Cubo A Mathematical Journal
Vol.05/Nº03 - OCTOBER 2003

N-Person Prisoners' Dilemmas

Miklos N. Szilagyi
Department of Electrical & Computer Engineering,
University of Arizona, Tucson, AZ 85721, USA

ABSTRACT. This review is an attempt to systematically present the problem of various N-person Prisoners' Dilemma games and some of their possible solutions. Thirteen characteristics of the game are discussed. The role of payoff curves, personalities, and neighborhood is investigated. We report computer simulation experiments based on our new agent-based simulation tool to model social situations for the case of large numbers of not necessarily rational decision-makers. Our model has a number of user-defined parameters such as the size and shape of the simulation environment, the definition of neighborhood, the payoff (reward/penalty) functions, the learning rules, the agents' personalities, and the initial conditions. We have performed a series of simulation experiments with various combinations of these parameters. Investigations of realistic (non-dyadic) situations in which agents have various personalities show interesting new results. For the case of Pavlovian agents the game has two non-trivial but remarkably regular solutions. For a wide range of initial conditions, the number of cooperators oscillates around a relatively small value. When the initial aggregate cooperation probability is above a certain value, the solutions tend to reach well-defined constant values that are dependent on the initial values. For other types of agents the solutions show interesting chaos-like behavior. Examples of non-uniform distributions and mixed personalities are also presented. All solutions strongly depend on the choice of parameter values. The paper provides some insight into the conditions of decentralized cooperation in spatially distributed populations of agents.

Keywords: agent-based simulation, cooperation, Prisoners' Dilemma.

1 Introduction

Prisoners' Dilemma is usually defined between two players (Rapoport and Chammah, 1965) and within game theory that assumes that the players act rationally. Realistic investigations of collective behavior, however, require a multi-person model of the game (Schelling, 1973) that serves as a mathematical formulation of what is wrong with human society (Hardin, 1968). This topic has great practical importance because its study may lead to a better understanding of the factors stimulating or inhibiting cooperative behavior within social systems. It recapitulates characteristics fundamental to almost every social intercourse.

Various aspects of the multi-person Prisoners' Dilemma have been investigated in the literature (Bixenstine et al. 1966, Weil 1966, Rapoport 1970, Kelley and Grzelak 1972, Hamburger 1973, Anderson 1974, Dawes 1975 and 1980, Bonacich et al. 1976, Goehring and Kahan 1976, Fox and Guyer 1978, Heckathorn 1988, Liebrand et al. 1992, Huberman and Glance 1993, Okada 1993, Komorita and Parks 1994, Schulz et al. 1994, Schroeder 1995, Nishihara 1997, Hegselmann 1998, Szilagyi 2001, Szidarovszky and Szilagyi 2002, Szilagyi and Szilagyi 2002) but there is still no consensus about its real meaning.

The participants of a Prisoners' Dilemma game may be persons, collectives of persons, organizations, any other decision-making entities, or even computer programs. They are usually called *agents*. The individual agents may cooperate with each other for the collective interest of their society or may defect, i. e., pursue their selfish interests. Their decisions to cooperate or defect will accumulate over time to produce a resulting collective order that will determine the success or failure of the society.

Formal models have been proposed to simulate collective phenomena (Oliver, 1993). Some of the models include computer simulation. Feinberg and Johnson (1990) simulated the effects of alternative strategies on achieving consensus for action. A computer simulation of temporary gatherings was presented by McPhail et al. (1992). Glance and Huberman (1993) used a thermodynamical model to investigate outbreaks of cooperation in a social system. Epstein and Axtell (1996) demonstrated that it is possible to build complex artificial societies based on simple participating agents.

Thousands of papers have been published about the two-agent iterated Prisoners' Dilemma game (Axelrod 1984, Marinoff 1992, Macy 1995, Messick and Liebrand 1995). The interest in investigating various strategies for pair-wise interactions in multi-agent Prisoners' Dilemma computer tournaments is amazing because – as Rapoport (1994) rightly noted – these "tournaments demonstrated neither evolution <nor> learning because nothing evolved and nothing was learned" in the succession of two-person games. Nevertheless, the obsession with these tournaments continues (Hoffmann, 2000). Even papers that claim the simulation of multi-agent games are usually based on dyadic interactions between the agents. A stochastic learning model was developed by Macy (1991) to explain critical states where threshold effects may cause shifting the system of agents from a defective equilibrium to a cooperative one. Nowak and May (1992) and Lloyd (1995) wrote simple computer programs that demonstrate the dynamics of deterministic social behavior based on pair-wise interactions between the participants.

Akimov and Soutchansky (1994) presented a multi-agent simulation (not a succession of two-person games) but their experiment was limited to six agents. Our own simulation tool (Szilagyi and Szilagyi 2000) was designed to simulate social dilemmas with a wide range of user-defined parameters. It is suitable for an unlimited number of agents with various personalities. We were able to perform interesting non-trivial experiments with this tool (Szilagyi 2001, Szilagyi and Szilagyi 2002).

This paper is an attempt to systematically present the problem of the N-person Prisoners' Dilemma and some of its possible solutions.

2 N-person dilemmas

The N-person Prisoners' Dilemma considers a situation when each of N agents has a choice between two actions: cooperating with each other for the "common good" or defecting (following their selfish short-term interests). As a result of its choice, each agent receives a reward or punishment (payoff) that is dependent on its choice as well as everybody else's (Figure 1).

The dilemma can be formulated by the following two statements (Dawes, 1980):

(1) Regardless of what the other agents do, each agent receives a higher payoff for defecting behavior than for cooperating behavior.

(2) All agents receive a lower payoff if all defect than if all cooperate.

If m of the N agents are cooperating and C(m) and D(m) are the payoffs to a cooperator and a defector, respectively, then the above conditions can be expressed as

$$D(m) > C(m+1) \tag{1}$$

and

$$C(N) > D(0) \tag{2}$$

C(0) and D(N) are undefined; therefore, the value of m is between 0 and N-1 in Eq. (1). The game has N+1 distinguishable outcomes: 0, 1, 2,...,N-1, N participants may choose cooperation.

At a first glance, it looks a well-defined problem. However, at least the following questions arise immediately:

1. Are the choices and actions of the agents simultaneous or distributed in time?
2. Can individual agents see and adapt to the actions of others?
3. Can they form coalitions?
4. What are the agents' goals in the game: to maximize their payoffs, to win a competition, to do better than their neighbors, to behave like the majority, or any other goal?
5. Is it a one-shot game or an iterated one? If it is an iterated game, how will the next action be determined?
6. What are the personalities of the agents? (Surely, different people react quite differently to the same conditions.) Can they change their personalities?
7. Can an agent refuse participation in the game?
8. Are the payoff curves the same for all agents?
9. What are the payoff curves?
10. How is the total payoff to all agents related to the number of cooperators?
11. How are the agents distributed in space? Can they move?
12. Do the agents interact with everyone else or just with their neighbors?
13. How is neighborhood defined?

With so many open questions it is obviously quite difficult to create a general classification scheme for the N-person Prisoners' Dilemma and there is a great variety of possible games. It is, in fact, a whole family of quite different games. Even in the case of a *uniform* game the number of possible variations is infinitely large because of the infinite variety of the payoff curves. In a uniform game the payoff curves are the same for all agents, they are monotonically increasing functions of the number of cooperators, and there is some minimum number of cooperators that can gain by their cooperative choice (Schelling, 1973). It is, however, desirable to investigate at least the most characteristic cases because each possible variation may represent an important social situation.

Let us first take a closer look at each of the thirteen questions listed above.

1) There is a huge difference between simultaneous actions and actions distributed in time. In the first case all agents see the same environment at the moment of their simultaneous action. In most social settings, however, agents act at different and non-correlated times. Therefore, each agent sees a slightly different world than another agent who acts at a slightly different time (Huberman and Glance, 1993). Simulation of this case is a more sophisticated task than that of the previous case.

2) Even if the agents' actions are distributed in time, they may or may not have information about the actions of others. You may look out of the window and see how many cars are on the road before deciding if you are going to drive your car or take a bus but you do not know how many children will be born next year before deciding if you are going to have another child.

3) Obviously, if you do not know the other participating agents, you cannot form coalitions with them. Even if you know all of them, it is not certain that you can communicate with them, let alone forming a coalition. However, coalitions may drastically change the outcome of the game.

4) The agents' goals in the game is a critical issue. The game is totally different if the goals are different. Note that in real-life situations different agents have different goals. It is also possible that the agents simply react to their and their neighbors' payoffs without specific goals.

5) The one-shot game is less interesting than an iterated one where the agents act repeatedly on the basis of their personalities, their neighbors' situations, and the payoffs received for their previous actions. The next choices are determined by updating schemes that are different for different agents.

6) The personalities of the agents constitute one of the most important characteristics of the game. The psychological literature on the impact of personalities in social dilemmas is summarized in Komorita & Parks (1994). It is possible but not easy to quantify personality profiles in the traditional psychological sense. We will use the term "personality" in the sense of decision heuristics (repeated-game strategies) in this work, to represent the fact that different agents react differently to the same stimulus from their environment. This is a rather primitive approach but it is still much better than the unjustified assumption of uniform response.

   Personalities are usually neglected in the literature. Szilagyi (2001) has considered N-person Prisoner's Dilemmas with various personalities of the participating agents. Different agents may have quite different personalities in the same experiment. The agents' personalities may also change in time based on the influences by other agents.

Personalities of the agents may represent genetic as well as cultural differences between them. The simplest and probably most important personality profiles are the following:

1. Pavlovian: its probability of taking a certain action p changes by an amount proportional to its reward or penalty from the environment. This personality is based on Pavlov's experiments and Thorndike's (1911) law: if an action is followed by a satisfactory state of affairs, then the tendency of the agent to produce that particular action is reinforced.

2. Stochastically predictable: p is a constant. Such an agent is not influenced by the environment at all. Special cases:
   a) Negatively stubborn: never takes this action (p=0)
   b) Positively stubborn: always takes this action (p=1)
   c) Unpredictable: acts randomly (p=0.5)

3. Accountant: p depends on the average reward for previous actions.

4. Conformist: imitates the action of the majority of its neighbors.

5. Greedy: imitates the neighbor with the highest reward.

Other profiles may include properties like aggression, sensitivity, devotion, etc.

7) The iterated game may considerably change if an agent may refuse participation in some iterations.

8) It is usually assumed that the game is uniform; therefore, the payoff curves are the same for all agents. This condition is, however, not always guaranteed.

9) When everything else is fixed, the payoff curves determine the game. There is an infinite variety of payoff curves. In addition, stochastic factors can be specified to represent stochastic responses from the environment. Zero stochastic factors mean a deterministic environment. Even in the almost trivial case when both payoff curves are straight lines and the stochastic factors are both zero, four parameters specify the environment. Attempts to describe it with a single variable are certainly too simplistic (Nowak & May 1992, Komorita 1976). As we will see, the relative position of the two payoff curves with respect to each other does not always determine the outcome of the game. Ordinal preference is not enough to represent the payoff functions: the actual amounts of reward and punishment may be as important as the relative situation of the two curves.

The N-person game is a *compound* game (it can be reduced to a series of two-person games) if and only if both payoff functions are linear (Hamburger, 1973). Therefore, a dyadic tournament where every agent plays 2-person games against each of the N-1 other agents represents only a very limited subset of the N-person game.

10) The total payoff to all agents is related to the number of cooperators but the maximum

collective payoff is usually not at maximum cooperation.

11) The agents may be distributed in space in many different ways. If there are fewer agents than locations in space or if more than one agent may occupy one location, then it is possible that the agents move around in space and their neighborhood constantly changes.

12) The agents may interact with everyone else or just with their neighbors. In the latter case they behave like cellular automata (Wolfram, 1994).

13) The number of neighborhood layers around each agent and the agent's location determine the number of its neighbors. The depth of agent A's neighborhood is defined as the maximum distance, in three orthogonal directions, that agent B can be from agent A and still be in its neighborhood. An agent at the edge or in the corner of the available space has fewer neighbors than one in the middle. The neighborhood may extend to the entire array of agents.

To make our task manageable, in the following we will assume that the game is uniform and iterated, the agents are distributed in and fully occupy a finite two-dimensional space, the updates are simultaneous, the agents have no goals, know nothing about each other, and they cannot refuse participation in any iteration. This restriction leaves the problem of payoff curves, personalities, and neighborhood open for investigation. These are the factors that are mostly neglected in the literature.

We will use computer simulation to demonstrate the role of these factors in the outcome of the game. If the parameters are selected appropriately, the simulation will exhibit behavior that is close enough to the behavior of real people when they are placed in a similar situation. It should be noted that even if only three factors are considered there is a huge number of different variations. Therefore, we can only show some characteristic examples in this paper.

3 The Model

We have developed an agent-based model for the investigation of social dilemmas with a large number of decision-makers operating in a stochastic environment (Szilagyi and Szilagyi, 2000). Our model has three distinctive new features:

(1) It is a genuine multi-agent model and it has nothing to do with repeated two-person games.

(2) It is a general framework for inquiry in which the properties of the environment as well as those of the agents are user-defined parameters and the number of interacting agents is theoretically unlimited.

(3) Although the analysis of rational agents may predict their behavior in some areas (e.g.,

economics), biological objects and even human beings are not always rational. It seems to us that human behavior can be best described as stochastic but influenced by personality characteristics. In view of this hypothesis, it becomes crucially important to investigate the role of personalities in Prisoners' Dilemma. Our agents have various distinct, user-defined personalities.

The participating agents are described as stochastic learning cellular automata, i.e., as combinations of cellular automata (Wolfram 1994, Hegselmann and Flache 1998) and stochastic learning automata (Narendra and Thathachar 1989, Flache and Macy 1996). The cellular automaton format describes the environment in which the agents interact. In our model, this environment is not limited to the agents' immediate neighbors: the agents may interact with all other agents simultaneously. Stochastic learning rules provide more powerful and realistic results than the deterministic rules usually used in cellular automata. Stochastic learning means that behavior is not determined but only shaped by its consequences, i.e., an action of the agent will be more probable but still not certain after a favorable response from the environment.

Szilagyi and Szilagyi (2000) describe the model in detail. We will only briefly explain its most important features here.

A realistic simulation model of a multi-person game must include a number of parameters that define the game to be simulated. Our model in its present form has the following user-defined parameters:

1) Size and shape of the simulation environment (array of agents).
2) Definition of neighborhood: the number of layers of agents around each agent that are considered its neighbors.
3) Payoff (reward/penalty) functions.
4) Updating schemes (learning rules) for the agents' subsequent actions.
5) Personalities.
6) Initial probabilities of cooperation.
7) Initial actions of the agents.

Our simulation environment is a two-dimensional array of the participating agents. Its size is limited only by the computer's virtual memory. The behavior of a few million interacting agents can easily be observed on the computer's screen.

There are two actions available to each agent, and each agent must choose between cooperation and defection. Each agent has a probability distribution for the two possible actions. The agents as stochastic learning cellular automata take actions according to their probabilities updated on the basis of the reward/penalty received from the environment for their previous actions, their neighbors' actions, and of the agents' personalities. The updating occurs simultaneously for all agents.

The updated probabilities lead to new decisions by the agents that are rewarded/penalized by the environment. With each iteration, the software tool draws the array of agents in a window on the computer's screen, with each agent in the array colored according to its most recent action. In an iterative game the aggregate cooperation proportion changes in time, i. e., over subsequent iterations. The experimenter can view and record the evolution of the society of agents as it changes in time. The outcome of the game depends on the personalities of the agents. For example, agents with short-term rationality will always choose defection, benevolent agents will ignore their short-term interests and will all cooperate, etc.

The updating scheme is different for different agents. Agents with completely different personalities can be allowed to interact with each other in the same experiment. Agents with various personalities and various initial states and actions can be placed anywhere in a two-dimensional array. A variety of personality profiles and their arbitrary combinations can be represented in the model.

The payoff (reward/penalty) functions are given as two curves: one (C) for a cooperator and another (D) for a defector. The payoff to each agent depends on its choice, on the distribution of other players among cooperators and defectors, and also on the properties of the environment. The payoff curves are functions of the ratio of cooperators to the total number of neighbors (Figure 1). The freedom of using arbitrary functions for the determination of the reward/penalty system makes it possible to simulate a wide range of dilemmas and other social situations, including those where the two curves intersect each other.

The number of neighborhood layers around each agent and the agent's location determine the number of its neighbors. We do not wrap around the boundaries; therefore, an agent in the corner of the array has fewer neighbors than one in the middle. The neighborhood may extend to the entire array of agents.

We wish to emphasize again that this is a genuine multi-agent model and it has nothing to do with repeated two-person games (Axelrod, 1984). It is well suited for simulating the behavior of artificial societies of large numbers of agents.

4 Pavlovian agents

It is realistic and interesting to consider Pavlovian agents first. Their response is stochastic but their probability of cooperation p changes by an amount proportional to their reward/punishment from the environment (the coefficient of proportionality is called the learning rate). These agents are primitive enough not to know anything about their rational choices but they have enough 'intelligence' to learn a behavior according to Thorndike's law. Kraines and Kraines (1989), Macy (1995), Flache and Hegselmann (1999) and others used such agents for the investigation of iterated two-person games. We will show below that it is possible to accurately predict the solutions of the multi-person Prisoners' Dilemma for such agents.

A linear updating scheme is used for these agents: the change in the probability of choosing the previously chosen action again is proportional to the reward/penalty received from the environment (payoff curves). Of course, the probabilities always remain in the interval between 0 and 1.

Let us assume that in a society of N Pavlovian agents the ratio of cooperators is x=m/N and the ratio of defectors is (1-x) at a certain time. The cooperators and the defectors are distributed randomly over the lattice. Then mC + (N-m)D is the total payoff received by the entire society and xC + (1-x)D is the average payoff to a single agent where C and D are the reward/penalty functions as defined earlier. This latter quantity is the so-called production function for the collective action of the society(Szilagyi, 2000). When the average payoff is zero, it is easy to think that nothing will happen and an equilibrium is reached. This is, however, not true. Indeed, this situation can only happen if either C = D = 0 or C and D have opposite signs. The first case means the two curves are crossing which contradicts the definition of Prisoners' Dilemma. In the second case evidently D is positive and C is negative; therefore, the defectors are rewarded and the cooperators are punished. As a result, the number of cooperators will decrease and we do not have an equilibrium.

Let us investigate what happens when the cooperators receive the same total payoff as the defectors, i.e.,

$$xC(x) = (1 - x)D(x) \tag{3}$$

(Szilagyi and Szilagyi, 2002). This may happen if C and D are both negative or both positive. In the first case, a small number of cooperators are punished big and a large number of defectors are punished little. This leads to a stable equilibrium at this point. In the second case, a large number of cooperators are rewarded slightly and a small number of defectors are rewarded greatly. This point corresponds to an unstable equilibrium.

If C and D are both linear functions of x, then the equilibrium equation is quadratic; if C and D are quadratic functions, then it is a cubic equation, etc. The equation generally has up to two real solutions. If both solutions are in the interval $0<x<1$, then both equilibria are present. We will denote these equilibrium solutions p_1 and p_2, so that $0<p_1<p_2<1$. The initial cooperation probability (which is set as a constant and uniform across all the agents) is p_0.

Let us consider the payoff curves shown in Figure 1. Suppose first that $p_0 < p_1$. Then there are few agents cooperating and many agents defecting. Those agents that happened to cooperate will be heavily punished, and their probability of cooperation will consequently go down substantially. As a result, some of the cooperators will become defectors. The agents that happened to defect will be punished somewhat, and their probability of cooperation will consequently go up. Because there are so many more defectors than cooperators, the aggregate effect is that the overall cooperation probability goes up toward p_1. A systematic formal analysis of equilibria of the Pavlovian learning model in the N-person game shows that in case of linear payoff functions and low initial cooperation rate p_0 the value of p_1 cannot exceed 50% (Szidarovszky and Szilagyi, 2002).

If $p_1 < p_0 < p_2$, three regions must be distinguished. Near the lower limit, the defectors and the cooperators are still both punished but there are more cooperators now who become defectors; therefore, the aggregate effect is that the overall cooperation probability goes down toward p_1. If the value of p_0 is chosen higher, we are in the region where the cooperators are punished but the defectors are rewarded. As a result, more will defect and the aggregate probability of cooperation again goes down toward p_1. When the value of p_0 is even higher, the defectors and the cooperators are both rewarded, but the defectors are rewarded more than the cooperators, so the proportion of cooperators will decrease and an equilibrium will be reached in this region or, if the aggregate probability reaches the region of mutual punishment, the equilibrium will occur at p_1 again.

The two cases above work together to keep the long-term aggregate cooperation proportion stable at p_1. However, since none of the agents are rewarded for always taking the same action (always cooperating or always defecting), the probability of cooperation for an individual agent varies according to the agent's own history of actions (and hence rewards). Over the long term, every single agent acquires a distinct cooperation probability depending on its own history of random actions. The amplitude of the aggregate oscillation depends on the size of the population: the larger the population, the more effectively the oscillation of each agent's actions is compensated for by the oscillation of all the other agents' actions.

When $p_2 < p_0$ there are many agents cooperating and a few agents defecting. The agents that cooperated are rewarded; at each iteration their cooperation probability tends toward 1. Since their cooperation probability is high, most of the cooperators continue to cooperate. After a few iterations their cooperation probability reaches 1, and they continue to be rewarded so they can never again defect. The few agents that happened to defect are also heavily rewarded; this encourages them to defect. The defectors still have a fairly high probability of cooperation, so at each iteration several of the defectors start to cooperate. (Note that there are defectors with high probability of cooperation and vice versa. What we cannot have is a defector with probability of cooperation consistently = 1 or a cooperator with probability of cooperation consistently = 0.) The few defectors that still continue to defect will be rewarded greatly for their defection; they eventually reach a probability of cooperation = 0 after which they will never cooperate for the duration of the simulation. After a while, the net result is that most of the agents are cooperating with probability 1 and are being continuously rewarded for doing so, and a few of the agents are always defecting, have cooperation probability 0, and are being continuously rewarded for doing so. Thus, a steady state is reached.

The two solutions are different from each other in three important ways:

1) The solution at p_1 is a stable equilibrium (attractor) with respect to the aggregate cooperation proportion while the solution at p_2 is an unstable equilibrium (repulsor).

2) The solution converges toward p_1 as an oscillation while it stabilizes exactly in the $p_2 < p_0$ case. This is because around p_1 the agents are punished no matter what they do and tend to change their cooperation probabilities over time. Therefore, these probabilities do not converge to zero or one for any individual agent. In the latter case, each agent in the steady state has a probability of cooperation of 0 or 1, and it is just the proportion of agents cooperating that determines the final aggregate cooperation proportion.

3) Initial aggregate cooperation proportions of $p_0 > p_2$ do not result in the aggregate cooperation proportion converging to 1, as you would expect if you think of p_2 as an unstable equilibrium. This is because, for an individual agent that started off as a defector, there is always some likelihood that the agent will continue to defect. This probability is initially small but continues to increase as the agent is always rewarded for defecting. If the number of agents is sufficiently large and p_0 is not too close to 1, then there will be some agents that continue to defect until their cooperation probability reaches zero due to the successive rewards they have received, and these agents will defect forever. The exception is if you start off with the aggregate cooperation proportion equal to 1. Then no agent starts as a defector and there is no chance of any of them defecting in the steady state.

The solutions can be predicted in a similar way for any situation. We have developed an algorithm that accurately predicts the final aggregate outcome for any combination of Pavlovian agents and any payoff functions (Szilagyi and Szilagyi, 2002).

Let us define the aggregate cooperation proportion $x(t)$ for iteration t as the ratio of the number of agents cooperating to the total number of agents. The algorithm computes $x(t)$ for any value of t when the array consists of a large number of agents and each agent is every other agent's neighbor. The initial value of $x(0)$ is given. If there are N agents, then $Nx(0)$ agents are initially cooperating and $N[1-x(0)]$ agents are initially defecting.

First, we take all of the agents in a given iteration of the simulation and distribute them into a set of groups called "rows," where each row represents agents that have exactly the same state. Two agents have the same state if and only if they have the same probability of cooperation and the same current action. We define a "Row" as an ordered triple indicating the proportion of the agents described, the probability of cooperation for these agents, and a Boolean value for the action of cooperation (1) or defection (0). Then we define a "Table" as an array containing all of the rows for certain iteration. Table(t) returns the table that describes iteration t. The sum of the proportions from each row of a table of course always equals 1, so that each agent is described exactly once. A Table is essentially a complete description of the state of all the agents in an iteration with the locations of the agents neglected.

To compute x(t) for any t, first compute Table(t). From Table(t), we can compute x(t) by summing up the proportions of agents in each row that describes cooperating agents.

Table(0) is as follows, based on the given information:

Row	Proportion of Agents (PA)	Probability of Cooperation (PC)	Action
1	x(0)	x(0)	1
2	1 - x(0)	x(0)	0

We can compute Table(t+1) from Table(t). This will give us Table(1) based on Table(0), Table(2) based on Table(1), and so on, then we compute x(t) as described above. For each Row in the old Table(t), construct two Rows in the new Table(t+1). Denote the ith Row in Table(t) as Row[i](t). The two new Rows in Table(t+1) will then be Row[2i](t+1) and Row[2i+1](t+1). To create them, we first compute the probability of cooperation PC(t+1) for both new Rows from that of the old Row by using the given update function. Denote the proportion of agents in Row[i](t) as PA[i](t). The two new Rows of Table(t+1) then will look like this:

Row	Proportion of Agents (PA)	Probability of Cooperation (PC)	Action
Row[2i](t+1)	{PA[i](t)}{PC(t+1)}	PC(t+1)	1
Row[2i+1](t+1)	{PA[i](t)}{1 - PC(t+1)}	PC(t+1)	0

Repeat this procedure for all values of i from 1 to the number of Rows in Table(t), and we obtain Table(t+1). As noted above, creating the series of Tables for an arbitrary number of iterations is sufficient to find the aggregate cooperation proportion x(t). As Table(0) has 2 Rows, and the number of Rows doubles whenever t is incremented, Table(t) has 2^(t+1) Rows. Therefore, this is an exponential algorithm but we were able to compute the value of x(t) for t=20 iterations in a couple of minutes.

The predictions are exact for an infinite number of agents but the experimental results of the simulation approximate the predictions very closely even for a few hundred agents and they are in complete agreement with the above qualitative explanation.

For the experiments reported in this paper the simulated societies have 10,000 agents each. The graphics output of Figure 2 shows the initial configuration for the case when the initial actions of all agents are random and their initial probability of cooperation is equal to 0.5. We see an approximately equal number of black (cooperator) and white (defector) spots. The initial state of the system is the decisive factor that determines its future state. This can be clearly seen from Figures 3 and 4 that show the evolution of a society for the case when the payoff curves are given by Figure 1. The graphs show the proportions of cooperating agents as functions of the number of iterations for different initial cooperation ratios.

For the payoff functions of Figure 1 the solutions are $p_1 = 0.180$ (stable attractor) and $p_2 = 0.695$ (unstable repulsor). Figure 3 refers to the case when the neighborhood is the entire collective of agents. When the initial cooperation ratio is below p_2, the solution of the game converges toward p_1 as an oscillation while it stabilizes exactly when the initial cooperation ratio is above p_2. As explained above, the latter case does not result in the aggregate cooperation proportion converging to 1.

The situation is different when the neighborhood is only one layer deep. In this case each agent has maximum eight neighbors whose behavior can influence its reward/penalty. Accordingly, the result is a more gradual dependence on the initial convergence ratio (Figure 4).

These results certainly satisfy the definition of chaos as "sensitive dependence on initial conditions" (Gleick, 1987). It means that a perturbation to the initial state of the system will cause the system to evolve into a different future state within a finite period of time. Thus, a very small difference in the initial cooperation ratio leads to totally different behaviors. This phenomenon satisfies the discussion of Eq. (3) above.

Naturally, the results are strongly dependent on the payoff functions. In case of Pavlovian agents the relative situation of the two payoff curves with respect to each other does not determine the outcome of the game. It is equally important to know the actual values of the payoff. For example, consider the simple payoff functions shown in Figure 1. If we shift the horizontal axis up and down, the following cases are possible:

a) Both curves are positive for any value of x. In this case only the unstable equilibrium is possible and the solution of the game depends on the value of this equilibrium and on the initial ratio of cooperators. When the initial cooperation ratio is below p_2, the solution of the game stabilizes at a lower value between zero and p_2. When the initial cooperation ratio is above p_2, the final stable ratio has a higher value between p_2 and one.

b) The D(x) curve is entirely positive but C(x) changes sign from negative to positive as the value of x grows. The situation is similar to case a). The only difference is that in this case the region where both C(x) and D(x) are positive may be too narrow to produce a solution other than total defection.

c) The most interesting case is when both C(x) and D(x) change sign. In this case both equilibria exist and we have the solutions discussed above (Figures 3 and 4).

d) The C(x) curve is entirely negative but D(x) changes sign from negative to positive as the value of x grows. Only the stable equilibrium exists. However, the region where both C(x) and D(x) are negative may be too narrow to produce a solution substantially different from total defection.

e) Both C(x) and D(x) are negative for all values of x. In this case only the stable equilibrium exists and the solution always converges to p_1.

Our experiments totally confirm these findings.

5 Experiments with Agents of Various Personalities

We have simulated the behavior of artificial societies for the case of various personalities of the participating agents. The stochastically predictable agents have, in fact, no personality. Therefore, we will not consider them in this study. For the other personality types, remarkably interesting patterns arise even when all agents have the same personality.

A) Conformist agents

The conformist agent imitates the action of the majority. If all agents are conformists and the neighborhood extends to the entire society of agents, then the outcome depends on the exact relationship between the initial number of cooperators and defectors: every agent will immediately imitate the majority and stay there. The behavior becomes quite interesting for the one-layer deep neighborhood. In this case, while the proportion of cooperators will not change substantially, their distribution will. Both cooperators (black spots) and defectors (white spots) will form mutually intertwined clusters (Figure 5).

B) Greedy agents

The greedy agent always imitates the behavior of the neighbor with the highest reward (this is the case investigated for dyadic interactions by Nowak and May, 1992). If all agents are greedy and the neighborhood extends to the entire organization, they will all defect immediately at the first iteration because they will all imitate the defectors who received higher rewards for their initial action than the cooperators. The situation is not so hopeless for a one-layer deep neighborhood but the behavior will stabilize with a relatively small number of cooperators. Interesting oscillating patterns arise when the payoff functions are those shown in Figure 6 (see Figures 7 and 8).

C) Accountants

The accountant's payoff depends on the average reward for its previous actions. If initially the number of cooperators is approximately equal to the number of defectors for a one-layer deep neighborhood, the result is universal defection because the defectors' payoff is always higher than that of the cooperators. If, however, the initial distribution is unequal, clusters will form. Agents situated at the borders of cooperative clusters will receive smaller and smaller payoffs. As a result, they will eventually defect, these clusters become smaller and smaller and after several thousand iterations universal defection takes over.

D) Non-uniform distributions and mixed personalities

More realistic simulations must take non-uniform distributions of different agents into account. Consider, for example, Pavlovian agents with the payoff functions of Figure 1 for the case when the initial actions of all agents are random but the society is equally divided into two parts: agents in the upper half initially defect, those in the lower half initially cooperate. Figure 9 shows the graphics output of the initial configuration for this case.

If the neighborhood is one layer deep, the upper half will be gradually infected with cooperators (Figure 10). As the neighborhood depth is increased, a protective layer is formed where no cooperation occurs (Figure 11). The situation is completely different when the neighborhood is the entire society of agents. In this case change starts in the lower region (Figure 12) and it gradually spreads into the entire society (Figure 13).

Figure 14 shows a situation when a single defector sits in the middle of the society. If all agents are greedy, the payoff functions are given by Figure 6, and the neighborhood is one layer deep, beautiful symmetric fractal patterns arise (Figure 15) that oscillate around a 29% cooperation rate. It is instructional to investigate the emergence of these patterns. As the D(x) curve is always above the C(x) curve, a layer of defectors will surround the lonely defector after the first iteration. After the second iteration, however, the further development depends on the actual shapes of the payoff curves. Accordingly, the result may be universal defection, a small stable defection pattern around the center, oscillation in the same region, or the symmetric oscillating pattern of Figure 15. If we allow a small number of individual defectors randomly distributed among cooperators, these patterns interact with each other and can produce other interesting patterns (Figure 16).

The number of variations is infinitely large. We can change all the parameters simultaneously and mix different personalities in arbitrary ways. Figure 17 shows situations similar to that of Figure 7 but with mixed personalities. Szilagyi (2001) reported additional interesting cases.

6 Conclusion

The multi-agent Prisoner's Dilemma game has non-trivial but remarkably regular solutions. The experiments performed with our new simulation tool for realistic situations when agents have various personalities show interesting new results. For the case of Pavlovian agents we found two distinctly different solutions. For a wide range of initial conditions, the number of cooperators in the society oscillates around a relatively small value. When the initial aggregate cooperation probability is above a certain value, the solutions tend to reach well-defined constant values that are dependent on the initial values. Universal defection occurs only when $p_1=0$ and $p_0 < p_2$ or when p_2 is the only solution and $p_0 << p_2$.

For other types of agents the solutions show interesting chaos-like behavior. All solutions strongly depend on the choice of parameter values. Our results show that a viable model for the study of N-person Prisoners' Dilemmas must be based on a more careful selection of parameters than that offered in the literature.

The paper provides some insight into the conditions of decentralized cooperation in spatially distributed populations of agents. However, many questions remain open. Future research will find answers to many of them. For example, we will learn the mechanism of cluster formation and the interactions of clusters with each other, the explanation of oscillatory behavior of greedy agents, the role of group size in the emergence of cooperation, etc. As a result, the study of N-person Prisoners' Dilemmas may lead us to a better understanding of some basic social dynamics, the emergence of social norms, and even may give us some insight into the possibility of changing human behavior.

Acknowledgment

This work was partially supported by the University of Arizona Foundation and the Office of the Vice President for Research of the University of Arizona.

List of Figures

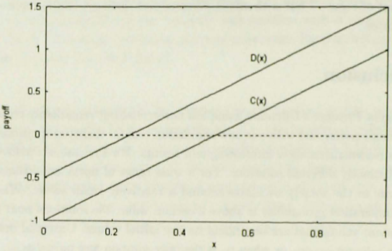

Figure 1

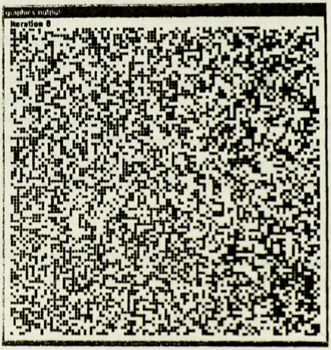

Figure 2

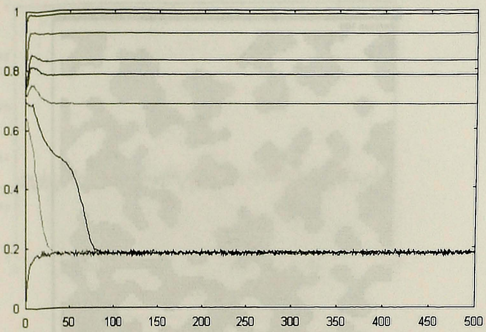

Figure 3

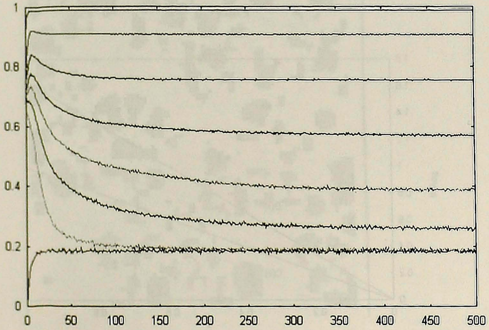

Figure 4

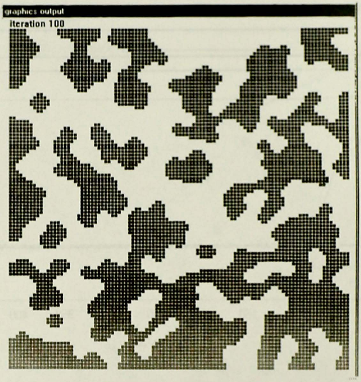

Figure 5

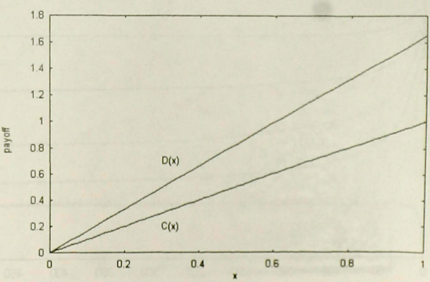

Figure 6

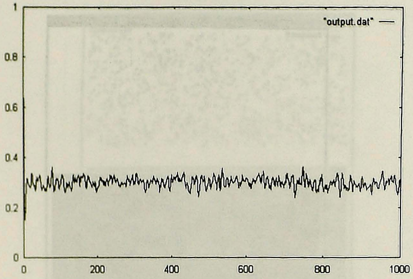

Figure 7

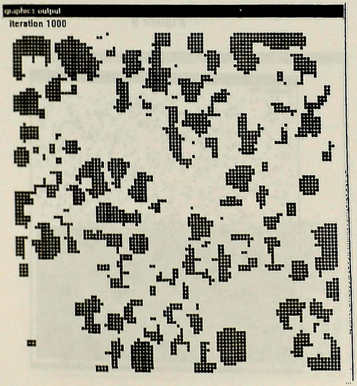

Figure 8

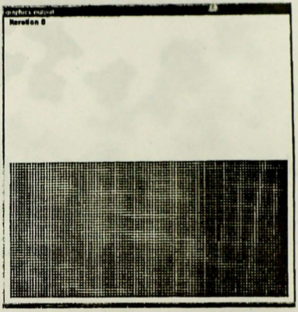

Figure 9

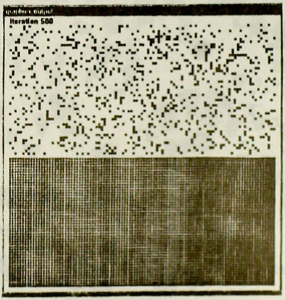

Figure 10

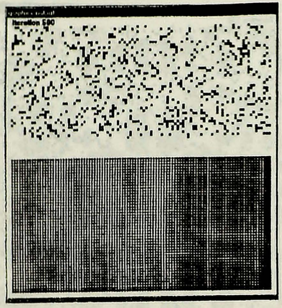

Figure 11

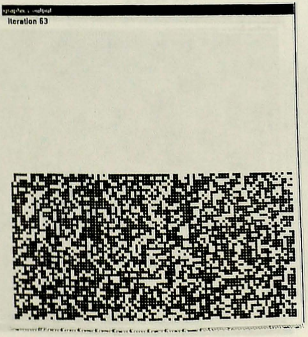

Figure 12

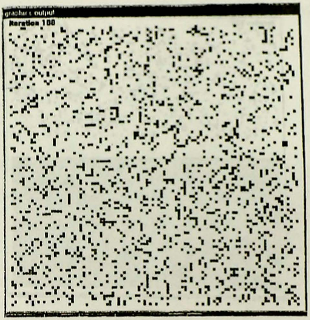

Figure 13

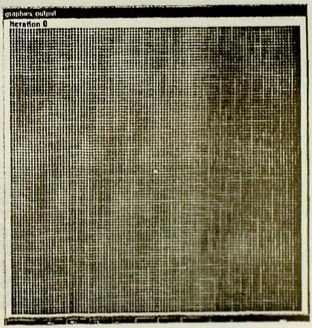

Figure 14

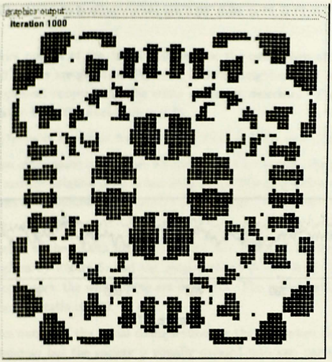

Figure 15

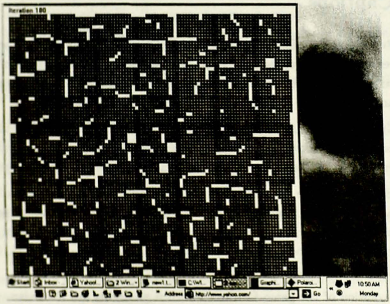

Figure 16

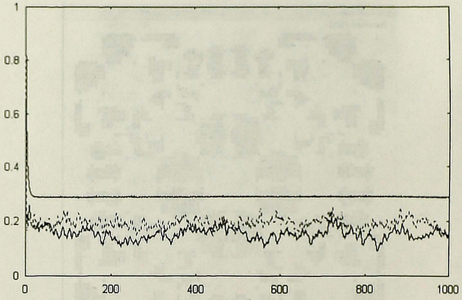

Figure 17

Figure Captions

Figure 1. Reward/penalty functions for defectors (D) and cooperators (C). The horizontal axis (x) represents the ratio of the number of cooperators to the total number of neighbors; the vertical axis is the reward/penalty provided by the environment. In this figure, D(x) = - 0.5 + 2 x and C(x) = -1 + 2 x.

Figure 2. Graphics output of the initial configuration for the case when the initial actions of all agents are random and their initial probabilities of cooperation are equal to 0.5. There is an approximately equal number of black (cooperator) and white (defector) spots.

Figure 3. Evolution of the game for the case when all agents are Pavlovian, the payoff curves are given by Figure 1, and the neighborhood is the entire collective of agents. The graphs show the proportions of cooperating agents as functions of the number of iterations. The initial cooperation ratios from top to bottom curves are 0.90, 0.80, 0.75, 0.73, 0.71, 0.69, 0.65 and 0.00, respectively.

Figure 4. Evolution of the game for the case when all agents are Pavlovian, the payoff curves are given by Figure 1, and the neighborhood is one layer deep. The graphs show the proportions of cooperating agents as functions of the number of iterations. The initial

cooperation ratios from top to bottom curves are 0.90, 0.80, 0.75, 0.73, 0.71, 0.69, 0.65 and 0.00, respectively.

Figure 5. Graphics output of the 100th iteration for the case when all agents are conformists, the payoff curves are given by Figure 1, and the neighborhood is one layer deep. The black spots represent cooperators, the white spots are defectors. The initial ratio of cooperation is equal to 0.50, the final ratio is 0.49.

Figure 6. Reward/penalty functions for the case of $D(x) = 1.65x$ and $C(x) = x$.

Figure 7. Evolution of the game for the case when all agents are greedy, the payoff curves are given by Figure 6, and the neighborhood is one layer deep. The graph shows the proportion of cooperating agents as a function of the number of iterations. The initial cooperation ratio is equal to 0.9.

Figure 8. A snapshot of the 1000th iteration for the case when all agents are greedy, the payoff curves are given by Figure 6, and the neighborhood is one layer deep. The black spots represent cooperators, the white spots are defectors. The initial ratio of cooperation is equal to 0.90, the final ratio is 0.29.

Figure 9. Graphics output of the initial configuration for the case when the initial actions of all agents are random but the society is equally divided into two parts: agents in the upper half initially defect while those in the lower half initially cooperate.

Figure 10. Graphics output of the 500th iteration for the case when the initial actions of all Pavlovian agents are random but the array of agents is equally divided into two parts: agents in the upper half initially defect while those in the lower half initially cooperate. The neighborhood is one layer deep.

Figure 11. Graphics output of the 500th iteration for the case of Figure 10 when the neighborhood is ten layers deep.

Figure 12. Graphics output of the 63rd iteration for the initial case of Figure 9 when the neighborhood is the entire society of agents.

Figure 13. Graphics output of the 100th iteration for the initial case of Figure 9 when the neighborhood is the entire society of agents.

Figure 14. Graphics output of the initial configuration for the case when a single defector sits in the middle of the society.

Figure 15. Snapshot of the 1000th iteration for the case when initially a single defector sits in the middle of a sea of greedy cooperators. The payoff functions are given in Figure 6. The neighborhood is one layer deep.

Figure 16. A snapshot of the 120th iteration for the case when all agents are greedy, the

payoff curves are given by $C(x) = 5x - 1$ and $D(x) = 5x - 0.5$, and the neighborhood is one layer deep. The initial ratio of cooperation is equal to 0.90, the final ratio oscillates between 0.91 and 0.92.

Figure 17. Evolution of the game for the case when the payoff curves are given by Figure 6 and the neighborhood is one layer deep. The graphs show the proportions of cooperating agents as functions of the number of iterations. The lower solid curve corresponds to the case when 97% of the agents are greedy, 3% are Pavlovian. For the middle dotted curve 97% of the agents are greedy, 3% are conformists. In case of the upper solid curve 45% of the agents are greedy, 45% of them are conformists and 10% are Pavlovian. The initial cooperation ratio is equal to 0.9.

References

- AKIMOV, V. & SOUTCHANSKY, M., *Automata simulation of N-person social dilemma games*, J. of Conflict Resolution, **38**(1), (1994), 138–148.
- ANDERSON, J.M., *A model for "The Tragedy of the Commons"*, IEEE Transactions on Systems, Man, and Cybernetics, (1974), 103–105.
- AXELROD, R., *The Evolution of Cooperation*, New York: Basic Books, 1984.
- BIXENSTINE, V.E., LEVITT, C.A. & WILSON, K.V., *Collaboration among six persons in a Prisoner's Dilemma game*, J. of Conflict Resolution, **10**(4), (1966), 488–496.
- BONACICH, P., SHURE, G.H., KAHAN, J.P. & MEEKER, R.J., *Cooperation and group size in the N-person Prisoners' Dilemma*, J. of Conflict Resolution, **20**(4), (1976), 687–706.
- DAWES, R.M., *Formal models of dilemmas in social decision-making*, In Human Judgment and Decision Processes (M.F. Kaplan and S. Schwartz, Eds.), Academic Press, New York, (1975), 87–107.
- DAWES, R.M., *Social dilemmas*, Ann. Rev. Psychol, **31** (1980), 169–193.
- EPSTEIN, J.M. & AXTELL, R., *Growing Artificial Societies*, Brookings Institution Press-The MIT Press, Washington-Cambridge-London, 1996.
- FEINBERG, W.E. & JOHNSON, N.R., *Radical leaders, moderate followers: Effects of alternative strategies on achieving consensus for action in simulated crowds*, J. Mathematical Sociology, **15** (1990), 91–115.

- Flache, A. & Hegselmann, R., *Rationality vs. learning in the evolution of solidarity networks: A theoretical comparison*, Computational and Mathematical Theory, **5**(2), (1999), 97–127.

- Flache, A. & Macy, M.W., *Weakness of strong ties: Collective action failure in a highly cohesive group*, J. Mathematical Sociology, **21** (1996), 3–28.

- Fox, J. & Guyer, M., *"Public" choice and cooperation in n-person Prisoner's Dilemma*, J. of Conflict Resolution, **22**(3), (1978), 469–481.

- Glance, N.S. & Huberman, B.A., *The outbreak of cooperation*, J. of Math. Sociol, **17**(4): (1993), 281–302.

- Gleick, J., *Chaos: Making a New Science*, New York: Viking, 1987.

- Goehring, D.J. & Kahan, J.P., *The uniform N-person Prisoner's Dilemma game*, J. of Conflict Resolution, **20**(1), (1976), 111–128.

- Hamburger, H., *N-person Prisoner's Dilemma*, J. of Mathematical Sociology, **3** (1973), 27–48.

- Hardin, G., *The tragedy of the Commons*, Science, **162** (1968), 1243–1248.

- Heckathorn, D.D., *Collective sanctions and the creation of Prisoner's Dilemma norms*, American Journal of Sociology, **94**(3), (1988), 535–562.

- Hegselmann, R., *Modeling social dynamics by cellular automata*, In Computer Modeling of Social Processes (W.B.G. Liebrand, M.A. Nowak and R. Hagselmann, Eds., Sage, London), (1998), 37–64.

- Hegselmann, R. & Flache, A., *Understanding complex social dynamics: A plea for cellular automata based modelling*, J. of Artificial Societies and Social Simulation, **1** (1998), 3.

- Hoffmann, R., *Twenty years on: The Evolution of Cooperation revisited*, J. of Artificial Societies and Social Simulation, **3** (2000), 2.

- Huberman, B.A. & Glance, N.S., *Evolutionary games and computer simulations*, Proc. Natl. Acad. Sci. USA, **90** (1993), 7716–7718.

- Kelley, H.H. & Grzelak, J., *Conflict between individual and common interest in an N-person relationship*, J. of Personality and Social Psychology, **21**(2), (1972), 190–197.

- Komorita, S.S., *A model of the n-person dilemma-type game*, Journal of Experimental Social Psychology, **12** (1976), 357–373.

- Komorita, S.S. & Parks, C.D., *Social Dilemmas*, Brown & Benchmark, Madison, WI, 1994.

- Kraines, D. & Kraines, V., *Pavlov and the Prisoner's Dilemma, Theory and Decision, **26** (1989), 47–79.*

- Liebrand, W.B.G., Messick, D.M. & Wilke, H.A.M. Eds., *Social Dilemmas: Theoretical Issues and Research Findings, Pergamon Press, Oxford – New York, 1992.*

- Lloyd, A.L., *Computing bouts of the Prisoner's Dilemma, Scientific American, **272(6)**, (1995), 110–115.*

- Macy, M.W., *Chains of cooperation: Threshold effects in collective action, Am. Sociol. Rev. **56** (1991), 730–747.*

- Macy, M.W., *Pavlov and the evolution of cooperation. An experimental test, Social Psychology Quarterly **58(2)**, (1995), 74–87.*

- Marinoff, L., *Maximizing expected utilities in the Prisoner's Dilemma, J. of Conflict Resolution **36(1)**, (1992), 183–216.*

- McPhail, C., Powers, W.T. & Tucker, C.W., *Simulating individual and collective action in temporary gatherings, Soc. Sci. Computer Rev, **10** (1992), 1–28.*

- Messick, D.M., & Liebrand, W.B.G., *Individual heuristics and the dynamics of cooperation in large groups, Psychological Review, **102** (1995), 131–145.*

- Narendra, K.S. & Thathachar, M.A.L., *Learning Automata (An Introduction), Prentice Hall, Englewood Cliffs, NJ, 1989.*

- Nishihara, K., *A resolution of N-person prisoners' dilemma, Economic Theory, **10** (1997), 531–540.*

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

- Okada, A., *The possibility of cooperation in an N-person Prisoners' Dilemma with institutional arrangements, Public Choice, **77** (1993), 629–656.*

- Oliver, P.E., *Formal models of collective action, Ann. Rev. Sociol, **19** (1993), 271–300.*

- RAPOPORT, A., *N-Person Game Theory, University of Michigan Press, Ann Arbor, MI, 1970.*

- RAPOPORT, A., *Editorial comments, J. of Conflict Resolution, **38**(1), (1994), 149–151.*

- RAPOPORT, A. & CHAMMAH, A.M., *Prisoner's Dilemma, University of Michigan Press, Ann Arbor, MI, 1965.*

- SCHELLING, T.C., *Hockey helmets, concealed weapons, and daylight saving, J. of Conflict Resolution, **17**(3), (1973), 381–428.*

- SCHROEDER, D.A., ED., *Social Dilemmas: Perspectives on Individuals and Groups, Praeger, Westport, CT, 1995.*

- SCHULZ, U., ALBERS, W. & MUELLER, U., EDS., *Social Dilemmas and Cooperation, Springer-Verlag, Berlin – Heidelberg – New York, 1994.*

- SZIDAROVSZKY, F. AND SZILAGYI, M.N., *An Analytical Study of the N-Person Prisoners' Dilemma, Southwest Journal of Pure and Applied Mathematics, 2002.*

- SZILAGYI, M.N., *Quantitative relationships between collective action and Prisoners' Dilemma, Systems Research and Behavioral Science, **17** (2000), 65–72.*

- SZILAGYI, M.N., *Solutions to Realistic Prisoners' Dilemma Games, Proceedings of the 2001 IEEE Systems, Man and Cybernetics Conference*, TA12/2 (2001), 841–846.

- SZILAGYI, M.N. AND SZILAGYI, Z.C., *A tool for simulated social experiments*, Simulation, **74**:1, (2000), 4–10.

- SZILAGYI, M.N. AND SZILAGYI, Z.C., *Nontrivial Solutions to the N-Person Prisoners' Dilemma*", Systems Research and Behavioral Science, **19** 3, (2002), 281–290.

- THORNDIKE, E.L., *Animal Intelligence*, Hafner, Darien, CT, 1911.

- WEIL, R.L., *The N-person Prisoner's Dilemma: Some theory and a computer-oriented approach*. Behavioral Science, **11** (1966), 227–234.

- WOLFRAM, S., *Cellular Automata and Complexity*, Addison-Wesley, 1994.