

This excerpt from

Adaptation in Natural and Artificial Systems.

John H. Holland.

© 1992 The MIT Press.

is provided in screen-viewable form for personal use only by members of MIT CogNet.

Unauthorized use or dissemination of this information is expressly forbidden.

If you have any questions about this material, please contact
cognetadmin@cognet.mit.edu.

1. The General Setting

1. INTRODUCTION

How does evolution produce increasingly fit organisms in environments which are highly uncertain for individual organisms?

What kinds of economic plan can upgrade an economy's performance in spite of the fact that relevant economic data and utility measures must be obtained as the economy develops?

How does an organism use its experience to modify its behavior in beneficial ways (i.e., how does it learn or "adapt under sensory guidance")?

How can computers be programmed so that problem-solving capabilities are built up by specifying "*what* is to be done" rather than "*how* to do it" ?

What control procedures can improve the efficiency of an ongoing process, when details of changing component interactions must be compiled and used concurrently?

Though these questions come from very different areas, it is striking how much they have in common. Each involves a problem of optimization made difficult by substantial complexity and uncertainty. The complexity makes discovery of the optimum a long, perhaps never-to-be-completed task, so the best among *tested* options must be exploited at every step. At the same time, the uncertainties must be reduced rapidly, so that knowledge of *available* options increases rapidly. More succinctly, information must be exploited as acquired so that performance improves apace. Problems with these characteristics are even more pervasive than the questions above would indicate. They occur at critical points in fields as diverse as evolution, ecology, psychology, economic planning, control, artificial intelligence, computational mathematics, sampling, and inference.

There is no collective name for such problems, but whenever the term *adaptation* (*ad* + *aptare*, to fit to) appears it consistently singles out the problems of interest. In this book the meaning of "adaptation" will be extended to encom-

pass the entire collection of problems. (Among rigorous studies of adaptation, Tsyarkin's [1971] usage comes closest to this in breadth, but he deliberately focuses on the man-made systems.) This extension, if taken seriously, entails a commitment to view adaptation as a fundamental process, appearing in a variety of guises but subject to unified study. Even at the outset there is a powerful warrant for this view. It comes from the observation that all variations of the problem give rise to the same fundamental questions.

- To what parts of its environment is the organism (system, organization) adapting?
- How does the environment act upon the adapting organism (system, organization)?
- What structures are undergoing adaptation?
- What are the mechanisms of adaptation?
- What part of the history of its interaction with the environment does the organism (system, organization) retain?
- What limits are there to the adaptive process?
- How are different (hypotheses about) adaptive processes to be compared?

Moreover, as we attempt to answer these questions in different contexts, essentially the same obstacles to adaptation appear again and again. They appear with different guises and names, but they have the same basic structure. For example, "nonlinearity," "false peak," and "epistatic effect" all designate versions of the same difficulty. In the next section we will look more closely at these obstacles; for now let it be noted that the study of adaptation is deeply concerned with the means of overcoming these obstacles.

Despite a wealth of data from many different fields and despite many insights, we are still a long way from a general understanding of adaptive mechanisms. The situation is much like that in the old tale of blind men examining an elephant—different aspects of adaptation acquire different emphases because of the points of contact. A specific feature will be prominent in one study, obscure in another. Useful and suggestive results remain in comparative isolation. Under such circumstances theory can be a powerful aid. Successful analysis separates incidental or "local" exaggerations from fundamental features. A broadly conceived analytic theory brings data and explanation into a coherent whole, providing opportunities for prediction and control. Indeed there is an important sense in which a good theory defines the objects with which it deals. It reveals their interactions, the methods of transforming and controlling them, and predictions of what will happen to them.

Theory will have a central role in all that follows, but only insofar as it illuminates practice. For natural systems, this means that theory must provide

techniques for prediction and control; for artificial systems, it must provide practical algorithms and strategies. Theory should help us to know more of the mechanisms of adaptation and of the conditions under which new adaptations arise. It should enable us to better understand the processes whereby an initially unorganized system acquires increasing self-control in complex environments. It should suggest procedures whereby actions acquired in one set of circumstances can be transferred to new circumstances. In short, theory should provide us with means of prediction and control not directly suggested by compilations of data or simple tinkering. The development here will be guided accordingly.

The fundamental questions listed above can serve as a starting point for a unified theory of adaptation, but the informal phrasing is a source of difficulty. With the given phrasing it is difficult to conceive of answers which would apply *unambiguously* to the full range of problems. Our first task, then, is to rephrase questions in a way which avoids ambiguity and encourages generality. We can avoid ambiguity by giving precise definitions to the terms appearing in the questions, and we can assure the desired generality if the terms are defined by embedding them in a common formal framework. Working within such a framework we can proceed with theoretical constructions which are of real help in answering the questions. This, in broad outline, is the approach we will take.

2. PRELIMINARY SURVEY

(Since we are operating outside of a formal framework in this chapter, some of the statements which follow will be susceptible of different, possibly conflicting interpretations. Precise versions will be formulated later.)

Just what are adaptation's salient features? We can see at once that adaptation, whatever its context, involves a progressive modification of some structure or structures. These structures constitute the grist of the adaptive process, being largely determined by the field of study. Careful observation of successive structural modifications generally reveals a basic set of structural modifiers or operators; repeated action of these operators yields the observed modification sequences. Table 1 presents a list of some typical structures along with the associated operators for several fields of interest.

A system undergoing adaptation is largely characterized by the mixture of operators acting on the structures at each stage. The set of factors controlling this changing mixture—the adaptive plan—constitutes the works of the system as far as its adaptive character is concerned. The adaptive plan determines just what structures arise in response to the environment, and the set of structures attainable

Table 1: Typical Structures and Operators

<i>Field</i>	<i>Structures</i>	<i>Operators</i>
Genetics	chromosomes	mutation, recombination, etc.
Economic planning	mixes of goods	production activities
Control	policies	Bayes's rule, successive approximation, etc.
Physiological psychology	cell assemblies	synapse modification
Game theory	strategies	rules for iterative approximation of optimal strategy
Artificial intelligence	programs	"learning" rules

by applying all possible operator sequences marks out the limits of the adaptive plan's domain of action. Since a given structure performs differently in different environments—the structure is more or less fit—it is the adaptive plan's task to produce structures which perform "well" (are fit) in the environment confronting it. "Adaptations" to the environment are persistent properties of the sequence of structures generated by the adaptive plan.

A precise statement of the adaptive plan's task serves as a key to uniform treatment. Three major components are associated in the task statement: (1) the environment, E , of the system undergoing adaptation, (2) the adaptive plan, τ , whereby the system's structure is modified to effect improvements, (3) a measure, μ , of performance, i.e., the fitness of the structures for the environment. (The formal framework developed in chapter 2 is built around these three components.) The crux of the problem for the plan τ is that initially it has incomplete information about which structures are most fit. To reduce this uncertainty the plan must test the performance of different structures in the environment. The "adaptiveness" of the plan enters when different environments cause different sequences of structures to be generated and tested.

In more detail and somewhat more formally: A characteristic of the environment can be unknown (from the adaptive plan's point of view) only if alternative outcomes of the plan's tests are allowed for. Each distinct combination of alternatives is a distinct environment E in which the plan may have to act. The set of all possible combinations of alternatives indicates the plan's initial uncertainty about the environment confronting it—the range of environments in which the plan should be able to act. This initial uncertainty about the environment will be formalized by designating a class \mathcal{E} of possible environments. The domain of action

of the adaptive plan will be formalized by designating a set α of attainable structures. The fact that different $E \in \mathcal{E}$ in general elicit different performances from a given structure $A \in \alpha$ means formally that there will be a different performance measure μ_E associated with each E . Each field of study is typified as much by its performance measures as by its structures and operators. For the fields mentioned in connection with examples of structures and operators, we have a corresponding list of performance measures:

Table 2: Typical Performance Measures

<i>Field</i>	<i>Performance Measure</i>
Genetics	Fitness
Economic planning	Utility
Control	Error functions
Physiological psychology	Performance rate (in some contexts, but often unspecified)
Game theory	Payoff
Artificial intelligence	Comparative efficiency (if specified at all)

The successive structural modifications dictated by a plan τ amount to a sequence or trajectory through the set α . For the plan to be adaptive the trajectory through α must depend upon which environment $E \in \mathcal{E}$ is present. Symbolizing the set of operators by Ω , this last can be stated another way by saying that the order of application of operators from Ω must depend upon E .

It is clear that the organization of α , the effects of the operators Ω upon structures in α , and the form of the performance measure μ_E all affect the difficulty of adaptation. Among the specific obstacles confronting an adaptive plan are the following:

1. α is large so that there are many alternatives to be tested.
2. The structures in α are complicated so that it is difficult to determine which substructures or components (if any) are responsible for good performance.
3. The performance measure μ_E is a complicated function with many interdependent parameters (e.g., it has many dimensions and is non-linear, exhibiting local optima, discontinuities, etc.).

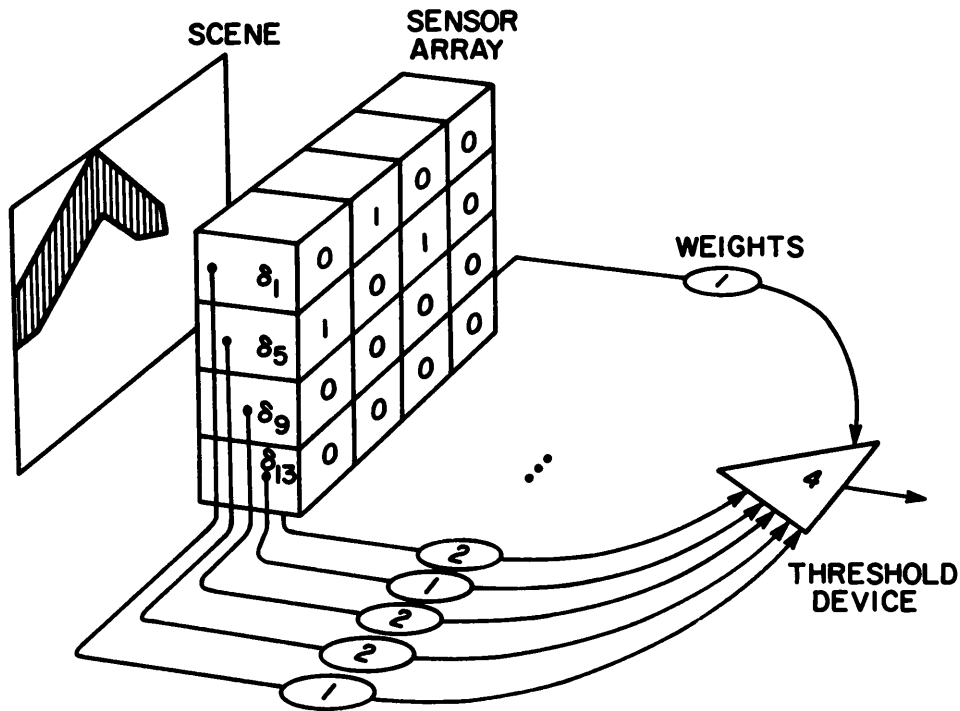
4. The performance measure varies over time and space so that given adaptations are only advantageous at certain places and times.
5. The environment E presents to τ a great flux of information (including performances) which must be filtered and sorted for relevance.

By describing these obstacles, and the adaptive plans meant to overcome them, within a general framework, we open the possibility of discovering plans useful in any situation requiring adaptation.

Before going further let us flesh out these abstractions by using them in the description of two distinct adaptive systems, one simple and artificial, the other complex and natural.

3. A SIMPLE ARTIFICIAL ADAPTIVE SYSTEM

The artificial adaptive system of this example is a pattern recognition device. (The device to be described has very limited capabilities; while this is important in applications, it does not detract from the device's usefulness as an illustration.) The information to be fed to the adaptive device is preprocessed by a rectangular array of sensors, a units high by b units wide. Each sensor is a threshold device which is activated when the light falling upon it exceeds a fixed threshold. Thus, when a "scene" is presented to the sensor array at some time t , each individual sensor is either "on" or "off" depending upon the amount of light reaching it. Let the activity of the i th sensor, $i = 1, 2, \dots, ab$, at time t be represented formally by the function $\delta_i(t)$, where $\delta_i(t) = 1$ if the sensor is "on" and $\delta_i(t) = 0$ if it is "off." A given scene thus gives rise to a configuration of ab "ones" and "zeros." All told there are 2^{ab} possible configurations of sensor activation; let C designate this set of possible configurations. It will be assumed that a particular subset of C_1 of C corresponds to (instances of) the pattern to be recognized. The particular subset involved, among the 2^{ab} possible, will be unknown to the adaptive device. (E.g., C_1 might consist of all configurations containing a connected X-shaped array of ones, or it might consist of all configurations containing as many ones as zeros, or it might be any one of the other 2^{ab} possible subsets of C .) This very large set of possibilities constitutes the class of possible environments \mathcal{E} ; it is the set of alternatives the adaptive plan must be prepared to handle. The adaptive device's task is to discover or "learn" which element of \mathcal{E} is in force by learning what configurations belong to C_1 . Then, when an arbitrary configuration is presented, the device can reliably indicate whether the configuration belongs to C_1 , thereby detecting an instance of the pattern.



The scene shown is classified as C^+ because $\sum_{i=1}^{10} w_i \delta_i(t) = \delta_1(t) + 2\delta_2(t) + 2\delta_3(t) + \delta_4(t) + 2\delta_5(t) + 4\delta_6(t) + \dots > 4$.

Fig. 1. A simple pattern recognizer

The particular pattern recognition device considered here—a linear threshold device—processes the input signals $\delta_i(t)$ by first multiplying each one by some weight w_i and then summing them to yield $\sum_{i=1}^n w_i \delta_i(t)$. When this sum exceeds a given fixed threshold K the input configuration will be said to be a member of the set C^+ , otherwise a member of the set C^- . (It should be clear that $C^+ \cup C^- = C$ and that $C^+ \cap C^-$ is empty, so that the linear threshold device partitions C into two classes.) More precisely C^+ is supposed to be an approximation to C_1 , so that when the sum exceeds the fixed threshold K , the device indicates (rightly or wrongly) that the input configuration is an instance of the pattern. The object of the adaptive plan, then, is to discover as rapidly as possible a set of

weights for which the partition (C^+, C^-) approximates the partition (C_1, C_0) , so that $C^+ \cong C_1$ and $C^- \cong C_0$. (This device, as noted earlier, is quite limited; there are many partitions (C_1, C_0) that can only be poorly approximated by (C^+, C^-) , no matter what set of weights is chosen.) Now, let $W = \{w_1, w_2, \dots, w_k\}$ be the set of possible values for the weights w_i ; that is, each $w_i \in W$, $i = 1, \dots, ab$. Thus, with a fixed threshold K , the set of attainable structures \mathcal{A} is the set of all ab -tuples, W^{ab} .

The natural performance measure, μ_E , relative to any particular partition $E \in \mathcal{E}$ is the proportion of all configurations correctly assigned (to C_1 and C_0). That is, μ_E maps each ab -tuple into the fraction of correct recognitions achieved thereby, a number in the interval $[0, 1]$, $\mu_E: W^{ab} \rightarrow [0, 1]$. (In this example the outcome of *each* test—"configuration correctly classified" or "configuration incorrectly classified"—will be treated as the plan's input. The same ab -tuple may have to be tested repeatedly to establish an estimate of its performance.)

A simple plan τ_0 for discovering the best set of weights in W^{ab} is to try various ab -tuples, either in some predetermined order or at random, estimating the performance of each in its turn; the best ab -tuple encountered up to a given point in time is saved for comparison with later trials—this "best-to-date" ab -tuple being replaced immediately by any better ab -tuple encountered in a later trial. It should be clear that this procedure must eventually uncover the "best" ab -tuple in W^{ab} . But note that even for $k = 10$ and $a = b = 10$, W^{ab} has 10^{100} elements. This is a poor augury for any plan which must exhaustively search W^{ab} . And that is exactly what the plan just described must undertake, since the outcome of earlier tests in no way affects the ordering of later tests.

Let's look at a (fairly standard) plan τ_{00} which *does* use the outcome of each test to help determine the next structure for testing. The basic idea of this plan is to change the weights whenever a presentation is misassigned so as to decrease the likelihood of similar misassignments in the future. In detail: Let the values in W be ordered in increasing magnitude so that $w_{j+1} > w_j$, $j = 1, 2, \dots, k-1$ (for instance, the weights might be located at uniform intervals so that $w_{j+1} = w_j + \Delta$). Then the algorithm proceeds according to the following prescription:

1. If the presentation at time t is assigned to C_0 when it should have been assigned to C_1 then, for each i such that $\delta_i(t) = 1$, replace the corresponding weight by the next highest weight (in the case of uniform intervals the new weight would be the old weight w_i increased by Δ , $w_i + \Delta$). Leave the other weights unchanged.

2. If the presentation at time t is assigned to C_1 instead of C_0 then, for each i such that $\delta_i(t) = 1$, replace the corresponding weight by the next lowest weight (for uniform intervals, the new weight is $w_i - \Delta$).

We cannot yet fruitfully discuss the merits of this plan in comparison to alternatives; we can only note that the order in which τ_{00} tests the elements of α does indeed depend upon the information it receives. That is, the trajectory through $\alpha = W^{ab}$ is conditional on the outcomes $\mu_B(A)$, $A \in \alpha$, of prior tests.

4. A COMPLEX NATURAL ADAPTIVE SYSTEM

Here we will look at biological adaptation via changes in genetic makeup—the first of a series of progressively more detailed examinations. This section will present only biological facts directly relevant to adaptation, with a caveat to the reader about the dangers of unintentional emphasis and oversimplification inherent in such a partial picture.

It is a familiar fact (but one we will delve into later) that every organism is an amalgam of characteristics determined by the genes in its chromosomes. Each gene has several forms or alternatives—*alleles*—producing differences in the set of characteristics associated with that gene. (E.g., certain strains of garden pea have a single gene which determines blossom color, one allele causing the blossom to be white, the other pink; bread mold has a gene which in normal form causes synthesis of vitamin B₁, but several mutant alleles of the gene are deficient in this ability; human sickle cell anemia results from an abnormal allele of one of the genes determining the structure of hemoglobin—interestingly enough, in environments where malaria is endemic, the abnormal allele can confer an advantage.) There are tens of thousands of genes in the chromosomes of a typical vertebrate, each of which (on the evidence available) has several alleles. Taking the set of attainable structures α to be the set of chromosomes obtained by making all possible combinations of alleles, we see that α contains on the order of $2^{10,000} \cong 10^{3000}$ structures for a typical vertebrate species (assuming 2 alleles for each of 10,000 genes). Even a very large population, say 10 billion individuals of that species, contains only a minuscule fraction of the possibilities.

The enormous number of possible genetic structures—*genotypes*—for a single vertebrate species is an indicator of the complexity of such systems, but it is only an indicator. The basic complexity of these systems comes from the interactions of the genes. To see just how extensive these interactions are, it is worth

looking briefly at some of the related biochemistry. Without going into detail, different alleles of the same gene produce related proteins, which in turn produce the variations in expressed characteristics associated with that gene. Typically these proteins (or combinations of them) are powerful biological catalysts called *enzymes*, capable of modifying reaction rates by factors of 10,000 and more. For this reason, genes exercise extensive control over the ongoing reactions in a cell—the enzymes they produce modulate ongoing reactions so strongly that they are the major determinants of the cell's form. Moreover, the products of any given enzyme-controlled reaction may, and generally do, enter into several subsequent reactions. Thus the effects of changes in a single enzyme are often widespread, causing gross changes in cell form and function. The human hereditary disorder called phenylketonuria results from an (undesirable) allele of a single gene; the presence of this allele has pronounced effects upon a whole battery of characteristics ranging from hair color and head size through intelligence. It is equally true that several genes may jointly determine a given characteristic, e.g., eye color in humans.

All of this adds considerably to the complexity of the system, but the greatest complexities come about because the effects of different enzymes are not additive—a phenomenon known as *epistasis*. For example, if a sequence of reactions depends upon several enzymes, for practical purposes the sequence does not proceed at all until all of the enzymes are present; subtraction of one enzyme stops the reaction completely. More complicated reactions involving positive and negative feedback are common, particularly those in which the output of a reaction sequence is a catalyst or inhibitor for some intermediate step of the reaction. The main point is that the effect of each allele depends strongly upon what other alleles are present and small changes can often produce large effects. The amalgam of observed characteristics—the *phenotype*—depends strongly upon these epistatic effects.

Because of epistasis there is no simple way to apportion credit to individual alleles for the performance of the resulting phenotype. What may be a good allele when coordinated with an appropriate set of alleles for other genes, can be disastrous in a different genetic context. Thus adaptation cannot be accomplished by selecting among the alleles for one gene independently of what alleles appear for other genes. The problem is like the problem of adjusting the “height,” “vertical linearity,” and “vertical hold” controls on a television set. A “best setting” for “height,” ignoring the settings of the other two controls, will be destroyed as soon as one attempts to better the setting of either of the other two controls. The problem is vexing enough when there are three interdependent controls, as anyone who

has attempted these adjustments can testify, but it pales in comparison to the genetic case where dozens or hundreds of interdependent alleles can be involved. Roughly, the difficulty of the problem increases by an order of magnitude for each additional gene when the interdependencies are intricate (but see the discussions in chapter 4 and pp. 160–61).

Given the pervasiveness of epistasis, adaptation via changes in genetic makeup becomes primarily a search for *coadapted* sets of alleles—alleles of different genes which together significantly augment the performance of the corresponding phenotype. (In chapter 4 the concept of a coadapted set of alleles will be generalized, under the term *schema*, to the point where it applies to the full range of adaptive systems.) It should be clear that coadaptation depends strongly upon the environment of the phenotype. The large coadapted set of alleles which produces gills in fish augments performance only in aquatic environments. This dependence of coadaptation upon characteristics of the environment gives rise to the notion of an *environmental niche*, taken here to mean a set of features of the environment which can be exploited by an appropriate organization of the phenotype. (This is a broader interpretation than the usual one which limits niche to those environmental features particularly exploited by a given species.) Examples of environmental niches fitting this interpretation are: (i) an oxygen-poor, sulfur-rich environment such as is found at the bottom of ponds with large amounts of decaying matter—a class of anaerobic bacteria, the thiobacilli, exploits this niche by means of a complex of enzymes enabling them to use sulfur in place of oxygen to carry out oxidation; (ii) the “bee-rich” environment exploited by the orchid *Ophrys apifera* which has a flower mimicking the bee closely enough to induce pollination via attempted copulation by the male bees; (iii) the environment rich in atmospheric vibrations in the frequency range of 50 to 50,000 cycles per second—the bones of the mammalian ear are a particular adaptation of parts of the reptilian jaw which aids in the detection of these vibrations, an adaptation which clearly must be coordinated with many other adaptations, including a sophisticated information-processing network, before it can improve an organism’s chances of survival. It is important to note that quite distinct coadapted sets of alleles can exploit the same environmental niche. Thus, the eye of aquatic mammals and the (functionally similar) eye of the octopus exploit the same environmental niche, but are due to coadapted sets of alleles of entirely unrelated sets of genes.

The various environmental niches $E \in \mathcal{E}$ define different opportunities for adaptation open to the genetic system. To exploit these opportunities the genetic system must select and use the sets of coadapted alleles which produce the appropriate phenotypic characteristics. The central question for genetic systems is: How

are initially unsuited structures transformed to (an observed range of) structures suited to a variety of environmental niches \mathcal{E} ? To attempt a general answer to this question we need a well-developed formal framework. The framework available at this point is insufficient, even for a careful description of a candidate adaptive plan τ for genetic systems, unlike the case of the simpler artificial system. A fortiori, questions about such adaptive plans, and critical questions about efficiency, must wait upon further development of the framework. We *can* explore here some of the requirements an adaptive plan τ must meet if it is to be relevant to data about genetics and evolution.

In beginning this exploration we can make good use of a concept from mathematical genetics. The action of the environment $E \in \mathcal{E}$ upon the phenotype (and thereby upon the genotype $A \in \mathcal{A}$) is typically summarized in mathematical studies of genetics by a single performance measure μ_E called *fitness*. Roughly, the fitness of a phenotype is the number of its offspring which survive to reproduce (precise definitions will be given later in connection with the appropriate formal models, see section 3.1). This measure rests upon a universal, and familiar, feature of biological systems: Every individual (phenotype) exists as a member of a population of similar individuals, a population constantly in flux because of the reproduction and death of the individuals comprising it. The fitness of an individual is clearly related to its influence upon the future development of the population. When many offspring of a given individual survive to reproduce, then many members of the resulting population, the “next generation,” will carry the alleles of that individual. Genotypes and phenotypes of the next generation will be influenced accordingly.

Fitness, viewed as a measure of the genotype’s influence upon the future, introduces a concept useful through the whole spectrum of adaptation. A good way to see this concept in wider context is to view the testing of genotypes as a sampling procedure. The sample space in this case is the set of all genotypes \mathcal{A} and the outcome of each sample is the performance μ_E of the corresponding phenotype. The general question associated with fitness, then, is: To what extent does the outcome $\mu_E(A)$ of a sample $A \in \mathcal{A}$ influence or alter the sampling plan τ (the kinds of samples to be taken in the future)? Looking backward instead of forward, we encounter a closely related question: How does the history of the outcomes of previous samples influence the current sampling plan? The answers to these questions go far toward determining the basic character of any adaptive process.

We have already seen that the answer to the first question, for genetic systems, is that the future influence of each individual $A \in \mathcal{A}$ is directly proportional to the sampled performance $\mu_E(A)$. This relation need not be so in general—

there are many well-established procedures for optimization, inference, mathematical learning, etc., where the relation between sampled performance and future sampling is quite different. Nevertheless reproduction in proportion to measured performance is an important concept which can be generalized to yield sampling plans—*reproductive plans*—applicable to any adaptive problem (including the broad class of problems where there is no natural notion of reproduction). Moreover, once reproductive plans have been defined in the formal framework, it can be proved that they are efficient (in a reasonable sense) over a very broad range of conditions.

A part of the answer to the second question, for genetic systems, comes from the observation that future populations can only develop via reproduction of individuals in the current population. Whatever history is retained must be represented in the current population. In particular, the population must serve as a summary of observed sample values (performances). The population thereby has the same relation to an adaptive process that the notion of (complete) state has to the laws of physics or the transition functions of automata theory. Knowing the population structure or state enables one to determine the future without any additional information about the past of the system. (That is, different sampling sequences which arrive at the same population will have exactly the same influence on the future.) The state concept has been used as a foundation stone for formal models in a wide variety of fields; in the formal development to follow generalizations of population structure will have this role.

An understanding of the two questions just posed leads to a deeper understanding of the requirements on a genetic adaptive plan. It also leads to an apparent dilemma. On the one hand, if offspring are simple duplicates of fit members of the population, fitness is preserved but there is no provision for improvement. On the other hand, letting offspring be produced by simple random variation (a process practically identical to enumeration) yields a maximum of new variants but makes no provision for retention of advances already made. The dilemma is sharpened by two biological facts: (1) In biological populations consisting of advanced organisms (say vertebrates) no two individuals possess identical chromosomes (barring identical twins and the like). This is so even if we look over many (all) successive generations. (2) In realistic cases, the overwhelming proportion of *possible* variants (all possible allele combinations, not just those observed) are incapable of surviving to produce offspring in the environments encountered. Thus, by observation (1), advances in fitness are not retained by simple duplication. At the same time, by observation (2), the observed lack of identity cannot result from simple random variation because extinction would almost certainly

follow in a single generation—variants chosen completely at random are almost certain to be sterile.

In attempting to see how this “dilemma” is resolved, we begin to encounter some of the deeper questions about adaptation. We can only hint at the dilemma’s resolution in this preliminary survey. Even a clear statement of the resolution requires a considerable formal structure, and proof that it is in fact a resolution requires still more effort. Much of the understanding hinges on posing and answering two questions closely related to the questions generated by the concept of fitness: How can an adaptive plan τ (specifically, here a plan for genetic systems) retain useful portions of its (rapidly growing) history along with advances already made? How is the adaptive plan τ to access and use its history (the portion stored) to increase the likelihood of fit variants ($A \in \mathcal{A}$ such that $\mu_E(A)$ is above average)? Once again these are questions relevant to the whole spectrum of fields mentioned at the outset.

The resolution of the dilemma lies in the action of the genetic operators Ω within the reproductive plan τ . The best-known genetic operators exhibit two properties strongly affecting this action: (1) The operators do not directly affect the size of the population—their main effect is to alter and redistribute alleles within the population. (The alleles in an individual typically come from more than one source in the previous generation, the result, for example, of the mating of parents in the case of vertebrates, or of transduction in the case of bacteria.) (2) The operators infrequently separate alleles which are close together on a chromosome. That is, alleles close together typically remain close together after the operators have acted.

Useful clues to the dilemma’s resolution emerge when we look at the effect of these operators in a simple reproductive plan, τ_1 . This plan can be thought of as unfolding through repeated application of a two-phase procedure: During phase one, additional copies of (some) individuals exhibiting above-average performance are added to the population while (some) individuals of subaverage performance are deleted. More carefully, each individual has an expected number of offspring, or rate of reproduction, proportional to its performance. (If the population is to be constant in size, the rates of reproduction must be “normalized” so that their average over the population at any time is 1.) During phase two, the genetic operators in Ω are applied, interchanging and modifying sets of alleles in the chromosomes of different individuals, so that the offspring are no longer identical to their progenitors. The result is a new, modified population. The process is iterated to produce successive generations of variants.

More formally, in an environment which assigns an observable performance

to each individual, τ_1 acts as follows: At the beginning of each time period t , the plan's accumulated information about the environment resides in a finite population $\alpha(t)$ selected from α . The most important part of this information is given by the discrete distributions which give the proportions of different sets of alleles in the population $\alpha(t)$. $\alpha(t)$ serves not only as the plan's repository of accumulated information, but also as the source of new variants which will give rise to $\alpha(t + 1)$. As indicated earlier, the formation of $\alpha(t + 1)$ proceeds in two phases. During the first phase, $\alpha(t)$ is modified to form $\alpha'(t)$ by copying each individual in $\alpha(t)$ a number of times dependent upon the individual's observed performance. The number of copies made will be determined stochastically so that the expected number of copies increases in proportion to observed performance. During the second phase, the operators are applied to the population $\alpha'(t)$, interchanging and modifying the sets of alleles, to produce the new generation $\alpha(t + 1)$.

One key to understanding τ_1 's resolution of the dilemma lies in observing what happens to small sets of adjacent alleles under its action. In particular, what happens if an adjacent set of alleles appears in several different chromosomes of above-average fitness and not elsewhere? Because each of the chromosomes will be duplicated an above-average number of times, the given alleles will occupy an increased proportion of the population after the duplication phase. This increased proportion will of course result whether or not the alleles had anything to do with the above-average fitness. The appearance of the alleles in the extra-fit chromosomes might be happenstance, but it is equally true that any correlation between the given selection of alleles and above-average fitness will be exploited by this action. Moreover, the more varied the chromosomes containing the alleles, the less likely it is that the alleles and above-average fitness are uncorrelated.

What happens now when the genetic operators Ω are applied to form the next generation? As indicated earlier, the closer alleles are to one another in the chromosome the less likely they are to be separated during the operator phase. Thus the operator phase usually transfers adjacent sets of genes as unit, placing them in new chromosomal contexts without disturbing them otherwise. These new contexts further test the sets of alleles for correlation with above-average fitness. If the selected set of alleles does indeed augment fitness, the chromosomes containing the set will again (on the average) be extra fit. On the other hand, if the prior associations were simply happenstance, sustained association with extra-fit chromosomes becomes increasingly less likely as the number of trials (new contexts) increases. The net effect of the genetic plan over several generations will be an increasing predominance of alleles and sets of alleles augmenting fitness in the given environment.

In observing what happens to small sets of genes under its action, we have seen one way in which the plan τ_1 preserves the history of its interactions with the environment. It also retains certain kinds of advances thereby, favoring structural components which have proved their worth by augmenting fitness. At the same time, since these components are continually tried in new contexts and combinations, stagnation is avoided. In brief, sets of alleles engendering above-average performance provide comparative success in reproduction for the chromosomes carrying them. This in turn assures that these alleles become predominant components of later generations of chromosomes. Though this description is sketchy, it does indicate that reproductive plans using genetic operators proceed in a way which is neither enumeration nor simple duplication of fit structures. The full story is both more intricate and more sophisticated. Because reproductive plans are provably efficient over a broad range of conditions, we will spend considerable time later unraveling the skeins of this story.

5. SOME GENERAL OBSERVATIONS

One point which comes through clearly from the examples is the enormous size of \mathcal{A} , even for a very modest system. This size has a fatal bearing on what is at first sight a candidate for a "universal" adaptive plan. The candidate, called τ_0 in the first example, and henceforth designated an *enumerative* plan, exhaustively tests the structures in \mathcal{A} . Enumerative plans are characterized by the fact that the order in which they test structures is unaffected by the outcome of previous tests. For example, the plan first generates and tests all structures attainable (from an initially given structure) by single applications of the basic operators, then all structures attainable by two applications of the operators, etc. The plan preserves the fittest structure it has encountered up to any given point in the process, replacing that structure immediately upon generating a structure which is still more fit. Thus, given enough time (and enough stability of the environment so that the fitness of structures does not change during the process) an enumerative plan is guaranteed to discover the structure most fit for any environment confronting it. The simplicity of this plan, together with the guarantee of discovering the most fit structure, would seem to make it a very important adaptive plan. Indeed enumerative plans have been repeatedly proposed and studied in most of the areas mentioned in section 1.1. They are often set forth in a form not obviously enumerative, particularly in evolutionary studies (mutation in the absence of other genetic operators), learning (simple trial-and-error), and artificial intelligence (random search).

However, in all but the most constrained situations, enumerative plans are a false lead.

The flaw, and it is a fatal one, asserts itself when we begin to ask, "How long is eventually?" To get some feeling for the answer we need only look back at the first example. For that very restricted system there were 10^{100} structures in α . In most cases of real interest, the number of possible structures vastly exceeds this number, and for natural systems like the genetic systems we have already seen that numbers like $2^{10,000} \cong 10^{3000}$ arise. If 10^{12} structures could be tried every second (the fastest computers proposed to date could not even add at this rate), it would take a year to test about $3 \cdot 10^{19}$ structures, or a time vastly exceeding the estimated age of the universe to test 10^{100} structures.

It is clear that an attempt to adapt by means of an enumerative plan is foredoomed in all but the simplest cases because of the enormous times involved. This extreme inefficiency makes enumerative plans uninteresting either as hypotheses about natural processes or as algorithms for artificial systems. It follows at once that an adaptive plan cannot be considered good simply because it will eventually produce fit structures for the environments confronting it; it must do so in a reasonable time span. What a "reasonable time span" is depends strongly on the environments (problems) under consideration, but in no case will it be a time large with respect to the age of the universe. This question of efficiency or "reasonable time span" is the pivotal point of the most serious contemporary challenge to evolutionary theory: Are the known genetic operators sufficient to account for the changes observed in the allotted geological intervals? There is of course evidence for the existence of adaptive plans much more efficient than enumeration. Arthur Samuel (1959) has written a computer program which learned to play tournament calibre checkers, and humans do manage to adapt to very complex environments in times considerably less than a century. It follows that a major part of any study of the adaptive process must be the discovery of factors which provide efficiency while retaining the "universality" (robustness) of enumeration. It does not take analysis to see that an enumerative plan is inefficient just because it always generates structures in the same order, regardless of the outcome of tests on those structures. The way to improvement lies in avoiding this constraint.

The foregoing points up again the critical nature of the adaptive plan's initial uncertainty about its environment, and the central role of the procedures it uses to store and access the history of its interactions with that environment. Since different structures perform differently in different environments, the plan's task is set by the aspects of the environment which are unknown to it initially. It must

generate structures which perform well (are fit) in the particular environment confronting it, and it must do this efficiently. Interest centers on *robust* adaptive plans—plans which are efficient over the range of environments \mathcal{E} they may encounter. Giving robustness precise definition and discovering something of the factors which make an adaptive plan robust is the formal distillation of questions about efficiency. Because efficiency is critical, the study of robustness has a central place in the formal development.

The discussion of genetic systems emphasized two general requirements bearing directly on robustness: (1) The adaptive plan must retain advances already made, along with portions of the history of previous plan-environment interactions. (2) The plan must use the retained history to increase the proportion of fit structures generated as the overall history lengthens. The same discussion also indicated the potential of a particular class of adaptive plans—the reproductive plans. One of the first tasks, after setting out the formal framework, will be to provide a general definition of this class of plans. Lifting the reproductive plans from the specific genetic context makes them useful across the full spectrum of fields in which adaptation has a role. This widened role for reproductive plans can be looked upon as a first validation of the formalism. A much more substantial validation follows closely upon the definition, when the general robustness of reproductive plans is proved via the formalism. Later we will see how reproductive plans using generalized genetic operators retain and exploit their histories. Throughout the development, reproductive plans using genetic operators will serve to illuminate key features of adaptation and, in the process, we will learn more of the robustness, wide applicability, and general sophistication of such plans.

Summarizing: This entire survey has been organized around the concept of an adaptive plan. The adaptive plan, progressively modifying structure by means of suitable operators, determines what structures are produced in response to the environment. The set of operators Ω and the domain of action of the adaptive plan \mathcal{A} (i.e., the attainable structures) determine the plan's options; the plan's objective is to produce structures which perform well in the environment E confronting it. The plan's initial uncertainty about the environment—its room for improvement—is reflected in the range of environments \mathcal{E} in which it may have to act. The related performance measures μ_E , $E \in \mathcal{E}$, change from environment to environment since the same structure performs differently in different environments. These objects lie at the center of the formal framework set out in chapter 2. Chapter 3 provides illustrations of the framework as applied to genetics, economics, game-playing, searches, pattern recognition, statistical inference, control, function optimization, and the central nervous system.

A brief look at the enormous times taken by enumerative plans to discover fit structures, even when the domain of action \mathcal{A} is greatly constrained, makes it clear that efficiency is a *sine qua non* of studies of adaptation. Efficiency acts as a cutting edge, shearing away plans “too slow” to serve as hypotheses about natural systems or as algorithms for artificial systems. Whether an adaptive plan is to serve as hypothesis or algorithm, information about its robustness—its efficiency in the environments \mathcal{E} —is critical. The latter part of this book will be much concerned with this topic. Chapter 4 introduces a critical tool for the investigation and construction of efficient adaptive plans—schemata. This generalization of coadapted sets of alleles provides an efficient way of defining and exploiting properties associated with above-average performance. Chapter 5 develops a criterion for measuring the efficiency with which adaptive plans *improve* average performance and then relates this criterion to the exploitation of schemata. Chapter 6 introduces generalized genetic plans and chapter 7 establishes their robustness. Chapter 8 studies mechanisms which enable genetic plans to use predictive modeling for flexible exploitation of the large fluxes of information provided by typical environments.

The emphasis throughout the book is on general principles which help to resolve the problems and questions raised in this chapter. One particular interest will be the solution of problems involving hundreds to hundreds of thousands of interdependent parameters and multitudes of local optima—problems which largely lie outside the prescriptions of present day computational mathematics.

This excerpt from

Adaptation in Natural and Artificial Systems.

John H. Holland.

© 1992 The MIT Press.

is provided in screen-viewable form for personal use only by members of MIT CogNet.

Unauthorized use or dissemination of this information is expressly forbidden.

If you have any questions about this material, please contact
cognetadmin@cognet.mit.edu.