from *SIAM News*, Volume 36, Number 8, October 2003

The World According to Network Theory

Linked: The New Science of Networks. *By Albert-László Barabási, Perseus Publishing, Cambridge, Massachusetts, 2002, 280 pages, including copious notes, $26.00; paperback version, Plume (Penguin Group), New York, 2003, with an afterword, $14.00.*

BOOK REVIEW

By Philip J. Davis

We are instructed by the film *My Fair Lady* that Hungarians put on great shows. Albert-László Barabási, a native of Transylvania who is now a professor of physics at Notre Dame, has produced in *Linked* a swashbuckling popularization of applied network theory. His book swept me into thinking that for better and for worse, we are all simply nodes in a network of rapidly expanding networks. Aldous Huxley in *Brave New World* divided us into alphas, betas, and gammas; George Orwell in *Animal Farm* set us up as animals. Now we are treated as nodes, and some of us are more nodal than others.

Barabási's mise en scène is grandiose; his prose is enthusiastic and seductive, at once expository and given to oral history: a lot of the "this was my train of thought during my two-hour drive from Notre Dame to Chicago" sort of thing. His knowledge of the historical and technical sides of his field is profound, and before we are through reading, we've been linked to the principal research nodes in the link business.

Barabási starts with the old game of "getting to the President in four handshakes," now widely known as determining your "Erdős number," or "six degrees of separation between us and everybody else on the planet." The book then runs through the spread of AIDS (it was written before SARS), the failure of power networks on the west coast in the U.S. (the book also predates this summer's power failure in much of the northeastern and midwestern U.S. and Canada), the Internet, ecosystems, terrorist networks, the recent Florida elections, the CEO of a corporation as the root of a tree (if not occasionally the root of great evil), designer drugs, DNA, the spread of gossip about Clinton's misbehavior. Barabási winds up with the Internet as a vast, autonomous, self-perpetuating "skin" that has already wrapped around us as tightly as a tortilla wraps refried beans. He creates a world according to network theory, a world that embraces just about everything.

Shortly after immersing myself in this eminently readable work, I sat down with the *Providence Journal* to a breakfast of scrambled eggs. Hundreds of networks were involved in this simple act. What in life, I mused, cannot be placed under the inexorable, universal laws of network theory, under the ubiquitous and mysterious power of the power law distributions? "All is number," Pythagoras taught, and since a network can in principle (that tricky phrase) be represented and manipulated by a 0–1 incidence matrix, I would call Barabási a philosophical neo-Pythagorean.

I sat down with the Providence Journal *to a breakfast of scrambled eggs. Hundreds of networks were involved in this simple act. What in life, I mused, cannot be placed under the inexorable, universal laws of network theory, under the ubiquitous and mysterious power of the power law distributions?*

Well, over the years we have had many all-embracing mathematics-driven visions. In the pure field, math has been seen as nothing but logic or structures. In applied mathematics, Laplace's Demon reduced everything to an initial-value problem for a vast system of ordinary differential equations. Much more recently, cybernetic (feedback and control) visions, visions of catastrophe theory and of complexity, visions of theories of everything, have all competed for their Warholian moments. Barabási conjures up a Demon that—like the genome project—will describe for us the dynamic topology of the Internet. When this is in hand, we will then be able to pass from mere description to intelligent action. At least that's the hope. My own vision is that all is matrix theory, but I haven't gone public with it.

Enough five-cent philosophizing. Let's get to a few specifics.

Linked is, in a great measure, a vivid history of the search for network-theoretic models that realistically describe certain social and biological phenomena of the sort mentioned in the third paragraph of this review. Description is paramount, while prediction and prescription go begging. Adhering to Hawking's principle for popularizations (each math symbol reduces the potential readership by a factor $r = ½$), the book presents essentially no explicit mathematics, other than a few formulas in notes.

A network consists of a set of nodes joined by links. Pure mathematicians prefer to call such things graphs. While the amount of graph theory developed in pure mathematics is enormous, hardly any of it is visible in the present work. While most of it is very likely irrelevant, there's plenty to say without it. We meet random (Erdős–Rényi) networks, Watts–Strogatz networks (nodes are placed in a circle and only a few neighbors are linked), but most importantly, we meet power law networks, which are also called scale-free networks. These are networks in which the number of nodes $n(k)$ with k links is proportional to $k^{-\gamma}$ for some positive γ. The question, of course, is how you arrive at n. Should I consider myself linked to the telemarketer who tried yesterday to sell me vinyl siding? It depends.

Besides nodes, networks may have "hubs," i.e., nodes that are popular or get preferential treatment. O'Hare in Chicago is a hub, while the airport in Manchester, New Hampshire, is merely a node. Harvard University attracts more money because it has more money; it is a hub. Brown University is a node. Hubs are associated with the popular "universal 80/20 law," ascribed to Italian economist V. Pareto (1896), which says that 80% of whatsoever is achieved by 20% of whomsoever: money, acclaim, goals, you name it. Random (Erdős–Rényi) networks have no hubs (they are democratic). Power law networks have hubs (they are oligarchic).

Nodes have "fitness," i.e., the ability to succeed in a competitive environment. If you say that fitness satisfies a Bose–Einstein distribution ($f(E) = 1/(A \exp(kE) - 1)$), then, as in the "Bose–Einstein condensation" phenomenon in gases, the result can be a winner-take-all situation. (Such nodes are absolute monarchies: Coca-Cola forty years ago.)

Opining that "nature normally hates power laws," Barabási reports with some amazement and considerable glee the discovery that everywhere he looks he finds networks with a power law structure: AIDS networks, the friendships of Hollywood actors, the Web, the network of sexual relationships among humans, and many more such. I'm not sure why Barabási should be amazed that power laws are all over the place when this has been known generally since George Zipf's work of the 1930s. With respect to networks, it's been known for quite some time already.

Anyway, Barabási writes that power laws are

"nature's unmistakable sign that chaos [i.e., randomness] is departing in favor of order. . . . They are the patent signature of self-organization in complex systems."

The power law distribution is thus the statistical "star" of this book. Those of us who have been mentally exhausted by the controversies surrounding interpretations of bell curve distributions will welcome the opportunity to focus our critical souls on the power law. Thus, the Pareto 80/20 "law" has been much questioned. Apart from the remark that power law networks reside in the fragile territory between chaos and order, there is little discussion of the generic physical or social conditions that give rise to power laws. But this is a minor distraction.

Networks exhibit robustness and vulnerability. Robustness is the ability of a system to survive under a wide range of unfavorable but random conditions. Vulnerability is the susceptability of a system to break down under a smart attack. Power law systems, harboring as they do many hubs, are paradoxically both robust and vulnerable. "The scale-free topology at AIDS's disposal," Barabási writes, "allowed the virus to spread and persist."

Numerous conclusions are asserted about the various types of networks, and I wonder whether these conclusions were reached empirically with an assist from computer simulations or whether precise (but idealized) theorems lie behind them. The author is not always clear. Some of the conclusions seem to be a matter of common sense. For example, if you want to discombobulate a scale-free network, just shoot out its hubs. The Pentagon, in dealing with Iraq, seemed to know all about this. In the opposite direction of combobulation, most authors vent frequent fury at their publishers because the publishers don't advertise their books; rather, they sock the cash available for publicity into their blockbuster hubs.

I suspect that, in adherence to Hawking's principle, the author has suppressed considerable mathematics. I've learned, for example, that the mathematical strategy behind the Google search engine is to set up (in principle at least) a very sparse Markoff transition matrix A whose states are the Web pages and whose transition probabilities $A[i,j]$ have been computed according to a certain rule. We then compute the principal eigenvector of a certain linear combination of A and J (all 1's) to produce a key parameter known as "the page rank." This is quite a job since the number of Web pages is of the order of several billion and grows daily. The sparseness of the matrix helps.

Once a page rank (independent of individual inquiries) is in place, a second algorithm, mostly of a bookkeeping nature, and whose details are largely proprietary, goes to work on a specific inquiry and in zip time drowns the inquirer in information. The time-varying PageRank (Larry Page and Sergei Brin)* is under attack; it's not democratic. Pareto is at work here: The rich, the smart, the talented, the beautiful (who are always among us) get most-favored-node treatment.

I wish the various centers that study network problems (such as the one Barabási runs) good luck with their studies. One grandiose scheme now on their front burner is to determine the cell organizations that lead to the "principles behind life." This project has led to the recognition of the importance of clustered or modular, scale-free networks whose clusterings can be measured by a "clustering coefficient."[†]

But my interest as a client of networks is more modest. A good fraction of *Linked* is devoted to the Internet. I should personally like to know how misinformation spreads on the Internet and how serious it is. I understand that there is a proposal[‡] to provide (in principle) each Web page with an "index of reliability." This strikes me as an updated inversion of the notorious *Index Librorum Prohibitorum* that was alive and well from 1557 to 1966. Yells and screams will emerge from the low-indexed majority. As an unforeseen consequence, a system indexed in this way might even build itself into the liar paradox, and if that event results in litigation, it might even take a ruling from the U.S. Supreme Court to straighten matters out.

---

[*]Interested readers can check out *The PageRank Citation Ranking: Bringing Order to the Web* by Lawrence Page, Sergei Brin, Rajeev Motwani, and Terry Winograd at http://www.actonvision.com/GooglePageRank1.html.

[†]If a node has k links and all of these nodes have N links among themselves, the clustering coefficient C of the node is defined as $2N/(k(k-1))$.

[‡] For more information, see "Of Course It's True; I Saw It On The Internet," by Graham and Metaxas, in the May 2003 issue of *Communications of the ACM*. You can download it from the Web for a fee if you have an ACM public access account.

Philip J. Davis, professor emeritus of applied mathematics at Brown University, is an independent writer, scholar, and lecturer. He lives in Providence, Rhode Island, and can be reached at philip_davis@brown.edu.