Only Connections?

Alan Baker
Department of Philosophy
Swarthmore College
abaker1@swarthmore.edu

"Only connect! That was the whole of her sermon. Only connect the prose and the passion, and both will be exalted, and human love will be seen at its height. Live in fragments no longer. Only connect, and the beast and the monk, robbed of the isolation that is life to either, will die." (E.M. Forster, *Howard's End*)

Figure 1: Kneiphof Island, Kaliningrad (2008)

Introduction: the Bridges of Königsberg

Our story begins in 1735, at which point in history the Russian port city of Kaliningrad, on the Baltic Sea, was part of Prussia and was known by its original Germanic name of Königsberg. Strategically position near the mouth of the Pregel River, Königsberg had been a bustling commercial port and trading center since its founding in the early middle ages. As the Pregel river flowed around Kneiphof and a second island, it divided the Königsberg into four distinct districts. Linking these districts together were no less than seven bridges, this number a testament to the thriving economic base of the city, as were their names: Blacksmith's Bridge, Connecting Bridge, Green Bridge, Merchant's Bridge, Wooden Bridge, High Bridge, and Honey Bridge.

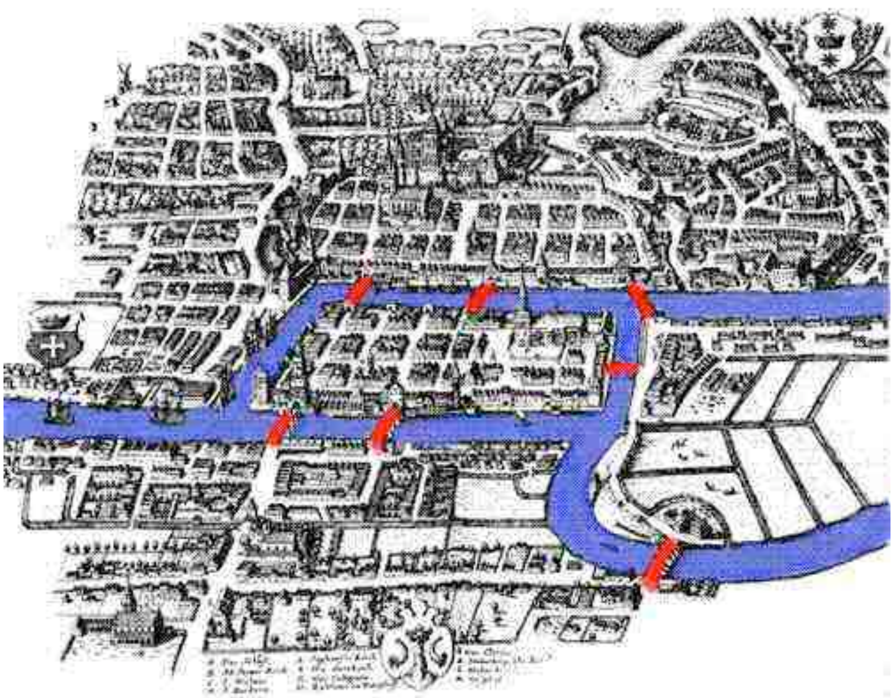

Figure 2: Kneiphof Island, Königsberg (1735)

According to legend, the residents of Königsberg – who liked to spend their Sunday afternoons strolling around the center of their beautiful city – became intrigued by the question of whether it was possible to complete a tour of the city crossing each of its seven bridges exactly once. Fortunately for them, Königsberg was not far from another important city on the Baltic, namely

St. Petersburg. And St. Petersburg was at this time home to Leonard Euler, one of the most prolific, versatile, and influential mathematical minds of his (or indeed any) age.

Euler was initially skeptical as to whether the problem was anything more than a relatively trivial brainteaser. In a letter to the Mayor of Danzig, who had asked him for a solution, Euler wrote

> "Thus you see, most noble Sir, how this type of solution bears little relationship to mathematics, and I do not understand why you expect a mathematician to produce it, rather than anyone else, for the solution is based on reason alone, and its discovery does not depend on any mathematical principle."

Despite this dismissive attitude, Euler continued to think about the problem. He eventually concluded that, although he was correct in his initial verdict that it bore "little relationship to mathematics," this said more about the impoverished scope of existing mathematics than it did about any unsuitability in principle of the problem to be addressed by mathematical methods. What Euler realized was that although the Königsberg Bridge problem was ostensibly geometrical in flavor, it differed from the canonical problems of Euclidean geometry in that *distance* was irrelevant. It clearly makes no difference to the essence of the problem whether Kneiphof Island is 1 mile wide or 100 miles wide, nor whether the Blacksmith's Bridge is longer or shorter than the Merchant's Bridge. The irrelevance of distance can be seen from Euler's own diagrammatic rendering of the situation in the paper he ended up writing on the puzzle of the bridges of Königsberg, in which no distances or other numerical scales are shown.

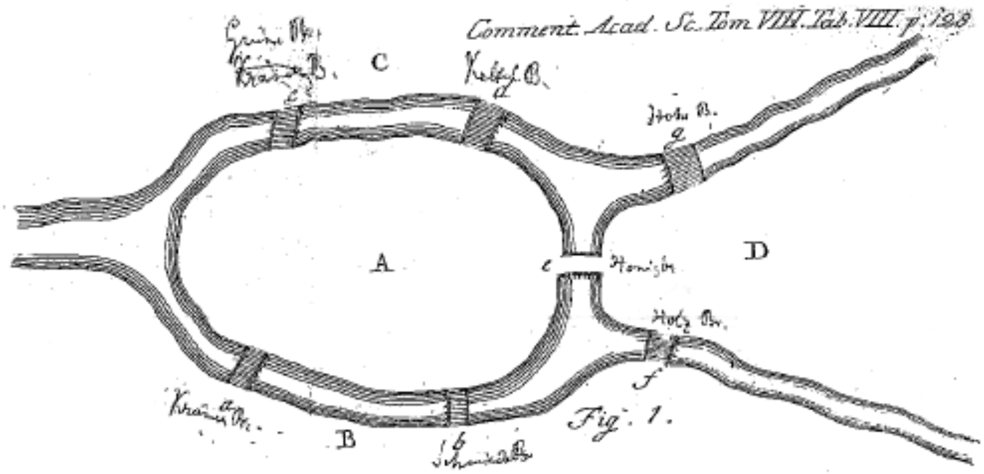

Figure 3: Euler's diagram of the Bridges of Königsberg (1736)

What was needed, therefore, was a form of geometry that abstracted away from any features associated with distance. No such mathematical theory had yet been developed, however Euler recalled some brief remarks that Leibniz had made over fifty years earlier that pointed in an apparently similar direction. As Euler ended up writing in the introduction to his 1736 paper,

> “In addition to that branch of geometry which is concerned with distances, and which has always received the greatest attention, there is another branch, hitherto almost unknown, which Leibniz first mentioned, calling it the geometry of position. This branch … does not involve distances, nor calculations made with them. It has not yet been satisfactorily determined what kinds of problem are relevant to this geometry of position, or what methods should be used in solving them. ”

Euler wrote this paper in Latin, thus “the geometry of position” was rendered as *Geometriam situs*. It was another century before the hellenized version of this term appeared in print, in Johann Benedict Listing’s 1847 treatise *Vorstudien zur Topologie*, and not until 1883 that the term “topology” was used, in the journal *Nature*, to distinguish “qualitative geometry from the ordinary geometry in which quantitative relations chiefly are treated." To this day, Euler's 1736 paper is widely viewed by topologists as the first substantive result in their field.

Graph Theory (I)

Remarkably enough, topology is not the only subfield of mathematics to trace its source to Euler's paper on the bridges of Königsberg. For the proof that Euler gives in this paper is also regarded as the seminal result in the subsequently developed field of *graph theory*. The term "graph" is a confusing one, since it tends to suggest to non-mathematicians the familiar technique of plotting one variable against another on an *x*-axis and *y*-axis. In the technical, mathematical sense, however, a graph is simply a collection of *nodes* (or vertices), together with a collection of *links* (or edges) that connect certain of these nodes to one another. An example is shown below of a graph with four nodes and seven links. A graph-theoretic property that will be important for later discussions is *degree*. The degree of a node is simply the number of links that connect to that node. (Thus the bottom node of the graph shown in Figure 4 has degree 3.)

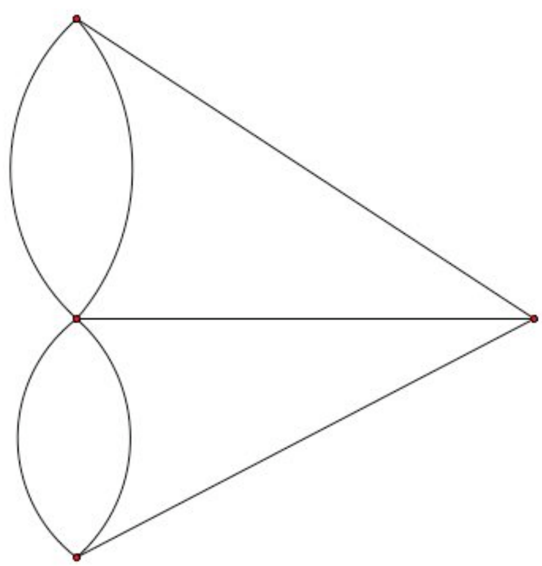

Figure 4: An example of a graph

What do graphs – in this mathematical sense – have to do with the Königsberg bridge problem? The connection involves *connections*. On the one hand, the links in graphs can be viewed as connections, and thus the graph shows which nodes are connected to which other nodes. On the other hand, what are bridges, fundamentally, other than structures connecting one otherwise inaccessible location to another? Nor is the connection here merely an idle curiosity. It turns out that the elements of the graph shown in Figure 4 can be matched up with features of the Königsberg bridge scenario, and thus the graph can be viewed as encapsulating, in abstract mathematical form, the essential features of the original, real-world problem.

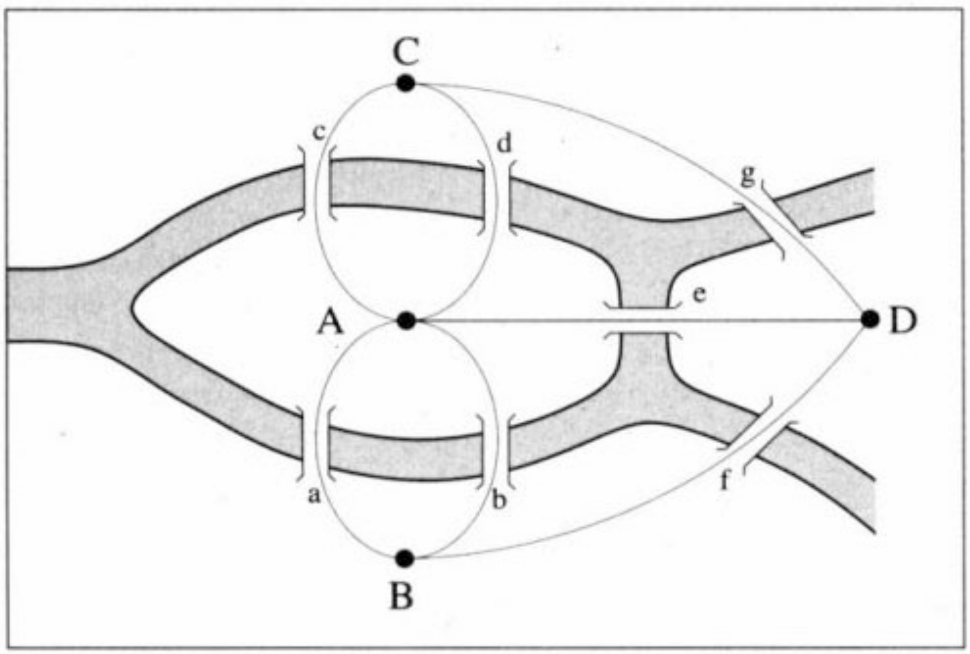

Figure 5: A graph superimposed on a map of the bridges of Königsberg

The received view, according to which Euler's 1736 paper on the Königsberg bridge problem marks the historical starting-point for the modern mathematical field of graph theory, has a couple of ironic features. Firstly, there is no evidence that Euler himself drew a graph at any point during the process of solving the Königsberg conundrum. There are certainly no graphs in the published version of the 1736 paper, nor have any been found in Euler's unpublished papers or in his extensive correspondence. Secondly, the subsequent development of graph theory from the late 19th Century to the late 20th Century proceeded in almost total isolation from the sort of 'real-world' problem that prompted Euler's initial foray into the field. The point, in other words, is that the link between graphs – as abstract mathematical representations – and systems of objects in the world – such as bridges, islands, and rivers – is more tenuous, historically speaking, than might be suggested by the neat overlay depicted in Figure 5 above.[1]

Graph Theory (II)

Until very recently, graph theory occupied a place at the 'pure' end of pure mathematics. Graphs were curious mathematical structures whose properties were explored as objects of interest in their own right, without much if any attention being paid to potential applications outside mathematics. Of course, the history of mathematics is rife with examples of apparently completely 'pure' subfields of mathematics, pursued for their own internal ends, subsequently finding – often utterly unexpectedly – application to the real world.[2] Thus the fact that graph theory was more-or-less completely divorced from applications during the bulk of its theoretical development did not mean that the properties of graphs that were focused on by mathematicians did not have significant real-world counterparts. For example, one crucial topological property of graphs is *connectedness*. A mathematical graph is *fully connected* if it is possible to move, via some sequence of links, from any node to any other node. Intuitively it makes sense that as the number of edges in a graph increases so does the chance of the graph being fully connected. A fairly simple model for graphs (called the Erdos-Renyi random graph model) assigns a fixed probability for each node to connect to each other node. If this probability is 0 then there are no edges at all (i.e. the degree of each node is 0), as shown in Figure 6(a) below, for a graph with 20 nodes.[3] If the probability of interconnection is 0.1, then the average node degree is 2, and clusters of interconnected nodes start to form. [Figure 6(b)] With a probability of interconnection of 0.2

---

1 To be sure, both of the points of irony just articulated can be expressed in a more nuanced, less stark fashion. Thus, although there are no graphs *per se* appearing in Euler's 1736 paper, there is definitely reasoning expressed in this paper which is 'graph-theoretical' in nature. And, although tracing graph theory back to the Königsberg bridge problem emphasizes an historical link to real-world problems, the motivating puzzle in this case is both socially constructed and, ultimately, of no great import even from the parochial point of view of the citizens of Königsberg themselves.

2 Two important examples involving 20th-century physics are non-Euclidean geometry (which was crucially involved in Einstein's development of his General Theory of Relativity), and group theory (which ended forming an important theoretical framework for particle physics).

3 Recall that the *degree* of a node is just the number of links that connect to that node.

the average node degree is 4; in this case there is a giant component that links together all but one of the graph's 20 nodes. [Figure 6(c)]

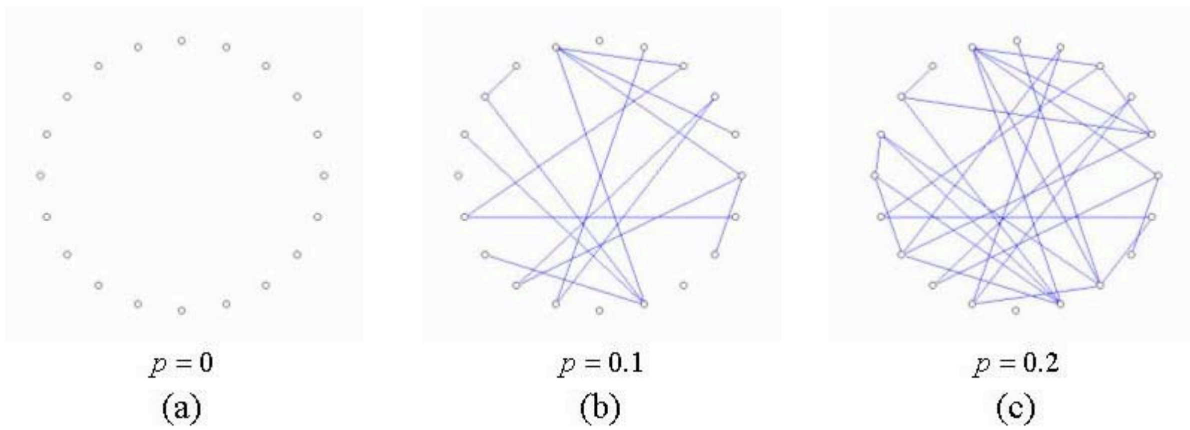

Figure 6: Random Graphs

Going back to the original puzzle that motivated Euler's 1736 paper, it is clear that connectedness plays a crucial role. Indeed the Königsberg bridge problem could not even have been posed unless the areas in question (i.e. islands and river banks) were all connected in some way by bridges.

The 'random graph' model presented above provides a neat way to explore the relationship between connectedness and average node degree. But although connectedness has clear importance for real-world systems, there is no reason to think that connections in such systems are formed at random as implied by the Erdos-Renyi model. However, quite independently of developments in graph theory, sociologists had begun in the mid-20th century to look at what they called “social graphs.” Unlike the graphs drawn by pure mathematicians, sociologists’ graphs are extracted from empirical data concerning real-world systems. The nodes in these social graphs represent people, and the links correspond to some specified social relationship, such as friendship, or kinship, or being work colleagues.

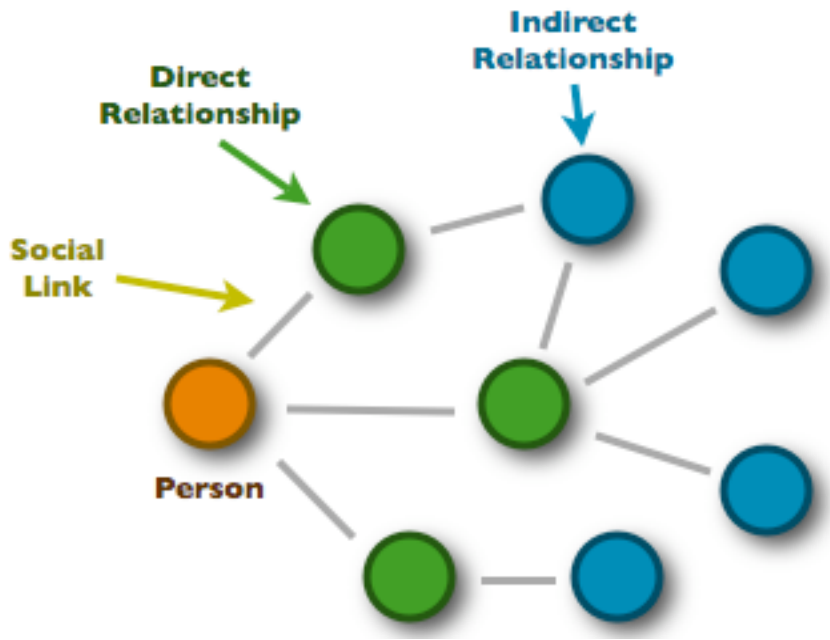

Figure 7: A social graph

A second difference, at least historically speaking, is that the social graphs actually constructed by sociologists tended to be relatively small. At first glance one might attribute this fact to the logistical constraints on human social interactions. For there are surely limits to the number of people with whom a given individual can be meaningful connected by social ties such as friendship. However, this cannot be the whole story because all that is constrained here is the number of *direct* relationships a person can have, in other words the number of other people who are 'one link away' in the corresponding social graph. There is no limit, in principle, to the number of people to whom an individual can be indirectly connected through some chain of links. This can be seen even in an artificially simplifed graph such as the one depicted in Figure 7. In this graph, no node is directly linked to more than four other nodes. So the maximum number of direct relationships a person can have is 4. But a person can have up to 16 indirect relationships that are 'one-step removed', 64 two-step indirect relationships, 256 three-step relationships, and so on. Another way of putting the point is that there is no limit to the number of nodes that can be present in a graph even if the number of links emanating from each node (i.e. the degree of each node) is tightly constrained.[4]

If many actual networks of human social interaction are large – think of the thousands of interacting nodes in a university or a corporation or a town – then why was this not reflected in at least some of the social graphs produced by sociologists? The problem, in essence, was data collection. For the pure mathematician using the Erdos-Renyi model, it makes little difference whether the graph being constructed has 20 nodes or 2,000 nodes. Since the links to and from each node are generated randomly (using whatever probability has been decided upon in advance), all nodes can be treated in the same way, and the graph can be built up quickly by iterating a simple algorithm. For the sociologist, by contrast, the links between nodes in a social graph can only be reliably determined by observing each individual node, i.e. person, in the network. This is a slow and laborious process, hence drawing a social network with 2,000 nodes might well take several years of fieldwork.[5]

The Birth of the Science of Networks

The history of the graph between Euler and the latter part of the 20th Century thus had a strangely disjunctive character. On the one hand, there were the abstract, generic, often large graphs studied by pure mathematicians. On the other hand, there were the data-driven, idiosyncractic, relatively small graphs studied by sociologists. What happened during the thirty year period between the late 1960's the late 1990's was a series of developments that led to a fusing of these two, hitherto very much distinct kinds of graph, and with it the birth of a new interdisciplinary field that has come to be known as *network science*.

---

[4] There is also active debate among sociologists concerning what limits there are, if any, on the number of nontrivial social relations of a given kind a particular person can have. Anthropologist Robin Dunbar has proposed 150 as the (approximate) upper bound on the number of direct, stable social relationships an individual person can have, using arguments based on human brain size in comparison to other primates.

[5] There are ways of compressing the burden of data collection, for example by asking people about their past and current social interactions. The problem is that people turn out not to be very reliable at generating lists of everyone with whom they interact, or at remembering details of the timing and frequency of past interactions. (This is another aspect of Dunbar's point about the 'cognitive load' of keeping track of social relationships.)

The gaping hole in the history of graph theory was the almost total ignorance about what large, 'real-world' graphs might look like: mathematicians had large graphs, and sociologists had real graphs, but no-one had both. That this hole was significant began to dawn on researchers in the late 1960's, prompted by revelations at both the empirical and theoretical ends of the spectrum.

At the empirical end, the evidence was anecdotal but also very suggestive. American psychologist Stanley Milgram published an influential paper in 1967 that described a series of experiments designed to learn more about the probability that two randomly selected people know each other. Packages were sent to randomly selected individuals in Omaha and Wichita with the name of a target person living in Boston. They were instructed to forward the package to someone they knew on a first-name basis who, in their opinion, would be more likely to know the target person. This process was repeated until the package reached its target. Of the 64 packages that did eventually reach their target, the average number of steps in the chain was just under six. This was later to give rise to the popular phrase “six degrees of separation.”[6] From a graph-theoretic perspective, Milgram’s quest was to find the *average path length* between the nodes of a social graph (in this case a graph that includes the entire U.S. population), where the path length between two nodes is simply the minimum number of links that must be traversed to get from one to another.

What started to become clear, on the basis of this and other 'localized' experiments is that large, real-world graphs are often *not* random graphs. For example people in social networks are generally more likely to forge links with those who are geographically closer to themselves. This leads to clusters of interlinking nodes, together with a smaller number of 'long-range' links that connect one local community to another. Many real networks also have the property that nodes are more likely to link to other nodes that are similar to them in various ways (a property known as *assortativity*), while other real networks have the opposite property, whereby nodes are more likely to link to other dissimilar nodes (i.e. *disassortativity*). Most of these properties are statistical in nature, and hence their presence in smaller networks is difficult or impossible to ascertain.

[6] It should be noted that Milgram himself never used this term (or made any general claim of this sort) in his own publications on this topic.

However, the real breakthrough in the study of real networks did not come until the late 1990's, and was the direct result of two technology-related developments. Firstly, large bodies of data became available in electronic form, including information about the structure of huge networks such as the Internet. For the first time, researchers had access to datasets about entire, large networks. Secondly, and relatedly, computer processing power had by this time progressed to the stage where datasets of this size and complexity could be efficiently analyzed within a reasonable amount of time. These developments quickly led to important new theoretical models for large networks, models that were for the first time informed by real data, and with it the first steps in the newly self-defined field of network science.

Network Science and the Privileging of Connection

Three characteristic features of network science during its first decade have been its focus on large networks, its interest in the statistical properties of networks, and its domination by researchers with backgrounds in theoretical physics.[7] There has also been an increasing trend to apply the network approach beyond the domain of narrowly physical networks such as systems of bridges or of power cables. This has seen network scientists bring their theoretical tools to bear on the study of complex systems in biology, computer science, medicine, ecology, linguistics, even literary studies. A perusal of recent publications in network science shows work being done on the mapping out and analysis of brain networks (neurons connected by synapses), food webs (predators connected to prey), communication networks (people connected by phone calls), citation networks (papers connected by citations), archaological networks (sites connected by transport routes), to mention just a few examples.

Nonetheless, the most significant feature of the network approach is undoubtedly its emphasis on **connections**. This emphasis is manifested in various ways. It is part of the rhetoric used by practitioners of network science, as can be seen in the titles of some of the semi-popular books that give an exposition of the field, such as Duncan Watts' *Six Degrees: the Science of a*

[7] The conjunction of these three aspects is not coincidental: as already mentioned, many statistical properties do not make sense in the context of small networks. And a statistics-based approach is common to many areas of theoretical physics, especially condensed matter physics, thermodynamics, and high energy physics.

Connected Age, and Lazlo Barabasi's *Linked: How Everything is Connected to Everything Else and What it Means*.[8] And it is also inherent in the very graphs that form the basic representational devices of network science. The nodes in these graphs have no features other than the way they are linked to other nodes: what makes two graphs the same is that they have the same patterns of links.[9]

As has already been mentioned, another characteristic feature of network science is its crossing of traditional disciplinary boundaries. An important question that may be raised in this context is whether there are any limits to the network approach. Just how broad is the notion of network, and its accompanying mathematical apparatus from graph theory? There are at least two ways of understanding these questions. Firstly, there is an issue of cogency: what kinds of phenomena in the world can legitimately be viewed as networks? Secondly, there is an issue of fruitfulness: are there topics or disciplinary areas for which the network approach, although possible, does not lead to any interesting insights?

The Meaning of 'Connection'

I shall begin by taking up the first of the two questions identified at the end of the previous section: what kinds of phenomena in the world can legitmately be viewed as networks? Given our earlier remarks about the centrality of connections in the network approach, it makes sense to reframe this question in these terms. What, then, can be counted as a connection?

Recall the title of Barabasi's book, *Linked: How Everything is Connected to Everything Else and What it Means*. At first glance, one might think that taking this claim literally threatens to reduce the network approach to triviality. For if it is indeed true that everything is connected to everything else, won't every graph of a real-world system have every node connected to every other node, so all graphs will look essentially the same? This conclusion would be correct if by "connected" here, Barabasi meant "*directly* connected." More plausible is that he is alluding to the distinction we made earlier (in the discussion of social graphs), between direct connections

---

[8] Barabasi, A.-L. (2003) *Linked: How Everything is Connected to Everything Else*, Plume.
Watts, D. (2004) *Six Degrees: the Science of a Connected Age*, Norton.
[9] Another way of putting this point is that graphs are purely *topological*.

and indirect connections. The latter, weaker claim is just that there is that for any two nodes there is some intermediate chain of links via which one can be connected to the other. And this is compatible with there being many different kinds of real-world graphs.[10]

But even to make progress on the question of whether or not everything (or anything!) is connected, we really need first to address the question of what we mean by "connection." What is it for two things in the world to be connected? It seems highly unlikely that there will be a precise and concise answer to this question, a neat definition that distinguishes genuine connections from their spurious counterparts. Rather than attempting anything this ambitious, I shall instead proceed by drawing up a classification scheme under which many (though as we shall see, not all) types of connection can be grouped together.

Classifying Connections

(I) Enduring Physical Link

This has a good claim to be considered the core notion of connection, at least in the context of networks. Returning to the original Königsberg puzzle, we can see that the links in the network are the bridges themselves – solid, physical structures that run continuously between one landmass and another. Other examples that feature in contemporary network science include roads (as connections between population centers, in transport networks), wires (as connections between routers, in the Internet), and axons (as connections between neurons, in brain networks). This sense of connection is also primary in the standard dictionary definition of the term (for example, "1. to join, link, or fasten together; unite or bind: *to connect the two cities by a bridge*").

(II) Transfer of Material

Another significant way in which two things can be connected is via the transfer of material between them. Examples include airline flights between airports (and here we often use the term

[10] Of course, even this weaker claim is somewhat debatable. It is far from clear, for example, that every human being is socially connected, even indirectly, with every other. One set of potential counterexamples are the inhabitants of North Sentinel Island, part of the Andaman Island chain in the Indian Ocean, who have thus far resisted attempts at contact by the outside world.

"connection" to describe such flights), and business transactions between corporations (in which material goods are transferred in return for payment, or for other goods). This sense of connection already involves an extension of the core notion of 'enduring physical link,' since although some material transfers occur via physical linkages (for example, cars traveling along roads, and oil traveling through pipelines), in cases such as the airline example there is no such link.

(III) Transfer of Information

Characteristic of many important networks in the modern world is the linking of nodes via transfer of information. Examples of this sort of connection include cell phone calls as links between individuals in a social network, radio signals as links between communication satellites, and URL hyperlinks as links between html pages on the world-wide web. This is also one of the important secondary senses of "connect" in dictionary definitions of the term (e.g. 2. to establish communication between; put in communication: *Operator, will you please connect me with Mr. Jones?*), though it should be noted that this moves even further away from the core notion of 'solid' physical linkage.

(IV) Colocation

Graphs of real systems are often constructed by drawing links between pairs of nodes that have been located in the same place at some point in time. Sometimes this colocation is literal and specific, as with coworkers working in the same office. Sometimes colocation is meant in a looser sense, as when actors are linked together on the basis of appearing in the same movie.[11] And sometimes the 'colocation' must be understood almost completely figuratively, as when network analyses of literary works link characters whose names are mentioned in the same chapter, or on the same page. Even where the colocation is literal, there will often be some further condition on the nature of the interaction for it to count as a genuine connection. For example, in ecological networks, for a link to be drawn between a predator and a prey it is necessary that they be colocated at some time, but clearly mere colocation is not enough![12]

[11] I take this to be a looser sense because it is possible for two actors to appear in the same movie without ever being physically located in the same place, for example if they do not share any scenes.
[12] A further complication with this example is that ecological networks typically use zoological categories such as species as nodes, rather than individual organisms.

Even this brief survey is enough to give a sense of the many different kinds of relations that can count as connections in different contexts. Whether it is the breadth of topics to which the network approach has been applied that has expanded the original core notion of 'enduring physical link,' or whether it is the innate flexibility in the notion of connection that has contributed to the diversity of applications of network science, is an interesting question but one that I shall not pursue here.

It is also important to be aware of the shortcomings of the above classification scheme. Firstly, the boundaries between the four categories are often blurry. For example, transfer of information (category III) often involves transfer of material (category II), as when a letter is delivered through the mail from one address to another. Similarly, transfer of material (category II) often takes place along enduring physical links (category I), as when cars travel along the highways that connect different cities.

Secondly, there are types of connections in well-established networks that do not seem to fall under any of the above four categories. One important example is the friendship relation in social networks. Even for sociologists, friendship is an elusive notion. Even though it is the canonical way of defining links in social networks, it does not fall clearly into any of the categories of connection that have been defined thus far. With the exception of enduring physical linkage, friendship may involve aspects of all of the remaining three categories (transfer of material, e.g. gift giving; transfer of information, e.g. phone conversations; colocation, e.g. spending time together). But none of these are sufficient in themselves to define friendship.

Another example of a type of connection that lies outside the above categories is citation in networks whose nodes are academic papers. Unlike in the case of friendship, the citation relation is definable very precisely. However it does not seem to fit neatly into any of our previous categories. The link here is not physical, nor does it involve the transfer of anything material, nor need the cited paper be at all similar to the paper in which it is cited. Citation involves the name (and publication details) of one paper being mentioned in another, but it would be a stretch to think of this as colocation, since one paper does not appear in its entirety at the same place as another. The best fit is probably with category III (transfer of information), although even here it is unclear that much – if any – of the informational *content* of one paper need be transferred to the other via the citation relation.

Is Everything Connected to Everything Else?

Armed with our – admittedly imperfect – classification scheme, are we any better placed to answer the question of whether everything is connected to everything else? Given the wide range of binary relations that can be (and are) counted as "connections", it is tempting to wonder whether there are in fact any substantive constraints on this notion, and – if not – whether we are therefore forced into the rather implausible view that everything is directly connected, in *some* sense, to everything else.

Furthermore, the situation is even worse than it might seem from the above survey. For networks often contain links that represent not *actual* connections of one of the four types but *potential* connections. For example, imagine a 'network' of communications satellites each of whose transmitters are pointed at certain other satellites in the network, but amongst whom no actual radio signals are passing. It is natural to treat these transmitter pointings as *connecting* one satellite to another. Yet there is no actual connection, just the potential for a connection of category III (transfer of information). A similar sort of example is an airline network some of whose nodes (airports) are connected by routes with flight numbers, where no planes are actually scheduled on the route.[13]

Faced with this diversity of connections, both actual and potential, the temptation is just to throw up our hands and admit that *any* well-defined binary relation can legitimately be counted as a connection in some context or other.[14] One problem with this sort of capitulation is philosophical. Wittgenstein made the point, in connection with knowledge, that if there is no criterion for distinguishing correct from incorrect applications of a term, then we risk losing our grip on meaning. If every relation is a connection, and any two things are connected, then what can we possibly *mean* by the term? Another problem is more practical. Consider the relation, defined between people, of having at least three letters in common between their respective last names. There is no problem in principle with drawing up a network where people are linked in

---

[13] Yet another example may be many cases of hyperlinks that connect webpages to each other. Many of these hyperlinks may never actually be 'clicked on' and hence remain potential rather than actual connections.

[14] And perhaps undefined / undefinable binary relations too ... .

virtue of this relation holding. Or consider a network in which people are linked together if they share a favorite color, animal, or ice-cream flavor. Such networks might have interesting topological properties, and thus (perhaps) repay mathematical investigation. But there remains a strong suspicion that we would not thereby learn anything interesting or important about the world from such networks. Why not, exactly? The obvious answer is that we have not picked out genuine connections. But this in turn suggests that not *every* definable relation should count as a connection.

What Connections Could Not Be

With an eye to this more general issue, let us pursue a little further the question of why the links in the above networks ought not to count as genuine connections. Here are a couple of suggestions why not. Firstly, links of the above sort are not relevant to the prediction and explanation of the behavior of the nodes: in other words, the presence or absence of such links does not *make any difference*. Drawing a link between Mr. Smith and Ms. Mishra on the basis of the overlap in the letters of their names does not tell you anything interesting about either of these individuals. Secondly, such links are arbitrary and do not correspond to any *natural* features of the world. Why pick on three letters of overlap as the key property, and why care about last names rather than first names? There does not seem to be any good answer to such questions. Much more would have to be said to flesh out these candidate conditions on 'connectionhood' before they could play an effective role in delineating this concept.[15] I won't try to do that here, but I do want to briefly consider a couple of immediate objections based on examples of connections that have already been mentioned, since they seem at first glance to be obvious counterexamples.[16]

The first objection concerns connections that are merely potential, such as the communication satellites' transmitters pointing at one another. How can 'merely' potential connections make a difference to anything? In some cases the potential for connection makes a

[15] One quick point: the term 'natural' as it is used here is not meant to imply that the associated property is unrelated to any human construct, whether physical, social, or conceptual.
[16] Another question concerns the relationship, if any, *between* these two conditions.

difference because it presupposes some other, actual connection in order for the potential to be there. For example, my potential to connect (in the sense of transfer of information) with another telephone user using a landline depends on the presence of this physical connection between one telephone and another. In other cases, however, including the satellite example, there is no facilitating physical connection. But even 'pure' potential connections of this sort can make a difference. For example, they can alter the behavior of agents that make use of a given network. (For example, I may take more risks while out hiking alone if I am carrying a cellphone with me, even if I never actually use the cellphone to make a call. The *potential* I have to connect to others here makes a difference to my actual behavior.)[17]

The second objection concerns apparently acceptable connections in actually studied networks that seem to be highly arbitrary and 'unnatural.' As has already been mentioned, the study of large social networks has been greatly facilitated in the past decade by the growth of various technological means for social interaction, such as mobile phones, email, and online social networking sites. For example, in a 2007 paper, J.-P. Onnela and co-authors analysed 18 weeks of (anonymized) call data from a mobile phone network that had almost 5 million users.[18] Their policy in constructing an associated network was to link two people together if there had been at least one reciprocated phone call between them over the 18 week period. Their argument for doing this was not that a single phone call in each direction constitutes in itself a meaningful social connection, but rather that it "serves as a signature of some work, family, leisure or service based relationship." Thus the mobile phone calls are acting as a *proxy* for a genuine social connection.

There are both obvious advantages and obvious disadvantages to networks that are built up out of proxy connections. One major advantage is that proxies can often be found that are precisely defined, relatively straightforward to measure, and for which large amounts of data are available. This allows for analysis to be carried out and predictions made that would be difficult or impossible if working with the 'target' social connections themselves. Another example of proxy connections are citations between academic papers. Presumably these are intended to

[17] Similarly, in game-theoretic contexts the possible 'moves' available to one player, even when not chosen, may affect the (actual) actions of the other players.
[18] J.-P. Onnela, J. Saramaki, J. Hyvonen, G. Szabo, D. Lazer, K. Kaski, J. Kertesz, and A.-L. Barabasi (2007) "Structure and Tie Strengths in Mobile Communication Networks," *Proc. Natl. Acad. Sci*., 104, 7332-7336.

stand as proxies for the intellectual influence of one paper on another. This is a difficult concept to pin down directly, whereas citation is a cut-and-dried, easily recognizable relation. The major disadvantage of proxy connections is that they turn out not to be reliable guides to the target connections that they are supposed to track. This unreliability may be manifested in two directions. A proxy may 'overgenerate,' in the sense of forming a link even when the underlying target connection is absent. This is a notorious problem for citation networks: it is quite common for one paper to cite another even if the latter has had no influence on the former (for example, if it is being cited to demonstrate the author's familiarity with the background literature, or because of a personal or professional connection between the two authors). A proxy may also 'undergenerate,' in the sense of failing to hold even when an actual target link is present. In the mobile phone network example, there are presumably huge numbers of genuine social relations that hold between individuals but do not result in reciprocated mobile phone calls (for example if one or both do not own a mobile phone, or if they live close to one another and do all their communication face to face).

Let us acknowledge that a connection such as "has made at least one reciprocated mobile phone call in the past 18 weeks" is difficult to classify as *natural*. Even so, it may still be possible to hang on to the naturalness condition for a connection to be genuine by building in an extra clause for proxy connections. Thus for a connection to count as genuine it must correspond to a natural feature of the world *or* be a proxy for some such natural feature. The idea is that this will let in relations such as reciprocated mobile phone calls and citations as genuine connections, since they are proxies for natural relations (namely social acquaintanceship, and intellectual influence), but will exclude relations such as shared letters between last names. Not only is this latter relation arbitrary and artificial, it also fails to track any genuine, natural connection.

Precision and the Illusion of Objectivity

The broadening of the concept of connection, combined with the inherently abstract nature of graphs and networks, has led to the increasing application of network science across traditional disciplinary and even divisional boundaries. In the passage quoted below, Duncan Watts (himself trained as a physicist) gives one perspective on this trend:

> "Physicists, it turns out, are almost perfectly suited to invading other people's disciplines, being not only extremely clever but also generally much less fussy than most about the problems they choose to study. Physicists tend to see themselves as the lords of the academic jungle, loftily regarding their own methods above the ken of anybody else and jealously guarding their own terrain. But their alter egos are closer to scavengers, happy to borrow ideas and techniques from anywhere if they seem like they might be useful, and delighted to stomp all over someone else's problem. As irritating as this attitude can be to everybody else, the arrival of the physicists into a previously non-physics area of research often presages a period of great discovery and excitement."
>
> [Duncan Watts, *Six Degrees*]

Historically, the inter-disciplinary application of networks began with other scientific fields such as biology and ecology, and with social sciences such as sociology and economics. However more recently there has also been a move – albeit to a lesser degree – to apply network science within various humanities-based disciplines.

There is a tendency for graphs of real-world systems to be drawn so that the links represent some precisely determinable relation, such as a rigid physical connection, a mobile phone call, or a webpage hyperlink. In some cases, for example power grids, this determinable connection is also the ultimate connection of interest. However in many cases it is not, and in general the further one moves from graph theory's 'home turf' of mathematical physics, the greater the use of proxy links to stand in for what are (purportedly) the underlying connections at issue. Thus in social networks, as we have seen, sharing a workplace or exchanging emails or phone calls may act as proxy relations for the core social relations such as friendship and cooperation. An even starker example is the literary network, which may be drawn up for a novel or play, and in which characters are linked on the basis of appearing in the same scene, chapter, or other sub-unit of the work. Here, presumably, the proxy relation is standing in for the various underlying social relations that are portrayed in the given literary work.

As was discussed earlier, particular examples of proxy links are often criticized for failing to accurately reflect the underlying connection that they are supposed to be standing in for. For example, not everyone has a substantive social relationship with every other person in their workplace. One response to this sort of criticism is to invoke the *statistical* standpoint that characterizes network science. From this standpoint, it does not matter if some people have no social relationship with some other people in their places of work. All that matters is that there is some (non-trivial) *positive correlation* between working alongside someone else and having some sort of substantive social relationship with them.

However, there is another pitfall arising from the prevalence of proxy links that has nothing to do with their reliability, namely that it encourages what might be termed 'the illusion of objectivity.' The objectivity I have in mind here is of two sorts. Firstly, the construction of graphs of real world systems, with neatly separated nodes and precise interconnections, produces networks which tend (often very rapidly) to acquire the status of **objects** in their own right. This reification of networks is encouraged by some of the terminology that is prevalent within the field of network science, in particular the habit of referring to graphs that represent actual systems in the world (such as an animal brain, or a power grid, or social community) as "real networks." A second meaning of objectivity that comes into play here is the more familiar notion of there being a privileged, observer-independent way of uncovering the truth about a particular phenomenon in the world. This facet of objectivity is fortified by the mathematical pedigree of network science (i.e. its origins in graph theory, as traced above). And it is also encouraged by the terminological practices of network science, especially the use of phrases such as "*the* network of the human brain" or "*the* structure of the Internet." This practice carries with it at least two potentially problematic implications concerning uniqueness. Firstly, that there is a unique structure that underlies each particular real-world system, and, secondly, that this unique structure has been correctly captured by the chosen network representation.

It is important to realize that these two uniqueness claims do not stand or fall together. The first is a metaphysical worry about whether the world itself has a unique structure. The second is an epistemological worry about what is delivered by the methodological practices of network science. As a brief illustration of the latter, consider again the example of a community of mobile phone users. There are many ways to move from the dataset of phone calls to a network, depending on what kind of phone communication is chosen as the basic link (e.g. one-way vs. reciprocated calls) and on whether some threshold is required before a link is drawn between two nodes (e.g. at least one call per week vs. at least three calls per week). Hence, even if there is such a thing as *the* underlying network for this mobile phone community, this does not imply that the given network is the only, or even the best, way of representing that network. Doing this requires defending the various decisions that are made in fixing on particular relations as proxies for the underlying connections of interest.

Return to Königsberg

It is the first, metaphysical worry, however, that I shall pursue in more detail. This is a more purely 'philosophical' worry, about whether reality itself is uniquely structured. What we are asking is whether systems of objects in the world have a structure that transcends our various ways of getting epistemological access to them. One might think that if the answer to this question is 'no' then this is only because we have extended the notion of connection in so many ways and in so many directions that it has become ambiguous and ill-defined. In other words, the reason why we might resist attributing a unique structure to a community of mobile phone users is because we no longer have a clear conception of what it is for one user to be connected, *really* connected to another. And this in turn is because network science has pushed the application of graph theory beyond its legitimate boundaries and into terrain that is vaguer and more metaphorical.

I think that this response is misguided. One way to see why is to return to the starting-point of our story, Königsberg. What could be more well-defined, connectionwise, or more precisely determinable than a bridge? According to the above view, therefore, the bridges of Königsberg must surely have a unique structure. This structure is correctly captured by the graph that was superimposed on the (simplified) map of Königsberg, as shown in Figure 5 (reproduced below).

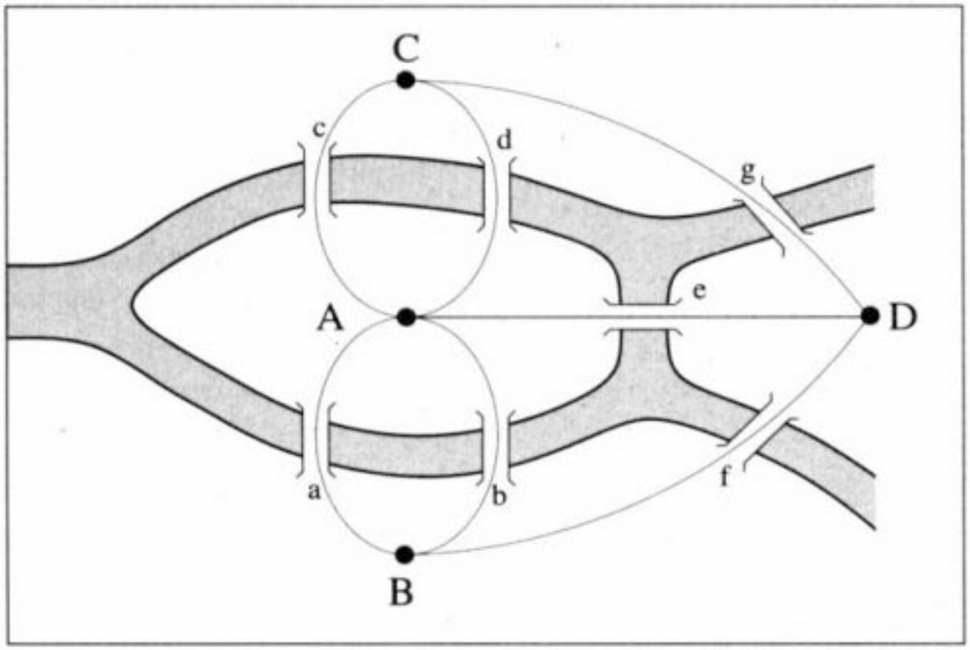

Figure 5: A graph superimposed on a map of the bridges of Königsberg

Isn't the resulting network therefore objectively correct? At first glance, this might seem like the right thing to say here. But at second glance things are not so clear.

None of Königsberg's original bridges have survived the city's tumultuous transition into modern-day Kaliningrad. But new bridges have been built, so a contemporary version of Euler's puzzle can still be posed: what might be called "the Kaliningrad Bridge Problem." Below is an aerial view of Kneiphof Island as it looks today.

Figure 8: Kneiphof Island, 2008, aerial view

What, then, is the network of Kaliningrad's bridges? Or, putting the same question a different way, what is the graph for the Kaliningrad Bridge Problem? This very simple question is surprisingly hard to answer in a definitive manner. The problem lies with the roadway that passes vertically through the left-hand side of the above photograph. Should this be counted as one link or two when we draw the graph? In other words, is this structure one bridge or two bridges? It carries a highway that passes over Kneiphof Island, but there is no way for vehicles to exit from highway to island or vice-versa. Does this bridge (or bridges) therefore *connect* Kneiphof to the mainland? It is unclear whether there are determinately correct answers to these questions. If not then this puts pressure on the idea that, even for systems with such (literally) concrete connections as bridges, there is necessarily such a thing as *the* structure of the system, or *the* network that underlies the system.[19]

[19] Another modern-day analog of the Königsberg Bridge Problem is the 'Manhattan River-Crossing Problem.' Ambiguities here include the status of tunnels, whether one-way bridges count as genuine connections, and whether crossings that allow only trains should be included.

The Sage of Königsberg

It is satisfying to be able to complete the circle – formed by the return to the geographical starting-point of our inquiry – by connecting these philosophical concerns to Königsberg's other great claim to intellectual fame. The only substantical building that still stands on Kneiphof Island, and one that appears in both Figure 1 and Figure 8, is Königsberg Cathedral. In a mausoleum adjoining the northeast corner of the Cathedral lies buried the philosopher Immanuel Kant, a lifelong resident of Königsberg, and a Professor at the University of Königsberg, which was itself located on Kneiphof Island.

Figure 9: Kant's grave, Kneiphof Island

This coincidental link between Kant and the Königsberg Bridge Problem, and thereby – indirectly – between Kant and network science, is ironic given the particular metaphysical and epistemological preoccupations to which Kant gives voice in his monumental work, *The Critique of Pure Reason*. A characteristic feature of Kant’s philosophy is a deep skepticism about our ability to come to know the nature of ‘things in themselves.’ Some modern interpreters see Kant as reacting to this by arguing that our knowledge of the world must therefore be based solely on the relational properties that things bear to other things, as opposed to intrinsic properties which they may (or may not) have in isolation.[20] Thus Kant sought epistemological safety in the relational. The irony of this is that our discussion of real networks, which can be thought of in some sense as the ultimate relational objects, suggests that even here the goal of objective knowledge may be illusory. If the world does not come 'pre-structured' in a determinate (or uniquely determinable) way then the networks that we construct, and the connections that we draw from one node to another, can never fully escape the influence of our own particular interests.

[20] Kant also had a relational view of space, something that Einstein claimed influenced his own development of the theory of relativity.