

Cooperation: The Ghost in the Machinery of Evolution

John L. Casti

SFI WORKING PAPER: 1993-12-077

SFI Working Papers contain accounts of scientific work of the author(s) and do not necessarily represent the views of the Santa Fe Institute. We accept papers intended for publication in peer-reviewed journals or proceedings volumes, but not papers that have already appeared in print. Except for papers by our external faculty, papers must be based on work done at SFI, inspired by an invited visit to or collaboration at SFI, or funded by an SFI grant.

©NOTICE: This working paper is included by permission of the contributing author(s) as a means to ensure timely distribution of the scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the author(s). It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author's copyright. These works may be reposted only with the explicit permission of the copyright holder.

www.santafe.edu



SANTA FE INSTITUTE

December 1993

Cooperation: The Ghost in the Machinery of Evolution

JOHN L. CASTI

Santa Fe Institute
Santa Fe, NM 87501, USA

To appear as Chapter 3 of *Cooperation and Conflict in General Evolutionary Processes*, J. Casti and A. Karlqvist, eds., Wiley, New York, 1994

CHAPTER 3

Cooperation: The Ghost in the Machinery of Evolution

JOHN L. CASTI

Santa Fe Institute
Santa Fe, NM 87501, USA

1. Neo-darwinism and the Problem of Competition

Neo-darwinian evolution can be described compactly by the formula

$$\text{adaptation} = \text{variation} + \text{heredity} + \text{selection}.$$

Here “variation” refers to the fact that individual organisms each have a different genetic endowment bequeathed to it by its parents, while “heredity” means that these genetic characteristics can be passed on to an organism’s own offspring. Since the pioneering work by Mendel on genetic inheritance, there has been little controversy in the evolutionary biology community about these two processes. But not so for the mysterious process of “selection,” whereby Nature bestows its favors on some and ignores others.

The traditional view of evolution handed down to us by Darwin can be encapsulated in two pithy phrases: “Nature, red in tooth and claw,” and “survival of the fittest.” These epigrams paint a picture of a cruel, hostile world indeed, one in which each organism has to battle literally to the death with every other organism in a kind of “all-against-all” war of attrition. Of course, if we open our eyes and just look around, the kind of world we see doesn’t match this picture,

Cooperation and Conflict in General Evolutionary Processes, edited by John L. Casti and Anders Karlqvist.

ISBN 0-471-59487-3 ©1994 John Wiley & Sons, Inc.

at all. Instead, what we see is a world in which there is at least as much cooperation as there is competition, and where living entities organizing themselves into social groups and communities in order to bolster their individual chances for survival.

By the tenets of a strictly darwinian view of the world, the emergence of cooperative behavior poses a real problem: If 'survival of the fittest' is indeed Nature's golden rule, how can it be that an individual's fitness is increased by a cooperative act, when by definition "cooperation" means foregoing benefits that you might have taken so that those benefits can go to another? We might term this the "Problem of Competition."

Modern sociobiologists have put together a variety of explanations for how such "altruistic" cooperative acts might arise in the natural course of evolutionary events. Basically, these mechanisms can be divided into four categories:

- *Group selection*— In his work on animal behavior, Konrad Lorenz noted that animals hardly ever engaged in behavior dangerous to their own species. Lorenz's explanation of why potentially harmful aggression in animals appeared to be confined to interspecies competition is that an individual within a group would be willing to suffer a personal loss in fitness only if that loss was more than compensated for by an increase in overall group fitness. But as the result of theoretical models, as well as ingenious alternative explanations, there is more or less universal agreement today that group selection is a pretty rare phenomenon, taking place only under very special circumstances.

- *Kin selection*—It's often observed that close relatives tend to look after each other more than they look after strangers, and the closer the relationship (e.g., identical twins vs. distant cousins), the greater the willingness to sacrifice. Experiments in the animal world involving the social insects like wasps and bees demonstrate how this kind of mechanism works, and it doesn't take too much exercise of the imagination to see how it would work for humans, as well.

- *Parental manipulation*—This is a type of enforced altruism in which a parent coerces a child to give help to another for the parent's benefit. A typical situation of this sort might arise, for instance, if a mother cat has a litter of, say, five kittens but can raise only three of them to maturity using her own resources. Then it would pay her (genetically speaking) to employ her position of authority to force some of her older offspring to devote a part of their resources to helping her raise the litter. She can do this in many ways, perhaps the most

common being a threat to withhold some of her attention from certain offspring if they refuse to help out. In Nature the strategy of parental manipulation often takes the form of cannibalism, in which the weaker members of the litter are sacrificed for the benefit of the stronger. Of course it might be argued that putting yourself on your brother's dinner plate hardly constitutes an "altruistic" act, in the sense that the term is normally used in polite conversation. But in Nature "altruism" means only an act that decreases your own fitness in order to enhance the fitness of another. So such an act of sacrifice is indeed altruistic, at least in Nature's dictionary.

At first glance it may appear that there is no real difference between parental manipulation and kin selection—they both involve the sacrifice of an individual for the benefit of another. However, there is one critical difference: In kin selection, one individual helps another because they share some genes; in parental manipulation, one person helps another for the benefit of a third party (the parent). So in parental manipulation, the fact that the two parties might share genes is incidental, although it often happens that they do. In practice, however, it may not be easy to distinguish between the two forms of altruism, and any given situation may involve both. In fact, it has been suggested that the main causal factor at work in the development of sterile castes in the social insects is parental manipulation and not kin selection. This is because when the queen sets up the nest, she chooses to make workers rather than reproductives by virtue of what she feeds her initial offspring. But this is still a matter of some controversy and the jury is out as to which of the two altruistic mechanisms is really at work in that situation.

• *Reciprocal altruism*—By far the largest share of altruistic acts, at least among humans, involve parties who are not related at all. Robert Trivers introduced the idea of reciprocal altruism to account for these sorts of sacrificial acts. In essence, the principle governing reciprocal altruism is "If you'll scratch my back, I'll scratch yours." Briefly, the claim is that individuals engage in altruistic acts because they expect that by doing so they will benefit by someone else's altruism toward them at sometime in the future. Note the very great difference here between an act of reciprocal altruism and an act of kin selection altruism. In the reciprocal case, the giver expects to see a direct return from a sacrifice; in the latter situation, the giver sees no direct reward but only the satisfaction of seeing his or her genes being given a better chance to make it into future generations.

The most convincing example of reciprocal altruism in Nature

seems to be the case of the "cleaner fish." Certain species of fish clean parasites off fish of a different species. This is a situation in which both parties gain: The cleaners get a hearty meal, while the fish being cleaned avoid the sores and diseases that would otherwise result from the parasites. The most remarkable aspect of this situation is that the cleaner fish are never eaten by those they're cleaning, even though this could easily happen. Furthermore, it's often the case that other types of fish try to imitate the cleaners, rushing in to bite big chunks off the fish being cleaned. In these cases, the big fish happily gobble up the pretenders despite the fact that the pretenders have developed high-level camouflage techniques to fool them. Since the cleaners and the cleaned have no genetic relationship at all, Trivers argues persuasively that this situation can be explained only as a case of reciprocal altruism. We'll return to a deeper consideration of reciprocal altruism later on when we consider the evolution of cooperative behavior.

In order to study how these various forms of cooperative behavior might emerge naturally in a population, theoreticians have noted that many of the encounters organisms have in their environment take the form of a game, in the sense in which mathematicians use that term. We devote the next section to an account of how one makes the transition from behaviors in Nature to games.

2. Games, Strategies and the Prisoner's Dilemma

The mathematical formulation of a game of strategy involves two or more players, each of whom has at his or her disposal a set of actions. Each play of the game consists of the players choosing an action, with the combined choice of actions leading to a payoff to each player. A strategy for each player is simply a rule by which the player selects an action for each play of the game.

The simplest type of game is the two-person game. In this situation, we have two players, call them Player I and Player II, who choose actions i and j , respectively. This results in a payoff a_{ij} to Player I and a payoff b_{ij} to Player II. If the game's payoff structure is such that

$$a_{ij} + b_{ij} = 0,$$

we have what is termed a *two-person, zero-sum game*. As it turns out, this is really the only broad class of games for which there is a satisfactory mathematical theory for how rational players will act so as to optimize their expected returns from playing the game. Here's a sketch of how this theory goes.

Suppose the components of the vector $p = (p_1, p_2, \dots, p_m)$ represent the probability that Player I chooses action $i, i = 1, 2, \dots, m$. We call p the *strategy vector* for Player I. Similarly, the vector $q = (q_1, q_2, \dots, q_n)$ is the strategy vector for Player II. Further, let r_{ij} be the payoff to Player I when he selects action i and Player II selects action j . Suppose the game is zero-sum. Then the payoff to Player II is simply $-r_{ij}$, and the *expected* return to Player I is just the sum

$$\sum_{i=1}^m \sum_{j=1}^n p_i q_j r_{ij}.$$

Player I clearly wants to choose his strategy vector p so as to maximize this quantity. Similarly, Player II tries to select the strategy vector q to minimize this sum. The question of the moment is to ask if there is a choice of strategy vectors for the two players that will result in

$$\max_p \min_q \sum_{i=1}^m \sum_{j=1}^n p_i q_j r_{ij} = \min_q \max_p \sum_{i=1}^m \sum_{j=1}^n p_i q_j r_{ij}. \quad (*)$$

In 1928, following up on work by Emile Borel, who analyzed this question in the special case when $m, n \leq 4$, and Ernst Zermelo, who conjectured that equality should hold in (*), John von Neumann proved the famous

MINIMAX THEOREM. *For a two-person, zero-sum game with payoff matrix $R = [r_{ij}]$, there exists a unique number V such that*

$$\max_p \min_q \sum_{i=1}^m \sum_{j=1}^n p_i q_j r_{ij} = V = \min_q \max_p \sum_{i=1}^m \sum_{j=1}^n p_i q_j r_{ij},$$

and strategy vectors p^* and q^* such that

$$\min_q \sum_{i=1}^m p_i^* q_j r_{ij} = V = \max_p \sum_{j=1}^n p_i q_j^* r_{ij}.$$

Here the quantity $V = V(p^*, q^*)$ is termed the value of the game.

The two-person, zero-sum game can be generalized in many directions: more than two players, nonzero-sum payoff structures, imperfect information about opponents actions, and so forth. For our purposes here, the most interesting extension is to consider payoff structures that

are not zero-sum. This means that $a_{ij} + b_{ij} \neq 0$. Such games are often termed *cooperative* or *mixed-motive games*.

It turns out that there are exactly 12 *symmetric* mixed-motive games, in which the two players can exchange roles without changing the outcome of the game. Of these, it can be shown that eight have equilibrium points corresponding to dominant strategies for both players, i.e., courses of action that remain unchanged regardless of what the other player is doing. Such games are, of course, uninteresting from a strategic point of view. So we now briefly look at each of the four remaining types of mixed-motive games.

Suppose, for simplicity, that each player has two actions at his disposal which, for reasons that will become apparent later, we'll label "C" and "D". The general form of the payoff matrix is then

		Player II	
		C	D
Player I	C	(R, R)	(S, T)
	D	(T, S)	(P, P)

where P, R, S and T are real numbers.

Each of the four "interesting" mixed-motive games has been extensively studied in the literature, and can be represented by a prototypical situation capturing the concepts peculiar to that particular type of game. These qualitatively different games can be classified by the relative magnitudes of the numbers P , R , S and T :

- *Leader*—($T > S > R > P$).
- *The Battle of the Sexes*—($S > T > R > P$).
- *Chicken*—($T > R > S > P$).
- *The Prisoner's Dilemma*—($T > R > P > S$).

Since our concern here will be with the last type of game, the Prisoner's Dilemma, we refer the reader wishing a more thorough discussion of the first three types of mixed-motive games to the volumes [1-2] cited in the References.

By far the most interesting and well-studied mixed-motive game is the famous situation involving two prisoners who are accused of a crime. Each prisoner has the option of concealing information from the police (C) or disclosing it (D). If they both conceal the information

(i.e., they cooperate), they will be acquitted with a payoff of 3 units to each. If one conceals while the other "squeals" to the police, the squealer receives the defector's reward of 5 units, while the payoff to the "martyr" is zero, reflecting his role in the obstruction of justice. Finally, if they both talk, they will each be convicted of a lesser crime, thereby receiving a payoff of only 1 unit apiece. The appropriate payoff matrix for the Prisoner's Dilemma game is then

		Prisoner II	
		C	D
Prisoner I	C	(3, 3)	(0, 5)
	D	(5, 0)	(1, 1)

(Note: Use of the symbols C and D to represent the possible actions by the players in these games is motivated by the usual interpretation of the actions in the Prisoner's Dilemma game. Here C represents "cooperating" with your pal and not confessing, whereas D signifies "defecting" and giving information to the police.)

The Prisoner's Dilemma is a real paradox. The minimax strategies intersect in the choice of mutual defection, which is also the only equilibrium point in the game. So neither prisoner has any reason to regret a minimax choice if the other also plays minimax. The minimax options are also dominant for both prisoners, since each receives a larger payoff by defecting than by cooperating when playing against either of the other player's choices. Thus, it appears to be in the best interests of each prisoner to defect—*regardless* of what the other player decides to do. But if both prisoners choose this individually rational action, the 2 units they each receive are less than the 3 units they could have obtained if they had chosen to remain silent.

The essence of the paradox in the Prisoner's Dilemma lies in the conflict between individual and collective rationality. According to individual rationality, it's better for a prisoner to defect and give information to the police. But, paradoxically, if both attempt to be "martyrs" and remain silent, they each wind up being better off. What's needed to ensure this better outcome for both players is some kind of selection principle based upon their collective interests. Perhaps the oldest and most well-known principle of this sort is the Golden Rule of Confucius: "Do unto others as you would have them do unto you."

Arguably the most important development in the theory of games

since von Neumann's proof of the Minimax Theorem was the notion of an *evolutionary stable strategy (ESS)*, introduced by Maynard Smith and Price in 1973. The basic idea behind an ESS comes from animal ecology.

Suppose we have a population of animals that interact with each other in a sequence of contests or "games." The contests involve two animals fighting for a scarce resource, say food. Each animal has a particular strategy that it always employs, and the result of a particular contest is determined by the strategies used by the two contestants. Following Maynard Smith and Price [3], we suppose there are only two pure strategies in the population, an aggressive strategy called *Hawk (H)*, and a non-aggressive strategy termed *Dove (D)*.

Payoffs in the Hawk-Dove game are measured in units of Darwinian fitness, i.e., the increase or decrease in the expected number of offspring. Suppose the victor of a contest receives V units of fitness, while an injury to either party reduces fitness by W units. Finally, there is the time and energy wasted in a "war of attrition," which costs each participant T units.

Each contest is one of three possible types: Hawk-Hawk, Hawk-Dove, or Dove-Dove. The outcomes are determined by the following rules:

- *Hawk-Hawk*: In this case we assume that the two participants each have an equal chance of winning the contest or getting injured. Since such a contest will be short, there is a negligible amount of time and energy wasted. Thus, an individual's expected payoff is $\frac{1}{2}(V - W)$. Generally, we assume that the cost of injury W is greater than the gain from victory V , so that there is some nontrivial danger associated with escalation.

- *Hawk-Dove*: The Dove will flee immediately in this type of encounter, thereby receiving a payoff of zero. The Hawk, on the other hand, then wins the contest without injury, receiving thereby a gain of V units of fitness.

- *Dove-Dove*: Again, in this case the two participants each have equal chances of winning. But the contest is likely to be a long one, with a lot of time and energy expended in posturing and maneuvering, before the issue is settled. Thus, the expected payoff to each party is $\frac{1}{2}V - T$.

The following payoff matrix summarizes the overall Hawk-Dove situation:

		Animal II	
		H	D
Animal I	H	$\frac{1}{2}(V - W)$	V
	D	0	$\frac{1}{2}V - T$

Note that here the matrix entries express the payoffs to Animal I, omitting by convention an explicit listing of the payoffs to Animal II—even though this is not a zero-sum game. But since the game is the same for both players, it's easy enough to fill-in the payoffs to Animal II if need be.

We have already assumed that any given *individual* is either a Hawk or a Dove. So the outcome of every encounter is rigidly fixed by the above rules. With this payoff structure, the problem is to determine the optimal distribution of strategies within the overall population, since it's clear that from an evolutionary point of view either a population of all Hawks or all Doves can easily be invaded by a mutant playing the other strategy. This is because too much fitness is lost due to injuries if there are too many Hawk-Hawk encounters, while too much time and energy are expended in protracted contests if the dominant encounters are between Doves. The evolutionary stable strategy (ESS) addresses the question of how the population should distribute itself in order to achieve an optimal balance between the two pure strategies.

Underlying the ESS is the following simple question: Is there a strategy such that if most members of the population adopt it, no mutant strategy can invade the population by natural selection? Any such strategy is what we call an ESS, since no mutant strategy confers greater Darwinian fitness on the individuals adopting it; consequently, the ESS cannot be invaded by any competing strategy. For this reason, an ESS is the strategy we would expect to see in Nature.

In the Hawk-Dove game, neither of the pure strategies is an ESS. To see this explicitly, a population of Hawks can be invaded by Doves because the expected payoff to the winner, $\frac{1}{2}(V - W)$, is less than zero, the payoff to a Dove battling a Hawk. Similarly, a population of Doves can be invaded by Hawks, since the payoff V to a Hawk fighting a Dove is greater than $\frac{1}{2}(V - T)$, the payoff to the winner of a Dove-Dove battle. So an ESS for this game must be a mixture of both Hawk and Dove.

To formalize the ESS concept, assume there are n pure strategies available, and that p_i represents the probability that the animal plays

pure strategy i , where, of course, $\sum_{i=1}^n p_i = 1$. Further, let $A \in R^{n \times n}$ be the payoff matrix for the game. Under these circumstances, the probability vector $p = (p_1, p_2, \dots, p_n)$ is an ESS strategy if it satisfies the following equilibrium and stability conditions:

(1) p is as good a reply against itself as any other strategy x ,

and

(2) if x is a best reply against p , then p is a better reply against x than x itself.

Mathematically, we can state these two conditions succinctly as

- I. $(p, Ap) \geq (x, Ap)$ for all $x \in S^n$ (*Equilibrium*),
- II. If $(p, Ap) = (x, Ap)$ for $x \neq p$, then $(p, Ax) > (x, Ax)$ (*Stability*).

Using these definitions, I'll leave it to the reader to verify that the mixed strategy

$$p_H = 1 + \frac{V - W}{2T + W},$$

$$p_D = \frac{W - V}{2T + W} (= 1 - p_H),$$

constitutes an ESS for the Hawk-Dove game. Inspection of this strategy offers mathematical support for the fairly obvious fact that as the reward for winning a contest outpaces the loss in fitness from injuries, it becomes increasingly more advantageous to be a Hawk.

With these game-theoretic ideas at our disposal, let's turn now to a consideration of how they have been employed to study how cooperative behavior can emerge in an evolutionary context.

3. TIT FOR TAT and the Evolution of Cooperation

The cornerstone of sociobiological reasoning is the claim that human behavior patterns, including what look on the surface like selfless acts of altruism, emerge out of genetically selfish actions. The relevance of the Prisoner's Dilemma game for sociobiology is evident. In the context of the Prisoner's Dilemma, we can translate the sociobiological thesis into the statement that the individually rational act of defection will always be preferred to the collectively rational choice of cooperation. The question then becomes: Can that situation can ever lead to a

population of cooperators? If there is no way for cooperative acts to emerge naturally out of self-interest, it's going to be very difficult for the sociobiologists to support their case.

Put in game-theoretic terms, the strategy of defecting on every play (ALL D) is an evolutionary stable strategy, since players who deviate from this strategy can never make inroads against a population of defectors. Or can they? Are there situations in which a less cutthroat course of action can ultimately get a foothold in a population of defectors? This was the question that political scientist Robert Axelrod set out to answer in one of the most intriguing psychological experiments carried out in recent years. The separate issues that Axelrod wanted to address were: (1) How can cooperation get started at all in a world of egoists? (2) Can individuals employing cooperative strategies survive better than their un-cooperative rivals? (3) Which cooperative strategies will do best, and how will they come to dominate?

Axelrod's key observation was to note that while ALL D, the strategy of always defecting, is uninvadable for a sequence of Prisoner's Dilemma interactions that is of known, fixed, and finite duration, there may be alternative ESS strategies if the number of interactions is not known by both parties in advance. So after having played a round of the Prisoner's Dilemma, if there is a nonzero chance that the game might continue for another round, then perhaps there is a nice strategy that is also ESS. Here by "nice" we mean a strategy that would not be the first to defect.

To put this speculation to the test, Axelrod invited a number of psychologists, mathematicians, political scientists, and computer experts to participate in a contest pitting different strategies against one another in a computer tournament [4]. The idea was for each participant to supply what he or she considered to be the best strategy for playing a sequence of Prisoner's Dilemma interactions, with the different strategies then competing against each other in a round-robin tournament. Fourteen competitors sent in strategies, which were in the form of computer programs. The ground rules allowed the programs to make use of any whatsoever information about past plays of the game. Furthermore, the programs didn't have to be deterministic, but were allowed to arrive at their choice of what to do by some kind of randomizing device if the player so desired. The only condition imposed was that the program ultimately come to a definite decision for each round of play: C or D. In addition to the submitted strategies, Axelrod also included the strategy RANDOM, which decided whether to cooperate or defect by, in effect, flipping a coin. In the tournament itself, every

program was made to engage every other (including a clone of itself) two hundred times, the entire experiment being carried out five times in order to smooth out statistical fluctuations in the random-number generator used for the nondeterministic strategies.

Using the payoff values $R = 3, S = 0, T = 5$ and $P = 1$, the strategy that won the tournament turned out to be the simplest. This was the three-line program describing the strategy TIT FOR TAT. It was offered by game theorist Anatol Rapoport, and consisted of the two rules: (1) cooperate on the first play; (2) thereafter, do whatever your opponent did on the previous round. That such a simple, straightforward strategy could prevail against so many seemingly far more complex and sophisticated rules for action seems nothing short of miraculous. The central lesson of this tournament was that in order for a strategy to succeed, it should be both nice and forgiving, i.e., it should be willing both to initiate and to reciprocate cooperation. Following a detailed analysis of the tournament, Axelrod decided to hold a second tournament to see if the lessons learned the first time around could be put into practice to develop even more effective cooperative strategies than TIT FOR TAT.

As prelude to the second tournament, Axelrod packaged up all the information and results from the first tournament and sent it to the various participants, asking them to submit revised strategies. He also opened up the tournament to outsiders by taking out ads in computer magazines, hoping to attract some programming fanatics who might take the time to devise truly ingenious strategies. Altogether Axelrod received sixty-two entries from around the world, including one from John Maynard Smith, mentioned earlier as the developer of the ideas of the evolutionary game and the ESS. The winner? Again it was Rapoport with TIT FOR TAT!

Even against this supposedly much stronger field, Rapoport's game-theoretic version of the Golden Rule was the hands-down winner. The general lesson that emerged from the second tournament was that not only is it important to be nice and forgiving, but it's also important to be both provokable and recognizable, i.e., you should get mad at defectors and retaliate quickly but without being vindictive, and you should be straightforward, avoiding the impression of being too complex. After extensive study of the results, Axelrod summarized the success of TIT FOR TAT in the following way:

TIT FOR TAT won the tournaments not by beating the other player but by eliciting behavior from the other player that allowed both to do well. . . . so in a non-zero sum world, you do not have to do better than the other player

to do well for yourself. This is especially true when you are interacting with many different players. . . . The other's success is virtually a pre-requisite for doing well yourself.

So what are the implications of these results for sociobiology and evolution of cooperative behavior?

If we think of the total points amassed by a strategy during the course of the tournament as its "fitness," interpret "fitness" to mean "the number of progeny in the next generation," and let "next generation" mean "next tournament," then what happens is that each tournament's results determine the environment for the next tournament. The fittest strategies then become more heavily represented in the population of strategies fighting it out in the next tournament. This interpretation leads to a kind of ecological adaptation without evolution, since no *new* species come into existence. Sociobiologists can take heart in this sort of interpretation of Axelrod's experiments because it shows that it's possible for phenotypically altruistic (cooperative) behavior to emerge out of individually selfish motives. It's important to emphasize here, though, that these results say nothing about the actual causal factors at work generating the individual motives. They could be genetic, as hard-core sociobiologists would love to argue, but there is nothing in Axelrod's work to say that they are. Nevertheless, the experiments do offer some support to the sociobiological explanation of cooperative behavior by means of reciprocal altruism.

Following his work on the evolution of cooperation, Axelrod carried out another set of experiments that also give succor to the sociobiologist's claim for an evolutionary development of standards of behavior, i.e., cultural norms. The basic idea is to use a souped-up version of the Prisoner's Dilemma in which the players had the choice not only of cooperation or defection, but also of punishing a defection or letting it pass. Player in the Norms Game are characterized by two qualities: his Boldness (B), which measures the risk they are willing to run in defecting; and Vengefulness (V), a measure of their inclination to punish defection. Strategies were assigned randomly to twenty players, with the first round of play lasting until each player had four opportunities to defect. At the end of the first generation, a strategy was given one offspring if its score was near average, two offspring if its score was at least one standard deviation above the mean, and no offspring if its score was more than one standard deviation below the mean. Furthermore, Axelrod also allowed for the emergence of new strategies through a process of mutation in such a way that about one new strategy emerged in each generation.

The results of the simulation showed that with enough time, all populations eventually converge to the collapse of the norm, i.e., V approaching zero. The reason appears to be that the players lack sufficient incentive to punish the defectors, i.e., nobody wants to play sheriff. As one way of enforcing the norm, Axelrod suggests a *metanorm*: direct vengeance not only against those who defect, but also against those who refuse to punish them. This is the kind of procedure we see in some totalitarian countries, where when a citizen is accused by the authorities of some real or imagined ideological transgression, others are called upon to pile their own denunciations onto the back of the hapless offender.

While these results are still in the preliminary stage, the Evolution of Cooperation Game and the Norms Game both provide some theoretical evidence in support of the idea that cooperative social behavior can emerge as the result of evolutionary processes involving individually selfish agents.

4. *Beyond TIT FOR TAT*

The overwhelming success of TIT FOR TAT in Axelrod's experiments raised a number of intriguing questions surrounding the stability of this particular strategy. For the most part, these questions can be divided into two qualitatively different types:

- A. Asking about whether there are strategies for playing the game *under the original rules* that are even better than TIT FOR TAT.
- B. Asking about the degree to which TIT FOR TAT retains its preeminent position if we change the ground rules of the game by, for example, admitting uncertainty into the identification of what the other player has actually done, changing the actual numbers used in the payoffs, or allowing the players to have more than two courses of action.

Here we examine studies that have looked at both of these questions.

In an attempt to discover strategies that perform even better than TIT FOR TAT under the original Axelrod rules of the game, Karl Sigmund and Martin Nowak reported an improved strategy for playing the game in [5]. This was the strategy they termed PAVLOV, which can be described very simply as "win-stay, lose-shift." What this means is that a PAVLOV player cooperates if and only if both players chose the

same action on the previous round; otherwise, the player defects. The terminology "Pavlov" comes from the fact that this strategy embodies an almost conditioned reflex-like response to the payoff: it repeats its previous move if it was rewarded with the *R* or *T* payoff, while it changes its action if it was punished by receiving only the *S* or *P* payoff.

The reasoning leading up to the discovery of PAVLOV's superiority to TIT FOR TAT starts by observing that TIT FOR TAT can suffer from random perturbations in two ways. First of all, a TIT FOR TAT population can be degraded by the random appearance of unconditional cooperators (ALL C), whose appearance then allows the growth of exploiters. This problem arises because TIT FOR TAT is not an evolutionary stable strategy. (Note: It can be shown, in fact, that *no* pure strategy can be ESS in the iterated Prisoner's Dilemma game.) And, secondly, an occasional mistake between two TIT FOR TAT players can result in a prolonged run of mutual backbiting, as the players get "out of synch" with each other and begin engaging in a sequence of mutual defections until another mistake occurs that neutralizes the previous misunderstanding.

PAVLOV has two major advantages over TIT FOR TAT: (1) a misunderstanding between two PAVLOV players causes only a single round of mutual defection followed by a return to joint cooperation; (2) PAVLOV cannot be invaded by unconditional cooperators (ALL C), thereby allowing exploiters to gain a foothold in the population. To the contrary, PAVLOV has no qualms about fleecing a sucker, once it has discovered (after a misunderstanding) that it need not fear any retaliation. So in this sense, cooperation based on PAVLOV is a safer bet than cooperation arising from TIT FOR TAT, since such cooperation is less prone to exploitation by the intermediary of an ALL C invasion.

In an extensive series of computer simulations, the work of Sigmund and Nowak uncovered the fact that over relatively long time horizons (number of generations in the population), there is a pronounced tendency for populations to move toward generally cooperative behavior. Moreover, as the time horizon gets longer, the cooperative behavior is exclusively composed of two strategies: PAVLOV and "Generous TIT FOR TAT," a strategy that cooperates after the opponent cooperates, but also cooperates with a certain probability after a defection, as well. As an example, after one million rounds, approximately 60 percent of the population consisted of cooperative strategies, with PAVLOV accounting for about two-thirds of the cooperators and Generous TIT FOR TAT another 5 percent or so. And when the number of generations reaches ten million, nearly 90 percent of the population

is cooperators, with 80 percent of the population playing PAVLOV and 10 percent Generous TIT FOR TAT. Readers wishing more details of these fascinating experiments should consult the original article [5].

We have just noted the problems that can arise in “no frills” TIT FOR TAT when there is a nonzero chance that the players will misunderstand the action taken by their opponent, thereby opening up the possibility for a long sequence of mutual defections. If a misunderstanding of this type occurs, the two players will alternate between three modes of behavior: (a) First, they will play the ordinary TIT FOR TAT actions (C, C) , but when a misunderstanding takes place, (b) they will shift to alternating (C, D) and (D, C) . Finally, the third possibility is (c) sequences of mutual defections (D, D) . The average likelihoods of the three modes (a), (b) and (c) are $1/4$, $1/2$ and $1/4$, respectively, leading to an average payoff of $9/4$.

In an attempt to develop strategies that are more resistant to misunderstanding, Kristian Lindgren employed genetic algorithms in [6] as a way of “evolving” strategies for playing the iterated Prisoner’s Dilemma game. What he discovered is that there exist deterministic (i.e., pure) strategies that are robust against noise (misunderstandings), unexploitable (in the sense that they cannot be invaded) and reach an average score of nearly 3—substantially better than ordinary TIT FOR TAT.

In Lindgren’s experiments, he coded different finite-memory strategies using the following binary scheme. An m -length history is given by the vector

$$h_m = (a_{m-1}, \dots, a_1, a_0),$$

where each a_i is either 0 or 1. Thus, since a deterministic strategy of memory m associates an action C or D to each history h_m , we can specify the strategy by a binary sequence

$$S = (A_0, A_1, \dots, A_{n-1}).$$

This strategy is like the “genetic code” for the strategy that chooses action A_k when history h_k turns up. So the length of such a “genome” is $n = 2^m$.

As an example of this scheme for strategies of memory length $m = 1$, the histories are labeled 0 or 1, corresponding to the opponent’s having defected or cooperated, respectively. Thus, the four memory-1 strategies are: $S_1 = (00)$, $S_2 = (01)$, $S_3 = (10)$ and $S_4 = (11)$. The strategy S_1 always defects (ALL D), S_2 cooperates only if history 1

turns up (TIT FOR TAT), S_3 does the opposite of S_2 (ANTI-TIT FOR TAT), and S_4 always cooperates (ALL C).

The graph in Figure 1 shows a simulation of thirty thousand generations, starting with a population equally distributed among the four memory-1 strategies. Here we see that the system reaches periods of stasis, where there is coexistence between TIT FOR TAT and ANTI-TIT FOR TAT. This stasis is then punctuated by a number of memory-2 strategies, and after a period of unstable behavior, the system slowly stabilizes when the strategy $S = (1001)$ increases in the population. (Note: The perceptive reader will recognize this strategy as PAVLOV: it cooperates only if both players took the same action in the previous round.) Another period of stasis then ensues, before being broken up by memory-3 and finally memory-4 strategies.

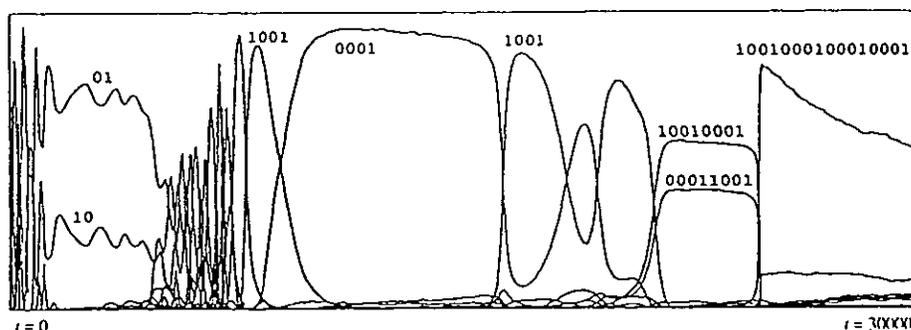


Figure 1. Evolution of Population Strategies

Figure 2 shows the average score s and number of genotypes n for the population generated in the simulation shown in Figure 1. Here we see that when the exploiting memory-2 strategy dominates, the average score drops down close to 1. But the final stasis populated by the ESS memory-4 strategy reaches a score of 2.91, close the maximum possible score of 3 that's attained by the best strategies in a noise-free environment.

There are many conclusions we can draw from Lindgren's experiments, one of the most important being that when there is a nonzero chance for a misunderstanding, there exists an unexploitable strategy that is cooperative. But the simulations suggest that the minimal memory needed for such a strategy is 4, i.e., the player should take account of both parties actions for the previous two rounds. Basically, defecting twice in response to a single defection is a strategy that cannot be exploited by intruders. Lindgren offers many other insights into

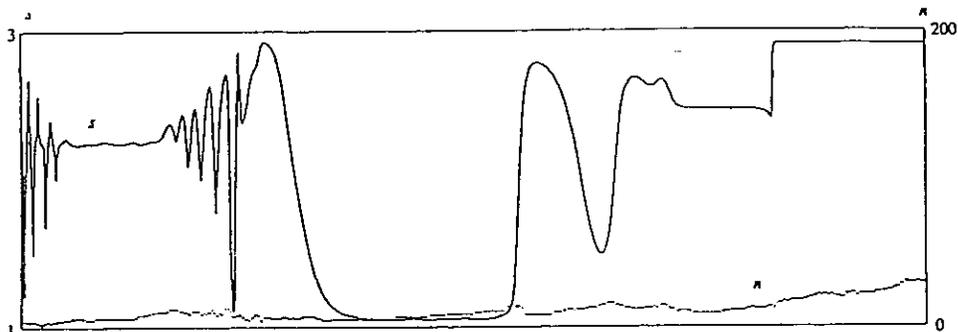


Figure 2. Average Scores and Number of Genotypes in the Population

the structure of "good" strategies for the Prisoner's Dilemma that we have no space to recount here. So again, the reader is referred to the source [6] for these details.

For the most part, the foregoing results have with minor exceptions adhered pretty much to Axelrod's original ground rules for playing the iterated Prisoner's Dilemma. Now we consider briefly a cluster of experiments involving major changes to the rules of the game, starting with work in which the players retain a memory of which players they have faced in the past and can choose whether or not to play the game again with particular players.

In [7], Stanley, Ashlock and Tesfatsion set up an iterated Prisoner's Dilemma ecology in which the agents are represented by finite automata. These agents interact with each other in the usual Prisoner's Dilemma way, with the notable exception that players choose and refuse potential game partners on the basis of continually updated expected payoffs. This choice mechanism gives players a way to increase their chance of encountering other cooperative players, as well as offering protection against defectors without having to become defectors themselves.

Under the choice/refusal option, ostracism of defectors occurs endogenously as an increasing number of players refuse their game offers. However, choice and refusal also permits clever "ripoff" players to home-in quickly on exploitable players and to establish parasitic partnerships. Simulation studies indicate that the emergence of cooperation tends to be accelerated by the choice/refusal option, but that the rate of emergence is a highly nonlinear function of the parameters characterizing the choice/refusal mechanism. Unfortunately, to chronicle the details of this mechanism and the resulting simulation studies

would involve a bit more space than this chapter can accommodate, so we again refer the reader to the paper [7] for the full story.

In the survey paper [8], Axelrod and Douglas Dion have summarized a number of studies in which the original iterated Prisoner's Dilemma game has been modified in one direction or another. Here is a brief account of their findings:

- *Interactions*—The original game specifies interactions between pairs of players. In some situations, though, it makes more sense to assume that interactions are among more than two players. In these games, players make a choice that they then play with all other players simultaneously. This leads to what's termed the *n-person iterated Prisoner's Dilemma*.

It has been shown that increasing the number of players makes cooperation more difficult. In both evolutionary and non-evolutionary settings cooperation can be part of an equilibrium of the game—but only if the 'shadow of the future' is long and/or the number of cooperators is large.

- *Choices*—If the players are allowed more than the two choices of cooperation or defection, the outcome of the game can also be significantly altered. For example, we have already seen the effect of allowing players the option of playing or not playing a particular partner. In a related study, it was shown that the stable equilibria of a nonsimultaneous move game are either ALL D or a form of conditional cooperation.

- *Payoffs*—Experiments with changing the payoff matrix also lead to interesting departures from the standard Axelrod results. One such result involved having the players play both the iterated Prisoner's Dilemma and Chicken. It turned out that under these conditions, TIT FOR TAT does well in this situation—but not as well in Chicken as the strategy PERMANENT RETALIATION, under which the player cooperates until the other player defects and then defects for the remainder of the game.

- *Noise*—In the work of Sigmund and Nowak, as well as in the studies of Lindgren, we have already seen the rather dramatic effect of adding faulty transmission of strategy choices (i.e., noise) into the picture. In one study, it was shown that in the presence of any amount of noise, two TIT FOR TAT players will in the long run achieve the same average payoffs as two interacting RANDOM players. Nevertheless, if the amount of noise is small, TIT FOR TAT can still perform well. For instance, when the noise level is only 1 percent, TIT FOR TAT still turned out to be the best strategy in one tournament. But when the

noise level went up to ten percent, TIT FOR TAT dropped down to sixth place among 21 strategies.

But there are many ways that noise can be introduced into the Prisoner's Dilemma setup, and it's difficult to make any general statements about when TIT FOR TAT can be "souped up" to still win out in these noisy environments. About the best that can be said is that if the amount of noise is small, unilateral generosity seems to be the best response. But for larger amounts of noise, there is a trade-off between the unnecessary conflict that can be avoided by being generous and the exploitation that generosity invites.

- *Shadow of the Future*—The expected length of the game plays a crucial role in strategic thinking in the Prisoner's Dilemma situation. So when the probability of the game continuing for one more round w is modified, so are the winning strategies in the computer tournaments we've been talking about here. The work by Stanley, Ashlock and Tesfatsion mentioned earlier involving the changing likelihood of encountering a given player again is an example of this sort of behavior-dependent play.

Other experiments have focused on changing the shadow of the future by using a stochastic payoff matrix. For instance, in some moves of this kind of game, the players face a normal Prisoner's Dilemma, while in other moves cooperation is the dominant choice. If the probability of surviving the current move is a function of the total payoff received to that time, an egoist may cooperate on the Prisoner's Dilemma moves of the game to ensure the presence of a "congenial" partner in other stages of play.

In general, it seems to be the case that extending the shadow of the future by whatever means tends to enhance the likelihood of cooperative behavior substantially. So as a guiding principle, if you want to promote cooperation, change the rules of the game so as to increase the probability that two players will meet again in future rounds of play.

- *Population Dynamics*—It turns out that if a population can be invaded by multiple mutants (i.e., those playing strategies different from the population as a whole), then there is no single strategy the population can adopt that is ESS. In other words, if there is some chance for future interaction, then for each strategy X , there is always a set of strategies $Z(X)$ such that X is not an ESS against Z . Moreover, no pure strategy whose behavior is determined solely by the history of the game is an ESS if the future is important enough. More specifically, if the shadow of the future $w > \min\{(T-R)/(T-P), (P-S)/(R-S)\}$,

then there can be no ESS for the iterated Prisoner's Dilemma.

This kind of result places even greater emphasis on the analysis of sets of strategies, since they show that the evolutionary pathway leading to cooperation can only be understood by looking at the set of possible competitors to a population playing a given strategy. The results considered earlier by Lindgren using genetic algorithms is one approach to studying this kind of issue. But there are others. What one tends to find is that strategies similar to TIT FOR TAT often emerge within a few dozen generations. But not always. Occasionally, strategies like WIN STAY, LOSE CHANGE appear that outperform TIT FOR TAT in a given environment appear as well. In general, about all one can say is that the situation is an unstable one, in the sense that what's best against one population can be worst against another. There's just no universal strategy that can do well under all circumstances.

- *Population Structure*—In his original studies, Axelrod found that cooperation can invade a population of egoists *if* the cooperative strategies invade in clusters. This raises the question of whether cooperation can evolve without such a population structure.

Part of the answer comes from considering an ESS. Since no strategy is ESS if the shadow of the future is important enough, it is possible for cooperation to emerge without clustering if the "right" combination of strategies is present. For example, if we have what's termed "assortative mating and meeting," in which players pick only fellow cooperators and shun defectors, then cooperation can emerge without clustering. In another direction, it has been shown that cooperation can get started without clustering if the payoffs to cooperative acts depend on the frequency with which those acts are performed.

This is just the briefest of reviews of the many ways in which the Prisoner's Dilemma game can be extended. For full details, together with references to the literature on these and many other extensions of the game, we again refer the reader to the survey article [8].

5. *Testing the Theory in the Real World*

The proof of any pudding is in the eating. And the various theories of cooperation embodied in the iterated Prisoner's Dilemma game is no exception: if we're going to take seriously the idea that cooperation can emerge in a world of egoists, we have to validate the theory with examples of such behavior in the real world of animals and humans.

This section discusses several studies of this sort. So on to the real world of cabbages and kings, or in this case, the world of spiders.

One of the most interesting tests of the ESS in nature was carried out by Susan Riechert, who studied the behavior of the common grass spider *A. aperta* in settling territorial disputes [9]. Riechert studied these spiders in two habitats that differed greatly in the availability of suitable locations for building webs—a desert grassland in New Mexico and a desert riparian area consisting of a woodland bordering a stream in Arizona, a region offering many more favorable locations for webs. While there is no room here to go into the details of how Riechert determined the actions available to the spider and assigned the various payoffs, her final conclusions are worth pondering. She discovered that the contest behavior for web sites in the riparian regions deviated substantially from the ESS predicted by the game-theoretic model. In particular, contrary to theory, a riparian spider does not withdraw from occupied territory when it encounters the owner of the web. Rather, they engage in a dispute that escalates to potentially injurious behavior. On the other hand, the behavior of grassland spiders does follow the ESS as predicted by the theory, with the time and energy they expend in fights varying with their probability of emerging victorious.

So, while the riparian spiders are less aggressive than their desert grassland cousins, just as ESS theory predicts, they are still somewhat more aggressive than they should be. This leads us to ask: Why does the behavior in these territorial disputes differ from the ESS for riparian spiders and not for their grassland cousins? Riechert gives an answer that will gladden the heart of any sociobiologist. She states:

If one assumes that the model is correct—that it has taken into account all the important parameters and includes all possible set of strategies—then there must be some biological explanation for the observed deviation. . . . One possibility is that the release from strong competition is a recent event and that there just has not been sufficient time for natural selection to operate on the behavioral traits to complete the expected change. . . . Finally, a major change in the wiring of *A. aperta*'s nervous system might be required to achieve the new ESS, and such a mutant may simply not have arisen yet.

Another example of the ESS in nature was reported by Maynard Smith in [10]. This involves six stickleback fish who were fed at both ends of a tank, the rate of feeding being twice as great at one end of the tank as at the other. The ESS is when no fish can gain by moving from one end of the tank to the other. This occurs when there are four fish at the end with the higher food supply and two fish at the other end.

In a set of experiments in which the fish are all dropped into the tank at random locations, Milinsky found that the fish distribute themselves statistically in accordance with the ESS. Individual fish, on the other hand, continue to move from one end of the tank to the other. This is to be expected, since in nature the relative attractiveness of patches of food will not remain constant over time. And when the experimenters switched the rate of supply between the two ends during the course of a given experiment, the fish again distributed themselves statistically as predicted by the ESS.

The spawning of sea bass provides a final example in nature of the how cooperative behavior can emerge in a world of egoists. These fish have both male and female sexual organs, form pairs, and can be said to take turns at being the high investment (laying eggs) and the low investment (providing sperm to fertilize the eggs) partner. Up to ten spawnings occur in a single day, with only a few eggs provided each time. Pairs of fish tend to break up if sex roles are not divided evenly. From studies in the wild, it appears that this type of cooperative behavior got started at a time when the sea bass was scarce as a species, leading to a high level of inbreeding. Such inbreeding implies a relatedness in the pairs, which would have initially promoted cooperation without the need for further relatedness. A more complete account of this phenomenon is found in the book [4]. Now let's turn our attention from general cooperation and the ESS to the specific case of real-world laboratory experiments involving the iterated Prisoner's Dilemma.

By now, there have been well over a thousand experiments based on the Prisoner's Dilemma and related games. So it would be futile to try to discuss even a small fraction of them here. We refer the reader to the excellent review articles on the topic [11], contenting ourselves with only a brief summary of a few of the more interesting cases.

Without a doubt, the most striking general finding of the Prisoner's Dilemma experiments is the *DD lock-in effect*. When the game is repeated many times, there is a pronounced tendency for long runs of D choices by both players to take place. Of course, even though the ALL D strategy is dominant—and even minimax—for this game, both players are better off if they both chose C instead of D. Strangely, though, in their influential 1957 book *Games and Decisions*, Luce and Raiffa stated that “We feel that in most cases an unarticulated collusion between the players will develop.” That this prediction has *not* been borne out in hundreds of experiments is a fact of which many game theorists still remain largely ignorant.

We can get some insight into the DD lock-in effect from the work of Rapoport and Chamah [12]. They found that three phases typically occur in a long series of plays of the Prisoner's Dilemma game: Initially, the proportion of C choices is typically slightly higher than 1/2. But this is followed by a rapid decline in cooperation, the so-called "sobering period." This period usually lasts for about 30 rounds, after which cooperative choices begin to increase slowly in frequency (the "recovery period"). The frequency of C choices generally then reaches about 60 percent by round 300.

The moderately high fraction of initial C choices is often attributed to an initial reservoir of goodwill or simply to a naïve lack of comprehension on the part of the subjects of the strategic structure of the game. The sobering period then reflects a decline in trust and trustworthiness, an increase in competitiveness, or perhaps just a dawning understanding of the structure of the payoff matrix. The recovery period can be thought of then as reflective of the slow and imperfect growth of the "unarticulated collusion" mentioned by Luce and Raiffa.

Another interesting outcome of experiments with human subjects in Prisoner's Dilemma situations involves programmed strategies. Experiments in this area pit subjects against another "player," who in reality is just a confederate of the experimenter programmed in advance to make predetermined sequences of choices.

Certainly the simplest of these kinds of experiments are those in which the confederate chooses either C or D on every trial. In the Prisoner's Dilemma situations, an ALL C strategy elicits much higher frequencies of C choices from subjects than ALL D. Of course, this is in line with expectations, since self-defense considerations force a subject to choose D against an ALL D player in order to avoid the sucker's payoff. What is surprising, however, is the large proportion of subjects who seize the opportunity to exploit an ALL C player by choosing to defect. And, in fact, the tendency for many subjects to exploit "pacifist strategies" has been confirmed in mixed-motive games other than the Prisoner's Dilemma.

One of the most striking and unexpected findings to emerge from early experiments with games like the Prisoner's Dilemma is the apparent tendency for females to exhibit much lower frequencies of C choices than males. In light of traditional sex roles, one might have expected just the opposite. As an illustration of this effect, in games involving two male players, the frequency of C choices was 59 percent. In the same study, it was only 34 percent when the two players were both female.

There are now over a hundred experiments on record showing this pronounced sex difference in mixed-motive games. But the effect is still shrouded in mystery. It has been claimed that the sex difference is an artifact, and there is some evidence that the effect disappears when the experimenter is a female. At present, no one seems to have been able to come up with a convincing explanation of the effect, but it's not unreasonable to assume that in Western industrial cultures, males are brought up to behave more boldly and with a greater willingness to take risks than females. This would then account for their greater likelihood of choosing the non-minimax C choice in the Prisoner's Dilemma in spite of the danger of receiving the sucker's payoff.

While there are many other effects of the above sort scattered throughout the experimental gaming literature, space constraints demand closing our discussion here. Again, the interested reader is urged to consult the literature on these matters cited in the chapter References.

References

1. Colman, A. *Game Theory and Experimental Games*. Oxford: Pergamon Press, 1982.
2. Casti, J. *Reality Rules-II: Picturing the World in Mathematics*. New York: Wiley, 1992.
3. Maynard Smith, J. and G. Price. "The Logic of Animal Conflicts." *Nature*, 246 (1973), 15-18.
4. Axelrod, R. *The Evolution of Cooperation*. New York: Basic Books, 1984.
5. Nowak, M. and K. Sigmund. "A Strategy of Win-Stay, Lose-Shift that Outperforms Tit-for-Tat in the Prisoner's Dilemma Game." *Nature*, 364 (1993), 56-58.
6. Lindgren, K. "Evolutionary Phenomena in Simple Dynamics," in *Artificial Life-II*, C. Langton et al, eds., Redwood City, CA: Addison-Wesley, 1992, pp. 295-312.
7. Stanley, A., D. Ashlock and L. Tesfatsion. "Iterated Prisoner's Dilemma with Choice and Refusal," preprint, June 1992.
8. Axelrod, R. and D. Dion. "The Further Evolution of Cooperation." *Science*, 242 (1988), 1385-1390.
9. Riechert, S. "Spider Fights as a Test of Evolutionary Game Theory." *American Scientist*, 74 (1986), 604-610.

10. Maynard Smith, J. *Evolution and the Theory of Games*. Cambridge: Cambridge University Press, 1982.

11. Schlenker, B. and T. Bonoma. "Fun and Games: The Validity of Games for the Study of Conflict." *Journal of Conflict Resolution*, 22 (1978), 7-38.

12. Rapoport, A. and A. Chammah. *Prisoner's Dilemma: A Study in Conflict and Cooperation*. Ann Arbor: University of Michigan Press, 1965.