Quantifying the Complexity of Flow Networks: How many roles are there?

ALEXANDER C. ZORACH AND ROBERT E. ULANOWICZ
University of Maryland Center for Environmental Science, Chesapeake Biological Laboratory, PO Box 38, 1 Williams Street, Solomons, Maryland 20688-0038; e-mail: cazort@cazort.net, ulan@cbl.umces.edu

Received July 26, 2002; revised November 19, 2002; accepted November 19, 2002

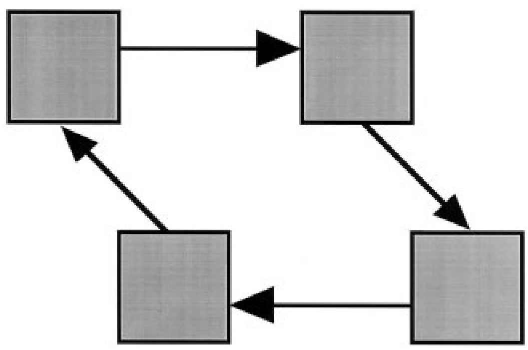
Weighted flow networks are structures that arise naturally when analyzing complex systems. The countable properties of unweighted networks are not easily generalized to weighted networks. One candidate measure of complexity is the number of roles, or specialized functions in a network. It is easy to identify the number of roles in a linear or cyclic unweighted network. There is only one logically consistent way to generalize the measures of nodes, flows, connectivity, and roles into weighted networks, and these generalizations are equivalent to indices derived from information theory and used by ecologists since the late seventies. Data from ecosystem networks suggests that ecosystems inhabit a narrow window of the parameter space defined by these measures.

Key Words: networks; information theory; connectivity; ecosystem, role

The shift in paradigm from conventional science to systems thinking has been characterized by a shift in focus from parts to wholes and from physical structures to processes. Capra [1] made the bold assertion that "Systems thinking is always process thinking." This shift is occurring in all fields of science, and changes in one field influence other disciplines. For example, advances in quantum physics have convinced those in other sciences of the importance of indeterminism and randomness. At the other end of the spectrum, ecology and the other macro-scale sciences have developed ideas that can be applied to the so-called hard sciences. The core of what ecologists have to offer other sciences is the flow network approach. Rather than limiting attention to fixed properties such as biomass, populations, or nutrient levels, ecologists have found that a more fecund theory can emerge from the study of changes in these variables across divisions of space, time, species, or other categories. A system is then described as a network of flows between nodes. In ecology, the flows often measure carbon transfers, but other nutrients can be measured as well.

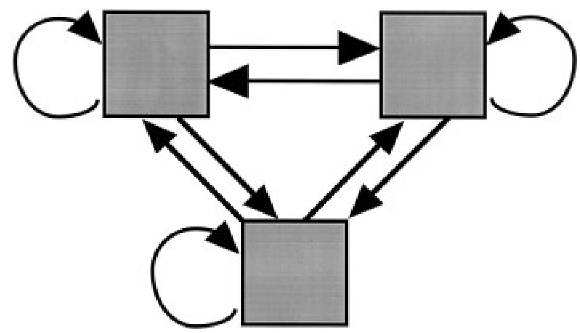
The applications of flow networks are numerous in fields such as ecology [2, 3], economics [4], and of course, engineering [5–7]. It should be noted that although weighted flow networks are identical in form to weighted digraphs, the convention in the literature is that digraph weights represent costs or lengths, and thus larger weights on an edge indicate *lesser* significance. In flow networks, larger weights on a flow represent larger flows and thus *greater* importance. The mathematical similarity of digraphs and

FIGURE 1



An unweighted cyclic network, with 4 nodes, 4 flows, a connectivity of 1 flow per node, and 4 roles.

FIGURE 2



An unweighted network with 3 nodes, 9 flows, a connectivity of 3 flows per node, and 1 role.

networks is superficial; the bulk of the literature on weighted digraphs is not useful in the analysis of weighted networks. The massive literature on neural networks is not of much help either; in spite of an awareness of the need to address the issue of weighting [8], many scientists still use binary on/off networks, partly because of the successes of such networks [9].

Ecologists, however, need the weights, and accordingly, most of the work on flow networks comes from ecology. Ecologists have developed a set of variables, based on information theory, that quantifies the growth and development of ecological or economic flow networks [10]; these variables have the potential to give quantitative expression to many of the qualitative observations by ecologists regarding the development of ecosystems. This theory is difficult to use, however, because its measures do not have direct interpretations in countable quantities. Rather than start with this theory, we will first look at countable properties of unweighted flow networks, and then derive generalizations of these measures that are equivalent to the information theoretic measures.

UNWEIGHTED NETWORKS

An unweighted flow network is a collection of nodes and directed flows of equal weight. One can divide the number of flows in the network by the number of nodes to obtain the connectivity, the average number of flows per node in the network. (Some use "connectivity" to refer to the fraction of possible flows that are realized; this measure has dimension flows per node squared [11]) In Figure 1, the connectivity is 1 flow per node, so on average, each node has one flow leaving and one entering. Because the network is symmetrical, the average connectivity is equal to the connectivity of each node.

Let us now make a few definitions: Let F be the number of flows. Let N be the number of nodes. Let $C = F/N$ be the connectivity, measured in flows/node.

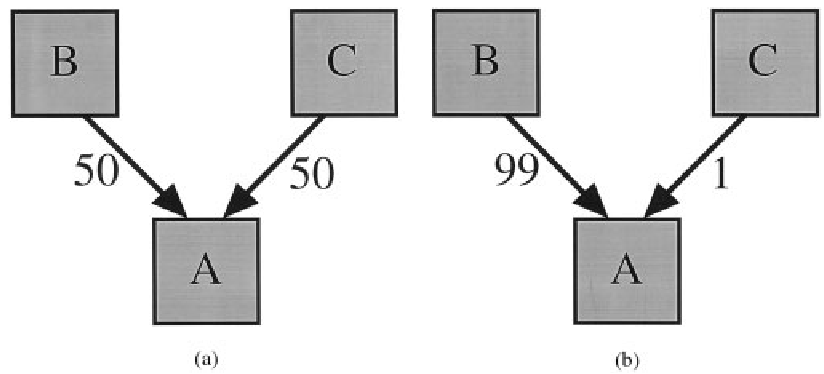
Already, with just F, N, and C, we have scientifically useful concepts. Numerous investigators in diverse fields have shown that network connectivity relates to the stability of a dynamical system. For example, Kauffman has related connectivity of boolean logic networks to their stability [12, 13]. May was one of the first to argue that the stability of a complex system is related to the connectance of a trophic web [11]; others disputed the form of the relationship, but generally agreed that connectivity relates to stability [14, 15].

We now introduce one more measure. Let $R = N/C = N^2/F = F/C^2$ be the number of roles in the network. A role is, loosely speaking, a specialized function: it is a group of nodes that takes its inputs from one source and passes them to a single destination. The source and destination can be a group of nodes as well. Although the dimensions of a role, nodes squared per flow, may seem unnatural, this definition of role corresponds well with our intuitive notion of a role for unweighted networks with equal connectivity for each node. Note that this definition of role is different from that proposed by other ecologists [16].

This measure for roles cannot be developed here because, except in special cases, it only makes sense when applied to weighted networks. To give an intuitive feel for a role, however, we consider a few examples of unweighted networks in which the idea of a role makes sense. In Figure 1, $R = 4$. This makes intuitive sense, because each node is playing a unique role in the network, taking its inputs from one source and passing them on to one destination.

In Figure 2, $R = 1$. This also makes sense, because no compartment is doing anything unique; each node receives from every node, and gives to each node. The compartments are indistinguishable with respect to where they re-

FIGURE 3



(a) Node A has an input connectivity of 2 flows per node; (b) If unweighted, A has an input connectivity of 2. Effectively, however, A's input connectivity is close to 1.

ceive or channel their flows. One can think of the nodes as distinct agents playing the same role.

WEIGHTED NETWORKS

The extent to which unweighted networks can be used to describe the real world is limited. In real systems, flows have unequal size, with sometimes extraordinary differences. Even when the flows themselves are equal in size, the importance of each nodes is often not: with equal flows, the throughput of a node with 10 inputs and outputs is greater than that of a node with only 1 input and output. Specific measures from food web analysis have been generalized into weighted networks [17], but so far the number of nodes, flows, and roles have not. In order to extend these measures to apply to weighted networks, we must weight our measures both by the size of the flows and by the influence of each node. We must do so in a manner consistent with the countable properties of unweighted networks and our notion of a role. We will start by looking at connectivity in simple weighted networks; analogous derivations exist for our other measures.

We now look at the input connectivity of node A in Figure 3. With unweighted measures, the input connectivity of A is identical for both networks: 2. In Figure 3(a) the flows from C to A and from B to A are identical in size. In Figure 3(b) the flows are different; if these networks represented physical systems, 3(b) would behave as if the input to A from B were the sole input; the influence of C would be negligible. If these were trophic exchanges between species in an ecosystem, species A would depend on B for its livelihood. If the compartments were neurons and the weights were strengths of the synaptic connections, the firing of A would essentially be determined by B. Intuitively, we want the effective connectivity of A in Figure 3(b) to be close to 1. We can do this by weighting the connections by their magnitudes.

We now introduce some notation. Let $(a_1|w_1, a_2|w_2, \ldots, a_n|w_n)$ be a weighted mean function, where the a_i are the values being considered and the w_i are the respective weights, normalized so that $\Sigma_i\, w_i = 1$. Let T_{ij} represent the flow from node i to node j. Let $T_{i\bullet} = \Sigma_j T_{ij}$ be the total flow out of node i, and $T_{\bullet j} = \Sigma_i T_{ij}$ be the total flow into node j. Let $T_{\bullet\bullet} = \Sigma_{ij} T_{ij}$ be the total sum of all flows. $T_{\bullet\bullet}$ is known as the total system throughput (TST), and we will call $T_{i\bullet}$ and $T_{\bullet j}$, respectively, the output and input throughputs of a node. We have two formulae for the connectivity of a given node: one for inflows and one for outflows:

$$C_{Node\ i\ Outflows} = \frac{1}{W_j\left(\frac{T_{ij}}{T_{i\bullet}}\middle|\frac{T_{ij}}{T_{i\bullet}}\right)} \qquad C_{Node\ j\ Inflows} = \frac{1}{W_i\left(\frac{T_{ij}}{T_{\bullet j}}\middle|\frac{T_{ij}}{T_{\bullet j}}\right)}$$

It may appear as though any mean will suffice; with any weighted mean we will have a connectivity consistent with the case of equal flows, and continuity with respect to the flows. The most obvious choices are a weighted arithmetic or geometric mean, but there are an infinitude of options. We will use the geometric mean; below, we show that the geometric mean is the only mean consistent with our notion of a role.

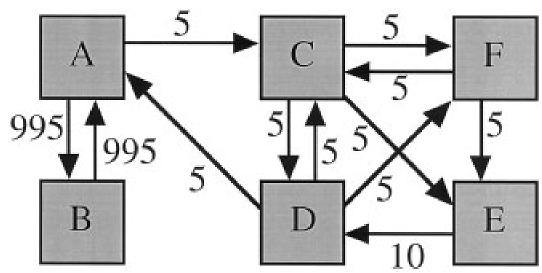
WEIGHTED GEOMETRIC MEAN

$$W(a_1|w_1,\ a_2|w_2) = a_1^{w_1}a_2^{w_2},\ W_i(a_i|w_i) = \prod_i a_i^{w_i}$$

We now must address the weighting of each node in the context of the whole system. As before, we look at a problematic network that illustrates our need for weighting.

In Figure 4, even if we had a measure that accounted for the weights of flows into and out of a given node, we could not simply average over all nodes, because some nodes

FIGURE 4



This network illustrates the need to weight the mean over all nodes. The two flows on the left constitute almost 98% of the TST and both flow between nodes of low connectivity. The other nodes have higher connectivity.

constitute a greater portion of the TST. In Figure 4, B has connectivity of 1, whereas the connectivity of C, D, E, and F are higher. If we did not weight the compartments, the connectivity of the system would be much greater than 1, due to the 4 nodes C, D, E, and F with high connectivity. But the system should be influenced mostly by the apparatus on the left, as these flows constitute almost 98% of the TST. Again we must use a weighted mean. The connectivities of each node and the unweighted vs. weighted connectivity measures of the entire network in Figure 4 are listed in Table 1.

We weight the contribution to total connectivity from each node by that node's weight in the total system. Because each node functions both as source and destination, there are two terms for each node. We weight a node's inflow connectivity as the fraction of its inflows of the TST, and similarly for outflows. Then we have two indices describing the network, one for inflows and one for outflows. To remain internally consistent, we combine these by the same mean that we chose earlier. In general, we weight the inflows and outflows equally, although if one were looking at connectivity in the context of a specific question, one could look exclusively at one term or use unequal weightings. In weighted networks, the input and output connectivity are often unequal, and when branching or condensing structures are present, the difference can be significant:

$$C = \prod_{i,j} \left(\frac{T_{ij}^2}{T_{i\cdot} T_{\cdot j}} \right)^{-(1/2)\cdot(T_{ij}/T_{\cdot\cdot})}$$

Note the $-(1/2)$ in the exponent from the reciprocal and mean.

This expression is equivalent to the effective connectivity as proposed by Ulanowicz [18]; he defines effective connectivity of a flow network as $e^{\Phi/2}$, where

$$\Phi = -\sum_{i,j} \frac{T_{ij}}{T_{\cdot\cdot}} \ln \frac{T_{ij}^2}{T_{i\cdot} T_{\cdot j}}.$$

We now return to our beginning discussion of networks and recall that we had four variables: F for flows, N for nodes, $C = F/N$ for connectivity, and $R = N/C$ for number of roles. Without deriving them here (one can follow reasoning similar to that above) we will state that we have

$$F = \prod_{i,j} \left(\frac{T_{ij}}{T_{\cdot\cdot}} \right)^{-(T_{ij}/T_{\cdot\cdot})}$$

which is the effective number of flows; note the negative in the exponent. This expression approaches the actual number of flows as the flows approach equal size and

$$N = C \cdot F = \prod_{i,j} \left(\frac{T_{\cdot\cdot}^2}{T_{i\cdot} T_{\cdot j}} \right)^{(1/2)\cdot(T_{ij}/T_{\cdot\cdot})}$$

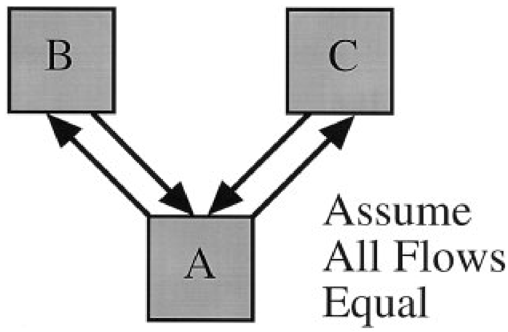
which is the effective number of nodes; this is a weighted mean of the normalized throughput of each node. Again, note the 1/2 in the exponent. We have now freed ourselves

Table 1

Measures of Figure 4

Node:	Input Conn:	Output Conn:	Weight of Node:	Unweighted Mean Connectivity:	
				Arithmetic:	1.83
A	1.03	1.03	0.48900	Geometric:	1.66
B	1.00	1.00	0.48655	Weighted Measures:	
C	3.00	3.00	0.00733		
D	1.89	1.89	0.00733	Effective # of Nodes:	2.28
E	2.00	2.00	0.00489	Effective # of Flows:	2.36
F	2.00	2.00	0.00489	Effective Connectivity:	1.04

FIGURE 5



Even though the flows are equal in influence, the nodes are not. Node A has more flows, and thus a greater relative throughput and greater importance in the context of the whole network.

from any direct reference to nodes while retaining a working definition of the number of nodes; we have abstracted the notion of a "node" to the point where it only makes sense to talk about a node with respect to a flow, for any node without attached flows makes no contribution to N.

At this point, one might object that this measure does not correspond to the countable value of nodes for certain equally weighted networks. This lack of correspondence happens in some examples because, although the flows are weighted equally, nodes have differing throughputs.

In Figure 5 there are only $2\sqrt{2} \approx 2.82$ effective nodes. This happens because node A has twice the throughput of B or C. When weighting is taken into account, the weight of A is twice that of B or C. As a node is defined only in reference to its adjacent flows, the nodes having greater throughputs thus "count" for more. Reassuringly, in the case of equal flows and equal throughputs for each node, the effective number of nodes is equal to the countable number of nodes.

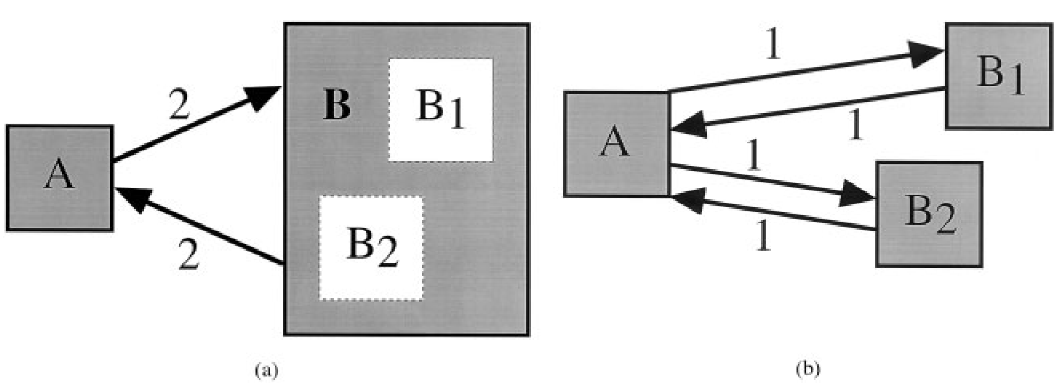
Still, the objection that our idea of effective nodes is too abstract and potentially flawed is a valid one. The importance of one node to a network can differ considerably from its relative throughput. A complete theory of flow networks must leave room for additional means of weighting the nodes. This question will be thoroughly addressed later.

The main objection remaining is our choice of mean. It may appear that our choice of the geometric mean is arbitrary. Given only the properties we have so far discussed, it is; however, there is one more property that our measures must satisfy. Our choice of mean is constrained by the properties of roles; we want the number of roles to correspond with our intuitive notions of a role. Notably, we require that the number of roles in a network should remain unchanged if we aggregate nodes that have the exact same inflows and outflows.

The number of roles should be 2 for each network in Figure 6. In Figure 6(a), A and B have unique roles; in Figure 6(b), A has a unique role, but B_1 and B_2 together fulfill one role. Thus, although the right network has roughly 2.82 effective nodes, it has two effective roles. Note that there is no issue of the two roles being weighted differently, as each role has the same throughput.

To understand this situation thoroughly, quickly note all the variables involved. $T_{\bullet\bullet} = 4$, and $T_{i\bullet}$ and $T_{\bullet j}$ are both 2 for A and 1 for each B. Now, regardless of what mean we use, all flows are equal, so $F = 4$. We can also look at the connectivity of each node; A has connectivity 2 and each B has connectivity 1. Thus we have $C = W(1|_{\frac{1}{4}}1, 1|_{\frac{1}{4}}1, 2|_{\frac{1}{2}}) = W(1|_{\frac{1}{2}}, 2|_{\frac{1}{2}})$, which is the unweighted mean of 1 and 2. If we can calculate C by some other method, we will know what mean should be used. We

FIGURE 6



(a) This two-node network can be seen as a simplification of the 3-node system in 6b, in which nodes B1 and B2 have been grouped into one node, B; (b) B1 and B2 are indistinguishable with respect to the source and destination of their flows and the distribution of flows among these sources and destinations. Thus the B's constitute one role, and the system only has two roles.

know that C, R, and F are related by the formula $R = F/C^2$. But we also know that $R = 2$.

Solving for C, we have $C = \sqrt{2}$, the geometric mean of 2 and 1. Thus, we must use the geometric mean. One can derive expressions for F, N, and C that use other means, but the introduction of roles constrains our options so that the geometric mean is our only choice. We now have a unique set of measures, and its fourth element:

$$R = N^2 \cdot F = \frac{N}{C} = \frac{F}{C^2} = \prod_{i,j} \left(\frac{T_{ij} T_{\bullet\bullet}}{T_{i\bullet} T_{\bullet j}} \right)^{(T_{ij}/T_{\bullet\bullet})}$$

which is the effective number of roles.

The effective number of roles is vital to our understanding of complexity, as it measures the degree to which the system has become differentiated into distinct functions. Note that our definition of role is process-based; it does not correspond to a well-defined group of nodes fulfilling a role; rather, it describes the number of different *functions* that are occurring in the network. We can construct networks in which the individual roles are well defined, but in general it is difficult to pinpoint a specific role in a flow network representing a natural system; while the measure is rooted in our concrete notions of a role, it has been generalized further. It is also interesting to note that the effective number of roles has an analog in information theory; taking the logarithm of R yields

$$\log R = \sum_{i,j} \frac{T_{ij}}{T_{\bullet\bullet}} \log \frac{T_{ij} T_{\bullet\bullet}}{T_{i\bullet} T_{\bullet j}} = AMI$$

This is the average mutual information (AMI) contained in the organization of flows in the network [10, 19]. This measure has been derived independently by measuring the information contained in the deviation of the observed probabilities of flows from the reference state of a network of random flows (equal, but scaled by the throughput of each node) [10]. To think of a network from this probabilistic perspective $T_{ij}/T_{\bullet\bullet}$, can be seen as an estimate of the probability that a quantum of currency would flow from i to j in a given time period. AMI has been proposed as a measure of complexity [10], but it has not met with much success in the scientific community because of its lack of a direct interpretation. Now the significance of the AMI is clear; it is the logarithm of the number of roles. Just as one exponentiates a traditional information measure to retrieve the number of choices that generated an entropy term, in flow networks one exponentiates the AMI to retrieve the number of roles. Although many have criticized information measures as arbitrary and meaningless, we now see that they are directly related to the specialization of a system.

APPLICATIONS

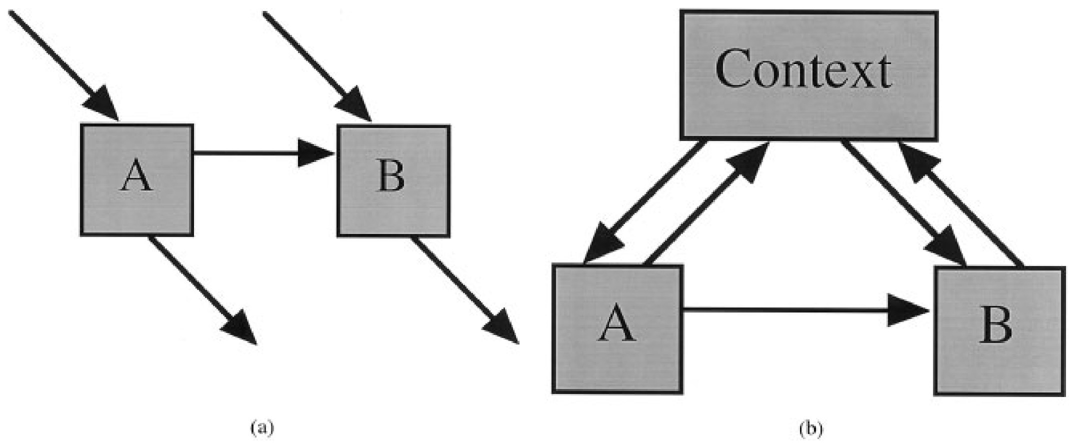
The four variables F, N, C, and R form a toolbox for analyzing complex systems from a new perspective. One obvious application is the study of stability, or adaptability, in terms of these variables. It has been argued that C is correlated with the stability of ecosystems and has an upper bound of ~3.015 in natural ecosystems [18]; extensive testing of this hypothesis as well as applications of these measures to the stability of economics or neural networks remains to be done. Other questions also arise: can these measures tell us anything about nonliving complex systems, such as convection cells or self-organizing chemical reactions? Or can these measures be applied to individual organisms or systems within organisms?

We first note that these measures can easily be applied to open systems. Our discussions so far have revolved around closed systems, but the extension to open systems is not difficult. The simplest way is to treat the context of the system as a node. Interestingly, this makes a linear flow network equivalent to a cyclic one; it is as if the system were participating in a cycle that extended beyond its boundaries. For example, the open system in Figure 7(a) could be represented as the closed system in Figure 7(b). This technique is adequate for describing systems in which inputs and outputs constitute a relatively small portion of TST; when this is not the case, adjustments must be made so that the external flows do not dominate the measures [18].

One potential weakness of the measures we have developed is that they depend on arbitrary choices of what constitutes a node. For instance, in an ecosystem, does one treat each species as a separate node? If not, which species does one group together? Viewed another way, however, these weaknesses are a potential strength, for it has been argued that measures of complexity must be dependent on the level of detail [20], and these measures certainly are. By looking at different scales, we can use these measures to evaluate the complexity present in different levels of organization of the system. A way to assess the effect of groupings is built into our measure R. When we aggregate nodes, R will decrease or stay constant. A small decrease means that the species have similar roles and little information is lost in the aggregation [21]. A large decrease is a sign that information is lost in the aggregation.

One can also categorize nodes in response to a specific question. For example, if we wanted to evaluate the response of an ecosystem to changing conditions (for example, varying rainfall, temperatures, or salinity levels), we could break up species into categories based on their susceptibility to one variable. A high value of C would then represent a greater diversity of flows per node with respect to categories of different susceptibility to that variable, and thus should correspond to the system's adaptability in response to that variable.

FIGURE 7



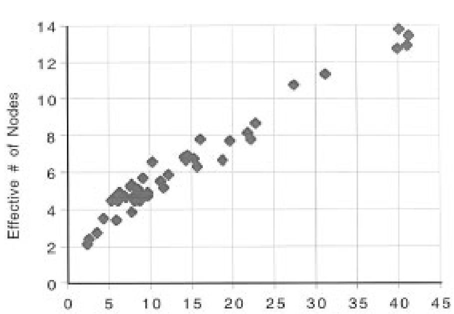
(a) An example of an open system described by a network; (b) The same system, recast as a closed system, with the system's context as a node.

To demonstrate the application of these measures to real networks, Figures 8 and 9 show data from 44 ecosystem networks of varying sizes [22]. Most networks measure carbon or energy flows, but some measure nitrogen or other elements. Figure 8 shows a strong correlation between effective nodes and effective flows. This correlation can be recast in terms of a restriction of the ratio F/N, which is C. Figure 9 demonstrates that, although the countable number of nodes and flows varies widely, natural ecosystems lie in a small window of vitality with respect to the variables C and R. Interestingly, the effective number of nodes and flows take on values less than 15 and 45, respectively, in spite of the fact that many of the networks have over 30 countable nodes and over 100 countable flows. There may be a connection between the bounds on C and others' observations of an upper bound about 3 for food web connectivity [23, 24], and similarly, a connection between the bounds on R and proposed upper bounds of about 5 for ecosystem trophic levels [25–28], although these diagrams are by no means a rigorous demonstration.

To demonstrate that this "window of vitality" is significant and not just an artifact of our mathematics, we generated and analyzed 100 random networks. These networks were constructed with a random number of nodes between 0 and 100, a random density of 0s in the connection matrix, and flows with random values between 0 and 1. Figure 10 shows the role and connectivity distribution of the random networks. Figure 11 compares the ecosystem networks to the random networks; the window of vitality that ecosystem networks are found in comprises a very small portion of the distribution of the randomly generated networks.

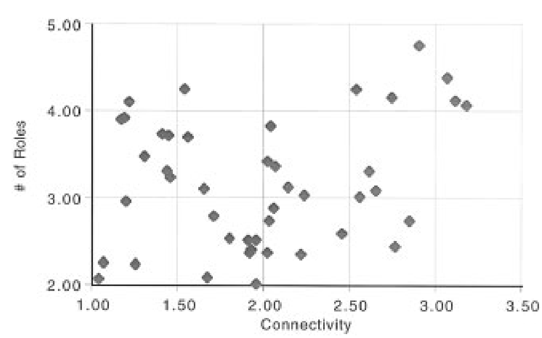
We mentioned earlier an additional problem that must be addressed: implicit in all of our measures is the assumption that the importance of a node is proportional solely to its throughput ($T_{i\bullet}T_{\bullet j}/T_{\bullet\bullet}^2$). We can discard this assumption

FIGURE 8



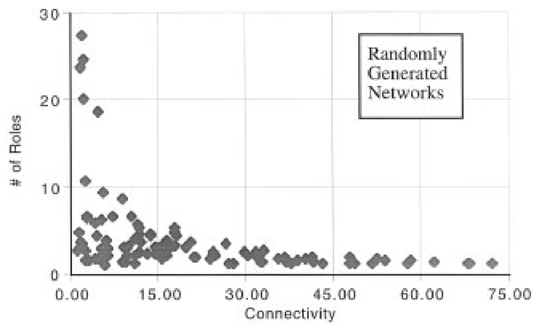
Effective # of nodes and flows for various ecosystem networks.

FIGURE 9



# of roles and connectivity for the same networks.

FIGURE 10



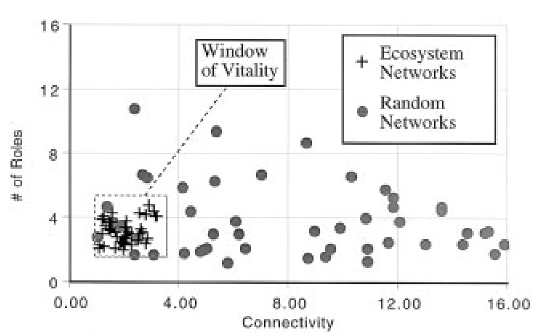
# of roles and connectivity for randomly generated networks.

and generalize our measures even further. We weight each node according to some importance; call it B_i for node i. We set these weights according to any other knowledge we have of the system. This technique has been used to account for biomass in ecosystems [29], but the weighting can be done on whatever basis one desires. The result is a new set of measures that incorporates this additional information. These generalized measures are presented in the Appendix.

CONCLUSION

The countable properties of unweighted flow networks can be successfully generalized to apply to weighted networks. The central measures are the numbers of flows and nodes, connectivity in flows per node, and the number of roles as measured by nodes divided by connectivity. Provided one makes reasonable assumptions regarding the weighting and the behavior of roles in special cases, these measures are unique, and are equivalent to information-theoretic measures previously used in ecosystem network analysis. These measures depend upon the level of observation and have the potential to measure the complexity of a wide variety of natural systems. Additional information about the system can be incorporated in the form of arbitrary weightings of the nodes. When we compare ecosystem networks to randomly generated networks, we find that ecosystems inhabit a narrow window of parameter space, with C between 1 and 3.25 and R between 2 and 5, whereas randomly generated networks often lie far outside this window. Complexity science has recently embraced the network approach, but the analysis of weighted networks has been uncommon. Now, however, we have the core of a quantitative theory of the organization of weighted flow networks—one that is consistent with information theory and common sense, and helps illuminate empirical data from ecology. By providing a glimpse of the analytical power that this theory has to offer, we hope to convince others that weighting leads in the end to a more elegant and fruitful analysis of networks in complex systems.

FIGURE 11



# of roles and connectivity for both randomly generated networks and ecosystem networks. Many of the random networks fall off the graph to the top left or bottom right. The ecosystem networks inhabit a very small range of parameter space relative to that occupied by the random networks.

ACKNOWLEDGMENTS

A. Zorach was supported by a grant from the Research Experience for Undergraduates Program of the National Science Foundation (Contract No. OCE-0139268) administered through the Maryland Sea Grant. R. Ulanowicz was supported in part by the National Science Foundation's Program on Biocomplexity (Contract No. DEB-9981328.), and the US Geological Survey Program for Across Trophic Levels Systems Simulation (ATLSS, Contract 1445CA09950093).

REFERENCES

1. Capra, F. The Web of Life; Doubleday: New York, 1996.
2. Baird, D.; Ulanowicz, R.E. The seasonal dynamics of the Chesapeake Bay ecosystem. Ecol Monogr 1989, 59, 329–364.
3. Wulff, F.; Field, J.G.; Mann, K.H., editors. Network Analysis in Marine Ecology. Coastal and Estuarine Studies Series; Springer-Verlag, Berlin, 1989.
4. Leontief, W. The Structure of the American Economy, 1919–1939. 2nd ed. Oxford University Press, New York, 1951.
5. Peusner, L. Studies in Network Thermodynamics; Elsevier, Amsterdam, 1986.
6. Ford, L.R., Jr; Fulkerson, D.R. Maximal flow through a network. Can J Math 1956, 8, 399–404.
7. Belady, C.; Kelkar, K.M.; Patankar, S.V. Improving productivity in electronic packaging with flow network modeling (FNM). Electronics Cooling, 1999. Online: http://www.electronics-cooling.com/html/1999_jan_article4.html.
8. Strogatz, S.H. Exploring complex networks. Nature 2001, 410, 268–276.

9. Amit, D.J. Modeling Brain Function: The World of Attractor Neural Networks; Cambridge University Press: New York, 1989.
10. Ulanowicz, R.E. Growth and Development: Ecosystem Phenomenology; Springer-Verlag: New York, 1986.
11. May, R.M. Will a large complex system be stable? Nature 1972, 238, 413–414.
12. Kauffman, S.A. Antichaos and adaptation. Sci Am 1991, 265, 78–84.
13. Kauffman, S.A. The Origins of Order; Oxford University Press: New York,1993.
14. Rejmanek, M.; Stary, P. Connectance in real biotic communities and critical values for stability of model ecosystems. Nature 1979, 280, 311–313.
15. Haydon, D. Pivotal assumptions determining the relationship between stability and complexity: An analytical synthesis of the stability-complexity debate. Am Nat 1994, 144, 14–29.
16. Johnson, J.C.; Borgatti, S.P.; Luczkovich, J.J.; Everett, M.G. Network role analysis in the study of food webs: An application of regular role coloration. J Soc Struct 2001, 2(3), Online: http://zeeb.library.cmu.edu:7850/JoSS/johnson/RoleAnalysis.html.
17. Bersier, L. Quantitative descriptors of food web matrices. Ecology 2002, 83(9), 2394–2407.
18. Ulanowicz, R.E. The balance between adaptability and adaptation. Bio Systems 2002, 64, 13–22.
19. Rutledge, R.W.; Basorre, B.L.; Mulholland, R.J. Ecological stability: an information theory viewpoint. J Theor Biol 1976, 57, 355–371.
20. Gell-Mann, M. What is complexity? Complexity 1995, 1(1), 16–19.
21. Hirata H.; Ulanowicz, R.E. Information theoretical analysis of the aggregation and hierarchical structure of ecological networks. J Theor Biol 1985, 116, 321–341.
22. Sources of these data can be found at http://cbl.umces.edu/~ulan/ntwk/netrefs.html.
23. Pimm, S.L. Foodwebs; Chapman and Hall: London, 1982; pp 219.
24. Wagensberg, J.; Garcia, A.; Sole, R.V. Connectivity and information transfer in flow networks: two magic numbers in ecology? Bull Math Biol 1990, 52, 733–740.
25. Hutchinson, G.E. Homage to Santa Rosalia, or why are there so many kinds of animals? Am Nat 1959, 93, 145–159.
26. Odum, H.T.; Pinkerton R.C. Time's speed regulator: the optimum efficiency for maximum power output in physical and biological systems. Am Sci 1955, 43, 331–343.
27. Slobodkin, L.B. Growth and Regulation of Animal Populations; Holt, Rinehart and Winston: New York, 1961.
28. Pimm, S.L.; Lawton, J.H. Number of trophic levels in ecological communities. Nature 1977, 268, 329–331.
29. Ulanowicz, R.E.; Abarca-Arenas, L.G. An informational synthesis of ecosystem structure and function. Ecol Model 1997, 95, 1–10.
30. Ulanowicz, R.E.; Norden, J.S. Symmetrical overhead in flow networks. Int J Systems Sci 1990, 21(2), 429–437.

APPENDIX

Formulae

F is the number of flows and N is the number of nodes, and connectivity $C = F/N$ is measured in flows/node.

Effective flows:

$$F = \prod_{i,j} \left(\frac{T_{ij}}{T_{\bullet\bullet}}\right)^{-(T_{ij}/T_{\bullet\bullet})}.$$

Effective number of nodes:

$$N = \prod_{i,j} \left(\frac{T_{\bullet\bullet}^2}{T_{i\bullet}T_{\bullet j}}\right)^{(1/2)\cdot(T_{ij}/T_{\bullet\bullet})}.$$

Effective connectivity:

$$C_{Total} = \prod_{i,j} \left(\frac{T_{ij}^2}{T_{i\bullet}T_{\bullet j}}\right)^{-(1/2)\cdot(T_{ij}/T_{\bullet\bullet})}.$$

Number of roles:

$$R = \prod_{i,j} \left(\frac{T_{ij}T_{\bullet\bullet}}{T_{i\bullet}T_{\bullet j}}\right)^{(T_{ij}/T_{\bullet\bullet})}.$$

ln R is the average mutual information (AMI) of the network.

ln F is the Shannon entropy of the normalized flows.

ln $C = \Phi/2$, where Φ is the overhead, as defined by Ulanowicz and Norden [30].

Generalized

Let B_i be the importance of node i. Let $B_\bullet = \sum_i B_i$. To derive these, we replace $T_{\bullet\bullet}^2/T_{i\bullet}T_{\bullet j}$ (previously used as the importance of each node) with $B_\bullet^2/B_iB_j$. Note that $B_iB_j/B_\bullet^2$ represents the flow from i to j in a "reference" network; this network yields the same values for these variables that an unweighted network would yield using the conventional measures.

Effective number of flows:

$$F = \prod_{i,j} \left(\frac{T_{ij}}{T_{\bullet\bullet}}\right)^{-(T_{ij}/T_{\bullet\bullet})}.$$

Effective number of nodes:

$$N = \prod_{i,j} \left(\frac{B_\bullet^2}{B_iB_j}\right)^{(1/2)\cdot(T_{ij}/T_{\bullet\bullet})}.$$

Effective connectivity:

$$C_{Total} = \prod_{i,j} \left(\frac{T_{\bullet\bullet}^2 B_\bullet^2}{T_{ij}^2 B_iB_j}\right)^{-(1/2)\cdot(T_{ij}/T_{\bullet\bullet})}.$$

Effective number of roles:

$$R = \prod_{i,j} \left(\frac{T_{ij}B_\bullet^2}{T_{\bullet\bullet}B_iB_j}\right)^{(T_{ij}/T_{\bullet\bullet})}.$$