Network representation and complex systems

Charles Rathkopf[1]

Abstract In this article, network science is discussed from a methodological perspective, and two central theses are defended. The first is that network science exploits the very properties that make a system complex. Rather than using idealization techniques to strip those properties away, as is standard practice in other areas of science, network science brings them to the fore, and uses them to furnish new forms of explanation. The second thesis is that network representations are particularly helpful in explaining the properties of non-decomposable systems. Where part-whole decomposition is not possible, network science provides a much-needed alternative method of compressing information about the behavior of complex systems, and does so without succumbing to problems associated with combinatorial explosion. The article concludes with a comparison between the uses of network representation analyzed in the main discussion, and an entirely distinct use of network representation that has recently been discussed in connection with mechanistic modeling.

1 Introduction

It is often remarked that network science is transforming our understanding of complex systems. In a recent opinion paper in *Nature Physics*, Albert-Laszlo Barabasi says: "One thing is increasingly clear: no theory of the cell, of social media, or of the

---

✉ Charles Rathkopf
charles.rathkopf@gmail.com

[1] Graduate Center at the City University of New York, 5674 Frist Center, Princeton, NJ 08544, USA

 Springer

Internet can ignore the profound network effects that their interconnectedness cause [sic]. Therefore, if we are ever to have a theory of complexity, it will sit on the shoulders of network theory" (Barabasi 2012, p. 15).

Although it is reasonable to doubt that any general theory of complexity is in fact forthcoming, Barabasi is right to think that network science has special relevance to the study of complex systems. One reason for this is that network models have been used to shed light on a bewildering variety of complex empirical phenomena, including the frequencies of protein–protein interactions, the social causes of obesity, the propagation of viruses through the Internet, and the neural correlates of Alzheimer's disease.[1] Another reason for this is that that network models offer a powerful way of representing and reasoning about the manner in which complex systems are interconnected.

In this article, network science is discussed from a methodological perspective, and two central theses are defended. The first is that network science exploits the very properties that make a system complex. Rather than using idealization techniques to strip those properties away, as is standard practice in other areas of science, network science brings them to the fore, and uses them to furnish new forms of explanation. This head-on approach to complexity is quite novel, especially in comparison with explanatory strategies that have been emphasized in recent philosophical accounts of complex systems science. The second central claim in this article is that network representations are particularly helpful in explaining the properties of *non-decomposable* systems. Where part-whole decomposition is not possible, network science provides an alternative method of explaining system behavior. Together, these two claims show that network science is, as Barabasi imagines, a powerful method for understanding the interconnectedness of complex systems, and one that breaks free of certain limitations inherent in other methods.

The discussion is organized as follows. In Sect. 2, some network science concepts are introduced and a pioneering model is presented in some detail. In Sect. 3, it is argued that this pioneering model, despite being highly abstract and statistical in nature, satisfies standard tests for explanatory relevance. In Sect. 4, an argument is given for the first of the two aforementioned central claims: that network science exploits complexity rather than avoiding it. In Sect. 5, an argument is given for the second central claim: that network representations are particularly helpful in explaining the properties of non-decomposable systems. In Sect. 6, a comparison is made between the perspective defended here, and that defended in a recent paper by Levy and Bechtel on the role of network representation in the study of mechanisms. Levy and Bechtel conceive of network analysis as a continuation of mechanistic analysis. I defend the view that much of network science should be seen instead as a departure from the mechanistic approach, and one that offers a completely distinct explanatory strategy.

---

[1] Canonical sources for these developments can be found, respectively, in Rual et al. (2005), Christakis and Fowler (2007), Pastor-Satorras and Vespignani (2001), and Supekar et al. (2008).

2 The distinctive strategy of network science

2.1 Representing networks

Network science is a systematic attempt to study properties that are hidden in the pattern of connections among the elements that compose complex systems. The principal mathematical tool of network science is graph theory. A graph consists of a set of points either on a plane, or in n-dimensional space, and a set of line segments, each of which either joins two points to one another, or joins one point back to itself (Gross and Yellen 2006). The canonical form of graph theoretic representation is known as an adjacency matrix, which is a matrix in which both rows and columns are labeled by an ordered list of elements. If a direct connection exists between two elements, their intersection is marked with a 1; otherwise, it is marked with a 0.

Adjacency matrix *A* in Fig. 1 represents a graph with six points, or *nodes* and seven connections, or *edges*. The same information can also be represented in a diagram that is easier to inspect visually.

It is important to keep in mind that Figs. 1 and 2 are equivalent from a graph theoretic perspective. In Fig. 2, no information is carried by the angles between lines or the apparent spatial position of the nodes.

Graph-theoretic representations can be used to model empirical systems. For example, each node might represent a protein type, and each edge might represent a protein–protein interaction that occurs in one species. Such a graph summarizes which proteins depend on which others to realize their biological function. Such a summary can yield predictions about unknown functions by relying on the fact that proteins with overlapping pathways tend to have a high degree of functional similarity (Schwikowski et al. 2000).

Fig. 1 Adjacency matrix for a 6-node graph

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

Fig. 2 The same system represented by the adjacency matrix in Fig. 1, here represented diagrammatically
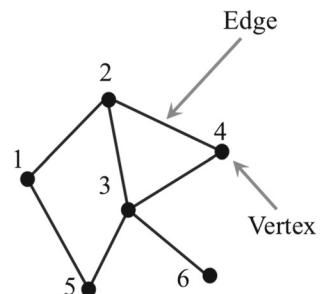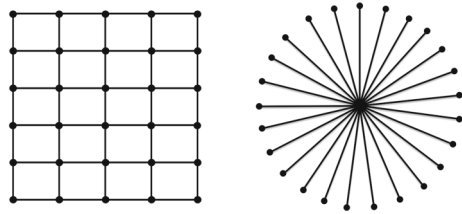
Fig. 3 A 25-node network configured as a square lattice, and another configured as a star network

2.2 Organization influences behavior: a clear case

There is a straightforward sense in which the behavior of a highly interconnected system depends on the way its inter-element connections are organized. The star network depicted in Fig. 3 offers a maximally simple example of this kind of dependence.

One network property that is often used in the analysis of complex systems is path length, which is defined as the number of edges on the shortest path between two vertices. In Fig. 3, it is easy to confirm visually that the longest path length in the star-shaped network is two. Since every node is connected to the central node, (other than the central node itself) one can trace a line between any pair of nodes in two steps or less. Compare this to the square lattice on the left. The path length along just one side of the lattice is four. Some path lengths in the graph are smaller; others larger. If we imagine that nodes represent computers, for example, it is easy to see that information originating at one computer can be distributed to the whole network in fewer steps in the star network than it can be in the lattice-like network. This is a toy example, but it shows that the pattern of organization among the elements of a network can influence the dynamical processes instantiated on that network. Now we consider a more scientific example.

2.3 An influential network model

In the paper that triggered the contemporary proliferation of network science, Watts and Strogatz (1998) use network representations to explore the dynamics of infectious diseases. Throughout most of its history, the mathematics of graphs had been studied under the guise of combinatorics, and had had little impact on empirical science.[2] For the most part, mathematicians had restricted their investigations to regular lattices (much like the square lattice in Fig. 3), in which each node is connected to a fixed number of its nearest neighbors, or random graphs, in which every possible pairwise connection is either drawn in or left blank at the flip of a coin. What intrigued Watts

---

[2] If we interpret "network science" loosely, its history is much longer. Sociologists have applied small networks to social phenomena since the work of Moreno in the 1930s (Easley and Kleinberg 2010), and operations researchers made use of graph theory in WWII (Cunningham et al. 1984). The very large-scale work described here, however, requires computational power that has only been available more recently. It should also be noted that the more recent history of network science has had a much broader impact than early work in the field. In 2008, the Thompson Reuters publication "Science Watch" ranked the Watts and Strogatz paper as the 6th most cited paper in the history of physics. On August 9th, 2014, the Thomson Reuters Web of Knowledge (apps.webofknowledge.com) shows that the paper has received 10,308 citations.

and Strogatz was the empirical finding that most of the network-like systems that had been studied in the real world displayed neither of these topological structures. Instead, empirical network-like systems display irregular clustering at the local level, and these local clusters are often connected to one another by means of a long-range bridge.[3] This kind of network structure had recently been discovered in systems as diverse as the nervous system of a nematode worm, a network of actors that had appeared in films together, and the power-grid of the Western United States.

Because this empirical data looked qualitatively different from the kinds of graphs typically studied in pure mathematics, Watts and Strogatz (among others) became interested in developing algorithms to construct networks that shared some of the statistical properties observed in empirical data. This algorithm subsequently became the model for many similar graph construction algorithms that have been studied with increasing enthusiasm since the original publication.

The algorithm they devised takes as its input a ring-lattice in which each node is connected to a fixed number of its closest neighbors on the ring.[4] The algorithm then considers each edge in turn, choosing with some fixed probability to rewire it to a new, randomly selected node. Watts and Strogatz decided to use this model to study how patterns of connections among people might influence disease spreading.[5] To do this, they adapted a traditional epidemiological model to fit the network context. That model is known as the SIR model of disease spreading, and it can be expressed as three differential equations in which a population is divided into susceptible (S), infected (I), and recovered (R) compartments. These quantities vary as a function of time, but are conserved, such that $S(t) + I(t) + R(t) = N$, where N is the total population size. One infected person is introduced into a susceptible population, and the disease then spreads to other people along the edges in the network. After a fixed number of time steps, each infected person is moved into the recovered compartment, and then enjoys a fixed period of immunity before becoming susceptible once again. The three equations that comprise the model are:

1. $dS/dt = \delta R - \beta SI$,
2. $dI/dt = \beta SI - \gamma I$,
3. $dR/dt = \gamma I - \delta R$,

where parameter β is the rate at which the infection is transmitted, parameter γ is the (memory-less) rate of recovery, and parameter δ is a fixed period of immunity acquired

---

[3] The term "long-range" refers to path-length in graph-theoretic, rather than physical space. The idea can be expressed intuitively in terms of a social network. In a graph of my social network, the set of all my friends are represented as being one unit removed from me. The set of all the friends of my friends (who are not also friends with me directly) are represented as being two units removed from me, and so on. A long-range connection is one that directly connects two nodes that would otherwise be separated by a long path, relative to the global average.

[4] The ring-format of the lattice was chosen for the computational convenience of periodic boundary conditions, and is not meant to have any representational significance.

[5] They also studied how the structure of networks influences the interactions of models of rational agents in a Prisoner's dilemma scenario. Although the results of this initial work also led to a huge amount of subsequent study, I focus on the epidemiological aspect of their work to avoid the special complications associated with models of rational decision-making.

after recovery.[6] By varying these biological parameter values, one can represent a variety of different diseases.

One of the assumptions of the traditional SIR model is that the population is randomly mixed. This assumption makes it possible to model the transmission probability β as a uniform random variable, each value of which describes the proportion of the population that becomes infected at one time-step in a computer simulation. This idealization, which is known as 'the homogenous mixing assumption,' is nearly equivalent to the hypothesis that the probability of coming into contact with a friend or neighbor is the same as the probability of coming into contact with someone randomly selected from the population. I say "nearly equivalent" because traditional compartmental models do not represent individual disease transmission events explicitly, so it is false, strictly speaking, to say that the model describes probabilities that are attached to individual transmissions. The value for β can be more accurately thought of as the average number of infection opportunities experienced by the whole population during a given time interval (Bansal et al. 2007). However, when we constrain the disease to spread along a pre-drawn graph of the contact structure, β ceases to be a uniform random variable and becomes instead a vector quantity that summarizes at each point in time precisely which contact opportunities exist.[7] To explore this new graph-theoretic version of the SIR model, computer simulation is required. On each simulation trial, only some infection opportunities result in successful transmission. This opportunity/success ratio is averaged over simulation trials, and the resulting quantity represents the transmission probability of the disease in a more fine-grained way than is possible in the original compartmental model.

It turns out that the dynamical properties of diseases, as represented by the SIR model, are highly sensitive to network topology. This makes intuitive sense. One might suspect that the dynamics of a disease spreading process would be different in an urban area packed full of people than they would be in a rural area in which large crowds are rare. The beauty of the Watts–Strogatz model is that it allows us to explore the impact of the contact structure on the disease dynamics in a quantitative way.

2.4 What the Watts–Strogatz model explains

What can be learned about real disease behavior from the Watts–Strogatz model? I will focus on just one lesson that involves two simple network properties. The first property is called the characteristic path length. The definition of path length provided above

---

[6] The form of these equations is slightly more complex than the one actually implemented by Watts and Strogatz. I have chosen to write this version of the model both because it is the most widely recognized version of the SIR model, and also because it allows me to define a biological parameter in Section 3.5, where I argue that the network epidemiology is reasonably autonomous from certain biological details. It is also worth noting that I am following epidemiological practice by referring to the model parameters as "rates." In fact they are just proportions and do not literally have a temporal component.

[7] Some texts use the term "population structure" to describe the disease topology of biological population. I prefer the term "contact structure" because the set of relevant connections—instances of physical proximity between two organisms that make transmission biologically possible—will be different for each population/disease pairing. "Contact structure" does a better job of avoiding the misleading implication that there is a single structure associated with each population.

described a relation between two nodes: the number of edges found on the shortest path between them. Characteristic path length is a generalization of this idea. It is defined as the average path length over all pairwise combinations of nodes in the graph. The second property is called the clustering coefficient of the graph, and it can be defined as the probability that two neighbors of a given node are themselves neighbors (where the term "neighbor" indicates a direct link). Somewhat more precisely, the clustering coefficient is a measure of the ratio of the number of three-node combinations that form triangles to the number of all possible three-node combinations. Empirically speaking, therefore, clustering is a measure of the cliquishness of a population.

Real human populations are known to be highly clustered. Theoretically, high degrees of clustering should prevent a disease from reaching a large proportion of the total population quickly. High levels of clustering imply that the possible routes of disease transmission are highly overlapping, so that a given chain of transmission is likely to follow a closed loop back to one of the already-infected individuals, rather than follow an open path toward a new susceptible individual.

Watts and Strogatz found empirically, however, that many networks had very short characteristic path lengths despite the fact that clustering coefficients remained high. This means that the number of transmission events required to traverse from any one node in the graph to any other was quite small. This was unexpected because in both the random graphs and the lattice-like graphs that were known to mathematicians, large clustering coefficients are always associated with large characteristic path lengths, and small clustering coefficients are always associated with small characteristic path lengths. By contrast, in the empirical data sets Watts and Strogatz inspected, they observed that clustering coefficient and path length were very often anti-correlated. Their algorithm was designed to construct graphs that matched the empirical data in precisely this respect: it provides a way of exploring the mathematical space in which clustering coefficient and characteristic path length diverge.

Watts and Strogatz found that these two properties diverge much more readily than expected. As the random, potentially long-range connections are introduced via the re-wiring algorithm, the path length of the resulting graph drops off precipitously while the clustering coefficient hovers near its maximal value.

Notice that the x-axis in Fig. 4 is plotted logarithmically so that the extremely steep descent into the low path length/high clustering regime is made visible. This shows that the introduction of only a very small fraction of long-range connections has a radical impact on the path-length, but very little impact on the clustering coefficient. Since path length can be thought of as a kind of distance, and since most of the graphs described by Fig. 4 have low path length, Watts and Strogatz named the phenomenon the *small-world property*, a term that has since gained prominence both in scientific and popular circles.

How does this relate to the dynamics of infectious disease? In a randomly connected contact structure, diseases can spread very quickly because characteristic path length is low.[8] Since it only takes a small number of long-range connections to turn a highly clustered graph with high path length into a small world graph with low path length,

---

[8] Specifically, it is about $ln(n)/ln(k)$, where n is the number of nodes in the graph, and k is the number of initially connected neighbors.

Fig. 4 Log-normal plot of clustering coefficient ($C(p)/C(0)$) and path length ($L(p)/L(0)$) as functions of rewiring probability. Both properties in the vertical axis are expressed as ratios because they are normalized to values taken from a randomized graph. Rewiring probability determines the proportion of long-range connections. From Watts and Strogatz (1998)
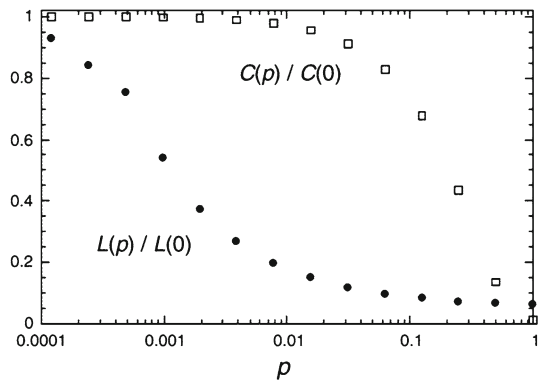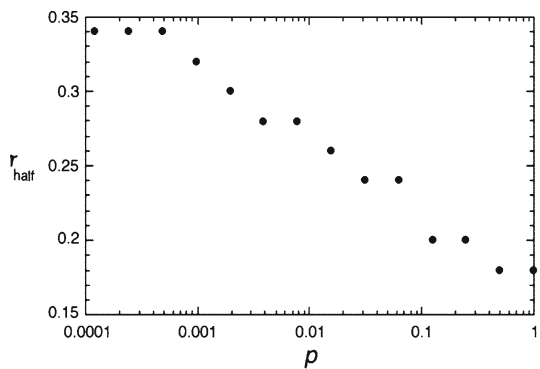

Fig. 5 Log-normal plot of the SIR r_{half} rate against rewiring probability, which determines the proportion of long-range connections. From Watts and Strogatz (1998)

one might expect that very small perturbations to the contact structure might have an enormous impact on the dynamics of a disease. Indeed, Watts and Strogatz observed that the very fast rate at which path length drops off results in surprising sensitivity of disease dynamics to increased rewiring probability. In particular, the value of β required for the disease to reach half the total population, (a quantity they call r_{half}) drops off steeply with increasing rewiring probability.

The x-axis in Fig. 5 is again plotted logarithmically so that the steep drop of the r_{half} rate is visible. The figure shows that a very small percent change in the number of long-range disease transmission events makes diseases with low critical infectiousness rates, which are otherwise easily contained, capable of generating massive epidemics. Moreover, these results carry implications for intervention strategies such as vaccination. Watts and Strogatz themselves interpret their results as evidence that if a disease can escape the quarantine-like effects of clustering early on in the spreading process, it will be able to spread just as easily as if there were no clustering at all.

3 Explaining generic properties

The Watts–Strogatz algorithm is intended to produce a class of graphs that share just two statistical properties found in a wide variety of empirical data sets. In most

other respects, the model is highly idealized, and by itself not sufficient to predict the dynamics of a specific historical epidemic. Nevertheless, the model has considerable explanatory value. To see this, let us first be explicit about the explanatory reasoning to which the model gives rise. It runs as follows: measured populations have the small-world property. Instantiations of the small-world property in empirical systems lead to fast disease spreading, despite the presence of high clustering coefficients. Therefore, many real diseases can spread rapidly despite the presence of high clustering coefficients.

This line of reasoning ought to be counted as explanatory because it passes both probabilistic and counterfactual tests for explanatory relevance. The probabilistic test, originally formulated by Hans Reichenbach and developed to particular effect by Wesley Salmon, involves two necessary conditions. For two variables A and B, if B is explanatorily relevant to A, then (i) the probability of A conditional on B is greater than the unconditional probability of A, and (ii) no third factor screens off A from B. In our case, the relevant inequality says: the probability that variable r_{half} takes on a certain value x, given information about the fraction of disease transmissions that are long-range, is greater than the unconditional probability that variable r_{half} takes on value x. Although the r_{half} rate of a given disease cannot be observed directly, it is clear that the influence of the contact structure on the r_{half} rate is a matter of the frequency with which people come into close physical contact with one another. This suggests that information about the particular causal factors relevant to the determination of the r_{half} rate are already subsumed by the information about the contact structure, and therefore do not screen off its influence. Moreover, model-fitting studies show that information about the contact structure does, under many conditions, improve the accuracy of models when they are used to retrodict epidemiological data (Bansal et al. 2007). So we have reason to believe that information about the structure of real populations carries explanatory information about the r_{half} rate of real diseases, whatever the actual value may be.

The other well-known criterion of explanatory relevance is counterfactual: if variable A had not taken value a, then variable B would not have taken value b (Woodward 2003). One of the counterfactuals prompted by the line of reasoning above is this: if human contact structures did not typically have the small-world property, then epidemics would not spread as quickly as they typically do during their early stages. This counterfactual is almost certainly true. If we could somehow intervene on the population in such a way as to increase the path-length of the contact structure, an epidemic would have fewer opportunities to spread. Alone, such an intervention does not necessitate that every disease would spread more slowly. We can imagine, for example, a situation in which some bacteria are willingly introduced into the country's water supply en masse. In this case, person to person contact of the relevant biological kind would no longer be the primary means by which the disease is transmitted, so contact structure would make little difference to the speed with which the disease spreads. However, if the intervention on the contact structure were "surgical" and all biological factors were held fixed, it would be mathematically impossible for the epidemic to spread as quickly as it would have done without the intervention.

One potential source of doubt about the explanatory status of the reasoning above is that it is highly abstract and idealized. It is therefore unclear whether this reasoning is sufficiently well grounded in the empirical target phenomenon. In particular, the reasoning above is not firmly entrenched in the causal-mechanical factors that underlie real instances of disease transmission.[9] I offer a two-fold response to this concern.

The first part of the response is that the target of the explanation under consideration is a generic feature of disease-like processes, rather than a feature of any particular epidemic. What we would like to understand is why disease-like processes in general tend to display faster spreading behavior when the structure through which they spread has the small world property than when it does not. If a model is to explain this very general fact, it should be able to reproduce the faster spreading behavior using only generic system properties. The model must not, in other words, rely on the contribution of system-specific parameters that pertain to only some peculiar instances of the target phenomenon.

Recently, many philosophical accounts of scientific explanation have incorporated similar arguments. I'll mention two. Weslake (2010) argues that when the targets of scientific explanation are generic facts, rather than events bound to a specific spatio-temporal location, explanations will be deeper to the extent that they do not limit their range of application by making essential reference to system-specific spatio-temporal parameters. On Weslake's view, as long as the statements in the explanans are true, abstraction of the kind present in the network model actually improves the quality of the explanation.

Huneman (2010) develops a theory of explanation in which abstract mathematical structures are shown to describe topological properties of an empirical phenomenon in a way that entails the explanandum behavior. In order to represent an empirical system with a topological structure such as a graph, one needs empirical justification. Naturally, justification depends on causal detail. However, justification and explanation are two distinct ends. Huneman argues that, conditional on the correct representation of topological properties, causal-mechanical details contribute little or nothing to the derivation of the target phenomenon. (Huneman 2010, p. 218).[10]

Huneman's informal definition of a topological explanation runs as follows. "When among the consequences of some topological properties, stands the behavior, property, or outcome we want to explain, then I say that we have given a topological explanation of our explanandum" (Huneman 2010, p. 215). Huneman also understands whole-graph properties as topological in his broad sense of that term.[11] "From now on, I call 'topological properties,' those properties that are either proper to subsets in a topological space or to some graphs and networks" (Huneman 2010, p. 217). The

---

[9] I do not mean to presume that all causal explanations are mechanistic. I discuss causal explanation in this section, and take up a comparison with mechanistic explanation in Sect. 6.

[10] Jones (2014) presents a similar view. However, the focus of that article is on the way that diagrammatic representations facilitate inferences about counterfactual structure. In large networks, very few interesting properties are detectable by means of visual inspection. What makes the representations useful is rather the fact that they can be studied algorithmically.

[11] It should be noted that Huneman is sensitive to the fact that topology and graph theory are distinct mathematical disciplines, but chooses to use the term "topological" in a broad sense that covers both disciplines.

explanation of generic disease behavior in terms of the r_{half} rate appears to count as a topological explanation on Huneman's view. However, not everything Huneman says about topological explanation applies to the view of network explanation presented here. Huneman strives to show that topological explanations are non-causal. For two reasons, I prefer to remain neutral on that score. First, the truth conditions for the claim that an explanation is non-causal are intimately bound up with questions about the metaphysics of causation, which is a topic far too broad to be taken up here. Second, and more importantly, the central concern in this article is to show that networks are particularly appropriate for representing and reasoning about complex distributed systems. The argument for that claim does not require us to take a stand on the philosophical question of whether all scientific explanations are causal. To preserve the broad appeal of this project, I therefore prefer not to take a side on that issue.

The second part of my response to the worry about abstraction highlights the fact that the small-world property in the model has a clear and measurable analogue in empirical data. While it is true that raw epidemiological data are noisy, and that it is necessary to extrapolate from them by means of data models, the data are gathered using standard means of demographic measurement. We already have mechanistic knowledge of the spatial limitations of disease transmission that indicates which kind of demographic measurements are relevant. Clearly, epidemic dynamics would not depend so directly on contact structure if an airborne virus could travel hundreds of miles through the air between infections. But we know that even in the case of airborne viruses, disease transmission depends on physical proximity. Moreover, we have mechanistic knowledge of the transmission process that tells us whether sharing a bus ride is more likely to result in transmission than sharing an office, for example. Given this knowledge, we can construct a model of contact structure based on surveys about personal relationships, and physical and demographic data on schools, workplaces, and transportation systems. For details about how these data are collected, see Rothman et al. (2008). This background knowledge about the causal processes underlying disease transmission is an essential part of the justification for applying the mathematical, graph-theoretic model to the empirical phenomena. Given that justification, the applied model should be counted as what Craver (2006) calls a "how-actually" model.[12]

A second source of doubt about the explanatory status of network models is the fact that when philosophers of science describe novel kinds of scientific explanation, the kinds they have in mind tend to line up with an established scientific discipline such as astrophysics or molecular biology. Since graph-theory is not itself a scientific discipline with unique empirical content, the set of scientific explanations it can be used to construct will not be sufficiently unified to deserve methodological commentary. The response to this worry is that in some cases, new mathematical techniques give rise to a distinctive kind of explanation that fails to line up with traditional disciplinary boundaries. This domain-general, representation-centric approach to explanation has been emphasized in a growing body of philosophical work.

---

[12] Although I am borrowing the terminology from Carl Craver's well-known article "When mechanistic models explain," I do not mean to imply that the small-world model should be counted as a mechanistic. Reasons for resistance are discussed in Sect. 6.

An excellent example can be found in a recent (2014) article by Paul Humphreys, in which he articulates an explanatory paradigm called "explanation as condition satisfaction." Condition-satisfaction explanations proceed by showing that the mathematical generating conditions for a given mathematical object (such as a probability distribution) are satisfied by the empirical generating conditions for an empirical phenomenon. It is then demonstrated that facts about the mathematical object account for non-trivial facts about the empirical phenomenon. According to Humphreys, one of the most important features of such explanations is that they "do not require the use of detailed models of system-specific processes […]" (Humphreys 2014, p. 1103). A number of other recent articles have also promoted that idea that certain kinds of mathematical representation generate epistemologically distinctive kinds of explanation that are not constrained to a particular empirical discipline. These include Batterman and Rice (2014) on minimal model explanation, Lange (2009) on dimensional explanation, and Ladyman et al. (2007) on renormalization group explanation. On the view defended here, network explanation deserves to be added to the list.

We have seen that network explanations are abstract, and that the explanatory inferences they make possible are not confined to a particular discipline. As recent philosophical literature has emphasized, neither of these two features should be viewed as a credible threat to the explanatory status of network models. In any case, despite their salience, neither of these two features is unique to network models. In the following two sections, I describe what I take to be the most distinctive features of network explanation. Taken together, these features illustrate a relatively new and relatively cohesive strategy for reasoning about a variety of complex distributed systems.

4 Network models embrace complexity

4.1 Embracing complexity

In this section, it is argued that network models embrace complexity, rather than shy away from it. In order to establish this claim, I first say what it means to embrace complexity, and show how network models do so. Then, I describe why this orientation toward complexity should be counted as an epistemologically significant development.

In order to understand the claim that network models embrace complexity, we begin with a characterization of complexity itself, as it relates to the phenomena of interest in network science. In this context, complex systems are those in which many discrete elements display a non-trivial pattern of interaction. If you want to understand the collective behavior of a complex system, you need to take that pattern of interaction into account. In other many-element systems, the pattern of inter-element interaction plays a less important role in determining system behavior, and can therefore be ignored or simplified. In ideal gases at equilibrium, for example, intermolecular interactions are stochastically independent. This fact provides justification for ignoring the precise features of the pattern, and allows us to represent the set of interactions with a single distribution function. This is a case in which randomness justifies the suppression of information about the pattern of interaction. In other cases, a highly ordered configu-

ration of interactions provides the needed justification. For example, the melting point of many crystals is determined by the ratio of the displacement of an atom from its lattice position to the distance to its nearest neighbor. Although crystals are almost never organized into perfect lattices, and distance to nearest neighbor varies from site to site, it can nevertheless be represented by a fixed parameter for many computational purposes. In both cases, successful models suppress information about the pattern of interaction by replacing it with an all-purpose fixed parameter.

When dealing with complex systems, such idealizations are not satisfactory. The particular pattern of inter-element interaction plays an ineliminable role in producing the collective behavior of interest. In such cases, it becomes necessary to include at least one variable that carries information about the precise pattern of inter-element interaction. Let us call such a variable a "pattern variable." To say that a network model embraces complexity is to say that it reproduces the collective behavior of interest by representing that behavior as a function of a pattern variable. Figure 5 depicts a canonical example of such a functional relationship: the pattern variable p captures the proportion of interactions that are long-range. The r_{half} rate is the collective behavior that epidemiologists would like to explain.

We now have an account of the sense in which network models embrace complexity. That account relies directly on the notion of a pattern variable, which has not been clearly defined. Naturally then, our next task is to construct such a definition.

4.2 The fixation criterion

Notice that the graph construction algorithm in the Watts and Strogatz model is designed such that the graph retains the same number of nodes and edges on each simulation trial. Moreover, the only biological parameter in the model is transmission versus lack of transmission. Even though these properties are held fixed on each trial, our explanatory variable—the proportion of long-range connections—remains free to vary.

This observation can be generalized to yield an account of what it is for a variable to represent a pattern of interaction. A variable will count as a pattern variable in the sense intended here if and only if holding fixed (i) the number of nodes (ii) the number of dyadic interactions per node and (iii) the dynamics of each interaction, is not sufficient to fix the value of the variable.

Conditions (i) and (ii) reflect the fact that information about the pattern of inter-element interaction cannot be captured by facts about the mere magnitude of the system. Condition (iii) shows that information about the pattern of interaction cannot be captured by facts about the composition and dynamics of local dyadic relations. Patterns of interaction provide information that is instead about the unique manner in which interactions are distributed across the entire collection of elements. Variables that satisfy the fixation criterion provide some direct insight into the properties that make the system complex. If those variables can be recruited to explain system behavior, then the explanation is one that exploits complexity for explanatory ends.

4.3 The epistemological significance of network models

At the outset of Sect. 4, it was suggested that the orientation toward complexity implicit in the use of network models is epistemologically significant. In this section, I give two reasons for this view.

The first reason has to do with the ability of network models to abstract away from the underlying theory of the individual system components. To see this, let us compare network and non-network approaches to answering the question "why do diseases sometimes move quickly through a population?" Before the Watts and Strogatz result, all known attempts to answer this question incorporated information about the biological properties of infectious diseases (information about the values of the parameters in the compartmental model above). A standard answer would cite the fact that the degree of infectiousness of the disease in question has surpassed a critical threshold. Specifically, the standard answer would cite the fact that a quantity known as the *basic reproductive number* is substantially greater than one. The basic reproductive number is a measure of the average number of new infections at each time step, and can be calculated approximately as the ratio of two biological parameters, β/γ. The Watts–Strogatz results provide an entirely different answer that is based on the organization of the contact structure, rather than its biological underpinnings. The network approach provides an avenue of understanding that is at least partially uncoupled from biological theory, and is therefore resistant to traditional accounts of inter-theoretic reduction.

The second reason that the network orientation toward complexity is epistemologically significant is that network models are far more *abundant* than traditional forms of scientific representation. Every element in the system is represented explicitly in the graph, as is every pairwise interaction. This abundance necessitates massive representations and considerable computational power. Since the number of elements in an adjacency matrix grows as the square of the number of system elements, network representations are subject to combinatorial explosion. For example, in one of the first large-scale, network-oriented studies of protein–protein interactions in humans, Rual et al. studied the interactions between 8,100 protein types, creating a space of over 65 million possible interactions. It is this capacity of graphs to explicitly represent every possible pairwise interaction between elements that sets it apart from other forms of representing complex systems.

In traditional approaches to many-element systems, the goal is to construct a model that is itself no longer complex and therefore (hopefully) tractable. Weisberg (2007) calls this model-building strategy "Galilean idealization," and argues that it is one of the three most prominent forms of idealization in the history of science. Although network models always involve idealization away from the empirical character of individual nodes, their head-on approach to complexity breaks free of the need to engage in Galilean idealizations concerning the number of discrete elements in a system, and thereby allows scientists to overcome what was previously a significant epistemic constraint.[13]

---

[13] This freedom from Galilean idealization is one reason why network representations can be viewed as epistemologically distinctive. The computational methods used to study them provide a kind of understanding that at least partially breaks free from the processing limitations intrinsic to human cognition.

5 Explaining non-decomposable systems

5.1 Defining near-decomposability

In Sect. 4, it was argued that network science confronts complexity in a head-on manner. In this section, it is argued that the head-on approach is especially valuable when applied to the study of non-decomposable systems.

A natural starting point for the argument is with a definition of *nearly-decomposable systems*. Strevens (2005) defined nearly-decomposable systems as ones in which "the short-term behavior of a system's components can be understood largely independently of the behavior of other components." In his landmark paper on complex systems, Herbert Simon gave a similar, but somewhat more subtle definition. "In a nearly-decomposable system, the short-run behavior of each of the component sub-systems is nearly independent of the short-run behavior of the other components, and in the long-run, the behavior of any one of the components depends only in an aggregate way on the behavior of other components" (Simon 1962, p. 474).

Where systems are nearly-decomposable, it is not necessary to represent every component in the system explicitly. In virtue of the independence of component sub-systems, scientists are free to develop a theory of each subsystem, and then compose the predictions of those theories in order to yield predictions about the behavior of the whole. According to both Simon and Strevens, the benefit of decomposition operations is that they relieve us of the burden of representing every element explicitly, and thereby save us from the troubles associated with combinatorial explosion. In short, therefore, when dealing with nearly-decomposable systems, some *compact* form of representation will be available.

With this conception of near-decomposability in mind, we can now define non-decomposability. Roughly, a non-decomposable system is one that is not even close to being nearly-decomposable. More precisely, a system is non-decomposable just in case the behavior of any given component part, even over a short time period, depends on the behavior of many other individual components.

This definition is to be interpreted in such a way that the term "component" can refer either to the basic elements in a system or to any collection of basic elements other than the entire set. Interpreted this way, the definition entails that within a non-decomposable system, no component subsystems exist that are either independent or nearly independent of one another. This consequence yields an appropriate contrast with the property of near-decomposability. Near-decomposability is an empirical property that provides justification for using compact representations. Non-decomposability, as defined here, is the corresponding empirical property that guarantees that compact representations will be empirically inadequate.

5.2 Epidemiological populations are non-decomposable

As we saw in the epidemiological case above, network models typically do require the representation of every element in the system, and since many representations involve thousands of elements, I called network science representations *abundant*. Now we

can see that this abundant character is by no means accidental. The alternative to using an abundant form of representation is to rely on part-whole decomposition, with the hope of securing a more compact and therefore more manageable representation. But no part-whole explanatory strategy could have answered the question of why disease-like phenomena spread faster in small-world structures than otherwise, precisely because the epidemiological case is an example of a non-decomposable system.

To see this, we must ask what units of decomposition would serve as component subsystems if the epidemiological system were decomposable. What about individual organisms? Might they serve as units of decomposition? While there is no general difficulty in individuating organisms, the possibility of individuation at this fundamental level is only a degenerate case of decomposability. As both Simon and Strevens have stressed, the goal of decomposition is precisely to find a level of description at which basic elements need not be represented individually, so that our representations do not grow to unwieldy proportions. In our epidemiological example, the only higher-level units available are the groups of organisms that correspond to the compartments in the SIR model. These groups cannot serve as units of decomposition, however, because they do not satisfy the criterion of short-term independence. The size of the whole population is conserved, and therefore any change to one compartment necessitates a corresponding change in another compartment that is, moreover, realized in the very next time-step. The short-term behavior of these groups will therefore be highly interdependent. As a result, they satisfy the crucial short-term dependence condition in the definition of non-decomposability.

Neither individuals nor compartmental groups are legitimate units of decomposition. Might there be a third way? Could we impose some ad hoc taxonomy that is not reflected in the model as it stands? For example, could we divide the population into clusters as defined above, and then predict the behavior of each cluster on a purely local basis? We cannot. The interconnected nature of graphs precludes such a strategy. Clusters can be connected to the rest of the graph in many ways, and the manner in which they are connected will have an immediate effect on the short-term behavior of the cluster itself.

Consider a case in which the cluster in question is composed of three nodes in the susceptible compartment, each of which has some non-zero disposition to acquire the disease upon contact with an infected individual. The cluster is either connected to the rest of the graph or it is not. If not, then the transmission probability drops to zero. If purely local considerations predict a positive probability of infection, but considerations of connectivity tell us that the true probability is zero, then the unit in question is not sufficiently independent from the rest of the population to serve as unit of decomposition. If the cluster is connected (which is the more realistic case for most biological populations) the susceptibility of each node depends on whether the external node to which it is connected is itself infected. Since we are assuming that all edges represent short-term interactions, this dependence obtains at a very short time interval. As a result, ad hoc taxonomies give us no reason to think that epidemiological populations might be decomposable after all. If we want to reason effectively about

system behavior, therefore, it is necessary to tackle the interconnectedness of the population more directly.[14]

From this discussion, it is already evident that there is a comparative sense in which the epistemic benefits of network representation are particularly helpful in the domain of non-decomposable systems. Other approaches to complex systems rely on part-whole decomposition, but such approaches fail when applied to non-decomposable systems. Network science therefore offers unique benefits in this area. But there is a more important and more interesting sense in which network representations have special applicability in the domain of non-decomposable systems. To articulate the idea, it will be helpful to introduce another model.

5.3 A model of urban traffic density

Consider the question of why some particular road in a city has the highest occurrence of traffic jams. One way of answering this question is to examine the kinds of institution to which the road provides access. The more traffic attracted by those institutions, the more traffic will appear on the roads that lead to it. This is known as the inherent travel demand on a road. There is another, network-oriented approach to answering this question that has recently become popular. The network approach determines probability that a road will become congested by examining its location relative to all other roads, regardless of the kinds of institution to which it provides access. In a recent paper, Wang et al. (2012) show that by incorporating information about how centrally a road is situated in graph-theoretic space, they can better predict traffic patterns in San Francisco and Boston.[15]

The extent to which a road segment occupies a central place in the city grid is measured in terms of a mathematical property of networks known as *betweenness*. Betweenness is a property of a single edge in a graph. To compute the betweenness of an edge, a search algorithm examines every pairwise combination of nodes in the graph, and finds the shortest path length between each pair. It then searches the resulting data structure to determine what proportion of those paths incorporate the road segment in question. That proportion is the betweenness of the edge. It is not difficult to see how this mathematical measure can be converted into an empirical one. Edges represent road segments, which are defined as stretches of road between legal intersections, and nodes represent the intersections themselves. The movement of automobiles is represented with a dynamical system not entirely unlike the one discussed in the epidemiological case above. At each time step in the simulation, a number is assigned to a road segment that represents the number of automobiles occupying that segment. Wang et al. show that the traffic density on a road segment

---

[14] It is important that many epidemiological models assume constancy of infection time. If the target population had, for example, a bi-modal distribution of infection times in which one cluster of organisms is disposed to become infected quickly and another is disposed to become infected slowly, it is likely that some of the collective dynamical behavior could be explained on the basis of a decompositional strategy.

[15] The mathematical measure is sometimes called betweenness and sometimes called betweenness centrality. I use the term "betweenness" to refer to the mathematical property and the term "road centrality" to refer to an empirical property of a road.

can be predicted better by modeling both the road's centrality and inherent travel demand than it can by modeling inherent travel demand alone.

5.4 Betweenness and complete non-decomposability

A closer look at the betweenness relation reveals quite a lot about the distinctive nature of network science. On the one hand, it is a purely local property that applies only to a single edge, and on the other hand, its value is only visible once a complete search of the graph has been conducted.

It is impossible for a single edge between two otherwise disconnected nodes to instantiate the property of betweenness because it is undefined where no other edges exist (on pain of division by zero). Logically speaking, the betweenness of an edge is a "collapsed" relational property between that edge and every other edge in the graph. As such, where betweenness is employed in an empirical setting, it carries information about the *complete* pattern of interaction in the system to which it is applied. This fact is epistemologically significant, and can help to illustrate the sense in which network representations have special applicability in the domain of non-decomposable systems.

Recall the definition of non-decomposable systems: a system is non-decomposable just in case the behavior of any given component part, even over a short time period, depends on the behavior of many other individual components. Above, I argued that the study of non-decomposable systems forces scientists to use abundant representations. Some representations are more abundant than others, and among the class of non-decomposable systems, we can define a subclass for which appropriate representations must be maximally abundant. Call these systems completely non-decomposable. To define this subclass, simply replace the term "many" in the definition of non-decomposable systems with the term "every." Where the behavior of every other component in the system must be consulted in order to predict the behavior of a single component, the task of predicting general system behavior is unusually daunting since, in order to carry out the task, one must face the problem of combinatorial explosion in its most radical form.

Now we are in position to see why properties such as betweenness make network representations particularly valuable in the domain of non-decomposable systems: even in the most extreme case of complete non-decomposability, where combinatorial explosion is maximal, properties such as betweenness offer us perfectly serviceable explanantia. In virtue of the fact that betweenness "collapses" information about the entire graph, it makes even completely non-decomposable systems epistemically accessible. Once we know that the centrality of a road segment in a city is predictive of local traffic density, the fact that the task of measuring centrality requires a computationally demanding global search of the city grid makes little difference to our ability to understand why the most central road is likely to have one of the highest values of traffic density. This shows that network measures such as betweenness are especially well-suited to modeling non-decomposable systems. They compress what appear at first to be hopelessly complex patterns of interaction into meaningful variables, and those variables often have considerable explanatory power.

I have been arguing that network representations are particularly well suited to modeling non-decomposable systems. The argument presented thus far can be strengthened a bit by noting one additional point about the nature of betweenness. To make the point, it is necessary to introduce another strategy for dealing with non-decomposable systems that is nicely articulated in Strevens (2005). Strevens describes a method he calls enion probability analysis, or EPA. Roughly, EPA is a strategy for aggregating probabilities at the level of individual elements in a non-decomposable system in order to make predictions about macrolevel dynamics. The case described above of computing internal energy in a physical particle system is an example of EPA. Given that kind of approach to aggregation, information about individual behaviors is lost. EPA does not, therefore, support inferences from the aggregated behaviors of individual elements to the behavior of any particular element within the aggregate. However, this kind of many-to-one influence is one of the hallmarks of a highly interconnected system, and it is complex patterns of influence like this that demand a novel approach. Thankfully, network representations do allow predictions of this kind. In Wang's model, the centrality of a road segment can, for example, be recruited to explain why a particular location in Boston has as much traffic as it does.

In highly non-decomposable systems, this kind of influence between the pattern of interaction, one the one hand, and some particular local behavior, on the other, makes a real difference. Aggregative approaches that fail to explicitly model inter-element interaction cannot capture that difference. Network representations can, and this gives us yet another reason to believe that they are particularly well suited to explaining non-decomposable systems.

6 Networks and mechanistic explanation

In conclusion, I would like to briefly compare the strategy suggested here with the strategy suggested in a recent paper by Levy and Bechtel (2013). Levy and Bechtel show how network representations can be used to model the organization of mechanisms. Their discussion focuses on gene expression in small-scale networks, where patterns of interaction appear to make a genuine explanatory difference. The philosophical aspect of their discussion emphasizes a contrast between two views about the role of system organization in mechanistic explanation. On the view they criticize, abstract principles of system organization can at best serve as templates for mechanistic explanation. The real explanatory work is done only when details regarding the material properties of the component parts (size, shape, chemical reactivity, etc) are filled in. Their own view, by contrast, is that some mechanistic explanations are actually driven by considerations of system organization. Sometimes, mechanistic explanations operate almost exclusively at a high level of abstraction, where very little in the way of detail regarding the material properties of system components is included.

The description of network science provided in this article has much in common with the view that Levy and Bechtel defend. But there is a difference of emphasis that might be misinterpreted as disagreement. Where Levy and Bechtel describe the use of network representations as an extension of the mechanistic project, the models described here have very little to do with the discovery of mechanisms. Nevertheless,

I see our views as more or less compatible. Levy and Bechtel have described one scientific application of network thinking, and I have described another.[16] Because the phenomena Levy and Bechtel focus on have only five or six components, the modeling efforts they describe do not face the problem of combinatorial explosion. The epistemological problem they are trying to solve is therefore quite different from the one articulated here. Where combinatorial explosion is not a problem, it is natural to envision the application of network representations as an extension of a mechanistic strategy.

Although I am therefore largely in agreement with Levy and Bechtel, I do not want to advocate some kind of mushy pluralism. There is a hard line to be drawn: the strategy presented in this article cannot be subsumed within the mechanistic framework because it applies principally to systems that are highly non-decomposable. As Bechtel and Richardson explain in their wonderful book on complexity, the mechanistic strategy for explanation assumes that near-decomposability obtains. The assumption is necessary because the mechanistic strategy depends on localizing particular activities to particular substructures. In order for the localization strategy to work, each substructure must contribute in a relatively consistent fashion to the behavior of the whole. If there are no mechanically distinct substructures that realize some particular "activity," then mechanistic models cannot explain (Bechtel and Richardson 1993, p. 203).

Further support for the claim that localization and decomposition are essential to the mechanistic program can be found by examining the criteria proposed in a 2010 paper by Bechtel and Abrahamsen. The goal of the paper is to incorporate dynamical and organizational features into the mechanistic account of explanation. Even in that context, their explicit characterization of mechanism reads, "A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization" (Bechtel and Abrahamsen 2010, p. 323). In order to generate a mechanistic explanation, therefore, one must be in position to individuate the relevant components and provide evidence that associates components with specific operations. That both of these goals must be achieved is supported by the observation that they are necessarily interdependent. Part of the evidence that a particular component is mechanistically relevant is the fact that it is responsible for carrying out a particular operation. Of course, one could simply stipulate that mechanistic explanation is possible without any commitment to identifying parts and operations, but that kind of bare stipulation threatens to take the normative bite out of the mechanistic program. Chemero and Silberstein make a similar point. They emphasize that if components cannot be identified, or if no stable role can be assigned to components, the mechanistic approach to explanation breaks down (Silberstein and Chemero 2013, p. 961).[17]

---

[16] As a matter of fact, Levy and Bechtel mention the Watts and Strogatz model as well as the concept of a small-world, apparently because the concept has played such a major role in helping to popularize network science. But they quickly set aside such "large-scale statistical models" and explicitly restrict their discussion to much smaller and non-statistical models.

[17] Thanks to an anonymous referee for alerting me to this article, which has a lot in common with the discussion in this section. Given the overlap in emphasis, it is worth noting three points of contrast. First, the Silberstein and Chemero article is aimed exclusively at explanations in biology and neuroscience. My aims are much broader. Second, it is my view that network models have the best chance at explaining in cases in which the individuation of nodes is unproblematic and can be relocated to the background (this point

This hard line suggests something like a division of labor. Recall the quote from Barabasi given at the outset. Barabasi claims that network science will become the foundation of a general theory of complex systems. Insofar as there are many mechanistic explanations of complex systems that do not involve the kind of network thinking that Levy and Bechtel discuss, Barabasi's claim is too strong. It would be wrong to call a modeling framework the foundation of a theory if many of the systems to which the theory applies cannot be captured within that modeling framework. We can weaken Barabasi's claim a bit by thinking about the principles of network science in terms of heuristics rather than fundamental theory. The mechanistic tradition in the philosophy of science has often construed the norms of mechanistic explanation in this way, and it may be useful to take a similar attitude toward network science.

The division of labor I envision is this. When we are dealing with large, robustly non-decomposable complex systems in which combinatorial explosion cannot be avoided, the network strategy described here can be seen as a defeasible heuristic for conducting research. For decomposable or nearly-decomposable complex systems, mechanistic research is an appropriate heuristic, despite the fact that in many cases it must be supplemented by network-like representations.[18]

A committed mechanist might object to this proposal on the grounds that the individual elements that compose a complex system must themselves be subject to mechanistic analysis before the kind of reasoning described here can be carried out. As Alan Baker has stressed in a recent paper, it is difficult to know how to apply network concepts without some theoretical framework that offers individuation principles for the phenomena that correspond to the nodes and edges in a graph (Baker 2012). A committed mechanist might build upon Baker's observation by claiming that it is necessary to have a mechanistic model of the individual components and relations in a system before any network representations can legitimately be applied. If this is right, then mechanistic explanation is a necessary condition on the successful application of

---

Footnote 17 continued

is discussed below). In systems neuroscience—which receives the bulk of the attention in the Silberstein and Chemero article - the task of individuation is often quite controversial. This is not to say that examples in the literature are inadequate, but to suggest that there are explanatory hurdles in neuroscience over and above those typically present in the social and epidemiological sciences. Third, Silberstein and Chemero are interested in showing that the style of explanation they see in systems neuroscience is something "more lawlike" than what the mechanists propose (p. 970). In my view, the relation between law-like and mechanistic forms of explanation is not mutually exclusive. The explanatory status of the cases I have described depends neither on a commitment to predictivism nor on the role of nomological necessity. Rather, they are cases in which we have empirical justification that a mathematical model accurately represents a particular empirical phenomenon. The accuracy of the representation justifies certain explanatory inferences. The question of how mathematics matches up with reality is both more general and more difficult than the question of how network representations explain. Good discussions of this topic that conform to the perspective defended here can be found in Pincock (2012) and Humphreys (2014).

[18] In fact, a position like this was anticipated by Bechtel and Richardson two decades ago. Describing a continuum of cases in which the pattern of interaction among elements plays an increasingly greater role in understanding system behavior, they conclude: "If organization becomes even more dominant in explaining the behavior of the system, and we appeal less to different and distinctive functions performed by the components, we reach a point where decomposition and localization have to be surrendered" (Bechtel and Richardson 1993, p. 199).

network representation. It follows, the objection goes, that network explanations are always extensions of mechanistic explanations.[19]

For the sake of argument, let us grant the premise that it is always necessary to have a mechanistic model of the individual components of a complex system before its organizational properties can be modeled. From this premise, however, the conclusion that network explanations are always extensions of mechanistic explanations does not follow. The modeling strategy described here has nothing to do with the kinds of decompositional analysis that are central to the mechanistic project. To see this, we must first remember that explanations are not generally transitive. So the fact that we have a mechanistic explanation of the individual components of a complex system is no guarantee that we have an explanation of the system itself, mechanistic or otherwise. To show that a mechanistic explanation at the systems level is available, more evidence is needed. Moreover, if by "mechanism" we have in mind the understanding of mechanism discussed in the literature on mechanistic explanation in biology, that additional evidence will have to point to the existence of an intermediate level of organization at which independent component subsystems can be identified. As was argued in Sect. 5.2, however, there are cases in which no such intermediate level exists. In those cases, mechanistic explanation is just impossible.

Although the argument offered here on behalf of the committed mechanist turns out not to be compelling, it does hint at the fact that in many cases, some mechanistic understanding of the components in a complex system is extremely valuable. It is worth reflecting on why this is so. Perhaps mechanistic approaches are particularly useful when the elements that compose a system are not yet well understood. If we attempt to apply network representations in cases where our understanding of the individual components is very poor, we are likely to misrepresent the system, and draw bad inferences as a result. But in traffic science, epidemiology, computer science, microeconomics, and other areas, we have quite a lot of mechanistic knowledge about individual interactions either because they are straightforwardly observable (paying for milk at the grocery store is a paradigmatic economic exchange) or because they are constructed by us (we construct the circuit boards that execute computations). But partly because that kind of knowledge is relatively unproblematic in comparison with other scientific concerns, we may relegate it to the background. There is a sense in which we can legitimately take such mechanistic knowledge for granted because it does not need to be explicitly represented when we construct new explanations, even if those explanations implicitly depend on the fact that certain mechanistic details are part of the fabric of our background knowledge. Perhaps it is in part because we already have an intimate mechanistic understanding of certain kinds of interaction in a complex system that we are in position to know that certain details can be left out

---

[19] Baker also argues that one of the dangers associated with network science is the unwarranted reification of networks. The fact that network representations are explanatorily useful does not entail that network-like empirical phenomena constitute a natural kind. This is one reason that I have stressed the epistemic, rather than ontological contribution of network science. In areas such as molecular biology, neuroscience, and perhaps even ecology, it is reasonable to treat networks as real representation-independent features of the empirical world. In many other areas, however, including those discussed in this article, it is far from clear that this realist attitude is appropriate. In any case, arguments for such a view would require premises that are not clearly supported by the present discussion.

(for certain purposes). In such cases, we stand to gain additional understanding by focusing on the unique pattern of interaction among a system's many elements.

References

Baker, A. (2012). Complexity, networks, and non-uniqueness. *Foundations of Science*, *18*(4), 1–19.

Bansal, S., Grenfell, B. T., & Meyers, L. A. (2007). When individual behaviour matters: Homogeneous and network models in epidemiology. *Journal of the Royal Society Interface*, *4*(16), 879–891.

Barabasi, A. L. (2012). The network takeover. *Nature Physics*, *8*(1), 14–16.

Batterman, R. W., & Rice, C. C. (2014). Minimal model explanation. *Philosophy of Science*, *81*(3), 349–376.

Bechtel, W., & Abrahamsen, A. (2010). Dynamic mechanistic explanation: Computational modeling of circadian rhythms as an exemplar for cognitive science. *Studies in History and Philosophy of Science Part A*, *41*(3), 321–333.

Bechtel, W., & Richardson, R. C. (1993). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Princeton, NJ: Princeton University Press.

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine*, *357*(4), 370–379.

Craver, C. F. (2006). When mechanistic models explain. *Synthese*, *153*(3), 355–376.

Cunningham, W. P., Freeman, D., & McCloskey, J. F. (1984). OR forum—Of radar and operations research: An appreciation of A. P. Rowe (1898–1976). *Operations Research*, *32*(4), 958–967.

Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets reasoning about a highly connected world*. New York: Cambridge University Press.

Gross, J. L., & Yellen, J. (2006). *Graph theory and its applications*. Boca Raton: CRC Press.

Humphreys, P. (2014). Explanation as condition satisfaction. *Philosophy of Science*, *81*(5), 1103–1116. doi:10.1086/677698.

Huneman, P. (2010). Topological explanations and robustness in biological sciences. *Synthese*, *177*(2), 213–245. doi:10.1007/s11229-010-9842-z.

Jones, N. (2014). Bowtie structures, pathway diagrams, and topological explanation. *Erkenntnis*, *79*(5), 1135–1155. doi:10.1007/s10670-014-9598-9.

Ladyman, J., Ross, D., Spurrett, D., & Collier, J. (2007). *Every thing must go: Metaphysics naturalized* (1st ed.). Oxford: Oxford University Press.

Lange, M. (2009). Dimensional explanations. *Noûs*, *43*(4), 742–775.

Levy, A., & Bechtel, W. (2013). Abstraction and the organization of mechanisms. *Philosophy of Science*, *80*(2), 241–261.

Pastor-Satorras, R., & Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Physical Review Letters*, *86*(14), 3200–3203.

Pincock, C. (2012). *Mathematics and scientific representation*. Oxford: Oxford University Press.

Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern epidemiology* (3rd ed.). Philadelphia: Lipincott, Williams, & Wilkins.

Rual, J.-F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., et al. (2005). Towards a proteome-scale map of the human protein-protein interaction network. *Nature*, *437*(7062), 1173–1178.

Schwikowski, B., Uetz, P., & Fields, S. (2000). A network of protein-protein interactions in yeast. *Nature Biotechnology*, *18*(12), 1257–1261.

Silberstein, M., & Chemero, A. (2013). Constraints on localization and decomposition as explanatory strategies in the biological sciences. *Philosophy of Science*, *80*(5), 958–970. doi:10.1086/674533.

Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, *106*(6), 467–482.

Strevens, M. (2005). How are the sciences of complex systems possible? *Philosophy of Science*, *72*(4), 531–556.

Supekar, K., Menon, V., Rubin, D., Musen, M., & Greicius, M. D. (2008). Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. *PLOS Computational Biology*, *4*(6), e1000100.

Wang, P., Hunter, T., Bayen, A. M., Schechtner, K., & González, M. C. (2012). Understanding road usage patterns in urban areas. *Scientific Reports*, *2*, 1–6.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*(6684), 440–442.

Weisberg, M. (2007). Three kinds of idealization. *The Journal of Philosophy*, *104*(12), 639–659.

Weslake, B. (2010). Explanatory depth. *Philosophy of Science*, *77*(2), 273–294.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.