

BOOK REVIEW

The Mathematical Intelligencer 14, No. 4 (Fall 1992), pp. 69–71

A Diary on Information Theory by Alfréd Rényi
Chichester: John Wiley & Sons, 1987; ix + 125 pp.
Hardcover, US\$54.95 (ISBN 0–471–90971–8)
Reviewed by Gregory J. Chaitin

Can the difficulty of an exam be measured by how many bits of information a student would need to pass it? This may not be so absurd in the encyclopedic subjects but in mathematics it doesn't make any sense since things follow from each other and, in principle, whoever knows the bases knows everything. All of the results of a mathematical theorem are in the axioms of mathematics in embryonic form, aren't they? *I will have to think this over some more.*

A. Rényi (*A Diary on Information Theory*, p. 31)

This remarkable quotation comes from Rényi's unfinished 1969 manuscript, written in the form of a fictitious student's diary. This "diary" comprises the bulk of Rényi's posthumous work, *A Diary on*

Copyright © 1992, Springer-Verlag New York Inc., reprinted by permission.

Information Theory, a stimulating introduction to information theory and an essay on the mathematical notion of information, a work left incomplete at Rényi's death in 1970 at the age of 49.

Alfréd Rényi was a member of the Hungarian Academy of Sciences. The *Diary*, as well as the material on information theory in his two books on probability theory [1, 2], attest to the importance he attached to the idea of information. This *Diary* also illustrates the importance that Rényi ascribed to wide-ranging nontechnical discussions of mathematical ideas as a way to interest students in mathematics. He believed the discussions served as vital teaching tools and stimuli for further research.

Rényi was part of the tidal wave of interest in information theory provoked by Claude Shannon's publications in the 1940s. The many papers with titles like "Information Theory, Photosynthesis, and Religion" actually published illustrate the tremendous and widespread initial interest in information theory.

When Rényi wrote his *Diary*, the initial wave of interest in information theory was dying out. In fact, Rényi was unaware of a second major wave of interest in information theory slowly beginning to gather momentum in the 1960s. At that time, Andrei Kolmogorov and I independently proposed a new **algorithmic** information theory to capture mathematically the notion of a random, patternless sequence as one that is algorithmically incompressible.

The development of this new information theory was not as dramatically abrupt as was the case with Shannon's version. It was not until the 1970s that I corrected the initial definitions. The initial definitions Kolmogorov and I proposed had serious technical deficiencies which led to great mathematical awkwardness. It turned out that a few changes in the definitions led to a revised algorithmic information theory whose elegant formulas closely mirror those in Shannon's original theory in a radically altered interpretation [3].

In the 1970s I also began to apply algorithmic information theory to extend and broaden Gödel's incompleteness theorem, culminating in the 1980s in an explicit constructive proof that there is randomness in arithmetic [4]. (For recent discussions of algorithmic information theory directed to the general scientific public, see [5–16].)

Rényi's *Diary* stops at the brink between Shannon's ensemble infor-

mation theory and the newer algorithmic information theory applying to individual sequences. With the benefit of hindsight, one can detect the germ of ideas that, if Rényi had pursued them properly, might have led him in the direction of algorithmic information theory.

Let us take the quotation at the head of this review. If Rényi had developed it properly, it might have led him to my insight that incompleteness can be obtained very naturally via metatheorems whose spirit can be summarized in the phrase, “a theorem cannot contain more information than the axioms from which it is deduced.” I think this new information-theoretic viewpoint makes incompleteness seem a much more menacing barrier than before.

A second instance occurs later in Rényi’s *Diary*, p. 41:

Therefore, the method of investigating the redundancy of a text by erasing and reconstruction is not appropriate. By this method, we would get a correct estimation of the real redundancy only if the reconstruction could be done by a computer. In that case, the meaning of the text wouldn’t be a factor because a computer wouldn’t understand it and could reconstruct it only by means of a dictionary and grammatical rules.

If Rényi could have formalized this, perhaps he might have discovered the complexity measure used in algorithmic information theory. (In algorithmic information theory, the complexity of a string or sequence of symbols is defined to be the size of the smallest computer program for calculating that string of symbols.)

So Rényi’s *Diary* balances on the edge between the old and the new versions of information theory. It also touches on connections between information theory and physics and biology that are still the subject of research [7, 8].

In what remains of this review, I would like to flesh out the above remarks by discussing Hilbert’s tenth problem in the light of algorithmic information theory. I will end with a few controversial remarks about the potential significance of these information-theoretic metamathematical results, and their connection with experimental mathematics and the quasi-empirical school of thought regarding the foundations of mathematics.

Consider a diophantine equation

$$P(k, x_1, x_2, \dots) = 0$$

with parameter k . Ask the question, “Does $P(k) = 0$ have a solution?”

Let

$$q = q_0q_1q_2 \dots$$

be the infinite bit string with

$$q_k = \begin{cases} 0 & \text{if } P(k) = 0 \text{ has no solution} \\ 1 & \text{if } P(k) = 0 \text{ has a solution.} \end{cases}$$

Let

$$q^n = q_0q_1 \dots q_{n-1}$$

be the string of the first n bits of the infinite string q , that is, the string of answers to the first n questions. Let $H(q^n)$ be the complexity of q^n , that is, the size in bits of the smallest program for computing q^n .

If Hilbert had been right and every mathematical question had a solution, then there would be a finite set of axioms from which one could deduce whether $P(k) = 0$ has a solution or not for each k . We would then have

$$H(q^n) \leq H(n) + c.$$

The c bits are the finite amount of information in our axioms, and this inequality asserts that if one is given n , using the axioms one can compute q^n , that is, decide which among the first n cases of the diophantine equation have solutions and which do not. Thus, the complexity $H(q^n)$ of answering the first n questions would be at most order of $\log n$ bits. We ignore the immense **time** it might take to deduce the answers from the axioms; we are concentrating on the amount of **information** involved.

In 1970, Yuri Matijasevič showed that there is no algorithm for deciding if a diophantine equation can be solved. However, if we are told the number m of equations $P(k) = 0$ with $k < n$ that have a solution, then we can eventually determine which do and which do not. This shows that

$$H(q^n) \leq H(n) + H(m) + c'$$

for some $m \leq n$, which implies that the complexity $H(q^n)$ of answering the first n questions **is still** at most order of $\log n$ bits. So from an information-theoretic point of view, Hilbert's tenth problem, while undecidable, does not look too difficult.

In 1987, I explicitly constructed [4] an exponential diophantine equation

$$L(k, x_1, x_2, \dots) = R(k, x_1, x_2, \dots)$$

with a parameter k . This equation gives complete randomness as follows. Ask the question, "Does $L(k) = R(k)$ have infinitely many solutions?" Now let

$$q = q_0q_1q_2 \dots$$

be the infinite bit string with

$$q_k = \begin{cases} 0 & \text{if } L(k) = R(k) \text{ has finitely many solutions} \\ 1 & \text{if } L(k) = R(k) \text{ has infinitely many solutions.} \end{cases}$$

As before, let

$$q^n = q_0q_1 \dots q_{n-1}$$

be the string of the first n bits of the infinite string q , that is, the string of answers to the first n questions. Let $H(q^n)$ be the complexity of q^n , that is, the size in bits of the smallest program for computing q^n . Now we have

$$H(q^n) \geq n - c'',$$

that is, the string of answers to the first n questions q^n is irreducible mathematical information and the infinite string of answers $q = q_0q_1q_2 \dots$ is now algorithmically random.

Surprisingly, Hilbert was wrong to assume that every mathematical question has a solution. The above exponential diophantine equation yields an infinite series of independent irreducible mathematical facts. It yields an infinite series of questions which reasoning is impotent to answer because the only way to answer these questions is to assume each individual answer as a new axiom! Here one can get out as theorems only what one explicitly puts in as axioms, and reasoning is completely useless! I think this information-theoretic approach to incompleteness makes incompleteness look much more natural and pervasive than has previously been the case. Algorithmic information theory

provides some theoretical justification for the experimental mathematics made possible by the computer and for the new quasi-empirical view of the philosophy of mathematics that is displacing the traditional formalist, logicist, and intuitionist positions [5].

References

1. Alfréd Rényi, Introduction to information theory, *Probability Theory*, Amsterdam: North-Holland (1970), 540–616.
2. Alfréd Rényi, Independence and information, *Foundations of Probability*, San Francisco: Holden-Day (1970), 146–157.
3. Gregory J. Chaitin, A theory of program size formally identical to information theory, *Information, Randomness & Incompleteness—Papers on Algorithmic Information Theory*, Second Edition, Singapore: World Scientific (1990), 113–128.
4. Gregory J. Chaitin, *Algorithmic Information Theory*, Cambridge: Cambridge University Press (1987).
5. John L. Casti, Proof or consequences, *Searching for Certainty*, New York: Morrow (1990), 323–403.
6. Gregory J. Chaitin, A random walk in arithmetic, *The New Scientist Guide to Chaos* (Nina Hall, ed.), Harmondsworth: Penguin (1991), 196–202.
7. David Ruelle, Complexity and Gödel’s theorem, *Chance and Chaos*, Princeton: Princeton University Press (1991), 143–149.
8. David Ruelle, Complexité et théorème de Gödel, *Hasard et Chaos*, Paris: Odile Jacob (1991), 189–196.
9. Luc Brisson and F. Walter Meyerstein, Que peut nous apprendre la science?, *Inventer L’Univers*, Paris: Les Belles Lettres (1991), 161–197.

10. Gregory J. Chaitin, Le hasard des nombres, *La Recherche* 22 (1991) no. 232, 610–615.
11. John A. Paulos, Complexity of programs, Gödel and his theorem, *Beyond Numeracy*, New York: Knopf (1991), 47–51, 95–97.
12. John D. Barrow, Chaotic axioms, *Theories of Everything*, Oxford: Clarendon Press (1991), 42–44.
13. Tor Nørretranders, Uendelige algoritmer, *Mærk Verden*, Denmark: Gyldendal (1991), 65–91.
14. Martin Gardner, Chaitin's omega, *Fractal Music, Hypercards and More...*, New York: Freeman (1992), 307–319.
15. Paul Davies, The unknowable, *The Mind of God*, New York: Simon & Schuster (1992), 128–134.
16. Gregory J. Chaitin, Zahlen und Zufall, *Naturwissenschaft und Weltbild* (Hans-Christian Reichel and Enrique Prat de la Riba, eds.), Vienna: Verlag Hölder–Pichler–Tempsky (1992), 30–44.

IBM Research Division
Yorktown Heights, NY 10598 USA