GRAPH THEORY IN PRACTICE: PART II

Brian Hayes

A reprint from

American Scientist

the magazine of Sigma Xi, the Scientific Research Society

GRAPH THEORY IN PRACTICE: PART II

Brian Hayes

Part I of this article, in the January–February issue, discussed some very large structures that can usefully be looked upon as mathematical graphs. In this context a graph is a set of vertices (which are usually represented as dots) and a set of edges (lines between the dots). One large object that can be described in this way is the World Wide Web; its 800 million pages are the vertices of a graph, and links from one page to another are the edges. A second example comes out of Hollywood: The vertices are 225,000 actors, and an edge connects any two actors who have appeared in a feature film together.

Although graph theory has a history of two centuries and more, only in recent years has it been applied routinely to structures like these, with many thousands or millions of vertices and edges. Studying such enormous graphs is by no means easy. The Hollywood collaboration graph just barely fits in the memory of a large computer. The Web, a few orders of magnitude larger, requires all the resources of the Internet to keep track of its tentacles. Certain other graphs are even bigger. The human acquaintanceship graph, with a vertex for every person on earth and edges linking all those who know each other, may never be recorded beyond a few small, sampled regions.

Even when the vertices and edges of a large graph can be catalogued in full detail, gaining a deeper understanding of the graph's structure still calls for something more. What's needed is a mathematical theory or model. Typically this takes the form of an algorithm for generating new graphs that share certain properties of the graph under examination: You understand the original graph by building structures that resemble it. Part II of this article looks at a few such models.

The Small World of Large Graphs

The various gigantic graphs that have lately attracted notice share other properties besides sheer size. In particular:

They tend to be sparse. The graphs have relatively few edges, considering their vast numbers of vertices. In a graph with n vertices, the maximum number of edges is $n(n–1)/2$, or roughly $n^2/2$. (I consider here only "simple" graphs, as opposed to multigraphs, where more than one edge can join a pair of vertices.) In large real-world graphs, the number of edges is generally closer to n than to $n^2/2$. For example, the Hollywood graph has 13 million edges connecting its 225,000 vertices. That sounds like a lot, but it falls far short of the 25 billion edges in a "complete graph," or "clique," where an edge joins every pair of vertices.

They tend to be clustered. In the World Wide Web, two pages that are linked to the same page have an elevated probability of including links to one another. Likewise among friends, if two people both know you, there's a higher-than-normal chance they also know each other. Thus the edges of the graph are not distributed uniformly but tend to form clumps or knots.

They tend to have a small diameter. The diameter of a graph is the longest shortest path across it, or in other words the length of the most direct route between the most distant vertices. Diameter is finite only for connected graphs—those that are all in one piece. A connected graph must have at least $n–1$ edges, and its largest possible diameter is $n–1$. At the opposite extreme, a complete graph, with $n^2/2$ edges, has a diameter of 1, since you can get from any vertex to any other in a single step. Graphs nearer to the minimum than the maximum number of edges might be expected to have a large diameter. Clustering could increase the diameter further still, since edges used up in creating local clumps leave fewer edges available for long-distance connections. Nevertheless, the diameter of the Web and other big graphs seems to hover around the logarithm of n, which is much smaller than n itself.

Graphs with the three properties of sparseness, clustering and small diameter have been termed "small-world" graphs, after the familiar cocktail-party experience of making a new acquaintance in a distant city and discovering you have a friend in common. The name was introduced by Duncan J. Watts and Steven H. Strogatz of Cornell University, whose 1998 paper in *Nature* discussed the Hollywood graph and several other examples. Watts, who has since moved on to the Santa Fe Institute and MIT, describes the graphs more fully in a recent book titled *Small Worlds*.

Brian Hayes is a former editor of American Scientist. *Address: 211 Dacian Avenue, Durham, NC 27701. Internet: bhayes@amsci.org*
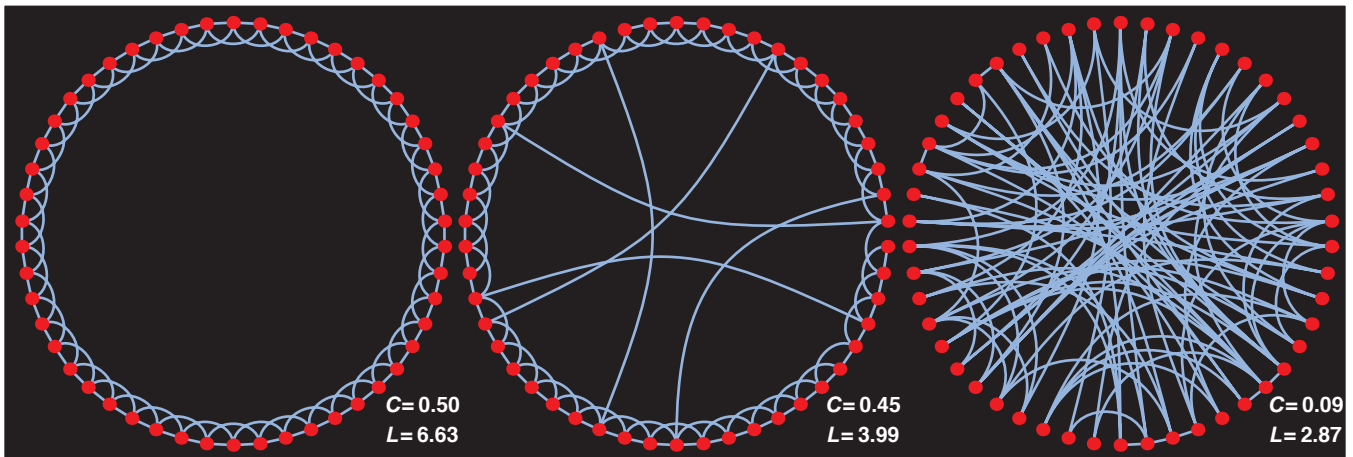
Figure 1. Watts-Strogatz model interpolates between a regular lattice *(left)* **and a random graph** *(right)***. Randomly rewiring just a few edges** *(center)* **reduces the average distance between nodes,** *L***, but has little effect on the clustering coefficient,** *C***. The result is a "small-world" graph.**

Lattices and Random Graphs

In trying to understand very large small-world graphs, two simplified models serve as useful points of reference.

The simplest of all graph models are lattices—highly regular graphs in which every vertex is joined to a few neighbors. The term lattice brings to mind a two-dimensional square grid, but a lattice can have other geometries. The minimal lattice is a one-dimensional structure, like a row of people holding hands. Bending this linear lattice around and joining the two ends creates a ring lattice, or cycle. A nearest-neighbor ring lattice with n vertices has n edges, and every vertex has degree 2, meaning that two edges meet there. When edges extend both to nearest neighbors and to next-nearest neighbors, the ring has $2n$ edges and vertices of degree 4.

How does the ring lattice perform as a model of small-world graphs? Not terribly well. It is suitably sparse, with just n edges in the nearest-neighbor case, and there is a sense in which it is highly clustered, since all the edges are "local." But the diameter is not small. The only way to travel in a ring is to go from neighbor to neighbor; the lattice is like a railroad line without an express track. The diameter of the nearest-neighbor ring is $n/2$, which is much larger than log n.

Whereas a lattice is a very orderly graph, the other benchmark model is maximally random. The graphs in this class were first studied around 1960 by the Hungarian mathematicians Paul Erdős and Alfred Rényi. To build one of their graphs, you start with a collection of n vertices and no edges. Then you make a sweep through the graph, considering every possible pairing of vertices, and in each case you either draw an edge with probability p or do nothing with probability $1-p$. The outcome of this process is easy to predict in the extreme cases: If $p=0$, the graph remains edgeless, and if $p=1$, the graph becomes a clique. Between the extremes, you can expect the graph to have about $pn^2/2$ edges, placed randomly and independently.

Erdős and Rényi proved a number of interesting results about these graphs. Most of the proofs are statements about "almost every" random graph; this sounds like a strangely vague manner of speaking for mathematical discourse, but it has a precise meaning. Saying that almost every random graph has some property Q means that as the size of the graph n goes to infinity, the probability of Q approaches 1. For example, Erdős and Rényi showed that if the edge probability p is greater than a certain threshold, then almost every random graph is connected. This doesn't mean you can't construct disconnected graphs if you set out to do so, but the random process has no chance of producing them when n approaches infinity.

As a model of small-world networks, the Erdős-Rényi random graph has some strengths. It can be made as dense or as sparse as necessary just by adjusting the edge probability p. And the diameter tends to be small (in some cases *too* small). But Erdős-Rényi graphs show no tendency to form clusters. They cannot, since the edges are placed independently, and neighbors of neighbors are no more likely to be linked than any other randomly chosen vertices.

Rewiring the Lattice

It's not really surprising that the lattice model and the Erdős-Rényi model fail to reproduce some features of networks such as the human friendship graph or the World Wide Web. After all, these real-world networks are neither entirely regular nor entirely random. People generally know their neighbors, but their circle of acquaintances is not confined to those who live next door, as the lattice model would imply. Conversely, links between pages on the Web are not created at random, as the Erdős-Rényi process requires.

Watts and Strogatz deal with these failures by a strategy that seems perfectly obvious once someone else has thought of it: They interpolate between the two models. They begin with a regular lattice, such as a ring, and then "rewire" some of

the edges to introduce a measure of randomness. Each edge in the original lattice is examined in turn, and is either left in place or else is redirected to another randomly chosen destination. The decision to rewire an edge is governed by a probability p, which can be adjusted over the range from 0 to 1. If p is equal to 0, then nothing happens; the lattice is unchanged. If p is equal to 1, the lattice is transformed into a random graph much like a product of the Erdős-Rényi procedure. The interesting range lies between these extremes.

In their analysis of the rewired graphs, Watts and Strogatz examined not the diameter—the shortest path between the most distant vertices—but the minimum path length L averaged over all pairs of vertices. They observed an abrupt transition in L as the rewiring probability increased. L is at its maximum in the regular lattice, but it falls steeply when just a few of the edges are rewired.

As a measure of clustering in their hybrid graphs, Watts and Strogatz defined a coefficient C. To calculate C, list all the neighbors of a vertex, count the edges that link those neighbors, and divide by the maximum number of edges that could possibly be drawn among the neighbors; then repeat these operations for all the vertices, and take the average. In contrast to the path length L, the clustering coefficient remains high until the rewiring probability is rather large. Hence over a wide range of p values, the graph is dominated by local connections between nearby nodes, but the few shortcuts provide efficient long-distance connections.

Watts and Strogatz tested their model against the Hollywood graph, which has $L=3.65$ and $C=0.79$. They created a rewiring model with the same number of vertices and edges and the same average degree. The model has an L value of 3.9 and a C coefficient in the range from 0.61 to 0.84, in good agreement with the actors' graph. The Erdős-Rényi model cannot do nearly as well: The

path length of $L=2.99$ is reasonably close, but the cliquishness score, at $C=0.00027$, is too small by three orders of magnitude.

Watts and Strogatz suggest several more applications of their model to natural or technological graphs, including a part of the U.S. electric power grid and the nervous system of the nematode worm *Caenorhabditis elegans*. Again, model and observation are consistent, although the agreement is not as impressive as it is with the Hollywood graph. More recently, Lada A. Adamic of the Xerox Palo Alto Research Center has shown that the Web also fits the small-world model.

As it happens, Watts and Strogatz were not the first to explore the effect of short-circuiting a ring graph. In the 1980s, Fan R. K. Chung (now at the University of California, San Diego), in collaborations with Michael R. Garey of AT&T Laboratories and Béla Bolobás of the University of Memphis, studied various ways of adding edges to cyclic graphs. They found cases where the diameter is proportional to log n.

Finding Your Way in a Small World
Although randomly rewired connections can dramatically shrink the diameter of a lattice, the shortcuts are not much use if you don't know where they are. Without an aerial view, how do you find the best path to an unknown destination? The question has recently been taken up by Jon Kleinberg of Cornell University. His answer leads to a refinement of the Watts-Strogatz model.

Kleinberg describes his model in the context of the social-network experiments carried out by Stanley Milgram in the 1960s. These experiments gave the first clear evidence of a small-world structure in the acquaintanceship graph. Milgram prepared envelopes addressed to a person in the Boston area and asked volunteers in Nebraska to pass the envelopes along to any acquaintance who might be able to get them closer to their des-
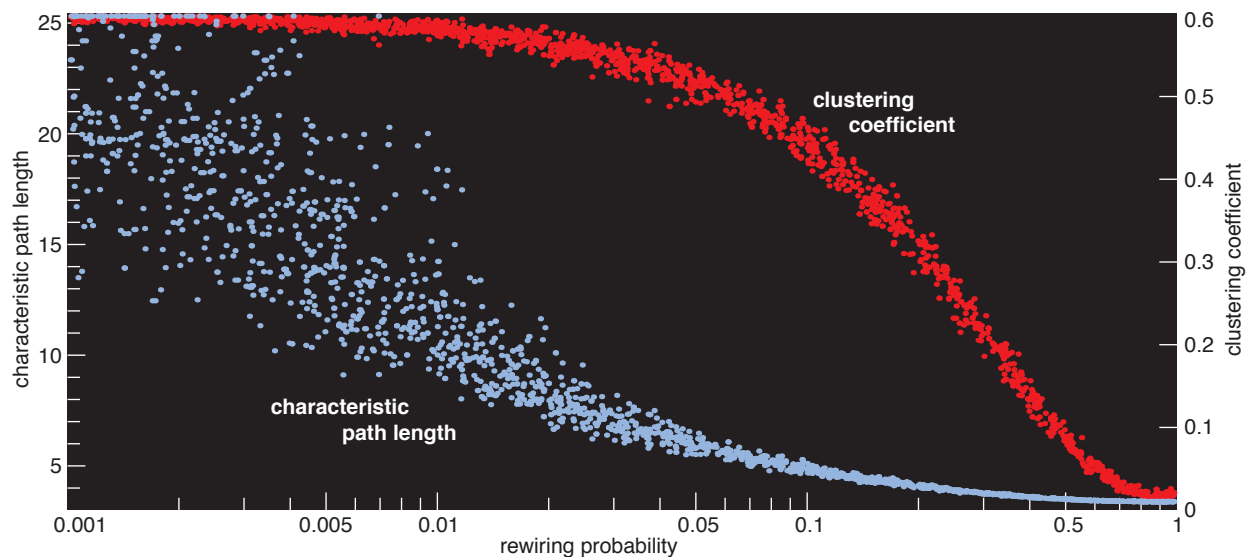

Figure 2. As the probability of rewiring increases in the Watts-Strogatz model, the characteristic path length falls off long before the clustering coefficient drops. Results are from 2,000 random graphs, each with 300 vertices and 900 edges.

tination. Many of the envelopes reached the recipient after passing through the hands of just a few intermediaries. What is surprising about this outcome is not that short paths exist but that people were able to find them so easily.

Kleinberg's model of this process begins with a two-dimensional square lattice, where each node is joined to its four nearest neighbors. Then long-distance interconnections are added, but not purely at random. For each vertex, all the possible destinations of a shortcut link are assigned a rank based on their distance from the source vertex. The probability of choosing a vertex at distance d is proportional to d^{-r}, where r is an additional parameter of the model. If r is set equal to 0, then destinations at all distances are chosen with uniform probability, and the model is just a two-dimensional version of the Watts-Strogatz model. If r is large, then only nearby destinations have any appreciable chance of being chosen, and the original structure of the lattice is hardly altered. The crucial parameter value turns out to be $r=2$, where the probability obeys an inverse-square law.

Graphs with $r=2$ are easy to get around in not because they have the smallest diameter (they don't) but because an algorithm exists for finding a short path through them. The algorithm is a simple "greedy" one. If you are asked to find a route from node a to node b, list all the edges emanating from a, and choose the one that takes you closest to b, as measured by lattice distance; then repeat the same procedure from this intermediate point, and continue until you reach the destination. Kleinberg proves that this algorithm is at its most efficient when the spectrum of edge lengths is determined by $r=2$, and also that no other algorithm performs better at any other value of r. When $r=0$, paths with fewer steps exist, but there is no way to find them; at large r, the optimum route is unlikely to be much shorter than a path following strictly local lattice links.

Is there any evidence that graphs we meet in everyday life have the convenient properties of a lattice augmented with inverse-square shortcuts? The success of Milgram's experiments might be taken to suggest that the acquaintanceship graph has such a structure, but the data are scanty. And the model is not easy to adapt to other contexts. It requires that a graph have a geometric substrate—some way of measuring distance other than counting edges in the graph itself. This kind of metric does exist in Milgram's experiments: People in Nebraska know which of their friends are nearer to Boston. Many other graphs, however, have no such geometric structure. The Web offers few clues to proximity; when you go searching for a site by clicking on links, you seldom know whether you're getting warmer or cooler.

Power Laws

Sparseness, clustering and small diameter are not the only distinctive properties of large real-world graphs. Another trait that has attracted notice is
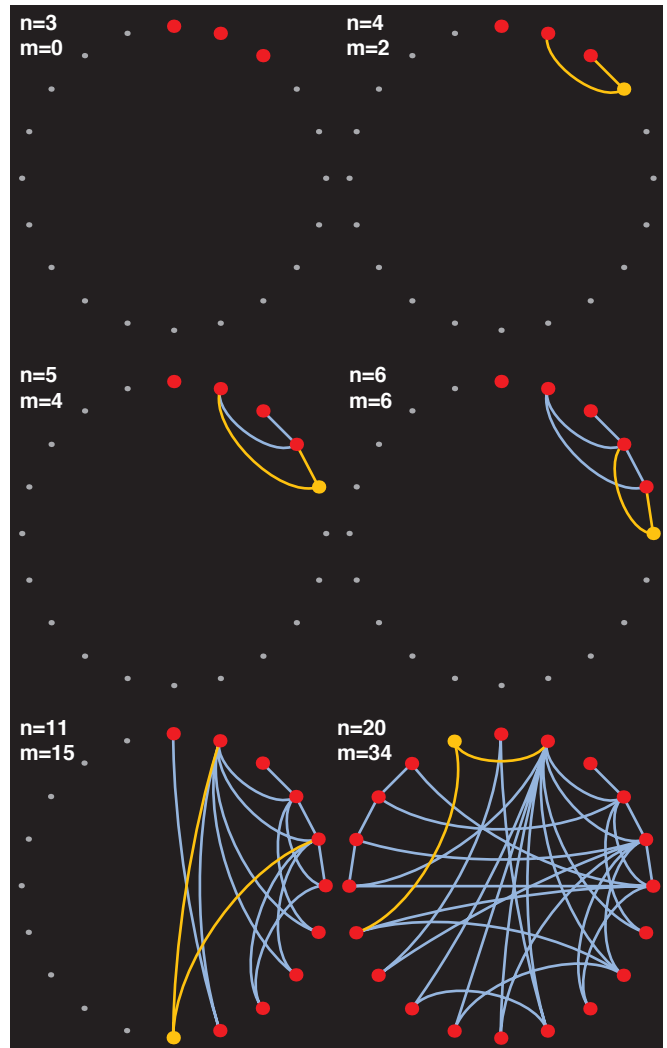

Figure 3. Model of Barabási, Albert and Jeong grows a graph by adding both vertices and edges. In the example shown here, each stage adds one new vertex and two new edges (yellow).

the degree sequence—the number of vertices with each possible number of edges from 0 to $n-1$.

A lattice has a simple degree sequence. All the vertices have the same number of edges, and so a plot of the degree sequence consists of a single sharp spike. Any randomness in the graph broadens this peak. In the limiting case of an Erdős-Rényi graph, the degree sequence has a Poisson distribution, which falls off exponentially away from the peak value. Because of this exponential decline, the probability of finding a vertex with k edges becomes negligibly small for large k.

There is evidence that real graphs such as the Web behave differently. The distribution of degrees is described by a power law rather than an exponential. That is, the number of vertices of degree k is given not by e^{-k} (an exponential) but by $k^{-\gamma}$ (a power law, where the power γ is a positive constant). The power-law distribution falls off more gradually than an exponential, allowing for vertices of very large degree.

Several research groups have independently discovered this property of the degree sequence.

A group that includes Kleinberg and several workers from the IBM Almaden Research Center found evidence of a power law in the Web, and so did Adamic and Bernardo A. Huberman of the Xerox Palo Alto Research Center. Michalis Faloutsos, Petros Faloutsos and Christos Faloutsos observed power-law relationships in the hardware substrate of the Internet.

Still another group that recognized a power law in the statistics of the Web consists of Albert-László Barabási of Notre Dame University and his colleagues Réka Albert and Hawoong Jeong. Barabási, Albert and Jeong set out to estimate the diameter of the Web (or, strictly speaking, the characteristic path length), and found that two randomly chosen pages are about 19 mouse-clicks apart. In the course of this analysis they tabulated the degree sequence, observing that the probability that a page has links to k other pages is approximately $k^{-2.45}$; the probability that k pages point to a given page is $k^{-2.1}$.

To explain the power-law degree sequence, Barabási, Albert and Jeong have proposed a new random graph model. They argue that other models fail to take into account two important attributes of the Web. In the first place, the Web is continually sprouting new pages, but most models are static: Although edges can be added or rearranged, the number of vertices never changes. Second, both the Erdős-Rényi and the Watts-Strogatz processes assume uniform probabilities when creating new edges, but this is not very realistic either. Barabási and his coworkers note that Web pages that already have many links are more likely to acquire still more links.

Like the Erdős-Rényi model, the Barabási model starts with n vertices and no edges, but the evolution is different. At every step the graph adds a single new vertex and m edges; all the new edges link the new vertex to some of the vertices already present. The probability that a given vertex will receive a new edge is proportional to the share of the total set of edges that the vertex already owns; hence the well-connected become still better-connected. After t steps, the graph has $n + t$ vertices and $m \times t$ edges. Growing according to these rules, the graph attains a statistical steady state: The shape of the distribution of node degrees does not change over time. The distribution is described by a power law with an exponent of 3; in other words, the probability of finding a vertex with k edges is proportional to k^{-3}.

Barabási and his colleagues have tested their model on several large graphs, including those discussed by Watts and Strogatz—the Hollywood graph, an electric power grid and the neural network of *C. elegans*. They find that both of the novel features of the model are essential to its success; eliminating either growth in the vertex set or preferential attachment of edges impairs the model's performance.

The straightforward way that the growth of a Barabási graph mimics the growth of the Web gives the model strong intuitive appeal, and yet the correspondence of theory and observation is not quite as close as one might hope. As noted above, the actual γ exponents for the Web are about 2.45 for outward links and 2.1 for inward links, significantly different from the model's prediction of 3. For some other graphs, such as the *C. elegans* network, the discrepancy is even greater. Various adjustments could tune the model for a better match, but they inevitably sacrifice some of its simplicity.

Alpha-Beta Graphs

The entire class of random graphs with a power-law degree sequence has been analyzed by Chung, William Aiello of AT&T Laboratories and Linyuan Lu of the University of Pennsylvania. Their model organizes all such graphs according to the values of two parameters, α and β.

When the degree sequence of a power-law graph is plotted on logarithmic scales, it forms a straight line. The parameters α and β specify this line; α is the point where the line intercepts the y axis, and β is the line's slope (or rather the negative of the slope). Thus α is the logarithm of the number of vertices of minimal degree, and β is the rate at which the logarithm of the number of vertices decreases as the degree increases. More formally, if $y(k)$ is the number of vertices of degree k, then the family of graphs defined by the model consists of all those graphs that satisfy the equation $y(k) = e^{\alpha}/k^{\beta}$, subject to the constraint that both k and $y(k)$ must be integers. For any fixed values of α and β, $y(k)$ specifies a finite set of graphs; a random alpha-beta graph is simply one chosen with uniform probability from this set.

The alpha-beta model is somewhat different from the models discussed above in that it does not directly specify the size of the graph in terms of n, the number of vertices. Another difference is that multiple edges between vertices are allowed, and so are "self-loops," or edges that start and end on the same vertex. (Both of these features are present in some natural graphs, such as the Web.)

Aiello, Chung and Lu explore how the properties of alpha-beta graphs vary as a function of β. A larger value of β corresponds to a steeper slope in the power law and hence to fewer nodes of large degree. If β is greater than a threshold value of approximately 3.4785, almost every alpha-beta graph consists of many small, disconnected components; there are not enough edges to hold the graph together. If β is greater than 1 but less than 3.4785, almost every graph has a giant component—a single large, connected piece that includes most of the vertices and edges in the graph. When β is less than 1, high-degree vertices are so abundant that almost every graph consists of one connected piece.

Aiello, Chung and Lu developed their model with a specific family of real-world graphs in mind, namely "call graphs" recording long-distance telephone traffic. As described in Part I, a

call graph has telephone numbers as its vertices, and an edge is drawn between two vertices whenever a call is placed between the corresponding numbers. Analysis of an actual call graph shows that the degree sequence is not quite a perfect power law, and yet the model captures important features of the graph. The best approximation has parameter values of roughly $\alpha = 17$ and $\beta = 2.1$. Since β lies between 1 and 3.4785, the model predicts that the graph is not connected but does have a giant component.

Graph Practice, in Theory

It may seem remarkable that the abstract and austere formalism of graph theory would prove useful in explaining such worldly phenomena as the architecture of the World Wide Web or the casting of Hollywood films or patterns of telephone calls. But graph theory is a branch of mathematics that has never been afraid to get its hands dirty with applications. Early on, it had close and fruitful encounters with organic chemistry and electrical engineering.

In another sense, though, the sudden blooming of graphs—the tendency to see them everywhere we look—is indeed new and noteworthy. For a long time, science has preferred to see the world through a grid of Cartesian coordinates, organizing everything around us according to row and column, rank and file, latitude and longitude. Although there have always been exceptions to this rectilinear prejudice—notably the treelike structures of taxonomists and grammarians—the simplicity of lattices has been hard to resist. And both lattices and trees have a convenient property that physicists call locality. The only way to move through a lattice or a tree is to crawl from one node to the next. There is no action at a distance; there are no secret wormholes connecting distant parts of the universe.

Graphs are a generalization of both lattices and trees. They admit more-flexible arrangements and less-regular connections in the way the world is put together. In graphs the principle of locality can be violated, as the shortcuts in the Watts-Strogatz model make explicit. Such subterranean channels make the world a harder place

to comprehend, but perhaps a more interesting place to inhabit.

Bibliography

Adamic, Lada A. 1999. The small world Web. http:// www.parc.xerox.com/istl/groups/iea/www/smallworld.html

Aiello, William, Fan Chung and Linyuan Lu. 1999. A random graph model for massive graphs. To appear in *Proceedings of the 32d Annual ACM Symposium on the Theory of Computing*. http://math.ucsd.edu/~fan/random.ps

Albert, Réka, Hawoong Jeong and Albert-László Barabási. 1999. Diameter of the World Wide Web. *Nature* 401:130–131. http://xxx.lanl.gov/abs/cond-mat/9907038

Barabási, Albert-László, and Réka Albert. 1999. Emergence of scaling in random networks. *Science* 286:509–512. http://www.nd.edu/~networks/science.pdf

Barabási, Albert-László, Réka Albert and Hawoong Jeong. 1999. Scale-free characteristics of random networks: The topology of the World Wide Web. http://www.nd.edu/~networks/proceeding.ps

Bollobás, Béla. 1985. *Random Graphs*. London: Academic Press.

Bollobás, B., and F. R. K. Chung. 1988. The diameter of a cycle plus a random matching. *SIAM Journal on Discrete Mathematics* 1:328–333.

Chung, F. R. K., and M. R. Garey. 1984. Diameter bounds for altered graphs. *Journal of Graph Theory* 8:511–534.

Erdős, P., and A. Rényi. 1960. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* 5:17–61.

Faloutsos, Michalis, Petros Faloutsos and Christos Faloutsos. 1999. On power-law relationships of the Internet topology. *Computer Communication Review* 29(4):251–262. http://www.acm.org/sigcomm/sigcomm99/papers/ session7-2.html

Huberman, Bernardo A., and Lada A. Adamic. 1999. Growth dynamics of the World-Wide Web. *Nature* 401:131. http://xxx.lanl.gov/abs/cond-mat/9901071

Kleinberg, Jon. 1999. The small-world phenomenon: an algorithmic perspective. Cornell Computer Science Technical Report 99-1776. http://www.cs.cornell.edu/home/kleinber/swn.ps

Kleinberg, Jon M., Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan and Andrew S. Tomkins. 1999. The Web as a graph: measurements, models, and methods. In *Proceedings of the Fifth International Conference on Computing and Combinatorics*, Tokyo July 26–28, 1999 (COCOON '99). Berlin: Springer-Verlag, pages 1–17.

Watts, Duncan J. 1999. *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton: Princeton University Press.

Watts, Duncan J., and Steven H. Strogatz. 1998. Collective dynamics of 'small-world' networks. *Nature* 393:440–442.