Chapter

15

What Is Life?

15.1 Defining Life

Definitions explain the meaning of a term by relating the defined term to other expressions in the language. For example, a definition of *acid* specifies the necessary and sufficient conditions that all, and only, acids share. More generally, definitions relate items in a language to other items in that language. Some of these other terms, in turn, may have their meanings explained through definitions. But at some point the chain of definitions must end. Some concepts must be understood without the help of other verbal formulae. So in semantics and psychology, it is now realized that our capacity to use concepts and refer to kinds need not depend on a grasp, implicit or explicit, on the necessary and sufficient conditions of membership of those categories. Humans have been able to use terms for chemical and physical kinds *(iron, liquid, salt, planet)* long before they understood the nature of those kinds. Though natural kinds may have essences, those essences are discovered not through the construction of definitions at the beginning of inquiry, but, if we are lucky, as the culmination of inquiry.

So biologists do not need a definition of *life* to help them recognize what they are talking about. But definitions are often useful. When categories overlap, or are easily confused with one another, the precision induced by definition is important, for definitions enable us to notice important distinctions that are easily overlooked. Confined as we are to the surface of a near-spherical globe, we can easily overlook the distinction between mass and weight, which is the interaction of mass and a gravitational field. So definitions that made this distinction explicit were important in the development of physics. As we saw in part 2, *gene* has been used to name very different kinds in biology; making these distinctions explicit avoids confusion. Similarly, different concepts of the organism may be important, and hence it is

important that the distinctions among them be explicitly marked. So definition is sometimes an important tool in theoretical advance.

However, we doubt that biology is currently impeded by biologists using *life* for distinct though related kinds. For example, we do not see how a definition of *life* is likely to help us with odd and hard-to-classify cases: prions, viruses, social insect colonies, or the much less plausible idea that the earth itself is a living system. The adequacy of the definition is settled by our view of the case, not vice versa. Consider for a moment the *Gaia hypothesis,* the idea (in one of its forms) that the earth itself, or perhaps just the biosphere, is a living organism. We see no useful role for a definition of life in evaluating this metaphor. In some very important ways, the earth is obviously unlike an organism. It is not the result of evolution through competition within an ancestral population of proto-Gaias. Nor does the biosphere result from a developmental cycle. The biosphere we have now will not produce a world-seed that grows into the biosphere of the earth at some later stage.

If we emphasize the typical histories of living things, then Gaia is not lifelike. But so what? Defenders of the Gaia idea emphasize the interconnections and reciprocal causal influences of living things with one another and the abiotic environment. These reciprocal interactions, they suggest, act like stabilizing or homeostatic mechanisms. There are both conceptual and empirical problems in evaluating this claim. As Kirchner (1991, 41) points out, in some respects, clearly life has not been homeostatic. Life, after all, radically altered the composition of the earth's atmosphere. So without an exact specification of the particular homeostatic mechanisms under consideration, the idea that the biosphere is a connected set of self-sustaining homeostatic mechanisms is too vague to evaluate. But even if it is made precise, the issue of whether the biosphere is alive is irrelevant. We do not need to detour through that question to evaluate the various Gaia hypotheses about the extent to which living systems and their environment change one another, the extent and ways in which these interactions are stabilized, or the extent to which these mutual changes make the earth more life-friendly.

So defining life is not a prerequisite for determining the scope of biology. The revival of interest in definitions of life has a different source: an interest in *universal biology.* All living systems on earth share many important properties. They are cells or are built from cells. Proteins play an essential role in the metabolism of all living things, and nucleic acids play an essential role in the process through which life gives rise to life. Replication and reproduction results in populations in competition, and natural selection on variation within those populations produces adaptation, sometimes complex adaptation. For all living things live in regimes in which natural selection is at work. But are these and other universal features of life on earth characteristic only

of life as we find it here and now? Or are some of these features truly universal: features of life anywhere, any time? Those interested in universal biology seek a characterization not just of life *as it happens to be,* but of life *as it must be.* For them, a definition of life is a specification of life's real essence (Bedau 1996; Langton 1996; Ray 1996).

It is worth pausing for a moment to remind ourselves just how ambitious this project really is. Biologists have always been interested in general principles. We have discussed plenty of candidates from ecology and evolutionary biology. It has been tough enough to find principles that are true of all life here and now. We have argued that adaptive and ecological hypotheses are best seen as hypotheses about particular clades, particular branches in the tree of life, not life as a whole. If that is right, then what price really universal biology: generalizations true not just of our life-world, but of any life-world?

Despite the ambition of the project, a number of biologists have explored the distinction between the specific features of life on earth and those features that life necessarily has. Gould, Kauffman, Goodwin, and Dawkins have deeply contrasting ideas on evolution, but they share this interest. We considered in chapter 12 both Gould's idea that the array of complex adaptations evolution on earth has produced is contingent, and his idea that the complexity of life tends to drift upward over time as a matter of statistical rather than evolutionary necessity. Gould's main emphasis is on the contingency of life's actual history. In contrast, Dawkins argues against a "historical accident" view of life's most central mechanisms. The most central features of both developmental biology and genetics are, he claims, features of universal biology. He argues against the possibility of *Lamarckian evolution,* at least if we understand Lamarckian evolution to involve the inheritance of only *adaptive* changes by the next generation. An organism's phenotype can certainly change its germ line genotype. For example, an organism may expose itself to mutagens in the environment, or act in ways that lower the efficiency of its DNA proofreading mechanisms. But that is not yet Lamarckian, for those changes in the stream of influence from parent to offspring do not make the offspring more likely to resemble the parent in this respect. A rat with a taste for nesting in nuclear reactors is unlikely to produce offspring with their DNA altered in such a way as to induce in them the same preference. Dawkins concedes that it is possible, though difficult, to imagine mechanisms in which the acquisition of a novel phenotypic trait changes the replicators responsible for the phenotype of the next generation in ways that make that novel phenotype reappear. In his view, it is much harder to imagine mechanisms that are sensitive to the distinction between adaptive and other novelties, and which make only adaptive changes more likely to reappear in the next generation (Dawkins 1986).

We are skeptical about Gould's ideas on contingency. We are also very wary of plausibility arguments for impossibility claims—"arguments from personal incredulity," as Dawkins himself has called them in a different context. Dawkins, after all, thinks that memes are replicators (13.6). Memes—if memes are taken to be the information content of ideas—do change, and sometimes adaptively, during the time they are in a particular interactor. If someone using a stone tool of a standard pattern discovers that grinding its edge on sandstone gives it a sharper cutting surface, that is a change in a specific meme token. It is a mutation, and one likely to be passed on because it is adaptive. So if interaction between phenotype and environment can improve a meme that is carried and transmitted, it is not obvious why Dawkins thinks that no similar mechanism could work with other replicators. Admittedly, if memes are replicators at all, they are late-model replicators. They are replicators that emerge deep into the history of a life-world. So perhaps the mechanisms that permit their evolution to be in this sense "Lamarckian" depend on a rich history of prior evolutionary change. But we do not see why this must be so. After all, the fidelity of genetic replication, and the sequestering of the germ line genes in many species, is itself the product of much evolution.

Despite our skepticism about these particular claims, we agree that there is a very good question lurking behind the idea of a universal biology. We seek not just an account of actual biology in all its diversity, but also an explanation of why that diversity is not greater still. However, we see two problems in asking for an explanation of the limits on life's diversity.

First, we should not conceptualize this question by contrasting chance with necessity. Consider, for example, David Raup's representation of possible and actual shell shapes. He shows that, to a first approximation, shell form can be represented as the outcome of only three different growth parameters. In light of this understanding, actual shells occupy a rather small region of the space of possible shells (for an elegant discussion, see Dawkins 1996, chap. 6). Why? Is this restriction a consequence of function, of subtle constraints on development, or of historical contingency? These are clearly difficult but interesting questions. But it is surely unlikely that most of the unoccupied region is literally impossible to occupy. It is equally unlikely that the occupied region is occupied through nothing but historical chance. Similarly, there are no species with three sexes, and that is no accident. As the literature on the evolution of sex makes clear, sex has a cost, and that cost would increase with the number of sexes. But should we infer that the evolution of three sexes is impossible? That would surely be rash: we can conceive of a developmental biology that might work with three sexes. Nuclear DNA has two parents, so we could have three if mitochondrial DNA came from a third. But an evolutionary trajectory leading to three sexes would be both available to a lineage and favored by selection only in very extraordinary circumstances. So, as Dennett (1995) has noted, contrasting historical accident with necessity is likely to be the wrong way of posing this problem. It is probably too crude a distinction to get at the questions that really interest us. Instead, we need some notion of a phenomenon's *improbability*. Bats evolved; no marsupial equivalent did. Is there some reason why a flying marsupial is less likely than a flying placental? Difficult though this question is to answer, it is surely a better question than asking whether a flying marsupial is impossible.

A second problem is the difficulty of testing conjectures about universal biology. This problem of testing is one of the fuels of the developing but over-hyped field of *artificial life*. One of the repeated themes of A-life literature is the "$N = 1$" problem, the problem of distinguishing between accidental and essential features of life with a sample size of one.

> Ideally, the science of biology should embrace all forms of life. However, in practice, it has been restricted to the study of a single instance of life, life on earth. Because biology is based on a sample size of one, we cannot know what features of life are peculiar to earth, and what features are general, characteristic of all life. (Ray 1996, 111)

One aim of A-life is to increase N, and in doing so, generate a definition of life that tells us which features of life are essential to life in and of itself. Just as "strong AI" claims that some computing systems housed in current or near-current computers are not mere simulations of thought, but instances of it, the defenders of "strong A-life" argue that some computer models of lifelike interactions are not simulations of life, but instances of it. They are alive.

The defenders of strong AI argue that a cognitive system is any system organized in the right way. Whether a system thinks is independent of its physical constitution. The essence of mind is form, organization, or function: some abstract property. Because the essential features of having a mental life are not tied to a specific physical implementation, thinking is *substrate-neutral*. Mental properties are functional properties, not physical ones. Strong A-life models itself on this line of argument. Being alive is substrate-neutral. Life is a feature of form, not matter. A living system is any system with the right organization or structure.

> Life is a property of form, not matter, a result of the organization of matter rather than something that inheres in the matter itself. Neither nucleotides nor amino acids nor any other carbon-chain molecule is alive—yet put them together in the right way, and the dynamic behavior that

> emerges out of their interactions is what we call life. It is effects, not things, upon which life is based—life is a kind of behavior, not a kind of stuff—and as such it is constituted of simpler behaviors, not simpler stuff. (Langton 1996, 53)

and therefore,

> it is possible to abstract the logical form of a machine from its physical hardware, it is natural to ask whether it is possible to abstract the logical form of an organism from its biochemical wetware. (Langton 1996, 55)

So in this view, the data structures in, for example, Thomas Ray's famous Tierra program are alive, not merely illustrations of life.

We see no merit at all in these claims. First, the form/matter distinction, the distinction on which the whole idea rests, is an untenable dichotomy. There is no single level of function or organization resting on a single level of matter. Rather, there is a cascade of increasingly or decreasingly abstract descriptions of any one system. In philosophy of psychology, the original home of the function/realization distinction, "two-levelism" has been powerfully criticized by William Lycan (1990). In David Marr's famous description of the structure of psychological theories (1980), there are at least two functional levels alone. The highest level describes the task that the psychological system accomplishes. In the case of vision, Marr claims that the task is to interpret the world in terms of moving, three-dimensional colored objects using patterns of stimulation of the retina as data. An intermediate level might describe how the system processes information in order to accomplish this task. It details the algorithms by which retinal patterns are transformed into representations of the world. The lowest level describes how these computational processes are physically implemented in the brain. Many authors have argued for a number of separate algorithmic or computational levels of description between the superficial level of task description and anything resembling a direct description of brain structure. Lycan has pointed out that much of what passes for a description of the "physical realization" of the mind is really a description of function. Synapses, the connections between brain cells, come in radically different forms, but for most purposes we can abstract away from this detail and describe them by their function: transferring excitation from one cell to the next.

Exactly the same multilevel picture applies to biological systems. For some purposes, a highly abstract, purely informational description of the genome may be appropriate. For others, we want to know in great physical detail the structure of the DNA molecule; for instance, in explaining its coiling

properties. Other needs will call for intermediate degrees of detail. There is a whole language of genetics—of introns and exons, of crossing-over, of gene duplication and gene repair—that is functional in abstracting away from the intricate details of molecular mechanisms (some of which are still, indeed, not known), but which is not wholly abstract. This is certainly not a language of form as opposed to matter. So the substrate neutrality thesis rests on a false dichotomy.

Moreover, we think that the idea that simulations are instances of life is an unnecessary hostage to fortune, for the importance of A-life models does not depend on the claim that they create life. The $N = 1$ problem is indeed a serious obstacle to the testing of conjectures in universal biology. But the $N = 1$ problem has been exaggerated, and in any case, the testing problem is not solved by deeming computer simulations to be alive. Of course exobiology would be great if we could do it; a genuinely independent life-world could scarcely fail to tell us much of importance about what is robust about biological process and what is not. But the problem of universal biology can be attacked here and now by the construction of distinct theories that have different implications for evolutionary, developmental, and ecological possibilities, and which can be tested by their application to the huge and varied experiment we actually have available. We do not have a wonderful array of theories that are well confirmed and empirically equivalent with respect to life on earth, but with different implications about how life might have been. $N = 1$ may begin to bite if and when we have to decide between empirically well-confirmed and locally equivalent theories: theories that make the same predictions about life here—predictions that are confirmed—but which make different predictions about what life might be like elsewhere. But we are yet to be indulged with such choices.

Evolutionary simulations will have an important role to play in constructing these theories of life's robust properties. Such models could test conditions under which particular developmental, genetic, ecological, or evolutionary phenomena would arise. Under what circumstances could a third sex evolve? Under what circumstances could variation be directed rather than random? Well-calibrated models that showed the evolution of exotic phenomena not observed in the natural world would be very suggestive indeed. But they can play that role as representations of biological processes, not manifestations of them.

In running simulations, we are trying to find out what those models predict, when those predictions are inaccessible to analytic techniques. The great virtue of these simulations is that one can play with various parameters and thus get a feel for which outcomes are robust under fine-scale changes

in the model, and which are not. Thus, for example, Nilsson's model of the evolution of eyes is impressive because the parameters are chosen conservatively, and yet eyes evolve, by geological standards, with great speed (Nilsson and Pelger 1994; see also Dawkins 1996, chap. 5). Simulations are important, and we will consider their message further in section 15.3. But nothing of what these models tell us depends on thinking of them as actually alive. Indeed, we think that the view that these programs are instances of life rather than representations of it trivializes the real questions that motivate universal biology. Consider, again, three sexes. We would like to know whether there are circumstances that would effectively select for three sexes. It is likely that only evolutionary modeling will advance our grip on this problem. But to do so, such models must be well calibrated. Their assumptions must be realistic. Suppose we were to accept that the data structures manipulated in a Tierra-like program were themselves alive. Suppose, further, that we accept that sex is defined not by the physical exchange of nucleic acids, but abstractly and functionally, as the A-life program urges. Sex, in this abstract conception, is information exchange. So any information exchange between token data structures before they are replicated is sex. There is no doubt that it is possible to develop models with three-way exchange of information between data structures. Hence, by this A-life definition, we could have life with three sexes. But this would be a trivial solution to the problem; it is too cheap. Unless the model faithfully represented the constraints on physically embodied living things—for example, constraints on development—it would not tell us what we wanted to know about the possibility of three sexes. If it did faithfully represent those constraints, we could learn what we wanted to know. But nothing would be added by insisting that the model manifests as well as represents life.

15.2 Universal Biology

So we interpret the project of defining life—investigating the extent to which features of the tree of life are historically specific to life here and now—as the program of universal biology. Until quite recently, the issue of universal biology was enmeshed with the issue of biological laws. Scientists and philosophers of science have often taken the main aim of scientific investigation to be the discovery of "laws of nature," such as Newton's laws of motion and of gravitational attraction. Ernest Rutherford, the famous New Zealand physicist who discovered that atoms are mostly empty space, thought that the discovery of such laws was an essential feature of science. Newton's "laws" turned out not to be laws after all, but nonetheless they

were the central exemplar of scientific discovery for two hundred years. Scientists record many particular facts: about the charge of particles, the structure of particular compounds, or the age and composition of a particular star. But their main task, in this conception, is to discover the universal principles that particular facts instantiate. Particular domains of science are characterized by distinctive laws, the laws that organize all the innumerable singular facts in each domain. For example, the laws of chemistry might be the general principles that specify the array of possible molecular structures while ruling out others as impossible.

One way of asking questions about biology's status as a science is to ask whether there are any *distinctly biological* laws of nature. To see what such laws might look like, consider von Baer's laws of embryology. In 1828 Karl Ernst von Baer suggested the following generalizations about development:

1. In development, generalized features appear before specialized ones.

2. Within major taxonomic groups, the embryos of different species resemble one another more in early development than they do in late development.

3. The embryos of higher species are like the embryos, but not the adults, of lower species.

4. The embryos of different taxonomic groups diverge progressively and do not recapitulate different levels of adult organization.

Suppose these or other generalizations turn out to hold true. A further question then arises: Are these generalizations reducible, in one of the senses we distinguished in section 6.1, to more general principles? That is, can they be incorporated within chemistry or physics as special cases of more general chemical or physical principles? The status of biology as a good and autonomous science has sometimes been tied to the existence of biological laws and their relation to the laws of more general disciplines. Biology, in this view, is an autonomous science in good repute only if biologists have discovered laws—and moreover, laws that are not just special cases of more fundamental principles. Physics-oriented philosophers of science such as J. J. C. Smart (1963) have suspected that biology is not in this sense a real science, but instead a technical discipline like civil engineering. In this view, biology merely explores the consequences of the operation of general physical and chemical principles in particular contexts. We saw in chapters 6 and 7 that the principles of Mendelian genetics are probably not reducible in any simple way to those of biochemistry. So if those principles counted as laws, they would form the subject matter of an autonomous discipline. However, as we

shall see, they probably do not count as laws by the classic criteria of lawlikeness.

Laws of nature have two features. First, they are exceptionless universal generalizations. The generalization

> On earth, all organisms have a particular genetic code in which four distinct bases specify twenty amino acids and a stop signal

is not a law of nature because it is spatiotemporally restricted, and it is not quite exceptionless (Dyer and Obar 1994, 73–74). Mendel's laws are not exceptionless either. Second, in a sense that no one has ever succeeded in making properly clear, laws of nature hold *necessarily*. Their truth is no accident. Consider the contrast between "No dense object 20 kilometers in diameter consists of chemically pure gold" and "No dense object 20 kilometers in diameter consists of chemically pure plutonium." The first statement may well be true. Quite likely, no large planetoid of chemically pure gold has ever formed. But if true, its truth is accidental. There is nothing about the way the universe works that debars such a possibility. The truth of the second statement, however, is no accident, as a lump of plutonium that large would be above critical mass and would blow apart. So while there could be a gold planetoid, there could not be a plutonium one. Hence only the second generalization is an application of a law of nature. So even if the "genetic code" were universal (perhaps because life here on earth is all the life there happens to have been), a specification of the codon/amino acid pairing is not a law of nature unless this pairing is the only pairing there could be, which it is not.

It is now widely accepted that in this sense, there are no biological laws of nature. Rosenberg argues that this follows from the fact that biological kinds are functionally rather than structurally identified. There are, he says, no interesting true generalizations about marine animals because of the great physical and structural heterogeneity of those animals (1994, 33–34). He takes biology to contrast with physics and chemistry in this respect. But it's very far from clear that he is right. To the contrary, the picture of physics and chemistry as scientific domains in which myriads of particular facts are organized by exceptionless laws, laws that it is our aim is to discover, may be wrong. It is arguable that this picture depends on an oversimplified view of those disciplines (Cartwright 1983, 1989). If so, then law-hunting is the wrong aim for universal biology. Universal biology will not consist of a set of exceptionless generalizations. And earthly biology cannot be segmented into a universal part—generalizations true of all life everywhere—and historically contingent, spatiotemporally restricted generalizations that happen to be true of life here and now.

Nonetheless, we agree that there is an itch to be scratched. We see two different routes by which universal biology can be pursued, one focusing on pattern and the other on mechanism. First, we can think of universal biology as a set of hypotheses—speculations might be a better term—about robust patterns in the history of life, here and now and in such other life-worlds as there may be. We have already seen one example of such a hypothesis. In section 12.1 we discussed Gould's claim that an increase in mean complexity over time depends on the fact that life starts close to the point of minimum complexity. If he is right, this is a pattern we would expect to see in most life-worlds. Dawkins has floated a much more ambitious set of ideas about robust patterns of complexity, outlining a series of complexity thresholds through which he expects all or most life histories to pass. Dawkins defends a "replicator-first" view of the origin of life, so for him, the first of these thresholds is the formation of a replicating molecule. The second he calls the "phenotype threshold," which is passed when replicators begin to increase not by virtue of their intrinsic chemical properties, but through phenotypic effects on their environment. A third critical threshold, in his view, is passed when replicators and their phenotypes become linked in teams; we might think of this as the invention of something like an organism (Dawkins 1995, 151–155).

These are very large scale hypotheses about robust evolutionary patterns. Many much more particular hypotheses have also been proposed. For example, Dennett (1995) discusses "forced moves" and "good tricks" in design space: adaptations we might reasonably expect to find in an independent experiment in life. If organisms that move and explore their environment evolve in a world, then vision will be a "good trick." It is clearly not inevitable: organisms need to be big enough to support eyes, and some kind of light-sensitive pigment must be available. But we certainly would not be surprised to find vision in an independent life-world. These specific pattern hypotheses are likely to be conditional rather than categorical. If avoiding Muller's ratchet explains the existence of sex, then we should not expect to find the equivalent of sex in a life-world unless creatures in it have segregated into distinct species and some of those species have small population sizes. For it is in such populations that mutations accumulate.

An alternative approach to universal biology focuses on mechanisms. We noted above Rosenberg's skepticism about biological laws, even earth-bound ones. We agree with his views on biological generalizations, but are not convinced that biology is distinctive in its lack of laws. Rosenberg himself thinks otherwise. He thinks that the physical sciences, dealing with a simpler and more structurally uniform domain, can still hope to discover simple universal

principles that underpin and explain "the buzzing, blooming confusion of nature" (Rosenberg 1994, 33, in turn borrowing from William James). Rosenberg takes this to mark an important difference between biology and other sciences. His pessimism about biology—his insistence that we can conceive of it only as a kind of useful instrument—seems to overlook the possibility that realist biology can be pursued not by seeking exceptionless general laws, but by discovering recurrent causal mechanisms.

Most obviously, natural selection itself will be a distinctive and critical mechanism operating in any living world, for the complex adaptive mechanisms distinctive of life can arise only by cumulative natural selection. We have emphasized that cumulative natural selection depends on more than variation, heritability, and differential fitness (2.2). So we cannot rule out the existence of semi-life-worlds, worlds in which replicators of a sort exist and interact with their environment in ways that enable them to gather the resources to replicate, but in which replication is so inaccurate, and the direction of selection so variable, that no complex structures have ever evolved. But these are precisely worlds in which we would be pushed to decide whether there was life or not.

Natural selection might not be the only universal or near-universal mechanism. It might turn out that the chemistry of life is inevitably carbon-based, so some biochemical mechanisms might be universal. At a larger scale, there might be universal aspects of development. Consider, for example, gastrulation, the first major reorganization of a developing animal embryo. In this process, the hollow ball of cells formed by the initial divisions of a fertilized egg folds to form a cup with an inner lining, beginning the more obvious process of cell differentiation. There are probably no important universal and exceptionless generalizations about gastrulation. Nonetheless, it is a conservative and conserved process. Gastrulation takes place in quite similar ways across animal life. It is a very important developmental mechanism, even if it is expressed somewhat differently in many developmental processes (Buss 1987, chap. 2). We doubt that gastrulation as such is likely to be a feature of a really universal biology, if such is to be had. But some developmental mechanisms might be. We might risk a modest wager that some form of developmental entrenchment—early aspects of development are increasingly difficult to change—will be a robust feature of life. So despite their antiquity, von Baer's laws may be part of a future exobiology!

Other claims about necessary mechanisms strike us as more suspect. Both Dawkins and Maynard Smith have floated the idea that the information transmitted from generation to generation must be digitally coded (Dawkins 1995; Maynard Smith 1996). They argue that if natural selection is to build

complex organisms, the fundamental mechanisms of replication must be digital. Complex living structures arise only through long histories of gradual change under cumulative selection, and that, in turn, requires high-fidelity replication. *Digital codes* can be replicated many times with high fidelity, for they are inherently far less ambiguous than *analog codes*. Analog codes, on the other hand, are impossible to replicate many times without critical degradation. A document photocopy chain a hundred links long will have an unreadable blur at the hundredth link. Send the same document through a hundred-link e-mail chain, however, and the first and the hundredth will probably be identical.

> An analog genetic system could be imagined. But we have already seen what happens to analog information when it is recopied over successive generations. It is Chinese Whispers. Boosted telephone systems, recopied tapes, photocopies of photocopies—analog systems are so vulnerable to cumulative degradation. Genes . . . can self-copy for ten million generations, and scarcely degrade at all. Darwinism works only because—apart from discrete mutations which natural selection either weeds out or preserves—the copying process is perfect. Only a digital genetic system is capable of sustaining Darwinism over eons of geological time. (Dawkins 1995, 19)

We discussed in chapter 5 our general worries about the idea of genes and genotypes as codes. Let us set these aside. We are still unconvinced of the digital encoding hypothesis. The fidelity of replication depends not only on the ease with which distinct characters in the code can be recognized for what they are, but on error correction systems as well. So even if analog replication has a higher error risk, if it is supported by good error detection and correction mechanisms, long chains of high-fidelity analog representations are possible. Thus if at each link in the photocopy chain, thousands of slightly varying copies are made, and only the best is retained for copying into the next link, then a long high-fidelity series is possible. Actual biology shows that this is no idle possibility. For, as Dawkins himself points out, a fertilized cell is no mere package of DNA. Cell differentiation in the early embryo depends on a series of chemical gradients in the fertile egg: from top to bottom, from front to back, and, often, from left to right. It is these gradients that cause different cells in the early embryo to differentiate from one another. If genes are digitally coded information, then chemical gradients are analog instructions telling cells where they are. Yet this information, this gradient, is reconstructed with high fidelity generation by generation.

Despite our skepticism about this hypothesis, we suspect that somewhat

more convincing arguments can be mounted for a universal biology of mechanism than of pattern, just because the basic mechanisms of life are more directly constrained by the physical and chemical basis of life. Thus the fact that energy is never converted with perfect efficiency has implications in ecology for the structure of food chains and communities. Big fierce animals are rare, for they can never harvest more than a smallish fraction of the energy potentially available to the primary producers (Colinvaux 1980).

15.3 Simulation and Emergence

Universal biology has been most consistently pursued in the field of artificial life, most importantly in the work of Stuart Kauffman (1993, 1995a,b). Much of this literature itself, and even more of the philosophical reflection on it, has focused on the issues of *emergence* and *self-organization.* The contribution this work makes to universal biology is the claim that there are both very general patterns and very general constraints that emerge out of the complexity of the organization of life. As we shall see, these constraints are often read as constraints on selection.

The idea of self-organization is the idea that living systems are inherently organized; organization arises spontaneously in the system itself rather than having to be imposed from the outside through the mechanism of selection. We shall see the importance of this idea in Kauffman's work shortly. This discussion of emergence links an empirical idea to a conceptual one. The central empirical idea defining emergence is that surprisingly complex system-level behavior can arise out of locally interacting simple units. Complexly behaving systems require neither complex parts nor central direction. The elements in A-life models are often quite simple units whose interactions are all governed by local rules—indeed, relatively simple local rules. But the behavior of the system as a whole is often adaptively complex. Some social insect colonies may provide natural examples of the phenomenon in question. Simply interacting simple creatures nonetheless produce complex, adaptive, and patterned behavior. So a good many of the most striking examples of A-life models can be seen as undercutting the idea that fancy systems must be built of fancy components. They show that complex system-level behavior may arise out of interacting simple components.

The conceptual idea is methodological. Since the interaction of the components determines system-level behavior, we will not get much of a handle on what the system will be like by studying the components in isolation. Understanding emergence as an empirical phenomenon will require new models of scientific explanation (see Burian and Richardson 1996; Clark 1996, 1997; Hendriks-Jansen 1996). We noted in section 7.3 that some ideas in developmental biology are thought to support a similar message.

An example might make these abstract points clearer. One good example is Reynolds's model of flocking behavior. He calls his simulated creatures "boids," and the rules they follow are very simple. Each acts

> to maintain a minimum distance from other objects in the environment, including other boids,
>
> to match velocities with boids in its neighborhood, and
>
> to move toward the perceived center of mass of the boids in its neighborhood. (Langton 1996, 66)

Despite the simplicity of these rules, boids simulate flocking rather well. Boids flow naturally around obstacles, and they show the illusion of coordination that we see in schools of fish and flocks of birds. So this example shows how creatures following very simple, locally cued behavioral rules could form flocks whose global behavior appears coordinated.

So some of these A-life simulations are very suggestive. But what, exactly, do they show? What is their evidential status? This question is particularly important in thinking about Kauffinan's work, for many see him as developing a picture of life that underplays the role of natural selection. His work is often presented both as showing restrictions on the power of natural selection and as showing that we do not need to invoke selection to explain order. Order arises "naturally."

Kauffinan's work exemplifies the idea that complex macroscopic organization can derive from the interactions of simple systems under local control. For example, Kauffman argues against a "replicator-first" version of the origin of life. He claims that two constraints make replicator-first views implausible. If the simplest bacteria are any guide, even first-generation replicators would have to be quite long, for short sequences would not exert phenotypic power over their propensity to be copied. Yet, despite the lack of evolved catalysts and evolved error-correcting machinery, these first longer sequences—the first sequences with phenotypic power—would have had to be replicated accurately enough to avoid an error catastrophe that would destroy the biological properties of their copies. Kauffinan doubts that this is possible (1993, 287–291; 1995a, 41–43). He defends instead a "metabolism-first" or a "cell-first" view of life's origins. When enough biochemicals are confined in a single system (and he suggests ways in which this might happen), the chance becomes quite high that there will be sufficient catalytic links between the individual constituents for the "soup" as a whole to

become "autocatalytic," sustaining itself without there being any element dedicated to replication. The properties of life emerge spontaneously at some threshold of complexity of the system as a whole without that system containing any element that plays a distinct role in its maintenance or replication. Life, in Kauffman's view, is an *emergent phenomenon:* it arises from relatively simple locally interacting constituents. It is a property of an ensemble, not of any special element within the ensemble.

We find these ideas on the origin of life interesting and suggestive, but Kauffman is probably best known for his ideas on evolution. These ideas are generated from very simple, very abstract models in which just two elements vary. N is the size of a population whose units vary in fitness. In these models, N is often thought of as a population of genes. K measures the "connectedness" between members of N. The more other units each unit of N interacts with, the greater is K. So K measures the extent to which the fate of each unit is determined locally: as K goes up, local control goes down. If we think of N as the genes in a genotype, K might measure the number of genes that determine whether a given gene is switched on or off. K's role can be modified by a third parameter, P, which modulates K. P measures the sensitivity of our target gene to its promoters and repressors. If P is high (near 1.0), the target gene's action is insensitive to its environment; for example, it will remain on unless all its inputs are telling it to turn off. If P is low (near 0.5), the target gene is sensitive to all its inputs, and high values of K will have a profound effect.

Kauffman derives some striking and lifelike general results from these models. Dennett (1995) suggests that we think of selection as an engine that, granted order of a certain kind—order with variation—generates design. One way to think of Kauffman's results—a way he often suggests—is to see them as showing that selection has rather more order to work with than we might have thought. At the beginning of life, selection would not have to build cells all the way up from amino acid biochemistry. Instead, richer and more complex structures would automatically arise and become available for selection. Given the size of gene populations in cells, selection would not have to build the whole array of differentiated cells from single-celled prototypes that varied only slightly from one another. For if K were low, but not too low, different gene activity patterns would automatically generate an array of cell types. In this sense, Kauffman thinks his models yield "order for free," not as a replacement for selection in the explanation of organic differentiation and adaptive design, but as a richer input to that process. This is at once a constraint on selection, for fewer apparently possible biological structures are really possible, and a boost to selection, for it makes it easier for selection to reach some regions of design space.

Even so, the most general and important result in the NK models is that connectedness damps down the effect of selection. Kauffman agrees that selection is central to the history of life. But he argues that it is effective only in certain adaptive landscapes. Selection, recall, can take a population from one phenotype to another only if the intermediate phenotypes are of intermediate fitness (2.2). Consider a rat population whose body weight averages about 1 kilogram, living in an environment in which rats would be better adapted if they weighed 2 kilos. If a 1.5 kilo rat is less fit than either a 1 kilo or a 2 kilo rat (too fat to run; too small to fight), then selection alone cannot edge the phenotype to 2 kilos, even if the 1 kilo average is disastrously less fit than the ideal 2 kilo rat. If (holding other aspects constant) there is a steady increase in fitness as weight approaches 2 kilos, the phenotype fitness landscape is *smooth.* If, instead of a smooth upward curve, when we plot weight against fitness we see a jagged curve with many rises and dips between 1 and 2 kilos, the fitness landscape is *rugged,* and there are many local optima.

How effective will selection be in these different fitness landscapes? This in part depends on a third factor. We have spoken so far of the fitness of *phenotypes*—in particular, of body weight. A further condition for effective selection is a reasonably systematic relationship between genotype and phenotype. Let's call the genotype of a rat that weighs 0.98 kilo **R. R*** is the genotype of a 1 kilo rat, and 1 kilo is the local optimum. If you are a rat in the range 0.8 to 1.2 kilos, you are best off being exactly 1 kilo. But can selection push a population of 0.98 kilo rats (with genotype **R**) to a population of 1 kilo rats (with genotype **R***)? Only if genotypes that are similar to **R** have a similar fitness, presumably because they have a similar phenotype with respect to body weight and other traits relevant to rat survival and reproduction. If a small variation in **R** (say, a change in one gene) produces a distinctly different phenotype, and hence a genotype of distinctly different fitness, then the fitness landscape is *uncorrelated.* Selection is ineffective if the fitness landscape is uncorrelated. If it is rugged but correlated, selection can at least push populations to local optima. If it is smooth and correlated, we can get to a global optimum. But in uncorrelated landscapes, selection takes us nowhere.

Recall that in Kauffman's models, K measures the connectedness of genes in a genome. K measures the number of genes that determine, say, whether a given gene is turned on. Now, according to Kauffman's models, if K gets too high, we should expect uncorrelated fitness landscapes. (Values above

3 or 4 are high unless a gene is relatively unresponsive to the genes connected to it—unless P is high.) So evolution under natural selection is possible only in a rather abstractly defined class of environments in which the linkage of the components is not too tight, and in which the fitness landscape is not too rugged.

What are we to make of these results? Some of them are genuinely striking. They seem to accord well with what we know of development. For instance, the models predict that early ontogeny should be more fixed than it is later because of the entrenchment of early mechanisms. Furthermore, Kauffman also argues that his models predict the pattern of the Cambrian radiation. We should expect evolutionary histories to be characterized by a "Cambrian explosion" pattern, with most diversity generated early and relatively less originating later. Kauffman's reasons depend in part on the broad developmental considerations we discussed in chapters 10 and 12, but also on the idea that many relatively low fitness peaks will be unoccupied early in an evolutionary radiation. As these peaks become occupied, it becomes harder to find a higher peak. Imagine, for example, a previously unoccupied region being penetrated for the first time by plant-eating insects. At first, many different varieties will find ways of making a living. But as time goes by, fewer and fewer changes will result in organisms whose lifestyles have not been pre-empted.

> The Cambrian explosion is like the earliest stages of the technological evolution of an entirely new invention, such as the bicycle. Recall the funny early forms: big front wheels, little back ones; little front wheels, big back ones. A flurry of forms branched out . . . giving rise to major and minor variants. Soon after a major innovation, discovery of profoundly different variations is easy. Later innovation is limited to modest improvements on increasingly optimised designs. (Kauffman 1995a, 13–14)

So a lot happens fast, then not much happens at all.

These models are very clearly suggestive, but we remain cautious. Note, for example, that Kauffman's "Cambrian pattern" is not the pattern of the actual Cambrian radiation. As we discussed in sections 12.2 and 12.3, the critical claim about the Cambrian is about *morphological* diversity, not *adaptive* diversity. Gould claims that the Cambrian saw morphological diversity at its maximum. He thinks that even after the Permian mass extinction, there was no comparable invention of new body organizations. In some views, Gould's picture reflects taxonomic practice rather than biological reality. But no one claims that the adaptation-building engine switched off after the Cambrian.

Many major adaptive complexes postdate that era. Yet Kauffman's claim is, as we have seen, one about adaptive evolution and adaptive complexes.

More generally, the very abstractness of these models makes their connection with real biological phenomena difficult to evaluate. A mean-spirited approach would be to argue that these models are like the "proofs" nineteenth-century scientists are alleged to have produced showing that bumblebees cannot fly. Consider, for example, the idea that genotypes are self-organized. Given their degree of interconnection, it is unlikely that selection could prevent mutation and other disruptions from "spreading genotypes more evenly over the fitness landscape" (Burian and Richardson 1996, 157–58). Yet genotypes are not just ordered and complex: they are very considerably differentiated from one another, and this in many ways must be the result of selection. The differences between primate genotypes may be partly due to drift, but surely many are the result of selection. So we *already know* that selection can change genotypes, despite their apparently high connectedness. That knowledge cuts across the model result that as the connectedness of a system goes up, and the number of elements in that system goes up, selection becomes increasingly ineffective. Kauffman's investigations into universal biology have discovered a "constraint on selection" that shows that most actual biology is impossible, and that much actual evolution has not happened.

But there is a more generous way of thinking about these models. They lead us to ask how evolution under natural selection dodges the apparent constraints that would seem to make it impossible. Is the number of effective units (the size of N) smaller than it would seem? Is effective connectivity less than it seems? It is often thought that if natural selection is to be effective, the phenotypes of organisms must be modular, with some traits able to vary independently of one another. So perhaps we should see these models as offering a hint that genotypes, too, are more modular than they seem. More generally, we should treat these models as "how possibly" explanations. Adaptationism's critics have often made the point that we should not conflate "how-possibly" explanations with "how-actually" explanations, and that point is well taken. Even so, how-possibly explanations are important. First, even in those areas in which we think we have approximately the right how-actually story, an expansion of the space of possible explanations is often useful, for competing explanations suggest critical tests. How-possibly explanations are still more important when we deal with puzzling phenomena. A how-possibly explanation of, say, the evolution of human language would be useful because we have no good grip on what intermediate forms of language

might be like and why they were adaptive. Language poses a *trajectory problem*. The same is true of the origin of life.

When it comes to universal biology, then, we are left shivering on the brink, nervous virgins wondering about sex. There is no denying the fascination of the problems posed. We would love to know which features of the tree of life are robust. Would relatively small, chemically possible changes in the DNA-RNA-protein transcription machinery preclude the evolution of sex? Would relatively small changes in mitochondrial inheritance make a third sex possible? Yet though speculating on these questions is fun, and simulation imposes some discipline on our speculation, we suspect that they remain empirically recalcitrant.

Further Reading

15.1 Putnam's work is primarily responsible for the insight that definition plays a relatively minor role in our grasp of concepts (see especially Putnam 1975, chapters 11 and 12). For an introduction to this view of our concepts and the way they relate to the world, see Devitt and Sterelny 1986. Griffiths (1997) applies these ideas specifically to concepts in biology. We rather doubt that the Gaia hypothesis has been worth all the ink spilled in its elaboration. But readers who think otherwise might find Joseph 1990, a very friendly overview and history of the hypothesis, enjoyable. Schneider and Boston 1991 is a well-balanced collection on the subject.

For a general introduction to A-life, see Emmeneche 1994. Chris Langton is a central figure in the development of A-life. He has edited a series of collections on A-life for the Sante Fe Institute for Studies in the Sciences of Complexity. These are published as *Artificial Life 1* to *Artificial Life N*, with a rapidly growing N. These volumes are very variable in content. Boden 1996 is a very useful anthology, partly but not wholly drawn from this series. Langton 1995 is also a good anthology, though much less philosophically oriented. For a fine critique of the form/matter dichotomy on which the strong A-life program rests, see Lycan 1990. Sober (1996) surveys this program in his typically lucid and sensible way. See also Sterelny 1997a.

15.2 Ernst Mayr has campaigned long and hard in defense of the idea that biology is both a good and an autonomous science (see the first two essays in Mayr 1988, and more recently, Mayr 1996). The relationship between biology and other sciences is central to the work of both Rosenberg (1985, 1994) and Dupré (1993). In very different ways, they both end up with the view that biology has a different character than physics and chemistry, though only

Rosenberg reads this as an indication that biology has a different status as a science. In the supplement to volume 64, number 4 (1997), *Philosophy of Science* has published an important symposium on laws in biology, with papers by Beatty, Brandon, Sober, and Mitchell. Weinert 1995 is a recent good collection on the general issue of laws of nature.

15.3 Kauffman's magnum opus (1993), as we have noted before, is very difficult. Kauffman 1995a is much more readable, though somewhat infested by musings on the meaning of life. In Kauffman 1995b, he gives a good short introduction to his views. There are good introductory discussions of his work in Emmeneche 1994, Depew and Weber 1995, Burian and Richardson 1996, and Weber and Depew 1996. There is a briefer introduction in Dennett 1995.