BNFO 591
INTRODUCTION TO HIGH PERFORMANCE COMPUTING IN BIOINFORMATICS AND THE LIFE SCIENCES

FORTRAN HOMEWORK 3: COMPUTING NETWORK DIAMETER OF AN ARBITRARY NETWORK

TARYNN M. WITTEN

Abstract. Network structures and their analysis are now quite common in bioinformatics. Networks have many properties. One of these properties is called *network diameter*. In this homework assignment you will construct a FORTRAN program to compute the network diameter of a network from node-node connectivity data.

1. Introduction

Reductionism, while it has its uses, also causes us to lose important information about the behavior of a system because breaking a system apart costs us information about how the "whole" organism functions. Living systems are complex. If we were to examine a large collection of different complex systems, we would find that complex systems have certain common or unifying characteristics:

- They demonstrate emergent behavior; behavior that cannot be inferred from a linear analysis of the behavior of the components.
- They contain many components that are dynamically interacting (feedback, controllers, detectors, effectors and rules). There is no master controller. The parts interact extensively at their local level with nearest neighbors.
- The components are diverse thereby leading to a significant diversity of information in the system.
- The components have surrendered some of their uniqueness or identity to serve as elements of the complex system. This is called dissolvence.
- All interactions of the components within the system and the system acting as a component in a higher hierarchy occur locally. There is no action at a distance.
- These interactions take place across a number of scale levels and they are arranged in a hierarchical structure where fine structure (scale) influences large-scale behavior.
- They are able to self-organize, to adapt and to evolve.

Complex systems have "emergent properties" meaning that a behavior that was not predicted from infinite knowledge of the parts emerges as part of the systems

Date: September 17, 2015.

behaviors. Living systems, whether they are cells or ecosystems, do not function like pieces of a jigsaw puzzle. Instead, they are often fuzzy or stochastic, with backup systems and redundancies that belie their true structure. And we pointed out that an understanding of these systems requires a different conceptual framework. Thus, in order to understand complex systems, we must understand them through a reverse engineering perspective rather than a reductionist perspective. One approach to gaining this understanding is through the use of network representations of living systems.

2. Networks & Graphs

Biological networks can be represented using the concept of a graph G that has nodes n_j (or vertices) and edges E_{ij} (a connection between node n_j and node n_i. From a biological perspective, we can consider the nodes to be genes or proteins and the edges as paths between them. We represent the overall network connectivities in a matrix format called the *adjacency matrix* which we denote with the symbol A. The elements of A are denoted a_{ij} and are simple; if node n_i is connected to node n_j, we enter the number one in the $(i,j)^{th}$ element of the matrix A otherwise we enter a zero. Observe that if n_i is connected to n_j then it follows that n_j is connected to n_i so that the matrix A is a symmetric matrix. In the case where there are multiple edges connecting the same nodes, we enter the number of edges. Thus, if two different edges connect n_i to node n_j, we would enter the number two. Nodes that are not connected to anything in the graph G are called *islands*. The edges can have *weights*, denoted w_{ij}, assigned to them where, for example, the weight value may correspond to a rate of reaction. An edge E_{ij} can also have a direction assigned to it. For example, if E_{12} represents the edge between nodes n_1 and n_2, we might denote the fact that n_1 is upstream of n_2 by $E_{1\to 2}$. An edge that does not have any direction assigned to it is said to be *undirected* whereas edges that have a direction assigned to them are called *directed* edges. Note that we can have other types of edges in a network. For example, *multi-edges* are multiple edges between nodes and *self-edges* occur when a node that is connected to itself. We illustrate an undirected longevity gene-protein network in Figure[1][84, 85]. With this simple set of definitions, we have some powerful tools with which to investigate the structure of a network and how it might inform us about the biological dynamics of the overall network. We begin with the concept of connectivity.

It is natural to conclude that the more edges going in and out of a node, the more likely that the given node is going to be of importance to the network. *Hubs* or nodes with large numbers of connections are known to play central roles in keeping complex networks connected. This is important when we consider, in an upcoming section, the concepts of *robustness*, *resilience* and *frailty* of a network. The number of connections going in and out of a node n_i is called the *connectivity* or *degree* k_i of the i^{th} node. Sometimes you will see the degree of a node expressed using $d(n_i)$. In mathematical terms $k_i = \sum_{j=1}^{N} a_{ij}$ where N is the number of nodes in the network and a_{ij} is the $(i,j)^{th}$ element of the adjacency matrix A. Computing the connectivity of a large set of nodes leaves us with nothing more than a frequency table and it is hard to interpret this string of numbers k_i, particularly if the number of nodes in the network is large. In order to assist us in understanding the connectivity structure of the network, we create a connectivity plot (Figure[2]). To do this we first count the number of nodes with a given connectivity k where

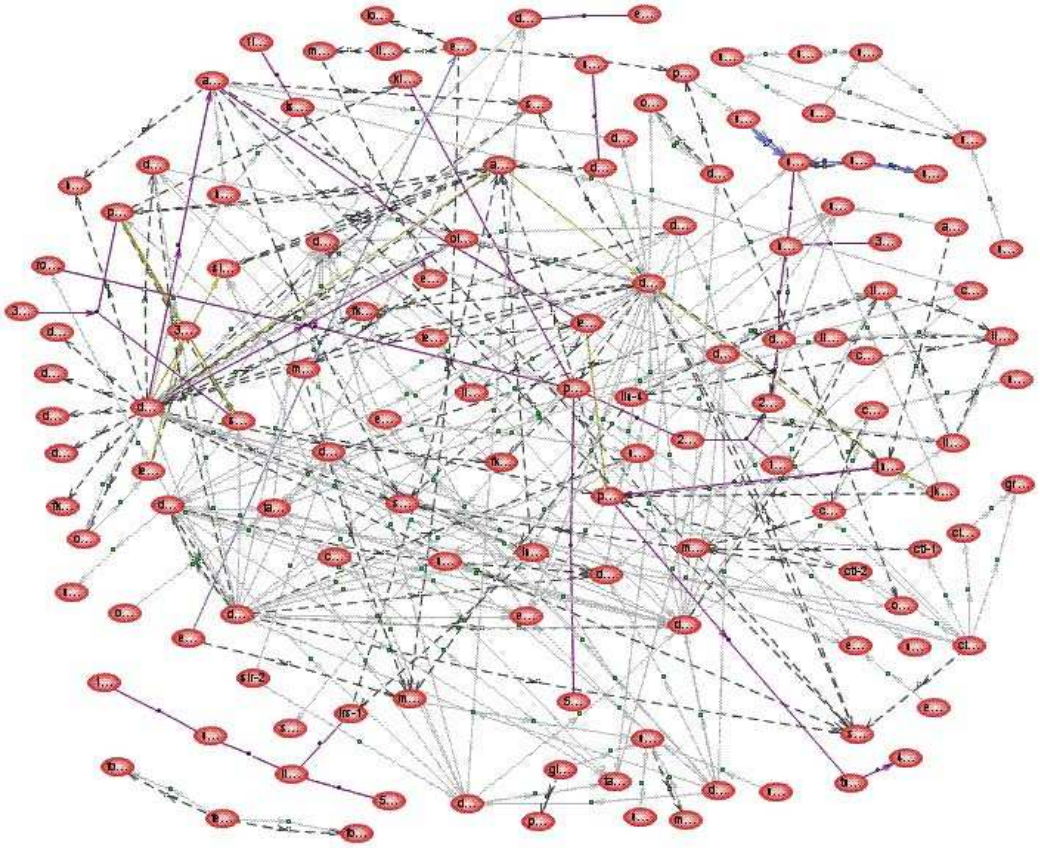

FIGURE 1. Illustration of a sample *C. elegans* longevity gene-protein network. See [85] for more details.

the connectivity varies from zero to the maximum connectivity value. The number of nodes with a given connectivity k_i is called the frequency of that connectivity and is denoted $f(k)$. Next, plot the frequency $f(k)$ versus the connectivity k. We illustrate this in Figure[2].

Studies of the statistical behavior of various network structures[6, 55] have shown that networks can have a small variety of overall topologies; random, regular, small-world and scale-free. Moreover, each of these network topologies has a classic pattern form for its degree distribution plot. *Random networks* are just what you would imagine them to be; nodes are randomly connected to each other. *Regular networks* can be thought of as lattices where there is a repetitive pattern of connections such as a grid. Small-world and scale-free networks are of greater interest because they have some fascinating underlying properties. Moreover, many real-world networks can be shown to be small world or scale free[2, 9, 34, 76, 77, 79, 80, 86]. A *small-world network* can be described as a network in which most nodes are not neighbors of one another, but most nodes can be reached from every other node by a small

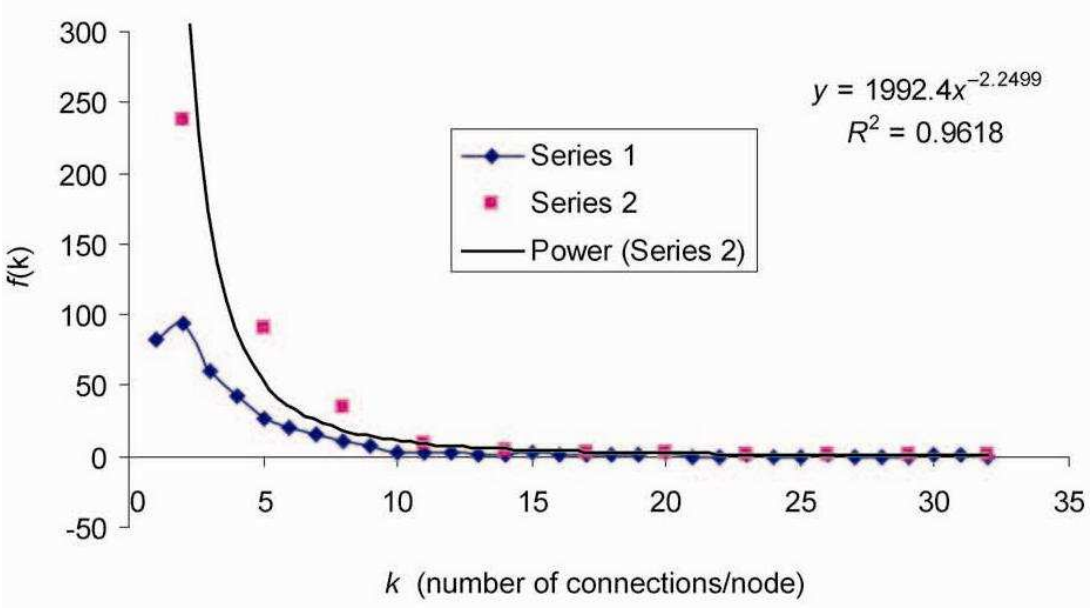

FIGURE 2. Illustration of a sample connectivity or *degree distribution* plot for the network in Figure[1] See [85] for more details. The rhombs represent the complete distribution. The squares are the data points binned into groups of three. The black solid line is the nonlinear regression line. Results are significant at $P < 0.05$

number of hops or steps[19, 30]. A *scale-free network* may appear to be a random network however, in a scale-free network the links between the nodes are preferentially attached to the most highly connected nodes thereby creating a greater frequency of links connected to a smaller number of nodes[3]. Because scale-free networks are ubiquitous and highly relevant to our discussion, let us look at them a bit more closely. In examining Figure[2] we see that not all nodes in the network have the same number of edges. If we divide the y-axis in Figure[2] by the total number of nodes in the network, call that N, then $\frac{f(k_i)}{N}$ represents the probability $P(k_i)$ that a randomly selected node has exactly k_i edges. In a randomly connected graph, the edges are placed at random and one can show that the majority of the nodes will have approximately the same connectivity which is close to the *average connectivity* $< k >$. In fact, it has been shown that the connectivities k in a random network follow a Poisson distribution with a peak at $< k >$.

What became interesting is that, for larger networks like gene, protein and metabolic networks, is that these networks did not follow the traditional Poisson probability distribution. Rather, they followed a probability distribution where the connectivity probability $P(k)$ was a power law of the form

$$P(k) = Bk^{-\gamma} \tag{1}$$

Observe that since $P(k)$ is a probability, when we sum over all of the values of k, the result had better add up to one. Thus, the parameter value of B is chosen so that this is true. We will not get into all of the varied aspects of scale-free networks[11, 21, 35, 71, 75]. However, how can we determine if we have a scale-free distribution?

It is hard to interpret a plot like that illustrated in Figure[2]. However, we observe that if we take the log of both sides of equation(1) the better the fit, the more linear the plot should be. Thus, networks whose connectivity structure follows a power law of the form $f(k) = Bk^{-\gamma}$ where B and γ are parameters to be estimated should look like negative slope lines if they are scale-free. The simplest way to estimate the parameters is to perform a linear regression on the log-log transformed $f(k)$ vs. k data, dropping the $k = 0$ data point because there are no connectivities. The more linear this curve, the more the connectivity behaves like a power law. We found that $B = 1992.4$ and $\gamma = 2.2499$ with an $r^2 = 0.969$. Thus, our *C. elegans* longevity gene-protein network[85] can be said to be a scale-free network.

Scale-free networks are also special in that they are built in a unique way. To build a scale-free network you start off with a set of N nodes in which each node in the network is connected to all of the other nodes. Next, to add a new node you make k connections to existing nodes in the network. However, whether a new node m is connected to an already existent network node n_i is determined by the degree of the given node n_i; the greater the degree of n_i the more likely m is going to be connected to n_i. In other words, the probability that node m will be connected to node n_i is given by

$$P(n_i) = \frac{d(n_i)}{\sum_j d(n_i)} \tag{2}$$

Notice that this connectivity algorithm means that if you are already very tightly connected in the network, then you are more likely to get even more connected in the network. Many scale-free networks have an exponent $\gamma \approx 3$. However, the exponent value very much depends upon the rule used for the probability of new node connection. Equation[2] is a very simple example. Given the large number of biological networks, particularly at the cellular level[43], that have been shown to be small-world formations, this suggests that tendency to create small-small world networks is a natural evolutionary pathway.

2.0.1. *Categorizing Small-World Networks.* Because of the unique nature of scale-free networks, a log-log connectivity plot is enough to let you know if you are dealing with a scale-free network or not. However, this does not work for other network forms. Because many biological systems also demonstrate small-world network behavior[9, 37, 38, 69, 76, 77], we briefly examine how to determine whether or not a network is a small-world network. Remember, a working definition of a small-world network is a network in which most of the neighbors of a node are neighbors themselves (thing regular network here, lattice structure for example). However, in addition to this property, the average number of connections between two chosen random nodes in the network n_i and n_j is small (similar to the properties of a randomly connected network).

In order help characterize small-world networks we introduce a few new network descriptors. The first is the *average path length* of a network. Path length is the

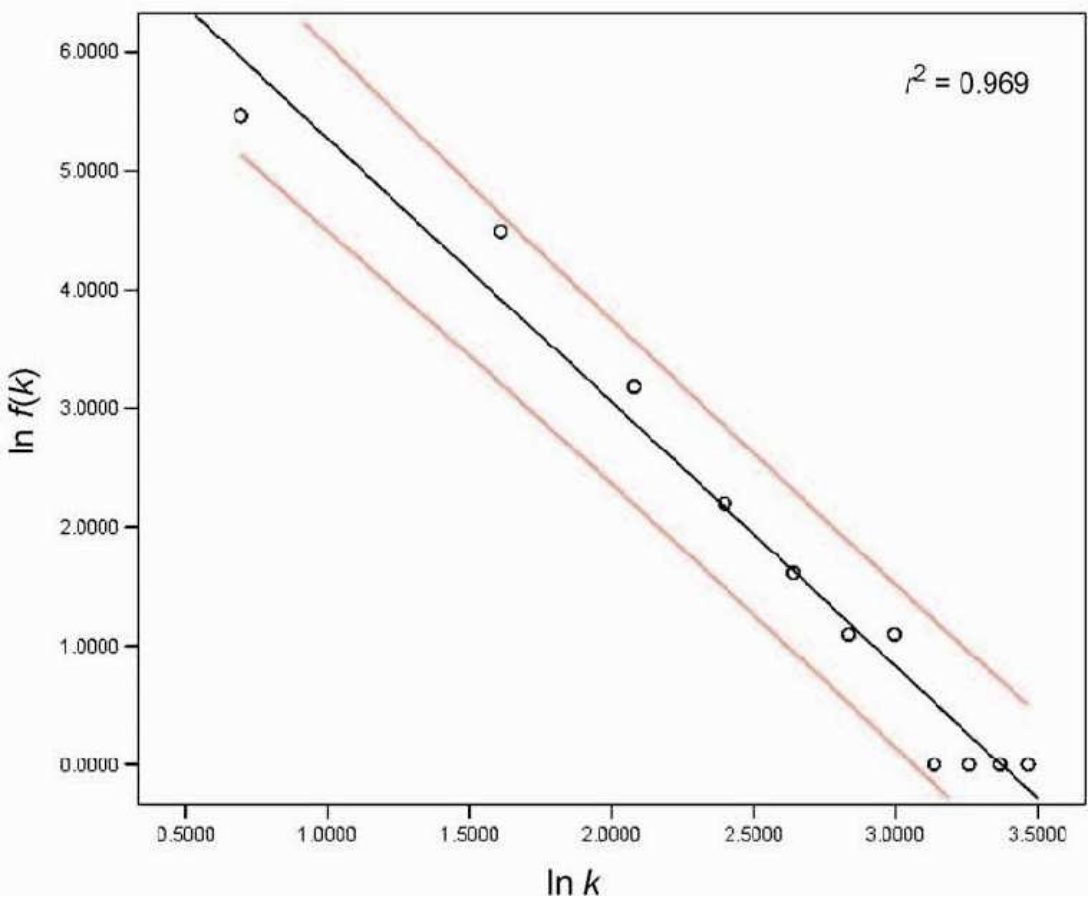

FIGURE 3. Illustration of the log-log data and regression curve through the data of Figure[2]. The outer lines are the 95% confidence interval boundaries for the linear regression estimate. See [85] for more details.

distance or number of edges between two nodes in the network. We use the idea of path length to construct the *minimum path length* between node n_i and node n_j and denoted it by ℓ_{ij}. We now introduce the concept of the *diameter* of a network. To do this we consider the example network in Figure[4].

Now, consider the adjacency matrix A derived from the network in Figure[4] and which is illustrated on the left hand side of Table[**??**]. The matrix A represents the number of length 1 paths between node n_i and node n_j. If we multiply $A \times A$ the entries that are non-zero represent the number of length 2 paths between each pair of nodes. If we repeat this process exactly $N-1$, then we get the total number of paths of length $1, 2, \ldots, N-1$ for the network between each pair of nodes in the network where N is the total number of nodes in the network. The right hand side of Table[**??**] illustrates $A \times A \times A \times A \times A$ which is the number of paths of length five between each pair of nodes. So, for example, there are 29 possible paths of length

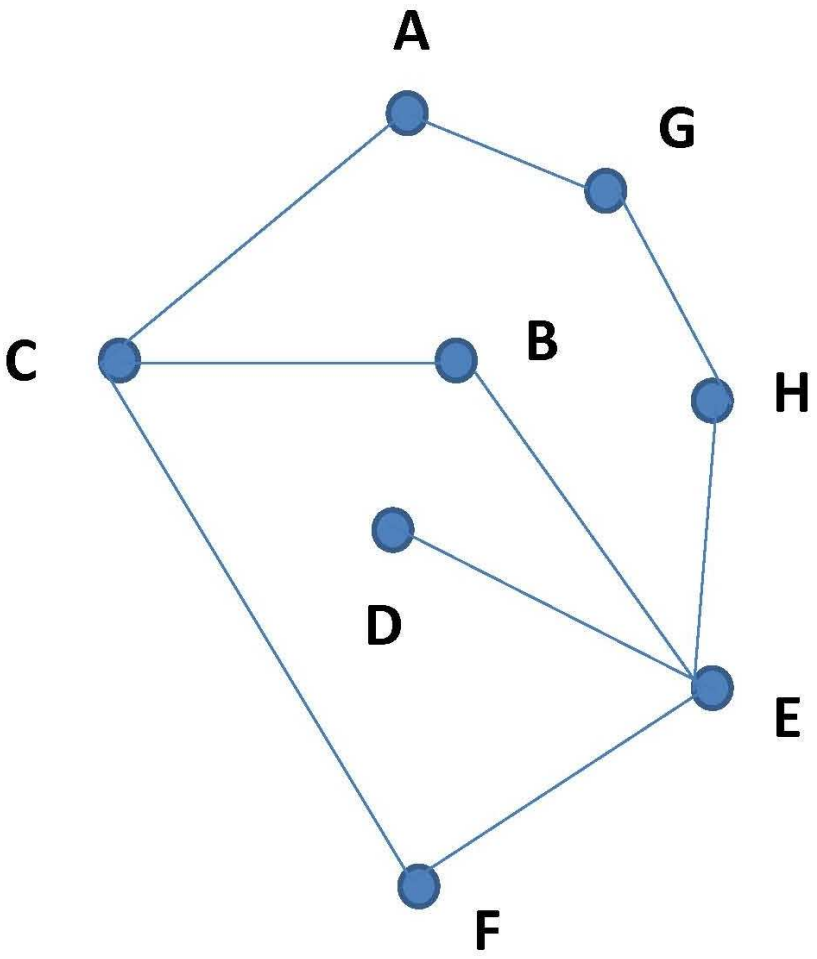

FIGURE 4. Illustration of a simple hypothetical network

5 between node C and node B. At some point between 1 and $N-1$ multiplies, every element in the matrix or its subsequent multiplies $A, A^2, \ldots, A^{N-1}, A^N$ will have been non-zero during the multiply sequence. The diameter is the minimum number of times the adjacency matrix A has to be multiplied by itself so that each entry has taken a value greater than 0 at least once during the multiply sequence.

It is obvious that, if you are statistically minded, diameter and path length measure spread of a network. They function kind of like variance and standard deviation. There are two other important network characteristics *assortativity* and *clustering*. These two properties function like the clustering variables of statistics, mean and median. They tell us how the network bunches together and in what ways. Definition of the clustering coefficient is not unique. A verbal description would say that the *clustering coefficient* C is the probability that any two randomly chosen neighbors of a given node in a network are also neighbors themselves. C is

	A	B	C	D	E	F	G	H		A	B	C	D	E	F	G	H
A	0	0	1	0	0	0	1	0		0	0	21	0	23	0	13	0
B	0	0	1	0	1	0	0	0		0	0	29	0	36	0	15	0
C	1	1	0	0	0	1	0	0		21	29	0	15	0	29	0	22
D	0	0	0	0	1	0	0	0		0	0	15	0	21	0	8	0
E	0	1	0	1	0	1	0	1		23	36	0	21	0	36	0	29
F	0	0	1	0	1	0	0	0		0	0	29	0	36	0	15	0
G	1	0	0	0	0	0	0	1		13	15	0	8	0	15	0	14
H	0	0	0	0	1	0	1	0		0	0	22	0	29	0	14	0

TABLE 1. In this table we illustrate a sample adjacency matrix A (left block of numbers). On the right, we have multiplied A times itself five times. The non-zero number in the entry represents the number of possible paths between node n_i and node n_j.

also called the *aggregation coefficient*. With a little thought, you can see that this construct is related to how many triangles there are in a given network diagram.

Clustering coefficients can be interpreted as measures of how close the local neighborhood of a node in as network is to being part of a *clique* (a sub-region of the larger network in which every node is connected to every other node). It has been observed that biological networks display high average clustering coefficients. This indicates a high level of network redundancy and cohesiveness. Further, in [**?**], it was demonstrated that clustering coefficients of many protein interaction networks have clustering coefficients that behave in the form

$$C(k) = \frac{B}{k^{\alpha}} \tag{3}$$

where B is a constant and α has values between 1 and 2. Observe that this implies that nodes with low connectivity k have clustering coefficients $C(k)$ that are large. This implies that they belong to very cohesive subnetworks. As k gets larger, $C(k)$ gets smaller so that higher degree nodes tend to have neighbors that are less connected to each other. Studies of many networks, including biological networks, have shown that network nodes tend to form groups or clusters that are characterized by high connectivities. Moreover, this type of formation tends to be greater than the average probability of a tie between nodes if you were to examine a randomly formed network. Because there are many different clustering definitions, we do not provide one specific one here but encourage the reader to visit[81] for a more complete discussion of the different types of clustering coefficient.

Assortativity is the property that a given node in the network will show a preference to attach itself to other nodes that are, in some way, similar to the given node. Again, there are a variety of different ways to define assortativity. One of the most common is to look at whether or not a node n_i with node degree d_i will attach itself to other nodes with similar degree values. Because these calculations, particularly for even medium-sized networks, can be very complex we do not provide formulae. Most software packages will calculate them for you. See Section[**??**] for software details.

2.0.2. *Node-Node Connectivities.* Another common network term is the *centrality* of a node. Centrality is a measure of the "position" or relative importance of a node in a network. In the literature there are four main measures of centrality of a node: *degree centrality*, *betweenness centrality*, *closeness centrality* and *eigenvector centrality*. From an aging-related perspective, understanding node centrality of the nodes in a network could lead to potential targets for pharmaceuticals that might help hinder disease progression or extend lifespan. The simplest of the centrality measures is degree centrality. *Degree centrality* of a node is defined by $C_D(n_i) = d(n_i)$. In other words, the degree centrality of a node n_i is simply the number of edges that are connected to the given node n_i. Obviously, this measures the change that a given node n_i in the network will receive something flowing along the network. In the case where the graph is directed, or we know the flow along the edges (upstream, downstream), we can define two new concepts C_{inD} and C_{outD} as the number of edges going in and out of n_i. These are called *indegree* and *outdegree* respectively. *Closeness centrality*, denoted $C_C(n_i)$ is the idea that the more central that a node is in a network, the lower its total distance is to all of the other nodes in the network. In other words, if a node n_i is very close to all of the nodes, then it should take a small number of edges to get to every other node in the network. From a biological perspective, closeness can be thought of as a measure of how long it will take to send a chemical or other biological signal out from n_i to all of the other nodes in the network. *Betweenness centrality*, denoted $C_B(n_i)$ looks at how often, in a network, a given node n_i acts as a bridge along the shortest path between two other nodes. From a biological perspective, knocking out a node with high betweenness centrality would force a signal to reroute itself along a path that was not the shortest path. Lastly, *Eigenvector centrality*, denoted $C_E(n_i)$ is a measure of the "influence" of a node in a network. Here, the idea is that not all connections between nodes are equal. That is, if a node is influential and it is connected to another node, it is likely that it will have more influence on that node than a node that is not that influential. These concepts are illustrated in Table [1] of [85] for the *C. elegans* longevity gene network illustrated in Figure [1] of this paper. An excellent discussion of the various concepts of centrality may be found in[10],[25]-[29],[56]-[58].

2.0.3. *Clusters and Hierarchies.* Earlier on we mentioned the concepts of clustering. Many biological networks, metabolic networks and protein interaction networks, demonstrate both clustering and scale-free properties[19, 41, 42, 50, 76, 77]. When examining network structures of this class of networks, we find that they are often modular and *hierarchical* in nature. That is, networks that exhibit the combination of small-worldness and clustering appear to be built out of modules that are themselves networks. One measure of the intrinsic hierarchical nature of a network is to make use of the mathematical result that deterministic scale-free networks that are hierarchical tend to have a clustering coefficient that goes as $C(k) \approx k^{-1}$. That is, if a node n_i as k connections, then it's clustering coefficient is approximately $\frac{1}{k}$. Thus, the higher a node's degree, the smaller its clustering coefficient. Moreover, the larger k gets, the more likely the clustering coefficient of the given node behaves as $\frac{1}{k}$. Studies of many biological systems have, indeed, shown that the networks demonstrate modularity[19, 23, 22, 24, 38, 53, 54, 57, 60, 61].

3. Closing Thoughts

In the previous sections we have introduced a large number of concepts and constructs that are based upon the premise that biological systems can be represented as network graphs. These concepts described how network nodes were interconnected and the consequences of certain specific classes of connectivity and network structure. At the 1982 Palo Alto American Mathematical Society, I presented a paper on representing aging using the model of network decay. Of course, in those days, network analysis was not what it is today and we had next to node of the genomic and network level data that we now have. However, even then, it was natural to consider aging as the temporal decay of an hypothetical organismal "aging network"[1, 14, 33, 32, 47, 65, 84, 87, 88]. How then may we apply these ideas to the study of aging?

While little is currently known about how aging-related networks evolve across the organism's lifespan, it is reasonable to assume that two possible changes can occur; inactivation of active nodes/activation of inactive nodes and loss of connectivity/increase in connectivity. Albert[4] addresses some of the theory behind evolving complex networks. How or why nodes become inactive or edges disappear is irrelevant here; just that they do[1, 14, 17, 87, 88]. It turns out that the structure of small-world networks, due to their hub connectivity, makes them vulnerable to targeted attacks aimed at specific hubs[40]. Attacks that knock out essential genes are knocking out the lifespan network because the organism dies when an essential gene is knocked out. Thus, essential genes are critical hub genes[82, 83, 85]. Small-world neural networks have been shown to exhibit short-term memory capability[16, 68]. This suggests that memory decay, such as that seen in Alzheimer's disease may be related to decay of brain neural network structure in such a way as to remove the small-worldness property of the memory network[9, 15, 23, 30, 37, 38, 39, 69]. Understanding patterns in network decomposition could lead to potential early AD detection and to potential pharmaceutical intervention at earlier points in the disease course.

Connectivity gain and loss also have implications when it comes to discussing the hierarchical modularity of aging-related network architectures. Loss of connectivity through inactivity of a node or through loss of an edge could unlink an entire module of importance. Thus, nodes that connect modules within a larger network are critical to the functioning of the network[36, 52, 62, 64]. Questions around the role of evolutionary processes in the development of network architectures of various organisms may be of importance in understanding how network architectures related to aging processes are constructed. Why are some components of a network redundant while others are not[5](see also all of the citations on reliability theory)? What is the role of backup subnetworks? What is the importance of robustness[8, 13, 31, 46, 59] and resilience[51]? Why are some networks more robust to attack, less fragile than others[66] or more frail[73]? How do we balance the need to adapt and evolve with robustness[18, 44, 48, 49]? What, if any, is the association of lifespan with network architecture? These and many other questions remain to be answered.

4. Homework Assignment FORTRAN 3

In this homework assignment you will construct a FORTRAN program to compute the network diameter of a network from node-node connectivity data. You will do it in the following steps. The goal of this assignment is **NOT** to write the most efficient program. In fact, just solve it, don't go for perfect. Better yet, you can be as sloppy as you want. Just get an answer. Later in the class you will revisit this program to improve it. So, just get a solution. Remember, if you obtain supporting information from other sources, you must include it as part of a bibliography to your handed in homework. For information on how to format your references, see the bibliography at the end of this homework handout.

First, however, you will need the following details about your data.

(1) You will need to create two datasets one called `testdata.dat` and the other called `experimental_data.dat`.
   (a) The file `testdata.dat` should contain the node-node connectivities for the network illustrated in Figure[4] of this handout. You will have to construct this file yourself. To do this, use the matrix on the left hand side of Table[1].
   (b) Your program should be able to read in the file `testdata.dat` data.
   (c) Your program should be able to detect the end of the file so that you do not have to input how many node-node pairs there are in the file. You may assume that there will not be more that $1,000$ for the purposes of this homework assignment. Bonus points for making your program not care and figure out if it reaches the end of file without the assumption of a maximum number of linkage pairs.
   (d) Your program should output all of the required output to an output file called `testoutput_yournamehere.dat`.
   (e) Your program should count the number of data pairs in `testdata.dat` and output that number to the output datafile.
   (f) Your program should count the number of unique nodes in the network and output that number to the output file.
   (g) Your program should store the unique node names in an array that can be accessed if needed.
   (h) Your program should then construct the matrix A described in the text discussion above Table[1] left hand side.
   (i) You should then, using the discussion in the text, compute the diameter of the given network (Figure[4]). For those of you who have never multiplied matrices together, get a book on basic matrix algebra or linear algebra and look it up.
   (j) For the purposes of `testdata.dat`, you should be able to reproduce the result given on the right hand side of Table[1].
   (k) To demonstrate that you have actually done this, you will need to output both the original matrix and the final matrix to the output file. Your matrices should look nice and square so you will need some sort of formatted output.
   (l) When you have completed the `testdata` analysis and verified that your program works correctly for this dataset, you must then read in the `Network of Matabolic Pathways` dataset that you have downloaded

from the Blackboard Homework 3 folder. However, you will have to pre-process it to make it readable by your program.

a: Write an AWK program to remove all of the commas after the last column

b: Your AWK program should output a space delimited file called `experimental_data.dat`.

c: This is your experimental dataset that you will use for the second part of the homework assignment.

(m) Repeat the previous computations using `experimental_data.dat`. You do not have to print out the matrices for the calculations. Just the final result. Which, if you did it correctly should be `Experimental Network Diameter = 3`.

(i) Your output file should contain your name, the date, the homework assignment name and the requisite information given in each part.

(ii) Each above item should be on one line and labeled appropriately. Don't just write a number. You should say something like `Network Diameter = 87`. I hope you didn't get 87 because if you did, you are way off and did something really incorrectly.

(iii) You should save your FORTRAN program file as `yourname_HW3.f90`. You will be needing it later-on in the course.

What Should You Hand In: You will be expected to hand in the following items

(1) A printout of your actual fortran program and a printout of the output of the program. You can combine them into one document if you want. I am old school.
(2) The printouts should be on actual paper and if it takes more than one page, you had better staple them. No paper clips!
(3) A bibliography of any resources you used to help you do the problem. There is no shame in using resources as long as you didn't just copy a program that was already on the net or in a book. **SHAME ON YOU** if you did! You won't learn if you don't attempt to do the problems.

Bonus Points For:

(1) If you repeat the above assignment and hand in your results in another programming language.
(2) If you use gnuPlot to draw and label the test network correctly.
(3) If you go whole hog and use gnuPlot to draw the experimental network correctly.

References

[1] Agoston, V., Csermely, P. & Pongor, S. (2005). Multiple, weak hits confuse complex systems. Phys. Rev. E., 71:051909
[2] Albert, R. (2005). Scale-free networks in cell biology. J. Cell Sci., 118: 4947-4957.
[3] Albert, R. (2006). General network theory. **LACUS Forum 32: Networks**. (eds). S.J. Hwang, W.J. Sullivan & A.R. Lommel. Houston, TX: Forum 32.
[4] Albert, R. & Barabási, A.-L. (2000). Topology of evolving complex networks: local events and universality. Phys. Rev. Lett., 85: 5234.

[5] Albert, R., DasGupta, B., Gitter, A., Gursoy, G., Paul, P. & Sontag, E. (2011). Computationally efficient measure of topological redundancy of biological and social networks. Phys. Rev. E, 84: 036117.
[6] Amarai, L.A.N., Scala, A., Barthélémy & Stanley, H.E. (2000). Classes of small world networks. Proc. Nat. Acad. Sci., 97 (21): 11149-11152.
[7] Bar-Yam, Y. (1997). **Dynamics of Complex Systems**. Reading, MA: Perseus Books.
[8] Barkai, N & Leibler, S. (1997). Robustness in simple biochemical networks. Nature, 387: 913-917.
[9] Basset, DS & Bullmore, E. (2006). Small-world brain networks. Neuroscientist, 12: 512-523.
[10] Borgatti, S.P. & Everett, M.G. (2005). A graph-theoretic perspective on centrality. Social Networks, 28: 466484.
[11] Bornholdt, S. & Schuster, H.G. (2003). **Handbook of Graphs and Networks: From the Genome to the Internet**. Germany: Wiley-VCH Publishing.
[12] Brown, J. & West, G.B. (eds.). (2000). **Scaling in Biology**. Oxford, England: Oxford University Press.
[13] Callaway, D.S., Newman, M.E.J., Strogatz, S.H. & Watts, D.J. (2000). Network robustness and fragility: peroclation on random graphs. Phys. Rev. Lett., 85: 5468-5471.
[14] Chan, K.P., Zhen, D. & Hui, P.M. (2004). Effects of aging and links removal on epidemic dynamics in scale-free networks. Int. J. Modern Phys. B, 18: 2534.
[15] Chow, J.Y., Davids, K., Hristovski, R., Araujo, D., & Passos, P. (2011). Nonlinear pedagogy: learning design for self-organizing neurobiological systems. New Ideas in Psychology, 29: 189 200. doi: 10.1016/j.newideapsych.2010.10.001 .
[16] Cohen, P. (2004). Small worlds key to memory. New Scientist, http://www.newscientist.com/article.ns?id=dn5012 .
[17] Csermely, P. (2004). Strong links are important but weak links stabilize them. Trends Biochem. Sci., 29: 331-334.
[18] Daniels, B.C., Chen, Y-J., Sethna, J.P., Gutenjunst, R.N. & Myers, C.R. (2008). Sloppiness, robustness and evolvability in systems biology. Current Opinion in Biotechnology, 19: 389 395.
[19] del Sol, A., Fujihashi, H. & O'Meara, P. (2005). Topology of small-world networks of protein-protein complex structures. Bioinformatics, 21 (8): 1311-1315.
[20] Elandt-Johnson, R. & Johnson, N. (1980/1999). **Survival Models and Data Analysis**. New York, NY: John Wiley & Sons, Inc.
[21] Emmert-Streib, F. & Glazko, G.V. (2011). Network biology: a direct approach to study biological function. Wiley Interdiscip Rev Syst Biol Med., 3(4):379-91. doi: 10.1002/wsbm.134
[22] Fadigas, I.S. & Pereira, H.B.B. (2013). A network approach based on cliques, Physica A, 392: 2576-2587.
[23] Feldt, S., Bonifazi, P. & Cossart, R. (2011). Dissecting functional connectivity of neuronal microcircuits: experimental and theoretical insights. Trends in Neurosciences, 34(5): 225-236.
[24] Fortney, K., Kotlyar, M. & Jurisica, I. (2010). Inferring the functions of longevity genes with modular subnetwork biomarkers of *Caenorhabditis elegans* aging. Genome Biol., 11, R13.
[25] Freeman,L.C. (1977). A set of measures of centrality based on betweeness. Sociometry, 40: 35-41.
[26] Freeman,L.C. (1979). Centrality in social networks: I. Conceptual clarification. Social Networks, 1: 215-239.
[27] Freeman,L.C. (1980). The gatekeeper, pair-dependency and structural centrality. Quality and Quantity, 14: 585-592.
[28] Freeman, L.C., Roeder, D. & Mulholland, R.R. (1980). Centrality in social networks. II. Experimental results. Social Networks, 2: 119-141.
[29] Freeman, L.C., Borgatti, S.P. & White, D.R. (1991). Centrality in valued graphs: a measure of betweenness based on network flow. Social Networks, 13: 141-154.
[30] Gallos, L.K., Makse, H.A. & Sigman, M. (2012). A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks. Proc. Nat. Acad. Sci., 109: 2825-2830.
[31] Gao, J., Buldyreve, S.V., Havlin, S. & Stanley, H.E. (2010). Robustness of a network of networks. http://arxiv.org/abs/1010.5829v1

[32] Gavrilov, L.A. & Gavrilova, N.S. (2006). Models of Systems Failure in Aging. In: P Michael Conn (Editor): **Handbook of Models for Human Aging**, Burlington, MA : Elsevier Academic Press. pp. 45-68.

[33] Gavrilov, L.A. & Gavrilova, N.S. (2004). Why We Fall Apart. Engineering's Reliability Theory Explains Human Aging. IEEE Spectrum. 41(9): 30-35.

[34] Greenbury, S.F., Johnston, I.G., Smith, M.A., Doye, J.P.K. & Louis, A.A. (2010). The effect of scale free topology on the robustness and evolvability of genetic regulatory networks. http://arxiv.org/PS_cache/arxiv/pdf/1005/1005.4342v1.pdf

[35] Gros, C. (2008). **Complex and Adaptive Dynamical systems: A Primer**. Berlin, Germany: Springer-Verlag Publishing.

[36] Han, J.D., Bertin, N., Hao, T., Goldberg, D.S., Berriz, G.F., Zhang, L.V., Dupuy, D., Walhout, A.J., Cusick, M.E., Roth, F.P. et al. (2004). Evidence for dynamically organized modularity in the yeast protein-protein interaction network. Nature, 430, 88-93.

[37] He, Y. Chen, Z. & Evans, A. (2008). Structural insights into aberrant topological patterns of large-scale cortical networks in Alzheimer's disease. J. Neurosci., 28: 4756-4766.

[38] He, Y. Chen, Z., Gong, G. & Evans, A. (2009). Neuronal networks in Alzheimer's disease. Neuroscientist, 15: 333-350.

[39] Hu, K., Van Someren, E.J.W., Shea, S.A. & Scheer, F.A.J.L. (2009). Reduction of scale invariance of activity fluctuations with aging and Alzheimer's diease: involvement of the circadian pacemaker. Proc. Nat. Acad. Sci., 106: 2490-2494.

[40] Huang, X., Gao, J., Buldyrev, S.V., Havlin, S. & Stanley, H.D. (2010). Robustness of interdependent networks under targeted attack. http://arxiv.org/abs/1010.5829v1 %bibitemHuang2012Huang, T., Zhang, J. Xu, Z.-P., Hu, L.-L., Chen, L., Shao, J.-L., Zhang, L., Kong, X.-Y., Cai, Y.-D. & Chou, K.-C. (2012). Deciphering the effects of gene deletion on yeast longevity using network and machine learning approaches. Biochimie, 94(4): 1017-1026.

[41] Jeong, H., Mason, S.P., Barabasi, A.L. & Oltvai, Z.N. (2001). Lethality and centrality in protein networks. Nature, 411(6833): 41-2.

[42] Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N. & Barabási, A.-L. (2000). The large-scale organization of metabolic networks. Nature, 407: 651-654.

[43] Képès, F. (2007). **Biological Networks**. New Jersey: World Scientific Publishers.

[44] Kirschner, M. & Gerhart, J. (1998). Evolvability. Proceedings of the National Academy of Science USA, 95: 8420-8427.

[45] Kirkwood, T.B.L. (1997). Network theory of aging. Exp. Gerontol., 32: 395-399.

[46] Kriete, A. (2013). Robustness and aging – a systems-level perspective. Biosystems, 112 (1): 37-48.

[47] Lemke, N., Heredia, F., Barcellos, C.K., Dos Reis, A.N. & Mombach, J.C. (2004). Essentiality and damage in metabolic networks. Bioinformatics, 20: 115-119.

[48] Lenski, R.E., Barrick, J.E. & Ofria, C. (2006). Balancing robustness and evolvability. PLOS Biology, 4 (12): e428.

[49] Lesne, A. (2008). Robustness: confronting lessons from physics and biology. Biological Reviews, 83: 509 532. doi: 10.1111/j.1469-185X.2008.00052.x .

[50] Li, D., Li, J. Ouyang, S., Wang, J., Wu, S., Wan, P., Zhu, Y., Xu, X. & He, F. (2006). Protein interaction networks of *Saccharomyces cerevisiae*, *Caenorhabditis elegans* and *Drosophila melanogaster*: Large-scale organization and robustness. Proteomics, 6: 456-461.

[51] Luthar, S.S., Cicchetti, D. & Becker, B. (2000). The construct of resilience: a critical evaluation and guidelines for future work. Child Development, 71 (3): 543 562.

[52] Ma, H.-W. & Zeng, A.P. (2003). Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. Bioinformatics, 19: 220-277.

[53] Ma, H.-W., Buer, J. & Zeng, A.-P. (2004). Hierarchical structure and modules in *Escherichia coli* transcriptional regulatory network revealed by a new top-down approach. BMC Bioinformatics, 5: 199. doi:10.1186/1471-2105-5-199 .

[54] Mones, E., Vicsek, L. & Vicsek, T. (2012). Hierarchy measure for complex networks. PLOS One,7(3):e33799. doi: 10.1371/journal.pone.0033799

[55] Newman, M.E.J., Barabsi, A.-L. & Watts, D.J. (2006). **The Structure and Dynamics of Networks**. Princeton: Princeton University Press.

[56] Opsahl,T., Agneessens, F. & Skvoretz, J. (2010). Node centrality in weighted networks: generalizing degree and shortest paths. Social Networks, 32 (3): 245-251.

[57] Padmanabhan, K., Wang, K. & Samatova, N.F. (2012). Functional annotation of hierarchical modularity. PLOS One, 7(4):e33744. doi: 10.1371/journal.pone.0033744
[58] Passos, J.F., Simillion, C., Hallinan, J., Wipat, A. & Zqlinicki, T. (2009). Cellular senescence: unraveling complexity. Age (Dordr). 31(4): 353-363.
[59] Pradhan, N., Dasgupta, S. & Sinha S. (2011). Modular organization enhances the robustness of attractor network dynamics. http://arxiv.org/abs/1101.5853v1
[60] Ravasz, E. & Barabási, A.-L. (2003). Hierarchical organization in complex networks. Phys. Rev. E, 67: doi: 10.1103/PhysRevE.67.026112.
[61] Ravasz, E. (2009). Detecting hierarchical modularity in biological networks. (in). **Computational Systems Biology: Methods in Molecular Biology**, v. 541, 145-160. (eds.). McDermott, J., Samudrala, R., Bumgarner, R.E., Montgomery, K. & Ireton, R. New York, N.Y.: Humana Press.
[62] Ravasz, E., Somera, A.L., Mongru, D.A., Oltvai, Z.N. & Barabasi, A.L. (2002). Hierarchical organization of modularity in metabolic networks. Science, 297: 1551-1555.
[63] Reed, W.J. (2006). A brief introduction to scale-free networks. Natural Resource Modeling, 19 (1): 3-13.
[64] Rives, A.W. & Galitski, T. (2003). Modular organization of cellular networks. Proc. Nat. Acad. Sci. USA, 100: 1128-1133.
[65] Saavedra, S., Reed-Tsochas, F. & Uzzi, B. (2008). Asymmetric disassembly and robustness in declining networks. Proc. Nat. Acad. Sci., 105: 16466-16471.
[66] Schneider, C.M., Araujo, N.A.M., Havlin, S., & Herrmann, H.J. (2011). Towards designing robust coupled networks. http://arxiv.org/abs/1106.3234
[67] von Neumann, J. (1956). Probabilistic logics and the synthesis of reliable organisms from unreliable components. (in). Shannon, C. (ed.). **Automata Studies**. Princeton, NJ: Princeton University Press.
[68] Sola, S. (2013). Self-sustained activity in a small-world network of excitable neurons. http://online.itp.ucsb.edu/online/brain04/solla/
[69] Stam, C.J., de Haan, W., Daffertshofer, A., Jones, B.F., Manshanden, I., van Cappellen, A.M., van Walsum, T., Montez, J.P.A., Verbunt, J.C., de Munck, B.W. & Scheltens, P. (2009). Graph theoretical analysis of magnetoencelphalographic functional connectivity in Alzheimer's disease. Brain, 132: 213-224.
[70] Strogatz, S.H. (1994). **Nonlinear Dynamics and Chaos**. New York, NY: Addison Wesley Publishers.
[71] Strogatz, S.H. (2001). Exploring complex networks. Nature, 408: 268-276.
[72] Strogatz, S.H. (2003). **Sync: How Order Emerges from Chaos in the Universe, Nature and Daily Life**. New York, N.Y.: Hyperion Books.
[73] Stromberg, S.P. & Carlson, J. (2006). Robustness and fragility in immuno-senescence. PLOS Comput. Biol., 2: e160.
[74] Thadakamalia, H.P., Kumara, S.R.T. & Albert, R. (2007). Complexity and large-scale networks. (in). **Operations Research and Management Science Handbook**. (ed.) A.R. Ravindran. Boca Raton, FL: CRC Press.
[75] Thadakamalia, H.P., Raghavan, U.N., Humara, S. & Albert, R. (2004). Survivability of multiagent-based supply networks: a topological perspective. Intelligent Systems, 19 (5): 24-31,
[76] Van Noort, V., Snel, B.,& Huynen, M.A. (2004). The yeast coexpression network has a small-world, scale-free architecture and can be explained by a simple model. EMBO Rep. 5(3): 280284. doi:10.1038/sj.embor.7400090.
[77] Wagner, A. & Fell, D. (2001). The small world inside large metabolic networks. Proc. Roy. Soc. Lond. B: 1803-1810.
[78] Wagner, A. (2005). **Robustness and Evolvability in Living Systems**. Princeton, NJ: Princeton University Press.
[79] Watts, D.J. & Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks. Nature, 393(6684): 440-442.
[80] Watts, D. J. (1999). **Small Worlds: The Dynamics of Networks Between Order and Randomness**. Princeton, NJ: Princeton University Press.
[81] http://en.wikipedia.org/wiki/Clustering_coefficient

[82] Witten, T.M. (1985a). Reliability theoretic methods and aging: Critical elements, hierarchies, and longevity—Interpreting survival curves, (in) **The Molecular Biology of Aging**. (eds.) A. Woodhead, A. Blackett, and R. Setlow. New York: Plenum Press.
[83] Witten, T.M. (1985b). A return to time, cells, systems and aging: III. Critical elements, hierarchies, and Gompertzian dynamics, Mech. Ageing and Dev., 32: 141-177.
[84] Witten, T.M. (2007). (M,R)-systems, (P,M,C)-nets, hierarchical decay and biological aging: Reminiscences of Robert Rosen. Chemistry and Biodiversity, 4 (10): 2332-2344.
[85] Witten, T.M. & Bonchev, D.G. (2007). Predicting aging/longevity-related genes in the nematode *C. elegans*. Chemistry and Biodiversity, 4: 2639-2655.
[86] Wuchty, S. (2001). Scale-free behavior in protein domain networks. Mol. Biol. Evol., 18(9): 1694-1702.
[87] Yates, F.E. & Benton, L.A. (1995a). Biological senescence: loss of integration and resilience. Can. J. Aging, 14: 106-130.
[88] Yates, F.E. & Benton, L.A. (1995b), Loss of integration and resiliency with age: a dissipative destruction. (In). **Handbook of Physiology and Aging**. Bethesda, MD: American Physiological Society. Section 11. Chapter 22. pp. 591-610.
[89] Zhang, A. (2009). **Protein Interaction Networks: Computational Analysis**. Cambridge, England: Cambridge University Press.

Center for the Study of Biological Complexity, Virginia Commonwealth University, Richmond, VA 23284-2030 US
E-mail address: tmwitten@vcu.edu