The Emergence of Social Organization in the Prisoner's Dilemma: How Context-Preservation and Other Factors Promote Cooperation

Michael D. Cohen
Rick L. Riolo
Robert Axelrod

SFI WORKING PAPER: 1999-01-002

SANTA FE INSTITUTE

The Emergence of Social Organization in the Prisoner's Dilemma: How Context-Preservation and other Factors Promote Cooperation

Michael D. Cohen Rick L. Riolo Robert Axelrod
The University of Michigan
Ann Arbor, MI 48109 USA
{mdc,rlriolo,axe}@umich.edu

16 Dec 1998

Abstract

While complex adaptive systems (CAS) theories focus primarily on phenomena such as systemic robustness against perturbation, self-organization, and on the emergence, transformation, and dissolution of organizational entities or action patterns, the metaphorical resonance of CAS work is not easily translated into careful scientific results. It can be very difficult to identify the right level at which to develop more precise theoretical generalizations with well-specified domains of applicability. And constructing experimental parameters that cleanly map to important, general constructs is usually not a simple exercise. This paper demonstrates an approach to this problem. We report results of agent-based simulation experiments in which the basic activity of the agents is to play the Iterated Prisoner's Dilemma with other agents. We systematically investigate how the emergence and maintenance of cooperation is affected by variations in three key dimensions: (1) strategy space from which the agents' strategies are selected, (2) the *interaction processes* that channel agents into interactions, and (3) the *adaptive processes* that govern the changes in agents' strategies over time. Overall, our experiments both confirm results which have been reported in the literature (e.g., that embedding agents in a 2 dimensional space can lead to the emergence of cooperation), and our results demonstrate surprising results (e.g., that high levels of cooperation can arise even when agents are randomly mixing, when the agents use simple deterministic strategies and update them using a kind of evolutionary algorithm). Our results also support a generalized view of "neighborhood" where the important factor is the degree to which the interaction processes lead to *context preservation,* independent of any particular topology. The preservation of context, even as agents are changing their strategies, acts as a "shadow of the adaptive future," resulting in sets of agents who are highly cooperative and resistant to invasion by cheaters.

1 Introduction

Many social scientists have an intuitive sense that the development of research on complex adaptive systems promises to illuminate major themes in the study of social organization that are not easily understood from more classical perspectives. Complex adaptive systems (CAS) theories appear to focus centrally on phenomena such as sensitive dependence on history, (and conversely) on systemic robustness against perturbation, on self-organization, and on the emergence, transformation, and dissolution of organizational entities or action patterns.

However, the metaphorical resonance of this work is not easily translated into careful scientific results. There are many studies of particular complex systems, either empirical or theoretical, that offer persuasive analogies to students of social systems. But there is still considerable difficulty in identifying the right level at which to develop more precise theoretical generalizations with well-specified domains of applicability. On the one hand, the causal mechanisms act at very low levels, involving the interaction and adaptation of individual agents–agents who are acting in a specific problem domain, using strategies appropriate to that particular domain. Thus the causal level is far too detailed to serve as a basis for building general theories. On the other hand, the dynamics of complex adaptive systems are a result of various nonlinear interactions of the underlying causal mechanisms, non-linearities which differ from domain to domain. Thus it also is hard to build general theories that apply across many domains using high level, domain-independent constructs.

An additional impediment to the discovery of useful, general theoretical concepts is the difficulty of constructing experimental parameters that cleanly map to important, general constructs. For example, one way to bias interactions between agents is to embed them in a 2 dimensional (2D) space, and allow them to interact only with their immediate neighbors. Note that if agents do not move (or move slowly relative to other dynamics in the model), the imposition of this neighborhood bias means that not only will each agent continue to interact with the same agents over long periods of time (often the intended effect of embedding agents in a 2D topology), but the neighbors of an agent also will be neighbors of each other. Thus effects attributed to one theoretical concept (preservation of neighbor relations) could be in part the result of another concept (neighbors' shared neighbors). This kind of confounding of potential theoretical concepts relative to experimentally manipulable parameters is quite common in CAS models, in part because it is difficult to avoid such confounding when building models out of socially plausible mechanisms.

This paper reports an approach to this problem in the context of one broad area of interest, namely the creation of theories which help us understand the conditions which lead to the emergence, maintenance, transformation and dissolution of population-level patterns of activity. These patterns of activity in turn can lead to effective social outcomes and high-level social structures. In particular, we are interested in how the emergence and maintenance of patterns of activity are affected by variations in the forces influencing *interaction patterns* between agents, by variations in the *adaptive processes* used by the agents and by interactions between these two dimensions.

In this paper we establish a systematic experimental framework for studying the emergence of social action patterns in populations of simulated agents. Our experimental design derives from three key facts:

1. The nature of the problem domain confronting the agents, and the types of strategies agents can use to cope with those problems, in part determines the resultant population-level patterns of activity.

2. Human beings live within arrangements of space, social labeling, and organizational roles that make some interactions much more likely than others.

3. Human beings[1] learn rapidly, if imperfectly, from comparisons of their own experience with that of others.

In this paper we report results of agent-based simulation experiments in which the basic activity of the agents is to play the Iterated Prisoner's Dilemma with other agents in a population. (However, our experimental design was created to apply to many other forms of action such as market exchanges, friendship formations, disease transmissions, or rumor diffusions.) We systematically investigate the effects of variations in three key dimensions:

- the *strategy space* from which the agents' strategies are selected;

- the *interaction processes* that channel agents into interactions with some agents and away from interactions with others; and

- the *adaptive processes* that govern the changes in agents' strategies over time.

The results of interest are the effects of these systematically varied factors on the emergence of population-level action patterns—in this case sustained epochs of widespread mutual cooperation. Overall, our experiments both confirm results which have been reported in the literature and demonstrate surprising

[1]Note that the model we study also may be applied to systems of non-human agents, e.g., by interpreting our learning mechanisms in ways appropriate to evolutionary mechanisms for adaptation.

results. For instance, our experiments confirm that embedding agents in a 2 dimensional space can lead to the emergence of cooperation. But our experiments surprisingly also show that high levels of cooperation can arise even when agents are randomly mixing, when the agents use simple deterministic strategies and update them using a kind evolutionary algorithm. In order to establish intuitions and a baseline understanding of the behavior of our agent populations, we examine in detail two cases, one that leads to all defections and a second that results in a high level of (nearly) stable cooperation. We then focus on one factor which influences the emergence of cooperation, namely the extent to which the various interaction processes preserve, to a greater or lesser degree, the context of interaction among agents and their descendants. By showing that cooperation emerges for a variety of interaction processes, we demonstrate support for a generalized view of the concept of "neighborhood," i.e., where the important factor is the degree to which the social (interaction) structures lead to *context preservation*, independent of a particular (e.g., 2D) topology.

The rest of this paper is organized as follows: Section 2 introduces the Iterated Prisoner's Dilemma and its relationship to dynamic social structures. Section 3 describes in detail our experimental design, implementation and performance measures. Section 4 describes the overall results of our experiment, in terms of our three basic experimental dimensions, and discusses some of the most notable trends and exceptions in those results. Sections 5 through 8 examine some experimental cases in more detail, in order to show how and why context preservation is so important for the emergence and maintenance of cooperation. The paper ends with a discussion of how this work relates to research done by others, and to studies to be done in the future, followed by a brief summary.

2 The Iterated Prisoner's Dilemma and Dynamic Social Structures

We have chosen the Iterated Prisoner's Dilemma (IPD) as one of the simplest settings in which we can study issues we believe to be fundamental. The game has been extensively analyzed from many perspectives in the social and computational sciences and is now fairly well known in the general intellectual community ([Axelrod, 1984], [Poundstone, 1992]). As an iterated interaction between two players, it requires each to choose, before knowing the choice of the other, whether to "cooperate" or to "defect." The utilities resulting for each player are shown in the following matrix, with the first number in each cell being the result for the player labelled *Row*, the second number for the player labeled *Column*.

The dilemma is fundamental to social life: how to sustain over time a pattern of cooperation between agents that may be quite beneficial to both, when in the short run it is always in the interest of each agent to defect, no matter what action the other agent may take. (It should be recalled that the actions labelled

		Column	
		Cooperate	Defect
Row	Cooperate	3, 3	0, 5
	Defect	5, 0	1, 1

Table 1: The Prisoner's Dilemma Payoff Matrix

"cooperate" may not be benign to third parties. The actors could be fixing prices or attacking their shared enemies just as well as they could be disposing properly of toxic waste or sharing socially valuable data.)

We regard a sustained regime of cooperation among a collection of agents as a kind of emergent social structure, in which the reciprocation of cooperative actions and the punishment of defection create and maintain the defining action tendencies of the members of the cooperative set. The agents' actions reinforce their roles and renew the resources that sustain the structure within which they act. This stance has a substantial resonance with sociological approaches such as Anthony Giddens' structuration [Giddens, 1984], and with anthropological dynamics for artifacts and words posited by Hutchins and Hazelhurst [Hutchins and Hazelhurst, 1990].

The emergence of a cooperative regime has to do with the spread of a community of self-reinforcing strategies. In a population of PD playing agents, such cooperative regimes may or may not arise, and if they do, they may be sustained or they may disintegrate. If they are sustained, they may include all the members of the population, or only some of them, and membership may vary over time. The rates of cooperative behavior occurring within such a regime may also vary. What determines these alternate histories of emergence is a combination of the nature of the strategies available to the agents, the pattern of interaction among the agents, and the processes operating that create, destroy and transform the agents. These observations lead us to the systematic study of a variety of interaction patterns and adaptation processes that are plausible in the social world.

This last point deserves emphasis: our interest is not simply in systems of agents whose strategies do not change, which is a widely studied case (cf. [Axelrod and Dion, 1988], [Nowak and Sigmund, 1989]). Instead we are interested in populations of adaptive agents who change their strategies as a result of the interactions they have with other agents. For this reason we define an agent as having both a strategy and a location in an interaction space, each of which may change over time. When agent and strategy are confounded, as happened in many early IPD studies, the separate contribution of agent-level interaction processes is obscured.

Finally, in our view there is an additional element required for "minimal realism": imperfection, (or "error", or "noise") at all stages of the processes. Real interactions may always involve failures to execute what was intended, misperceptions of the actions of others, misjudgments of the experience of others, and errors in understanding and emulating the strategies of others. We want to be sure that the results we report are robust, and do not depend on the pristine exactitude of one particular computer simulation setting. For this reason we incorporate several sources of error or noise into our models, and study what differences occur when those sources are present or absent. We shall see that the presence of error in judging the performance of, and in copying the strategies of, others will play a key role in many of the results to be reported.

3 Experimental Details

This section describes the model we studied, including the basic model elements common to all our experimental variations, as well as the three key dimensions we have systematically varied in this study, namely:

1. The strategy space (StratSp), with 2 variants;
2. The interaction processes (IntProc), with 6 variants; and
3. The adaptive processes (AdpProc), with 3 variants.

Thus our basic experiment consists of $2 * 6 * 3 = 36$ experimental conditions, which we call *cases.* The 36 combinations of experimental factors are summarized in Table 2, which lists the abbreviated names for the variables (experimental conditions) which appear in figures and tables in this paper, along with short accounts of the reason for the acronyms.

This section also describes the measures used to track and characterize the behavior of the model when run under those various conditions.

The basic model consists of a population of individuals playing the IPD with each other. The population size is always 256. The basic time step is a *learning period*, most often referred to as just a *period* throughout this paper[2]. For period 0, the initial population is generated by evenly distributing the agents throughout the strategy space, as described in Section 3.1. In each subsequent period each individual plays a small sample of others agents from the population; the exact number and selection procedure is determined by the interaction process being used, as described in Section 3.2. When two agents play, they play exactly four rounds (iterations) of the IPD. As a result of these plays, each individual accumulates an average payoff per move for the period, which is treated as its "fitness" or "score." After all agents have played, agents are changed by the

[2]We refer to the basic time step as a learning period to emphasize an interpretation of the model as a collection of agents who change by learning in a social context. Under an alternative, more biologically inspired interpretration of the model, in which agents change as a result of evolutionary processes, the basic time step would best be called a "generation."

adaptive process being used for the run, as described in Section 3.3. Then the next period commences.

A single *run* of the model lasts 2500 periods (after the initial one), and various measures are recorded for the run, as described later in this section. The result of a run is a *history* of agent activity and change, as well as a history of resultant population-level aggregate measures. Because these models include many stochastic processes, we generate 30 histories for each case (i.e., for each combination of experimental parameter values), each run beginning with a different seed for the random number generator.

Abstractly, a run of the model, to generate a single history, proceeds as follows:

```
Population(0) = GenerateRandomPopulation(StratSp)
  for each Period T
    for each individual A from Population(T)
      for 4 IPD games
        X = GetOtherToPlay(IntProc)
        PlayPrisonersDilemma( A, X, 4 Moves )
        UpdateCummulativePayoffs( A, X )
      endfor
    endfor
    NormalizePayoffs
    Population(T+1) = ApplyAdaptiveProcesses( Population(T) )
  endfor
```

Note that we picked four moves per game because under the general conditions we are using, the shadow of the future imposed by that game length makes the attainment of cooperation difficult but not impossible [Riolo, 1997a]. This level of difficulty allows us to discriminate population-level performance between more and less challenging combinations of interaction patterns, adaptive processes and strategy spaces. Also note that since strategies are restricted to the simple memory-1 (i,p,q) type (as described in the next section), individuals cannot make use of knowledge of the fixed number of moves, so they don't degenerate into always defecting (because of a known approaching last-play).

The experiments were carried out using software written using the Swarm simulation package, a free, objective-C based system available from the Santa Fe Institute[3] The full experiment reported here involves 36 cases, each with 30 replicated histories, for a total of 1080 histories involving 256 individuals learning for 2500 periods. Since each individual plays 4 games with 4 other

[3] For information about Swarm, see http://www.santafe.edu/projects/swarm. A copy of the software used to carry out the experiments reported here is available on request from the authors.

agents, this makes a total of $256 \times 4 \times 4 \times 2500 \times 36 \times 30 = 11,059,200,000$ agent moves; and since each move involves a choice by both the Row and Column players, the total number of choices made is approximately 22 billion. The full experiment took about $2-3$ days, with runs distributed over a set of shared, low to medium speed (1997) HP-unix workstations, using the Drone experimental control package[4].

3.1 The PD Strategy Spaces

To represent the variety of strategies that players may adopt in playing the IPD we study two variants of the (i, p, q) framework which has been employed in many other studies ([Nowak and Sigmund, 1989], [Nowak and May, 1992], [Riolo, 1997a]). In this general framework, strategies may vary in p, the probability of cooperation on the next move after the other player has cooperated, and q, the probability of cooperation on the next move after the other play has defected. Since on the first move of the interaction there is no history to draw on (and since the number of plays is finite), we also require i, the player's probability of cooperation on the initial move of a game.

For one variant strategy space we study, which we call the *Binary* strategy space, the agents are restricted to using one of these four (i, p, q) combinations:

- $i = p = 1, q = 1$: Always cooperate (all-C).
- $i = p = 1, q = 0$: Mirror opponent's last action, i.e., Tit-for-Tat (TFT).
- $i = p = 0, q = 1$: Anti-Tit-For-Tat (aTFT).
- $i = p = 0, q = 0$: Always defect (all-D).

Note that these Binary strategies are deterministic, so that when two agents meet, we can immediately calculate the outcome of any number of plays. (Table 3 shows the payoffs for all combinations of these four agent types playing four-round games.) The Binary strategy space may be considered a noiseless, error-free strategy space, in that agents' moves are communicated perfectly. For experiments using the Binary strategy space, the initial population of agents created at period 0 is divided equally and randomly among the four possible strategy types. As a run proceeds, the fraction of the population in each (i, p, q) combination varies, but no new combinations are created.

We also study a *Continuous* strategy space, in which the agents' (i, p, q) values may vary independently over the entire range $[0, 1]$ for each of i, p and q. In this case, the initial population of (256) agents is spread evenly across the (p, q) space, one agent situated at each of 256 combinations of 16 equally spaced levels of p ($p = 1/32, 3/32, ..., 31/32$) and q ($q = 1/32, 3/32, ..., 31/32$).

[4] The Drone package is available from http://www.pscs.umich.edu/Software/Contents.html .

Strategy Space (StratSp)	C	Continuous; i,p,q ranges over $[0, 1]$; "noisy"
	B	Binary; i=p,q restricted to 4 alternatives (all-C,aTFT,TFT,all-D)
Interaction Processes (IntProc)	2DK	2-Dimensions, Keeping locations for entire run. Keep same 4 (N-E-W-S) neighbors for entire run. Neighbors are symmetrical.
	FRNE	Fixed Random Network of neighbors, each agent with Equal number of neighbors (4). Keep for entire run. Neighbors are symmetrical.
	FRN	Fixed Random Network of neighbors, chosen at random at period 0 (as in RWR) but keep for entire run. Asymmetrical neighbors, an average of 8 per agent (choose 4, chosen 4 times) but with variation in the number across agents.
	Tag	Each agent has an arbitrary Tag value in [0,1]; Agents have bias to play other agents with similar tags, i.e. social labels.
	2DS	2-Dimensions used to determine 4 NEWS neighbors, but neighbors are shuffled each period. Thus 4 new symmetrical neighbors each period.
	RWR	Random-With-Replacement choice of neighbors, re-chosen at each period. Approximate complete mixing, expecting 8 assymetrical neighbors (with variation in that number).
Adaptive Processes (AdpProc)	Imit	Imitation. Copy (no errors) neighbor with highest average payoff, if payoff strictly greater than own.
	BMGA	Best-Met-GA. Copy (with errors) best performing of agents met, when better than agent. Two possible errors: copy when not better 10% of the time; Sometimes make copying errors.
	1FGA	1-Fixed-GA. Copy (with errors) a randomly chosen other agent from entire population, when other better. Two possible errors: copy when not better 10% of the time; sometimes make copy errors.

Table 2: Table of Experimental Conditions and abbreviated names.

In this case the i value in each agent is initially set to its p value. But with the Continuous strategy space, as a run proceeds not only can the fraction of agents using each (i, p, q) combination vary, but for some adaptive processes (those with "noisy" copying) agents with *new* combinations of (i, p, q) are created while other combinations are lost from the population (perhaps to be recreated later).

It should be noted that even in the Binary world, where *individuals* have values of i, p and q that are only 0 or 1, it is still true that the *population* of agents has average values of these variables that are rational numbers in the interval $[0, 1]$. We will make extensive use of these population level values of p and q in our analyses of our simulations' dynamics.

It is also worth noting that the (i, p, q) values can be viewed as a memory-1 summary of the characteristic behavior generated by any possible underlying strategy. For example, the "genotype" of a strategy may be represented as a Moore machine (as in [Miller, 1996] [Hoffmann and Waring, 1996]), but the "phenotype" (behavior) of an agent can be summarized (more or less accurately) as a probability (i) of cooperating on first move and a probability of cooperating after the other cooperates (p) or defects (q). Further, the population average (i, p, q) values are well defined even if the underlying strategies are Moore machines (or other mechanisms) and the agents themselves do not track the (i, p, q) values of each other.

Mixes of PD Strategies

A key table for many calculations presented later in the paper is Table 3. It shows the 4 move scores and total score attained for the pure (deterministic) strategy shown in the row when it meets the pure strategy shown as a column.

	Other's Strategy							
	all-C		TFT		aTFT		all-D	
Own Strategy	pay / move	sum	pay / move	sum	pay / move	sum	pay / move	sum
all-C	3333	12	3333	12	0000	0	0000	0
TFT	3333	12	3333	12	0153	9	0111	3
aTFT	5555	20	5103	9	1313	8	1000	1
all-D	5555	20	5111	8	1555	16	1111	4

Table 3: Individual move and four-move total payoffs resulting from meetings of all combinations of the four deterministic pq-strategies.

Some notes on this table will help us in later exposition. We can observe that TFT does as well as, or better than, all-C against any mix of others. It does better than aTFT except when there are many all-Cs around. All-Cs however, receive very low scores with aTFT or all-D, and do well only with TFT (or their own kind). All-D is better than TFT in any mix of others that doesn't have very much TFT.

In a world of just TFT and all-D, the "break even" proportion for TFT is 0.2. Above that proportion, TFT's expected score will be higher than all-D's. Below it, the reverse will be true. Some algebra shows that there is no mix with all four types present that gives all four types the same expected score from random interactions. Therefore, the strategy transformation processes used in our studies must produce changes in the frequency of some strategy(ies) in the initial phase of our histories, since we begin with equal numbers of strategy types, placed at random locations in any interaction structure we investigate.

Note that success in the Prisoner's Dilemma does not necessarily come from scoring more in each encounter than does the other player[5]. This is a very difficult principle to hold on to. When all-D meets TFT, all-D will come away with the relatively higher score, but both agents have low absolute scores. When TFT meets a TFT or an all-C, they tie, but at a high absolute level. What counts is a high average across all the others an agent encounters. It is possible to have the highest average without "winning" a single encounter! Indeed, this is just what happened in the computer tournament organized by [Axelrod, 1984] which was won by TFT although TFT can never get a higher score than the other strategy in a pairwise interaction.

Since the overall success of a strategy depends on its performance over the mix of others encountered, there is no strategy that is best against all possible populations of others. Holding aside error processes, in a world of all-Ds a single TFT is doomed. And in a (deterministic) world of all TFTs, a single all-D is doomed. It is precisely because the mix of others encountered is so crucial that variation in interaction structures can play such a large role in the emergence of cooperative regimes. The dynamics of the system are not directly determined by global proportions of strategy types, but rather by who is meeting whom on a local scale, and how the agents adapt to the resulting experience.

3.2 Interaction Processes

Each period, each agent chooses 4 other agents to play during that period. The way agents select others to play is controlled by the *interaction process* imposed on the agents. That is, each interaction process determines (perhaps probabilistically) which agents play each other and (for some adaptive processes) which other agents a given agent may learn from when deciding whether to change its strategy. Thus each interaction process establishes an interaction topology over

[5]Strictly speaking, TFT, all-D and the like are strategies, which determine the events in an IPD, in contrast to agents, which use strategies to accumulate payoff. In this paper the terms agent and strategy are often used interchangably, since the agents in our model have just one strategy at a time.

the agents, which can be thought of as an elementary social structure, shaping the interaction patterns between agents.

In this section we describe six interaction processes. These variants were chosen to cover a range of interaction patterns, from 2DK, which imposes strong constraints, to RWR, which allows for complete mixing of agents. Table 2 is a summary of the interaction process variants we are reporting on in this paper, as well as abbreviations and short description for each variant.

Note that for all variants, an agent never selects itself for play. And while an agent may (except for 2DK, 2DS and FRNE) select another agent twice, in general that will be a low frequency event, since there are 255 other agents to choose from.

Also note that we define an agent A's *neighbors* (or neighborhood) as the agents A plays in a given period. For some interaction processes (2DK, FRNE, FRN) an agent's neighbors remain the same across periods. (Of course, the strategies those agents use may change over time.) For other interaction processes (2DS, RWR), an agent's neighbors (whom it plays) are changed every period[6]. For the remaining interaction process (Tag), the neighbors of A are not fixed, but in general an agent will tend to select neighbors (others to play) from a pool of like-Tagged others, a pool much smaller than the entire population, greatly increasing the odds that at least *some* of its neighbors will be the same from period to period.

The six interaction processes we studied are:

RWR. (Random-With-Replacement.) This interaction process approximates a kind of complete-mixing model often utilized in analytic population-level models. As such it is perhaps the most commonly studied interaction process. At each period each agent selects 4 other agents as its neighbors and plays the IPD with them. Since all agents do the same, on average each agent plays 8 others (the 4 it chooses, and, on-average, the 4 others who choose it), but there is considerable variation across agents in the number of neighbors they have. Since neighbors are re-chosen every period, RWR provides no continuity in an agent's neighborhood.

2DK. (2-Dimensional topology, Keep neighbors.) In contrast to RWR, 2DK provides complete continuity in neighborhoods. At the start of each run, agents are placed at random on a 16×16 torus[7]. Each agent then plays its 4 NEWS (North-East-West-South) neighbors. Since those neighbors also have 4 NEWS neighbors (i.e., the neighbor relation is symmetrical), each agent plays exactly 4 other agents, playing each twice during a period. The agents never move, so they keep the same neighbors throughout a run. This kind of 2-D interaction

[6] Occassionaly a neighbor agent at period T will be re-selected in the next period, $T + 1$, but this will be quite rare whenever the neighborhood size is small relative to the population size, as it is in the experiments described here.

[7] That is, agents are assigned strategies (i, p, q values) as described in Section 3.1, and then the agents are placed randomly on the torus.

pattern is the only interaction process besides RWR that is well studied in the IPD literature (cf. [Axelrod, 1984] pp158-168, [Hoffmann and Waring, 1996], [Epstein-97]), and it is known to promote cooperation under a variety of conditions.

FRNE. (Fixed Random Network, Equal number of neighbors.) This interaction process is designed to be like 2DK in that each agent has exactly 4 other agents as neighbors (symmetric neighbors), but each agent's neighbors are chosen at random, so that there are none of the correlations between neighbors-of-neighbors that are induced by a 2DK-type topology. As in 2DK, the agents keep their neighbors for the entire run. Thus at the start of each run, each agent A randomly selects 4 other agents to be its neighbors, and they in turn choose A to be one of their neighbors. After all agents have 4 symmetric neighbors, the neighbors are randomly mixed to ensure there are no correlations induced by the order in which agents choose others[8].

FRN. (Fixed Random Network.) FRN is designed to distinguish the effects of two factors that are confounded when changing from 2DK to RWR: (1) the number of different neighbors each agent has, and (2) whether or not the neighbors are retained over time. With FRN, at the start of each run, each agent picks 4 others to be its neighbors. Thus as with RWR, on average each agent plays 8 others, and there is considerable variation across agents in the number of neighbors they have. However, with FRN, agents retain the same neighbors for the entire run (as they do for 2DK and FRNE). Relative to FRNE, then, the FRN interaction process is a move further away from 2DK, in that each agent has a fixed set of neighbors for an entire run (as in 2DK and FRNE), but unlike 2DK and FRNE the neighborhood relation is not symmetrical, in that if A chooses B to be a neighbor, there is only a tiny (4/255) chance that B will also choose A.

2DS. (2-Dimensional topology, Shuffle each period.) The 2DS interaction process also was designed to distinguish the effects of different factors that are confounded when changing from 2DK to RWR. For 2DS, as with 2DK, at the start of each run agents are placed randomly in a 2-dimensional torus, and like 2DK they choose and play their 4 NEWS neighbors. However, unlike 2DK, in 2DS runs the agents are randomly re-located (shuffled) each period. Thus like 2DK (and FRNE), each agent has exactly 4 neighbors, which it plays twice each period. But like RWR, the agents choose new neighbors each period.

Tag. (Use arbitrary Tags to bias partner selection.) The basic idea of the "Tag" interaction process is that agents select others to play (to be neighbors) who have similar "tags." Thus tags may be thought of as a kind of social label.

[8] The randomization is done as follows: After all agents have 4 symmetric neighbors, each agent randomly swaps a neighbor with 256 randomly chosen other agents, with each swap subject to the FRNE constraints that each agent have 4 unique neighbors and each neighbor relation is symetric.

There are many ways agents could use tags to select which other agents to play ([Riolo, 1992], [Holland, 1993], [Lindgren and Nordahl, 1994], [Holland, 1995]). For our model, the method of biasing the selection of "neighbors" was implemented exactly as described in [Riolo, 1997b] and [Riolo, 1997a]. The goal was to model individuals who each period wander around randomly, looking for opportunities to play the IPD with "acceptable" partners. The tags are real numbers in the range $[0, 1]$, and they are assigned at random to the initial agents at period 0. When a wandering individual A meets another individual B, A and B look at each other's tags (which can be thought of as external markings or behavior patterns) and then decide whether they both want to play each other or not. If they both choose to play, they play in the usual way (i.e., they play 4 rounds and tally up their scores). If either chooses not to play, then A continues to wander, looking for an acceptable partner. However, there is a cost of searching which reduces A's fitness in proportion to the number of searches it does. Also, an individual must play a minimum number of IPD games each period, so if after 4 of meetings A still has not found an acceptable partner, then A gives up being picky and plays the next individual it meets.

In some ways this tag selection mechanism resembles schemes in which individuals are located in a one dimensional (1-D) space, biasing their choice of partners to favor nearby individuals (e.g., [Oliphant, 1994], [Hoffmann and Waring, 1996]). However, there are some significant differences between the tag mechanism described here and 1-D space-based biases. First, note that the tag-space as used here is not a scarce resource in the way space-based biases usually are. That is, models that use space to bias interactions usually assume only one or a few individuals can be at any given distance away from an individual A. In the tag based scheme used here, since any number of agents can have a particular tag, and since candidates are picked at random from the whole population, all other agents could be equally near or far. Also note that if an individual fails to find an acceptable partner after a small number of searches (4 in these experiments), the individual reverts to selecting a partner at random from the whole population. Finally, in this scheme there is a search cost which increases for individuals who for one reason or another are "far" in tag-space from the rest of the population[9].

3.3 Adaptive Processes

An *adaptive process* controls how agents adapt (learn) over time. In the studies reported here, adaptation is synchronous: the adaptive process is applied after all the agents have played all games with all others they are going to play in the current period. Thus all agents change in parallel at the same virtual instant[10].

[9] The search cost is $0.02 \times c \times P$, where $c = 4$ is the number of others rejected while searching for an acceptable neighbor, and P is the payoff from the IPD game the agent eventually plays.

[10] In some situations synchronous updating can lead to non-robust dynamics (cf. [Huberman and Glance, 1993]), because we include conditional strategies (e.g., Tit-For-Tat)

All three adaptive processes use an agent's per-move average payoff as its score (or "fitness" in the parlance of evolutionary algorithms) for determining whether it will change or not. Thus an agent's score can range between 0.0 and 5.0, given the PD payoff matrix we used for these experiments. For all three processes we study, each agent is given one chance to adapt each period. Whether an agent changes, and if so how it changes, depends on the specific adaptive process being used, as described below. In general, though, the higher an agent's score, the less likely it is to change, whereas the lower the score, the more likely it is to change.

Imitation. Imitation is based on the learning heuristic "if an agent I played is doing better than me, copy that agent's strategy because it is likely to work better than my current, poorly performing strategy." Using Imitation, an agent copies *perfectly* the strategy of some other better-performing agent it has played in the current period. Imitation is a commonly studied adaptive process in the IPD literature, either explicitly (cf. [Nakamura, Matsuda and Iwasa, 1997]) or implicitly, via replicator dynamics (cf. [Nowak and May, 1992]). In our model of Imitation, each agent compares its score to the score of all its neighbors, i.e., the agents it played during the current period. If its score is greater than or equal to the best of its neighbors, it remains unchanged this period. If its score is less than the score of one or more others, it copies the strategy (and Tag, if it is being used) of the neighbor with the best score. If there are several tied for best, one is chosen at random.

BMGA. (Best-Met Genetic Algorithm hybrid.) BMGA was designed to model Imitation with occasional error. BMGA is similar to Imitation in that an agent first compares its score to the scores of its neighbors, i.e., the agents it played during the current period, and then it copies the strategy (and perhaps tag) of the agent with the better score, if any. (If there is a tie between an agent's score and the best score of all its neighbors, the agent "copies" its own strategy.) However the BMGA adaptive mechanism introduces two sources of possible copying errors (or noise):

1. A *comparison* error. When an agent compares its score to another, it may make an error of comparison and conclude that the other has a higher score than its own when it does not (or vice versa).
2. A *copy* error. When the agent copies the strategy (and tag) from the winning agent (perhaps itself), it may make an error and mis-copy part of the strategy (or the tag).

These sources of error are analogous[11] to processes found in Genetic Algorithms [Mitchell, 1996]: (1) tournament selection of a winning strategy, and

and because there are many sources of "noise" in our model (during neighbor selection, during play and during adaptation) synchonous updating does not lead to the kind of rigid Cellular Automata-like dynamics that can sometimes be seen.

[11] For an early example of social learning processes as Genetic Algorithms see [Cohen, 1982].

(2) possible mutation of the values being copied. In particular, first agent A's neighbor N with the best score is selected (with ties between neighbors broken at random). If A and N do not have the same score, a tournament is carried out between A and N, in which 90% of the time the individual with the higher score is picked as the winner, but 10% of the time the one with the lower score is picked. If A and N have the same score, A is declared the winner. Agent A then copies the strategy of the winner (which may be itself), but the copying is done with a possibility of error (mutation). In particular, *each* of the "genes" (the i, p, q and tag values) have a fixed probability of being mutated. For Continuous Strategy experiments, the mutation rate was 0.1 per gene. If a continuous gene is mutated, mean 0, standard deviation 0.4 Gaussian noise is added to the gene. If the mutated value goes out of range (e.g., is greater than 1), the value is set to that limiting value. For Binary Strategy experiments, the mutation rate was 0.0399 per gene[12], and only the p, q and tag genes were mutated. If a Binary gene is mutated, its value is toggled (0 to 1 or vice versa); when the p gene is mutated, the i value is set to the new p value.

In general, these error rates (for the tournament and for mutation) were chosen to so that the rates are not so low as to lead to no changes, but not so high as to prevent the agents from "remembering" anything[13].

1FGA. (1-Fixed Agent Genetic Algorithm.) The 1FGA adaptive process is designed to be a *global learning* version of BMGA. That is, in contrast to BMGA, where each agent only learns from agents it has played (i.e., its neighbors), with 1FGA an agent can learn from *any* agent in the population[14]. The 1FGA adaptive process we use is similar to the update mechanism used by [Hoffmann and Waring, 1996]. The basic idea is to give each agent one chance to adapt each period, as with the Imitation and BMGA mechanisms, but to use a GA-like algorithm to carry out the changes[15]. As with the other adaptive algorithms we use, for 1FGA each agent A gets one chance each period to instigate an update event. Given an A, another agent B is chosen at random from the *entire* population (except A itself). As with BMGA, a tournament is held between A and B, and a winner is chosen with the same chance of comparison error. (Again ties go to agent A, the instigator.) Once the winner is chosen, its strategy (and perhaps tag) is copied into A, with a possibility for copy errors (mutations) just as done with BMGA.

[12] This value makes the overall variance per gene from the mutation process the same as in the Continuous Strategy experiments.

[13] See [Riolo, 1997a] for studies on the effects of changing these and other Genetic Algorithm related parameters.

[14] In fact, for the small neighborhood sizes used in the experiments described in this paper, agents using 1FGA will almost always be comparing their performance to agents they have *not* played during the current period.

[15] 1FGA is a GA-like algorithm that can be used in conjunction with interaction processes like 2DK. Without the inheritance mechanism that comes with the "1-fixed" aspect of 1FGA, there would be problems finding where to put offspring of randomly mated pairs.

Note one key difference between BMGA and 1FGA is that with BMGA, the "other" agent for the tournament is chosen just from A's neighbors, i.e., agents it played in the current period. So in some respects BMGA is similar to their local-GA/local-play cases described in [Hoffmann and Waring, 1996]. On the other hand, for 1FGA, the other agent is selected from the population at large, which is in some respects related to the global GA case studied in [Hoffmann and Waring, 1996]. However those comparisons are not perfect: for example, Hoffman and Waring select the other agent with probability proportional to fitness and then recombine its strategy with A, whereas we just use a (perhaps mutated) copy of the winner of a tournament[16].

3.4 Performance measures

Since we are interested in the emergence and maintenance of regimes of mutual cooperation among the agents, an important measure is the average score (fitness) of all the agents at each period. A population average at or near 1.0 indicates that the population is defecting almost all the time, whereas a value near 3.0 indicates the agents are nearly always cooperating with each other.

However, note that while population average payoffs could[17] range from 1.0 to 3.0, in our experiments the average usually ranged from (roughly) 1.1 to 2.7 (with a few exceptions, to be discussed below). Even for populations that are quite cooperative with each other, the population average payoffs don't reach the highest possible values (i.e., at or close to 3.0) for two reasons. First, for Continuous strategy spaces, the "noisiness" in the strategies leads to "misunderstandings" which lead to instabilities in the play of the game. For example, for populations which are generally cooperative, most mistakes will be accidental defections that will generally lower the scores. Second, for both strategy spaces, the BMGA and 1FGA adaptive algorithms have a small error rate, which leads to a small but constant injection of heterogeneity into the population, again leading to instability in the overall rate of cooperation. This heterogeneity of the agents' strategies, combined with the stochastic nature of the strategies, keeps the average payoffs below the maximum possible. The exceptions are found when agents are using Binary strategies and Imitation (no error) adaption. In these cases, depending on the adaptive algorithm, the population average can sometimes reach 3.0 and sometimes fall to 1.0.

A second simple aggregate measure, which indicates the long-run performance of a population, is the population-average score averaged over a number

[16] There are other differences as well which make comparing and "docking" our results with their's complicated. For instance, for Hoffman and Waring's global interaction cases each agent plays more and more of the other agents, whereas for our closest case to global mixing, RWR, our agents play only a sample of the other agents.

[17] In fact, the average payoff could be somewhat less than 1.0. If a population using the Tag interaction process was almost always defecting *and* if the agents were usually refusing others, thus incurring the Tag search cost, the defector's payoff (1) less that tag search cost could result in an average net score less than 1.

of periods. In particular, we calculate the mean of the single-period population averages, averaged over the last 1000 periods of each run (in order to avoid any effects of initial conditions). We then calculated the average, over all 30 runs of each case, of this within end-of-run average performance, to yield the long-run average performance that can be expected for each experimental condition. These aggregate values are listed in the column labelled "Mean Payoff" in Tables 4 and 5. These aggregate Mean Payoff values also are shown in histogram form in Figures 1 and 2. (The results in those figures will be discussed in Section 4.)

While these aggregate measures of population cooperation are useful, they miss many of the differences in the *dynamic paths* the various experimental conditions generate. In Sections 5– 7 there are a number of plots of population average payoff versus period, for all or part of runs from a variety of experimental cases. These plots show a wide variety of dynamics exhibited in those cases, from very stable through various levels of near-stability and on to aperiodic jumping between regimes of high cooperation and intervals of virtually all defection. To try to capture some of these different history dynamics, we have defined several additional measures, which are included for the 36 experimental cases covered in Tables 4 and 5. These measures and their definitions are as follows:

- **NumCC**: The number of runs out of the 30 for a case in which the population average score reached 2.3 for at least one period. A threshold of 2.3 was chosen because given the noisiness of (most of) our experimental conditions, a population can reach and sustain such an average score only by a significant amount of mutual cooperation [Riolo, 1997a].

- **Gen To Hi**: This is the number of periods for the population average to reach the threshold score (2.3). This indicates how well an experimental condition promotes the emergence of cooperation from an initial sea of non-cooperators.

- **Frac High**: This is a measure of the fraction of the periods the population average spends above the 2.3 threshold, *after* the threshold is first reached. Thus a high value indicates the population is able to continuously *maintain* a high level of cooperation, whereas lower levels means the population is *not* able to maintain a highly cooperative regime.

- **Frac Low**: This is a measure of the fraction of the periods the population average spends *below* a population average of 1.7, *after* the high (2.3) threshold is first reached. Falling below an population average of 1.7 indicates the population is generally defecting. Thus a high value for Frac Low indicates the population spends most time in high-defection regimes, even after it has attained high cooperation. A low value for this indicates the population can avoid the worst "crashes" in cooperation, once it has emerged during a run.

- **In-Run StdDev**: This measures the standard deviation of the single-period population average payoff over the last 1000 periods of a run. The average of the these values then is averaged over the 30 runs for a given experimental case. This value provides a convenient single measure of the amount of variation in population average payoff *within* a run, measured over the same interval for which the run-average is computed. Of course, this measure cannot distinguish all types of variation, e.g., between cases that generate a few larger variations versus cases that generate a higher proportion of small variations.)

While these measures don't capture all interesting differences between individual histories, they do convey many contrasts between cases, and they are much easier to examine than looking at 30 histories from each of 36 cases! In particular, these measures are often useful in providing an indication of what might be causing a high or low overall level of cooperation to be generated by an experimental condition. And at the least, extreme values in these measures are flags that indicate that the detailed histories of the runs for a case might be of particular interest or use.

4 Overall Results

Tables 4 and 5 show the overall aggregrate performance of the model for the 18 Continous strategy space and 18 Binary strategy space cases, respectively. Figures 1 and 2 show the Mean Payoff in the form of histograms, again for the Continuous and Binary strategy space cases, respectively.

There are a number of general observations that are supported by the data presented in those tables and figures. Some of the most striking trends and observations include:

- Overall, the interaction processes which best promote cooperation are those in which agents retain the same neighbors throughout each run—2DK, FRNE and, to a lesser extent, FRN. We call these *context preserving* interaction processes, in that each agent's interaction (social) context is preserved over time. Section 6 describes in more detail how these context preserving interaction processes lead to the emergence and maintenance of high levels of cooperation, and discusses the reasons for the differences among these cases.

- In general, the interaction processes which lead to the least amount of social organization (cooperation) are 2DS and RWR. These results are the obverse of the high cooperation obtained with the context preserving interaction processes, since 2DS and RWR are the interaction processes in which the population is mixed the most, with agents switching neighbors every period. Some notable exceptions to this are the Binary-1FGA cases, which have high Mean Payoffs for both 2DS and RWR.

Int Proc	Adaptive Process	Mean Payoff (sd)	Num Make CC	Gen to High	Frac In High	Frac In Low	In-Run StdDev (sd)
2DK	1FGA	2.025 (.069)	30	127	.213	.172	0.313 (.045)
	BMGAS	2.554 (.009)	30	26	.998	.000	0.075 (.006)
	ImitS	2.213 (.352)	15	27	.868	.000	0.016 (.003)
FRNE	1FGA	2.035 (.089)	30	122	.226	.161	0.292 (.049)
	BMGAS	2.572 (.007)	30	26	.996	.000	0.077 (.007)
	ImitS	2.176 (.507)	14	26	.901	.000	0.017 (.014)
FRN	1FGA	1.884 (.120)	30	162	.182	.311	0.379 (.044)
	BMGAS	2.476 (.026)	30	40	.939	.003	0.124 (.062)
	ImitS	1.362 (.434)	3	12	1.00	.000	0.009 (.004)
Tag	1FGA	2.198 (.057)	30	80	.380	.064	0.248 (.058)
	BGMAS	1.613 (.277)	30	331	.117	.499	0.469 (.093)
	ImitS	1.573 (.187)	0	-	-	-	0.012 (.004)
2DS	1FGA	1.484 (.086)	30	695	.040	.764	0.273 (.080)
	BMGAS	1.089 (.003)	1	978	.001	.977	0.024 (.009)
	ImitS	1.096 (.013)	0	-	-	-	0.006 (.000)
RWR	1FGA	1.502 (.109)	30	612	.055	.729	0.332 (.064)
	BMGAS	1.098 (.036)	9	1384	.021	.883	0.063 (.097)
	ImitS	1.104 (.032)	0	-	-	-	0.006 (.001)

Table 4: For Continuous strategy cases, the Mean Payoff column shows the mean and standard deviation of the population average fitness over last 1000 periods. The numbers in parentheses are the standard deviation for the associated mean. See text for definitions of the labels on the other columns.

Int Proc	Adaptive Process	Mean Payoff (sd)	Num Make CC	Gen to High	Frac In High	Frac In Low	In-Run StdDev (sd)
2DK	1FGA	2.560 (.013)	30	19	.914	.001	0.173 (.017)
	BMGAS	2.552 (.006)	30	10	.970	.000	0.122 (.006)
	ImitS	3.000 (.000)	30	7	1.00	.000	0.000 (.000)
FRNE	1FGA	2.558 (.015)	30	21	.915	.000	0.171 (.012)
	BMGAS	2.564 (.007)	30	9	.968	.000	0.127 (.007)
	ImitS	2.991 (.022)	30	6	1.00	.000	0.000 (.000)
FRN	1FGA	2.691 (.008)	30	22	.990	.000	0.120 (.010)
	BMGAS	2.629 (.010)	30	14	.913	.006	0.229 (.016)
	ImitS	1.869 (1.01)	13	7	1.00	.000	0.000 (.000)
Tag	1FGA	2.652 (.010)	30	15	.975	.000	0.132 (.010)
	BGMAS	1.449 (.186)	30	255	.191	.763	0.554 (.170)
	ImitS	1.133 (.507)	2	6	1.00	.000	0.000 (,000)
2DS	1FGA	2.522 (.024)	30	26	.867	.000	0.197 (.020)
	BMGAS	2.053 (.128)	30	95	.443	.280	0.532 (.064)
	ImitS	1.000 (.000)	0	-	-	-	0.000 (.000)
RWR	1FGA	2.685 (.0009)	30	54	.985	.000	0.127 (.013)
	BMGAS	1.175 (.099)	13	972	.191	.763	0.109 (.182)
	ImitS	1.000 (.000)	0	-	-	-	0.000 (.000)

Table 5: For Binary strategy cases, the Mean Payoff column shows the mean and standard deviation of the population average fitness over last 1000 periods. The numbers in parentheses are the standard deviation for the associated mean. See text for definitions of the labels on the other columns.

- The Tag interaction process generally is intermediate in its ability to promote cooperation, with Mean Payoffs generally lower than the Mean Payoffs for the three context preserving interaction processes mentioned above (2DK, FRNE, FRN), but generally with Mean Payoff above those obtained for the interaction processes with high mixing rates (2DS and RWR). One exception to this rule is the Binary-Tag-BMGA case, which seems to lead to a lower than expected Mean Payoff. Another even more striking exception is the Binary-Tag-Imitation (B-Tag-Imit) case, in which case the Mean Payoff is 1.00, indicating all defection all the time[18].

- The cases fall into three rough classes: (1) Cases which result in relatively high levels of more or less stable cooperation, e.g., C-2DK-BMGA, C-FRN-1FGA, and many others, which generate high average payoff for all runs, and which then never (or seldom) fall below the 2.3 payoff threshold, e.g., FracInHigh is over 0.9 and the FracInLow is 0.0; (2) Cases which never (or rarely) achieve high cooperation, e.g., C-RWR-ImitS, B-2DS-BMGA, and others, which have low average payoff, and seldom or ever get over the fitness threshold of 2.3, and if they do, fall right back into mutual defection; and (3) Cases in which the population flips back and forth between regimes of high cooperation and regimes of mostly defection, e.g., C-Tag-BMGAS, C-2DK-1FGA and others, which have a moderate overall average payoff, and have FracInHi and FracInLow levels substantially greater than zero.

- In general, 1FGA appears to be the most robust adaptive process, closely followed by BMGA. For the Continuous strategy space, 1FGA and BMGA lead to the same overall average Mean Payoff, with 1FGA leading to more cooperation than BMGA in three interaction process cases (Tag, 2DS, RWR), whereas BMGA leads to higher scores in the three other cases (2DK, FRNE, FRN). But for the Binary strategy spaces, 1FGA and BMGA lead to the (statistically) equivalent scores for three cases (2DK, FRNE, FRN) while 1FGA leads to considerably better scores in the other three cases (Tag, 2DS, RWR). Thus BMGA seems to do best when used in conjunction with interaction processes in which agents keep their same neighbors throughout a run, whereas 1FGA does well in the other cases.

- The Imitation adaptive process generally leads to the least cooperation. There are two exceptions to this: Imitation with Binary 2DK and FRNE. In these two cases, Imitation is the best adaptive process, and in fact these cases yield the best Mean Payoffs overall. In short, the problem with Imitation is that all variety in the agents' strategies is lost, so that with Imitation the population becomes stuck at a given performance level, usually with low cooperation, but in a few cases with high cooperation.

[18]Note that in this Tag case, there is no loss of payoff due to search costs once equilibrium is reached. This is because as a result of Imitation, all agents have the same tag, just as they have the same (all-D) strategy.

- The use of a Binary Strategy Space leads to more extreme results than does the Continuous Strategy Space. For example, the two highest Mean Payoffs at or near 3.0 were observered for the Binary cases of Imitation-2DK and FRNE. And at the other extreme, the lowest Mean Payoffs, 1.0 were produced by the Binary cases of Imitation-2DS and Imitation-RWR. To some extent these cases are greater (lower) than the corresponding Continuous cases because of the inherent noisiness of the Continuous strategies; unless the Continuous strategy reaches exactly the (i, p, q) boundaries, there will be occassional errors in the responses, leading to lower or higher scores than obtained with the corresponding deterministic (Binary) strategies. Another view is that the Binary Strategies give clearer signals, which enables the agents to respond in ways appropriate to the situation. This clearer signaling may enable populations to self-organize into cooperation (or defection) faster or more stably than when signals are less clear. This may explain why, for example, B-2DS-BMGA achieves much higher levels of cooperation than C-2DS-BMGA.

The above is only a partial list of the major trends and notable exceptions we have observed. The experiments which generated the aggregrate measures displayed in Tables 4 and 5 are a rich source of information about what influences the emergence and maintenance of cooperation in the IPD populations we studied. One of our major goals is to explain why the given combinations of experimental conditions produce the regularities, and the surprises, seen in the measures of cooperation shown in those tables. In the rest of this paper, we focus on one of the most important factors—the *preservation of context*—that contributes to the emergence of cooperation in our experiments. However, given the many nonlinearities in Tables 4 and 5, it is clear our experimental dimensions do not cleanly separate all important factors. Issues that await later analysis include (a) the ways in which the generation and maintenance of *variety* in the population of strategies is critical for the emergence of cooperation; (b) how interactions between the strategy spaces, interaction processes and adaptive processes generate variety and affect the dynamics of these systems; (c) the importance of *vigilance* for the maintenance of cooperation; and (d) how some combinations of our experimental parameters interact to work for or against the pressures toward defection which are inherent in the fundamental dynamics of populations playing the IPD.

5 The Basic Dynamics of Defection

This section examines a case in which cooperation never emerges, in order to establish several general points about the dynamics of our populations of IPD playing agents. These form the background against which the variations explored in the subsequent sections will stand out more clearly.

As an exemplar of the fundamental forces at work, we will use an experimental condition that implemented noisy strategy execution ("Continuous"), with agents on a 2-dimensional torus being randomly relocated each period by the interaction process ("2-D Shuffle"), and the Best-among-those-you-Met (with possible errors) adaptive process (BMGA). Thus our abbreviation for the condition is C-2DS-BMGA.

Figure 3 plots the history of the population average score per move, for a single population of 256 agents over 2500 learning periods. Figure 4 plots the same measure in greater detail for the first 50 time periods. There are several notable points in the figures:

- The initial average score per move (also called "average fitness") is shown in Figure 4 to be 2.25. This is true in every history within every experimental condition, because our initialization of the agents' strategies always yields an average (i, p, q) of $(.5, .5, .5)$, distributed randomly over whatever interaction process is being used. This is exactly what a calculation predicts as the expected fitness from the play of randomly initialized agents, interacting randomly.

- The population average score per move declines immediately and sharply. This also is true in every population for every experimental condition, no matter where the population may end up in the long run. The reason for this sharp decline is as follows: Besides many other strategies, there are many "suckers" (all-C players), i.e., individuals who have (i,p,q) at or near 1.0, who tend to cooperate no matter what the other agent does. There are also many "meanies" (all-D players), i.e, individuals who have (i,p,q) at or near 0.0, who tend to defect no matter what the other does. As a result, some all-D players achieve relatively high scores by taking advantage of neighbors who (as a result of the initial random distribution) are suckers, leading to rapid spread and dominance by the (near) all-D players. But as the number of all-D players increases and the suckers die off, so the the average payoff falls rapidly to $1.1 - 1.3$, i.e., near the mutual defection payoff expected from a population of mostly all-D players.

- Because the initial value and decline are so uniform, the population average score of 2.3 offers a convenient definition of a level of payoff that can be called improved relative to the initial configuration. As can be seen in Tables 4 and 5 (column 1, Mean Payoff), many populations cannot maintain this level of cooperation. In fact, for some experimental cases many populations never even reach this level (see column 2, NumCC), and of those that do, many cannot maintain a high level of cooperation (see column 4, "FracInHi," i.e., fraction of periods in which average payoff stays over 2.3 once that level is reached).

- The population average payoff value varies somewhat from one period to the next, even if the fundamental level of cooperation is quite stable. For example, in this experimental condition (B-2DS-BMGA) all the histories have an average value over the whole 2500 periods that is indistinguishable from the run history shown in Figure 3, as indicated by the very low standard deviation (0.003) of the Mean Payoff over the 30 runs for the case. But the variation over a few periods within one history is relatively large, as indicated by the In-Run StdDev (0.024), compared to that variation across the histories. This period-to-period variation is found in most of our experimental conditions and has three principal sources: (1) the noisy nature of strategy execution in the Continuous strategy space cases; (2) the change in an agent's neighbors, in those interaction process cases in which neighbors change (Tag, 2DS, RWR); and (3) the fluctuation of strategies resulting from effects of the "errors" (in comparing or copying) in the BMGA and 1FGA adaptive processes.

- Finally, although the population average score per move is low in this run, it never falls to the minimum sustainable average value of 1.0. This is also an effect of the three sources of noise described above. Similarly, most cases which ultimately reach high average payoff values have a "ceiling effect" that corresponds to the "floor" seen here, so they seldom reach the theoretical sustained maximum of 3.0.

An alternate picture of the forces producing population histories such as that in Figure 3 and 4 is given by Figure 5. It shows a diagram of movement in p-q space, a diagram that will be one of our standard tools in the rest of this paper. The entire figure shows a range of combinations of the the population average values of $0.0 \leq p \leq 1.0$ and $0.0 \leq q \leq 0.6$. Note that these are measured at the population level, and so are well-defined even in the Binary strategy space conditions that force each individual agent to have (i, p, q) values that are either 0 or 1. Also note that we only show q up to 0.6 because the population average q value is always well below that high level. Finally, note that we are ignoring i in this analysis, because for the most part, i is highly correlated with p[19]. (For the Binary cases, i always is forced to equal p.) While an analysis of the movement of i may be enlightening, that will have to await future studies.

For each small region (there are 20-by-20 "bins") of the space in the diagram we have determined all the periods in which the population average values of p and q were in that region. For each such we have found the average of the p and q values in the immediately succeeding period. The line originating at the center of each region shows the average one period movement of pq for populations that were ever in that region. The data shown were collected from all the thirty histories of the C-2DS-BMGA condition.

Figure 5 shows a feature found in all such diagrams based on the complete population history—a cascade from the center of p-q space toward the lower-left

[19] This correlation might indicate the Continuous strategies are approximating their Binary counterparts. For example, for p low, i tends to be low, approximating the Binary Tit-For-Tat strategy.

corner, where $p = 0$ and $q = 0$. This is the trace of the collapse from the initial state of the population, in which agents (p, q) values are evenly distributed across the entire pq-space. In that initial environment agents with strategies closer to all-D ($p = 0$, $q = 0$) do better than more cooperative strategies. This spread of lower p and q through the initially disorganized population produces the sharp movements shown. This movement corresponds to the rapid decline in the population average payoff described earlier in this section.

It is important to note that the map of forces in pq-space is contingent on the interaction structure of the population, as well as on the average and distribution of p and q in the population. Thus later in a history, when the population might have developed more correlated interaction patterns, the movement from a particular p-q cell might be different. As it happens, none of the populations in this condition ever revisit the area of p-q space that is associated with the initial state. But the fact that the map of forces depends on the interaction structure is just what makes a pq map so useful for us. We can show how different interaction structures reshape the forces acting on various population p and q levels.

A second feature of this pq-space figure is the large amount of total time the process spends in the low p and low q region of the space. This is indicated by the large circles in the corresponding cells[20]. The reason the populations spend so much time in that region of pq-space is indicated by the direction of the vectors in that region. In the cells with the largest circles, where most time is spent, the vectors are of virtually zero length, indicating there is little pressure to move from those pq states. Further, in the nearby cells the vectors all point back toward the large circles, indicating "restoring forces" that tend to push the populations back to low p and q if they should happen to drift randomly to higher values. Thus that reqion with the largest circles can be considered a basin of attraction for the populations run under the B-2DS-BMGA conditions. (There are a few disorganized vectors, some pointing away from the large basins of attraction at low p, q, which might indicate that if the population *could* more often get to the higher p values, it might move into a basin of attraction in a different (presumably higher p) region of the space. We tested for this by collecting data from 10 runs in which the entire initial population were TFT ($p = 1, q = 0$) players. Even in this case the population rapidly fell to the all-D corner of pq-space and stayed there.)

[20]The largest circle corrsponds to the cell in which the population spent the most time during the run; the areas of other circles are scaled relative to that cell.

6 The Basic Dynamics of Stable Cooperation via Context Preservation

The basic picture portrayed in the previous section corresponds to the war of all against all in Hobbes *Leviathan*, in which a typical life is described as "solitary, poor, nasty, brutish, and short." These regimes of nearly complete mutual defection occur frequently in our experiments. And the corresponding social situations are well known in the real world: conditions in which efforts to establish mutual cooperation collapse and cannot be revived, even though all would be better off if they could.

Hence a central concern of social science is to understand the conditions that allow cooperative regimes to be established and maintained. In Figure 6 we see the history of a population that does recover from the typical initial collapse to institute long-lasting cooperation. The conditions that produced this history are identical to those producing the all-D regime reported previously, with one exception: the locations of the agents on the 2-dimensional torus are not shuffled between periods. Instead agents "keep" their locations in the interaction structure, so the condition is labelled C-2DK-BMGA.

Figure 6 (and Figure 7, the first 50 periods of the run) shows some of the typical features that were listed earlier:

- The evenly distributed start produces an initial collapse of cooperation, and therefore a rapid decline in population average score per move.
- The eventual "equilibrium" population average score is stable, but it exhibits substantial period-to-period fluctuation (indicated by an In-Run StdDev of 0.75) due to the various sources of "noise" in the model under these conditions (i.e., noise in play and in adaptation).
- The overall average score does not vary much from run to run (a Mean Payoff standard deviation of 0.009). There would appear to be little dependence on the random variation in initial conditions in this case.

However, the change from the 2DS interaction process (in the earlier C-2DS-BMGA case) to the 2DK interaction process used in this C-2DK-BMGA case also results in a number of differences:

- Most notably, the equilibrium per move score is dramatically higher: for C-2DS-BMGA it was 1.089, but for C-2DK-BMGA it is 2.554. An effective cooperative regime has been established and maintained in the C-2DK-BMGA case. Although individual defecting strategies with low p and q are being continuously created (by the BMGA adaptive mechanism), they are held in check and do not take over the population.
- The initial collapse, though dire, is not as deep as the corresponding collapse in Figure 4. and it is followed by a rapid rebound that is not evident in the 2-dimensional shuffle (2DS) interaction process.

- Period-to-period variation is actually greater in this case than in the earlier one: for C-2DS-BMGA the In-Run StdDev is 0.024, but for the C-2DK-BMGA case the In-Run StdDev is 0.075.

Figure 8 shows the same set of pq-space forces associated with the collapse in the C-2DS-BMGA case, but for the C-2DK-BMGA case a remarkably different pattern dominates the majority of every population's history. From the low values of p and q that follow the collapse, the populations are firmly swept to right along the bottom of the pq phase space. That is, values of p increase rapidly, making the agents more likely to reciprocate cooperation, while values of q remain low, leaving the agents likely to answer defection with defection. What is taking over the population is therefore a strategy of the Tit-For-Tat (TFT) type.

How does this happen? In Figure 5 the all-D (low p and q) type that followed the initial collapse for the C-2DS-BMGA case was effectively self-perpetuating. But in this C-2DK-BMGA case there seems to be a strong force pushing the population toward higher p values. The answer lies in the difference contributed by the "keep" property of this experimental condition.

Consider, for simplicity of presentation, the purified case in which a sea of all-Ds (p=0, q=0) is blemished by a single TFT that has arisen through an error of copying the p value. From Table 3 one can calculate that this strategy will score 0.75 per move, while the all-Ds it meets are averaging (8+4+4+4)/(4∗4) = 1.25. Holding aside errors of comparison, that TFT will immediately disappear. More rarely, however, such an accidentally created TFT will encounter another agent that also has an accidentally created TFT strategy. With errors of copying at .1, four encounters for each agent, and 256 agents in the population, such a "double-mutation" neighborhood is in fact fairly likely to occur somewhere in the population. Now the TFTs each score $(12 + 3 + 3 + 3)/(4 * 4) = 1.3125$, while their all-D neighbors get only $(8 + 4 + 4 + 4)/16 = 1.25$. In the BMGA adaptive process, this difference in scores means that nearly all the agents who encountered the TFTs will switch their strategies to TFT in the following period. The seed of two accidental TFT's will have grown in one period to eight TFT's. And they will remain together in following period. Every TFT will again have at least one TFT encounter, and the logic will repeat itself. Islands of TFT will expand and rapidly take over the population.

However, when the population has acheived a very high density of Tit-For-Tat, a slightly different dynamic can be discerned. There is a modest tendency for average q (i.e., the probability an agent will cooperate after the other player has defected) to rise in such a population, as seen in the direction of the vectors along the right edge of Figure 8, the pq-space diagram for the C-2DK-BMGA case. This is because in a world where any other agent's strategy has high probability of reciprocating cooperation (very high average p), and where one's neighbors are persistent rather than transient, there is no cost to an individual when errors of copying result in higher values of q. Put anthropomorphically, if you are TFT and your neighbors also are all Tit-For-Tat, you cooperate fully with them and so you can drop the habit of reciprocating accidental defections. In a tie the BMGA adaptative mechanism retains the agent's existing strategy (perhaps mutated slightly). So once created, a high q value will be carried along in a benevolent neighborhood, which—in the third phase of this story—most neighborhoods are. Over time, then, the population will "*drift*" to higer q values[21] (by a process analogous to genetic drift in biological settings).

Again for simplicity of exposition, consider the extreme case of an all-C strategy ($p = 1$, $q = 1$) that has three TFT neighbors ($p = 1$, $q = 0$) and one all-D neighbor ($p = 0$, $q = 0$) that has arisen from a TFT through a copying error on the p component of the strategy. From Table 3 one can calculate the per move score that the all-C strategy will produce: $(12+12+12+0)/(4*4) = 2.25$. The all-D strategy, if it is surrounded by TFTs, will yield 20 against the endlessly forgiving all-C, but only 8 with its other 3 TFT neighbors, for a 2.75 per move average. Thus the agent following the all-C strategy will conclude in the adaptive process that all-D is the best strategy it has seen and (usually) will convert, as will the three TFT neighbors of the all-D, each of which will have experienced $(12 + 12 + 12 + 3)/16 = 2.4375$.

So in one period, a single all-D that has arisen next to an all-C within a TFT background will convert all four of its neighbors, creating a small island of all-D. Because the 2DK condition preserves neighborhood relations, the following period will force the all-D strategies to experience the neighborhood they have created. Each all-D will have at least one all-D neighbor. If no further all-C "food" is available, the remaining neighbors all will be TFT. The all-D's will yield no more than $(4 + 8 + 8 + 8)/16 = 1.75$. Any of the TFT neighbors will get $(3 + 3 + 12 + 12)/16 = 1.875$. Therefore the agents who had converted to all-D will convert back to TFT in the very next period. In the following period the all-D they had imitated will also disappear. The island of all-D strategies that had been triggered by the conjunction of a new all-D with an all-C will be "repaired" by the dominant TFT type, or—in an alternate metaphor—the all-D bubble will have burst.

Although this description is cast for simplicity in terms of the pure strategies and therefore describes the deterministic (Binary) condition, it approximates correctly the dynamics also occuring in the Continuous (noisy) execution condition plotted in this section's figures. With this account it becomes apparent why Figure 5 has the circulation pattern seen in its lower right corner: the lack of pressure against all-C when p is very high allows a drift to higher average q. Once all-Cs are sufficiently dense, all-Ds arising by copying errors can

[21]This drift probably is, in part, a result of forces similar to those that led to an equilibrium of "generous TFT" strategies (i.e., strategies with $p \geq 0.9$ and $q \simeq 0.3$) in the model described by [Nowak and May, 1992]. That model involved a population of heterogenous agents, each playing a strategy selected from a fixed set of (p,q) strategies. Each agent played an infinite length IPD with all others in the population, after which the frequency of use of any particular strategy in the population was changed by applying a simple replicator dynamic.

temporarily gain converts, producing movement down and to the left, as seen around $p = .8, q = .3$. As q falls, however, conversion of all-Ds to TFTs drives p upward again.

Thus for these parameter values (especially error rates, payoff matrix, and numbers of others encountered) we find a steady *self-repairing* regime of cooperation, maintained by a population of agents mostly adhering to the TFT strategy. The population is vulnerable because of the tendency for TFT to drift into all-C, allowing a niche for all-Ds to arise and prosper by "eating" the all-Cs. However, once an all-D has "eaten its seed corn" (the all-C suckers), it cannot spread because under the 2DK interaction process and BMGA adaptive process, the all-D is left in area that consists of a few all-D converts, surrounded by a background of TFT players.

This cycle is a constrained version of a cycle from all-D to TFT to TFT/all-C mix, and back to all-D which has been noted in populations consisting of a few fixed strategies [Nowak and Sigmund, 1989], as well as in populations with continuously variable strategies (cf. [Riolo, 1997b] or [Nowak and Sigmund, 1998]). In this C-2DK-BMGA case the cycle is constrained to a generally high level of cooperation because the combination of the BMGA adaptive process and the context preservation inherent in the 2DK interaction process result in the rapid clearing out of all-C by all-D strategies, which are in turn rapidly replaced by TFT stratagies. Cases which are less *vigilant* about eliminating the potential for all-D infestations are described in Section 8.

A key factor underlying this tight, high-cooperation dynamic, which differs so sharply from the all-D equilibrium of 2DShuffle condition, is the preservation of interaction structure across learning periods that occurs in the 2DKeep condition. Indeed, that is the only difference between the two situations. This property, which we call *context-preservation* makes an enormous contribution to the emergence and maintenance of mutual cooperation. It creates a kind of "shadow of the *adaptive* future." Axelord [Axelrod, 1984] used the notion of the shadow of the future to account for the increase in expected benefits of cooperation that a player would experience if the Prisoner's Dilemma with another player was likely to last for many iterations. In the present comparison there is no variation in how long the game lasts. It is always four moves with each agent encountered. But when agents remain in their 2-D locations over time, the descendants of their current strategies will be playing with the descendants of their neighbors' current strategies. So an effect very like the shadow of the future can occur across learning periods with adapting strategies. This is the key to the expansion of accidental TFT islands, to the drift to higher density of all-Cs among TFTs, and to the self-repair in which TFTs reabsorb brief bubbles of all-D.

Context-preservation in a 2-dimensional space can induce an inter-period shadow of the future strong enough to overcome the forces that sustained an all-D regime in the C-2DS-BMGA condition. Our result on this point aligns with and generalizes a recent finding of Epstein [Epstein-97] which he calls "Demographic Prisoner's Dilemma." He also found that in an adapting population with low mobility in a 2-dimensional space, an accidental pattern of cooperation could diffuse and become established. In the next section we will show that our our results and Epstein's lead to the broader insight captured by what we label "context-preservation." The key attribute of the condition is not the structure of 2-dimensional space. Rather, the key is the relative stability of agents' neighborhoods over time, as induced by the underlying interaction process.

7 Context Preservation is More General than 2-D

Figure 9 (one row from Figure 4) and Table 6 show the average payoffs (over the last 1000 periods) for populations of Continuous strategies, being updated by the BMGA adaptive process, for the six interaction process we have studied. These show the level of cooperation achieved with 2DK and FRNE is almost the same, and the level with FRN is only slightly lower. Using 2DS and RWR interaction processes, cooperation virtually never emerges, whereas with the Tag interaction process, an intermediate level of cooperation is achieved.

Interaction Process	Mean Payoff (sd)	In-Run StdDev (sd)
2DK	2.554 (.009)	0.075 (.006)
FRNE	2.572 (.007)	0.077 (.007)
FRN	2.476 (.026)	0.124 (.064)
Tag	1.613 (0.277)	0.469 (.093)
2DS	1.089 (.003)	0.024 (.009)
RWR	1.098 (.032)	0.063 (.097)

Table 6: A Comparison of Average Payoff from all interaction processes, for the Continuous strategy space and BMGA adaptive process.

Note that there are two pairs of interaction process conditions that differ from one another in just one dimension, namely the interaction process used:

- 2DK and 2DS: With these interaction processes each agent plays exactly four other agents each period (two games with each). The only difference is that in the 2DK case, the agents play the same four others throughout a run, whereas with 2DS the agents pick four new neighbors to play (at random) each period.

- FRN and RWR: In each of these cases, each agent choose four others at random to be neighbors. Thus on average each agent plays 8 other agents (with some variation in the number played across the agents): four others the agent chooses and, on average, four others who choose the agent. The only difference between these cases is that for FRN the agents play the same neighbors throughout a run, whereas for RWR the agents choose new neighbors each period.

In each of these pairings, the context preserving interaction processes (2DK and FRN) result in very high levels of cooperation, whereas the non-perserving cases (2DS and RWR) result in virtually no cooperation.

Also note that the FRNE case produces virtually the same level of cooperation as does the 2DK case, despite the fact that in FRNE interaction process does not use a simple spatial topology as 2DK does. That is, the only difference between these cases is that for 2DK, an agent's neighbors are it's North-East-West-South (NEWS) neighbors, whereas for FRNE, each agent's 4 neighbors are chosen at random. (For both cases, neighbors choose each other, and remain neighbors for an entire run.) This means that with FRNE there are none of the correlations between neighbors-of-neighbors that are induced by a 2DK-type "spatial" topology.

Together these results strongly support our claim that context preservation, not spatial correlation of neighborhoods, is responsible for the observed inter-period "shadow of the adaptive future" that allows the noisy (Continuous) strategy and BMGA adaptive mechanism to sustain cooperation when the interaction process is 2DK, FRNE or FRN, but not when it is 2DS or RWR.

Interaction Process	Mean Payoff
2DK	2.48
FRNE	2.48
FRN	2.15
Tag	1.77
2DS	1.09
RWR	1.09

Table 7: A Comparison of the level of cooperation (Mean Payoff) for the six interaction processes, averaged over all cases (strategy spaces and adaptive processes).

Furthermore, the sixth interaction process, Tag, also is (in general) intermediate in its ability to achieve cooperation, for most cases achieving overall Mean

Payoff rates below the three strong context preserving interaction processes (2DK, FRNE, FRN) and above the mixing cases (2DS, RWR). This intermediacy is not surprising, given the way Tag selection of neighbors works. That is, while 2DK, FRNE and FRN guarantee the exact same neighbors each period, Tag only leads to an increased probability that neighbors will be the same. This is because with the Tag interaction process, neighbors will tend to be chosen from a pool of like-Tagged agents, i.e., from a pool considerably smaller than the entire population size. Thus while some of an agent's neighbors will change from period to period, depending on the number of agents with the same tag, there is a good chance that at least some neighbors will be the same from period to period. In particular, the probability of having neighbors overlap over time will generally be higher than in the 2DS and RWR cases, where neighbors are chosen at random from the entire population.

The comparisons between interaction processes made thus far have focused on just one method of changing strategies (BMGA) and one strategy type (Continuous). Table 7 shows the level of cooperation (average payoff) acheived by the six interaction processes, averaged over both strategy spaces (Binary and Continuous) and all adaptive processes. Again, the strongly context preserving interaction processes (2DK,FRNE,FRN) generate higher levels of cooperation than the non-preserving cases, 2DS and RWR. Further, when the other methods of strategy change are considered (Imitation and 1FGA), the pairwise comparisons listed above also are found again, i.e., 2DK equal to FRNE, 2DK greater than 2DS, and FRN greater than RWR. And while results using Tag-based interaction processes are more variable, showing the effects of interactions with other experimental dimensions, on average over all cases the Tag interaction process results in levels of cooperation that are between those of the strong context preserving cases (2DK, FRNE, FRN) and the non-preservers (2DS, RWR).

In sum, from all these results it is clear that context-preservation is a general and powerful enabling factor for cooperative regimes throughout our experimental data.

8 The Dynamics of Unstable Cooperation

As mentioned in Section 4 and seen in Tables 4 and 5 some cases lead to unstable cooperation: the population cycles endlessly from predominantly defectors to a regime of high cooperation, and back to dominance by defectors. For example, C-FRN-BMGA is a case in which the population has generally high levels of cooperation (see Table 6), but occasionally the population falls to lower payoff levels, as indicated by the FracInHigh for C-FNR-BMGA of 0.939, and as seen in Figure 10. On the other extreme, C-Tag-BMGA is a case in which the level of cooperation is moderate to low (1.613), but the population always achieves an average payoff over the 2.3 threshold, which it can stay above only 0.117 of the time after achieving it. The oscillations in average payoff for C-Tag-BMGA are clearly seen in Figure 11, in which the population seems to flip rapidly back and forth between (unstable) periods of average fitness above 2.0 and periods of all-defection (fitness of about 1.05).

The precise reasons why different cases lead to different dynamics is a result of complex interactions between the three different experimental conditions we are manipulating. It is worth noting here, though, that the underlying unstable dynamic is related to the cycle from defectors to TFT to TFT/all-C mixes and back to defectors, as described briefly in Section 6. This is illustrated by Figures 12 and 13, which show pq-Phase plots for the C-FRN-BMGA and C-Tag-BMGA cases, respectively. These figures clearly show how the population cycles through the average pq-space, with the FRN case centered at high p,q values (and so high cooperation and payoff), whereas the Tag case is centered at low p,q values, with occassional forays into the more cooperative regimes. The effects of strategy spaces, interaction processes and adaptive processes on the dynamics of cooperation in IPD populations are high priorities for future analysis.

9 Discussion

We have seen that the ability of populations to achieve and maintain cooperation is profoundly affected by the processes of interaction that determine "who meets whom." The results reported here show as well that other factors, including both the strategy spaces and adaptive processes used by the agents, also have important effects.

There have been other recent research reports that share our emphasis on *comparative* examination of forces affecting the dynamics of populations of interacting agents, including [Kephart, 1994], [Hoffmann and Waring, 1996], and [Klos, 1997]. The work most closely related to the present investigation is that of Hoffman and Waring. However, there are a number of differences between their experimental setup and the one we have employed. This has the unfortunate consequence of making it difficult compare our results exactly with theirs– though it does have the fortunate consequence that when our observations are roughly parallel we have more hope that the findings are robust.

Hoffmann and Waring simulate agents playing Prisoner's Dilemma with their strategies represented as one- and two-state Moore machines. Their agents' interactions last for fifty iterations. They are located on a one-dimensional ring. The authors do not vary this basic structure, but rather they compare varying "radii" of interaction neighborhoods from the two immediate neighbors up to all N others on the ring. They use only one strategy change process, but again they vary the radius of the neighborhood within which learning agents make comparisons of strategy results.

Their strategy change process is a kind of genetic (evolutionary) algorithm. The algorithm they described blends (crosses over) the Moore machine representation of a located parent with that of a mate selected from the learning neighborhood with a bias toward the better performing strategies. Thus it roughly resembles our compare-to-best-met (BMGA) adaptive process in those cases where learning and interaction neighborhoods are the same and are small relative to N (31 in the cases that best compare to ours). And it partially resembles some of our compare-to-random-other (1FGA) adaptive process when the learning neighborhood is the whole population.

Hoffmann and Waring report that their simulations typically had the bursty dynamics we have often seen, with sharp transitions occuring between cooperative and defecting regimes. They describe a build up of exploitable strategies once a population is heavily cooperative, leading to eventual outbreak of high-scoring by mutant defecting strategies and then to a shift into the opposite basin of attraction. The cycle they sketch is very like the one we have observed.

Hoffmann and Waring also study the effects of varying the degree of localization of interaction and of strategy learning. Localization of interaction is ambiguous in its effects, but localization of learning favors the maintenance of cooperative regimes.

The clearest comparisons can be made with our context-preserving results. In our deterministic (Binary) strategy execution conditions all context preserving histories maintain high levels of cooperation for both the best-met (BMGA) and random-other (1FGA) strategy change processes. So in our Binary cases we have little variance available for comparison. But it is arguable that our noisy strategy execution conditions are the more comparable cases. With p and q values available between 0 and 1 our strategies exhibit more unpredictable variety, as should happen in Hoffmann and Waring's populations composed from the 26 possible two-state Moore machines. In our conditions that have noisy strategy, context-preservation, and errors of comparison and copying, the relatively local BMGA change process does better than the relatively global 1FGA. This holds for all three available pairwise comparisons (2DK, FRN, FRNE). Thus there is a second parallel of our findings and those of Hoffmann and Waring: localized social learning leads to more cooperation than less localized learning. By establishing a shadow of the adaptive future, localized social learning—we might speculate—offers some additional opportunities to "catch" misleading lessons from noisy experience, before they diffuse widely through the population.

Klos [Klos, 1997] makes a related comparison. He reimplements an earlier simulation of iterated Prisoner's Dilemma by Miller [Miller, 1996], altering the genetic algorithm used for strategy change from one using the global population (with bias to the successful) to one using only the same local neighborhood in which interaction occurred. He also finds that cooperative regimes are favored by the more local process of strategy change, which is similar to some of our results.

We have just begun to understand the many regularities and exceptions evident in the aggregrate results shown in Tables 4 and 5. To make further progress will require much additional analysis, both at the aggregate level and most likely at more detailed levels. For example, there may be correlations between the amount of drift toward non-provocable strategies (higher q) and the size of the resulting defector infestation (which in turn causes a drop in cooperation). If so, differences in these correlations across our experimental cases may help us understand how the underlying processes interact to generate the differently shaped cycles in p-q space, e.g., as seen in Figures 12 and 13. More generally, additional analysis will help us understand how interaction and adaptive processes lead to the emergence, maintenance, transformation and dissolution of population-level patterns of activity—in particular, activiy which results in effective social outcomes and structures.

10 Summary

In this paper we have applied a general framework for studying the emergence of patterns of activity—social organization—to the study of cooperation in populations playing the Iterated Prisoner's Dilemma. In particular, we examined the effect of systematically varying three key parameters: (1) the strategy spaces available to agents, (2) the interaction processes that determine which agents interact, and (3) the adaptive processes that determine how agents change strategies as a result of their experience. Thus we reported on aggregate performance measures for 36 combinations of parameters varied over these three dimensions. There clearly are simple regularities in our results, but there are also non-linear interactions producing unexpected high and low levels of cooperation that remain to be fully explained.

Overall, our experiments have confirmed some observations well established in the literature, while also highlighting some surprising results. For example, our results confirm that embedding agents in simple spatial (2D) topologies can promote the emergence of cooperation. On the other hand, we unexpectedly observed high levels of cooperation in populations which agents were randomly mixed between games, when the agents used simple Binary (deterministic) strategies and a kind of global evolutionary algorithm as an adaptive process.

We also observed the emergence of well-known fundamental cycle in the aggregate behavior of populations playing the IPD, in which average level of cooperation oscillates between higher and lower levels. The extent and nature of the oscillations were found to vary in a complicated way as a function the our basic experimental parameters. The oscillations result from a multi-step process in which: (1) defectors are replaced by TFT strategies, which get their start by finding themselves in a fortuitous cluster with other TFT players; (2) the population can drift to include more and more non-provocable (overly cooperative) strategies, since they are safe in a TFT environment; and (3) defectors begin to proliferate, feeding on the overly-cooperative strategies in the population. Under some experimental conditions, the cooperating population quickly repairs itself, e.g., when defectors "eat" the overally-cooperative strategies before there are many of them. In other cases, the population drifts far toward non-provocability, so that when defectors invade, the large supply of "food" allows the defectors to rapidly over-run the population, leading to a deep collapse in the level cooperation.

In this paper we focused on one important factor that contributes to the emergence and maintenance of cooperation, namely the degree to which the interaction processes *preserve the context,* or neighborhood, in which the agent's (and their descendants) act. In general, we found that conditions that strongly preserved context led to the highest levels of cooperation. Further, the concept of context or neighborhood was found to be more general than just 2D spatial topologies. Other interaction processes—for example our FRN condition in which agents have a fixed, but random set of neighbors throughout a run—also produce high levels of cooperation. Similarly, under most conditions, agents who identify neighbors probabilistically, based on arbitrary "tags," also are able to attain higher levels of cooperation than agents who are randomly mixed.

The kind of self-repair mentioned above was found to be prevalent in cases that involve strongly context-preserving interaction processes. The preservation of agent context over time results in a *shadow of the adaptive future*, which acts as a force pushing agents to be more TFT-like, resulting in high levels of cooperation for all the context preserving interaction processes we studied.

Another important factor for the emergence of cooperation is related to the ways in which *variety* in strategies is generated, maintained and constrained. For instance, in general the Imitation adaptive process produces the least amount of cooperation in our experiments. Note that (error-free) Imitation rapidly eliminates all variety of strategies in a population, which means that a population most often will get stuck in very uncooperative regimes. However, in rare situations Imitation will result in very high cooperation (in fact the best over all our cases); this happens in cases where for other reasons it is so easy to achieve cooparation that Imitation is the fastest way to drive the population to full cooperation and sustain it through a lack of variability.

Other regularities are also apparent in the data reported in this paper. Further progess in understanding these regularities, and the surprising exceptions to them, will require additional analysis, both at the aggregate level and most likely at more detailed levels of activity.

Finally, we observe that the core of the system's dynamics cannot be separated from what the agents are doing. If we were using our framework to study marriage formation, HIV transmission, or exchanges of sugar for spice, those activities would imply other arrays of action possibilities and other results of the interactions among the types of agents. Our hope is that some of the results we obtain about effects of variation in interaction patterns and adaptation processes will apply to many different task environments. We see reasons to expect this. An accumulation of results from applying this framework to a range of interaction situations at least will provide a basis for deciding whether this is so. If we are fortunate, the results will provide the basis for more precise generalizations about the mechanisms that promote emergence of social action patterns in populations of agents. If it is not so, then a major extension of the conventional notion of "theory" may be required to bring conceptual order to the study of complex adaptive systems.

Acknowledgements. This research was supported in part by the Intel Corporation, by the Santa Fe Institute and by a grant through the Santa Fe Institute, ONR N00014-95-1-0975, from the Office of Naval Research, acting in cooperation with the Defense Advanced Research Projects Agency. Additional support was provided by the University of Michigan Program for the Study of Complex Systems. The authors also thank the other memebers of the UM BACH group, Arthur Burks, John Holland and Carl Simon, for many fruitful discussions.

References

[Axelrod, 1984] Axelrod, Robert. *The Evolution of Cooperation.* Basic Books. New York.

[Axelrod and Dion, 1988] Axelrod, Robert and Dion, Douglas. "The Further Evolution of Cooperation." *Science, 242,* 1385-1390.

[Cohen, 1982] Cohen, Michael D. "The Power of Parallel Thinking." *Economic Behavior and Organization, 2,* pp 285-306.

[Epstein-97] Epstein, Joshua M. "Zones of Cooperation in the Demographic Prisoner's Dilemma." Santa Fe Insitute Working Paper 97-12-094 E, 1997.

[Giddens, 1984] Giddens, Anthony. *The Constitution of Society: Outline of the Theory of Structuration.* Cambridge, Polity Press, 1984.

[Hoffmann and Waring, 1996] Hoffmann, Robert and Waring, Nigel. "The Localization of Interaction and Learning in the Repeated Prisoner's Dilemma." Sante Fe Institute Working Paper 96-08-064. Santa Fe, NM.

[Holland, 1993] Holland, John. "The Effect of Labels (Tags) on Social Interactions." Santa Fe Institute Working Paper 93-10-064. Santa Fe, NM.

[Holland, 1995] Holland, J. H. *Hidden Order.* MIT Press, Cambridge MA.

[Huberman and Glance, 1993] Huberman, Bernardo A. and Glance, Natalie S. "Evolutionary Games and Computer Simulations." *Proc. Natl. Acad. Sciences (USA), 90,* 7716-18.

[Hutchins and Hazelhurst, 1990] Hutchins, E. and Hazelhurst, B. "Learning in the Cultural Process." In C.G. Langton et al(eds), *Artificial Life II: Studies in the Sciences of Complexity,* Addison-Wesley, Reading MA.

[Kephart, 1994] Kephart, Jeffrey O. "How Topology Affects Population Dynamics." In *Artificial Life III,* pp447-463. C. Langton (ed.) Wiley, NY.

[Klos, 1997] Klos, Tomas B. "Spatially Coevolving Automata Play the Repeated Prisoner's Dilemma." In *Simulating Social Phenomena,* pp. 153-159, Conte, Rosaria, et al. (eds.), Berlin, Springer Verlag, 1997.

[Lindgren and Nordahl, 1994] Lindgren, Kristian and Nordahl, Mats G. "Artificial Food Webs." In *Artificial Life III,* edited by C.G.Langton. Volume XVII of the Santa Fe Institute Studies in the Science of Complexity. Addison-Wesley.

[Lomborg, 1996] Lomborg, Bjorn. "Nucleus and Shield: The Evolution of Social Structure in the Iterated Prisoner's Dilemma." *American Sociological Review, 61,* 278-307.

[Miller, 1996] Miller, John H. "The Coevolution of Automata in the Repeated Prisoner's Dilemma." *Journal of Economic Behavior and Organization, 29 (1),* pp 87-112 (January 1996).

[Mitchell, 1996] Mitchell, Melanie. *An Introduction to Genetic Algorithms.* MIT Press, Cambridge, MA.

[Nakamura, Matsuda and Iwasa, 1997] Nakamura, M, Matsuda, H. and Iwasa, Y. "The Evolution of Cooperation in a Lattice Structured Population." *Journal of Theoretical Biology, 184,* pp. 65-81, 1997.

[Nowak and May, 1992] Nowak, Martin and May, R.M. "Tit for tat in heterogeneous populations." *Nature, 355,* pp250-253.

[Nowak and Sigmund, 1989] Nowak, Martin and Sigmund, Karl. "Oscillations in the Evolution of Reciprocity." *J. Theor. Biology, 137,* 21-26.

[Nowak and Sigmund, 1998] Nowak, Martin and Sigmund, Karl. "Evolution of indirect reciprocity by image scoring." *Nature, 393,* 573-576.

[Oliphant, 1994] Oliphant, Michael. "Evoloving Cooperation in the Non-Iterated Prisoner's Dilemma: The Importance of Spatial Organization." In *Artificial Life IV,* pp349-352. Christopher G. Langton (ed.) Addison-Wesley, Reading MA.

[Poundstone, 1992] Poundstone, William. *Prisoner's Dilemma.* Doubleday, New York.

[Riolo, 1992] Riolo, Rick L. “Survial of the Fittest Bits.” *Scientific American,* July 1992, pp114-116.

[Riolo, 1997a] Riolo, Rick L. “The Effects of Tag-Mediated Selection of Partners in Evolving Populations Playing the Iterated Prisoner’s Dilemma.” Santa Fe Institute Working Paper 97-02-016. Sante Fe, NM.

[Riolo, 1997b] Riolo, Rick L. “The Effects and Evolution of Tag-Mediated Selection of Partners in Populations Playing the Iterated Prisoner’s Dilemma.” In *Proc. of Inter. Conf. Genetic Algorithms (ICGA-97),* Thomas Bäck (ed.), pp 378-385. Morgan Kaufmann, San Fransico.

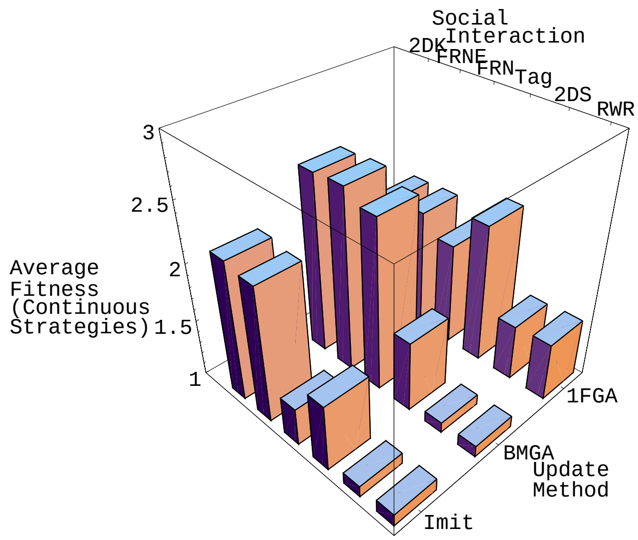

Figure 1: Average Fitness over last 1000 periods for Continuous strategy cases.

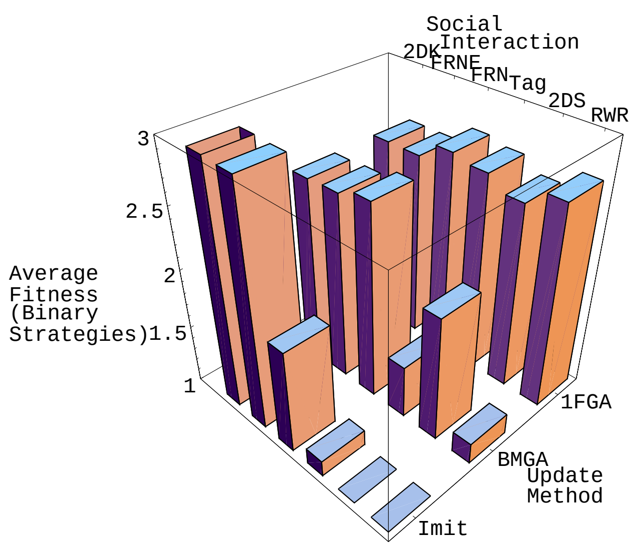

Figure 2: Average Fitness over last 1000 periods for Binary strategy cases.

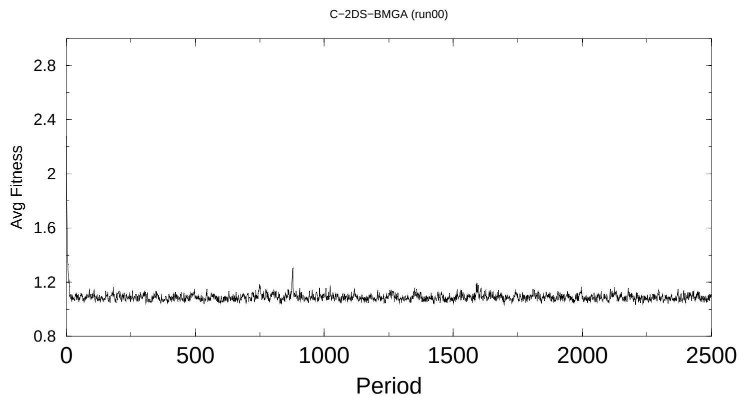

Figure 3: Average fitness versus period for a C-2DS-BMGA run.

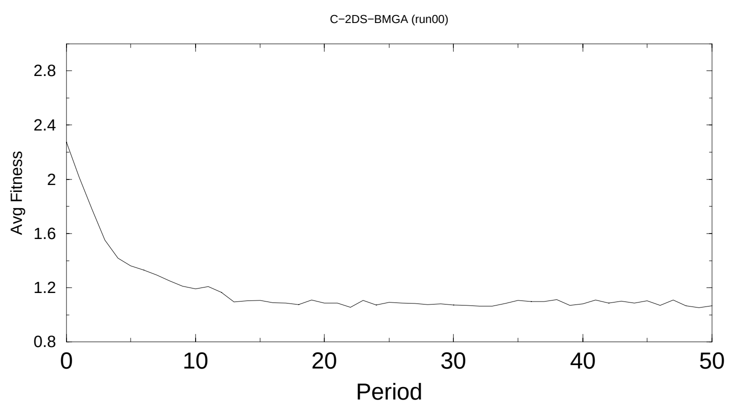

Figure 4: Average fitness for first 50 periods of the C-2DS-BMGA run shown in Figure 3.

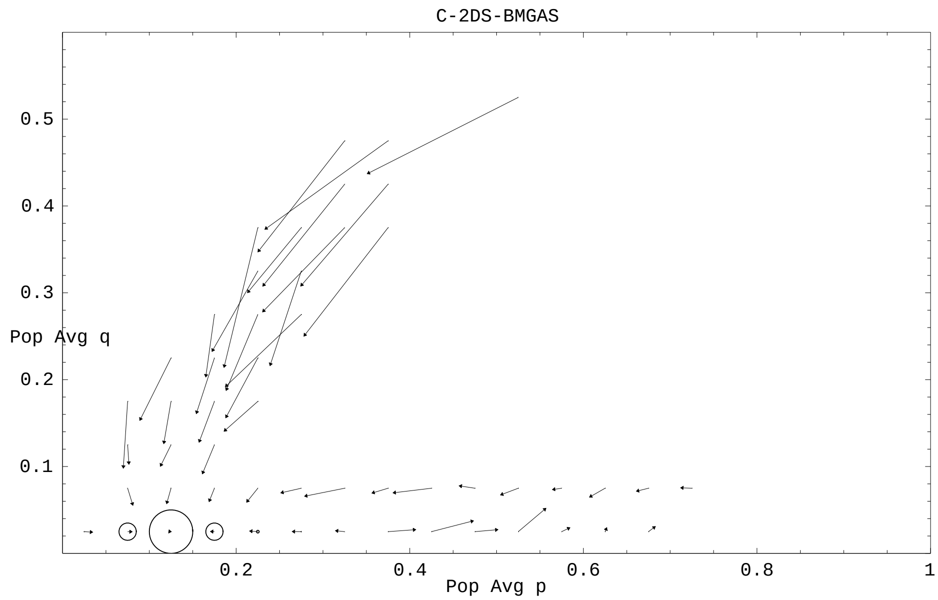

Figure 5: A "pq-Phase plot" for all runs of the C-2DS-BMGA case. The size of each circle indicates the number of periods the population average (p,q) remains in that (p,q) bin, and the vectors indicates the average direction the population average (p,q) changes when it is in that p,q-bin at period T.

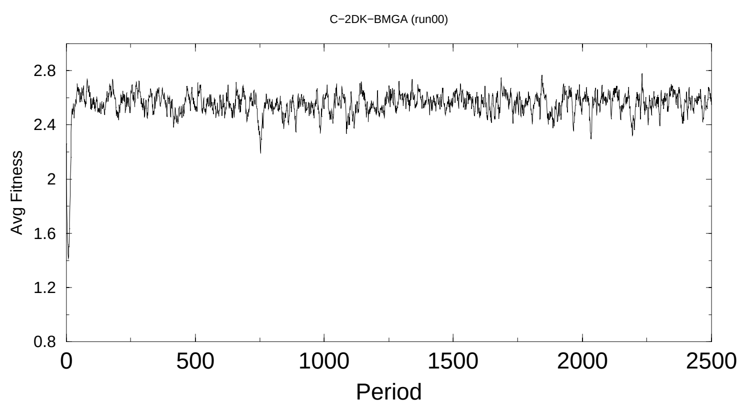

Figure 6: Average fitness versus period for a C-2DK-BMGA run.

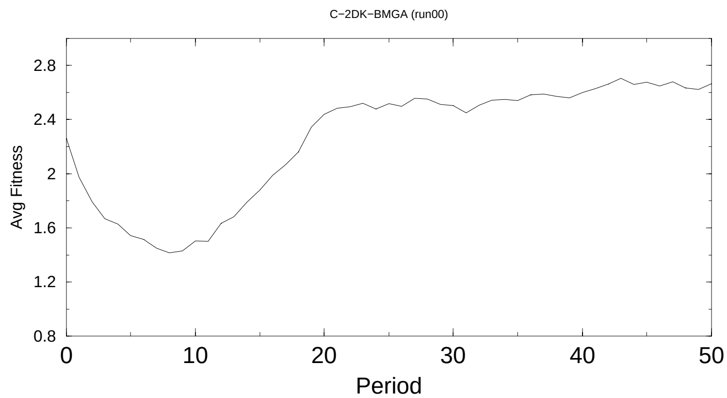

Figure 7: Average fitness versus first 50 periods of C-2DK-BMGA run shown in Figure 6.

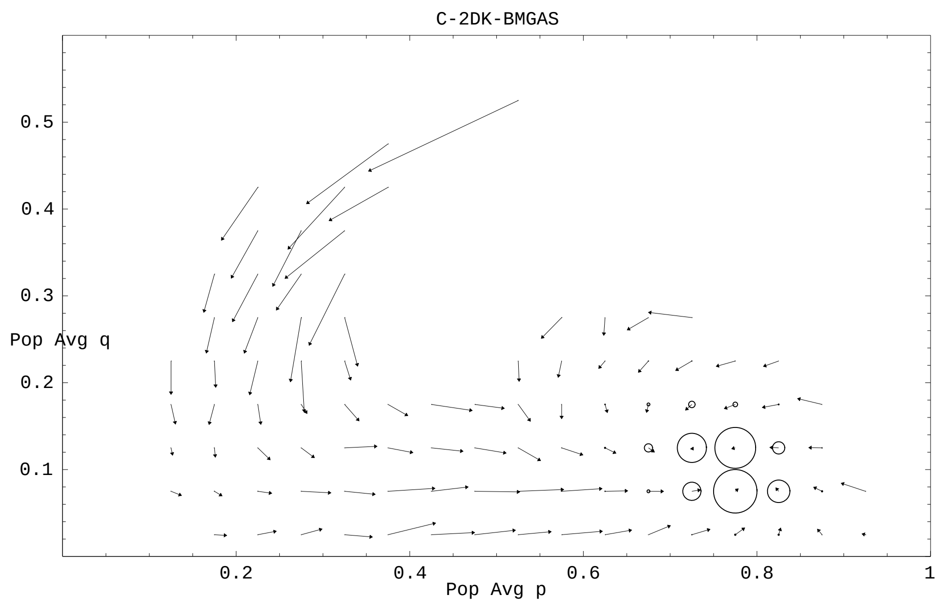

Figure 8: A "pq-Phase plot" for all runs of the C-2DK-BMGA case. The size of each circle indicates the number of periods the population average (p,q) remains in that (p,q) bin, and the vectors indicates the average direction the population average (p,q) changes when it is in that p,q-bin at period T.

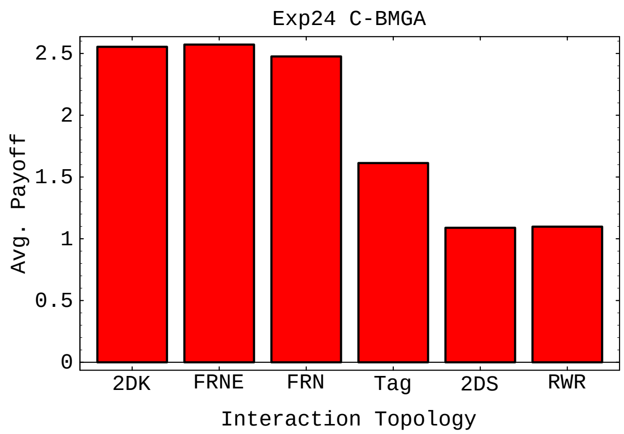

Figure 9: Average payoff (over last 1000 periods) of populations of Continuous strategies, updated by the BMGA adaptive process, for different interaction processes.

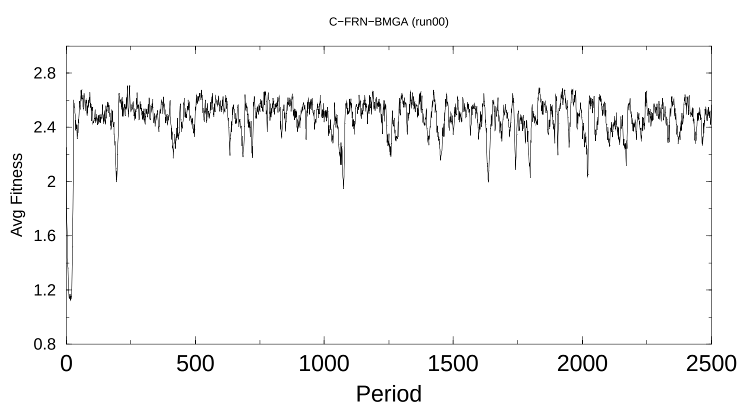

Figure 10: Average fitness versus period for a C-FRN-BMGA run.

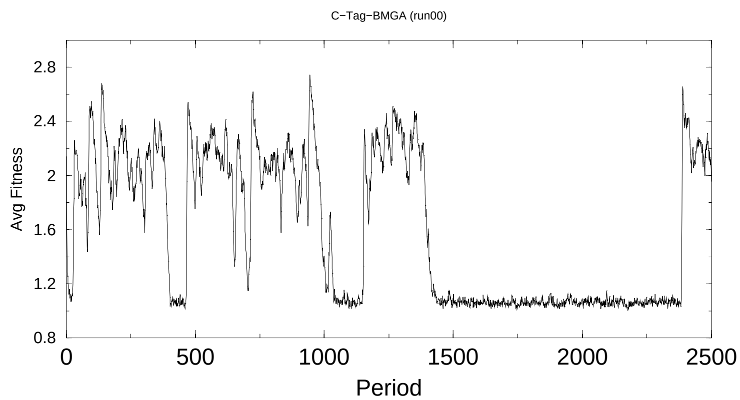

Figure 11: Average fitness versus period for a C-Tag-BMGA run.

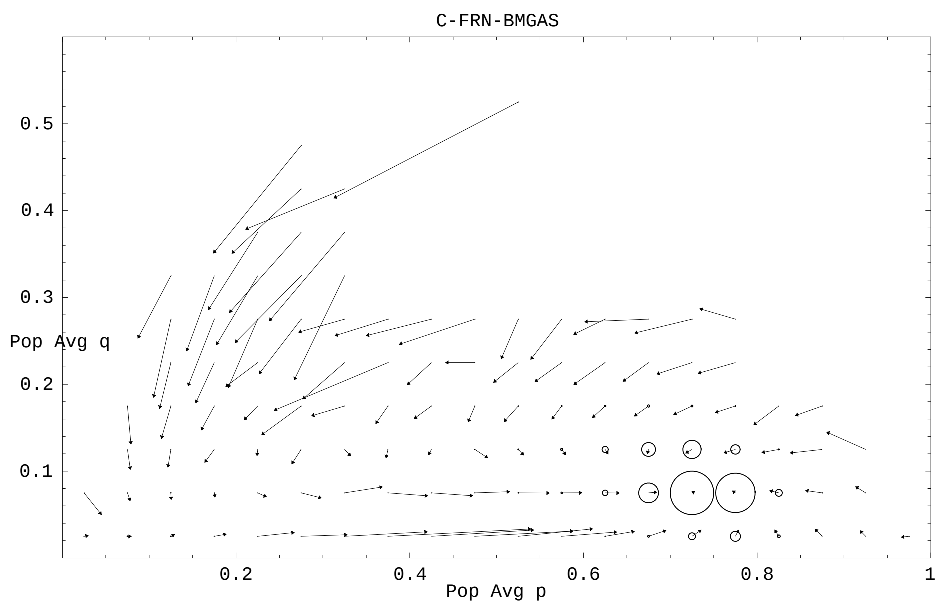

Figure 12: A "pq-Phase plot" for all runs of the C-FRN-BMGA case. The size of each circle indicates the number of periods the population average (p,q) remains in that (p,q) bin, and the vectors indicates the average direction the population average (p,q) changes when it is in that p,q-bin at period T.

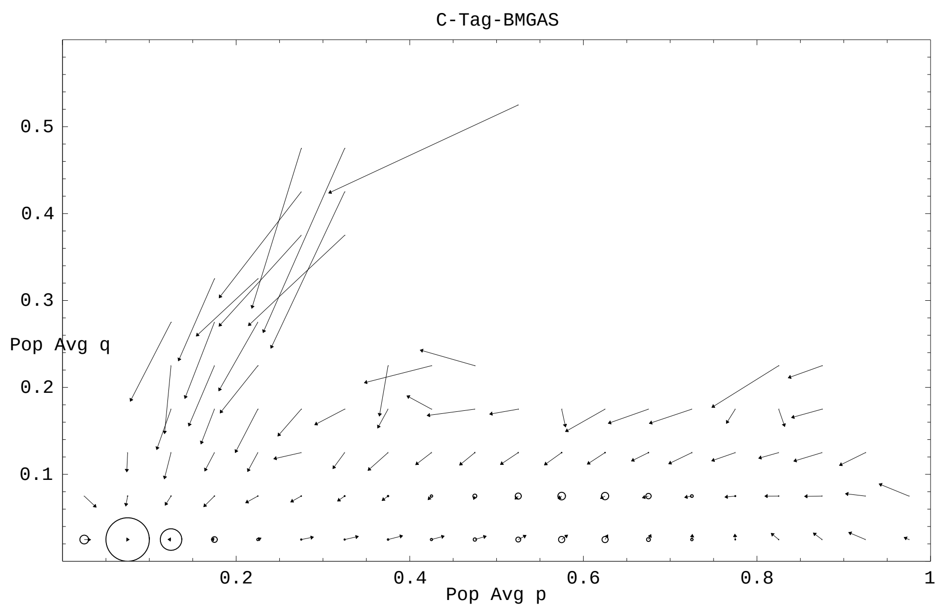

Figure 13: A "pq-Phase plot" for all runs of the C-Tag-BMGA case. The size of each circle indicates the number of periods the population average (p,q) remains in that (p,q) bin, and the vectors indicates the average direction the population average (p,q) changes when it is in that p,q-bin at period T.