On Random and Hard-to-Describe Numbers

Charles H. Bennett
IBM Watson Research Center Yorktown Heights, NY 10598, USA
3 May 1979

The first essay discusses, in nontechnical terms, the paradox implicit in defining a random integer as one without remarkable properties, and the resolution of that paradox at the cost of making randomness a property which most integers have but can't be proved to have. The second essay briefly reviews the search for randomness in the the digit sequences of natural irrational numbers like π and artificial ones like Champernowne's $C = 0.12345678910111213...$, and discusses at length Chaitin's definable-but-uncomputable number Ω, whose digit sequence is so random that no betting strategy could succeed against it. Other, Cabalistic properties of Ω are pointed out for the first time.

1 Berry's Paradox and the Unprovability of Randomness

The number 1,101,121 is unusual in that it is, or appears to be, the number named by the expression *the first number not nameable in under ten words*. However, since the italicized expression has only nine words, there is an inconsistency in regarding it as a name for 1,101,121 or any other number. This paradox, a variant of one due to Russell and Berry[1], shows that the concept of nameability or definability is too vague and powerful to be used without restriction. Because of it, the "function" $N(x) =$ *the number of English words required to name the integer* x must be regarded as ill-defined for all but finitely many x. Martin Gardner[2] has pointed out that a similar paradox arises when one attempts to classify numbers as "interesting" or "dull": there can be no dull numbers, because, if there were, the first of them would be interesting on that account.

Berry's paradox can be avoided and tamed by restricting nameability to mean describability as output of an algorithm or computer program. Consider the function $C(x) =$ *the number of bits in the smallest program to compute the integer* x. A (binary) integer p is said to be a program to compute x when some standard universal computer, given p as its sole input, computes x as its sole output, afterward halting. In this case there can be no doubt that p indeed describes x. Since every integer admits such a description, $C(x)$ is well-defined for all x. However, to avoid Berry's paradox, it must be concluded that the function $C(x)$ is itself uncomputable. For if $C(x)$ were computable, one could design a contradictory program q to find and print out the least number x for which $C(x)$ exceeded the number of bits in q.

Returning to the question of interesting and dull numbers, an interesting number may without paradox be defined as one computable by a program with fewer bits than the number itself. This short description would attest some special feature of the number, by which it could be distinguished from the general run of numbers. A dull or "random" number, on the other hand, would be one that is algorithmically incompressible. Obviously, most numbers are random in this sense, since, for any n, there are more than twice as many $\leq n$ -bit numbers as ($\leq n$–1) -bit numbers available to serve as shorter descriptions. Using this definition of randomness, G. Chaitin[3] demonstrated the following surprising fact, a form of Gödel's incompleteness theorem: *although most numbers are random, only finitely many of them can be proved random within a given consistent axiomatic system.* In particular, a system whose axioms and rules of inference require about n bits to describe cannot prove the randomness of any number much longer than n bits. If the system could prove randomness for a number much longer than n bits, the *first* such proof (first, that is, in an unending enumeration of all proofs obtainable by repeated application of the axioms and rules of inference) could be manipulated into a contradiction: an approximately n-bit program to find and print out the specific random number mentioned in this proof, a number whose smallest program is by definition considerably longer than n bits.

2 The Search for a "Random" Real Number

It has been conjectured that the decimal expansions of irrational numbers such as, π, e, and $\sqrt{2}$ are random in the sense of being "normal"[4] i.e. that each digit 0 through 9, and indeed each block of digits of any length, occurs with equal asymptotic frequency. It is easy to show that no rational number is normal to any base, and that almost all irrational numbers are normal to every base; but the normality of these most famous irrational numbers remains open. The question cannot be settled by any finite amount of statistical evidence, since an ultimately normal number might begin abnormally (e.g. $e = 2.718281828...$), or *vice versa*. Existing evidence[5] shows no significant departures from randomness in π. e also appears to be normal, though there is some evidence for other statistical irregularities[6].

In contrast to π, whose random-appearing digit sequence mocks the attempt to prove it so, the following very non-random number:

$$C = 0.12345678910111213141516171819202122232425262728293031...$$

is nevertheless provably normal, to base 10. This number, invented by D. G. Champernowne[7], consists of the decimal integers written in increasing order (Benoit Mandelbrot has pointed out another number of this sort, whose base-2 normality is implicit in an earlier paper by N. Wiener[8]). Departures from equidistribution are large in the initial portion of Champernowne's or Wiener's

number, but approach zero as the count is extended over more and more of the sequence. It is apparently not known whether these numbers are normal to every base.

Although the digit sequence of π may be random in the sense of being normal, it is definitely not random in the sense of being unpredictable: a good gambler betting against it would eventually infer its rule and thereafter always win, and only a very inept gambler could lose many bets against Champernowne's number. Is there a sequence so random that no computable betting strategy, betting against it at fair odds, can win an infinite gain? Any number that is random in this strong sense is also normal to every base. It is a basic result of probability theory that almost all real numbers are random in this strong sense[9], but here again we are seeking a *specific* random number.

There is, of course, a sense in which no specifically definable real number can be random. Since there are uncountably many real numbers but only countably many definitions, the mere fact that a real number is definable makes it atypical of real numbers in general. Here, however, we are only seeking a number whose atypicality is unrecognizable by constructive means. In particular, the number we are seeking must not be computable from its definition; since if it were, that would already imply a perfect betting strategy. One may define an uncomputable real number K in terms of the halting problem[1] for programs on some standard universal computer or programming language, setting the n'th binary digit of K to 1 or 0 according to whether the n'th program halts. Although the resulting digit sequence is indeed uncomputable, a gambler could nevertheless make infinite profit betting against it, by betting only on solvable cases of the halting problem, of which there are infinitely many. G. Chaitin[10] discovered a real number which is uncomputable in the stronger sense needed:

Ω = *the halting probability of a universal computer whose program is generated randomly, by tossing a fair coin whenever the computer requests another bit of input.*

Clearly, once the universal computer or programming language is specified, Ω is a well defined real number between zero and one. For typical programming languages like Fortran, Ω will be nearer one than zero, since a program generated at random is more likely to halt immediately (e.g. due to a syntax error) than to loop. However, it can be shown that after the first few digits Ω would look

[1] The halting problem, i.e. the problem of distinguishing programs that come to a spontaneous halt from those that run on indefinitely, is the classic unsolvable problem of computability theory. At first sight the problem might seem solvable since, if a program halts, that fact can certainly be demonstrated by running the program long enough. Moreover there are many programs which can easily be proven to halt or not to halt even without running the program. The difficulty comes not in solving particular cases, but in solving the problem in general. It can be shown that there is no effective prescription for deciding how long to run a program that waits long enough to reveal the halting of all halting programs, nor any consistent system of axioms strong enough to prove the non-halting of all non-halting ones. The unsolvability of the halting problem can be derived from and indeed is equivalent to the fact that most random integers can't be proved random.

quite random, far more than Champernowne's number.

Ω has three related properties that make it unusual:

1. It encodes the halting problem in a very compact form. Knowing its first few thousand digits would in principle permit the solution of all interesting finitely refutable mathematical conjectures.

2. It is algorithmically incompressible: there exists a constant c such that the first n bits of Ω are never expressible as the output of a program smaller than $n - c$ bits.

3. No computable gambling scheme can make infinite profit betting against it.

Ω encodes the halting problem, but in a much more compact form than K: knowing its first n bits is sufficient to solve the halting problem for any program up to n bits in length. Suppose one wishes to solve the halting problem for a particular n-bit program p. The program p corresponds to a particular sequence of n coin tosses having probability 2^{-n}, and, if it halts, contributes this amount of probability to the total halting probability Ω. Let Ω_n represent the known first n bits of Ω, so that

$$\Omega_n < \Omega \leq \Omega_n + 2^{-n}.$$

In order to decide the halting of p, begin an unending but systematic search for *all* programs that halt, of whatever length, running first one program then another for longer and longer times (cf. Fig. 1) until enough halting programs have been found to account for more than Ω_n of the total halting probability[2]. Then either p is among the programs that have halted so far, or else it will never halt, since its subsequent halting would drive the total halting probability above its known upper bound of $\Omega_n + 2^{-n}$. Note that there is apparently no way of using Ω to solve the halting problem for one n-bit program without solving the halting problem for all other $\leq n$-bit programs at the same time.

Most of the famous unproved conjectures of mathematics (Fermat's conjecture, Goldbach's conjecture, the extended Riemann hypothesis, and, until recently, the four-color problem) are conjectures of the nonexistence of something, and would be refuted by a single finite counterexample. Fermat's conjecture, for example, would be refuted by finding a solution to the equation $x^n + y^n = z^n$ in positive integers with $n > 2$; Riemann's hypothesis by finding a misplaced zero of the zeta function. Such conjectures are equivalent to the assertion that some program, which searches systematically for the allegedly nonexistent object, will never halt.

[2] If Ω were a terminating binary rational, the expansion ending in infinitely many ones should be used, making $\Omega_n < \Omega = \Omega_n + 2^{-n}$. In fact, this problem never arises, since, as will be proved presently, Ω is irrational, and so lies strictly between Ω_n and $\Omega_n + 2^{-n}$

Interesting conjectures of this sort are generally sufficiently simple to describe that they can be encoded in the halting of *small* programs, a few thousands or tens of thousands of bits long. Thus only the first few thousand digits of Ω would be needed in principle to solve these outstanding "finitely refutable" conjectures as well as any others of comparable simplicity that might be thought of in the future.[3]

An important class of statements decidable by Ω are statements of the form "proposition p is provable in axiomatic system $\mathbf{A}$". As indicated in the first essay, given a description of the axioms and rules of inference of $\mathbf{A}$, it is possible to effectively enumerate all possible proofs within the system, and hence all provable statements. Assuming that the proposition p and system $\mathbf{A}$ together require n bits to describe, there is a certain program of about n bits which will halt if and only if p is provable in $\mathbf{A}$. Thus, for any proposition p and axioms $\mathbf{A}$ simple enough to be "interesting", the first few thousand bits of Ω suffice to decide among the three possibilities: p is provable in $\mathbf{A}$, p is refutable in $\mathbf{A}$, or p is independent of $\mathbf{A}$. Another consequence of the enumerability of proofs,

[3]Some well-known conjectures, e.g. that π is normal, or that there are infinitely many twin primes (consecutive odd primes like 3 and 5 or 17 and 19), or that there are only finitely many primes of the form $2^n + 1$, are not in principle decidable one way or the other by any finite amount of direct evidence. Perhaps the most important conjecture of this sort is the $P \neq NP$ conjecture in computational complexity theory, which holds that there are problems for which the validity of a guessed solution can be tested quickly, but for which solutions cannot be found quickly. Logically, such higher level conjectures involve multiple quantifiers such as $\forall\exists$ "for infinitely many", while finitely refutable conjectures, i.e. those equivalent to the statement that a certain program will not halt, involve only a single quantifier $\forall$ "for all". Although higher level conjectures cannot be directly decided by Ω, there is good reason to believe that many of them, including most of the interesting ones, could be decided indirectly, as logical consequences of stronger, finitely refutable conjectures. For example, many twin primes are known, and empirical evidence indicates that the spacing between them grows rather slowly. Thus the twin prime conjecture may be viewed as an unnecessarily weak form of a stronger but still probably true assertion about the spacing of twin primes, say that there is always at least one pair of twin primes between 10^n and 10^{n+1}. This stronger conjecture would be decided by the early digits of Ω, since it is equivalent to the nonhalting of a simple program that looks for an excessively large gap the distribution of twin primes. Conversely, the assertion that there are only finitely many primes of the form $2^n + 1$ may be viewed as an unnecessarily weak form of the assertion that there are fewer than, say, 10^{100} such primes, or some other easily-named large number (in fact only six are known). Like the strengthened form of the twin prime conjecture, this assertion is equivalent to the non-halting of a simple program, one that looks for the 10^{100}'th prime of the specified form. Similarly, the normality of π and the inequality of P and NP would follow from stronger, finitely refutable statements supported by the same evidence as the original conjectures. In all these cases the finitely refutable statement is obtained by assuming a generous but computable bound on one of the original statement's existential quantifiers. Aside from these conjectures, which probably follow from finitely refutable ones, there are some mathematical statements that definitely cannot be reduced to halting problems. Typical of these are some statements about Ω itself, e.g. "the j'th bit of Ω is a 1". By virtue of the incompressibility of Ω, the first n members of this family of two-quantifier statements[10] cannot be decided by any algorithm smaller than about n bits. This implies, incidentally, that no irrational number can efficiently encode the decision of all higher level statements the way Ω encodes the decision of all finitely refutable ones.

as mentioned earlier, is Chaitin's form of Gödel's theorem: even though most integers are algorithmically random, no axiomatic system describable in n bits can prove randomness for integers much larger than n bits. Ω provides a strong converse to this theorem: its first n bits constitute a sufficient "axiom" to decide the randomness of all integers of $n+1$ bits or less. The procedure for doing this is essentially the same as that used to solve the halting problem: to find whether a given $n+1$ bit integer x is algorithmically random, use Ω_n as described earlier to find all $\leq n$-bit programs that halt. If none of these has x as its output, then by definition x is algorithmically random.

Let us now return to the senses in which Ω itself is random: its incompressibility and the impossibility of successfully gambling against it. It may appear strange that Ω can contain so much information about the halting problem and yet be computationally indistinguishable from a meaningless random sequence generated by tossing a coin. In fact, Ω is a totally informative message, a message which appears random because all redundancy has been squeezed out of it, a message which tells us only things we don't already know.

To show that Ω is incompressible, let p be a program that for some n computes Ω_n, the first n bits of Ω. This program may be altered, increasing its size by at most c bits (c a constant independent of n), so that instead of printing Ω_n it finds and prints out the first algorithmically random $(n+1)$-bit number, as explained above. This would be a contradiction unless the original program p were at least $n-c$ bits long.

No finitely describable computable gambling scheme can win an infinite profit betting against the bits of Ω. Let G be a gambling scheme, describable in g bits, and able to multiply the gambler's initial capital 2^k fold by betting on some number n of initial bits of Ω. Without loss of generality we may suppose that the scheme includes a specification of the desired gain 2^k, and that it quits as soon as this gain is achieved, making no further bets. One may imagine the same gambling scheme applied to other inputs besides Ω. On most of them it would fail to achieve its goal, but on some it would succeed. Indeed one may use G to enumerate *all* the finite inputs on which G would quit successfully. This set has total probability 2^{-k} or less, of which 2^{-n} is contributed by Ω_n.

It can be shown[10] that about $n-k$ bits suffice to locate Ω_n within this enumeration of successful inputs. Therefore Ω_n can be computed by a program of approximately $g+n-k$ bits. This means, in turn, that k cannot be much greater than g without violating the incompressibility of Ω. Therefore no g-bit gambling scheme betting on Ω can multiply the initial capital by more than about 2^g, the amount one would win by simply knowing g bits of Ω and betting only on those bits.

Throughout history philosophers and mystics 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 Bible, the Koran, the mythical secret books of Hermes Trismegistus, and the medieval Jewish Cabala have been so regarded. 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 and deeper questions than the user wishes to ask. Like God the esoteric book is simple yet undescribable, omniscient, and transforms all who know It. The use of classical texts to fortell mundane events is considered supersitious nowadays, yet, in another sense, science is in quest of its own Cabala, a concise set of natural laws which would explain all phenomena. In mathematics, where no set of axioms can hope to prove all true statements, the goal might be a concise axiomatization of all "interesting" true statements.

Ω 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. It embodies an enormous amount of wisdom in a very small space, inasmuch as its first few thousand 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, including all interesting finitely-refutable conjectures. Its wisdom is useless precisely *because* it is universal: the only known way of extracting from Ω the solution to one halting problem, say the Fermat conjecture, is by embarking on a vast computation that would at the same time yield solutions to all other equally simply-stated halting problems, a computation far too large to be carried out in practice. Ironically, although Ω cannot be computed, it might accidentally be generated by a random process, e.g. a series of coin tosses, or an avalanche that left its digits spelled out in the pattern of boulders on a mountainside. The initial few digits of Ω are thus probably already recorded somewhere in the universe. Unfortunately, no mortal discoverer of this treasure could verify its authenticity or make practical use of it.

The author has received reliable information, from a Source who wishes to remain anonymous, that the decimal expansion of Ω begins

$\Omega = 0.9999998020554253273471801908...$

References

1. Whitehead, A.N., and Russell, B., *Principia Mathematica*, Vol. **1**, Cambridge University Press, London (1925) p. 61.

2. Gardner, M., *Sci. Amer.* **198** (1958) No. 1, p. 92.

3. Chaitin, G., *J. Assoc. Comput. Mach.* **21** (1974) 403; *Sci. Amer.* **232** (1975) No. 5, p. 47.

4. Borel, E., *Rend. Circ. Mat. Palermo* **27** (1909) 247.

5. Pathria, R.K., *Mathematics of Computation* **16** (1962) 188.

6. Stoneham, R.G., *Amer. Math. Monthly* **72** (1965) 483.

7. Champernowne, D.G., *J. Lond. Math. Soc.* **8** (1933) 254.
8. Wiener, N., *Acta Math.* **55** (1930) 117, eq. 11.03.
9. Martin-Löf, P., *Information and Control* **9** (1966) 602.
10. Schnorr, C.P., *Math. Systems Theory* **5** (1971) 246.