp. 225–230.
- [20] ———, *The definition of random sequences*, *Information and Control*, vol. 9 (1966), pp. 602–619.
- [21] H. ROGERS, *Theory of recursive functions and effective computability*, McGraw-Hill, New York, 1967.
- [22] C. P. SCHNORR, *Zufälligkeit und Wahrscheinlichkeit*, *Lecture Notes in Mathematics*, vol. 218, Springer-Verlag, Berlin, 1971.
- [23] C. SMORYŃSKI, *The incompleteness theorems*, *Handbook of mathematical logic* (J. Barwise, editor), North-Holland, Amsterdam, 1977, pp. 821–865.
- [24] R. I. SOARE, *Recursively enumerable sets and degrees*, Springer-Verlag, Berlin, 1986.
- [25] T. TYMOCZKO (editor), *New directions in the philosophy of mathematics*, Birkhäuser, Boston, Massachusetts, 1986.
- [26] P. M. B. VITANYI and M. LI, *Two decades of applied Kolmogorov complexity*, Technical Report TR-08-88, Computer Center, Harvard University, Cambridge, Massachusetts, 1988.
- [27] J. FORD, *Directions in classical chaos*, *Directions in chaos*, vol. 1 (B. L. Hao, editor), World Scientific Publishing, Singapore, 1987, pp. 1–16.

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF AMSTERDAM
1018 WB AMSTERDAM, THE NETHERLANDS