HARVARD UNIVERSITY
JOHN F. KENNEDY SCHOOL OF GOVERNMENT

Faculty Research Working Papers Series

The Parable of the Hare and the Tortoise: Small Worlds, Diversity, and System Performance

David Lazer and Allan Friedman

October 2005

RWP05-058

The Parable of the Hare and the Tortoise:
Small Worlds, Diversity, and System Performance

David Lazer
(david_lazer@harvard.edu)

Allan Friedman
(allan_friedman@ksg.harvard.edu)

Kennedy School of Government
Harvard University
Cambridge, MA 02138

This material is based upon work supported by the National Science Foundation under Grant No. 0131923. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF). The authors also acknowledge helpful feedback from Robert Axelrod, Robert Behn, Michael Macy, Robert Putnam, Ezra Zuckerman, as well as participants in the Faculty Research and Work in Progress Seminars at the Kennedy School.

The Parable of the Hare and the Tortoise:

Small Worlds, Diversity, and System Performance

David Lazer and Allan Friedman

Whether as team members brainstorming, or cultures experimenting with new technologies, problem solvers communicate and share ideas. This paper examines how the structure of these communication networks can affect system-level performance. We present an agent-based model of information sharing, where the less successful emulate the more successful. Results suggest that where agents are dealing with a complex problem, the more efficient the network at disseminating information, and the higher the velocity of information over that network, the better the short run and lower the long run performance of the system. The dynamic underlying this result is that an inefficient network is better at exploration than an efficient network, supporting a more thorough search for solutions in the long run. This suggests that the efficient network is the hare—the fast starter—and the poorly connected network is the tortoise— slow at the start of the race, but ultimately triumphant.

Networks provide the architecture of systemic adaptation. In any social system, the units of that system seek success—whether those criteria for success are determined by evolutionary processes or by (boundedly rational) goal-seeking processes. Networks provide the architecture of success in social systems because information about success that flows through the network provides one of the primary mechanisms for improvement in the system. This paper explores the implications of this architecture for system-level success (e.g., Putnam 2000), rather than nodal-level success (e.g., Burt 1995). In particular, using computational experiments, it demonstrates how network architecture affects the aggregation of problem solving within a system. It finds that small world (small number of degrees of separation) networks (the hare) are better at quickly converging on the best solution that initially exists in the network. However, small worlds perform very poorly in the long run, because they so quickly squeeze diversity out of the system. Larger world networks (the tortoise) perform better in the long run because they explore more of the solution space. More generally, we find that the smaller the world, the better the system in the short run, and the worse it is in the long run. Finally, we find that the velocity of information is negatively related to performance, especially in highly connected worlds.

We argue that these findings are robust across a wide variety of settings, offering illustrations from the diffusion of pre-written history agricultural innovations and creative groups.

The first section of the paper examines the role that networks play in the diffusion of information, focusing in particular how networks affect the aggregation of information. The second part of the paper develops a simulation model, basing the solution space on NK models (Kauffman 1995), to understand the implications of emulation processes, where it is assumed that actors emulate the most successful other actors in the space. The third part of the paper discusses the broader relevance of these ideas, and the conclusion examines potential directions for future research.

Network configuration, information diffusion, and systemic performance

There is ample research that highlights the role of social networks in providing information to individuals. Diffusion of innovation research (Rogers 2003) illustrates how the well connected tend to be early adopters. Acquiring information quickly, in turn, provides advantage to individuals. Thus, for example, Granovetter's landmark research (1973) highlights the role of relationships in funneling information to individuals about employment opportunities, and Burt's (1995) research focuses on the competitive advantage that individuals/businesses gain if they bridge otherwise unconnected groups.

There is less research on systemic network structure and systemic-level success. Further, the individual-level results do not necessarily scale to the systemic level, because many of the individual-level results may reflect competitive advantage in zero sum games (e.g., finding a job). Putnam's (1993) enormously influential work found a strong relationship between associational affiliations (presumably correlated with density of networks) and the success of regional governments. He later cautioned about the social and economic declines which could come from the erosion of those connections (2000). Granovetter (1973), while primarily focused at the nodal level, did present a comparison of two communities— one with more casual social interactions than the other—and argued that the former is better equipped to overcome collective action problems. There is also a recent vein of research studying small group networks suggesting that denser ties among group members is correlated with success (examples include Baldwin, Bedell, and Johnson 1997; Cummings 2004; Reagans and Zuckerman 2001; see Katz et al 2004 for a review). It is not clear, however, what role information dissemination plays in any of these results, as compared to other processes that might mediate the relationship between network configuration and success (e.g., control and coordination).

The relationship between network ties and collective outcomes has not been shown to be entirely positive, however. Putnam (2000) acknowledged the dark side of social capital when bonding ties insulate a community. Uzzi and Spiro (2005) found an inverse U relationship between the density of ties among people who work on Broadway and the overall economic success of Broadway, suggesting that a network can be too dense.

There are, however, compelling reasons to believe that network structure does affect informational factors that, in turn, affect systemic success. The diffusion literature (Rogers 2003; Coleman et al. 1966) largely focuses on the diffusion of practices (e.g., penicillin in Coleman's study) that improve the well-being of individuals (and not at the cost of other individuals). Of course, not all innovations are welfare improving, and the informational cascade literature (Banerjee 1992; Bikhchandani, Hirshleifer, and Welch, 1992) highlights how non-welfare improving practices (fads) can spread through the system under the assumption that people observe adoption decisions and not success of actors. Strang and Macy (2001) extend this essential finding to the circumstance where actors can observe success of other actors but not the reason for that success. This extension models circumstances in which public adoption decisions may outweigh private signals people have that (if aggregated) would point the opposite direction. The information cascade literature highlights the more general issue of information aggregation within networks—how well does the system pool together unique signals?

A critical question is then how different network structures affect the aggregation of information and problem solving. The enormous growth of interest in how information flows through macro social structures, following Watts and Strogatz (1998) and Watts (1999), highlights the need to understand the underlying process. It is therefore striking that there is little research at the intersection of information aggregation and social networks (for an exception, see Watts 2002). The diffusion and social influence literatures (e.g., Friedkin 1998; Lazer 2001) highlights the potential role of the network in a cascade—e.g., emulation occurs because two people have a relationship. The literature on institutional isomorphism (DiMaggio and Powell 1983) highlights a necessary corollary to this mimetic process—that imitation reduces the diversity present in a system. Diversity, in a setting where agents are continually searching for improvement, defines the set of options for agents. Following this, as has been found in a variety of settings—democratic deliberation (Sunstein 2003); small group performance (Nemeth 1985); entrepreneurial systems (Florida 2002); and elite decision making (George 1971; Janis 1971)— reduced diversity can result in reduced performance. There is thus a potential trade-off between maintaining diversity and mechanisms that diffuse information and thus set into motion these mimetic processes (see figure 1).

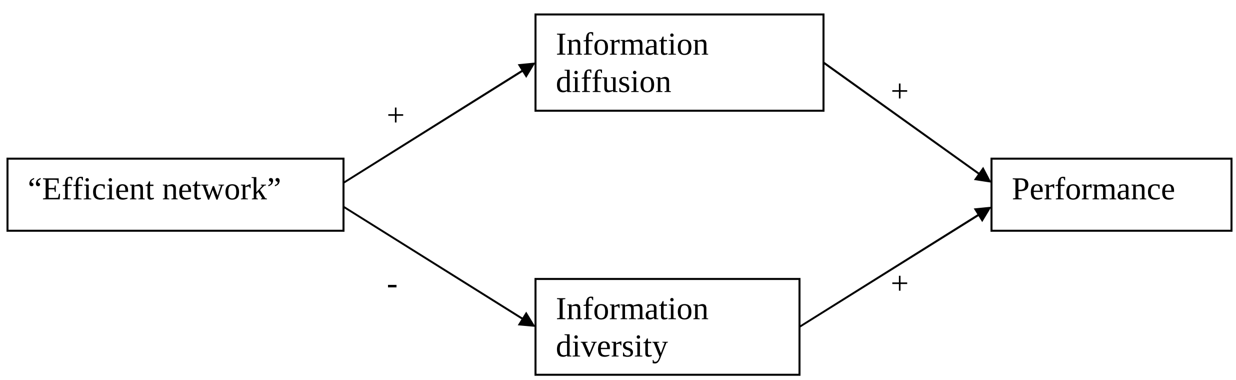

Figure 1: Trade-off between information diffusion and diversity

Our objective is to explore this trade-off through computational models of information diffusion in a variety of network structures.

Research Design

We begin with a simple scenario to motivate our model: imagine a world full of people trying to solve the same problem. The way they tackle that problem is through two mechanisms: (1) imitation of others more successful than they, where they can only see the solutions and success of people they are connected to; and (2) tinkering with their current solution. Actors seek to better their own position, but in a non-zero-sum game that is neither explicitly competitive or collaborative. This scenario, we would argue, captures the essence of a non-trivial set of human phenomena, whether it is about people trying to succeed within an organization, businesses selecting strategies, or governments seeking to improve public policies. The question of information aggregation shifts to a question of search aggregation, where the key issue is how to best aggregate the searches of many agents. The architecture of that aggregation process, and the focus of this paper, is the network connecting the agents.

There are thus three components to this computational model: (1) the problem the agents are trying to solve; (2) the rules governing agent behavior; and (3) the configuration of the network(s). We discuss each in turn, and then summarize how we will analyze this model.

The problem space

From the perspective of this analysis, the critical characteristic of the problem space is its topography. In a "simple" problem space there is only one optimum, where incremental changes by agents will eventually yield that optimum. Our assumption is that most problems of interest in the world are characterized by far messier problem spaces, with many local optima that are vastly inferior to the global optimum. There are some standard complex problems that computer scientists use to test search algorithms (e.g., the traveling salesman problem), however, we choose to follow Levinthal (1997), Gavetti and Levinthal (2000), Rivkin (2000), Carroll and Burton (2000) and several others in using the "NK" model to produce problem spaces. The NK problem space is named for the two parameters that are used to randomly generate problem spaces. It was originally developed by an evolutionary biologist, Stuart Kauffman (1995) to model *epistasis*—where genetic traits interacted so as to produce fitness that was greater (or lesser) than the sum of the contributions of the individual traits. From our perspective, its key features are (1) that we can produce an arbitrary number of problem spaces from these two parameters, and (2) that we can "tune" how rugged the space is through the two parameters (more on this below).

An NK space is represented as a string of N numbers, usually 0 and 1, and each string has an overall score that represents how "high" the peak is. Each digit contributes to the score of the total string, but that contribution may be dependent on other digits. For example, in Figure 2, the first digit's contribution to the score will change if either the first digit, the third digit or the fifth digit changes. The contributions of each digit are then averaged to form the score of the entire string.

1 1 *0* 1 *0* 1 1 1

001 → .5453
010 → .44893
011 → .8374
100 → .2037
…

Figure 2: Computing an NK score. Each digit contributes to the final score, and each digit's contribution is dependent on K other digits. Above, the first digit adds 0.2307 to the sum of all digits, which will be averaged to produce the NK score.

It is not possible to draw an NK landscape because it is N dimensional; figures 3a, 3b, and 3c offered a stylized representation of the impact of manipulating K. K = 0 creates a simple problem space with a single optimum; and K= N-1 creates a maximally rugged landscape, where the performance of any given solution in the space offers no signal as to the quality of adjacent solutions. Of greater interest is the universe of spaces in between, where there are local optima, but where the quality of adjacent solutions is correlated. There is no best way to find the optimal point in an unknown rugged space, but a local maximum can be easily found by searching up hill. However, there is no mechanism to prove whether or not a local maximum is the global maximum without measuring every other peak.

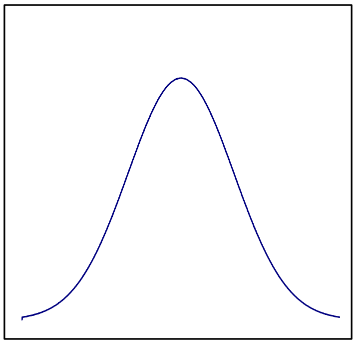

Figure 3a: A simple problem space, similar to $K= 0$

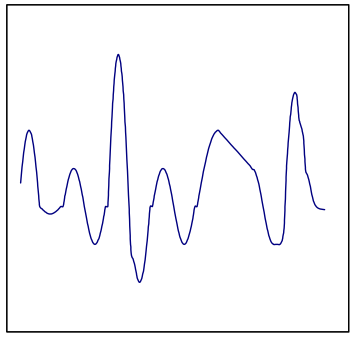

Figure 3b: A complex rugged space with local maxima and minima, similar to $0 < K < N - 1$

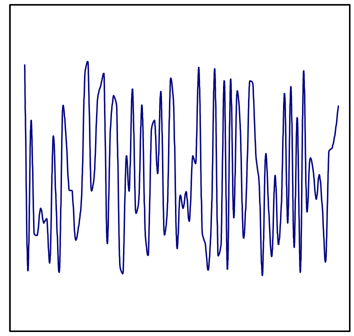

Figure 3c: A chaotic space where the value of every point is inpendent of adjacent points, similar to $K = N - 1$

In the computational experiments summarized below, we examine how the mean score from each agent's NK string changes over time. Because different NK spaces will have different structures, and some will have larger global maxima (high score) than others, we normalize the score against the global maximum for that space with the ratio r_{NK}. This produces a distribution like that seen in the thick line in Figure 4. In most interesting problems, most solutions are (very) bad, and only a few are good. To achieve this distribution of quality, we raise this ratio to the eighth power. This monotonic transformation does not alter the ranking of solutions, but distributes them numerically to show that some are better than others more clearly. This gives us a distribution like that seen in Figure 4, and allows us to see the results of distinct models below.

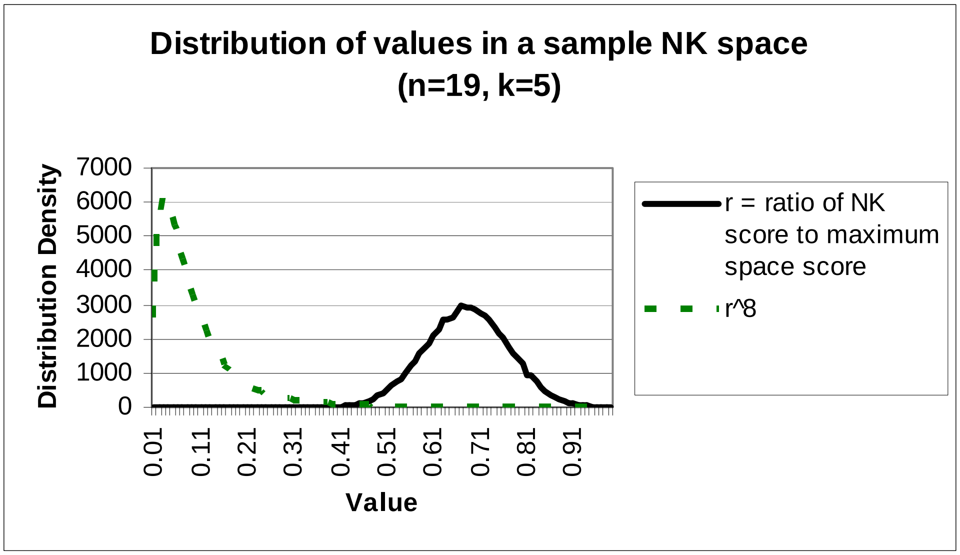

Figure 4: Density curves of the value distribution for an NK space. We use the monotonic transformation r^8 to skew normalized scores to demonstrate the relative value of higher scores.

The scores reported are the ratios, r_{NK}^{8}, derived for each NK space.

Agents' behavioral rules

Each "world" is made up of a number of agents, where each agent is connected to a subset of other agents. The agents sit in the space, and the score of each agent is the score of the point on which they sit. The underlying assumption is that the space is vastly largely than the search capabilities of any agent. The objective of the simulation is to understand how each agent looks around the neighborhood of its current solution to produce novel (if incremental) solutions, versus emulating the solutions of successful actors that they are connected to.

We assume that each round in the simulation, if an actor sees another actor who is performing better, they adopt that other actor's solution. Further, we assume that if an actor does not see another actor who is performing better, they examine the impact of randomly changing one digit of their status quo solution, and if that change results in an improvement they change their solution. Following March (2001), we would view the former as "exploitation"—taking advantage of what the system already "knows"—and the latter as "exploration"—attempting something novel.

Thus, an agent will mimic other successful agents, and when there is no one to mimic, they will attempt to adapt. New successful adaptations will subsequently be copied by neighbors, and so on. If no agent (or subset of agents) is isolated, then all will eventually converge on the same solution.

Network configuration

One of the central questions we examine below is the impact of changing the network structure on the performance of the system, given the aforementioned search behaviors of the constituent agents. In the simulations below, we examine four types of networks: a linear network, a fully connected network, a variety of random networks, and a variety of small world networks. We describe each in turn below

Linear Network: A linear network is simply a set of nodes where each node, except for two, communicates with two other nodes (it is assumed in all of these networks that communication is two way). The nodes, and their relationships, may thus be arrayed in a linear fashion (see figure 5a). A linear network produces the maximum degree of separation between an average pair of nodes possible in a fully connected network.

Fully connected network: A fully connected network is one in which every node communicates with every other node (see figure 5b).

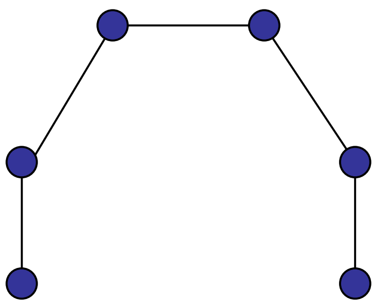

Figure 5a: A linear network

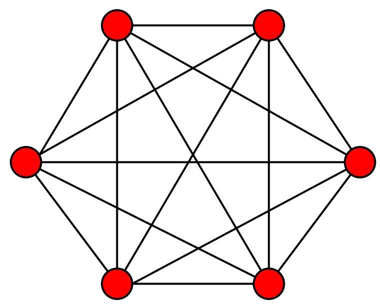

Figure 5b: A fully-connected network

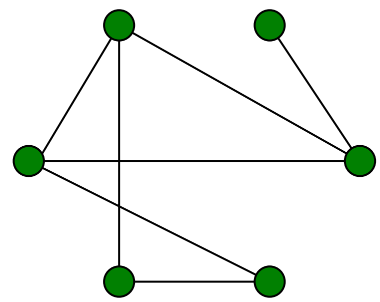

Figure 5c: A random network (p=.3)

Random network: The random network (Erdos and Renyi 1959) is created by creating a link between any two nodes with probability p (we examine below the impact of varying p). Figure 5c offers an illustration of a randomly generated network.

Small world network: In order to examine the importance of average path length, we use Watts and Strogatz' small world model (1998). By rewiring a defined lattice (in this case, where each node talks to its immediate four neighbors—see figure 6a), the number of links is held constant, and thus density is held constant. As figure 6b illustrates, switching a local link to a long distance link can significantly lower average path distance. More generally, as the probability of rewiring increases, the average path distance between two nodes will drop rapidly, making the network into a "small world."

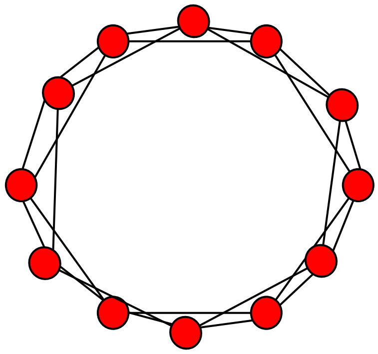

Figure 6a: A regular lattice with each node connected to its 4 closest neighbors, with an average path length of 1.909

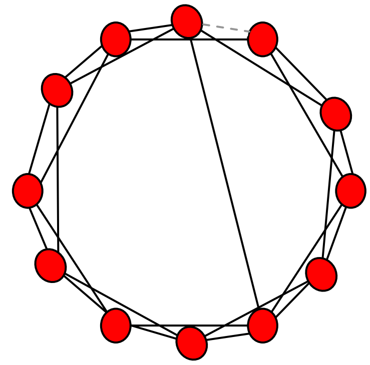

Figure 6b: The lattice with one link rewired to connect to a distant node, shortening the average path length to 1.833. In a 100-node lattice, a single rewired link can reduce average distance by over 20%.

Analysis

The model is implemented in Repast[1], an object-oriented software library for agent-based modeling. For each simulation, we assume a population of 100 agents who interact for discrete time steps: the population is large enough for diffusion processes to occur, but small enough to credibly represent real world phenomena such as organizations. The solutions of the initial population of agents are randomly generated, and those agents are randomly placed in networks with pre-existing configurations. For the NK model, we assume N=19,[2] and K=5, except where noted below. The interlinking factor of 5 produces a space that is moderately rugged—with many local optima, but where the quality of proximate solutions are highly correlated.

For each simulation, the population is run to the point that it converges on a single solution.[3] For each network and set of parameter values, the simulation is run 1000 times. For each set of 1000 simulations, we show the average performance of the set of simulations (measured by the average performance of agents).

For each set of simulations with Population = 100, N = 19, and K = 5, we use the same random seed of 1000 starting points of solutions. In appendix A, we produce a summary of the long run performance of these simulations, indicating how often each network came in first, second, etc. In the analyses summarized below, we examine the impact of varying the configuration of the network, the ruggedness of the problem space, and the "velocity" of diffusion.

[1] Repast is available here: http://repast.sourceforge.net/ Simulation code was written in Java and is available at http://allan.friedmans.org/networks/

[2] A problem space of 524,288 possible solutions.

[3] When all nodes are connected in a complete graph, all nodes will converge. In situations with isolated cliques or nodes, models are run until every clique has converged on a solution.

RESULTS

We analyze the impact of varying three components of the above model: the configuration of the network, the ruggedness of the problem space the agents are scaling, and the velocity of information.

Configuration

We compare the performance over time of each of the network archetypes summarized above: the linear network, the fully connected network, the random network, and the small world network.

Linear vs Fully connected networks: The extreme opposite archetypical networks summarized above are the linear and fully connected networks. Figure 7 plots the average performance of 1000 linear networks versus fully connected networks over time.

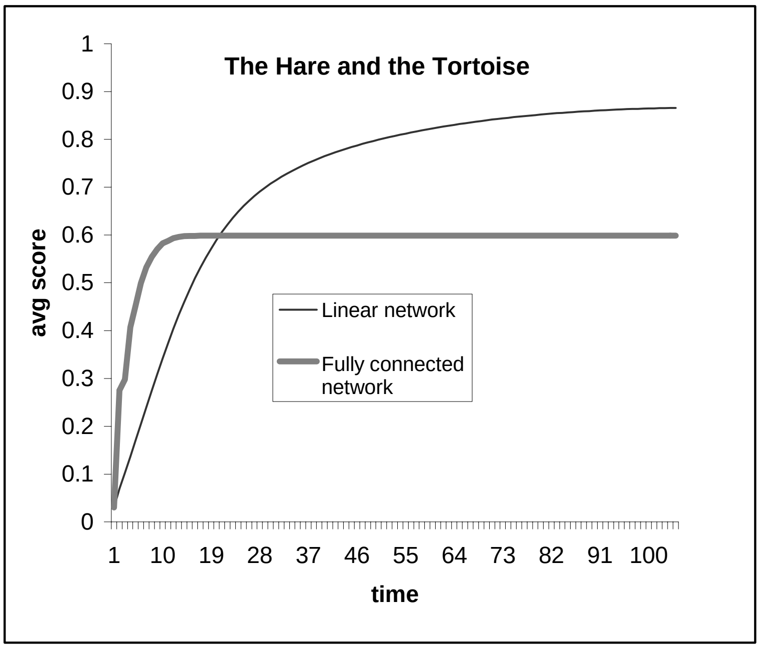

Figure 7: The fully connected network finds a good solution quickly, but the linear network finds a better solution on the long run.

Figure 7 demonstrates a clear pattern: the fully connected network outperforms the linear network in the short run, but in the long run (like the tortoise), the linear network performs significantly better. The reason for this pattern is captured in figure 8, which plots the average number of distinct solutions in the system over time in the linear and fully connected networks.

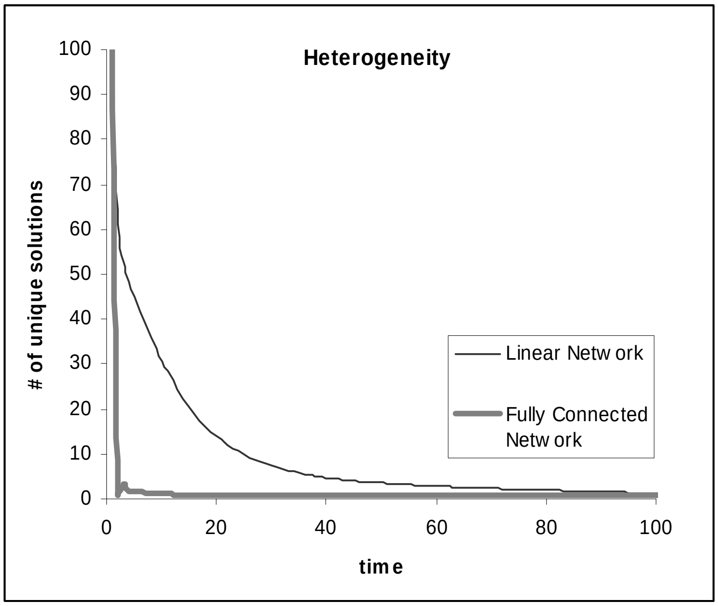

Figure 8: The number of unique solutions held by agents plummets in the fully connected graph, while the linear network maintains heterogeneity for longer.

Notably, the full network quickly drives out heterogeneity. In fact, there are only one or two strategies after the first round—because 99 agents converge on the strategy of the best performing agent, at which point the agents explore (because their performance is identical) and then again converge, and so on, resulting a the "bouncing" observed in the figure above. The system can, at best, only find the best local optimum that is reachable by climbing uphill from the best strategy that exists in the population when it is initially seeded. The linear network, however, drives out diversity far more slowly, allowing exploration around a number of the better strategies that exists in the initial population set.

Random networks: In order to further explore this dynamic, we examine the performance of random networks over time. Figure 9 plots the average long run performance of random networks against the density of the random network for all random networks generated, as well as for the subset of random networks that were a single component (i.e., where there was a path that leads between any two nodes).

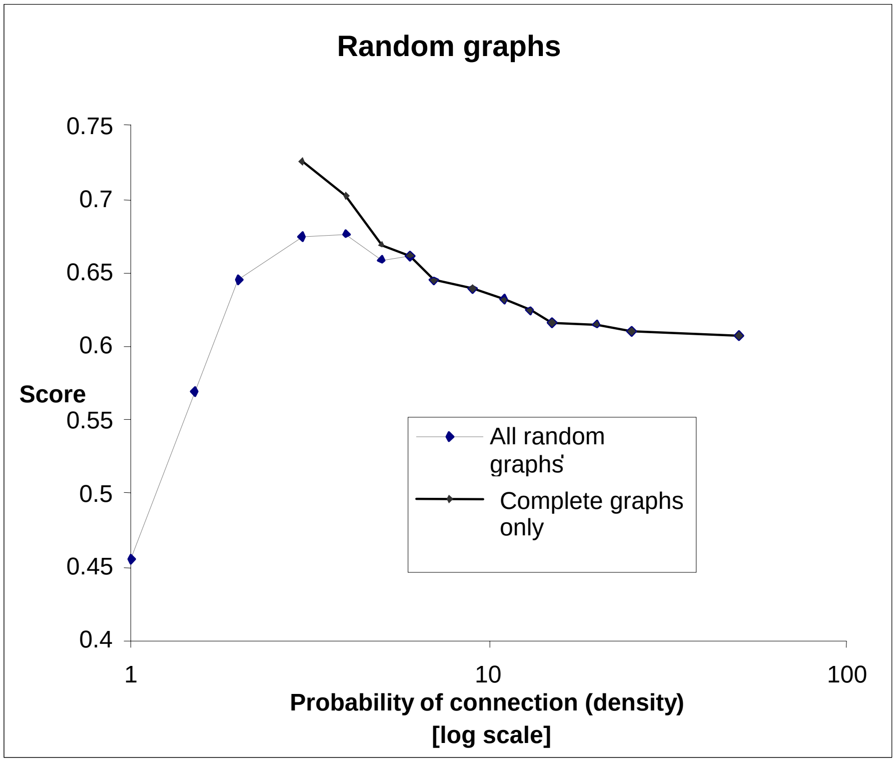

Figure 9: Simulation results appear to show a curvilinear relationship between density and social outcomes. This is an artifact of sparse stochastic networks which are not completely connected. Looking only at graphs with not isolated cliques, performance declines with density.

The random network demonstrates a distinct inverted U shape—very low densities and very high densities are low performing, and the optimum performance is at in between densities. The censored data reveal the underlying process. At very low densities, the network breaks up into multiple components. For a given structure, smaller populations perform less well than larger populations because they start from a smaller set of possible solutions.[4] The very low densities thus are made up of a number of small (and thus poorly performing) subgroups. However, if one only looks at the random networks that are a single component, the inverse U shape disappears—because the less dense networks are better at preserving diversity.

Small world networks: In order to test the impact of average path distance, we examine the average long run performance of a variety of small world networks, varying the probability of rewiring, where density remains constant but path distance declines with increased rewiring. As with the random network, this rewiring process can produce isolated components. Discarding these components produces the desired effect, but leaves a potentially biased sample. Any rewiring process that does not leave isolated cliques is more likely to have shortcuts, possibly over-estimating the effect of rewiring. To get around the unpredictability of the Watts-Strogatz model, Newman and Watts (1999) developed an alternate small world model where shortcuts are added between two

[4] Results on population size are available from the authors upon request.

random points in a lattice, rather than being rewired. Figure 10 presents average long run score as a function of the number of added shortcuts.

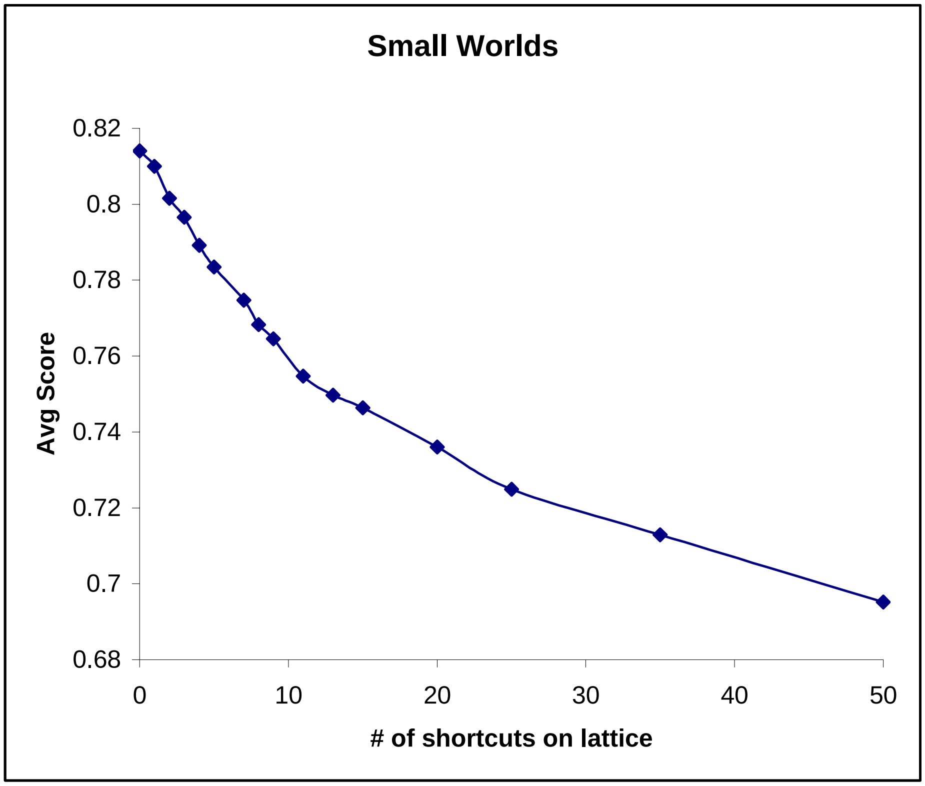

Figure 10: As the network becomes more of a small world by adding shortcuts, the average social outcome declines. Note that this model implements (Newman and Watts 1999) instead of (Watts and Strogatz 1998).

Strikingly, and consistent with the preceding results, the average long run scores decline monotonically with the increase in probability of rewiring.

Ruggedness of problem space

One critical question is whether the results above are contingent on the configuration of the problem space. In order to examine whether this is the case, we manipulate the K parameter of the NK space, setting it to 0 (producing a space with a single optimum). Figure 11 summarizes the relative average performances of the linear and fully connected networks.

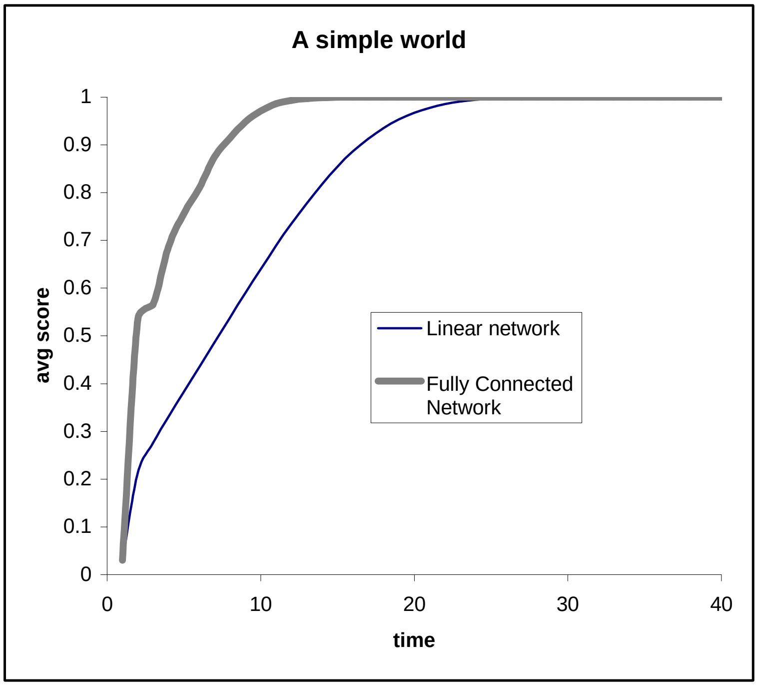

Figure 11: In a simple world, every network finds the global optimum, but the more connected network finds it faster.

Figure 11 highlights that the preceding results are contingent on the ruggedness of the problem space. Both the linear and fully connected networks find the global maximum in the long run, but the fully connected network gets there much faster. The essential dynamic in the fully connected network is the same as in the more complex world: all agents are pulled up to the strategy of the highest performing agent at the beginning of the simulation. The key difference, of course, is that in the simple world the system does not get caught in a local optimum.[5]

Velocity of information

The results above suggest that the configuration of the network has a strong effect on the balance between exploration and exploitation within the system. The focus of the above analysis was on the structure of the network and the problem. Here we look at an additional dimension: the velocity of information in the system. In particular, in the above models, it is assumed that the opportunity to emulate successful others occurs every round. Here we assume that this opportunity may come less often—that most of the time agents are "muddling" around with incremental exploration, and that periodically agents look at the agents that they are connected to, and emulate any agents (if any) that are performing better than they. We label the frequency that agents look around as the *velocity of information*. We model this in two fashions: (1) where communication is synchronous (every agent looks around at the same time); and (2) where communication is asynchronous (where every agent looks around with the same

[5] Notably, even a small degree of ruggedness confounds the fully connected network, which is (on average) beaten by the linear network when K = 1.

average frequency, but where the actual timing for each agent is stochastic). Formally, we represent velocity with the parameter v, where for the synchronous model, all agents emulate successful others every v rounds; and for the asynchronous model, there is a 1/v probability of emulating successful others each round for each agent. Figure 12 summarizes the average performance over time of the fully connected and linear networks for asynchronous and synchronous communication for v = 10.

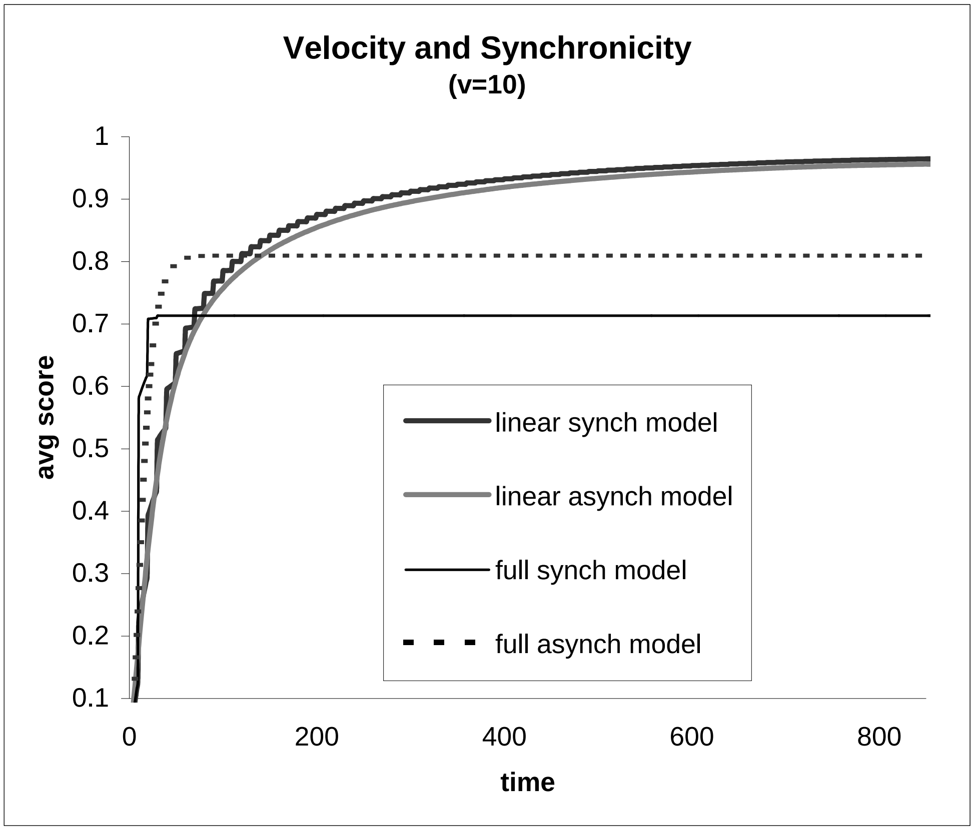

Figure 12: When communication is irregular, overall scores improve. An asynchronous model is good for a well-connected network, but the average solution of the linear network is slightly better using synchronous communication.

Figure 12 highlights a number of striking patterns. First, slowing down the velocity (increasing v) of information uniformly increases long run performance for both networks (at the expense of short run performance). Second, there are striking differences in effects between synchronous and asynchronous communication, which, in turn, are contingent on network structure. For the fully connected network, asynchronous communication performs notably better than synchronous communication. For the linear network, this pattern is reversed—the synchronous linear network on average outperforms the asynchronous linear network. Plots of the average number of strategies in existence in the network offer some insight into the underlying processes (see figure 13).

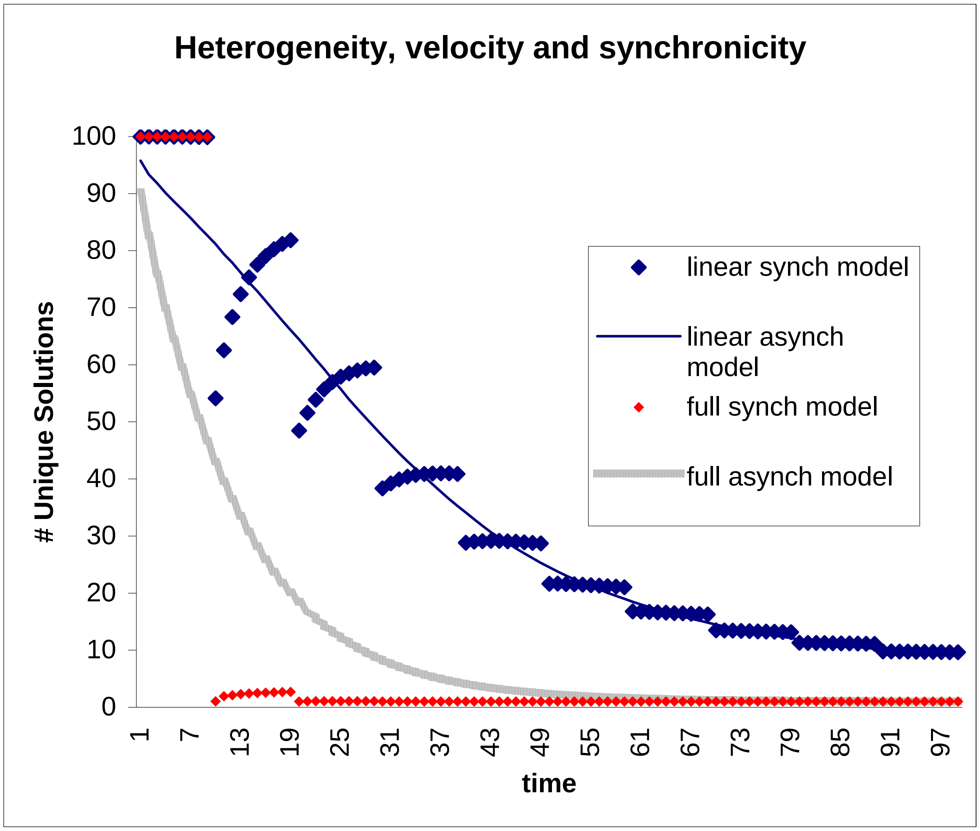

Figure 13: Asynchronous communication makes networks behave as if they were less connected. Synchronous communication drives irregular behavior on a linear network displays because each agent must explore independently in periods of no communication. The full network, on the other hand, converges within two synchronous cycles.

Figure 13 reveals that in the fully connected network, synchronous communication drives out diversity much more rapidly than asynchronous. Specifically, in the 10th round, when there is a spike in performance, all heterogeneity is eliminated as all agents converge on a single strategy. This strategy is then used as a foundation for experimentation the following 10 rounds by all agents, at which point the system then converges again. In contrast, diversity is squeezed out more slowly when there is asynchronous communication, providing for greater exploration and greater long run performance by the system.

The reason for the (smaller but statistically significant) superiority of synchronous linear networks to asynchronous linear networks is less clear—neither is clearly superior at maintaining diversity. Our conjecture is that synchronous communication is more effective in the linear network at pruning truly unpromising paths toward good solutions, by eliminating the very worst strategies every 10 rounds, and thus causing convergence on an exploration based on a smaller number of promising paths toward good solutions.

Discussion

It is beyond the scope of this paper to rigorously test the implications of this model against real world data. However, a number of existing studies have results consistent with this model. Here we will briefly highlight two studies at the opposite ends of the micro to macro spectrum: Diamond's (1999) analysis of the competition among civilizations, and Leenders et al. (2003) study of the role of connectedness on product development teams. We also discuss briefly the relationship between this model and evolutionary-based models.

Competition among civilizations: Diamond (1999), in *Guns, Germs, and Steel*, poses the question: why, over the last several centuries, did European-based civilization "triumph" over civilizations from other continents? His extended answer focuses on the locational advantages of Europe over other continents, an important component of which was the advantages of the architecture of the informational networks of Europe. Examining the archaeological evidence, he points out that there was very little diffusion of agricultural innovations in the Americas and Africa, because the axis of those continents was largely north-south, but compatible climates (generally) run east-west. Eurasia thus provided a larger band of connected compatible climates, and thus vastly larger network to produce and diffuse information. The second question is why did Europe eventually dominate Asia? His answer: the networks of Asia were too efficient, because of the smoothness of the geography. The more rugged topography of Europe impeded but did not fully obstruct communication, which in turn fostered greater heterogeneity and thus greater exploration.

Creativity on teams: Building on the tradition of Janis' (1982) work on Groupthink, and Nemeth's (1985) work on minority views in groups, Leenders et al. (2003) show a curvilinear relationship between success at new product development and connectedness, and a negative relationship between success and centralization of team communications. Their analysis is strongly compatible with our results above: low levels of connectedness mean that ideas do not flow within the group, however high levels of connectedness result in premature convergence on a solution—i.e., inadequate exploration. Our results (in particular, figure 9) also suggest an important contingency in their findings (and other analyses that have found an inverse U relationship between performance and density of connections, such as Uzzi and Spiro (2005)): that the poor performance of poorly connected groups *is contingent on the presence of multiple components*.

Evolution and evolutionary modeling: A style of programming called genetic algorithms is based on a different but related model of diffusion to the social emulation process we examine here (Holland 1975). In particular, it is based on an evolutionary metaphor, where traits are based on genetic profiles, and more successful strategies are more likely to reproduce. When strategies reproduce, they recombine to produce new strategies (with occasional mutation).[6] This modeling approach has been used to study social phenomena (e.g., see Axelrod 1987) where it is plausible that such evolutionary processes are at work. It has also been used with significant success to "breed" programs to solve problems (see www. genetic-programming.com). A key aspect of successful applications is the structure of interaction among competing programs. Recent findings suggest that a superior approach incorporates the division of the population of competing programs into subpopulations within which there is competition and reproduction, and between which there is only occasional interaction (Koza et al. 2003a; 2003b).

[6] In future analyses we will explore the analog to cross-over in our model—errors in emulation, which would yield strategies that are a mix of the strategies of the copier and the copied.

Consistent with the simulations above, a single integrated population of programs produces mediocre and stable homogeneity quickly. Dividing the population into cliques allows the exploration of a variety of strategies, where the occasional intermingling, in turn, allows the recombination of those alternative successful approaches into even more successful solutions.

The process of natural evolution itself depends on a reduction in communication to evince rapid change, through allopatric speciation. Evolutionary theorists argue that large integrated populations can damp out new mutations, and eliminate speciation (Gould 1980).

Conclusion: the dark side of the small world

We take it as a stylized, but indisputable, fact that we live in a smaller world today than ever before, where there are more distant linkages which rapidly spread information from one corner of the globe to the other, and, further, that the velocity with which information flows over those networks has increased enormously just in the last decade. We also observe that the last decade has witnessed a massive surge of interest in networks, both within the academy (Borgatti and Foster 2003), and within popular culture. There has been a proliferation of services to increase the efficiency with which we exploit our personal networks, and armies of consultants have emerged to improve the efficiency of organizational networks. Silo's are to be eliminated; the boundaryless organization is the wave of the future. Technologies that enable distant actors to access each other's knowledge are to be utilized (Wenger, McDermott, and Snyder 2002). Even the most hierarchical of organizations, the military, is ostensibly shifting to a more flexible, network-based system (Arquilla and Ronfeldt 2001), with bottom-up, boundary-spanning, web-based knowledge systems (Baum 2005).

Our results suggest a potential downside in the rush to connectedness: that it will homogenize our world and limit our collective future menu of choices. We conducted a series of computational experiments in which all agents were solving same problem, where those agents followed a simple behavioral rule: copy those more successful than you (exploitation), otherwise tinker with your current solution (exploration). We manipulated (1) the structure of the network (who agents could emulate); and (2) the velocity of information—i.e., the frequency with which emulation occurred. Appendix A summarizes a pairwise comparison of all of the networks we studied, where the smaller the world and the faster the velocity of information, the worse the long run performance of that world. Remarkably, the highest performing network in the long run was the *slowest velocity linear network*—the network with the least efficient mechanisms for diffusing information among fully connected networks. Thus, for example, the fast fully connected network outperformed the slow linear network only 9 times out of 1000 experiments. We also found that in the highly connected world for a given velocity of information, synchronized communication (e.g., one might think of small group meetings, conferences, listservs) caused performance to deteriorate, and in the poorly connected world, actually improved performance.

This model is highly extensible, and we would point to a number of directions for future research. The first is to continue looking for the "optimal" network for a given situation, where the trade-off between time and score can made explicit with discounting functions to emphasize early success, reflecting the fact that, in many networked interactions, time is not unbounded.

Another avenue of exploration would be to examine the impact of changes in the emulation process. For example, agents may be imbued with more complex copying behavior. Imitation is seldom perfect, and introducing errors in copying would both hamper the spread of successful solutions and generate “accidental” experimentation. Perhaps they are more likely to copy a “strong” tie located locally than a weak tie across the network, or vice versa. Alternatively, agents can populate a world where innovation is rewarded as a goal itself, such that they are less likely to copy a popular successful solution; risk-averse agents may be unwilling to explore at all and only copy.

Finally, the landscapes that agents are “climbing” might differ in subtle ways, so that their topography is similar, but not identical, reducing, but not eliminating the benefits to emulation and the pressure toward homogeneity.

We therefore assert that our contribution is two-fold. First are the provocative findings that are summarized above. These findings, as with any model, are driven by our assumptions, and thus raise far more questions than they answer. These questions do not undercut our results as much as they highlight the extensibility of a fairly simple modeling framework. We believe that our second contribution is therefore a simple baseline model with which to study the very general problem of the role of network structure in the aggregation of problem solving.

Appendix A: Pairwise comparison of network models

The data presented above summarizes the average score of a range of network models on the same set of 1000 NK spaces with a set of fixed starting points on each space. Since the only variation was across network structure, we can compare each of the 1000 initial configurations with the results of other models. Note that, for the small world, the random network and the asynchronous velocity models, the use of a random seed in the network generation or model execution produces non-deterministic results. The following table summarizes those comparisons.

	Full	Line	Lattice	SmallWorld (10 added)	Line synch	Line asynch	Full synch	Full asynch	Random (p=.07)
Full		26	50	88	9	9	276	35	138
Line	845		468	605	73	105	715	501	787
SW_0	762	216		478	57	68	631	386	684
SW_10	632	132	222		27	44	524	277	557
Line synch	916	578	710	785		260	846	727	893
Line asynch	918	572	703	789	188		841	716	891
Full synch	616	155	244	341	23	45		241	524
Ful asynch	715	256	370	465	63	80	630		652
Rand_07	320	54	83	136	11	14	336	84	

Table A: Number of times that the row network has a *higher* score than the column network on the same set of problem spaces and starting points.

Diagonal elements do not sum to 1000 because of ties. Ties occur most frequently when both models reach the global maximum.

References:

Arquilla, John, and David Ronfeldt. 2001. Networks and Netwars: The Future of Terror, Crime, and Militancy. Santa Monica, CA: Rand.

Axelrod, Robert. 1987. The evolution of strategies in the iterated Prisoners' Dilemma. In L. Davis, editor, *Genetic Algorithms and Simulated Annealing*. Morgan Kaufmann, Los Altos, CA.

Baldwin, Timothy T., Michael Bedell, and Jonathan Johnson. 1997. The social fabric of a team-based M.B.A. program: Network effects on student satisfaction and performance. Academy of Management Journal, 40: 1369-1397

Banerjee, Abhijit V. 1992. "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107/3: 797-817.

Baum, Dan. January 17, 2005. "Battle Lessons: What the Generals Don't Know." *New Yorker*.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. "A Theory of Fads, Fashions, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100/5: 992-1026.

Borgatti, Stephen P., and Pacey C Foster,. 2003. The network paradigm in organizational research: A review and typology. *Journal of Management*. 29(6): 991-1013

Burt, Ronald. 1995. *Structural Holes: The Social Structure of Competition*. Belknap Press.

Carroll, Tim, and Richard M. Burton. 2000. "Organizations and Complexity: Searching for the Edge of Chaos" *Computational and Mathematical Organization Theory;* 6/ 4: 319-337.

Coleman, James S., Ernest Q. Campbell, Carol Hobson, James McBartland, Alexander Mood, Frederick Wernfield and Robert York. 1966. *Equality of Economic Opportunity*. Washington, DC: U.S. Government Printing Office.

Cummings, Jonathan. 2004. "Work Groups, Structural Diversity, and Knowledge Sharing in a Global Organization." *Management Science* 50/3. 352-364.

Diamond, Jared. 1999. *Guns, Germs, and Steel*, W.W. Norton.

DiMaggio, Paul, and Woody Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological* Review (April 1983).

Florida, Richard. 2002. *Rise of the Creative Class*. Perseus Book Group.

Friedkin, Noah E.1998. A Structural Theory of Social Influence. Cambridge: Cambridge University Press

Gavetti Giovanni, and Daniel Levinthal. 2000. "Looking forward and looking backward: Cognitive and experiential search." *Administrative Science Quarterly* 45 113-137

George, Alexander L., "The Case for Multiple Advocacy in Making Foreign Policy," *American Political Science Review*, vol. 66, no. 3 (Sept. 1972).

Gould, Stephen Jay. 1980. *The Panda's Thumb.* New York: W.W. Norton & Co.

Granovetter, Mark. 1973. The strength of weak ties. *American Journal of Sociology*, 78, 1360-1380.

Holland, John H. 1975. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor.

Janis, Irving *Groupthink*, 2d ed., Houghton Mifflin, Boston, 1982.

Katz, Nancy, David Lazer, Holly Arrow, and Noshir Contractor "The Network Perspective on Teams," *Small Group Research*, June 2004

Kauffman, Stuart. 1995. *At Home in the Universe*.

Koza, John R., Martin A. Keane, and Matthew J. Streeter. 2003a. Evolving inventions. *Scientific American*. February 2003. 288(2) 52 – 59.

Koza, John R., Martin A. Keane, Matthew J. Streeter, William Mydlowec, Jessen Yu, and Guido Lanza. 2003b. *Genetic Programming IV: Routine Human-Competitive Machine Intelligence*. Kluwer Academic Publishers.

Lazer, David. 2001. "The Co-evolution of Individual and Network," *Journal of Mathematical Sociology*, January, 69-108.

Leenders, Roger Th.A.J., Jo M.L. van Engelen, and Jan Kratzer, 2003, "Virtuality, Communication, and New Product Team Creativity: A Social

Network Perspective.” *Journal of Engineering and Technology Management*, 20: 69-92.

Levinthal, Daniel. A. 1997. Adaptation on rugged landscapes. *Management Science*. 43 934–950.

March, James. 1991. “Exploration and exploitation in organizational learning.” *Organization Science*, 2(1): 71-87

Nemeth, Charlan J. 1985 `Dissent, Group Process, and Creativity' *Advances in Group Processes* Vol.2 pp.57-75.

Newman, Mark E.J., and Duncan Watts. 1999 “Renormalization group analysis of the small-world network model.” *Physics Letters A* 263, 341–346.

Putnam, Robert. 1993. *Making Democracy Work* Princeton University Press, Princeton

Putnam, Robert. 2000. *Bowling Alone*. New York: Simon & Schuster

Reagans, Ray E., Ezra Zuckerman. 2001. “Networks, Diversity, and Performance: The Social Capital of Corporate R&D Teams.” *Organization Science*.

Rivkin, Jan W. 2000. Imitation of complex strategies. *Management Science*. 46 824–844.

Rogers, Everett M. (2003). *Diffusion of Innovations* (5th ed.). New York: Free Press

Strang, David, and Michael Macy. 2001. “‘In Search of Excellence’: Fads, Success Stories, and Adaptive Emulation” *American Journal of Sociology*, 106.

Sunstein, Cass. 2003. *Why Societies Need Dissent*. Cambridge, MA: Harvard University Press.

Uzzi, Brian, and Jarrett Spiro. 2005. “Collaboration and Creativity: The Small World Problem.” Forthcoming American Journal of Sociology

Watts, Duncan. 2002. “A simple model of global cascades on random networks” Proceedings of the National Academy of Sciences USA 99, 5766-5771.

Watts, Duncan, and Steven Strogatts. 1998, Collective dynamics of 'small-world' networks, Nature 393:440-42.

Wenger, Etienne, Richard McDermott, and William M. Snyder. 2002. *Cultivating Communities of Practice*. Cambridge, MA: Harvard Business School.