8

The Architecture of Complexity: Hierarchic Systems

In this chapter I should like to report on some things we have been learning about particular kinds of complex systems encountered in various sciences. The developments I shall discuss arose in the context of specific phenomena, but the theoretical formulations themselves make little reference to details of structure. Instead they refer primarily to the complexity of the systems under view without specifying the exact content of that complexity. Because of their abstractness, the theories may have relevance—application would be too strong a term—to many kinds of complex systems observed in the social, biological, and physical sciences.

In recounting these developments, I shall avoid technical detail, which can generally be found elsewhere. I shall describe each theory in the particular context in which it arose. Then I shall cite some examples of complex systems, from areas of science other than the initial application, to which the theoretical framework appears relevant. In doing so, I shall make reference to areas of knowledge where I am not expert—perhaps not even literate. The reader will have little difficulty, I am sure, in distinguishing instances based on idle fancy or sheer ignorance from instances that cast some light on the ways in which complexity exhibits itself wherever it is found in nature.

I shall not undertake a formal definition of "complex systems."[1] Roughly, by a complex system I mean one made up of a large number of

---

This chapter is a revision of a paper with the same title, reprinted with permission from *Proceedings of the American Philosophical Society,* 106(December 1962): 467–482.

[1] W. Weaver, in "Science and Complexity," *American Scientist,* 36(1948):536, has distinguished two kinds of complexity, disorganized and organized. We shall be concerned primarily with organized complexity.

parts that have many interactions. As we saw in the last chapter, in such systems the whole is more than the sum of the parts in the weak but important pragmatic sense that, given the properties of the parts and the laws of their interaction, it is not a trivial matter to infer the properties of the whole.[2]

The four sections that follow discuss four aspects of complexity. The first offers some comments on the frequency with which complexity takes the form of hierarchy—the complex system being composed of subsystems that in turn have their own subsystems, and so on. The second section theorizes about the relation between the structure of a complex system and the time required for it to emerge through evolutionary processes; specifically it argues that hierarchic systems will evolve far more quickly than nonhierarchic systems of comparable size. The third section explores the dynamic properties of hierarchically organized systems and shows how they can be decomposed into subsystems in order to analyze their behavior. The fourth section examines the relation between complex systems and their descriptions.

Thus my central theme is that complexity frequently takes the form of hierarchy and that hierarchic systems have some common properties independent of their specific content. Hierarchy, I shall argue, is one of the central structural schemes that the architect of complexity uses.

Hierarchic Systems

By a *hierarchic system,* or hierarchy, I mean a system that is composed of interrelated subsystems, each of the latter being in turn hierarchic in structure until we reach some lowest level of elementary subsystem. In most systems in nature it is somewhat arbitrary as to where we leave off the partitioning and what subsystems we take as elementary. Physics makes much use of the concept of "elementary particle," although particles have a disconcerting tendency not to remain elementary very long. Only a couple of generations ago the atoms themselves were elementary particles; today to the nuclear physicist they are complex systems. For certain purposes of astronomy whole stars, or even galaxies, can be regarded as elementary subsystems. In one kind of biological research a cell may be treated as an elementary subsystem; in another, a protein molecule; in still another, an amino acid residue.

Just why a scientist has a right to treat as elementary a subsystem that is in fact exceedingly complex is one of the questions we shall take up. For the moment we shall accept the fact that scientists do this all the time and that, if they are careful scientists, they usually get away with it.

Etymologically the word "hierarchy" has had a narrower meaning than I am giving it here. The term has generally been used to refer to a complex system in which each of the subsystems is subordinated by an authority relation to the system it belongs to. More exactly, in a hierarchic formal organization each system consists of a "boss" and a set of subordinate subsystems. Each of the subsystems has a "boss" who is the immediate subordinate of the boss of the system. We shall want to consider systems in which the relations among subsystems are more complex than in the formal organizational hierarchy just described. We shall want to include systems in which there is no relation of subordination among subsystems. (In fact even in human organizations the formal hierarchy exists only on paper; the real flesh-and-blood organization has many interpart relations other than the lines of formal authority.) For lack of a better term I shall use "hierarchy" in the broader sense introduced in the previous paragraphs to refer to all complex systems analyzable into successive sets of subsystems and speak of "formal hierarchy" when I want to refer to the more specialized concept.[3]

---

[2] See also John R. Platt, "Properties of Large Molecules that Go beyond the Properties of Their Chemical Sub-groups," *Journal of Theoretical Biology,* *1*(1961):342–358. Since the reductionism-holism issue is a major *cause de guerre* between scientists and humanists, perhaps we might even hope that peace could be negotiated between the two cultures along the lines of the compromise just suggested. As I go along, I shall have a little to say about complexity in the arts as well as in the natural sciences. I must emphasize the pragmatism of my holism to distinguish it sharply from the position taken by W. M. Elsasser in *The Physical Foundation of Biology* (New York: Pergamon Press, 1958).

[3] The mathematical term "partitioning" will not do for what I call here a hierarchy; for the set of subsystems and the successive subsets in each of these define the partitioning, independent of any systems of relations among the subsets. By "hierarchy" I mean the partitioning in conjunction with the relations that hold among its parts.

Social Systems

I have already given an example of one kind of hierarchy that is frequently encountered in the social sciences—a formal organization. Business firms, governments, and universities all have a clearly visible parts-within-parts structure. But formal organizations are not the only, or even the most common, kind of social hierarchy. Almost all societies have elementary units called families, which may be grouped into villages or tribes, and these into larger groupings, and so on. If we make a chart of social interactions, of who talks to whom, the clusters of dense interaction in the chart will identify a rather well-defined hierarchic structure. The groupings in this structure may be defined operationally by some measure of frequency of interaction in this sociometric matrix.

Biological and Physical Systems

The hierarchical structure of biological systems is a familiar fact. Taking the cell as the building block, we find cells organized into tissues, tissues into organs, organs into systems. Within the cell are well-defined subsystems—for example, nucleus, cell membrane, microsomes, and mitochondria.

The hierarchic structure of many physical systems is equally clear-cut. I have already mentioned the two main series. At the microscopic level we have elementary particles, atoms, molecules, and macromolecules. At the macroscopic level we have satellite systems, planetary systems, galaxies. Matter is distributed throughout space in a strikingly nonuniform fashion. The most nearly random distributions we find, gases, are not random distributions of elementary particles but random distributions of complex systems, that is, molecules.

A considerable range of structural types is subsumed under the term "hierarchy" as I have defined it. By this definition a diamond is hierarchic, for it is a crystal structure of carbon atoms that can be further decomposed into protons, neutrons, and electrons. However, it is a very "flat" hierarchy, in which the number of first-order subsystems belonging to the crystal can be indefinitely large. A volume of molecular gas is a flat hierarchy in the same sense. In ordinary usage we tend to reserve the word "hierarchy" for a system that is divided into a *small or moderate number* of subsystems, each of which may be further subdivided. Hence we do not ordinarily think of or refer to a diamond or a gas as a hierarchic structure. Similarly a linear polymer is simply a chain, which may be very long, of identical subparts, the monomers. At the molecular level it is a very flat hierarchy.

In discussing formal organizations, the number of subordinates who report directly to a single boss is called his *span of control.* I shall speak analogously of the *span* of a system, by which I shall mean the number of subsystems into which it is partitioned. Thus a hierarchic system is flat at a given level if it has a wide span at that level. A diamond has a wide span at the crystal level but not at the next level down, the atomic level.

In most of our theory construction in the following sections we shall focus our attention on hierarchies of moderate span, but from time to time I shall comment on the extent to which the theories might or might not be expected to apply to very flat hierarchies.

There is one important difference between the physical and biological hierarchies, on the one hand, and social hierarchies, on the other. Most physical and biological hierarchies are described in spatial terms. We detect the organelles in a cell in the way we detect the raisins in a cake—they are "visibly" differentiated substructures localized spatially in the larger structure. On the other hand, we propose to identify social hierarchies not by observing who lives close to whom but by observing who interacts with whom. These two points of view can be reconciled by defining hierarchy in terms of intensity of interaction, but observing that in most biological and physical systems relatively intense interaction implies relative spatial propinquity. One of the interesting characteristics of nerve cells and telephone wires is that they permit very specific strong interactions at great distances. To the extent that interactions are channeled through specialized communications and transportation systems, spatial propinquity becomes less determinative of structure.

Symbolic Systems

One very important class of systems has been omitted from my examples thus far: systems of human symbolic production. A book is a hierarchy in the sense in which I am using that term. It is generally divided into chapters, the chapters into sections, the sections into paragraphs, the paragraphs into sentences, the sentences into clauses and phrases, the clauses and phrases into words. We may take the words as our elementary units, or further subdivide them, as the linguist often does, into smaller units. If the book is narrative in character, it may divide into "episodes" instead of sections, but divisions there will be.

The hierarchic structure of music, based on such units as movements, parts, themes, phrases, is well known. The hierarchic structure of products of the pictorial arts is more difficult to characterize, but I shall have something to say about it later.

The Evolution of Complex Systems

Let me introduce the topic of evolution with a parable. There once were two watchmakers, named Hora and Tempus, who manufactured very fine watches. Both of them were highly regarded, and the phones in their workshops rang frequently—new customers were constantly calling them. However, Hora prospered, while Tempus became poorer and poorer and finally lost his shop. What was the reason?

The watches the men made consisted of about 1,000 parts each. Tempus had so constructed his that if he had one partly assembled and had to put it down—to answer the phone, say—it immediately fell to pieces and had to be reassembled from the elements. The better the customers liked his watches, the more they phoned him and the more difficult it became for him to find enough uninterrupted time to finish a watch.

The watches that Hora made were no less complex than those of Tempus. But he had designed them so that he could put together subassemblies of about ten elements each. Ten of these subassemblies, again, could be put together into a larger subassembly; and a system of ten of the latter subassemblies constituted the whole watch. Hence, when Hora had to put down a partly assembled watch to answer the phone, he lost only a small part of his work, and he assembled his watches in only a fraction of the man-hours it took Tempus.

It is rather easy to make a quantitative analysis of the relative difficulty of the tasks of Tempus and Hora: suppose the probability that an interruption will occur, while a part is being added to an incomplete assembly, is p. Then the probability that Tempus can complete a watch he has started without interruption is $(1 - p)^{1000}$—a very small number unless p

is 0.001 or less. Each interruption will cost on the average the time to assemble $1/p$ parts (the expected number assembled before interruption). On the other hand, Hora has to complete 111 subassemblies of ten parts each. The probability that he will not be interrupted while completing any one of these is $(1 - p)^{10}$, and each interruption will cost only about the time required to assemble five parts.[4]

Now if p is about 0.01—that is, there is one chance in a hundred that either watchmaker will be interrupted while adding any one part to an assembly—then a straightforward calculation shows that it will take Tempus on the average about four thousand times as long to assemble a watch as Hora.

We arrive at the estimate as follows:

1. Hora must make 111 times as many complete assemblies per watch as Tempus; but
2. Tempus will lose on the average 20 times as much work for each interrupted assembly as Hora (100 parts, on the average, as against 5); and
3. Tempus will complete an assembly only 44 times per million attempts ($0.99^{1000} = 44 \times 10^{-6}$), while Hora will complete nine out of ten ($0.99^{10} = 9 \times 10^{-1}$). Hence Tempus will have to make 20,000 as many attempts per completed assembly as Hora. $(9 \times 10^{-1})/(44 \times 10^{-6}) = 2 \times 10^4$. Multiplying these three ratios, we get

---

[4] The speculations on speed of evolution were first suggested by H. Jacobson's application of information theory to estimating the time required for biological evolution. See his paper "Information, Reproduction, and the Origin of Life," in *American Scientist*, 43(January 1955):119–127. From thermodynamic considerations it is possible to estimate the amount of increase in entropy that occurs when a complex system decomposes into its elements. (See for example, R. B. Setlow and E. C. Pollard, *Molecular Biophysics* (Reading, Mass.: Addison-Wesley, 1962), pp. 63–65, and references cited there.) But entropy is the logarithm of a probability; hence information, the negative of entropy, can be interpreted as the logarithm of the reciprocal of the probability—the "improbability," so to speak. The essential idea in Jacobson's model is that the expected time required for the system to reach a particular state is inversely proportional to the probability of the state—hence it increases exponentially with the amount of information (negentropy) of the state.

Following this line of argument, but not introducing the notion of levels and stable subassemblies, Jacobson arrived at estimates of the time required for evolution so large as to make the event rather improbable. Our analysis, carried through in the same way, but with attention to the stable intermediate forms, produces very much smaller estimates.

$$1/111 \times 100/5 \times 0.99^{10}/0.99^{1000}$$
$$= 1/111 \times 20 \times 20.000 \sim 4,000.$$

Biological Evolution

What lessons can we draw from our parable for biological evolution? Let us interpret a partially completed subassembly of k elementary parts as the coexistence of k parts in a small volume—ignoring their relative orientations. The model assumes that parts are entering the volume at a constant rate but that there is a constant probability, p, that the part will be dispersed before another is added, unless the assembly reaches a stable state. These assumptions are not particularly realistic. They undoubtedly underestimate the decrease in probability of achieving the assembly with increase in the size of the assembly. Hence the assumptions understate—probably by a large factor—the relative advantage of a hierarchic structure.

Although we cannot therefore take the numerical estimate seriously, the lesson for biological evolution is quite clear and direct. The time required for the evolution of a complex form from simple elements depends critically on the numbers and distribution of potential intermediate stable forms. In particular, if there exists a hierarchy of potential stable "subassemblies," with about the same span, s, at each level of the hierarchy, then the time required for a subassembly can be expected to be about the same at each level—that is, proportional to $1/(1 - p)^s$. The time required for the assembly of a system of n elements will be proportional to $\log_s n$, that is, to the number of levels in the system. One would say—with more illustrative than literal intent—that the time required for the evolution of multi-celled organisms from single-celled organisms might be of the same order of magnitude as the time required for the evolution of single-celled organisms from macromolecules. The same argument could be applied to the evolution of proteins from amino acids, of molecules from atoms, of atoms from elementary particles.

A whole host of objections to this oversimplified scheme will occur, I am sure, to every working biologist, chemist, and physicist. Before turning to matters I know more about, I shall lay at rest four of these problems, leaving the remainder to the attention of the specialists.

First, in spite of the overtones of the watchmaker parable, the theory assumes no teleological mechanism. The complex forms can arise from the simple ones by purely random processes. (I shall propose another model in a moment that shows this clearly.) Direction is provided to the scheme by the stability of the complex forms, once these come into existence. But this is nothing more than survival of the fittest—that is, of the stable.

Second, not all large systems appear hierarchical. For example, most polymers—such as nylon—are simply linear chains of large numbers of identical components, the monomers. However, for present purposes we can simply regard such a structure as a hierarchy with a span of one—the limiting case; for a chain of any length represents a state of relative equilibrium.[5]

Third, the evolutionary process does not violate the second law of thermodynamics. The evolution of complex systems from simple elements implies nothing, one way or the other, about the change in entropy of the entire system. If the process absorbs free energy, the complex system will have a smaller entropy than the elements; if it releases free energy, the opposite will be true. The former alternative is the one that holds for most biological systems, and the net inflow of free energy has to be supplied from the sun or some other source if the second law of thermodynamics is not to be violated. For the evolutionary process we are describing, the equilibria of the intermediate states need have only local and not global stability, and they may be stable only in the steady state—that is, as long as there is an external source of free energy that may be drawn upon.[6]

---

[5] There is a well-developed theory of polymer size, based on models of random assembly. See, for example, P. J. Flory, *Principles of Polymer Chemistry* (Ithaca: Cornell University Press, 1953), chapter 8. Since *all* subassemblies in the polymerization theory are stable, limitation of molecular growth depends on "poisoning" of terminal groups by impurities or formation of cycles rather than upon disruption of partially formed chains.

[6] This point has been made many times before, but it cannot be emphasized too strongly. For further discussion, see Setlow and Pollard, *Molecular Biophysics*, pp. 49–64; E. Schrödinger, *What Is Life?* (Cambridge: Cambridge University Press, 1945); and H. Linschitz, "The Information Content of a Bacterial Cell," in H. Quastler (ed.), *Information Theory in Biology* (Urbana: University of Illinois Press, 1953), pp. 251–262.

Because organisms are not energetically closed systems, there is no way to deduce the direction, much less the rate, of evolution from classical thermodynamic considerations. All estimates indicate that the amount of entropy, measured in physical units, involved in the formation of a one-celled biological organism is trivially small—about -10^{-11} cal/degree.[7] The "improbability" of evolution has nothing to do with this quantity of entropy, which is produced by every bacterial cell every generation. The irrelevance of quantity of information, in this sense, to speed of evolution can also be seen from the fact that exactly as much information is required to "copy" a cell through the reproductive process as to produce the first cell through evolution.

The fact of the existence of stable intermediate forms exercises a powerful effect on the evolution of complex forms that may be likened to the dramatic effect of catalysts upon reaction rates and steady-state distribution of reaction products in open systems.[8] In neither case does the entropy change provide us with a guide to system behavior.

Evolution of Multi-Cellular Organisms

We must consider a fourth objection to the watchmaker metaphor. However convincing a model the metaphor may provide for the evolution of atomic and molecular systems, and even uni-cellular organisms, it does not appear to fit the history of multi-cellular organisms. The metaphor assumes that complex systems are formed by combining sets of simpler systems, but this is not the way in which multi-cellular organisms have evolved. Although bacteria may, in fact, have been produced by a merging of mitochondria with the cells they inhabited, multi-cellular organisms have evolved through multiplication and specialization of the cells of a single system, rather than through the merging of previously independent subsystems.

Lest we dismiss the metaphor too quickly, however, we should observe that systems that evolve through specialization acquire the same kind of boxes-within-boxes structure (e.g., a digestive system consisting of mouth, larynx, esophagus, stomach, small and large intestines, colon; or a circulatory system consisting of a heart, arteries, veins, and capillaries) as is acquired by systems that evolve by assembly of simpler systems. The next main section of this chapter deals with nearly decomposable systems. It proposes that it is not assembly from components, per se, but hierarchic structure produced *either* by assembly or specialization, that provides the potential for rapid evolution.

The claim is that the potential for rapid evolution exists in any complex system that consists of a set of stable subsystems, each operating nearly independently of the detailed processes going on within the other subsystems, hence influenced mainly by the net inputs and outputs of the other subsystems. If the near-decomposability condition is met, the efficiency of one component (hence its contribution to the organism's fitness) does not depend on the detailed structure of other components.

Before examining this claim in detail, however, I should like to discuss briefly some non-biological applications of the watchmaker metaphor to illustrate the important advantages that hierarchic systems enjoy in other circumstances.

Problem Solving as Natural Selection

Hierarchy, as well as processes akin to natural selection, appear in human problem solving, a domain that has no obvious connection with biological evolution. Consider, for example, the task of discovering the proof for a difficult theorem. The process can be—and often has been—described as a search through a maze. Starting with the axioms and previously proved theorems, various transformations allowed by the rules of the mathematical systems are attempted, to obtain new expressions. These are modified in turn until, with persistence and good fortune, a sequence or path of transformations is discovered that leads to the goal.

The process ordinarily involves much trial and error. Various paths are tried; some are abandoned, others are pushed further. Before a solution is found, many paths of the maze may be explored. The more difficult and novel the problem, the greater is likely to be the amount of trial and error required to find a solution. At the same time the trial and error is not completely random or blind; it is in fact rather highly selective. The

---

[7] See Linschitz, "The Information Content." This quantity, 10^{-11} cal/degree, corresponds to about 10^{11} bits of information.

[8] See H. Kaeser, "Some Physico-chemical Aspects of Biological Organization," appendix, pp. 191–249, in C. H. Waddington, *The Strategy of the Genes* (London: George Allen and Unwin, 1957).

new expressions that are obtained by transforming given ones are examined to see whether they represent progress toward the goal. Indications of progress spur further search in the same direction; lack of progress signals the abandonment of a line of search. Problem solving requires *selective* trial and error.[9]

A little reflection reveals that cues signaling progress play the same role in the problem-solving process that stable intermediate forms play in the biological evolutionary process. In fact we can take over the watchmaker parable and apply it also to problem solving. In problem solving, a partial result that represents recognizable progress toward the goal plays the role of stable subassembly.

Suppose that the task is to open a safe whose lock has 10 dials, each with 100 possible settings, numbered from 0 to 99. How long will it take to open the safe by a blind trial-and-error search for the correct setting? Since there are 100^{10} possible settings, we may expect to examine about one half of these, on the average, before finding the correct one—that is, 50 billion billion settings. Suppose, however, that the safe is defective, so that a click can be heard when any one dial is turned to the correct setting. Now each dial can be adjusted independently and does not need to be touched again while the others are being set. The total number of settings that have to be tried is only 10 × 50, or 500. The task of opening the safe has been altered, by the cues the clicks provide, from a practically impossible one to a trivial one.[10]

---

[9] See A. Newell, J. C. Shaw, and H. A. Simon, "Empirical Explorations of the Logic Theory Machine," *Proceedings of the 1957 Western Joint Computer Conference*, February 1957 (New York: Institute of Radio Engineers); "Chess-Playing Programs and the Problem of Complexity," *IBM Journal of Research and Development*, 2(October 1958):320–335; and for a similar view of problem solving, W. R. Ashby, "Design for an Intelligence Amplifier," pp. 215–233 in C. E. Shannon and J. McCarthy, *Automata Studies* (Princeton: Princeton University Press, 1956).

[10] The clicking safe example was supplied by D. P. Simon. Ashby, "Design for an Intelligence Amplifier," p. 230, has called the selectivity involved in situations of this kind "selection by components." The even greater reduction in time produced by hierarchization in the clicking safe example, as compared with the watchmaker's metaphor, is due to the fact that a random *search* for the correct combination is involved in the former case, while in the latter the parts come together in the right order. It is not clear which of these metaphors provides the better model for biological evolution, but we may be sure that the watchmaker's

A considerable amount has been learned in the past thirty years about the nature of the mazes that represent common human problem-solving tasks—proving theorems, solving puzzles, playing chess, making investments, balancing assembly lines, to mention a few. All that we have learned about these mazes points to the same conclusion: that human problem solving, from the most blundering to the most insightful, involves nothing more than varying mixtures of trial and error and selectivity. The selectivity derives from various rules of thumb, or heuristics, that suggest which paths should be tried first and which leads are promising. We do not need to postulate processes more sophisticated than those involved in organic evolution to explain how enormous problem mazes are cut down to quite reasonable size (see also chapters 3 and 4).[11]

The Sources of Selectivity

When we examine the sources from which the problem-solving system, or the evolving system, as the case may be, derives its selectivity, we discover that selectivity can always be equated with some kind of feedback of information from the environment.

Let us consider the case of problem solving first. There are two basic kinds of selectivity. One we have already noted: various paths are tried out, the consequences of following them are noted, and this information is used to guide further search. In the same way in organic evolution various complexes come into being, at least evanescently, and those that are stable provide new building blocks for further construction. It is this information about stable configurations, and not free energy or negentropy from the sun, that guides the process of evolution and provides the selectivity that is essential to account for its rapidity.

The second source of selectivity in problem solving is previous experience. We see this particularly clearly when the problem to be solved is

---

metaphor gives an exceedingly conservative estimate of the savings due to hierarchization. The safe may give an excessively high estimate because it assumes all possible arrangements of the elements to be equally probable. For an application of a variant of the watchmaker and the clicking safe arguments to structure at the molecular level, see J. D. Watson, *Molecular Biology of the Gene*, 3rd ed. (Menlo Park, CA: W. A. Benjamin, 1976), pp. 107–108.

[11] A. Newell and H. A. Simon, "Computer Simulation of Human Thinking," *Science*, *134*(December 22, 1961):2011–2017.

similar to one that has been solved before. Then, by simply trying again the paths that led to the earlier solution, or their analogues, trial-and-error search is greatly reduced or altogether eliminated.

What corresponds to this latter kind of information in organic evolution? The closest analogue is reproduction. Once we reach the level of self-reproducing systems, a complex system, when it has once been achieved, can be multiplied indefinitely. Reproduction in fact allows the inheritance of acquired characteristics, but at the level of genetic material, of course; that is, only characteristics acquired by the genes can be inherited. We shall return to the topic of reproduction in the final section of this essay.

On Empires and Empire Building

We have not exhausted the categories of complex systems to which the watchmaker argument can reasonably be applied. Philip assembled his Macedonian empire and gave it to his son, to be later combined with the Persian subassembly and others into Alexander's greater system. On Alexander's death his empire did not crumble to dust but fragmented into some of the major subsystems that had composed it.

The watchmaker argument implies that if one would be Alexander, one should be born into a world where large stable political systems already exist. Where this condition was not fulfilled, as on the Scythian and Indian frontiers, Alexander found empire building a slippery business. So too, T. E. Lawrence's organizing of the Arabian revolt against the Turks was limited by the character of his largest stable building blocks, the separate, suspicious desert tribes.

The profession of history places a greater value upon the validated particular fact than upon tendentious generalization. I shall not elaborate upon my fancy therefore but shall leave it to historians to decide whether anything can be learned for the interpretation of history from an abstract theory of hierarchic complex systems.

Conclusion: The Evolutionary Explanation of Hierarchy

We have shown thus far that complex systems will evolve from simple systems much more rapidly if there are stable intermediate forms than if there are not. The resulting complex forms in the former case will be hierarchic. We have only to turn the argument around to explain the observed predominance of hierarchies among the complex systems nature presents to us. Among possible complex forms, hierarchies are the ones that have the time to evolve. The hypothesis that complexity will be hierarchic makes no distinction among very flat hierarchies, like crystals and tissues and polymers, and the intermediate forms. Indeed in the complex systems we encounter in nature examples of both forms are prominent. A more complete theory than the one we have developed here would presumably have something to say about the determinants of width of span in these systems.

Nearly Decomposable Systems

In hierarchic systems we can distinguish between the interactions *among* subsystems, on the one hand, and the interactions *within* subsystems—that is, among the parts of those subsystems—on the other. The interactions at the different levels may be, and often will be, of different orders of magnitude. In a formal organization there will generally be more interaction, on the average, between two employees who are members of the same department than between two employees from different departments. In organic substances intermolecular forces will generally be weaker than molecular forces, and molecular forces weaker than nuclear forces.

In a rare gas the intermolecular forces will be negligible compared to those binding the molecules—we can treat the individual particles for many purposes as if they were independent of each other. We can describe such a system as *decomposable* into the subsystems comprised of the individual particles. As the gas becomes denser, molecular interactions become more significant. But over some range we can treat the decomposable case as a limit and as a first approximation. We can use a theory of perfect gases, for example, to describe approximately the behavior of actual gases if they are not too dense. As a second approximation we may move to a theory of *nearly decomposable* systems, in which the interactions among the subsystems are weak but not negligible.

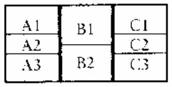
At least some kinds of hierarchic systems can be approximated successfully as nearly decomposable systems. The main theoretical findings from the approach can be summed up in two propositions: (1) in a nearly decomposable system the short-run behavior of each of the component subsystems is approximately independent of the short-run behavior of the other components; (2) in the long run the behavior of any one of the components depends in only an aggregate way on the behavior of the other components.

Let me provide a very concrete simple example of a nearly decomposable system.[12] Consider a building whose outside walls provide perfect thermal insulation from the environment. We shall take these walls as the boundary of our system. The building is divided into a large number of rooms, the walls between them being good, but not perfect, insulators. The walls between rooms are the boundaries of our major subsystems. Each room is divided by partitions into a number of cubicles, but the partitions are poor insulators. A thermometer hangs in each cubicle. Suppose that at the time of our first observation of the system there is a wide variation in temperature from cubicle to cubicle and from room to room—the various cubicles within the building are in a state of thermal disequilibrium. When we take new temperature readings several hours later, what shall we find? There will be very little variation in temperature among the cubicles within each single room, but there may still be large temperature variations *among* rooms. When we take readings again several days later, we find an almost uniform temperature throughout the building; the temperature differences among rooms have virtually disappeared.

We can describe the process of equilibrium formally by setting up the usual equations of heat flow. The equations can be represented by the matrix of their coefficients, r_{ij}, where r_{ij} is the rate at which heat flows from the ith cubicle to the jth cubicle per degree difference in their temperatures. If cubicles i and j do not have a common wall, r_{ij} will be zero. If cubicles i and j have a common wall and are in the same room, r_{ij} will be large. If cubicles i and j are separated by the wall of a room, r_{ij} will be nonzero but small. Hence, by grouping together all the cubicles that are in the same room, we can arrange the matrix of coefficients so that all its large elements lie inside a string of square submatrices along the main diagonal. All the elements outside these diagonal squares will be either zero or small (see figure 7). We may take some small number, ε, as the upper bound of the extradiagonal elements. We shall call a matrix having these properties a *nearly decomposable matrix*.

	A1	A2	A3	B1	B2	C1	C2	C3
A1	—	100	—	2	—	—	—	—
A2	100	—	100	1	1	—	—	—
A3	—	100	—	—	2	—	—	—
B1	2	1	—	—	100	2	1	—
B2	—	1	2	100	—	—	1	2
C1	—	—	—	2	—	—	100	—
C2	—	—	—	1	1	100	—	100
C3	—	—	—	—	2	—	100	—

Figure 7
A hypothetical nearly decomposable system. In terms of the heat-exchange example of the text. A1, A2, and A3 may be interpreted as cubicles in one room, B1 and B2 as cubicles in a second room, and C1, C2, and C3 as cubicles in a third. The matrix entries then are the heat diffusion coefficients between cubicles:



Now it has been proved that a dynamic system that can be described by a nearly decomposable matrix has the properties, stated earlier, of a nearly decomposable system. In our simple example of heat flow this means that in the short run each room will reach an equilibrium temperature (an average of the initial temperatures of its offices) nearly independently of the others and that each room will remain approximately in a state of equilibrium over the longer period during which an over-all temperature equilibrium is being established throughout the building.

---

[12] This discussion of near decomposability is based upon H. A. Simon and A. Ando, "Aggregation of Variables in Dynamic Systems," *Econometrica*, 29 (April 1961):111–138. The example is drawn from the same source, pp. 117–118. For subsequent development and applications of the theory see P. J. Courtois, *Decomposability: Queueing and Computer System Applications* (New York, NY: Academic Press, 1977); Y. Iwasaki and H. A. Simon, "Causality and Model Abstraction," *Artificial Intelligence*, 67(1994):143–194; and D. F. Rogers and R. D. Plante, "Estimating Equilibrium Probabilities for Band Diagonal Markov Chains Using Aggregation and Disaggregation Techniques," *Computers in Operations Research*, 20(1993):857–877.

After the intra-room short-run equilibria have been reached, a single thermometer in each room will be adequate to describe the dynamic behavior of the entire system—separate thermometers in each cubicle will be superfluous.

Near Decomposability of Social Systems

As a glance at figure 7 shows, near decomposability is a rather strong property for a matrix to possess, and the matrices that have this property will describe very special dynamic systems—vanishingly few systems out of all those that are thinkable. How few they will be depends of course on how good an approximation we insist upon. If we demand that epsilon be very small, correspondingly few dynamic systems will fit the definition. But we have already seen that in the natural world nearly decomposable systems are far from rare. On the contrary, systems in which each variable is linked with almost equal strength with almost all other parts of the system are far rarer and less typical.

In economic dynamics the main variables are the prices and quantities of commodities. It is empirically true that the price of any given commodity and the rate at which it is exchanged depend to a significant extent only on the prices and quantities of a few other commodities, together with a few other aggregate magnitudes, like the average price level or some over-all measure of economic activity. The large linkage coefficients are associated in general with the main flows of raw materials and semifinished products within and between industries. An input-output matrix of the economy, giving the magnitudes of these flows, reveals the nearly decomposable structure of the system—with one qualification. There is a consumption subsystem of the economy that is linked strongly to variables in most of the other subsystems. Hence we have to modify our notions of decomposability slightly to accommodate the special role of the consumption subsystem in our analysis of the dynamic behavior of the economy.

In the dynamics of social systems, where members of a system communicate with and influence other members, near decomposability is generally very prominent. This is most obvious in formal organizations, where the formal authority relation connects each member of the organization with one immediate superior and with a small number of subordinates.

Of course many communications in organizations follow other channels than the lines of formal authority. But most of these channels lead from any particular individual to a very limited number of his superiors, subordinates, and associates. Hence departmental boundaries play very much the same role as the walls in our heat example.

Physicochemical Systems

In the complex systems familiar in biological chemistry, a similar structure is clearly visible. Take the atomic nuclei in such a system as the elementary parts of the system, and construct a matrix of bond strengths between elements. There will be matrix elements of quite different orders of magnitude. The largest will generally correspond to the covalent bonds, the next to the ionic bonds, the third group to hydrogen bonds, still smaller linkages to van der Waals forces.[13] If we select an epsilon just a little smaller than the magnitude of a covalent bond, the system will decompose into subsystems—the constituent molecules. The smaller linkages will correspond to the intermolecular bonds.

It is well known that high-energy, high-frequency vibrations are associated with the smaller physical subsystems and low-frequency vibrations with the larger systems into which the subsystems are assembled. For example, the radiation frequencies associated with molecular vibrations are much lower than those associated with the vibrations of the planetary electrons of the atoms; the latter in turn are lower than those associated with nuclear processes.[14] Molecular systems are nearly decomposable systems, with the short-run dynamics relating to the internal structures of

---

[13] For a survey of the several classes of molecular and intermolecular forces, and their dissociation energies, see Setlow and Pollard, *Molecular Biophysics*, chapter 6. The energies of typical covalent bonds are of the order of 80–100 k cal/mole, of the hydrogen bonds, 10 k cal/mole. Ionic bonds generally lie between these two levels; the bonds due to van der Waals forces are lower in energy.

[14] Typical wave numbers for vibrations associated with various systems (the wave number is the reciprocal of wave length, hence proportional to frequency):
Steel wire under tension—10^{-10} to 10^{-9} cm^{-1}
Molecular rotations—10^{0} to 10^{2} cm^{-1}
Molecular vibrations—10^{2} to 10^{3} cm^{-1}
Planetary electrons—10^{4} to 10^{5} cm^{-1}
Nuclear rotations—10^{9} to 10^{10} cm^{-1}
Nuclear surface vibrations—10^{11} to 10^{12} cm^{-1}

the subsystems and the long-run dynamics to the interactions of these subsystems.

A number of the important approximations employed in physics depend for their validity on the near decomposability of the systems studied. The theory of the thermodynamics of irreversible processes, for example, requires the assumption of macroscopic disequilibrium but microscopic equilibrium, exactly the situation described in our heat-exchange example.[15] Similarly computations in quantum mechanics are often handled by treating weak interactions as producing perturbations on a system of strong interactions.

Some Observations on Hierarchic Span

To understand why the span of hierarchies is sometimes very broad—as in crystals—and sometimes narrow, we need to examine more detail of the interactions. In general the critical consideration is the extent to which interaction between two (or a few) subsystems excludes interaction of these subsystems with the others. Let us examine first some physical examples.

Consider a gas of identical molecules, each of which can form covalent bonds in certain ways with others. Let us suppose that we can associate with each atom a specific number of bonds that it is capable of maintaining simultaneously. (This number is obviously related to the number we usually call its valence.) Now suppose that two atoms join and that we can also associate with the combination a specific number of external bonds it is capable of maintaining. If this number is the same as the number associated with the individual atoms, the bonding process can go on indefinitely—the atoms can form crystals or polymers of indefinite extent. If the number of bonds of which the composite is capable is less than the number associated with each of the parts, then the process of agglomeration must come to a halt.

We need only mention some elementary examples. Ordinary gases show no tendency to agglomerate, because the multiple bonding of atoms "uses up" their capacity to interact. While each oxygen atom has a valence of two, the O_2 molecules have a zero valence. Contrariwise, indefinite chains of single-bonded carbon atoms can be built up, because a chain of any number of such atoms, each with two side groups, has a valence of exactly two.

Now what happens if we have a system of elements that possess both strong and weak interaction capacities and whose strong bonds are exhaustible through combination? Subsystems will form, until all the capacity for strong interaction is utilized in their construction. Then these subsystems will be linked by the weaker second-order bonds into larger systems. For example, a water molecule has essentially a valence of zero—all the potential covalent bonds are fully occupied by the interaction of hydrogen and oxygen molecules. But the geometry of the molecule creates an electric dipole that permits weak interaction between the water and salts dissolved in it—whence such phenomena as its electrolytic conductivity.[16]

Similarly it has been observed that, although electrical forces are much stronger than gravitational forces, the latter are far more important than the former for systems on an astronomical scale. The explanation of course is that the electrical forces, being bipolar, are all "used up" in the linkages of the smaller subsystems and that significant net balances of positive or negative charges are not generally found in regions of macroscopic size.

In social as in physical systems there are generally limits on the simultaneous interaction of large numbers of subsystems. In the social case these limits are related to the fact that a human being is more nearly a serial than a parallel information-processing system. He or she can carry on only one conversation at a time, and although this does not limit the size of the audience to which a mass communication can be addressed, it does limit the number of people simultaneously involved in most other forms of social interaction. Apart from requirements of direct interactions, most roles impose tasks and responsibilities that are time consuming. One cannot, for example, enact the role of "friend" with large numbers of other people.

It is probably true that in social as in physical systems the higher-frequency dynamics are associated with the subsystems and the lower-

---

[15] S. R. de Groot, *Thermodynamics of Irreversible Processes* (New York: Interscience Publishers, 1951), pp. 11–12.

[16] See, for example, L. Pauling, *General Chemistry* (San Francisco: W. H. Freeman, 2nd ed., 1953), chapter 15.

frequency dynamics with the larger systems. It is generally believed, for example, that the relevant planning horizon of executives is longer, the higher their location in the organizational hierarchy. It is probably also true that both the average duration of an interaction between executives and the average interval between interactions are greater at higher than lower levels.

Summary: Near Decomposability

We have seen that hierarchies have the property of near decomposability. Intracomponent linkages are generally stronger than intercomponent linkages. This fact has the effect of separating the high-frequency dynamics of a hierarchy—involving the internal structure of the components—from the low-frequency dynamics—involving interaction among components. We shall turn next to some important consequences of this separation for the description and comprehension of complex systems.

Biological Evolution Revisited

Having examined the properties of nearly-decomposable systems, we can now complete our discussion of the evolution of multi-cellular organisms through specialization of tissues and organs. An organ performs a specific set of functions, each usually requiring continual interaction among its component parts (a sequence of chemical reactions, say, each step employing a particular enzyme for its execution). It draws raw materials from other parts of the organism and delivers products to other parts, but these input and output processes depend only in an aggregate way on what is occurring within each specific organ. Like a business firm in an economic market, each organ can perform its functions in blissful ignorance of the detail of activity in other organs, with which it is connected by the digestive, circulatory, and excretory systems and other transport channels.

Expressing the matter slightly differently, changes within an organ affect the other parts of the organism mainly by changing the relation between the quantities of outputs they produce and the inputs they require (that is, their overall efficiency). Thus, biological organisms are nearly-decomposable: the interactions *within* units at any level are rapid and intense in comparison with the interactions *between* units at the same level. Inventories of various substances, held in the circulatory system or in special tissues, slow down and buffer effects of each unit on the others. In the short run, single units (e.g., single organs) operate nearly independently of the detail of operation of the other units.

Within the Darwinian framework of natural selection there is no way in which the fitness (efficiency) of individual tissues or organs can be separately evaluated; fitness is measured by the number of offspring of the entire organism. Evolution is like a complex experiment, with fitness as the sole dependent variable, and the structures of the individual genes as independent variables. The goal of the process is to compare the contribution to total fitness of alternative forms (alleles) of each gene—and of combinations of these alternatives for sets of genes.

If, in fact, the fitness of a particular gene depended on which alleles of all the other genes it was combined with, the combinatorics, involving tens of thousands of genes in complex organisms, would be staggering, and the problem of measuring the contribution of a particular allele to fitness would be overwhelming.[17]

With near-decomposability, we can assume that the *relative* efficiency of two different designs for the same organ (e.g., two different gene sequences with the same function) is approximately independent of which variants of other organs are present in the organism. The total fitness is essentially an additive measure of the nearly independent contributions of the individual organs. Essentially, we obtain the advantages of the clicking safe: the "correct" setting of each dial (the genes governing one organ's processes) can be determined independently of how the other dials are currently set. The search is for effective sets of organs instead of effective sets of individual genes.

Enough is known today about the architecture of the genome to be reasonably certain that it has a hierarchical control structure mapping reasonably closely to the hierarchy of processes in the organism.[18] Of

---

[17] With only two alleles for each of N genes, 2^N alternatives would have to be evaluated by selection. This is equivalent, in the watchmaker metaphor, to assembling 2^N parts without interruption. For an organism with even a thousand genes, say, change by natural selection would be extremely slow, even on a geological scale.

[18] F. Jacob and J. Monod, "Genetic Regulatory Mechanisms in the Synthesis of Proteins," *Molecular Biology*, 3(1961):318–56.

course, this is a gross simplification of the total picture in any actual organism. In addition to the genes that operate in particular organs (turned on and off by control genes), there are also the genes that determine the more general metabolic processes that are found within all the cells. But these common intra-cellular processes are at the cell level of the hierarchy, below the level of tissues and organs, and again the corresponding genes can be supposed to operate nearly independently of those that control specialized processes in specific organs.[19]

The Description of Complexity

If you ask a person to draw a complex object—such as a human face—he will almost always proceed in a hierarchic fashion.[20] First he will outline the face. Then he will add or insert features: eyes, nose, mouth, ears, hair. If asked to elaborate, he will begin to develop details for each of the features—pupils, eyelids, lashes for the eyes, and so on—until he reaches the limits of his anatomical knowledge. His information about the object is arranged hierarchically in memory, like a topical outline.

When information is put in outline form, it is easy to include information about the relations among the major parts and information about the internal relations of parts in each of the suboutlines. Detailed information about the relations of subparts belonging to different parts has no place in the outline and is likely to be lost. The loss of such information and the preservation mainly of information about hierarchic order is a salient characteristic that distinguishes the drawings of a child or someone untrained in representation from the drawing of a trained artist. (I am speaking of an artist who is striving for representation.)

---

[19] How hierarchical architectures of these kinds can be introduced into the genetic algorithms discussed in chapter 7, in order to speed up their rates of learning or evolution, is discussed by John H. Holland in *Adaptation in Natural and Artificial Systems* (Ann Arbor, MI: The University of Michigan Press, 1975). See especially pp. 167–168 and 152–153.

[20] George A. Miller has collected protocols from subjects who were given the task of drawing faces and finds that they behave in the manner described here (private communication). See also E. H. Gombrich, *Art and Illusion* (New York: Pantheon Books, 1960), pp. 291–296.

Near Decomposability and Comprehensibility

From our discussion of the dynamic properties of nearly decomposable systems, we have seen that comparatively little information is lost by representing them as hierarchies. Subparts belonging to different parts only interact in an aggregative fashion—the detail of their interaction can be ignored. In studying the interaction of two large molecules, generally we do not need to consider in detail the interactions of nuclei of the atoms belonging to the one molecule with the nuclei of the atoms belonging to the other. In studying the interaction of two nations, we do not need to study in detail the interactions of each citizen of the first with each citizen of the second.

The fact then that many complex systems have a nearly decomposable, hierarchic structure is a major facilitating factor enabling us to understand, describe, and even "see" such systems and their parts. Or perhaps the proposition should be put the other way round. If there are important systems in the world that are complex without being hierarchic, they may to a considerable extent escape our observation and understanding. Analysis of their behavior would involve such detailed knowledge and calculation of the interactions of their elementary parts that it would be beyond our capacities of memory or computation.[21]

---

[21] I believe the fallacy in the central thesis of W. M. Elsasser's *The Physical Foundation of Biology,* mentioned earlier, lies in his ignoring the simplification in description of complex systems that derives from their hierarchic structure. Thus (p. 155):

> If we now apply similar arguments to the coupling of enzymatic reactions with the substratum of protein molecules, we see that over a sufficient period of time, the information corresponding to the structural details of these molecules will be communicated to the dynamics of the cell, to higher levels of organization as it were, and may influence such dynamics. While this reasoning is only qualitative, it lends credence to the assumption that in the living organism, unlike the inorganic crystal, the effects of microscopic structure cannot be simply averaged out; as time goes on this influence will pervade the behavior of the cell "at all levels."

But from our discussion of near decomposability it would appear that those aspects of microstructure that control the slow developmental aspects of organismic dynamics can be separated out from the aspects that control the more rapid cellular metabolic processes. For this reason we should not despair of unraveling the web of causes. See also J. R. Platt's review of Elsasser's book in *Perspectives in Biology and Medicine,* 2(1959):243–245.

I shall not try to settle which is chicken and which is egg: whether we are able to understand the world because it is hierarchic or whether it appears hierarchic because those aspects of it which are not elude our understanding and observation. I have already given some reasons for supposing that the former is at least half the truth—that evolving complexity would tend to be hierarchic—but it may not be the whole truth.

Simple Descriptions of Complex Systems

One might suppose that the description of a complex system would itself be a complex structure of symbols—and indeed it may be just that. But there is no conservation law that requires that the description be as cumbersome as the object described. A trivial example will show how a system can be described economically. Suppose the system is a two-dimensional array like this:

$$\begin{matrix}
A & B & M & N & R & S & H & I \\
C & D & O & P & T & U & J & K \\
M & N & A & B & H & I & R & S \\
O & P & C & D & J & K & T & U \\
R & S & H & I & A & B & M & N \\
T & U & J & K & C & D & O & P \\
H & I & R & S & M & N & A & B \\
J & K & T & U & O & P & C & D
\end{matrix}$$

Let us call the array $\begin{vmatrix} AB \\ CD \end{vmatrix}$ a, the array $\begin{vmatrix} MN \\ OP \end{vmatrix}$ m, the array $\begin{vmatrix} RS \\ TU \end{vmatrix}$ r, and the array $\begin{vmatrix} HI \\ JK \end{vmatrix}$ h. Let us call the array $\begin{vmatrix} am \\ ma \end{vmatrix}$ w, and the array $\begin{vmatrix} rh \\ hr \end{vmatrix}$ x. Then the entire array is simply $\begin{vmatrix} wx \\ xw \end{vmatrix}$. While the original structure consisted of 64 symbols, it requires only 35 to write down its description:

$$S = \begin{matrix} wx \\ xw \end{matrix}$$

$$w = \begin{matrix} am \\ ma \end{matrix} \qquad x = \begin{matrix} rh \\ hr \end{matrix}$$

$$a = \begin{matrix} AB \\ CD \end{matrix} \qquad m = \begin{matrix} MN \\ OP \end{matrix} \qquad r = \begin{matrix} RS \\ TU \end{matrix} \qquad h = \begin{matrix} HI \\ JK \end{matrix}$$

We achieve the abbreviation by making use of the redundancy in the original structure. Since the pattern $\begin{vmatrix} AB \\ CD \end{vmatrix}$, for example, occurs four times in the total pattern, it is economical to represent it by the single symbol, a.

If a complex structure is completely unredundant—if no aspect of its structure can be inferred from any other—then it is its own simplest description. We can exhibit it, but we cannot describe it by a simpler structure. The hierarchic structures we have been discussing have a high degree of redundancy, hence can often be described in economical terms. The redundancy takes a number of forms, of which I shall mention three:

1. Hierarchic systems are usually composed of only a few different kinds of subsystems in various combinations and arrangements. A familiar example is the proteins, their multitudinous variety arising from arrangements of only twenty different amino acids. Similarly the ninety-odd elements provide all the kinds of building blocks needed for an infinite variety of molecules. Hence we can construct our description from a restricted alphabet of elementary terms corresponding to the basic set of elementary subsystems from which the complex system is generated.
2. Hierarchic systems are, as we have seen, often nearly decomposable. Hence only aggregative properties of their parts enter into the description of the interactions of those parts. A generalization of the notion of near decomposability might be called the "empty world hypothesis"—most things are only weakly connected with most other things; for a tolerable description of reality only a tiny fraction of all possible interactions needs to be taken into account. By adopting a descriptive language that allows the absence of something to go unmentioned, a nearly empty world can be described quite concisely. Mother Hubbard did not have to check off the list of possible contents to say that her cupboard was bare.
3. By appropriate "recoding," the redundancy that is present but unobvious in the structure of a complex system can often be made patent. The commonest recoding of descriptions of dynamic systems consists in replacing a description of the time path with a description of a differential law that generates that path. The simplicity resides in a constant relation between the state of the system at any given time and the state of the system a short time later. Thus the structure of the sequence 1 3 5 7 9 11 . . . is most simply expressed by observing that each member is obtained by adding 2 to the previous one. But this is the sequence that Galileo found to describe the velocity at the end of successive time intervals of a ball rolling down an inclined plane.

It is a familiar proposition that the task of science is to make use of the world's redundancy to describe that world simply. I shall not pursue the general methodological point here, but I shall instead take a closer look at two main types of description that seem to be available to us in seeking an understanding of complex systems. I shall call these *state description* and *process description*, respectively.

State Descriptions and Process Descriptions

"A circle is the locus of all points equidistant from a given point." "To construct a circle, rotate a compass with one arm fixed until the other arm has returned to its starting point." It is implicit in Euclid that if you carry out the process specified in the second sentence, you will produce an object that satisfies the definition of the first. The first sentence is a state description of a circle; the second, a process description.

These two modes of apprehending structures are the warp and weft of our experience. Pictures, blueprints, most diagrams, and chemical structural formulas are state descriptions. Recipes, differential equations, and equations for chemical reactions are process descriptions. The former characterize the world as sensed; they provide the criteria for identifying objects, often by modeling the objects themselves. The latter characterize the world as acted upon; they provide the means for producing or generating objects having the desired characteristics.

The distinction between the world as sensed and the world as acted upon defines the basic condition for the survival of adaptive organisms. The organism must develop correlations between goals in the sensed world and actions in the world of process. When they are made conscious and verbalized, these correlations correspond to what we usually call means-ends analysis. Given a desired state of affairs and an existing state of affairs, the task of an adaptive organism is to find the difference between these two states and then to find the correlating process that will erase the difference.[22]

Thus problem solving requires continual translation between the state and process descriptions of the same complex reality. Plato, in the *Meno*,

argued that all learning is remembering. He could not otherwise explain how we can discover or recognize the answer to a problem unless we already know the answer.[23] Our dual relation to the world is the source and solution of the paradox. We pose a problem by giving the state description of the solution. The task is to discover a sequence of processes that will produce the goal state from an initial state. Translation from the process description to the state description enables us to recognize when we have succeeded. The solution is genuinely new to us—and we do not need Plato's theory of remembering to explain how we recognize it.

There is now a growing body of evidence that the activity called human problem solving is basically a form of means-ends analysis that aims at discovering a process description of the path that leads to a desired goal. The general paradigm is: Given a blueprint, to find the corresponding recipe. Much of the activity of science is an application of that paradigm: Given the description of some natural phenomena, to find the differential equations for processes that will produce the phenomena.

The Description of Complexity in Self-Reproducing Systems

The problem of finding relatively simple descriptions for complex systems is of interest not only for an understanding of human knowledge of the world but also for an explanation of how a complex system can reproduce itself. In my discussion of the evolution of complex systems, I touched only briefly on the role of self-reproduction.

Atoms of high atomic weight and complex inorganic molecules are witnesses to the fact that the evolution of complexity does not imply self-reproduction. If evolution of complexity from simplicity is sufficiently probable, it will occur repeatedly; the statistical equilibrium of the system will find a large fraction of the elementary particles participating in complex systems.

If, however, the existence of a particular complex form increased the probability of the creation of another form just like it, the equilibrium between complexes and components could be greatly altered in favor of the former. If we have a description of an object that is sufficiently clear

---

[22] See H. A. Simon and A. Newell, "Simulation of Human Thinking," in M. Greenberger (ed.), *Management and the Computer of the Future* (New York: Wiley, 1962), pp. 95–114, esp. pp. 110 ff.

[23] *The Works of Plato*, B. Jowett, translator (New York: Dial Press, 1936), vol. 3, pp. 26–35. See H. A. Simon, "Bradie on Polanyi on the Meno Paradox," *Philosophy of Science*, 43(1976):147–150.

and complete, we can reproduce the object from the description. Whatever the exact mechanism of reproduction, the description provides us with the necessary information.

Now we have seen that the descriptions of complex systems can take many forms. In particular we can have state descriptions, or we can have process descriptions—blueprints or recipes. Reproductive processes could be built around either of these sources of information. Perhaps the simplest possibility is for the complex system to serve as a description of itself—a template on which a copy can be formed. One of the most plausible current theories, for example, of the reproduction of deoxyribonucleic acid (DNA) proposes that a DNA molecule, in the form of a double helix of matching parts (each essentially a "negative" of the other), unwinds to allow each half of the helix to serve as a template on which a new matching half can form.

On the other hand, our current knowledge of how DNA controls the metabolism of the organism suggests that reproduction by template is only one of the processes involved. According to the prevailing theory, DNA serves as a template both for itself and for the related substance ribonucleic acid (RNA). RNA in turn serves as a template for protein. But proteins—according to current knowledge—guide the organism's metabolism not by the template method but by serving as catalysts to govern reaction rates in the cell. While RNA is a blueprint for protein, protein is a recipe for metabolism.[24]

Ontogeny Recapitulates Phylogeny

The DNA in the chromosomes of an organism contains some, and perhaps most, of the information that is needed to determine its development and activity. We have seen that, if current theories are even approximately correct, the information is recorded not as a state description of the organism but as a series of "instructions" for the construction and maintenance of the organism from nutrient materials. I have already used the metaphor of a recipe; I could equally well compare it with a computer program, which is also a sequence of instructions governing the construction of symbolic structures. Let me spin out some of the consequences of the latter comparison.

If genetic material is a program—viewed in its relation to the organism—it is a program with special and peculiar properties. First, it is a self-reproducing program; we have already considered its possible copying mechanism. Second, it is a program that has developed by Darwinian evolution. On the basis of our watchmaker's argument, we may assert that many of its ancestors were also viable programs—programs for the subassemblies.

Are there any other conjectures we can make about the structure of this program? There is a well-known generalization in biology that is verbally so neat that we would be reluctant to give it up even if the facts did not support it: ontogeny recapitulates phylogeny. The individual organism in its development goes through stages that resemble some of its ancestral forms. The fact that the human embryo develops gill bars and then modifies them for other purposes is a familiar particular belonging to the generalization. Biologists today like to emphasize the qualifications of the principle—that ontogeny recapitulates only the grossest aspects of phylogeny, and these only crudely. These qualifications should not make us lose sight of the fact that the generalization does hold in rough approximation—it does summarize a very significant set of facts about the organism's development. How can we interpret these facts?

One way to solve a complex problem is to reduce it to a problem previously solved—to show what steps lead from the earlier solution to a solution of the new problem. If around the turn of the century we wanted to instruct a workman to make an automobile, perhaps the simplest way would have been to tell him how to modify a wagon by removing the singletree and adding a motor and transmission. Similarly a genetic program could be altered in the course of evolution by adding new processes that would modify a simpler form into a more complex one—to construct a gastrula, take a blastula and alter it!

The genetic description of a single cell may therefore take a quite different form from the genetic description that assembles cells into a multicelled organism. Multiplication by cell division would require as a

---

[24] C. B. Anfinsen, *The Molecular Basis of Evolution* (New York: Wiley, 1959), chapters 3 and 10, will qualify this sketchy, oversimplified account. For an imaginative discussion of some mechanisms of process description that could govern molecular structure, see H. H. Pattee, "On the Origin of Macromolecular Sequences," *Biophysical Journal*, 1 (1961):683–710.

minimum a state description (the DNA, say), and a simple "interpretive process"—to use the term from computer language—that copies this description as a part of the larger copying process of cell division. But such a mechanism clearly would not suffice for the differentiation of cells in development. It appears more natural to conceptualize that mechanism as based on a process description and a somewhat more complex interpretive process that produces the adult organism in a sequence of stages, each new stage in development representing the effect of an operator upon the previous one.

It is harder to conceptualize the interrelation of these two descriptions. Interrelated they must be, for enough has been learned of gene-enzyme mechanisms to show that these play a major role in development as in cell metabolism. The single clue we obtain from our earlier discussion is that the description may itself be hierarchical, or nearly decomposable, in structure, the lower levels governing the fast, "high-frequency" dynamics of the individual cell and the higher-level interactions governing the slow, "low-frequency" dynamics of the developing multicellular organism.

There is a rapidly growing body of evidence that the genetic program is organized in this way.[25] To the extent that we can differentiate the genetic information that governs cell metabolism from the genetic information that governs the development of differentiated cells in the multicellular organization, we simplify enormously—as we have already seen—our

task of theoretical description. But I have perhaps pressed this speculation far enough.

The generalization that we might expect ontogeny partially to recapitulate phylogeny in evolving systems whose descriptions are stored in a process language has applications outside the realm of biology. It can be applied as readily, for example, to the transmission of knowledge in the educational process. In most subjects, particularly in the rapidly advancing sciences, the progress from elementary to advanced courses is to a considerable extent a progress through the conceptual history of the science itself. Fortunately the recapitulation is seldom literal—any more than it is in the biological case. We do not teach the phlogiston theory in chemistry in order later to correct it. (I am not sure I could not cite examples in other subjects where we do exactly that.) But curriculum revisions that rid us of the accumulations of the past are infrequent and painful. Nor are they always desirable—partial recapitulation may, in many instances, provide the most expeditious route to advanced knowledge.

Summary: The Description of Complexity

How complex or simple a structure is depends critically upon the way in which we describe it. Most of the complex structures found in the world are enormously redundant, and we can use this redundancy to simplify their description. But to use it, to achieve the simplification, we must find the right representation.

The notion of substituting a process description for a state description of nature has played a central role in the development of modern science. Dynamic laws, expressed in the form of systems of differential or difference equations, have in a large number of cases provided the clue for the simple description of the complex. In the preceding paragraphs I have tried to show that this characteristic of scientific inquiry is not accidental or superficial. The correlation between state description and process description is basic to the functioning of any adaptive organism, to its capacity for acting purposefully upon its environment. Our present-day understanding of genetic mechanisms suggests that even in describing itself the multicellular organism finds a process description—a genetically encoded program—to be the parsimonious and useful representation.

---

[25] For extensive discussion of these matters, see J. D. Watson, *op. cit.*, especially chapters 8 and 14. For a review of some of the early evidence, see P. E. Hartman, "Transduction: A Comparative Review," in W. D. McElroy and B. Glass (eds.), *The Chemical Basis of Heredity* (Baltimore: Johns Hopkins Press, 1957), pp. 442–454. Evidence for differential activity of genes in different tissues and at different stages of development is discussed by J. G. Gall, "Chromosomal Differentiation," in W. D. McElroy and B. Glass (eds.), *The Chemical Basis of Development* (Baltimore: Johns Hopkins Press, 1958), pp. 103–135. Finally, a model very like that proposed here has been independently, and far more fully, outlined by J. R. Platt, "A 'Book Model' of Genetic Information Transfer in Cells and Tissues," in M. Kasha and B. Pullman (eds.), *Horizons in Biochemistry* (New York: Academic Press, 1962), pp. 167–187. Of course this kind of mechanism is not the only one in which development could be controlled by a process description. Induction, in the form envisaged in Spemann's organizer theory, is based on a process description in which metabolites in already formed tissue control the next stages of development.

Conclusion

Our speculations have carried us over a rather alarming array of topics, but that is the price we must pay if we wish to seek properties common to many sorts of complex systems. My thesis has been that one path to the construction of a nontrivial theory of complex systems is by way of a theory of hierarchy. Empirically a large proportion of the complex systems we observe in nature exhibit hierarchic structure. On theoretical grounds we could expect complex systems to be hierarchies in a world in which complexity had to evolve from simplicity. In their dynamics hierarchies have a property, near decomposability, that greatly simplifies their behavior. Near decomposability also simplifies the description of a complex system and makes it easier to understand how the information needed for the development or reproduction of the system can be stored in reasonable compass.

In science and engineering the study of "systems" is an increasingly popular activity. Its popularity is more a response to a pressing need for synthesizing and analyzing complexity than it is to any large development of a body of knowledge and technique for dealing with complexity. If this popularity is to be more than a fad, necessity will have to mother invention and provide substance to go with the name. The explorations reviewed here represent one particular direction of search for such substance.

Name Index