Toward the Measurement of Ecological Integrity

Robert E. Ulanowicz

The notion of ecosystem integrity has repeatedly appeared in the legislation of Canada and the United States in the absence of serious consideration as to exactly what the term means. The charge to the Global Integrity Project has been to give concrete definition to "ecological integrity" and to suggest how the concept might be implemented (Westra 1994).

Accordingly, the Project has identified the concept of ecological integrity with at least four attributes: (1) system "health," that is, the continued successful functioning of the community; (2) the capacity to withstand stress; (3) an undiminished "optimum capacity" for the greatest possible ongoing developmental options; and (4) the continued ability for ongoing change and development, unconstrained by human interruptions.

The main question to be considered here is whether these attributes can be anchored in sound ecological theory. By addressing this issue, I also hope to address the larger controversy of whether the current body of ecological theory is indeed sound—for critics of the concept of ecological integrity in particular, and ecosystem theory in general, are indeed numerous and highly articulate. Shrader-Fréchette and McCoy (1993), for example, maintain that "[e]cology is not a science enough to provide testable laws." Furthermore, Shrader-Fréchette asserts that "[i]ntegrity is stipulative" and therefore possesses no phenomenological or operational character (personal communication to L. Westra).

Similarly, Mark Sagoff (chapter 4, this volume) contends that "[s]cience is too changeable, insecure and internally conflicted to be able to provide any guidance for environmental morality or public policy." Both Shrader-Fréchette and Sagoff take their cues from Peters (1991), who in his Critique for Ecology chided ecologists for indulging in theory that is (1) not operational, (2) tautologous and therefore nontestable, and (3) not predictive.

These are serious charges leveled by respected individuals and are not to
be taken lightly. Most will acknowledge that there are credible reasons for such misgivings. But is ecology actually in such dire straits that the best its practitioners can do is to stipulate characteristics that cannot be quantified or made elements of testable theory? The majority of those who make up the Global Integrity Project are not so pessimistic, but it remains to respond to these challenges in a succinct but convincing manner. We offer three points by way of apology:

1. Many critics of ecology theory assume an unnecessary dichotomy between determinism and stochasticity as these characteristics pertain to ecology. Indeed, this “error of the excluded middle” is widespread throughout science in general. Most practicing scientists believe that strict Newtonian causality prevails over macroscopic phenomena, and that indeterminacy, if it exists at all, is confined to the netherworld of molecular and quantum phenomena. Philosophers, such as Charles Saunders Peirce and Karl Popper, however, believe the world to be causally open at all levels of the hierarchy. Furthermore, Popper (1990) warns that we will never achieve an “evolutionary theory of knowledge” unless we first revise our views on basic causality. He advocates a shift from the Newtonian notion of “force” to a more general concept that he calls “propensity.” Elsewhere, I have attempted to elaborate a new view of causation in living systems (Ulanowicz 1997, 1999) and have suggested ways to make the idea of propensity operational in ecology and the other life sciences (Ulanowicz 1996).

2. Many detractors of ecology cling to an unwarranted refusal to admit directionality into ecological discourse. The ambiguities surrounding whether directionality is admissible into evolutionary theory are somehow thought to be universal. This extrapolation, of course, is consistent with the modernist view that only material and efficient causes exist in nature—whence any perceived directionality must be epiphenomenal or illusory. According to this viewpoint, to admit of essential direction anywhere in the living realm is to take the first, irreversible step down the slippery slope into the depths of teleology! Mayr (1992), among others, countered by asserting that biology without directionality is puerile. Again, it is the “error of the excluded middle” not to consider directionality, like Mayr’s “teleonomy,” that are distinctly nonteleological in essence. Elsewhere, I have posited the existence of a “telos” in ecosystem development as a weak form of final cause that falls far short of teleology (Ulanowicz 1997).

3. Skeptics are not convinced that ecological integrity can become an operational tool. They choose to ignore the obvious successes that quantitative indices, such as Karr’s (1986) index of biological integrity (IBI), have
contributed to ecological management. Perhaps they disregard IBI as being “stupulative” because they feel that its connections with basic science are not as explicit as might be desired. But other quantitative assessments of ecological integrity exhibit strong links to the corpus of accepted science and phenomenology (e.g., Svirezhev and Svirejeva-Hopkins 1998, Jørgensen 1997).

Space does not permit a full exposition of all three apologies, and the reader interested in either of the first two counterassertions is urged to consult the references cited. The remainder of this chapter will bear upon the final point, and, in particular, will elaborate current efforts to invoke the theories of networks, information, and economics to identify tools that will give palpable quantitative form to all four aspects of ecological integrity.

Quantifying Ecosystem Performance

As defined above, ecological integrity explicitly subsumes the notion of "health." Costanza (1992) identifies the health of an ecosystem as its ability to sustain its structure and function over time in the face of external stress (Mageau et al. 1995). The function, or vigor, of a system relates to its overall level of activity in processing material and energy. Its structure, or organization, refers to how effectively its various processes are linked to each other. Together vigor and organization specify the system’s level of performance as pertains to aspect 1 under the definition of ecological integrity. The third point, resilience to perturbation, corresponds to aspect 2.

The task now before us is to quantify the levels of vigor, organization, and potential for resilience exhibited by any arbitrary ecosystem. Toward this end we will assume that these attributes are identifiable properties of the network of trophic exchanges of material or energy that occur within the system. Inevitably, some will object that the magnitudes of palpable flows do not make explicit the multitude of signs and intricate behaviors that influence and guide those flows. This is true; however, such detailed phenomena do function as implicit constraints that modulate and direct the perceived exchanges. It will be argued here how the effective magnitudes of these constraints can be measured in the absence of any knowledge of their explicit details.

We begin by identifying four different types of exchanges that occur in an ecosystem (Figure 6.1). The rate at which taxon $i$ contributes as prey to the sustenance of predator $j$ will be denoted by $T_{ij}$, where $i,j = 1,2,3, \ldots, n$. The large majority of transfers in an ecosystem usually consist of such intrasystem exchanges, but all ecosystems also are open to the external
The four types of exchanges used to quantify ecosystem flow networks. (\(T_{ij}\)) internal flow from arbitrary compartment \(i\) to any other taxon, \(j\), of an \(n\)-compartment system. (\(T_{0i}\)) External inputs to the arbitrary compartment \(i\). (\(T_{ik+i}\)) Export of usable resources from unit \(i\). (\(T_{ik+i}\)) Dissipation of resources from system element \(i\).

world. There are three categories of exogenous exchanges: (1) inputs from the external world, signified by \(T_{0i}\), (2) exports of medium that can be used by systems of similar size, denoted by \(T_{ik+i}\), and (3) dissipation of medium into a form that is useless or of marginal value elsewhere, signified by \(T_{ik+i}\).

Once all four types of flows pertaining to all the important taxa of an ecosystem have been measured, the values may be either arrayed into matrices and vectors or depicted in schematic form as "box and arrow" diagrams. Figure 6.2 is a schematic example of the trophic exchanges of carbon (in mgC/m²/yr) among the 36 most significant taxa of the mesohaline (moderately salty) ecosystem of Chesapeake Bay (Baird and Ulanowicz 1989). Sets of such measurements allow one to quantify the degrees of vigor, organization, and resilience inherent in the prototype ecosystems.

Of the three attributes, vigor is easiest to quantify. One measure of system vigor would be the simple sum of all the individual exchanges that occur in the system—a quantity known in economic theory as the "total system throughput," \(T\), where

\[
T = \sum_{ij} T_{ij}
\]

The total system throughput is related to (but not identical to) the ubiquitous "gross domestic product" that is a common item in the daily news.
FIGURE 6.2

Estimated flows of carbon (mg/m²/y) among the 36 principal components of the mesohaline Chesapeake Bay ecosystem (Baird and Ulanowicz 1989). Circles represent autotrophic system elements (plants); hexagons, heterotrophic taxa (fauna); and “birdhouses,” nonliving storages. Numbers inside each box are the standing stocks in mg/m².
Organization requires somewhat more effort to quantify. We take as a working definition of organization, "the degree of constraint that guides a typical flow in the system." That is, medium leaving a particular taxon does not flow arbitrarily to just any other taxon. Only a limited number of other taxa are capable of consuming a given source. The actual distributions of flow from various prey to their specific predators are the outcomes of organism morphologies and behaviors. As a result, when the system is viewed on the whole, certain pathways will tend to dominate the network of exchanges. Constraints that channel flows along these major routes usually are the outcomes of interspecific competitions and/or indirect mutualisms (autocatalysis) among the participating taxa.

Autocatalysis as an agency that drives ecosystem development (the increase of organization) deserves more attention than can be paid it here. Suffice it to mention that, when autocatalysis acts against a background of indeterminate kinetics, it comes to exhibit properties that are decidedly nonmechanical in nature (Ulanowicz 1977). Among other traits, autocatalysis is a source for what physicists call "symmetry-breaking" that establishes the direction in which the system will develop.

How, then, to quantify the constraints that engender system organization? As with so many approaches in mathematics, one first considers the opposite of what is being defined, namely, freedom or indeterminacy. Thus information theory, as pioneered by Boltzmann (1872) and Shannon (1948), begins by postulating that the indeterminacy of phenomenon $A_i$ is related to the probability of its occurring, $p(A_i)$, through the expression

$$[-k \log p(A_i)]$$

where $k$ is a scalar constant, and the base of the logarithm is taken to be 2. Because $p(A_i)$ is a number between 0 and 1, the indeterminacy is sure to be a positive quantity (the logarithm of a fraction is a negative number), and it becomes very large when $A_i$ is rare or very small when the outcome is almost certainly $A_i$.

In writing $p(A_i)$, we do not take explicit account of any constraints upon the occurrence of $A_i$. If, however, $A_i$ happens to be in proximity to some other event, $B_j$, and the effect of $B_j$ is to constrain, to a degree, the probability of $A_i$, then the conditional probability of $A_i$ under the influence of $B_j$ is written $p(A_i|B_j)$. The revised indeterminacy thus becomes,

$$[-k \log p(A_i|B_j)]$$

On average, the indeterminacy of unconstrained $A_i$ should exceed its indeterminacy when under the constraint of $B_j$, so that the difference,
should measure the degree of constraint that $B_j$ exerts upon $A_i$. Because the difference between two logarithms may be expressed as the logarithm of the quotient, this last expression may be rewritten,

$$k \log \left( \frac{p(A_i|B_j)}{p(A_i)} \right)$$

Those familiar with probabilities will recognize that this last expression is related via Bayes' theorem to the (symmetrical) joint probability that $A_i$ and $B_j$ will co-occur:

$$\frac{p(A_i|B_j)}{p(A_i)} = \frac{p(A_i, B_j)}{p(A_i)} = \frac{p(B_j|A_i)}{p(B_j)}$$

Whence, the measure of the constraint of $B_j$ upon $A_i$ is equal to that which $A_i$ exerts upon $B_j$. One thus may speak of the mutual constraint that $A_i$ and $B_j$ exert upon each other, as given by the expression,

$$k \log \left( \frac{p(A_i, B_j)}{p(A_i)p(B_j)} \right)$$

This last expression pertains to any arbitrary pair of events $A_i$ and $B_j$. We are more concerned, however, with the overall degree of constraint present in the system as a whole. Accordingly, we calculate the average mutual constraint by multiplying each pairwise mutual constraint by the probability of its occurrence $p(A_i, B_j)$ and summing the products over all combinations of $i$ and $j$. Because constraint is operationally indistinguishable from information, we call the result of this averaging process the average mutual information (AMI) and write it as

$$AMI = k \sum_i \sum_j p(A_i, B_j) \log \left( \frac{p(A_i, B_j)}{p(A_i)p(B_j)} \right)$$

It remains to specify how $A_i$ and $B_j$ might relate to a network of trophic exchanges. We therefore identify $A_i$ with the event, "a quantum of medium leaves taxon $i$" and $B_j$ with "a quantum of medium enters taxon $j". This allows us to estimate all the probabilities constituting AMI in terms of the intercompartamental flows, $T_{ij}$, and the total system throughput, $T$, as follows:

$$p(A_i, B_j) = \frac{T_{ij}}{T}$$
\begin{align*}
p(A_i) &= \left( \frac{\sum T_{ij}}{T} \right) \\
p(B_j) &= \left( \frac{\sum T_{ij}}{T} \right) \\
\text{and} \\
p(B_j \mid A_i) &= \frac{T_{ij}}{\sum_q T_{iq}}
\end{align*}

Substituting these estimators into the expression for AMI yields the formula

\[ \text{AMI} = k \sum_{i=0}^{n} \sum_{j=1}^{m} \left( \frac{\sum T_{ij}}{T} \right) \log \left( \frac{T_{ij}}{\sum_k \sum_q T_{iq}} \right) \]

Information theory guarantees that AMI \( \geq 0 \).

It bears repeating that whenever all the \( T_{ij} \) are known from the network of trophic exchanges, it then becomes a straightforward task to calculate an AMI for the system. That this measure does indeed quantify the degree of constraint in a system can be seen from Figure 6.3. All three networks of four "taxa" have identical total system throughputs (96 units). In Figure 6.3a, however, there is minimal constraint about where medium currently in a particular box will flow next. It is equally likely to flow to any of the three other taxa, or to remain where it is. The AMI of this configuration is identically zero. In Figure 6.3b outputs from any box are constrained to flow to only two other boxes. The measure of this constraint is \( 1k \) bits. (A "bit" is the unit associated with the resolution of a single binary indeterminacy [fork].) In Figure 6.3c the system is maximally constrained. Medium leaving any taxon may flow to only one other taxon. The AMI rises to \( 2k \) bits.

It is awkward to express the values of AMI in terms of multiples of the scalar constant, \( k \). The usual practice in information theory is to choose a base for the logarithms (2, \( e \), or 10), set \( k = 1 \), and call the resultant units "bits," "nats," or "hartleys," respectively. By retaining the scalar constant, however, we wish to emphasize that it can be used to impart physical units to the information index that it multiplies (Tribus and McIrvine 1971). Accordingly, we choose to normalize AMI in a very natural way by setting \( k = T \), i.e., we use the system vigor to scale its corresponding measure of
organization. To highlight that the resultant product is a physical measure, we rename it the system ascendency (Ulanowicz 1980) and denote it simply by $A$. In the ascendency we have incorporated both aspects of system performance,

$$A = T \times \sum_{i,j} \left( \frac{T_{ij}}{T} \right) \log \left( \frac{\sum_{i} T_{ij}^T}{\sum_{j} T_{ij} \sum_{i} T_{ji}} \right)$$

or

$$Ascendancy = Vigor \times Organization$$

It can be demonstrated that the magnitude of system ascendency is abetted by greater speciation, more specialization, further internalization, and increased cycling (Ulanowicz 1980). These four attributes represent groupings
of the 24 system properties that Eugene Odum (1969) identified with more highly developed ecosystems. We arrive, then, at the following phenomenological principle:

In the absence of major perturbations, ecosystems exhibit a propensity toward configurations of ever greater ascendency.

Quantifying Potential for Resilience

It was noted above how the AMI was maximal for the configuration depicted in Figure 6.3c. Such a statement was possible, because information theory reveals that the AMI possesses an upper bound. Without going into detail (see Ulanowicz and Norden 1990), an appropriate upper bound on the ascendency may be taken as

\[ C = -\sum_{i} \sum_{j} T_{ij} \log \left( \frac{T_{ij}}{T} \right) \]

where \( C \) is termed the system capacity. Information theory guarantees that \( C \geq A \geq 0 \). This allows one to define a system's overhead as \( \Phi = C - A \), where \( \Phi \geq 0 \). In terms of the system flows, the overhead works out to be

\[ \Phi = -\sum_{i,j} T_{ij} \log \left( \frac{T_{ij}^2}{\sum_j T_{ij} \sum_i T_{ij}} \right) \]

The system overhead is complementary to the ascendency. Whereas ascendency gauges how well the internal constraints cause the system to perform, the system overhead gauges all the inchoate, inefficient, and redundant degrees of freedom that the system retains.

To recapitulate what has just been accomplished, the system capacity can be parsed into two distinct and complementary components—the ascendency and the overhead (Figure 6.4). Whenever external conditions are relatively benign, there is a propensity for the system ascendency to grow at the expense of its overhead, i.e., the ecosystem “ratchets” itself toward ever higher and more efficient performance.

The real world is never so benign, however. External perturbations always arise, and a system with very high performance is “brittle” and subject to collapse (Holling 1986). It can successfully adapt to novel perturbation only if it retains a sufficient repertoire of degrees of freedom. That is, overhead, which is inimical to performance under undisturbed conditions, becomes essential if the system is to recover from trauma. Overhead becomes a prerequisite for system resilience, creativity, and persistence (Ulanowicz
Graphical representation of the segregation of development capacity (complexity) into disordered and organized components (overhead and ascendency, respectively).

1980, 1997). Engineers have always been cognizant of the antagonism between efficiency and reliability, and that complementarity is given quantitative form in the relationship between ascendency and overhead.

Quantifying Integrity

The countervailing attributes of ascendency and overhead nonetheless sum to yield the system capacity. In other words, the capacity subsumes the system's capabilities to perform well, to be resilient, and to develop in a creative fashion. We recall points 3 and 4 in the definition of integrity: an undiminished "optimum capacity" for the greatest possible ongoing developmental options, and the continued ability for ongoing change and development, unconstrained by human interruptions. We recognize, therefore, in the decomposition of the capacity, all four elements of ecosystem integrity, namely,

\[
C = A + \Phi
\]

\[
\begin{align*}
\text{Capacity for development} &= \\
\text{Healthy Performance} + \text{Resilience and creativity}
\end{align*}
\]

(3 and 4) \hspace{1cm} (1) \hspace{1cm} (2)

If one has access to sufficient data to assemble a network of trophic interactions, one is equipped to assess ecosystem integrity in a quantitative way. Wulff and Ulanowicz (1989), for example, constructed networks of the ecosystems of the Baltic Sea and the Chesapeake Bay, respectively, with the intention of using the networks to compare their relative trophic status. The picture that emerged from the exercise was somewhat surprising. Despite
being much more vigorous ($T$ in Chesapeake is about fourfold that of the Baltic), the Chesapeake showed visible signs of heavy impact. The flow diversity ($C/T$) remained higher in the Baltic ecosystem, which also exhibited a higher relative proportion of ascendancy ($A/C$). The evidence for impact on the Chesapeake ecosystem was no surprise to those who study and manage its resources, but the fact that the oligohaline Baltic seemed to have more integrity than the mesohaline Chesapeake raised numerous eyebrows among the Baltic Sea research community.

A final observation on capacity: The diversity of exchanges that make up the system capacity obviously is limited by the diversity of the participating taxa. Hence, any decrement in the biodiversity is likely to have detrimental consequences for the system capacity, i.e., its integrity. Ascendancy theory, therefore, provides a theoretical link between the preservation of biodiversity and the maintenance of biotic integrity (which in its turn encompasses both system functioning and sustainability).

**Summary and Conclusions**

Despite the protests of its detractors, ecosystem integrity has already become an operational concept. The IBI is currently being used to evaluate the integrity of numerous aquatic ecosystems, and plans are afoot to extend the measure to cover marine, estuarine, and terrestrial communities.

Ecosystem ascendancy and its related measures derive quite naturally from a mix of disciplines, tightly weaving the concept into the existing web of phenomenological sciences—most especially, thermodynamics. Ascendancy measures have already been employed to assess the relative integrities of real systems, and one can anticipate many more such comparisons as new software makes the task of assembling trophic networks ever easier (Christensen and Pauly 1997).

That ascendancy variables fit the notion of ecosystem integrity hand-in-glove derives in part from a certain overlap in the goals of those who developed both concepts. The correspondence between the capacity and its components and the several aspects of ecological integrity permits one to assess quantitatively how the aspects of integrity stand in relation to one another in any particular system.

The foregoing considerations on the relationship between ascendancy and integrity did not include a riposte to the criticism that ecological theory deals only in tautologies and does not provide testable predictions. Ascendancy as described above is a highly aggregated measure. In order to apply the concept to more specific hypotheses, the notion must be elaborated along several lines. Toward this end, ascendancy has recently been extended
to apply to systems that vary in time (Pahl-Wostl 1992) and space (Ulanowicz 2000). Furthermore, ascendency has been revised to incorporate species stocks in a natural way (Ulanowicz and Abarca 1997).

This latter advance has allowed the use of the ascendency formula to investigate problems of nutrient limitation and control. For example, if $B_{ik}$ is the amount of mass of element $k$ in taxon $i$, and $T_{ijk}$ is the amount of $k$ flowing from $i$ to $j$, then one may calculate the sensitivities of the ascendency with respect to either of these variables in the hope that such sensitivities will provide clues in the search for limiting nutrients and processes (Ulanowicz and Baird 1999).

In particular, when one calculates

$$\left( \frac{\partial A}{\partial B_k} \right)$$

the derivative of the ascendency with respect to each nutrient stock, the nutrient that is present in least proportion in each taxon emerges as the one with the greatest sensitivity. That is, the ascendency principle can be shown to subsume Liebig's Law of the Minimum. Demonstrating that a general principle encompasses a more specific, acknowledged law is certainly a point in favor of the covering law. But there is more.

Liebig's Law provides no clue as to what the limiting source of a limiting element might be. By calculating the sensitivity

$$\left( \frac{\partial A}{\partial T_{ijk}} \right)$$

one is able to identify the limiting source as the one that is being depleted at the fastest relative rate. Often, this source is not identical to the one that is supplying the limiting element at the fastest rate (Ulanowicz and Baird 1999). Thus the identification of the limiting source is a "prediction" of the ascendency principle that can be tested experimentally.

So long as one is careful to build new theory with identifiable connections to the existing body of science, the likelihood remains good that progress will not become mired in untestable tautologies. In view of what has been reported here, the prospect for a robust ecosystems theory no longer should be regarded with pessimism, but rather with hope.

ACKNOWLEDGMENTS

The author wishes to acknowledge partial support for this work from the U.S. Environmental Protection Agency Multiscale Experimental Ecosystem
Research Center (Contract R.819640) and the U.S. Geological Survey Program for Across: Trophic Levels System Simulation (Contract 1445CAC09500093).

REFERENCES


Toward the Measurement of Ecological Integrity


