Communication and Structured Correlation

Elliott Wagner

Abstract Philosophers and social scientists have recently turned to Lewis sender–receiver games to provide an account of how lexical terms can acquire meaning through an evolutionary process. However, the evolution of meaning is contingent on both the particular sender–receiver game played and the choice of evolutionary dynamic. In this paper I explore some differences between models that presume an infinitely large and randomly mixed population and models in which a finite number of agents communicate with their neighbors in a social network. My results show that communication with neighbors is more conducive to the evolution of meaning than communication with strangers. Additionally, I show that the behavior of the system is highly dependent on the topological structure of the social network. I argue that a specific class of networks—small world graphs—is especially conducive to the evolution of meaning. This is because small world graphs have a short characteristic path length while still maintaining a high degree of correlation between neighbors. Since many actual social networks, such as friendship networks and nervous systems, are conjectured to be small world structures, these results indicate that these networks are quite hospitable to the efficient evolution of meaning.

1 Introduction

Evolutionary game theory has become a useful tool for scholars attempting to explain the emergence and persistence of social phenomena such as cooperation,

E. Wagner (✉)
Department of Logic and Philosophy of Science, University of California, 3151 Social Science Plaza A, Irvine, CA 92697-5100, USA
e-mail: elliottw@uci.edu

fairness, and communication. It is common for this sort of model to presume an infinitely large, well-mixed population of agents interacting at random. In fact, this is an assumption of the replicator dynamic, which is one of the most frequently used evolutionary rules in the philosophical literature. Many other influential models of social behavior rely on a dynamical framework known as either Herrnstein or Roth and Erev Reinforcement Learning. Although reinforcement learning does not directly assume an infinite population, it has been proved that its long-run behavior approximates that of the replicator dynamic (Beggs 2005).

Unfortunately, these results do not always transfer to models using other plausible evolutionary rules. Indeed, it is sometimes the case that results obtained in large population models do not apply to arguably more realistic models in which agents are embedded in a network and only interact with their neighbors. For example, defection in the prisoner's dilemma always takes over the entire population under the replicator dynamic, but cooperation can persist when played on a network (Pollock 1989; Nowak and May 1992). Cooperation in the stag hunt is also more prevalent when the players are embedded in a network. Skyrms (2004) has found that over 99% of spatial stag hunts ended in stag hunting compared to far fewer with the replicator dynamic. Alexander and Skyrms (1999) have likewise shown that although fair division in the Nash bargaining game arises in a mere 60% of trials using the replicator dynamic, it is the virtually unique solution when the game is played with neighbors on a lattice. Similarly, Zollman (2005) has demonstrated that the behavior of agents playing a sender–receiver game with their neighbors on a lattice differs substantially from large population models of the same game. Under the replicator dynamic the entire population converges to the same strategy, but on a lattice many stable regions of agents adopt their own communication conventions.

In many ways this paper continues Zollman's (2005) analysis of the evolution of communication on a lattice. I expand on his work by investigating variations of the standard sender–receiver game that do not always converge to perfect communication in replicator dynamic models. As will be shown, interaction with neighbors in a social network is more conducive to the emergence of perfect communication than the replicator dynamic. I also consider a wide class of network topologies. It turns out that the behavior of the system is highly contingent upon the structure of the network. Even slight modifications to the connections in a two dimensional lattice can cause large changes in the population's collective behavior. In particular, if a small number of edges are rewired, then regions of agents playing certain undesirable strategies can persist on the network. Random rewiring of edges can also drastically affect the number of regions that form.

2 Sender–Receiver Games

Sender–receiver games were originally introduced by David Lewis in *Convention* (1969) to explain how linguistic terms can acquire meaning through nothing but the coordination of speakers and hearers. The framework for sender–receiver games is remarkably simple. The most austere variety consists of two players, two states,

called s_1 and s_2, two messages, called m_1 and m_2, and two actions, called a_1 and a_2. Before play begins, the state of the world is chosen at random by nature. The first player, called the sender, observes the state that nature chose and sends one of the two messages to the second player, called the receiver. The receiver, who is ignorant of the state, observes the message and then chooses an action. Each action is the correct response to exactly one of the states. If the receiver performs the correct action for the state of the world that obtained, then both players receive a payoff of 1. Otherwise, both players receive no payoff.

For this sort of game, a pure strategy for the sender is a function that maps each state to a message. Similarly, a strategy for the receiver is a function that maps each message to an action. In this two state, two message, two act game, there are six strategy profiles that constitute pure Nash equilibria. Two of these profiles are perfectly communicative. If the players adopt one of these two profiles, then they are guaranteed to receive a positive payoff regardless of the state that obtains. These two equilibria are called signaling systems. The other four Nash equilibria are less desirable. These babbling equilibria convey no information. A babbling equilibrium is one in which the sender always sends the same message regardless of the state and the receiver always performs the same action regardless of the message. Agents playing a babbling equilibrium perform no better than chance.

In *Convention*, Lewis implied that the messages in a sender–receiver game gain meaning when the players adopt a signaling system strategy. Therefore, if we can explain how the use of signaling system strategies can emerge and persist in an evolutionary setting, we can explain how messages can acquire meaning. To be sure, this framework is quite simple—perhaps too simple to apply to human communication. Nevertheless it is appropriate for modeling less complex communication systems, such as those employed by social bacteria, neurons, or non-human animals. Domestic chickens (*Gallus gallus domesticus*), for instance, identify aerial and terrestrial predators with two distinct alarm calls (Searcy and Nowicki 2005). There is a correct response to each type of predator. The optimal response to an aerial predator is to run for cover while crouching, but the optimal response to a terrestrial predator is to stand tall and scan the horizontal plane. This interaction between chicken senders and receivers mirrors the form of a Lewis sender–receiver game.

Much is known about the evolution of populations playing this sort of two state game in which the states are equiprobable. Skyrms (1996) has shown that under the replicator dynamic, which is meant to simulate a large, well-mixed population of players that interact in pairs chosen at random, the population always converges to a signaling system equilibrium. Huttegger (2007) has recently proved this result analytically. Additionally, Zollman (2005) has shown that agents playing this game on a grid lattice and updating their strategies by imitating their best neighbor will always converge to regional signaling systems. Each agent adopts a signaling system strategy, but the entire lattice does not necessarily adopt the same strategy. In Zollman's setup agents within each stable region perfectly communicate with each other, but they completely miscommunicate with agents inhabiting another region using a different signaling system.

However, it has only lately become known that such nice results do not necessarily hold for even slight modifications of this very simple game. Huttegger (2007) has proven that if the states are not equiprobable, then there is a chance that the population will converge to a non-communicative babbling equilibrium. Using computer simulations, I found approximately 41% of runs using the discrete-time replicator dynamic ended with the population playing a babbling equilibrium when $p(s_1) = .9$ and $p(s_2) = .1$.

Additionally, another sort of Nash equilibrium exists in sender–receiver games with n states, n messages, and n actions when $n > 2$. This type of equilibrium, called partial pooling, conveys some information about the state, but is not perfectly communicative like a signaling system. Huttegger (2007) proved that if $n > 2$, then a population evolving according to the replicator dynamics has a chance to converge to a partial pooling equilibrium. Using computer simulations of the discrete time replicator dynamic, it has been estimated that approximately 4.7% of initial populations converge to partial pooling (Huttegger et al. 2009).

3 Communication About More Than Two States

3.1 Simulation Results

Since network interaction aids the emergence of fair bargaining and cooperation, a natural question to ask is whether interaction with neighbors also aids the evolution of perfect communication under these unfavorable circumstances (non-equiprobable states and more than two states). The short answer is yes. Interaction with neighbors increases the likelihood of perfect communication. But the full story is rather subtle.

Following Alexander and Skyrms (1999) and Zollman (2005) the first model I consider will be 10,000 agents arranged on a 100 by 100 two dimensional lattice mapped to a torus. Each agent's neighbors are given by the Moore-8 neighborhood (an agent's neighbors are those vertices directly to her NW, N, NE, E, SE, S, SW, W), although the qualitative behavior of this system is identical to the case in which the agents only interact with their four von Neumann neighbors (N, E, S, W). Before play begins each agent is randomly assigned both a sending strategy and a receiving strategy. An iteration of the dynamic consists of two steps. First, each agent plays with each of her neighbors twice—once as sender and once as receiver. Summing the payoffs from these 16 encounters and then dividing this sum by the number of games played gives the agent's average payoff for the round. Each agent observes the average payoff of each of her neighbors and then adopts both the sending and receiving strategies of her most successful neighbor, so long as her most successful neighbor received a higher payoff than herself. Otherwise the agent does not switch strategies. Ties are broken by coin flips. This evolutionary rule is known as the imitate-the-best dynamic.

Consider the case in which there are more than two possible states of nature. Recall that with the replicator dynamic 4.7% of trials fell into partial pooling equilibria when there were three possible states. For the network interaction model

described above, all simulations (out of ten thousand total) converged to regional signaling systems. Inefficient partial pooling strategies were driven to extinction.

The impact of network interaction is even more pronounced as the number of states increases. When there are four states, approximately 24% of trials converge to a signaling system using the replicator dynamics. However, 200 out of 200 trials of 1,000,000 agents arranged on a 1,000 by 1,000 two dimensional lattice converged to regional signaling systems. For an n state sender–receiver game, there are n^{2n} possible strategy profiles, but only $n!$ of these profiles are signaling systems. So, in the random starting configuration of the network, each agent will be using a signaling strategy with probability $\frac{n!}{n^{2n}}$. Therefore, the number of agents on the network must grow exponentially in n for it to be even modestly likely that some agent will begin the simulation by playing a signaling system strategy. However, these simulation results indicate that for any n, the lattice can be made sufficiently large so that the system is virtually guaranteed to converge to regional signaling systems. Although, as is true of any theoretical system with exponential growth, such a lattice will be unrealistically large for even moderately sized n. For example, with $n = 10$, a 10^7 by 10^7 lattice is expected only to contain about 3.6 agents using signaling system strategies.

3.2 Analysis

All observed simulations of the three state, three message, three act sender–receiver game played on the Moore-8 toroidal lattice ended with partial pooling strategies driven to extinction. But is it ever possible for a region of agents playing partial pooling strategies to stably coexist with signaling systems on the network? Since an agent on the interior of a region can only imitate its own strategy, the key to a region's stability is interaction along its frontiers. A partial pooling region will be stable just in case each partial pooler on the frontier imitates an interior partial pooler instead of a frontier signaling system. As shown in Fig. 1, there are three relevant shapes that a frontier between regions can take. In order to investigate the potential stability of partial pooling it is necessary to consider the stability of each frontier structures in turn.

In the first case a frontier partial pooler is adjacent to three fellow partial poolers and five signalers. At least one of these partial poolers is on the interior of the region and interacts solely with other partial poolers to earn an average payoff of $\frac{2}{3}$. The frontier partial pooler is also adjacent to a frontier signaler that interacts with seven

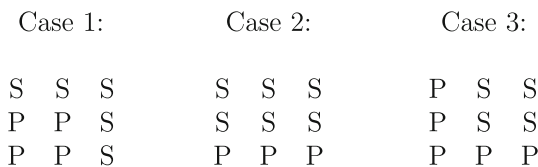| Case 1:			Case 2:			Case 3:		
S	S	S	S	S	S	P	S	S
P	P	S	S	S	S	P	S	S
P	P	S	P	P	P	P	P	P

Fig. 1 The three relevant shapes that the border between two regions can take. Agents using a signaling system strategy are denoted with the letter S. Agents using a partial pooling strategy are denoted with the letter P

other signaling systems and only a single partial pooler. Suppose this signaling system completely miscommunicates with the partial pooling strategy (Fig. 2). Then the frontier signaler earns nothing when interacting with a partial pooler, but earns the maximum payoff of one when either sending to or receiving from another signaler. This gives the frontier signaler an average payoff of $\frac{7}{8}$ which is greater than the $\frac{2}{3}$ that an interior partial pooler earns. Consequently the frontier partial pooler will imitate the frontier signaler and the region of partial pooling will not be stable.

In the second case a frontier partial pooler is adjacent to five fellow partial poolers and three signalers. The region is stable if an interior partial pooler earns a greater payoff than a frontier signaler. In this case each frontier signaler is adjacent to three partial poolers and five other signalers, so each frontier signaler earns an average payoff of $\frac{5}{8}$. This is less than the $\frac{2}{3}$ earned by an interior partial pooler. Therefore, the frontier partial pooler will imitate the interior partial pooler and the region will be stable.

In the third case a frontier partial pooler is adjacent to seven fellow partial poolers and a single frontier signaler. This frontier signaler performs quite poorly since she interacts with only three other signalers. She earns an average payoff of $\frac{3}{8}$ which is less than the $\frac{2}{3}$ earned by an interior partial pooler. Thus the frontier partial pooler will imitate an interior partial pooler and the corner between regions will be stable.

Notice, however, that the stability of the frontier in case two, and hence the persistence of partial pooling on the lattice, relies on total miscommunication between agents in the two regions. For example, suppose that some information is exchanged between the boundary agents. If an agent playing a signaling system strategy and an agent playing a partial pooling strategy communicate correctly about just one of the three states, then the frontier signaler's payoff becomes $\frac{6}{8}$. This payoff is greater than the $\frac{2}{3}$ that the interior partial pooler earns, so the partial pooler on the frontier will imitate the frontier signaling system. This imitation destabilizes the region of partial poolers. So regions of partial pooling strategies can persist on the network provided that they only share borders with the signaling system that they completely miscommunicate with.

This analysis shows that partial pooling can coexist stably with signaling systems on a network only if two rather unlikely conditions are fulfilled. No partial pooler can be adjacent to a signaling system with which it partially communicates. And the region of partial poolers cannot include a corner surrounded by signaling systems on five out of eight sides. The second condition can be satisfied if the region of partial poolers either occupies a slice of the toroidal lattice or entirely surrounds a rectangular region of signalers. Although a theoretic possibility, the likelihood of
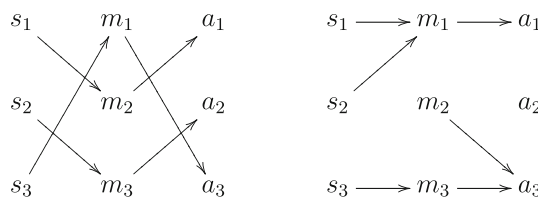

Fig. 2 A signaling system and a partial pooling equilibrium that exchange no information

these two conditions being simultaneously fulfilled for a region of partial poolers is quite small. This explains the fact that partial pooling equilibria are never seen in simulations executed on a sufficiently large lattice.

4 Communication When Nature is Biased

4.1 Simulation Results

The situation is more complicated for two-state sender–receiver games in which nature flips a biased coin to determine the state. As Fig. 3 shows, as $p(s_1)$ increases from .55 to .95, the number of discrete-time replicator dynamic simulations that don't converge to signaling systems grows in a uniform manner. On the other hand, all observed imitate-the-best simulations on a 100×100 Moore-8 toroidal lattice converged to regional signaling systems unless $p(s_1) \geq .9$. When $p(s_1) = .95$, 156 out of 300 simulations ended in a stable state in which a region of agents using a babbling strategy was surrounded by a region of agents using a signaling system. But, these babbling regions were always small. Of the 156 trials that ended with babblers, an average of 99.311% of agents were using signaling system strategies.

Similarly, when $p(s_1) = .9$, 123 of 300 simulations ended in a stable state in which not all agents had adopted signaling system strategies, but here the non-signaling system being played was not a babbling equilibrium. Instead it was a non-Nash equilibrium strategy of the form shown in Fig. 4. This half babbling strategy sends the same message regardless of the state, but discriminates between messages
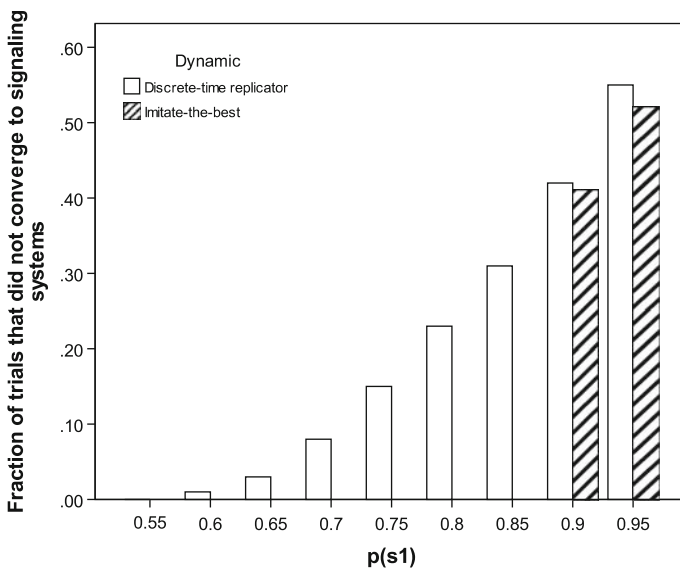

Fig. 3 Proportion of simulations that did not converge to signaling systems under the replicator and imitate-the-best dynamics. Imitate-the-best simulations were performed on a 100×100 Moore-8 toroidal lattice
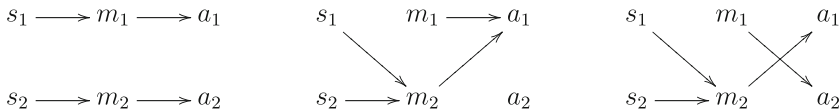
Fig. 4 A signaling system and the babbling and half babbling strategies that miscommunicate with it

when choosing an action. The regions of half babblers seen in simulations always perfectly miscommunicate when receiving messages from the surrounding signaling systems. That is, they earn absolutely no payoff when receiving a message from an adjacent signaling system. It is, perhaps, surprising to see the emergence of non-Nash play in the sender–receiver game. The replicator dynamics, for example, never converge to half babbling strategies because they are not Nash equilibria of the sender–receiver game. But it has been shown that the imitate-the-best dynamic can lead to cooperation in the prisoner's dilemma persisting on the Moore-8 lattice (Nowak and May 1992), so this is not the sole instance of imitate-the-best leading to non-Nash equilibrium behavior.

These simulation results suggest that the interaction environment of the Moore-8 toroidal lattice is more hospitable to the emergence of perfect communication than the well-mixed environment presumed by the replicator dynamic in two ways. First, unless the disparity between $p(s_1)$ and $p(s_2)$ is quite large, babbling equilibria and half babbling strategies are driven to extinction on the lattice. Second, even when non-signaling system strategies persist on the lattice, they are only seen in small regions. The lattice overwhelmingly adopts signaling systems, and only small regions maintain inefficient communication strategies.

4.2 Analysis

The fact that non-signaling system strategies are never observed in simulations on a lattice until the disparity between state probabilities becomes quite large indicates that there may be a particular threshold of $p(s_1)$ such that babbling and half babbling strategies are unstable below the threshold but become stable above. This is indeed the case. Just like in the analysis of the three state sender receiver game, interaction along the frontiers between regions determines the stability of the regions. The three relevant frontier structures to consider are the same as in Fig. 1, except here the role of partial pooling is played by a half babbling strategy.

Suppose that $p(s_1) = \alpha$ and $p(s_2) = 1 - \alpha$. Let the signaling system and half babbling strategies shown in Fig. 4 be the particular ones played. Then a frontier signaler will earn 1 when sending to or receiving from another signaler, 0 when sending to a half babbler, and $1 - \alpha$ when receiving from a half babbler. An interior half babbler interacts only with fellow half babblers and hence receives an average payoff of α. In the first case (as shown in Fig. 1), the frontier signaler interacts with seven signalers and one half babbler to earn an average payoff of $\frac{15-\alpha}{16}$. Consequently, the frontier half babbler will imitate the interior half babbler only if $\alpha > \frac{15}{17} \approx .8824$. This explains why stable rectangles of half babblers were seen in simulations when $p(s_1) = .9$.

In the second case, a frontier half babbler is adjacent to five fellow half babblers and three signaling systems. Each frontier signaler interacts with five other signaling systems and three half babblers to earn an average payoff of $\frac{13-3\alpha}{16}$. Therefore, the frontier half babbler will imitate an interior babbler instead of a frontier signaler just in case $\alpha > \frac{13}{19} \approx .6842$.

In the third case, a frontier half babbler is adjacent to seven fellow half babblers and a single signaler. This frontier signaler interacts with three signalers and five half babblers to earn an average payoff of $\frac{11-6\alpha}{16}$. So this frontier half babbler in the corner between regions will imitate an interior half babbler if $\alpha > \frac{1}{2}$. But this case is rather redundant because a region that includes this sort of corner and is large enough to contain an interior half babbler must also include a frontier in the shape of case two. So for a convex region of half babblers with a case two shaped corner to be stable, it must be the case that $\alpha > \frac{13}{19}$.

This analysis can be carried out in a similar fashion to determine the conditions under which babbling equilibria become stable. A frontier signaler earns $1 - \alpha$ when receiving from a babbler and earns α when sending to a babbler. It turns out that a case one corner of babblers is stable if $\alpha > \frac{15}{16}$. A case two edge is stable if $\alpha > \frac{13}{16}$. And a region of babblers with a case three corner is also stable if $\alpha > \frac{13}{16}$.

Now it is possible to completely characterize the conditions under which regions playing non-signaling system strategies can be stable on the Moore-8 lattice. When $p(s_1) > \frac{13}{19} \approx .6842$, regions of half babblers that do not include a case one style corner become stable. This sort of region is either an entire slice of the toroidal lattice or completely surrounds a rectangular region of signalers. No stable regions of this sort were ever seen in simulations. When $p(s_1) > \frac{13}{16} = .8125$, regions of the same shape populated by babblers becomes stable. This type of region was also never seen in simulations. When $p(s_1) > \frac{15}{17} \approx .8824$, regions of half babblers that include a case one corner become stable. Regions of this sort were often seen in simulations when $p(s_1) = .9$. When $p(s_1) > \frac{15}{16} = .9375$, the same type of region populated by babblers becomes stable. This sort of region was often seen in simulations when $p(s_2) = .95$.

It is worth noting that this sort of analysis can be extended beyond models of agents arranged on a lattice playing with their Moore-8 neighbors. It can be applied to any sort of graph structure in which each agent has the same number of neighbors. For example, on a lattice with the von Neumann neighborhood (4 neighbors), babbling becomes stable when $p(s_1) > \frac{7}{8}$. But, as the following pair of theorems illustrates, not all graph structures provide such a favorable environment for the emergence of perfect communication.

Theorem 1 *If agents that update their strategies according to the imitate-the-best dynamic play a sender–receiver game with their neighbors on a regular graph of degree k, then regions of agents using babbling equilibria can be stable only if either $p(s_1) > \frac{k+1}{2k}$ or $p(s_2) > \frac{k+1}{2k}$.*

Proof Without loss of generality, suppose $p(s_1) > p(s_2)$. Let $p(s_1) = \alpha$ and $p(s_2) = 1 - \alpha$. For a region of babblers to be resistant to invasion by signaling systems, each frontier babbler must imitate an interior babbler instead of a frontier signaler. This requires that interior babblers perform better than each frontier

signaler. An interior babbler earns an average payoff of α in each round. A frontier signaler performs worst when adjacent to $k - 1$ babblers and a single other signaler. In this case the frontier signaler earns an average payoff of α when sending to a babbler, $1 - \alpha$ when receiving from a babbler, and 1 when either sending or receiving from a signaler. This gives an average payoff of $\frac{\alpha(k-1)+(1-\alpha)(k-1)+2}{2k} = \frac{k+1}{2k}$. Therefore regions of babbling can stably coexist with regions of signaling only if either $p(s_1) > \frac{k+1}{2k}$ or $p(s_2) > \frac{k+1}{2k}$.   □

This theorem shows that there are some regular graph structures on which a region of babblers can be stable even if nature is only slightly biased. Furthermore, as the degree of the graph increases, and hence the number of interactions each agent participates in increases, the disparity between $p(s_1)$ and $p(s_2)$ required for the stability of babbling decreases. But the good news here for the emergence perfect communication is that, regardless of the network structure, babbling regions are never stable when the states are equiprobable. However, as the next theorem shows, this pleasant result does not hold for regions of half babblers.

Theorem 2 *When the states of nature are equiprobable, regions of half babbling agents can be stable on some regular graphs of degree k when $k > 3$.*

Proof Suppose the states of nature are equiprobable so that $p(s_1) = p(s_2) = .5$. Then an interior half babbler will communicate only with other half babblers to earn an average payoff of .5. A frontier signaler earns 0 with sending to a half babbler, .5 when receiving from a half babbler, and 1 when either sending to or receiving from a signaler. In the worst case for the frontier signaler, she may be adjacent to $k - 1$ half babblers and only a single other signaler. In this case she earns an average payoff of $\frac{2+.5(k-1)}{2k}$. The region of half babblers will be stable only if each frontier half babbler imitates the interior babbler instead of a frontier signaler. This will occur if $.5 > \frac{2+.5(k-1)}{2k}$. Therefore the region will be stable if $k > 3$.   □

So not only can some regular graphs sustain babbling agents when nature is only slightly biased, but some regular graphs, such as the one shown in Fig. 5, can sustain half babbling agents even when the states are equiprobable. Graph structures that are extremely inhospitable towards signalers, like the one in Fig. 5 or the ones invoked in the proofs of the above theorems, all have the peculiar property that agents on the frontier of a region of signalers are perfectly anti-correlated. That is, agents on the frontier of the signaling region never interact with each other. Instead of interacting with each other, these agents are fated to interact with many uncommunicative babblers or half babblers. In contrast to this situation, agents on the frontiers of regions on the Moore-8 lattice often interact with each other. And, not so coincidentally, the Moore-8 lattice provides a favorable environment for the emergence of perfect communication.

The observation that correlation of encounters between agents along the frontiers aids the emergence of perfect communication can be seen as an instance of the more general moral from game theoretic analysis that positive correlation of encounters between agents playing the same strategy increases the likelihood of cooperation and coordination. This lesson was perhaps first pointed out by Hamilton (1971) who observed that cooperation can emerge and persist in a population playing a
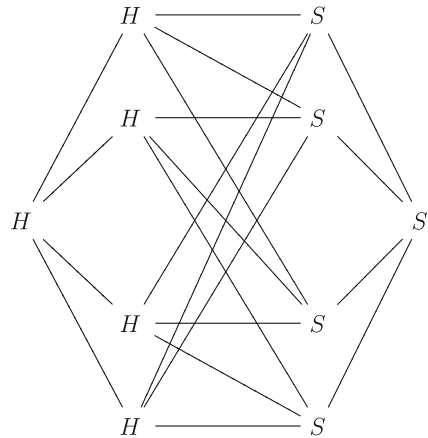
prisoner's dilemma provided that encounters are non-random in such a way as to make individuals more likely to pair with others who use the same strategy. In the prisoner's dilemma this non-random assortment of strategy types causes the expected fitness of cooperators to increase and the fitness of defectors to decrease. Although the game being played here is not the prisoner's dilemma, the graph structure imposed by agents interacting with their Moore-8 lattice neighbors generates just this sort of positive correlation along the frontiers between regions. The Moore-8 lattice structure guarantees that a signaler on the boundary between a region of signalers and a region of babblers plays at least six of her sixteen games with other signalers.

This correlation of strategy types along the frontiers is caused by correlation between neighbors. A signaler on the frontier must be adjacent to an interior signaler. Each agent adjacent to this interior signaler must also be a signaler. Consequently, since on the Moore-8 lattice the frontier signaler and the interior signaler share at least four of the same neighbors, the frontier signaler will often be found interacting with other signalers. On the Moore-8 lattice the neighbors of any particular agent are often neighbors of each other, and this correlation between neighbors causes a corresponding correlation of strategy types along the frontiers between regions.

On the other hand, in an arbitrary graph a signaler on the frontier between a group of signalers and a group of babblers may be adjacent to only a single other signaler. If this is the case, then the frontier signaler will always imitate the interior signaler despite the fact that the frontier signaler will herself perform poorly because she may be adjacent to many uncommunicative babblers. And because the frontier signaler performs so poorly, her uncommunicative neighbors will have less incentive to imitate her perfectly communicative signaling system strategy.

Lack of correlation between neighbors in a social network means that there is less positive correlation between agents playing signaling systems on the frontiers. This lack of correlation between frontier signalers depresses the value of $p(s_1)$ for which babbling and half babbling can stabilize. Graph structures that impose a high degree

of correlation between neighbors will therefore be more conducive to the emergence of signaling systems than graph structures that only weakly correlate neighbors.

5 Communication in Small Worlds

Correlation between neighbors is the feature of network structure that is responsible for destabilizing regions of inefficient communication. However, different graph topologies provide different amounts of correlation. A natural next step in this investigation is to systematically analyze how slight alterations of the graph topology, and hence small changes to the correlation between neighbors, impacts the emergence and stability of signaling systems. But before this can be done it is necessary to chose both a method of measuring correlation between neighbors and a method of generating graphs that exhibit a controlled amount of correlation.

A graph's clustering coefficient is one way to measure correlation between neighbors in a network. The clustering coefficient γ_v of a vertex v in a graph is the fraction of possible edges in the neighborhood of v that actually occur in the neighborhood of v. More precisely, $\gamma_v = |E(\Gamma_v)|/\binom{k_v}{2}$ where $|E(\Gamma_v)|$ is the number of edges in the neighborhood of v and k_v is the number of vertices adjacent to v. So γ_v measures the probability that vertices adjacent to v are also adjacent to each other. The clustering coefficient γ of an entire graph is γ_v averaged over all vertices in the graph.

Another useful metric for studying graph structures is characteristic path length, denoted L. Following Watts (1999) I'll define L as the median of the means of the shortest path lengths connecting each vertex in the graph to all other vertices. Finding L for a large graph is quite computationally intensive. So, again following Watts (1999) in this paper I'll use a technique due to Huber (1996) that chooses an appropriately sized random sample of vertices to compute L within an acceptable margin of error.

Both these metrics have proven useful to mathematicians and sociologists investigating so-called small-world graphs. If a graph represents a social network in which edges connect acquaintances, then the clustering coefficient measures the degree to which a person's acquaintances are acquainted with each other. And in this interpretation the characteristic path length of a graph approximates the average number of acquaintances that separate any two individuals of the social network. Small-world graphs are those graphs that have a large clustering coefficient and a short characteristic path length. It is conjectured that many real-world interaction networks are small-world graphs. Popular examples include the Internet Movie Database (the Kevin Bacon Game) and co-authorship networks (Erdös numbers). Also Watts and Strogatz (1998) have shown that the neural connections of the worm *C. elegans* exhibit small-world properties, and it is believed that the same goes for many other neural systems.

To study small-world phenomena, Watts introduced a parameterized family of graphs called β-graphs. A β-graph is generated by starting with a ring in which each vertex is connected to its k nearest neighbors. Then, for every vertex v, each of the $\frac{k}{2}$ leftmost edges is rewired to a random vertex u with probability β. The vertex u is

chosen uniformly at random from the whole graph omitting redundant connections and self-loops. For low values of β, β-graphs are highly structured, but as β increases, the resulting graphs become more and more random. Watts studied β-graphs because there is a range of β values for which the resulting graphs are highly clustered and yet have short characteristic path lengths. That is, some β-graphs exhibit small-world properties, as shown by Fig. 6.

Since β-graphs transition from highly structured (much correlation between neighbors) to completely random (very little correlation between neighbors) they provide an excellent way to systematically study how the amount of correlation between neighbors impacts the emergence of communication in sender–receiver games. On a ring in which each vertex is connected to its 8 nearest neighbors, half babbling becomes stable when $p(s_1) > \frac{15}{17}$. Thus, if a sender–receiver game with $p(s_1) = .88$ and $p(s_2) = .12$ is played by 10,000 agents on a β-graph with $k = 8$, all agents will adopt signaling system strategies when $\beta = 0$. As β increases, however, the graph becomes less structured. Therefore, as the analysis in the last section suggests, half babbling should become stable and not all simulations will converge to perfect communication. Furthermore, if it is truly correlation between neighbors that is responsible for destabilizing non-signaling system regions, then the proportion of agents on the network playing signaling systems should follow the clustering coefficient of the graph as β ranges from 0 to 1. This is indeed what is seen in simulations. As Fig. 7 shows, the proportion of agents playing signaling systems closely follows γ as β increases.

One might wonder if something about this result is particular to starting with a ring in which every vertex is connected to its k nearest neighbors. After all, I began this paper by pointing out that all too often evolutionary game theory results are dependent on a specific interaction dynamic or on a specific graph topology. For high β, β-graphs are largely random, but for low β, the properties of β-graphs are

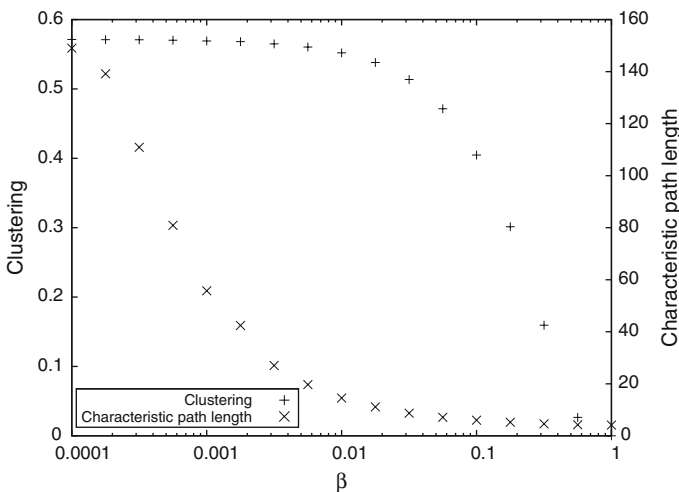**Fig. 6** L and γ vs. β for β-graphs with 3,000 vertices and average degree 8. Each data point is an average of 20 realizations of the β-graph construction algorithm
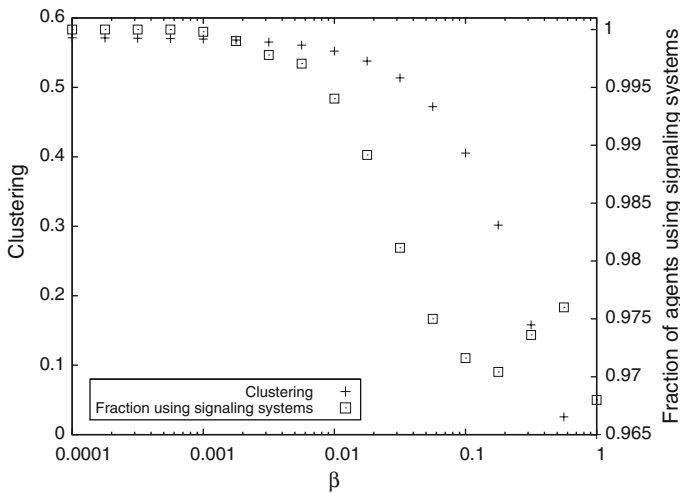
Fig. 7 γ and fraction of agents using signaling system strategies vs. β for β-graphs with 10,000 vertices and average degree 8. $p(s_1) = .88$ and $p(s_2) = .12$. Each point is an average of 20 realizations of the β-graph construction algorithm

dominated by the properties of the ring. So, one might also consider performing the β-graph random rewiring process on the most frequently looked at topology: a two dimensional toroidal lattice with edges given by the Moore-8 neighborhood. Remember that non-signaling system strategies were only seen on the Moore-8 lattice when $p(s_1) \geq .9$. Hence, for sender–receiver games with $p(s_1) = .88$ and $p(s_2) = .12$, when $\beta = 0$, the entire network should adopt signaling system strategies. As β increases, however, regions of babbling and half babbling should become stable. Just like on the ring, this is actually what is seen in simulations (Fig. 8). As β increases, the proportion of agents playing perfectly communicative signaling systems very closely follows the clustering coefficient of the graph. These results strongly indicate that it is indeed correlation between neighbors that aids the emergence of communication when the states are not equiprobable.

As mentioned earlier, one of the most surprising results from Zollman (2005) is that when sender–receiver games are played on a two dimensional toroidal lattice, multiple stable regions of communication emerge. Within each region all agents play the same signaling system. But, at least in the two state sender–receiver game, agents from a region playing one signaling system utterly miscommunicate with agents from another region playing a different signaling system. So there is complete communication within regions, but no communication between regions. In general, this result holds as network topology is altered. However, different structures with the same number of vertices can lead to remarkably different numbers of regions evolving. For example, out of 1,000 simulations on a Moore-8 toroidal lattice with 10,000 agents playing the two state sender–receiver game with equiprobable states, an average of 9 stable regions existed when the evolution stabilized. But, on a 10,000 node cycle, an average of 462 stable regions were observed. And on a 10,000 node wheel, the entire networked evolved to play the
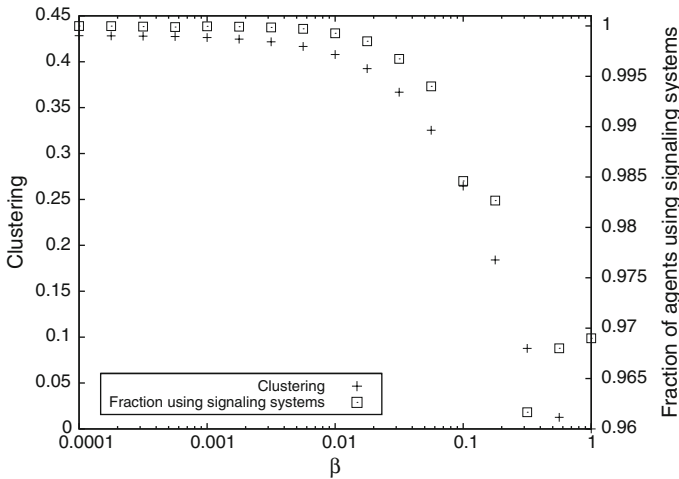
Fig. 8 γ and fraction of agents using signaling system strategies vs. β for Moore-8 toroidal lattices with edges rewired using the β-graph randomization process. Each graph has 10,000 vertices. $p(s_1) = .88$ and $p(s_2) = .12$. Each point is an average of 20 realizations of the β-graph rewiring algorithm

same signaling system on each of 1,000 out of 1,000 simulations. These results clearly show that in addition to the network topology influencing the likelihood of arriving at perfect communication, it can also drastically affect the number of regions that exist when the evolution stabilizes.

It turns out that the characteristic path length of the graph controls the number of stable regions that emerge. As β increases, the number of stable regions closely follows L (Fig. 9). Remember that the characteristic path length is a measure of the average distance between agents on the graph. So, if L is short, like in the wheel or the Moore-8 toroidal lattice, then the first signaling system to establish itself will rapidly spread and take over the entire network. However, if L is long, like in the ring, it can take many steps of the imitate-the-best dynamic for the signaling system strategy to spread. During this time, other clusters of different signaling systems are likely to form.

So, the clustering coefficient of the graph influences the likelihood that agents will play signaling system strategies when the evolution stabilizes. And the characteristic path length of the graph influences the number of stable regions of agents all playing the same strategy. An interesting ramification of these two observations is that there exists a class of graphs in which very few regions emerge and that these regions are all perfectly communicative—even when the states are not equiprobable. This is the class of graphs that are highly clustered and have short characteristic path lengths. Namely, the class of small-world graphs. It turns out that small-worlds, like our own social networks and the organization of neural systems, are very conducive to highly efficient communication. Networks with longer characteristic path lengths have more regions, and hence more miscommunication between agents on the borders. Networks with less clustering are more prone to being inhabited by uncommunicative babblers when evolution stabilizes. It is only
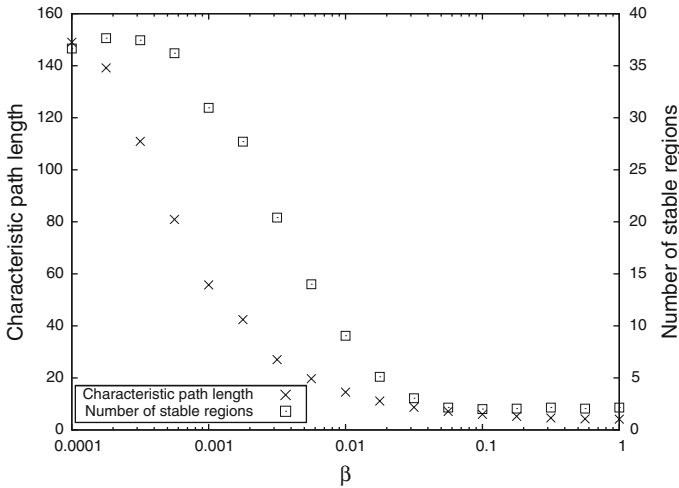
Fig. 9 L and number of stable regions observed when the evolutionary dynamic stabilized vs. β for β-graphs with 3,000 vertices of average degree 8. $p(s_1) = p(s_2) = .5$. Each point is an average of 20 realizations of the β-graph construction algorithm

in the small-world graphs that both babblers are eliminated and that the number of regions is minimized.

6 Conclusion

Throughout this paper I have investigated the behavior of agents playing a simple communication game with their neighbors in a social network. Not only do the results in this sort of model differ greatly from those obtained through models that presume an infinitely large, well-mixed population, but the results can also differ greatly depending on the particular network topology that the agents inhabit. Furthermore, slight changes in structure can cause large changes in collective behavior. Witness the number of stable signaling system regions on β-graphs, for example (Fig. 9). With a .2% chance of randomly rewiring each edge, an average of 22.1 regions formed. By increasing the chance of rewiring to only 3%, the average number of regions decreased to 2.55.

I have also identified one type of correlation present in models of networked interaction that is not present in random interaction models. Namely, correlation between neighbors. It is this type of correlation that is responsible for destabilizing otherwise stable regions of inefficient communication. This is the form of correlation that makes interaction through social networks a more hospitable environment for the emergence of perfect communication in sender–receiver games when nature is biased than, say, the well-mixed environment of a test tube. Due to high correlation between neighbors and short characteristic path length, one family of graphs—small worlds—is especially hospitable to the emergence of efficient

signaling. Fortunately for aspiring communicators, many real-life social interactions are conjectured to take place through small world network structures.

References

Alexander, J. M., & Skyrms, B. (1999). Bargaining with neighbors: Is justice contagious? *Journal of Philosophy, 96*, 588–598.

Beggs, A. (2005). On the convergence of reinforcement learning. *Journal of Economic Theory, 122*, 1–36.

Hamilton, W. D. (1971). Selection of selfish and altruistic behavior in some extreme models. In J. Eisenberg & W. Dillon (Eds.), *Man and beast: Comparative social behavior* (pp. 57–91). Washington, DC: Smithsonian Press.

Huber, M. (1996). *Estimating the average shortest path length in a graph*. Technical report, Cornell University.

Huttegger, S. (2007). Evolution and the explanation of meaning. *Philosophy of Science, 74*, 1–27.

Huttegger, S., Skyrms, B., Smead, R., & Zollman, K. (2009). Evolutionary dynamics of lewis signaling games: Signaling systems vs. partial pooling. *Synthese*. doi:10.1007/s11229-009-9477-0.

Lewis, D. (1969). *Convention: A philosophical study*. Oxford: Blackwell.

Nowak, M. A., & May, R. M. (1992). Evolutionary games and spatial chaos. *Nature, 359*, 826–829.

Pollock, G. B. (1989). Evolutionary stability on a viscous lattice. *Social Networks, 11*, 175–212.

Searcy, W. A., & Nowicki, S. (2005). *The evolution of animal communication: Reliability and deception in signaling systems*. Princeton: Princeton University Press.

Skyrms, B. (1996). *Evolution of the social contract*. New York: Cambridge University Press.

Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. New York: Cambridge University Press.

Watts, D. J. (1999). *Small worlds: The dynamics of networks between order and randomness*. Princeton: Princeton University Press.

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

Zollman, K. J. (2005). Talking to neighbors: The evolution of regional meaning. *Philosophy of Science, 72*, 69–85.