

NewScientist



13 APRIL 2002 No 2338 WEEKLY £2.30 US\$3.95

BORN TO BE EATEN

Early humans weren't hunters, they were lunch

HOW CAN A SINGLE LAW GOVERN OUR SEX LIVES, THE PROTEINS IN OUR BODIES, MOVIE STARS AND SUPERCOOL ATOMS? NATURE IS TELLING US SOMETHING...



ALL THE WORLD'S A NET

What do the proteins in your body, the Internet, a cool collection of atoms and sexual networks have in common? One man thinks he has the answer, and it's going to transform the way we view the world. David Cohen reports

ALL researchers dream of making a discovery that will transform their field. Albert-László Barabási can go one better. In just three years, his discovery has started making waves in fields as diverse as ecology, molecular biology, computer science and quantum physics.

It all began when he found that sites on the Web form a network with unique mathematical properties. In itself, this may not seem very profound, but it soon emerged that these properties were not unique to the Web. We are surrounded by networks: social, sexual and professional. Ecosystems are networks, and even our bodies—and the pathogens that lay us low—are kept alive by networks of chemicals. Barabási and others have found that many of these networks have the same architecture as the Web. They grow in much the same way and have the same strengths and weaknesses: understand one and you start to understand them all. Universal mathematical laws are rare in biology but, without meaning to, Barabási seems to have uncovered one.

Born in Romania and educated in Hungary, Barabási is now a professor of physics at the University of Notre Dame in Indiana. Until a few years ago, he was preoccupied with arcane fields such as the fractal nature of surfaces and the dynamics of granular materials such as sand. To understand all these fields needs a heavy dose of statistics, which is Barabási's forte. He also had a long interest in complex networks, but information about them was sparse. By 1998, however, the tools for interrogating the Web had reached a level of sophistication that enabled Barabási to go exploring.



'A SCALE-FREE DISTRIBUTION POINTS TO THE FACT THAT THE WEB IS DOMINATED BY A FEW, HIGHLY CONNECTED SITES'

For theoreticians, the established way to model complex networks is with a random network. Make some dots on a page and start drawing lines between them at random. You end up with a network in which, on average, all the dots—or “nodes”—have the same number of links. Now count the number of nodes with one link, two links and so on, and plot these numbers on a graph. You end up with a well-known distribution—a bell curve (see Graph below).

This is what Barabási expected to find when he and his colleagues Réka Albert and Hawoong Jeong started studying the Web. They sent a software robot crawling around the Web to analyse the links between websites. But when they looked at the architecture of the Web and plotted the distribution of sites and the numbers of links to them, something strange happened. “It became

clear we were looking at a more complex situation than that described by random networks,” Barabási says.

There was no bell curve. Instead, the Web had lots of sites with a few links, a few sites with a medium number of links and a very few sites with loads of connections. This produced an ever-decreasing curve characteristic of what physicists call a power law (*New Scientist*, 8 November 1997, p 30). Gone was the average number of links—or scale—of the bell curve. Instead, announced Barabási, the Web was a “scale-free” network.

“This distribution,” says Barabási, “points to the fact that the Web’s structure is dominated by a few, highly connected sites.” He calls these sites “hubs”—classic examples are Yahoo and Napster—which have developed because they offer short cuts to the information we want.

A curious property of this architecture is that it takes only a few clicks to get from one site to any other on the Web. “On average, the journey from one Web page to any other can be made in just 19 clicks,” he says. This shows that the Web is a type of “small world”, a concept made popular by John Guare in his play *Six Degrees of Separation*.

In turn, Guare based his work on the idea that a message between any two individuals on the planet would only need to pass through an average of six intermediaries (*New Scientist*, 4 December 1999, p 24). This small-world property is essential to future growth because it means that as more sites come online, the Web will stay easy to navigate. Even if it grows by 1000 per cent, Barabási calculates that websites would still be separated by an average of only 21 clicks.

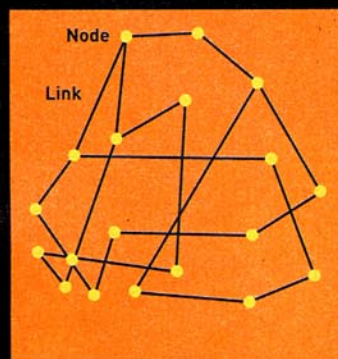
At first, Barabási thought his scale-free structure was unique to the Web. But he soon discovered the same pattern in other networks, such as the Kevin Bacon game (www.cs.virginia.edu/oracle). Picture all the world’s actors as nodes with links between them when they’ve appeared together in a movie. The aim is to link an actor to Bacon through the smallest number of other actors. Barabási found that the actors’ network is dominated by a few, usually famous actors, such as Bacon, who appear as hubs because they’ve made so many films.

Since then, numerous other networks have been added to the list of the scale-free, not least the network of computers that underlies the Web itself—the Internet. In biology, the grids of interacting proteins and chemicals that keep cells in good working order are scale-free. Food webs—the networks of who eats whom in various ecosystems—are built around “hub species” that eat large numbers of different prey species (*New Scientist*, 18 August 2001, p 30). And in human society, the network of scientists who’ve worked together is scale-free, as is the way they cite each other’s research. Even the web of human sexual contacts turns out to be scale-free.

So Barabási’s work has begun to expose a pattern of organisation that crops up time and again in natural and artificial worlds. Somehow, the collective actions of individual agents—be they websites or proteins—generate networks that conform to a single, well-defined mathematical formula. And every agent in all these systems seems to share the same behaviours.

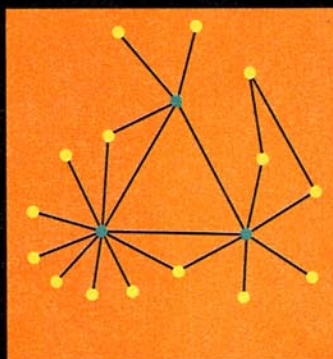
It didn’t take Barabási and his team long

Random network



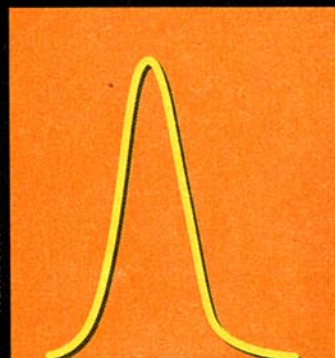
All nodes have about the same number of links. The average number (at the peak of the bell curve) gives the network a “scale”

Scale-free network



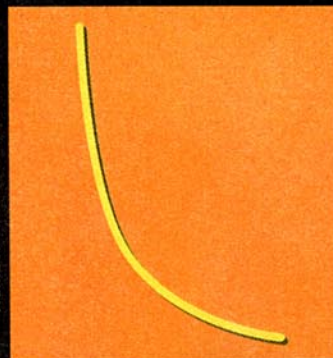
It makes no real sense to talk about scale or an average number of links. Scale-free networks have many nodes with few links and a few nodes with many links

Number of nodes with k links



Number of links (k)

Number of nodes with k links



Number of links (k)

to pin down these shared features. They found two vital ingredients. First, a scale-free network must be growing—so the Web needs new pages to be added every day, and the actors' network needs a constant supply of raw talent. Second, these new recruits must show some form of preference as they attach to the network. So, for example, new websites want to be picked up by popular sites, such as Yahoo, to increase their traffic. And ambitious actors want to appear in films with established stars, rather than unknown B-movie actors. In general, then, the highly connected tend to become even more connected or, if you like, "the rich get richer".

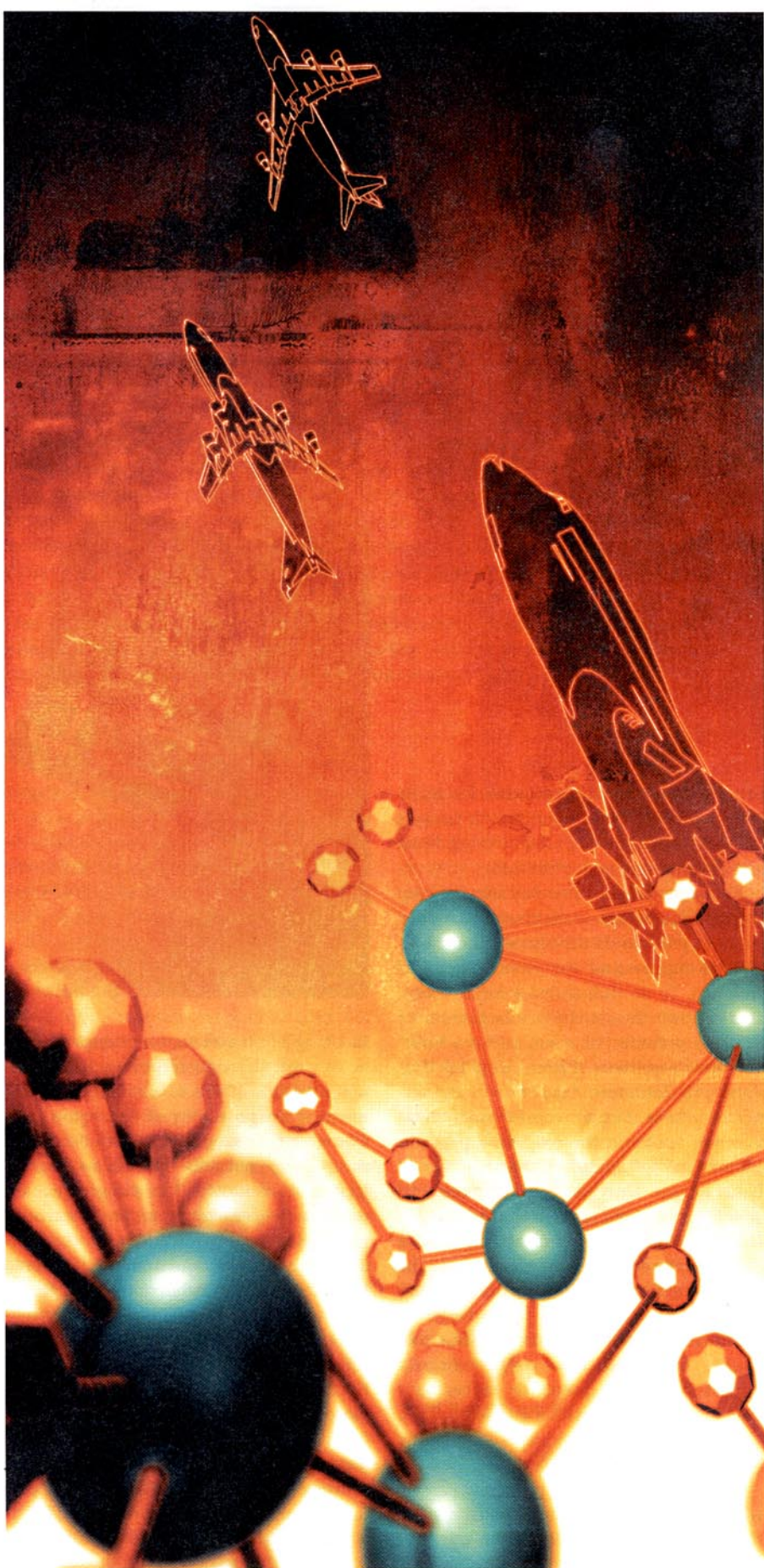
For some scale-free networks, the preferences at work are not clear. It's absurd, for example, to think that prey species choose to be eaten by a predator with a particularly varied diet. Nonetheless, solving this puzzle will undoubtedly improve our understanding of how ecosystems evolve. With proteins, one candidate for this "preference" mechanism is gene duplication—a rare occurrence during cell division when genes are copied twice. Every time this happens, all the proteins that interact with the duplicated protein gain another link.

Robust yet vulnerable

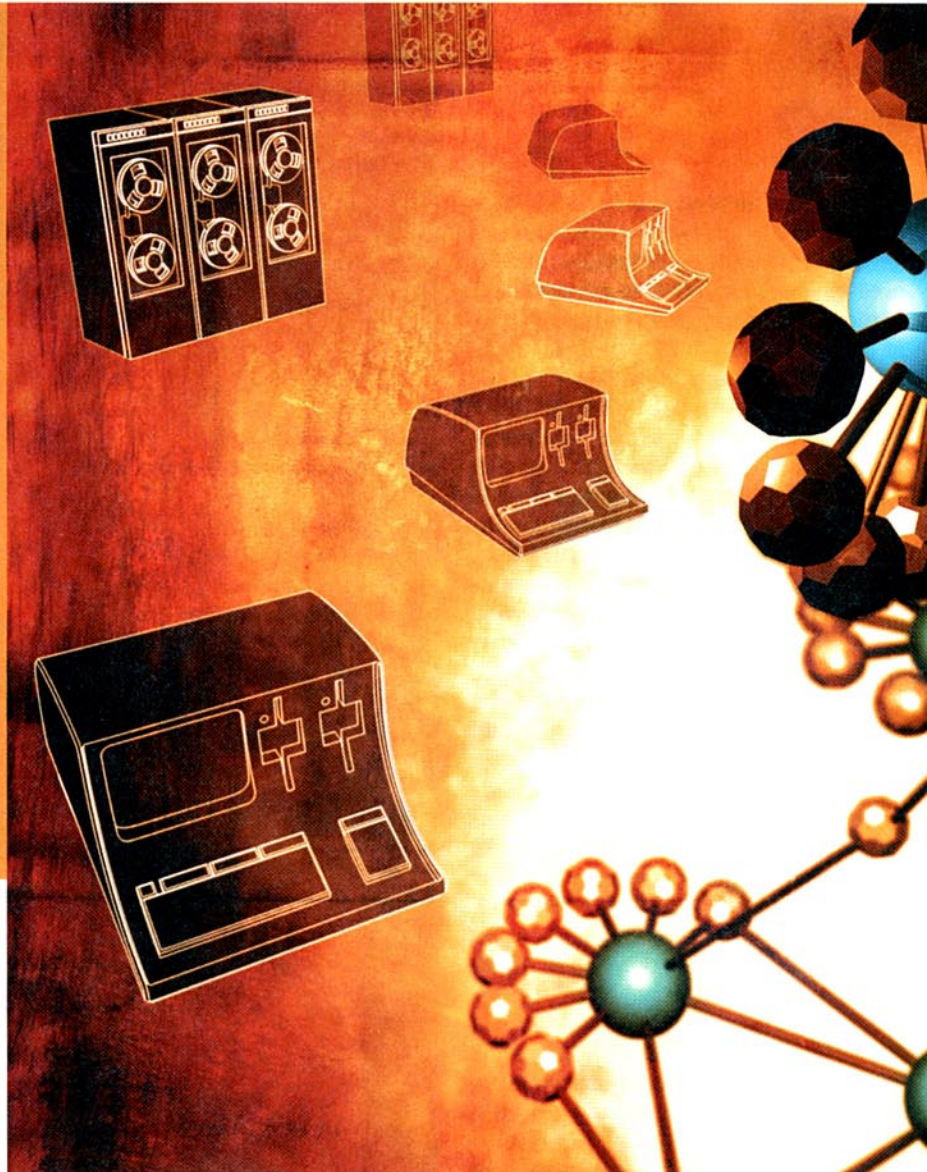
If the discovery that scale-free networks are everywhere is presenting us with new answers and questions about the world, so too are their properties. These networks are robust and vulnerable at the same time. Barabási, Albert and Jeong subjected a scale-free network to two types of attack. In one, they hit individual nodes at random, while in the other they only took out the hubs—the highly connected nodes in the network.

Random networks are highly susceptible to indiscriminate attacks. As more and more nodes are destroyed, the number of steps needed to get from one node to another increases steadily. By contrast, scale-free networks are robust in the face of such attacks. Even with 5 per cent of the nodes obliterated, the performance of the network is unaffected.

With highly targeted attacks, random networks decay in the same way as with indiscriminate attacks, but scale-free networks fare much worse. Once 5 per cent of the hubs have been removed, the number of steps needed to cross the network doubles. "This shows that scale-free networks, in general, are highly vulnerable to intelligent



'SCALE-FREE NETWORKS MAY HAVE SUCCEEDED NOT ONLY BECAUSE THEY ARE ROBUST, BUT ALSO BECAUSE THEY ALLOW VARIATION TO TAKE PLACE'



attack," says Barabási. It exposes the Internet's Achilles's heel—the hubs. "If hackers wanted to, they could probably bring down the Internet very easily," he adds.

The same vulnerability may also show up in protein networks—with disastrous results. In 1979, *p53* was the first gene to be identified as suppressing the development of tumours. To do this, it codes for a protein that controls the activity of a large number of other proteins. "It seems *p53* is a hub," says Bert Vogelstein of Johns Hopkins University in Baltimore. "It is one of the few genes whose failure causes such catastrophic results in the cell." In a paper in *Nature* (vol 408, p 307), Vogelstein, David Lane of the University of Dundee and Arnold Levine from Rockefeller University in New York likened the failure of *p53* in a cell—and the development of cancer—to the collapse of a hub on the Web and the subsequent crash.

Viewing protein networks as scale-free could help develop more realistic approaches to treatment for cancer, says Vogelstein. But he stresses it's still early days. "We're far away from understanding all the biochemical interactions in a cell," he says.

For any disease, seeing proteins as actors within a larger play will help drug designers to aim their chemicals in such a way that they don't disrupt the whole performance.

Blocking a hub protein, for example, would be very risky because of the large number of potential side effects it could cause, not to mention the possibility of destroying cells. Conversely, there may be times when you want to wipe out cells. Barabási and his colleagues have shown that the protein network in *Helicobacter pylori*, the bacterium thought to cause peptic ulcers, is scale-free. Knocking out hub proteins in this bug could be a good way to disable or even kill it.

At the evolutionary level, scale-free networks may have succeeded not only because they are robust in the face of random errors, but also because they allow variation to take place. Proteins with only a few connections could mutate or be lost entirely without damaging the health of the organism. Some of these mutations could give it an advantage, allowing it to outcompete its rivals.

Perhaps the most surprising property of scale-free networks emerged last year and is

changing our understanding of the way diseases spread among humans. Once again the story begins with the Net, when Alessandro Vespignani of the International Centre for Theoretical Physics in Trieste and Romualdo Pastor-Satorras of the Technical University of Catalonia in Barcelona decided to look at how computer viruses spread across the Net.

According to epidemiologists, a virus must reach a certain level of virulence for an outbreak to occur. Below this "epidemic threshold", the virus is not infectious enough to spread quickly and dies out. The higher above the threshold it is, the faster it will spread. But when Vespignani charted the movement of his software virus across the Internet, he got a shock. "There is no such threshold for an outbreak to occur," he says. "The hubs propagate viruses so efficiently that even a weak virus will spread rapidly through these nodes." This discovery has profound implications not only for the Net,

Imperfect world

REAL-WORLD problems have a habit of tripping up mathematical models, and scale-free networks are no exception.

Take the network of airports and flight routes. Newcomers to the airline business will want to connect to a hub—an airport with a huge number of connecting flights. But what happens when a hub is so busy it becomes saturated, as it has in the US? The newcomer must then choose a less well-connected airport. “Many real-world situations that at first seem to be scale-free networks in fact turn out not to be,” says Luís Amaral of Boston University. This is because constraints on their behaviour restrict how they can evolve (*Physical Review Letters*, vol 88, p 138,701).

There are other imperfections. When a new site connects to the Web, its owners must convince the best “hub” sites to link to it. But because they cannot see the whole Web they may not choose the most popular hubs. Or an actor may not crop up in the Kevin Bacon game because of a personal preference for low-budget films rather than blockbusters. These kinds of events inevitably upset the perfect “preference mechanism” that is assumed for creating the ideal scale-free network. “The ramifications of this could change our understanding of scale-free networks,” Amaral says.

Barabási says these issues do not detract from the importance of his ideas. “Of course you can destroy the scale-free state if you impose systematic limitations on the nodes,” he says. “Each

system has to be treated separately. It is not anybody’s goal to show that all networks are scale-free.”

One limitation that Barabási has already addressed is competition between nodes. If the accumulation of links over time is the only important factor as networks grow, then the oldest nodes would always have the most links. But we know from the Web that some latecomers—such as Yahoo—have become much bigger hubs than some older sites. In biological parlance, they appear to be “fitter” than other sites at making new links. So Barabási added a variable to his formulae to represent that fitness.

At this point, Barabási’s colleague Ginestra Bianconi made an extraordinary leap. She imagined nodes as energy levels, the links between them as particles in a gas moving between energy levels. Fitness determined the absolute energy of each node—the fitter a node, the lower its energy.

When she allowed the model to evolve, most of the particles connected to the fittest node—they fell into the lowest energy level (*Physical Review Letters*, vol 86, p 5632). This, to a physicist, is precisely what happens during the formation of a Bose-Einstein condensate, a bizarre quantum state of matter that forms when atoms are cooled to within a whisker of absolute zero. It’s ironic that correcting the scale-free model to make it agree with the everyday should reveal the rarest of quantum phenomena.

but also for human disease. “It is a breakthrough in the understanding of a certain class of epidemics,” Vespignani says (*Physical Review Letters*, vol 86, p 3200).

This discovery gave Fredrik Liljeros, a sociologist at the University of Stockholm, the impetus to look at how human diseases spread. Studies by Vespignani and Barabási had concluded that the way HIV spreads through populations is similar to the spread of viruses on the Net. So Liljeros chose to look at sexually transmitted diseases.

Sexual networking

He and his colleagues studied the sexual habits of 2900 Swedes. It came as no surprise that a few “hubs” had lots more sexual partners than the rest. But Liljeros also recognised this distribution of partners as a mark of a scale-free network. “Maybe people become more attractive the more partners they get,” he says. If so, it looks strangely like the preference mechanism needed to create a scale-free network.

Normally with a new vaccine, public health officials aim for blanket immunisation of at-risk people, setting a target for the percentage to be immunised. That percentage depends on the epidemic threshold of the disease-causing microbe. Liljeros’s findings suggest that this approach could have little or no impact. “In diseases such as AIDS,

targeting the most promiscuous individuals is the crucial factor,” says Liljeros. “We can attempt to stop the spread of a virus by blindly vaccinating huge groups, but without treating these key individuals we may never bring it under control.”

It’s common sense that a programme of vaccination against sexually transmitted diseases should try to reach the most promiscuous individuals first. But the idea that health campaigns may be utterly worthless if they miss these people is a shock. For Vespignani, this mixture of the obvious and the unexpected shows the real value of the scale-free revolution. It gives a mathematical form to common-sense concepts—which means theories can be tested and results understood. And it leads to important, non-obvious results. “Nobody would have thought that there is no immunisation threshold in such networks,” he says.

We can expect more surprises like this in future. How significant those surprises will be is hard to say: the concept of scale-free networks is only three years old, after all. Yet it’s difficult to conceive that a theory which predicts the behaviours of both a collection of inanimate chemicals and a group of thinking humans is not telling us something profound about nature.

The idea has already been jumped upon by AIDS researchers, computer network

designers and ecologists. For Barabási it’s this pervasiveness that gives scale-free networks their significance. “It is not that they are creating a revolution in any single field,” he says. “Rather, they prompt us to use the same tools, methods and approach to study very disparate systems. They allow us to see in a new and very similar perspective all the nodes and links around us.”

If there is one way that scale-free networks are destined to make their mark, it could be in helping us to understand emergence: the idea that many interacting agents following simple rules can collectively produce complex behaviours. So thousands of ants can produce a thriving colony that behaves like a single organism.

The challenge facing scientists now is to work out how the rules governing individual agents relate to large-scale behaviours. Scale-free networks give us the beginnings of a mathematical way to study that relationship. They are unlikely to be the whole answer, but they are at least a start. □

David Cohen is a writer based in London

Further reading: *Linked* by Albert-László Barabási is published in the US this month by Perseus. *Nexus* by Mark Buchanan is published in the US in May by W. W. Norton. In Britain, it’s called *Small Worlds* and is published in June by Weidenfeld & Nicolson