THE LANGUAGE OF LIFE

David Berlinski

April 1985
CP-85-20

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
2361 Laxenburg, Austria

FOREWORD

This paper represents the written version of a lecture given at IIASA in September 1984 under the auspices of the Science & Technology and the Regional Issues projects. In its current form it will appear as a chapter in the forthcoming IIASA book, ***Complexity, Language and Life: Mathematical Approaches***, J. Casti and A. Karlqvist, eds.

Boris Segerstahl
Leader
Science & Technology Program

ABSTRACT

This paper explores the idea that life comprises a language-like system. The arguments are carried out against the background of the neo-Darwinian theory of evolution. The principal conclusion is the dilemma that if life is a language-like system, then certain concepts are missing from the Darwinian paradigm; if not, then Darwinian thought is suspicious in the sense that its principles do not naturally apply to cognate disciplines.

AFFILIATION

David Berlinski is a Researcher at the Institut des Hautes E'tudes Scientifique, Bur-sur-yvette, Paris, France.

The Language of Life

David Berlinski

Introduction

In the spring of 1984, I delivered two lectures at IIASA under the title *The Language of Life*. Dianne Goodwin was kind enough to prepare a verbatim transcript of my talks; I have used the months since then to purge the written record of what I said of its incoherence, vagrant inaccuracies, and general slovenliness.

This chapter is at once long and terse – an unhappy combination, and one that makes severe demands of the reader. Many arguments are highly compressed and must be elaborated before they appear convincing. I have not hesitated to make use of mathematical concepts in expressing myself; but I draw no mathematical conclusions. I thus run the risk of alienating the general reader even as I antagonize the mathematician. For these reasons, it may be helpful if in this introduction I endeavor to place this chapter in a somewhat wider personal and intellectual context.

As it stands, *The Language of Life* represents a draft of one-third of a larger work entitled *Language, Life and Logic*. Another part of that more ambitious project was delivered at IIASA two years ago as a set of lectures. The written record of those lectures, which I hope to publish separately as a working paper, is entitled *Classification and its Discontents*.

My aim in *Language, Life and Logic* is to explore a certain complicated common ground that holds between language, on the one hand, and the graphic arts, on the other. These are the classic systems of representation of the human imagination. In both, there is a curious division between the system's syntactic and semantic structures: a theory, for example, consists of a finite set of sentences, the sentences of words; paint and then pigment comprise a painting; and yet, words and sentences, paints and pigments, manage, somehow, to cohere and, then, in a miraculous act of self-transcendence, to make contact with a distinct and different external world. The problems of theoretical biology, it might seem, have nothing much to do with issues that arise in the philosophy of language or the philosophy of art. Not so. A gene comprises a linear array of nucleotides that under certain conditions expresses a protein or set of proteins. The proteins, in turn, are organized to form a structure as complicated as a moose or a mouse. The nucleotides are plainly alphabetic or typographic in character; the organism itself is rich, complex, complete, continuous, unlike an alphabet. How is it, then, that such typographic structures as DNA manage to express so much that is not typographic at all? This is a question quite similar to questions that might be raised about language itself, or works of the graphic arts; and when it is pursued, certain metaphors and quite peculiar images begin drifting from one subject to the other. There is the notion of meaning, of course, which is common to language, art, and life; but also the idea that life is itself a language-like system; or that art is organic. The relations of satisfaction, representation, and expression, while formally distinct, of course, nonetheless display points of contact. In order to explain how it is that a painting may represent a face, for example, one has recourse to the notion of a metaphor, a concept from the philosophy of language and linguistics; to make sense of gene expression, one deals in concepts such as code, codon, information, and regulation. In a general way, a theory, a painting, and a gene belong to the class of interpreted or significant typographic objects. It is for this reason that it has seemed to me profitable to explore some of the concerns of theoretical biology and the philosophy of art and language in a single volume.

Within the context of *this* chapter, my aim is to explore the ramifications of a controlling metaphor: the idea that life comprises a language-like system. I do this against the background of the neo-Darwinian theory of evolution – the most global and comprehensive scheme of thought in theoretical biology. My argument at its most general is constructed as a dilemma: if life is a language-like system, then certain concepts are missing from Darwinian thought; if not, then Darwinian thought is suspicious in the sense that its principles do not naturally apply to cognate disciplines. The intellectual pattern to this chapter is thus one of movement between two unyielding points, a kind of whiplash.

Part One establishes the historical and contemporary background to Darwinian thought; and makes the argument that much of biology cannot be reduced to physics. In Part Two, I consider the confluence of certain concepts: distance in the metric spaces of organisms and of strings, metric spaces in phase, complexity, simplicity, Kolmogorov complexity, the ideas of a weak theory, and a language-like system. Part Three plays off concepts of probability against the hypothesis that molecular biological words are high in Kolmogorov complexity – with results that are inconclusive. In Part Four, I examine evolution or biological change as a process involving paths of proteins. The discussion is set in the mathematical contexts of ergodic theory and information theory. In many respects, the classical concepts of information and entropy are most natural in discussing topics such as the generation of protein paths by means of stochastic devices; but there is a connection between Kolmogorov complexity and entropy in the sense of information theory, which remains to be explored. Almost all of Part Four represents a tentative exploration of concepts that require, and will no doubt receive, a far fuller mathematical treatment.

Many of the points that I make in this paper I first discussed with M. P. Schutzenberger in Paris in 1979 and 1980. Indeed, it was our intention and hope to publish jointly a monograph on theoretical biology. This has not come to pass. Still, to the extent that my ideas are interesting, they are his; to the extent that they are not, they are mine.

John Casti read the penultimate draft of this essay and discovered any number of embarrassing errors. I am grateful for his stern advice, which I have endeavored to heed.

PART ONE

A System of Belief

The natural thought that theoretical biology comprises a kind of intellectual Lapland owes much to the idea that biology itself is somehow a derivative science, an analogue to automotive engineering or dairy management, and, in any case, devoid of those special principles that lend to the physical or chemical sciences their striking mahogany lustre. This is the position for which J.J.C. Smart (1963) provided a classic argument in *Philosophy and Scientific Realism*.[1] Analytic philosophers, for the most part, agree that nothing in the nature of things compels them to learn organic chemistry; Feyerabend, Putnam, and Kuhn have wondered whether *any* discipline can properly be reduced to anything at all; and, then, whether anything is ever scientific, at least in the old-fashioned and honorific sense of that term.[2] Naive physicists – the only kind – are all too happy to hear that among the sciences physics occupies a position of prominence denied, say, to urban affairs or agronomy. The result is *reductionism from the top down*, a crude but still violently vigorous flower in the philosophy of science. The physicist or philosopher, with his eye fixed on the primacy of physics, thus needs to sense in the other sciences – sociology, neurophysiology, macrame, whatever – intimations of physics, however faint. This is easy enough in the case of biochemistry: chemistry is physics once removed; biochemistry, physics at a double distance. Doing biochemistry, the theoretician is applying merely the principles of chemistry to living systems: like the Pope, his is a reflected radiance.

In 1831, the German chemist Uriel Wohler synthesized urea, purely an organic compound – the chief ingredient in urine, actually – from a handful of chemicals that he took from his stock and a revolting mixture of dried horse blood. It was thus that organic chemistry was created: an inauspicious beginning, but important, nonetheless, if only because so many European chemists were convinced that the attempt to synthesize an organic compound would end inevitably in failure. The daring idea that all of life – I am quoting from James Watson's textbook (1965), *The Molecular Biology of the Gene* – will ultimately be understood in terms of the "coordinative interaction of large and small molecules" is now a commonplace among molecular biologists, a fixed point in the wandering system of their theories and beliefs. The contrary thesis, that living creatures go quite beyond the reach of chemistry, biochemists regard with the alarmed contempt they reserve for ideas they are prepared to dismiss but not discuss. Francis Crick, for example, devotes fully a third of his little monograph, *Of Molecules and Men*, to a denunciation of vitalism almost ecclesiastical in its forthrightness and utter lack of detail.[3] Like other men, molecular biologists evidently derive some satisfaction from imagining that the orthodoxy they espouse is ceaselessly under attack.

Theoretical biologists still cast their limpid and untroubled gaze over a world organized in its largest aspects by Darwinian concepts; and so do high-school instructors in biology – hardly a group one would think much inclined to the idea of the survival of the fittest; but unlike the theory of relativity, which Einstein introduced to a baffled and uncomprehending world in 1905, the Darwinian theory of evolution has never quite achieved canonical status in contemporary thought, however much like a cold wind over water its influence may have been felt in economics, sociology, or political science. Curiously enough, while molecular genetics provides an interpretation for certain Darwinian concepts – those differences between organisms that Darwin observed but could not explain – the Darwinian theory resists reformulation in terms either of chemistry or physics. This is a point apt to engender controversy. Woodger, Hempel, Nagel, and Quine cast reduction as a logical relationship: given two theories, the first may *directly* be reduced to the second when a mapping of its descriptive apparatus and domain of interpretation allows the first to be derived from the second. I am ignoring details, now. The standard and, indeed, the sole example of reduction successfully achieved involves the derivation of thermodynamics from statistical mechanics. In recent years, philosophers have come to regard direct reduction with some unhappiness. There are problems in the interpretation of historical terms: the Newtonian concept of mass, for example; and theories that once seemed cut from the same cloth now appear alarmingly incommensurable. Kenneth Schaffner has provided a somewhat more elaborate account of reduction: his definition runs to five points.[4] By a *corrected theory*, he means a theory logically revived to bring it into conformity with current interpretations: Newton upgraded, for example. His general scheme for reduction, then, is this:

(1) All of the terms in the corrected theory must be matched to terms in the reducing theory – a requirement of *completeness*.
(2) The corrected theory must be deducible from the larger theory, given the existence of suitable reduction functions – a requirement of *derivability*.
(3) The corrected theory must indicate why the original theory was incorrect – a requirement of epistemological *insight*.
(4) The original theory must be explicable in terms of the reducing theory – a requirement of *cogency*.
(5) The original and corrected theories must resemble each other – a requirement of intellectural *symmetry*.

In the case of theoretical biology, to speak crisply of deriving, say, molecular biology from biochemistry is rather like endeavoring to cut steel with butter: there is a certain innocence to the idea that molecular biology has anything like a discernable logical structure. What one actually sees is a mass of descriptive detail, a bewildering plethora of hypothetical mechanisms, much by way of anecdotal evidence, a few tiresome concepts, and an array of metaphors drawn from physics, chemistry, information theory, and cybernetics. The definition of reduction just cited is, in addition, incomplete, its flagrant inapplicability aside. In Mendelian genetics, the concept of a gene is theoretical, and genes figure in that theory as abstract entities. To what should they be pegged in molecular genetics in order to reduce the first theory to the second? DNA, quite plainly, but how much of the stuff counts as a gene? "Just (enough) to act as a unit of function," argues Michael Ruse, a philosopher whose commitment to prevailing orthodoxy is a model of steadfastness.[5] The functions that he has in mind are biochemical: the capacity to generate polypeptides; but to my way of thinking, the reduction achieved thus is illicit. In biochemistry, the notion of a unit of function is otiose, unneeded elsewhere. To the extent that molecular genetics is biochemistry, it does not reflect completely Mendelian genetics; to the extent that it does, it is not biochemistry, but biochemistry beefed-up by extrinsic concepts, a conceptual padded shoulder. What holds in a limited way for molecular genetics holds in a much larger way for molecular biology. Concepts such as code and codon, information, complexity, replication, self-organization, stability, negative entropy (grotesque on any reckoning), transformation, regulation, feedback, and control – the stuff required to make molecular biology work – are scarcely biochemical: the biochemist following some placid metabolic pathway need never appeal to them. Population genetics, to pursue the argument outward toward increasing generality, is a refined and abstract version of Darwin's theory of natural selection, applied directly to an imaginary population of genes: selection pressures act directly on the molecules themselves, a high wind that cuts through the flesh of life to reach its buzzing core. Has one achieved anything like a reduction of Darwinian thought to theories that are ***essentially*** biochemical, or even vaguely physical? Hardly. The usual Darwinian concepts of fitness and selection appear unvaryingly in place. These are ideas, it goes without saying, that do not figure in standard accounts of biochemistry, which very sensibly treat of valences and bonding angles, enzymes and metabolic pathways, fats and polymers – anything but fitness and natural selection. To Schaffner's list of five, then, I would add a sixth: ***no reduction by means of inflation*** – a contingent and cautionary restriction that, for the time being at least, enforces a stern separation between biology and mathematical physics.

The Darwinian theory of evolution is the great, global organizing principle of biology, however much molecular biologists may occupy themselves locally in determining nucleotide sequences, synthesizing enzymes, or cloning frogs. Those biologists who look forward to the withering away of biology in favor of biochemistry and then physics are inevitably neo-Darwinians, and the fact that *this* theory – *their* theory – is impervious to reduction they count as an innocent inconsistency. If mathematical physics offers a vision of reality at its most comprehensive, the Darwinian theory of evolution, like psychoanalysis, Marxism, or the Catholic Faith, comprises, instead, a *system of belief*. Like Hell itself, which is said to be protected by walls that are seven miles thick, each such system looks especially sturdy from the inside. Standing at dead center, most people have considerable difficulty in imagining that an outside exists at all.

The Historical Background

Charles Darwin completed his masterpiece, *On the Origin of Species*, in 1859. He was then forty-nine, ten years younger than the century, and not a man inclined to hasty publication. In the early 1830s, he had journeyed to the islands of the South Atlantic as a naturalist aboard *H.M.S. The Beagle*. The stunning diversity of plant and animal life that he saw there impressed him deeply. Prevailing biological thought had held that each species is somehow fixed and unalterable. Looking backward in time along a line of dogs, it is dogs all the way. Five years in the South Atlantic suggested otherwise to Darwin. The great shambling tortoises of the Galapagos, surely the saddest of all sea-going creatures, and countless subspecies of the common finch, seemed to exhibit a pattern in which the spokes of geographic variation all radiated back to a common point of origin. The detailed sketches that Darwin made of the Galapagos Finch, which he later published in *On the Origin of Species*, show what caught his eye. Separated by only a few hundred miles of choppy ocean, each subspecies of the finch belongs to a single family; and yet, Darwin noted, one group of birds had developed a short, stubby beak; another, living northward, a long, pointed, rather Austrian sort of nose. The variations among the finch were hardly arbitrary: birds that needed long noses got them. By 1837, Darwin realized that what held for the finch might hold for the rest of life and this, in turn, suggested the dramatic hypothesis that far from being fixed and frozen, the species that now swarm over the surface of the Earth *evolved* from species that had come before in a continuous, phylogenetic, saxophone-like slide.

What Darwin lacked in 1837 was a theory to account for speciation, but the great ideas of fitness and natural selection evidently came to him before 1842, for by 1843 he had prepared a version of his vision, and committed it to print in the event of his death. He then sat on his results in an immensely slow, self-satisfied, thoroughly constipated way until news reached him that A.R. Wallace was about to make known *his* theory of evolution. Wallace, so far as I know, had never traveled to the South Atlantic, sensibly choosing, instead, to collect data in the East Indies, and, yet, considering the same problem that had earlier vexed Darwin, he had hit on precisely Darwin's explanation. The idea that Wallace might hog the glory was too much for the melancholic Darwin: he lumbered into print just months ahead of his rival; but in science, as elsewhere, even seconds count.

The theory that Darwin proposed to account for biological change is a conceptual mechanism of only three parts. It involves, in the first instance, the observation that living creatures vary naturally. Each dog is a member of a common species and thus dog-like to the bone; but every dog is doggish in his own way: some are fast, others slow, some charming, and others bad from the first, suitable only for crime. Darwin wrote before the mechanism of genetic transmission was understood, but he inclined to the view that variations in the plant and animal kingdoms arise by *chance*, and are then passed downward from fathers to sons.

The biological world, Darwin observed, striking now for the second point to his three-part explanation, is arranged so that what is needed for survival is generally in short supply: food, water, space, tenure. Competition thus ensues, with every living thing scrambling to get his share and keep it. The struggle for life favors those organisms whose variations give them a competitive edge. Such is the notion of *fitness*. Fast feet make for fitness among the rabbits, even as a feathery layer of oiled down makes the Siberian swan a fitter foul. At any time, those creatures fitter than others are more likely to survive and reproduce. The winnowing in life effected by competition Darwin termed *natural selection*.

Working backward, Darwin argued that present forms of life, various and wonderful as they are, arose from common ancestors; working forward, that biological change, the transformation of one species to another, is the result of small increments that accumulate, step by inexorable step, across the generations, until natural selection recreates a species entirely. The Darwinian mechanism is both random and determinate. Variations occur without plan or purpose – the luck of the draw; but Nature, like the House, is aggressive; organized to cash in on the odds.

The Central Dogma

Everything that lives, lives just once. To pass from fathers to sons is to pass from a copy to a copy. This is not quite immortality, even if carried on forever, but it counts for something, as every parent knows. The higher organisms reproduce themselves sexually, of course, and every copy is copied from a double template. Bacteria manage the matter alone, and so do the cells within a complex organism, which often continue to grow and reproduce after their host has perished, unaware, for a brief time, of the gloomy catastrophe taking place around them. It is possible, I suppose, that each bacterial cell contains a tiny copy of itself, with the copy carrying yet another copy; biologists of the early eighteenth century, irritated and baffled by the mystery of it all, actually thought of reproduction in these terms: peering into crude, brass-rimmed microscopes, they persuaded themselves that on the thin, stained glass, they actually saw a homunculus; the more diligent proceeded to sketch what they seemed to see. The theory that emerged had the great virtue of being intellectually repugnant. Much more likely, at least on the grounds of reasonableness and common sense, is the idea that the bacterial cell contains what Erwin Schroedinger called a *code script* – a sort of cellular secretary organizing and recording the gross and microscopic features of the cell. Such a code script would be logically bound to double duty. As the cell divides in two, it, too, would have to divide without remainder, doubling itself to accommodate two bacterial cells where formerly there was only one. Divided, and thus doubled without loss, each code script would require powers sufficient to organize anew the whole of each bacterial cell. The code script that Schroedinger (1945) anticipated in his moving and remarkable book, *What is Life*? – he wrote in the 1940s – turns out to be DNA, a long and sinewy molecule shaped rather like a spiral in two strands. The strands themselves are made of stiff sugars, and stuck in the sugars, like beads in a sticky string, are certain chemical bases: adenine, cytosine, guanine, and thymine: A, C, T, and G, in the now universal abbreviation of biochemists. It is the alternation of these bases along the backbone of DNA that allows the molecule to store information.

One bacterial cell splits in two: each is a copy of the first. All that physically passes from cell to cell is a strand of DNA: the message that each generation sends faithfully into the future is impalpable, abstract almost, a kind of hidden hum against the coarse wet plops of reproduction, gestation, and birth itself. James Watson and Francis Crick provided the correct description of the chemical structure of DNA in 1952. They knew, as everyone did, that somehow the bacterial cell in replication sends messages to each of its immediate descendents. They did not know how. As it turned out, the chemical structure of DNA, once elaborated, suggests irresistably a mechanism both for self-replication and the transmission of information. In the cell itself, strands of DNA are woven around each other and by an ingenious twist of biochemistry matched antagonistically: A with T, and C with G. At reproduction, the cell splits the double strand of DNA. Each half floats for a time, a gently waving genetic filament; chemical bonds are then repaired as each base fastens to a new antagonist, one simply picked from the ambient broth of the cell and clung to, as in a single's bar. The process complete, there are now two strands of double-stranded DNA where before there was only one.

What this account does not provide is a description of the machinery by which the genetic code actually organizes a pair of new cells. To the biochemist, the bacterial cell appears as a kind of small sac enclosing an actively throbbing biochemical factory; its products are proteins chiefly – long and complex molecules composed, in their turn, of twenty amino acids. The order and composition of the amino acids along a given chain determines which protein is which. The bacterial cell somehow contains a complete record of the right proteins, as well as the instructions required to assemble them directly. The sense of genetic identity that marks *E. Coli* as *E. Coli* and not some other bug must thus be expressed in the amino acids by means of information stored in the nucleotides.

The four nucleotides, we now know, are grouped in a triplet code of 64 codons or operating units. A particular codon is composed of three nucleotides. The amino acids are matched to the codons: C–G–A, for example, to arginnine. In the translation of genetic information from DNA to the proteins, the linear ordering of the codons themselves serves to induce a corresponding linear ordering first onto an intermediary, messenger RNA, and then onto the amino acids themselves – this via yet another messenger, transfer RNA. The sequential arrangement of the amino acids finally fixes the chemical configuration of the cell.

Molecular biologists often allude to the steps so described as ***the Central Dogma***, a queer choice of words for a science.

The dour Austrian monk, Gregor Mendel, founded the science of genetics on purely a theoretical notion of a gene, which he likened to a bead on a string. In DNA, one looks on genetics bare: the ultimate unit of genetic information is the nucleotide. All that makes for difference, and hence for charm, in the natural world, and which is not the product of culture, art, artifice, accident, or hard work, all this, which is brilliantly expressed in maleable flesh, is a matter of an ordering of four biochemical letters along two ropey strands of an immemorial acid.

The Central Dogma describes genetic replication; but the concepts that it scouts plainly illuminate Darwinian theory from within. Whether as the result of radiation or chemical accident, letters in the genetic code may be scrambled; one letter shifted for another; entire codons replaced, deleted, or altered. These are genetic mutations: arbitrary, because unpredictable; and yet enduring, because they are variations in the *genetic* message. The theory by which Darwin proposed to account for the origin of species and the nature of biological diversity now admits of expression in a single English sentence. Evolution, or biological change, so the revised, the *neo*-Darwinian theory, runs, is the result of natural selection working on random mutations.

PART TWO

Evolutionary Theories

The popular view of evolution tends to be a tight shot on a tame subject: the dinosaur, who did not make it; the shark, who did; but the maturation of an organism is itself much like the evolution of a species; only our intimate acquaintance with its precise and unhesitating character suggests, misleadingly, I think, that the two processes differ in degree of freedom. Psychology, economics, urban affairs, anthropology, political science, and history also describe processes that begin in a state of satisfying and undemanding simplicity, and end later with everything complex, unfathomable, chaotic. The contrast to physics is sobering.

The dynamics of evolutionary theories are often divided into two conceptual stages. In economics, there are macro- and micro-economic theories, aggregate demand versus the theory of the firm; within linguistics, language at the continuous level of speech, and language some levels below, discrete, a matter of the concatenation of words or morphemes. Biology, too, is double-tiered: above, the organism prances; unseen, below, at a separate level, its life is organized around the alphabetic nucleotides.

Metric Spaces

By a metric space S I mean a space upon which a function

$$d: S \times S \rightarrow R_+$$

has been defined, assigning to each pair of points s, s' in S a nonnegative real number – the *distance* $d(s, s')$ – and satisfying the usual axioms:

$$d(s, s') = 0 \Longleftrightarrow s = s' \; ; \tag{9.1}$$

$$d(s, s') = d(s', s) \; ; \tag{9.2}$$

$$d(s, s') + d(s', s'') \geq d(s, s'') \; . \tag{9.3}$$

Double metrics

The distance between organisms

The disciplines of comparative anatomy and systematic zoology classify creatures into ever-larger sets and sets of sets: individuals (dogs, say), species, genera, families, orders, classes, phyla, taxa, and kingdoms. The classification itself forms an algebraic lattice, with individuals acting as the system's atoms. Comparative anatomists and zoologists bring an exquisitely refined and elaborate intuition to the task of sorting the various biological creatures into appropriate categories: the obvious cases leap to the eye; at the margins of the system, where the whale resides, difficult matters are decided by reference to historical and comparative anatomy, parallel structure, common organization, biological traits, and, often, levels of biological achievement. If the image of a lattice is for the moment taken literally, then each level of the lattice, from the atoms upward, comprises a set or ***ensemble***: of individuals, in the first instance, of sets of individuals, in the second. An ensemble at any distinct level of the lattice, I assume, satisfies equations (9.1)–(9.3), and counts thus as a metric space.

The distance between strings

DNA is a string drawn from a four-letter alphabet; proteins are strings of fixed length composed of 20 amino acids; as such, both strings belong to a wider family of string-like objects: computer programs written in a given language, the sentences of a natural language, formal systems; and acquire by osmosis a distinct conceptual and mathematical structure. It makes little difference whether strings of DNA or strings of amino acids are taken as fundamental; and, in any case, I often alternate between the two. By an ***alphabet*** A I mean a fixed and finite collection of elementary entities called ***words***; by the ***universe of strings*** over a finite alphabet, the set of all finite sequences A^* whose elements lie in A.

The natural distance between words $W = w_1 \ldots w_m$, $V = v_1 \ldots v_n$ ($W, V \in A$) is $|n| + |n| - 2 \times |k|$, where k is the maximum of the length of a word $U = u_1 \ldots u_l$, which is a subword both of W and V. For example, let W = cadbabbd, V = xcaaba. An appropriate U is U = caab; hence $\hat{D}(W,V) = 8 + 6 - 2 \times k$.

Grantham (1974) has proposed a definition of distance in a Euclidean metric space of proteins based on properties of composition, polarity, and volume; but the theory of evolution suggests that changes in biological strings come about through mutations – random flash points at which letters are scrambled. Some strings may change in a large-hearted way, with whole blocks of letters wheeling and shifting like cavalry horses; but the least mechanism to which these operations may be resolved is the simple one of erasure and substitution – deletion and insertion. The elementary processes of evolution at the molecular level lend to the natural metric a certain simple plausibility in the face of fancy competition. $T = A^*$, then, is a ***typographic metric space***; d_T, its natural distance.

Metric spaces in phase

M and M^*, suppose, are two metric spaces; $g: M \to M^*$ assigns to each point p in M a distinct point P^* in M^*. M and M^* are ***in phase*** under g if g acts roughly to preserve distances: for any $\xi > 0$, there exists a $\varphi > 0$, such that for all p and q in M

$$d_M(p, q) < \varphi \to d_{M^*}[g(p), g(q)] < \xi \ .$$

g is thus ***uniformly continuous*** on M; φ is, of course, a function of ξ. It often happens that a particular mapping between metric spaces is especially natural – for reasons that are not mathematical. The English alphabet, for example, makes for two metric spaces: strings of letters, sets of words. Strings of letters are close if they agree in spelling; words if they agree in meaning. Small typographic changes give rise to large differences in meaning: these metric spaces are not in phase. This observation is often regarded as a paradox in the context of theoretical biology. In an important and influential article, King and Wilson recount evidence showing that chimpanzee and human polypeptide sequences are more than

99 percent identical; the *species* appear further apart than a comparative analysis of their polypeptide chains might otherwise suggest.[6]

Complexity

Complexity and simplicity, like Yin and Yang, are metaphysical duals; except for a vagrant connection to intuition, it hardly makes a difference what is called which. Mathematicians and philosophers are interested in complexity for their own ends; so are theoretical biologists, who in their better moments are quite capable of evincing a sense of Heraclitian awe when confronted with the intricacies of the protozoan swim bladder. Simple counting principles often seem as if they might provide a general scheme for the measurement of complexity. Suppose that X is a nonempty set of objects and that A, B, C,... are constructed from the elements of X by certain specified operations – concatenation, for example. Can we not then say that the complexity $C(x)$ of any object is a measure of the number of its distinct elements and the separate and specifiable relations between them? $C(x)$ would be a monotonically increasing function of the square of the number of distinct elements in any given construction. Simple, no? And intuitively satisfying?

Apparently not. Label the parts of an ordinary watch in an obvious alphabetic fashion; and the binary relations between its parts as well. The watch when working, let me suppose, has a complexity measured at C; but so, then, does the watch when not working – when not assembled, in fact, binary relations being free for the asking. Examples of this sort, when extended and made precise, suggest ultimately that any complex object belongs to an embarrassingly large equivalence class of objects precisely equal in point of complexity.

Statistical mechanical complexity

A system of identical particles moving within a fixed, bounded, and finite volume of space constitutes a *configuration*; never having seen the blue smoke from a cigar spontaneously collect in but one corner of a warm room, the thoughtful physicist – pipe, slippers, Beagle-eyes, an air of earnest confusion – concludes that not all configurations are equally probable; yet if there are N configurations $\Pr(N_i) = N_i / N$ – this for each i. This incompatibility between what one sees and what one gets is known as *Boltzmann's paradox*, an unhappy name if only because no real paradox is forthcoming; but an unhappiness nonetheless. Distinct configurations, Boltzmann argued, may be grouped into *states*; what the altogether more elegant Gibbs called *ensembles*. Within thermodynamics – statistical mechanics from above – the entropy S of a system appears perpetually in the ascendancy and tends inexorably to a maximum; statistically, Boltzmann reasoned, S is thus proportional to

$$S = k \log W \; ; \tag{9.4}$$

where k is Boltzmann's constant, and W a measure of those configurations compatible with a given state – *complexions* as they are called in old-fashioned texts. Configurations are alike in point of probability: not so complexions; the probability of finding a mechanical system in a given state is proportional to the number of distinct complexions realizing that state. At equilibrium, the complexions are at a maximum; and so, too, the entropy, which functions as a kind of ectoplasmic measure of *randomness* or *disorder*.

Complexity under a classification

Statistical mechanics has a good point to its credit, and implies a second. Certain states of a physical system may be multiply realized; their number, if counted, makes for a measure of sorts. What is measured within statistical mechanics is plainly not complexity; the description of entropy as disorder serves only to explain the whole business to the baffled undergraduate, with the explanation rapidly withdrawn by the time he enters graduate school. Still, I am struck by the extent to which the mathematical definition of entropy is made possible by an enterprising reorganization of the way in which mechanical systems are classified; in assessing complexity, a concept with a brutish family resemblance to disorder, the classification may well come first.

An example? Of course. I shall pass glowing colored slides about shortly. Consider the set of all functions $f : R^n \rightarrow R$. Those smooth functions whose critical points are nondegenerate are known as *Morse functions* and are at once open, dense, and locally stable in $C^\infty(R^n, R)$. Any Morse function may be expressed in canonical form: if x is a critical point of f, there exists a number k such that in a neighborhood of x, and after a suitable change in coordinates,

$$f(x) = x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_n^2 \ . \qquad (9.5)$$

Such is Morse's lemma. Their mathematical docility suggests that the Morse functions are simple, if anything is; but the Morse functions are simple *because* they are Morse functions, and not Morse functions because they are simple; simplicity is a derivative quality, like color, contingent upon a classification, and unremarked otherwise.

The concept of a degenerate singularity makes for a simple classification on the space of smooth functions $C^\infty(R^n, R)$; but a set of objects may be simple under a classification even if the classification is itself unpleasantly complex. Writing some years ago, Smale asked whether there exists a least Baire set U in the space of all dynamical systems Dyn(M) on a compact manifold M, whose elements might be qualitatively described "by discrete numerical and algebraic invariants".[7] The question as posed admitted of a simple answer: no. What is needed, Smale later concluded, is a sequence of nested subsets U_i [Dyn(M)], where k is relatively small, U_i open, and U_k dense. As i increases, more of Dyn(M) is swallowed; as i decreases, stability and regularity properties come to the fore. It is for U_1 that Axiom A is satisfied, nonwandering sets are finite, and the transversality condition is met. U_1 thus consists of "the simplest, best-behaved, nontrivial class of dynamical systems"; but nothing in Smale's organization of Dyn(M) is simple at all.

A set is *absolutely simple* under a classification if it is at once open, dense, and locally stable; under this definition simplicity does not come in degrees. Often, suitable sets turn out to be merely of the first Baire category, the best one can do; sets that are dense need not be stable, and vice versa. First category sets and sets of measure zero coincide in the case of countable sets; but not beyond.

From the point of view of statistical mechanics, simplicity and complexity are concepts that involve configurations; complexity under a classification is a matter of routine: what is complex is singular, unusual. These notions may be brought into alignment – but only for a certain class of objects. An object A is *dissective* only when it may be decomposed to a finite stock of parts in a finite number of steps. The mammalian eye is a dissective structure; so is the whole of a mouse, a moose, or a mole; but curves and concepts, the real numbers, the coast of Britain, sea-green sea-waves, and, perhaps, the entire bizarre universe of elementary particles, are indissective. A dissective object is thus composed of its parts taken together under a certain distinctive relationship. Say that A is composed of a_1, $a_2, \ldots, a_n$ under R. By a relational alternative to R I mean a single permutation of the parts of A. If A, for example, contains but two parts, a and b, say, under the relationship $R(a, b)$, $R(b, a)$ is a relational alternative to R – the only one in fact. Given R, I denote by R^* the full set of all relational alternatives to R. If A is dissective it is R^* that forms its complexion class: the set of all sets of its parts under all and only their relational alternatives.

An elementary partition of a complexion class splits the class as a whole into equivalence classes; relative to a partition, complexity and simplicity are attributes of equivalence classes, and are judged simply by size. To the extent that $[E_i]$ is larger than $[E_j]$, it is simpler as well; and vice versa. Almost all structures in theoretical biology may be dissected to a finite, although very large, base; in this sense, biological complexity and simplicity have pliant finite measures.

The mammalian eye, for example, is a dissective structure. Its parts (on one level of dissection, at least) are proteins, which are arranged in various delicate and precise ways. I am ignoring, now, any dynamic considerations and thinking instead of the mammalian eye as a static object. The complexion class to the mammalian eye consists of all and only those rearrangements of proteins that comprise relational alternatives to the mammalian eye itself.

What makes an eye distinctively an eye, rather than some assembly of jelly-like proteins, is obviously the fact that it is capable of sight. This invocation of function sounds an unavoidably Aristotelian note; but without some concept such as function or purpose, theoretical biology loses much of its point. Let me partition the relational alternatives to the mammalian eye into equivalence classes on the simple basis of function. In the full complexion class, those structures that are capable of sight fall to one side; and those that are blind and stare sightlessly, fall to the other. Complexity and simplicity appear as matters of relative size: the larger the equivalence class, the simpler the structures. Given the delicacy of the mammalian eye, most of its relational alternatives will be incapable of sight; like the Morse functions, these complexions are simple structures; but again, simple because they are sightless, and not sightless because they are simple.

Complexity in strings

Of the 2^n binary sequences of length n, some, such as

$$0, 0, 0, 0, 0, \ldots \tag{9.6}$$

seem simpler than others,

$$0, 0, 1, 0, 1, \ldots \tag{9.7}$$

for example; yet the most natural probability distribution over the space of n-place binary strings assigns to both the same probability: 2^{-n}. It goes against the grain, mine, at any rate, to reckon (9.6) as likely as (9.7), especially when n is large; but nothing in the sequences themselves indicates obviously the point of distinction.

The goal of science, René Thom has suggested, is to reduce the arbitrariness of description; substitute data for description, and the apothegm gains my assent. A law of nature is data made compact: $F = ma$, said once and for all, the whole of an observed or observable world compressed into just four symbols. A series of observations compactly described is rational; if rational, not random. This curious but compelling chain of deductions prompted Kolmogorov to argue that randomness in binary sequences or strings might be measured by the degree to which such strings admit of a simpler description.[8] In following this line, Kolmogorov took the first step toward severing information theory from its unwholesome connection to the theory of probability. If S is a binary string its length is measured in bits: an n-place binary string is n bits long. By a *simpler description* of S, Kolmogorov meant a string D shorter than S such that D describes S by acting as the input to a fixed computer that generates S. Strings that cannot be compactly described are *complex*, *random*, or *information-rich*; strings that can, are not; of these adjectives, only the second preserves even a vagrant connection between the concept that it connotes and what is being measured. This rather inelegant idea makes plain the felt difference between a string of n 0s, and a mixed string. Sequence (9.6), for example, may be expressed by a program, speaking loosely, whose length is $\log_2 n + C$. If $n = 32$, $\log_2 n = 5$: the relevant instruction is simply to write or compute 0 2^5 times. C measures what little is needed to carry out the instructions; $32 - 5 = 27$, the compactness of the program. The shortest program that computes a mixed sequence such as sequence (9.7), by way of contrast, may well be close to 32 bits in length: to compute the sequence, the computer must first store it precisely.

The details? They have been changing since Kolmogorov first spoke, oracle-like, on the subject in a note published in 1967; like a snake engulfing an egg, the theory of recursive functions is engaged in swallowing algorithmic information theory, a development that I deplore, but accept as inevitable. Consider the set of all n-place binary strings A^* over a binary alphabet A and let TM be a fixed computer – a Turing machine, say; g is a general input–output function on TM mapping strings onto strings. The *complexity* of a string S of length n is the length of the shortest binary string D that generates S under TM by means of g. Whatever the complexity of S, D will plainly be maximally complex, and, hence, entirely random. Otherwise, it would not be the shortest description of S. All finite length strings quite obviously have a finite measure of complexity; and only finitely many distinct strings of the same length have the same finite measure of complexity. Quite surprisingly, the decision problem for complexity is recursively unsolvable; this result follows almost directly from the unsolvability of the halting problem for Turing machines. Like truth, randomness is a property that remains ineluctably resistant to recursive specification.

If all else fails, a binary sequence of length n may be generated by a binary sequence of length n: there are 2^n such algorithms, and $2^{n-1} - 2$ algorithms shorter than this. On any reasonable interpretation of complexity, algorithms within a fixed integer k of n itself must be reckoned random or complex or nearly so. Thus $2^{n-k-l} - 2/2^n$ algorithms have a complexity less than $n - k$; and are hence nonrandom or simple. If $k = 10$, this ratio is roughly 1 in 1000; of 1000 binary sequences of length n, only one can be compressed into a program more than ten bits shorter than itself. Hence:

Theorem 9.1 The set of random sequences of length n in the space A^* of all binary sequences of length n is generic in A^*.

These random sequences are simple under a classification because they are typical, but complex in a stronger and more absolute sense because they are random or information-rich. In this context, genericity is a *finite* measure of size. The number of purely random strings grows exponentially with n, of course. If most binary sequences are random, the appearance of sequence (9.6) prompts a natural stochastic surprise: sequences such as (9.7) are what one expects. The definition of Kolmogorov complexity may be directly extended to recursively enumerable sets; sets of strings especially, and hence languages.

Language-like Systems

When it comes to language, there is syntax and semantics. Phonetics is the province of the specialist; pragmatics remains a pale albino dwarf. To semantics belongs the concept of meaning; to syntax, the concept of a well-formed formula or a grammatical sentence. The reference to logic is happy if only because it highlights the fact that language-like systems go beyond the natural languages. Any language no doubt exists primarily to convey meaning; but meaning in mathematics is a matter of a model – an extrinsic object.

The construction of strings within a language-like system involves concatenating or associating simpler strings; any finite string may be dissected to a finite set of least elements. Going up, concatenation; going down, finite dissection; retrograde motion of this sort suggests that language-like systems on this level be represented algebraically as semigroups. Let A be any nonempty set of objects – words, for example, or letters, or numbers. A has the structure of a semigroup if there exists a mapping $A \times A \rightarrow A$ such that for all a, b, and c in A

$$(a \circ b) \circ c = a \circ (b \circ c) .$$

In English words go over to sentences from left to right; in Hebrew, from right to left; but in any case, one step at a time. Let A be a finite set of words now, with words understood implicitly as the least elements of a natural language; and let A^* be the set of all finite sequences $(a_1, \ldots, a_n)$ whose elements $\{a_1, \ldots, a_n\}$ lie in A. To endow A^* with the structure of a semigroup, it suffices to define an associative mapping $A^* \times A^* \rightarrow A^*$: easy enough. If

$$S_1 = (a_1, \ldots, a_m)$$

and

$$S_2 = (b_1, \dots, b_n) \ ,$$

then

$$S_1 \circ S_2 = (c_1, \dots, c_{m+n}) \ ,$$

where

$$c_i = a_i ; \ c_{m+j} = b_j \ ,$$
$$i = 1, 2, \dots, m; \ j = 1, 2, \dots, n \ .$$

A^* is at once a *free-semigroup* over a finite alphabet and a *universal language*: no sequences are left out.

Almost all language-like systems are large in the sense that they have many distinct strings. Meditating on the matter in the late 1950s, and regularly thereafter, Noam Chomsky argued that every natural language is infinite by virtue of its recursive mechanisms – conjunction and alternation, for example – and, simultaneously, that such mechanisms are recursive by virtue of the fact that every natural language is infinite. Both halves to this argument, taken together, describe a closed circle in space. Whatever the truth, language-like systems, if they are infinite, are countably infinite and no bigger.[9]

Going further toward a definition of a language-like system involves the badlands beyond triviality. Linguistics, the French linguist Maurice Gross once provocatively remarked, admits of but a single class of crucial experiments. Native speakers of a given language are able to determine whether a given sentence is grammatical. Experiments of this sort exist because no language-like system encompasses the whole of a set of strings drawn on a finite alphabet – a curious and interesting *fact*, which the sheer concept of communication might otherwise not suggest. The distinction between grammatical and ungrammatical strings induces a primitive classification on a language-like system; and reflects an even stronger principle of fastidiousness: the vast majority of language-like strings are not grammatical at all and represent syntactic gibberish. The fastidiousness of language-like systems is yet again a fact: it would be easy, if unrewarding, to design an artificial language in which most strings were grammatical. From the point of view of grammar, the strings of a natural language are complex under the classification of strings into grammatical and ungrammatical sets. With the strings arrayed in front of the mind's bleak and rheumy eye, in ascending order, by length, with sets of strings stacked like an inverted pyramid, the grammatical strings in a language-like system appear as nothing more than a thin smudge; they are thus *complex* under this classification because they are singular, unusual. The origins of this bit of natural history are to be discovered, no doubt, in the algorithmic properties of the human brain: in order to store a natural language, the brain must first represent it – in the form of recursive rules, for example. This suggests that language-like systems are low in point of Kolmogorov complexity; and from this point of view, *simple*.

A natural language, I have already observed, realizes two metric spaces (cf. p 240); but the informal example that I gave involved the concept of meaning, and not grammar. No matter: the point carries over to the case at hand – and comprises the third of three queer natural facts that nothing in the concepts of grammar or communication obviously implies. Thus, let T be a typographic metric space of strings under the natural metric; the same set of strings comprises a second metric space under the degenerate distance function d^*: if s and s' are both grammatical, $d^*(s, s') = 0$; if not, $d^*(s, s') = \infty$. These are the ***natural*** and (degenerate) ***grammatical*** metric spaces of a language-like system. In a language-like system, natural and grammatical metric spaces are plainly *not* in phase.

Two models of generation

Linguistics is a rebarbative, hair shirt of a subject; and grammar a vexing property. Linguists, for reasons of their own, are often interested in the weakest of generative devices that specify all and only the sentences of a natural language.

Representation by grammar

A ***phrase structure grammar*** is a quadruple $G = (A, T, S, P)$, where A is some finite alphabet of symbols; T, a distinguished subset of A – the set of so-called terminal symbols; S, a distinguished initial symbol; and P, a finite set of production rules of the form $u \to v$; u is a nonempty set of nonterminal symbols, and v some specified string of characters. The set of all strings of terminal symbols constitutes a ***phrase structure language*** – a proper subset of the set of all strings A^* defined over A.

By a *context-free* production rule, I mean one in which u may occur in any context – in effect, a rule in which u figures in isolation. Correspondingly, there are context-free grammars.

Example 9.1 Let $A = (a, b), T = (a, b)$, and P be the two rules $S \to ab$; $S \to aSb$. This grammar generates all and only the strings of the form $a^n b^n$.

Representation by systems of equations

Consider the context-free grammar G whose production rules are $S \to aSa$, and $S \to c$, where $T = (a, c)$, and S is an initial symbol. Let the variable f_i range over terminal symbols. The action of the production rules may be mimicked by an equation:

$$S = f_1 + f_2 + \cdots + f_n \quad ,$$

where addition is construed as set theoretic union. For G,

$$S = aSa + c \quad .$$

Replacing S by $S^0 = c$,

$$S^{(1)} = aca + c \quad .$$

This process repeated ultimately yields a system of equations

$$S^{(1)} = aca + c$$

$$S^{(2)} = a(aca + c)a + c = a^2ca^2 + aca + a$$

$$\vdots$$

$$S^{(n)} = a^n ca^n + \cdots + aca + c = \sum_{i=0}^{i=n} a^i ca^i \ .$$

At the limit, the solution $s^{(\infty)} = \sum_{i=0}^{i=\infty} a^i ca^i$ is given by a formal power series in noncommutative variables.[10]

A language-like system has *formal support* when each and every string in the system may be described by a single algorithm; only for context-free languages may grammars and systems of equations be balanced against each other. Elsewhere, the situation is darker. There is a sense, however, in which these two representations exhaust the possibilities for the description of structured and infinitary objects; and correspond, in the Metaphysical Large, to the alternatives confronting an imaginary Deity in creating the observable world.

Weak Theories

The vitalist believes that life cannot be explained in terms of physics or chemistry. In the nineteenth century, in Germany and France, at least, his was the dominant voice before Darwin; and natural philosophers, such as Cuivier or von Baer, or Geoffrey St. Hilaire, dismissed mechanism with a kind of troubled confidence that suggests, in retrospect, a combination of assurance and wistfulness. Orthodoxies have subsequently reversed themselves with no real gain in credibility. David Hull, in surveying this issue, concludes that neither mechanism nor vitalism is plausible, given the uninspiring precision with which each position is usually cast.[11] *D'Accord.* To the extent that the refutation of vitalism involves the reduction of biological to physical reasoning, the effort involved appears to me misguided, and reflects a discreditable, almost oriental, desire for the Unity of Opposites. On the standard view of reduction, the sciences collapse downward until they hit physics: *Rez-de-Chausee*; but our intellectual experience *is* divided: mathematics, physics, biology, the social sciences. Each science extends sideways for some time and then simply stops. The ardent empiricist, surveying the contemporary scene, might well incline to scientific polytheism, with mathematics under the influence of an austere Artin-like figure, and biology directed by a God much like Wotan: furious, bluff, subtle, devious, and illiterate.

Still, the philosopher of science is bound to wonder why so many philosophers have remained partial to the reductionist vision, and hence to mechanistic thought in biology. David Armstrong, J.J.C. Smart, Michael Ruse, and even the usually cagey W.v.O. Quine, call on elegance to explain their attachment. Were the sciences irreducibly striated, one set of laws would cover physics, another biology, and still a third, economics and urban affairs, with the whole business resembling nothing so much as a parfait in several lurid and violently clashing colors. This is an aesthetic argument, and none the worse for that, but surely none the better either. If elegance is inadequate as a motive, intellectual anxiety, realized unconsciously, is not.

Vitalism commences from the conviction that nothing in our experience is much like the life that ripples and bubbles so abundantly over an entire planet, and nowhere else, apparently. Now mathematical physics is not only the pre-eminent discipline of our time – it is where the laws are. Evolutionary theories in biology are ***weak*** in the sense that they are not directly sustained by the authority of physics; and, worse, weaker still in being ***counterphysical***. Thermodynamic arguments count against the very existence of the structures that they are meant to explain. Fact heavy, law poor, such theories remain surprisingly resistant to confirmation. Were biology an aspect merely of physics, the sceptic would get short shrift: there, the answer to whether what works, works, is simply that it does.

Science is unavoidably general. To say that copper conducts electricity is weakly to imply the counterfactual conditional that were anything much like copper it would conduct electricity as well. It has often appeared to philosophers of science that specifying what it means for something – an x, say – to be much like copper inevitably comes to claiming that, among other things, x conducts electricity. Still, the similarity in structure between two domains of discourse – computer programs and natural languages, for example – may be obvious on grounds other than the fact that they share the same laws.

When I speak of a theory, I follow the logician's lead: a theory consists of a consistent set of sentences in a given language; the set-theoretic or algebraic structures in which a theory is satisfied comprise its models. Two models that share the same structure are isomorphic and hence elementarily equivalent in the sense that they satisfy the same sets of sentences. What I am after is a weaker notion entirely – partial similarity in structure. I know of no way, unfortunately, to define this concept so that the definition applies equably to biology, and, say, geology; I suspect, in fact, that partial similarity in structure will require a definition with indefinitely many separate clauses. Whatever the details, similarity in structure is bound to be a matter of degree, so that it makes sense of sorts to say of two models that they are at a certain distance, one from the other. In this way a family $\{M_i\}$, $i = 1, 2, \ldots$ of (possibly) first-order models may be given an appropriate and empirical metric structure.

Suppose that T is a theory holding in M; and let M^* be a model at some fixed distance from M. By the ***symmetric difference*** T/T^* of T I mean the number of formulas T^* of T that fail to hold in M^* when T is interpreted in M^*.

A theory T is general, I shall say, if for any $\varepsilon > 0$, there exists a $\delta > 0$ (a function of ε, of course) such that

$$\mathrm{d}(M, M^*) < \delta \rightarrow T/T^* < \varepsilon \tag{9.8}$$

Generality in my sense is a kind of stability; and as Dr Johnson remarks, the soul must ultimately repose in the stability of the truth.

To see an analogy between the operations of life, on the one hand, and the operations of language, on the other, is to raise the question whether the laws of biology have a natural and legitimate interpretation in linguistic terms. I am myself indifferent to the fate of the Darwinian theory, and perfectly prepared to believe, along with Wickramasinghe and the luckless Hoyle, that life originated in outer space, or that the Universe-as-a-Whole is alive and breathing stertorously;

but if Darwinian theories work in life, they should work elsewhere – in language-like systems, I should think. Should they fail there, this may be taken as evidence for the inadequacy of Darwinian theories, or as evidence for the inadequacy of the analogy that prompted the comparison in the first place.

I stress this point if only because it has so often been misunderstood.

Life as a language-like system

It was von Neumann who gave to the idea that life is like language a part of its curious current cachet. The last years of his life he devoted to a vast and clumsy orchestration of cellular automata, showing in a partial fashion that when properly programmed they could, like abstract elephants, reproduce themselves. Some years before, McCulloch and Pitts had constructed a series of neural nets in order to simulate simple reflex action; Kleene demonstrated that their nets had the power of finite automata and were capable of realizing the class of regular events; von Neumann's automata had the full power of Turing machines. Michael Arbib, E.F. Codd, G.T. Herman, A. Lindenmayer, and many others, have carried this work forward, with results that asymptotically approach utter irrelevance.[12] Yet the analogy between living systems and living languages has not lost any of its brassy charm. There is information, of course, which is apparently what the genes store; replication, coding; messages abound in the bacterial cell, with *E. Coli*, in particular, busy as a telephone switchboard. So striking has the appropriation of terminology become, that some biologists now see the processes of life, in all their grandeur, as the effort of a badly protected and vulnerable bit of genetic material to keep talking for all eternity.

Unlike an argument, an analogy stands or falls in point of plausibility; good arguments in favor of bad analogies are infinitely less persuasive than bad arguments in favor of good analogies. Certainly the proteins, to stick with one class of chemicals, may be decomposed to a finite base – the 20 amino acids. The precise, delicate, dance-like steps that are involved in their formation suggest, moreover, that they satisfy some operation as abstract as concatenation. On the other hand, the number of possible proteins, although large, is finite; but one of the joys of analogical reasoning is the vagueness with which the line between success or failure may be drawn.

The grammatical strings of a language-like system are low in Kolmogorov complexity, and so are not random. Such is the fastidiousness of a language-like system. What of the proteins? If they are random, it makes little sense to think of them as biological words or sentences. Jacques Monod, whose metaphysical attitude toward biology suggested nothing so much as a kind of chirpy bleakness, drew attention to the random character of the proteins in *Le hasard et la necessite*; his argument has been gravely accepted by many molecular biologists.[13] In fact, the evidence leading to his conclusion is fragmentary; the standards of randomness to which he appealed, imprecise. Thus it struck Monod that knowing, say, 249 amino acid residues in a chain 250 residues in length, one could yet not predict the last member of the chain; much the same is true for English sentences, of course; it is, in any case, simply untrue that protein strands exhibit such wanton degrees of freedom. Within protein chemistry, there are many instances of what appear to be strong internal regularities: palindromic patterns, for example.

Nonetheless, I am in sympathy with Monod to this extent: it is unlikely that the analogy between life and language will be profitably pursued on the atomistic level of the nucleic acids or the proteins themselves.

PART THREE

Arguments Good and Bad

The theory of evolution is haunted by an image and an observation: the first, that of the hapless chimpanzee, typewriter-bound, endeavoring, quite by chance, to strike off the first twenty lines of Hamlet's soliloquy; the second, the comment of an anonymous Jansenist logician, who remarked, quite sensibly, "that it would be sheer folly to bet even ten coppers against 10000 gold pieces that a child arranging at random a printer's supply of letters would compose the first twenty lines of Virgil's ***Aneid***". Image and observation do not quite cohere into a single argument: it is clear in neither case ***how*** the imagined stochastic experiment is to stop. Still, the notion of randomness yet lies at the center of evolutionary thought, and there it sits, toad-like and croaking. On the simplest and most intuitive conception of probability, what can occur is weighted against the background of what might occur: five diamonds: all other combinations of the cards. In poker, there are 2598960 five-card hands, but only 5148 flushes. It is their ratio that one might expect to observe as cards are actually dealt; but in the longest of long runs, the passage to the limit gives content to the intuitive idea that a number of successive trials will converge to a particular real number: 0.002, for example, if flushes are being counted.

One of the curiosities of the very notion of probability is the inescapability of the improbable. The laws of thermodynamics, to take a notorious example, are anisotropic: they go in one direction; downhill, as it happens, a circumstance with what appears to be overwhelming personal support. Statistical mechanics provides a brilliant and persuasive explanation for thermodynamic laws; yet Poincaré demonstrated, in an absurdly easy proof, that any statistical mechanical configuration, of whatever degree of implausibility – k molecules of gas, for example, occupying $1/V$ of the total volume V of a finite and bounded container – is bound to recur, in all its vividness, poignant symmetry, and complexity, given enough time. Physicists often explain the discrepancy between thermodynamics and statistical mechanics by arguing that the time involved is very long. No doubt.

The evolution of life on this planet is, as Darwin realized, not a hurried affair. Early on, Darwinian biologists got rid of the theological limits set to the age of the Earth by Bishop Ussher and others in the seventeenth century; the scale within which Darwinian evolution might have worked is bounded by perhaps five billion years. Nineteenth century biologists assumed that whatever else one might say about Darwinian biology, it would not fail for lack of time; this thesis twentieth century biologists have carried over intact.

Five billion years is apt to seem long if one is counting the minutes; but it is not long enough to sample on a point by point basis a space whose cardinality is roughly 10^{15} – touching base with a new point at every second, say; and yet there are 20^{250} possible proteins – a number larger by far than the expected life of the universe measured in seconds. In a space of this size, the odds against discovering a specific protein – fishing it from an urn, say – are prohibitive: 1 in 20^{250}.

I spoke hastily just now of a *specific* protein: if *any* protein will do, the odds improve: in a uniform probability space $\{a_i\}$, $\Pr(a_1 \vee a_2 \cdots \vee a_i) = 1$. The distinguished British biologist Peter Medawar has seized upon this point, and commenced happily to trot, but in what I think is the wrong direction.[14] "Biologists," he writes, "in certain moods are apt to say that organisms are madly improbable objects or that evolution is a device for generating high degrees of improbability. I am uneasy about this entire line of thought," he continues, "for the following reason:

> Everyone will concede that in the games of whist or bridge any one particular hand is just as unlikely to turn up as any other. If I pick up and inspect a particular hand and then declare myself utterly amazed that such a hand should have been dealt to me, considering the fantastic odds against it, I should be told by those who have steeped themselves in mathematical reasoning that its probability cannot be measured retrospectively, but only against a prior expectation ... For much the same reason, it seems to me profitless to speak of natural selection's 'generating improbability' ... it is silly to be thunderstruck by the evolution of organ A if we should have been just as thunderstruck by a turn of events that had led to the evolution of B or C instead."

Medawar is roughly right about probability: the fallacy to which he refers is the *error of retrospective specification*; and consists precisely in reading back into an original sample space information revealed only on the realization of a particular event. In poker, a deal distributes n hands of equal probability: 1 in 2598960, as it happens. This sample space is retrospectively specified if one hand in particular is contrasted with the full set of 2598959 hands that remain, and probabilities assigned to the partition so created; what appears initially as one among equiprobable events becomes under retrospective specification an improbable event in a sample space of only two points. It is embarrassing for an author to point such things out. Still, Medawar is wrong in the general conclusions that he draws from this paragraph. Card sharps and statisticians are little interested in the set of all five-card sequences. In poker, sequences are *initially* partitioned into equivalence classes of uneven size: a royal straight flush, of which there are four, a straight flush, four of a kind, a full house, a straight, three of a kind, two pairs, and, then, finally, whatever is left – the vast majority. There are four ways to achieve a royal straight flush; many more ways in which to realize a full house. Since they are specified in advance, partitions in poker carry no taint of retrospection; and plainly, in poker there is only a rough correlation between the internal character of sequences within a partition and their payoffs: what is important here, as elsewhere, is the classification, which is very largely arbitrary.

Medawar's argument, on its face, thus involves rather an uninspiring mistake, but it is not yet a mistake in evolutionary thought. The human eye, a chastened Medawar might argue, turning his back on his own analogy between life and the cards, represents one arrangement of its constituents: any other might have done as well. In admiring the structure that results, we suffer from misplaced awe, like a toad contemplating a dog. Does this argument carry conviction eye-wise? Is it reasonable to suppose that any other arrangement of the eye's constituents would result in an eye? In anything at all? The question sounds an unavoidably Aristotelian note: an eye is an organ with a specific function – sight, most obviously; an eye-like configuration does not count as an eye unless it can see. To frame the discussion thus is to answer the question immediately, at least on the level of intuition; but what I have said must not be confused with an *argument* in refutation.

Viable proteins

Linguistics is possible if only because human beings have strong and reliable intuitions about natural languages. The polypeptides are alien strings, accessible only through an arduous act of the biochemical imagination. Grammar effects a segregation of strings in a language-like system; beyond grammar, aloof, untouchable, there is *meaning*; the two concepts do not coincide. Some grammatical strings, in a natural language, at least, are grammatical and meaningless; others, meaningful but ungrammatical; but meaning and grammar belong together, yoked pairs in the same corner of some dimly understood conceptual space. An algebraic system of strings in which no distinctions of meaning and grammar are recognized is profligate; and pointless because of its profligacy.

In a preanalytic sense, the concept of meaning indicates a kind of coherence; and has a usefulness of application in domains other than language. A life well-spent is meaningful: its parts and patterns are ordered; full with life, biological creatures are filled with meaning, a kind of blunt, irrefrangible purpose; in death, this meaning disappears, and what is left, the corpse and its grim constituents, appears all at once to lose the integrity of the creature itself, and becomes, instead, a thing among other things, an object merely. To the vitalist, living creatures instantiate some unique property that remains stubbornly unseen elsewhere – in the domain of objects studied by mathematical physics, for example; in death, this property vanishes, like a fluid evaporating. In mechanistic thought, the passage from life to death is rather like a phase transition, a singularity of sorts in the trajectory of the organism, a disabling and permanent catastrophe, that reflects, as it must, only a change in the constituents of the organism, a variation in its underlying pattern. The concept of a complexion, which figures in statistical mechanics, provides a useful measure of meaning. The complexion set to a biological organism represents those relational alternatives of its biological parts that correspond to living systems. The unalterable fact that living systems die and hence do not persist indicates that some of their complexions fail to preserve life and hence meaning; in fact, the number of meaningless complexions must be significant: most of the arbitrary rearrangements of a complex organ – a mammal, say – result in nothing more than a botch – a circumstance with which every surgeon is familiar. The Central Dogma of molecular biology establishes a relationship between strings of nucleotides and strings of proteins; to the extent that the whole of a biological organism may be resolved into its protein-like parts, the Central Dogma establishes a larger, more indirect, relationship between molecular biological order and order in the larger sense of life. This relationship has an inverse: if only certain forms of life have meaning, this, too, is reflected, as it must be, in the universe of molecular biological strings – on the level of string *ensembles*, for example. If certain protein ensembles are meaningful, and not others, this suggests, but does not imply, that the same distinction is palpable on the level of the individual proteins themselves. The term *viable* I mean as a biological coordinate to the Siamese concepts of meaning and grammar; a protein is viable only when it achieves a certain minimum level of biological organization and usefulness. What level? What kind of organization? Usefulness in what respect and to what degree? Who knows?

Full loads, fair loads, fair samples

In a natural language, sentences decompose to words; words to letters. Grammatical constraints hold weakly at the level of English words. The set of all word-like combinations of English letters of fixed length n, I shall say, make up a ***full load***; the set of all grammatical words, a ***fair load***. Within molecular biology, a full load corresponds to all possible proteins of normal length: a set whose cardinality is 20^{250}. To the fair loads in English correspond the viable proteins in molecular biology. How large is the biological fair load? Again, who knows? Whatever its ultimate size, those proteins that have already been synthesized in the course of biological history are viable if anything is: nothing succeeds like success. This set is a ***fair sample*** of a fair load. Its size Murray Eden calculates at 20^{52}. The task that he sets himself is the infinitely delicate one of drawing inferences about the fair load from its fair sample.[15]

Between the fair sample of a fair load, and the fair load itself, is the difference between what is and what might be; between the fair load and the full load, the difference between biology and mathematics. In English, the difference between the fair load and the full load is as absolute as death. Any two words of English thus resemble each other more than they are likely to resemble a word generated at random from the letters of the English alphabet. In the case of the polypeptides, Murray Eden writes:

> Two hypotheses suggest themselves. Either functionally useful proteins are very common to this space, so that almost any polypeptide one is likely to find has a useful function to perform, or else the topology appropriate to this protein space is an important feature of the exploration: that is, there exists certain strong regularities for finding paths through this space.

In asking whether the viable proteins are common in the space of all polypeptides, Eden is asking, in effect, whether the fair sample is marked by discernable statistical regularities. "We cannot now discard the first hypothesis," he adds, "but there is certain evidence which seems to be against it: if all polypeptide chains were useful proteins, we would expect that existing proteins would exhibit very different distributions of amino acids." Statistical tests appear to show that pairs of proteins are drawn from a common stock. His example involves the alpha and beta human hemoglobin chain. One form of hemoglobin has 146 amino acid residues, the other 140. The two chains may be set down, side by side, and matched, residue by residue. They agree at 61 points; there are 76 points at which they differ, and 9 points at which no match is possible because the chains are not of the same length. It is plausible that one chain was derived from the other, or that both were derived from a common ancestor. What is curious about these pairs of proteins, however, is the fact that even though the chains do not agree completely in the order of their amino acids, they do agree in their ***distribution***; reason enough, Eden argues, to suppose that the proteins themselves are drawn from a statistically significant fair sample.

The criticism of this historically important argument, I leave as an exercise.

Delicate inferences

In *What is Life?*, Schroedinger argued that living systems must have recourse to what he dubbed an "aperiodic crystal" in order to store information. Crystals are repetitive, regular, and information poor; the order of a living system is specific, irregular, information rich. There is a certain splendid effulgence to the vocabulary of theoretical biology that it would be uncharitable not to cherish. H.P. Yockey identifies order with Kolmogorov complexity; and so does R.M. Thompson, a mathematician who in writing on theoretical biology alternates between information theory and a pious endeavor to communicate to the reader his appreciation for the many faces of Krishna.[16] On the other hand, G.J. Chaitin and R.M. Bennett identify biological order with algorithmic simplicity. A division of intuition on so fundamental a point may suggest a degree of conceptual confusion approaching the schizophrenic.

If biological words are characterized by a high degree of Kolmogorov complexity, could time and chance have combined to discover a structure comparable, say, to cytochrome c or any of the modern hemoglobin chains? This is the question raised by the redoubtable H.P. Yockey: the problem as posed has but two parameters.[17] In the beginning the primeval soup, which I always imagine as rather a viscous, Borscht-like fluid, contained perhaps 10^{44} amino acid molecules. There is, inevitably, an element of fantasy to all quantitative calculations of this sort. At each second, over the course of 1×10^9 years, an indefatigable stochastic Deity arranges and then rearranges the 10^{44} amino acid residues in sequences whose length $N = 101$. There are

$$20^{101} \text{ or } 2.535 \times 10^{131} \tag{9.9}$$

such sequences. The odds against discovering any one in particular thus stand at 1 in 2.535×10^{131}. Not all residues, however, are equally probable. Save for a very large set of strings of small probability, the number of sequences of length N is

$$a^{NH} , \tag{9.10}$$

where

$$H = - \sum_{j=1}^{n} p_j \log_a p_j \ . \tag{9.11}$$

Here p_j measures the probability of the jth residue, and $a = 2$, so that H is measured in bits.

In the end – the details are not important to my argument – Yockey concludes that

$$H = 4.153 \text{ bits/residue} \ ; \tag{9.12}$$

the number of 101 place sequences is

$$2^{4.153 \times 101} = 1.8067 \times 10^{126} \ . \tag{9.13}$$

"Information theory," he remarks, "shows that, in this case, the actual number of sequences is smaller than the total possible number by a factor of 10^5". Now there are, in all, 3.8×10^{61} families of cytochrome c sequences; in order to obtain any one of them by chance, Yockey argues, it would be necessary to repeat an elementary stochastic experiment 3.15×10^{58} times on 10^8 separate planets "in order to have a reasonable expectation of selecting at least once a member of the ensemble of 3.8×10^{61} cytochrome c sequences in only ten of them".

From nothing, nothing, the Darwinian doubters have always claimed; and I have been there with the best of them; but *this* argument, couched as it is largely within the algorithmic theory of complexity stands on what seems to me dubious ground:

(1) A binary string is random to the extent that its shortest program is roughly of the same length as the string itself; this definition trades only in counting bits. Now, the impulse to assert that contemporary proteins are random owes much, I think, to the rather primitive idea that life, if complex, requires complex constituents or atoms; I have suggested something similar in arguing that the proteins inherit a grammatical distinction from the structures that they constitute. Kolmogorov complexity, however, is ill-defined on any level of biological organization past the molecular; but even if a mammal or a mollusk could be represented as a binary string, nothing suggests that those strings would be high in Kolmogorov complexity. Quite the contrary. Life in the large, on the level of the organism itself, is organized with what appears to be brisk algorithmic efficiency. Living creatures are simple in the sense of Kolmogorov complexity; but complex under the classification of their complexions. In this sense, they behave much as a language-like system. This observation is compatible with the thesis that protein strings are, nonetheless, high in Kolmogorov complexity; but it is compatible, too, with the contrary thesis that protein strings reflect the complexity of life by means of their *organization* and not their complexity. Nothing in the concept of Kolmogorov complexity measures the algorithmic organization of a string or set of strings; two equally complex strings may well differ in their *time complexity* to the extent that only one is polynomially bounded.

(2) The difference between the space of available proteins, and the small subset actually chosen by evolution, makes for a trite contrast; yet what lends to cytochrome c its position of statistical distinction? "Because of the very fundamental function of the cytochromes," Yockey writes, " ... the histones and other proteins, which are believed to be of very ancient and even precellular origin, one cannot relax the *specificity* requirement derived from cytochrome c" [emphasis added]. In generous conversation, Yockey has amplified this point by suggesting that the specific protein chains necessary for life correspond to the set of words in a language – fair and not full loads; a curious remark inasmuch as words in a natural language are low, and not high, in Kolmogorov complexity. Still, I am sympathetic to the drift of this line; but the difficulty goes beyond the problems of an imperfect analogy. Certain classes of proteins, Yockey argues, are necessary for life. Such are the information-rich, complex strands; other strands are specific in the limited sense that they are statistically unlikely: "only a tiny fraction of the (available) sequences will carry specificity." It follows by **Theorem 9.1** (p 245) that specificity and complexity are not the same thing: the *set* of complex strands (of a given length) is in the majority; their emergence is probabilistically favorable, indeed, unavoidable. Cytochrome c, considered simply as a complex protein, is no more likely to appear than any other complex protein; but no less likely either. Having discovered cytochrome c, quite by chance,

Life might have made do with any other protein of comparable complexity. If by specificity, Yockey means statistical unlikelihood in a uniform sample space – the space of all complex proteins, for example – his surprise at the emergence of cytochrome c is attributable to retrospective specification; if not, what then is specificity, the mysterious middle term to his argument? If the specific proteins have some independent description, Yockey does not provide it; and their size, apart from suggesting that it is low, he does not calculate.

PART FOUR

Der Prozess

The evidence in favor of the thesis that proteins are random sequences of amino acids is exiguous; and random words may well be grouped into nonrandom sequences. This suggests that the close study of the statistical properties of certain proteins may involve a kind of dense conceptual myopia, something that reflects a passionate absorption in minutiae. The *process* by which evolution in strings takes place, on the other hand, is macroscopic and global, an energetic probabilistic swarming over sample spaces that are never specified by means of mechanisms that are never clarified.

Biological paths

Life loiters over two metric spaces. The first is alphabetic; the second, zoological. Evolution comprises a drama in the large, at the zoological level; but the Central Dogma requires that any change in the large be mirrored by an alphabetic change, and so the process is doubled as it is divided. To talk blithely of evolution in strings is to assume the completion of the two first steps in biological evolution: the emergence of life-like systems from inorganic matter; and the adventitious creation of the modern biological system of replication and genetic information. An explanation of these steps I cede to the forces of the Night: my more limited concern is with evolution as a process that takes place once the genetic machinery is throbbing moistly. In evolution at the molecular level, one amino acid is dropped from a protein string, another is inserted: ***make way!, move over!, get out!, get lost!***, to cast the operations in easily understood terms; even if the process is more complicated, it may mathematically be resolved into discrete and finite steps. Whatever the details, proteins change over time; and the changes leading to their creation may be regarded as a *path* $P = p_1, p_2, \ldots, p_n$ or protein *sequence*. Suppose that A comprises the full stock of 20 amino acids; $A/$, the set of all *words* of amino acids precisely 250 points in length; and A^*, the set of all finite sequences drawn over $A/$. I assume – an *assumption* note! – that A^* has the structure of a language-like system under the binary and associative operation of protein *concatenation*, where concatenation has precisely its usual linguistic meaning.

Stochastic processes

Let S be a system and X the set of its states or configurations. State transitions are represented by a transformation T: $X \to X$, an artifice expressing the action of the system's laws of evolution. If $T_{s+t} = T_s T_t$, $[T_t \in R]$ is a *flow*, or *group action* of R on X.

On the Darwinian theory, evolution is at its secret heart stochastic; it is natural, therefore, to specialize the concept of a process to the case in which X is a measure space, T a measure preserving transformation. This is the domain chiefly of ergodic theory. Its underlying, indeed, fundamental, object is a probability space (X, B, u), where X is a set of states, B a σ-algebra of measurable subsets of X, and u a countably additive nonnegative set function on B. $u(X)$ is, of course, 1. Let T be an invertible injection from X onto X; if $u(T^{-1}E) = u(E)$ for all E in B, T is a *measure-preserving transformation*; the system (X, B, u, T), a *basic probability space*.[18]

By the *orbit* of a measure-preserving transformation T, I mean the extended history of a single point x under T from the infinite past to the infinite future: a trajectory from void to void. Artificially truncated at x, the system is in an initial state or condition. A real valued function $f: X \to R$, whose values correspond to $f(x)$, $f(Tx)$, $f(T^2x)$, ... acts to measure a system along its orbit; the class of such measurements is defined only to the extent that f is itself measurable:

$$\bar{f}(x) = \lim_{N \to \infty} 1/N \sum_{k=0}^{N-1} f(T^k x) \tag{9.14}$$

is thus the *time mean* of the system;

$$\int_X f \, du \tag{9.15}$$

its *space mean*: systems in which the two coincide for every measurable function are *ergodic*.

Example 9.2 Let A be an alphabet of n symbols $a_1, a_2, \ldots, a_n$, with probabilities $p_1, p_2, \ldots, p_n$, such that $p_i > 0$, and $\sum p_i = 1$. The product space n^z consists of the set of all two-sided sequences in n; the various probabilities assigned to each sequence induce a measure u on n^z. The *shift transformation* $(fx)_n = x_{n+1}$ is measure preserving; the system that results is a finite-valued stationary stochastic process with identically distributed terms.

Example 9.3 Let $M = (a_{ij})$ be an $n \times n$ stochastic matrix. Let $p = (p_1, \ldots, p_n)$ be a row probability vector fixed by M:

$$pM = p \ .$$

Keep the product space and shift transformation from **Example 9.2**; U_M may be extended to a countably additive measure on the algebra generated by cylinder sets; by the Caratheodory–Hopf theorem, U_M thus forms a measure on the Borel Fields of n^z.

Example 9.2 models, say, a doubly infinite series of coin flips, each with probability of one-half; **Example 9.3**, a regular Markov chain, where p measures the a

priori probability of each symbol, M, the transition probabilities from one symbol to another.

Entropy

Consider a *source* consisting of a finite alphabet A and an associated string of symbols, ... $x_0 x_1 x_2$ ... , where each x_i is an element of A. Symbols appear in sequence with a fixed probability p_i; if the probabilities are independent, the *average entropy per symbol* is

$$H = -\sum_{i=1}^{n} p_i \log_2 p_i \quad . \tag{9.16}$$

H is at its maximum if each $p_i = 1/n$. In general, the probability that a particular symbol appears in sequence may depend on symbols that have gone before. This is true if the source is a finite-state Markov device. Let A^I be the ensemble of all doubly infinite sequences drawn on A; the cross section on A^I of sequences that coincide at a finite number of points $a_i = x_{t_i}$, where t_i represents any set of integers, is a *cylinder set*. Now if A contains k letters, the number of n-term sequences over A is k^n; and each sequence is a cylinder in the larger space A^I. It is the cylinder that has a fixed probability $\Pr(C)$: the set of all n-term sequences represents a finite probability space, k^n points in size. The *average amount of information per symbol* sent out by a source of this sort is

$$H_k = -1/k \sum_{C \in C_k} \Pr(C) \log_2 \Pr(C) \quad ; \tag{9.17}$$

the entropy of the source itself

$$H = \lim_{k \to \infty} -1/k \sum_{C \in C_k} \Pr(C) \log_2 \Pr(C) \quad . \tag{9.18}$$

The concept of a source may be specialized to the case of a measure-preserving system under ergodic constraints.[19]

The Shannon–Macmillan theorem

A source puts out sequences; at any given time, there will be only finitely many – A^n in fact, if A is a finite alphabet, and n is the length of each sequence. Finite length sequences are cylinders in the infinite probability space determined by the source; they inherit a probability structure. If n is sufficiently large, there exists an arbitrarily small ξ and $\delta > 0$ such that the n-term sequences may be separated into two groups. For the first

$$|\log_2 \Pr(C)/n + H| < \xi \quad ;$$

for the second,

$$\sum \Pr(C_i) < \delta \quad .$$

This is Shannon's theorem, a result in mathematics that appears to add an author in regular periods. In any case, sequences of the first group are characterized by the fact that $(1/n) \log_2 \Pr(C)$ is arbitrarily close to $-H$. The probability of any such sequence C_i is thus 2^{-nH}; the number of such sequences is 2^{nH}, and comprises a very small share of the total number $a^n = 2^{n \log_2 a}$ of available sequences: a happy result. In coding a channel of communication, attention need be directed only to a tiny sample of the output.

The Stochastic Structure of an Evolutionary Source

In considering evolution as a stochastic process, the object of study becomes biological paths; and not biological words – stray proteins, say, or bits of nucleic acids. The full set of paths in evolutionary space comprises an infinitely large set of strings, if only because evolution appears unbounded as a natural language. The sheer stress on the notion of randomness in popular accounts of evolutionary thought suggests at first that something like a pure Bernoulli process may underlie the whole business, an extended coin flip by means of a coin with 20^{250} separate faces. This is obviously absurd. Evolution is a process by which an ensemble of strings changes over time. Each string is composed of points – amino acids, in fact; the probability that any particular point will change is arbitrary but low; there is little likelihood that *all* points in a string will change simultaneously. Transition probabilities in a neighborhood N of a set of proteins E are thus *concentrated* in that neighborbood.

If an ensemble of proteins occupies a certain finite set of states A_i, its evolution comprises a finite state *Markov process* – a stochastic source satisfying the hypothesis of the Shannon–Macmillan theorem.

Trapping problems

The entropy of a source is a measure of its stochastic character: H at its maximum represents a high degree of uncertainty; all messages are equally probable. The hypothesis of the neo-Darwinian theory is that evolutionary sources are largely random. What this means is, in fact, not entirely clear; but it surely implies that H is relatively large. Let H_{Max} thus be the *imagined entropy* of an evolutionary source. "If the process of manufacturing messages", Chomsky and Miller remark, "were completely random, the product would bear little resemblance to actual utterances in a natural language".[20] Going backward, if the utterances of a natural language are regular, their source is not random. To the extent that a fair sample of evolutionary paths is regular, the fair load is regular as well. A source *specifically designed* to generate the fair load of protein paths has thus an entropy H

$$H < H_{\text{Max}} \tag{9.19}$$

In itself, this is neither controversial nor surprising; if the degree of protein regularity is small, the difference between H and H_{Max} is negligible; if large, an evolutionary source *over-generates*. The real issue is a matter of degree, a question of finesse. Linguistics, of course, suggests that if H is very much lower than H_{Max}, over-generation becomes inordinate; a stochastic source cannot, in general, converge on any *natural* language whose complexity is beyond the recursive capacity of finite-state automata; but while life may be a language-like system, it is not necessarily like a natural language, Chinese, say, or even Esperanto.

Hemoglobin chains

Statistical entropy is a measure of uncertainty; and a measure, too, of the number of alternative messages – my use of the word is metaphoric – that a stochastic source may generate. H at a relative maximum indicates that a source may send out multiple messages, a kind of energetic babble; at a relative minimum, H is ***constrained*** – by the rules of grammar, for example, or the laws of logic. In the case of life, Murray Eden observes, path lengths between proteins are most obviously limited by time, and evolution must be achieved within bounds set by the number of generations in the history of an organism. Meandering paths between proteins are temporally inaccessible. The alpha and beta hemoglobins, Eden argues, were derived by a process of evolution, one from the other; a path between the two sequences must thus exist. Eden calculates that this path at its shortest requires something like 120 separate steps, where each step involves a specific point mutation.[21] The population size of hemoglobin proteins – the fair sample – is, he estimates, 10^6; the rate of mutation 10^{-6}. Each step in this path corresponds to a positive gain in fitness: movement upward along a local gradient of relative perfection. In a conservative sense, it would take roughly 2700000 generations to convert a population of 10^6 alpha hemoglobin chains to a population of beta hemoglobin chains. So far, so good.

If certain paths, of whatever length, are inaccessible to life, a stochastic source is *occluded*. This is certainly what the hypothesis that life is a language-like system implies; it is implied, too, by the fact that contemporary hemoglobin chains exhibit relatively little ***variance***: certain possible paths are deficiently viable, ungrammatical in a sense. Nature, in passing from one chain to another, has evidently rather a small target in mind.

2700000 generations for the evolution of a protein is short: twice that number is long. Let k be a point midway between these numbers. If HG comprises an initial set of 10^6 hemoglobin chains, P_{HG} is the full load of protein paths $[P_{ij}]$, whose initial terms P_{1j} all lie within HG. The number of ***targeted*** protein paths in P_{HG} is small: so much to do, so much to see. The number of targeted protein paths that reach their target within k generations is vastly smaller: so much to do, so little time. If the entropy of a stochastic source is great, the untargeted meandering paths are apt to be favored; by contraposition, this implies that the source entropy for an evolutionary system is rather low; constrained, in fact, by the choice of time and targets; but what expresses these constraints, the Darwinian theory does not say.

Weizenbaum Theory [22]

It is a peculiarity of molecular biological strings that, like the elements of a natural language, they realize two spaces. These are spaces with distinct and ***different*** metrics: there is no reason to suppose that they are in phase. Evolution as a process works most directly on biological organisms, which must perish or persevere in the face of circumstance. To the extent that evolution is a process by which organisms converge over time to some local (or global) optimal, the processes of convergence that are sketched broadly in life must have some substantial echo at the molecular biological level, where words and strings hold sway. The relationship between metric spaces that this pattern exemplifies is quite general – the province, in fact, of Weizenbaum theory. Thus let M and N be two metric spaces, each with its own natural metric; points in M are labeled $t_1, t_2, \ldots, t_n$; points in N, $e_1, e_2, \ldots, e_n$; $f: M \to N$ is a mapping between points in M and points in N – a bijection, to make matters trivially simple. M and N are arbitrary, and admit of obvious specification:

(1) M is a typographic metric space; N, the space of biological organisms (see p 240).

(2) M is a typographic metric space under the natural metric on words; N, the same space under distance defined in terms of meaning or grammar (see pp 245–248).

(3) M is a typographic metric space; N, a space of algorithms.

Thus f might map linear sequences of DNA or proteins, or sets of such sequences, onto organisms, or sets of organisms; equally, f might map a linear string of letters onto a sentence, with a fixed meaning in a natural language; or onto an algorithm in a given computer language such as Algol; then, too, f might map fixed strings in an assembly language onto a computer program. In each of these cases, f does not preserve metrics; M and N are not necessarily in phase.

In addition to the natural metric on M, there exists an *induced metric* $d_{N(i)}$ on M defined by the following relationship:

$$d_{N(i)}(f^{-1}e, f^{-1}e') = d_N(e, e') \ . \tag{9.20}$$

The Weizenbaum experiment

To specify a Weizenbaum experiment, it is necessary to provide M with a *probability transition system* Pr determining for each point t in M the probability that t will change to t'; and an *initial probability distribution* Pr_0. A *distinguished element* $e^* \in N$ is fixed from the first. Within the context of molecular biology, transition probabilities are focused on relatively nearby strings – this because point mutations result in string-like changes of a short typographic distance. In a biological Weizenbaum experiment, this fact is respected to the extent that the typographic metric space and the probability transition system are mutually in accord: probabilities follow typographic neighborhoods. Elsewhere, probabilities and distances are adjusted accordingly.

A point t_0 is selected in accordance with the initial probability distribution Pr_0 over M. The distance $d_{N(0)}$ from $f(t_0)$ to e^* is measured; the system engaged for $i = 1, 2, 3, \ldots$; as t_{i-1} moves to t_i, the distance $d_{N(i)}$ between $f(t_i)$ and e^* is recorded. The outcome of the Weizenbaum experiment is the sequence

$$d_{N(0)}, d_{N(1)}, \ldots, d_{N(n)} \ .$$

The Weizenbaum experiment is *successful* if:

Condition W For $d_{N(0)}$ at an *average distance* from e^* the sequence $\{d_{N(i)}\}$ converges to a neighborhood of 0.

Condition W, when met, implies that $\{d_{N(i)}\}$ is both stable and oriented. The graph of a sequence of points constitutes a trajectory; the set of trajectories in N that are at once stable and oriented is of measure zero. A successful Weizenbaum experiment thus establishes that $\Pr(M)$ *cannot* be arbitrary with respect to its *induced* metric structure. In particular, points that are far in the *induced* metric have small transition probabilities: those probabilities that count must be concentrated on nearby objects – nearby in the sense of the induced metric. On the other hand, transition probabilities over molecular biological strings are, on the neo-Darwinian theory, focused on neighborhoods that are nearby in a natural metric.

It is perhaps for this reason that, with the exception of life itself, no one has ever seen a successful Weizenbaum experiment.

Eigenvalues of natural selection

In Darwinian thought, the effects of randomness are played off against what biologists call the *constructive* effects of natural selection, a mechanism that philosophers have long regarded with sullen suspicion. Wishing to know why a species that represents nothing more than a persistent snore throughout the long night of evolution should suddenly (or slowly) develop a novel characteristic, the philosopher will learn from the definition of natural selection only that those characteristics that are relatively fit are relatively fit in virtue of the fact that they have survived, and that those characteristics that have survived have survived in virtue of the fact that they are relatively fit. This is not an intellectual exercise calculated to inspire confidence.

Natural selection is a force-like concept; and, as such, acts locally if it acts at all. Mathematicians often assume that evolution proceeds over a multidimensional fitness surface, something that resembles a series of hills and valleys; a great deal that is theoretically unacceptable is often hidden in a description of its topology. But I am anticipating my own argument. In speaking of locality, I mean to evoke the physicists's unhappiness at action at a distance. Strings that are far apart should be weak in mutual influence; this is a spatial constraint. Then again, no string should be influenced by a string that does not yet exist. This is a temporal constraint, a rule against *deferred success*. The historical development of a complex organ such as the mammalian ear involved obviously a very long sequence of precise historical changes. Comparative anatomy suggests that the reptilian jaw actually migrated earward in the course of evolution. It is very difficult to understand why each of a series of partial changes in the anatomy of the reptilian jaw should have resulted in a net increase in fitness *before* the advent of the mammalian ear. Certain genes within the bacterial cell, to take another example, "are organized into larger units under the control of an operator, with the genes linearly arranged in the order in which the enzymes to which they give rise are utilized in a particular metabolic pathway".[23] The genetic steps required to organize an operon cluster do not "confer any selective advantage to the phenotype so that individual steps are independent".[23] The rule against deferred success functions as a prophylactic against the emergence of teleological or Aristotelian thought in theoretical biology.[23]

I have pictured evolution on the molecular level as a process involving paths; natural selection acts to induce a statistical drift on some paths, and not others; those paths involving a positive gain in fitness are favored. At any particular time, at any particular place, one has an ensemble E of protein strings, embedded, so to speak, in an underlying probabilistic structure, a measure-preserving system, to keep to the concepts already introduced. To this structure, natural selection is grafted, and acts, presumably in virtue of a property that may be represented by the action of a real-valued, measure-theoretic function: thus $f(x)$, $f(Tx)$, $f(T^2x)$, ... are successive local calculations of fitness under the action of the system's transformation, the ***eigenvalues*** to the system. Suppose now we consider a finite-state system consisting of an alphabet of 26 letters; and the set of sequences k places in length. There are, of course, 26^k such sequences. Each letter $a_i \in A$ occurs with a fixed and independent probability p_i. The shift transformation moves a given string one place to the left. In effect, this system is simply the finite-state stationary process with identically distributed terms mentioned in the example already discussed; and may be represented as a linear array of k squares. An initial probability distribution fixes the configuration of the system for the first (integral) moment; at each subsequent step, every square changes: the odds in favor of any particular letter appearing are 1/26. If doubly infinite in extent, this system models the play of k 26-sided dice continued from the indefinite past to the indefinite future.

What are the chances, one might ask (with a marked lack of breathlessness in my own case), that a system of this sort – a pure Bernoulli process – could converge on a particular sentence of English? Following Mannfred Eigen, let us suppose that the sentence in question is TAKE ADVANTAGE OF MISTAKES, so that k is 23; this is the ***target sentence*** $- S$.

Even here, poised between irrelevance and imprecision, delicate and important biological questions arise.[24] Thus, while it makes sense of sorts to say that for every string, there exists a target – there would be many target sentences – it makes far less sense to say, as Eigen does, that there exists a target for every string – just one, in fact. Fixed in advance, a target so singular would seem suspiciously like a goal and hence ***streng verboten*** in evolutionary thought. How might such a target be represented and by what means might its influence be transmitted to strings? These are not trivial questions.

In any event, nothing in Eigen's own example quite indicates why a stochastic system with a target sentence, however defined, should stop when it has reached its goal. This, however, is a trivial defect, easily made good by the construction of an ***evaluation measure***. Suppose, for the sake of simplicity, that fitness involves only a mapping from strings to 0 and 1: at S, $f(S) = 1$, elsewhere, f is 0. An evaluation measure serves to size up strings in point of fitness as they appear: at S, where $f(S) = 1$, it orders the system simply to stop; at all other strings, the command is to mush on.

Stochastic device, target sentence, fitness function, and evaluation measure, taken as a quartet, comprise an ***Eigen system***. The enterprising Professor William R. Bennett Jr has calculated that an Eigen system would require a virtually infinite amount of time to reach even a simple target sentence – a number roughly a trillion times greater than the life of the universe In the same spirit, Murray

Eden has figured that life would require something like 10^{13} blubbery tons of *E. Coli* "if one expected to find a single ordered gene pair in 5 billion years". The trouble is not simply one of finding the right letters: it is also the problem of not losing them once they are found.

What more, then, is needed? The opportunity, Stephen Jay Gould remarks, for the system to capitalize on its ***partial successes***. Curiously enough, this is Eigen's answer as well, a bizarre example of independent origin and convergent confusion. As Eigen works though his example, his system is designed to retain those random changes that fit the target sentence. Looking at the record of Eigen's own simulation, we see that quite by chance the letter A appears in the first generation in the right place on the sequence. It stays intact, A'ish so to speak, for the rest of the simulation. When an E pops up, it, too, gets glued to the system.

The result is an ***advanced Eigen system***, and an improvement over the hopelessly slow Eigen system already described. Under the advanced Eigen system, fitness is no longer an all or nothing affair; f thus takes values, let us say, between 0 and 1. Scanning every new string, the evaluation measure selects those strings s_i such that $f(s_i) > f(s_{i-1})$. These the system retains until it finds a string superior in point of fitness. The result is a sequence of strings the ***ascends in fitness***. At S, as before, the system stops.

An advanced Eigen system may well reach a target system in rather a short time: unfortunately, in theoretical biology, as elsewhere, the question is not whether but ***how***. To the extent that fitness is purely a local property, it is difficult to understand why ***every*** ascending sequence should necessarily converge to a neighborhood of 1, and hence indirectly toward S. A string that only partially conforms to S is locally no fitter than a string that remains resolutely unlike S. On the other hand, if each of the ascending sequences converge to S it is very hard to see that fitness is a local property, and hard thus to understand what it is that an evaluation measure manages to measure. What is unacceptable is the obvious and tantalizing idea that an evaluation measure judges fitness by calculating the ***distance*** between random strings and a target sentence: distance is not a local property; an evaluation measure so constructed would plainly be responding to signals sent from the Beyond, a clear case of action at a distance. The problem of discovering a target sentence remains unchanged, hopeless. In fact, this is precisely what the advanced Eigen system actually measures, since an arbitrary sentence in which A appears in the second position is judged fit only because it is closer to the target sentence than it might otherwise be. When the matter is carefully explained, theoretical biologists understand at once that the very concept of a target sentence constitutes a beery and uninvited guest in evolutionary thought. I have taken the argument a step further by insisting that evaluation measures themselves be purely local.

Need I insist that the situation is made no better if instead of a specific target sentence I talk of systems set for success when they reach any sentence whatsover? I suppose, since it may at first appear easier to design a system that by randomly changing letters, in what Eigen hopefully calls the evolution game, approximates an arbitrary English sentence instead of just one. The illusion of ease is ill-gotten, of course: a target sentence is a minor stand-in for a major concept. If no particular target sentence is fixed in advance, then any sentence of English, once reached, makes for success. Simply to stop, the system must have an abstract characterization of all the English sentences. Of these, there are infinitely many. A system bouncing briskly from one set of random permutations to another, no less than the linguist or logician, thus requires nothing less than a grammar of the English language if it is not to keep babbling forever.

I have described grammars in terms of the notion of formal support; these concepts receive no definition in Darwinian theory.

Notes

[1] See my review of Michael Ruse's *The Philosophy of Biology*, in *Philosophy of Science* 41, Number 4, December 1974, for a discussion of Smart's position, and related issues.

[2] See, for example, Thomas Kuhn (1970) *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press), a book which has prompted a vast secondary literature and much merited soul-searching among analytic philosophers of science.

[3] "Those who are oppressed by their own reputations," Dr Johnson remarks, "will perhaps not be comforted by hearing that their cares are unnecessary." Francis Crick (1966) *Of Molecules and Men* (Seattle: University of Washington Press).

[4] I discuss reductionism from the perspective of atomistic theories in Berlinski (forthcoming) *The Rise of Differential Topology* (Boston: Birkhaeuser Boston). See also Kenneth Schaffner (1967) Approaches to reduction. *Philosophy of Science* 34 (1): 137–47.

[5] Michael Ruse has argued for his thoroughly incoherent position in Ruse (1973) *The Philosohy of Biology* (London: Hutchinson). The concept of evolution was, of course, in the European air for at least a century before Darwin wrote. European biologists are yet unreconciled to Darwin. In this regard, see Pierre Grasse (1977) *Evolution of Living Organisms* (New York: Academic Press). The facts of molecular biology, it is worth stressing, are not in dispute: it is their interpretation that remains clouded. The central role of DNA, in particular, has troubled many thoughtful observers. "To attribute such powers to a single substance", Grasse remarks "however complicated its molecular structure, is in my view aberrant."

[6] M.C. King and A.C. Wilson (1975) Evolution at two levels in humans and chimpanzees. *Science* 88 (4184).

[7] S. Smale (1980) *The Mathematics of Time* (New York: Springer).

[8] A. Kolmogorov (1967) Logical basis for information theory and probability theory. *IEEE Transactions on Information Theory* IT - 14 (5). I have patterned my discussion on: G.J. Chaitin (1974) Information-theoretic computational complexity. *IEEE Transactions on Information Theory*, IT - 20 (1). The interested reader should consult Chaitin's other papers, and relevant papers by Solovay. Chaitin's bibliography may be consulted for details.

[9] See, for example, Noam Chomsky (1972) *Language and Mind* (New York: Harcourt, Brace Javonovich).

[10] The idea of representing context-free languages by means of a system of equations in noncommutative variables is due to M.P. Schutzenberger. See M. Gross (1972) *Mathematical Models in Linguistics* (New Jersey: Prentice-Hall) for details. I am inclined to think that the Deity, in creating the observable world, hesitated between programming or painting the whole business. As a programmer, he would have chosen a set of recursive rules; as a painter, a system of simultaneous equations.

[11] David Hull (1974) *Philosophy of Biological Sciences* (New Jersey: Prentice-Hall), although staid, contains a competent discussion of many of these issues.

[12] See, for example, L. Lofgren (1975) On the formalizability of learning and evolution, in Suppes, Henkin, Joja, and Mosil (Eds) ***Logic, Methodology and Philosophy of Science*** (Amsterdam: North-Holland).

[13] J. Monod (1971) ***Chance and Necessity*** (New York: Alfred Knopf).

[14] See Peter Medawar (1977) ***The Life Sciences*** (London: Wildwood House).

[15] Murray Eden (1967) Inadequacies of neo-Darwinian evolution as a scientific theory, in P. Moorhead and M. Kaplan (Eds) ***Mathematical Challenges to Neo-Darwinism*** (Philadelphia: The Wistar Institute Press).

[16] R.M. Thompson (1981) ***Mechanistic and Non-Mechanistic Science*** (Lynbrook, New York: Bala Books).

[17] H.P. Yockey (1977) A calculation of the probability of spontaneous biogenesis by information theory. ***Journal of Theoretical Biology*** 67.

[18] See K. Petersen (1983) ***Ergodic Theory*** (Cambridge: Cambridge University Press) for details.

[19] My discussion follows that of A.I. Khinchine (1957) ***Mathematical Foundations of Information Theory*** (New York: Dover Publications).

[20] N. Chomsky and G. Miller (1963) Finitary models of language use, in Luce, Bush, and Galanter (Eds) ***Handbook of Mathematical Psychology*** (New York: John Wiley & Sons).

[21] Eden, *op. cit.*

[22] The idea of the Weizenbaum experiment is due to M.P. Schutzenberger.

[23] Eden, *op. cit.*

[24] See, for example, Eigen (1971) Self-organization of matter and the evolution of biological macromolecules. ***Die Naturwissenschaften*** 10. Together with Ruth Winker, Eigen has recently (1981) published a popular account of his thought under the title ***The Laws of the Game*** (New York: Harper & Row).

References

Grantham, R. (1974) ***Science*, 185**, 62.

Schroedinger, E. (1945) ***What is Life?*** (New York: Macmillan).

Smart, J.J.C. (1963) ***Philosophy and Scientific Realism*** (London: Routledge and Kegan Paul).

Watson, J.D. (1965) ***Molecular Biology of the Gene*** (New York: Benjamin) p 67.