# Virtual Round Table on ten leading questions for network research

*The following discussion is an edited summary of the public debate started during the conference "Growing Networks and Graphs in Statistical Physics, Finance, Biology and Social Systems" held in Rome in September 2003. Drafts documents were circulated electronically among experts in the field and additions and follow-up to the original discussion have been included. Among the scientists participating to the discussion L.A.N. Amaral, A. Barrat, A.L. Barabasi, G. Caldarelli, P. De Los Rios, A. Erzan, B. Kahng, R. Mantegna, J.F.F. Mendes, R. Pastor-Satorras, A. Vespignani are acknowledged for their contributions and editing.*

The last few years have witnessed a tremendous activity devoted to the characterization and understanding of networked systems. Indeed large complex networks arise in a vast number of natural and artificial systems. Ecosystems consist of species whose interdependency can be mapped into intricate food webs. Social systems may be represented by graphs describing various interactions among individuals. The Internet and the World-Wide-Web (WWW) are prototypical examples of self-organized networks emerging in the technological world. Large infrastructures such as power grids and the air transportation network are critical networked systems of our modern society. Finally, the living cell is not an exception either, its organization and function being the outcome of a complex web of interactions among genes, proteins and other molecules.

For a long time all these systems have been considered as haphazard set of points and connections, mathematically framed in the random graph paradigm. This situation has radically changed in the last five years, during which the study of complex networks has received a boost from the ever-increasing availability of large data sets and the increasing computer power for storage and manipulation. In particular, mapping projects of the WWW and the physical Internet offered the first chance to study the topology of large complex networks. Gradually, other maps followed describing many networks of practical interest in social science, critical infrastructures and biology. Researchers thus have started to have a systematic look at these large data sets, searching for hidden regularities and patterns that can be considered as manifestations of underlying laws governing the dynamics and evolution of these complex systems. Indeed, when studying the structure of complex networks, one finds out that in spite of the apparent complexity and randomness of the underlying systems, clear patterns and regularities emerge, which can be expressed in mathematical and statistical fashion.

Specifically, many of these systems show the small-world property, which implies that in the network the average topological distance between the various nodes increases very slowly with the number of nodes (logarithmically or even slower). A particularly important finding is the realization that many networks are characterized by the statistical abundance of "hubs"; i.e. nodes with a large number of connections to other elements. This feature has its mathematical roots in the observation that the number of elements with $k$ links follows a power-law distribution, indicating the lack of any characteristic scale. This has allowed the identification of the class of scale-free networks whose topological features turn out to be extremely relevant in assessing the physical properties of the system as a whole, such as its robustness to damages or vulnerability to malicious attack.

The attempt to model and understand the origin of the observed topological properties of real networks has led to a radical change of perspective, shifting the focus from static graphs, aiming to reproduce the structure of the network in a certain moment, to modeling network evolution. This new approach is the outcome of the realization that most complex networks are the result of a growth process. As a result, we currently view networks as dynamical systems that evolve through the subsequent addition and deletion of vertices and edges. The set of dynamical rules defining these processes thus outlines the dynamical theory required in order to understand the macroscopic properties of networks. This methodology that is akin to the statistical physics approach to complex phenomena appears as a revolutionary

path in our understanding of networked systems and provides new techniques to approach conceptual and practical problems in this field.

While the advances that we have witnessed in the past few years were truly amazing, both in its impact on basic science and practical implications, they have highlighted the incompleteness of our knowledge as well. We are therefore in the position to ask a series of important questions that require the community's attention. Our goal here is to formulate some of these questions, offering a loose guide to the community and ourselves. We should emphasize that we are aware of the fact that advances in all sciences are often induced by the ability of its practitioners to ask novel questions. Thus these questions should by no means be seen as a way to limit new ideas, or to channel our thinking into narrowly defined directions. We feel, however, that formulating these questions would offer a valuable benchmark to both practitioners and those interested in network theory, thus it is worthwhile elaborating on them.

**- Are there formal ways of classifying the structure of different growing models?**

Many networks models, or classes of models, have been recently formulated and empirically studied by numerical simulations or approximate analytical methods. While this corresponds to a great advancement in the modeling and representation of networks, a rigorous understanding of the topology of these models is far less developed. Questions concerning the universality of some topological properties, the correlations introduced by the dynamical process and the interplay between clustering, hierarchies and centrality in networks are still only partially answered. A full and general understanding of this issue amounts to the development of rigorous methods to uncover the mathematical structure of growing networks.

**- Are there further statistical distributions that can provide insights on the structure and classification of complex networks?**

In the study of complex networks a definite set of statistical distributions and observable quantities are customarily used in the real data analysis and in model characterization and validation. These distributions usually rely on the analysis of the network's degree spectrum. The degree distribution $P(k)$ of vertices, the clustering coefficient and the degree-degree correlations expressed as the joint probability $P(k, k')$ of having an edge between two vertices of degree $k$ and $k'$ provide a general classification of the connectivity pattern properties. The ever-increasing evidence in networks for the presence of communities, motifs and modular ordering, however, calls for the developing of new ways to quantitatively characterize these features in precise mathematical terms.

**- Why are most networks modular?**

The hierarchical nature of networks goes hand in hand with the existence of a modular architecture. In this perspective, networks can be fragmented in different groups of interconnected elements, or modules, each one being responsible for different functions, quantitatively identified by highly interlinked communities of elements. Modules can be repeated at different hierarchical levels and interconnected via the "hubs" of the system. How modularity emerges across many different networks and how it can be reconciled with the other properties of networks are basic questions of network theory.

**- Are there universal features of network dynamics?**

Networks are not only specified by their topology but also by the dynamics of information or traffic flow taking place along the links. For instance, the heterogeneity in the intensity of connections may be very important in understanding social systems. Analogously, the amount of traffic characterizing the connections of communication systems or large transport infrastructure is fundamental for a full description of these networks. The final aim is a mathematical characterization that might uncover very general principles describing the networks' dynamics.

**- How do the dynamical processes taking place on a network shape the network topology?**

The network provides the substrate on top of which the dynamical behavior of the system must unfold. At the same time, however, the various dynamical processes are expected to affect the network's evolution. Dynamics, traffic and the underlying topology are therefore mutually correlated and it is very important to define appropriate quantities and measures capable of capturing how all these ingredients participate in the formation of complex networks. To carry out this task, we need to develop large empirical datasets that simultaneously capture the topology of the network and the time-resolved dynamics taking place on it.

**- What are the evolutionary mechanisms that shape the topology of biological networks?**

While in technological and large infrastructure networks it is possible to uncover the fundamental dynamical rules governing the network evolution, in biological systems this task is much more difficult. In particular, the role of evolution and selection in shaping biological networks (with emphasis on cell biology) is still unclear, especially if we want a quantitative dynamical implementation of evolutionary principles in network modeling.

**- How to quantify the interaction between networks of different character (networks of networks)?**
Most networks are interconnected among them forming networks of networks. This is the case of the Internet, where diverse networks interconnect with each other and communicate through a common protocol. In more complicated situations, however, the interconnected networks are very different in nature and their interactions and reciprocal influence have been completely ignored. For instance, the energy and power distribution networks represent a physical layer over which many critical infrastructures, such as information systems, transportation networks and other public services are lying. In the biological world, the gene network is interconnected with the protein-protein interaction network and the metabolic network. The understanding and characterization of the complicate set of regulatory and feedback mechanisms connecting various networks is probably one of the most ambitious tasks in network research. In this case we are also in need of novel quantities and mathematical tools tailored to describe and model the structure of networks of networks.

**- Is it possible to develop tools to address in a systematic fashion the robustness and vulnerability of large technological and infrastructural networks?**
Complex networks react in different ways to different perturbations. In general they are robust to random damages but weak to attacks targeting some key elements of the system. Furthermore, networks have a dynamics as well as structure. For instance, in power grids, which transport energy, the failure of one component increases the burden on other elements, potentially overloading them, and disrupting their functions. In this way failures may cascade through the network, causing far more disruption that one would expect from the initial cause. A systematic theory of network resilience and robustness needs to address both local (individual failures) and global vulnerabilities (cascading failures).

**- How to characterize small networks?**
Many real world networks are far from being large scale objects well described by statistical measures. While statistical tools may still be used to characterize ensembles of these objects, we also need to look for a detailed description of their particular structure. Concepts such as scaling behavior or average properties are not well defined for a small network and new unifying concepts and mathematical tools need to be devised.

**- Why are social networks all assortative, while all biological and technological networks disassortative?**
A series of recent measurements indicate the social networks are fundamentally different from all other networks: while in social networks there is a tendency for the hubs to be linked together (assortativity), in biological and technological network the hubs show the opposite tendency, being primarily connected to less connected nodes. On one end it is not clear if this is a universal property (the measurements do seem indicate this, however). On the other end, the origin of this difference is not understood. Is there a generic explanation for the observed pattern, or does it represents a feature that needs to be addressed in each network individually?

The above list represents an attempt to limit ourselves to ten questions that we consider as potentially playing an important role in the field at this point. Their order does not reflect their relative importance, but a gradual shift from the more theoretical to the predominantly practical and applied problems. The list cannot be exhaustive and naturally it should be seen just as a starting point, highlighting some of the challenges in front of us. They have been collected through discussions and debates with several colleagues, thus they may be considered as expressing a (albeit inherently imperfect) consensus of the network community.