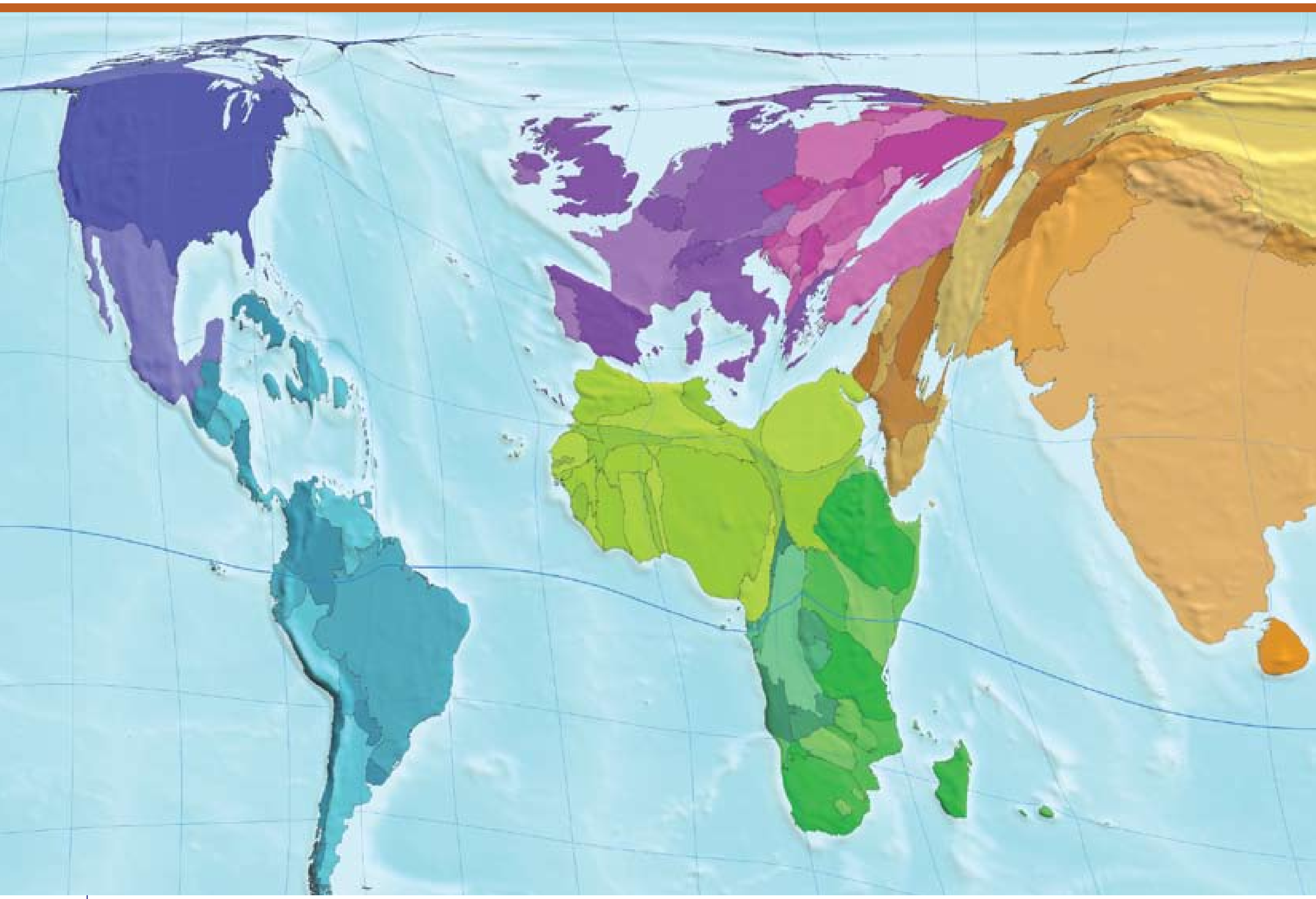


MARK NEWMAN: Exploring the Physics of Connection

BY JULIE J. REHMEYER

"All men are caught in an inescapable network of mutuality," Martin Luther King said in 1963. "Whatever affects one directly affects all indirectly. I can never be what I ought to be until you are what you ought to be…This is the interrelated structure of reality."

As profound as King's statement is, it's not exactly scientific. Mark Newman pondered this same interconnectedness in the late 1990s—with a physicist's eye. Yes, he thought, networks are big and important and have great moral implications, but how do they work?

Newman's thoughts were inspired in

THE ATLAS OF THE REAL WORLD. © 2008 DANIEL DORLING, MARK NEWMAN, ANNA BARFORD. THAMES & HUDSON, LONDON AND NEW YORK.

In this cartogram, Newman and his colleagues use population density to shrink or grow each country, illustrating world population distribution.

the classic Santa Fe Institute way: teatime conversations with someone in an entirely different field. Postdoctoral Fellow Duncan Watts told him about his thesis, where he showed that the neurons of the worm *C. elegans*, the power grid of the western U.S., and the collaborations of film actors all formed networks that were alike.

Their chats made Newman realize that Watts' network theory might be able to shed new light on the web of connections between human beings. Then he could go beyond the observation that people influence one another to work out *how* they do so. That understanding might allow him to predict things people need to know, like how fast the flu will spread, who's a terrorist, and how robust the Internet is.

Newman became one of the first physicists to apply network theory to social connections. And as his collaboration with Watts blossomed, so did network theory.

Newman was ready for just this kind of project. He had arrived at SFI as a refugee from traditional physics. Statistical physics, he'd come to believe, had become a victim of its own success. The important problems it could easily solve had been worked out long ago. Now, a theoretician like him had to either labor for decades to chip away at the big, fundamental problems, or settle for secondary issues. But Newman wanted to answer questions that mattered, and he didn't want to wait decades for the answers.

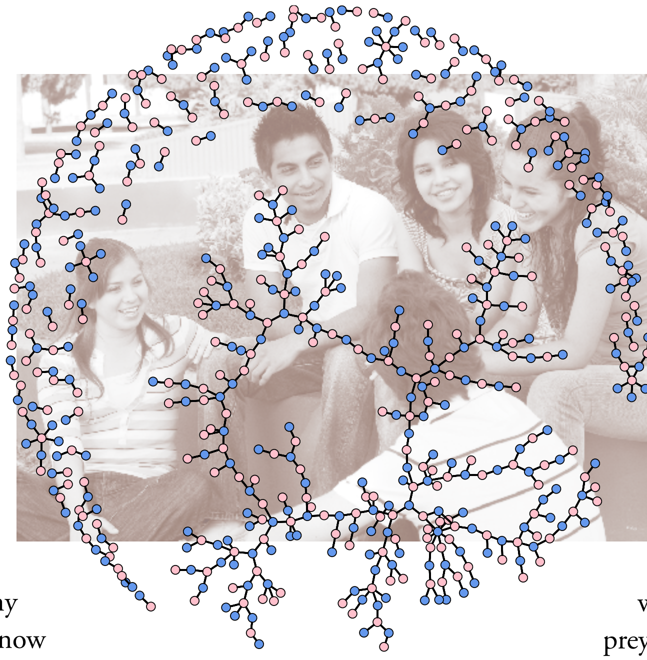
Network theory was littered with rich theoretical questions with important practical payoffs. For example, when a hospital in Evansville, Indiana, experienced an outbreak of pneumonia, the Center for Disease Control (CDC) turned to Newman and his collaborator Lauren Ancel Meyers. The CDC had tested all the hospital's patients and staff for the illness. It knew each patient's ward, roommates, and which doctor and nurse treated them. Could Meyers and Newman figure out how the illness had spread, and how to stop it?

The pair built a network in which each patient and staff member was a node, and those nodes were connected when two people were known to have had contact. Then they worked out how the bug must have spread. It was astonishing: The data showed that patients were nearly certain to pick up the bug any time they had contact with an infected staff member. An infected staff member, then, was enormously dangerous, spreading the disease from ward to ward throughout the hospital. And yet the staff almost never seemed sick. Even when they were carrying the bug, they kept the symptoms at bay. To control the outbreak, Newman and Meyers realized, the hospital had to treat the staff.

Network theory was the key to cracking the case. "There's a long history of mathematical work in epidemiology," Newman says. "People had considered how a disease affects an individual, and they had also considered how diseases spread using the connections between individuals. What had received rather little attention were

NORMAN JOHNSON

In some of his earlier work with networks, Newman mapped this network of dating patterns in a U.S. high school. The nodes are students, color-coded blue for boys and pink for girls, and the connections show who dated whom.



ISTOCKPHOTO.COM

the patterns of connections between individuals. If you don't know the patterns of connections, then it's hopeless to predict how quickly the disease will spread or how many people will get infected." He's now applying similar methods to understanding the spread of HIV, using data from random phone surveys of people's sex lives.

To solve such puzzles, Newman draws on years of work developing a theoretical understanding of how networks function. He started decoding the secrets of networks through a community that was close to home: scientists. Scientists leave a paper trail of their collaborations through joint publications, and Newman used that data to create a big network graph that encapsulated how scientists work together. And that picture looked like…well, a nasty, giant hairball.

To make sense of it, Newman had to build some mathematical tools. He started by asking how spread out the network of scientists is, picking two random scientists—say, a Finnish ecologist and an Australian high-energy physicist—and connecting them as directly as possible through their chain of collaborations. How many "degrees of separation" are there? He calculated that it was usually just five or six.

To probe deeper, Newman considered how we informally make sense of networks. The social networks in high schools, for example, are dominated by cliques: the jocks, the geeks, the Goths. So Newman developed a clique-detector, a tool to detect subcommunities that are tightly interconnected. When he tested it on collaborations among physicists, the communities he found corresponded to the discipline's traditional subfields, like astrophysics or particle physics. This suggests that traditional subfields are a good reflection of what people actually do.

Divisions like astrophysics and particle physics are still pretty crude. Among astrophysicists, for example, there are observational astrophysicists, and among those there are radio astronomers, and so on. So recently, Newman has worked with SFI Postdoctoral Fellow Aaron Clauset and SFI Professor Cris Moore to create a tool that automatically detects whole hierarchies of communities.

In addition, their tool does something extraordinary: predicts missing links. It computes statistics about the hierarchical structure of the subcommunities, and then it generates thousands of other simulated networks with different links but the same structure. If a particular link is common in these "sister" networks, the researchers figure it's likely to be in the original network too. This can be invaluable when studying, say, food webs, where the links show predator-prey relationships and validating each link can take months of fieldwork. The tool successfully predicted missing links in three real-world networks: a food web, a terrorist network, and metabolic interactions.

"Mark is singular in combining theoretical strength with fearlessness about wrestling with real-world data," Moore says. "In networks, there have been theoretical models that aren't very realistic, and there are also people with huge datasets they don't know what to do with. No one has done more to build a bridge between them than Mark. He's helped make the field of networks both theoretically deep and grounded in real data."

Newman's focus on data, Moore says, has been influential at SFI even outside of network theory. In the early days at SFI, researchers were fascinated by simple models that generated complex behavior that looked qualitatively similar to the real world, but they struggled to figure out how to make those connections quantitative. As Newman's postdoctoral fellowship at SFI extended into a research faculty position and then regular visits, his focus on testing his tools on data has led the way in a broad shift at the Institute toward a data-driven approach, even in fields far removed from network theory.

Newman has also formed a bridge between sociologists, physicists, and computer scientists in network theory. In the last 10 years, physicists and computer scientists have developed revolutionary techniques, but most don't know the rich ideas the sociologists have come up with over decades. Newman takes sociological ideas and makes them quantitative. "Mark is like an archaeologist," says Albert-László Barabási, a fellow network scientist at the University of Notre Dame. "He finds these pieces of gold spread around the communities." Now Newman is writing a textbook on network theory that draws together sociology, physics, and computer science.

"Mark is singular in combining theoretical strength with fearlessness about wrestling with real-world data," Moore says.

His work goes beyond network theory. A friend once sent him a postcard showing "A Texan's View of the United States." Texas dominated the middle, and all the other states were squished around the edges. California was marked "uninhabitable," and New England, "Damn Yankee Land." Newman chuckled over it and set it aside, but it got him thinking about people's internal maps of the world. For example, people often visualize Michigan as being in the middle of the U.S., even though it's far closer to the East Coast. Was there some rigorous way, he wondered, to depict these distorted maps?

He filed the question away for a few years, until he and SFI Postdoctoral Fellow Michael Gastner were collaborating on a project to map out the physical location of computers on the Internet. Computers tended to be where the people were, so their map was essentially a map of population. But then they wondered, did some areas have more computers per person than others?

One way to tell would be to squish the areas of the map with few people and expand the areas with more people, creating a map with uniform population density. If computers were plotted on that map, variations would show up.

To do this, they imagined that each person in the U.S. was a molecule of gas. If you release a gas in a room, it will spread out according to well-understood statistical rules. Newman and Gastner applied those rules to people, allowing them to warp the boundaries of their states with them as they moved. In the resulting maps, the crowded west coast and northeast were swollen and the sparsely populated Rocky Mountain and high plains states shriveled.

And smack dab in the middle of the map was Michigan, just as people tend to imagine. Newman realized he was looking at the mental map he'd been envisioning for years!

Newman and Gastner applied their technique to create a "cartogram" of the 2004 national election data. Typical maps of the election results make it seem as if George W. Bush won by a landslide, with a great mass of red, Republican-voting states in the middle of the country and a much smaller area of blue, Democratic-voting states on the west coast and in the northeast. In Newman and Gastner's cartogram, though, the states form a nearly even balance between red and blue.

Recently, Newman teamed up with a group at the University of Sheffield to apply his method to all kinds of international data. In a cartogram of people affected by natural disasters, China, India, and Southeast Asia mushroom, while the U.S. shrinks to near-invisibility. On the other hand, the U.S. is bloated in a cartogram of extinct species, sharing the stage with the tiny island of Mauritius. Newman and his collaborators published these cartograms last October in a book titled *The Atlas of the Real World.*

Colleagues describe Newman as a giant in the field, and they talk about the enormous impact his work has had. Newman himself is just grateful. "It's rewarding to be working in a field where people actually care about what you're doing." ◀

Julie J. Rehmeyer was SFI's very first undergraduate intern. She went on to do graduate work in mathematics at MIT and to teach at St. John's College. She is now a freelance writer and the mathematics columnist for Science News.