Emergence of a Signaling Network with "*Probe and Adjust*"

Simon M. Huttegger
University of California, Irvine

Brian Skyrms[1]
University of California, Irvine
Stanford University

Introduction: Once individuals have developed a system of signals, in one way or another, they may spontaneously assemble into a signaling network. The network structure achieved may depend both on the payoffs involved and on the kind of adaptive dynamics driving the evolution of the network. We focus here on one example due to Bala and Goyal (2000), in which a ring network has strong distinguishing properties. The ring structure is both optimal for all involved, and the unique structure of strict Nash equilibria in the associated network formation game.

Bala and Goyal prove that a myopically rational dynamics of *Best Response with Inertia* leads to this ring structure (in a sense to be made precise later.) This dynamics requires individuals who best respond to know the whole network configuration in the previous round, as well as the payoff structure of the game. These requirements for the application of best response may not be too onerous in small groups but – at least in certain cases – may presume too much in larger settings. Even in small groups in the

[1] Skyrms research partially supported by AFOSR grant #FA9550-08-1-0389.

economics laboratory, where the experimenter makes sure that this information is available, *Best Response with Inertia* does not seem to fit the empirical data well (Berninghaus et al.). Nevertheless, individuals do tend to learn the ring structure. It is therefore interesting for two reasons to explore alternative adaptive dynamics – to see whether dynamics with a lower informational requirement can learn the ring, and to generate candidates for empirical test.

Here we address the first question. Is there a plausible low-information adaptive dynamics that learns the ring? We focus on a low-rationality, payoff-based, dynamics of *Probe and Adjust*, introduced in Skyrms 2010. This dynamics only requires individuals to know their own actions and payoffs, and to remember their actions and payoffs from the last round. Nevertheless, individuals using this dynamics also learn the ring configuration. These results have broader application, both to other network structures and to broader classes of games.

The Bala-Goyal Ring Game: Individuals get private information by observing the world. Each gets a different piece of information. Information is valuable. An individual can pay to connect to another and get her information. The individual who pays does not give any information; it only goes from payee to payer. The payer gets not only the information from private observations of those whom she pays, but also that which they have gotten from subscribing to others for their information. Information flows freely and without degradation along the links so established. It flows in one direction, from payee to payer. We assume that information flow is fast, relative to any adjustment of the network structure.

More precisely, each of a finite number of individuals can make links to any set of other individuals. Making a connection has a cost, c. Each individual has information with a value, v. If individual 1 has a link to another individual 2, individual 1 gets all the information that 2 has including all information gotten from links with other players. That is, let *x be connected to y* be the ancestral of *x has a link to y* together with the stipulation that x is connected to x. Then a player has all her own original information and all the information of players to whom she is connected.

If the cost of subscribing to someone's information is too high, then it won't pay for anyone to do it. But let's suppose that that the cost of establishing a connection is less than the value of each piece of information, $c<v$. Then connections certainly make sense. We assume that any individual can make as many connections as she wishes. This model can be viewed as a game, with an individual's strategy being a decision of what connections to make. It could be none, all, or some. The game has multiple equilibria, but one is special. This is the ring. (a.k.a. the circle, the wheel) The ring structure in this game is special in two ways. The first is that it is *strict,* the second that it is *efficient.* It is a strict equilibrium in that someone who unilaterally deviates from such a structure finds herself worse off. It is *Pareto efficient* in that there is no way to change it to make

someone better off without making someone worse off. In fact, there is no way at all to make anyone better off. Everyone has the highest possible payoff that they could get in any network structure. The key to both these properties is that information flows freely around the ring, so that for the price of one connection a player gets all the information that there is.

Consider a player in such a ring who changes her strategy. She could establish additional links, in which case she pays more and gets no more information. She could break her link, in which case she would forego the cost but get no information. She could break the link and establish one or more new ones, but every way to do that would deliver less than total information. Every deviation leaves her worse off. That is to say that the ring is a *strict* Nash equilibrium of the game.Now suppose that, starting from the ring, there is some lucky guy that everyone else would like to make better off. There is nothing they can do! He is already getting all the information at the cost of one link. They can't alter their links so as to give him more information, since he is already getting it all. Only he can avoid the cost of the link by breaking it – that is, not visiting anyone – but then he gets no information at all. The ring is *efficient.*

These rather special properties of the ring depend on rather special properties of the model – one way flow of information, flow without informational decay or degradation, subscriber pays all the costs. Various modifications of these assumptions lead to a rich array of situations with different properties. For the moment, however, we will focus on the ring.

Best Response with Inertia: Bala and Goyal propose a myopically rational adaptive dynamics that learns the ring. Most of the time an agent just keeps the network connections that she has the last time – this is *inertia*. But with some small probability, e, she wakes up and chooses a set of connections that is optimal against the network configuration of her associates the previous time – this is *best response*. The best response probabilities are independent between trials and between players. (Thus for small e, simultaneous or subsequent best responses are highly unlikely.) If there are ties for best response, the individual chooses between the tied set of connections at random.

Since, for small e, simultaneous or subsequent best responses are highly unlikely, Bala and Goyal analyze a simpler process in which nature, with small probability, chooses one player to best respond. First, since the ring is a strict Nash equilibrium configuration, no best response will lead a player to depart from a ring. By definition, deviation by a single player would leave him worse off. Next, since the ring is the unique strict Nash equilibrium, it is the *only* network configuration with this property. Players will exit any other configuration with positive probability. This is true even of Nash equilibria that are not strict, because tie-breaking for best response can lead to an exit.

The whole process, consisting of all players using best response with inertia, is thus a Markov chain with ring configurations as the unique absorbing states. Next, Bala and Goyal show that there is a best response path from any state of the system to a ring. That is, there is a sequence of network configurations starting from the original state and ending in a ring, such that each change is the result of a best response by a player. It follows from the definition of best response with inertia that every such path has a positive probability. In this simplified theory, it then follows from standard Markov chain theory that from any starting point, players will reach a ring structure with probability one.

Probe and Adjust: We now only assume that players know their payoffs and can remember them for one period. They need not know the previous configuration of the network or the structure of the game. We retain *inertia*. Most of the time players just keep doing the same thing. But with some small probability, e, a player *probes.* A probe consists of choosing some set of connections at random and trying them out. If a player probes, she compares the payoff obtained on the probe with the payoff obtained on the previous move. If it is higher, she sticks with the connections tried on the probe. If it is lower, she goes back to the previous connections. If there is a tie, she chooses between the alternatives with equal probability. Probe probabilities are independent, just like best-response probabilities in the previous dynamics.

Analysis of Probe and Adjust: If the probe probability is small, it is very unlikely to have simultaneous or subsequent probes. So (as before) we start by analyzing a simpler process. With small probability, nature chooses a player at random to probe and adjust. That player probes and then adjusts according to the results of the probe, while all other individuals keep doing the same thing. Since nothing happens most of the time, we can just look at the cases where there are probes. So we consider this even simpler *process S*:

> *Nature chooses a player at random to probe. This player probes in one round and reacts in another. Repeat.*

This process contains an embedded Markov chain, consisting of the even times. Probes are omitted but reactions to them are included. As before, rings are absorbing states because they are strict Nash equilibria, and the only absorbing states because they are the unique strict Nash equilibria. Probes may lead away from non-strict Nash configurations, but they cannot lead away from strict Nash configurations.

That there is a positive probability of transition from any state to a ring is a consequence of Bala and Goyal's proof that there is a best response from any state to a ring, together with the fact that if there is a best response a probe will -- with positive probability – lead to its being taken. The embedded Markov chain leads to a ring with probability one, just as before. That means that process S spends most of its time in a ring configuration, except for occasional unproductive probes that can discover nothing better.

In the original Probe and Adjust dynamics, simultaneous or subsequent probes, although unlikely, are not impossible. We now remove our simplifying assumptions, and address the effect of such unlikely perturbations on the stability of the ring. Suppose that from a ring, many players probe simultaneously. Unless all probe and hit another ring structure, probes will lead to worse payoffs, and all will go back to the same ring. These simultaneous probes don't change anything. But suppose from a ring, player 1 probes thus lowering the payoff of player 2, and immediately player 2 probes and gets a higher payoff than in the previous round (although lower than her payoff in the ring). This can lead away from the ring.

Strings of subsequent probes can, with small probability, lead away from the ring. The question for the real Probe and Adjust dynamics is whether getting away from the ring to another state in this way is more or less probable than getting back from that state to the ring. Consider the scenario for departure from the ring just discussed. At t1 player 1 probes, gets a worse payoff and causes a worse payoff for player 2. At t2 player 1 goes back to her original connect, and player 2 probes and gets a better payoff. At t3 player 2 adopts her probe strategy. This is a scenario of probability order of e^2. To get back to the ring, it suffices that player 2 then probes her original strategy. This has a probability of order e. For small e, the probability of getting from the alternative state to the ring is much greater than the transition in the opposite direction. Basically the same argument works in general: At t0 we have a ring. The player (or players) that probe at t1 return to their original strategies at t2. Then we suppose that there are subsequent probes, either sequential or simultaneous or some combination of them - it doesn't matter. To get m players to switch from their original circle strategies then takes *at least* m+1 probes. (order e^(m+1) probability) To get back only requires the m players to simultaneously probe their original strategies. (order e^m probability). We can conclude that (i) Probe and Adjust dynamics learns the ring and (ii) for infrequent probes, it then spends most of its time in a ring configuration.

Probe and Adjust in the Short Run: In order to get a sense of how long it takes Probe and Adjust to learn the ring in the Bala-Goyal game, we conducted some numerical simulations with three players. The results of these simulations are shown in Figure 1. In these simulations, the cost for establishing a link was set to 0.6. The cost parameter could be varied and still lead to the same simulations as long as establishing links does not get too costly or too cheap ($0<c<1$), for adjustment exclusively rests on the ordinal relationships of the payoffs. When faced with a tie, players are assumed to choose equivalent strategies with equal probabilities.

Figure 1 shows two simulations, one with a probe rate of 0.1 and the other with a probe rate of 0.01. There were 10,000 trials each. The data points show the first time of hitting the ring configuration. If each player probes every tenth round on average, convergence to the ring appears to be very rapid. About 87% of all trial hit the ring before round 500. By round 1,000, almost all trials did hit the ring at least once. If each player only probes on average once in hundred rounds, convergence appears to be somewhat slower, although by 10,000 rounds all trials have hit the circle. Looking at individual trials, once players choose according to the ring configuration, switching stably to a different configuration requires at least two players to probe consecutively and is a rare event. Even if it happens, the players typically switch back to the ring fairly rapidly.

Related Literature: On the *best response with inertia* dynamics for network formation see also Watts (2001), Jackson and Watts (2002), Goyal (2007), Jackson (2008). An alternative, low rationality model of network formation based on *reinforcement learning* was introduced in Skyrms and Pemantle (2000). See also Pemantle and Skyrms (2004a,b), Skyrms and Pemantle (2010). Extensive simulations show that this dynamics does not learn the ring in the game discussed in this paper. That slightly modified *fictitious play* dynamics reliably learns the ring is shown in Huttegger and Skyrms (2008). There we ask whether there is a low-rationality, low-knowledge, payoff-based dynamic that learns the ring – a question that is given an affirmative answer here. The *probe and adjust* dynamics discussed here is related to a class of payoff-based learning dynamics studied in Marden et. al. (2009) and Young (2009). It is somewhat simpler than these in that it only requires agents to remember the past round.

There is a growing experimental literature on network formation. See Berninghaus et. al. 2007, Callander and Plott 2005, Falk and Kosfeld 2003. Experiments find that subjects do learn the ring in the game discussed here. They also tend to show that subjects are not using rules like *best response with inertia* or *fictitious play*. Rather there is a noisier learning process in which individual at a strict equilibrium may temporarily leave it and then come back (Berninghaus et. al.). It would be interesting to see which learning rules best fit the empirical data.

Alternative network games with different payoff structures support star network configurations as strict equilibria. See Bala and Goyal(2000), Berninghaus et. al. (2007), Hojman and Szeidl (2008). *Probe and Adjust* dynamics also has application in these cases, as well as well as in more general classes of games. We plan to explore this elsewhere.

References:

Bala, V. and Goyal, S. “A Noncooperative Model of Network Formation.” *Econometrica* 68: 1181-1231, 2000.

Berninghaus, S., K.-M. Ehrhart, M. Ott and B. Vogt. “Evolution of Networks – an Experimental Analysis.” *Journal of Evolutionary Economics* 17: 317-347, 2007.

Callander, S. and C. R. Plott. “Principles of Network Development and Evolution: An Experimental Study.” *Journal of Public Economics* 89: 1469-1495, 2005.

Falk, A. and M. Kosfeld. “It’s All About Connections: Evidence on Network Formation.” IEW Working Paper 146. University of Zurich, 2003.

Goyal, S. *Connections: An Introduction to the Economics of Networks.* Princeton: Princeton University Press, 2007.

Hojman, D. A. and A. Szeidl. “Core and Periphery in Networks.” *Journal of Economic Theory.* 139: 295-309, 2008.

Huttegger, S. and B. Skyrms. “Emergence of Information Transfer by Inductive Learning.” *Studia Logica* 89: 237-256, 2008.

Jackson, M. *Social and Economic Networks.* Princeton: Princeton University Press, 2008.

Jackson, M. and A. Watts. “On the Formation of Interaction Networks in Social Coordination Games.” *Games and Economic Behavior* 41: 265-291, 2002.

Kosfeld, M. “Economic Networks in the Laboratory.” *Review of Network Economics* 3: 20-41, 2004.

Marden, J. P., H. P. Young, G. Arslan and J. S. Shamma. “Payoff-based dynamics for Multiplayer Weakly Acyclic Games.” *SIAM Journal on Control and Optimization* 48: 373-396, 2009.

Pemantle, R. and B. Skyrms. “Network Formation by Reinforcement Learning: The Long and the Medium Run.” *Mathematical Social Sciences* 48: 315-327, 2004a.

Pemantle, R. and B. Skyrms. “Time to Absorption in Discounted Reinforcement Models” *Stochastic Processes and Their Applications* 109: 1-12, 2004b.

Skyrms, B. *Signals: Evolution, Learning and Information.* Oxford: Oxford University Press, 2010.

Skyrms, B. and R. Pemantle. “A Dynamic Model of Social Network Formation.” *Proceedings of the National Academy of Sciences of the USA.* 97: 9340-9346, 2000.(Reprinted with a postscript in In *Adaptive Networks* Ed. Thilo Gross and Hiroki Sayama. Springer Verlag 2009, 231-251.)

Skyrms, B. and R. Pemantle "Learning to Network" In *Probability in Science*. Ed. E. Eells and J. Fetzer. Berlin: Springer, 2010.

Watts, A. "A Dynamic Model of Network Formation." *Games and Economic Behavior* 34: 331-341, 2001.

Young, H. P. "Learning by Trial and Error." *Games and Economic Behavior* 65: 626-643, 2009.

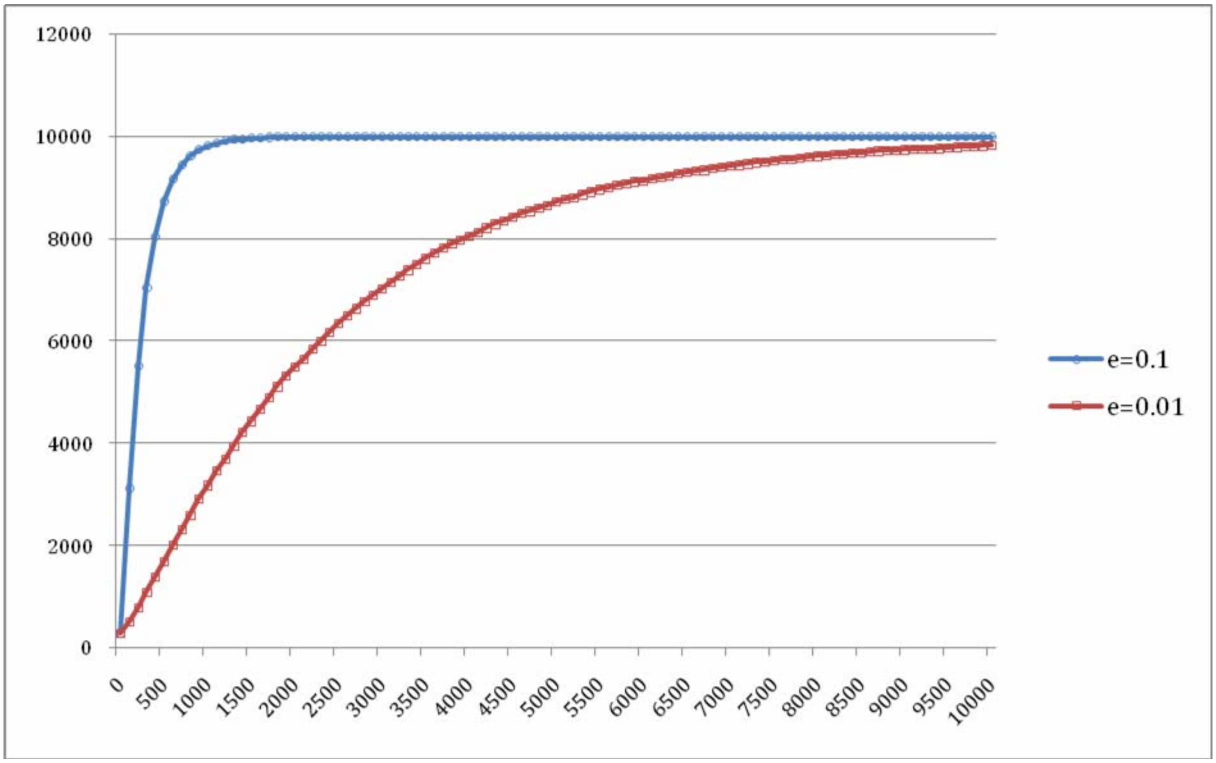
12000
10000
8000
6000
4000
2000
0
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
6000
6500
7000
7500
8000
8500
9000
9500
10000
e=0.1
e=0.01

Figure 1.