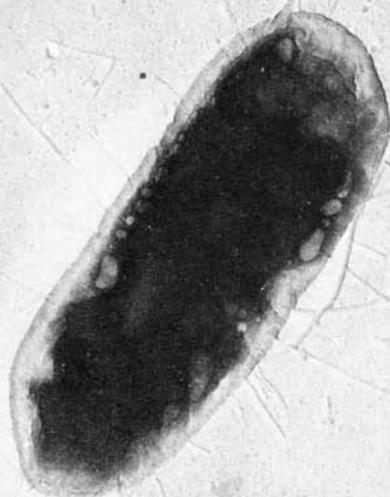


The Bhaktivedanta Institute
Monograph Series Number 2

Demonstration By Information Theory That Life Cannot Arise From Matter

by
Richard L. Thompson
(Sadāputa Dāsa Adhikārī)



**Demonstration By Information Theory
That Life Cannot Arise From Matter**

**Demonstration by Information Theory
That Life Cannot Arise From Matter***

by

**Richard Thompson, Ph.D.
(Sadāputa Dāsa Adhikārī)**

The Bhaktivedanta Institute Monograph Series

Number 2

***This monograph forms part of a forthcoming book, *The Origin of Life and Matter*, by Thoudam D. Singh, Michael Marchetti, and Richard Thompson.**

Readers interested in the subject matter of this monograph are invited to send correspondence to the Bhaktivedanta Institute at the following addresses.

70 Commonwealth Avenue
Boston, Massachusetts 02116
U.S.A. (617) 266-8369

Hare Krishna Land
Juhu, Bombay 400 054
India (Phone: 57-9373)

Bhaktivedanta Gurukula and
Institute for Higher Studies
Bhaktivedanta Swami Marg
Vrindavana, Mathura
India

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without permission in writing from the Publisher.

© 1977 Bhaktivedanta Book Trust

Printed in the United States of America

The Bhaktivedanta Institute

Founder — Ācārya

His Divine Grace A. C. Bhaktivedanta Swami Prabhupāda

Editorial Board.

Thoudam D. Singh, Ph.D.
(Svarūpa Dāmodara Dāsa Brahmācārī)

Robert Corens
(Rupānuga Dāsa Adhikārī)

Michael Marchetti, Ph.D.
(Mādhava Dāsa Brahmācārī)

Richard Thompson, Ph.D.
(Sadāputa Dāsa Adhikārī)

Editor.

William Deadwyler, M.A.
(Ravindra Svarūpa Dāsa Adhikārī)

Dedicated to His Divine Grace

A. C. Bhaktivedanta Swami Prabhupāda

*om ajñāna-timirāndhasya jñānāñjana-śalākayā
cakṣur unmlitaṁ yena tasmai śrī-gurave namaḥ*

About Bhaktivendanta Institute

Bhaktivendanta Institute is a center for advanced study and research into the Vedic scientific knowledge concerning the nature of consciousness and the self. The Institute is the academic division of the International Society for Krishna Consciousness. It consists of a body of scientists and scholars who have recognized the unique value of the teachings of Krishna Consciousness brought to the West by His Divine Grace A. C. Bhaktivendanta Swami Prabhupāda. The main purpose of the Institute is to explore the implications of the Vedic knowledge as it bears on all features of human culture, and to present its findings in courses, lectures, monographs, books, and a quarterly journal, *Sa-vijñānam*.

The Institute presents modern science and other fields of knowledge in the light of Vaiṣṇava philosophy and tradition, providing a new perspective on reality quite different from that of our modern educational systems. One reason for the increasing interest of young intellectuals in Śrīla Prabhupāda's teachings is doubtlessly the growing awareness that in spite of great scientific and technological advancements, the real goal of human life has somehow been missed. The philosophy of Bhaktivendanta Institute provides a meaningful answer to this concern, based on the fundamental premise that life—not matter—is the basis of the world we perceive.

The central doctrine of modern science is that all phenomena, including those of life and consciousness, can be fully explained and understood by recourse to matter alone. The dictum that "life is a manifestation of matter" is, indeed, the ultimate rational for the entire civilization of material aggrandizement. The Vedas, on the other hand, teach that conscious life is original, fundamental, and eternal. This is the essence of *Bhagavad-gītā*—"aham sarvasya prabhavo mattaḥ sarvaṁ pravartate." (10.8) On this fundamental and critical point, modern science and Vedic knowledge find themselves opposed.

Bhaktivendanta Institute is dedicated to disseminating this most fundamental knowledge throughout the world. The Institute is clearly demonstrating that the Vedic version is not a matter simply of "faith" or "belief", but is scientific in the strict sense of the term. Although many of its features may appear difficult to verify experimentally, others have direct implications concerning what

we may expect to observe. Thus, this view should serve as a stimulating challenge to the truly scientific spirit to go beyond the very restrictive framework imposed on our scientific understanding of nature over the last two hundred years. Modern science began as an experiment to see how far nature could be explained without invoking God. But the purpose of Bhaktivedanta Institute is to introduce Vedic knowledge on a genuinely scientific basis for the first time in the history of this modern scientific age.

CONTENTS

Introduction	1
Section 1. The laws of nature, as understood by modern physics, have a low information content.	4
Section 2. The bodily structures of living organisms have a high information content.	17
Section 3. No system of high information content can evolve from a system of low information content.	29
Section 4. What is the origin of life? Some implications of this analysis.	39
Appendix. Mathematical derivations.	46
Notes	57
Bibliography	59

Introduction

In this paper we will show how the mathematical analysis of information content demonstrates that the laws of nature as understood by modern science are not sufficient to account for the life we see about us. The basic argument is this: The laws of nature and the corresponding mathematical models of physical reality can all be described by a few simple equations and other numerical expressions. This means that they possess a low information content. On the other hand, the intricate and variegated forms of living organisms possess a high information content. It can be shown that configurations of high information content cannot arise with substantial probabilities, in models defined by mathematical expressions of low information content. It therefore follows that life forms could not arise by the action of the kind of natural laws or fundamental causal agents which are considered in modern science.

This argument has the particular feature of bypassing the proposition made by evolutionary theorists that even though the steps leading to life are improbable, they are nonetheless likely to happen, given the immense spans of geological time available. We show that no period of time from zero to billions of billions of years will suffice to make the evolution of life from matter by chance and natural selection a probable affair. Rather, we show that the probability of the evolution of life forms over periods many times that of the estimated 4.5-billion-year age of the earth is bounded by numerical limits of the order of $64^{-80,000}$, an almost infinitesimal number. This implies that the entire history of the earth would have to be repeated over and over again on the order of $64^{80,000}$ times in order for there to be a substantial chance that complex living entities would evolve even once.

An intuitive reason for the impossibility of organic evolution lies at the basis of these figures. It can be shown that the process of natural selection—the alleged mechanism of evolution—must have specific direction in order to bring about the evolution of complex living organisms. Without this direction, this process is unable to discriminate among random events (mutations) in such a way as to bring complex order out of chaos. For this reason, the standard argument that evolution will occur, given large enough time spans, is false. Natural selection lacks direction because the

basic cause of this selection, namely the laws of nature, has too low a level of information content to specify such direction.

More generally, it is shown that in a mathematical system, any information which can be extracted from the system after transformations corresponding to the passage of time must have been built into the system in the first place. Random processes cannot contribute anything. It follows that mathematical systems are in principle unable to explain the origin of highly variegated and complex entities, such as the forms of living organisms. At best, such systems can merely account for complex order here and now by postulating that an equivalent complex order was present at some earlier time or was transported into the system from outside. This does not account for the origin of such order, but simply confronts us either with an infinite regress or with an eternal source of order containing the information necessary to specify all life forms.

This indicates that the original scientific program of describing the world by mathematical systems is too limited and can therefore only serve to impede our understanding of nature. This program is based on the conviction that the simple regularities observed by physicists and chemists in experiments with inanimate matter will suffice to account for every phenomenon in the world. But since we show here that this program cannot succeed in accounting for the origin of life, we suggest that the alternative view—that life, rather than matter, is the fundamental cause of all observed phenomena—must be seriously considered.

The program of science during the last two or three hundred years has been to reduce life to matter and deny the existence of any higher living principle transcendental to matter. In this program, the idea that matter is both simple and conceivable has been essential. But if, as shown here, it is necessary to attribute to matter all the characteristic features of life in order to explain the origin of life from matter, then it may be concluded that this program has failed. It then makes more sense to reverse the scientists' program and admit that since the cause of life must possess the properties of life, life itself must be the fundamental causative principle, and matter must be a derived phenomenon originating in life. This approach to nature is consistent with the ancient understanding, accepted by the seminal thinkers of both East and West, that the ultimate causative principle is the Supreme Living Being, and that all phenomena are derived from Him. It should also open up a wealth of new avenues of scientific exploration.

In Sections 1, 2, and 3 the basic arguments are developed. The implications of these arguments are described in Section 4. The mathematical definitions and derivations are included in the appendix.

**The laws of nature,
as understood by modern physics,
have a low information content.**

One of the fundamental principles of modern physics has been that the laws of nature can be described by very simple and general mathematical relationships. This is perhaps epitomized by the example of Albert Einstein, who strove during the major portion of his life to find a "unified field theory" which would derive all the forces and laws of interaction of the physics of his day from a single, simple unifying principle.

In physics, the laws of nature are generally studied only in highly restricted and simplified situations. This is done primarily for the practical reason that in more complex situations the mathematical analysis quickly becomes so difficult that no one can carry it out. The equations of quantum mechanics, for example, can be solved only for a system containing two particles at most. For larger systems approximations must be used, and these become very difficult to handle even in the case of the diatomic hydrogen molecule (which has four particles). For still larger systems, such as simple organic molecules, only guesses and conjectures can be made about the predictions of quantum mechanical theory. For large organic molecules such as those found in living cells, it is hopeless even to attempt to give a quantum mechanical description.

Nonetheless, the theories of physics are regarded as being, at least in principle, complete and universal descriptions of the phenomena of nature. It is believed that the reality of nature (or that which is really there) can be described by systems of numbers, and that all of the phenomena of nature can be calculated from certain mathematical relationships between numbers which are called "the laws of nature." Furthermore, it is accepted that these fundamental laws are simple in form and can be ascertained by human scientific investigation.

This assumption is very striking from the viewpoint of mathematics, where it is appreciated that theoretical models possessing any number of axioms, or basic laws, can exist. Indeed, it has been shown that any model of the natural number system will be logi-

cally ambiguous unless it contains an infinity of basic axiomatic laws. (This is Gödel's incompleteness theorem.) The assumption of the physicists that a very few basic laws will suffice to describe nature is therefore very restrictive, even though it is a practical necessity if the goal of physics—the determination of these laws—is to be realizable.

It is now an almost universal belief among scientists that this basic program of physics has been successfully carried out, at least for phenomena involving moderate masses, temperatures, and velocities. Specifically, it is believed that all of the phenomena of chemistry follow the known physical laws, and that all of the phenomena of life can be reduced to chemistry (and, hence, ultimately to physics). Thus the biochemist James Watson declares that with the development of quantum mechanics, "the various empirical laws about how chemical bonds are formed were put on a firm theoretical basis. It was realized that all chemical bonds, weak as well as strong, were based on electrostatic forces."¹ In his description of the basic goal of molecular biology, he states, "we see not only that the laws of chemistry are sufficient for understanding protein structure, but also that they are consistent with all known hereditary phenomena. Complete certainty now exists among essentially all biochemists that the other characteristics of living organisms (for example, . . . the hearing and memory processes) will all be completely understood in terms of the coordinative interactions of small and large molecules."²

The known laws of quantum mechanics must indeed be very remarkable entities, since it is proposed that all the features and characteristics of life depend on these laws. Let us therefore try to write them down.

First of all, let us consider the laws of nature in classical physics. These can be summed up by the following equations:

$$\frac{dq_j}{dt} = \frac{\partial H(p_1, \dots, p_n; q_1, \dots, q_n)}{\partial p_j} \quad (1)$$

$$\frac{dp_j}{dt} = -\frac{\partial H(p_1, \dots, p_n; q_1, \dots, q_n)}{\partial q_j} \quad (2)$$

In classical physics, the state of a physical system at any given time is completely described by the position coordinates, q_j , and

momentum coordinates, p_j . This is a set of $6N$ numbers for a system of N material particles, and equations (1) and (2) describe how they change with the passage of time. The function, H , which is called the Hamiltonian, is generally given by a simple formula in classical physics:

$$H = \sum_{j=1}^n p_j^2 / (2m_j) + V(q_1, \dots, q_n) \quad (3)$$

$$V = \sum_{i < j} A_{ij} / |\bar{r}_i - \bar{r}_j| \quad (4)$$

$$|\bar{r}_i - \bar{r}_j| = \sqrt{(q_{3i} - q_{3j})^2 + (q_{3i+1} - q_{3j+1})^2 + (q_{3i+2} - q_{3j+2})^2} \quad (5)$$

We have written out these formulas in full in order to show how very simple the laws of classical physics are. This is quite literally the full extent of the laws of nature as understood in classical physics up to the time of Maxwell. According to those who adhered purely to the scientific philosophy that nature could be completely described by mathematical laws, all of the phenomena of nature are consequences of equations (1) through (5) and the initial values of the q_j 's and p_j 's at some arbitrary starting time, $t = 0$.

This philosophy was given its initial impetus in the 18th century by Isaac Newton. He summed it up as follows: "I . . . suspect that [the phenomena of nature] may all depend upon certain forces by which the particles of bodies . . . are either mutually impelled towards one another and cohere in regular figures, or are repelled and recede from one another."³ Equation (4) specifies these forces, which attract if A_{ij} is positive and repel if it is negative. The same view was expressed more recently in the 19th century by the physicist Hermann von Helmholtz: "The task of physical science is to reduce all phenomena of nature to forces of attraction and repulsion, the intensity of which is dependent only upon the mutual distance of material bodies. Only if this problem is solved are we sure that nature is conceivable."⁴ Needless to say, these scientists include life as a "phenomenon of nature."

With the advent of Maxwell's electromagnetic theory, Einstein's theory of relativity, and the theory of quantum mechanics, the simple view of nature summed up by equations (1) through (5) underwent a considerable change. However, the basic sentiment

expressed by Helmholtz that natural phenomena should be reducible to the interplay of elementary material forces has been retained. In the present dominant theory—quantum mechanics—physical systems are still described by arrangements of numbers, although the particle coordinates of classical physics have given way to Hilbert space vectors. The laws of transformation of these numbers are still given by brief equations which may be written down in a few lines.

In quantum mechanics, the basic equation of motion for a physical system is the Schrödinger equation:

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi \quad (6)$$

Here, the state or exact physical description of the physical system is given by the Hilbert space vector, ψ , which can be represented in various alternative ways as a mathematical function or as a sequence of numbers. The Hamiltonian function, H , has been adopted from classical physics and now appears as an operator capable of acting on ψ to produce a new vector. In analogy with equation (3), H could be given by:

$$H = \sum_{j=1}^n \frac{-\hbar^2}{2m_j} \frac{\partial^2}{\partial q_j^2} + V(q_1, \dots, q_n), \quad (7)$$

where V is the same as in equation (4).

Equations (6) and (7), along with (4) and (5) and an initial value for ψ at the time, $t = 0$, completely specify the quantum mechanical picture of a physical system of $n/3$ particles moving according to the attractive and repulsive forces given by V .

In further developments of the theory of quantum mechanics, things become somewhat more complicated. In addition to equation (4), various other terms are added to V to represent different kinds of forces believed to be acting in physical systems. These include terms for "spin" and electromagnetic interactions. Also, the basic form of H in equation (7) is sometimes modified in various ways, as happens in relativistic quantum mechanics and quantum field theory. It remains true, however, that the Hamiltonian for any system which is supposed to represent the fundamental laws of nature can be expressed by very brief formulas. When the abbreviated notations in these formulas are written out

in full, as done above for the classical case, the resulting equations will be found to occupy a few lines at most.

$$\begin{aligned}
 \text{(a)} \quad H\Psi &= i\hbar \frac{\partial}{\partial t} \Psi \\
 \text{(b)} \quad H &= \\
 &\sum_n \frac{-\hbar^2 c^2 \frac{\partial^2}{\partial q_n^2} + \eta_n^2 q_n^2}{2} + \sum_k \frac{-\hbar^2 \nabla_k^2}{2m_k} \\
 &+ \sum_k \frac{i\hbar e_k}{m_k c} \bar{A}(\bar{Q}_k) \cdot \nabla_k + \sum_k \frac{e_k^2}{m_k c^2} |\bar{A}(\bar{Q}_k)|^2 \\
 &- \sum_k \frac{e_k}{2m_k c} \bar{\sigma}_k \cdot \nabla_k \cdot \bar{A}(\bar{Q}_k) + \sum_{i,j} \frac{e_i e_j - G m_i m_j}{|\bar{Q}_i - \bar{Q}_j|} \\
 \bar{A} &= \sum_n q_n \bar{A}_n
 \end{aligned}$$

Figure 1. The Laws of Physics Underlying Chemistry

This is true in particular of the physical model of chemical interactions referred to above by Watson as being sufficient for a complete understanding of life. The Hamiltonian for this model should include terms for electric forces, spin interactions, and electromagnetic interactions (plus gravity). This Hamiltonian is illustrated in figure 1.⁵

It is our thesis that a system of equations as brief and simple in form as these cannot possess sufficient power of discrimination to summon forth from a chaos of randomly distributed atomic particles the complex and variegated world of life we see about us. The theory of the origin of life from inanimate matter invokes two processes: chance and natural selection. The idea is that "chance" will provide various combinations of molecules which may or may not be useful in living organisms, and "natural selection" will pick out those which are useful and eliminate those which are not. Geneticists such as R. Fisher have argued statistically that even if natural selection only slightly favors one combination or form over another, in a sufficient length of time the favored form will nonetheless be found to predominate over the unfavored one.⁶

However, natural selection must have some direction if it is consistently to choose certain material configurations out of the myriads of configurations possible. The local selective advantages within particular populations must add together (as in a vector sum of vectors added tail to head) to a general trend from primordial soup to higher organisms. Ultimately, the fundamental laws of nature must provide this direction. At least, this must be true if nature is indeed to run in accordance with such laws.

It is very hard to see, however, why "forces of attraction and repulsion . . . dependent only upon the mutual distance of material bodies" should select trees, amoebas, bumblebees, or human beings in favor of other possible material configurations, such as inert globs or blotches. Enhancing the theoretical picture with spin interactions following the Pauli matrices, or electromagnetic fields composed of harmonic oscillators, does not seem to add more plausibility to the idea that natural selection could do this. We will argue that the very brevity of the laws of nature as they are expressed in physical theories makes them unsuitable for selecting the complex forms of living organisms from an initial state of molecular chaos, no matter how much time is allowed for the process. Basically, we shall show that in order for a set of natural laws to select a complex form out of a random distribution of matter, these laws must possess a corresponding level of complexity themselves. This will imply that the Hamiltonian for a system in the quantum mechanical formulation would require many pages of symbols to write down (as many as fifty at the very least) in order for that system to select configurations with the complexity of living organisms. In other words, the known laws of physics are insufficient to account for the origin of life, and in order for a system of physical laws to do so, its sheer size and complexity would make it impossible for the human mind to handle. (We should point out that high complexity will be shown to be a necessary but not sufficient requirement for a system to exhibit the evolution of complex forms.)

As we have already pointed out, the equations of motion cannot be solved exactly for systems containing more than two particles in either the classical or the quantum mechanical theories of physics. Indeed, in the theory of relativistic fields these equations have given rise to difficulties even in the case of zero particles (the so-called vacuum state). Since systems capable of describing living organisms must contain enormous numbers of particles on the order of 10^{23} , we can see that it can never be practical to study

the nature of such organisms by explicitly solving the equations of motion. However, such solutions must exist in principle in order for the theory to be valid at all. The procedure of scientists, then, is to establish the existence of solutions by abstract reasoning and to demonstrate that these solutions could be calculated in principle. Then they attempt to apply and verify the theory by making logical deductions about the properties of the solutions without actually seeing them. We shall proceed on the assumption that the required solutions always do exist and could be calculated.

We shall formally measure the complexity or information content of a theory as the length of the shortest computer program which can numerically solve the equations of motion for the theory to within any desired degree of accuracy. The Schrödinger equation (equation (6)) can be solved in principle by a simple numerical algorithm. Consequently, the information content of a theory having this equation as its basic equation of motion is nearly proportional to the number of symbols needed to write out the Hamiltonian for that theory in terms of the specific programming language. For consistency, all estimates of information content will be referred to a fixed programming language. The information content of a configuration of matter, such as the body of a living entity, can also be estimated as the length of the shortest program which will generate a complete numerical description of the configuration. We shall use this measure of information content to provide a clearcut numerical demonstration that the known laws of physics, or any system of laws of a similar nature, should fail, even in principle, to account for the origin of life.

In a mathematical model of a physical system, two other ingredients are needed along with the laws of motion. These are the initial conditions and boundary conditions for the system. Normally, a physical system will be confined within a certain fixed volume of space: in order to describe the events within the system, it is necessary to describe the physical conditions along the boundary of this volume during the time in which the system is being studied. Also, in most physical models the system is considered during a finite time interval, $0 \leq t \leq t_1$. It is therefore necessary to specify the physical state of the system at the beginning of this time period, $t = 0$.

In a theory of the origin of life, the initial conditions should describe a "primordial" situation possessing a very low degree of organization, if any. For example, most theories of the chemical

origin of life postulate that life arose from a "primordial soup" consisting of a mixture of water and simple compounds such as CO_2 , CH_4 , N_2 , NH_3 and H_2S and a reducing atmosphere composed mainly of CH_4 and NH_3 .⁷ This mixture of chemicals is presumed to receive radiation from the sun, to receive supplies of gases from the earth (volcanic venting), and to radiate heat and light into outer space. This sums up the initial and boundary conditions for this model.

As another model, one could start with the supposed origin of the solar system from a cloud of gas. The initial conditions would then consist of a description of the initial gas cloud, and the boundary would correspond to an unlimited vacuum surrounding this cloud (if we ignore the influence of distant stars). Then, according to theory, the laws of nature would presumably first generate the solar system, complete with primordial soup, and then generate life in the soup. Or, one might consider a model of the universe as a whole, such as the "big bang" theory, which features a superhot soup of subatomic particles as its initial condition.

In any case, the idea behind all theories of the origin of life from matter is that one only has to propose a simple set of conditions to hold in the beginning. After all, the idea of the theory is to "explain" all the features of life, and the greater the intricacy of the specifications required for the initial conditions, the less complete the explanation becomes. One then has to explain the origin of the intricate initial conditions in terms of some still earlier state of affairs.

A typical initial condition for a model of the origin of life will consist of an ensemble of possible initial states, such as one of the ensembles of statistical mechanics or some simple combination of these. This certainly holds true for the models referred to above. Such an ensemble can be specified by a simple equation or a brief set of equations. For example, in the theory of quantum mechanics one of the standard thermodynamic ensembles, known as the canonical ensemble, is given by the equation,

$$\rho_0 = K^{-1} \exp(-H/kt) \quad (8)$$

where H is a Hamiltonian operator.⁸

In this equation, ρ_0 is called a density matrix. It corresponds to a collection or ensemble of quantum mechanical states, each of which has a statistical weight assigned to it which gives the proba-

bility that that particular state will be found to hold in the physical system. This means that the initial state of the system is left ambiguous or undefined to a very large extent. In fact, one of the basic principles of statistical mechanics is the assignment of equal probabilities to all initial states satisfying certain simple restrictions, such as a certain particle density or a certain range of energies. (This is how the micro-canonical ensemble is defined.) In this case, the state of the system is completely ambiguous, apart from the requirement that these restrictions should be satisfied.

The idea of the theory of evolution is that out of this chaos or ambiguity, the laws of nature should be able to "select" molecular configurations capable of exhibiting all the phenomena of life. Another way of looking at this is the following: The laws of nature should have the ability to generate life forms from most of the possible initial physical states which satisfy certain simple physical requirements. This would assure that if one of these states were chosen "at random" as the initial condition of the system, life forms would probably appear in the system at some later time.

The boundary conditions for the physical model should also be definable in simple terms. At most, one would expect energy or material particles in a simple form to pass back and forth across the boundary. One can define these conditions by specifying external electromagnetic fields, and also probability distributions for incoming matter, such as cosmic rays. This can also be done by means of simple equations. In some models the boundary might consist of reflecting or non-interactive walls, or a limitless void, or there may be no need for explicit boundary conditions at all. Basically, any interaction between the system and matter or energy beyond its boundary should be describable in simple statistical terms, as can be done, for example, for cosmic rays or the influx of solar radiation. Boundary conditions, like initial conditions, must be relatively simple: if intricate specifications were required for the boundary conditions, one would need to explain their origin also.

Once the initial and boundary conditions are defined, the state of the physical system at each subsequent time, t , between 0 and t_1 must be determined by the natural laws of the system. The basic situation is summed up in figure 2. A theory of the origin of life from matter should require that in some range of times, perhaps in the range of three or four billion years or so, these states will exhibit the molecular configurations characteristic of living entities with a reasonably high probability. Let us consider how

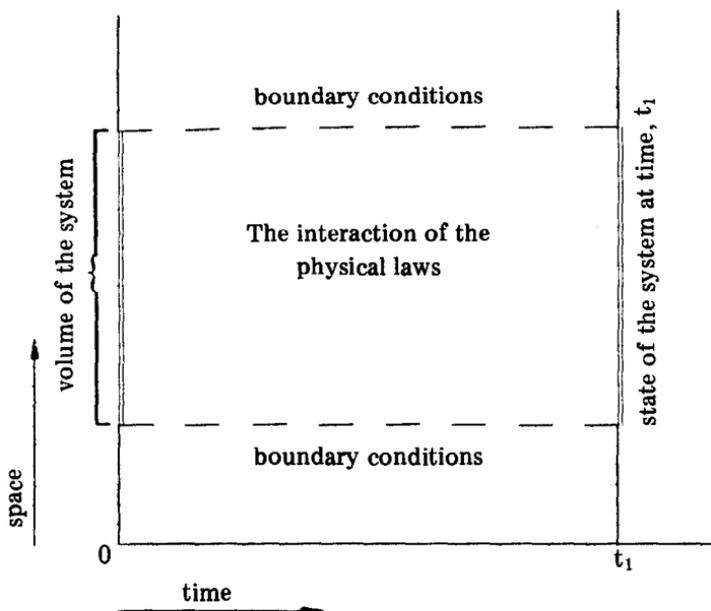


Figure 2. The basic form of a physical model

these configurations may be mathematically represented in the physical states. We will do this for the theory of quantum mechanics, in particular. In the theory of quantum mechanics, any entity or material configuration one might want to observe in the system is represented by a mathematical operator called an "observable." Let us denote the entity to be observed as B and denote the observable to which it corresponds by \hat{B} . Then the probability of finding B in the system represented by the density matrix, ρ , is given by,

$$\text{Prob}(B) = \text{Trace}(\rho \hat{B}) \quad (9)$$

All that we want to point out about this formula is that it can also be calculated by a brief set of computer instructions.

We are interested in the case where B is a particular configuration of atoms representing a molecule or collection of molecules. One way of describing such a configuration is by means of a numerical code indicating which atoms should be found in close proximity to which other atoms—that is, indicating the pattern of chemical bonds which characterizes the configuration. If X repre-

sents such a configuration, then the observable, $B(X)$, corresponding to it can be expressed by a simple formula (see the appendix).

Thus, the boundary conditions, initial conditions, and laws of nature determine the density matrix, ρ , describing the physical state of affairs at time, t_1 . From ρ we can (at least in principle) calculate the probability, $M(X)$, that any particular configuration described by a code, X , can be found somewhere in the system at time, t_1 :

$$M(X) = \text{Trace } [\rho B(X)] \quad (10)$$

The probability function, $M(X)$, determines what configurations can be expected to have evolved in the system by the time, t_1 . If $M(X)$ is very large, X can be expected to have evolved, but if it is very small, the evolution of X is unlikely. If we measure the information content of a number or a mathematical function as the length of the shortest computer program which will calculate it, then it follows that the information content, $L(M)$, of the function, M , is no greater than the total length of all the various calculations we have just described. This should correspond to about three or four pages of tightly packed programming instructions at the most. (We envision this program as being written in some standard programming language with provisions for handling numbers with an arbitrary number of significant digits.)

We shall show in a later section that $M(X)$ must be exceedingly small if the information content, $L(X)$, of X is very much greater than the information content, $L(M)$, of M plus a certain fixed constant. The difference between the information in X and that supplied by the system (represented by M) must be made up by pure chance, and therefore the probability of X goes down exponentially with this difference. This shows that the evolution of life forms in such a system is exceedingly unlikely, for it stands to reason that the length of the shortest program needed to calculate the essential molecular structures of even a "primitive" living organism should be a great deal longer than three or four pages. This is argued in detail in the next section.

We should note the ways in which "chance" enters into the expression, $M(X)$, for the probability of finding X . In both the initial conditions and the boundary conditions, there is randomness due to the thermodynamic ensembles expressing them. This is the source of the randomness postulated for the "random mutations" of the theory of evolution, which are supposed to be due to

chance molecular collisions and collisions with cosmic rays. Also, randomness is built into the basic structure of quantum mechanics, since the quantum mechanical states are statistical distributions. Both the mutations and the natural selection of the theory of evolution are built into these physical models, as we would expect.

As a final point, we would like to mention one other model that exhibits all of the features involved in the theory of evolution. This is the cellular automaton model of John von Neumann. In this model, we suppose that a small automaton, which can be in any one of a finite number of states, is situated on each square of a large two-dimensional lattice. The state of the system is given by specifying the states of each small automaton in the lattice. The system changes with time in the following way: Let t be a small, fixed time interval. At the end of each successive time interval, t , each automaton changes to a new state in a way which depends only on the states of the automatons on the squares immediately adjacent to it. The transition function which determines these changes can be specified by a few lines of instructions.⁹

Von Neumann has shown that a self-reproducing machine capable of performing complex calculations (a universal Turing machine) can be constructed in this system in the form of an arrangement of a large number of these small automatons. His idea was to prove that the property of self-reproduction characteristic of living organisms can be exhibited by a mechanical system. This was intended as one further step in demonstrating that life is a mechanical process governed by mathematical laws.

It is therefore interesting to inquire whether self-reproducing systems of the sort which von Neumann considered could evolve in his cellular automaton system in a sufficient length of time. The law of transformation of the system with time should determine the natural selection of different forms in preference to others within this system by the same sort of processes of competition envisioned in ordinary evolutionary theory. The random mutations of evolutionary theory could be easily introduced in the form of a Markov process which would make random changes in the states of the small automatons.

However, such evolution could not be expected to occur. Owing to the great simplicity of von Neumann's system, the function corresponding to $M(X)$ for the system can be expressed by a very brief list of equations. Von Neumann's self-reproducing machines, on the other hand, are very complicated, and he re-

quired several hundred pages to describe them in his book. It therefore seems reasonable to suppose that their information content greatly exceeds that of M. As we shall see in Section 3, this rules out the evolution of such configurations within the system.

**The bodily structures of living organisms
have a high information content.**

We shall consider the bodily structures of living organisms in three aspects: the visible bodily form (phenotype), the genetic DNA code (genotype) and other molecular structures of cells, and the behavioral patterns of organisms. Under the last heading, we shall be particularly concerned with human behavior and its by-products, such as language, literature, technology, and scientific theories.

First, let us consider the molecular structures of cells. One of the most thoroughly studied organisms is the bacterium *Escherichia coli*, a unicellular organism about 500 times smaller than the average cells of higher plants and animals.¹⁰ It is one of the smallest and simplest of all living organisms. Yet it is estimated that a single *E. coli* cell contains between 3,000 and 6,000 different types of molecules. Among these are some 2,000 to 3,000 different kinds of proteins with an average molecular weight of 40,000, as well as a single DNA molecule with a molecular weight of 2.5×10^9 . This molecule of DNA is believed to contain coded instructions for the construction of all the other molecules. James Watson, one of the foremost authorities on molecular biology, admits that these molecules do not obey any simple structural rule:

Most of these macromolecules are not being actively studied, since their overwhelming complexity has forced chemists to concentrate on relatively few of them. Thus we must immediately admit that the structure of a cell will never be understood in the same way as that of water or glucose molecules. Not only will the exact structures of most macromolecules remain unsolved, but their relative locations within cells can be only vaguely known.¹¹

It is believed that protein molecules may be described as chains of amino acid molecules, which may be of twenty different varieties. A typical protein molecule in an *E. coli* cell will contain some 300 of these amino acid subunits. Since each subunit may be any one of twenty different amino acids, this means that the

number of possible protein molecules of typical size is about 20^{300} . The information content of a typical protein molecule in an *E. coli* cell is therefore bounded by an upper limit of $L(X) = 300 \log_2 20 \approx 1290$ bits. Since there are an estimated 2,000 to 3,000 different proteins of this kind in a cell, the total information content for cellular protein is bounded by an upper limit of between 2,580,000 and 3,870,000 bits.

The DNA molecules of cells are thought to be helical chains composed of successive pairs of DNA bases. There are four different kinds of bases—adenine (A), thymine (T), guanine (G), and cytosine (C)—and these are limited to forming four different kinds of base pairs: A-T, T-A, C-G, and G-C. The DNA molecules within a cell are believed to contain coded instructions for the structure of each protein molecule within the cell. Each group of three base pairs along the DNA chain specifies a specific amino acid in a corresponding protein molecule, or it codes for the termination of a protein chain. Each DNA base pair has a molecular weight of about 660, and so a group of three pairs has a molecular weight of about 2,000. Since there is a molecular weight of about 2.5×10^9 for the (haploid) *E. coli* DNA, this means that the genetic code for this organism will consist of some 1.2×10^6 triplets of bases. Since each triplet can discriminate between 21 alternatives (20 amino acid types and a stop code), this gives us an upper bound of about 1.2×10^6 times $\log_2 21$, or some 5.2×10^6 bits, for the genetic information content of an *E. coli* cell.

Yet the *E. coli* is a very simple cell. In the cells of higher plants and animals, much larger amounts of DNA are found than in *E. coli*. It is estimated that mammalian cells contain some 800 times as much DNA, yielding an upper limit of some 4.2×10^9 bits for the genetic information content of these cells.¹² Some idea of the size of these numbers can be obtained by considering the number of pages of print that would be required to write down these amounts of genetic coding in full. In a typical book there are some 70 characters per line and 40 lines per page. This gives us about 1.7×10^4 bits per page if we use an alphabet of 64 characters. (Since $64 = 2^6$, we can code 6 bits per character.) At this rate it would take about 300 pages to write down the coding for *E. coli*, and about 240,000 pages to write down the coding for a mammalian cell.

One of the dogmas of modern molecular biology (called the central dogma, in fact) is that all the information needed to specify a cell is contained in the cell's DNA coding, and that this

coded information is not changed except by random mutations. Whether this is true or not, the amount of DNA in cells should give some idea of the amount of information they contain. According to the central dogma, the figures we have derived thus far should provide upper limits for cell information content. We would now like to estimate reasonable lower limits. To do this, we shall need to consider the variety and complexity of the structures cells are found to contain. Since some genes in the chromosomes of higher animals apparently exist in multiple copies, it would seem that their total genetic information content should be lower than the figures suggested by their total DNA content. However, based on the great complexity of the structures that can be seen in vertebrate cells under the microscope, Watson estimates that such cells should be "at least 20 to 50 times more complex genetically than *E. coli*."¹³ By this he means that such cells should contain coding for at least 20 to 50 times as many protein macromolecules as *E. coli*. We would need between 6,000 to 15,000 pages to write down this amount of coding, based on our 300-page estimate for *E. coli*. This corresponds to between 1.0 and 2.6×10^8 bits of information, and between 80,000 and 200,000 proteins, with an average of 300 amino acid subunits apiece.

In order to determine a lower limit for information content, let us consider how the genetic information is to be distributed along the length of the chromosomes in a cell of a higher plant or animal. The total DNA coding of such a cell should be representable by a binary number of 4.2×10^9 digits, or, using an alphabet of 64 characters, by a character string of 7×10^8 characters. Let us call this character string X . We can divide X into N blocks, Y_1, \dots, Y_N , each of which has a length of $M = 7 \times 10^8 / N$ characters. Call the string made up of the first k of these blocks (going from left to right) X_k , for each integer, $k = 1, \dots, N$.

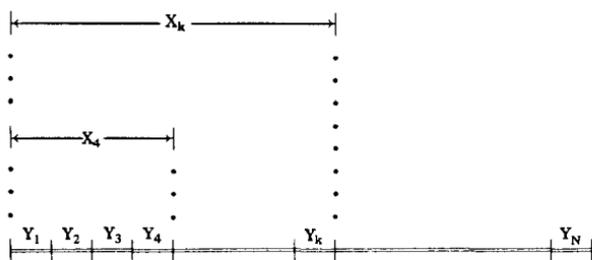


Figure 3. Division of the genetic coding sequence, X , into a series of adjacent blocks.

The following inequality relates the total information content of X with the incremental amounts of information required to express each successive Y_k in terms of X_{k-1} .

$$L(X) \geq L(Y_1|I) + L(Y_2|X_1,I) + \dots + L(Y_k|X_{k-1},I) \\ + \dots + L(Y_N|X_{N-1},I) - Nc \quad (11)$$

The term, $L(Y_k|X_{k-1},I)$, is the amount of information needed to calculate the string, Y_k , given that we already know the string, X_{k-1} . (Note that $X_1 = Y_1$ and $X = X_N$.) Specifically, this term is the length in characters of the shortest program that will calculate Y_k using X_{k-1} as a possible source of data. The term, $L(X)$, is the total information content of the string, X , in characters. The constant, c , is an artifact of the mathematical derivation of this inequality, and is no more than about 47 characters.

Inequality (11) can be rephrased as follows:

$$\frac{L(X)}{7 \times 10^8} \geq \frac{\text{Average of } L(Y_k|X_{k-1},I) - 47}{k = 1, \dots, N} \quad (12)$$

Suppose, for the sake of argument, that $L(X) \leq 10^5$ characters. Then, if we set M , the length of the blocks, Y_k , equal to 7,000, we find that on the average, $L(Y_k|X_{k-1},I) \leq 48$ for $k = 1, \dots, N$. ($N = 10^5$) This means that a typical block, Y_k , of 7,000 characters can be calculated exactly from X_{k-1} by means of a program no more than 48 characters in length.

Let us consider what this means. We have postulated that the information content of the genetic coding string, X , is only 1/7,000 of its maximum possible value. This implies that, on the average, each successive block of 7,000 characters along X can be exactly calculated from the previous blocks by a computer program of no more than 48 characters in length. Now, a block of 7,000 characters corresponds to several genes specifying protein macromolecules. For example, nearly 33 genes specifying proteins of 300 amino acid subunits in length can be coded with 7,000 characters. Since the amino acid sequences of such proteins appear very complicated, it is hard to believe that they could be consistently transformed from one to another by such short programs.

These proteins play very sophisticated roles in the metabolism of cells. The exact way they function is far from known at the

present time. However, it is known that they are able to behave like small computers. Here is an example taken almost at random from Watson:¹⁴ In *E. coli* there occurs a sequence of chemical reactions that convert the compound threonine into isoleucine in five steps. The first step in this sequence can occur only with the aid of the enzyme L-threonine deaminase, a protein macromolecule. When the final product, isoleucine, has reached sufficiently high concentrations, it interacts with the enzyme molecules in such a way that they no longer catalyze the first reaction of the series. This prevents the manufacture of more of the product than needed in the cell. We may note that the enzyme is so structured that it catalyzes only one specific reaction. While not affecting the rates of other chemical reactions at all, such biological enzymes are known for their ability to cause certain chemical reactions to occur millions of times faster than they will occur under laboratory conditions in which the enzyme is not present. We may also note that this enzyme is inhibited specifically by isoleucine at the proper level of concentration, and not by any other chemical that would normally be present in the cell (for otherwise this control system wouldn't work.) Many enzymes of this sort, which interact with other molecules according to a logical scheme and perform very specific chemical operations in a highly efficient way, are evidently required if the cell is to function as a highly versatile and complex automaton. It would therefore seem very remarkable if the typical block of 33 such enzymes coded on X could be exactly calculated from the previous blocks in the chain (namely, X_k) by a program of only 48 characters. This would be like expecting the first few hundred decimal digits of π or $\sqrt{2}$ to code for an *E. coli* enzyme.

Yet we have estimated that there may be as many as 200,000 such enzymes in a mammalian cell. A block of genes could be simply computed from the previous blocks in the chain if it were a duplicate copy or very near copy of the genes in the previous blocks. But we would not expect this to be possible very often for genes performing functions distinct from those in the previous blocks. We conclude that we must be incorrect in our assumption that the total complexity, $L(X)$, of the genetic code for a mammalian cell is less than 10^5 characters. We therefore propose that this figure can be taken as a reasonable lower bound for $L(X)$.

$$\text{Lower bound of } L(X) > 10^5 \text{ characters.} \quad (13)$$

This lower bound corresponds to 1/7,000 of the upper bound of $L(X)$. Watson¹⁵ indicates that some of the genes (but, presumably, by no means all) in cells of higher animals may appear in as many as 100 to 1,000 redundant copies, but this could account for only about 1/70 of the redundancy our lower bound allows. Therefore it might also be reasonable to consider a lower bound of 7×10^6 for $L(X)$. As we have noted, Watson proposes that the cells of higher animals must have some 20 to 50 times as many distinct protein macromolecules as *E. coli*. If these proteins can be specified in a chain only by programs with an average length of no less than 57 characters, then the minimal value of $L(X)$ will be between 8×10^5 and 2×10^6 . [This follows from inequality (11) with $57 \leq \text{Average of } L(Y_k | X_{k-1}, I)$ and 80,000 to 200,000 for N . In this case we let X represent the sequence of codes for the different proteins.]

In order to give a further indication of the type of complexity we are dealing with in living cells, let us now describe some of the different categories of structure that organisms display. We have ruled out the possibility that very short programs can transform the sequence of genes in X_{k-1} into the sequence in Y_k . Actually, for the overwhelming majority of sequences of characters,

$$L(Y|X,I) \simeq L(Y|I) \simeq L(Y) \simeq \text{length of } Y. \quad (14)$$

For proteins of 300 amino acid units in length, we have length of $Y = 300 \log_{64} 20 \simeq 215$ characters. A very elementary theorem about the information function, L , indicates that no more than $1/64^{(215-57)}$ of the total number of Y 's of 215 characters can be calculated from a given X by programs of 57 characters or less. Almost all sequences of symbols of a given length can be specified most briefly by simply writing them out in full. Of those that remain, the overwhelming majority can be most simply specified by writing them out nearly in full and indicating some brief calculations that will compute the remainder of the sequence.

In view of this, consider the following hierarchy of structures exhibited by living organisms.

(a) *The chemical reactions involved in cellular metabolism.* These involve respiration, the synthesis of various chemicals needed in the cell from food molecules, photosynthesis in plants, and the processes involved in the orderly breakdown of different molecules. It would appear that most of the genetic coding of *E. coli* must be devoted to this, since these bacterial cells do very

little but grow and divide in half. Even though *E. coli* is one of the simplest of organisms, its metabolic interactions are so intricate that "the exact way in which all these transformations . . . occur is enormously complex, and most biochemists concern themselves with studying (or even knowing about!) only a small fraction of the total interactions."¹⁶ That these interactions must be governed by a complex system of logic rivaling the most sophisticated programs of modern electronic computers is certainly indicated by the descriptions given in Watson's book.

(b) *The morphological structures of single cells.* The *E. coli* cell appears to possess a relatively simple gross structure, but many cells, even among the bacteria and protozoa, possess very intricate structures. For example, the cilia of protozoa such as the paramecium have been shown to possess a very intricate structure that elegantly exploits certain mechanical principles to produce a synchronized rowing machine.¹⁷ Even the *E. coli* cell possesses some remarkable mechanical features. This bacterium propels itself through the water by means of a spiral flagellum that is spun about its axis by a motor built into the bacterial cell wall.¹⁸ This motor is complete with drive shaft and some kind of rotating disks, but the principle underlying its operation is still unknown. It seems very doubtful that structures such as these can be calculated by computer programs of 57 or so characters, or that they can be transformed by such programs into, say, the coding for the enzymes involved in the Krebs cycle of cellular respiration. (Note that the preceding sentence is 225 characters in length.)

(c) *The different types of cells involved in the organs of multicellular organisms, and their different functions.* We can easily write down a long list of different types of cells appearing in different bodily organs. These include muscle cells, nerve cells, bone cells, different kinds of blood cells, glandular cells, liver cells, epithelial cells, etc. The study of a particular kind of cell can be a whole academic subject in itself, and doctoral dissertations are frequently devoted to the study of a detail of a detail of the structure and function of such cells. The complete instructions for constructing all these different cells must be contained in the genetic coding of any higher organism, at least according to the understanding of modern biology. There must also be instructions controlling the development of these various types of cells during the growth of the embryo.

(d) *The structure and function of different types of organs in higher plants and animals.* This is an aspect of (c). The different

organs of the body perform a vast array of complicated functions, most of which are either unknown or poorly understood. Examples include the disease fighting system of the blood, the image-producing eye and its retina, the brain, the endocrine gland system, and the heart and circulatory system. It also seems very doubtful that the coding for different features of these various organs could be transformed into the coding for the exact specifications of other organs by computer programs of a few characters in length.

As two contrasting examples of such organs, let us consider the eye and the feather of a bird. Both of these examples have long troubled the more thoughtful students of the theory of evolution because it is very hard to see how they could have originated by natural selection and mutation. Thus, Charles Darwin said, "I remember well the time when the thought of the eye made me cold all over, but I have got over this stage of the complaint, and now small trifling particulars of structure often make me very uncomfortable. The sight of a feather in a peacock's tail, whenever I gaze at it, makes me sick!"¹⁹

A brief consideration of the structures making up the eye will give some indication of why Darwin felt as he did. These include the lens, the muscles supporting it, the iris, the cornea, the retina, the nerve connections, and the muscle system that moves the eyeball. Many of these structures are very intricate. For example, the iris contains a muscle system for opening and closing the pupil; the lens must be transparent and shaped so as to focus a sharp image on the retina; the retina contains systems of cells and nerves designed to detect different elementary visual patterns such as lines and edges; the light-sensitive cells contain complex chemical systems designed to respond to different colors of light; and so forth. All these structures must be coded into the genetic specifications of the organism, and it is hard to see how the codes for such diverse structures could be transformed, one into another, by simple computer programs. On the contrary, it seems much more reasonable to suppose that the shortest set of instructions defining the eye should contain detailed descriptions of each of these structures individually. In other words, the eye should have a high information content.

The feather is also a very complicated structure. Alfred Russell Wallace, the co-founder with Darwin of the natural selection theory, later rejected this theory because he felt that it could account neither for the nature of man nor for the amazing struc-

tures of living organisms in the plant and animal kingdoms. One of his examples is the bird's feather, which, as he points out in his book *The World of Life*, contains many different structural features of great intricacy.²⁰ We may also note that, superimposed on this basic structure, many feathers display very precise and detailed colored designs. The peacock feather, which so troubled Darwin, contains a multicolored image of an eye, set off from its background by a systematic thickening and thinning of the smaller strands of feather that make it up. These structures should also have a high information content.

(e) *The behavior of animals other than man*. Many different patterns of complex behavior are exhibited by lower animals. We may note, for example, the social systems of bees and ants, the spinning of spider webs, and the transcontinental migrations of birds. These different behavioral patterns are thought to be built into the genetic material of the organisms, and, as such, they must be coded in terms of a series of logical "if-then" instructions equivalent to the program of a computer. It might be an interesting challenge for a student of animal behavior to try to write computer programs that would duplicate the behavior of a particular animal. As a clue for the difficulty this would involve, consider the problem of pattern recognition. Many types of birds are highly discriminating as to the coloring and physical shape of other birds with which they relate in the course of their life. This means that they are able to make fine distinctions between various complex patterns of color and form. However, it has proven very difficult to write computer programs that can discriminate between printed letters.

We can therefore see that the coded instructions that specify animal behavior may be expected to be very complex. Also, we can see no particular reason to suspect that these instructions can be specified by a simple transformation of the instructions for any of the structures or systems of structures in categories (a) through (d). Rather, we should expect that these different categories of information should be independent, in the sense that knowledge of one of them, say the code for the antibody system, would not provide much knowledge about another, such as the coding for transcontinental migration. In symbols this means that we should expect $L(X) \simeq L(X|Y)$ and $L(Y) \simeq L(Y|X)$ if X and Y are the two different instruction codes.

(f) *Human behavior*. This is perhaps the most complex topic of all, and is at present quite beyond the reach of the reduction-

istic speculations of modern science. Most features of human behavior are believed to be independent of the genetic coding system of heredity and are thought to be transmitted from generation to generation by cultural learning processes. The evolution of human behavior is therefore regarded as something different from organic evolution and is called "cultural evolution," although it is presumed to occur according to the same laws as organic evolution. Indeed, modern science recognizes no laws other than the blind mathematical laws of molecular interaction.

Since human culture is an aspect of life, we may legitimately consider it in our estimate of the complexity or information content of living systems. From the reductionistic view of man as a system of molecules, it follows that the information of human culture must be coded in different bodies in some sort of molecular patterns, whether DNA or not. We may gain some idea of the amount of information involved in human culture by considering the numbers of volumes of books to be found in different libraries. To be sure, there is a great deal of redundancy in these books. We wonder, however, whether the total information content of the brain of a single learned person could be as low as our minimum estimate of the information content of a higher living system: 10^5 characters, or about 36 pages.

In order to rule out the bizarre thought that the information content of human knowledge could be this low, we can again perform the breakdown of total information into coded segments as described in inequality (11). If X now represents the text of a series of works of human literature, involving such fields as science, music, and philosophy, we could regard Y_k as a particular passage from one of these texts. If $L(X)$ is very low, then $L(Y_k|X_{k-1}, I)$ will have to be very low on the average, just as in our previous analysis. This would mean that Y_k can be computed from X_{k-1} by a very short program of, say, 48 characters or so. That *Principia Mathematica* can be transformed in this way into the first act of *Macbeth*, or that these together can be thus transformed into the *Eroica* symphony seems very doubtful, to say the least. We therefore conclude that the information content of human culture must be a great deal higher than 10^5 characters.

As one final point, consider again the breakdown of the total genetic code sequence into segments, Y_1, \dots, Y_N . This time, let the lengths of the segments be 700 characters, so that each Y_k specifies about three genes. As before, X_k represents the sequence

formed by joining together Y_1, Y_2, \dots, Y_k . Then there is *one particular program*, G_w , with an index $w = 1, 2, 3, \dots$, which will transform each X_k into Y_{k+1} .²¹ That is, for each k between 1 and N there is a number, w_k , for which

$$Y_{k+1} = G_{w_k}(X_k). \quad (15)$$

This program can be specified very simply in one line.

Now, if $L(X) \leq 10^5$, it turns out that for over 90% of the segments, Y_k , we have

$$Y_{k+1} = G_1(X_k). \quad (16)$$

This implies that the simple transformation, G_1 , serves as a magic formula which by itself is able to generate most of the genetic coding of a human being by repeated application. For most of the remaining segments, G_2 or G_3 will transform X_k into Y_{k+1} .

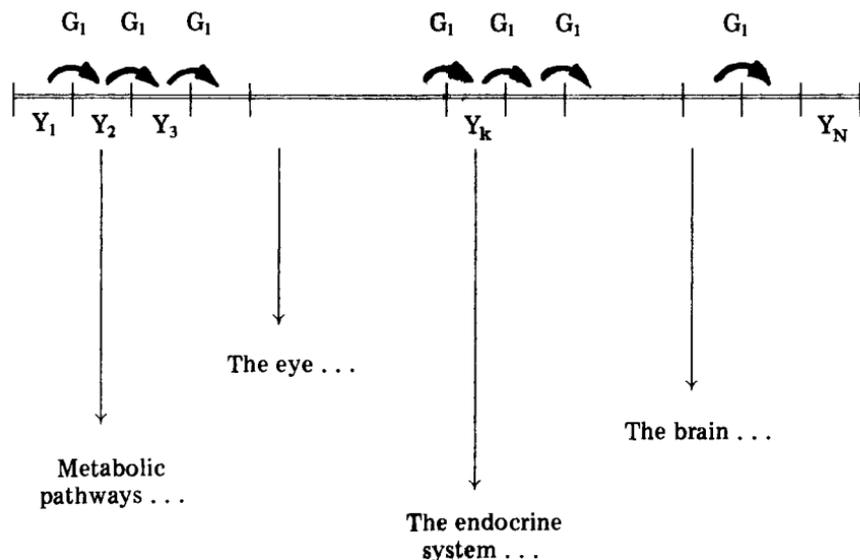


Figure 4. The magic formula, G_1 . Can repeated applications of this simple one-line program generate large stretches (over 90%) of the genetic code? This is implied if this code has a low information content. [$L(X) \leq 10^5$ char.]

If we consider the great variety of information that must be encoded in X , it seems very implausible indeed that great stretches of this information could be simply reeled off by repeatedly applying the transformation, G_1 . This would mean that successive applications of G_1 would generate coding for eyes, hands, the brain, the liver, metabolic pathways, and so on, as we have already described. It seems far more reasonable to suppose that $L(X)$ cannot in fact be as low as 10^5 characters.

**No system of high information content
can evolve from a system
of low information content.**

In section 1, we demonstrated that the models of modern physics for the course of events within the universe have a low information content. Our basic format for a physical model is as follows: We have a system, S , in which events take place in accordance with a set of natural laws, F . Generally F will consist of a differential equation or set of differential equations. S will have initial conditions, S_0 , and boundary conditions, B . These will generally be described by probability distributions, s_0 , and b , which are defined by ensembles of statistical physics such as the Gibbsian canonical ensemble. The probability distribution, s_t , for the system, S , at the time, t , can be calculated, at least in principle, from the equations for s , b , and F by a short computer program. By "short" we mean that this program can be written down in full using no more than four or five pages of solid print (at 70 characters/line and 40 lines/page.)

From the probability distribution, s_t , we can calculate the probability, $M(X)$, that a particular molecular configuration described by a code, X , will be found somewhere within the system, S , at the time, t . $M(X)$ is the probability that the configuration, X , will have "evolved" within the system by the time, t . Given X , the value of $M(X)$ can be calculated from s_t by a few more lines of programming instructions. Thus, the total information content of the function, $M(X)$, i.e. the value of $L(M)$, can be estimated to be no more than five or six pages of computer instructions. This corresponds to no more than 14,000 to 16,800 characters of information:

$$\text{Upper bound for } L(M) \leq 16,800 \text{ characters.} \quad (17)$$

On the other hand, in Section 2, we estimated that the lower bound for the configuration, X , describing the genetic information for the cell of a higher animal could be no less than:

Lower Bound for $L(X) \geq 100,000$ characters. (18)

The implications of this for the theory of evolution are given by the following inequality:

$$M(X) \leq 64^c + \log T + L(M) - L(X), \quad (19)$$

where c is a constant of 32 characters and T is given by:

$T =$ The maximum number of distinct configurations, X , that can appear within one given state of the physical system, S . (20)

This T is no greater than the volume of S divided by the minimum volume within S occupied by a configuration, X . For the volume of S let us take 10^{12} cubic kilometers as an upper bound, since this is greater than the volume of the Earth plus atmosphere. Since we are interested in molecular configurations, let us take one cubic Ångström unit as the minimum volume of a configuration. This gives us $T = 10^{51}$, so that $\log T \approx 28$ using the base of 64.

$$M(X) \leq 64^c + 28 + 16,800 - 100,000 \ll 64^{-80,000} \quad (21)$$

Inequality (21) says that the probability that the configuration, X , will be found anywhere within the volume of the system, S , at the time, t , can be no greater than one out of $64^{80,000}$. Suppose that t is 4.5 billion years—the estimated age of the Earth according to present geological theories. Multiplying $M(X)$ in inequality (21) by $365 \times 4.5 \times 10^9$, we can see that the probability of finding X in the system at the end of any day in a 4.5-billion-year period is also much less than $64^{-80,000}$. We conclude that the entire course of events in S covering a 4.5-billion-year period would have to be repeated over and over again at least $64^{80,000}$ times in order for there to be a reasonable expectation that the configuration X would be seen in S even once. (This is in accordance with what is known as the frequency interpretation of probabilities.) Since $64^{80,000}$ is an enormous number, to say the least, it is reasonable to say that X will not evolve within the system S in 4.5 billion years.

Now, X could represent the genetic code for a higher animal, and S could be a physical model of an earth sized physical system

which might be considered as a candidate model for the evolution of life. Our conclusion is that no higher animal will ever evolve in such a system in any realistic span of time. For our time, t , we could have chosen any time period for which $L(t)$ is small, for the only property of t that affects our considerations is the number of characters needed to express t in the program for calculating the events in S . This means that our t could be anything from one year to billions of billions of years and the conclusion expressed in inequality (21) would still hold.

It is often said that while it is highly unlikely that any given complicated structural part of a living organism will arise by chance, still, given the immense spans of geological time, such structures are bound to arise sooner or later. When they do arise they will be preserved and propagated by the forces of natural selection, and in this way evolution will occur, even though it may seem improbable. The analysis presented here proves that this actually cannot happen in a system governed by simple natural laws and possessing simple initial and boundary conditions. The reason for this is easy to understand. The process of natural selection can only select simple configurations if the basic physical laws on which it runs are simple. Thus, even though various sequences of random change may occur over vast spans of time, simple processes of physical interaction will only be able to select simple patterns from among them, no matter how long the time span.

For this reason, the theory of evolution can be seen to fail at the very weak point that has been criticized by so many students of the theory. Many observers have noted that "natural selection" has never been adequately defined in evolutionary theory, either in Darwin's original version or in the more recent "synthetic theory." According to Darwin, natural selection means "survival of the fittest," but no one, unfortunately, has been able to define which creatures are "fittest," except by saying that they are the ones which survive. A similar problem plagues the modern definition of natural selection as "differential reproduction." The problem is that evolutionists have always had an intuitive feeling that the ordinary interactions between living organisms would serve to select those patterns which, over long periods of time, would transform simple molecular arrangements into modern higher life forms. This intuition was based on commonplace examples. For example, a mutation producing longer legs in a deer might well be selected because it would enable the animal to escape from predators more easily. However, this is no reason to believe that na-

tural selection would possess the discriminating power needed to guide the development of a world of plants and animals from an inanimate primeval slime. What we have shown here is that in a system governed by simple natural laws, no process, whether it be natural selection or any other imagined principle of evolutionary development, is sufficient to do this.

Let us now examine these arguments further in order to understand the basic principle behind them. First of all, note that the upper bound on $M(X)$ given by inequality (21) will become reasonably large only if we increase the complexity of our system of natural laws to the point where $L(M)$ is nearly equal to $100,000 - 28 - c$. Now, 100,000 characters corresponds to about 36 solid pages of coding. Modern science has never reached the point of even considering that nature may run according to laws of this complexity, and it is unlikely that such considerations could ever be practical, given the limitations of the human mind. At any rate, the basic laws of present-day physics are based on the Schrödinger equation and a few basic potentials. They are far less complicated than this.

Of course, $L(M)$ might be increased by increasing the complexity of the initial conditions or the boundary conditions of the system, S . But this would amount to building instructions for the genetic code, X , into the initial or boundary conditions. This would certainly violate the requirements of a theory of evolution. If instructions for X are built into the boundary conditions, then as time passes in S we will at some point find instructions for X coming across the boundary into S . In this case it would be valid to say that X has originated from outside of S and been transmitted into S in some form. But this is not evolution. Likewise, if instructions for X were built into the initial conditions, it would be valid to say that X in some form was already present in S in the beginning. But this is not evolution either. Also, if we code instructions for X into the laws of transformation of S , then we are saying that X is an inherent part of the fundamental absolute nature of things. This also is not evolution.

This point can be made clearer by the following stronger version of the inequality (19):

$$M(X) \leq 64^{c + \log T - L(X|M)} \quad (22)$$

This inequality shows that there is a reasonable probability of

finding X in S at time, t , only if $L(X|M)$, the amount of information needed to specify X given M , is not very much greater than $c + \log T$. This constant may be taken to be no more than 60 characters, since $c = 32$ characters in this case. This essentially means that most of the information needed to code X must already be there in M .

In a study of actual evolution, we would not be dealing with simply one genetic code for one organism. Rather, we would expect to find that many different kinds of plants and animals had evolved, and these may be designated by X_1, \dots, X_K , where K will certainly be in the thousands, if not millions. Since $L(X_k|M)$ would have to be small for each X_k in order for it to have a reasonable probability of being found in S , it follows that $L(X_1 X_2 \dots X_K|M)$ would have to be fairly small also—certainly no greater than $K(c + \log T)$. For a selection of 1,000 organisms, this comes out to 60,000, or less than our minimum estimate of 10^5 for a single higher organism. This means that M would have to contain coding in some form for most of the information required to specify all these organisms.

In this connection the basic inequality,

$$L(X) \leq L(X|M) + L(M) \quad (23)$$

indicates that $L(M)$ would have to be quite large, since, as we have pointed out in Section 2, the great diversity of form and function among the different species of life indicates that the total information content, $L(X)$, for a selection, $X = X_1 X_2 \dots X_K$, of many different species should be very high.

Our essential point, then, is that in a system operating according to a set of physical laws, the only forms at all likely to appear in the course of time are those coded into M . This means that these forms are coded into the initial conditions of S , or that coded instructions for them are transmitted across the boundary at some time, or that they are coded into the laws of transformation themselves. In any case, they cannot be said to evolve; rather, they are built into the system in the first place, or they come from outside.

We would like to emphasize that a complex code picked at random cannot be expected to convey a specific complex message. For example, if cosmic rays enter S across its boundary, and their energies and directions of motion are given by a simple statistical distribution, then the system, S , cannot be expected to extract a

large amount of specific information from them in the form of a particular organism with genetic code, X. This is what inequality (22) implies. In order for a large amount of information to be obtainable from the cosmic rays, their statistical distribution function, b , would have to be complex. Also, it could not be just any complex function, but would have to be specifically chosen so that $L(X|b)$ was small. In other words, a large amount of specific information for X would have to be coded into the cosmic ray stream.

This shows that the vague notion that "little bits of negentropy from the sun" can add up over a long time period to yield a living organism of high information content is mistaken. In popular literature advocating the theory of chemical evolution, the negative of the thermodynamic quantity called entropy has been interpreted as a measure of information content. This measure is commonly referred to as "negentropy." Since the biosphere of the earth is thermodynamically an open system, it is possible for the negentropy of a physical entity in the biosphere to increase at the expense of a decrease in the negentropy of the sun, without violating the second law of thermodynamics. This has been interpreted to mean that information is transmitted in the sunlight passing from the sun to the Earth, and that this information can add up to define the structures of living organisms.²² This argument is fallacious because the negative of entropy is not a measure of information content. For example, a crystal at absolute zero has the highest possible negentropy value (of zero), but it presumably has a very low information content, since a simple program will specify the periodic arrangement of atoms in a crystal lattice. We can see from the analysis of information content that the negentropy argument is wrong. In order for the structural patterns of organisms to be extracted from sunlight, they would in great measure have to be coded into the sunlight in a very specific form.

This possibility is a stronger version of the panspermia hypothesis of Svante Arrhenius. According to this theory, life may never have had an origin, and it may have been spread throughout the universe in the form of spores that could travel in outer space from one favorable planet to another. Arrhenius thought that these spores were primitive and would evolve once they reached their given planet. We can see from this analysis, however, that "primitive" spores would not do—the spores would have to contain instructions for all of the organisms that later evolved. Owing to a variety of serious drawbacks, the theory proposed by Arrhen-

ius never had much standing in science, and these stronger requirements are not likely to make it much more acceptable. Another idea, sometimes discussed in science fiction stories, is that higher intelligences might transmit instructions for organisms across space in the form of coded radio beams.

Both the hypothesis that instructions for the design of living organisms were transmitted across the boundary of S and the hypothesis that they were contained in the initial conditions for S fail to account for the origin of these instructions. Essentially, they amount to the hypothesis that life is eternally existent as a material phenomenon. This suffers from the defect that although the action of simple physical laws cannot, as we have seen, generate information, it can certainly obliterate it. It is common experience that many physical processes tend to destroy order and information, both on a small and on a large scale. This makes it very hard to see how high levels of information content could be maintained eternally in a material system, and there indeed seem to be no theories as yet developed that seriously propose this.

One further consideration involving inequalities (19), (21), and (22) is the following: Are there so many possible varieties within a particular species that even though it is highly unlikely that any one of them would evolve in S, it is nonetheless reasonably probable that at least one out of the total collection would evolve? If that were the case, one could still speak of the evolution of that species as being a possibility. The many varieties of the organism could be more or less trivial variations all possessing the basic features entitling them to be referred to as examples of that species. After all, geneticists have observed that no two individuals of a given species have identical genotypes (with the possible exception of identical twins).

We can easily show, however, that this is not possible. Basically, in order to sufficiently increase the probability of finding at least one genetic code, X, out of our collection, C, of codes, we would have to make C so large that all the members of C could not possibly share in common the basic distinguishing traits of a particular kind of organism. In fact, C would have to be so large that only a very small percentage of its members could even be expected to possess the basic features of life as we know it. Most of the codes in C would simply be arbitrary patterns not corresponding to any kind of viable living organism.

In order to show this, let us first estimate how large C would have to be. Consider a collection, C, of N different varieties of an

organism, each of which has a complexity level of $L(X) \geq 10^5$ characters. By inequality (21) the probability of finding even one of these varieties in S at time, t , is bounded by,

$$\text{Prob}(C) \ll N 64^{-80,000} . \quad (24)$$

In order for $\text{Prob}(C)$ to be reasonably large, we would have to have $N > 64^{80,000}$. In comparison, we may note that 10^{75} is a very generous upper bound on the total number of organisms that could have lived on the earth in a 4.5-billion-year-period.²³ Even though life on the earth shows an impressive diversity of form, naturalists have estimated that there are only some two to three million species of living plants and animals (including micro-organisms).²⁴ Also, although one writer has estimated the total number of fossil forms at ten million, only about a hundred thousand of these have been actually described by paleontologists.²⁵ These figures should provide some perspective on the magnitude of the number, $64^{80,000}$, which we are considering.

Let us specifically examine the subclass, C_K , of codes in C with information content, $L(X|M) = K$. Our bounds, (17) and (18), in conjunction with inequality (23) show that we should consider values of $K \geq 100,000 - 16,800$. There are at most 64^K codes in total having the particular information content of K . Yet from inequality (22) we can see that the number, N_K , of codes in C_K must be

$$N_K \geq 64^{K - 60 - A} , \quad (25)$$

in order to have $\text{Prob}(C_K)$ as great as 64^{-A} . What this means is that in order for $\text{Prob}(C_K)$ to be reasonably large, the number of codes in C_K must be of nearly the same order of magnitude as the total number of codes having information content, K .

Should we expect such a large proportion of the codes of information content, K , to specify living organisms? We can easily see the answer to be no. Consider that for each viable organism there are vast numbers of defective forms that are not viable. Each gene code for a viable organism is subject to many thousands of different fatal mutations; and n kinds of fatal mutations can be combined in 2^n different ways, most of which should specify non-viable forms. Beyond this there should be enormous numbers of forms which do not correspond to any living organism at all, and

most of which follow no coherent pattern that would convey any meaning to us or appear at all familiar.

The class, C_K , should therefore have a substantially lower order of magnitude than the total class of codes of information content, K . Since C is composed at most of a few billion C_K 's for K 's between our lower and upper bounds on genetic information content, it follows that $\text{Prob}(C)$ must be very small if C is restricted to the class of life forms. [$\text{Prob}(C)$ is the sum of the different $\text{Prob}(C_K)$'s.]

Further light is shed on this matter by the concept of information shared "in common" by a set of codes. We would expect all the members of a particular category of organisms to share in common certain information which defines that category of form. For example, if there is any meaning to the term "horse," there must be certain specific information common to all horses and not shared fully with any other forms. The arguments in Section 2 indicate that for a higher animal, this characteristic defining information should have a content of at least 10^5 characters.

For a class of forms of fixed information content, there must be a trade-off between the number of forms in the class and the amount of information they can share in common. The totality of codes of information content, K , share no information in common, and subclasses of nearly the same order of magnitude as this class (such as C_K) can share very little. For this reason, we would not expect such subclasses to correspond to well-defined complex categories of forms, such as "horses" or even "vertebrates."

The concept of shared information is formally defined in the appendix. The following inequality gives a bound on the probability of finding a configuration, X , belonging to a certain class, C , in the physical system at time, t .

$$\text{Prob}(\text{Some } X \text{ in } C \text{ exists in } S \text{ at time, } t) \leq 64^{c' + \log T + L(M) - K}. \quad (26)$$

Here it is assumed that C is a collection of codes, all of which possess K characters of specific information in common. The class of horses should share in common the information for all the essential features that go into the structure of a horse. As described in Section 2, this includes information for eyes, legs, circulatory systems, etc., and should have an information content of at least 10^5 characters. Substituting 10^5 for K , we can conclude that no representative of this class of forms is likely to evolve.

Thus, the chances that a particular form or reasonably close variant of that form will evolve within a physical system are extremely small unless the information specifying that form is essentially built into the initial conditions, boundary conditions, or natural laws of the system.²⁶ Multiplying by 10^{75} (our upper bound on numbers of organisms), we can conclude that the chances are exceedingly small that any life form will evolve which has ever existed on the earth in 4.5 billion years and which has a greater information content than $L(M) + \log T + 75 \log 10 + c'$.

**What is the origin of life?
Some implications of this analysis.**

We have argued that the information for the structures of living organisms must either be built into the initial and boundary conditions of the system, or must be built into the system's fundamental laws. In the first case, we are left with the question of where the information came from. It must have had some origin, either on the other side of the boundary or at a time previous to the initial time, t_0 . In the second case, we are confronted with a system of absolute, invariant laws that encode all of the detailed information for thousands of species of living beings, for human civilizations, symphonies, etc. In either case, we have a world picture with a very dissatisfying degree of incompleteness or disunity. Either the information for the intricate and harmonious world we perceive is coming from an unknown source, or it is coded into a kind of abstract cosmic laundry list of laws that satisfy (due to their very irreducibility) no unifying principle.

We have already rejected the alternative that the information for life forms originated in the boundary or initial conditions of the system. Essentially, this alternative leaves us with an infinite regress that is worse than no explanation at all: as we push the boundaries of the system further and further back in space and time we must continue to find the same intricate information coded in some kind of material configurations. This picture of things is especially unsatisfactory because the transformations of matter occurring in nature have a strong tendency to obliterate information. We will not consider this picture further here.

The other alternative is that the information is built into the fundamental natural laws. It is more or less typical for scientists to believe that the laws they know at the moment are the ultimate laws—and this is especially prevalent today in the life-sciences. There is, however, much evidence that the laws known to science today are by no means complete or final. Even in the realm of biochemistry there is increasing evidence of phenomena that do not fit into the current theories. These phenomena give support to the idea that many laws operating in nature are still unknown.

For example, the research of C. L. Kervran in France has

shown that many chemical elements can be created and destroyed within the bodies of living organisms in the course of their ordinary metabolism.²⁷ He describes these phenomena as "biological transmutations" wherein organisms can transform one element into another by nuclear reactions without resorting to the high energies and temperatures these reactions require according to our present understanding. The transmutations are of interest because they directly indicate the existence of as yet unknown laws which are intimately involved with the phenomena of life. This analysis implies that the search for such laws should prove to be a very fertile field of future scientific investigation. Far from their being "impossible," we should expect to find many such remarkable phenomena. (That biochemists have failed to observe these phenomena in spite of their extensive chemical analyses of living organisms, demonstrates the blinding effect that narrow adherence to one limited hypothesis can have on human understanding.)

Let us explore more thoroughly the idea that the information for life forms is stored in an extensive system of what could be called "higher order laws." Fundamental laws are conceived of in modern science as being invariant in time and all-pervading in space. This is somewhat mysterious, but it becomes more so if we require that these laws encompass large amounts of mathematically irreducible information specifying living organisms, cultures, and so forth. We must suppose that large numbers of invariant, all-pervading laws in some sense "exist." Yet the very fact that they possess a high information content means that they cannot be mathematically reduced to any unified scheme.

We are thus confronted with an invisible and intractable collection of absolute rules that just so happen to specify a variegated world of living beings. This alternative is also unsatisfactory, yet we cannot avoid it as long as we adhere to the basic mathematical approach to the study of nature outlined in Section 1.

We are now forced to reject both of our original alternatives for explaining life as entirely unsatisfactory. Neither constitutes what we could call an explanation. Is there any other recourse? We suggest that a reasonable and satisfying alternative can be found only if we are willing to go beyond the strictly mathematical approach to nature. We can do this by introducing "life" as a basic principle, in and of itself, which is not capable of exhaustive mathematical description.

The idea is that "life" is something irreducible—something that cannot be explained in terms of combinations of other non-living

things. An analogy would be the concept of electric charge, which in current physical theories is also a fundamental entity that cannot be expressed in terms of anything else. In this picture, the laws of nature cannot be finally represented by fixed mathematical equations. This doesn't mean, of course, that equations cannot exist which approximate these laws. We may expect to find many different laws operating in nature. However, these should exist as derived entities depending ultimately on the higher, non-mathematical principle of life. This picture also requires that the physical combinations of elements which form the bodies of living organisms—and which are commonly regarded as constituting life itself—must be regarded instead as byproducts of life.

Since our analysis shows that the final cause underlying the phenomena of nature must contain information for all the forms of living beings, we propose that this final cause is itself a primordial living being lying completely outside the realm of mathematical describability. All of the temporary, numerically describable patterns of the material manifestation arise from this being. The individual living entities found in nature may be considered minute quantized parts of this being, just as electrons are quantized particles of electric charge. As certain electrical properties are characteristic of electric charge, likewise complex and yet harmonious activity—that is, intelligence—is characteristic both of this primordial being and of the derived quantized parts.

We are considering here the question of what is the character of the absolute truth, or the fundamental, underlying cause of phenomena. The driving motive behind science—at least in its purer forms—has always been to arrive at an understanding of this ultimate cause. (This is of course denied by the philosophy of positivism, which holds that science is and can be nothing more than the search for regular patterns in the welter of sense data.²⁸) The implication of our analysis is that no coherent, unified view of nature is in fact possible as long as we accept the hypothesis that nature is mathematically describable. Although this hypothesis has dominated science for the last two to three hundred years, we would like to suggest that it is, in fact, unnecessary. We are proposing that a valid and fruitful alternative exists which circumvents the frustrating impasse inherent in this viewpoint.

From the point of view of scientific procedure, it is perfectly valid to consider our hypothesis, especially if we hope to obtain actual knowledge of reality from scientific study. Even from the point of view of positivism, this hypothesis may be regarded as an

admissible proposal that will bear implications for the expected ordering of sense data. With a strictly numerical approach to empirical data, we can only hope to deal with numerical patterns and mathematical relationships. From such a viewpoint, the primordial living being can be manifest only in terms of more and more complicated and irreducible patterns of material interaction. Here, then, is one falsifiable prediction of this hypothesis: we may expect to find unlimited patterns and relationships in the world that defy reduction to any finite scheme. (This, of course, has been the gist of this paper.)

Philosophically, the problem of unity in nature is solved by the proposal that the absolute truth is a purposeful living being. Traditionally, scientists have felt that there is unity and harmony in nature and have sought to find this unity in the extreme simplicity of the natural laws. Thus, Einstein sought to find a unified field theory, which would express the laws of both gravity and electromagnetism in one formula. Such unity is ruled out if the natural laws have a high level of complexity.

However, this sought-after unity can be attained if we consider sentient life itself irreducible and absolute. The harmony in diversity lies in the purposeful comprehension and direction of the variegated cosmic manifestations by the ultimate sentient being. We are dealing here with the question of what should be regarded as the irreducible absolute truth: we are basically forced to choose between a senseless multitude of arbitrary rules and conditions and a higher unifying principle lying (necessarily) outside the grasp of mathematical expression. That this higher principle should have the character of life follows from simple empirical observation: we have been led to the consideration of complex laws and conditions from the observation of life.

One of the basic aspects of the reductionistic approach to understanding nature is that it eliminates "meaning" and "purpose" from the picture altogether. According to this approach, simple physical causation is to be put in the place of intelligent purpose. However, since we have eliminated *simple* physical causation as a possible cause of life, we propose that "meaning" and "purpose" should be reinstated. This requires the existence of conscious intelligence as the fundamental causal principle in nature. This not only makes sense out of the otherwise meaningless display of complex form without purpose, but it also ties in the theme of consciousness in a unified way.

The phenomenon of consciousness fits neatly into place if we

suppose that life is an absolute principle. In Western scientific thought there has been great difficulty in fitting consciousness into nature in any unified way. The standard view of science has been that consciousness is simply a phenomenon of material interaction—or a dangling epiphenomenon. However, no one has ever been able to indicate how the interaction of insentient entities can produce conscious awareness. In fact, consciousness seems to have an inherently irreducible nature.

This problem is solved if we accept conscious self-awareness as an inherent feature of the irreducible absolute truth. Consciousness can then be seen to occupy a natural role as a fundamental unifying characteristic of the absolute living being. The individualized consciousness of the multitudes of living beings can be understood in a simple and unified way if these beings are regarded as minute quantized parts of the single absolute being. These minute parts derive all of their properties, including that of conscious awareness, from this absolute source. The bodies of the living organisms may be regarded as carriers of these “quanta of life.”

In this paper our basic argument, drawn from information theory, supports the understanding that the unified absolute cause lying behind nature must be an irreducible living being. The other concept we have introduced—that of a class of quantized individual parts of this being—does not directly follow from this argument. One might suppose that the one, absolute living being had simply generated a series of automatons out of material elements. However, the concept of the quantum of life follows naturally from these considerations: Many individual conscious beings (such as ourselves) *do* exist, and consciousness is not reducible to a combination of material elements. Since consciousness plays a natural unifying role as an inherent feature of the absolute truth, the many conscious beings may be most simply seen as its quantized parts.

A more detailed analysis of consciousness will be presented in other papers of this series. We would simply like to stress here that all of these concepts—far from being unscientific, or even anti-scientific—are potential subjects for truly scientific investigation and understanding. We might also point out that we are not violating the principle—known as Occam’s razor—that one should not multiply entities beyond necessity. We feel that the picture we have outlined here is more coherent and economical than the alternative picture of thousands of piecemeal laws.

As a final point, we should say something about the nature of

chance. Thus far we have referred to “chance” without giving it any explicit definition, and we have manipulated it mathematically according to standard formulas. In this way we have calculated that the chance that a given form will evolve goes down exponentially with the difference between its information content and the information content of the system.

What is the meaning of chance? There are two basic conceptions of chance. According to the first, chance is a measure of the ignorance of an observer about some event. According to this view, the event itself is perfectly definite, but the observer refers to it in terms of chance because his knowledge is imperfect. In this view, chance, is subjective and may vary from one observer to another. According to the second view, there actually exist chance events, which are by nature undetermined and which obey only statistical laws.

The first viewpoint is represented by the so-called classical theory of probability, in which equal chances are assigned to alternative events when there is no reason to expect one of them to happen in favor of another. This is clearly a means of describing one’s ignorance numerically. The second viewpoint is represented by the relative-frequency interpretation of probability, in which the probability of an event is measured by repeating the circumstances of the event many times and seeing how often it comes up. The standard example is the flipping of a coin. Here it is often supposed that the event itself is random. An example is the interpretation of the state functions of quantum mechanics, which are believed to give an inherently random description of phenomena such as radioactive decay and the movement of atomic particles.

The analysis presented here strongly favors the first point of view, since the inequalities show a reciprocal balance between information content and probability. As the gap between the information needed for Y and the information provided by the system goes up, the probability of finding Y correspondingly goes down. Thus, probability can be seen as measuring a lack of information.

It is therefore interesting to note that in an analysis of several standard theories of probability, the mathematician T. L. Fine found that none gave a reasonable characterization of the second interpretation of probability.²⁹ Circular reasoning and unavoidable contradictions were found in each proposed method of describing “chance” as an actual phenomenon of nature. Thus, only

the first point of view seems to provide a tenable interpretation of chance.

According to the world picture presented here, we expect all phenomena to be of an essentially deterministic character. The statistical nature of the present quantum mechanical laws can be seen as a symptom of their incompleteness. In general, we expect that natural laws can be only approximately described by mathematical equations, and it is not surprising for the mathematical approximations to contain a statistical element reflecting their incomplete nature. The laws themselves can be understood as manifestations of the primordial living being.

Thus we conclude that life exists as the absolute source of the material manifestation, and that the process by which the physical forms of living entities are created is quite different from the process envisioned by the theory of evolution. The slogan of evolutionary theory has been, "Life comes from matter." This analysis shows that this is wrong. The real slogan should rather be: "Both life and matter come from life."

Appendix. Mathematical derivations.

In order to define information content, we must first establish a fixed "computer," C. By

$$C(P|f_1, \dots, f_n) = X \quad (27)$$

we mean the following: The computer, C, executes the program, P, yielding an output of X. The functions, f_1, \dots, f_n , are externally supplied functions that may be used by the program, P, in the course of its calculations (subroutines). X may be either a number or a function. If X is a function, it is understood that the arguments of X must be supplied as inputs to the program, P, which calculates X.

We shall assume that all programs are to be written in a standard programming language with an alphabet of 64 characters. We will specify a number of special features of this language as we need them. Programs can be coded as binary numbers. If P is such a number, *define* $l(P)$ to be the number of binary digits in P (including leading zeros.) We will refer to the lengths of programs both in characters and in binary digits, or "bits." (One character equals 6 bits.)

For the purpose of calculating solutions to equations arising in physics, the programming language may be assumed to have facility for expressing standard mathematical operations. It is also convenient to assume that this language can refer to numbers with arbitrarily large numbers of digits.

Define the *information content*, $L(X|f_1, \dots, f_n)$, of X given f_1, \dots, f_n to be the length of the shortest program (or programs) that will compute X, given that f_1, \dots, f_n are available to refer to during the calculation.

In the previous sections, we measured information content, L, in characters, since this seems more easy to visualize. In this section we shall always measure L in binary bits, unless otherwise mentioned. This is more convenient in the mathematical derivations.

This method of measuring information content was first devised by the Russian mathematician Kolmogorof.³⁰ The subject

has been developed further by G. Chaitin.³¹ Another name we shall use for information content is *complexity*.

The first thing to note about L is that relatively few X's have a low information content. The number of X's with information content, $L(X|f_1, \dots, f_n)$, less than k is no more than 2^k . This is because there are no more than 2^k programs, P , with $l(P)$ less than k . It is interesting to note that this remains true no matter how many functions, f_1, \dots, f_n , are available for reference by P .

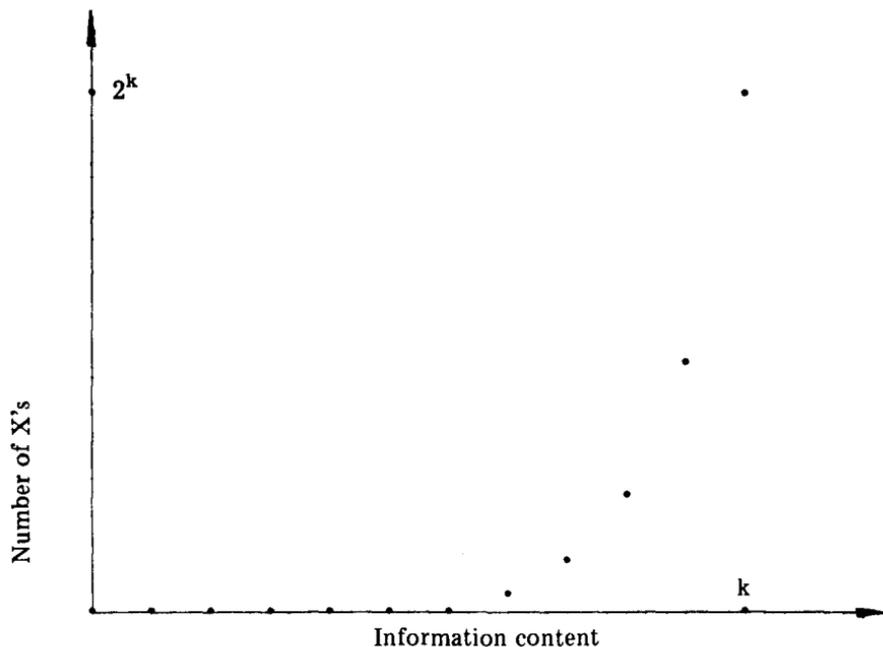


Figure 5. Upper bound on the number of X's with information content less than k .

Define $s[R(\cdot), w]$ to be the w th number, X , in numerical order, for which the statement, $R(X)$, is true. This is considered to be defined only if there is such an X . We shall assume that the function, $s(\cdot, \cdot)$, which is defined in this way may be referred to in the programming language. (It could otherwise be expressed by means of a "do loop.") $R(X)$ must be a statement that can be expressed in the computer language. We shall be interested in statements of the form, $A \geq B$, where A and B are numbers or expressions for the calculation of numbers.

We shall need some provision in the programming language for coding information in the most compact way possible. One way of doing this is to adopt the following convention for the coding of positive integers: A positive integer, j , can be coded in the form, $n; b_1 b_2 \dots b_n$, where n is the number of binary bits (0 or 1) needed to write j in binary notation and $b_1 b_2 \dots b_n$ is the actual binary expression for j . Both n and the semicolon are coded in terms of characters, each of which requires six binary bits to code. This enables us to encode j in such a way that if $j < 2^m$ then

$$L(j) \leq m + L(m) + 6. \quad (28)$$

The '6' corresponds to the coding for the semicolon.

We shall also need a method whereby programs can be joined together. Let us make the convention that a program must be in the form of a numerical constant, $xx \dots x$, a function, $f(\dots)$, or a string of symbols surrounded by parentheses, (\dots) . We shall call a function or variable symbol in a program *free* if it is used but not calculated in the program. Suppose x is a free function or variable symbol appearing in the program, P . P might be of the form, $(\dots x \dots x \dots x \dots)$, with several appearances of x . If x is a variable symbol, let Q be a program that calculates a numerical constant, and if x is a function symbol requiring n variables, let Q be a program computing a function of n variables. Interpret $(\dots Qx \dots x \dots x \dots)$ as the program obtained by executing P and using Q to compute a value for x whenever x is used. This notation unambiguously establishes Q as a subroutine for P .

As a result of this convention we have the basic relationship,

$$L(Y) \leq L(Y|X) + L(X). \quad (29)$$

Using these conventions, we can derive the following basic theorem:

Proposition 1. Let $M(X)$ be a function mapping non-negative integers into non-negative real numbers for which,

$$\sum_X M(X) \leq T. \quad (30)$$

Then it follows that,

$$M(X) \leq 2^c + 1 \log_2 T - L(X|M). \quad (31)$$

The constant, c , is no more than 191 bits, or 32 characters.

Proof Let k be a positive integer, and let N be the number of X 's for which $M(X) > 2^{-k}T$. By equation (30) we know that $N(2^{-k}T)$ is less than T . It follows that if $M(X) > 2^{-k}T$, then we can write,

$$X = s[M(\cdot) > 2^{-k}T, w] \quad (32)$$

for some $1 \leq w \leq N < 2^k$. If we write $k; b_1 b_2 \dots b_k$ for w , this gives us an expression for X which requires 17 characters (including multiplication and exponentiation symbols) and k binary bits and which refers to k , T , and M as unknowns. Therefore,

$$L(X|M, k, T) \leq 6 \times 17 + k. \quad (33)$$

This further implies that,

$$L(X|M) \leq 6 \times 17 + L(T) + L(k) + k. \quad (34)$$

Let $b = L(X|M) - L(T) - 103$. Write $L_m(j) = \max [L(1), L(2), \dots, L(j)]$. If we pick $k = b - L_m(b)$, then it follows that $L(k) + k \leq b$. Since $103 = 6 \times 17 + 1$, this means that (34) is violated for this choice of k . This means that $M(X) \leq 2^{-k}T$, or

$$M(X) \leq 2^{103 + L(T) + L_m(b) + 1} T - L(X|M). \quad (35)$$

For the constant, c , let us pick some number bigger than the sum of the first three terms in this exponent of 2. We expect to take a value of 10^{51} for T , in which case $L(T) = L(10^{51}) = 5 \times 6$. We will not be worried about X 's with $L(X)$ much bigger than about 2^{40} . If $b < 2^{40}$, then $L_m(b) \leq 3 \times 6 + 40$. Adding these together, we get $c = 191$ bits, or less than 32 characters. Q.E.D.

Corollary 1. Suppose that the function, M , is the same as in proposition 1. Then,

$$M(X) \leq 2^{c + \log_2 T + L(M) - L(X)}. \quad (36)$$

Proof This follows from proposition 1 and (29). Q.E.D.

Corollary 2. Suppose that M is the same as in proposition 1. Suppose that G is a function mapping non-negative integers to non-negative integers. Then,

$$\sum_{X: G(X) = Y} M(X) \leq 2^{c' + \log_2 T + L(M) - L(Y|G)}. \quad (37)$$

Proof Define $M_1(Y)$ to be the left-hand side of (37). Then,

$$\sum_Y M_1(Y) = \sum_X M(X) \leq T. \quad (38)$$

If we perform the derivation of (36) with G as an externally supplied function, we obtain:

$$M_1(Y) \leq 2^{c + \log_2 T + L(M_1|G) - L(Y|G)}. \quad (39)$$

The left-hand side of (37) gives us an expression for M_1 in terms of M , G , and c_1 bits worth of additional programming to express the sum. This means that $L(M_1|G) \leq c_1 + L(M)$. The desired inequality follows if we take $c' = c_1 + c$. Q.E.D.

Definition. The numbers, X_1, \dots, X_m , shall be said to possess K bits of information *in common* if there is a function, G , such that $G(X_1) = G(X_2) = \dots = G(X_m) = Y$, and $L(Y|G) \geq K$.

The idea behind this definition is that the function, G , represents a program or process of analysis intended to extract information from the X 's. If the same Y can be extracted from all of the X 's, then they are said to possess in common the information in this Y . $L(Y|G)$ is used to measure the amount of this information instead of $L(Y)$ because no information that Y obtains from G itself could be said to belong to the X 's.

In the reductionistic view of the world, according to which a human being is regarded as a computing machine, the action of a human observer may be taken as an example of the function, G .

Thus, a series of books would be said to have the statement Y in common if an observer would be able to write down Y after examining any one of the books. Of course, any aspects of Y originating in the observer himself might not be actually shared by these books.

Using this definition, Corollary 2 can be rewritten as follows: If X_1, \dots, X_m have K bits of information in common, then,

$$M(X_1) + \dots + M(X_m) \leq 2^{c'} + \log_2 T + L(M) - K. \quad (40)$$

Here is yet another way of looking at this basic situation: Suppose that $R(X, Y)$ is an equivalence relation on the non-negative integers. For example, if X and Y are regarded as descriptions of living organisms, then $R(X, Y)$ might mean "X and Y both belong to the same species of life." The relation, R, divides the X's into disjoint equivalence classes, C_1, C_2, \dots, C_n . Let us define the *simplest member* of the class, C_j , to be that X in C_j which has the smallest value of $L(X|R)$. (If there are several such X's, pick the numerically smallest.) Also define the *minimum amount of information* in C_j to be equal to this value of $L(X|R)$. Call this I_j . In our example, I_j is the information content of the simplest individual organism in the species, C_j . We use $L(X|R)$ for this instead of $L(X)$ because this information should be independent of the process of observation, R, which determines whether or not two individuals belong to the same class.

The following theorem holds for these equivalence classes:

Corollary 3. Let M be the same as in proposition 1, and let R be an equivalence relation, as described above. Then for each j,

$$\sum_{X \text{ in } C_j} M(X) \leq 2^{c''} + \log_2 T + L(M) - I_j \quad (41)$$

Proof Define $G(X) = s[R(\cdot, X), 1]$. Then $R(X, Y)$ holds if and only if $G(X) = G(Y)$. It follows that $I_j \leq L(Y|R) \leq L(Y|G) + c_2$ for Y in C_j . By corollary 2 and $c'' = c' + c_2$ we obtain the inequality, (41). The constant, c_2 , is the number of bits needed to express R in terms of G. This can be done by the statement, " $R(X, Y)$ iff $G(X) = G(Y)$." Q.E.D.

Let us briefly indicate some of the applications of proposition 1 and its corollaries. Consider a quantum mechanical system with variables for the position of N atomic nuclei corresponding to a

variety of different elements. (There will also be spin variables, electromagnetic variables, and variables for the positions of electrons.) The bonding together of various atoms to form molecules can be expressed in terms of the relative distance between these various nuclei. A necessary condition that two atoms be bonded together is that their corresponding nuclei be situated within a certain distance of one another. This distance will depend on the type of atoms and bond being considered. The pattern of bonds making up a particular molecule can be encoded into the digits of a binary number, X , in some standard way. Using this coding scheme and the data for interatomic bonding distances, one can define a function, $B_X(\bar{q}_1, \dots, \bar{q}_N)$, which will equal 1 if the nuclear positions, $\bar{q}_1, \dots, \bar{q}_N$, are so arranged that some $\bar{q}_{i_1}, \dots, \bar{q}_{i_k}$ among them satisfy the spacing conditions necessary for the molecule described by X . This function will otherwise be set equal to 0.

The function, B_X , provides us with an operator that enables us to calculate the probability, $M(X, \psi)$, that the molecule described by X exists within the physical system when that system is in the quantum mechanical state, ψ . (Or, rather, it provides us with an upper bound on this probability, since we are dealing only with a necessary condition for the existence of the molecule.) Specifically, we have,

$$M(X, \psi) = \int \psi^* (\bar{Q}) \psi (\bar{Q}) B_X(\bar{q}_1, \dots, \bar{q}_N) d\bar{Q}, \quad (42)$$

where the integral is taken over all of the variables of ψ .

For any given arrangement of the variables, $\bar{q}_1, \dots, \bar{q}_N$, there will be only a limited number of molecular configurations, X , that can be found among them. That is, there should be a number, T , for which,

$$\sum_X B_X(\bar{q}_1, \dots, \bar{q}_N) \leq T \quad (43)$$

for all arrangements of $\bar{q}_1, \dots, \bar{q}_N$.

In order to obtain an estimate of T , let us make one further restriction of B_X . Suppose that $B_X(\bar{q}_1, \dots, \bar{q}_N)$ shall equal 1 only if there is a collection, $\bar{q}_{i_1}, \dots, \bar{q}_{i_k}$, of these variables that satisfy the conditions for X but are not within bonding distance of any other \bar{q}_i 's. Then $B_X = 1$ will mean that the molecule represented by X is present without being part of a larger molecule. This means that each \bar{q}_i can be part of at most one configuration cor-

responding to an X with $B_X(\bar{q}_1, \dots, \bar{q}_N) = 1$. Therefore, $T \leq N$. In order to estimate N , let us consider that Avogadro's number, the number of molecules in one molecular weight of a substance, is about 6×10^{26} . For the entire earth we can estimate that there will be no more than 10^{51} atoms. These values indicate what we might expect of T in biological applications of proposition 1.

The function, $M(X, \psi)$, can be calculated using (42) and the calculations for ψ in terms of natural laws, initial conditions, and boundary conditions. We can also calculate the probability of finding X in a system described by a mixture of states. If the mixture consists of the states, ψ_i , with probability, a_i , then this probability is given by $\sum a_i M(X, \psi_i)$. Thus, the total information content of the function, $M(X)$, which bounds the probability of finding the configuration, X , in the system at a given time, t , is given by:

$$L(M) \leq L(\text{initial conditions}) + L(\text{boundary conditions}) + L(\text{natural laws}) + L(t) + L(B) + \text{constant.} \quad (44)$$

The constant equals the number of symbols needed to express (42), plus some other odds and ends. B refers to B_X , with X considered as a variable. The time, t , is counted since it enters into the calculation of the state of the system. $L(t)$ will be no more than the number of characters needed to express t in base 32. This is small for t ranging from 0 to billions of billions of years.

Our next step is to derive a theorem that will make it possible to estimate a lower bound for $L(X)$. Define the function, $I(n, X)$, as follows: $I(n, X)$ is the number of programs, P , for which $C(P) = XY$ for some Y , and $l(P) = n$. By XY we mean the concatenation of the two strings, X and Y . For completeness we also define $I(n, \emptyset)$ to be the number of programs, P , for which $C(P) = Y$ for some Y , and $l(P) = n$. Clearly, $I(n, \emptyset) \leq 2^n$.

Proposition 2. If X and Y are two non-negative integers, then

$$I(n, XY) \leq I(n, X) 2^{c - L(Y|X, I, n)}. \quad (45)$$

The constant, c , is no more than 43 characters.

Proof It follows from the definition of I that,

$$\sum_{Y: l(Y) = h} I(n, XY) \leq I(n, X). \quad (46)$$

Define $M(Y) = x(1(Y),h) I(n,XY)$ and $T = I(n,X)$. Here $x(a,b)$, the indicator function, is taken to be 1 if $a=b$ and 0 otherwise. We shall assume that the computer language can refer to this function.

With these definitions we can see that the conditions of proposition 1 apply. If we carry out the steps of the proof of proposition 1, we find that inequality (35) becomes

$$I(n,XY) \leq I(n,X) 2^{L_m(b) + 6 \times 39 + 1 - L(Y|X,I,n)} \quad (47)$$

as long as $1(Y) = h$. Here we assume that symbols for concatenation of strings and exponentiation to powers of 2 are included in the programming language. We also assume that the language can handle numbers of base 32, and that h is expressed by a number of this kind of no more than four digits. As we are interested in cases where $L(Y|X,I,n) < 32^4$, we may take four characters for $L_m(b)$. This gives us a figure of about 43 characters for c .³² Q.E.D.

Now, suppose that we have Y_1, \dots, Y_N , with $1(Y_1), \dots, 1(Y_N) < 32^4$. Let $X_j = Y_1 Y_2 \dots Y_j$, for $j = 1, \dots, N$. Write $X_0 = \emptyset$.

Corollary 4. Let $n = L(X_N)$. Then,

$$L(X_N) \geq \sum_{j=1}^N [L(Y_j|X_{j-1},I,n) - c]. \quad (48)$$

Proof By proposition 2,

$$I(n, X_N) \leq I(n, \emptyset) 2^{\sum_{j=0}^{N-1} [c - L(Y_{j+1}|X_j, I, n)]} \quad (49)$$

as long as $L(Y_{j+1}|X_j, I, n) \leq 32^4$. Since $n = L(X_N)$, we have $I(n, X_N) \geq 1$. We also have $I(n, \emptyset) \leq 2^n$. As a result, n plus the exponent of 2 in (49) must be non-negative. Q.E.D.

The purpose of corollary 4 is to provide a means of estimating a lower bound for $L(X_N)$ for a long string of digits, X_N . The idea is to break X_N into a number of segments, Y_j . Corollary 4 indicates that if $L(X_N)$ is very low, then the average value of $L(Y_j|X_{j-1}, I, n)$ must be very close to c . This means that Y_j can be calculated from X_{j-1} by means of a program using n and I which can be written down in little more than $c = 43$ characters. If $n = L(X_N) < 32^4$ then we can write,

$$L(X_N) \geq \sum_{j=1}^N [L(Y_j|X_{j-1}, I) - c'] , \quad (50)$$

where $c' = c + L(n)$ is no more than 4 characters larger than c . This is the equation we used in Section 2.

If X_N is the genetic coding sequence for a living organism, especially a higher plant or animal, then the Y_j 's will correspond to genetic coding for different organs, tissues, and other features of the organism. Since it seems highly implausible that these features can be mapped into one another by transformations requiring no more than about c symbols to write down, we conclude (as described in Section 2) that $L(X_N)$ must be quite large. This certainly agrees with the intuition that the living organisms must be highly complex in structure.

Let us examine this arrangement in greater detail. Define

$$K(X, Y) = \min \{ k : I(n, XY) > 2^{-k} I(n, X) \} . \quad (51)$$

It follows that,

$$I(n, X_{j+1}) \leq I(n, X_j) 2^{1 - K(X_j, Y_{j+1})} . \quad (52)$$

If we let $K_j = K(X_{j-1}, Y_j)$, then the reasoning used in the proof of corollary 4 gives us:

$$L(X_N) \geq \sum_{j=1}^N (K_j - 1) , \quad (53)$$

when we choose $n = L(X_N)$.

Define the function, $F_w(h, n, X)$ by,

$$F_w(h, n, X) = s[x(1(\cdot), h) \times I(n, X \cdot) > 2^{-1(w)} I(n, X), w] . \quad (54)$$

The dots (\cdot) indicate that $Y = F_w(h, n, X)$ is the w th Y in order satisfying the indicated inequality. As we have pointed out in the proofs of propositions 1 and 2, there are less than 2^k Y 's for which $1(Y) = h$ and $I(n, XY) > 2^{-k} I(n, X)$. If k is chosen to be $k = K(X, Y)$, then this inequality is satisfied. Therefore, if we are given X and Y with $1(Y) = h$, we can write,

$$Y = F_w(h, n, X) , \quad (55)$$

where w is a binary integer of $k = K(X,Y)$ digits. [This may include leading zeros. We want $1(w) = k$ in (54).]

Summing this up, we obtain the following result:

Proposition 3. Suppose that $1(Y_1) = 1(Y_2) = \dots = 1(Y_N) = h$ and let $X_j = Y_1 Y_2 \dots Y_j$. Then there are numbers, K_1, \dots, K_N and binary integers, w_1, \dots, w_N for which $1(w_j) = K_j$ and,

$$Y_j = F_{w_j}(h, n, X_{j-1}), \quad (56)$$

and

$$L(X_N) = n \geq \sum_{j=1}^N (K_j - 1). \quad (57)$$

Proof This is proven above. Q.E.D.

From (51) we can see that $K(X,Y)$ must be positive, since $I(n,XY) \leq I(n,X)$. Therefore each K_j in proposition 3 must be positive. Let us consider what happens when $L(X_N) = aN$ for some $0 < a \leq 1$. This means that the average value of K_j will be no more than $a+1$. This implies that the number of K_j 's which equal 1 must be at least $N(1-a)$. For these K_j 's we will have $w_j = 1$, since w_j cannot be zero. It follows from this that at least $N(1-a)$ Y_j 's will be given by $F_1(h, n, X_{j-1})$. If some of the K_j 's are very large, even larger numbers of them must equal 1 in order to preserve the average, and hence even more Y_j 's will be generated by this formula. Also, if $K_j = 2$, then Y_j is given by F_1, F_2 , or F_3 in (56).

Let us apply these results to the case where X_N is the genetic coding for an organism and the Y_j 's are a series of segments of this code. It certainly seems unlikely that so many Y_j 's could be generated from X_{j-1} by such a simple transformation as F_1 . F is given by equation (54), and I is defined just before proposition 2. The function, I , depends only on the properties of the computer language, and so it seems quite unreasonable to suppose that it could have the property, in conjunction with equation (54), of transforming the genetic coding for, say, the liver into the coding for the eye. For this reason, we conclude that $L(X_N)$ cannot be very small.³³

Notes

1. Watson, *Molecular Biology of the Gene*, pg. 105.
2. Watson, pg. 67.
3. *Newton's Principia*, pg. lxviii.
4. von Helmholtz, *Über die Erhaltung der Kraft*, pg. 6.
5. The quantum mechanical equations are taken from von Neumann, *Mathematical Foundations of Quantum Mechanics*, and Messiah, *Quantum Mechanics*.
6. Fisher, *The Genetical Theory of Natural Selection*.
7. See, for example, Oparin, *The Origin of Life* and Orgel, *The Origins of Life*.
8. This is discussed in Tolman, *Statistical Mechanics*.
9. von Neumann, *Theory of Self-Reproducing Automata*.
10. The data for *E. Coli* is taken from Watson, chap. 3.
11. Watson, pg. 85.
12. Watson, pg. 507.
13. Watson, pg. 508.
14. Watson, pg. 461.
15. Watson, pg. 540.
16. Watson, pg. 94.
17. Satir, "How Cilia Move," *Sci Amer*, Oct. 1974.
18. Berg, "How Bacteria Swim," *Sci Amer*, Aug. 1975.
19. Macbeth, *Darwin Retried*, pg. 101.
20. Alfred Russel Wallace, *The World of Life*.
21. Specifically, $G_w(X) = F_w[4200, L(X_N), X]$, where F is given in equation (54) of the appendix. [$L(X_N)$ should be in bits here.]
22. This argument can be found in Orgel, *The Origins of Life*.
23. This is more than the number of cubic Ångström \times microseconds in the volume of the biosphere \times 4.5 billion years.
24. Percy E. Raymond, "The First Animals and Plants" in Preston Cloud, ed., *Adventures in Earth History*, pg. 667.
25. Normal D. Newell, "The Nature of the Fossil Record" in Preston Cloud, ed., *Adventures in Earth History*, pg. 649.
26. This neglects the small constants, c , c' , etc., which appear in the various inequalities. These are small and leave the basic conclusion unaffected. They are artifacts of the mathematical derivations and do not seem conceptually necessary.
27. Kervran, *Biological Transmutations*.
28. The doctrine of positivism holds that one can speak meaningfully only of sense perceptions, and that all other categories are mere verbiage. This view is described by the physicist Y. Freundlich as follows: "To us, the statement that trains have wheels when they are not in the station (when we are not sensing them) means that at the station they will have wheels."

This is, to us, a very satisfying solution, for having thus defined existence we proceed to speak of wheels on the train even when it is not in the station. In general, attributing a property to a system means that certain predictions about the system can be made." ("Mind, Matter, and Physicists," *Found. of Phys.*, Vol. 2, No. 2/3, 1972, pgs. 130-1.) Taken literally, this doctrine implies that the statement "Man evolved from a primate ancestor" means that certain bones may be seen in certain museums. This is a far from satisfying solution to us, and, we believe, to many others. It seems far more reasonable to suppose that science has simply failed insofar as it cannot deal with statements about actual reality.

29. Fine, *Theories of Probability*.
30. Kolmogorof, "Logical Basis for Information Theory and Probability Theory," *IEEE Trans. IT-14* (1968), pgs. 662-4.
31. Chaitin, "Information Theoretic Computational Complexity," *IEEE Trans. IT-20* (1974), pgs. 10-15.
32. Write $\lg(x)$ for the number of digits in base 32 needed to write down x . Then we have $L_m(b) \leq 6 \times \lg(b) \leq 6 \times \lg[L(Y|X, I, n)]$. Then we can replace the exponent of 2 in (45) by:

$$1 + 6 \times 39 + 6 \times \lg[L(Y|X, I, n)] - L(Y|X, I, n) \quad (58)$$

This is more complicated, but it enables us to avoid worrying about the bound of 32^4 on $L(Y|X, I, n)$. This can also be carried over to (48) in corollary 4.

33. We should note that the function, I , may be taken to be a recursive function. To insure this, we need only require that the computer language have a criterion for rejecting programs that may fail to produce their output in a finite time. This can easily be done in a system that is suited for the calculations required by modern physics.

Bibliography

- Berg, Howard C. "How Bacteria Swim." *Scientific American*, Vol. 233 No. 2, August 1975, pp. 36-44.
- Chaitin, G.J. "Information Theoretic Computational Complexity." *IEEE Trans. IT-20* (1974), pp. 10-15.
- Cloud, Preston, ed., *Adventures in Earth History*. San Francisco: W.H. Freeman & Co., 1970.
- Fine, Terrence L. *Theories of Probability*. New York and London: Academic Press, 1973.
- Fisher, Ronald. *The Genetical Theory of Natural Selection*, 2nd ed. New York: Dover, 1958.
- Freundlich, Y. "Mind, Matter, and Physicists." *Foundations of Physics*, Vol. 2, No. 2/3, 1972.
- Kervran C. Louis. *Biological Transmutations*. Binghamton, New York: Swan House Publishing Co., 1972.
- Kolmogorof, A.N. "Logical Basis for Information Theory and Probability Theory." *IEEE Trans. IT-14* (1968), pp. 662-664.
- Macbeth, Norman. *Darwin Retried: An Appeal to Reason*. Boston: Gambit, 1971.
- Messiah, Albert. *Quantum Mechanics*. New York: Interscience Pub., 1961-63.
- Newton's Principia*, Motte, Andrew, trans. New York: Daniel Adee, 1846.
- Oparin, A.I. *The Origin of Life*, 2nd ed. New York: Dover, 1953.
- Orgel, Leslie E. *The Origins of Life*. New York: Wiley, 1973.
- Satir, Peter. "How Cilia Move." *Scientific American*, Vol. 231, No. 4, October 1974, pp. 44-52.
- Tolman, Richard C. *The Principles of Statistical Mechanics*. Oxford: Clarendon Press, 1938.
- von Helmholtz, H. *Über die Erhaltung der Kraft*, Ostwald's Klassiker der Exakten Wissenschaften Nr. 1, 1847.
- von Neumann, John. *The Mathematical Foundations of Quantum Mechanics*. Princeton: Princeton University Press, 1955.
- von Neumann, John. *Theory of Self-Reproducing Automata*, ed. Arthur Burkes. Urbana, Illinois: University of Illinois Press, 1966.
- Wallace, Alfred Russell. *The World of Life*. New York: Moffat, 1911.
- Watson, James D. *Molecular Biology of the Gene*, 2nd ed. New York W.A. Benjamin, 1970.

About the Author

Sadāputa dāsa (Richard Thompson) was born in Binghamton, New York, on February 4, 1947. In 1969 he earned his B.S. degree in mathematics from the State University of New York at Binghamton, and in 1970 he earned his M.A. in mathematics from Syracuse University. After receiving a National Science Fellowship in 1970, he completed his Ph.D. in mathematics at Cornell University in June of 1974, specializing in probability theory and statistical mechanics. His dissertation has been published as Memoir number 150 of the American Mathematical Society, "Equilibrium States on Thin Energy Shells."

Throughout his studies, the author was struck by the lack of any meaningful foundation to reality in modern scientific theories. His dissatisfaction with this culminated in 1970, when he studied the reduction of man to a Turing machine, a kind of abstract clockwork with a few moving parts. Surely, he felt, the truth must be something different from this. Consequently, he began to study many different philosophies, with a view to finding a practical route to higher knowledge.

In 1972 he discovered some of the books of His Divine Grace A.C. Bhaktivedanta Swami Prabhupāda in a book store in Ithaca, New York, and was struck by the beauty of their conceptions and the clarity of their presentation. Later he met the disciples of Śrīla Prabhupāda at the Rādhā-Krishna Temple in New York City. Here, he found, was a deeply meaningful philosophy capable of practical application in day-to-day life. He became formally initiated as Śrīla Prabhupāda's disciple in 1975 at the temple of Śrī Śrī Gaur-Nitai in Atlanta, Georgia. He is now a full-time member of Bhaktivedanta Institute for Higher Studies.