Group Decision and Negotiation **11**: 449–467, 2002

On Adaptive Emergence of Trust Behavior in the Game of Stag Hunt

CHRISTINA FANG, STEVEN ORLA KIMBROUGH, ANNAPURNA VALLURI AND ZHIQIANG ZHENG
The Wharton School, University of Pennsylvania, Philadelphia, PA 19104 USA

STEFANO PACE
School of Management, University of Bocconi, Italy

Abstract

We study the emergence of trust behavior at both the individual and the population levels. At the individual level, in contrast to prior research that views trust as a fixed trait, we model the emergence of trust or cooperation as a result of trial and error learning by a computer algorithm borrowed from the field of artificial intelligence (Watkins 1989). We show that trust can indeed arise as a result of trial and error learning. Emergence of trust at the population level is modeled by a grid-world consisting of cells of individual agents, a technique known as spatialization in evolutionary game theory. We show that, under a wide range of assumptions, trusting individuals tend to take over the population and trust becomes a systematic property. At both individual and population levels, therefore, we argue that trust behaviors will often emerge as a result of learning.

Key words: cooperative game, evolutionary game theory, reinforcement learning, simulation spatial game, Stag Hunt game, trust

1. Introduction

Allan Greenspan recently, and unexceptionally, underscored the critical nature of trust to our social and economic way of life, "[T]rust is at the root of any economic system based on mutually beneficial exchange . . . if a significant number of people violated the trust upon which our interactions are based, our economy would be swamped into immobility" (Greenspan 1999). Trust, or at least the principle of give and take, is a pervasive element in social exchange.

Despite its centrality, trust is a concept fraught with ambiguities and even controversies. The concept means different things to researchers in different fields or even sub-fields. Indeed, it is not entirely clear that there is a single concept to be found. Sociologists see trust as embedded in the larger concept of social capital (Adler and Woo 2002), while social psychologists interested in the same concept refer to 'emotional states and involuntary non-verbal behavior' as trust. Even organization researchers find it hard to agree on a consistent set of definitions (Zaheer *et al*. 1996). In fact, more than 30 different definitions of trust have been found in a recent survey of the related literature (McKnight and Chervany 1998).

Following the tradition of evolutionary game theory, we operationalize trust as the propensity to behave cooperatively in the absence of other behavioral indicators (Macy 1996). While we acknowledge that cooperation may not always be the result of trust, we believe that this operationalization is defensible because trust cannot be said to have existed if there is no manifested cooperation. In addition, equating trust with cooperation has been at the heart of a long established convention in both evolutionary and experimental game theories. As such, studying cooperative behavior represents a basic but important first step towards empirical understanding of trust. However, since we cannot refute the objection that the mechanisms we study do not produce trust, but merely its functional equivalence (Granovetter 1985), we refrain from making definitive statements about trust itself.

By way of framing this paper, we report results from a series of experiments that *pertain* to the origin and emergence of trust (or trust-like) behavior. By behavior we mean action that indicates or at least mimics the behavior of trusting individuals. We broadly take a game-theoretic perspective, in contrast to a philosophical analysis or a social-psychological study of trust. In certain games, it is natural and accepted to label some strategies as cooperative or trusting, and other strategies as evidencing lack of cooperation or trust. We simply follow that tradition.

In this paper, we investigate three main research questions related to the emergence of trust (i.e., trust-like behavior) at two levels of analysis. First, will individuals modeled as computer agents learn to behave cooperatively as a result of trial and error learning? We model this possibility of process or emergent trust by a computer algorithm borrowed from the field of artificial intelligence (Sutton and Barto 1998, Watkins 1989). Second, is this learning effective, do the learners benefit by it? We examine the performance implications of learning against different types of opponents adopting various strategies. For example, will the agent learn to recognize an opponent playing Tit-for-Tat and behave cooperatively as a result? Third, if trust can indeed be learned by individuals, will it spread throughout a society and emerge as a property of the entire system? We model this evolution of trust at the population level by a grid-world consisting of cells of individual agents, a technique known as spatialization in evolutionary game theory.

Our investigations are embedded in the context of a well-known game called *stag hunt*. This game has long been of interest because when humans encounter it, play is naturally described as either depicting or not depicting trust, depending on what the players do. At the least, there is apparently trusting behavior and apparently non-trusting behavior. Our focus in this paper is on whether rather simple artificial agents will display ‘trust behavior’ – behavior corresponding to apparently trusting behavior of humans playing the same game – when they play the game of stag-hunt. To that end, we discuss next the game of stag hunt and what in it that counts as trust behavior. From there, we move on to a discussion of our experiments with artificial agents playing the game.

2. Learning to Trust in a Game of Stag Hunt

The game of stag hunt (SH) is also known as the trust dilemma game (Grim *et al*. 1999). Before we set out to describe it, we briefly discuss a close kin of it that has been widely studied. That is the game of prisoner's dilemma (PD), a game that highlights the stark conflict that may exist between what is best for the society as a whole and the "rational" pursuit of individual needs. The payoff matrix of PD is shown in Table 1. The unique Nash equilibrium of the game (marked by *) occurs when both actors end up defecting (i.e., when both choose an apparently non-cooperative strategy).

Using iterated prisoner's dilemma (IPD, repeated plays of the game between fixed players), researchers have consistently found that cooperation, or trust, will evolve to be the norm under a broad range of conditions (e.g., Axelrod 1980; Grim *et al*. 1999; Nowak and Sigmund 1993). This basic model has therefore become the "*E. coli* of social psychology", and has been extensively applied in theoretical biology, economics, and sociology during the past thirty years. Perhaps too extensively, for according to behavioral ecologist David Stephens, researchers have been "trying to shoehorn every example of cooperative behavior into this Prisoner's Dilemma since 1981" (Morrell 1995).

Prisoner's dilemma (Macy and Skvoretz 1996) represents but one plausible model of social interaction in which the pursuit of individual self-interest will lead actors away from a mutually beneficial ("cooperative" or "trusting") outcome. Many social interactions involve the possibility of fruitful cooperation and do so under a regime other than PD. Because these situations have been comparatively under-investigated, we turn our attention to a different and much less studied game, namely the stag hunt (SH) game.

Stag hunt takes its name from a passage in Rousseau emphasizing that each individual involved in a collective hunt for a deer may abandon his post in pursuit of a rabbit adequate merely for his individual needs (Grim *et al*. 1999):

> When it came to tracking down a deer, everyone realized that he should remain dependably at his post, but if a hare happened to pass within reach of one of them, he undoubtedly would not have hesitated to run off after it and after catching his prey, he would have troubled himself little about causing his companions to lose theirs. (Rousseau, Discourse on the Origin of Inequality, 1755)

Here, the study of trust is embedded in a context rather different from prisoner's dilemma. No individual is strong enough to subdue a stag by himself, but it takes only one hunter to catch a hare. Everyone prefers a stag to a hare, and a hare to nothing at all (which is what a player will end up with if he remains in the hunt for a stag and his partner, the counter-player, runs off chasing hares). In this game mutual cooperation takes on the highest value for each player; everything is fine as long as the other player does not defect. Cooperation against defection, however, remains far inferior to defection against either cooperation or defection. As seen in Table 2, there are two Nash equilibria. One is the cooperative (trusting) outcome of mutually staying in the hunt for a stag. The other outcome is (non-trusting) mutual defection.

Table 1. Payoff matrix of a game of Prisoner's dilemma

	Cooperate	Defect
Cooperate	3,3	0,5
Defect	5,0	1,1*

Table 2. Payoff matrix of a game of stag hunt

	Cooperate	Defect
Cooperate	5,5*	0,3
Defect	3,0	1,1*

Our subsequent investigations are concerned with this trust game.

3. Experiments in Individual-Level Learning

At the individual level, we model agents who are able to learn in repeated games and who may then learn (apparently) trusting behavior. We model, or simulate, this individual learning process by an algorithm known as Q-learning in artificial intelligence (Sandholm and Crites 1995; Watkins 1989). Our simulation consists of two players playing the game of stag hunt iteratively for a specified number of times. To begin, we fix the strategies of one of the players [the opponent] and examine how the other player [the learner] adapts. We hypothesize that a cooperative (apparently trusting) outcome will be learned by the learner if the opponent also acts cooperatively.

Q-learning is widely used in artificial intelligence research (Hu and Wellman 1998; Littman 1994; Sutton and Barto 1998). It is part of the family of reinforcement learning algorithms, inspired by learning theory in psychology, in which the tendency to choose an action in a given state is strengthened if it leads to positive results, weakened if the results are unfavorable. This algorithm specifies a function, the Q-function, which depends on a pair of state and action variables that keep track of the value for the agent of taking a particular action in a given state. There are only four possible outcomes in any single play of a 2-player stag hunt game. Each player independently has 2 available actions: to cooperate [C] or to defect [D]. When the player has a memory of the outcome of one (the most recent) play of a game, the Q-function is a 4-by-2 table with 8 cells as shown in Table 3. (Given each of the four outcomes of a game, the Q-function has to decide which of two strategies [C] or [D] to play next.). Here we assume an agent has the capability of remembering the outcome of only the previous round of play. If, however, the agent has a memory capacity of two (plays), the number of states would increase by a multiple of the number of possible outcomes. For example, if the memory capacity were two, the number of possible states would be 16 (4×4).

Table 3. Q function with 4 states and 2 actions

State/action pairs	Cooperate [C]	Defect [D]
Both cooperate [CC]	X	X
One cooperate; the other defect [CD]	X	X
One defect; the other cooperate [DC]	X	Y
Both defect [DD]	X	X

In Table 3, Y corresponds to the expected value of defecting while in state DC (i.e., during the last game the learner defects but the opponent cooperates). In other words, the value of the cell Q(DC, D) equals Y. (Note: [CD] is shorthand for the learner cooperates and the opponent defects, while [DC] indicates that the learner defects and the opponent cooperates.)

As entries in the Q table store the value of taking a particular action given an observed state from the previous iteration, learning the relative magnitude of these values is key to effective adaptation. Such learning can occur through repeated exposure to a problem, when an agent explores iteratively the consequences of alternative courses of action. In particular, the value associated with each state-action pair (s, a) is updated using the following algorithm (Watkins 1989):

$$Q(s, a) \leftarrow (1 - a)\, Q(s, a) + a * (R + g \max aeA{,}Q(s\hat{a}^{\grave{A}\grave{U}}, a\hat{a}^{\grave{A}\grave{U}}))\,.\,.\,. \qquad (1)$$

where â and g are the learning rate and the discount factor, respectively. sâ is the next state that occurs when the action a is taken while in state s, and aâ is the action, in the set of possible actions, in state sâ with the maximum value, and R is the reward received by the agent for the actions taken.[1]

As more games are played, the initially arbitrary beliefs in the Q-function table are updated to reflect new pieces of information. In each game, the agent chooses probabilistically[2] the preferred action, observes the state of the game and the associated payoffs, and uses this information to update his beliefs about the value of taking the previous action. In this way, he learns over time that certain states are better than others. Since the cooperative outcome yields the maximum payoff, we might expect the learner to gradually choose to cooperate over time when playing against a cooperative opponent. Such emergence of trust or cooperative behavior, however, may be critically dependent on the strategies adopted by the opponent. For instance, if the opponent has a fixed strategy of always defecting, then an intelligent learner should learn not to cooperate since doing so will always earn him a payoff of 0. Therefore, to provide a more realistic picture of the dyadic dynamics, we need to account for the opponent.

To simplify, we fix the strategies of the opponent such that he behaves quite predictably. In general, any fixed strategy can be characterized by triples <I, c, d>, where I indicates initial play, c the learner's response to his opponent's cooperative move in the previous round and d his response to defection by the opponent in the previous round; assuming that the players have a memory of one. For example, the most commonly studied strategy, Tit-for-Tat, is described as <1, 1, 0>. An agent following such a strategy will always cooperate in the first round, and later continue to cooperate unless the opponent defects. Therefore, we can study the efficacy of learning by pitting our learner against opponents with each of the eight possible strategies stated in a binary fashion in Table 4.[3]

Two modeling clarifications need to be mentioned here. First, there are other relevant computational models of learning, notably genetic algorithms and neural networks. For example, Macy and Skvoetz (1996) use genetic algorithms to model trust in an iterated prisoner's dilemma (IPD). These so-called evolutionary approaches differ from Q-learning in that they directly search the space of possible policies for one with a high probability of

Table 4. Eight Fixed Strategies [0 indicates Defect; 1 indicates Cooperation]

Initial action	If observe cooperation	If observe defection	Reactive strategy
0	0	0	Always defect
0	0	1	Suspicious doormat
0	1	0	Suspicious tit for tat
0	1	1	Suspicious Quaker
1	0	0	Deceptive defector
1	0	1	Gullible doormat
1	1	0	Tit-for-Tat (TFT)
1	1	1	Quaker

winning against the opponent. Learning therefore is said to occur off line through extensive training. Q-learning, on the other hand, learns while interacting with the environment without conducting an explicit search over possible sequences of future states and actions (Sutton and Barto 1998). Second, learning in a 2-by-2 game encompasses more than just a stationary game against nature since the behavior of the opponent may change as well. In multi-agent learning tasks in the context of a 2-by-2 game, one can either explicitly model the opponent or simply consider the opponent as a part of a non-stationary environment. While Hu and Wellman (1998) prove that Q-learning algorithms taking into account opponents do converge to Nash equilibria under specific conditions, Banerjee, Mukherjee, and Sen (2000) show that those who do not model opponents often out-perform those who do in the long run in their 2-by-2 game. Since the latter approach is far simpler and has been shown to work well, we have chosen not to explicitly model opponents' behavior.

Our results for individual learning are based on simulations of 500 plays against an opponent for each strategy in Table 4, averaged over 5000 games. The learning rate α is set at 0.999954 and set to decay at the same rate. The discount factor γ is fixed at 0.9. The softmax temperature τ is initially set at 30.0598 and reduces incrementally to 0.0598, such that eventually the search becomes almost greedy. We have chosen these parameter values after the values used by other studies (e.g., Sandholm and Crites 1995).[4]

3.1. Simple learning against a fixed opponent

The Q-learning mechanism ought to find reasonable strategies when playing against a fixed opponent. Therefore, as a validation test and for purposes of benchmarking, we begin with a series of experiments involving simple Q-learning by a learner against an opponent playing a fixed strategy in the iterated stag hunt (ISH) game. Table 5 summarizes the frequency of occurrence across of the four possible states of play at the end of the 500 rounds of play averaged over 5000 runs, in which the learner plays against opponents with different fixed strategies. The opponents' strategies are indicated in the topmost row of the table (see Table 4 for definitions). The row headings (in the left-most column) indicate possible outcomes, e.g., CD means learner cooperates and opponent defects. Indeed, after playing

Table 5. Evolution of trust in 2-player game of stag hunt

	Always defect	Suspicious doormat	Suspicious tit for tat	Suspicious Quaker	Deceptive/suspicious defector	Gullible doormat	Tit for Tat	Quaker
CC	0	0.593	1	1	0	0.45	1	1
CD	0	0	0	0	0.006	0	0	0
DC	0	0.083	0	0	0	0.05	0	0
DD	1	0.324	0	0	0.994	0.5	0	0

500 games, cooperation emerges successfully in matches in which the learner is pitted against relatively trusting opponents such as those playing Tit-for-Tat or Quaker. The learners quickly (apparently) realize that it is in their mutual best interest to play cooperatively. The cooperative outcome [CC] becomes the dominant state and by the end of the 500 plays, it is observed 100% of the time. Trusting behavior, however, is not beneficial to the learner if the opponent tends to defect. Not surprisingly, the learner learns to defect against opponents playing an all defect strategy. This statement is corroborated by the fact that the defecting state [DD] occurs 100% of the time.

Moreover, our learner fairs well against opponents playing tricky strategies such as Gullible doormat, and suspicious doormat. A tricky strategy is one in which the opponent defects when the learner cooperates, and cooperates when the learner defects. Optimal play against such a strategy is not completely apparent to our players. Nevertheless, it is obvious that the learner will be exploited if found in a state where he cooperates and the opponent defects. Results in Table 5 substantiate the reasoning that the learner seems to be learning, since he never finds himself being exploited. In fact, the learner's actions seem to match the actions of the opponent quite accurately with a fixed strategy of "Gullible doormat", given that only 5% of the outcomes result in the state DC.

It is instructive to plot the evolution of various outcome states over time against opponents playing three types of fixed strategies: Tit-for-Tat, Always Defect and Suspicious Defector.

In Figure 1, where the opponent plays Tit-for-Tat, we see that the learner eventually learns to cooperate with the opponent and attains the higher payoffs associated with the cooperative outcome. In Figure 2, where the opponent always defects, the learner learns to avoid being taken advantage of and always defects.

In Figure 3, the learner encounters a tricky opponent, namely, a gullible doormat (defect after seeing cooperation and cooperate after seeing defection). However, the learner quite intelligently learns to play against him. He avoids the state [CD] where he is taken advantage of, and 95% of the time chooses the same action as his opponent.

In a nutshell, we have shown that in cases in which cooperation is in the mutual interest of both players, our learner has no trouble identifying and carrying out the cooperative solution. Moreover, even when the learner encounters an opponent playing a tricky strategy (gullible doormat), he adapts well. Such learning is quite effective, despite the simplicity and informational paucity of the learning algorithm and the fact that the agents start with limited knowledge of the external environment.

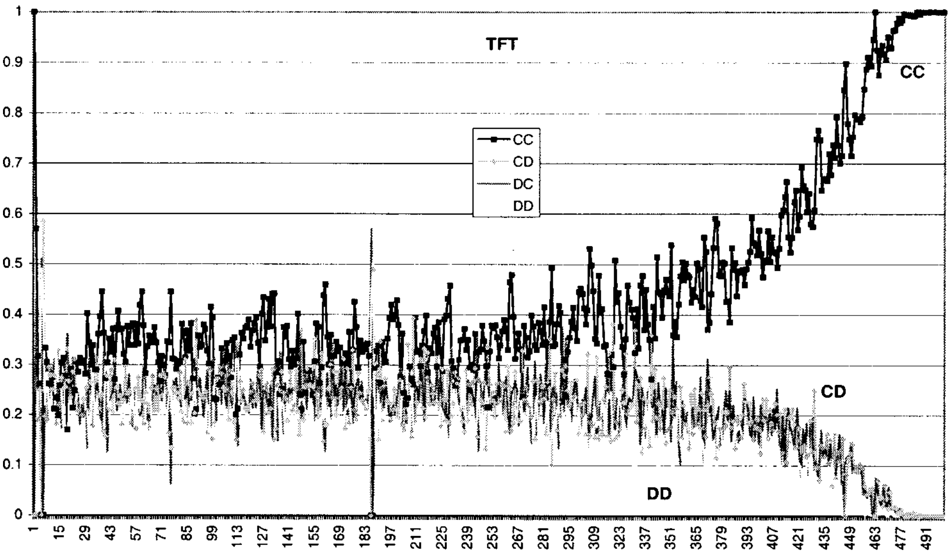

Figure 1. Evolution of Trust for fixed opponent playing TFT.

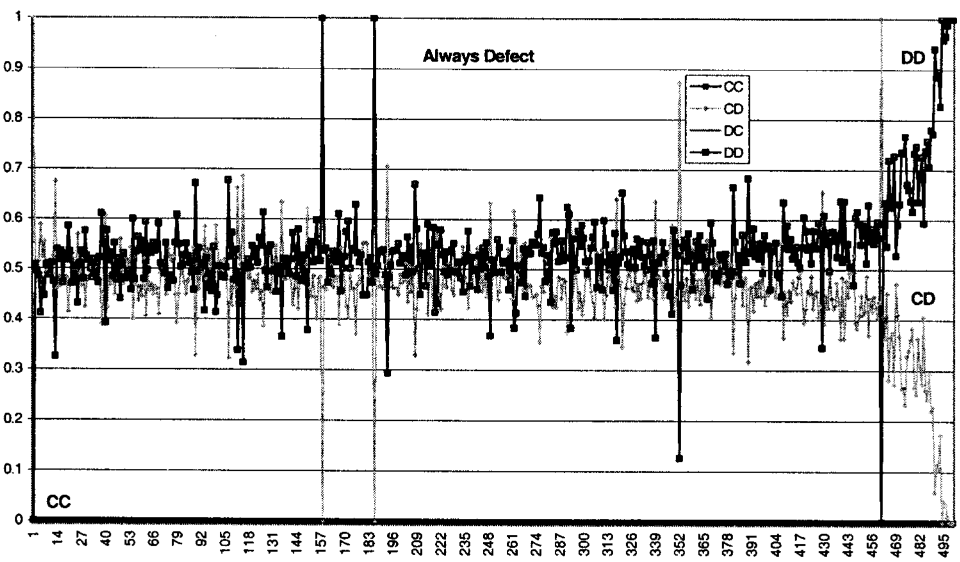

Figure 2. Evolution of Trust for fixed opponent playing Always Defect.

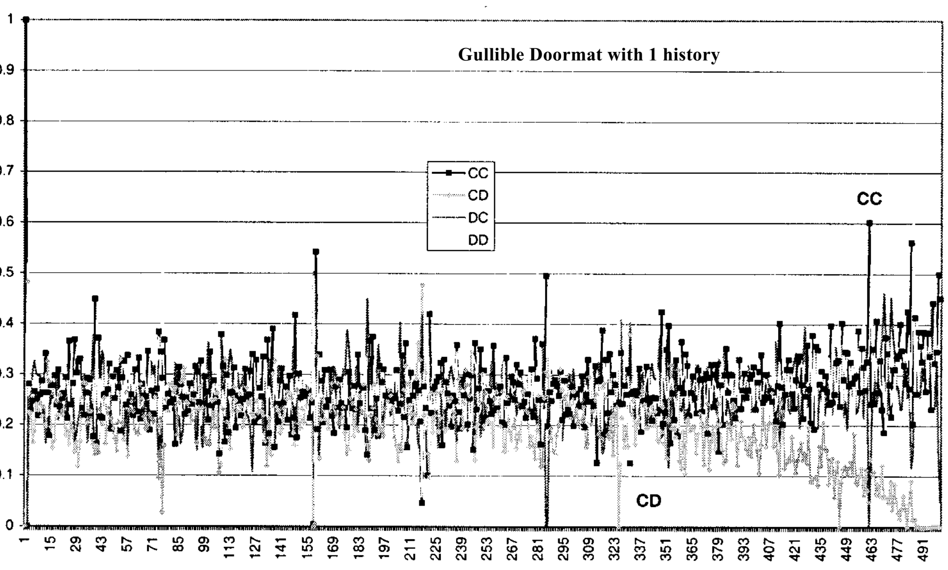

Figure 3. Evolution of Trust for fixed opponent playing Gullible Doormat.

3.2. Learning with More Memory Capacity

In the experiments of the previous section, both the learner and the opponent have a memory capacity of 1. They can only remember what happened in the last round, which becomes the basis for their current action. It is reasonable to hypothesize that a 'smarter' agent, one with a larger memory capacity, will outperform his more mnemonically challenged counterparts. In particular, we would expect this intelligence to manifest itself to good effect in games in which the opponent's strategy is more complicated. Sandholm and Crites (1995), for instance, find that agents with longer memories tend to fair better than those with shorter ones. They use history-lengths of 2 and 3 to study learning against a Tit-for-Tat opponent in the iterated Prisoner's Dilemma game. To investigate this conjecture, we endow agents with a slightly greater memory (history) size of 4 units – meaning that the learning agents can remember what transpires in the previous 4 episodes. In addition to studying the impact of history length on learning against a Tit-for-Tat (TFT) opponent, we also study play against a tricky opponent (gullible doormat). In the case of playing against a TFT opponent, we find that the number of episodes it takes for the pair to reach a cooperative equilibrium is much lower when the learner can remember more. It takes only 180 episodes whereas it takes almost the entire 500 episodes if the agents have a memory capacity of 1. More memory in this context clearly implies better performance. However, in the case of the opponent playing the gullible doormat strategy, more memory is not very helpful. Although the learner never finds himself being exploited (state [CD]), he is unable to accurately judge the next action of the opponent; a capability he has in the case of less memory.

3.3. Learning against another learner

The experiments so far assume that there is only one agent who is learning; the strategy of the opponent is fixed. What will the collective outcome be if both parties learn simultaneously? On one hand, if both agents can learn to realize the Pareto-enhancing feature of the cooperative outcome, games played by two learners will converge to the social optimal more quickly than before. On the other hand, since both agents are learning, one agent's action will alter the environment faced by the other, presenting the agents with a non-stationary environment. In other words, there is less stability because now both agents have to constantly adjust to the other, as in a real game.

As shown in Figure 4, we find that in about 73% of the runs the agents learn to achieve the mutually beneficial cooperative outcome (state [CC]), while 13% of the time they end up in the equilibrium state [DD]. Interestingly, the two states ([CC] and [DD]) correspond to the two Nash equilibria of the stag hunt game, and learners emerge to these states 86% of the time. Traditional game theory is unable to predict which of the two Nash equilibria would be sustained. However, our computational approach gives some insight on the greater possibility of the emergence of the cooperative outcome. Moreover, we can still conclude that two self-interested agents can often produce a Pareto-efficient outcome that maximizes the welfare of the pair together, as seen in Figure 4.

4. Population Level Learning

The evolution of trust is quintessentially a problem beyond the simplicity of a 2-by-2 game. We now turn to learning to trust with multiple individuals and examine how trust

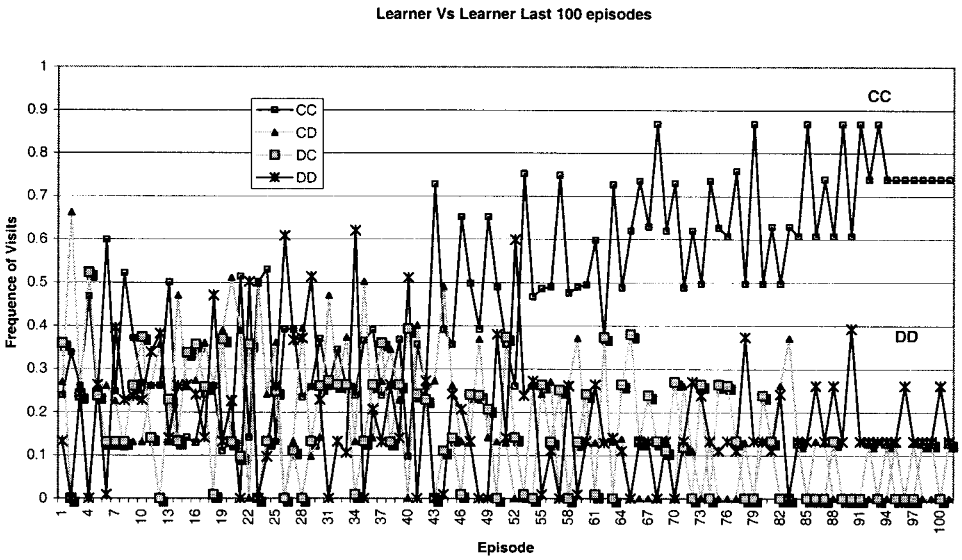

Figure 4. Evolution of trust with two Learners.

can evolve into a systemic property. At the population level, we are primarily interested in whether and how trust can spread throughout the entire population. Specifically, we want to see how this evolution is influenced by (1) the initial percentage of trusting individuals and (2) the initial distribution of them.

To model the emergence of trust at the system level, we use a natural extension of the 2-by-2 game by embedding it in a spatial framework. In this set of experiments, we develop an 81-by-81 grid to represent the space of a population. Individual agents, represented by cells in a spatialized grid, behave entirely in terms of simple game-theoretic strategies and motivations specifiable in terms of simple matrices (Grim *et al*. 1999; Picker 1997). We shall see how a behavior resembling cooperation and even generosity can arise as a dominant pattern of interaction in a population of individuals primarily driven by self-interests. A given cell is occupied in the next round by a player of the type that received the highest payoff among the nine cells centered on the cell in question. This model is a natural interpretation of the Darwinian model of adaptive success generating reproductive success.

Specifically, each player interacts with his immediate eight neighbors (Table 6). Without loss of generality, the central player "C" interacts with his eight adjacent neighbors from N1 to N8. He plays the free standing 2-by-2 game with each of his eight neighbors, but plays the same strategy per round, either the strategy of cooperation or defection. In short, he is either a pure cooperator or a pure defector and his sphere of interaction is restricted to immediate neighbors. Player C's payoff is determined from the payoff function defined by the stand-alone 2-by-2 game of stag hunt, and the plays of each of his neighbors. His total payoff is computed by summing payoffs from eight games; and the payoffs of his neighbors are determined similarly. In the next round, each agent then surveys his neighbors. For instance, player C observes his own payoff and the payoffs of his eight neighbors of the previous round. He then chooses the action (coorporate or defect) that yields the highest payoff among the nine cells. As such, an action with higher performance replaces the initial action, which is selected out of the evolutionary struggle.

Trust in this context is measured by the *type* of players, as defined by the 2-period history of his plays (Epstein and Axtell 1996; Picker 1997). For instance, if the cell player has cooperated in both the previous round and the current round, we code his strategy as 1. All four possible 2-period strategies can be summarized in Table 7.

Table 6. A representation of spatial games

	N1	N2	N3	
	N4	**C**	N5	
	N6	N7	N8	

Table 7. Types and their corresponding strategies

Type	Strategy
1	Two consecutive rounds of cooperation
2	Two consecutive rounds of defection
3	Switch from cooperate to defect
4	Switch from defect to cooperate

Type 1 player is defined as a "trust" type player since under this strategy a player chooses to cooperate consecutively. With this set of simulation tools, we set out to investigate whether the players will gradually adopt the "trust" strategies over time, despite an initially random distribution of strategy types. If the population is increasingly taken over by players with a genetically wired "trust" strategy, then we are able to show that trust has emerged as a systemic property of the population.

The game stops when all the cells converge to one of the two possible equilibria: [DD] or [CC], which are the two Nash equilibria. We define the first time the game reaches a uniform action (all "C" or all "D") for all cells as the *converging point*, and call the number of steps/games to reach this point the *converging steps*.

4.1. Impact of the initial density of trusting individuals

We quantify "trust" by defining a trust density index. Density of trust in a particular period is defined as the percentage of trust type players among all the players. We've defined a trust type player as one who takes a cooperative action in two consecutive periods. In our grid world, there are 6561 (= 81X81) players in each period. So in a particular period,

$$\textit{Trust Density Index} = \textit{Number of players who are "trust" type}/6561 \quad (2)$$

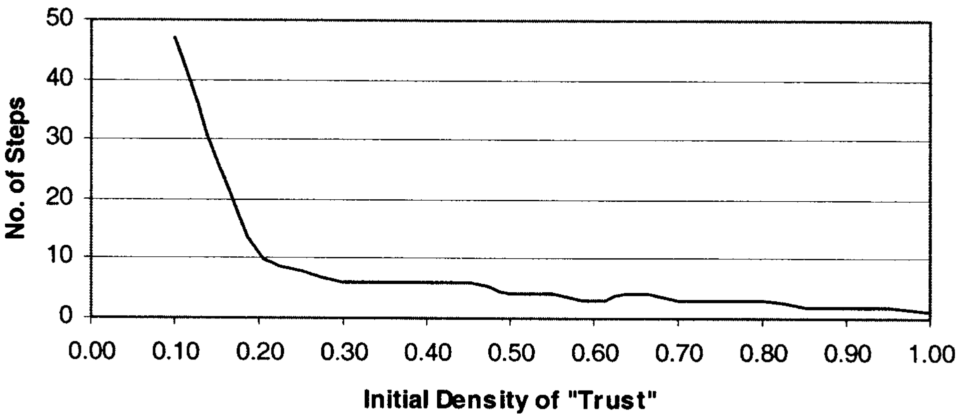

Figure 5. Impact of Density on Speed of Convergence.

We experimented by initializing the population randomly with different trust density levels. We plot the speed of convergence as a function of different trust densities in Figure 5. In general, the more cooperative the initial population, the faster it converges to an all-trust society (every-one in the population belong to the "trust" type.) Figure 5 also suggests that when the initial density is below a certain point, it's highly unlikely for the population to converge to an all-trust one (the curve is not tangent to the Y axis).

While Figure 5 suggests an optimistic outcome for societal evolution of trust, it is intriguing why so many societies (initial distributions of trust behavior) tend to evolve toward cooperation. In particular, societies with little initial trust seem able to overcome this seemingly insurmountable barrier of suspicion.

In Table 8, we show the actual density of the first 10 steps for six density levels. These societies often need to go through "a valley of death" before trusting behaviors take over. 'Valley of death' refers to the process of the whole society rapidly degenerating towards mutual distrust, and later its dramatic reversal, wherein the remaining trusting agents propagate and take over the entire population. This illustrates very well the temporal tradeoff inherent in many natural phenomena – initially the act of defection pays as it exploits the other trusting players. Notice that in the first couple of steps, the density of three levels (0.25, 0.15, 0.10) declines quickly and the population is quickly depleted of trusting players. However, this depletion results in a corresponding decline in the number of "bad elements" before trusting individuals take over again by reaping superior payoffs. Moreover, the reversal process takes almost 30 generations when initial density is at a low 0.1 level.

4.2. *Impact of initial distribution of trust*

From the results of the previous section, we find that the convergence path and speed depend heavily on the initial density of trusting individuals in the population. We further find that the speed of convergence might be completely different between populations with the same initial density but different distributions of trust. We ran ten rounds with the

Table 8. First Ten Generations of Evolution

Steps	Density = 1.00	Density = 0.50	Density = 0.25	Density = 0.15	Density = 0.1	Density = 0.05
1	1.00	0.5	0.25	0.15	0.1	0.05
2	1.00	0.59	0.09	0.01	0.01	0.03
3	1.00	0.76	0.26	0.01	0.01	0.02
4	1.00	0.94	0.53	0.03	0.01	0.01
5	1.00	0.99	0.76	0.07	0.01	0
6	1.00	1	0.92	0.14	0.01	0
7	1.00	1	0.98	0.21	0.01	0
8	1.00	1	1	0.3	0.01	0
9	1.00	1	1	0.4	0.01	0
10	1.00	1	1	0.48	0.01	0

Table 9. Number of steps to converge in ten rounds with initial density = 0.10

Round	1	2	3	4	5	6	7	8	9	10
Steps	26	32	42	37	27	25	39	21	32	13

same initial density of 0.10 and show the number of steps needed before convergence in Table 9. In short, there is a large variance in the speed of convergence.

This clearly suggests that the speed of convergence not only depends on the initial density, but also on how the trust type players are distributed over the space. In this sub-analysis, we further examine the effects of location of these trust-type players. We hypothesize that a population with an even distribution of trusting individuals may behave quite differently from segregated populations with the same trust density overall. In order to test this conjecture, we need to develop an index to measure how evenly trust is distributed across the grid.

First, we slice the 81-by-81 grid into 9 blocks (see Table 10), with each block containing 81 players (9 × 9). After initialization, each block contains a certain number of "trust" type players. If trust is evenly distributed, the nine blocks should contain roughly the same number of trust agents. So any deviation of the number of trust type players from this benchmark indicates an uneven distribution of trust.

Therefore, we can measure the evenness of the initial distribution by an index of deviation. This can be done in two steps: (1) First we compute the number of trust type players for each of the nine blocks (by doing so we get nine numbers). (2) Then we compute the standard deviation of these nine numbers and use it as the measure of the evenness of the distribution. The higher the deviation, the less evenly trust is distributed.

Index Trust Distribution = StandardDeviation (Block1, Block2, . . . Block9) (3)

The effect of the initial trust distribution on the evolutionary dynamic is shown in Figure 6, which plots the relationship between evenness of distribution of the trusting types and the speed of convergence, under three different density levels of 0.5, 0.25 and 0.10. The three curves exhibit the same pattern: the higher the deviation, the slower the convergence (more steps to reach convergence). In other words, a more evenly distributed society moves more quickly towards a completely trusting society. Segregation, an extreme case of uneven distribution, impedes the evolution of trust of a society.

It is not difficult to see why distribution matters. A lone cooperator will be exploited by the surrounding defectors and succumb. However, four cooperators in a block can conceivably hold their own, because each interacts with three cooperators; a defector, as an outsider, can reach and exploit at most two. If the bonus for cheating is not too large, clus-

Table 10. Blocks to Measure Evenness in Distribution

Block 1	Block 2	Block 3
Block 4	Block 5	Block 6
Block 7	Block 8	Block 9

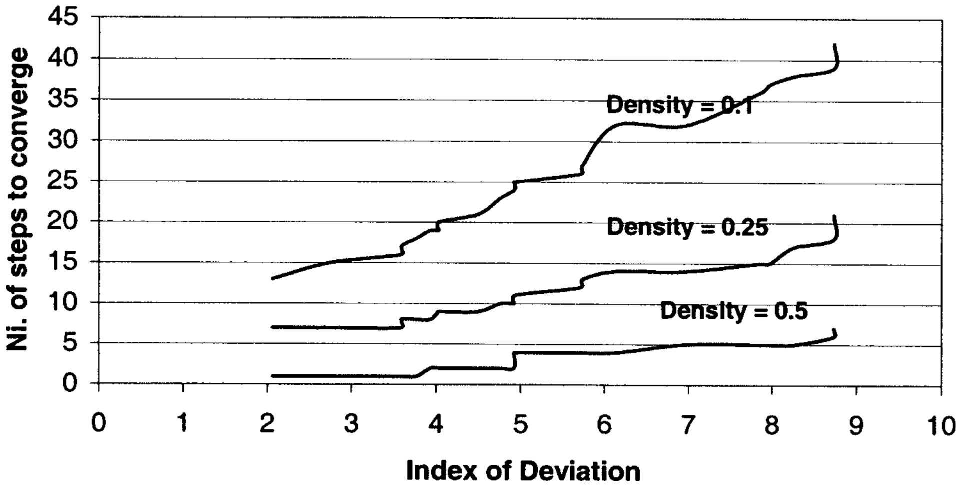

Figure 6. Impact of Distribution on Speed of Convergence.

ters of cooperators will grow. As in any path dependent process, the actual evolution of such spatial societies is sensitive to initial values. However, the long-term average of the final composition of the population is highly predictable.

In sum, these results show that trust indeed diffuses under general conditions. The artificial society we simulate here does evolve towards cooperative behavior without external intervention. As such, trust eventually becomes a systemic property of the entire population, despite the fact that initially it only characterizes a small subset of the population. By exhibiting trusting behaviors, both the society as a whole and individuals themselves reap higher payoffs, ensuring that these behaviors will be passed on to the next generation.

5. Discussion and Conclusion

The subject of trust has been the focus of much attention both in theory and in applications. The experiments we report here are admittedly exploratory and much remains to be done. It is intriguing, however, that trust behavior emerges so pervasively in the simulations we undertake. With this in mind, we offer a few comments by way of framing this work and its significance.

Trust is naturally studied in the context of games, or strategic situations, which themselves have been studied from broadly three perspectives. First, classical game theory and its modern descendants have tended to view game theory as a branch of applied mathematics. Games are formalized (strategic form, extensive form, etc.), axiom systems are presented (e.g., utility theory, common knowledge), solution concepts are conceived (equilibria), and results are derived analytically. The concern about this approach is and has always been that these "highly mathematical analyses have proposed rationality requirements, that people and firms are probably not smart enough to satisfy in everyday decisions" (Camerer 1997). In response, a rich and flourishing field, called behavioral game theory, has arisen "which aims to describe actual behavior, is driven by empirical observation (mostly experiments), and charts a middle course between over-rational equilibrium analyses and under-rational adaptive analyses" (Camerer 1997).

Our study lies in the third approach, which Camerer calls "under-rational adaptive analyses" employing "adaptive and evolutionary approaches [that] use very simple models – mostly developed to describe nonhuman animals – in which players may not realize they are playing a game at all." Indeed, since our agents are so lacking in what anyone would call rationality, of what possible relevance is their behavior to the study of trust?

Classical as well as behavioral game theoretical analyses have shared in common an outlook in which the game players have a psychology of the sort postulated in everyday thought and language. Philosophers call this folk psychology. Consistent with folk psychology, the artificial intelligence community has often sought to model human (and intelligent agent) behavior in the BDI style (beliefs, desires, and intentions). Behavioral game theory seeks to explain behavior in games by appealing to realistic beliefs, desires, intentions, and computational capabilities of the players. While classical game theory can be understood as investigating behavior under (often unrealistically) ideal cognitive conditions, it too assumes a broadly folk psychological outlook.

In contrast, adaptive or evolutionary game theory has focused on the behavior of simple algorithms and strategies as they play out in strategic contexts. Typically, as is the case for our agents, the players cannot by any stretch of imagination be granted beliefs, intentions, or desires. But they can, and do, play games that are studied from the classical as well as behavioral perspectives. We find that in the stag hunt game behavior emerges that, *had the players been BDI agents*, would be described as trusting or cooperative.[5] These findings are interesting and significant in a number of ways.

1. We study stag hunt, a game that has been studied relatively little compared to prisoner's dilemma (Cohen *et al*. 1998). It represents an alternative but equally interesting context in which to investigate cooperation.
2. It is interesting that, as in the case of prisoner's dilemma, "cooperative" behavior occurs robustly.
3. Our computational approach sheds light on the emergence of the cooperative outcome when both players are learning. Traditional game theory, however, is unable to predict which of the two Nash equilibria (of the stag-hunt game) would be sustained.
4. The dynamics of stag-hunt (spatial) populations are surprising, particularly the "valley of death" phenomenon and the adverse effect of segregation (uneven distribution of trusting players) on the evolution of society.
5. Simulation as a methodology is also a form of experimental mathematics. That cooperative behavior emerges in these experiments (in the absence of anything like cognition or rationality) may *suggest* that any cooperative behavior emerging may simply be a reflection of the underlying power arrangements and "correlation of forces" in a particular strategic situation.
6. The previous remark suggests further that at least some of the purportedly rational

behavior of BDI agents such as ourselves might be explained by adaptive or evolutionary processes, which produce seemingly cooperative behavior in our game of stag hunt (See (Skyrms 1996) for a similar suggestion.).

Of course, much remains to be investigated. What we have begun here needs a broader and deeper study. It is especially important to investigate the space of plausible adaptive mechanisms. In this regard, computational studies of reinforcement learning (broadly construed) have begun to connect with the tradition of classical game theory as it has developed an interest in algorithms for game dynamics (Fudenberg and Levine 1998; Weibull 1995; Young 1998). Our paper represents a first step towards this synthesis.

At the individual level, we find that cooperative behaviors can emerge purely as a result of trial and error learning. Trust emerges almost autonomously, without any need for central intervention. This finding has several interesting applications to organizations and businesses in general. For instance, a long-standing discourse in organization theory is the tension between centralization and decentralization. Centralization has the potential to minimize redundancy and waste but runs the risk of over-intervention. Decentralization, however, delegates decision-making power to local entities but in an attempt to optimize selfishly, may be accused of sacrificing global welfare. Our paper provides empirical support to the benefits of decentralization by showing that mutual and seemingly selfish adjustment by subunits does indeed lead to greater global welfare over time. More importantly, this adjustment is achieved without the presence of a centralized authority. This implies that there is no need for complete knowledge of the environment as well.

More specifically, many dynamic interactions in business exhibit qualities that are similar to the game of stag hunt. For instance, in joint ventures, two distinct firms have to learn to co-exist and achieve pre-defined goals. While there is certainly common interest in seeing the venture pay off, a conflict of interest also exists when players have incentive to shirk. In a similar vein, in an e-commerce context, the interactions between the ‘bricks’ unit and the ‘clicks’ unit also have the distinctive character of a trust game. While they need to cooperate with each other to further the organizational goal, these subunits are also locked in a bitter internal battle as cannibalization becomes inevitable. These diverse business decisions, in particular, simply lack the egoistic character of the Prisoner’s dilemma. Instead, they share important common features with a game of stag hunt. As such, they may prove to be potentially fruitful applications of our research.

At a population level, we find that while trust eventually spreads to the entire population via selection, it does so with an interesting temporal pattern: the valley of death. This finding may have important normative implications from a central planner’s perspective. In particular, it is important for the planner not to be “deceived” by the initial dying down of cooperation. In a sense, this initial drop in trust is a necessary price to be paid for the eventual taking over of the population. As such, while intervention in any form is not needed in general, it should be especially avoided during this valley of death.

To conclude, in this paper, we examine the adaptive emergence of trust by showing how it can evolve at both the individual and the population levels. Although highly simplified and admittedly stylized, our simulation produces some very general and interesting observations.

Acknowledgement

This material is based upon work supported by, or in part by, NSF grant number 9709548.

Notes

1. For an introduction to the family of reinforcement learning models, see *Reinforcement Learning* by Sutton and Barto (1998).
2. The choice of action is guided by the so-called Softmax exploration method to prevent pre-mature locking in into local optima (Sandholm and Crites 1995 for details). Essentially, the method ensures that all actions have a positive probability of being chosen in any given round. The degree of greediness in search is tuned by a temperature parameter τ, with smaller values of τ representing a greedier search process.
3. We do not consider Pavlov, a recently proposed strategy by Nowak and Sigmund (1993). It cooperates on the first move and thereafter repeats its previous move if the opponent cooperated; otherwise it switches.
4. We use the payoffs shown in Table 2 as rewards in our simulation. However, we have also conducted the experiments by altering the payoffs in the table. We have chosen not to report them as our qualitative results have not changed. As such, our results are not sensitive to the particular payoff values we have specified.
5. Analogous results have also been found for Iterated Prisoner's Dilemma.

References

Adler, P., and S. W. Kwon. (2002). "Social Capital: Prospects for a New Concept," *Academy of Management Review* 27 (1), 17–40.

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

Banerjiee, Bikramjit, Rajatish Mukherjee, and Sandip Sen. (2000). "Learning Mutual Trust," in Working Notes of AGENTS-00 Workshop on Deception, *Fraud and Trust in Agent Societies*. Barcelona, Spain, 9–14.

Banerjiee, Bikramjit, Sandip Debnath, Sandip Sen. (2000). "Combining Multiple Perspectives," in *Proceedings of the International Conference on Machine Learning 2000*, June 29–July 2. Stanford University, CA, 33–40.

Camerer, Colin F. (1997). "Progress in Behavioral Game Theory," *Journal of Economic Perspectives* 10 (4), 167–188.

Camerer, C., and T. Ho. (1999). "Experience-Weighted Attraction Learning in Normal Form Games," *Econometrica* 67.

Cohen, Michael D., Rick L. Riolo, and Robert Axelrod. (1998). "The Emergence of Social Organization in the Prisoner's Dilemma: How Context-Preservation and Other Factors Promote Cooperation," Working Paper in the University of Michigan.

Epstein, R., and R. Axtell. (1996). *Growing Artificial Societies*. MIT Press.

Erev, Ido, and Alvin E. Roth. (1998). "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *The American Economic Review* 88 (4), 848–881.

Fudenberg, Drew, and David K. Levine. (1998). *The Theory of Learning in Games*. Cambridge, MA: MIT Press.

Greenspan, A. (1999). Harvard University Commencement Address.

Grim, P., Gary Mar, and Paul St. Denis. (1999). *The Philosophical Computer: Exploratory Essays in Philosophical Computer Modeling*. Cambridge, MA: MIT Press.

Hannan, M., and J. Freeman. (1977). "The Population Ecology of Organization," *American Journal of Sociology* 82, 929–964.

Holland, John H. (1995). "Can There be a Unified Theory of Complex Systems?" *in* M. Morowitz and J. Singer (eds.), *The Mind, the Brain, and CAS*. SFI Studies in the Sciences of Complexity, Vol. XXII, Addison-Wesley.

Hu, Junling, and Michael P. Wellman. (1998). "Multiagent Reinforcement Learning: Theoretical Framework and an Algorithm," *in Proceedings of the Fifteenth National Conference on Artificial Intelligence*. Menlo Park, CA, 746–752.

Kandori, Michihiro, George J. Mailath, and Rafael Rob. (1993). "Learning, Mutation, and Long Run Equilibriain Games," *Econometrica* 61 (1), 29–56.

Lave, C., and J. March. (1975). *An Introduction to Models in the Social Science*. Harper & Row.

Littman, M. (1994). "Markov Games as a Framework for Multi-Agent Reinforcement Learning," *in Proceedings of the Eleventh International Conference on Machine Learning*. Morgan Kaufmann, San Francisco, CA, 157–163.

Macy, Michael W. (1996). "Natural Selection and Social Learning in Prisoner's Dilemma: Co-Adaptation with Genetic Algorithms and Artificial Neural Networks," *Sociological Methods and Research* 25, 103–137.

Macy, Michael W. (1997). "Identity, Interest and Emergent Rationality: An Evolutionary Synthesis," *Rationality and Society* 9, 427–448.

Macy, M., and J. Skvoretz. (1998). "The Evolution of Trust and Cooperation among Strangers: A Computational Model," *American Sociological Review*.

McEvily, Bill, and Akbar Zaheer. (1999). "Bridging Ties: A Source of Firm Heterogeneity in Competitive Capabilities," *Strategic Management Journal* 20, 1133–1156.

McKnight, D. Harrison, Cummings L. L., and Chervany N. L. (1998). "Initial Trust Formation in New Organizational Relationships," *Academy of Management Review* 23 (3), 472–490.

Miller, John H. (1993). "The Evolution of Automata in the Repeated Prisoner's Dilemma." Working Paper, Santa Fe Institute, Economics Research Program.

Morrell, V. (1995). "Cowardly Lions Confound Cooperation Theory," *Science* 269, 1216–1217.

Nowak, M., and K. Sigmund. (1993). "A Strategy of Win-Stay, Lose-Shift that Outperforms Tit for Tat in the Prisoner Dilemma Game," *Nature* 364, 56–58.

Picker, R. (1997). "Simple Games in a Complex World: a Generative Approach to the Adoption of Norms," *The University of Chicago Law Review* 64, 1224–1257.

Rousseau, Jean-Jacques. (1755). *Discourse on the origin of inequality*. Translated by Franklin Philip, Original title: *Discours sur l'origine et les fondements de l'inégalité parmi les hommes*. Oxford University Press, 1994.

Sandholm, Thomas W., and Robert H. Crites. (1995). "Multiagent Reinforcement Learning in Iterated Prisoner's Dilemma," *Biosystems* 37, 147–166.

Schelling, Thomas. (1978). *Micromotives and Macrobehavior*. Norton.

Sen, Sandip, Mahendra Sekaran, and John Hale. (1994). "Learning to Coordinate without Sharing Information," *in Proceedings of the Twelfth National Conference on Artificial Intelligence*. AAAI, Seattle, Washington.

Skyrms, Brian. (1996). *Evolution of the Social Contract*. Cambridge, UK: Cambridge University Press.

Sutton, R., and A.Barto. (1998). *Reinforcement Learning: An Introduction*. MIT Press.

Warnick, J. (2000). "Learning to Trust," Working Paper, University of Pittsburgh.

Watkins, C. (1989). "Learning from Delayed Rewards," Ph.D. Thesis, University of Cambridge, England.

Weibull, Jorgen W. (1995) *Evolutionary Game Theory*." Cambridge, MA: MIT Press.

Young, H. Peyton. (1998). *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*." Princeton, NJ: Princeton University Press.

Zaheer, A., B. McEvily, and V. Perrone. (1998) "Does trust matter? Exploring the Effects of Inter-Organizational and Interpersonal Trust on Performance," *Organization Science* 9 (2), 141–159.

Zaheer, Akbar, Bill McEvily, and Vincenzo Perrone. (1998). "The Strategic Value of Buyer-Supplier Relationships," *International Journal of Purchasing and Materials* 4 (August).