

FUNCTIONAL SIGNIFICANCE OF BIOLOGICAL VARIABILITY

■ MICHAEL CONRAD

Department of Computer and Communication Sciences,
The University of Michigan,
Ann Arbor, MI 48109, U.S.A.

The variability of biological matter contributes to both its adaptability and reliability. To represent the structure of this variability we treat a complete biological system (e.g. community and environment) as a system with sets of states and certain (unknown) probabilities governing the state to state transitions. Adaptability (defined operationally in terms of the maximum tolerable uncertainty of the environment) consists of behavioral uncertainty, ability to anticipate the environment, and indifference to the environment. It may also be decomposed into components associated with genetic, organismic, population, and community levels of organization. Considerations of adequate design suggest that adaptability tends to fall to its lowest allowable value in the course of evolution. This means that any change in adaptability associated with one level or unit of organization tends to be compensated by opposite changes in the adaptability associated with other levels or units, or by opposite changes in the indifference to the environment. The analysis shows that the adaptability is not independent of reliability, and that each functionally distinct state consists of : (1) finer states which mediate the processing of information about the environment; (2) redundant sets of such states; and (3) informationally equivalent states associated with macroscopically equivalent microdescriptions.

1. Introduction. The ability of biological systems to cope with the unexpected disturbances of their environment is without doubt a *sine qua non* for their stable function. This ability, which we call adaptability, is amenable to a quite general characterization, at least insofar as the uncertainty of the environment is amenable to a quite general characterization. In this paper we outline this characterization and its implications for the ramification and dissipation of disturbances in biological systems. The model is based on the idea that any biological system and its environment can be described as a system with sets of states and certain (generally unknown) transition probabilities governing the

state-to-state transitions (cf. Conrad, 1972a, 1972b, 1975, 1976). To turn this into a theory of adaptability we ask the following questions:

1. Given the statistical behavior of the environment, what is the statistical behavior of the biological system?
2. What correlates the behavior of biological system and environment, i.e. how does the statistical behavior of the former depend on its ability to process information about the latter?
3. How are the statistical processes of the biological system allocated to its various levels of organization?
4. How do the statistical characteristics at various levels of organization relate to the various modes of biological adaptability, e.g. gene pool diversity, physiological, developmental and behavioral plasticity, culturability at the population level?

Finally, we use the theory to classify the various components of biological variability on the basis of their functional significance and indicate how the various forms of variability relate to global features of biological organization. Indeed, the theory can be regarded as a theory of the functional significance of biological variability.

2. Review of the Basic Model. We suppose that both the biological system and its environment have sets of possible states and that the number of states is finite. We also suppose that if we prepare a large number of replicas of the system identically we can determine state-to-state transition probabilities. More formally, we say that the scheme of transition probabilities is given by

$$\Omega = \{p[\alpha^u(t+1), \beta^v(t+1) \mid \alpha^r(t), \beta^s(t)]\}, \quad (1)$$

where α^i is the i th state of the biological system, β^j is the j th state of the environment, u and v run over the index set of the former, r and s run over the index set of the latter, and t is the time. For simplicity we assume a discrete time scale.

We can also write partial schemes for the biological system and environment, viz.

$$\omega = \{p[\alpha^u(t+1) \mid \alpha^r(t), \beta^s(t)]\}, \quad (2a)$$

$$\omega' = \{p[\beta^v(t+1) \mid \alpha^r(t), \beta^s(t)]\}, \quad (2b)$$

where we use the prime notation to distinguish the environment scheme.

The uncertainty of the environment scheme is given by its entropy (cf. Shannon and Weaver, 1962)

$$H(\omega') = -\sum p[\alpha^r(t), \beta^s(t)] p[\beta^v(t+1) \mid \alpha^r(t), \beta^s(t)] \log p[\beta^v(t+1) \mid \alpha^r(t), \beta^s(t)], \quad (3)$$

where the sum runs over r, s, v and the entropy of the scheme is weighted by the probability of the initial states (cf. Khinchin, 1957). $H(\omega)$, the behavioral uncertainty of the biological system, is given by a precisely analogous formula.

We can also define a conditional entropy for the behavior of the biological system given the behavior of the environment. This is

$$H(\omega|\omega') = - \sum p[\alpha^r(t), \beta^s(t), \beta^v(t+1)] p[\alpha^u(t+1) | \alpha^r(t), \beta^s(t), \beta^v(t+1)] \log p[\alpha^u(t+1) | \alpha^r(t), \beta^s(t), \beta^v(t+1)], \quad (4)$$

where the sum runs over r, s, u, v . Similarly we can define $H(\omega'|\omega)$, the conditional entropy of the environment transition given the behavior of the biological system.

The following identity is well known from information theory (Khinchin, 1957)

$$H(\omega) - H(\omega|\omega') + H(\omega'|\omega) = H(\omega'). \quad (5)$$

The left hand side of the equation characterizes the uncertainties associated with the system in an environment of uncertainty $H(\omega')$. However, the adaptability is the system's potential ability to cope with environmental uncertainty. The simplest (indeed most naive) way of determining this is by determining experimentally the most uncertain environment in which the system can remain alive indefinitely. Denoting the uncertainty of this environment by $H(\hat{\omega}')$, we can rewrite the above identity as

$$H(\hat{\omega}) - H(\hat{\omega}|\hat{\omega}') + H(\hat{\omega}'|\hat{\omega}) = H(\hat{\omega}'), \quad (6)$$

where the hats indicate that the entropies are defined over the transition probabilities which characterize the system when it is operating under the greatest possible statistical stress. These can be regarded as "potential" entropies associated with the system.

By definition the condition for the system's remaining alive is

$$H(\hat{\omega}) - H(\hat{\omega}|\hat{\omega}') + H(\hat{\omega}'|\hat{\omega}) \geq H(\omega'), \quad (7)$$

where $H(\omega')$ is the actual uncertainty of the environment.

We note that our definition differs from capacity as ordinarily defined in information theory, i.e. it is not the maximum of the rate of information transfer. The latter definition is quite useful in adaptability theory, but not as experimentally tractable. Actually the original definition requires some sharpening to make it tractable from the operational point of view. Individual organisms (at least metazoans) inevitably go into a dead state and therefore their adaptability should be defined in terms of the uncertainty of the most uncertain environment which they are capable of surviving during some nonterminal stage of their life. For the population the idea of a nonliving state is reasonable since this corresponds to a state of extinction. For the entire

community the choice of nonliving states is really too strong and in practice it is better to use the idea of states of a given climax (i.e. states of an edaphic climax rather than dead states). The definition of the adaptability of organisms, populations or other units in the context of the adaptability of the community as a whole raises separate problems and is discussed below.

Since we are dealing with biological systems and since adaptability is always costly, we can go one step further and make the plausible assumption that biological systems tend to dispense with excess adaptability in the course of evolution. Thus (7) may be written

$$H(\hat{\omega}) - H(\hat{\omega}|\hat{\omega}') + H(\hat{\omega}'|\hat{\omega}) \rightarrow H(\omega'), \quad (8)$$

where the arrow expresses a tendency to an equality.

To interpret (8) note that

(i) $H(\hat{\omega})$ is the potential uncertainty in the behavior of the system, i.e. the measure of its repertoire of possible behaviors.

(ii) $H(\hat{\omega}|\hat{\omega}')$ is the potential uncertainty in the behavior of the system given the behavior of the environment. This is the measure of the extent to which the system is decorrelated from the environment, and therefore decreases as the ability of the system to anticipate the environment increases.

(iii) $H(\hat{\omega}'|\hat{\omega})$ is the potential uncertainty in the behavior of the environment given the behavior of the system. This is normally interpreted as error, but we will call it the indifference of the system to the environment (since it is sometimes a form of protection and at other times a costly sacrifice of information or resources).

Equation (8) thus says: *the potential behavioral uncertainty of the system less the potential ability to anticipate the environment tends to equal the behavioral uncertainty of the environment, but may be less if the system is indifferent to certain features of the environment* (cf. Conrad, 1972a, 1972b).

Note that the different significance of the indifference and anticipation terms comes from the fact that the system and environment are not on an equal footing (since one is alive and the other not). Anticipation of the environment is based either on the fact that the behavior of the system matches that of the environment (because of evolution or learning) or on the fact that the system states consist of finer states (considered later) which makes it possible to process information about the environment.

The particular forms which adaptability can assume will be specified in the next section. In each case it will be clear that excess adaptability would be costly. The assumption that excess adaptability tends to disappear in the course of evolution is an optimal design assumption of the type discussed by Rashevsky (1960) and also Rosen (1967). This type of assumption, which is a logical con-

sequence of the theory of evolution, plays an important role in the analysis of biological adaptability.

3. The Structured Model. Now suppose that we want to describe our biological system in terms of subsystems at various levels of organization. For example, if the system is a (biotic) community, it is convenient to describe it in terms of populations and organisms. Furthermore, each organism consists of subsystems at various levels of organization (e.g. organ, cell, genome). To keep things simple, however, we will dissect the organism into only two parts, viz. genome (by which we mean DNA sequence) and phenome (by which we mean all its other properties).

To introduce this kind of structure into the model we describe the state of the community in terms of variables characterizing each of the subsystems. In particular we denote the state of subsystem i at level j by subscripted Greek letters, $\alpha_{ij}^a, \alpha_{ij}^b, \dots$. Since we are considering four levels of organization $\alpha_{13}^a, \alpha_{13}^b, \dots$ represent states of the community; $\alpha_{i2}^a, \alpha_{i2}^b, \dots$ states of population i ; $\alpha_{i1}^a, \alpha_{i1}^b, \dots$ states of organism i ; $\alpha_{(2i-1)0}^a, \alpha_{(2i-1)0}^b, \dots$ states of the genome of organism i ; and $\alpha_{(2i)0}^a, \alpha_{(2i)0}^b, \dots$ states of the phenome of organism i . These states are specified as follows:

- (i) The state of the community is the species composition and foodweb structure;
- (ii) the state of the population is the number of organisms it contains and their relative positions;
- (iii) the state of the organism is its pattern of gene activation;
- (iv) the state of the phenome is the physiological state of the organism;
- (v) the state of the genome is its DNA sequence.

We note that the parameters of the physiological state include organismic behavior (e.g. brain state) and morphology.

With these conventions we can now describe the state of the community in terms of a many-tuple of subsystem states, e.g.

$$\alpha^a(t) = (\alpha_{13}^b(t), \dots, \alpha_{hk}^c(t), \dots, \alpha_{(2n)0}^d(t)) \equiv \bigcap_{i,j} \alpha_{ij}^f(t), \tag{9}$$

where $0 \leq k < 3$, $\alpha_{(2n)0}^d$ is the phenome state of organism n , the ordering is a matter of convention, and the replacement is possible for some choice of values for the superscripts. Similarly we can define the environment in terms of states of local regions

$$\beta^a(t) = (\beta_{10}^b(t), \dots, \beta_{m0}^c(t)) \equiv \bigcap_h \beta_{h0}^g(t), \tag{10}$$

where β_{10}^b is the b th state of region 1, β_{m0}^c is the c th state of region m , we adopt the convention that the regions are at level 0, and the replacement is

again possible for some choice of values for the superscripts. The environment states are specified in terms of macroscopic parameters of the environment, e.g. temperature, pressure, chemical composition, radiation intensity, and so forth.

The next step is to define (local) transition schemes associated with each subsystem or local region. The subsystem schemes are given by

$$\omega_{pq} = \{p[\alpha_{pq}^u(t+1) \mid \bigcap_{i,j} \alpha_{ij}^f(t), \bigcap_h \beta_{h0}^g(t)]\} \tag{11}$$

and the local region schemes by

$$\omega'_{r0} = \{p[\beta_{r0}^v(t+1) \mid \bigcap_{i,j} \alpha_{ij}^f(t), \bigcap_h \beta_{h0}^g(t)]\}. \tag{12}$$

If we are dealing with the most uncertain allowable environment we denote the region scheme by $\hat{\omega}_{r0}$ and the subsystem scheme by $\hat{\omega}_{pq}$.

The natural question is: how do these local schemes relate to the original community and environment schemes, i.e. to $\hat{\omega}$ and $\hat{\omega}'$? The answer is: the product of all the local schemes equals the set of probabilities of the *joint occurrence of the possible subsystem transitions*. More formally:

$$\omega = \prod_{i,j} \omega_{ij} = \{p[\bigcap_{i,j} \alpha_{ij}^d(t+1) \mid \bigcap_{i,j} \alpha_{ij}^f(t), \bigcap_h \beta_{h0}^g(t)]\}. \tag{13}$$

Similarly the environment scheme is given by

$$\omega' = \prod_h \omega'_{h0} = \{p[\bigcap_h \beta_{h0}^e(t+1) \mid \bigcap_{i,j} \alpha_{ij}^f(t), \bigcap_h \beta_{h0}^g(t)]\}. \tag{14}$$

We emphasize that this product definition is unusual (though the natural one in this case) since it involves a joint probability rather than a product of probabilities.

Equation (8) (the evolutionary tendency) may now be written

$$H(\prod_{i,j} \hat{\omega}_{ij}) - H(\prod_{i,j} \hat{\omega}_{ij} \mid \prod_h \hat{\omega}'_{h0}) + H(\prod_h \hat{\omega}'_{h0} \mid \prod_{i,j} \hat{\omega}_{ij}) \rightarrow H(\prod_h \hat{\omega}'_{h0}). \tag{15}$$

This is somewhat opaque in its present form. According to the algebra of entropies, entropies of product schemes can be replaced by sums of entropies and conditional entropies of the individual scheme. Thus we might write

$$\begin{aligned} & H(\hat{\omega}_{13}) + H(\prod_{\substack{i,j \\ j \neq 3}} \hat{\omega}_{ij} \mid \hat{\omega}_{13}) - H(\hat{\omega}_{13} \mid \prod_h \hat{\omega}'_{h0}) - H(\prod_{\substack{i,j \\ j \neq 3}} \hat{\omega}_{ij} \mid \prod_h \hat{\omega}'_{h0}, \hat{\omega}_{13}) \\ & + H(\prod_h \hat{\omega}'_{h0} \mid \prod_{i,j} \hat{\omega}_{ij}) \rightarrow H(\prod_h \hat{\omega}'_{h0}), \end{aligned} \tag{16}$$

where we express the entropy of the complete community in terms of its top level entropy and an entropy for all lower levels conditioned on the behavior of the top level. If there is no lower level uncertainty and the environmental

variability is associated with the energy input, the community must be capable of rerouting the variable portion of the energy flow, either by shunting it to detritus pathway organisms (where the variations are less harmful) or by taking advantage of compensating variations in energy input (Conrad, 1972c). If this is not possible or if the variability is associated with other features of the environment, the community will suffer extinction, introduction of new species from the outside, or will tend (in the long run) to considerable speciation.

The particular expansion, (16), gives some insight into the adaptability function of biological variability. However, there are many other possible expansions. To get a picture of the “adaptability dynamics” of the whole system we introduce a device to expand (15) which puts all subsystems on an equal footing. This device, which we will call the effective entropy, is simply the sum of the unconditional and all the possible conditional entropies of the subsystem, but suitably normalized (by normalizing the linear combination of all the possible expansions). For example, we may write

$$H(\hat{\omega}_{pq}\hat{\omega}_{rs}) = \frac{1}{2}[H(\hat{\omega}_{pq}) + H(\hat{\omega}_{rs}|\hat{\omega}_{pq}) + H(\hat{\omega}_{rs}) + H(\hat{\omega}_{pq}|\hat{\omega}_{rs})], \tag{17}$$

in which case the effective entropies are

$$H_e(\hat{\omega}_{pq}) = \frac{1}{2}[H(\hat{\omega}_{pq}) + H(\hat{\omega}_{pq}|\hat{\omega}_{rs})], \tag{18a}$$

$$H_e(\hat{\omega}_{rs}) = \frac{1}{2}[H(\hat{\omega}_{rs}) + H(\hat{\omega}_{rs}|\hat{\omega}_{pq})]. \tag{18b}$$

The expressions for the effective entropies rapidly become complicated as we go to higher-order expansions. The salient point, however, is that they always consist of two parts: (1) an unconditional term, to be called the *modifiability*, and (2) all possible conditional terms, to be called the *independence* terms.

In terms of effective entropies (15) becomes

$$\sum_{i,j} H_e(\hat{\omega}_{ij}) - \sum_{i,j} H_e(\hat{\omega}_{ij} | \prod_h \hat{\omega}'_{h0}) + \sum_h H_e(\hat{\omega}'_{h0} | \prod_{i,j} \hat{\omega}_{ij}) \rightarrow \sum_h H_e(\hat{\omega}'_{h0}). \tag{19}$$

We note the meaning of the terms. The modifiability terms $H(\hat{\omega}_{ij})$ describe the behavioral uncertainties associated with the various subsystems. These include: potentiality for species change and reroutability of energy flow at the community level; culturability and topographical plasticity at the population level; developmental plasticity at the organism level; physiological and behavioral plasticity at the phenome level; and gene pool diversity at the genetic level. The independence terms $H(\hat{\omega}_{ij} | . . .)$ describe the degree to which the different plasticities are independent and therefore the extent to which they make a distinct contribution to adaptability. The modifiability parts of the anticipation terms describe the extent to which the subsystem anticipates the environment; while the independence part describes the degree to which this anticipation is correlated to behavior of other subsystems, at the same or

different levels. The indifference terms, as before, consist of both a selective and a nonselective part. However, it is clear that populations or organisms not located in some region will be indifferent to disturbance in that region (from which it follows that spatial avoidance is a component of adaptability). Also, organisms are more indifferent to the environment if they are less specialized, i.e. have many sources of food, few requirements, or are insensitive to certain physical factors.

According to (19), *change in adaptability associated with one subsystem tends to be compensated by opposite change in adaptability associated with other subsystems, at the same or different levels, or by opposite change in the indifference to the environment.*

4. The Problem of Causality. The picture developed so far is useful for dealing with the broad interplay of the various different forms of adaptability. However, it hides an important inner layer of informational processes. The reason can be seen directly from our fundamental identity (5). The term $H(\omega) - H(\omega|\omega')$ is the information which the behavior of the environment provides about the contemporaneous behavior of the biological system (e.g. biotic community). The identically equal term $H(\omega') - H(\omega'|\omega)$ represents the information which the biological system (e.g. biotic community) provides about the contemporaneous behavior of its environment. But of course this type of instantaneous information transfer raises problems of causality.

The origin of these problems can be seen from the simple consideration that our fundamental identity assumes real significance only if

$$H(\omega) - H(\omega|\omega') > 0, \quad (20a)$$

$$H(\omega') - H(\omega'|\omega) > 0. \quad (20b)$$

According to our original definitions (3 and 4) this is possible (for example in the case of [20b]) only if one of the following conditions hold for some values of v, r, s, u :

$$p[\alpha^r(t), \beta^s(t)] \neq p[\alpha^r(t), \beta^s(t), \alpha^u(t+1)], \quad (21a)$$

$$p[\beta^v(t+1) | \alpha^r(t), \beta^s(t)] \neq p[\beta^v(t+1) | \alpha^r(t), \beta^s(t), \alpha^u(t+1)]. \quad (21b)$$

Actually, the two conditions are not independent, for the failure of (21a) implies the failure of (21b). This can be seen by rewriting (21a) as

$$p[\alpha^u(t+1) | \alpha^r(t), \beta^s(t)] \neq 1, \quad (22)$$

which just means that the present state of the ecosystem must not completely determine the future state of the biota, for otherwise the amount of information which the biota transition could provide would certainly be zero. Assuming the

contrary, i.e. that condition (21a) fails but that condition (21b) holds, we can write

$$p[\alpha^u(t+1) \mid \alpha^r(t), \beta^s(t)] = 1. \tag{23}$$

Now consider the identity:

$$\begin{aligned} & p[\alpha^r(t), \beta^s(t), \alpha^u(t+1)]p[\beta^v(t+1) \mid \alpha^r(t), \beta^s(t), \alpha^u(t+1)] \\ &= p[\alpha^r(t), \beta^s(t)]p[\beta^v(t+1) \mid \alpha^r(t), \beta^s(t)]p[\alpha^u(t+1) \mid \beta^v(t+1), \alpha^r(t), \beta^s(t)]. \end{aligned} \tag{24}$$

If (21a) fails this may be replaced by

$$\begin{aligned} & p[\beta^v(t+1) \mid \alpha^r(t), \beta^s(t), \alpha^u(t+1)] \\ &= p[\beta^v(t+1) \mid \alpha^r(t), \beta^s(t)]p[\alpha^u(t+1) \mid \beta^v(t+1), \alpha^r(t), \beta^s(t)], \end{aligned} \tag{25}$$

which implies that (21b) holds only if $p[\alpha^u(t+1) \mid \beta^v(t+1), \alpha^r(t), \beta^s(t)]$ is not equal to one or zero. But this contradicts (23), and is therefore inconsistent with the assumption.

Now suppose that condition (21b) fails. In this case (24) may be replaced by

$$p[\alpha^r(t), \beta^s(t), \alpha^u(t+1)] = p[\alpha^r(t), \beta^s(t)]p[\alpha^u(t+1) \mid \beta^v(t+1), \alpha^r(t), \beta^s(t)], \tag{26}$$

which is consistent with condition (21a). Thus the failure of the second condition (21b) does not necessarily imply the failure of the first (21a).

The second condition is the one which expresses the causality violation; for it implies that the state of the biota at time $t+1$ affects or is affected by the state of the environment at time t . The connection between the two conditions means that the behavior of the biotic community provides information about the contemporaneous behavior of the environment only if the former is not completely determined by the prior state of the ecosystem (community and environment), but that even if this condition is met there are two possible sources of information. The first results directly from the time development of the ecosystem itself; for if we begin by specifying $\alpha^r(t)$, $\beta^s(t)$, and $\alpha^u(t+1)$ it is clear that specifying $\beta^v(t+1)$ will in general add to our information. The second is concomitant to the correlations between the states of the biota and environment which seem to be associated with violations of causality. If the time development of the ecosystem is completely determined there can be no causality violation. But if there is no causality violation the time development of the ecosystem may of course still be indeterminate.

The conclusion is that our formalism is consistent with causality in the sense that it is possible for the equations to be nontrivial in the absence of any instantaneous transfer of influence from environment to biota. However, this leaves open the important question, why would the state of the biota ever assume a state appropriate to the state of the environment? In other words, if

we want to deal with the mechanism of correlation we have to expand the formalism, and in such a way that condition (21b) assumes a form consistent with causality.

5. Information Transfer Picture. The basic idea is that the biota is organized in such a way that certain details of its behavior, important for information transfer, are dissipated away and can therefore be discarded in our description. The violation of causality concomitant to the correlation process (condition 21b) is a consequence of this discarding of irrelevant details and is thus only apparent.

The biota states discussed so far are functionally distinct, in the sense that we assume that they are associated with different reproduction probabilities for the various organisms in the system. Each such functionally distinct state consists of a finer set of states which changes on a finer, in reality continuous, time scale. All these finer states are equivalent as regards the system's ability to function in a given environment, but may be quite different as regards its future behavior. Thus the correlation process results from the fact that the finer state at some earlier time serves as a message to the biota about the environment state at some later time. This is possible because the changes in the environment often take place in a series of correlated steps, the first of which only affects which of its finer states the biota assumes and therefore has no significant functional effect on the system. For example, cloudiness is a sign of impending rain, but the effect of cloudiness is in general not significant. This is the type of detail we discard, and which we are justified in discarding since it is soon completely forgotten by the biota in any case.

6. Formal Aspects of Information Transfer. The basic components of the correlation process are actually a bit more complicated than indicated above. To make these explicit we first write

$$\alpha^u = \{\alpha^{u1}, \dots, \alpha^{un}\}, \quad (27)$$

where each of the α^{ui} are macroscopically distinct but identical as regards each organism's probability for reproducing. These are the finer states. The transition scheme is now

$$\underline{\Omega} = \{p[\alpha^{ui}(t+\tau), \beta^v(t+\tau), \alpha^{gm}(t+\tau_B), \beta^h(t+\tau_B), \alpha^{ek}(t+\tau_E), \beta^f(t+\tau_E) | \alpha^{rj}(t), \beta^s(t)]\}, \quad (28)$$

where τ is a time interval, τ_E is a time interval appropriate to the time scale of the environment, τ_B is a time interval appropriate to the time scale of the biota, $\tau_E < \tau_B < \tau$, and superscripts run over the respective index sets associated with each of the finer sets of states. This is the analog of our original

transition law (1), but extended to include intermediate states. We can also write analogs of the marginal transition law. In particular, the marginal transition law of the environment is given by

$$\omega' = \{p[\beta^v(t + \tau), \beta^h(t + \tau_B), \beta^f(t + \tau_E) \mid \alpha^{rj}(t), \beta^s(t)]\}, \tag{29}$$

from which it follows that the analog of our condition for information transfer (21b) is

$$p[\beta^v(t + \tau), \beta^h(t + \tau_B), \beta^f(t + \tau_E) \mid \alpha^{rj}(t), \beta^s(t)] \neq p[\beta^v(t + \tau), \beta^h(t + \tau_B), \beta^f(t + \tau_E) \mid \alpha^{ui}(t + \tau), \alpha^{gm}(t + \tau_B), \alpha^{ek}(t + \tau_E), \alpha^{rj}(t), \beta^s(t)] \tag{30}$$

for some values of the superscripts. This still violates causality since what is significant is the correlation between biota and environment at time $t + \tau$.

Now suppose that the environment state at the earlier time often serves as a message to the biota about the environment state at the later time. This is possible if:

- (i) $\beta^v(t + \tau)$ and $\beta^h(t + \tau_B)$ are strongly correlated to $\beta^f(t + \tau_E)$;
- (ii) $\alpha^{gm}(t + \tau_B)$ is strongly correlated to $\beta^f(t + \tau_E)$;
- (iii) $\alpha^{ui}(t + \tau)$ is strongly correlated to $\alpha^{gm}(t + \tau_B)$.

Also, we suppose that the biota is in a state which is essentially similar to its initial state when it receives the message and that its state is essentially similar to its final state after it processes the message. Formally,

- (iv) $\alpha^{ek}(t + \tau_E)$ belongs to α^r ;
- (v) $\alpha^{gm}(t + \tau_B)$ belongs to α^u .

Conditions (i)–(v) make it possible to approximate $\underline{\Omega}$ by

$$\{p[\alpha^{gm}(t + \tau_B), \beta^f(t + \tau_E) \mid \alpha^{rj}(t), \beta^s(t)]\}. \tag{31}$$

This is because these conditions imply

$$p[\alpha^{ui}(t + \tau), \beta^v(t + \tau) \mid \alpha^{ek}(t + \tau_E), \beta^f(t + \tau_E), \alpha^{gm}(t + \tau_B), \beta^h(t + \tau_B), \alpha^{rj}(t), \beta^s(t)] \tag{32}$$

is always either high or low and that $\alpha^{ek}(t + \tau_E)$ and $\beta^h(t + \tau_B)$ are both redundant. Thus we can rewrite our condition for information transfer as

$$p[\beta^f(t + \tau_E) \mid \alpha^r(t), \beta^s(t)] \neq p[\beta^f(t + \tau_E) \mid \alpha^r(t), \beta^s(t), \alpha^{gm}(t + \tau_B)]. \tag{33}$$

Moreover, according to condition (i) the measurement of $\beta^f(t + \tau_E)$ and $\beta^v(t + \tau)$

are basically redundant. Thus suppose that we throw out the first of these measurements and take instead the second. Then

$$p[\beta^v(t + \tau) \mid \alpha^r(t), \beta^s(t)] \neq p[\beta^v(t + \tau) \mid \alpha^r(t), \beta^s(t), \alpha^{gm}(t + \tau_B)]. \quad (34)$$

This definitely does not violate causality, nor is the inequality inconsistent with any of our conditions. Furthermore, we may regain the original condition (21b) by utilizing (v) to replace $\alpha^{gm}(t + \tau_B)$ with $\alpha^u(t + \tau)$. Thus the original condition appears to violate causality because we have ignored details of behavior associated with preliminary changes in the environment and the recognition of these preliminary changes on the part of biological systems.

Notice that the correlation process does not avoid error; for in general the preliminary response of the biota is not the most suitable biota state relative to the environment state. However, the functional importance of the error is reduced to a minimum. Indeed, it is just this difference in the functional importance of error which justifies our assumption that certain details are dissipated. In other words, as with all other forms of irreversibility, the loss of information concomitant to the correlation process would be difficult to discern if examined in too much detail.

Naturally we could make the above description more refined by considering more intermediate instants of time. The results would be exactly the same, except that we could then consider the relation between the various forms of adaptability and different time scales of environmental change.

7. The Connection between Adaptability and Reliability. The transfer and processing of information about the environment raises problems of reliability. According to (8) and (19) the adaptability tends to decrease to the uncertainty of the environment. The natural question is: is it possible to approach this lower limit without too great an increase in $H(\hat{\phi} \mid \phi)$? This is important, for this term represents the indifference (or error) and therefore must not become too large.

The answer to this question is given by information theory for the case of communication channels (cf. Shannon and Weaver, 1949) and computing networks (Winograd and Cowan, 1963). Essentially, it is possible to reduce the error to a minimum by adding sufficient temporal or spatial redundancy, with the assumption that the encoding and decoding devices are reliable (in communication systems) or that the functions performed by these devices can become increasingly complex without any increase in the probability of error (in computing systems, where the devices must be included as part of the system). In our case we include all parts of the system. Thus the conclusion is that it is possible to approach an equality without unacceptable increase in error, but that this is in general concomitant to an increase in the absolute magnitude of both the behavioral uncertainty and anticipation terms. The

increase, however, must be expressed in terms of the enlarged formalism, i.e. either in terms of our enlarged version of (8)

$$[H(\underline{\omega}) - H(\underline{\omega}|\underline{\omega}')] + H(\underline{\omega}'|\underline{\omega}) \rightarrow H(\underline{\omega}'), \quad (35)$$

or of its hierarchical version (19):

$$[\sum_{i,j} H_e(\underline{\omega}_{ij}) - \sum_{i,j} H_e(\underline{\omega}_{ij} | \prod_h \underline{\omega}'_{h0})] + \sum_h H_e(\underline{\omega}'_{h0} | \prod_{i,j} \underline{\omega}_{ij}) \rightarrow \sum_h H_e(\underline{\omega}'_{h0}). \quad (36)$$

It is the terms in bracketed expressions in these equations whose magnitudes increase. This is because it is only the finer set of states (e.g. $\{\alpha^{u^1}, \dots, \alpha^{u^n}\}$) associated with each functionally distinct state (e.g. α^u) which has to expand. The reason is that it is these states whose changes are concomitant to the transfer and processing of information and which therefore must be ordered in some redundancy structure.

The essence of the redundancy structure is that each informationally distinct fine state consists of a set of macroscopically distinct but informationally equivalent states, with the property that atypical states of this set arising from perturbation always fall back into a subset of typical states. In classical information theory this "falling back" occurs because the states are organized (either spatially or temporally) in the form of error correcting codes. However, it is also possible for informationally distinct states to consist of numerous macroscopically equivalent states, organized in such a way that the falling back is associated with maximization of entropy or minimization of energy. This is perhaps of considerable biological importance for it enables the falling back to occur spontaneously and without any error correcting computations. We note that in any case each of our finer states must consist of even finer, macroscopically equivalent states, since biological systems are open systems and what we recognize as a state on any functional basis must necessarily be an ensemble of microstates. We also note that the problem of reliability is amenable to still another solution, viz. increase in the physical precision of the components. This is, in fact, the solution which plays the most important role in present day technical systems. It is connected with the fact that certain quantum mechanical structures are more resistant to thermal perturbation than others (cf. Pattee, 1968).

8. The Components of Biological Variability. We can now classify the fundamental components of biological variability. These are:

- (i) *Functionally distinct states.* These are states optimally suited to different macroscopic environments.
- (ii) *Informationally distinct states.* These are functionally equivalent states representing responses to different changes in the environment state.
- (iii) *Macroscopically distinct, but informationally equivalent states.* These are

redundant sets of informationally equivalent states, allowing for some error correction capability.

(iv) *Macroscopically equivalent states*. These are the ensemble of microstates corresponding to each macroscopically distinct state. They may also allow for reliability, but on the basis of self-organizing energy and entropy processes.

The functionally distinct states may be described in terms of functionally distinct states at various levels of biological organization, i.e. in terms of gross variables ordinarily associated with these levels of organization. The finer states may also be broken up on this basis. When the variability of the functionally distinct or finer states is conditioned on the behavior of the environment, it may be interpreted in terms of anticipation of the environment or in terms of its contribution to the reliability of the anticipation process. Ultimately this is based on the ability (acquired either through evolution or learning) to recognize correlations in the environment. When the variability of the environment is conditioned on the behavior of the biota, it represents indifference or error. When the error is high, this means that the biota is exhibiting more or less pathological behavior. Any contribution which such behavior makes to the variability of behavior is in effect a nonfunctional contribution. We should also mention that unreliable behavior may have functional significance in some cases, viz. in competitive interactions among members of the community. In these cases the magnitudes of the behavioral uncertainty and anticipation terms increases, but with no increase in indifference and no increase in overall adaptability. The resulting internally generated uncertainty serves to increase the environmental uncertainty of the competitor.

The variability, or uncertainty of behavior, must be distinguished from the diversity of behavior. Naturally, the behavior of a biological system may be quite diverse but at the same time not very variable.

The considerations in this and the previous section sharpen our generalizations about adaptability. The evolutionary tendencies for minimization and compensation continue to hold, but we can now add that the allocation of adaptability cannot be separated from the problem of reliability. This is because decreases in both the variability of functionally distinct states and increases in the anticipation term (the reduction of such variability conditioned on the environment) must in general be associated with increases in the ensemble of finer states, if the error is to be kept small.

9. Further Implications of the Model. The model has a number of implications for the ramification of disturbance in biological systems. The main ones are:

(i) In order to survive the disturbance (without intolerable injury) the system (e.g. community or population) must be able to assume a behavioral mode suitable to the disturbance, or must be isolated from it, either spatially,

by some protective structure, or by virtue of extra functional capabilities (e.g. extra metabolic machinery) which reduces its needs. The implication of equation (8) is that the ensemble of required behavioral modes is at least as great as the ensemble of disturbances to which the system is not indifferent. However, the uncertainty associated with the former must be greater if the system does not anticipate the environment reliably, either because of the inappropriateness of its behavioral pattern or because it fails to use certain minor environmental variations to predict the future behavior of the environment. The total adaptability of the system thus consists of three components: the behavioral ensemble, the ability to anticipate, and the indifference. Moreover, evolutionary considerations suggest that it is essentially a conserved quantity in environments of constant uncertainty.

(ii) The disturbance ramifies more widely in the system if it pushes more of the compartments (or subsystems) into alternate modes of behavior. According to equation (19) the ramification of disturbance decreases as the various compartments at the same or different levels are more independent. In this case a given total amount of observable variability of the system makes a bigger contribution to the adaptability (where the total variability is calculated by adding the variabilities of the individual compartments). However, again according to equation (19), this independence does not mean that information (e.g. about the disturbance) does not flow between compartments or levels. This is because such information flow is included in the anticipation term.

(iii) Point (ii) above suggests that centralized systems are less adaptable for a given total apparent modifiability than decentralized systems. With increasing system centralization the system must either develop special, relatively independent organs of adaptability, decrease niche breadth, or expend more energy coping with environmental uncertainty.

The above remarks seem to agree with a number of concrete situations. Microorganisms (as a class) are capable of genetic and physiological (more precisely, cellular) plasticity as well as culturability at the population level. Higher plants, because of their larger size and longer generation times, are not so amenable to culturability. However, their open growth systems and relatively simple morphologies allow for high developmental and genetic plasticity. Higher animals, with more complex morphologies and closed growth systems, are not so amenable to developmental and genetic plasticity. However, they compensate for these restrictions with a highly developed neuromotor system (and therefore behavioral plasticity, including learning), highly developed immune systems, and social organization. Or, turning the argument around, we may regard the requirements for such compensation as the main selective force concomitant to the rapid development of the higher nervous system in man in relatively recent times.

In general, increasing system interdependence (or integration of subsystems) is associated with specializations of labor within the system, and therefore with greater efficiency. Thus point (iii) suggests that the advantages of increased specialization of labor are counterbalanced by the increased costs of adaptability, either in terms of requiring greater total modifiability or in terms of requiring the system to support special organs of adaptability. The theory therefore predicts that as the harshness of the environment increases (i.e. as it is more difficult to extract energy or as it requires more energy expenditure to function), the advantages of more specialization of labor (for extracting energy) will increase, but the difficulty of paying for the concomitant increased costs of adaptability will also increase. Thus, at some critical point of increasing environmental harshness or increasing statistical perturbation for a given degree of harshness, internally specialized organisms should give way to simpler forms with much less costly forms of adaptability.

Notice that as the modifiabilities at different levels of organization become more independent the cost of adaptability decreases. Indeed, one way for an internally specialized system to develop new modes of adaptability is to develop another level of organization, for example, from the cellular to the multicellular. Later the multicellular, or organism level, can also specialize internally. The importance of independence of levels gives some internal justification for describing biological systems in terms of variables associated with particular levels of organization, although a really complete description must also take into account influences deriving from other levels.

10. Conclusions. The treatment of biological adaptability sketched here is based on the structural analysis of biological variability. The measure of adaptability is the maximum tolerable uncertainty of the environment. What this tolerance represents is the potential ensemble of modes of behavior of the system (for which the system remains alive) plus indifference to the environment based on mechanisms such as spatial separation or thermal isolation. Actually, each of the states of the system is really an equivalence class of finer states, associated with signal transmission, and even finer states which are concomitant to the inevitable flow of matter through the system. Over and above these "signal states" there is imposed a redundancy which allows for reliable information processing in the presence of noise. This may be temporal, as in classical communication systems, or spatial, i.e. concomitant to redundancy of components.

Equations (8) and (35) describe the relations among these different types of states but without explicit consideration of system structure. Equations (19) and (36) take level structure into account and therefore describe the relations in terms of variables ordinarily used to describe biological systems. The real import of these equations is that biological systems are characterizable

by a measurable quantity—the adaptability—which is roughly conserved in the course of evolution. This is important because it means that any contractions in one form of adaptability must be compensated by expansions in some other. Furthermore, by imposing certain extra constraints (costs and advantages of different forms of adaptability) we can turn the spectrum of adaptabilities in different types of biological systems into an optimization problem.

I would like to close out by comparing the statistical approach to the dynamical approach. In a certain sense the statistical approach is the more modest, at least in the sense that it does not assume a detailed knowledge of the biological world or give detailed answers. But by giving up the attempt to deal with all these we can answer certain questions which are inaccessible to dynamical models.

The most general question is: what sorts of biological systems are possible and which are chimerical. In principle we can treat this as an optimization problem, assuming that we in fact know the costs and advantages associated with the different forms of adaptability. Since each dynamical model is already a particular system, it could hardly be expected to deal with this question. Indeed the problem with dynamical models is that they deal with extinction, but never with the origin of new species (cf. Margalef, 1973).

Of course one would expect that the fundamental equations of physics should hide within them the answers to this question. However, dynamical models in biology are never set up using the fundamental principles of physics, for clearly the equations would be totally intractable. Rather, they are ad hoc models whose validity is at best limited to a limited class of systems for a limited amount of time. The statistical approach, on the other hand, deals with relations between system and environment which are always valid.

The statistical and dynamical approach, however, are not unrelated. From the dynamical standpoint a system is stable to a certain class of perturbations if it forgets these perturbations; and it is structurally stable if its qualitative behavior forgets any perturbation to its “law” of motion (see e.g., Rosen, 1970). But from the statistical point of view we say that a system which either forgets or tolerates such perturbations is adaptable. The dynamically stable forms of biological organization must thus be those which satisfy the requirements of adaptability and reliability. Indeed, we can regard the functional requirements of adaptability as conditions on the solution curves of a dynamical system and the class of equation structures which satisfy these conditions as the class of admissible biological organizations. This is important because it opens up the possibility of reformulating our questions about the relation between structural complexity and adaptability in terms of more tractable questions about the relation between structural complexity and stability.

I would like to thank Robert Rosen for a number of stimulating remarks, including remarks on the functional approach to biological problems.

LITERATURE

- Conrad, M. 1972a. "Statistical and Hierarchical Aspects of Biological Organization." In *Towards a Theoretical Biology*, Vol. 4, pp. 189–221. C. H. Waddington, ed., Edinburgh: Edinburgh University Press. Also available as Center for Theoretical Studies preprint, No. CTS-B-70-1, Coral Gables, 1970.
- . 1972b. "Can There Be a Theory of Fitness?" *Int. J. Neurosci.*, **3**, 125–134.
- . 1972c. "Stability of Foodwebs and Its Relation to Species Diversity." *J. Theor. Biol.*, **34**, 325–335.
- . 1975. "Analyzing Ecosystem Adaptability." *Math. Biosci.*, **27**, 213–230.
- . 1976. "Patterns of Biological Control in Ecosystems." To appear in *Systems Analysis and Simulation in Ecology*, Vol. 4, pp. 431–456. B. C. Patten, ed. New York: Academic Press. Also available as University of Tübingen preprint, "Patterns of Biological Control," February, 1973.
- Khinchin, A. I. 1957. *Mathematical Foundations of Information Theory*. New York: Dover.
- Margalef, R. 1973. "Some Critical Remarks on the Usual Approaches to Ecological Modeling." *Inv. Pesq.*, **37**, 621–640.
- Pattee, H. H. 1968. "The Physical Basis of Coding and Reliability in Biological Evolution." In *Towards a Theoretical Biology*, Vol. 1, pp. 67–93. C. H. Waddington, ed. Edinburgh: Edinburgh University Press.
- Rashevsky, N. 1960. *Mathematical Biophysics, Physico-Chemical Foundations of Biology*, Vol. 2. New York: Dover.
- Rosen, R. 1967. *Optimality Principles in Biology*. New York and London: Butterworth.
- . 1970. *Dynamical System Theory in Biology*, Vol. 1. New York: Wiley-Interscience.
- Shannon, C. E. and Weaver, W. 1949. *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.
- Winograd, S. and Cowan, J. D. 1963. *Reliable Computation in the Presence of Noise*. Cambridge, Mass.: MIT Press.