DISCUSSION PAPER:
THE QUALITATIVE ANALYSIS OF PARTIALLY SPECIFIED SYSTEMS

Richard Levins
Department of Biology
University of Chicago
Chicago, Illinois 60637

The most difficult general problem of contemporary science is how to deal with complex systems as wholes. Most of the training of scientists, especially in the United States and Great Britain, is in the opposite direction. We are taught to isolate parts of a problem and to answer the question “What is this system?” by telling what it is made of. The dramatic advances in science in our generation have almost all been in areas where such an approach is practicable. The notable stagnations have been in areas of complex systems approached in pieces.

It is now a commonplace, at least in ecology, that systems are complex and that the one-step linear causality is a poor predictor of ultimate outcome. Consider, for example, the problem of providing more food for hungry people. Since insects destroy a significant portion of the world’s crops, and since insecticides can be shown in the laboratory to kill insect pests, it is a plausible inference that the use of insecticides will control insects and increase food available to the hungry. Furthermore, to avoid side effects, laboratory tests may show that insecticides such as heptachlor are relatively nontoxic to mammals. Therefore, it is reasonable to expect that the use of such insecticides would reduce insect pests, increase yields, and alleviate hunger.

But often it does not work that way. First, the application of insecticide does not necessarily control the insect pest, for at least three reasons:

1. Any insect killed by insecticide is that much less food for the predators of the pest. This in itself reduces the predator population, so that the end result is a shift in the cause of death of the pest—more are poisoned, fewer are eaten—but not in the numbers.

2. The insecticide directly reduces the predators of the pest.

3. Natural selection in the target population rapidly builds up resistance to the insecticide. In general, an insecticide is physiologically effective for two to ten years.

The side effects may also behave in unexpected ways. The relatively safe heptachlor may be transformed into highly toxic substances under field conditions where the action of sunlight in the presence of a vast ensemble of organic and inorganic substances promotes reactions that do not occur in the simple laboratory test.

Finally, even the obvious expectation that increased food production alleviates hunger proves false. The whole domain of agricultural economics, grain prices, trade agreements, credits for farmers, land concentration, and speculation intervenes between the harvesting of a crop and its consumption.

Similar problems of complex interactions have arisen in ecology, medicine, economics, administration, and other disciplines. This has led to an interest in complexity per se and the exploration of strategies to deal with complexity. Three general approaches have emerged.

1. Statistical-biometrical. Here the system is treated as a black box, and its
behavior is described in terms of the patterns of variation and covariation of the component variables. In conjunction with other approaches its value is in the testing of hypotheses about a system's structure or in posing subproblems. Taken alone, it helps in the prediction of the behavior of very similar systems but not in understanding.

2. In affluent technological societies with a generally reductionist philosophy, the dominant approach is through engineering models. Derived from either the cybernetic school or "systems analysis," these are attempts to measure all the links in a system, write all the equations, measure all the parameters, and either solve the equations analytically or simulate the process on a computer to obtain numerical results.

3. Standing outside of academic science through most of its history, the Marxist tradition has always emphasized complexity itself as an object of study and has stressed interconnection, wholeness, qualitative relations, multiple causality, the unity of structure and process, and the frequently contra intuitive results of contradictory processes. It is the major source of my own research.

The complex networks of biological systems are only partially specified for various reasons:

1. When the number of links is large and each one difficult to measure, the complete description of a given system might be the work of several lifetimes. Thus, while we know that a predator eats its prey, the rate of change of predators as a function of the number of prey depends on its hunting strategy; the effect of hunger on feeding rate; the possibilities of learning to recognize types of prey; food preferences when several prey species are available; reproductive physiology, which establishes the relation between food intake, numbers of offspring, and physical parameters of the environment; and the genetic heterogeneity of predator and prey species with respect to all of these components.

2. While some variables of a system are readily measurable, others are too vaguely defined and yet very real. Thus, in case 3 the levels of blood glucose, insulin, and adrenalin and the kinetic constants describing their rates of breakdown are definable and measurable, but the subjective symptoms of hypoglycemia, adrenalin release, and external stress are not. To insist upon a precise description would therefore mean to exclude these aspects from the system. Yet they are important in its dynamics and often the central issue. We will show later than despite the vagueness of some of the psychological components a qualitative analysis permits a close integration of these with physiology.

3. Often the question at issue is not the interpretation of a particular system but of a class of systems. We may want to know such things as: how many species of birds, more or less as similar to each other, like the warblers of New England, can coexist in a stable community? How does an ecosystem stratified into discrete trophic levels of predators and prey differ in its behavior from one in which a herbivore and its predator have predators in common? How is the stability of a system affected when it grows in the number of variables but not the rules of construction? For problems of this sort, numerical specification is not very helpful.

4. The partial specification may sometimes be closer to the biological reality than a complete one. For instance, it is unlikely that the genotype specifies the complete circuit diagram of the central nervous system. Rather, it seems to specify certain general rules of construction such as connectivity, range of an individual neuron, distribution of thresholds, and ratio of excitatory to inhibitory neurons, for whole regions. We want to know which rules of construction would permit what kinds of behavior.
5. In a deeper sense, no system is ever completely specified. Thus we may observe that a species' rate of increase is reduced when population density increases. This suggests an equation in which the rate of growth, $1/N \frac{dN}{dt}$, is a decreasing function of $N$. But in fact, population size rarely affects the rate of increase directly. There is an indirect pathway, often of many links, from population size to the accumulation of waste products in the medium, to the physiological state of the organism. Between any two variables it is always possible to insert intermediate variables, while each variable itself is the lumping of a heterogeneous ensemble of genotypes, ages, and so on. The qualitative analysis should be capable of indicating when variables may be lumped and when to insert intermediate variables. The network to which we apply our mathematics is already the result of conscious or unconscious qualitative decisions; that is, theoretical interpretation.

A formal mathematical development of the qualitative theory is in preparation for publication elsewhere. Here my purpose is to outline the argument, support it with a heuristic framework, and present the results in a form that can be easily applied.

The first step is to show the partial equivalence of a system of differential equations, a matrix, and a graph. Consider a set of variables $x_i$ and the equations for the rate of change of each:

$$\frac{dx_i}{dt} = f_i(x_1, x_2, x_3 \cdots x_n).$$

At equilibrium, $f_i = 0$ for all $i$. Then, differentiating (1) with respect to each variable we get

$$\frac{\partial}{\partial x_j} \frac{dx_i}{dt} = \frac{\partial f_i}{\partial x_j}$$

evaluated at equilibrium. If $x_j$ does not appear at all in $f_i$, this will be zero. These $\frac{\partial f_i}{\partial x_j}$ terms are the elements $a_{ij}$ if the matrix $A$, an $n \times n$ square array in which $a_{ij}$ is the element in the $i^{th}$ row and $j^{th}$ column:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}$$

The diagonal elements $a_{ii}$ are of special importance. Many biological variables are self-reproducing, and may be represented by equations of the form

$$\frac{dx_i}{dt} = x_i g_i(x_1, x_2, x_3 \cdots x_n)$$

in which $g_i$ does not contain $x_i$ explicitly. Then

$$\frac{\partial}{\partial x_i} \frac{dx_i}{dt} = x_i \frac{\partial g_i}{\partial x_i} + g_i$$

But since $g_i$ does not include $x_i$ at all, the first term on the right side is zero, and since we are evaluating these terms at equilibrium, $g_i$ is also zero. Hence for this kind of variable the diagonal element of the matrix,
The symbols of loop analysis. $X_1$ and $X_2$ are variables. The arrow from $X_1$ to $X_2$ indicates a positive effect ($a_{21} > 0$), and the line ending in a circle indicates negative effect ($a_{12} < 0$). $X_2$ also has a negative effect on itself ($a_{22} < 0$).

$$a_{ii} = 0.$$ \hspace{1cm} (6)

But when a variable is not self-reproducing, when it is produced or introduced into a system at a rate that does not vanish when the variable itself is absent, while the rate of removal of the variable depends on its own concentration, this is no longer the case. Thus, the concentration of usable phosphate in a lake may follow the equation

$$\frac{dP}{dt} = I - bP$$ \hspace{1cm} (7)

where $I$ is the input from streams and $b$ is the rate of phosphate uptake by algae, removal by outflow, or transformation into unusable forms by inorganic processes. Then

$$\frac{\partial}{\partial p} \frac{dP}{dt} = -b$$ \hspace{1cm} (8)

and the diagonal element is negative. The same thing happens in a chain of chemical transformations where the rate of formation depends on the concentration of the precursor.

Thus the matrix is derived from the set of differential equations, but represents it only locally, near an equilibrium point.

The next step is to draw the graph of the matrix. Let each variable $x_i$ be represented by a vertex of the graph, and let the line from $x_j$ and $x_i$ be equivalent to $a_{ij}$. If $a_{ij} = 0$ there is no line from $x_j$ to $x_i$. Since $a_{ij}$ need not equal $a_{ji}$, each pair of points may be connected by zero, one or two oriented lines. Further, since we often know only the sign of $a_{ij}$ but not its value, we distinguish positive and negative lines by the symbols in Figure 1.

A variable that is self-damped is represented by a loop of length 1 as shown in Figure 1.

Corresponding to every square matrix there is also a determinant, which has a numerical value formed by a sum of products of the elements of the matrix. Thus, for example, the $2 \times 2$ matrix has the determinant

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$ \hspace{1cm} (9)

Note that the value of the determinant is equal to the product of the loops of length 1 minus the loop of length 2. This result can be generalized to all determinants: a determinant of order $K$ has the numerical value

$$D^{(k)} = \sum (-1)^{k-n}L(m, k)$$ \hspace{1cm} (10)

where $L(m, k)$ is a product of $m$ disjunct loops totaling $k$ elements. Since the prin-
principal diagonal yields a product of \( k \) disjunct loops, its sign is always positive, while a single loop of length \( k \) has the coefficient \((-1)^{k-1}\).*

It will be convenient to transform this value into the measure of the feedback of a matrix, which will be defined as

\[
F_k = (-1)^{k+1}D^{(k)}
\]

or

\[
F_k = \sum(-1)^{m+1}L(m, k).
\]

This measure has the following properties: if in a given product of \( m \) loops all loops are negative, then the \( m \) minus signs times \((-1)^{m+1}\) contributes a negative term to \( F_k \), and if all loops in a graph are negative \( F_k \) is negative. A single positive loop results in positive feedback, two positive loops in a product restore the negative feedback, and so on. Thus, when combining loops into product to get feedback, positive and negative behave opposite to the way they do in ordinary multiplication.

**The Local Stability of Systems**

In order to understand the conditions for local stability we first take a detour to a simple discrete-time, single variable cybernetic mechanism. Suppose that \( x \) is a variable that changes at discrete intervals \( \Delta t \) according to the rule

\[
x(t + \Delta t) = x(t) + a\Delta tx(t).
\]

This system has an equilibrium point at \( x = 0 \). But if it is not at equilibrium its behavior depends on the value of \( a\Delta t \). If \( a\Delta t \) is zero (that is, \( a = 0 \)), the system never moves, and it can be described as having neutral or passive equilibrium. If \( a \) is positive, we have a system with positive feedback. If \( x(t) \) is initially positive it will increase, if initially negative it will decrease, and in both cases it moves away from zero, which is therefore an unstable equilibrium. If \( a\Delta t \) is a small negative number (<1), \( x(t) \) will approach zero asymptotically from either side. Thus if \( a\Delta t = \frac{1}{3} \) and we start at \( x(t) = 1 \), the successive values are \( 1, \frac{2}{3}, \frac{1}{3}, \frac{2}{3}, \ldots \).

If the \(-a\Delta t = \frac{1}{3}\), the successive values starting at \( x(+1) = 1 \) are \( 1, -\frac{1}{3}, -\frac{2}{3}, -\frac{2}{3}, \ldots \). As the \(-a\Delta t \) term increases beyond 2, we get unstable oscillations: for \(-a\Delta t = 3\), the process would go from 1 to \(-2, +4, -8, +16 \ldots \) Norbert Weiner therefore defined as a measure of instability the product of the reaction rate, \( a \), times the time lag, \( \Delta t \).

This simple model could be complicated in many ways. We could make \( a \) a function of \( x \), the equation could depend on \( x(t - \Delta t) \) and other past values, we could have it a system of many variables. But already the essential features of interest to our argument have appeared:

There are two kinds of instability in systems. When positive feedback predominates, the system moves out from equilibrium in the same direction as its

* This relation between determinants and loops has apparently been discovered independently a number of times but does not appear in the generally available literature. While lecturing on these results I was directed to the earlier work of Mason.\(^5\) His work was directed toward developing algorithms for the easier computation of electrical circuits, rather than as an instrument for qualitative understanding. Therefore, even if I had followed the engineering literature \( I \) would have missed its significance. Similarly, Sewell Wright used the technique of "path coefficients" to find in-breeding measures.
initial displacement, while if there is excessive negative feedback with long enough time lags there is oscillatory instability.  

In systems of differential equations there are no explicit time lags. However, time lags sneak in indirectly through the effects of a variable on itself by way of other variables; that is, through the loops in the system of various lengths.

The analogy to the discrete-time cybernetic unit does not enter into the proof of the conditions for local stability but instead makes them more understandable.  
The first necessary condition for stability is that in a system of \( n \) variables

\[
F_k < 0 \quad (14)
\]

for all \( k \leq n \). That is, at each level \( k \) negative feedback must outweigh positive feedback. It follows immediately from this that the following systems shown in Figure 2 are unstable:

![Figure 2](image)

**Figure 2.** Unstable systems with nonnegative feedback. *Left:* the feedback at level 2, \( F_2 \), is positive, since it comes entirely from the \( X_1, X_3 \) loop. *Right:* \( F_3 = 0 \). If \( X_4 \) were self-damped, instead of \( X_3 \), then we would have \( F_3 < 0 \), since \( F_3 \) is a product of disjunct loops of lengths 1 and 2.

Each \( F_k \) term is the sum of the feedbacks from all the subsystems of \( k \) variables formed by fixing \( n-k \) variables at their equilibrium levels and considering only the remaining \( k \) as variables. It therefore follows that if \( F_k < 0 \) at least one subsystem of order \( k \) has negative feedback. Another conclusion is that if a system is unstable because feedback is positive at some level \( k \), no changes in loops of length greater than \( k \) can stabilize the system.

The second condition for stability, that negative feedback with long-time lags cannot be too large compared to the shorter-loop negative feedback, is more difficult to describe. It is a reinterpretation of the Routh-Hurwitz algorithm (see Gantmacher'). This condition is expressed as a sequence of expressions all of which must be positive. For our purposes the first of these is sufficient:

\[
F_1 F_2 + F_3 > 0. \quad (15)
\]

If the previous condition is satisfied, \( F_1, F_2, \) and \( F_3 \) are all negative. Then the condition requires the feedback at level three to be less than the product of the negative feedbacks at the lower levels. But we can carry the analysis one step further: \( F_3 \) contains loops of length 3 and also the products of disjunct loops of shorter length. The product \( F_2 F_3 \) does not involve the loops of length 3. But it includes all products of loops of length 1 with loops of length 2 (disjunct or not) and of length 1 with products of pairs of disjunct loops of length 1. The final result can be expressed in (16), where the subscripts refer to the elements included in a loop and different subscripts are distinct:

\[
\Sigma - L_i^2 L_j + L_i L_{ij} + L_{ijk} > 0. \quad (16)
\]
FIGURE 3. Unstable systems with excessive negative feedback. In both cases, the negative loop of length three is not balanced by negative feedback from conjunct loops of lengths 1 and 2.

Thus the following systems are unstable (see Figure 3).

There has been much discussion in recent years as to whether complex systems are more stable or less so than simple ones. In 1955, Robert MacArthur proposed that complexity promotes stability, and this generalization has been accepted within ecology and also in discussions of the environmental crisis. However, in 1968 Levins showed that as a community of competitors increases in the number of species it eventually becomes unstable, and this point is reached when the number of species is roughly the reciprocal of the average value \(a_{ij}a_{ji}\). Gardner and Ashby used computer simulation to show that if the interaction coefficients \(a_{ij}\) are chosen from random distributions, the probability of stability decreases from nearly one to almost zero when \(n\) increases through values of roughly \(1/\text{var}(a_{ij})\). And Robert May also argued analytically that large systems are usually unstable.

In the argument that follows we use (16) to show what happens as a system increases in size or connectedness. Suppose that a system has \(n\) variables all of which are self-damped with \(a_{ij} = -1\). Suppose further that the probability of a link between any pair of variables is \(p\), and that if they are connected the magnitude of the interaction averages \(-a\). Finally, we assume that all connections are independent of each other. Then the first term in equation (16), \(-\sum L_i^2 L_j\), is equal to the number of ordered pairs of elements or \(n(n-1)\). The second term is \(-n(n-1)p^2a^2\), while the number of loops of length three is \(\frac{1}{3}n(n-1)(n-2)p^3\) and the whole stability requirement becomes

\[
n(n-1) - n(n-1)p^2a^2 - \frac{1}{3}n(n-1)(n-2)p^3a^3 > 0. \tag{17}
\]

Since \(n(n-1)\) factors out, we are left with

\[
1 - p^2a^2 - (n-2)p^3a^3 > 0. \tag{18}
\]

But for \(n\) large enough, this will certainly be violated. This is a robust result. If we change the model so that whenever \(x_i\) is connected to \(x_j\) there is a reciprocal reaction of opposite sign but otherwise links are independent, we get

\[
1 + pb^2 - (n-2)p^3a^3 > 0. \tag{19}
\]

The point is, the loop of length three involves three distinct elements and the num-
ber of such combinations is of the order \( n^4 \), whereas the terms that involve lower order feedbacks have conjunct loops involving only \( n^2 \) combinations. Similar stability criteria involving feedbacks of greater length behave in the same way.

Consider now a crude model of the central nervous system in which a region of many neurons consists of excitatory and inhibitory links in such a way that on the average, inhibitory links exceed excitatory links. The self-damping term depends on the physiology of the separate neurons. The interaction intensity \(-a\) is a sigmoid function of the general level of sensory input. Then for a given number of neurons in the region, \( n \), and probability of connection, \( p \), there is a threshold input intensity below which the mean firing rate of neurons is stable and above which the system will oscillate. (Wilson and Cowan,\(^8\) using simulation and a different mathematical formulation, showed that intensity of stimulus may be coded by the frequency of oscillation. Their approach is more useful for a detailed investigation of neural nets, while the method described here gives a more intuitive over-all understanding.)

Instead of holding probability \( p \) of connection fixed, we define the mean connectivity \( c = p(n - 1) \), equation 3 becomes approximately, for large \( n \),

\[
n(n - 1) - c^2a^2 - c^3a^3 > 0.
\]

If \( c \) is held fixed, the system is eventually (for large enough \( n \)) stable. But for any given \( n \), a high enough connectivity makes the system unstable.

The general conclusion that systems of many variables with high connectivity are likely to be unstable (and that the higher the connectivity the greater the preponderance of long loops and of very low frequency, long-term oscillations) is rich in implications for a number of fields.

1. Instability must not be confused with lack of persistence. In cellular metabolism and in the central nervous system it is rather interpretable as spontaneous activity. In cells, where the connectivity of biochemical networks is relatively low, this results in periodic behavior, cell cycles. In the central nervous system it is possible that regions of high connectivity yield cycles so long that they are never completed. This simulates randomness and results in a certain unpredictability of neural phenomena.

2. Only where variables are self-reproducing does instability suggest extinction. This relation has made it possible to interpret the number of coexisting species in communities of different structures (Levins and colleagues,\(^4\) Vandermeer\(^7\) . . .).

3. A system involves variables the reactions of which proceed at commensurate rates. Those which are much slower are held as constant parameters, whereas the variables of interest equilibrate; those which are much faster are treated as already at their moving equilibria. They are therefore replaced by functions of the main variables. This sheds a new light on the significance of enzymes. It is not merely that they increase the rates of reactions by many orders of magnitude and make new reactions possible. By so doing, they simplify the biochemical network of cells. The network of enzymatic reactions at high velocity is very much less interconnected than the network of slow interactions among thousands of molecular types. The relative simplicity results in either stable or rapidly oscillating systems upon which natural selection may act. But once the enzymes are killed the slow reactions creep back into the system; the very high connectivity results in nonperiodic instability; dust returns to dust not by assuming the same chemical composition as its surroundings but by having its own kinetics merge in a sea of commensurate reaction rates.
Levins: Discussion Paper 13

THE RESPONSE OF EQUILIBRIUM LEVEL TO PARAMETER CHANGE

A system may be at equilibrium and yet change. If the parameters of a system that appear as constants in the equations in fact change slowly compared to the changes of the variables themselves, we can have a system that follows a moving equilibrium. We therefore want to know what happens to the equilibrium levels in a system when the parameters of that system—temperature, input rate, genetic properties, kinetic constants—are altered.

Consider the system of equations we examined before, but with the parameters explicitly included:

\[
\frac{dx_i}{dt} = f_i(x_1, x_2, \ldots, x_n, c_1, c_2, \ldots, c_m) \tag{21}
\]

where the \( c_n \) are constant parameters. At equilibrium the right hand side of \( f_i \) is zero and we can differentiate the system with respect to a single parameter, say \( c_h \), to obtain the new system of linear equations in \( \partial x_j / \partial c_h \):

\[
\sum \frac{\partial f_i}{\partial x_j} \frac{\partial x_j}{\partial c_h} + \frac{\partial f_i}{\partial c_h} = 0 \tag{22}
\]

The \( \partial f_i / \partial x_j \) are the familiar \( a_{ij} \) of the previous sections. The \( \partial x_j / \partial c_h \) are new variables, while \( \partial f_i / \partial c_h \) is a constant in each equation. Therefore, this system can be solved directly by determinants. For instance,

\[
\frac{\partial x_1}{\partial c_h} = \left| \begin{array}{c}
-\frac{\partial f_1}{\partial c_1} a_2 a_3 \cdots a_n \\
-\frac{\partial f_2}{\partial c_1} a_2 a_3 \cdots a_n \\
\vdots \\
-\frac{\partial f_n}{\partial c_1} a_2 a_3 \cdots a_n \\
\end{array} \right| \left| \begin{array}{c}
a_1 a_2 a_3 \cdots a_n \\
a_1 a_2 a_3 \cdots a_n \\
\vdots \\
a_1 a_2 a_3 \cdots a_n \\
\end{array} \right| \tag{23}
\]

That is, the solution for \( \partial x_1 / \partial c_h \) is found by substituting the column \( -\partial f_i / \partial c_h \) for the \( i^{th} \) column of the determinant of the system and dividing this determinant by the determinant of the whole. These results can be reexpressed in terms of feedback and pathways in the system, as follows: Let \( p_{ij}^{(k)} \) be a simple open path (one that does not cross itself) from variable \( j \) to variable \( i \) in the system which includes \( k \) elements (and therefore \( k - 1 \) links). Define the complementary subsystem to \( p_{ij}^{(k)} \) as the subsystem formed by only those elements which do not appear in \( p_{ij}^{(k)} \). Finally, let \( F_{n-k}^{(p-1)} \) be the feedback of the complementary subsystem. For formal reasons we must define a subsystem of zero elements as having feedback

\[ F_0 = -1 \tag{24} \]

and an open path from an element to itself

\[ p_{ii}^{(1)} = 1. \tag{25} \]

Then we can express our final result as follows:

\[
\frac{\partial x_i}{\partial c_h} = \sum_{j,k} \frac{\partial f_i}{\partial c_h} p_{ij}^{(k)} F_{n-k}^{(p-1)} / F_n. \tag{26}
\]

That is, \( \partial x_i / \partial c_h \) is the sum of the products of \( \partial f_j / \partial c_h \) times the products along each path from \( x_j \) to \( x_i \), each multiplied by the feedback of the complementary subsystem and all divided by the feedback of the whole. Of special interest are those cases where \( c_h \) is simply the input to variable \( h \). Then only \( \partial F_h / \partial c_h \) is different from zero (in fact, equals 1). Since \( F_n \) is the feedback of the whole system,
which is taken to be stable, it is negative. If the complementary subsystem has zero feedback then a path has no effect. Finally, if the complementary subsystem has net positive feedback, then $\frac{\partial x_i}{\partial c_h}$ has a sign opposite to that of the path.

We illustrate these principles with a few simple examples in FIGURE 4.

**FIGURE 4.** Equilibrium changes in two-variable systems. Top: $X_1$ has zero complement, and is therefore unchanged by direct input. Center: the complement is positive. Therefore, input to $X_1$ affects the equilibrium in the direction opposite to its own sign. Only 4c shows the common-sense effect that adding $X_1$ increases its level.

Here an input to $x_1$ has as its complement the system $[x_2]$ with zero feedback. Therefore, adding $x_1$ does not change the level of $x_1$ at all. But adding $x_1$ has a positive effect on $x_2$, which adsorbs the whole increase. An input to $x_2$ will increase $x_2$ (the complementary subsystem $[x_1]$ has negative feedback) and reduces $x_1$. In FIGURE 4 (center), $x_2$ has positive feedback. Now increasing the input to $x_1$ actually reduces the level of $x_1$. This comes about because the increase in $x_1$ increases $x_2$, but $x_2$ has positive feedback (i.e., is autocatalytic) and increases even more. Finally, the greatly increased $x_2$ reduces $x_1$. In FIGURE 4 (bottom), $x_1$ and $x_2$ are both self-damped.

Therefore, increasing the input to either one increases its own equilibrium level. But if $x_1$ is increased, then so is $x_2$, and the two variables will show a positive correlation in time or space, whereas increasing the input to $x_2$ reduces $x_1$ and results in a negative correlation. Suppose now that $x_2$ is sensitive to temperature, whereas $x_1$ is sensitive to soil pH. Variation in soil pH results in positive correlation between $x_1$ and $x_2$, temperature variation in negative correlation, and a combination of the two could give very confusing results.

We now proceed to consider a few semirealistic cases. It must be emphasized that these are not being offered as full models of the situations in question but only as illustrations of how qualitative analysis may be applied.

**Case 1. Control of Insect Pests of Crops with Insecticides.** Here $P_1$ is the crop plant that is self-damping because of crowding; $H_1$ is the herbivore, the pest species that eats the plant; $P_a$ is the specialized parasitoid, usually a wasp, which kills only $H_1$; $P_r$ is a generalized predatory insect or spider that eats $H_1$ and also $H_2$, a herbivore that feeds on other plants. This scheme is shown in FIGURE 5, where $I =$ insecticide.
An increase in the level of insecticide use, $I$, has a direct negative effect on $P_a$, $P_r$, and $H_1$. But the direct path $[I, H_1]$ has as its complement the subsystem

which has zero feedback because the parasitoid is isolated. Therefore this path has no effect. The other two paths, $[I, P_a, H_1]$ and $[I, P_r, H_1]$ are both positive (a product of two negative links) but only the first of these has a nonvanishing complement. Therefore, the final result of adding insecticide at a new, constant dosage is to increase the herbivore species. Any events further along the path $(P_r, H_s, P_3)$ likewise leave $H_1$ unaltered in numbers. For instance, if $H_2$ increases this increases $P_r$, which eats more $H_1$, reduces the food supply for $P_a$, and results finally in a shift in the cause of death of $H_1$ but not its numbers. The only way to affect $H_1$ is through $P_a$. Thus, if a species is partly controlled by a specialized predator, its equilibrium level is very stable, and control through its specialized predator is more effective than by way of generalized insectivores or insecticides.

Of course, this argument holds only if the system remains at equilibrium. High enough dosages of insecticide can destroy the equilibrium and wipe out the species.

**Case 2. Contamination of a Lake.** Here we consider a system with two nutrients that are washed into the lake from outside, nitrate ($N$) and phosphate ($P$). We distinguish two kinds of algae: the green algae use phosphate and nitrate and are sensitive to a toxin released by the blue-greens.

The blue-greens release nitrate into the lake but depend on phosphate (Figure 6).

If additional nitrate is washed into the lake, the effect on nitrate level depends

![Figure 5. A model of a cultivated field community with Plant ($P_3$), herbivores ($H_1$ and $H_2$), predators ($P_r$ and $P_2$), a parasite ($P_a$), and insecticide ($I$).](image)

**Figure 6.** Nutrients and organisms in a lake. A two-nutrient, two-alga community, with $N =$ nitrate, $P =$ phosphate, $G =$ green algae and $BG =$ bluegreen algae.
on the feedback of the complementary system $P$, $BG$, $G$. This consists of a single positive loop of length three. Therefore, adding nitrate results in a decrease in the nitrate level, an increase in green algae at the expense of blue-greens, and an unaltered level of phosphate. The addition of phosphate increases the blue-green algae, reduces the green algae, and increases the level of nitrate. The phosphate level itself is unchanged. If we now add a herbivore that can eat green but not blue-green algae, the graph becomes as shown in Figure 7.

![Figure 7](image)

**FIGURE 7.** Here we have added the herbivore $H$, which eats only the green algae. (See Figure 6.)

Now the complement to $N$ is a system with negative feedback:

![Diagram]

Thus now an increase in nitrate input raises the nitrate level at equilibrium, increases the herbivore population, and has no effect on the other variables. An increase in phosphate input increases nitrate and blue-green levels as before and leaves phosphate unchanged. But instead of decreasing the level of green algae, it reduces the herbivore population and leaves the green algae unchanged. Finally, the addition of toxic substances that kill the herbivores (a negative input to $H$) will increase green algae, reduce the blue-green and nitrate levels, and leave phosphate unaltered.

These cases are simplified models modified from a large-scale modeling of aquatic systems that I am working on in collaboration with Patricia Lane of the Kellogg Biological Laboratory.

**Case 3.** This model is not intended as a realistic description of hypoglycemic diabetic mechanisms. Instead, it is intended to illustrate the possibility of integrating physiological and psychological processes. It would be impossible at present if we insisted on full quantification.

In this model we include the variables insulin ($I$), blood glucose ($G$), and adrenalin ($E$), and we lump the subjective states associated with stress, hypoglycemia,
and high adrenalin levels into the single variable $A$. This is not necessary; a more serious analysis would make more distinctions among subjective states.

All of the causal links are the familiar ones. Glucose increases insulin while insulin reduces glucose; the stressed subjective state can come from external stimuli, high levels of adrenalin, or hypoglycemia; stress calls out adrenalin; and all of the variables are self-damped. The self-damping of adrenalin and insulin depends on the rate of metabolic breakdown of these substances and are therefore physiological parameters of the individual. The self-damping of the glucose depends on the equilibrium level of insulin and the effectiveness of the insulin. The self-damping of the subjective state $A$ is more difficult to describe. It is the recovery from anxiety, the "pulling oneself together" that depends on the individual psychology and also can be influenced by tranquilizers and other drugs. The model is shown in Figure 8.

First, consider the consequences of increasing the rate of input of glucose either by eating or by increased conversion of glycogen. Then, following the rule of Equation 6, the level of anxiety will decrease (both by the direct path and through $E$). The level of adrenalin will decrease. But the effect on the level of glucose itself and on insulin will be positive only if the subsystem $[E, A]$ has a net negative feedback. The loop of length two is positive, but may be outweighed by the product of the self-damping terms for $E$ and $A$. The self-damping term of $A$ is especially variable from person to person. In an anxious person for whom recovery from stress is slow, the self-damping of $A$ may be weak enough so that the $E, A$ subsystem has net positive feedback. Then we have the anomalous physiological results that increasing sugar input reduces the equilibrium level of blood sugar and reduces the insulin level, and that increasing insulin input results in an increase in equilibrium blood sugar and a reduction of insulin. However, the role of insulin in increasing adrenalin and anxiety levels is unaffected.

This particular response, due to the positive feedback of the $E, A$ subsystem, may be reversed if either tranquilizers or changes in psychological mechanisms strengthen the self-damping of $A$. But it need not be necessarily of psychogenic origin. Any mechanism that reduces the rate of breakdown of adrenalin, increases its release for a given hypothalamic stimulus, or increases the sensitivity of the subjective state to adrenalin would also increase the positive feedback of the subsystem.

\begin{figure}
\centering
\includegraphics{figure8}
\caption{A model of part of carbohydrate metabolism. $E =$ epinephrine concentration; $G =$ blood glucose; $I =$ insulin; $A =$ level of anxiety.}
\end{figure}
But if a reduction in glucose release increases anxiety in a typical hypoglycemic bout, laboratory tests during the anxiety state should show low blood sugar. If the primary event, however, is an increase in external stress, the increase in $A$ results, via adrenalin, in an increase in blood sugar. Now laboratory tests would show a positive correlation between anxiety states and blood sugar. Since external stress and reduced glucose intake may themselves be positively correlated (for instance, coming into the hospital for a fasting glucose test itself causes anxiety), the opposite effects are confounded, resulting in normal or irregular, ambiguous laboratory findings.

If persons with primary pancreatic insufficiency also suffer from slow damping of anxiety, the typical diabetic symptoms may fail to appear; the weakening of the pancreas will be misread by a secular increase in blood insulin and fall in blood sugar, while, on the other hand, pancreatic tumors can result in falling insulin levels, rising blood sugar, and greater anxiety.

So far we have examined changes in the equilibrium level. Two other properties of the system are worth noting: first, the $E, A$ system may be unstable because of positive feedback exceeding negative feedback when the self-damping of $A$ is weak. However, it is embedded in a larger system including blood glucose, insulin, and other variables, and the whole may be stable. What happens is that when the $A, E$ component is disturbed, variations in sugar level are part of the mechanism that reestablishes equilibrium. Therefore, if a major effort to stabilize blood sugar succeeds in preventing variation, blood sugar ceases to be a variable in the system and the $A, E$ subsystem becomes an isolated whole system that is unstable. Hence, fluctuations in blood sugar act as a cushion for variations in psychological stage, and an overzealous stabilization can have serious psychological consequences.

But if the $A, E$ system may have either positive or negative feedback it can also have values close to zero. When it is near neutral (or passive) stability, it itself stabilizes the equilibrium level of blood sugar.

A second problem relates to the overall stability of the whole system. Oscillatory instability may arise if the negative loop of length 3, $AG, A, E$, is strong enough. Equation 16 states this condition more precisely. From our model, the condition for stability suggests that an unstable, oscillatory system may occur if the emotional state $A$ is especially sensitive to fluctuations in blood sugar or adrenalin, if the self-damping (rate of breakdown) of adrenalin or insulin is slow, or if the response of blood sugar to these substances is weak.

Since the stability condition involves inequality relations among all the parameters, the locus of an abnormal response need not be a single lesion. Instead, diffuse small differences in the constants of the system can result in pathologies that are impossible to locate at a specific point. In the end we will have to recognize "positive feedback at level two" or "excess negative feedback of length three" as the primary pathologies, and develop the techniques for their measurement and management.

Since the last example given refers to a medical condition, it is necessary to repeat that it is not presented as a model of diabetes but rather as an approach to a complex problem that I hope to be able to follow up some day in collaboration with medical researchers.

**Evolutionary Processes**

Although systems that represent very different physical realities may be represented by graphs in the same way, they will evolve very differently. In the systems of interacting physiological processes within the organism, natural selection acts
only by way of the effect on the whole. Thus, if we know the adaptive significance of the kinetic constants, the effect on the probability of survival and reproduction of anxiety or high blood sugar or slow fluctuations in these, we could identify the direction of its evolution. The significance may be a function of average levels, stability properties, or transient events after perturbation. There is no ground \textit{a priori} to expect selection to favor any particular property. For example, dynamic stability may be either advantageous or harmful compared to low-frequency oscillations.

In ecological communities such as those described in cases 1 and 2, some of the variables are species undergoing selection in terms of their own population dynamics, whereas the kinetic constants of phosphate and nitrate are not under the control of these ions themselves. There will be selection within the planktonic species for those traits which increase the rates of survival and reproduction for their bearers. Whether, however, this increases the population level of the species, reduces it, or leaves it unchanged depends on the complementary subsystem. For instance, in case 1 a higher reproductive rate in the herbivore \( H_1 \) will be selected for within \( H_1 \), but since its complementary function is zero, the level of \( H_1 \) will remain unchanged. On the other hand, a reduced probability of being parasitized is a "kinetic constant" that depends on both the parasitoid and its host, a shared parameter. Since it appears as a negative change in the equation for the rate of growth of the parasitoid population, selection in the host will be represented as a direct positive input to \( H_1 \) and also as a negative input to the parasitoid. The path from the parasitoid to its host is also negative, and the complement of this path has ordinary negative feedback. Thus both genes for parasite avoidance and for higher reproductive rate will be selected for, but the former will increase the herbivore population while the latter leaves it unchanged.

Selection within the species is determined by the relative fitness of genotypes within the species. But the consequences of that selection for the species or for the whole community depends on the network structure as a whole. Traits that may have equal selective value within the species may appear in quite different ways in the network. There is no necessary relation between selection within a species and the consequences for the whole (or even for that species). In particular, there is no justification for the belief that selection results in greater efficiency or stability or in any other civic virtue.

The problem of environmental management posed in case 1 is a combination of organismic and community types of evolution. While each species is evolving according to its own population dynamics, the agronomist has control of several parameters (represented in the example by insecticide level) that are adjusted with the aim of maximizing some whole-system goal that, depending on the broader context, may be yield, profit, or insecticide sales. In any case, natural selection continues in even the most unnatural situations and may work either in the same direction as artificial selection and environmental management, or to thwart it. It must always be taken into account.

\textbf{Conclusions}

We have argued that complex systems may be studied qualitatively by an examination of network properties of the wholes. This procedure may be the only one available in partially specified systems. It is also advantageous for the understanding of what is taking place even when precise measurement is possible, and indeed helps determine what should be measured. Our emphasis on the holistic and qualitative is intended not as a complete program for science but as a corrective for the one-sided analytical quantitative approach that is still dominant.
REFERENCES