TO A MATHEMATICAL DEFINITION OF “LIFE”

ACM SICACT News, No. 4
(January 1970), pp. 12–18

G. J. Chaitin

Abstract

“Life” and its “evolution” are fundamental concepts that have not yet been formulated in precise mathematical terms, although some efforts in this direction have been made. We suggest a possible point of departure for a mathematical definition of “life.” This definition is based on the computer and is closely related to recent analyses of “inductive inference” and “randomness.” A living being is a unity; it is simpler to view a living organism as a whole than as the sum of its parts. If we want to compute a complete description of the region of space-time that is a living being, the program will be smaller in size if the calculation is done all together, than if it is done by independently calculating descriptions of parts of the region and then putting them together.

1. The Problem

"Life" and its "evolution" from the lifeless are fundamental concepts of science. According to Darwin and his followers, we can expect living organisms to evolve under very general conditions. Yet this theory has never been formulated in precise mathematical terms. Supposing Darwin is right, it should be possible to formulate a general definition of "life" and to prove that under certain conditions we can expect it to "evolve." If mathematics can be made out of Darwin, then we will have added something basic to mathematics; while if it cannot, then Darwin must be wrong, and life remains a miracle which has not been explained by science.

The point is that the view that life has spontaneously evolved, and the very concept of life itself, are very general concepts, which it should be possible to study without getting involved in, for example, the details of quantum chemistry. We can idealize the laws of physics and simplify them and make them complete, and then study the resulting universe. It is necessary to do two things in order to study the evolution of life within our model universe. First of all, we must define "life," we must characterize a living organism in a precise fashion. At the same time it should become clear what the complexity of an organism is, and how to distinguish primitive forms of life from advanced forms. Then we must study our universe in the light of the definition. Will an evolutionary process occur? What is the expected time for a certain level of complexity to be reached? Or can we show that life will probably not evolve?

2. Previous Work

Von Neumann devoted much attention to the analysis of fundamental biological questions from a mathematical point of view.[1] He considered

---

[1] See in particular his fifth lecture delivered at the University of Illinois in December of 1949, "Re-evaluation of the problem of complicated automata—Problems of hierarchy and evolution," and his unfinished *The Theory of Automata: Construction, Reproduction, Homogeneity.* Both are posthumously published in von Neumann (1966).

a universe consisting of an infinite plane divided into squares. Time is quantized, and at any moment each square is in one of 29 states. The state of a square at any time depends only on its previous state and the previous states of its four neighboring squares. The universe is homogeneous; the state transitions of all squares are governed by the same law. It is a deterministic universe. Von Neumann showed that a self-reproducing general-purpose computer can exist in his model universe.

A large amount of work on these questions has been done since von Neumann's initial investigations, and a complete bibliography would be quite lengthy. We may mention Moore (1962), Arbib (1966,1967), and Codd (1968).

The point of departure of all this work has been the identification of "life" with "self-reproduction," and this identification has both helped and hindered. It has helped, because it has not allowed fundamental conceptual difficulties to tie up work, but has instead permitted much that is very interesting to be accomplished. But it has hindered because, in the end, these fundamental difficulties must be faced. At present the problem has evidenced itself as a question of "good taste." As von Neumann remarks,[2] good taste is required in building one's universe. If its elementary parts are assumed to be very powerful, self-reproduction is immediate. Arbib (1966) is an intermediate case.

What is the relation between self-reproduction and life? A man may be sterile, but no one would doubt he is alive. Children are not identical to their parents. Self-reproduction is not exact; if it were, evolution would be impossible. What's more, a crystal reproduces itself, yet we would not consider it to have much life. As von Neumann comments,[3] the matter is the other way around. We can deduce self-reproduction as a property which must be possessed by many living beings, if we ask ourselves what kinds of living beings are likely to be around. Obviously, a species that did not reproduce would die out. Thus, if we ask what kinds of living organisms are likely to evolve, we can draw conclusions concerning self-reproduction.

---

[2] See pages 76–77 of von Neumann (1966).

[3] See page 78 of von Neumann (1966).

3. Simplicity and Complexity

"Complexity" is a concept whose importance and vagueness von Neumann emphasized many times in his lectures.[4] Due to the work of Solomonoff, Kolmogorov, Chaitin, Martin-Löf, Willis, and Loveland, we now understand this concept a great deal better than it was understood while von Neumann worked. Obviously, to understand the evolution of the complexity of living beings from primitive, simple life to today's very complex organisms, we need to make precise a measure of complexity. But it also seems that perhaps a precise concept of complexity will enable us to define "living organism" in an exact and general fashion. Before suggesting the manner in which this may perhaps be done, we shall review the recent developments which have converted "simplicity" and "complexity" into precise concepts.

We start by summarizing Solomonoff's work.[5] Solomonoff proposes the following model of the predicament of the scientist. A scientist is continually observing increasingly larger initial segments of an infinite sequence of 0's and 1's. This is his experimental data. He tries to find computer programs which compute infinite binary sequences which begin with the observed sequence. These are his theories. In order to predict his future observations, he could use any of the theories. But there will always be one theory that predicts that all succeeding observations will be 1's, as well as others that take more account of the previous observations. Which of the infinitely many theories should he use to make the prediction? According to Solomonoff, the principle that the simplest theory is the best should guide him.[6] What is the simplicity of a theory in the present context? It is the size of the computer program. Larger computer programs embody more complex theories, and smaller programs embody simpler theories.

Willis has further studied the above proposal, and also has introduced the idea of a hierarchy of finite approximations to it. To my

---

[4] See especially pages 78–80 of von Neumann (1966).

[5] The earliest generally available appearance in print of Solomonoff's ideas of which we are aware is Minsky's summary of them on pages 41–43 of Minsky (1962). A more recent reference is Solomonoff (1964).

[6] Solomonoff actually proposes weighing together all the theories into the prediction, giving the simplest theories the largest weight.

knowledge, however, the success which predictions made on this basis will have has not been made completely clear.

We must discuss a more technical aspect of Solomonoff's work. He realized that the simplicity of theories, and thus also the predictions, will depend on the computer which one is using. Let us consider only computers whose programs are finite binary sequences, and measure the size of a binary sequence by its length. Let us denote by $C(T)$ the complexity of a theory T. By definition, $C(T)$ is the size of the smallest program which makes our computer C compute T. Solomonoff showed that there are "optimal" binary computers C that have the property that for any other binary computer C', $C(T) \leq C'(T) + d$, for all T. Here d is a constant that depends on C and C', not on T. Thus, these are the most efficient binary computers, for their programs are shortest. Any two of these optimal binary computers C_1 and C_2 result in almost the same complexity measure, for from $C_1(T) \leq C_2(T) + d_{12}$ and $C_2(T) \leq C_1(T) + d_{21}$, it follows that the difference between $C_1(T)$ and $C_2(T)$ is bounded. The optimal binary computers are transparent theoretically, they are enormously convenient from the technical point of view. What's more, their optimality makes them a very natural choice.[7] Kolmogorov and Chaitin later independently hit upon the same kind of computer in their search for a suitable computer upon which to base a definition of "randomness."

However, the naturalness and technical convenience of the Solomonoff approach should not blind us to the fact that it is by no means the only possible one. Chaitin first based his definition of randomness on Turing machines, taking as the complexity measure the number of states in the machine, and he later used bounded-transfer Turing machines. Although these computers are quite different, they lead to similar definitions of randomness. Later it became clear that using the usual 3-tape-symbol Turing machine and taking its size to be the number of states leads to a complexity measure $C_3(T)$ which is asymptotically just a Solomonoff measure $C(T)$ with its scale changed: $C(T)$ is asymptotic to $2C_3(T) \log_2 C_3(T)$. It appears that people interested in computers may still study other complexity measures, but to apply

---

[7] Solomonoff's approach to the size of programs has been extended in Chaitin (1969a) to the speed of programs.

these concepts of simplicity/complexity it is at present most convenient to use Solomonoff measures.

We now turn to Kolmogorov's and Chaitin's proposed definition of randomness or patternlessness. Let us consider once more the scientist confronted by experimental data, a long binary sequence. This time he in not interested in predicting future observations, but only in determining if there is a pattern in his observations, if there is a simple theory that explains them. If he found a way of compressing his observations into a short computer program which makes the computer calculate them, he would say that the sequence follows a law, that it has pattern. But if there is no short program, then the sequence has no pattern—it is random. That is to say, the complexity $C(S)$ of a finite binary sequence S is the size of the smallest program which makes the computer calculate it. Those binary sequences S of a given length n for which $C(S)$ is greatest are the most complex binary sequences of length n, the random or patternless ones. This is a general formulation of the definition. If we use one of Solomonoff's optimal binary computers, this definition becomes even clearer. Most binary sequences of any given length n require programs of about length n. These are the patternless or random sequences. Those binary sequences which can be compressed into programs appreciably shorter than themselves are the sequences which have pattern. Chaitin and Martin-Löf have studied the statistical properties of these sequences, and Loveland has compared several variants of the definition.

This completes our summary of the new rigorous meaning which has been given to simplicity/complexity—the complexity of something is the size of the smallest program which computes it or a complete description of it. Simpler things require smaller programs. We have emphasized here the relation between these concepts and the philosophy of the scientific method. In the theory of computing the word "complexity" is usually applied to the speed of programs or the amount of auxiliary storage they need for scratch-work. These are completely different meanings of complexity. When one speaks of a simple scientific theory, one refers to the fact that few arbitrary choices have been made in specifying the theoretical structure, not to the rapidity with which predictions can be made.

4. What is Life?

Let us once again consider a scientist in a hypothetical situation. He wishes to understand a universe very different from his own which he has been observing. As he observes it, he comes eventually to distinguish certain objects. These are highly interdependent regions of the universe he is observing, so much so, that he comes to view them as wholes. Unlike a gas, which consists of independent particles that do not interact, these regions of the universe are unities, and for this reason he has distinguished them as single entities.

We believe that the most fundamental property of living organisms is the enormous interdependence between their components. A living being is a unity; it is much simpler to view it as a whole than as the sum of parts. That is to say, if we want to compute a complete description of a region of space-time that is a living being, the program will be smaller in size if the calculation is done all together, than if it is done by independently calculating descriptions of parts of the region and then putting them together. What is the complexity of a living being, how can we distinguish primitive life from complex forms? The interdependence in a primitive unicellular organism is far less than that in a human being.

A living being is indeed a unity. All the atoms in it cooperate and work together. If Mr. Smith is afraid of missing the train to his office, all his incredibly many molecules, all his organs, all his cells, will be cooperating so that he finishes breakfast quickly and runs to the train station. If you cut the leg of an animal, all of it will cooperate to escape from you, or to attack you and scare you away, in order to protect its leg. Later the wound will heal. How different from what happens if you cut the leg of a table. The whole table will neither come to the defense of its leg, nor will it help it to heal. In the more intelligent living creatures, there is also a very great deal of interdependence between an animal's past experience and its present behavior; that is to say, it learns, its behavior changes with time depending on its experiences. Such enormous interdependence must be a monstrously rare occurrence in a universe, unless it has evolved gradually.

In summary, the case is the whole versus the sum of its parts. If both are equally complex, the parts are independent (do not interact).

If the whole is very much simpler than the sum of its parts, we have the interdependence that characterizes a living being.[8] Note finally that we have introduced something new into the study of the size of programs (= complexity). Before we compared the sizes of programs that calculate different things. Now we are interested in comparing the sizes of programs that calculate the same things in different ways. That is to say, the method by which a calculation is done is now of importance to us; in the previous section it was not.

5. Numerical Examples

In this paper, unfortunately, we can only suggest a possible point of departure for a mathematical definition of life. A great amount of work must be done; it is not even clear what is the formal mathematical counterpart to the informal definition of the previous section. A possibility is sketched here.

Consider a computer C_1 which accepts programs P which are binary sequences consisting of a number of subsequences $B, C, P_1, \ldots, P_k, A$.

B, the leftmost subsequence, is a program for breaking the remainder of P into $C, P_1, \ldots, P_k$, and A. B is self-describing; it starts with a binary sequence which results from writing the length of B in base-two notation, doubling each of its bits, and then placing a pair of unequal bits at the right end. Also, B is not allowed to see whether any of the remaining bits of P are 0's or 1's, only to separate them into groups.[9]

C is the description of a computer C_2. For example, C_2 could be one of Solomonoff's optimal binary computers, or a computer which emits the program without processing it.

$P_1, \ldots, P_k$ are programs which are processed by k different copies of the computer C_2. $R_1, \ldots, R_k$ are the resulting outputs. These outputs would be regions of space-time, a space-time which, like von Neumann's, has been cut up into little cubes with a finite number of states.

A is a program for adding together $R_1, \ldots, R_k$ to produce R, a single region of space-time. A merely juxtaposes the intermediate results

---

[8] The whole cannot be more complex than the sum of its parts, because one of the ways of looking at it is as the sum of its parts, and this bounds its complexity.

[9] The awkwardness of this part of the definition is apparently its chief defect.

$R_1, \ldots, R_k$ (perhaps with some overlapping); it is not allowed to change any of the intermediate results. In the examples below, we shall only compute regions R which are one-dimensional strings of 0's and 1's, so that A need only indicate that R is the concatenation of $R_1, \ldots, R_k$, in that order.

R is the output of the computer C_1 produced by processing the program P.

We now define a family of complexity measures $C(d, R)$, the complexity of a region R of space-time when it is viewed as the sum of independent regions of diameter not greater than d. $C(d, R)$ is the length of the smallest program P which makes the computer C_1 output R, among all those P such that the intermediate results R_1 to R_k are all less than or equal to d in diameter. $C(d, R)$ where d equals the diameter of R is to within a bounded difference just the usual Solomonoff complexity measure. But as d decreases, we may be forced to forget any patterns in R that are more than d in diameter, and the complexity $C(d, R)$ increases.

We present below a table with four examples. In each of the four cases, R is a 1-dimensional region, a binary sequence of length n. R_1 is a random binary sequence of length n ("gas"). R_2 consists of n repetitions of 1 ("crystal"). The left half of R_3 is a random binary sequence of length $n/2$. The right half of R_3 is produced by rotating the left half about R_3's midpoint ("bilateral symmetry"). R_4 consists of two identical copies of a random binary sequence of length $n/2$ ("twins").

$C(d, R) =$ approx. ?	$R = R_1$ "gas"	$R = R_2$ "crystal"	$R = R_3$ "bilateral symmetry"	$R = R_4$ "twins"
$d = n$	n	$\log_2 n$ Note 1	$n/2$	$n/2$
$d = n/k$ ($k > 1$ fixed, n large)	n	$k \log_2 n$ Notes 1,2	$n - (n/2k)$ Note 2	n Note 2
$d = 1$	n	n	n	n

Note 1. This supposes that n is represented in base-two notation by a random binary sequence. These values are too high in those rare cases where this is not true.

Note 2. These are conjectured values. We can only show that $C(d, R)$ is approximately less than or equal to these values.

Bibliography

- Arbib, M. A. (1962). "Simple self-reproducing automata," *Information and Control.*
- Arbib, M. A. (1967). "Automata theory and development: Part 1," *Journal of Theoretical Biology.*
- Arbib, M. A. "Self-reproducing automata—some implications for theoretical biology."
- Biological Science Curriculum Study. (1968). *Biological Science: Molecules to Man,* Houghton Mifflin Co.
- Chaitin, G. J. (1966). "On the length of programs for computing finite binary sequences," *Journal of the Association for Computing Machinery.*
- Chaitin, G. J. (1969a). "On the length of programs for computing finite binary sequences: Statistical considerations," *ibid.*
- Chaitin, G. J. (1969b). "On the simplicity and speed of programs for computing infinite sets of natural numbers," *ibid.*
- Chaitin, G. J. (1970). "On the difficulty of computations," *IEEE Transactions on Information Theory.*
- Codd, E. F. (1968). *Cellular Automata.* Academic Press.
- Kolmogorov, A. N. (1965). "Three approaches to the definition of the concept 'amount of information'," *Problemy Peredachi Informatsii.*
- Kolmogorov, A. N. (1968). "Logical basis for information theory and probability theory," *IEEE Transactions on Information Theory.*

- Loveland, D. W. "A variant of the Kolmogorov concept of complexity," report 69-4, Math. Dept., Carnegie-Mellon University.
- Loveland, D. W. (1969). "On minimal program complexity measures," *Conference Record of the ACM Symposium on Theory of Computing,* May 1969.
- Martin-Löf, P. (1966). "The definition of random sequences," *Information and Control.*
- Minsky, M. L. (1962). "Problems of formulation for artificial intelligence," *Mathematical Problems in the Biological Science,* American Math. Society.
- Moore, E. F. (1962). "Machine models of self-reproduction," *ibid.*
- von Neumann, J. (1966). *Theory of Self-Reproducing Automata.* (Edited by A. W. Burks.) University of Illinois Press.
- Solomonoff, R. J. (1964). "A formal theory of inductive inference," *Information and Control.*
- Willis, D. G. (1969). "Computational complexity and probability constructions," Stanford University.