Eur. Phys. J. B **38**, 147–162 (2004)
DOI: 10.1140/epjb/e2004-00110-5

THE EUROPEAN PHYSICAL JOURNAL B

Complex networks

Augmenting the framework for the study of complex systems

L.A.N. Amaral[a] and J.M. Ottino

Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL 60208, USA

Received 12 November 2003
Published online 14 May 2004 –

Abstract. We briefly describe the toolkit used for studying complex systems: nonlinear dynamics, statistical physics, and network theory. We place particular emphasis on network theory—the topic of this special issue—and its importance in augmenting the framework for the quantitative study of complex systems. In order to illustrate the main issues, we briefly review several areas where network theory has led to significant developments in our understanding of complex systems. Specifically, we discuss changes, arising from network theory, in our understanding of (i) the Internet and other communication networks, (ii) the structure of natural ecosystems, (iii) the spread of diseases and information, (iv) the structure of cellular signalling networks, and (v) infrastructure robustness. Finally, we discuss how complexity requires both new tools *and* an augmentation of the conceptual framework—including an expanded definition of what is meant by a "quantitative prediction."

PACS. 89.75.Fb Structures and organization in complex systems – 89.75.Da Systems obeying scaling laws

1 Introduction

What do metabolic pathways and ecosystems, the Internet, and propagation of HIV infection have in common? Until a few years ago, the answer would have been very little. The first two examples are biological and shaped by evolution, the third is a human creation, and the fourth is an unwieldy mixture of biology and sociological components. However, in the last few years the answer that has emerged is that they all share similar network architectures. Seemingly out of nowhere, in the span of a few years, network theory has become one of the most visible pieces of the body of knowledge that can be applied to the description, analysis, and understanding of complex systems. New applications are developed at an ever-increasing rate and the promise for future growth is high: Network theory is now an essential ingredient in the study of complex systems.

However, before delving into networks themselves it is important to put the overall subject in context, attempt a definition of complexity itself, and present a brief review of the other tools that are used in the analysis of complex systems.

The discussion has to start with two important distinctions: First a differentiation between what is complex and what is merely complicated. Second a differentiation between the complexity of the dynamics generated by simple systems and that of complex systems [1].

Simple systems have a small number of components which act according to well understood laws—Consider what is perhaps the prototypical simple system; the pendulum. The number of parts is small, in fact, one. The system can be described in terms of well-known laws—Newton's equations. The example of the pendulum raises an important point: The need to distinguish between complex systems and complex dynamics: It takes little for a simple system such as the pendulum to generate "complex" dynamics. A forced pendulum—with gravity being a periodic function of time—is chaotic. In fact one can argue that the driven pendulum contains everything that one needs to know about chaos; the entire dynamical systems textbook by Baker and Gollub [2] is built around this theme. And a pendulum hanging from another pendulum—a double pendulum—is also chaotic (Fig. 1a).

Complicated systems have a large number of components which have well-defined roles and are governed by well-understood rules—A Boeing 747-400 has, excluding fasteners, 3×10^6 parts (Fig. 1b). In complicated systems, such as the Boeing, parts have to work in unison to accomplish a function. One key defect (in one of the many critical parts) brings the entire system to a halt. This is why redundancy is built into the design when system failure is not an option. More importantly, complicated

[a] e-mail: `amaral@northwestern.edu`

Fig. 1. Simple, complicated and complex systems. (a) The double pendulum—a pendulum hanging from another pendulum—is an example of a simple system. All parts can be well characterized and the equations describing their motion are also well known. (b) The Boeing 747-400 has on excess of 3×10^6 parts. (c) A flock of migrating geese. The Boeing 747-400 is not a complex system because all its parts have strictly defined roles and prescribed interactions. This is typical of complicated systems, in which robustness is achieved through redundancy, i.e., including several copies of the same part in parallel. In contrast, for complex systems, robustness is achieved by enabling the parts to adapt and adopt different roles. The migrating geese provide a good example of such strategy, the ubiquitous "V" formations of the migrating geese are not static structures with a leader at the head, instead the structures are fluid with a number of birds occupying the head position at different times.

systems have a limited range of responses to environmental changes. Even the most advanced mechanical chronometers can only adjust to a small range of changes in temperature, pressure and humidity before they loose accuracy. And a Boeing without its crew is not able to do much of anything to adjust to something extraordinary.

Complex systems typically have a large number of components which may act according to rules that may change over time and that may not be well understood; the connectivity of the components may be quite plastic and roles may be fluid—Contrast the Boeing 747-400 with a flock of migrating geese (Figs. 1b–c). Superficially, the geese are all similar and the flock has far fewer members than the Boeing has parts, so one might be tempted to think that the Boeing is more complex than the flock of geese. However, the flock of migrating geese is an *adaptable* system, which the Boeing is not. The flock responds to changes in the environment—that is indeed why it migrates—moreover, and unlike what one may naively guess, the migrating geese self-organize without the need for a leader or maestro to tell the rest of the flock what to do. This is clearly revealed by observing the dynamic unrepeated patterns generated by the geese as they adjust their flying formations. Roles in the flock are fluid and one goose at the head of the formation will quickly be replaced by another. This feature of the flock gives it a great deal of robustness as no single goose is essential for the flock's success during the migration.

The stock market, a termite colony, cities, or the human brain, are also complex. As for the flock of geese, the number of parts is not the critical issue. The key characteristic is adaptability—the systems respond to external conditions.

Complex systems: Self-organization and emergence—It is far from trivial to come up with an all-encompassing definition of complex systems. Nevertheless let us attempt one: A complex system is a system with a large number of elements, building blocks or agents, capable of interacting with each other and with their environment. The interaction between elements may occur only with immediate neighbors or with distant ones; the agents can be all identical or different; they may move in space or occupy fixed positions, and can be in one of two states or of multiple states. The common characteristic of all complex systems is that they display organization without any *external* organizing principle being applied. The whole is much more that the sum of its parts.

Examples of complex systems are among some of the most elusive and fascinating questions investigated by scientists nowadays: how consciousness arises out of the interactions of the neurons in the brain and between the brain and its environment, how humans create and learn societal rules, or how DNA orchestrates processes in our cells.

Organization of the manuscript—In Section 2 we describe the challenges faced when studying complex systems and describe how scientists from many different areas have responded to these challenges. We then, in Section 3, describe the toolkit used for studying complex systems: nonlinear dynamics, statistical physics, and network theory.

Nonlinear interactions, one of the greatest challenges in the study of complex systems, are at the core of the emergence of qualitatively different states, new states that are not mere combinations of the states of the individual units comprising the system. Indeed, the role of nonlinear dynamics in the understanding of complex systems has been important for more than two decades. Statistical physics provides the study of complex systems with, on one hand, techniques particularly suited for study of systems with a large number of units and, on the other hand, with two fundamental concepts for the quantitative characterization of complex systems—scaling and universality.

Notwithstanding the importance of nonlinear dynamics and statistical physics, we place particular emphasis on network theory due to it being the central topic of this special issue and to the explosive rate of advance that the field has experienced in the last five years. In Section 4 we review some of the most significant advances in our understanding of the mechanisms responsible for the emergence of real-world networks.

In Section 5, we briefly review several cases where research on networks has lead to new insights into the emergence, organization and behavior of complex systems. Specifically, we summarize a selection of results on four topics: (i) the Internet and other communication networks, (ii) the structure of natural ecosystems, (iii) the spread of diseases and information, (iv) the structure of cellular signalling networks, and (v) infrasrtucture robustness.

Finally, in Section 6, we discuss the need for a "new", or at very least, expanded definition of the meaning of prediction in the context of the study of complex systems. We end with a short discussion on the challenges ahead for researchers from different areas in contributing to the creation of a theory of complex systems.

2 Challenges in the study of complex systems

The challenges one faces when studying a complex system occur at various levels:

The nature of the units—Complex systems typically comprise a large number of units, however, unlike the situation in many physical and chemical problems, the units need not to be neither structureless nor identical.

Challenges at the unit level:
- Units have complex internal structures;
- Units are not identical;
- Units do not have strictly defined roles.

The nature of the interactions— Complex systems typically have units that interact strongly, often in a nonlinear fashion. Moreover, there are frequently stochastic components to the interaction and external noise acting on the system. An additional and crucial challenge is posed by the fact that the units are connected in a complex web of interactions that may be mostly unknown.

Challenges at the interaction level:
- Nonlinear interactions;
- Noise;
- Complex network of interactions.

The nature of the forcing or energy input—Complex systems are typically out-of-equilibrium. For example, living organisms are in a constant struggle with their environment to remain in a particular out-of-equilibrium state, namely alive. Social and economic systems are also driven out-of-equilibrium systems; a new technology changes the balance of power between companies, a terrorist attack changes economic expectations, etc.

Challenges at the forcing level:
- Poorly characterized distribution of external perturbations;
- Poorly characterized temporal and spatial correlations of external perturbations;
- Non-stationarity of external perturbations.

3 Tools for the study of complex systems

In a rough sense, the current toolbox used in tackling complex systems involves three main areas: (i) nonlinear dynamics, (ii) statistical physics, including discrete models, and (iii) network theory. Elements of nonlinear dynamics should be familiar to most of the readers of this journal. The one with perhaps the greatest degree of novelty—because of the recent nature (and the speed) of most of the significant advances—is network theory. First, however, we will quickly comment on (i) and (ii).

3.1 Nonlinear dynamics and chaos

Nonlinear dynamics and chaos in deterministic systems are now an integral part of science and engineering. The theoretical foundations are on firm mathematical footing and there are well agreed upon mathematical definitions of chaos, many of them formally equivalent. However, because of its relative novelty and, in many case, counterintuitive nature, there are still many misconceptions about chaos and its implications. Extreme sensitivity to initial conditions does not mean that prediction is impossible. Memory of initial conditions is lost within attractors but the attractor itself may be extremely robust. In particular chaotic does not mean unstable.

Chaos means that simple systems are capable of producing complex outputs. Simple 1D mappings can do this—the logistic equation being the most celebrated example. The flip side is that complex looking outputs need not have complex or even complicated origins; seemingly random-looking outputs can be due to deterministic causes. Many techniques have been developed to analyze signals and to determine if fluctuations stem from deterministic components.

Nonlinear dynamics is now firmly embedded throughout research; applications arise in virtually all branches of engineering and physics—from quantum physics to celestial mechanics. There are numerous applications in geophysics, physiology and neurophysiology [3]. Even subapplications have developed into full-fledged areas. For example, mixing is one of the most successful areas of applications of nonlinear dynamics [4].

It is clear that nonlinear dynamics does not exist in isolation but it is now a platform competency in science and engineering. This does not mean that all theoretical questions have been answered and that all ideas are uncontroversial. For example there is significant discussion about the presence of chaos in physics and the role it may play in determining the universe's "arrow of time", the irreversible flow from the past to the future.

3.2 Statistical physics: Universality and scaling

Of the three revolutionary new areas of physics born at the turn of the 20th century—statistical physics, relativity, and quantum mechanics—it is fair to say that statistical physics has been the one that has least caught people's imagination. The reason may be that, on the surface, statistical physics most resembles pre-20th century physics. However, statistical physics brought three very important conceptual and technical advances:

1. It lead to a new conception of prediction (we shall have more to say about this later in the paper).
2. It circumvented classical mechanics and the impossibility to solve the three-body problem by tackling the many-body problem. In doing so, it casted solutions in terms of ensembles.
3. It introduced the concept of discrete models—ranging from the Ising model to cellular automata [5] and agent-based models [6].

In the 1960s and 1970s, fundamental advances occurred in our understanding of phase transitions and critical phenomena leading to the development of two important new concepts: *universality* and *scaling* [7,8]. The finding, in physical systems, of universal properties that are independent of the specific form of the interactions gives rise to the intriguing hypothesis that universal laws or results may also be present in complex social, economic and biological systems (see Fig. 2).

Indeed, it has recently come to be appreciated that many complex systems obey universal laws that are independent of the microscopic details: Findings in one system translate into understanding of the behavior of many others. For example, fluctuations in physiologic outputs of healthy individuals display universal degree of correlations [9–12], as do fluctuations of financial assets [13–16]. Similarly, it has been recently shown that scaling and universality hold for a broad range of human organizations [17–23].

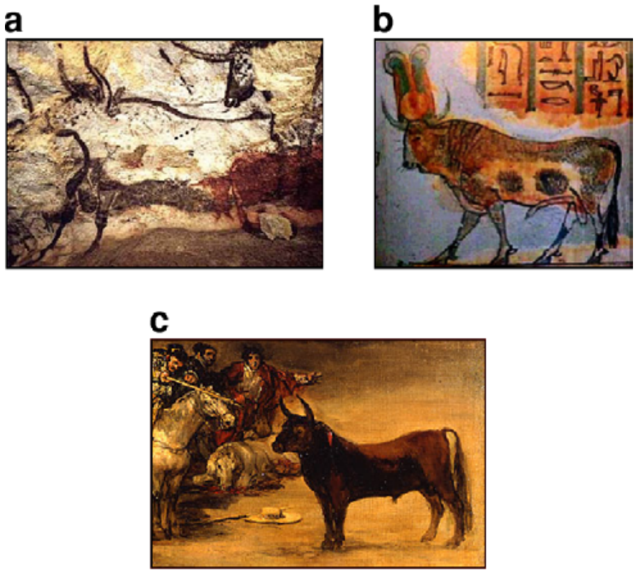

Fig. 2. Visualizing universality. (a) Lascaux cave paintings, beginning of Magdalenian Age (approximately 15,000 to 13,000 B.C.); (b) Apis bull, Egypt (3,000–500 B.C.); (c) Bullfight: Suerte de vara (detail), Francisco de Goya y Lucientes (1824); oil on canvas (50 × 61 cm), The J. Paul Getty Museum, Los Angeles. Despite the difference in details, styles, and medium, all images are easily identified as depictions of bulls. Clearly, all images capture the essential characteristics of the animal. However, for a computer program, the task of classifying the subject matter of all pictures as being identical is far from trivial. The concept of universality in statistical physics and complex systems may aspires to the same goal as such a computer program would: to capture the essence of different systems and to classify them into distinct classes.

3.2.1 Scaling

The *scaling hypothesis* which arose in the context of the study of critical phenomena led to two categories of predictions, both of which have been remarkably well verified by a wealth of experimental data on diverse systems. The first category is a set of relations, called *scaling laws*, that serve to relate the various critical-point exponents characterizing the singular behavior of the order parameter and of response functions. The second category is *data collapsing*.

The predictions of the scaling hypothesis are supported by a wide range of experimental work, and also by numerous calculations on model systems. Moreover, the general principles of scale invariance have proved useful in interpreting a number of other phenomena, ranging from elementary particle physics [24] and galaxy structure [25] to finance and sociology [11,23,26].

3.2.2 Universality

Another fundamental concept arising from the study of critical phenomena is *universality*. For systems in the same universality class, exponents and scaling functions are the same in the vicinity of the critical point. This fact suggests than when studying a given problem, one may pick the most tractable system to study and the results one obtains will hold for all other systems in the same universality class [7,8]. The problem, clearly, is to identify which systems belong to a given universality class.

The universality of critical behavior motivated the search for the features of the microscopic inter-particle force which are important for determining critical-point exponents and scaling functions, and which ones are unimportant. These questions were answered by numerous works on the renormalization group [27]. These studies led to the idea that when the scale changes, the equations which describe the system also change accordingly and that in the macroscopic limit only a few "relevant" features remain. When one uncovers universality in a given system, it means that some profound, usually simple, mechanisms are at work. This conceptual framework has guided many physicists into forays in interdisciplinary research yielding insights across seemingly dissimilar disciplines.

3.2.3 Discrete models

Discrete-space and discrete-time modeling is based on the assumption that some phenomena can and should be modeled directly in terms of computer programs (algorithms) rather than in terms of equations. Cellular automata—which can be traced to John von Neumann [28] and Stanislaw Ulam [29] and were further developed and popularized in Conway's "game of life" [30] and, more recently, Wolfram [5]—are the simplest example of discrete time and space models that were developed with the computer in mind.

Examples of the application of cellular automata exist in physical, chemical, biological and social sciences; they can be as simple as propagation of fire and simple predator-prey models between a handful of species and as complex as the evolution of artificial societies. The central idea is to have agents that live on the cells of regular d-dimensional lattices and interact with each other according to prescribed rules. The basic building blocks may be identical or may differ in important characteristics; moreover these characteristics may change over time, as the agents adapt to their environment and learn from their experiences—see e.g. Epstein and Axtell [6] in the context of the social sciences.

Discrete, or agent-based, modeling has been extremely successful because of the intuition-building capabilities it provides and the speed with which it permits the investigation of multiple scenarios. For this reason discrete modeling has led in some cases to a replacement of equation-based approaches in disciplines such as ecology, traffic optimization, supply networks, and behavior-based economics.

4 Networks

The third element in the toolbox is networks. A network is a system of nodes with connecting links. Once one adopts this viewpoint, networks appear everywhere [31–35]. Consider the following examples from the biological science:

- food webs, a network of species connected by trophic interactions [49–53],
- autonomous nervous systems of complex organisms, a network of neurons connected by synapses [54,55],
- gene regulation networks, a network of genes connected by cross-regulation interactions [56–58],
- protein networks, a network of protein connected by participation in the same protein complexes [59–61],
- metabolic networks, a network of metabolites connected by chemical reactions [62–64].

Social networks are also ubiquitous [36–43]: Individuals exchanging e-mails is one example [44,45]. Person A sends and e-mail to B; if B replies A and B are connected. Other clear-cut examples are the Internet, a network of servers, and the World Wide Web, a network of web pages connected by hyperlinks [46–48].

The two limiting network topologies typically considered are: (i) d-dimensional graphs—a lattice, for example—where every node connects with a well-defined set of closest neighbors, and (ii) random graphs, where every node has the same probability of being connected to any other node. Quantities used to quantitatively describe networks include [69]:

- The minimum number of links that must be traversed to travel from node i to node j is called the shortest path length or distance between i and j. A graph is connected if any node can be reached from any other node; otherwise the graph is disconnected. The average path length is the average of the minimum number of steps necessary to connect any two nodes in a connected network.
- The local clustering is (roughly) the number of actual links in a local sub-network divided by the number of possible links. It quantifies the fact that if Person A is good friend with both B and C, then there is a good chance B and C are also friends [70].
- The degree distribution, $p(k)$, which is the probability of finding a node with k links. In a lattice $p(k)$ is a delta-Kronecker function while in a random graph it is a Poisson distribution.

Real networks, however, are not well described by either model [31–35,66]. Real networks are both clustered (high degree of local connectivity) and small-worlds (it takes only a small number of steps to connect any two nodes).

4.1 Network theory: A short history

The surge of interest in networks is recent, however, the history of network has a distinguished past: The birth of network (or graph) theory links together two famous mathematicians: Euler and Erdös. The "conception" of the theory is universally attributed to Euler [65] and his solution of the celebrated Königsberg bridge puzzle. As stated in Euler's manuscript: *"In the town of Königsberg in Prussia there is an island A, called "Kneiphoff", with the two branches of the river (Pregel) flowing around it. There are seven bridges, a, b, c, d, e, f, and g, crossing the two branches. The question is whether a person can plan a walk in such a way that he will cross each of these bridges once but not more than once. [...] On the basis of the above I formulated the following very general problem for myself: Given any configuration of the river and the branches into which it may divide, as well as any number of bridges, to determine whether or not it is possible to cross each bridge exactly once."*

Euler's solution of the Königsberg bridge puzzle developed naturally from his formulation of the problem, once again showing that formulation of a problem is as important, if not more than, the solution itself. Euler noticed

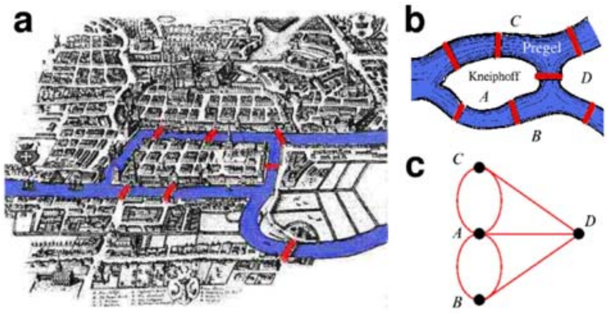

Fig. 3. The Königsberg bridge puzzle [65]. (a) The town of Königsberg, now Kaliningrad, Russia, had at the time seven bridges connecting the island of Kneiphoff to the margins of the river Pregel. (b) Schematic representation of the area with the bridges. (c) Euler's representation of the problem. Euler realized that physical distance was of no importance in this problem, only topology matters. For this reason the bridges can be represented as links in a graph connecting nodes representing the different margins and islands.

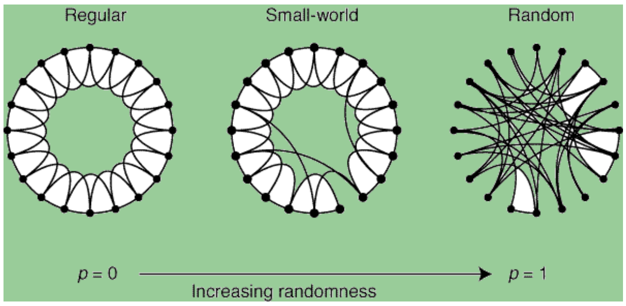

Fig. 4. A minimal model for generating small-world networks. Watts and Strogatz construct networks that exhibit the small-world phenomenon by randomizing a fraction p of the links connecting nodes in an ordered lattice. In the case displayed, the ordered lattice is one-dimensional with 4 connections per node. After [70].

that physical distance is of no importance in this problem and represented the topological constraints of the problem in the form of a graph—a set of nodes and the set of links connecting pairs of nodes (Fig. 3). Euler divided the nodes into odd and even based on the parity of the degree of the node, that is, the number of links directly connected to the node. He then demonstrated that:

1. The sum of degrees of the nodes of a graph is even;
2. Every graph must have an even number of odd nodes.

These results enabled Euler to show that:

1. If the number of odd nodes is greater than 2 no Euler walk exists—a Euler walk being a walk between two arbitrary nodes for which every link in the graph appears exactly once;
2. If the number of odd nodes is 2, Euler walks exist starting at either of the odd nodes;
3. With no odd nodes, Euler walks can start at an arbitrary node.

Therefore, since all four nodes in the Königsberg bridge problem are odd, Euler demonstrated that there was no solution to the puzzle, that is, there was no path transversing each bridge only once. Euler's work was of seminal importance because it identified topology as the key issue of the problem, thus enabling his later work on topology and the establishment of, e.g. relations among the numbers of edges, vertices and faces of polyhedrons.

If the conception of network theory is due to Euler, its "delivery" is due in great part to Erdös. As in Euler's case, Erdös interest on network theory was initially linked to a social puzzle: What is the structure of social networks? This problem was formalized by Kochen and Pool in the 50's, leading them to the definition of random graphs [66]—graphs in which the existence of a link between any pair of nodes has probability p. Erdös, in collaboration with Rényi, pursued the theoretical analysis of the properties of random graphs obtaining a number of important results, including the identification of the percolation threshold—that is, the average number of links per node necessary in order for a random graph to be fully connected—or the typical number of intermediate links in the shortest path between any two nodes in the graph.

4.2 Small-world networks

Kochen and Pool's work, which was widely circulated in preprint form before it finally was published in 1981 [66], was a precursor to experimental work that lead to the discovery of the so-called six-degrees of separation phenomenon, later popularized in a homonym play by John Guare. The six-degree of separation phenomenon is typically referred to in the scientific literature as the small-world phenomenon [67,68].

A recurrent characteristic of networks in complex systems is the small-world phenomenon, which is defined by the co-existence of two apparently incompatible conditions, (i) the number of intermediaries between any pair of nodes in the network is quite small—typically referred to as the six-degrees of separation phenomenon—and (ii) the large local "cliquishness" or redundancy of the network—i.e., the large overlap of the circles of neighbors of two network neighbors. The latter property is typical of ordered lattices, while the former is typical of random graphs [69].

Recently, Watts and Strogatz [70] proposed a minimal model for the emergence of the small-world phenomenon in simple networks. In their model, small-world networks emerge as the result of randomly rewiring a fraction p of the links in a d-dimensional lattice (Fig. 4). The parameter p enables one to continuously interpolate between the two limiting cases of a regular lattice ($p = 0$) and a random graph ($p = 1$).

Watts and Strogatz probed the structure of their small-world network model and of real networks via two quantities: (i) the mean shortest distance L between all pairs of nodes in the network, and (ii) the mean clustering coefficient C of the nodes in the network. For a d-dimensional lattice one has $L \sim N^{1/d}$ and $C = O(1)$, where N is the

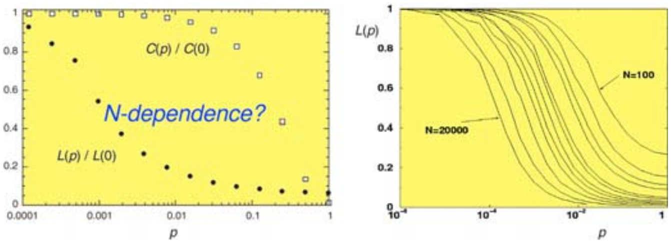

Fig. 5. Ubiquity of small-world networks. (a) Dependence of L and C on p for the small-world model of Watts and Strogatz. The emergence of the small-world regime is clear for $p > 0.01$, as $L(p)$ quickly converges to the random graph value, while $C(p)$ remains in the ordered graph range. After [70]. (b) Dependence of L on p for different network sizes. The numerical results show that the emergence of the small-world regime occurs for a value of p that approaches zero as N diverges [71,72]. After [72].

number of nodes in the network. In contrast, for a random graph one has $L \sim \ln N$ and $C \sim 1/N$. Figure 5a shows the dependence of L and C on p for the for the small-world model of Watts and Strogatz. The emergence of the small-world regime is clear for $p > 0.01$, as L quickly converges to the random graph value, while C remains in the ordered graph range, this two characteristics defining a small-world network. Watts and Strogatz [70] found clear evidence of the small-world phenomenon in (a) the electric-power grid for Southern California, (b) the network of movie-actor collaborations, and (c) the neuronal network of the worm *C. elegans*.

A question is prompted by the results of Figure 5a: "Under which conditions does the small-world regime emerge?" Specifically, does the small-world behavior emerge for a finite value of p when N approaches the thermodynamic limit? [71]. Numerical results and theoretical arguments show that the emergence of the small-world regime occurs for a value of p that approaches zero as N diverges [71,72]; cf. Figure 5b. The implications of this finding are quite important: Consider a system for which there is a finite probability p of random connections. It then follows that independently of the value of p, the network will be in the small-world regime for systems with size $N \sim 1/p$, the reason being that to have a finite number of random links, i.e., that Np must be of $O(1)$. This implies that most large networks are small-worlds. Importantly, the nodes will be "un-aware" of this fact as the vast majority of them has no long-range connections [71].

4.3 Scale-free networks

An important characteristic of a graph that is not taken into consideration in the small-world model of Watts and Strogatz is the degree distribution, i.e., the distribution of number of connections of the nodes in the network. The Erdös-Rényi class of random graphs has a Poisson degree distribution [69], while lattice-like networks have even more strongly peaked distributions—a perfectly ordered lattice has a delta-Dirac degree distribution. Similarly, the small-world networks generated by the Watts and Strogatz model also have peaked, single-scale, degree

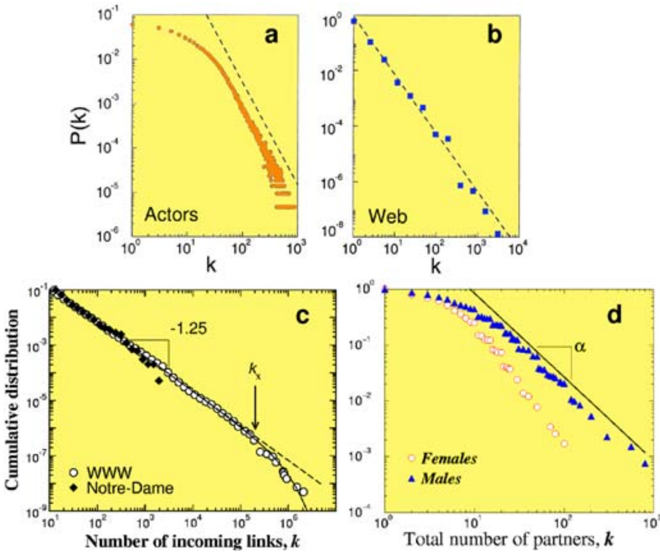

Fig. 6. Ubiquity of scale-free networks. Double logarithm plot of (a) the degree distribution of the network of movie-actor collaborations (each node corresponds to an actor and links between actors indicated that they collaborated on at least one movie); (b) the degree distribution of the webpages in the `nd.edu` domain (each node is a webpage and links between webpages indicate hyperlinks pointing to the other webpage). After [73]. (c) Degree distribution of the WWW. Note the truncation of the power law regime. After [78]. (d) Distribution of number of sexual partners for Swedish females and males. Note the power law decay in the tails of the distributions. After [39].

distributions, i.e., one can clearly identify a typical degree of the nodes comprising the network.

Against this theoretical background, Barabási and coworkers found that a number of real-world networks have a scale-free degree distribution with tails that decay as a power law [47,73]. As shown in Figures 6a–c, the network of movie-actor collaborations, the webpages in the `nd.edu` domain, and the power grid of Southern California, all appear to obey distributions that decay in the tail as a power law [73]. Moreover, other networks such as the network of citations of scientific papers also are reported to be scale-free [74,75].

Barabási and Albert [73] suggested that scale-free networks emerge in the context of growing network in which new nodes connect preferentially to the most connected nodes already in the network. Specifically,

$$p_i(n+1) = \frac{k_i(n)}{\sum_{i=-n_0+1}^{n} k_i(n)}, \qquad (1)$$

where n is the time and number of nodes added to the network, n_0 is the number of initial nodes in the network at time zero, k_i is the degree of node i and $p_i(n+1)$ is the probability of a new node, added at time $n+1$ linking to node i. As illustrated in Figure 7, as time ticks by the degree distribution of the nodes becomes more and more heterogeneous since the nodes with higher degree are the most likely to be the ones new nodes link to.

Significantly, scale-free networks provide extremely efficient communication and navigability as one can easily

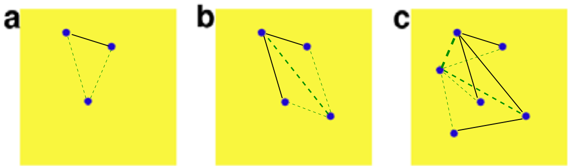

Fig. 7. Three stages in the time evolution of a minimal model for generating scale-free networks [73]. (a) We start with a network comprising two nodes linked by a bi-directional connection (black full line). Then, we add a new node which can link to either of the existing nodes (green dashed line). Because both existing nodes have degree one, there is an equal probability of linked to each of them (which is represented by the thickness of the dashed line). (b) At the following time step, we add a new node to the network. As before, this node can link to any of the existing nodes. However, now the probability of linking to each of the existing nodes is no longer identical because one of the nodes has higher degree than the others. (c) As times goes by, a heterogeneous degree distribution emerges because nodes with higher degree have a higher probability of being linked to new nodes.

reach any other node in the network by sending information through the "hubs", the highly-connected nodes. The efficiency of the scale-free topology and the existence of a simple mechanism leading to the emergence of this topology led many researchers to believe in the *complete* ubiquity of scale-free network. As it often happens, one finds what one is looking for (Figs. 8a–c).

Note that scale-free networks are a subset of all small-world networks because (i) the mean distance between the nodes in the network increases extremely slowly with the size of the network [73,76], and (ii) the clustering coefficient is larger than for random networks.

4.4 Classes of small-world networks

An important aspect question prompted by the work of Barabási and Albert is how to connect the findings of Watts and Strogatz on small-world networks with the the new finding of scale-free structures. Specifically, one may ask "Under what conditions will growing networks be scale-free?" or, more to the point, "Under what conditions will the action of the preferential attachment mechanism be hindered?" Recall that preferential attachment gives rise to a scale-free degree distribution in growing networks [73], hence if preferential attachment is not the only factor determining the linking of incoming nodes one may observe other topologies. Amaral and co-workers have demonstrated that preferential attachment can be hindered by at least three classes of factors:

Aging—This effect can be illustrated with the network of actors. In time, every actress or actor stops acting. For the network, this implies that even a very highly connected node will eventually stop receiving new links. The node may still be part of the network and contributing to network statistics, but it no longer receives links. The aging of the nodes thus limits the preferential attachment preventing a scale-free distribution of degrees from emerging [77].

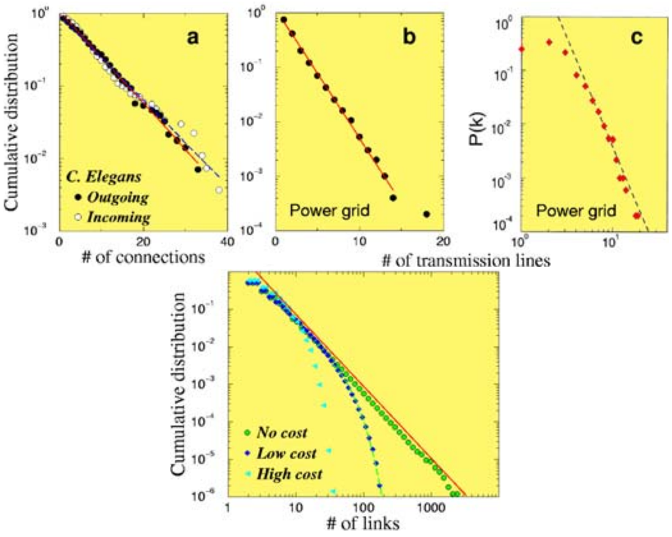

Fig. 8. Evidence for existence of single scale networks. (a) Degree distribution of the nematode *C. elegans*. Each of the 302 neurons of *C. elegans* and their connections has been mapped. Note that the plot is semi-logarithmic, so a straight line indicates an exponential dependence. After [77]. (b) Semi-logarithmic plot of the cumulative degree distribution of the power grid of Southern California (each node is a transmission station and links are power lines connecting the stations). After [77]. (c) Double-logarithmic the degree distribution of the power grid of Southern California. After [73]. Note that the power law fit provides a poor description of the data while the exponential fit matches the data remarkably well for all degree values. (d) Truncation of scale-free degree of the nodes by adding constraints to the model of reference [73]. Effect of cost of adding links on the degree distribution. These results indicate that the cost of adding links also leads to a cut-off of the power-law regime in the degree distribution, and that for a sufficiently large cost the power-law regime disappears altogether. After [77].

Cost of adding links and limited capacity—This effect can be illustrated with the network of world airports. For reasons of efficiency, commercial airlines prefer to have a small number of hubs through which many routes connect. To first approximation, this is indeed what happens for individual airlines, but when we consider all airlines together, it becomes physically impossible for an airport to become a hub to all airlines. Due to space and time constraints, each airport will limit the number of landings/departures per hour, and the number of passengers in transit. Hence, physical costs of adding links and limited capacity of a node will limit the number of possible links attaching to a given node [77].

Limits on information and access—This effect can be illustrated with the selection of outgoing links from a webpage in the world-wide-web: Even though there is no meaningful cost associated with including a hyperlink to a given webpage in one's own webpage, there may be constraints effectively blocking the inclusion of some webpages, no matter how popular and well connected they may been.

An example of such constraints is distinct interest areas—a webpage on granular mixing is unlikely to include links to webpages discussing religion [78].

These different constraints can be formalized by adding terms to 1. Specifically,

$$p_i(n+1) = \frac{k_i(n) f(k_i(n), n, i, ...)}{\sum_{i=-n_0+1}^{n} k_i(n) f(k_i(n), n, i, ...)}, \qquad (2)$$

where $f(k_i(n), n, i, ...)$ is a cost function that may depende on the degree of node i, on its age, on time, and on a number of other factors. As can be seen in Figure 8, the presence of constraints leads to a cut-off of the power-law regime in the degree distribution, and that for a sufficiently strong constraints the power-law regime disappears altogether [77]. Empirical data suggest the existence of three classes of small-world networks [77]: (a) *scale-free* networks; (b) *broad-scale* or truncated scale-free networks, characterized by a degree distribution that has a power-law regime followed by a sharp cut-off that is not due to the finite size of the network; (c) *single-scale* networks, characterized by a degree distribution with a fast decaying tail, such as exponential or Gaussian. It is important to note again that scale-free networks are small-world networks but the inverse may not be true!

4.5 Network modelling

The previous sections merely skim the surface of all the work being done in modelling the structure and emergence of complex networks. This is an extremely active field of research with hundreds of papers having been written, and published, in the last four years. Clearly, this paper is not the forum to discuss in detail all of those modelling efforts and we direct the readers to the excellent reviews [32,34,35,79] and books [31,80] available.

5 Brief review of a selection of complex systems' research with a network focus

5.1 The topology of the Internet and the dynamics of Internet traffic

The Internet [81,82]—a large communication network that now connects more than 10^8 hosts—is a prime example of a self-organizing complex system [83–86], having grown mostly in the absence of centralized control or direction. In this network, information is transferred in the form of packets from the sender to the receiver via routers, computers which are specialized to transfer packets to another router "closer" to the receiver (Fig. 9a). A router decides the route of the packet using only local information obtained from its interaction with neighboring routers, not by following instructions from a centralized server. Thus, it can be viewed as a simple and autonomous entity. A router stores packets in its finite queue and processes them sequentially. However, if the queue overflows

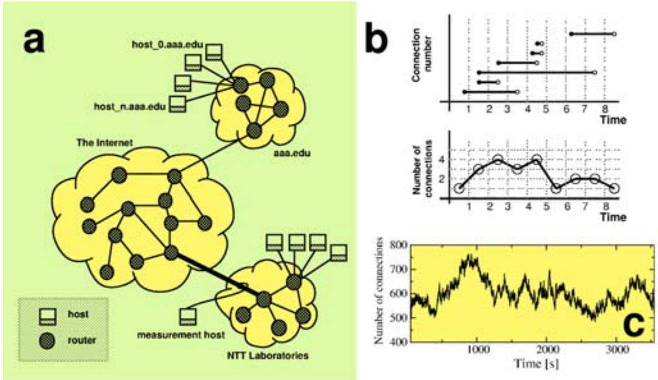

Fig. 9. Measuring Internet traffic. (a) Schematic representation of Internet topology. The Internet comprises links and nodes (routers and hosts), which are connected in a scale-free structure [46,106]. Suppose that one downloads a file on a web page in a host at `aaa.edu` from a host in NTT laboratories. In this case, a connection is established between the two hosts, and the packets encoding the file travel from the source host to the destination host via routers. The connection is based on a feedback control, so that the duration of the connection for transferring the file strongly depends on the current network traffic conditions; if the network is congested, it takes longer to finish the transfer. (b) Measuring the number of connections. In the top panel, each line represents a connection passing through the observation link. For example, the first connection starts at 0.5 s (black circle) and finishes at 3.5 s (open circle). We count the number of connections within a one second interval and obtain a time series of the number of connections (bottom panel). (c) A typical data set, showing the number of connections between NTT laboratories and the Internet for each of the 3,600 one second intervals between 13:00 and 14:00 on July 19, 2001. After [97].

due to excess demand, the router will discard incoming packets, a situation corresponding to congestion. A router can only control incoming traffic by discarding arriving packets, so that in order to know and adjust to the current network traffic condition, each host has the ability to control its traffic by using a feedback-based flow control [87] for the communication between the sender and the receiver.

Even though the rules controlling traffic flow were programmed by humans, the dynamics of Internet traffic [88–95] are difficult to predict due to the complex interactions between routers, the flow control mechanisms, and the diversity of applications running in the Internet. Moreover, the traffic flow is also highly correlated with human activity.

A number of studies has probed the topology of the Internet and its implications for traffic dynamics. It has been reported that Internet traffic fluctuations are statistically self-similar [88–90] and that the traffic displays two separate phases, congested and non-congested. At the point separating the two phases, traffic fluctuations are characterized by $1/f$–correlations [91–95].

Fukuda and co-workers recorded every packet flowing through the link between NTT laboratories in Tokyo and the Internet during 4 days in July 2001 (Figs. 9b, c). They found that time series of number of connections are non-stationary, and are characterized by different mean values depending on the observation period; the mean number of connections is $\approx$50 connections/s during the night, and $\approx$700 connections/s during the day. Moreover, they found that the distribution of durations of stationary periods [96] display tails decaying as power-laws for both night and day—implying that the duration of stationary periods has no typical scale [97].

Eriksen et al. [98] studied the spectral properties of a diffusion process taking place on the Internet network focusing on the slowest decaying modes. These modes enabled them to identify modules in the topology of the Internet. Eriksen et al. were able to map those modules to individual countries. For example, the slowest decaying mode corresponds to a diffusion current flowing from Russia to US military sites [98].

Barthélémy et al. analyzed data from the French national "Renater" network which comprises 30 interconnected routers located in difference regions of France and is used by approximately 2 million individuals [99]. They found that the Internet flow is strongly localized: most of the traffic takes place on a spanning network connecting a small number of routers which can be classified either as "active centers," which are gathering information, or "databases," which provide information. Interestingly, Barthélémy and co-workers also found that the Internet activity of a region increases with the number of published papers by laboratories of that region [99].

A number of groups have also demonstrated that the Internet displays a number of properties that distinguishes it from random graphs: wiring redundancy and clustering, [46,100–102], non-trivial eigenvalue spectra of the connectivity matrix [46], and a scale-free degree distribution [46,101–104].

Vespignani, Pastor-Satorras and collaborators undertook a systematic study of the evolution of the topology of the Internet by analyzing Internet maps collected by the National Laboratory for Applied Network Research (NLANR) for the period 1997–2000 [105]. Their studies demonstrated that most quantities characterizing the topology of the network have reached stationary values and that the Internet has an hierarchical structure reflected in non-trivial scale-free betweenness and degree correlations functions [106]. They also found that the time evolution of the Internet topology reveals the presence of growth dynamics with aging features [107].

5.2 The topology of natural ecosystems

Species in natural ecosystems are organized into complex webs [49]. Ecologists have studied these webs from the perspective of network theory. Every species in the ecosystem being a node in a network and the existence of a trophic link—i.e., a prey-predator relationship—between two species indicating the existence of a *directed* link between them. We are far from this ideal, but understanding the structure of these food webs should be of fundamental importance in guiding policy decisions concerning, for example, the recommended limits on consumption of fish with unacceptable levels of pollutants, the selection of areas for establishment of protected ecosystems, or the management of boundary areas between protected ecosystems and agro-businesses.

The study of such questions is extremely challenging for a number of reasons. First, the characterization of the topology of a given ecosystem is a very cumbersome and expensive task, which a priori may be of value only for the particular environment considered. Second, the precise modelling of the nonlinear interactions between the numerous individuals belonging to each of the many species comprising the ecosystem and their separation from stochastic external variables (such as the climate) affecting the ecosystem may be impossible.

Recently, Amaral and co-workers studied the *topology* of food webs from a number of distinct environments—including freshwater habitats, marine-freshwater interfaces, deserts, and tropical islands—and found that this topology may be identical across environments and described by simple analytical expressions [51,52,108]. This finding is demonstrated in Figure 10, where, as an example, we present the distributions of number of prey and number of predators for the species comprising twelve distinct food webs [108].

In the same spirit, a recent paper in Nature reports on a study of food webs as transportation networks [109]. The underlying idea is that the directionality of the links (pointing from prey to predator) defines a "flow" of resources—energy, nutrients, prey—between the nodes of the network. Because every species feeds directly or indirectly on environmental resources, food webs are connected (that is, every species can be reached by starting from an additional "source" node representing the environment. This fact enabled Garlaschelli et al. [109] to define a spanning tree on any food web—i.e., a loop-less subset of the links of the web such that, starting from the environment, every species can be reached. Importantly, they find that those spanning trees are characterized by universal scaling relations.

These results are of great practical and fundamental importance because they support the hypothesis that scaling and universality hold for ecosystems—i.e., food webs display universal patterns in the way trophic relations are established despite *apparently* significantly differences in factors such as environment (e.g. marine versus terrestrial), ecosystem assembly, and past history. This fact suggests that a *general* treatment of the problems considered in environmental engineering, with reasonable caveats, may be within reach.

5.3 Spread of epidemics in complex networks

Some infectious agents are beneficial—the spread of ideas, technologies, or domesticated animals and some kinds of plants. Others clearly are not: human or animal infectious diseases [26,110], email virus and other computer virus [111,112]. A third class may involve those with no

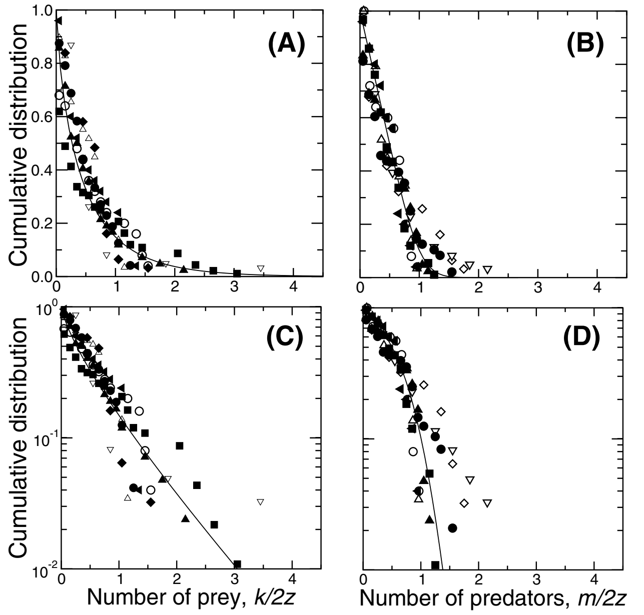

Fig. 10. Test of the "scaling hypothesis" that the distributions of number of prey (predators) have the same functional form for food webs from different environments. (a) Cumulative distribution $P_{\rm prey}$ of the scaled number of prey $k/2z$ for eight distinct food webs (see [108] for details), where z is the average number of trophic links per specie. The solid line is the analytical prediction derived in [51]. The data "collapse" onto a single curve that agrees well with the analytical results derived in [51]. (b) Cumulative distribution $P_{\rm pred}$ of the scaled number of predators $m/2z$ for the same eight webs as in a. The solid lines are the analytical predictions of semi-logarithmic plot of the scaled distributions of (c) number of prey, and (d) number of predators. After [108].

clear connotation, such as fads or rumours [113]. Regardless of the particular case, spreading processes share a number of properties which make them amenable to a generalized analysis. A number of general results have indeed been derived, including the fact that in the steady state regime a system can be in one of two phases—no epidemic or endemic disease—depending on the value of the so-called reproduction rate R of the infection. For $R < 1$ there is no epidemic, while for $R > 1$ the infection becomes endemic.

For the mean-field case, in which all units have the same number of contacts with (a random sample of other) units—the reproduction rate is proportional to the number of contacts c, to the probability of transmission β of the disease for each contact between infected and susceptible units, and the mean duration of the infectious state D. For the more realistic case in which the number of contacts per unit is not constant but follows some distribution, R is given by

$$R = c\beta D \left(1 + \frac{{\sigma_k}^2}{\bar{k}}\right) , \qquad (3)$$

where $\bar{k}$ and σ_k are the mean and standard deviation of the number of contacts, respectively. The dependence of R on c and β immediately suggests a strategy for preventing an epidemic: The reduction of the number of contacts between infected and susceptible individuals. One way to achieve this is through immunization, which reduces the number of susceptible individuals; another is to just decrease contacts by behavior modification.

Recently, Pastor-Satorras and Vespignani [111] demonstrated through numerical simulations and analytical calculations that scale-free networks do not have an epidemic threshold, i.e., $R > 1$ for $\beta > 0$. This seemingly counterintuitive result is particularly significant because network such as the web of sexual contacts [39] and the Internet [46] display a high degree of local redundancy and a power law decaying degree distribution (Fig. 6). As equation (3) implies, the absence of a finite infectability threshold is due to the unbounded variance of the degree in scale-free networks.

The same authors have also reported that the random uniform immunization of individuals does not lead to the eradication of infections in all complex networks. Namely, networks with scale-free degree distributions do not acquire global immunity from major epidemic outbreaks even in the presence of unrealistically high densities of randomly immunized individuals [112]. Pastor-Satorras and Vespignani showed that successful immunization strategies can be developed only by taking into account the heterogeneity of scale-free networks. In particular, targeted immunization schemes based on the nodes' degree can sharply lower the network's vulnerability to epidemic attacks. Similar results have been reported by other groups [114,115]

Eguíluz and Klemm have analyzed the spreading of infections in scale-free networks with high clustering and degree correlations, as found for the Internet [116]. They reported that the degree correlations restore a finite infectability threshold. However, Vazquez et al. [117] demonstrated that the finite threshold reported by Eguíluz and Klemm is simply a by-product of the fact that the model of reference [116] generates a one-dimension network.

Liebovitch et al. [118] analyzed the timing of the arrival of email viruses at different computers as a way of probing the structural and dynamical properties of the Internet. They found that the intervals between the arrival of four different strains of email viruses have a power law distribution and that there are positive correlations between these intervals [118]. Newman and coworkers [119] have studied the role of the structure of electronic mail networks [44,45] on the spread of computer viruses and have extracted the implications of this structure for the understanding and prevention of computer virus epidemics. Specifically, their study demonstrated that random immunization has little effect on virus spreading while targeted immunization can have a significant impact [119].

5.4 Cellular networks

The complexity of the web of nonlinear interactions between genes, proteins and the environment necessitates the development of simplified models to illuminate biological function. As Vogelstein et al. [120] wrote recently: "*How can the vast number of activating signals, covalent and non-covalent modifications, and downstream regulators of p53 be put into context? One way to understand the p53 network is to compare it to the Internet. [...] An appreciation of the existence and complexity of cellular networks should enable more rational design and interpretation of experiments in the future, and should allow more realistic approaches to treatment.*"

A number of recent studies have indeed started to highlight *the existence and complexity* of cellular networks. Oltvai, Barabási and co-workers performed a systematic analysis of the metabolic networks of 43 organisms representing all three domains of life [62]. They found that, despite significant variation in their individual constituents and pathways, these metabolic networks have the same topological scaling properties and show striking similarities to the inherent organization of complex non-biological systems. They concluded that metabolic organization is not only identical for all living organisms, but also complies with the design principles of robust and error-tolerant scale-free networks, and may represent a common blueprint for the large-scale organization of interactions among all cellular constituents [62].

The same group also studied the protein-protein interaction network for two organisms, the yeast *S. cerevisiae* and the bacterium *H. pylori* [59]. They found that the network of protein interactions for these two organisms form a highly inhomogeneous scale-free network in which a few highly connected proteins play a central role in mediating interactions among numerous, less connected proteins.

Further, Jeong et al. [59] tested the importance of the different proteins for the survival of the yeast by mutating its genome. For random mutations, they found that removal does not affect the overall topology of the network. However, they found that the likelihood that removal of a protein will prove lethal correlates with the number of interactions the protein has. For example, although proteins with five or fewer links constitute about 93% of the total number of proteins, they found that only about 21% of them are essential. By contrast, only some 0.7% of the yeast proteins with known phenotypic profiles have more than 15 links, but a single deletion of 62% of these proves lethal. This implies that highly connected proteins are three times more likely to be essential than proteins with only a small number of links to other proteins.

In order to uncover the structural design principles of complex networks, Uri Alon and co-workers defined network motifs, patterns of interconnections occurring in real networks at numbers that are significantly higher than those in randomized networks [57]. They found motifs in networks from biochemistry, neurobiology, ecology, and engineering. Remarkably, the motifs shared by ecological food webs were distinct from the motifs shared by the genetic networks of *E. coli* and *S. cerevisiae* or from those found in the World Wide Web. Similar motifs were found in networks that perform information processing, even though they describe elements as different as biomolecules within a cell and synaptic connections between neurons in the worm *C. elegans*.

Specifically, the two transcription networks and the neuronal connectivity network of *C. elegans* show the same motifs: a three-node motif termed "feedforward loop" and a four-node motif termed "bi-fan". The feedforward loop motif, in particular, may play a functional role in information processing. One possible function of this circuit is to activate output only if the input signal is persistent and to allow a rapid deactivation when the input goes off. Many of the input nodes in the neural feedforward loops are sensory neurons, which may require this type of information processing to reject transient input fluctuations that are inherent in a variable or noisy environment.

5.5 Infrastructure robustness

On September 11, 2000, a late Summer storm hit Chicago leading to the closing of O'Hare airport. Five thousand people were left stranded in the airport. Massive cancellations of incoming and departing flights caused ripple-effect delays at airports across the USA. On August 15, 2003, four otherwise innocuous events on the northeast Ohio power grid—including the failures of a coal-fired generator and an automated warning system—combined to create a catastrophic situation that darkened a huge patch of the eastern United States and Canada. Parts of New York City were left without electricity for several days! At 3am on September 28, 2003, 57 million italians—the entire country except for Sardinia—found themselves without electricity. The problem was blamed on a series of failures on power lines from Switzerland and France due to the bad weather and on a tree felled by the storms.

As theses cases amply illustrate, the robustness of critical infrastructures in particular, and complex networks in general, is a matter of the great importance. Recent work on network theory has started to address the question of the robustness of complex networks to failure and directed attack. By robustness of the network we mean the ability of the network to sustain a giant component. The problem will be quite different depending on the existence or not of cascading events. For example, in the cases considered earlier, the nodes have limited capacity and thus one observes cascading events, as the failure of a node leads to a redistribution of the load in the nodes of the network and the consequent failure of other nodes.

Albert and co-workers studied the robustness of scale-free and single-scale networks to failure of nodes in the absence of cascading effects [121]. They found that scale-free networks were more robust to failure than single-scale networks. However, they also found that scale-free networks were very susceptible to attack of the highest degree nodes [121]. Similar calculations were performed by Broder and co-workers for the world-wide-web graph [122]. These results provoked a number of theoretical efforts aiming to characterize the percolation transition controlled by the critical fraction of nodes needed to be removed before the giant component breaks-up [123–125]. The case of cascading failure was first considered by Watts [113] for node failure, while Holme has considered both link and node failure with cascading [126,127].

6 The meaning of prediction and the study of complex systems

Much discussion and debate, not always useful, has arisen when evaluating the fruits of a complex systems approach to problems. In our view, much of the disagreement is due to overly restrictive views of what is meant by prediction and what the limits to prediction are.

In order to put this question into perspective, let us examine the usual meaning of prediction in the natural sciences, the "Newtonian" definition, put forward in its strongest form by Laplace and the standard under which most scientists still operate today. In Newtonian physics one is able to predict the future and post-dict the past of any system for which one knows the position and velocity of all particles. A modern perspective reveals a number of deficiencies.

First, it does not take into consideration computability issues. These are of two kinds. Assume one wants to compute the state of the entire universe, wouldn't the "computer" be part of the system? Clearly, one cannot possibly model the behavior of the entire Universe, as that would not leaves us with any material substrate with which to store the information or with which to perform the computation, for the same reason that one cannot draw a map that contains every detail of the real-world as the map must then be part of the map itself. Even if one would consider only a subset of the Universe, say, the water in a glass, one would still have to take into consideration the influence of the rest of the Universe on the water. One could easily model such influence as 'noise' acting on the system but that noise would destroy our ability to implement the Laplacian goal of predicting *exactly* the position and velocity of all particles.

Second, the study of deterministic nonlinear systems has clearly demonstrated the impossibility in exactly predicting of the velocities and positions of even simple systems interacting nonlinearly. The extreme dependence on initial conditions of chaotic nonlinear systems implies that in order to predict the positions and velocities of the units comprising a system interacting nonlinearly one would need to to be able to measure initial velocities and positions exactly, a clearly unattainable goal even without considering quantum effects.

Moreover, even if Newtonian prediction was possible, it would not, in our view, convey in an enlightening and conceptual-building way the *relevant* information about the system. Consider again the water in a glass; it is clear that one can in principle determine the macroscopic state of the system—solid, liquid, or gas—and its temperature, volume and pressure from a complete description of the positions and velocity of the $O(10^{23})$ particles composing the system. However, why would one *want* to do this? Clearly, the values of the macroscopic thermodynamic quantities provide a considerably more parsimonious description of the system. And, unarguably, the thermodynamic description of the system permits a deeper insight into the behavior of the system than the Newtonian approach of calculating forces and determining trajectories of all particles.

A more relevant class of what constitutes prediction in the context of the study of complex systems originated with developments in our understanding of phase transitions and critical phenomena. Close to the critical point most details of the system become irrelevant and the behavior of the systems is determined by a small number of relevant parameters and "mechanisms". For this reason, systems that may be very different in their details are actually described by the exact same scaling functions and sets of exponents [7,8]. A striking example of this type of prediction is the derivation of so-called allometric relationships: For example, the functional relationship between an organism's mass and its metabolic rate holds for organisms varying in mass over 27 orders of magnitude [128–130].

7 Concluding remarks

A crude simplification, not without truth, ascribes the study of simple systems to the natural sciences and the study of complicated systems to engineering. Systems that we have classified here as complex, have been left to the social sciences, to ecology, and to medicine. The underlying belief is that for simple and complicated systems one can discover quantitative "laws" while for complex systems one can only hope to obtain qualitative "lessons" or heuristics [131].

The common characteristic of all complex systems is that they display organization without any external organizing principle being applied; a central characteristic is adaptability. The topic has clearly captured the attention of physicists, statistical physicists in particular.

For engineers the conceptual conflict may arise from the fact that the hallmark of complex systems is adaptability and emergence: No one designed the web, the US power grid, or the metabolic processes within a cell. Engineering is not about letting systems be. The etymology of "engineer," both the verb and the noun, is revealing: *ingenitor*, contriver, *ingenire*, to contrive, as in to engineer a scheme. Engineering has a purpose and end result. Engineering is about convergence, assembling pieces that work in specific ways, optimum design and consistency of operation; the central metaphor is a clock. Complex systems, on the other hand, are about adaptation, self-organization and continuous improvement; the best metaphor may be an ecosystem. It is robustness and failure where both camps merge [132]. However, a successful merge will require augmenting the conceptual framework, even to the point of reshaping what one means by prediction.

The conceptual challenge for social science is that quantitative laws may actually govern the dynamics and organization of complex systems, the result of many interactions resulting in a clearly determinable outcome and that these results are not in any way contradictory to the existence of free will. For engineering, the challenge is the realization that many systems—for example, the web, the power grid, and the metabolic pathways in a cell—are not the result of a single design but an evolution and merging of designs, and that in many instances, self-organization can be used in profitable ways. It is apparent that continuous and open exchange of ideas among all groups will lead to higher level of understanding of complex systems and the development of new theory and tools.

We thank J.S. Andrade Jr., A. Arenas, A. Diaz-Guilera, M. Barthélémy, C. Edling, R. Guimerà, V. Hatzimanikatis, F. Liljeros, A.A. Moreira, S. Mossa, A. Scala, and A. Vespignani for many stimulating discussions.

References

1. J.M. Ottino, AIChE J. **49**, 292 (2003)
2. G.L. Baker, J.P. Gollub, *Chaotic Dynamics: An Introduction* (Cambridge University Press, Cambridge, UK, 1990)
3. L. Glass, M.C. MacKey, *From Clocks to Chaos: The Rhythms of Life* (Princeton University Press, Princeton, NJ, 1988)
4. J.M. Ottino, *The Kinematics of Mixing: Stretching, Chaos, and Transport* (Cambridge University Press, Cambridge, UK, 1989)
5. S. Wolfram, *A New Kind of Science* (Champaign, IL, Wolfram Media, 2002)
6. J.M. Epstein, R.L. Axtell, *Growing Artificial Societies: Social Science from the Bottom Up* (Cambridge, MA, MIT Press, 1996)
7. H.E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, Oxford, 1971)
8. H.E. Stanley, Rev. Mod. Phys. **71**, S358–S364 (1999) [Special Issue for the Centennial of the American Physical Society]
9. C.-K. Peng, S. Havlin, H.E. Stanley, A.L. Goldberger, Chaos **5**, 82 (1995)
10. P.Ch. Ivanov, L.A.N. Amaral, A.L. Goldberger et al., Nature **399**, 461 (1999)
11. L.A.N. Amaral, P.Ch. Ivanov, N. Aoyagi et al., Phys. Rev. Lett. **86**, 6026 (2001)
12. A.L. Goldberger, L.A.N. Amaral, J.M. Hausdorff et al., Proc. Nat. Acad. Sci. USA **99 Supp. 1**, 2466 (2002)
13. U.A. Muller, M.M. Dacorogna, R.B. Olsen et al., J. Banking Fin. **14**, 1189 (1990)
14. A. Pagan, J. Empirical Finance **3**, 15 (1996)
15. P. Gopikrishnan, V. Plerou, L.A.N. Amaral et al., Phys. Rev. E **60**, 5305 (1999)
16. V. Plerou, P. Gopikrishnan, L.A.N. Amaral et al., Phys. Rev. E **60**, 6519 (1999)
17. M.H.R. Stanley, L.A.N. Amaral, S.V. Buldyrev et al., Nature **379**, 804 (1996)
18. L.A.N. Amaral, S.V. Buldyrev, S. Havlin et al., J. Phys. I France **7**, 621 (1997)
19. Y. Lee, L.A.N. Amaral, D. Canning et al., Phys. Rev. Lett. **81**, 3275 (1998)
20. L.A.N. Amaral, S.V. Buldyrev, S. Havlin et al., Phys. Rev. Lett. **80**, 1385 (1998)
21. V. Plerou, L.A.N. Amaral, P. Gopikrishnan et al., Nature **400**, 433 (1999)
22. L.A.N. Amaral, P. Gopikrishnan, K. Matia et al., Scientometrics **51**, 9 (2001)
23. H.E. Stanley, L.A.N. Amaral, S.V. Buldyrev et al., Proc. Nat. Acad. Sci. USA **99 Supp. 1**, 2561 (2002)
24. R. Jackiw, Phys. Today **25**, 23 (1972)
25. P.J.E. Peebles, *The Large-Scale Structure of the Universe* (Princeton University Press, Princeton, NJ, 1980)
26. F. Liljeros, C.R. Edling, L.A.N. Amaral, Microbes Infect. **5**, 189 (2003)
27. J.J. Binney, N.J. Dowrick, A.J. Fisher, M.E.J. Newman, *The Theory of Critical Phenomena: An Introduction to the Renormalization Group* (Oxford University Press, Oxford, 1992)
28. J. von Newmann, *Theory of self-reproducing automata*, edited by A.W. Burks (University of Illinois Press, Urbana and London, 1966)
29. S. Ulam, in Proc. Int. Congress Math. **2**, 264 (1950)
30. M. Gardner, Scientific American **223**, 120 (1970)
31. D.J. Watts, *Small Worlds: The Dynamics of Networks Between Order and Randomness* (Princeton University Press, Princeton, NJ, 1999)
32. M.E.J. Newman, J. Stat. Phys. **101**, 819 (2000)
33. S.H. Strogatz, Nature **410**, 268 (2001)
34. R. Albert, A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002)
35. S.N. Dorogovtsev, J.F.F. Mendes, Adv. Phys. **51**, 1079 (2002)
36. S. Wasserman, K. Faust, *Social Network Analysis* (Cambridge University Press, Cambridge, UK, 1994)
37. A.F.J. van Raan, Nature **347**, 626 (1990)
38. M.E.J. Newman, P. Natl. Acad. Sci. USA **98**, 404 (2001)
39. F. Liljeros, C.R. Edling, L.A.N. Amaral et al., Nature **411**, 907 (2001)
40. E.M. Jin, M. Girvan, M.E.J. Newman, Phys. Rev. E **64**, 046132 (2001)
41. M. Girvan, M.E.J. Newman, P. Natl. Acad. Sci. USA **99**, 7821 (2002)
42. D.J. Watts, P.S. Dodds, M.E.J. Newman, Science **296**, 1302 (2002)
43. M.E.J. Newman, Ego-centered networks and the ripple effect. Soc. Networks. **25**, 83 (2003)
44. H. Ebel, L.-I. Mielsch, S. Bornholdt, Phys. Rev. E **66**, 035103(R) (2002)
45. R. Guimera, L. Danon, A. Diaz-Guilera, F. Giralt, A. Arenas, *Self-similar community structure in organisations*, `arXiv:cond-mat/0211498`
46. M. Faloutsos, P. Faloutsos, C. Faloutsos, ACM SIGCOMM '99, Comput. Commun. Rev. **29**, 251 (1999)
47. R. Albert, H. Jeong, A.-L. Barabási, Nature **401**, 130 (1999)
48. B.A. Huberman, L.A. Adamic, Nature **401**, 131 (1999)
49. S.L. Pimm, J.H. Lawton, J.E. Cohen, Nature **350**, 669 (1991)
50. R.T. Paine, Nature **355**, 73 (1992)
51. J. Camacho, R. Guimerà, L.A.N. Amaral, Phys. Rev. E **65**, 030901(R) (2002)
52. J. Camacho, R. Guimerà, L.A.N. Amaral, Phys. Rev. Lett. **88**, 228102 (2002)

53. J.A. Dunne, R.J. Williams, N.D. Martinez, P. Natl. Acad. Sci. USA **99**, 12917 (2002)
54. C. Koch, G. Laurent, Science **284**, 96 (1999)
55. L.F. Lago-Fernandez, R. Huerta, F. Corbacho, J.A. Siguenza, Phys. Rev. Lett. **84**, 2758 (2000)
56. U. Alon, M.G. Surette, N. Barkai, S. Leibler, Nature **397**, 168 (1999)
57. R. Milo, S. Shen-Orr, S. Itzkovitz et al., Science **298**, 824 (2002)
58. Z.N. Oltvai, A.-L. Barabási, Science **298**, 763 (2002)
59. H. Jeong, S.P. Mason, A.-L. Barabási, Z.N. Oltvai, Nature **411**, 41 (2001)
60. A.C. Gavin et al., Nature **415**, 141 (2002)
61. Y. Ho et al., Nature **415**, 180 (2002)
62. H. Jeong, B. Tombor, R. Albert et al., Nature **407**, 651 (2000)
63. A. Wagner, D.A. Fell, P. Roy. Soc. London B-Biol. Sci. **268**, 1803 (2001)
64. E. Ravasz, A.L. Somera, D.A. Mongru et al., Science **297**, 1551 (2002)
65. L. Euler, *Solutio problematis ad geometriam situs pertinentis* (1736)
66. *The Small World*, edited by M. Kochen (Norwood, NJ, Ablex, 1989)
67. S. Milgram, Psychol. Today **1**, 61 (1967)
68. J. Travers, S. Milgram, Sociometry **32**, 425 (1969)
69. B. Bollobás, *Random Graphs* (London, Academic Press, 1985)
70. D.J. Watts, S.H. Strogatz, Nature **393**, 440 (1998)
71. M. Barthélemy, L.A.N. Amaral, Phys. Rev. Lett. **82**, 3180 (1999)
72. A. Barrat, M. Weigt, Eur. Phys. J. B **13**, 547 (2000)
73. A.-L. Barabási, R. Albert, Science **286**, 509 (1999)
74. P.O. Seglen, J. Am. Soc. Inf. Sci. **43**, 628 (1992)
75. S. Redner, Eur. Phys. J. B **4**, 131 (1998)
76. R. Cohen, S. Havlin, Phys. Rev. Lett. **90**, 058701 (2003)
77. L.A.N. Amaral, A. Scala, M. Barthélemy, H.E. Stanley, Proc. Nat. Acad. Sci. USA **97**, 11149 (2000)
78. S. Mossa, M. Barthélemy, H.E. Stanley, L.A.N. Amaral, Phys. Rev. Lett. **88**, 138701 (2002)
79. M.E.J. Newman, SIAM Rev. **45**, 167 (2003)
80. S.N. Dorogovtsev, J.F.F. Mendes, *Evolution of Networks: From Biological nets to the Internet and WWW* (Oxford University Press, Oxford, 2003)
81. A. Tanenbaum, *Computer Networks*, 3rd edn. (Prentice Hall, Englewood Cliffs, 1996)
82. D. Comer, *Internetworking with TCP/IP,* Vol. 1, 4th edn. (Prentice Hall, Englewood Cliffs, 2000)
83. P.W. Anderson, K.J. Arrow, D. Pines, *The Economy as an Evolving Complex System*, SFI Studies in the Sciences of Complexity (Addison Wesley Longman, 1988)
84. W. Brian Arthur, S.N. Durlauf, D.A. Lane, *The Economy as an Evolving Complex System II*, SFI Studies in the Sciences of Complexity (Addison Wesley Longman, 1997)
85. A. Bunde, J. Kropp, H.J. Schellnhuber, *The Science of Disasters: Climate Disruptions, Heart Attacks, and Market Crashes* (Springer-Verlag, Berlin, 2002)
86. *Unsolved Problem on Noise 1999*, edited by D. Abbott, L.B. Kish, AIP Conf. Proc. **511** (American Institute of Physics, Melville NY, 2000)
87. J. Postel, *Transmission Control Protocol.* RFC 793 (1981)
88. W.E. Leland, M.S. Taqqu, W. Willinger, D.V. Wilson, IEEE/ACM Trans. Network **2**, 1 (1994)
89. V. Paxson, S. Floyd IEEE/ACM Trans. Network **3**, 226 (1995)
90. M. Crovella, A. Bestavros, IEEE/ACM Trans. Network **5**, 835 (1997)
91. I. Csabai, J. Phys. A **27**, 417 (1994)
92. M. Takayasu, H. Takayasu, T. Sato, Physica A **233**, 924 (1996)
93. M. Takayasu et al., Physica A **277**, 248 (2000)
94. K. Fukuda et al., Physica A **287**, 289 (2000)
95. K. Fukuda et al., in *Traffic and Granular Flow 2001*, edited by Y. Sugiyama et al. (Springer, Heidelberg, 2002)
96. P. Bernaola-Glaván, P.Ch. Ivanov, L.A.N. Amaral, H.E. Stanley, Phys. Rev. Lett. **87**, 168105 (2001)
97. K. Fukuda, L.A.N. Amaral, H.E. Stanley, Europhys. Lett. **62**, 189 (2003)
98. K.A. Eriksen, I. Simonsen, S. Maslov, K. Sneppen, Phys. Rev. Lett. **90**, 148701 (2003)
99. M. Barthélémy, B. Gondran, E. Guichard, Physica A **319**, 633 (2003)
100. J.-J. Pansiot, D. Grad, ACM Comp. Comm. Rev. **28**, 41 (1998)
101. R. Govindan, A. Reddy, in *Proceedings of IEEE INFOCOM'97* (1997), p. 850
102. H. Chou, A *note on power-laws on Internet topology*, `arXiv:cs.NI/0012019` (2000)
103. G. Caldarelli, R. Marchetti, L. Pietronero, Europhys. Lett. **52**, 386 (2000)
104. S.-H. Yook, H. Jeong, A.-L. Barabási, Proc. Nat. Ac. Sci. USA **99**, 13382 (2002)
105. The National Laboratory for Applied Network Research (NLANR), sponsored by the National Science Foundation, provides Internet routing related information based on border gateway protocol data (see `http://moat.nlanr.net/`)
106. R. Pastor-Satorras, A. Vázquez, A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001)
107. A. Vázquez, R. Pastor-Satorras, A. Vespignani Phys. Rev. E **65**, 066130 (2002)
108. J. Camacho, R. Guimerà, D.B. Stouffer, L.A.N. Amaral, *Quantitative patterns in the structure of model and empirical food webs*, Ecol. Mon. (submitted, 2004)
109. D. Garlaschelli, G. Caldarelli, L. Pietronero, Nature **423**, 165 (2003)
110. R.M. Anderson, R.M. May, *Infectious Diseases in Humans* (Oxford University Press, Oxford, 1992)
111. R. Pastor-Satorras, A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001)
112. R. Pastor-Satorras, A, Vespignani, Phys. Rev. E **65**, 036104 (2002)
113. D.J. Watts, Proc. Nat. Ac. Sci. USA **99**, 5766 (2002)
114. R. Cohen, S. Havlin, D. ben-Avraham, *Efficient Immunization Strategies for Computer Networks and Populations*, `arXiv:cond-mat/0207387`
115. Z. Dezso, A.-L. Barabási, Phys. Rev. E **65**, 055103 (2002)
116. V.M. Eguíluz, K. Klemm, Phys. Rev. Lett. **89**, 108701 (2002)
117. A. Vazquez et al., Phys. Rev. E **67**, 046111 (2003)
118. L.S. Liebovitch, I.B. Schwartz, Phys. Rev. E **68**, 017101 (2003)
119. M.E.J. Newman, S. Forrest, J. Balthrop, Phys. Rev. E **66**, 035101 (2002)
120. B. Vogelstein, D. Lane, A.J. Levine, Nature **408**, 307 (2000)

121. R. Albert, H. Jeong, A.-L. Barábasi, Nature **406**, 378 (2000)
122. A. Broder et al., Computer Networks **33**, 309 (2000)
123. R. Cohen, K. Erez, D. Ben-Avrham, S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000)
124. D.S. Callaway, M.E.J. Newman, S.H. Strogatz, D.J. Watts, Phys. Rev. Lett. **85**, 5468 (2000)
125. R. Cohen, K. Erez, D. Ben-Avrham, S. Havlin, Phys. Rev. Lett. **86**, 3682 (2001)
126. P. Holme, B.J. Kim, Phys. Rev. E **65**, 066109 (2002)
127. P. Holme, Phys. Rev. E **66**, 036119 (2002)
128. T. McMahon, Science **179**, 1201 (1973)
129. G.B. West, J.H. Brown, B.J. Enquist, Science **276**, 122 (1997)
130. J.R. Banavar, A. Maritan, A. Rinaldo, Nature **399**, 130 (1999)
131. L.P. Kadanoff, Physics Today **54**, 34 (2001)
132. J.M. Ottino, Nature **427**, 399 (2004)