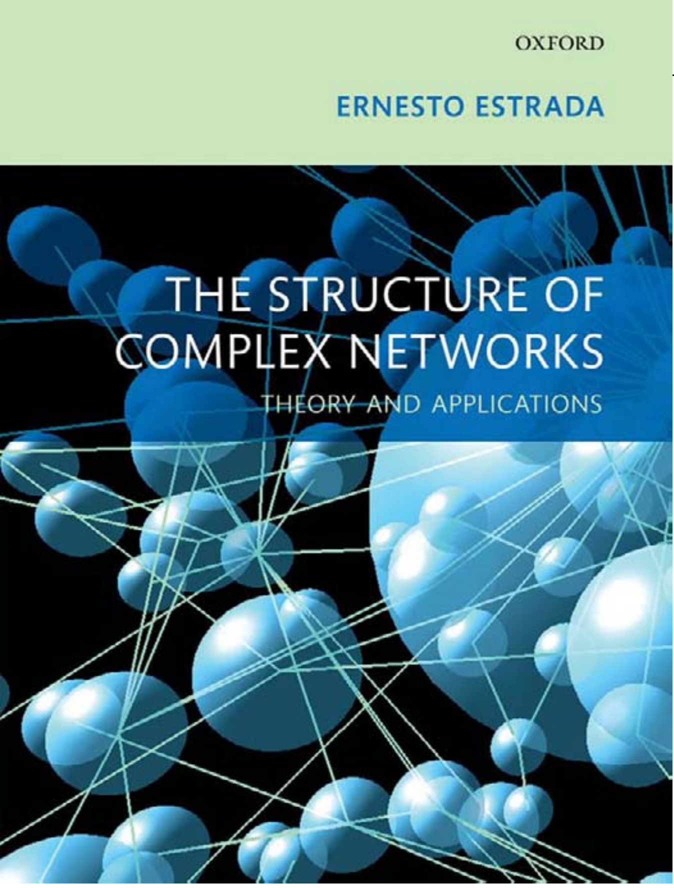
OXFORD
ERNESTO ESTRADA
THE STRUCTURE OF COMPLEX NETWORKS
THEORY AND APPLICATIONS

THE STRUCTURE OF COMPLEX NETWORKS

This page intentionally left blank

The Structure of Complex Networks

Theory and Applications

ERNESTO ESTRADA

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

First published 2011

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data available

Typeset by SPI Publisher Services, Pondicherry, India
Printed in Great Britain
on acid-free paper by
CPI Group (UK) Ltd, Croydon, CR0 4YY

ISBN 978–0–19–959175–6

10 9 8 7 6 5 4 3 2 1

To Gissell, Doris, and Puri

This page intentionally left blank

Preface

"If you want to understand function, study structure".
F. H. C. Crick

The idea of writing this book arose in Venice in the summer of 2009, while I was attending the scientific conference NetSci09. This international workshop on network science was mainly attended by physicists who were studying the structure and dynamics of complex networks. I had previously attended a workshop in Rio de Janeiro devoted to spectral graph theory and its applications, which was mainly attended by mathematicians. Although the two areas are very much overlapped, the language, approaches, and results presented there were almost disjointed. It is frequently read that the study of complex networks is a truly interdisciplinary research field. However, it is hard to find a book that really reflects this character, unifying the languages and results used in graph theory and network science. The situation is even more complicated if we consider that many practitioners of network science work in application areas such as chemistry, ecology, molecular and cellular biology, engineering, and socio-economic sciences. The result of this *melting pot* is that many results are duplicated in different areas, usually with different denominations, and more critically there is a lack of unification of the field to form an independent scientific discipline that can be properly called 'network science'.

This book is a modest contribution to the development of a unified network theory. I have tried to express in a common language the results from *algebraic, topological, metrical*, and *extremal* graph theory, with the concepts developed in statistical mechanics and molecular physics, and enriched with many ideas and formulations developed in mathematical chemistry, biology, and social sciences. All these concepts are first mathematically formulated, and then explained and illustrated with simple examples from artificial and real-world networks. This comprises the first part of this volume, which is specifically devoted to the development of network theory. The second part is dedicated to applications in different fields, including the study of biochemical and molecular biology networks, anatomical networks, ecological networks, and socioeconomic networks. The idea of these chapters is not to describe a few results previously reported in the literature but rather to critically analyse the role of network theory in tackling real problems in these fields. Then examples of successful applications of network theory in these fields are illustrated, while erroneous applications and interpretations are critically analysed—the overall idea being to present a unified picture of the field of network science. That is, I consider these application areas as being at the intersection of the scientific disciplines and network theory where the problems arise.

Why study the structure of networks? The quote from Francis H. C. Crick, co-discoverer of the structure of DNA, and Nobel laureate, is self-explanatory.

This is the real ethos of this book. Of course, it is hard to define what we understand by ‘structure’, and a small section is devoted to a discussion of this in the introductory chapter. However, even without a complete understanding of this concept, architects, scientists, and artists have shaped our world as we see it today. Children usually open their toys to ‘see their internal structures’ in an attempt to understand how they work. Knowing the structure of complex networks does not guarantee that we automatically understand how they work, but without this knowledge it is certainly impossible to advance our understanding of their functioning. This book is devoted to the characterisation of the structure of networks based on a combination of mathematical approaches and physical analogies. This characterisation is still incomplete, and more effort is needed in the unification of apparently disconnected concepts. The reader is encouraged to take this route to advance to a definitive theory of complex networks. Finally, I have tried to correct errors and misconceptions found here and there across all fields in which networks are studied, from mathematics to biology. However, I would be pleased to be notified of anything which might require correction.

This book was written mainly in Glasgow (UK), but some parts were written in Remedios (Cuba) and Santiago de Compostela (Spain). Throughout this time I had the support of many colleagues and friends, as well as my family, although mentioning all of them would be an impossible task. However, I would like to express my gratitude to all my past and present collaborators—those who have shared datasets, software, and figures which are used in this book, or who have contributed in some way with one or other part of the process of writing: A. Allendes, U. Alon, N. A. Alves, C. Atilgan, S. R. Aylward, G. Bagler, A.-L. Barabási, V. Batagelj, M. Benzi, A. G. Bunn, D. Bassett, P. A. Bates, M. Boguñá, C. Cagatay Bilgin, P. Chebotarev, J. Crofts, J. A. Dunne, L. da Fontoura Costa, J. de los Rios, K.-I. Goh, L. H. Greene, R. Guimerà, N. Hatano, Y. He, D. J. Higham, E. Koonin, D.-S. Lee, M. Martin, R. Milo, J. Moody, M. E. J. Newman, P. Pereira, A. Perna, E. Ravasz, J. A. Rodríguez-Velazquez, R. Singh, R. V. Solé, O. Sporn, S. J. Stuart, M. Takayasu, H. Takayasu, V. van Noort, E. Vargas, and D. J. Watts. I also want to thank all the members of my family—in particular, Puri—for their infinite patience. Last but not least, I thank S. Adlung, A. Warman, and C. Charles at Oxford University Press for their effective support.

Ernesto Estrada
Glasgow, 3 February, 2011

Contents

Part I Theory 1

1 Introduction 3
1.1 What are networks? 3
1.2 When did the story begin? 4
1.3 A formal definition 6
1.4 Why do we study their structure? 9
1.5 How can we speak their 'language'? 14

2 Adjacency relations in networks 25
2.1 Node adjacency relationship 25
2.2 Degree distributions 27
2.3 Degree–degree correlations 31
2.4 Link adjacency and beyond 36
2.5 Discrete Laplacian operator 37
2.6 Spectral properties of networks 39

3 Metric and topological structure of networks 47
3.1 Shortest path distance 47
3.2 Resistance distance 50
3.3 Generalized network distances 55
3.4 Topological structure of networks 57

4 Fragments (subgraphs) in complex networks 73
4.1 Network 'graphlets' 74
4.2 Network motifs 76
4.3 Closed walks and subgraphs 78
4.4 Counting small undirected subgraphs analytically 79
4.5 Fragment ratios 82

5 Accounting for all parts (subgraphs) 89
5.1 The Estrada index 90
5.2 Parameterised Estrada index 95
5.3 Network entropy and free energy 97
5.4 Network returnability 103
5.5 A brief on matrix functions 106

6 Communicability functions in networks 109
6.1 Network communicability 109
6.2 Communicability in directed and weighted networks 113
6.3 Generalised communicability functions 114
6.4 Communicability based on network vibrations 115

7 Centrality measures 121
7.1 Degree centrality 121
7.2 Centrality beyond nearest-neighbours 125
7.3 Closeness centrality 140
7.4 Betweenness centrality 142
7.5 Information centrality 146

8 Global network invariants 149
8.1 Adjacency-based invariants 149
8.2 Distance-based invariants 159
8.3 Clumpiness of central nodes 161
8.4 Self-similarity of complex networks 166

9 Expansion and network classes 171
9.1 Network expansion 171
9.2 Universal topological classes of network 173
9.3 Topological classification of real-world networks 181

10 Community structure of networks 187
10.1 Network partition methods 190
10.2 Methods based on link centrality 198
10.3 Methods based on modularity 203
10.4 Clustering based on similarity 206
10.5 Communities based on communicability 209

11 Network bipartivity 215
11.1 Bipartivity measures 216
11.2 How do we find bipartitions in networks? 220
11.3 Bipartitions in undirected networks 224
11.4 Bipartitions in directed networks 227

12 Random models of networks 233
12.1 ‘Classical’ random networks 233
12.2 Small-world random networks 238
12.3 ‘Scale-free’ networks 243
12.4 Random geometric networks 250
12.5 Other random models 252

Part II Applications 257

13 Genetic networks 259
13.1 A primer on gene networks 259
13.2 Topological properties of gene coexpression networks 263
13.3 Topological properties of gene regulatory networks 267
13.4 Other gene networks 274

14 Protein residue networks 277
14.1 A primer on protein structure 277
14.2 Protein residue networks 280
14.3 Small-worldness of residue networks 282
14.4 Degree distributions of residue networks 285
14.5 Centrality in residue networks 287
14.6 Global topological properties 295

15 Protein–protein interaction networks 301
15.1 A primer on protein–protein interactions 301
15.2 Global structure of PPI networks 304
15.3 Subgraph-based techniques in PPI networks 311
15.4 Bipartite structures in PPI networks 316
15.5 Local structure of PPI networks 319

16 The structure of reaction networks 325
16.1 Organic chemical reaction networks 326
16.2 Reaction networks in the atmospheres of astronomical bodies 330
16.3 Metabolic networks 333

17 Anatomical networks 345
17.1 Intercellular networks 345
17.2 Vascular networks 353
17.3 Brain networks 355

18 Ecological networks 373
18.1 Global topological properties 374
18.2 Searching for keystone species 383
18.3 Landscape networks 389
18.4 Other network-based ecological studies 395

19 Social and economic networks 399
19.1 Brief on socioeconomic network data 400
19.2 Global topology of social networks 401
19.3 Social network communities 408
19.4 Centrality in social networks 412
19.5 Economic networks 415

Appendix 423

A.1 Brain networks 423
A.2 Ecological networks 423
A.3 Informational networks 424
A.4 PPI networks 424
A.5 Protein residue networks 425
A.6 Social and economic networks 425
A.7 Technological networks 426
A.8 Transcription networks 426
A.9 Software 426

References 429
Index 457

Theory

I

It is the theory which decides what we can observe.

Albert Einstein

This page intentionally left blank

Introduction

1

The impossibility of separating the nomenclature of a science from the science itself, is owing to this, that every branch of physical science must consist of three things; the series of facts which are the objects of the science, the ideas which represent these facts, and the words by which these ideas are expressed. Like three impressions of the same seal, the word ought to produce the idea, and the idea to be a picture of the fact.

Antoine-Laurent de Lavoisier,
Elements of Chemistry (1790)

1.1 What are networks?

The concept of networks is very intuitive to everyone in modern society. As soon as we see this word we think of a group of interconnected items. According to the *Oxford English Dictionary* (2010) the word 'network' appeared for the first time in the English language in 1560, in the Geneva Bible, Exodus, xxvii.4: 'And thou shalt make unto it a grate like networke of brass.' In this case it refers to 'a net-like arrangement of threads, wires, etc.', though it is also recorded in 1658 as referring to reticulate *patterns* in animals and plants (*OED*, 2010). New uses for the word were introduced subsequently in 1839 for rivers and canals, in 1869 for railways, in 1883 for distributions of electrical cables, in 1914 for wireless broadcasting, and in 1986 to refer to the Internet, among many others. Currently, the *OED* (2010) defines a network as an 'arrangement of intersecting horizontal and vertical lines' or 'a group or system of interconnected people or things', including the following examples: a complex system of railways, roads, or other; a group of people who exchange information for professional or social purposes; a group of broadcasting stations that connect for the simultaneous broadcast of a programme; a number of interconnected computers, machines, or operations; a system of connected electrical conductors. The generality of this concept ensures that 'network' is among the most commonly used words in current language. For instance, a search for the word via Google (on 3 August 2010) produced 884 million items. Figure 1.1 illustrates the comparative results of a search for other frequently used words, such as 'food,' 'football,' 'sex', and 'coffee'.

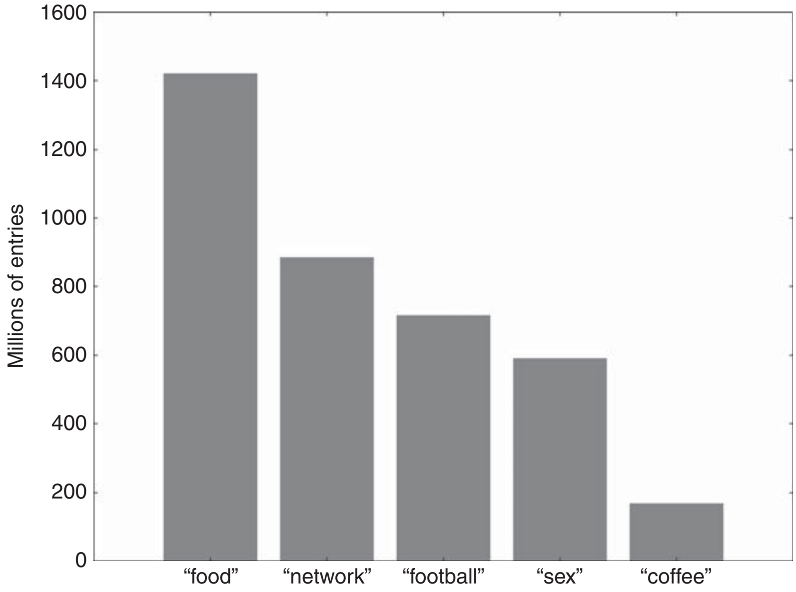

Fig. 1.1
Relevance of the word 'network' in current language. Results of a search via Google (on 3 August 2010) of some popular words in the English language.

According to the previous semantic definitions, the property of being interconnected appears as the most salient 'structural' characteristic of a network. Obviously, in order to have a system of interconnections we need a group of 'entities' and a way by which these items can be connected. For instance, in a system of railways the entities to be connected are stations, and the way they are connected to each other is with railway lines. In general, entities can represent stations, individuals, computers, cities, concepts, operations, and so on, and connections can represent railway lines, roads, cables, human relationships, hyperlinks, and so forth. Therefore, a way of generalizing the concept of networks in an abstract way is as follows.

A *network (graph)* is a diagrammatic representation of a system. It consists of *nodes (vertices)*, which represent the entities of the system. Pairs of nodes are joined by *links (edges)*, which represent a particular kind of interconnection between those entities.

1.2 When did the story begin?

In mathematics the study of networks is known as 'graph theory'. We will use the words 'network' and 'graph' indistinctly in this book, with a preference for the first as a more intuitive and modern use of the concept. The story of network theory begins with the work published by Leonhard Euler in 1736: *Solutio problematic as geometriam situs pertinentis* (*The Solution of a Problem Relating to the Theory of Position*) (Euler, 1736). The problem in question was formulated as follows:

> In Königsberg in Prussia, there is an island A, called the Kneiphof; the river which surrounds it is divided into two branches, as can be seen [in Figure 1.2]. . . and these branches are crossed by seven bridges, a, b, c, d, e, f, and g. Concerning these bridges, it was asked whether anyone could arrange a route in such a way that he would cross each bridge once and only once.

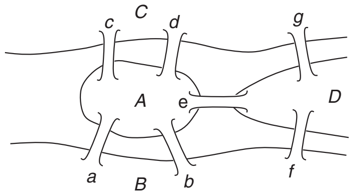

Fig. 1.2
The Königsberg bridges. Schematic illustration of the bridges in Königsberg.

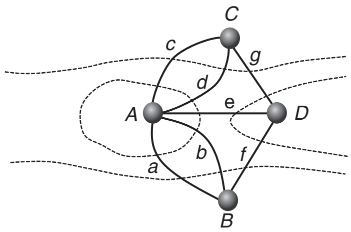

Fig. 1.3
Graph of the Königsberg bridges. Superposition of a multigraph (see further) over the schematic illustration of the bridges in Königsberg.

Despite Euler's solving this problem without using a picture representing the network of bridges in Königsberg, he reformulated it in a way that was the equivalent of what is now referred to as a 'graph'. A representation derived from Euler's formulation of Königsberg's bridges problem is illustrated in Figure 1.3 by continuous lines which are superposed on the scheme of the islands and bridges of Figure 1.2. His paper is therefore recognised as the very first one in which the concepts and techniques of modern graph (network) theory were used. As a consequence of this paper a new branch of mathematics was born in 1736. For an account on the life and work of Euler in particular respect to the Königsberg's bridges problem, the reader is referred to the works of Assad (2007), Gribkovskaia et al. (2007), and Mallion (2007).

The history of graph theory from 1736 to 1936 is well documented by Biggs et al. (1976), and the reader is referred to this book for further details. There is one very important moral to be derived from Euler's original work, which we would like to remark on here. What Euler demonstrated is the existence of certain types of 'real-world' problems that cannot necessarily be solved by using mathematical tools such as geometry and calculus. That is, navigating through Königsberg's bridges without crossing any of them more than once depends neither on the length of the bridges nor on any of their geometric features. It depends only on the way in which the bridges are interconnected—their connectivity. This historical message is important for those who even today deny the value of topological concepts for solving important real-world problems ranging from molecular branching to the position of individuals in their social networks.

A few other historical remarks are added here to provide an idea of the interdisciplinary nature of network/graph theory since the very early days of its birth. The word 'graph' arises in a completely interdisciplinary context. In 1878 the English mathematician James J. Sylvester wrote a paper entitled 'Chemistry and Algebra', which was published in *Nature* (Sylvester, 1877–78). Here the word 'graph' derived from the word used by chemists at that time for the terminology of molecular structure. The word appeared for the first time in the following passage:

> Every invariant and covariant thus becomes expressible by a graph precisely identical with a Kekulean diagram or chemicograph.

Many examples of graph drawing have been collected by Kruja et al. (2002), including family trees in manuscripts of the Middle Ages, and 'squares of opposition' used to depict the relations between propositions and syllogisms in teaching logic, from the same period. One very early application of networks for representing complex relationships was the publication of François Quesnay's *Tableau Économique* (Quesnay, 1758), in which a circular flow of financial funds in an economy is represented with a network. In other areas, such as the study of social (Freeman, 2004) and ecological relationships, there were also important uses of network theory in early attempts at representing such complex systems. For instance, in 1912, William Dwight Pierce published a book entitled *The Insect Enemies of the Cotton Boll Weevil*, in which we find a network of the 'enemies' of the cotton plant which at the same time are attacked by their own 'enemies', forming a directed network in an ecological system (Pierce et al., 1912). Several such fields have maintained a long tradition in the use of network theory, having today entire areas of study such as chemical graph theory (Balaban, 1976; Trinajstić, 1992) social networks theory (Wasserman and Faust, 1994), and the study of food webs (Dunne, 2005).

Lastly, it should be mentioned that the transition from these pioneering works to the modern theory of networks was very much guided and influenced by Frank Harary, who is recognised as the 'father' of modern graph theory. The multidisciplinary character of his work can be considered as pioneering and inspiring because, as written in his obituary in 2005, he 'authored/co-authored more than 700 scholarly papers in areas as diverse as Anthropology, Biology, Chemistry, Computer Science, Geography, Linguistics, Music, Physics, Political Science, Psychology, Social Science, and, of course Mathematics, which brought forth the usefulness of Graph Theory in scientific thought' (Ranjan, 2005). Among his other books are *Graph Theory and Theoretical Physics* (1968), which he edited, *Structural Models in Anthropology* (1984) and *Island Networks* (1996), co-authored with Hage, and his classic *Graph Theory* (1969).

1.3 A formal definition

Although the working definition of networks presented in Section 1.1 accounts for the general concept, it fails in accounting for the different ways in which pairs of nodes can be linked. For instance, every pair of different nodes can be connected by a simple segment of a line. In this case we encounter the *simple network* as illustrated in Figure 1.4 (top left). Links in a network can also have directionality. In this case, a *directed link* begins in a given node and ends in another, and the corresponding network is known as a *directed network* (bottom right of the figure). In addition, more than one link can exist between a pair of nodes, known as multi-links, and some links can join a node to itself, known as *self-loops*. These networks are called *pseudonetworks* (pseudographs) (see Figure 1.4). For more details and examples of these types of network, the reader is referred to Harary (1969).

Finally, links can have some weights, which are real (in general positive) numbers assigned to them. In this case we have *weighted networks* (Barrat et al., 2004; Newman, 2004b). Such weights can represent a variety of concepts in the real world, such as strength of interactions, the flow of information, the number of contacts in a period of time, and so on. The combination of all these types of link is possible if we consider a weighted directed network like that illustrated in Figure 1.5 (bottom right), which is the most general representation of a network.

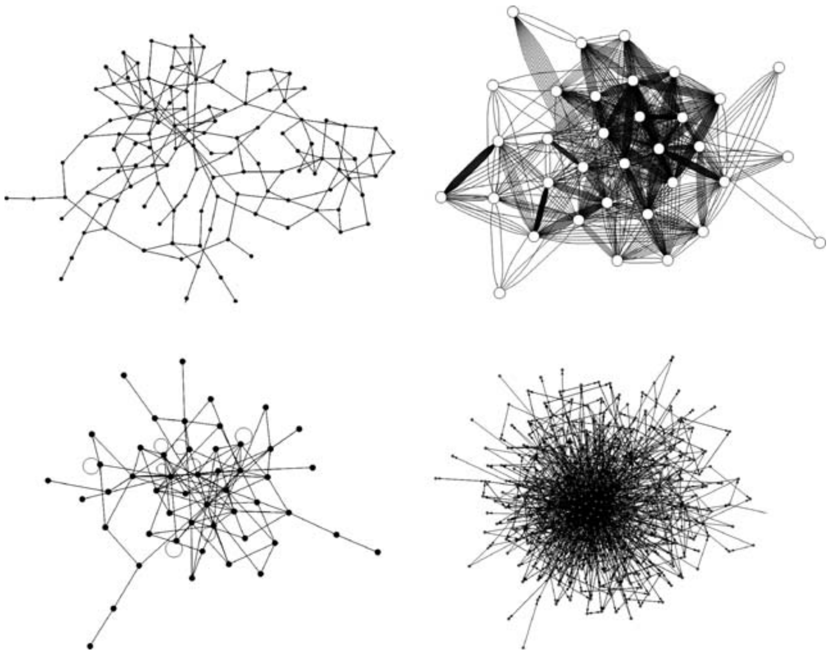

Fig. 1.4
Different kinds of network. (Top left) A simple network representing an electronic circuit. (Top right) A multi-network representing the frequency of interactions in a technical research group at a West Virginia university made by an observer every half-hour during one five-day work-week. (Bottom left) A network with self-loops representing the protein–protein interaction network of the herpes virus associated with *Kaposi sarcoma*. (Bottom right) A directed network representing the metabolic network of *Helicobacter pylori* (connected component only).

With these ingredients we will present some formal definitions of networks. Let us begin by considering a finite set $V = \{v_1, v_2, \cdots, v_n\}$ of unspecified elements, and let $V \otimes V$ be the set of all ordered pairs $[v_i, v_j]$ of the elements of V. A relation on the set V is any subset $E \subseteq V \otimes V$. The relation E is *symmetric* if $[v_i, v_j] \in E$ implies $[v_j, v_i] \in E$, and it is *reflexive* if $\forall v \in V$, $[v, v] \in E$. The relation E is *antireflexive* if $[v_i, v_j] \in E$ implies $v_i \neq v_j$. Then we have the following definition for simple and directed networks (Gutman and Polanski, 1987).

A *simple network* is the pair $G = (V, E)$, where V is a *finite set of nodes* and E is a *symmetric and antireflexive relation* on V. In a *directed network* the relation E is *non-symmetric*.

The previous definition does not allow the presence of multiple links and self-loops. Consequently, we introduce the following more general definition of network.

A *network* is the triple $G = (V, E, f)$, where V is a *finite set of nodes*, $E \subseteq V \otimes V = \{e_1, e_2, \cdots, e_m\}$ is a set of links, and f is a *mapping* which associates some elements of E to a pair of elements of V, such as that if $v_i \in V$ and $v_j \in V$, then $f : e_p \to [v_i, v_j]$ and $f : e_q \to [v_j, v_i]$. A *weighted network* is defined by replacing the set of links E by a set of link weights $W = \{w_1, w_2, \cdots, w_m\}$, such that $w_i \in \Re$. Then, a weighted network is defined by $G = (V, W, f)$.

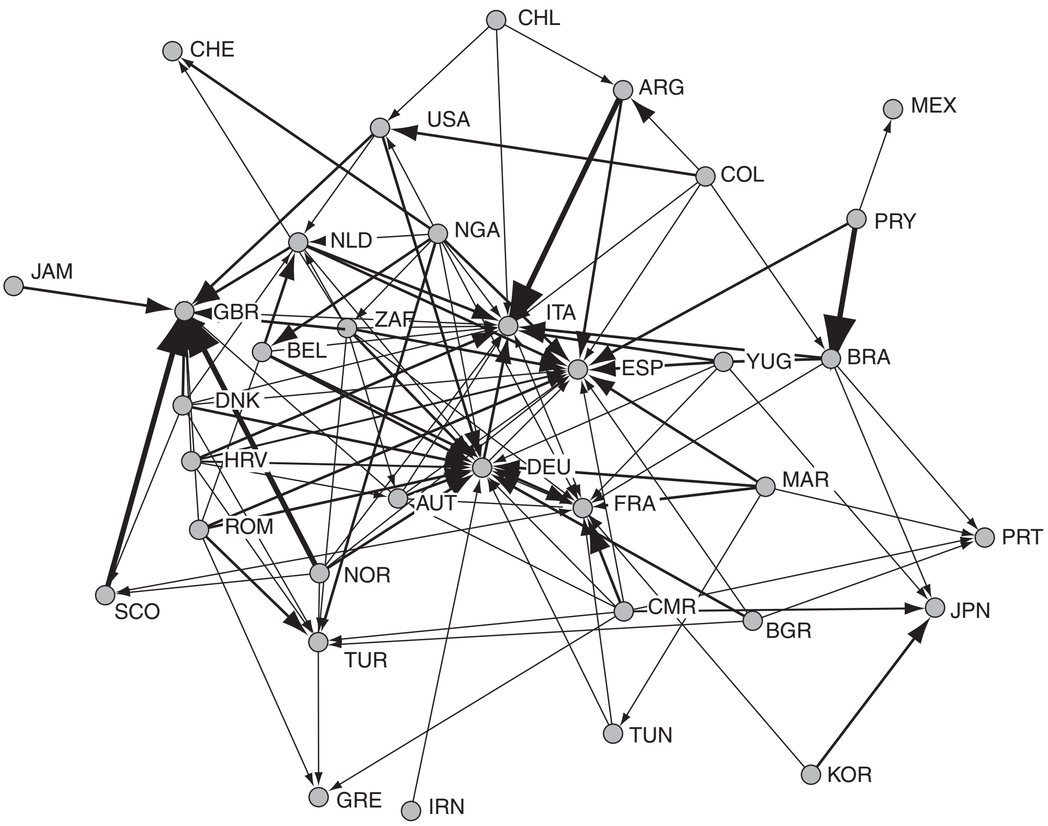

Fig. 1.5
Weighted directed network. Representation of the transfer of football players between countries after the 1998 World Cup in France. Links are drawn, with thickness proportional to the number of players transferred between two countries.

For instance, let $V = \{v_1, v_2, v_3\}$, $E = \{e_1, e_2, e_3, e_4, e_5, e_6\}$, and $f : e_1 \to [v_1, v_1]$, $f : e_2 \to [v_3, v_3]$, $f : e_3 \to [v_1, v_3]$, $f : e_4 \to [v_2, v_1]$, $f : e_5 \to [v_2, v_3]$, $f : e_6 \to [v_3, v_2]$, then we have the network illustrated in Fig. 1.3 (left). If we consider the following mapping:

$$f : 1.5 \to [v_1, v_1]$$

$$f : 0.9 \to [v_3, v_3]$$

$$f : 0.6 \to [v_1, v_3]$$

$$f : 2.1 \to [v_2, v_1]$$

$$f : 0.1 \to [v_2, v_3]$$

$$f : 1.2 \to [v_3, v_2]$$

we obtain the network illustrated in Figure 1.6 (right).

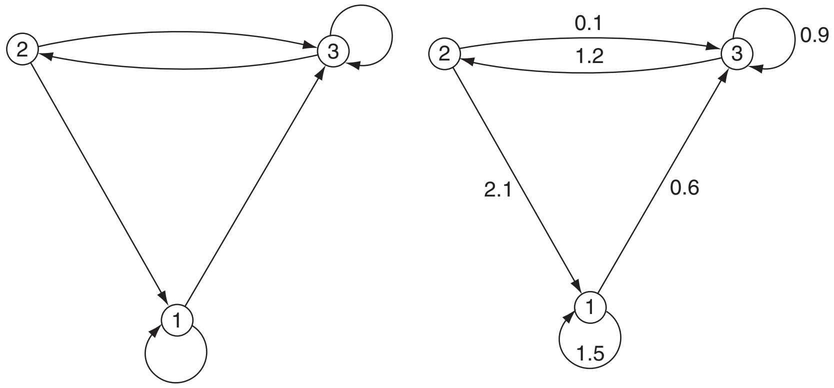

Fig. 1.6
Pseudonetwork and weighted network. The networks produced by using the general definition of networks presented in the text.

1.4 Why do we study their structure?

Complex networks are the skeletons of complex systems. Complex systems are composed of interconnected parts which display soe properties that are not obvious from the properties of the individual parts. One of the characteristics of the structure of these systems is their networked character, which justifies the use of networks in their representation in a natural way. In their essay ‘Models of structure’, Cottrell and Pettifor (2000) have written:

> There are three main frontiers of science today. First, the science of the very large, i.e. cosmology, the study of the universe. Second, the science of the very small, the elementary particles of matter. Third, and by far the largest, is the science of the very complex, which includes chemistry, condensed-matter physics, materials science, and principles of engineering through geology, biology, and perhaps even psychology and the social and economic sciences.

Then, the natural place to look for networks is in this vast region ranging from chemistry to socioeconomical sciences.

As we have seen in the historical account of network theory in this chapter, chemists have been well aware of the importance of the structure of networks representing molecules as a way to understand and predict their properties. To provide an idea we illustrate two polycyclic aromatic compounds in Figure 1.7. The first represents the molecule of benzo[a]pyrene, which is present in tobacco smoke and constitutes its major mutagen with metabolites which are both mutagenic and highly carcinogenic. By replacing the ring denoted as E in benzo[a]pyrene to the place below rings A and B, a new molecule is obtained. This polycyclic aromatic compound is called perylene, and it is known to be non-carcinogenic.

These two isomeric molecules are represented by networks of only 20 nodes, which reduces the ‘complexity’ of their analysis. However, this number can be increased to thousands when biological macromolecules, such as proteins, are represented by networks. The lesson that can be learned from this single example is that the *structure* of a network can determine many, if not all, of the properties of the complex system represented by it. It is believed that network theory can help in many important areas of molecular sciences, including the rational design of new therapeutic drugs (Csermely et al., 2005).

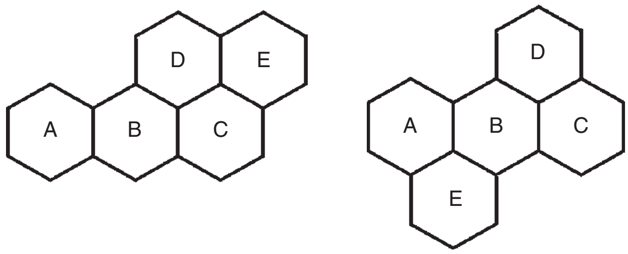

Fig. 1.7
Molecular isomeric networks. Network representation of two polycyclic aromatic compounds. The first corresponds to benzo[a]pyrene—a highly carcinogenic compound found in tobacco smoke. The second corresponds to perylene, which is a non-carcinogenic compound.

According to the *Oxford English Dictionary* (2010), 'structure' is defined as 'the arrangement of and relations between the parts or elements of something complex.' Similarly, the Cambridge Advanced Learner's Dictionary (Walter and Woodford, 2005) defines it as 'the way in which the parts of a system or object are arranged or organized, or a system arranged in this way.' The aim of our book, therefore, is to study the way in which nodes and links are arranged or organized in a network—though initially this appears to be a very easy task which probably does not merit an entire volume. In order to obtain a better idea about our subject, let us take a look at the networks which are depicted in Figure 1.8, which have 34 nodes. The structure of the first one is easily described: every pair of nodes is connected to each other. With this information we can uniquely describe this network, and can obtain precise information about all its properties. In order to describe the second network we have to say that all nodes are connected to two other nodes, except two nodes which are connected to only one node. This also uniquely describes this network, and we can know all its properties. The structure of the third network is more difficult to describe. In this case we can say that every node is connected to three other nodes. However, there are 453,090,162,062,723 networks with 34 nodes that share this property. Despite of this we can say a lot about this network with the information used to describe it. That is, there are many 'structural' properties which are shared by all networks that have the property of having every node connected to the other three. In a similar way, we can refer to the fourth and fifth networks. The fourth one has 34 nodes and 33 links—a type of structure known as a tree, which has many organizational properties in common. The fifth network has 34 nodes and 78 links, and it has been generated by a random process in which nodes are linked with certain probability. If we define such a process—the number of nodes and the linking probability—we can generate networks which asymptotically share many structural properties. We will study these types of network and some of their properties in the various chapters of this book.

Now let us try to describe, in a simple way, the last network, which represents the social interactions between individuals in a real-world system. It has 34 nodes and 78 links, but it does not appear to be as random as the previous one. It is not a tree-like, nor does it display regularity in the number of links per nodes, and it is far removed from the structures represented by the first two networks in Figure 1.8. Then, in order to describe the 'structure' of this

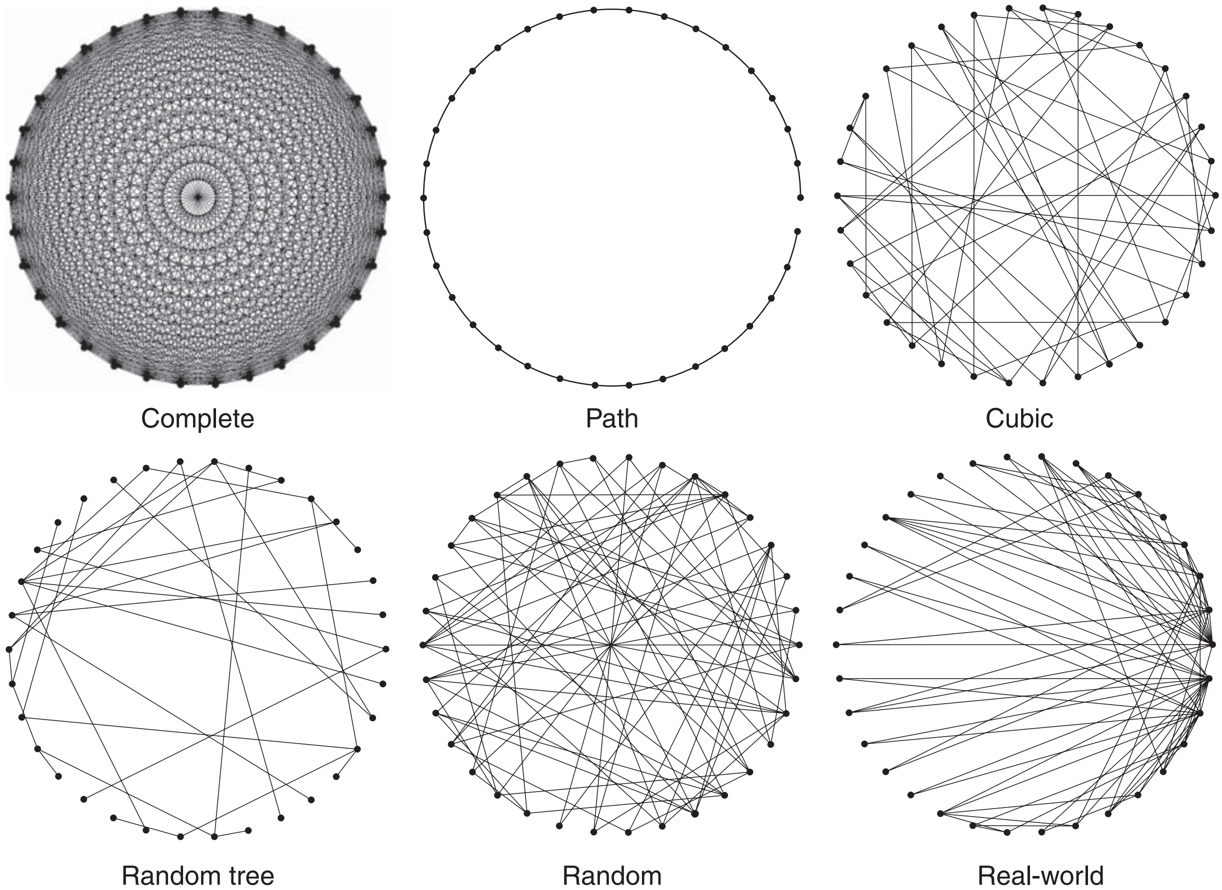

Fig. 1.8
Networks and complexity. Several types of network which can be 'defined' by using different levels of complexity.

network we need to add more and more information. For instance, we can say that the network:

(i) has *exponential degree distribution* and is strongly *disassortative* (see Chapter 2);
(ii) has small *average path length* (see Chapter 3);
(iii) has a high *clustering coefficient* (see Chapter 4);
(iv) has two nodes with high *degree*, *closeness*, and *betweenness centrality* (see Chapter 7);
(*v*) has *spectral scaling* with positive and negative deviations (see Chapter 9);
(vi) has two main *communities* (see Chapter 10);
(vii) displays small *bipartivity* (see Chapter 11), and so on.

The list of properties for this network can be considerably extended by using several of the parameters we will describe in this book, and others that are