A Network Model of Knowledge Acquisition

A. Martin Buoncristiani
Patricia Buoncristiani
Thinking and Learning in Concert
310 Charity Lane, Newport News, Virginia
23602, USA
ausaTLC@aol.com
www.thinkingandlearninginconcert.org

Abstract
Starting from the idea that organization of knowledge in an individual's mind can be considered a network, we use recent results from network science to speculate on how information can be organized in the brain. A hierarchy of three increasingly sophisticated network systems -- a random network, a small world network and a scale free network can be used to model the increasing sophistication of knowledge networks as they evolve during learning. These ideas are useful to teachers seeking to help their students organize their new knowledge.

Key words: education, cognition, teaching, learning, innovation

1. Introduction

The U. S. National Academies of Sciences and Engineering commissioned a survey of recent research in a variety of fields studying the brain and its functions. The theme of this study was how this knowledge can be used by educators. A report on the findings of this survey called "How People Learn" was published in 2000 (Bransford, Pellegrino & Donovan 2000). This report describes an emerging understanding that effective learning requires an individual's introspection – a metacognition or thinking about thinking. One focus of this metacognition is on the organization of newly acquired knowledge.

One of the three principle findings of this report encourages teachers to help their students

> "... understand facts and ideas in a context of a conceptual framework and to organize their knowledge in ways that facilitate retrieval and application"

Each learner must make a conscious effort to organize the information learned so that information becomes flexible enough to use in new or different circumstances. Without this organization, knowledge remains isolated and inert. Understanding how to perform this organization is not an innate characteristic of humans; it must be learned. Teachers who can help their students with this organization will increase the efficiency of their students' learning and thinking.

The major difficulty in carrying out this charge is that each child's mind is accessible only to that child. Science can study an individual's brain, locating where different mental activities are activated, but science cannot delve into an individual's mind (Damasio 2001). Consequently, we cannot now measure how knowledge is stored in someone else's mind. It would help if we had some model of the possible ways that a person could organize knowledge in their mind.

Teachers have the knowledge they expect to transmit to their students embedded in their own conceptual framework -- their own mental construct. They can impart the knowledge they have but they cannot transfer the construct in which that knowledge resides. A mental construct is personal and depends on the state of ones' mind when the knowledge was acquired and when that knowledge was used. So how can we help our students find a construct? To shed some light on this question we propose a structure for knowledge organization based on recent results in network science. This model will provide a generalized setting in which a mental construct can be discussed.

The organization of information is an old question. Diderot and d'Alembert had considered how human knowledge can be organized and how that might reflect on the way knowledge is organized in the human brain (Diderot & d'Alembert 1751). The result was a hierarchical structure called the *"Tree of Knowledge"* which has been added to from time to time. That early work has been revisited recently because it bears on how one might construct a knowledge network using hypertext (Rockwell 1999). While it is useful to consider that work as an example of a network structure for knowledge external to a human brain, we are striving to understand the network construct within an individual mind.

In this paper we introduce a simple (and very speculative) model of the organization of knowledge in a human's mind. The model is based on what we have learned recently about the science of networks. While this model lacks the rigor of a fully developed scientific theory it does provide insight into the possible stages of development of the mind as a storehouse of knowledge and can be used to inform teachers of the organizational processes going on in their student's mind.

2. The Emergence of Network Science

A new area of scientific inquiry focused on the study of networks blossomed during the last decade of the 20th century. It grew out of our need to understand the structure and dynamics of the many complex networks that are vital to our well being. This science developed from disparate results in a variety of fields (mathematics, physics, sociology, psychology and engineering) over the last 75 years and has been applied to understand the behavior of networks for communication, distribution of resources, epidemics, management in organizations and, of course, most famously the Internet. There are two popular books describing network science, written by important contributors to the field (Watts 2003), (Barabási 2002) and a there is a comprehensive technical review paper (Albert & Barabási 2001)

Perhaps the most familiar result in this field is the work on social networks of psychologist Stanley Milgram who endeavored to understand the connectivity of human social interactions (Travers & Milgram 1969). He performed several experiments in which individuals in one community (in Kansas) endeavored to make contact with a remote target individual (in Massachusetts) using the social network of their friends and their friends' friends. The resulting notion of "six degrees of separation" implies that an individual in a particular population can contact another individual through a directed chain of, on average, six contacts.

2.1 Properties of a Network

A network is a collection of objects called nodes or vertices, connected in pairs by lines called links or arcs, Figure 1. The linked objects can be anything from chemical reactions in complex biological processes to computers on the World Wide Web. One of the easiest networks to visualize is the social network of your friends. You and each of your friends are nodes connected by links representing friendship. Your node is connected to the node of each of your friends. In addition, the node of each of your friends is connected to their friends, some, but not all, of whom are also your friends. As the network expands to more and more layers of friends it comes to cover the globe. The surprising result of Milgram was that despite the wide range and apparent complexity of a social network of friends it is still possible to establish contact between two individuals with relatively few intermediaries.

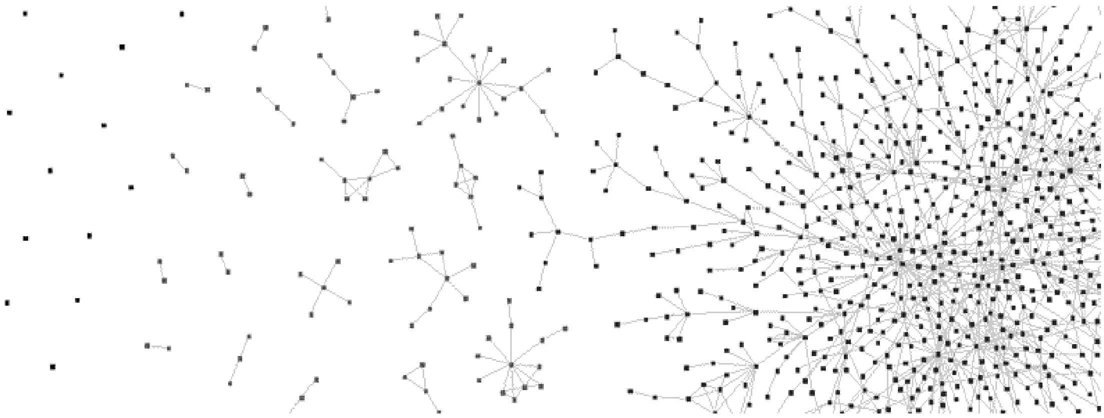

Figure 1. A network is a collection of objects called nodes or vertices connected by lines called links or arcs. In this diagram each point represents an individual node in the network and each line represents a link. The network shown grows more complex from left to right. The source of this figure was http://www.orgnet.com/prevent.html.

Despite the complexity of a network there are a few characteristic properties which help to describe its structure. One of these properties is the average length of the links which serves as a measure of penetration of links into the network. A second property is the cluster coefficient which measures how tightly the linked nodes are clustered together. In a network of friends the cluster coefficient is determined by how many of your friends are friends.

Figure 2 illustrates the idea of a cluster coefficient. Part (a) of the figure shows an individual (the starred node) and four friends (linked by solid lines). Two friends are the star's nearest neighbors and two are the next nearest neighbors. If each node had the same pattern of friends (nearest and next nearest neighbors) then the star's friends who were friends among themselves would be linked with the dotted lines shown in part (b). Thus three of the star's friends are friends. There are six possible ways that the four friends could be linked. The cluster coefficient is the ratio of the actual number of friend's friends to the largest possible number of links (3/6, in this case). In complex networks the determination ot the cluster coefficient is carried out by computer.

Figure 2. This is a simple illustration of the cluster coefficient. Part (a) illustrates an individual (the starred node) and four friends (linked by solid lines). Two friends are the star's nearest neighbors and two are the next nearest neighbors. If each node had the same pattern of friends (nearest and next nearest neighbors) then the star's friends who were friends would be linked with the dotted lines. Thus three of the star's friends are friends.

There are three stages in the development of a network science. The seminal work on network science was carried out by Paul Erdös and Alfréd Rényi beginning in the 1950's. Their work on the first large scale network studied the generation of networks by randomly assigning links. Figure 3 shows a random network.

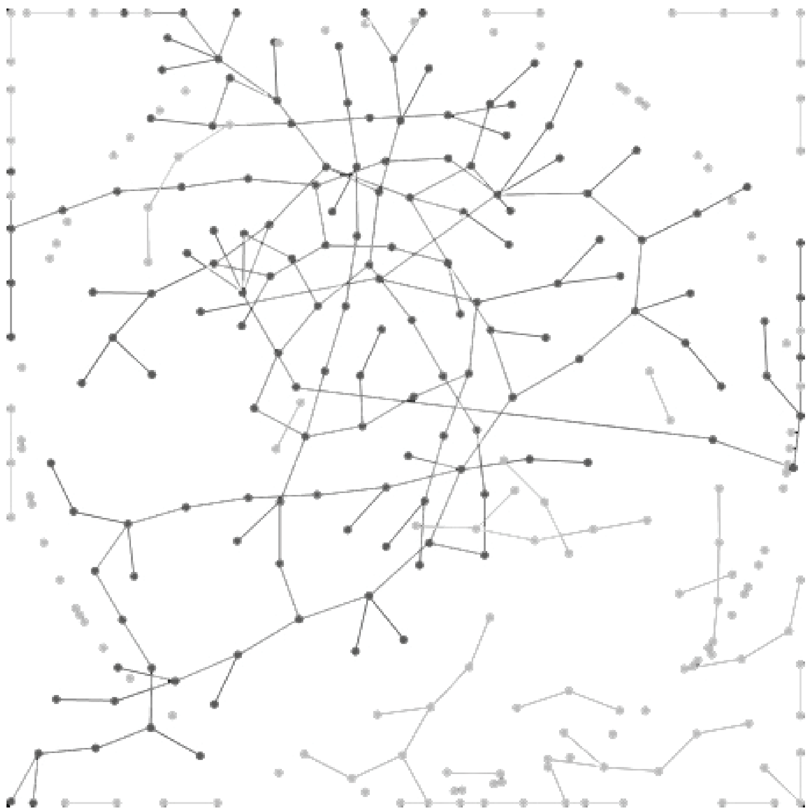

Figure 3. This is a depiction of a 300 node random network. Three hundred unconnected nodes initially uniformly distributed on a circle are randomly connected. The largest connected component of the network is indicated by darker nodes and links and it has 153 nodes. This figure was generated using software developed by Uri Wilenskyof Northwestern University and is incorporated in NetLogo: http://ccl.northwestern.edu/netlogo/models/.

The next significant contribution to network science came in the late 1990's and arose because of the ability of computers to explore complex network structures. Duncan Watts and Steve Strogatz sought to model the results of Stanley Milgram "small world" system. They recognized that real networks are not entirely random but have an element of order. So they developed a one dimensional model of a completely ordered network where each node is connected to its nearest neighbor and next nearest neighbor (like the example in Figure 2). They then systematically disordered the model by reconnecting links randomly, see Figure 3. They discovered "small world networks" have a characteristic relationship between the average link length and the cluster coefficient. They studied a large number of diverse networks have the same characteristics.

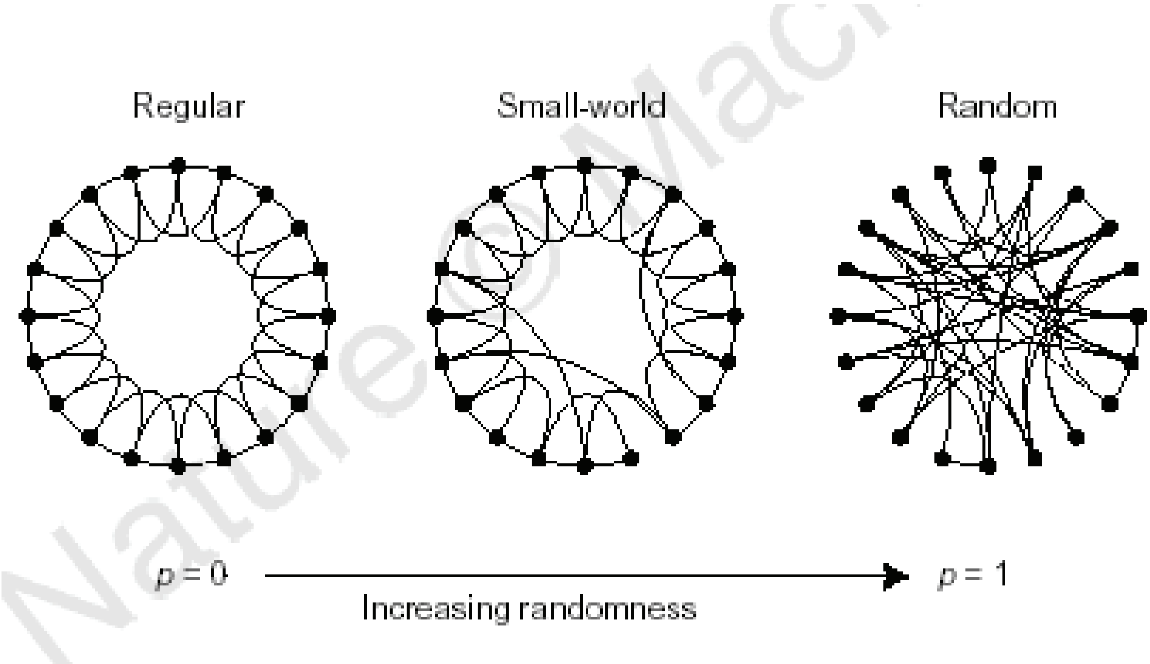

Figure 3. This figure illustrates the idea of Duncan Watts and Steve Strogatz who developed the first model to describe the results of Stanley Milgram on "six degrees of separation". A twenty node model is depicted so that the links can be clearly seen. In fact, the actual model involves thousands of nodes. The rule for connecting nodes in the organized model is the same as in Figure 2. Links are then systematically reconnected at random and at each stage the average path length and cluster coefficient are determined. This figure is from (Watts & Strogatz 1998).

The third stage in network science development came from the group of Albert-László Barabási at the University of Notre Dame in the early 2000's. Using the computer to crawl through large network systems they found that there was a tendency for links to cluster around specific nodes called hubs (Barabási A.-L., Albert R. 1999). There is a preference for a few nodes to have a very large number of links and a preference for new links to be made to these hubs. This preferential attachment is some times compared to "The rich get richer" Figure 4 shows a typical network with preferential attachment. In the figure, the hub nodes are shown as larger. In a network with this sort of "preferential attachment" there are a large number of nodes with a few links and a few nodes with a large number of links.

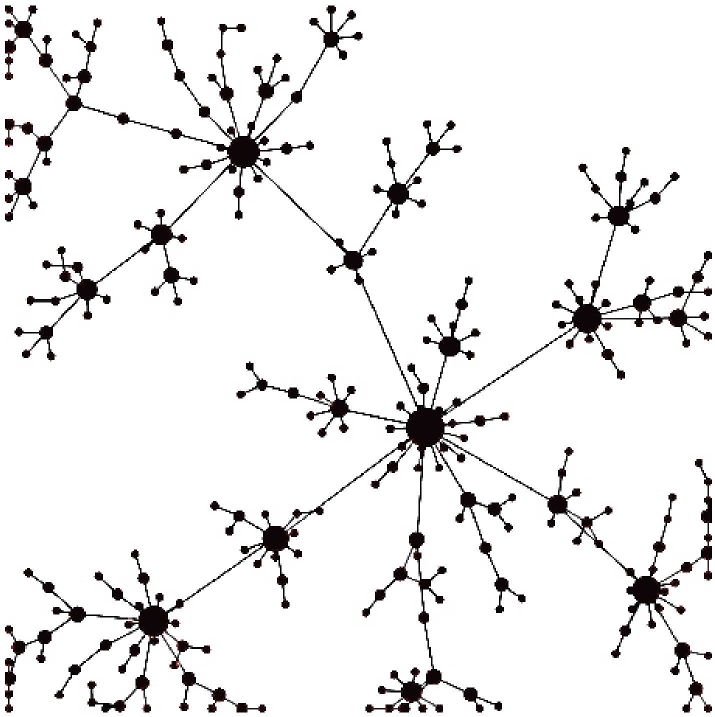

Figure 4. This figure shows a typical network with preferential attachment. In the figure, the hub nodes are shown as larger, their size proportional to the number of links. In a network with this sort of "preferential attachment" there are a large number of nodes with a few links and a few nodes with a large number of links. This figure was generated using software developed by Uri Wilensky of Northwestern University and is incorporated in NetLogo: http://ccl.northwestern.edu/netlogo/models/.

2.2 Order and Randomness in Networks

We can understand "six degrees of separation" phenomenon from the work of Watts and Strogatz (Watts & Strogatz 1998) who constructed a simple computer simulation of Milgram's "small world network" -- the social structure of our friends. In their analysis, one part of this network configuration comes from the structure of our lives, where we live, work, go to church, etc.; and another part comes from choices we make such as organizations to which we voluntarily belong. So in the social network setting an individual's friends are a combination of the order imposed by the structure in their life and the randomness of their own volition. Sociologists refer to this as a competition between structure and agency. The same sort of competition between order and disorder is characteristic of all small world networks.

The most impressive result of this new science is that dramatically different types of networks exhibit common properties. It is our premise here that it is reasonable to consider an individual's mental construct for knowledge to be a network and consequently, that it shares these common properties. This understanding is useful to the extent that it elucidate the organization of knowledge.

The technical details of network science are not so important here. What is important is that there is a hierarchy of three increasingly sophisticated networks -- a random network, a small world network and a scale free network – and these can be used to model the increasing sophistication of knowledge networks. We describe this hierarchy in sequel as we develop a model for knowledge organization.

3. A Model of Knowledge Acquisition

We can think of the context of one's knowledge as a network – a system of objects (called nodes) connected by links. Each bit of knowledge in the brain can be represented as a node in a network and the connection among individual bits of knowledge can be represented by links. The resulting pattern of nodes and links forms a network – the network of an individual's knowledge, their mental construct.

The knowledge network of any two individuals can be expected to be as different as their fingerprints. And just as there are common characteristics for fingerprint features, there should be some set of common organizing principles governing how knowledge networks are formed and if we can understand these principles we may be better able to guide the evolution of the mind's acquisition and organization of new knowledge. Indeed, this is the imperative of the primary findings of "How People Learn".

The following is a hypothetical scenario of such an evolution based on the nascent science of networks. While we may not now be able to prove this hypothesized development, it does provide a tantalizing and plausible story for how knowledge grows within the mind. It illustrates the unsatisfactory nature of the simple organizational schemes and it reinforces the need for personal effort in learning to achieve sophisticated learning.

When a child first begins to learn he/she is flooded with sense data; each new bit of information appears to be the same as every other bit since there is no basis to discriminate between them. Since there is no discrimination among the bits they are initially linked at random. The early learning network is a random network. It is featureless, with no organization of the information it contains.

As more information becomes lodged in children's minds the network topology changes as connections between bits of information become recognized and linked purposefully. It is transformed from a random network to a "small world network". Order begins to emerge from randomness. As more connections are discovered between bits of information they begin to cluster. In a small world network there are clusters of knowledge but also a few nodes in each cluster have long range links, a remnant of the random network stage. These long range links provide a connection among clusters.

As new information is added to the knowledge network its topology changes again. The attachment of links begins to prefer those nodes which have already established more links. These preferred nodes, called hubs, form the centers about which information is organized. This organization forms a scale free network. The chief characteristic of this stage is the appearance of hubs – centers about which knowledge is arranged without any bounds.

Thus the evolution of the organization of knowledge can be seen to pass through three stages. In early learning, knowledge is organized at random and the landscape is featureless. Later, as connections among the pieces of information are discovered, some bits of information begin to cluster while other bits still retain a random element. Finally, the sophisticated learner begins to organize information around hubs.

Table 1. Development of the Knowledge Network

Early Learning	Middle Learning	Polished Learning
Random Network	Small World Network	Scale Free Network
Structureless	Clusters	Hubs
Bounded range	Bounded Range	Unbounded Range

An important characteristic of network structure that gives some insight into knowledge acquisition involves the distribution of links among the nodes, described by a degree distribution. This is a function showing how links are distributed among the nodes expressed as the number of nodes as a function of the number of links. For a random network the degree distribution is a Poisson distribution. That means that there is a mean number of links attached to nodes and a maximum and minimum number of links. These bounds on the number of links to any one node impose a limit on the complexity of the network. The degree distribution for a small world network is approximately a Poisson distribution so the same limitations apply. On the other hand, the scale free network has a power law (Pareto) distribution so there is no limit to the number of links to nodes. Thus some hubs can have an unlimited number of links and the resulting network has no limit on its complexity imposed by the network structure. There certainly are limits on the complexity of a knowledge network imposed by the brain itself. This justifies thinking of the scale free network as a network of the sophisticated knower.

4. Conclusion

This simple view of the evolution of an individual's organization of knowledge suggests a strategy for teachers. In aiding students to transform from the Early to the Middle stage teachers should assist students in recognizing connections among bits of information they are learning. In assisting the transition to Sophisticated Learning teachers need to help students identify hubs of knowledge and fit what they are learning into these hubs. This description of the transition from an early learner to a sophisticated learner in stages could apply when a sophisticated learner takes on a new subject.

The model presented here is consistent with the findings of the report on "How People Learn". That report discourages the teaching disconnected facts. That mode of teaching promotes the organization of knowledge in a random network, the least sophisticated level of knowledge organization. The report also discourages the teaching in separated disciplines. It favors instead the teaching of concepts across disciplines corresponding to the organization of using hubs that extend over fields.

We have introduced an idea about how knowledge can be organized in an individual's mind in three stages of increasing complexity. The idea is useful in thinking about how people learn and may be useful for teachers who seek to help their students organize their new knowledge into a conceptual framework. There remains work to be done to validate this idea and to study the process of growth a knowledge network.

5. Bibliography

Albert R., Barabási A.L. (2001) *"Statistical Mechanics of Complex Networks"*, Review of Modern Physics, Vol. 74, No. 1. pp.47-97.

Barabási A.-L. (2002) "Linked: How Everything is Connected to Everything Else and What It Means for Business, Science and Everyday Life", Plume (Penguin Group), New York

Barabási A.-L., Albert R. (1999) "Emergence of Scaling in Random Networks", Science, Vol 286, No. 5439, pp. 509-512.

Bransford J, Pellegrino J., Donovan S. (2000) *How People Learn*, National Academy of Sciences, National Academy Press, Washington, DC.

Damasio A. R. (2001) *"How the Brain Creates the Mind"*, Scientific American Special Publication on the Brain.

Diderot D., d'Alembert J. *"The Tree of Knowledge"* first appeared in "*Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers"* (Encyclopedia of the sciences, arts and crafts), an early encyclopedia published in France beginning in 1751.

Rockwell G. (1999) *"Diderot and Hypertext."* unpublished essay presented in January 1999 to the *McMaster Association for Eighteenth-Century Studies.* http://www.geoffreyrockwell.com/publications/Diderot.Hypertext.pdf\ .

Travers J., Milgram S. (1969)"An Experimental Study of the Small World Problem" Sociometry, Vol. 32(4), pp. 425-443.

Watts D. J., Strogatz S. H. (1998) *"Collective Dynamics of 'Small World' Networks"*, Nature, Vol. 393, No. 4, pp 440-442.

Watts D. J. (2003) "*Six Degrees: The Science of a Connected Age*", W. W. Norton & Co. Inc., New York.