1

Introduction to Social Networks

Understanding the link between individual behavior and population-level phenomena is a long-standing challenge in ecology and evolutionary biology (Lima and Zollner 1996; Sutherland 1996). Behavior is expressed as a response to intrinsic and extrinsic factors, including an individual's physical and social environment, the latter made up of nonrandom and heterogeneous social interactions (Krause and Ruxton 2002). That is, individuals are part of a network of inter-individual associations that vary in strength, type, and dynamics. The structure of this social network has far-reaching implications for the ecology and evolution of individuals, populations, and species. For example, the social network supports a diverse array of behaviors that will be influenced by its structure, including: finding and choosing a sexual partner, developing and maintaining cooperative relationships, and engaging in foraging and anti-predator behavior. Such behavior is manifested at the population level in the form of, for example, habitat use, disease transmission, information flow, and mating systems, and forms the basis for evolutionary processes including adapting to changing environments, sexual selection, and speciation. Improving our ability to scale up from the individual to the population by establishing why certain patterns of association develop and how inter-individual association patterns affect population-level structure will revolutionize our understanding of the function, evolution, and implications of social organization.

Across the animal kingdom there is immense diversity in social behavior. Social interactions differ in their type (they might be cooperative, antagonistic, or sexual, for example) as well as their frequency and duration; social bonds may last for years or just minutes or seconds. Which type of interaction occurs and with what frequency and duration will depend on factors such as dominance, body size, sex, age, and parasite load of the participating individuals. This raises the question of how we deal with multiple interactions and complex interaction patterns that can arise even if the number of participants is relatively small. Interestingly, sociologists started addressing this question more than sixty years ago when looking at human interaction patterns, and this literature in combination with recent advances in areas such as statistical physics has provided us with a powerful set of tools for the analysis of animal social networks. These tools make it possible to calculate quantitative metrics describing social structure across different scales of organization, from

the individual to the population. The aim of this book is to explore some of the techniques of network analysis that might be applied to a study of animal social structure.

1.1 WHAT IS A NETWORK?

The essential elements of a network are "nodes" and "edges." In a graphical representation of a network, each node is represented by a symbol, and every interaction (of whatever sort) between two nodes is represented by a line (edge) drawn between them. In the context of a social network, each node would normally represent an individual animal (though see later in this chapter for some alternate approaches) and each edge would represent some measured social interaction or association. For example, figure 1.1 represents the social network for a population of bottlenose dolphins, *Tursiops truncatus*, in New Zealand (Lusseau 2003). In figure 1.1 each filled circle (node) represents an individual dolphin and the connections (edges) between them indicate a certain frequency of social contact over a six-year period. This is the type of network we wish to explore in this book. As we will see as our exploration continues, much of the quantitative analysis of animal social networks is performed not on a graphical representation of interactions but on a matrix of values that conveys the same information. Both the graph and the matrix are representations of the same network.

It should not be a surprise to learn that there are many systems, in many walks of life, that can be thought of as a collection of pair-wise connections between objects. Some types of network are very familiar. Probably all of us regularly tap into a telephone network on which we may simply and quickly be connected to pretty much anywhere in the world without giving it much thought. Other technological systems such as electrical power grids (Xu et al. 2004), transport systems (Sen et al. 2003), and the World Wide Web (Tadic 2001) are all quite naturally considered as a network.

Many people have discovered that network theory may provide novel insight into the local and global properties of a system of interconnected objects that is not possible from considering either the interactions between pairs of agents in isolation or from a study of the average properties of the system as a whole. This has lead to researchers studying networks across a range of systems to gain understanding both of their structure and of some of the consequences of this structure. For example, applications of network theory to technological systems include optimizing the efficiency of telephone communication systems and analyzing the vulnerability of power grids to the loss of a power station.

Mathematicians and statistical physicists have made important contributions to the network literature, providing concrete results on the properties of

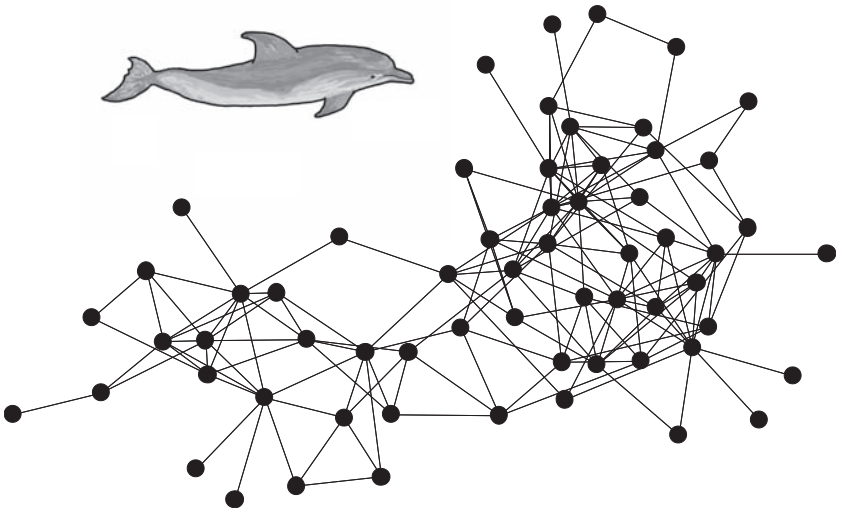

Figure 1.1. Social network of a bottlenose dolphin (*Tursiops truncatus*) population in New Zealand (Lusseau 2003). The network consists of *nodes* (drawn as filled circles here) that denote individuals, and *edges* (straight lines) that symbolize some form of relation. In this network, an edge is drawn between two animals if they were seen together more often than expected by chance. A total of 64 adult dolphins are interconnected by 159 edges.

certain large networks with random assignments of edges to nodes (Erdös and Rényi 1959; Bollobás 1985) and unearthing new paradigms for the characterization of the structure of complex networks and some of the processes that might occur on them (for excellent reviews of the world of networks from a physics perspective, see Albert and Barabási 2002; Newman 2003a; Boccaletti et al. 2006).

The networks approach has also been embraced by biologists interested in unraveling the interplay between cell function and the intricate web of interactions between genes, proteins, and other molecules involved in the regulation of cell activity. They are developing a general framework in which the biological functions of a cell can be understood by examining the structure of its interacting components (Kollmann et al. 2005), enabling them to move beyond "parts lists" of a system and to understand how its components interact to produce complex patterns and behaviors (Jasny and Ray 2003). For example, networks have been used to understand how selective forces have acted on the function of metabolic pathways (Rausher, Miller, and Tiffin 1999) and how gene regulatory networks shape patterns of development (von Dassow et al. 2000; MacCarthy, Seymour, and Pomiankowski 2003). A similar approach

has been applied at other levels of organization (Proulx Promislow, and Phillips 2005; May 2006). For example, biologists have investigated how cells and organs interact by studying neuronal networks (e.g., Laughlin and Sejnowski 2003), and have considered the structure and stability of ecological systems by plotting trophic interactions between species in the form of a food web (Sole and Montoya 2001; Dunne, Williams, and Martinez 2002). By comparison, relatively few biologists have built and analyzed animal social networks. We will of course be discussing their work throughout the book.

All this wealth of interest from various parties is both a good thing and a potentially bad thing for the budding network analyst who wishes to construct and analyze the structure of an animal social network. On the plus side, there are now many methods and measures that might be brought to bear, and many sources of novel methodology. On the minus side, it must be realized that some of the methods and results derived by one community of researchers do not necessarily translate directly to the analysis of all networks. For example, many of the results obtained by the mathematics and physics community are only applicable for networks with a very large number of nodes. A network such as the internet has so many nodes that its statistical properties may very accurately be approximated by some of the statistical models developed in the physical sciences. Looking for the same properties in a network with a few tens of nodes may, on the other hand, not be so well advised. The type of data that is collected can also have a large effect on how one should go about analyzing a network. Sen et al. (2003) studied the network structure of the Indian railway system. Here the edges represent physical connections (train tracks) between stations, so we can be reasonably confident that the network is an accurate representation of the real system. In contrast, social animals often live in fission–fusion societies (Krause and Ruxton 2002) and their social networks have to be inferred from observations of interactions been individuals or within groups of individuals. This creates a number of methodological issues associated with the sampling effort required to get a representative picture of the "real" network structure. It is essential that we take such factors into consideration when we are exploring and analyzing animal social networks.

1.2 SOCIAL NETWORKS AND RELATED METHODS

Social network theory has its origins in a number of different fields of research on humans. It goes back to the work of psychologists and sociologists in the 1930s who applied elements of mathematical graph theory to human relationships (e.g., Moreno 1934; Lewin 1951), and has mostly been concerned with the scenario in which each node represents a single person and each edge some interaction or relationship between two people. A great deal of progress has been made in the analysis and modeling of human social networks in the past

twenty or thirty years, made possible by the advent of readily available and cheap computing power, which enables randomization tests and other simulation techniques to calculate more sophisticated measures of social structure and to bring some much needed statistical rigor to the field. The books by Wasserman and Faust (1994), Scott (2000), and Carrington, Scott, and Wasserman (2005) provide an excellent account, from various angles, of many of the methods that have arisen in the social sciences, and we will refer to these sources frequently throughout the book. In recent times social network theory has also received important impulses from the physics community, which have contributed a number of important theoretical advances such as the small-worlds concept (Watts and Strogatz 1998), algorithms for community detection in networks, and qualitatively new insights into the spread of information through populations (Boccaletti et al. 2006).

Network theory provides a formal framework for the study of complex social relationships. Human social networks have been used to investigate a range of topics. These include the spread of HIV (Potterat et al. 2002), the interconnectedness of company boards of directors (Battiston, Weisbuch, and Bonabeau 2003; Battiston and Catanzaro 2004), and the spread of rumors (Moreno, Nekovee, and Pacheco 2004). In contrast to studies on human social networks, the use of network theory to study the social organization of animal groups or populations is still relatively uncommon (Sade et al. 1988; Connor, Heithaus, and Barre 1999; Fewell 2003; Lusseau 2003; Croft, Krause, and James 2004a; Cross et al. 2004; Flack et al. 2006). Perhaps not surprisingly, some of the earliest applications of ideas developed to study human social networks to other animals came in primatology, though such studies generally did not involve statistical validation of the observed patterns (Sade and Dow 1994). More recent studies (Lusseau 2003; Croft, Krause, and James 2004a) have compared quantitative network measures against null models, or used methods inspired by developments in network theory from the mathematics and physics literature (Lusseau and Newman 2004; Wolf et al. 2007) to relate heterogeneities in animal network structure to the biology of their system. However, despite the vast number of studies in the animal behavior literature that have collected information on interactions or associations between pairs of animals, very few investigations have used a network approach to analyze them. We believe that network theory may offer an exciting method to analyze both new and old data sets, which could provide insights into the structure of animal societies not possible with traditional methods.

At this point it seems appropriate to deal with any nagging doubts in the minds of some readers that this is all something you have seen before, just dressed up in new terminology. Surely, you might be thinking, the matrix of pair-wise interactions is nothing more than an association matrix (Whitehead 1997), and its visualization a sociogram (e.g., Zimen 1982; Sade 1989). Don't we already look for collections of closely associated animals in an association

matrix using cluster analysis such as Ward's or unweighted pair group method with arithmetic mean (UPGMA) methods (Whitehead 1999)? Well, the answer is a simple yes. An association matrix and a sociogram are indeed different names for a social network. So what is new, then?

The principal advantage, as we see it, of using the network approach to probe the structure of animal societies is that it allows us to tap into a very wide range of measures and approaches that are, as we have hinted, still being developed in parallel in several disciplines, and to apply these all under the umbrella of a single description of the data and the associated analytic tools. Thus we might learn tricks and methods from all manner of sources that might help us unravel what the important structural elements are in our animal social system, and what biological or other factors might be driving that structure. Of course, the real appeal of any approach that amalgamates many interactions is that in principle we can probe structure on all scales from the individual to the population. We then need robust measures that describe the properties of individuals, communities, and populations; our belief is that there are many methods for achieving this that come under the networks umbrella, that just happen to have been developed in the social or physical sciences. In addition, the analysis and visualization of social connections can often be rolled into a single computer program. Network theory therefore offers an "all in one" package that allows us to move between different levels of social complexity, and to tap into new analytic tools.

As we have already mentioned, the vast majority of social networks employ one node to represent a single individual, and each edge represents some form of interaction or association between two individuals (see figure 1.1). Furthermore, each edge in a social network represents the same type of interaction or association. The animal social networks we will analyse for most of this book fall into this category. Many of the systems we will study in this book are "fission–fusion" societies, in which animals frequently leave or join groups (Krause and Ruxton 2002); examples include species of ungulates, primates, cetaceans, fish, and insects. To investigate the fine-scale structure of social networks in such systems, we need to be able to identify individual animals. However, for some species this is not possible (or too time consuming) due to problems associated with identifying, capturing, or recapturing individuals. In such instances it can be useful to identify categories of individuals and consider interactions between them (rather than the individuals themselves). We will illustrate now how these interactions can be considered as a network.

In a study of a captive wolf pack (*Canis lupus*), Zimen (1982) made observations on social interactions in a 6-hectare enclosure over a 10-year period (figure 1.2). During this time juveniles matured into adults and the social status of individuals changed; for example the rank of alpha male was occupied by six different animals. The study looked at a number of different behaviors,

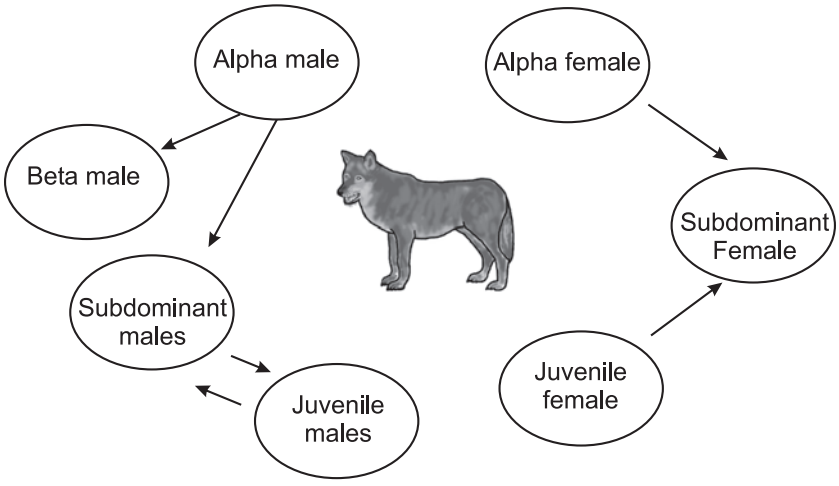

Figure 1.2. Sociogram of "following" events in a captive wolf pack (redrawn from Zimen [1982]). A total of 49 individuals were monitored over a 10-year period during which the social status of individuals changed. The sociogram summarizes the aggressive "following" interactions between individuals belonging to different social categories; arrows point toward the class of animal forced to keep its distance.

one of which was termed "following," which involves a dominant individual forcing a subdominant to keep a certain distance.

Figure 1.2 depicts the "following" behavior in terms of a network where nodes represent classes of animals, not individuals. It clearly shows that following occurs within, but not between, the sexes. Note also that the wolf network in figure 1.2 has directed interactions (i.e., the edges connecting the nodes in the network have arrows that illustrate "who followed whom"). For example juvenile females followed subordinate females but subordinate females did not follow juvenile females, so the interaction is represented with a directed edge going from juvenile females to subordinate females. Based on the information in the network, we could now formulate hypotheses to explore further the social behavior of wolves. For instance, we could ask whether this type of sexually segregated behavior is generally the case with aggressive interactions, or is specific to "following" behavior.

Another alternative approach to network construction that avoids the need for individually recognizable animals is to use nodes to represent behaviors in the population and categorize the individuals that are important for regulating the behaviors as the edges. Fewell (2003) adopted this approach to construct a network (figure 1.3) of the control of pollen foraging in a colony of honeybees (*Apis melifera*). Visualizing behaviors as a network makes it possible to

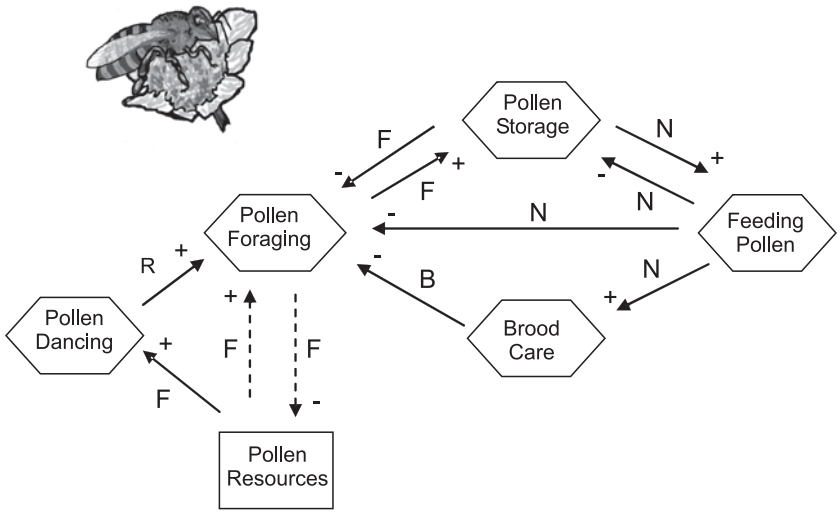

Figure 1.3. A process-oriented network depicting pollen-foraging behavior in the honeybee (redrawn from Fewell [2003]). In this network the nodes represent tasks within the colony and edges between them are the individuals that transmit the information: F, forager; N, nurse; B, brood; R, recruits. Both positive (+) and negative (−) feedback are indicated.

investigate how each caste in the colony contributes to overall system function. Fewell's (2003) network explains the modulation of foraging behavior within the colony. As foragers place pollen in the pollen cells, they receive information on the amount of pollen stored, which feeds back negatively on pollen foraging (i.e., the less pollen stored, the more foraging occurs). Nurse bees remove the pollen from the cells and feed it to the developing brood, and give access to the pollen foragers. Receiving access to pollen from a nurse bee feeds back negatively on foraging behavior (i.e., the more feeding, the less foraging occurs). When hunger levels in the brood are high they produce a hunger pheromone that encourages foraging behavior; brood care reduces hunger levels. Finally pollen dancers provide information on pollen availability and location, and dancing also elicits recruitment to foraging by workers that are not actively engaged in foraging.

1.3 OUR MOTIVATION FOR WRITING THIS BOOK

There are several reasons why we think it is timely to write a book exploring networks from the perspective of researchers working with groups of individually recognizable animals. First and foremost, we believe that a networks

approach has great potential as a means to perform a range of quantitative analyses on animal social structure on all levels from individuals to populations. As we will demonstrate in the following chapters, a networks approach allows us to assign quantitative measures of social structure to individuals and populations. These measures open up exciting opportunities for data analysis. For example, they can be analyzed in the context of measured attributes of individuals such as morphology (e.g., body size), behavior (boldness, say), or reproductive success, plus inter-individual measures such as relatedness, all of which will help us shed new light on the mechanisms and functions underpinning, and underpinned by, animal social structure. The fundamental point is that understanding network structure will potentially tell us substantially more about the individual and population than will information on individual attributes alone or interactions between individuals in isolation.

One of the messages that we hope will ring clearly from this book is that animal social network analysis is a work in progress. Much of what we might think of doing is still beyond us, but a second aim of this book is to stimulate interest in developing the remaining tools needed to make the approach an indispensable tool in the behavioral sciences. Before we list some of the potential uses of networks, let us whet the appetite with an illustration of the sort of thing that can already be done. To this end we made up a data set for an imaginary species (*Commenticius perfectus*) that we hope illustrates a number of useful techniques. In the logbook, or even loaded into a database or spreadsheet, a relational data set can look like an impenetrable tangle of numbers, with no discernable pattern to it. One of the first steps (and often a rather major one) in making sense of network data is to plot it. As we will explore more fully in chapter 3, the way we plot it can make a big difference.

Figure 1.4a shows that even for our very simple example, with only 20 nodes (animals) and 35 edges (interactions or associations) a network can look just as featureless as the original data. However, many computer packages exist that can very easily be instructed to lay out the nodes and edges in a way that often looks much more appealing. Figure 1.4b shows the same network, but laid out using a technique called spring embedding (see chapter 3). From this it is immediately clear that all animals are somehow interconnected into a single network, but that not all animals are equal in terms of the number of connections they have, for example, or in terms of whether they occupy central or peripheral positions in the overall network.

In our experience, this first simple visualization of hard-won relational data often has a large impact on the researcher, and almost invariably induces interesting questions and hypotheses about why the network structure is the way it is. We can often help ourselves by adding individual identities and attribute data such as sex or body size (fig. 1.4c). This information helps us to understand how phenotypic attributes influence who is connected to whom in the network and which individuals are central and which peripheral.

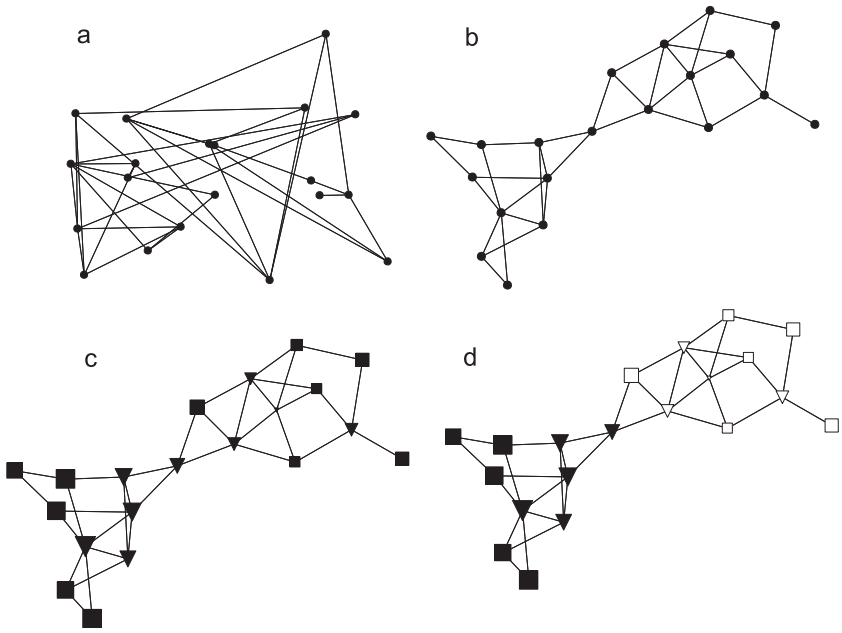

Figure 1.4. Four visualizations of the same animal social network for a fictitious species, *Commenticius perfectus*: (a) random layout, (b) spring embedding, (c) with node shape and size depicting animal attributes: shape indicates sex (square = male) and size indicates body length, (d) with node shading indicating community membership (the other node features [size and shape] are as in c).

Once we start to pose questions of our network, then we need to quantify its structure. We can calculate a whole range of descriptive statistics that measure the local and global properties of the structure of our network, including measures of the overall network size, how interconnected the local network neighborhoods are, how interconnected the network is globally, how central or peripheral different classes of individuals are (e.g., male versus female), and so on. A detailed explanation of some of these measures can be found in chapter 4. We can also look for substructures—so-called communities—in our network and we can identify those individuals that interconnect communities and relate their key position in the network to some of their attributes (fig. 1.4d). Finally, we may want to compare the structure of our *C. perfectus* network with other networks, perhaps for the same population under different environmental conditions or based on different behavioral interactions, or indeed with other populations or species, to gain insight into the generality of the observed patterns or the influence of external factors (such as the environment) on the network of social interactions. An exploration of the options for comparing networks can be found in chapter 7.

So what biological questions might we be able to address with a networks approach? We believe that there are many, but here is a brief discussion of a few of the possible avenues of research.

Social network analysis allows us to focus on individuals and examine the influence of the network on individual behavior. The notion that the social environment has fitness consequences for the expression of individual behavior is central to game theory (Maynard Smith 1982), and theoretical work on behavioral strategies has demonstrated the importance of interaction patterns (Nowak and May 1992; Nowak, Bonhoeffer, and May 1994). The network approach puts individual behavior in the context of the population social structure, thereby helping us to understand the evolution of behavioral strategies. For example, Ohtsuki et al. (2006) showed that the evolution of cooperation is strongly dependent on the fine-structure of social networks. They found that in networks representing perfect social mixing (i.e., each individual is equally likely to interact with each of the others), defectors benefit from exploiting cooperative individuals. However, if individuals do not live in a perfectly mixed world, then selection may favor cooperators when the average number of connections per individual is sufficiently small. In other words, the network structure may determine whether a cooperative strategy is able to persist in a population. The network approach thus puts individual behavior in the context of the population, thereby helping us to understand the evolution of behavioral strategies.

It has been recognized for some time that considering the social environment in which behavior is expressed is important in the context of signaling (McGregor and Dabelsteen 1996). There is a growing literature on "communication networks" that connect an individual signaler to each of its receivers. The approach has provided important insight into the design and function of signals and the role of the social environment on signal evolution. These communication networks are defined for each individual under investigation. There is obviously great potential in coupling a social networks approach with that of communication networks. One area that may be particularly productive is to look at the transmission of information across a social network spanning multiple communication networks (and which therefore contains individuals that are not in direct communication range).

A networks approach also has the potential to illuminate the influence of individual behavior on network structure and function. For example, Flack et al. 2006 studied the construction of social niches in primates. Using "knockout" experiments on a captive group of pigtailed macaques (*Macaca nemestrina*), they demonstrated the importance of a small number of individuals in the population that performed policing behavior (intervention during conflicts). When the policing individuals were absent from the network, social niches destabilized, with group members forming smaller, less-diverse and less-interconnected networks across a range of behaviors including play and sitting

in direct contact. There is obviously great appeal in extending this approach across different taxa and behaviors.

Probably the main strength of the network approach is its potential to address population-level or cross-population-level problems by building up complex social structures from individual-level interactions. Thus the network approach bridges the gap between the individual and population, and this is highly relevant to modeling population-level processes. For example, a network approach allows us to identify who is connected to whom in the population, information that helps formulate hypotheses as to who will learn from whom (Latora and Marchiori 2001), or who will infect whom with a disease (Watts and Strogatz 1998; Corner, Pfeiffer, and Morris 2003; Cross et al. 2004). An understanding of the fine-scale interaction patterns between individuals revealed by social networks is a major advance on the assumption of random interactions by traditional epidemiological models of disease transmission or models of social learning. In a similar vein, we now understand that many social systems contain elements of self-organization in that behavioral interactions strongly influence population-level processes that in turn will feed back on the individual (Camazine et al. 2001). It would seem likely that a networks approach can help to make both parts of the process, individual→population and population→individual, more tractable.

Another motivation for writing this book is that we believe that there are plenty of behavioral scientists who are using a graphical network approach, but not following it up with quantitative analysis, and others who have relational data sets which could fruitfully be analyzed using network methods. We would like to encourage anyone in either camp to have a go. It is remarkable how commonly behavioral biologists use a quasi-network approach to their study system. At a recent conference on animal behavior, we found that 25 percent of the posters showed data in the form of networks. Interconnecting individual animals with lines is a popular approach in studies on birds and mammals (particularly primates; see figure 1.5 for an example). However, often the authors appear not to be fully aware of the power of the network approach and the wealth of information that can be extracted from their data set in this way, provided the data were collected in an appropriate form (see chapter 2). Frequently networks are used purely as a graphics tool but not an analysis tool. This discrepancy between the widespread use of a network representation and the lack of, or superficial use of, statistical analysis is what prompted us to write this book.

We aim in writing this book to offer a synthesis of some of the very large networks literature that is most appropriate for behavioral biologists. Despite the wealth of literature on networks in the field of sociology and the physical and mathematical sciences, we felt that a book devoted to the needs of biologists is merited for a number of reasons. Many of the publications on network theory in the physics and mathematics literature are highly technical

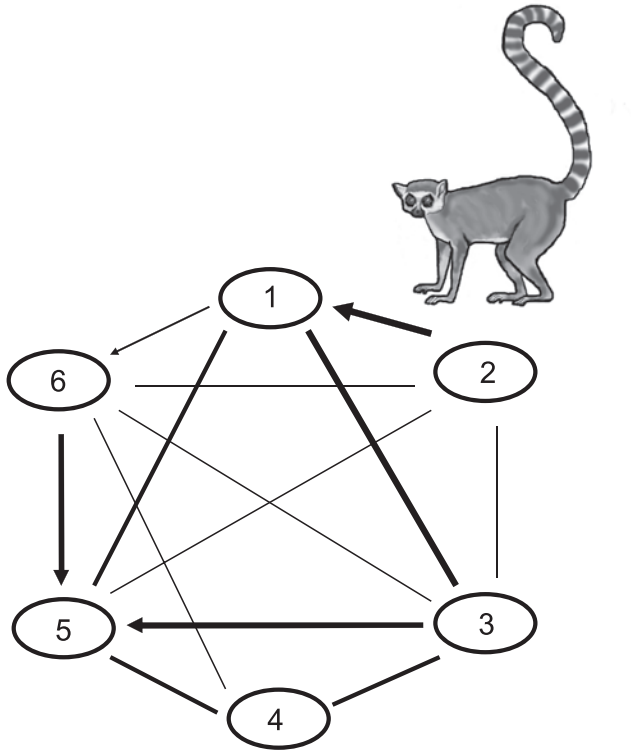

Figure 1.5. Sociogram of infant licking among mothers in ring-tailed lemurs (*Lemur catta*) (redrawn from Nakamichi and Koyama [2000]). Line thickness indicates frequency of interaction and arrows indicate nonsymmetric relationships (when one mother licking another mother's infant occurs significantly more often than expected by chance).

and difficult to grasp for the uninitiated reader. We have tried to provide an introduction to the concepts and tools of network theory suitable for novices of the subject. Though some of the literature in the social sciences is more digestible, there are differences in approach that must be addressed. Biologists in the behavioral sciences are usually interested not only in the mechanism but also the function and evolutionary origin of behavior. For instance, a biologist might want to understand what selection pressures have led to the evolution of particular network structures (i.e., interaction patterns between individuals). A comparative study of networks of different populations of the same (or closely related) species might be one way of addressing such a question. Such an approach can lead to the development of questions fundamentally different from those of interest to sociologists and psychologists. Furthermore, the

issues involved in experimental design and data collection—a topic to which chapter 2 of this book is dedicated—differ considerably between sociology and behavioral biology. It is crucial for behavioral biologists to replicate their experimental trials to test for the generality of the observed patterns and ideally to manipulate their system experimentally to investigate the underlying mechanisms and functions. We must also be aware that our sampling regime might influence network structure—a problem that we will return to at various points in later chapters. It is our hope that this book will go some way toward helping colleagues address these problems in future studies on animal social networks.

Our final aim is to apply the occasional "over-enthusiasm filter." Some of the physics literature on networks is concerned with data sets much larger than anything a behavioral biologist is likely to encounter and thus subject to different statistical approximations. We have tried to provide the reader with network theory that is "ready to use" for sample sizes that biologists are likely to work with. To illustrate this point, many if not most biologists collecting data on social networks are unlikely to obtain data sets large enough to facilitate a realistic test of scale-free properties of their system (see chapter 4 for details). The latter is a topic of intense interest among physicists and mathematicians (see Barabasi and Bonabeau 2003) among others, but unlikely to become a major issue for most behavioral biologists who observe and collect information on a few tens or hundreds of primates, cetaceans, or ungulates.

1.4 AN OUTLINE OF THIS BOOK

This book is aimed primarily at behavioral biologists who wish to collect data on the social organization of animal groups or populations in captivity or in the field. It is meant to give practical advice on how to design a study and how to collect, present, analyze, and interpret data using a social network approach. Although each section and each chapter should allow the reader to complete a certain analytical process that is useful in its own right, we have designed the book so that topics of increasing complexity are covered. Some readers may feel that all they need is some information on descriptive statistics. Others may be keen to run simulations and use or further develop algorithms that detect community structures. Therefore we hope that this book has something to offer to beginners as well as to established scientists in this field of research who share our interest in and fascination with social networks.

Wherever possible we will point the reader to one of a number of software packages that facilitate the calculation of descriptive network statistics and the running of tests and simulations (see box 1.1). These packages make the network approach more accessible to a wide audience. Though the book is not a guide to using a particular analysis or graphics package, we have made

considerable use of UCINET and NETDRAW to illustrate the text. From our experience both programs are easy to use and suitable for exploring networks in animals. Free trial versions can be obtained for both programs via the Internet. There are many other software packages out there that will do the same job (see Huisman and van Duijn [2005] for a review).

Box 1.1

An Overview of the Programs Referred to in the Book

In a recent book, Huisman and van Duijn (2005) provided a comprehensive review of both commercial and free packages for analyzing social networks. The appendix in Scott (2000) also contains a summary of many packages. To illustrate the potential of social network analysis, we have used UCINET (Borgatti, Everett, and Freeman 2002), which is probably one of the more frequently used software package for the analysis of human social network data (Huisman and van Duijn 2005). Other programs and packages have also proved very useful. Here we present an overview of the main programs featured in this book.

UCINET is a comprehensive package for the analysis of social networks (Borgatti, Everett, and Freeman 2002). It can read and write to a range of different format text files in addition to Excel files. It is able to deal with very large data sets—networks that contain up to 32,767 nodes (individuals), far more than most of us are ever likely to be able to collect data on. UCINET offers a range of network analysis methods and procedures, including many of the techniques described in the following chapters in this book. Overall UCINET is a very useful and user-friendly package. Integrated with UCINET is the NETDRAW program (see below) for drawing social networks.

NETDRAW is a free program written by Steve Borgatti for visualizing social network data in 2D space. The visualizations are very flexible, and use different algorithms to display the network in a range of formats. It can handle multiple relations at the same time, and node attributes can be used to set colors, shapes, and sizes of nodes. NETDRAW also has some analysis procedures, the most useful of which is perhaps the possibility to visualize communities (see chapter 6) in the network. The network images can be saved in a range of formats including metafile, jpg, gif, and bitmap. There are a number of options for importing files into NETDRAW.

SOCPROG is a series of MATLAB programs written by Hal Whitehead for analyzing data on the social structure, population structure, and movements of identified individuals. SOCPROG works through a user interface, and most procedures can be executed at the click of a button without knowledge of MATLAB. For those not running MATLAB on their computer, there is also a compiled version that will run without MATLAB.

POPTOOLS is a set of tools for analysis of matrix population models, simulation of stochastic processes, and calculation of Monte Carlo and bootstrap statistics in Excel 97 or above. The interface for the program is self-explanatory, and POPTOOLS also comes with worked demonstrations that are extremely useful. If the matrix exceeds the size limit allowed by Excel (255×255 cells, due to the limit on the number of columns in a worksheet), it is possible to use tab-delimited text files to input the distance matrices (more information can be obtained from the help files in POPTOOLS).

To help you find your way through the book we have outlined a number of steps that might be followed when taking a network approach:

1. Formulate an initial question: All scientific work starts with a question for which we want an answer. Whether your particular question is amenable to the use of a network approach depends on a number of criteria. You must be able to recognize individuals (see point 2 below) and observe interactions between them. In most cases you will have to make repeated observations on the same individuals to build up information on interaction patterns within the group (chapter 2).

2. Determine a method for identifying individuals (chapter 2): The construction of networks is usually dependent on the identification of individuals or at least categories of individuals (e.g., castes in social insects). A number of techniques for marking individuals are discussed in chapter 2.

3. Choose a measure of interaction and a research design (chapter 2): The choice of measure for interactions will dependent on the type of investigation. Social interactions between individuals often involve multiple sensory channels (e.g., visual, acoustic, mechanical stimuli). You also need to decide on the number of individuals that you want to monitor and how often and for how long to observe them.

4. Define each interaction measure (chapter 2): What constitutes an interaction and the precise nature of the interaction requires careful observation

and definition to standardize the data set and make it reproducible. In many cases co-membership in a group, or some other measure of association, may be a perfectly useable proxy for a pair-wise interaction.

5. Select the appropriate recording methods (chapter 2): Observations can be made in various ways (e.g., continuous observation, point sampling, event sampling), and the choice made can have important implications for the quantity and quality of data collected. Field studies are often constrained by animal movements (movement to and from the study site) and mortality, and need to be carefully designed from the outset to make recording representative social networks possible.

6. Organize the data (chapter 2): Information on social interactions between individuals needs to be organized into a matrix for data analysis.

7. Consider sample size (chapter 3): The amount of data that need to be recorded for a study will obviously depend on the question that we want to answer and on how dynamic the study system is. For example, if interactions between individuals are relatively stable, we may be able to get at the social structure with relatively few observations.

8. Construct and visualize the social network (chapter 3): It is often helpful to run a pilot study to test whether the experimental design and data-recording techniques produce the expected results. For this purpose it is useful to monitor how the data set for the network builds up over time and to run some preliminary analyses. A pilot study will usually show very quickly whether for a given sample size a realistic number of repeated observations on individuals can be made that results in meaningful data before the entire study is completed. Looking at the networks can be very helpful in this context. This is the place to jump in if you already have some data and want to use network theory to analyze them.

9. Perform detailed network analysis (chapters 4–7): A number of quantitative metrics can be calculated that describe social structure across different scales of organization, from the individual to the population. Most of the descriptive statistics (chapter 4) can be computed relatively quickly within computer packages. More advanced techniques require statistical tests, some but not all of which are available through computer packages (chapters 5–7).

10. Interpret network measures (chapter 5): Often it is very helpful to compare the observed network to a randomized network that provides a null hypothesis. There are a number of different randomization techniques that need to be carefully distinguished because the choice has an important influence on the results and interpretation of our observed data.

11. Search for sub-structures (chapter 6): A closer look at the fine-structure of networks can help identify subunits (so-called communities) and individuals

interconnecting sub-units. Both types of information can be very useful in formulating testable hypotheses.

12. Compare networks (chapter 7): Comparing the same set of individuals and the interaction patterns that interconnect them under different ecological conditions can provide important information on the social organization of a population. Likewise we can compare the interaction patterns of closely related species or different populations of the same species that are exposed to different ecological conditions. This type of analysis may provide us with insights into the evolution of social organization.