Networks, Dynamics, and the Small-World Phenomenon[1]

Duncan J. Watts
Santa Fe Institute

The small-world phenomenon formalized in this article as the coincidence of high local clustering and short global separation, is shown to be a general feature of sparse, decentralized networks that are neither completely ordered nor completely random. Networks of this kind have received little attention, yet they appear to be widespread in the social and natural sciences, as is indicated here by three distinct examples. Furthermore, small admixtures of randomness to an otherwise ordered network can have a dramatic impact on its dynamical, as well as structural, properties—a feature illustrated by a simple model of disease transmission.

INTRODUCTION

The small-world phenomenon (Milgram 1967; Pool and Kochen 1978) has long been an object of popular fascination and anecdotal report. The experience of meeting a complete stranger with whom we have apparently little in common and finding unexpectedly that we share a mutual acquaintance is one with which most of us are familiar—"It's a small world!" we say. More generally, most people have at least heard of the idea that any two individuals, selected randomly from almost anywhere on the planet, are "connected" via a chain of no more than six intermediate acquaintances, a notion made popular by the Broadway play (and later movie) *Six Degrees of Separation* (Guare 1990).

But is this phenomenon merely the confluence of unlikely coincidence and curious anecdote, or is it actually indicative of the underlying structure of modern social networks and, hence, not unlikely at all? Furthermore, if the small-world phenomenon does turn out to be a deep feature

[1] The author gratefully acknowledges the guidance of Steven Strogatz during the period in which this research was conducted and the extensive comments of Harrison White on an earlier draft of this article. Research was supported in part by the National Science Foundation through a grant to Steven Strogatz. Direct correspondence to Duncan Watts, Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501. E-mail: duncan@santafe.edu

0002-9602/2000/10502-0005$02.50

of the social world, in what other contexts can it arise (such as telecommunications and neural networks), and what mechanism drives it? Finally, does its presence in the real world have any implications for the dynamical properties of the networks in which it occurs?

An explanation of the phenomenon, and more generally, a framework for examining the properties of networks consisting of very many components, is of general sociological interest: Many social metrics, such as status (Harary 1959; Burt 1982) and power (Coleman 1973), and social processes, such as the diffusion of innovations (Rogers 1995) and transmission of influence (Friedkin 1990), are usefully represented in terms of networks of relationships between social actors, be they individuals, organizations, or nations. Indeed, the theory of social networks is one that has seen extensive development over the past three decades, yielding multiple measures both of individual significance, such as centrality (Freeman 1979, 1982; Friedkin 1991), and of network efficiency (Yamaguchi 1994*a*), which may elucidate nonobvious phenomena such as "key players" in an organization or its optimal structure for, say, information diffusion. Frequently, however, this research assumes linear models of social processes,[2] such as Markov models of diffusion, and is generally applied to networks that consist of a relatively small number of components.[3] While many of the measures defined in the literature can in principle be applied to networks of arbitrary size and structure, the computational costs of doing so may be prohibitive (such as for Freeman's [1979] betweenness centrality), and the benefits are at any rate unclear if the process of interest is inherently *nonlinear,* as is the case for information (or disease) contagion models involving threshold (Granovetter 1978; Arthur and Lane 1993) or refractory (Murray 1993, chap. 19) effects. Hence, the problem of analyzing efficiently the structure of extremely large networks (in which components may easily number in the hundreds of thousands, or more), and modeling the effects of structure on nonlinear dynamical processes, remains relatively unexplored. This article presents one possible approach to these very general problems by postulating a model of network formation that is sufficiently flexible to account for a wide variety of interesting cases (although not all cases, by any means, as will be pointed out) and that,

[2] "Social process" here refers to the person-to-person dynamics occurring on the network. Such *local* dynamical models are thus distinct from *global* models, such as classical homogeneous mixing models of diffusion, which are generally nonlinear (see e.g., Yamaguchi 1994*b*), but which do not model the underlying network explicitly.

[3] Test cases appearing in the literature typically use small numbers for ease of computation (e.g., Friedkin [1991] and Yamaguchi [1994*a,* 1994*b*] consider networks of five and seven elements respectively) and documented empirical examples rarely exceed $O(100)$ nodes.

in turn, suggests which structural parameters are the appropriate ones to study.

The problem of relating network structure to dynamics is then illustrated with a very simple model of disease spreading, which implies that the specified structural parameters are significant in determining the dynamics (usually in a highly nonlinear fashion), although probably insufficient to describe it completely. It should be pointed out that networks can affect a system's dynamical behavior in what might be termed an active and a passive sense; and that it is the passive sense that is investigated here. Active implies that the network is a device to be manipulated consciously for an actor's own ends; passive implies that the network connections themselves, in concert with blind dynamical rules, determine the global behavior of the system. The active sense has been investigated, for example, by Granovetter (1973, 1974) in the context of finding a job and Burt (1992) for maximizing social capital. The passive sense—with which this article is concerned—has been explored in systems as varied as biological oscillators (Strogatz and Stewart 1993), neural networks (Crick and Koch 1998), genetic-control networks (Kauffman 1969), epidemiology (Hess 1996*a,* 1996*b;* Longini 1988; Kretzschmar and Morris 1996), and game theory (Nowak and May 1993; Herz 1994; Cohen, Riolo, and Axelrod 1999). Before addressing any of these questions, however, it is necessary to agree on precisely what is meant by the small-world phenomenon, what is already known about it, and why such a thing should be surprising in the first place.

PROPERTIES OF RANDOM-CLUSTERED NETWORKS

What Is the Small-World Phenomenon?

What do we mean when we say the world is "small"? In general, there is no precise answer, but in this article, "small" means that almost every element of the network is somehow "close" to almost every other element, even those that are perceived as likely to be far away.[4] This disjuncture between reality and perception is what makes the small-world phenomenon surprising to us. But why should we perceive the world to be anything other than small in the first place? The answer to this is fourfold:

1. The network is numerically large in the sense that the world contains $n \gg 1$ people. In the real world, n is on the order of billions.
2. The network is sparse in the sense that each person is connected to an average of only k other people, which is, at most, on the order of thou-

[4] Precise definitions of "small" and "far away" require some additional terminology that is developed below.

sands (Kochen 1989)—hundreds of thousands of times smaller than the population of the planet.

3. The network is decentralized in that there is no dominant central vertex to which most other vertices are directly connected. This implies a stronger condition than sparseness: not only must the average degree k be much less than n, but the *maximal* degree k_{max} over all vertices must also be much less than n.
4. The network is highly clustered, in that most friendship circles are strongly overlapping. That is, we expect that many of our friends are friends also of each other.

All four criteria are necessary for the small-world phenomenon to be remarkable. If the world did not contain many people, then it would not be surprising if they were all closely associated (as in a small town). If most people knew most other people then, once again, it would not be surprising to find that two strangers had an acquaintance in common. If the network were highly centralized—say a star—then an obvious short path would exist through the center of the star between all pairs of vertices. Finally, if the network were not clustered—that is, if each person chose their friends independently of any of their friends' choices—then it follows from random-graph theory (Bollobás 1985) that most people would be only a few degrees of separation apart even for very large n.[5]

But are these criteria satisfied by the real world? Given that the population of the planet is several billion and that even the most generous estimates of how many acquaintances an average person can have (Kochen 1989) is only a few thousand, then the first two criteria are likely to be satisfied. The last two conditions are harder to be sure of and certainly harder to measure, but they also seem quite reasonable in the light of everyday experience. Some people are clearly more significant players than others, but even the most gregarious individuals are constrained by time and energy to know only a tiny fraction of the entire population. What significance these individuals have must be due to other more subtle and interesting reasons. Finally, while it might be difficult to determine in practice how many of a given person's friends are also friends with each other, and even more difficult to measure this for a large population, common sense tells us that whatever this fraction is, it is much larger than that which we would expect for a randomly connected network.[6]

The first evidence that the world might indeed be small was presented over 30 years ago by the psychologist Stanley Milgram (1967). Milgram

[5] In fact, a random graph is a close approximation to the *smallest possible* graph for any given n and k (where $k_{max} \ll n$ and the variance in k is not too large).

[6] To be more explicit, if the world *were* randomly connected, then one's acquaintances would be just as likely to come from a different country, occupation, and socioeconomic class as one's own. Clearly this is not the case in real life.

initiated a number of chain letters with sources in Kansas and Nebraska, to be sent to one of two targets in Boston. Each source was given the name of the target and some demographic information about them but was instructed that they could only send the letter to someone they knew by first name. If they did not know the target directly (a remote possibility), the idea was to send it to whichever of their friends they considered was most likely to. This procedure was then to be repeated, generating a chain of recipients that either reached the target or else petered out due to apathy. Of the chains that did complete, Milgram found that the median number of links in the chain was about six, thus giving rise to the famous phrase, "six degrees of separation." A later work of Milgram's (Korte and Milgram 1970) found similar results for senders and recipients in different racial subgroups, thus bolstering the claim that the world was not just small within particular socioeconomic categories but was, perhaps, small universally.

Although the first theoretical examination of the small-world phenomenon, by Pool and Kochen (1978), did not appear in published form until well after Milgram's experiment, the ideas had been in circulation for some 10 years beforehand. Pool and Kochen posed the problem in terms of the probability (p_i) that two randomly selected elements of a network would be connected via a shortest path consisting of i intermediaries. They calculated expected values of p_i under a variety of assumptions about local network structure and stratification. They concluded, as had Milgram, that the world was probably a small one, in the sense that randomly selected pairs could generally be connected by chains of only a few intermediaries. However, their assumptions concerning network structure and the independence of connections were so restrictive that they declined to place much weight on their hypothesis. Little progress has been made on this work since, and Pool and Kochen's conclusions remain essentially unchanged (Kochen 1989).

Another starting point for theoretical investigation of the small-world phenomenon was the study of random-biased nets, developed in the 1950s and 1960s by Anotol Rapoport and his colleagues at the University of Chicago. Motivated by the desire to understand the spread of infectious diseases, Solomonoff and Rapoport (1951) calculated the expected fraction of a randomly mixed population to be infected by a small initially infected seed. Rapoport then determined the corresponding fractions to be infected in populations where network connections exhibited increasing levels of local redundancy due to effects such as homophily, symmetry of edges, and triad-closure bias (Rapoport 1953*a,* 1953*b,* 1957, 1963; Foster, Rapoport, and Orwant 1963). More sophisticated approximations were developed subsequently by Fararo and Sunshine (1964) and, later, Skvoretz (1985) to account for differentiation of ties as well as vertices. However,

as was the case with Pool and Kochen, all these approximations focused on altering the *local structure* of the network and so showed only that significant changes in global structure will result from correspondingly significant changes in local structure. This statement is not contradicted here. Rather, what is new is that equally significant changes in global structure can result from changes in local structure that are so minute as to be effectively undetectable at the local level. This is an important distinction, as it is at the local level—and only at the local level—that individuals in a network make measurements.

Formalization of the Small-World Phenomenon

In order to make the requisite notions precise, some definitions are borrowed from graph theory. For simplicity, the networks considered here will be represented as connected graphs, consisting solely of undifferentiated vertices and unweighted, undirected edges.[7] All graphs must also satisfy the sparseness conditions specified above.

The first statistic of interest, for a given graph, is the *characteristic path length* (L), defined here as the average number of edges that must be traversed in the shortest path between any two pairs of vertices in the graph.[8] In terms of Milgram's experiment, L would be the chain length averaged over all possible sources in the network *and* all possible targets. L then is a measure of the global structure of the graph (because, in general, determining the shortest path length between any two vertices requires information about the entire graph). By contrast, the *clustering coefficient* (C) is a measure of the local graph structure. Specifically, if a vertex v has k_v immediate neighbors, then this neighborhood defines a subgraph in which at most $k_v(k_v - 1)/2$ edges can exist (if the neighborhood is fully connected). C_v is then the fraction of this maximum that is realized in v's actual neighborhood, and C is this fraction averaged over all vertices in the graph.[9] Equivalently, C can be regarded as the probabil-

[7] These assumptions are unrealistic in general, as many networks of interest in both the social and natural sciences are composed of weighted and directed relationships. However, generalizations of the resulting graph statistics to account for these added complexities—although straightforward in principle—may depend on the particular application at hand. Therefore, for the purpose of constructing a broadly relevant framework, undirected, unweighted graphs are the natural starting point.

[8] Variants of L have appeared in other contexts as diverse as the status of individuals in an organization (Harary 1959), the floor plans of buildings (March and Steadman 1971), the efficiency of communications networks (Chung 1986), and even the properties of chemical compounds (Wiener 1947; Rouvray 1986).

[9] Local clustering, or variants thereof, has also appeared in the literature as a measure of network structure, originally in Davis (1967).

ity that two vertices (u, v) will be connected, given that each is also connected to a "mutual friend" (w).

All the graphs considered in this article will be characterized in terms of these two statistics. But in order to contextualize the results—to decide, in effect, what is "small" and what is "large," what counts as "clustered" and what does not—it is necessary to determine the *ranges* over which L and C can vary. Three constraints are imposed upon this exercise:

1. The population size (n) is fixed.
2. The average degree k of the vertices is also fixed such that the graph is sparse ($k \ll n$) but sufficiently dense to have a wide range of possible structures ($k \gg 1$).[10]
3. The graph must be *connected* in the sense that any vertex can be reached from any other vertex by traversing a finite number of edges.

Fixing n and k enables valid comparisons to be made between many different graph structures. Clearly, the largest value that C can attain for *any* connected graph is $C = 1$, for a *complete graph* ($k = n - 1$). Conversely, the minimum conceivable value of C is $C = 0$ for an *empty graph* ($k = 0$). These two graphs also have extremal length properties. This, however, is not a very instructive comparison, as it is obvious that clustering and length will change as more and more edges are added to any graph. A more interesting question is how these statistics can change simply by rearranging a fixed number of edges among a fixed number of vertices. The sparseness conditions focus our attention on the most interesting terrain from the perspective of a wide range of applications in both the social and natural sciences. That is, the network is sufficiently well connected to admit rich structure, yet each element is confined to operate within a local environment that encompasses only a tiny fraction of the entire system. Finally, by insisting that all graphs be connected, L is guaranteed to be a truly global statistic. Hence, comparisons of characteristic path length are valid comparisons of global structure. Bearing in mind these conditions, the following questions present themselves:

1. What is the most clustered graph possible, and what is its characteristic path length?

[10] Because, in the models considered here, fluctuations in vertex degree (k_v) are roughly normally distributed around k, this condition becomes in practice the more strict condition mentioned earlier: ($k_{max} \ll n$). This is a qualitatively stronger constraint than $k \ll n$ because it precludes not only densely connected graphs, but also star-like graphs, which have $L \approx 2$, regardless of n, by virtue of one (or a few) highly centralized vertices. That star graphs are ignored in this analysis is not to imply that they are uninteresting, simply that our concern here is with *decentralized* graphs—a reasonable restriction if one considers large enough n (where no one member could possibly know all others).

2. What graph has the lowest possible characteristic path length, and what is its clustering coefficient?
3. What do these results imply about the relationship between the clustering coefficient and characteristic path length of a sparse graph?

Turning first to clustering, a significant insight is that, although a connected graph can only attain the maximal value of $C = 1$ when $k = (n - 1)$, even a very sparse graph may have a clustering coefficient that is, in practice, indistinguishable from the complete case. The most clustered, sparse graph possible is what might be termed the *caveman graph,* which consists of $n/(k + 1)$ isolated cliques or "caves": that is, clusters of $(k + 1)$ vertices within which all vertices are connected to all others but between which no edges exist at all. It is easy to see that this graph has $C = 1$, on a par with a complete graph. However, it fails another required condition—that all graphs must be *connected.* Fortunately, global connectivity is an easy property to achieve in this case, simply by extracting one edge from each clique and using it to connect to a neighboring clique such that all cliques eventually form a single, unbroken loop. This *connected caveman graph* (fig. 1) can be shown to have a clustering coefficient of

$$C_{caveman} \approx 1 - \frac{6}{(k^2 - 1)}, \tag{1}$$

which approaches 1 as k becomes sufficiently large (without violating $k \ll n$; a detailed derivation of equation 1 is given in Watts [1999, chap. 4]).[11]

One can also calculate the corresponding characteristic path length for large n and k (see Watts 1999, chap. 4):

$$L_{caveman} \approx \frac{n}{2(k + 1)}. \tag{2}$$

Note that, for $n \gg k$, L must necessarily be large and also increases linearly with increasing n. Hence, the connected caveman graph can be used as a benchmark for a "large, highly clustered graph."[12]

[11] The connected caveman graph does not, in fact, have the highest possible clustering coefficient for fixed n and k. For instance, the last edge required to complete the ring in figure 1 is not required for connectivity and so can remain in its clique, thus marginally increasing C. Other even more clustered constructions may also be possible. Nevertheless, no graph can be constructed whose clustering exceeds that of the connected caveman graph by more than $O(1/k^2)$, which becomes vanishingly small as k increases.

[12] Of course, sparse, connected networks with larger L can be constructed (trivially, by severing one of the between-cluster edges to form a line of clusters instead of a ring). Such changes, however, do not affect the essential structural properties of the network—that is, its linear scaling properties—and so are not of concern here. Other, more elaborate constructions (such as a large, dense cluster trailed by a long line of vertices) are ruled out by the same regularity requirement that excluded star-like graphs from consideration. Hence, within the bounds of the model, the connected caveman graph is a plausible (albeit approximate) upper limit for L.

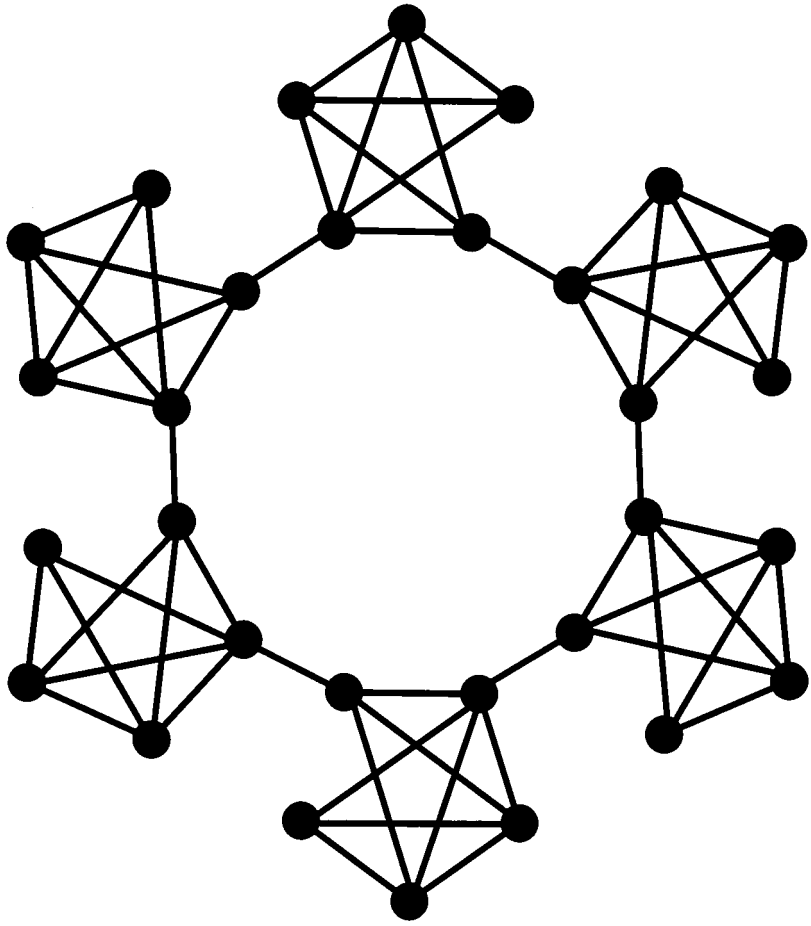

FIG. 1.—Schematic of the connected caveman construction

At the other extreme, no general, realizable structure can be shown to exhibit minimal characteristic path length for arbitrary n and k (Cerf et al. 1974; Bollobás 1985), but a good approximation to the theoretical lower bound is realized by a random graph (Bollobás 1985), where $kn/2$ out of all possible $n(n-1)/2$, edges are chosen at random and with equal probability. Precise formulas do not exist for L and C of a random graph, but in the limits of large n and k, the corresponding asymptotic approximations are[13]

[13] For an argument justifying the approximation for L, see Bollobás (1985) and Bollobás and Chung (1988). The asymptotic expression for C_{random} can easily be derived by

$$L_{random} \sim \frac{\ln(n)}{\ln(k)}, \tag{3}$$

and

$$C_{random} \sim \frac{k}{n}. \tag{4}$$

Note that not only is $L_{random} \ll L_{caveman}$ for any $n \gg 1$, but that the scaling of L_{random} is logarithmic with respect to n instead of linear. Hence, as n becomes larger, the discrepancy between the two extremes in length becomes increasingly pronounced (linearly in n). Note also that the sparse-graph condition ($k \ll n$) implies that C_{random} is very small. Hence (recalling the probabilistic interpretation of the clustering coefficient), C can be thought of as a simple measure of *order* in a graph—graphs with $C \gg k/n$ (like the connected caveman graph) are considered *locally ordered* (in the sense that vertices with at least one mutually adjacent vertex are likely to be themselves adjacent), and random graphs are, naturally, disordered.

The intuition that one might draw from these results is that highly clustered or locally ordered graphs necessarily have long characteristic path lengths, and conversely, graphs with short characteristic path lengths have clustering that is vanishingly small in the limit of large n. This is a reasonable intuition but is at odds with the (so far anecdotal) claim that the world can be small and still be highly clustered. In the absence of definitive data for the whole world, an alternative test of the small-world problem is to determine the minimum conditions that are both necessary and sufficient for the world to be small. The approach adopted here is to introduce a family of graphs that interpolates approximately between the two extremes discussed above and then to examine the intermediate region for evidence of small-world effects.

A Theory of Length Contraction in Sparse Networks

For this purpose, it is natural to consider a model that captures, in some abstract sense, the formation of social connections. A number of social network theorists have utilized the concept of a "social space" in which people exist as points separated by distances that can be measured according to some appropriately defined metric (see, e.g., Davidson 1983). Unfortunately, this approach often runs into treacherous waters due to

considering that a neighbor of any given vertex (v) has an expected probability ($k_v - 1)/n$ of being adjacent to another vertex (u) in v's neighborhood.

the inherent difficulty both of characterizing the space (which is all but unknown) and defining the metric (equally so). The following three assumptions avoid these difficulties:

1. All networks can be represented solely in terms of the connections between their elements, assuming that whatever combination of factors makes people more or less likely to associate with each other is accounted for by the distribution of those associations that actually form.
2. All connections are symmetric and of equal significance. That is, a definition of what is required in order to "know" someone is defined such that either two people know each other or they do not.[14]
3. The likelihood of a new connection being created is determined, to some variable extent, by the already existing pattern of connections.

Exactly how existing connections determine new ones is a big part of the mystery. One might imagine a world in which people only become acquainted through introduction by one or more mutual friends. It is easy to see that a mechanism such as this leads inevitably to a locally ordered world (in the sense of $C \gg k/n$), the extreme case being the caveman world. At the other extreme, one might also imagine a world in which new friendships are made autonomously and at random, without regard for current friendships.[15] The end product of this tie formation process is naturally a random graph. Of course, the real world lies somewhere between these two extremes but precisely where is anybody's guess. Hence, rather than assuming some specific functional relationship between current and future friendships, let us examine a whole universe of possible "worlds" that lie between the ordered and random extremes. One way to do this in a precise and explicit fashion is through a graph construction algorithm that embodies the following features (also shown graphically in fig. 2):

1. At the ordered extreme, the propensity of two unrelated people (meaning they share no mutual friends) to be connected is very small. Once they have just one friend in common, however, their propensity to be acquainted immediately becomes very high and stays that way regardless of how many additional mutual friends they may have. In worlds like this one, it is almost a certainty that the only people anyone will ever connect to are those with whom they share at least one mutual friend. So, plotting "propensity to become friends" against "fraction of current mutual friends," the propensity starts near zero, rises very rapidly to some relatively large number (which can be normalized to one), and then plateaus.

[14] Unlike the one-way connection that often exists between, e.g., a professor and a student, or a celebrity and a fan.

[15] We may be seeing the beginnings of such a world already in the proliferation of Internet "chat lines," where complete strangers can meet, interact, and sometimes even end up marrying.

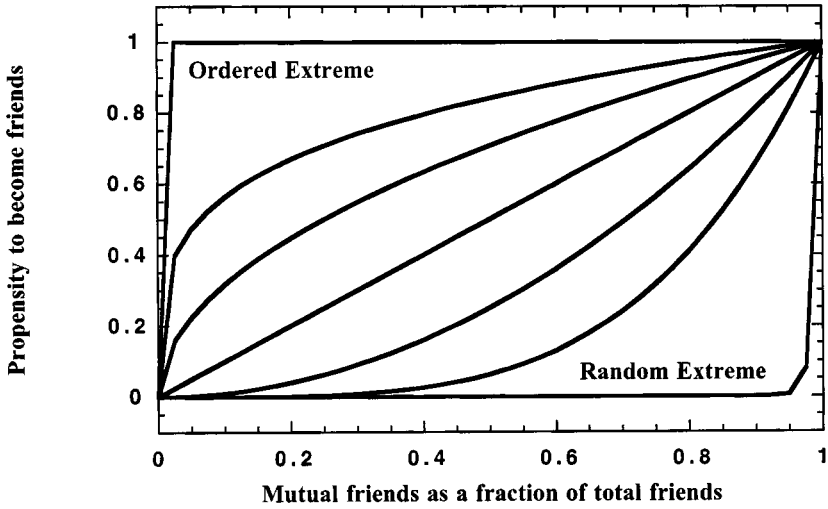

FIG. 2.—Family of functions representing the propensity of strangers to meet, given the fraction of the friends that they currently share.

2. At the random extreme, no one has much propensity to connect to anyone in particular. In this sort of world, the "propensity versus mutual friends" curve remains near zero up until the point where all friends are mutual friends.[16]
3. In between these two extremes, the propensity curve can take any one of an infinite number of intermediate forms, specified by a single, tunable parameter, where it is important only that the dependency be smooth and monotonically increasing with respect to increasing mutual friends. These conditions are satisfied by the following construction:

$$R_{i,j} = \begin{cases} 1 & m_{i,j} \geq k \\ \left[\dfrac{m_{i,j}}{k}\right]^{\alpha}(1 - p) + p & k > m_{i,j} > 0, \\ p & m_{i,j} = 0 \end{cases} \tag{5}$$

where $R_{i,j}$ is a measure of vertex i's propensity to connect to vertex j (zero if they are already connected), $m_{i,j}$ is the number of vertices that are adjacent both to i and j, k is the average degree of the graph, p is a baseline

[16] The rapid jump from near zero to near one in fig. 2 is necessitated by continuity conditions but can be rationalized in modeling terms by considering that in such a situation the two parties must have *all* their friends in common, in which case it is reasonable to assume that they cannot avoid meeting.

propensity for an edge (i, j) to exist (set at $p = 10^{-10}$),[17] and α is a tunable parameter, $0 \leq \alpha \leq \infty$.[18]

Numerical simulation.—Equation (5) is, in principle, an abstract representation of a graph or rather, for each value of α, an enormous (but finite) number of *potential* graphs that share certain statistical characteristics. However, it seems unlikely that the properties of these graphs, and how those properties vary with α, can be derived in any precise analytical sense. This is an important point, because a great deal of work in graph theory is analytical. However, analytical approaches are generally confined to cases where either n is small and the rules for constructing the graph are strictly deterministic or n is so large that it can be treated as infinite and the rules are strictly random. Both of these extremes exhibit certain properties that simplify the situation, thus enabling analytical descriptions. The case presented by equation (5), however, falls squarely into the messy no-man's-land between these two extremes: n is large but not infinite, and the rule for constructing edges is partly deterministic and partly random, where even the balance between determinism and randomness varies as α changes. The only manner in which such a model can be analyzed is through a rigorous process of computer-based, numerical simulation.[19] Adopting this approach, equation (5) now forms the basis for a construction algorithm,[20] which builds a graph of specified n, k, and α.[21]

A problem that immediately rises with this "α-model" is that, for small α, the resulting graphs tend overwhelmingly to consist of small, isolated, and densely internally connected components. This results from a *start-up problem*—initially no edges exist in the graph, so p dominates $R_{i,j}$ in equation (5), and edges form randomly until, by chance, two edges share a vertex in common. At this point, because α is small, the two vertices that share a mutual "friend" will almost certainly become connected at the expense of expanding their friendship networks into new territory. The fraction of pairs of vertices that are members of the same connected component grows only linearly with k, so connected graphs cannot be

[17] The actual value of p is not important so long as it is small enough that no random edges can be expected for $\alpha = 0$ (i.e., $p \ll 2/n[n-1]$). More specifically, the numerical results with respect to α do change as p changes, but this dependency disappears when the model is recast in terms of the model-independent parameter ϕ, introduced below. That is, $\phi(\alpha)$ varies with p, but $L(\phi)$ and $C(\phi)$ do not.

[18] Note that α has no physical or social significance—it is simply a parameter that enables the model to generate graphs ranging from highly ordered to highly random.

[19] All work described was conducted on either a SUN\ Sparc 20 or a DEC\ alpha 500 running *C* under a UNIX operating system.

[20] The technical details of the algorithm are provided in the appendix.

[21] Unless otherwise stated, the parameters used to generate the results presented here are: $n = 1{,}000$, $k = 10$.

generated for small α without violating the sparse graph condition. Disconnected graphs pose a problem because they necessarily have $L = \infty$, and this makes them hard to compare with connected graphs or even each other. One way to resolve this dilemma is to build in a connected substrate before commencing the algorithm, thus ensuring that all subsequently constructed graphs will be connected. A potentially serious objection to such a step is that the properties of the resulting graphs may be so dominated by the presence of the substrate that any conclusions drawn from the model will fail to be sufficiently general to be of interest. The following constraints on the choice of substrate minimize (although do not remove) this concern:

1. It must exhibit *minimal structure,* in that no vertices are to be identified as special. This eliminates structures like stars, trees, and chains, which have centers, roots, and end-points, respectively.
2. It must be *minimally connected.* That is, it must contain no more edges than necessary to connect the graph in a manner consistent with condition 1.

The only structure that satisfies both these criteria is a topological ring. One advantage of this choice of substrate is that for sufficiently small α, it results in graphs that resemble the connected caveman limit described above; that is, densely intraconnected clusters strung loosely together in a ring (with its attendant linear length-scaling properties). It is less obvious that the random limit also can be attained with this additional structure built in. Nevertheless, numerical evidence suggests that this is precisely the case. Hence, the ring substrate not only ensures connectivity, but also allows the model to interpolate between roughly the desired limits. More important, we will see that the results generated by the corresponding model exhibit sufficiently generic features in the intermediate regime that quite general conditions can be specified under which small-world networks should arise.

The clearest way to see this is to measure L and C for the α-model, for fixed n and k, over a range of $0 \leq \alpha \leq 20$. The following functional similarities between $L(\alpha)$ and $C(\alpha)$ are revealed (figs. 3 and 4):[22]

1. For large α, both statistics approach their expected random-graph values.
2. At $\alpha = 0$, both L and C are high relative to their random-graph limits and increase (as α increases) to a distinct maximum at small α.
3. Both statistics exhibit a sharp transition from their maximum values to their large-α limits.

[22] Each of the points in fig. 5 is the average value of the relevant statistic over 100 random realizations of the construction algorithm for the corresponding value of α.

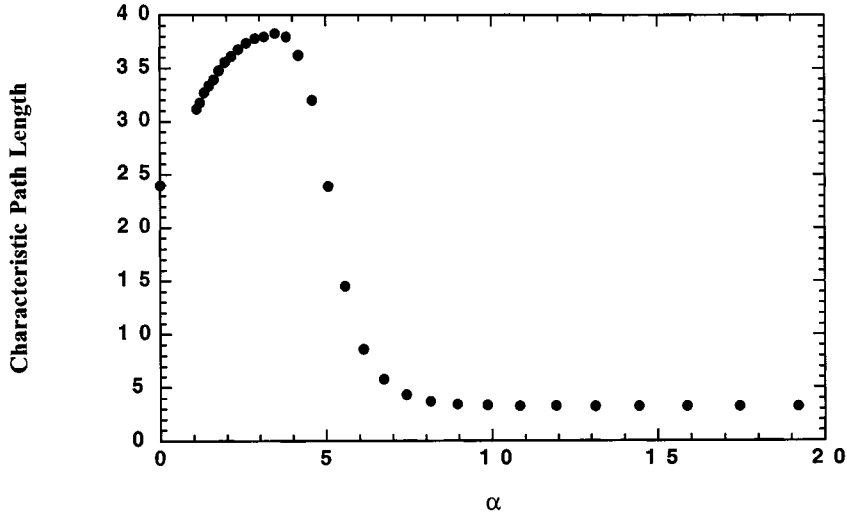

FIG. 3.—Characteristic path length (L) as a function of α for the α-model defined by equation 5 ($n = 1{,}000$, $k = 10$).

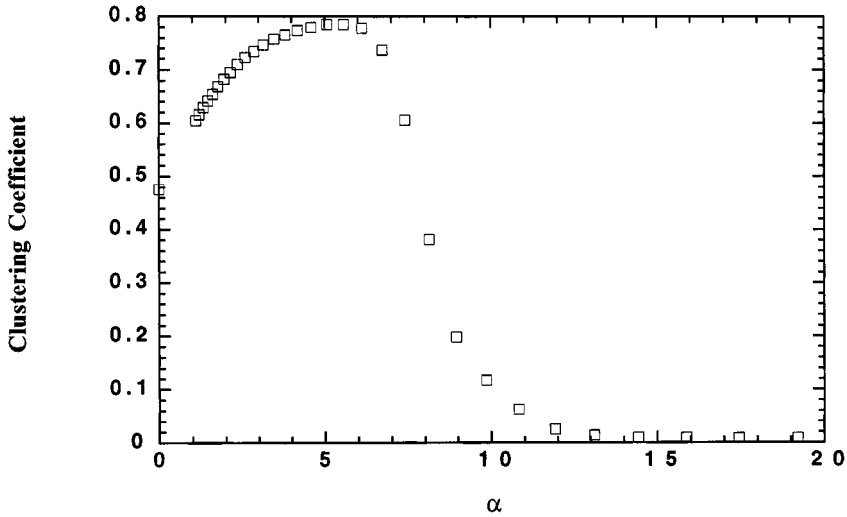

FIG. 4.—Clustering coefficient (C) as a function of α for the α-model defined by equation 5 ($n = 1{,}000$, $k = 10$).

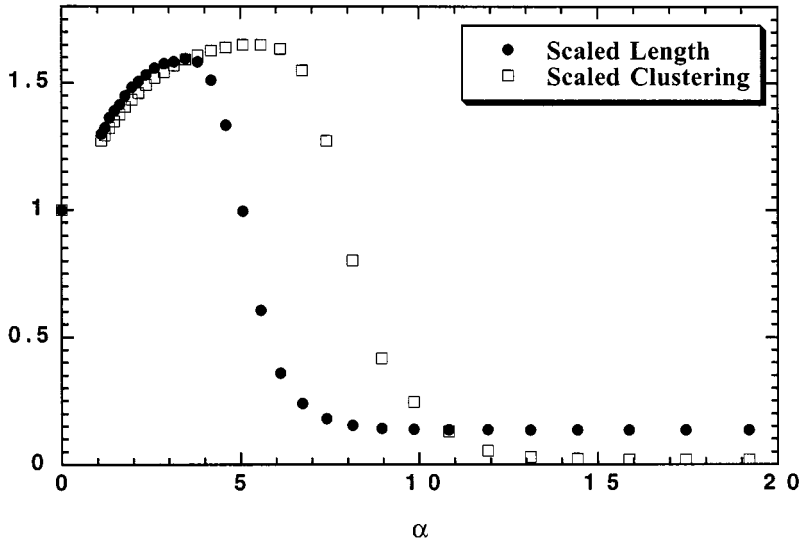

FIG. 5.—Comparison of $L(\alpha)$ and $C(\alpha)$, scaled by their corresponding values at $\alpha = 0$.

So far, the model appears to conform to the earlier statement that highly clustered graphs have large characteristic path lengths and, conversely, that graphs with short characteristic path lengths are necessarily poorly clustered.[23] However, as figure 5 indicates, $L(\alpha)$ and $C(\alpha)$ exhibit one important functional difference: the transition from large to small clustering coefficients occurs at a larger value of α than the equivalent length transition. The upshot of this disparity is that there exists a class of graphs in this region of α for which characteristic path length is small but clustering is high.[24] Thus the small-world phenomenon can be cast in graph-theoretic terms as the *coexistence* of high clustering and small characteristic path length.

DEFINITION 1.—*A small-world graph is a large-*n, *sparsely connected, decentralized graph* ($n \gg k_{max} \gg 1$) *that exhibits a characteristic path*

[23] This conclusion is bolstered by the additional observation that, in the small-α regime, L scales linearly with respect to n and logarithmically with respect to n for large α, in correspondence with eqq. (2) and (3).

[24] These results remain qualitatively the same for a wide range of n and k, strongly suggesting that they are true for *all* n and k, with the usual caveat $n \gg k \gg 1$. Of course, in practice, these inequalities are imprecise, the effective limits being: if k is too small (in this case $k \to 2$), the substrate *will* dominate the results; and if k is too large ($k \to n$), then all topologies will be equivalent.

length close to that of an equivalent random graph ($L \approx L_{random}$), *yet with a clustering coefficient much greater* ($C \gg C_{random}$).

This definition does not depend on the specifics of the graph-construction algorithm—in fact, it can be applied to any graph regardless of its construction. Nevertheless, the definition is only *interesting* if the phenomenon it describes can also be shown to be independent of the specifics of the model—in particular, the substrate. The reason for this is obvious: networks in the real world are no more likely to be constructed on ring substrates than they are to be completely ordered or completely random. This potential shortfall in the theory can be addressed in two ways. First, a variety of different substrates can be tested and their results compared with those generated above. If small-world graphs are still attainable over a significant interval of α values, then there is reason to think that they constitute a robust class of graphs. Second, a theoretical understanding of length contraction in partly ordered, partly random graphs may shed some light on the existence of small-world graphs and help to specify *model-independent* conditions that, if satisfied, will yield small-world graphs.

The first approach is straightforward but tedious. The same model has been tested with a number of other substrates—a two-dimensional lattice, a Cayley tree, and a random substrate. All the substrates surveyed exhibit qualitatively different properties from those of a ring, and also each other, yet α-graphs based on all substrates invariably yield small-world graphs over an extended interval of α.[25] The second approach—a theoretical explanation of small-world graphs—is presented in the next section.

Shortcuts and contractions.—Drawing intuition from the results of the α-model, it is now possible to develop an explanation of small-world graphs in terms of a parameter that is independent of the particular model used to construct them. To motivate this approach, note that, according to equation (5) for $\alpha = 0$, newly created edges are virtually guaranteed to complete at least one triad.[26] Two vertices connected by any such edge must necessarily have been separated by a path of length two, prior to the addition of the new edge. Hence, the addition of a new edge to an α-graph at $\alpha = 0$ contributes little in the way of length contractions, as it can connect only pairs of vertices that are already "close." In random graphs, however, this condition no longer applies, and vertices that are widely separated are as likely to become connected as those that are near neighbors. These observations lead to the following definitions:

[25] A detailed description of these substrates and the corresponding model properties is presented in Watts (1999, chap. 3).

[26] This is aside from the negligible propensity p to make a random connection.

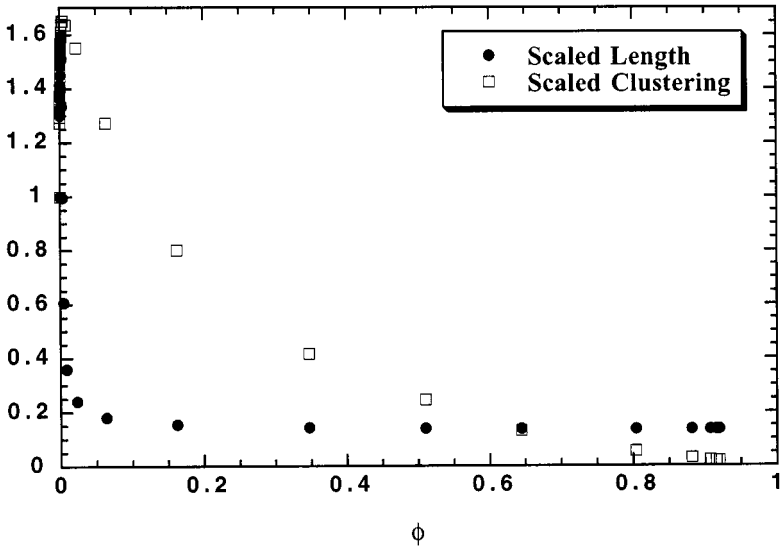

FIG. 6.—Comparison of $L(\phi)$ and $C(\phi)$, scaled by their corresponding values at $\phi_0 = \phi(\alpha = 0) = 0$.

DEFINITION 2.—*The range* r *of an edge is the distance between the two vertices that the edge in question connects when the edge itself has been deleted. Equivalently, range can be thought of as the second-shortest path length between two connected vertices (where the shortest path length is necessarily one).*

DEFINITION 3.—*A shortcut is any edge with a range* $r > 2$.

DEFINITION 4.—*The parameter* ϕ *is the fraction of all edges in the graph that are shortcuts.*

Figure 5 can now be replotted against ϕ instead of α. The result, in figure 6, demonstrates not only the previous result that it is possible for highly clustered graphs to have small characteristic path lengths, but also that this happens principally for small ϕ.[27]

[27] For very small ϕ (too small to resolve on the linear scale of fig. 6), there is an apparent increase in both L and C as ϕ increases—analogous to the humps in fig. 5. The basis of this effect is that, for small but nonzero α, an increase in α results in higher local clustering, and so previously overlapping neighborhoods can become distinct, yielding edges with $r > 2$ that do not connect previously distant parts of the graph. This is indeed a practical problem with detecting shortcuts for this particular construction algorithm, but it occurs only at very small ϕ—below the value at which the small-world phenomenon is relevant. Hence, it does not affect any of the results stated here.

The intuitive explanation for this additional observation is that, for small ϕ, the characteristic path length of the graph is large. Hence, the introduction of a single shortcut is likely to connect vertices that were previously widely separated. This shortcut then contracts the distance not only between that pair of vertices, but also between their immediate neighborhoods, their neighborhoods' neighborhoods, and so on. Thus, one single shortcut can potentially have a highly nonlinear impact on L. By contrast, the clustering coefficient C is only reduced in a single neighborhood—as the result of one less triad being formed—and so the decrease in C can be at most linear in ϕ. This nonlinear (global) versus linear (local) impact of shortcuts enables large C to coexist with small L at small values of ϕ. But once ϕ becomes large, L has already decreased to a small value, and so subsequent shortcuts can do little to reduce it further. Thus $L(\phi)$ must approach its random graph limit asymptotically, as shown in figure 6.

The parameter ϕ turns out to have explanatory power even beyond the specific construction presented above. Other models of graphs that interpolate between ordered and random limits also exhibit length and clustering properties that can be understood in terms of shortcuts, the key requirement being that shortcuts be permitted to connect vertices that are separated by distances on the order of the size of the entire graph. Thus, it can be conjectured that any graph with the property $n \gg k_{max} \gg 1$, which exhibits (*a*) a clustering coefficient $C \gg k/n$ and (*b*) a small fraction of long-range shortcuts, will be a small-world graph.

If this conjecture is true, then small-world graphs can be realized by a great many construction algorithms, of which the α-model is but one. However, it also suggests that there are many kinds of partly ordered, partly random graphs in which the small-world phenomenon will not occur. The key criterion that the small fraction of introduced shortcuts be "long range" is really equivalent to the statement that new connections be determined without regard to any kind of *external length scale* imposed upon the graph (such as by explicitly disallowing connections to be made between vertices that are separated by greater than a certain physical distance). It is the independence of any external length scale that enables a tiny fraction of shortcuts to collapse the characteristic path length of the system to near its asymptotic, random-graph value, without significantly reducing the corresponding clustering coefficient. This constraint implies that the small-world phenomenon is unlikely to be exhibited by networks whose connectivity is determined solely by physical forces, which imply corresponding length scales.

Although the above conjecture appears to be a sufficient condition for the existence of small-world networks, it turns out that it is not necessary. In other words, it is possible to contract distances in large graphs with a

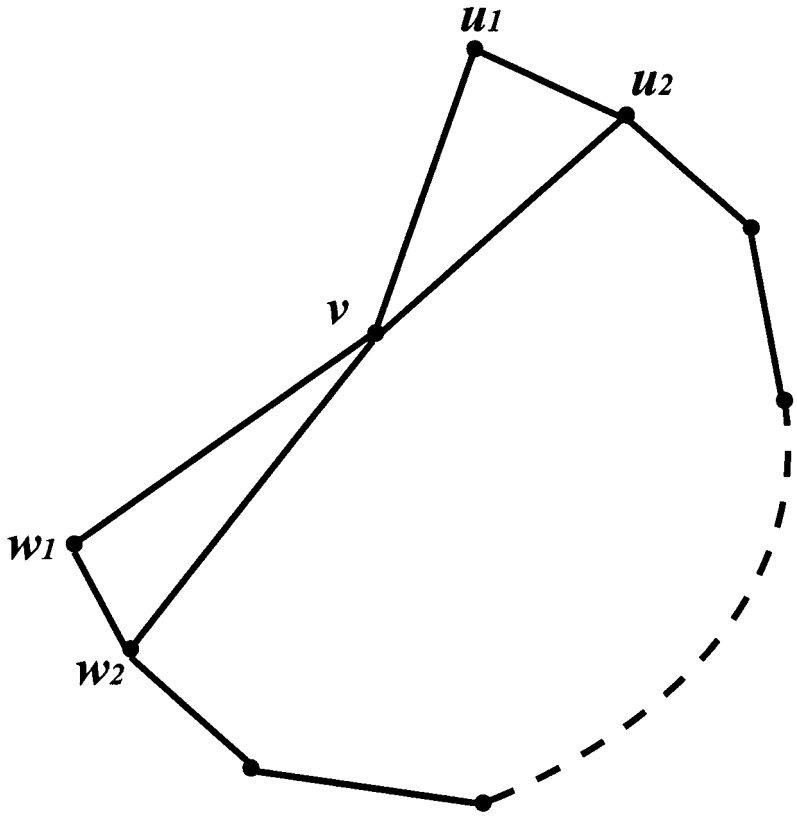

FIG. 7.—Schematic of a contraction: vertex v contracts the path length between groups u and w, but no edges are shortcuts.

negligible effect on the clustering, without using any shortcuts at all. The simplest example of such a situation is detailed in figure 7, from which it is obvious that groups of vertices are being brought closer together by virtue of a single common *member* but that none of the edges involved is a shortcut. Fortunately, a simple modification of the definition of a shortcut is adequate to capture this new scenario:

DEFINITION 5.—*A contraction occurs when the second-shortest path length between two vertices, sharing a common neighbor, is greater than two. In other words, a contraction is a pair of vertices that share one and only one common neighbor.*

DEFINITION 6.—*By extension,* ψ *can be defined as the fraction of pairs of vertices with common neighbors that are contractions.*

The results displayed in figure 7 can be expressed in terms of ψ (see

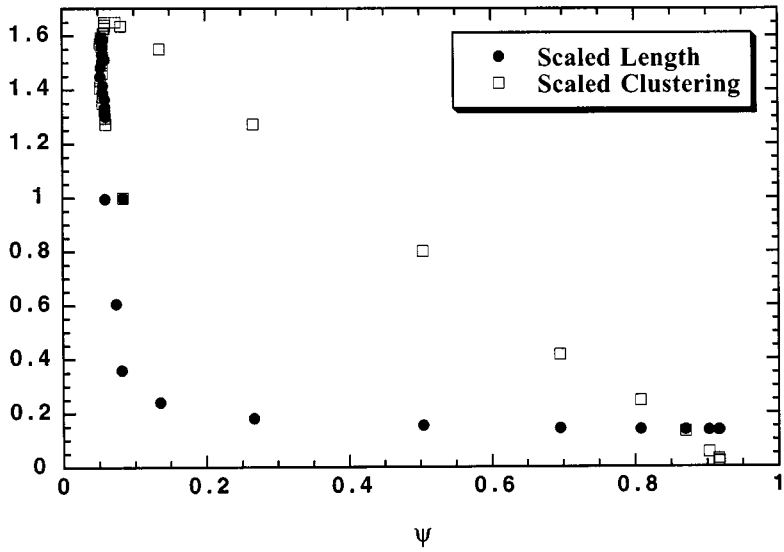

FIG. 8.—Comparison of $L(\psi)$ and $C(\psi)$, scaled by their corresponding values at $\psi_0 = \psi(\alpha = 0) \approx .08$.

fig. 8), which again shows the coexistence of small L and large C over a large range of ψ.[28] It follows from these definitions that a shortcut is simply a special case of a contraction in which one (or both) of the "groups" consists of a single vertex. Long-range contractions are thus a more general mechanism than shortcuts for generating small-world networks. In principle though, both shortcuts and contractions achieve essentially the same end—to connect what would otherwise be distant parts of a large, sparse graph with a large characteristic path length. Hence, because shortcuts are conceptually simpler and require less computational effort, all results in this article are stated in terms of ϕ, bearing in mind that they can be recast in terms of ψ if necessary.

Evidence of the Small-World Phenomenon in Real Networks

Milgram's experiment suggested that the idea of the small-world phenomenon is at least plausible for a real social network the size of the U.S. population. But the precise mapping out of social networks of this magnitude is a practical impossibility, due to the ambiguities inherent in defining

[28] Note that, unlike ϕ, ψ is never zero, even for the most clustered graphs.

both a member of the network and also what constitutes a friendship. Furthermore, tracing experiments like Milgram's face a number of technical obstacles, which he outlines (Milgram 1967); the main problem being that one can never be sure that the chain of intermediaries actually traced between two people is the shortest one possible.[29] We can now understand, in terms of shortcuts, at least one reason why this may be so. As noted in the previous section, small-world graphs occur for values of ϕ at which most vertices have no shortcuts at all. For example, for $n = 1{,}000$ and $k = 10$, $\phi = .01$ is sufficient for L to be indistinguishable from that of a random graph. If only one in a hundred edges is a shortcut, however, then (for $k = 10$) about 90% of all vertices will have no shortcuts in their local neighborhood. This absence of global information at the local scale poses significant problems for effective tracing of shortest path lengths in a network, as such an exercise requires knowledge not just of one's friends, but of one's friends' friends, and so on. If shortcuts exist, but only outside of one's local network vicinity, it becomes extremely difficult to utilize them consciously and thus construct an optimal path.

For these reasons, along with the practical difficulty associated with the empirical estimation of even local parameters like k and C (Kochen 1989), direct resolution of the small-world phenomenon in the actual social world seems unlikely. However, there is nothing about the small-world graph definition above that demands the graph in question represent a social network. In fact, one of the useful aspects of the corresponding conjecture is that it is quite general, specifying neither the nature of the vertices and edges, nor a particular construction algorithm required to build the graph. If it is true, then many real networks, satisfying the required n and k conditions, should turn out to be small-world networks. Examples of large, sparse graphs are easy to think of (neural networks, large organizations, citation databases, etc.) but difficult to obtain in the required format where both vertex and edge sets are precisely defined and completely documented. Nevertheless, three scientifically interesting examples are presented below.

The first example is the collaboration graph of feature-film actors. In the *actor collaboration graph* (Tjaden 1997), a vertex is defined as a cast member of any feature film registered on the Internet movie database,[30]

[29] On the other hand, as White (1970) points out, longer chains tend not to complete, biasing the sample toward shorter chains and leading to a corresponding underestimate of length.

[30] The Internet Movie Database (http://www.us.imdb.com) lists the cast members of all films, of all nationalities, since 1898. The graph studied in this article is actually the largest connected compenent of the entire graph, consisting of about 90% of all actors listed in the IMDB as of April 1997.

TABLE 1

CHARACTERISTIC PATH LENGTH (L) AND CLUSTERING COEFFICIENT (C) FOR THREE REAL NETWORKS

	L_{Actual}	L_{Random}	C_{Actual}	C_{Random}
Movie actors	3.65	2.99	.79	.00027
Power grid	18.7	12.4	.080	.005
C. elegans	2.65	2.25	.28	.05

and an edge represents two actors appearing in the same movie. This structure is interesting, as it is a simple case of a large ($n = 226{,}000$), sparse ($k = 61$) social network. It is also reminiscent of the collaboration graph of mathematicians that is traditionally centered on Paul Erdös (Grossman and Ion 1995).[31] The second example—the *western states power graph* ($n = 4{,}941$, $k = 2.94$)—represents the power-transmission grid of the western United States and is relevant to the efficiency and robustness of power networks (Phadke and Thorp 1988). Vertices represent generators, transformers, and substations, and edges represent high-voltage transmission lines between them. The final example is that of the neural network of the nematode *C. elegans* (White, Thompson, and Brenner 1986; Achacoso and Yamamoto 1992)—the sole example of a completely mapped neural network. For the *C. elegans graph* ($n = 282$; $k = 14$), an edge joins two neurons if they are connected by either a synapse or a gap junction. All edges are treated as undirected, and all vertices as identical, recognizing that these are crude approximations from a biological perspective.

Table 1 shows a comparison between L and C for each of these graphs and also L and C for random graphs with the same n and k. Note that, in each case, the characteristic path length is close to that of the equivalent random graph, yet the clustering coefficient is consistently much greater. A graphical way to view the same relationship is presented in figure 9. Here, L is plotted versus C for the three real graphs, and also for equivalent (that is, with same n and k) connected caveman graphs, where in each case, the statistics have been normalized by their corresponding random-graph values. From this picture, it is clear that not only are the real networks statistically distinct from both their random and caveman equivalents, but they are all distinct in the *same way*. Furthermore, their

[31] This collaboration graph would also be an interesting case to examine. Unfortunately, the only data available is that in the immediate neighborhood of Paul Erdös, and this is not sufficient to draw any conclusions about its global structure.

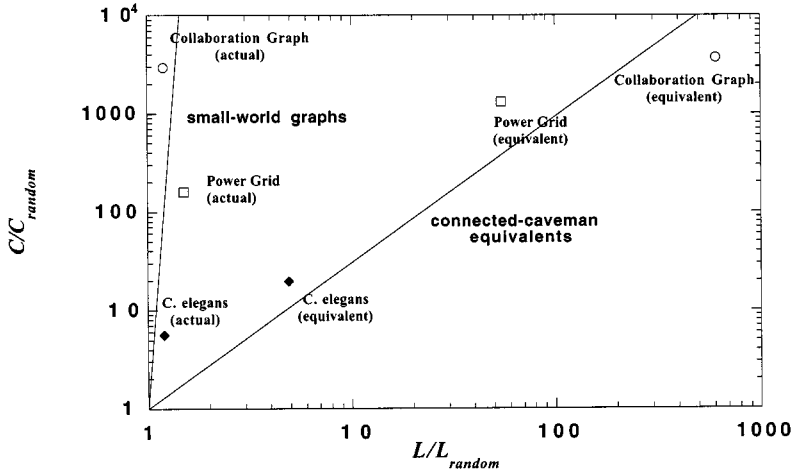

FIG. 9.—*L* and *C* statistics for the three real networks (the actor collaboration graph, the western states power graph, and the *C. elegans* graph) and their connected caveman equivalents (same *n* and *k*), scaled by their values for equivalent random graphs.

combination of small *L* and large *C* cannot be explained as a linear interpolation between the two extremes—in other words, clustering and length do not vary in a commensurate fashion. As figure 10 shows, however, this combination of properties can be replicated by the α-model. It is clear that, as α increases, the α-statistics remain at first clustered along the (high-*L,* high-*C*) diagonal and then sharply depart from the diagonal to become small-world graphs, decreasing rapidly in characteristic path length while remaining almost constant in terms of the clustering coefficient. As α increases to reach the clustering transition, *C* decreases rapidly for *L* fixed near its asymptotic limit, until the random limit is reached at large α. These results indicate that the small-world phenomenon is not just a property of an abstract class of hypothetical graphs, but arises in real networks. Furthermore, it is not specific to a particular kind of network or restricted to a certain size range. The three graphs examined span a range of three orders of magnitude in *n* and represent completely different actual networks, yet all are small-world graphs in the sense defined above. Finally, not only are all three graphs small-world graphs in the broad sense of high clustering coexisting with small characteristic path length, but the relationship between *L* and *C* in the real graphs is consistent with the corresponding statistics of the α-graph model.

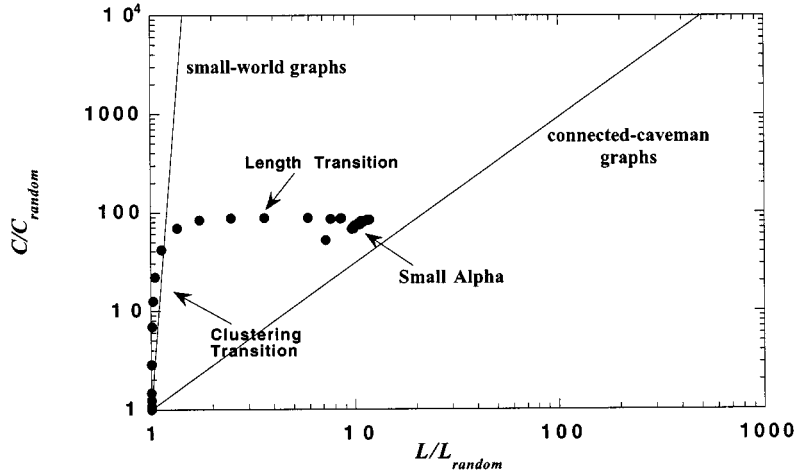

FIG. 10.—L and C statistics for the α-model ($n = 1{,}000$, $k = 10$), scaled by their values at the random-graph limit (in practice, this is taken as $\alpha = 20$).

DYNAMICAL SYSTEMS ON SMALL-WORLD GRAPHS

Having established that a set of relatively tiny perturbations to the local structure of a highly clustered graph can have a dramatic impact upon its global structural properties, it is natural to ask whether or not the same changes can also affect the behavior of dynamical systems that are coupled according to such a graph. This is a topic that is directly relevant to the social sciences: the role of social structure in generating globally observable, dynamical features. So far, structure has been treated as an autonomous feature of networks and defined narrowly in terms of sparse, undirected graphs. This has paid off by yielding some robust statements about the small-world properties of a general class of graphs that are partly ordered and partly random, and that seem to reflect some of the features of real networks. However, a greater issue is to understand the relationship between structure and dynamics.

As such, the following discussion builds not only on issues of network structure, but also upon a whole literature of distributed dynamical systems, in which systems are often assumed to be completely connected (the most tractable case). Where sparsely coupled systems are considered, the relevant coupling topology is usually assumed to be either completely ordered (e.g., a low-dimensional cubic lattice) or completely random. As emphasized earlier, real networks are likely to be sparse and may combine significant elements both of order and randomness, with resultant proper-

ties, like the small-world phenomenon, that cannot be captured by either of these approximations. Whether or not this structural oversight turns out also to be an oversight from the perspective of dynamical systems is not obvious. This article offers a more specific way to cast the issue at hand: do significant new phenomena emerge when the connectivity of distributed dynamical systems are modeled on the graphs presented earlier?

Disease Spreading in Structured Populations

A simple case of a dynamical system is that of a disease spreading from a small seed of initiators into a much larger population whose structure is prescribed by some underlying graph.[32] The bulk of previous work on the spread of diseases focuses on populations in which uniform mixing is assumed between elements (see, e.g., Murray 1991). This is an important assumption because it enables population structure to be ignored, thus greatly simplifying the analysis. Some work, however, has grappled with the issue of population structure. Kareiva (1990) reviews a number of such attempts (which he calls "stepping-stone" models), and May and colleagues (Hassell, Comins, and May 1994; May 1995) have considered various parasite-host problems on two-dimensional grids of discrete but homogeneous patches. Both Kareiva and May conclude that the introduction of spatial structure can significantly affect both the population size and its susceptibility to parasites and disease. A similar approach has been used by Hess (1996*a,* 1996*b*) to compare virus transmission among subpopulations that are connected according to various simple topologies such as a ring and a star. Also, Sattenspiel and Simon (1988) have considered a detailed model of the spread of an infectious disease in a structured population in which different connective arrangements between the subpopulations are compared. Finally, Longini (1988) has utilized real airline network data in order to model the 1968 global outbreak of influenza. None of this work, however, goes on to treat the global dynamical properties of the system explicitly as a function of the structure. The models of Sattenspiel, Simon, and also Hess do consider different types of connectivity between subpopulations, but they consider isolated topologies, as opposed to a continuum, and their choices reflect those cases which, in the context of this article, are extremal (such as a ring versus a random graph). Extremal cases are certainly natural to consider, but if it is true that real

[32] Here I discuss disease spreading because of its obvious public health relevance, but qualitatively similar dynamics could describe the spread of other kinds of contagion such as ideas, rumors, fashion, or even crime (Gladwell 1996).

networks exhibit important properties of both ordered and random networks, then it is important to consider the intermediate regime as well.

Dynamics as a function of structure.—In this sense, the work most closely related to the approach taken here is that by Kretzschmar and Morris (1996), who analyze the spread of a disease—both in terms of extent and time scale—as a function of the overall concurrency of relationships in the population. Roughly speaking, they examine a family of graphs that interpolates between a world of exclusively monogamous (but randomly formed) relationships and an unconstrained random graph, in which concurrent relationships are likely to form. They determine that increased concurrency of relationships significantly increases the extent of the disease and its rate of spread, even when the total number of relationships in the population is held constant. Essentially, this result stems from the increasing connectedness of extremely sparse ($k = 1$) graphs: serial monogamy yields almost completely disconnected graphs (in the sense that no connected components larger than dyads can exist), but random graphs with the same number of edges exhibit relatively large connected components. Hence, as Kretzschmar and Morris conclude, it is the size of the largest connected component that drives the spread of disease across the population. The approach here also examines the effects on disease spread of changing the distribution of a fixed number of edges over a fixed number of vertices. It is different, however, in that Morris and Kretschmar consider changing concurrency in an otherwise randomly mixing population with $k = 1$, while here the amount of randomness is varied in a connected population with $k \gg 1$. Hence, any observed differences in the dynamical properties of the system must be driven by more subtle features of the topology than connectedness. In fact, as we shall see, there is much about the system that cannot be reduced to any single structural characteristic of the underlying graphs—a warning signal for dealing with any more complicated dynamics.

In order to emphasize the role of population structure, the subsequent analysis is restricted to a simplified model of disease spreading in which each element of the population is in one of three states: *susceptible, infected,* or *removed.* At each discrete point in time (t), every infected element can infect each one of its neighbors with probability (p), the *infectiousness.* Any newly infected elements remain infected for one time step,[33]

[33] The time period (τ) for which an infected agent remains infectious can be set to one without loss of generality. The reason is that p and τ do not vary independently. In fact, the dynamics for any given τ can be reproduced with $\tau = 1$ merely by rescaling p. In other words, there is complete equivalence (in this model) between being exposed to a more infectious disease for a short period of time and a less infectious disease for a long period of time. This feature of the model greatly simplifies its analysis.

after which they are removed permanently from the population (presumably by immunity or death) and so play no further part in proceedings. Hence, if at $t = 0$ a single element is infected (by some external influence), then at some later time, one of two things must have happened:

1. The disease will have run its course and died out, infecting some fraction of the population and leaving the remaining fraction F_s uninfected,

or

2. The whole population will have been infected ($F_s = 0$) in some characteristic time T.

The natural question to ask then is whether or not the structure of the population, expressed in terms of the α-graph model presented earlier, has any effect on F_s and T of the related system. Furthermore, if so, can the functional forms of $F_s(\phi)$ and $T(\phi)$ be understood in terms of our structural statistics $L(\phi)$ and $C(\phi)$?

Results.—The model has two parameters that can vary independently of each other: the infectiousness p, which determines the local dynamics; and the fraction of shortcuts ϕ. We are now in a position to compare some results for an entire range of topologies as a means of answering the two questions stated above within the very narrow context of this specific dynamical system. Figure 11 shows the steady-state fraction of susceptibles F_s versus p for the two extreme values of ϕ, where graphs of parameters $n = 1{,}000$, $k = 10$ have been used to determine the couplings of the system. Three distinct regions are apparent:

1. For $p \lesssim 1/(k - 1) \approx .11$ all topologies (i.e., all values of ϕ) yield the same result: the disease infects only a negligible fraction of the population before dying out. This is a *trivial* steady state, because nothing happens that can distinguish between different topologies.
2. For $.11 \lesssim p \lesssim .5$ different topologies yield different F_s.
3. For $p \gtrsim .5$, all topologies once again yield the same end result, only this time it is a *nontrivial* steady state because the disease has taken over the entire population.

There is nothing more to say about region 1, but regions 2 and 3 deserve some extra attention. Region 2 is confusing: there appears to be some significant relationship between structure and dynamics, but its mechanism is not transparent. Figure 11 suggests that in region 2 the fraction of the population that will become infected with a disease of some specified p depends significantly on ϕ, and this is shown more explicitly in figure 12 for a particular value of p. In epidemiological terms, this dependence on ϕ is equivalent to the statement that the impact of a disease depends not just on how infectious it is, but also on the connective topology of the population. This message is not new in the epidemiological literature, es-

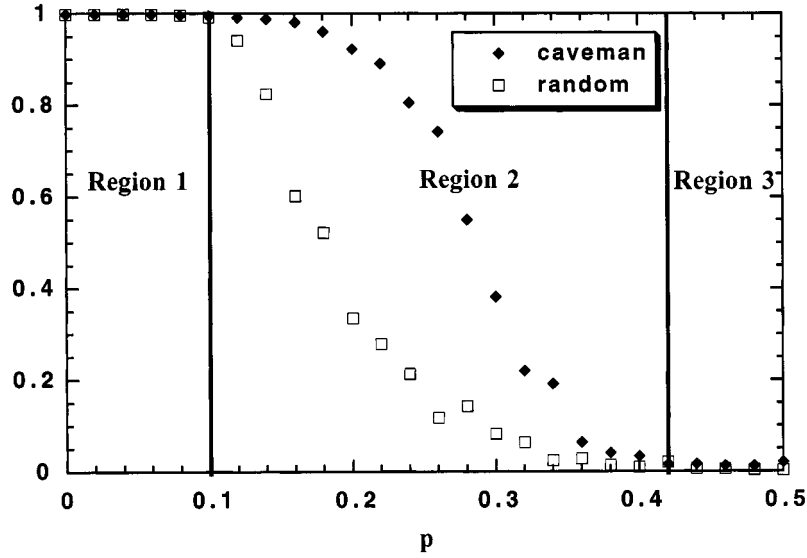

FIG. 11.—Fraction of uninfected survivors (F_s) versus infectiousness (p) for disease spreading dynamics on a network generated by the α-model at clustered and random extremes.

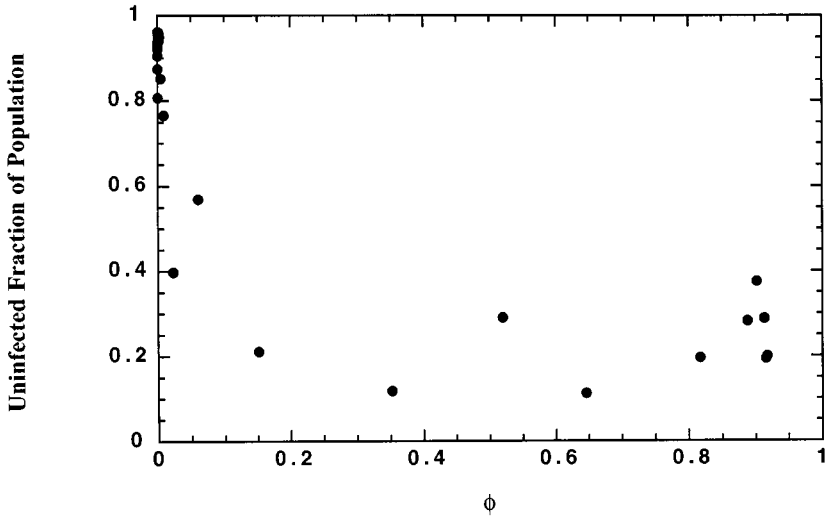

FIG. 12.—F_s versus ϕ for $p = .24$

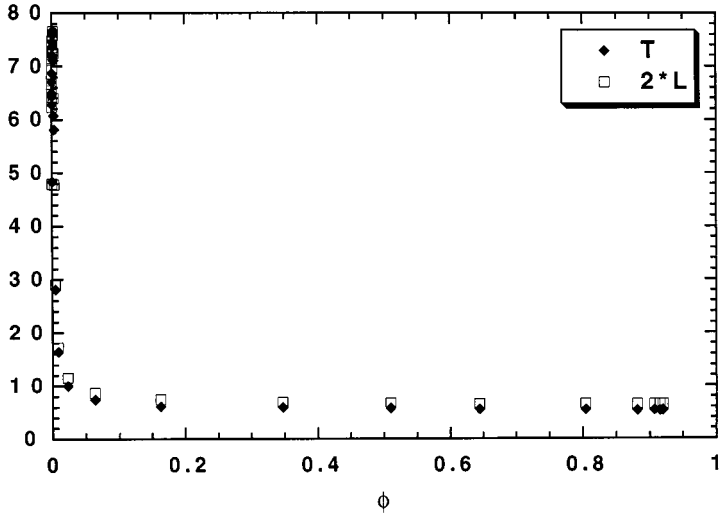

FIG. 13.—Time taken to reach a steady state (T) versus fraction of shortcuts (ϕ) in the underlying α-graph for the disease spreading model.

pecially that of sexually transmitted diseases. What is new is the idea that the structural changes required can be very subtle. Unlike the difference between a chain and a star graph or a ring lattice and a random graph, the difference between a big- and a small-world graph can be a matter of only a few randomly rewired edges—a change that is effectively undetectable at the level of individual vertices. Unfortunately, it is difficult to be any more precise than this, as the functional form of $F_s(\phi)$ is not clear. Nevertheless, even this broad-brush observation has implications for the way we think about social diseases, which are often perceived as confined to isolated subgroups of a population. The message here is that the highly clustered nature of small-world graphs can lead to the intuition that a given disease is "far away" when, on the contrary, it is effectively very close.

Region 3 is a simpler matter to examine, as the disease always takes over the entire population regardless of its connective topology. It is merely a question of how long this takes. However, figure 13 shows that in this region the time taken to reach the steady-state $F_s = 0$ varies dramatically as a function of ϕ and that disease can spread on a small-world graph far more rapidly than in a connected caveman graph and almost as fast as in a random graph. In fact, figure 13 also shows that when $p = 1$, $T(\phi)$ bears a close functional relationship to $L(\phi)$.[34] Hence, in this parameter

[34] For $p = 1$, this functional similarity might be expected, as T is just the maximum distance from the initially infected vertex to any other vertex in the graph. For a

regime at least, the transient time for the dynamics to reach the same, nontrivial steady state bears a simple and obvious relationship to the structure of the underlying graph. That is, shorter characteristic path length implies faster spreading of the disease. In a real-world scenario, where an epidemic can be responded to, the time scale on which it spreads becomes a crucial factor. A striking consequence of this result is that, in a small-world graph, the characteristic time scale has become very small, but the clustering C is still large. As with the extent of disease spread in region 2, the change in structure in region 3 that causes the disease to spread much faster may not be observable at a local level.

Although the relationship between population structure and disease dynamics is not always clear, even for this simple model, it does appear as if the gross features of the dynamics are dominated by the characteristic path length of the underlying graph. This is by no means universally true of dynamical systems on graphs. For instance, when the transmission of a behavioral trait exacts a cost, such as for a prisoner's dilemma model of cooperation, a high degree of reciprocity may be necessary for the trait to survive in the population (Axelrod 1984; Boyd and Richerson 1988; Cohen, Riolo, and Axelrod 1999). In the case of small groups of cooperators struggling for survival in a sea of defectors, Axelrod (1984) notes that cooperators must interact preferentially, which in network terms is equivalent to high local clustering. In a highly clustered graph, cooperators located in the same cluster can survive—even thrive—in the midst of a noncooperative majority. Conversely, in a random graph (with negligible clustering), any small group of initial cooperators will be eroded from the periphery, as each peripheral cooperator will be interacting predominantly with defectors, thus failing to reap the benefits of reciprocity. As with disease spreading, there is a transition between these two extremes, but in this case, the dynamics tend to be dominated by the clustering coefficient rather than by the characteristic path length.[35]

This result suggests an interesting role for small-world architectures, which by virtue of their short characteristic path length and high cluster-

---

perfect ring structure $T = D$, where D is the diameter of the graph and $D = 2L$. Hence, for even the approximate ring structure generated for an α-graph with $\phi = 0$, we could expect that $T \approx 2L$. Furthermore, for *any* graph, it is necessarily true that $L \leq T \leq D$. For random graphs, where the number of vertices at a distance d from any vertex generally grows exponentially (Bollobás 1985), $D < 2L$ so we could also expect that $L \leq T \leq 2L$ for *any* value of ϕ. Thus T should be related to L through nothing more than a multiplicative factor $1 \leq c(\phi) \leq 2$, which is not significant on the scale of the changes occuring in both statistics as a function of ϕ. Hence, the two curves in figure 13 might be expected to look very much alike, as they do.

[35] There are a number of subtleties to this result which are explained more completely in Watts (1999, chap. 8).

ing coefficient, can support the rapid dissemination of information without necessarily compromising behavior that is individually costly but beneficial when reciprocated. Furthermore, the more general idea of optimizing an architecture to satisfy two or more opposing constraints may prove a useful concept in the design or modification of large networks such as organizations.

SUMMARY

This article examines a particular class of graphs that interpolates between highly ordered and highly random limits. A significant feature of these graphs is that the presence of a very small fraction of long-range shortcuts can lead to the coexistence of high local clustering and a small global length scale. The superposition of these apparently contradictory properties is a graph-theoretic formalization of the small-world phenomenon. The motivation for the small-world phenomenon comes from social networks, but it turns out to be a much more general effect that arises under quite weak conditions in large, sparse, partly ordered and partly random networks. Its existence is not predicted by current network theories, yet it seems likely to arise in a wide variety of real networks, especially in social, biological, and technological systems. One consequence of this result is that it is highly likely that the phenomenon exists in the real social world—a notion currently supported by only limited data but consistent with anecdotal experience.

In addition to their interesting structural properties, small-world graphs are also relevant to the social and natural sciences through their effect on the globally emergent features of dynamical systems. Specifically, distributed dynamical systems can exhibit dramatically different behavior on small-world networks—an effect that may have implications in fields as diverse as public health and organizational behavior and design.

APPENDIX

The algorithm for constructing a graph according to equation (5) proceeds as follows:

1. Fix a vertex i.
2. For every other vertex j, compute $R_{i,j}$ according to equation (5), with the additional constraint that $R_{i,j} = 0$ if i and j are already connected.
3. Sum the $R_{i,j}$ over all j, and normalize each to obtain variables $P_{i,j} = R_{i,j}/(\sum_{l \neq i} R_{i,l})$. Then, since $\sum_j P_{i,j} = 1$, we can interpret $P_{i,j}$ as the probability that i will connect to j. Furthermore, we can interpret $P_{i,j}$ geometrically as follows: divide the unit interval $(0, 1)$ into $n - 1$ half-open subintervals with length $P_{i,j}$ $\forall$ $j \neq i$.

4. A uniform random variable is then generated on (0, 1). It must fall into one of the subintervals, say the one corresponding to j_*.
5. Connect i to j_*.

This procedure is then repeated until the predetermined number of edges ($M = kn/2$) has been constructed. The vertices i are chosen in random order, but once a vertex has been allowed to "choose" a new neighbor, it may not choose again until all other vertices have taken their turn. However, vertices may be "chosen" arbitrarily often, and this leads to a nonzero variance in the degree k. But the fact that all vertices are forced to make one new connection before any others are allowed to choose a second time ensures at least that no vertices will be isolated (as long as $k \geq 2$).

REFERENCES

Achacoso, T. B., and W. S. Yamamoto. 1992. *AY's Neuroanatomy of* C. elegans *for Computation.* Boca Raton, Fla.: CRC Press.

Arthur, W. Brian, and David Lane. 1993. "Information Contagion." *Structural Change and Economic Dynamics* 4 (1): 81–104.

Axelrod, Robert. 1984. *The Evolution of Cooperation.* New York: Basic Books.

Bollobás, Bela. 1985. *Random Graphs.* London: Academic Press.

Bollobás, Bela, and Fan R. K. Chung. 1988. "The Diameter of a Cycle Plus a Random Matching." *SIAM Journal of Discrete Mathematics* 1 (3): 328–33.

Boyd, Robert, and Peter J. Richerson. 1988. "The Evolution of Reciprocity in Sizable Groups." *Journal of Theoretical Biology* 132:337–56.

Burt, Ronald S. 1982. *Toward a Structural Theory of Action.* New York: Academic Press.

———. 1992. *Structural Holes: The Social Structure of Competition.* Cambridge, Mass.: Harvard University Press.

Cerf, V. G., D. D. Cowan, R. C. Mullin, and R. G. Stanton. 1974. "A Lower Bound on the Average Shortest Path Length in Regular Graphs." *Networks* 4:335–42.

Cohen, Michael D., Rick L. Riolo, and Robert Axelrod. 1999. "The Emergence of Social Organization in the Prisoner's Dilemma: How Context-Preservation and Other Factors Promote Cooperation." Working Paper 99-01-002. Santa Fe Institute.

Coleman, James S. 1973. *The Mathematics of Collective Action.* Chicago: Aldine.

Chung, Fan R. K. 1986. "Diameters of Communications Networks." Pp. 1–18 in *Mathematics of Information Processing,* vol. 34 in Proceedings of Symposia in Applied Mathematics. Providence, R.I.: American Mathematical Society.

Crick, Francis, and Christof Koch. 1998. "Constraints on Cortical and Thalamic Projections: The No-Strong-Loops Hypothesis." *Nature* 391:245–50.

Davidson, Mark L. 1983. *Multidimensional Scaling.* New York: Wiley.

Davis, James A. 1967. "Clustering and Structural Balance in Graphs." *Human Relations* 20:181–87.

Fararo, Thomas J., and M. Sunshine. 1964. *A Study of a Biased Friendship Net.* Syracuse, N.Y.: Syracuse University Press.

Foster, Caxton C., Anatol Rapoport, and Carol J. Orwant. 1963. "A Study of Large Sociogram: Estimation of Free Parameters." *Behavioral Science* 8:56–65.

Freeman, Linton C. 1979. "Centrality in Social Networks: Conceptual Clarification." *Social Networks* 1:215–39.

———. 1982. "Centered Graphs and the Structure of Ego Networks." *Mathematical Social Sciences* 3:291–304.
Friedkin, Noah E. 1990. "Social Networks in Structural Equation Models." *Social Psychology Quarterly* 53 (4): 316–28.
———. 1991. "Theoretical Foundations for Centrality Measures." *American Journal of Sociology* 96 (6): 1478–1504.
Gladwell, Malcolm. 1996. "The Tipping Point." *New Yorker,* June 3, pp. 32–38.
Granovetter, Mark S. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78 (6): 1360–80.
———. 1974. *Getting a Job: A Study of Contacts and Careers.* Cambridge, Mass.: Harvard University Press.
———. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83 (6): 1420–43.
Grossman, Jerry W., and Patrick D. F. Ion. 1995. "On a Portion of the Well-Known Collaboration Graph." *Congressus Numeratium* 108:129–31.
Guare, John. 1990. *Six Degrees of Separation: A Play.* New York: Vintage Books.
Harary, Frank. 1959. "Status and Contrastatus." *Sociometry* 22:23–43.
Hassell, Michael P., Hugh N. Comins, and Robert M. May. 1994. "Species Coexistence and Self-Organizing Spatial Dynamics." *Nature* 370:290–92.
Herz, Andreas V. 1994. "Collective Phenomena in Spatially-Extended Evolutionary Games." *Journal of Theoretical Biology* 169:65–87.
Hess, George. 1996*a*. "Disease in Metapopulation Models: Implications for Conservation." *Ecology* 77 (5): 1617–32.
———. 1996*b*. "Linking Extinction to Connectivity and Habitat Destruction in Metapopulation Models." *American Naturalist* 148 (1): 226–36.
Kareiva, Peter. 1990. "Population Dynamics in Spatially Complex Environments: Theory and Data." *Philosophical Transactions of the Royal Society of London* 330: 175–90.
Kauffman, Stuart A. 1969. "Metabolic Stability and Epigenesis in Randomly Constructed Genetic Nets." *Journal of Theoretical Biology* 22:437–67.
Kochen, Manfred. 1989. "Toward Structural Sociodynamics." Pp. 52–64 in *The Small World,* edited by Manfred Kochen. Norwood N.J.: Ablex.
Korte, Charles, and Stanley Milgram. 1970. "Acquaintance Linking between White and Negro Populations: Application of the Small World Problem." *Journal of Personality and Social Psychology* 15:101–18.
Kretzschmar, Mirjam, and Martina Morris. 1996. "Measures of Concurrency in Networks and the Spread of Infectious Disease." *Mathematical Biosciences* 133:165–95.
Longini, Ira M., Jr. 1988. "A Mathematical Model for Predicting the Geographic Spread of New Infectious Agents." *Mathematical Biosciences* 90:367–83.
March, L., and P. Steadman. 1971. Chapter 14 in *The Geometry of Environment.* London: RIBA Publications.
May, Robert M. 1995. "Necessity and Chance: Deterministic Chaos in Ecology and Evolution." *Bulletin of the American Mathematical Society* 32:291–308.
Milgram, Stanley. 1967. "The Small World Problem." *Psychology Today* 2:60–67.
Murray, James D. 1991. *Mathematical Biology,* 2d ed. Berlin: Springer Verlag.
———. 1993. *Mathematical Biology,* 2d ed. Berlin: Springer Verlag.
Nowak, Martin A., and Robert M. May. 1993. "The Spatial Dilemmas of Evolution." *International Journal of Bifurcations and Chaos* 3 (1): 35–78.
Phadke, Arun G., and James S. Thorp. 1988. *Computer Relaying for Power Systems.* New York: John Wiley.
Pool, Ithiel de Sola, and Manfred Kochen. 1978. "Contacts and Influence." *Social Networks* 1:1–48.
Rapoport, Anatol. 1953*a*. "Spread of Information through a Population with Socio-

Structural Bias: I. Assumption of Transitivity." *Bulletin of Mathematical Biophysics* 15:523–33.

———. 1953*b*. "Spread of Information through a Population with Socio-Structural Bias: II. Various Model with Partial Transitivity." *Bulletin of Mathematical Biophysics* 15:535–46.

———. 1957. "A Contribution to the Theory of Random and Biased Nets." *Bulletin of Mathematical Biophysics* 19:257–71.

———. 1963. "Mathematical Models of Social Interaction." Pp. 493–579 in *Handbook of Mathematical Psychology,* vol. 2. Edited by R. R. Bush, R. D. Luce, and E. Galanter. New York: Wiley.

Rouvray, Dennis H. 1986. "Predicting Chemistry from Topology." *Scientific American* (September): 40–47.

Sattenspiel, Lisa, and Carl P. Simon. 1988. "The Spread and Persistence of Infectious Diseases in Structured Populations." *Mathematical Biosciences* 90:341–66.

Skvoretz, John. 1985. "Random and Biased Networks: Simulations and Approximations." *Social Networks* 7:225–61.

Solmonoff, Ray, and Anatol Rapoport. 1951. "Connectivity of Random Nets." *Bulletin of Mathematical Biophysics* 13:107–17.

Strogatz, Steven H., and Ian Stewart. 1993. "Coupled Oscillators and Biological Synchronization." *Scientific American* 269 (6): 102–9.

Tjaden, Brett. 1997. E-mail to author, May 14.

Watts, D. J. 1999. *Small Worlds: The Dynamics of Networks between Order and Randomness.* Princeton, N.J.: Princeton University Press.

White, Harrison. 1970. "Search Parameters for the Small World Problem." *Social Forces* 49:259–64.

White, J., E. Southgate, J. N. Thompson, and S. Brenner. 1986. "The Structure of the Nervous System of the Nematode *Caenorhabditis elegans."Philosophical Transactions of the Royal Society of London,* Series B, 314:1–340.

Wiener, Harry. 1947. "Structural Determination of Paraffin Boiling Points." *Journal of the American Chemistry Society* 69:17–20.

Yamaguchi, Kazuo. 1994*a*. "The Flow of Information through Social Networks Diagonal-Free Measures of Inefficiency and the Structural Determinants of Inefficiency." *Social Networks* 16:57–86.

———. 1994*b*. "Some Accelerated Failure-Time Regressive Models Derived from Diffusion Process Models: An Application to Network Diffusion Analysis." *Sociological Methodology* 24:267–300.