

# Spatial Evolution of Automata in the Prisoners' Dilemma

Oliver Kirchkamp\*

30. June 1999

## Abstract

We apply the idea of evolution to a spatial model. Prisoners' dilemmas or coordination games are played repeatedly within neighbourhoods where players do not optimise but instead copy successful strategies.

Discriminative behaviour of players is introduced representing strategies as small automata, which can be in different states against different neighbours. Extensive simulations show that cooperation persists even in a stochastic environment, that players do not always coordinate on risk dominant equilibria in  $2 \times 2$  coordination games, and that success among surviving strategies may differ.

We present two analytical models that help understanding of these phenomena.

Keywords: Evolutionary Game Theory, Networks, Prisoners' Dilemma, Coordination Games, Overlapping Generations. JEL-Code: C63, C73, D62, D83, R12, R13.

---

\*University of Mannheim, SFB 504; L 13,15; D-68131 Mannheim

Support by the Deutsche Forschungsgemeinschaft through SFB 303 and SFB 504 is gratefully acknowledged. I thank George Mailath, Georg Nöldeke, Karl Schlag, Avner Shaked, Bryan Routledge, several seminar participants and two anonymous referees for helpful comments.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>A continuous model</b>	<b>4</b>
<b>3</b>	<b>A discrete spatial model</b>	<b>6</b>
3.1	Space and time . . . . .	6
3.2	Strategies . . . . .	7
3.3	Relevant history for learning . . . . .	8
3.4	Learning . . . . .	8
3.5	Initial State of the Population . . . . .	8
<b>4</b>	<b>A simple discrete model with only two strategies</b>	<b>9</b>
<b>5</b>	<b>Simulation results</b>	<b>11</b>
5.1	Representation of the Results . . . . .	11
5.2	The case of non-discriminative strategies . . . . .	11
5.3	The case of discriminative strategies . . . . .	13
5.3.1	Synchronous interaction and short memory . . . . .	13
5.3.2	Asynchronous interaction and long memory . . . . .	15
5.4	Coordination Games . . . . .	15
5.5	The Influence of Locality . . . . .	16
5.6	Other Dimensions . . . . .	17
<b>6</b>	<b>Concluding Remarks</b>	<b>17</b>
<b>A</b>	<b>Proof of Proposition 1</b>	<b>19</b>

# 1 Introduction

Evolutionary models have in the last years become increasingly popular to study spatial phenomena. These models frequently concentrate on prisoners' dilemmas and coordination games and often share the following assumptions: Players may not discriminate among neighbours, have small memory and interact and learn synchronously. As we will see in the following, the impact of these assumptions may be crucial.

We use a spatial model to describe a world where agents interact with each other to different degrees. An example for a spatial prisoners' dilemma are firms located in space and competing for customers which are placed among these firms. Firms could set high prices, aiming at a monopoly profit if neighbouring firms set high prices as well. Firms could also set low prices hoping to undercut prices of other firms and, thus, stimulating demand for their own product.<sup>1</sup>

In a world without space the above example is a trivial case of Bertrand competition. In a spatial world where firms do not all share the same market, the problem is less trivial. Since Hotelling (1929) a lot of spatial models have been specified and solved. Most of these models assume perfectly rationality. We, however, study a simple evolutionary process. Strategies grow as long as they are successful, unsuccessful strategies vanish from the population.<sup>2</sup> This dynamics has a biological and a social interpretation. Biologically it models a birth process where successful strategies produce more offspring and a death process where unsuccessful strategies are more likely to die soon. In social contexts we assume successful strategies more likely to be copied while unsuccessful strategies are more likely to be abandoned. The evolutionary dynamics will be applied in a spatial context — immobile agents play games with their neighbours and imitate successful neighbours' strategies.

One of the issues we will discuss is the timing of interaction and learning in these models. Most of the literature<sup>3</sup> assumes interaction and learning assumed to be synchronised. We will compare this case with random and asynchronous learning and interaction. Another issue that we will discuss is memory of players and their ability to react or to discriminate. The literature<sup>4</sup> usually assumes that players may not react on their oppo-

---

<sup>1</sup>An example for a spatial coordination game can be found similarly: Neighbouring firms could be interested in finding common standards, meeting in the same market etc.

<sup>2</sup>See e.g. Maynard Smith and Price (1973) for static concepts and Taylor and Jonker (1978) and Zeeman (1981) for a dynamic model)

<sup>3</sup>Axelrod (1984), May and Nowak (1992, 1993), Lindgren and Nordahl (1994), Eshel, Samuelson and Shaked (1998).

<sup>4</sup>May and Nowak (1992, 1993), Eshel, Samuelson and Shaked (1998).

ment's past actions and may not distinguish among their neighbours. We will investigate the impact of this assumption, modelling player's strategies as small automata.

As a result, we find that a common property of non-spatial evolutionary models, equal payoffs of surviving strategies, vanishes in space when players may discriminate among their neighbours. We present simulation results and give a motivation with a simplified discrete model which can be solved analytically. We compare simulation results with a simple continuous model which reliably approximates the discrete model both for prisoner's dilemmas and for coordination games. We identify and study parameters that determine the choice of strategies in prisoners' dilemmas and coordination games. We find that player choose equilibria in coordination games following a criterion that is between risk dominance and Pareto dominance.

We will study a model where a network structure is given, where players are immobile and learn through imitation. Let us briefly mention some related models that take different approaches.

A model where the network structure emerges endogenously is studied by Ashlock, Stanley and Tesfatsion (1994) who allow players to 'refuse to play' with undesired opponents. Their model presupposes no spatial structure *ex ante* — determining the structure endogenously is one way to allow for discriminative behaviour. In our model we take a different approach: The interaction structure is fixed and strategies are small automata which also allow players to treat different opponents differently.

Models where players are mobile are introduced by Sakoda (1971) and Schelling (1971) as well as Hegselmann (1994). These models assume that some properties of an agent (e.g. colour of skin as in Sakoda or a risk-class as in Hegselmann) are invariably fixed but that agents can move. These models offer an explanation for segregation of a population into patches of agents with different types.

Ellison (1993) studies a model where players play coordination games against their neighbours and optimise myopically. As a result these players quickly coordinate on the risk-dominant equilibrium. We will see in section 5.4 below that myopic imitation may choose a different equilibrium.

The assumptions of a given network structure, immobile players and learning through imitation are shared by Axelrod (1984, p. 158ff), May and Nowak (1992, 1993), Bonhoeffer, May and Nowak (1993), Lindgren and Nordahl (1994) and Eshel, Samuelson and Shaked (1998). What is different in our approach is the introduction of discriminative strategies together with asynchronous behaviour.

Axelrod (1984, p. 158ff) analyses a cellular automaton where all cells are occupied by players that are equipped with different strategies. Players obtain each period payoffs from a tournament (which consists of 200 repetitions of the underlying stage game) against all their neighbours. Between tournaments players copy successful strategies from their neighbourhood. This process gives rise to complex patterns of different strategies. Axelrod finds that most of the surviving strategies are strategies that are also successful with global evolution. Nevertheless some of the locally surviving strategies show only intermediate success in Axelrod's global model. These are strategies that get along well with themselves but are only moderately successful against others. They can be successful in the spatial model because they typically form clusters where they only play against themselves.

A similar model has been studied by Lindgren and Nordahl (1994), who in contrast to Axelrod do not analyse evolution of a set of pre-specified strategies, but instead allow strategies to mutate and, hence, allow for a larger number of potential strategies. Similar to Axelrod, Lindgren and Nordahl assume that reproductive success of a strategy is determined by the average payoff of an infinite repetition of a given game against a player's neighbours.

While Axelrod and Lindgren and Nordahl study a population with a large number of possible strategies, May and Nowak (1992) analyse a population playing a prisoners' dilemma where players can behave either cooperatively or non-cooperatively. They study various initial configurations and analyse the patterns of behaviour that evolve with synchronous interaction. Bonnhoeffer, May and Nowak (1993) consider several modifications of this model: They compare various learning rules, discrete versus continuous time and various geometries of interaction.

Eshel, Samuelson and Shaked (1998) derive analytically the behaviour of a particular cellular automaton. The price they have to pay for the beauty of the analytical result is the restriction to a small set of parameters. They have to assume a particular topology and neighbourhood structure. They find that there is a range of prisoners' dilemmas where increasing the gains from cooperation reduces the amount of players who actually cooperate. It is not clear how their result depends on the particular set of parameters they analyse.

In contrast to May and Nowak and Bonnhoeffer, May and Nowak we allow for discriminative behaviour representing strategies as small automata that are identical for a player, but possibly in different states against different neighbours. Introducing discriminative behaviour might seem similar to the repeated game strategies used by Axelrod and by Lindgren and Nordahl. However, there is an important difference. Most of the

literature, including Axelrod and also Lindgren and Nordahl, assumes that updates of players' strategies as well as players' interactions occur synchronously. As we shall see below such a synchronisation has substantial impact on the evolutionary dynamics. In a biological application synchronisation is naturally motivated through seasonal reproduction patterns. In social or economic contexts synchronisations are harder to justify. We, therefore, introduce discriminative strategies without requiring synchronisation.

That introducing asynchronous interaction and evolution might matter has also been mentioned by Huberman and Glance (1993) who argue that cooperation might be extinguished through introduction of asynchronous behaviour. Bonhoeffer, May and Nowak (1993), however, argue that introducing asynchronous behaviour matters only little. As will be seen in the following not the mere fact of introducing asynchronous behaviour determines the amount of cooperation. It is the way how asynchronous behaviour is introduced that determines persistence or breakdown of cooperation.

While the cellular automaton model of a population and the introduction of discriminative behaviour incorporates more real life flavour, it is on the other hand more difficult to analyse. We replace analytical beauty by extensive simulations. We carried out about 60 000 simulations on tori ranging from  $80 \times 80$  up to  $160 \times 160$  and continuing from 1000 to 1000 000 periods. We find that most of the results vary only slightly and in an intuitive way with the parametrisation of the model.

Section 2 describes a continuous model that can be solved analytically and that explains growth of strategies in a spatial but continuous world. The spatial model will be presented in section 3. In section 4 we start with a simplified version, a discrete model that exemplifies permanent exploitation among strategies — a property that is important for spatial evolution. Simulation results will be presented in section 5. Section 5.2 connects our results to the findings of May and Nowak (1992, 1993) comparing several models with non-discriminative strategies. Section 5.3 introduces discriminative behaviour, section 5.4 extends the analysis to coordination games and sections 5.5 and 5.6 discuss robustness of the results. Section 6 finally draws some conclusions.

## 2 A continuous model

In the following we concentrate on symmetric  $2 \times 2$  games and in particular on prisoners' dilemmas. Stage game strategies are named  $C$  and  $D$ . We represent the space of all

symmetric prisoners' dilemmas with the game shown in figure 1.<sup>5</sup>

[Figure 1 about here.]

To make the interpretation of the results presented in section 5 below easier, we will start with a continuous model that aims to explain growth of clusters of strategies.

We will first study a model where patches of homogeneous strategies, that we will call 'clusters', are relatively large. We will then extend this model to capture smaller clusters too.

We assume that players are continuously distributed along a line and interact and learn within a radius  $r$ . Players play the game described in figure 1 and copy the strategy with the highest average payoff in their learning neighbourhood. Let us first assume that the world consists of only two clusters that are both larger than  $2r$ .

Let us consider a player at position  $\hat{x}$  between two of these clusters. All players at position  $x < \hat{x}$  play  $C$ , all players at  $x > \hat{x}$  play  $D$ . A player that lives at position  $x$  meets another  $C$  with the following probability:

$$p_C(x) = \begin{cases} 1 & \text{if } x \leq \hat{x} - r \\ \frac{r + \hat{x} - x}{2r} & \text{if } x \in (\hat{x} - r, \hat{x} + r) \\ 0 & \text{if } x \geq \hat{x} + r \end{cases} \quad (1)$$

The probability to meet a  $D$  is  $1 - p_C(x)$ . Average payoffs  $u_C$  and  $u_D$  for the two strategies can be calculated for the player at  $\hat{x}$  as

$$u_C(\hat{x}) = \frac{3g + h}{4} \quad u_D(\hat{x}) = \frac{1}{4} \quad (2)$$

A player at position  $\hat{x}$  will imitate a  $C$  if  $g > (1 - h)/3$  and a  $D$  if  $g < (1 - h)/3$ . In either case clusters will grow until one strategy disappears completely.

In the above paragraph we have studied a world where strategies were allocated in two big clusters each. In the following we will assume the world consists of a sequence of alternating clusters of  $C$ s and  $D$ s. The  $C$  playing clusters have radius  $n_C r$  where  $n_C \in [1, 2]$  and the  $D$  playing clusters have radius  $n_D r$  where  $n_D \in [1, 2]$ .<sup>6</sup> A player who lives precisely between two clusters sees for both clusters a proportion of

$$F(n) = \frac{1}{4}(5 - 4n + n^2) \quad (3)$$

---

<sup>5</sup>All the dynamics of population behaviour that are discussed in the following are invariant to transformations of payoffs like adding constants or multiplying with a positive number.

<sup>6</sup>The precise analysis of even smaller clusters becomes really tedious. We think that it is reasonable to assume that our approach approximates the behaviour of smaller clusters where  $n < 1$ .

(for  $n = n_C$  and  $n = n_D$ ) playing against the respective opposite strategy. The remaining  $1 - F(n)$  play against their fellows which are in the same cluster. Average payoffs are

$$u_C(\hat{x}) = (1 - F(n_C))g + F(n_C)h \quad u_D(\hat{x}) = F(n_D) \quad (4)$$

The difference  $u_C(\hat{x}) - u_D(\hat{x})$  is a quadratic function of  $n_C$  and  $n_D$ :

$$u_C(\hat{x}) - u_D(\hat{x}) = g + \frac{1}{4}(h - g)(5 - 4n_C + n_C^2) - \frac{1}{4}(5 - 4n_D + n_D^2) \quad (5)$$

If  $u_C(\hat{x}) - u_D(\hat{x})$  is positive then  $n_C$  will grow while  $n_D$  will shrink and vice versa. For a given prisoners' dilemma the region where  $u_C(\hat{x}) > u_D(\hat{x})$  is the interior of an ellipsis around  $(n_C, n_D) = (2, 2)$ . Assume that a population starts from a distribution of clusters inside the ellipsis where  $u_C(\hat{x}) > u_D(\hat{x})$ . Then  $n_C$  will grow until at boundary of the ellipsis  $u_C(\hat{x}) = u_D(\hat{x})$  where we have a stable small cluster of  $D$ s next to a large cluster of  $C$ s.

Solving  $u_C(\hat{x}) > u_D(\hat{x})$  for  $g$  we find that player  $\hat{x}$  will become a  $C$  if

$$g > \frac{-(5 - 4n_D + n_D^2) + h(5 - 4n_C + n_C^2)}{1 - 4n_C + n_C^2} \quad (6)$$

Similar to equation 2 also equation 6 gives us a condition that determines when clusters of  $C$  or  $D$  grow, shrink or remain stable. We will use these conditions below to represent our results.

In the following we will move to a discrete model that will be described in section 3. Players are of discrete size, we will most of the time consider a two dimensional network and the size of the clusters is determined endogenously. We will see that the continuous model of the current section gives a fairly good approximation of the properties of the discrete model that we will study in the following.

## 3 A discrete spatial model

### 3.1 Space and time

Players are assumed to live on a torus<sup>7</sup>. The torus is represented as a rectangular network, e.g. a huge checkerboard. The edges of this network are pasted together. Each cell on the network is occupied by one player. Using a torus instead of a plain checkerboard avoids boundary effects. Most simulations were carried out on a two-dimensional torus of size

---

<sup>7</sup>In section 5.6 we investigate some other topologies where players are located on a circle, on a cube, or on a hypercube in four dimensions.



$80 \times 80$ . The results presented in sections 5.5 and 5.6 show that neither the exact size of the network nor the dimension matter significantly.

In contrast to models of global evolution and interaction we assume players to interact only with their neighbours and learn only from their neighbours. Each player has an ‘interaction neighbourhood’ of radius  $r_I$  as shown in figure 2 which determines the set of the player’s possible opponents.

[Figure 2 about here.]

In the same way we construct a ‘learning neighbourhood’ with a similar structure, but with a possibly different radius  $r_L$ . Each player observes payoffs and strategies within the learning neighbourhood.

We will compare synchronous with asynchronous interaction and learning. With synchronous interaction players interact in each period with all their neighbours. With asynchronous interaction a random draw decides in each period for each pair of neighbours whether the mutual interaction takes place. Hence, with asynchronous interaction in different periods the composition of opponents may be different. With synchronous learning all members of a population will learn simultaneously every  $t_L$  periods. With asynchronous learning, strategies each have an individual random ‘lifetime’  $t_L$ . As soon as the lifetime of the strategy expires, the individual player learns and possibly adopts a new strategy with again a random lifetime which may differ from the previous one.

## 3.2 Strategies

Strategies may be discriminative or non-discriminative, i.e. players may or may not be able to treat different neighbours differently. With non-discriminative strategies players either play  $C$  against all neighbours or they play  $D$  against all of them. Discriminative strategies are modelled as Moore automata with less than three states.<sup>8</sup> We have done several simulations with more complex automata and found that results do not change much when the number of states is increased to three or four.

---

<sup>8</sup>Each ‘state’ of a Moore automaton is described with the help of a stage game strategy ( $C$  or  $D$ ) and a transition function to either the same or any other of the automaton’s states. This transition depends on the opponent’s stage game strategy. Each automaton has one ‘initial state’ which the automaton enters when it is used for the first time. There are 2 automata with only one state (one of them initially plays  $C$  and remains in this state whatever the opponent does; the other always plays  $D$ ). There are 26 automata with one or two states. For example ‘grim’ is a two-state automaton. In the initial state it plays  $C$ . It stays there unless the opponent plays  $D$ . Then the automaton switches to the second state, where it plays  $D$  and stays there forever. Other popular two-state automata include ‘tit-for-tat’, ‘tat-for-tit’, etc.

### 3.3 Relevant history for learning

To choose a strategy players have to learn. To do that they use information available in their neighbourhood: Their neighbours' average payoffs (per interaction) and their strategies. We compare two extreme cases: Learning players may consider only today's payoffs when calculating average success of a neighbour's strategy or learning players consider payoffs from all periods while the neighbour used the same strategy as today. The first case will be called 'short memory' and the second 'long memory'.

### 3.4 Learning

We will compare two learning rules. Both are based on the assumption that players are not fully rational, they are not able or not willing to analyse games, and they do not try to predict their opponents' behaviour. Instead, behaviour is guided by imitation of successful strategies. 'Copy best player' copies the strategy of the most successful player in the neighbourhood.<sup>9</sup> 'Copy best strategy', instead, copies the strategy that is on average (per interaction) most successful in the neighbourhood.<sup>10</sup> Players that continue to use their old strategy after learning took place continue to use it in which states it currently is against the different neighbours. Players that learn a new strategy start using this strategy in the initial state.

### 3.5 Initial State of the Population

The network is initialised randomly. First relative frequencies of the available strategies are determined randomly (following an equal distribution over the simplex of relative frequencies) and then for each location in the network an initial strategy is selected according to the above relative frequencies of strategies. Hence, simulations start from very different initial configurations.

---

<sup>9</sup>If more than one player has highest average payoff, a random player is picked, unless the learning player himself or herself is already one of the 'best players'. In the latter case the learning player's strategy remains constant.

<sup>10</sup>If more than one strategy has highest average payoff, a random strategy is picked, unless the learning player himself or herself uses already one of the 'best strategies'. In the latter case the learning player's strategy remains constant.

## 4 A simple discrete model with only two strategies

Before we start presenting simulation results let us first discuss a simpler model that allows us to understand an interesting property of the discrete model, mutual exploitation.

For the simplified model that we discuss in this section we make the following assumptions:

1. There are only two strategies present in the population: Grim and blinker.

Grim is a strategy that starts with  $C$  and continues to play  $C$  unless the opponent plays at least one  $D$ . From then on grim will play  $D$  forever. Blinker alternates between  $C$  and  $D$  all the time.

2. Players interact with each of their opponents in each period exactly once.
3. Players do not observe their neighbours' payoffs but average payoffs of strategies over the whole population. They use the 'copy best strategy' learning rule, i.e. when they learn they copy the strategy with the highest average payoff (per interaction) in the population.
4. Players will learn very rarely. The individual learning probability in each period is  $\epsilon$ . We assume that  $\epsilon \rightarrow 0$ .

The most crucial departure from our simulation model is assumption 3: Learning is based on average payoffs of the whole population and not of the individual neighbourhood. This simplifies the analysis drastically. Simulations below show that the assumption is harmless.

We will denote possible states of grim with  $G^C$  and  $G_D$ , depending on whether grim is in its cooperative or its non-cooperative state. Similarly we denote possible states of blinker with  $B^C$  or  $B_D$ . The average payoff per interaction of grim and blinker in a given period is denoted with  $\bar{u}_G$  and  $\bar{u}_B$ . In the following we will describe an evolutionary dynamics of a population that has a cyclical structure. We define a cyclical equilibrium as follows:

**Definition 1 (cyclical equilibrium)** *A cyclical equilibrium is a sequence of states of a population that, once such a sequence is reached, it repeats forever.*

In the model that we consider below, all of our cycles have length one or two. We will not only consider cycles of complete states of the population, but also cycles of average payoffs

or cycles of relative frequencies of strategies or pairs of strategies. We will denote a cycle of any two objects  $a$  and  $b$  with the symbol  $\lceil ab \rceil$ . When we mention several cycles in the same context, like  $\lceil ab \rceil$  and  $\lceil cd \rceil$  then we want to say that  $a$  happens in the same period as  $c$  while  $b$  happens together with  $d$ . A cycle of length one can be denoted as  $\lceil a \rceil$ . When we mention  $\lceil a \rceil$  together with  $\lceil bc \rceil$  we want to say that a population alternates between a state where  $a$  happens together with  $b$  and another state where  $a$  happens together with  $c$ .

**Proposition 1** *Under the above assumptions there are four possible classes of cyclical equilibria.*

1. *A population that contains any proportion of  $\lceil \langle G^C, G^C \rangle \rceil$  and  $\lceil \langle G_D, G_D \rangle \rceil$  is in a cyclical equilibrium.*
2. *A population that contains any proportion of the cycles of pairs  $\lceil \langle B^C, B^C \rangle \langle B_D, B_D \rangle \rceil$ ,  $\lceil \langle B_D, B_D \rangle \langle B^C, B^C \rangle \rceil$  and  $\lceil \langle B^C, B_D \rangle \rceil$  is in a cyclical equilibrium.*
3. *There are cyclical equilibria where in both periods average payoffs of grim and blinker are equal.*
4. *There is a single cyclical equilibrium where both grim and blinker are present and have different payoffs at least in some periods.*

*This equilibrium has the property that payoffs alternate between  $\bar{u}_G = \epsilon \cdot (1 + 2h)/2 < \bar{u}_B = \epsilon/2$  and  $\bar{u}_G = 1/2 > \bar{u}_B = (g + h)/2$ .<sup>11</sup> 1/4 of the population consists of pairs  $\lceil \langle G_D, G_D \rangle \rceil$ , 1/2 of the population consists of a cycle of pairs  $\lceil \langle G_D, B_D \rangle \langle G_D, B^C \rangle \rceil$  and 1/4 of the population consists of a cycle of pairs  $\lceil \langle B_D, B_D \rangle \langle B^C, B^C \rangle \rceil$ .*

The proof is given in appendix A.

The interesting case is case 4. It shows how one strategy earns substantially less than average payoff but still does not die out. Both strategies replicate in exactly the same number of periods and at the same speed alternating in having more than average payoff. For reproductive success it does not matter that one of the two strategies is only an  $\epsilon$  more successful when it replicates, while the other replicates with a substantial payoff difference.

We will come back to this argument in the discussion of figure 8 below.

---

<sup>11</sup> $g$  and  $h$  are parameters of the underlying game (see figure 1).

## 5 Simulation results

### 5.1 Representation of the Results

To explain the representation of the results let us look at figure 3.

[Figure 3 about here.]

The figure shows the results of 800 different simulations, each for a different (randomly chosen) game, and each lasting for 2000 periods. The parameterisation for games follows figure 1. Values  $g$  and  $h$  are chosen such that most of these games are prisoners' dilemmas. Relative frequencies of  $CC$ ,  $CD$ , and  $DD$  are represented as circles of different sizes in each of the three diagrams (one circle for each of the 800 simulations). The position of the circle indicates the payoffs  $g$  and  $h$  and, hence, the game. The size of the circle in the first diagram is proportional to the average number of observed  $CC$ s during the last 50 periods of the simulation, in the second diagram proportional to  $CD$  (and  $DC$  likewise) and in the third diagram proportional to  $DD$ . If the frequency of  $CC$ ,  $CD$  or  $DD$  is zero, no circle is drawn.

Following the discussion in section 2 we represent the games where players at a boundary between two clusters are indifferent with gray lines. The solid gray line corresponds to the case  $n_C = n_D \geq 2$ , the dashed gray lines represent  $n_C = n_D = 1.5$  and  $n_C = n_D = 1.25$ . Borderline players will start cooperating above the line, below they will play  $D$ . If clusters shrink ( $n_C$  and  $n_D$  are smaller) the region where players at the boundary between two clusters are indifferent moves upwards.

For prisoners' dilemma games this means that the smaller clusters are, the smaller is (ceteris paribus) the set of prisoners' dilemmas where we should expect cooperation. As long as cluster size remains constant for all games, we should expect the borderline of cooperative behaviour to have a linear shape as described by inequality 6. Our simulations in the following show that this borderline has indeed an almost linear shape.

### 5.2 The case of non-discriminative strategies

In the following section we study a simple setup where strategies are non-discriminative, i.e. players play either  $C$  against all their neighbours or they play  $D$  against all of them. We will first apply the model of Nowak and May to a larger variety of games assuming short memory, synchronous interaction and synchronous learning. We will then follow a critique raised by Huberman and Glance who argue that cooperation, which survives in Nowak and May's model, may disappear with a stochastic context. We will replicate their

finding but we will also point out that their argument requires an additional assumption. There may be room for cooperation even within a stochastic framework, as long as player's memory is large enough.

In figure 3 we extend Nowak and May's analysis to a larger variety of payoffs. We see that the cooperative area is small and close to  $g = 1$  and  $h = 0$ , i.e. close to a range of payoffs where cooperation does not cost too much. The borderline between cooperation and non-cooperation follows the approximation given in section 2 for a cluster size of about  $1.2r$ . Local dynamics with non-discriminative strategies explain deviation from the Nash equilibrium only to a small extent.

A small number of simulations yield mutual cooperation for smaller values of  $h$ . Closer inspection shows that this is the case for particular initial configurations. Most of the simulations in this area lead to populations playing  $DD$  all the time (see the bottom left part of figure 3).

Huberman and Glance (1993) criticise May and Nowak's model arguing that due to the synchronous dynamics the network might run into cycles that are unstable against small perturbations. They suggest random sequential learning to model more realistic timing. While in each period all possible interactions take place, only one single player learns in a given period. The order in which players learn is determined randomly. In a particular setting<sup>12</sup>, where May and Nowak found cooperation, random sequential learning leads to general non-cooperation. They claim that synchronisation is an artificial assumption and can hardly be justified.

[Figure 4 about here.]

In figure 4 we find that they are right. It is indeed sufficient to go only a small step<sup>13</sup> into the direction of the model proposed by Huberman and Glance to notice that cooperation dies out almost completely with short memory and asynchronous interaction.

However, the assumption of short memory made by Huberman and Glance is crucial. Once long memory is introduced (see bottom part of figure 4) we find that in both cases, with synchronous or with asynchronous interaction, a range of games where cooperation survives.

---

<sup>12</sup>The initial configuration, game, learning rule etc. corresponds to figure 3.a–c of May and Nowak (1992).

<sup>13</sup>Huberman and Glance give each period at most one player an opportunity to change the strategy and players interact each period synchronously, i.e. each possible interaction takes place with certainty. In our setting each possible interaction takes place with probability  $1/2$  and players learn randomly after  $20 \dots 28$  periods. As we observe in figure 4, the precise way how an asynchronous environment is modelled does not matter. Cooperation dies out both in Huberman and Glance's and in our simulations.

We have the following intuition for this phenomenon: A cooperative player can only learn to become a  $D$  when there is at least one  $D$  in the neighbourhood. In such a period, however,  $C$ s are particularly unsuccessful. A  $C$  with long memory may keep in mind that sometimes there were more  $C$ s being more fortunate and finds  $C$  still more attractive than  $D$ . A  $C$  with short memory only looks at current payoffs, which are bad for  $C$  at the moment and, hence, is more likely to become a  $D$ . Hence, with short memory, cooperation is more vulnerable.

The above simulations were done with the learning rule ‘copy best player’ since this was also the learning rule used by Nowak and May. This particular choice is not essential. In figure 5 we show the results for the ‘copy best strategy’ learning rule. It turns out that changing the learning rule has little impact on the relative frequency of mutual cooperation.

[Figure 5 about here.]

### 5.3 The case of discriminative strategies

The non-discriminative strategies discussed above forced players to treat their neighbours without distinction. Either they had to play  $C$  or they had to play  $D$  against all of them. We will now introduce discriminative behaviour. Players’ strategies will be described as small (two-state) automata. Each player uses only one automaton, but this automaton can be in different states against different neighbours. Results change substantially now.

We begin in section 5.3.1 with a simple model: Synchronous interaction and learning from short memory. Starting with this approach allows us to understand more easily an interesting property of spatial evolution: mutual exploitation. In section 5.3.2 we proceed to a model where players interact stochastically and learn from long memory.<sup>14</sup> The dynamics in the latter model may be harder to analyse, but are also more robust and sometimes more convincing.

#### 5.3.1 Synchronous interaction and short memory

In the following section we assume that learning is asynchronous but that interaction is still synchronised. Players learn from short memory. Let us first look at figure 6 which shows simulation results for a population that uses discriminative strategies. Players use the learning rule ‘copy best strategy’ as defined in section 3.4.

[Figure 6 about here.]

---

<sup>14</sup>The intermediate case of stochastic interaction with short memory is described in Kirchkamp (1995).

We find a larger range of games where cooperation survives with discriminative strategies (in figure 6) than with non-discriminative strategies (see figure 5). Discriminative strategies, for instance members of a cluster of tit-for-tats, are able to cooperate with their kin in their neighbourhood, but will not be put-upon by  $D$ s. Notice that the range of games where  $CC$  prevails is mainly determined by the value of  $g$ , i.e. by the payoff for  $CC$ . The value of  $h$  for  $CD$  does not have much influence. Reproduction mainly occurs upon successful  $CC$ s when past  $CD$ s are already forgotten. This will change when we introduce long memory in section 5.3.2.

In section 4 we introduced a simple analytical model that could explain how a discrete world supports mutual exploitation. In the following paragraph we will corroborate this finding with simulation results. We will consider a population that initially contains all 26 discriminative strategies in random proportions. Two of them, *grim* and *blinker*, despite their different average payoff, appear together for a wide range of games.

[Figure 7 about here.]

The left part of figure 7 shows that both *grim* and *blinker* survive for a wide range of payoffs. In the right part we see that *grim*'s payoff is higher than average for most games while *blinker*'s payoff is smaller than the average population payoff. We did not observe this phenomenon with non-discriminative strategies, the reason being that in some games all agents played  $C$ , in the others they all played  $D$ . Hence, they all had the same payoff.

Notice that the above described heterogeneity of payoffs differs from the heterogeneity of strategies pointed out by Lindgren and Nordahl (1994). They find that spatial evolution yields in contrast to global evolution starting from a homogeneous initial state in the long run a heterogeneous state in the sense that several different strategies are distributed in 'frozen patchy states' over the network. In the current and in the following section we show two related things: First, even a population which has not reached a static state (and will possibly never reach one) can exhibit stable properties, e.g. relative frequencies of different strategies that may achieve a stable level. Second, in addition to strategies that are distributed heterogeneously over the network, also payoffs of strategies are heterogeneous.

As in the simple model in section 4 we find a symbiosis of two automata where one gains on average substantially more payoff. Figure 8 shows that the *grim* strategies alternate between a small amount of support for their environment, playing  $CD$ , and substantial exploitation with  $DC$ .

[Figure 8 about here.]

For the above result we assumed synchronous interaction and short memory to facilitate comparison with the simple dynamics in section 4. We will see below that neither syn-



chronous interaction nor short memory are crucial for the above phenomenon. We skip the case of asynchronous interaction with short memory (which is very similar to the above case and which is discussed in Kirchkamp (1995)) and turn immediately to asynchronous interaction with long memory in the next section.

### 5.3.2 Asynchronous interaction and long memory

In section 5.3.1 we have analysed an evolutionary process with learning from short memory and synchronous interaction. In the current section we will assume asynchronous interaction and long memory. Figure 9 shows relative frequencies of  $CC$ ,  $CD$  and  $DD$ . With long memory the frequency of  $CD$  is smaller for most games than with short memory (see figure 6). Further the range of games where  $CC$  is played has a different shape.

[Figure 9 about here.]

Let us first explain the smaller amount of  $CD$ . With long memory past losses remain in memory. What was successful with short memory, alternating between exploiting another strategy a lot and supporting it only a little, no longer ensures its replication. As a result we find in figure 10 that a strategy which could permanently be exploited with short memory, blinker, now disappears for most prisoners' dilemmas. Exploitation, however, persists. In most games where  $CD$  is played, grim still obtains more than average payoff.

[Figure 10 about here.]

Let us now turn to the shape of the range of games where cooperation survives. With short memory this range of games was mainly determined by the value of  $g$ , i.e. by the payoff for  $CC$ . With long memory, however, the experience of  $CD$ s remains in memory until reproduction occurs, and hence counts for the choice of strategies.

## 5.4 Coordination Games

The same analysis can be applied not only to prisoners' dilemmas, but also to coordination games. The parameterisation for symmetric coordination games that we have chosen is shown in figure 11.

[Figure 11 about here.]

The game has two pure equilibria, one where both players play  $C$ , and another where both players play  $D$ . Both players would prefer to coordinate on one of the two. Harsanyi and Selten (1988) propose two criteria to select among these equilibria. Risk dominance

and Pareto dominance. Unfortunately these two criteria do not always pick the same equilibrium.

Kandori, Mailath and Rob (1993) and Young (1993) suggest that in a global model where players optimise myopically, evolution selects the risk dominant equilibrium in the very long run. Ellison (1993) studies a spatial model where players optimise myopically and, hence, select the risk dominant equilibrium even faster than in the model of Kandori, Mailath and Rob (1993) or Young (1993).

In section 2 we have already studied a simplified model of a continuum of players and considered the situation of a learning player whose left neighbours all play  $C$  and whose right neighbours all play  $D$ . A myopically optimising player calculates that chances to meet either a  $C$  or a  $D$  are equally  $1/2$  and therefore selects the risk dominant equilibrium. A player who, however, imitates successful strategies will in the same situation become a  $C$  if

$$g > (1 + h) \frac{n - 4}{4 + n} \quad (7)$$

where  $n \leq 2$  describes the size of the cluster. Clusters with a diameter larger than 2 times the neighbourhood radius can be treated as  $n = 2$ . Since coordination games favour large clusters we concentrate on the case  $n = 2$ .<sup>15</sup> The gray line in figure 12 describes the set of games where  $g = -(1 + h)/3$ , i.e. the case where for  $n = 2$  we have  $g = (1 + h) \cdot (n - 4)/(4 + n)$ .

[Figure 12 about here.]

It is almost precisely the line  $g = -(1 + h)/3$  which divides the region where only  $C$  is played from the region where only  $D$  is played.  $CD$ 's which fail to coordinate are extremely rare. The continuous model described in section 2 explains the behaviour of the discrete population substantially better than Pareto dominance or risk dominance.

## 5.5 The Influence of Locality

So far we assumed that learning and interaction was within a neighbourhood of radius 1. Figure 13 shows the amount of mutual cooperation for larger neighbourhoods.

[Figure 13 about here.]

Starting from  $r_I = r_L = 1$  we enlarge neighbourhoods for learning and interaction simultaneously until we reach global interaction and learning. With less locality the range

---

<sup>15</sup>In a prisoners' dilemma clusters are often small since only small groups of defectors are successful. As soon as they start growing they kill themselves.

of games where  $CC$  is played becomes gradually smaller. We have done further simulations where we keep the learning radius  $r_L$  constant and vary the interaction radius  $r_I$  or vice versa and found that both locality of evolution and locality of interaction smoothly increase the range of games where cooperation prevails.

## 5.6 Other Dimensions

Beyond the results that we presented above we have carried out many further simulations, each showing that the effects that we pointed out here persist for several modifications of the model. One possible modification is the change from a two-dimensional torus to tori with other dimensions.

[Figure 14 about here.]

Figure 14 shows the relative frequency of mutual cooperation in one- to four-dimensional networks. All networks have 4096 players and all neighbourhoods have almost the same number (80...125) of opponents. We see that the dimension of the network has almost no effect on the size of the cooperative area.

## 6 Concluding Remarks

We have analysed a simple model of spatial evolution where agents repeatedly play prisoners' dilemmas and coordination games.

We investigate the range of games where cooperative strategies survive. One might expect cooperation to be easier to sustain if players use simple strategies. Players that are capable of using more complex strategies might move quicker to a more rational solution. We find, however, that more complexity fosters cooperation.

We further observe inequality of payoffs among coexisting strategies. While global evolution predicts that strategies with lower than average payoff die out, local evolution explains symbiosis of strategies where one strategy alternates between slightly feeding and substantially exploiting the other.

We respond to a criticism raised by Huberman and Glance (1993) who argue that asynchronous timing might eliminate cooperation completely. We find that their argument crucially depends on their assumptions about memory of players. As soon as players have long memory cooperation survives both with asynchronous as well as with synchronous timing.

We further extend the analysis of Huberman and Glance and Nowak, Bonnhoeffer and May (1993) allowing not only for asynchronous learning but also for asynchronous interaction. We find that this avoids that the evolutionary process gets stuck in odd patterns of behaviour.

We find that some properties of a discrete local model can be approximated by a continuous one that is easier to analyse. The approximation of the continuous model is surprisingly accurate in particular for coordination games.

## References

- ASHLOCK, D., E. A. STANLEY, AND L. TEFATSION (1994): “Iterated Prisoner’s Dilemma with Choice and Refusal of Partners,” in *Artificial Life III*, ed. by C. G. Langton, no. XVII in SFI Studies in the Sciences of Complexity, pp. 131 – 167. Addison-Wesley.
- AXELROD, R. (1984): *The evolution of cooperation*. Basic Books, New York.
- BONHOEFFER, S., R. M. MAY, AND M. A. NOWAK (1993): “More Spatial Games,” *International Journal of Bifurcation and Chaos*, 4, 33–56.
- ELLISON, G. (1993): “Learning, Local Interaction, and Coordination,” *Econometrica*, 61, 1047–1071.
- ESHEL, I., L. SAMUELSON, AND A. SHAKED (1998): “Altruists, Egoists, and Hooligans in a Local Interaction Model,” *The American Economic Review*, 88, 157–179.
- HARSANYI, J. C., AND R. SELTEN (1988): *A General Theory of Equilibrium Selection in Games*. The MIT Press, Cambridge, Massachusetts.
- HEGSELMANN, R. (1994): “Zur Selbstorganisation von Solidarnetzwerken unter Ungleichheiten,” in *Wirtschaftsethische Perspektiven I*, ed. by K. Homann, no. 228/I in Schriften des Vereins für Socialpolitik, Gesellschaft für Wirtschafts- und Sozialwissenschaften, Neue Folge, pp. 105–129. Duncker & Humblot, Berlin.
- HOTELLING, H. (1929): “Stability in Competition,” *Economic journal*, 39, 41–57.
- HUBERMAN, B. A., AND N. S. GLANCE (1993): “Evolutionary games and computer simulations,” *Proceedings of the National Academy of Sciences*, 90, 7716–7718.

- KANDORI, M., G. G. MAILATH, AND R. ROB (1993): “Learning, Mutation, and Long Run Equilibria in Games,” *Econometrica*, 61, 29–56.
- KIRCHKAMP, O. (1995): “Spatial Evolution of Automata in the Prisoners’ Dilemma,” Discussion Paper B–330, SFB 303, Rheinische Friedrich Wilhelms Universität Bonn.
- LINDGREEN, K., AND M. G. NORDAHL (1994): “Evolutionary dynamics of spatial games,” *Physica D*, 75, 292–309.
- MAY, R. M., AND M. A. NOWAK (1992): “Evolutionary Games and Spatial Chaos,” *Nature*, 359, 826–829.
- (1993): “The Spatial Dilemmas of Evolution,” *International Journal of Bifurcation and Chaos*, 3, 35–78.
- MAYNARD SMITH, J., AND G. R. PRICE (1973): “The Logic of Animal Conflict,” *Nature*, 273, 15–18.
- SAKODA, J. M. (1971): “The Checkerboard Model of Social Interaction,” *Journal of Mathematical Sociology*, 1, 119–132.
- SCHELLING, T. (1971): “Dynamic Models of Segregation,” *Journal of Mathematical Sociology*, 1, 143–186.
- TAYLOR, P. D., AND L. B. JONKER (1978): “Evolutionary stable strategies and game dynamics,” *Math. Biosci.*, 40, 145–156.
- YOUNG, H. P. (1993): “The Evolution of Conventions,” *Econometrica*, 61, 57–84.
- ZEEMAN, E. (1981): “Dynamics of the evolution of animal conflicts,” *Journal of Theoretical Biology*, 89, 249–270.

## A Proof of Proposition 1

In the following we will denote a pair of two automata  $a$  and  $b$  with the symbol  $\langle a, b \rangle$ . For notational convenience we will make no distinction between  $\langle a, b \rangle$  and  $\langle b, a \rangle$

- 1 If there are only grims in the population, then only  $\langle G^C, G^C \rangle$ ,  $\langle G^C, G_D \rangle$  and  $\langle G_D, G_D \rangle$  are possible.  $\langle G^C, G_D \rangle$  is unstable, thus, only  $\langle G^C, G^C \rangle$  and  $\langle G_D, G_D \rangle$  will remain in the population. Given any relative frequency of  $\langle G^C, G^C \rangle$  and  $\langle G_D, G_D \rangle$  no player will ever learn, since there are no alternative strategies to observe.

- 2 If there are only blinkers in the population, then only  $[\langle B^C, B^C \rangle \langle B_D, B_D \rangle]$ ,  $[\langle B_D, B_D \rangle \langle B^C, B^C \rangle]$  and  $[\langle B^C, B_D \rangle]$  are possible. All of these cycles of pairs are stable. No player will ever learn, since there are no alternative strategies to observe.
- 3 Trivially a state where possibly payoffs cycle, but nevertheless in all periods grim and blinker achieve the same payoff is an equilibrium, since no player ever has any reason to learn. An example for such an equilibrium is a population where  $1/4$  are cycles  $[\langle B^C, B^C \rangle \langle B_D, B_D \rangle]$ ,  $1/4$  are cycles  $[\langle B_D, B_D \rangle \langle B^C, B^C \rangle]$ ,  $(1+h)/(4g)$  are  $[\langle G^C, G^C \rangle]$  and  $(2g-1-h)/(4g)$  are  $[\langle G_D, G_D \rangle]$  where  $g$  and  $h$  are parameters of the game in figure 1 (of course not for all, but at least for some values of  $g$  and  $h$  the above relative frequencies are all in  $[0, 1]$ ). In this case average payoffs of blinkers and grims is in each period  $(1+h)/2$ .

This equilibrium is not stable. Change the relative frequency of a cycle that contains blinkers only slightly and payoffs of blinkers start cycling around the average population payoff. In the discussion of case 4 which is described below we will see that once such a state is reached, the cycling becomes stronger and stronger until the equilibrium of case 4 is reached.

- 4  $\langle G^C, G^C \rangle$ ,  $\langle G_D, G_D \rangle$ ,  $\langle G_D, B_D \rangle$ ,  $\langle B_D, B_D \rangle$ ,  $\langle B^C, B^C \rangle$  all belong to a stable cycle in the sense that after a limited number of transitions each pair will return to its original state. Unstable pairs are the remaining ones  $\langle G^C, G_D \rangle$ ,  $\langle G^C, B^C \rangle$  and  $\langle G^C, B_D \rangle$ . Unstable pairs can only be present in a cyclical equilibrium if they are introduced through learning. Since we take the limit of the the learning rate  $\epsilon \rightarrow 0$ , the relative frequency of unstable pairs in the population will be 0 as well. Thus, in a cyclical equilibrium we have only stable pairs.

All stable pairs belong to cycles of length either one or two. Therefore average payoffs in a cyclical equilibrium also form a cycle of length either one or two.

A cyclical equilibrium with both grim and blinker present in the population must have average payoffs of grim and blinker belonging to a cycle of length two. Otherwise one strategy would have always more than average payoff, which together with our learning rule, copy always the best average strategy, leads to the extinction of the other strategy. Let us without loss of generality assume that average payoffs form a cycle of length two with  $[\bar{u}_B > \bar{u}_G, \bar{u}_G > \bar{u}_B]$ . Thus, in the first period of each cycle  $\bar{u}_B > \bar{u}_G$ , in the second  $\bar{u}_G > \bar{u}_B$ .

[Figure 15 about here.]

Figure 15 shows now the dynamics of stable pairs. Let us give one example how to determine the transitions. There is some flow away from the cycle  $[\langle B^C, B^C \rangle \langle B_D, B_D \rangle]$ . In the second period of the cycle, when  $\bar{u}_G > \bar{u}_B$ , a  $\langle B_D, B_D \rangle$  has with probability  $\epsilon$  an opportunity to learn, which will then lead (in the first period of the next cycle) to a pair  $\langle G^C, B^C \rangle$ . The pair  $\langle G^C, B^C \rangle$  is an unstable pair which becomes in the second period of the current cycle a  $\langle G_D, B_D \rangle$ . The latter is a member of the stable cycle  $[\langle G_D, B^C \rangle \langle G_D, B_D \rangle]$ . Since we have two blinkers in the pair  $\langle B_D, B_D \rangle$  the total flow away from the cycle  $[\langle B^C, B^C \rangle \langle B_D, B_D \rangle]$  has speed  $2\epsilon$ . Since we consider the case  $\epsilon \rightarrow 0$  we can neglect the case that both members of a pair learn simultaneously.

When we consider all transitions given in figure 15 we see that only the three cycles  $[\langle G_D, G_D \rangle]$ ,  $[\langle G_D, B_D \rangle \langle G_D, B^C \rangle]$  and  $[\langle B_D, B_D \rangle \langle B^C, B^C \rangle]$  will remain in the population in the long run. Given the flow between these three cycles we see that  $1/4$  of  $[\langle G_D, G_D \rangle]$ ,  $1/2$  of  $[\langle G_D, B_D \rangle \langle G_D, B^C \rangle]$  and  $1/4$  of  $[\langle B_D, B_D \rangle \langle B^C, B^C \rangle]$  are a stable equilibrium.

Now let us look at the payoffs. In the first period of a cycle almost all pairs play  $DD$ , i.e. they get, given the game in figure 1, a payoff of zero. Nevertheless a few pairs are always present which actually learned a new strategy. Thus, we have  $\epsilon/2$  of  $\langle G^C, G_D \rangle$  and  $\epsilon/2$  of  $\langle G^C, B_D \rangle$ . Therefore payoffs are in the first period of a cycle as follows:

$$\bar{u}_G = \frac{\epsilon}{2}(1 + 2h) < \bar{u}_B = \frac{\epsilon}{2}$$

In the second period of a cycle learning players are negligible:

$$\bar{u}_G = \frac{1}{2} > \bar{u}_B = \frac{1}{2}(g + h)$$

Thus, indeed a cycle of payoffs like  $[\bar{u}_B > \bar{u}_G, \bar{u}_G > \bar{u}_B]$  exists. □

## List of Figures

1	The parameterisation of the underlying prisoners' dilemma game. . . . .	23
2	Neighbourhoods with interaction radius 1, 2 and 3 . . . . .	24
3	Frequency of $CC$ , $CD$ , and $DD$ with non-discriminative strategies . . . .	25
4	Non-discriminative strategies with 'copy best player'. . . . .	26
5	Non-discriminative strategies with 'copy best strategy'. . . . .	27
6	Discriminative strategies . . . . .	28
7	Frequency and success of grim and blinker with synchronous interaction and short memory . . . . .	29
8	Frequency of $CC$ , $CD$ , and $DD$ encountered by grim . . . . .	30
9	Discriminative strategies with asynchronous interaction and learning from long memory . . . . .	31
10	Frequency and success of grim and blinker with asynchronous interaction and long memory . . . . .	32
11	Parameterisation for coordination games. . . . .	33
12	Frequency of $CC$ , $CD$ , and $DD$ in coordination games . . . . .	34
13	Frequency of $CC$ depending on the size of the neighbourhood . . . . .	35
14	Frequency of $CC$ in 1...4-dimensional networks . . . . .	36
15	Transitions of cycles of stable pairs . . . . .	37



		Player $B_D$	
		$C$	$D$
Player $I$	$C$	$g$ $g$	$1$ $h$
	$D$	$h$ $1$	$0$ $0$

Figure 1: The parameterisation of the underlying prisoners' dilemma game. The game is a prisoners' dilemma for  $0 < g < 1$ ,  $h < 0$  and  $g > \frac{1}{2} + \frac{1}{2}h$ . We require  $g < 1$  and  $h < 0$  to ensure that  $C$  is a dominated strategy of the stage game. Then  $0 < g$  and  $g > \frac{1}{2} + \frac{1}{2}h$  make certain that  $CC$  Pareto dominates both  $DD$  and cycling between  $CD$  and  $DC$ .

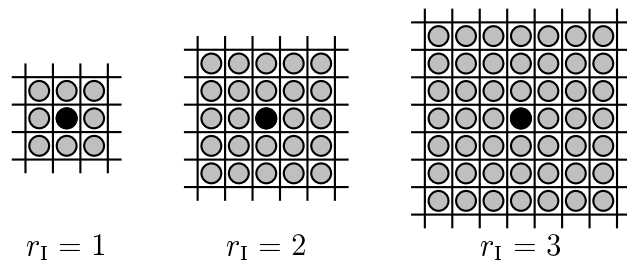


Figure 2: Neighbourhoods with interaction radius 1, 2 and 3

Possible interaction partners of a player are described as gray circles while the player itself is represented as a black circle.

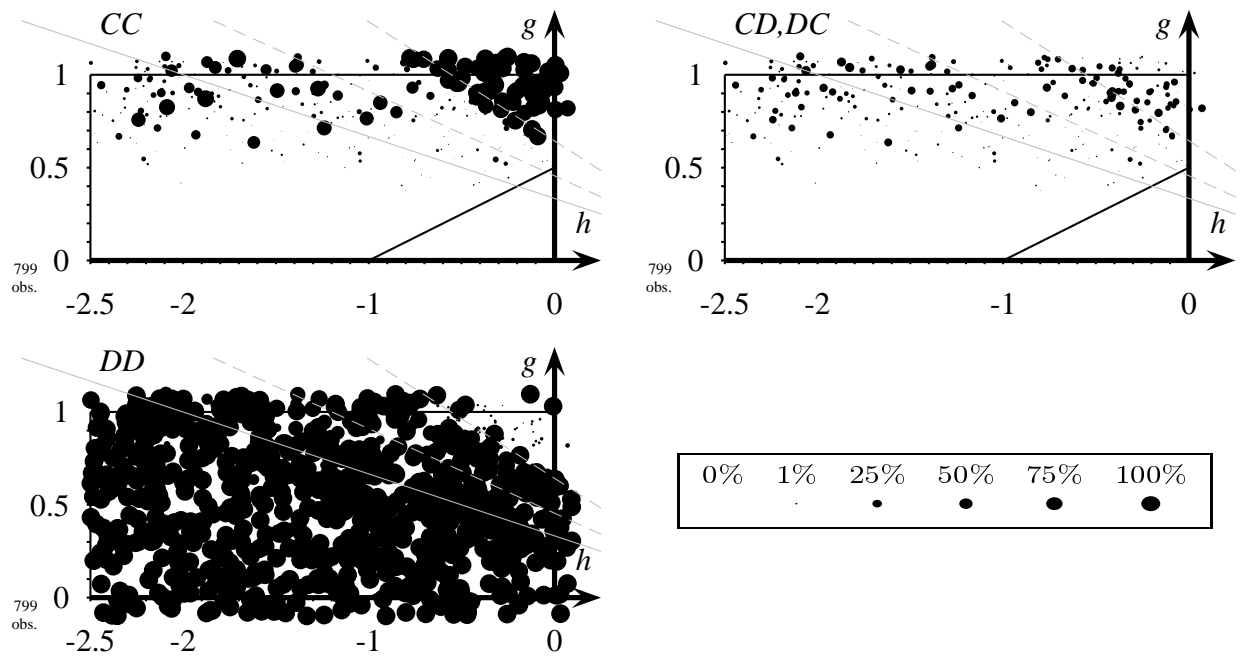


Figure 3: Frequency of *CC*, *CD*, and *DD* with non-discriminative strategies  
Non-discriminative strategies,  $p_I = 1$ ,  $t_L = 1$ , copy-best-player, short memory,  $r_L = r_I = 1$ , network= $80 \times 80$ .

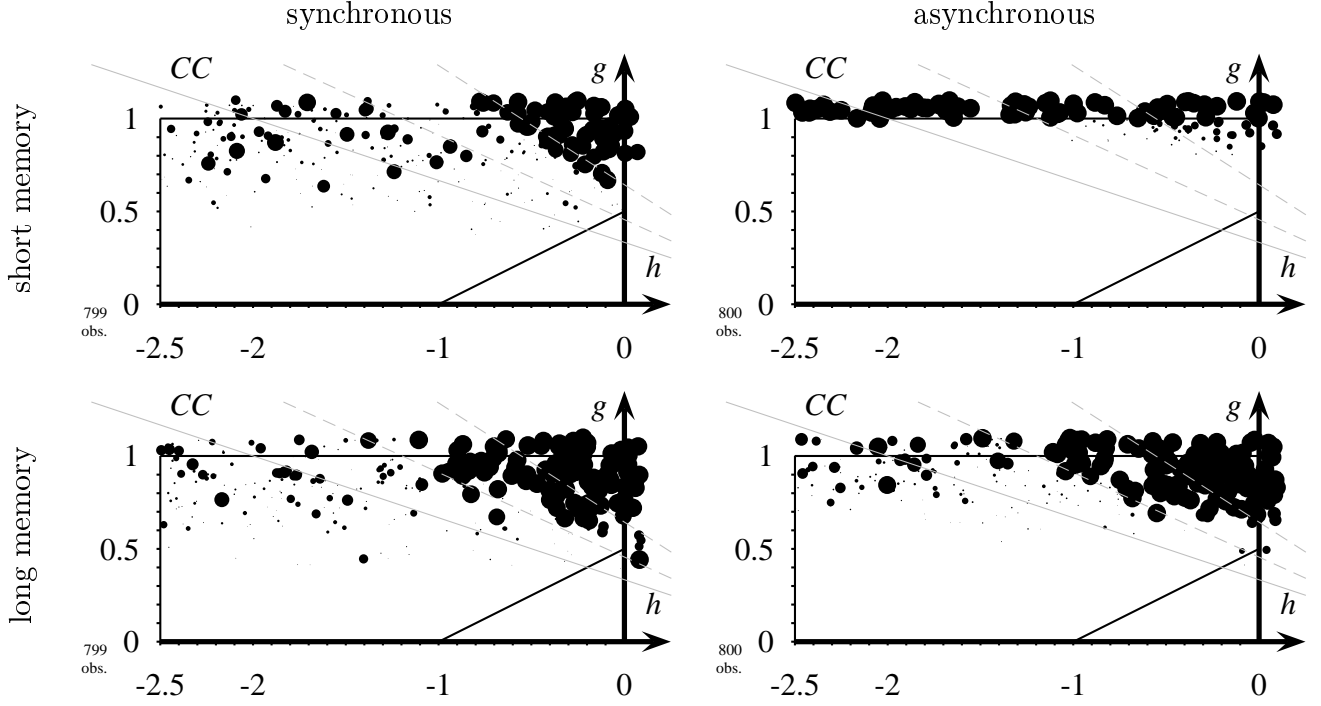


Figure 4: Non-discriminative strategies with ‘copy best player’.  
 (‘Copy best player’ learning rule, network= $80 \times 80$ ,  $r = 1$ )

Asynchronous learning eliminates cooperation only with ‘short memory’

To save space we display only the relative frequency of mutual cooperation for a variety of models. The top left diagram displays the relative frequency of cooperation in May and Nowak’s model. We have seen this diagram already in figure 3. The top right diagram shows the same model, but now with asynchronous learning and interaction as an approximation to Huberman and Glance. Cooperation dies out. The two diagrams at the bottom show the case of long memory where cooperation survives.

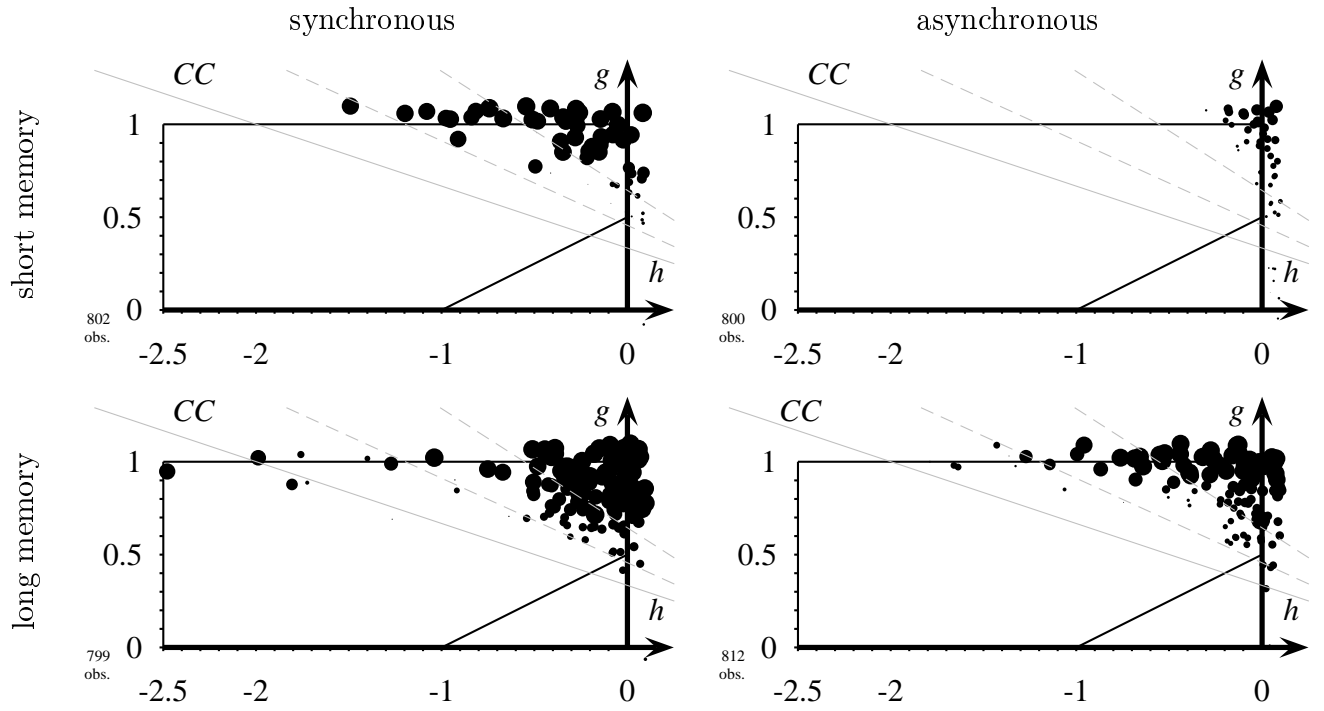
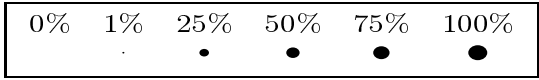


Figure 5: Non-discriminative strategies with ‘copy best strategy’.  
 (‘Copy best strategy’ learning rule, network= $80 \times 80$ ,  $r = 1$ )

The only difference to figure 4 is that here players use ‘copy best strategy’ as a learning rule. Again cooperation is eliminated only with ‘short memory’.



Discriminative strategies,  $p_I = 1$ ,  $t_L \in \{10, \dots, 14\}$ , copy-best-strategy, short memory,  $r_L = r_I = 1$ , network=80  $\times$  80.

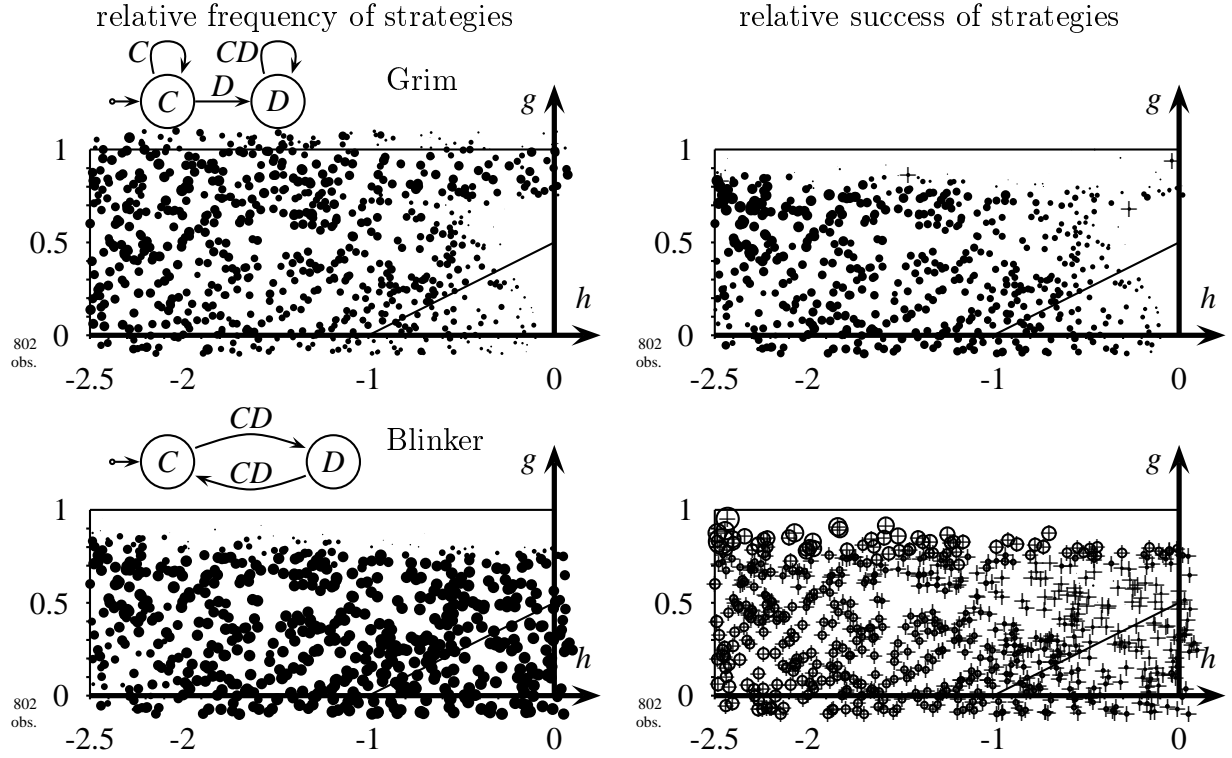


Figure 7: Frequency and success of grim and blinker with synchronous interaction and short memory

Discriminative strategies,  $p_I = 1$ ,  $t_L \in \{10, \dots, 14\}$ , copy-best-strategy, short memory,  $r_L = r_I = 1$ , network= $80 \times 80$ .

Frequencies are shown in the left part of the figure as explained in section 5.1. Relative success is shown in the right part of the figure in the following way: Average payoffs of strategies are calculated over the last 50 periods of a simulation. Black circles indicate that a strategy's average payoff is larger than population payoff. Crossed white circles indicate that the strategy's average payoff is smaller than average population payoff. Sizes of the circles are proportional to the payoff difference. When all automata use the same stage game strategy average payoffs per interaction for all automata are identical and no circles and crosses are shown.

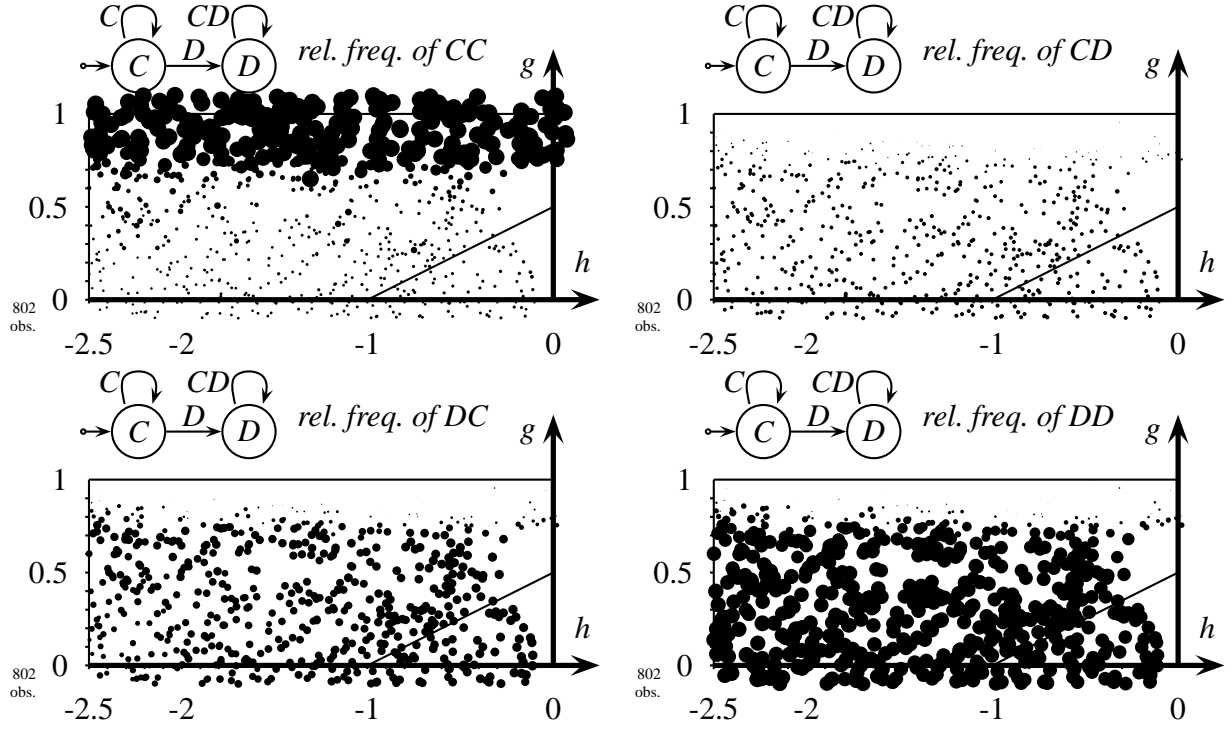


Figure 8: Frequency of  $CC$ ,  $CD$ , and  $DD$  encountered by grim  
 Grim strategies alternate between a small amount of support for their environment, playing  $CD$ , and substantial exploitation with  $DC$   
 discriminative strategies,  $p_I = 1$ ,  $t_L \in \{10, \dots, 14\}$ , copy-best-strategy, short memory,  $r_L = r_I = 1$ , network= $80 \times 80$ , sizes of the circles are proportional to relative frequencies of strategies.



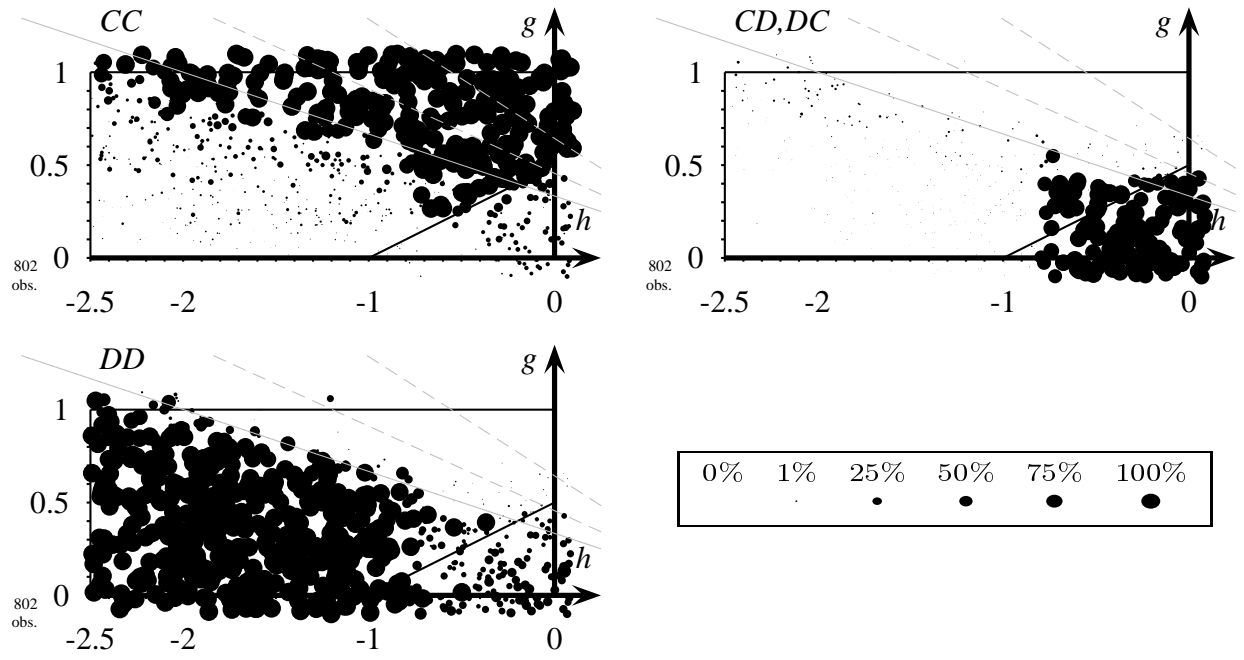


Figure 9: Discriminative strategies with asynchronous interaction and learning from long memory

Discriminative strategies,  $p_I = \frac{1}{2}$ ,  $t_L \in \{20, \dots, 28\}$ , copy-best-strategy, long memory,  $r_L = r_I = 1$ , network= $80 \times 80$ .

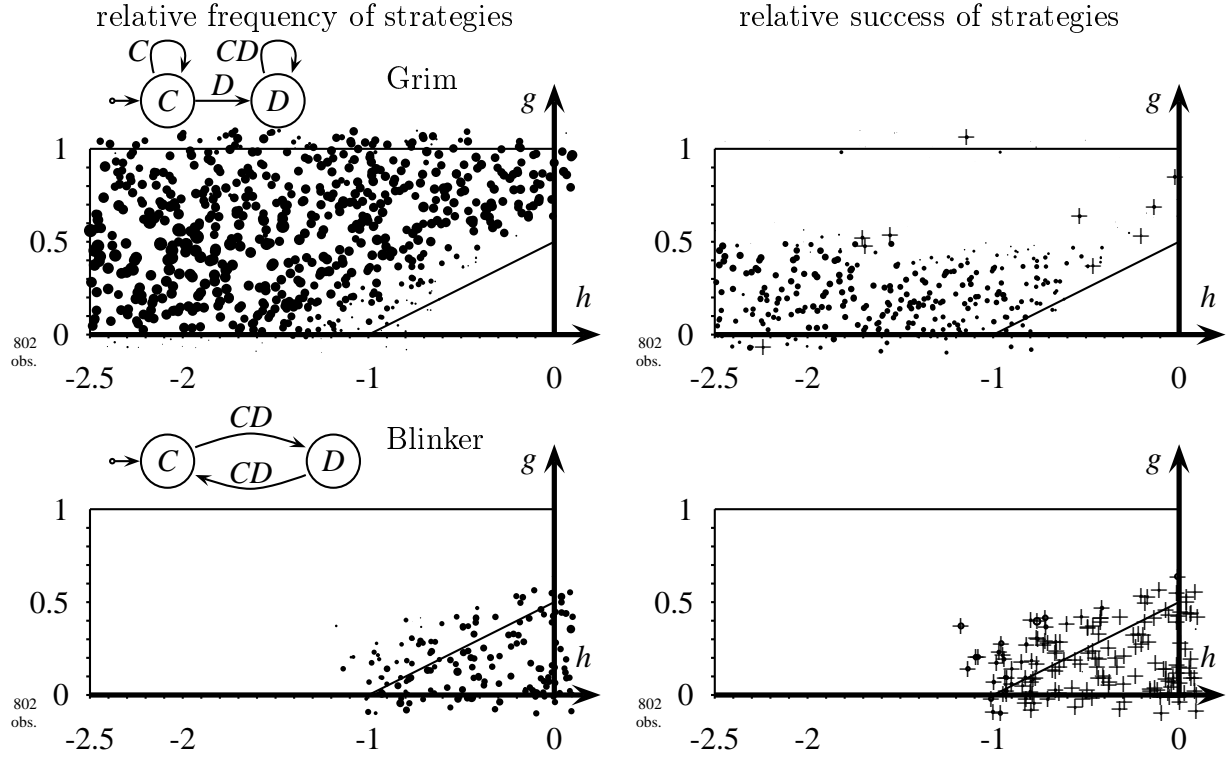


Figure 10: Frequency and success of grim and blinker with asynchronous interaction and long memory

Discriminative strategies,  $p_I = \frac{1}{2}$ ,  $t_L \in \{20, \dots, 28\}$ , copy-best-strategy, short memory,  $r_L = r_I = 1$ , network= $80 \times 80$ .

Frequencies are shown in the left part of the figure as explained in section 5.1. Relative success is shown in the right part. Average payoffs of strategies are calculated over the last 50 periods of a simulation. Black circles indicate that a strategy's average payoff is larger than population payoff. Crossed white circles indicate that the strategy's average payoff is smaller than average population payoff. Sizes of the circles are proportional to the payoff difference. When all automata use the same stage game strategy average payoffs per interaction for all automata are identical and no circles and crosses are shown.

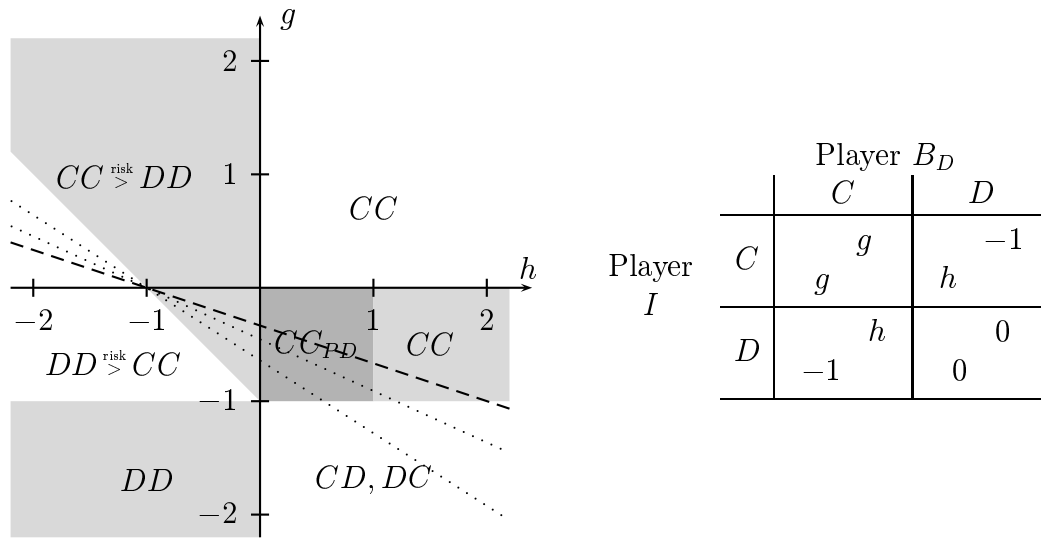


Figure 11: Parameterisation for coordination games.

The game is a coordination game for  $g > -1$  and  $h < 0$ . In this case the Pareto dominant equilibrium is  $CC$  if  $g > 0$  and  $DD$  otherwise. The risk dominant equilibrium is  $CC$  if  $g > -1 - h$  (denoted  $CC^{risk} > DD$  in the figure) and  $DD$  otherwise (denoted  $DD^{risk} > CC$  in the figure).

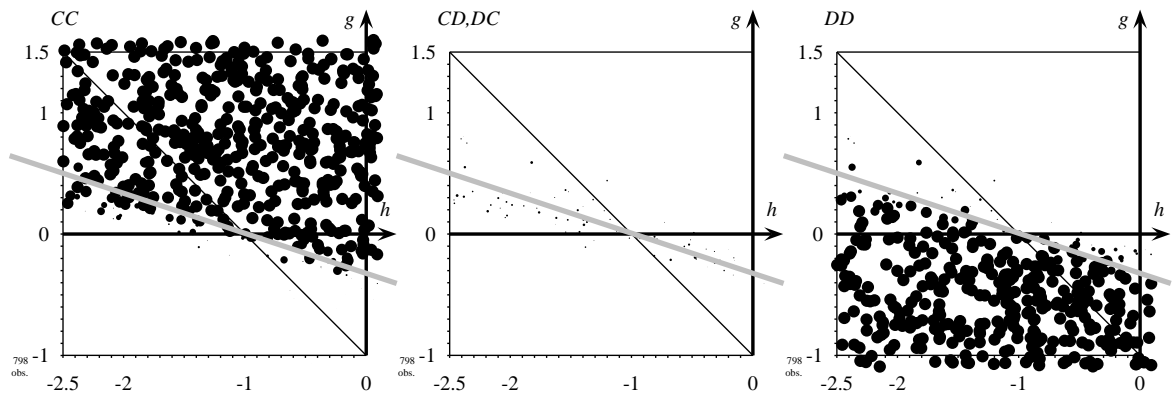


Figure 12: Frequency of  $CC$ ,  $CD$ , and  $DD$  in coordination games  
Discriminative strategies,  $p_I = \frac{1}{2}$ ,  $t_L \in \{20, \dots, 28\}$ , copy-best-strategy, long memory,  $r_L = r_I = 1$ , network= $80 \times 80$ .

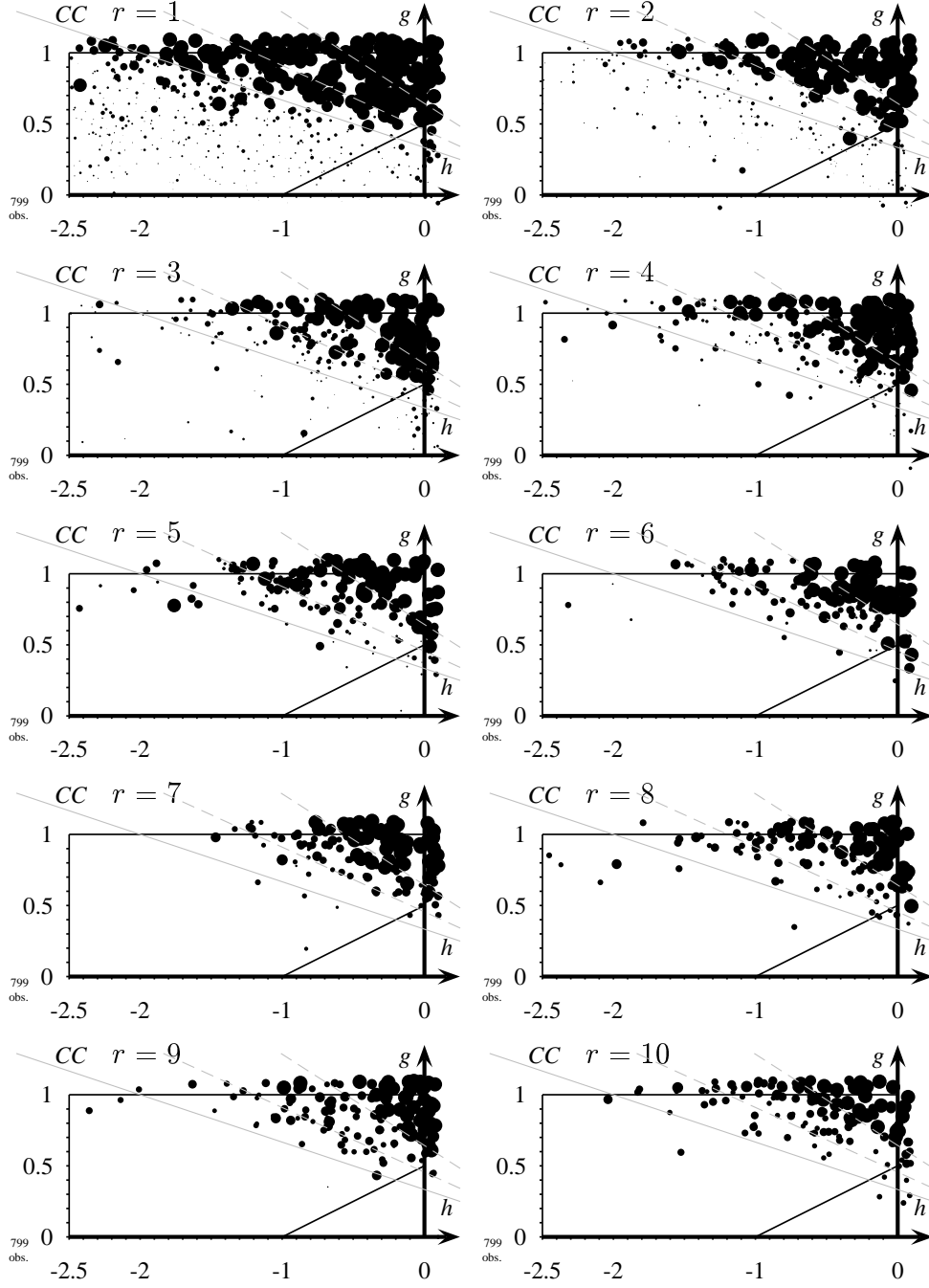


Figure 13: Frequency of  $CC$  depending on the size of the neighbourhood  $r = r_I = r_L$ , discriminative strategies,  $p_I = \frac{1}{2}$ ,  $t_L \in \{20, 28\}$ , copy-best-strategy, long memory, network= $21 \times 21$ . The population is smaller ( $21 \times 21$ ) than in the other simulations. This allows us to move all the way to a global model.

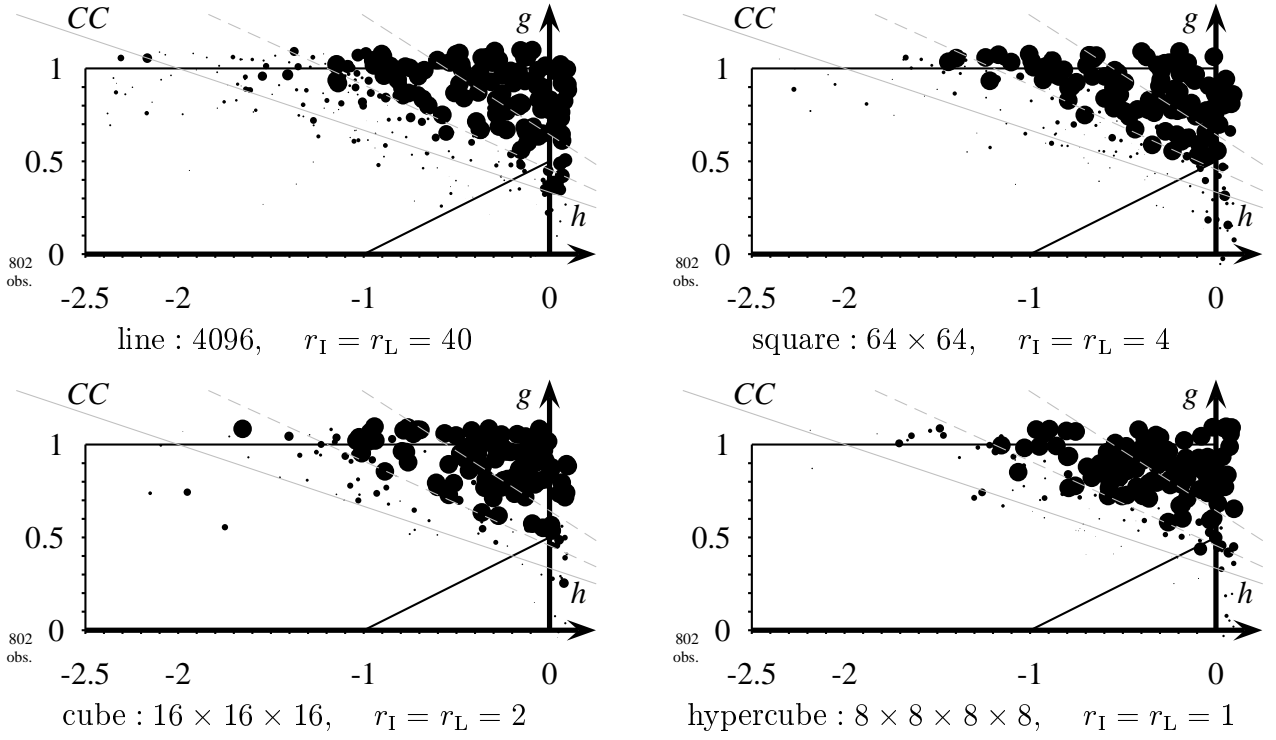


Figure 14: Frequency of  $CC$  in 1...4-dimensional networks

Discriminative strategies,  $p_I = \frac{1}{2}$ ,  $t_L \in \{20, \dots, 28\}$ , copy-best-strategy, long memory.

Notice that neighbourhoods are about 10 times as large than in the above simulations. This explains that the cooperative area is smaller than in figure 9.

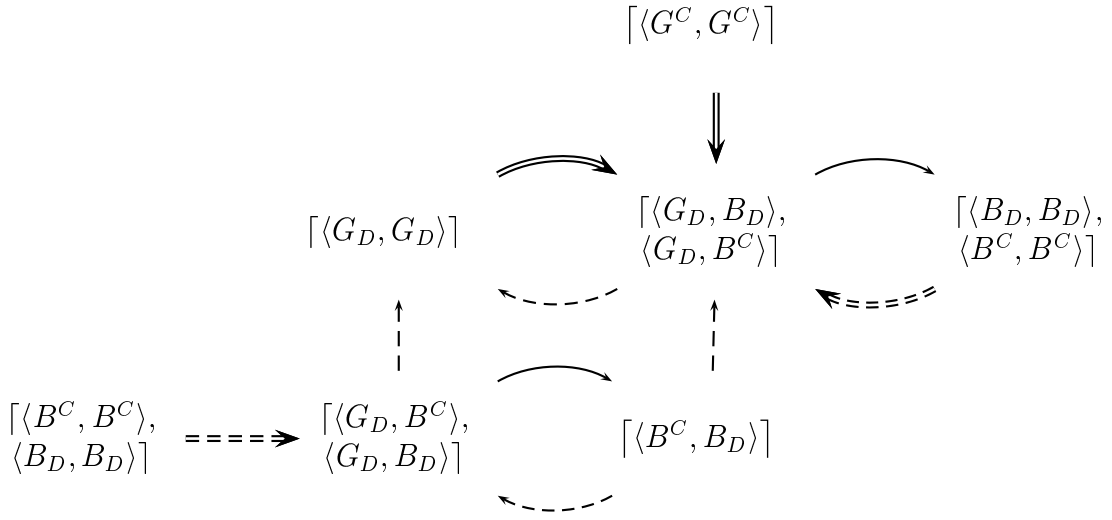


Figure 15: Transitions of cycles of stable pairs

Transitions of cycles of stable pairs if payoffs cycle like  $[\bar{u}_B > \bar{u}_G, \bar{u}_G > \bar{u}_B]$ .

Transitions with speed  $2\epsilon$  are shown with double lines. Speed  $\epsilon$  is shown with single lines. Transitions that take place in the second period are shown with dashed (single or double) lines. The first period is shown with solid lines.