Cooperation in a Multi-Dimensional Local Interaction Model.

Alexander F. Tieman[*], Harold Houba[†] and Gerard van der Laan[‡]

Abstract

We consider a local interaction model with a population on an h dimensional torus, in which in each round of play a random player gets a learning draw. This player plays a $k+1$ action stage game with players in his neighborhood, compares his own average payoff with the average payoff of the neighbors he played against and updates his action based on this comparison. Individuals use the update rule 'Win Cooperate, Lose Defect', a multi-player variant of Tit-for-Tat.

We prove that there are exactly $k+1$ stable states and that all of these can be reached with positive probability, for any dimension h of the torus. Furthermore, we prove that when $k+1=2$, both stable states will be reached with probability $\frac{1}{2}$. For $k+1>2$ we provide some insight in the probability of reaching each of the stable states by presenting simulation results.

Keywords: *Evolution, Local Interaction, Cooperation, Prisoner's Dilemma.*

JEL-code: C78.

[*]Tinbergen Institute and Free University, Department of Econometrics, De Boelelaan 1105, 1081HV Amsterdam, The Netherlands. E-mail: xtieman@econ.vu.nl, Phone: +31-20-4446022, Fax: +31-20-4446020, URL: http://www.econ.vu.nl/~xtieman/ .

[†]Free University.

[‡]Tinbergen Institute and Free University.

1 Introduction.

Cooperative behavior is often observed in situations in which there is repeated interaction between individuals. This is also the case in situations in which one of the individuals would be strictly better of if he did not cooperate with his opponent as is the case in for instance the Prisoner's Dilemma. Many repeated game models have been devised to get a grip on this kind of behavior. These models often take (hyper)rationality of the agents in the game as their point of departure. Sometimes the driving force behind the models is some kind of revealed preference argument: If an individual cooperates with his opponent, although he seemingly has nothing to gain from that, it must be the case that he has an unobservable preference for cooperating over other actions.

Recently, a different approach is taken by papers in the field of evolutionary game theory, see e.g. Eshel, Samuelson & Shaked (1996), Schlag (1996), Binmore & Samuelson (1992, 1997), the imaginary discussion between representatives of different economic views in Selten (1991) or the survey article by Van der Laan & Tieman (1998). This paper provides such an alternative model. We abandon the assumption that agents are rational utility-maximizers. Instead we propose that the agents follow simple behavioral rules or heuristics in updating their action, as is seen often in the field of behavioral game theory (see e.g. Camerer (1997)). Furthermore this paper is in the line of *local interaction models*, see also e.g. Ellison (1993, 1995). The population consists of agents located on a torus. Agents get to play the stage game only with agents in a subgroup of the population, called their *neighbors*. The group of neighbors consists of all agents located in positions on the torus directly adjacent to the location of the agent. The group of neighbors is different for each agent, although there may be substantial overlaps between the groups of neighbors of different agents.

At each stage game a player plays one of $k+1$ possible actions, labelled from the non-cooperative action 0 to the cooperative action k. Based upon the outcomes of the stage game an agent plays with his neighbors, he updates the action to play in the next stage game by comparing his average payoff of playing the game with all his neighbors to the average payoff his neighbors got from playing the game with him. Putted in the terminology of a Prisoner's Dilemma (PD), the update rule can be presented as follows. If the outcome for the agent is relatively high, he will act more cooperatively in the next round. If the outcome is relatively low, he will tend to display less cooperative behavior in the next round of play. This update rule is referred to as *'Win Cooperate, Lose Defect'* (WCLD). It is often encountered in experiments, both in economics (see e.g. Offerman, Sonnemans & Schram (1996)) and in sociology (see e.g. Messick & Liebrand (1995)). This update rule is a multi-player variant of the well know *Tit-for-Tat* (TfT) update rule. According to WCLD, a player who gets a low payoff knows that some of his neighbors defected on him and will therefore play non-cooperatively in the next round of play. A player that plays the non-cooperative action himself and gets a high payoff in this way, reasons that some of his neighbors are playing cooperatively and will then switch to cooperative behavior. The update rule WCLD can therefore be considered even more forgiving than TfT, since a player using WCLD will tend to play cooperatively in the next round of play when some of his neighbors played cooperatively in the last round of play, while a straightforward multi-player version of TfT would require all his neighbors to play cooperatively in order for him to switch to the cooperative action. [1]

An alternative rationale for this update rule can be found in the literature on aspiration levels (see e.g. Palomino & Vega-Redondo (1996), Rabin (1993), Van Lange et al (1997, 1997), Thibaut & Kelley (1959) or Kelley & Thibaut (1978)). An *aspiration level* is the minimum payoff an agent desires in order to play a certain action. If his payoff falls short of his aspiration level, he will play a certain other action. The update rule WCLD above, is the result of setting the aspiration level to be the average payoff of the neighbors an agent plays against. The aspiration level is thus endogenous. Whenever the agent gets a higher payoff than the average of his neighbors, he will tend to cooperation. When his payoff falls short of this number, he will play less cooperatively in the next round of play.

We find that there are exactly $k+1$ stable states in the model, that can all be reached with positive probability for any dimension $h \geq 1$ of the torus. Thus we know the model will always converge. In a stable state all players use the same action. Furthermore we show that for $k+1=2$, both stable states will be reached with probability $\frac{1}{2}$. We provide some insight in the probability of reaching a particular stable state when $k+1>2$ by presenting simulation results. We also provide a discussion on the relation between convergence time and the dimension of the torus. Finally, we look at the effect of a single deviation from a stable state to give an idea of the influence a mutant can have.

[1] Note that there are several different multi-player versions of Tit-for-Tat. In the version we consider, all opponents of an agent have to switch back to cooperative behavior after defection in order to let the agent himself switch back to cooperative behavior.

2 The Model.

We consider a population of N^h players located on a h dimensional set, being a *generalized torus*, i.e. each player is uniquely determined by a location $x = (x_1, \ldots, x_h)$, with $x_l \in \{1, \ldots, N\}$, $l = 1, \ldots, h$, and has the $3^h - 1$ players on locations $y = (y_1, \ldots, y_h) \in \{1, \ldots, N\}^h$, with $y_l = (x_l - 1) \bmod N, x_l, (x_l + 1) \bmod N$, $l = 1, \ldots, h$, except $y = x$ as his direct neighbors. Each player interacts only with his direct neighbors. For $h = 1$ this means players are located on a circle and have two neighbors, while for $h = 2$ players are located on a standard torus and have eight neighbors each.

The stage game is a symmetric $(k+1) \times (k+1)$ bimatrix game (A, B) characterized by payoff matrices $A = (a_{i,j})_{i,j=0}^{k}$ and $B = A'$. In the stage game a player playing action i and faced with an opponent playing action j will receive payoff $a_{i,j}$. The opponent will receive payoff $b_{i,j} = (A')_{i,j} = a_{j,i}$. We restrict attention to the case in which the payoff matrix has a Prisoner's Dilemma-like structure.

Definition 1 *For $k \geq 1$, the $(k+1) \times (k+1)$ payoff matrix A with $a_{i+1,j} < a_{i,j} < a_{i,j+1}$, $i, j = 0, \ldots, k-1$, $a_{i+1,i+1} > a_{i,i}$, $i = 0, \ldots, k-1$ and $2a_{i+1,i+1} > a_{i,j} + a_{j,i}$, $i = 0, \ldots, k-1$, $j = 0, \ldots, k$, is called a PD payoff matrix.*

Thus in a PD payoff matrix the action pair $(0,0)$ is the unique Nash equilibrium, and the action pair (k,k) is the unique Pareto efficient outcome. Furthermore action i, $i = 0, \ldots, k-1$, dominates action $i+1$. In section 4 we consider equidistant payoff matrices, a strict subset of the class of PD payoff matrices. Furthermore, we focus attention on the class of supermodular payoff matrices, which partially overlaps with the class of PD payoff matrices and which includes models of many interesting economic applications.

In the model each player plays an action i, $i = 0, \ldots, k$. There are infinitely many rounds of play, labelled $t = 0, 1, 2, \ldots$. Initially (at $t = 0$) each player is assigned an action at random with probability $\frac{1}{k+1}$ on each action. In each subsequent round of play a random individual in the population is selected. This individual is called the subject and he gets the possibility to update his action, a so called *learning draw*. From the set of the subject's neighbors a subset consisting of r, $r \in \{1, \ldots, 3^h - 1\}$, different players is chosen at random. The subject plays the same stage game r times, once against each of the r different individuals in the set. There are two versions of this model we consider, one in which r is a random number and one in which r is fixed to be $3^h - 1$, thus letting the subject play exactly once against all of his neighbors. After playing the stage game r times, the subject compares the average payoff he got from playing this game r times, π_{self}, to the average payoff his r selected neighbors got from each playing this game once with the subject, defined as π_{nbs}. [2] Based on the comparison of the payoffs, the subject updates his action using the update rule 'Win Cooperate, Lose Defect'. Whenever $\pi_{self} > \pi_{nbs}$, the subject sets his action i to $i+1 \leq k$, so in the next round of play he will play action $i+1$. When $\pi_{self} > \pi_{nbs}$ and the current action of the subject is k, this action remains unchanged. When $\pi_{self} < \pi_{nbs}$, the subject updates his action i to $i-1 \geq 0$. In case $\pi_{self} < \pi_{nbs}$ and the subject's current action is 0, the subject doesn't change his action. When both payoffs are exactly equal, the subject will stick to the action he is playing at present in the next round.

Given the selection of r neighbors, we define the $(k+1)$-dimensional neighborhood vector $\mu \in [0,1]^{k+1}$, by $\mu_j = \frac{n_j}{r}$, with n_j the number of selected neighbors playing action j, $j = 0, \ldots, k$. A subject using action i who compares π_{self} to π_{nbs} has to determine whether

$$\pi_{self} \gtreqless \pi_{nbs} \Leftrightarrow \pi_{self} - \pi_{nbs} \gtreqless 0 \Leftrightarrow \frac{1}{r}\sum_{j=0}^{k} n_j a_{i,j} - \frac{1}{r}\sum_{j=0}^{k} n_j a_{j,i} \gtreqless 0 \Leftrightarrow$$

$$\sum_{j=0}^{k} \frac{n_j}{r}\left(a_{i,j} - a_{j,i}\right) = \left(A - A'\right)_i \cdot \mu \gtreqless 0$$

where $(A-A')_i$ is the i-th row of the matrix $A - A'$. A subject currently playing action i, will thus play action $b(i)$ in the next period, where

$$b(i) = \begin{cases} i+1, \text{when } \ (A-A')_i \cdot \mu > 0 \text{ and } i < k, \\ i-1, \text{when } \ (A-A')_i \cdot \mu < 0 \text{ and } i > 0, \\ i, \qquad \text{all other cases.} \end{cases}$$

3 Main Results.

The main results of this paper are stated in the following Theorem.

Theorem 2 *For any integer $h \geq 1$, for any number $k+1 \geq 2$ of possible actions in the bimatrix stage game (A, A') with A a PD payoff matrix, there are exactly $k+1$ stable states of the model. These are the states in which all players use the same action i, $i = 0, \ldots, k$. Each of these stable states can be reached with positive probability for almost all initial states, both when at each iteration the*

[2]Note that we implicitly assume that different payoffs can be added up. This is justified in the case that payoffs are profits and players are risk-neutral.

number of neighbors the chosen subject interacts with is a random number $r \in \{1, \ldots, 3^h - 1\}$, *and when at each iteration the chosen subject interacts with all neighbors, i.e.* $r = 3^h - 1$ *always.*

Before we prove this Theorem, we provide some intuition for it by considering the case of N players located on a circle, who play a Prisoner's Dilemma as stage game and where a player who gets the learning draw plays against his two neighbors, i.e. $h = 1$, $k + 1 = 2$ and r fixed equal to 2. We take an arbitrary initial situation, not being a stable state. In this situation there will be several 1-clusters, i.e. clusters of adjacent players using action 1, and the same number of 0-clusters, i.e. clusters of adjacent players using action 0. The dynamics are *boundary preserving*, that is, an individual can change his action only when one of his neighbors uses an action different from his own (see e.g. Eshel, Sansone & Shaked (1996)). To see this, all we need is the observation that if both neighbors of a subject play the same action as the subject, all players get exactly the same payoff in all of the games that are played. Therefore the comparison of the subject's average payoff to the average payoff of his neighbors will always result in equality. Thus the subject will not change his action. It is only when at least one of the neighbors of the subject uses a different action than the subject itself, i.e. when at least one of the neighbors belongs to a different cluster than the subject, that there is the possibility of differences in average payoffs between the subject and his neighbors. Such situations only arise at the boundaries of clusters. At a boundary either the subject does not change his action and the clusters remain unchanged or the subject changes his action, and consequently the subject switches clusters. The length of the cluster the subject belonged to decreases by one and the length of the cluster the subject switches to increases by one. This argument implies that the number of clusters never increases.

In order to show convergence to a stable state in which all individuals in the population use the same action we now have to show that there are sample paths with positive probability that lead to a decrease of the number of clusters to 1. First, we will focus on the path to a state in which all individuals play action 1. In an arbitrary initial situation 0-clusters will either have length 1 or length greater than 1. [3] We look at the latter case first. When the 0-player on the boundary of (one of) the 0-cluster(s) gets the learning draw, his situation looks

[3] In the non-generic case where the 0-cluster is initialized at length N, i.e. in which all individuals are endowed with action 0, the initial state is stable and no player will ever change his action.

like that of the middle player in

$$\ldots\ 1\ 0\ 0\ \ldots$$

where a 1 represents a player using action 1 and a 0 represents a player using action 0. In this case the 0-player realizes an average payoff of $\frac{1}{2}(a_{0,0}+a_{0,1})$, while the average payoff of his neighbors is $\frac{1}{2}(a_{0,0}+a_{1,0})$. By the assumption $a_{i+1,j} < a_{i,j} < a_{i,j+1}$ with $i=j=0$ on the payoff matrix A, the former payoff is larger than the latter. Consequently the 0-player will update his action and become a 1-player in the next round. Thus there is always a positive probability that 0-clusters shrink. To see if they can actually disappear, we have to consider the situation in which the 0-cluster consists of a single player. In this situation the environment for the 0-player is different than above, namely

$$\ldots\ 1\ 0\ 1\ \ldots$$

Here the average payoff to the 0-player is $a_{0,1}$, while the average payoff of his neighbors is $a_{1,0}$. By assumption, the former payoff is strictly larger than the latter, resulting in a switch to action 1 of the 0-player. Thus we have shown that the number of 0-clusters can decrease. Applying this argument to all 0-clusters initially present in the population, results in a population in which all players use action 1. This environment arises with positive probability and when it is reached, no player will ever switch to action 0 again.

A similar argument goes in order to construct a path leading to a stable state in which all players use action 0. In this case we look at the situation

$$\ldots\ 0\ 1\ 1\ \ldots,$$

where the 1-player in the middle will get an average payoff $\frac{1}{2}(a_{1,0}+a_{1,1})$ and his neighbors will get an average payoff $\frac{1}{2}(a_{0,1}+a_{1,1})$, resulting in the 1-player switching to action 0. In the situation

$$\ldots\ 0\ 1\ 0\ \ldots,$$

the average payoff to the 1-player is $a_{1,0}$, while the average payoff of his neighbors is $a_{0,1}$. In this case the 1-player also updates his action to 0. This way a cluster of 1-players can disappear. With positive probability this happens to all 1-clusters initially present in the population, resulting in a stable state of the population in which all players use action 0. At any time t when the population has not converged to a stable state yet, when all clusters have length at least 2, there will be the same number of 0-players on a boundary as there are 1-players on a boundary. As a direct consequence, the probability that a 1-player on the boundary gets the learning draw before a 0-player on the boundary does, is $\frac{1}{2}$. Thus the probability that a 1-cluster shrinks is $\frac{1}{2}$. When there are clusters of length 1 present in the population, this probability can be anywhere between $\frac{1}{3}$ and $\frac{2}{3}$, depending on the exact state of the population.

In the process of explaining the paths to both stable states, we have argued that there are two stable states. When the population is not in one of these states, there are always at least two boundaries present in the population, separating a 1-cluster from a 0-cluster. In this setting, any player on a boundary that gets the learning draw will change his current action. Therefore, none of the other population states is stable. With this intuition in place, we now proof the following Lemma.

Lemma 3 *For a $(k+1) \times (k+1)$ symmetric bimatrix stage game (A, A'), with A a PD payoff matrix, every i-player, $i < k$, facing a set of opponents containing at least one j-player, $i < j \leq k$, and not containing any l-player, $0 \leq l < i$, will change his action into $i+1$ when this player gets the learning draw.*

Likewise, for the same stage game, every i-player, $i > 0$, facing a set of opponents containing at least one j-player, $0 \leq j < i$, and not containing any l-player, $i < l \leq k$ will change his action into $i-1$ when this player gets the learning draw.

Proof

When an i-player, $i < k$, facing a set of opponents containing at least one j-player, $i < j \leq k$, and not containing any l-player, $0 \leq l < i$, plays the stage game, his neighborhood vector looks like $\mu = (0, \ldots, 0, \mu_i, \mu_{i+1}, \ldots, \mu_k)^T$, with at least one of the $\mu_j > 0$, $j = i+1, \ldots, k$, and his payoff is

$$A_i \cdot \mu = 0 + \sum_{j=i}^{k} a_{i,j}\mu_j,$$

and hence it is a weighted average of the payoffs $a_{i,i}, a_{i,i+1}, a_{i,i+2}, \ldots, a_{i,k}$. The average payoff of his opponents is

$$A'_i \cdot \mu = 0 + \sum_{j=i}^{k} a_{j,i}\mu_j,$$

a weighted average of $a_{k,i}, \ldots, a_{i+2,i}, a_{i+1,i}, a_{i,i}$. The former payoff is always larger than the latter, since iterated application of the inequality $a_{i+1,j} < a_{i,j} < a_{i,j+1}$

gives us $a_{k,i} < \ldots < a_{i+2,i} < a_{i+1,i} < a_{i,i} < a_{i,i+1} < a_{i,i+2} < \ldots < a_{i,k}$. So, $(A - A')_i \cdot \mu > 0$ and, hence, according to WCLD an i-player becomes an $(i+1)$-player.

A similar argument can be applied to prove the second statement in the Lemma. $\square$

We are now ready to give the proof of Theorem 2.

Proof of Theorem 2.

The structure of this proof is as follows. First we argue that a population state in which all players use the same action, is stable. Then we show that no population state where multiple actions are present, can be stable. We conclude the proof by constructing a path to any one stable state.

We consider an i-player as subject on the h dimensional torus. If this player has only neighbors using action i, he will not change his action, since his average payoff will be exactly equal to the average payoff of his neighbors. In other words, the dynamics are boundary preserving. If the subject has a number of neighbors that use an action different from his own, i.e. if the subject is on the boundary, the subject has to compare payoffs. In this case there is always a positive probability that one of the r, $r \in \{1, 2, \ldots, 3^h - 1\}$, neighbors in his set of opponents is playing an action $j \neq i$. When this happens, chances are that the subject will not realize a payoff equal to that of his neighbors. Now the subject will either change his action or he is already playing one of the actions 0 or k and realizes a payoff lower respectively higher than the average payoff of the players in his set of opponents. However, Lemma 3 states that the latter cannot happen. Whenever a 0 (k)-player faces a set of opponents not only containing 0 (k)-players, he will update to action 1 ($k-1$). Thus we have shown that no state where multiple actions are present in the population is a stable state and that all states where there is only one action present in the population are stable states. Since there are $k+1$ possible actions in the stage game, the only $k+1$ stable states are those in which all players in the population use action i, $i = 0, \ldots, k$. Now consider the following sequence of learning draws to reach the stable state in which all players use action i, $i = 0, \ldots, k$. First give the learning draw sequentially to all players on a boundary using action 0. According to Lemma 3 these players will update their action to 1, whenever their set of opponents contains at least one player not using action 0. Note that after the 0-players that were originally on a boundary have updated, other 0-players will be on a boundary. Thus by sequentially selecting all 0-players on the boundary, we let all

0-players in the population update to action 1. When all 0-players have done so, give the learning draw sequentially to all 1-players. In absence of any 0-players, 1-players will update their action to action 2. When all 1-players have done so, give the learning draw to all 2-players, and so forth, until there are no players left using an action $j < i$. Then give the learning draw sequentially to all k-players on a boundary. They will update to action $k-1$ according to Lemma 3. When there are no k-players left, start giving the learning draw to $k-1$-players on the boundary, and so forth, until there are no players left who use an action $j > i$ and we have reached the stable state in which all players use action i. This particular sequence of learning draws occurs with positive probability. $\square$

Note that there are many more paths than the one described in the proof that lead to stable states. Therefore it is not necessarily the case that the occurrence of a stable state becomes increasingly rare when the number of actions $k+1$ increases. In the next section we show that convergence times do increase with k, but not dramatically.

Theorem 4 *For any integer $h \geq 1$, for $k+1 = 2$ and for the bimatrix game (A, A') with A a PD payoff matrix as stage game (i.e. the standard Prisoner's Dilemma), the a priori (i.e. before initialization of the population state) probability of convergence to each one of the two stable states is exactly $\frac{1}{2}$. This is true when r is a random number in each iteration, as well as when $r = 3^h - 1$.*

Proof.

For $k+1 = 2$, Lemma 3 shows that a 0-player, facing a set of opponents containing at least one 1-player, will change his action and become a 1-player. Similarly, the Lemma shows that a 1-player, facing a set of opponents containing at least one 0-player, will change his action and become a 0-player himself.

Now we show the existence of a path with positive probability that leads to extinction of all 0-players. Suppose there are $M,\ 0 < M < N^h$, 0-players present in the population. Let the learning draw go to each 0-player at a boundary sequentially. This happens with positive probability. Thus there is positive probability that at time $T_0 \geq M+1$, there will not be any 0-players left in the population and thus that the population has reached the stable state in which all players use action 1. Starting from the same state, a sequence of learning draws that all go to 1-players on a boundary, leads to extinction of action 1 in $T_1 \geq N^h - M + 1$ [4] rounds. Of course there a many more paths with positive

[4] If r is fixed at $3^h - 1$, the equality sign applies, i.e. $T_0 = M+1$ and $T_1 = N^h - M + 1$.

probability leading to one of the two stable states. Note that only in the two initial states equal to the stable states, no path that leads to the other stable state with positive probability exists.

For every path to an all-0 state from an initial state with M 0-players at locations $x^1, \ldots, x^M$ and $N^h - M$ 1-players at the other locations, there will be a similar path to an all-1 state starting from an initial state with M 1-players at these same locations $x^1, \ldots, x^M$ and $N^h - M$ 0-players at the other locations. Since the distribution over initial population states is symmetric, the a priori probabilities of converging to either stable state are equal and since these are the only two stable states in the case $k = 1$, it must be that Pr(all-0 state) = Pr(all-1 state) $= \frac{1}{2}$. We conclude by stating that for fixed $r = 3^h - 1$, similar reasoning leads to the same results. $\square$

4 Simulations.

We simulated [5] the above model in one and two dimensions ($h = 1$ and $h = 2$). We were especially interested in the behavior of the model with more than two possible actions. The interest in this behavior lies in the question how often each of these stable states will be reached. We look both at situations in which the population starts at time $t = 0$ in a random state and at situations in which the initial population state is not random.

First we initialize the simulation with a random population state, i.e. each player is assigned a random action l, $l = 0, \ldots, k$, where $\Pr(0) = \Pr(1) = \ldots = \Pr(k) = \frac{1}{k+1}$. Then we start round 1. A player is chosen at random (uniformly) from the population and gets the learning draw. Since the dynamics are boundary preserving, choosing a player that is not on a boundary does not result in a change of the population state. Therefore, in the simulations, we do not bother selecting an agent at the boundary. [6] The selected agent plays the $(k+1) \times (k+1)$ bimatrix stage game (A, A') against all his neighbors (we take r fixed at $3^h - 1$), looks at the outcomes and updates his action through the update rule WCLD. The entries of the PD payoff matrix A are taken to be $a_{i,j} = a_{0,0} - a \cdot i + 2a \cdot j$.

We let each simulation run until convergence to a stable state occurs. We study the effect of the dimension h and the number of different action available $k + 1$. In the case $h = 2$ we restrict simulations to a 4×4 torus, because of computing power restrictions. For each combination of parameter values we run

[5] The simulations were performed in Borland's *TurboPascal 7.0* on a Pentium 166 Mhz machine.

[6] Note that convergence time would decrease dramatically when we selected only agents on the boundary.

$1,000$ simulations and look at the frequency distribution of the outcomes over the $k+1$ stable states. These distributions are reported in tables 1 to 4 in appendix A for the one dimensional case and in tables 10 to 13 in appendix A for the two dimensional case.

We draw attention to several interesting features of the results. The first one is that the frequency distribution over stable states is similar for $h=1$ and $h=2$. In Theorem 4 this was already shown to be the case for $k=1$, but the simulations suggest it to be true for larger values of k as well. The second feature is that convergence times increase dramatically when h increases. This effect can be clarified by looking (for the case $k=1$) at the number of neighbors the last i-player, $i=0,1$ has in a population where all other players use action j, $j=0,1$, $j \neq i$. This player has 3^h-1 neighbors playing an action different from his own. Whenever he gets the learning draw, he will switch his action and from then on the population will be in a stable j state. However, when one of his neighbors gets the learning draw, this player will switch actions and the population state moves away from a stable state. Now note that the probability that a neighbor of the i-player gets the learning draw is 3^h-1 times as large as the probability that the i-player gets the learning draw. Thus the probability of moving away from a stable state is always larger than that of reaching the stable state, but when h increases the numbers get more unfavorable for reaching the stable state. This effect accounts for the rapidly increasing convergence time when h increases. In fact, the increase in convergence time is exponential in h.

A third interesting feature is that the frequency distributions are all symmetric. This is caused by fact that, on top of assuming that A is a PD payoff matrix, we have taken the matrix to be *equidistant* over actions, i.e. $a_{i+1,i+1} = a_{i,i}+a$, $a_{i+1,i} = a_{i,i}-a$, $a_{i,i+1} = a_{i,i}+2a$, $a>0$, resulting in $a_{i,j} = a_{0,0}-a\cdot i+2a\cdot j$. Note that such a PD equidistant payoff matrix is supermodular as defined in Topkis (1979). When we change the stage game into a game without equidistant payoffs, the frequency distributions are not neccesarily symmetric anymore, as is illustrated in section 4.1.

The last feature we discuss here is the effect of a single mutation on a population in stable state. To get a grip on this effect, we ran simulations with as initial state a state in which all but one player play the same action. We look at the extreme case in which all but one players play action 0 and one player plays action k and at the other extreme case in which all but one players play action k and one player plays action 0. See tables 5 and 6 in the appendix for some results for $h=1$ and tables 14 and 15 for some results for $h=2$. We see that in both the one and the two dimensional case, the distribution is very much concentrated on the higher labelled actions when the initial state is the state in which all but one player use action k, while the distribution is concentrated on the lower labelled actions when the initial state is the state in which all but one player set action 0. However, it is clear that a single mutation has considerable impact. The large effect a single mutation has on a population in stable state can be explained by the above arguments on convergence times. When a single mutant enters the population, it becomes more likely that a move away from the stable state will occur in the following rounds than that the same stable state will reappear in the next rounds. However, from the results we clearly see that the effect of a single mutation in a population in which all but one player set action $i,\ i = 0, k$, is limited in the sense that the stable states that are ultimately reached are states close to the original all-i state.

4.1 Supermodular Games.

A particular interesting class of games is the class of supermodular games introduced by Topkis (1979) and further explored by Milgrom & Roberts (1990, 1995), which includes models of oligopolistic competition, macroeconomic coordination failure, bank runs, R&D competition, pretrial bargaining, Bertrand price competition and many other interesting models. A game is supermodular when the two dimensions of the problem are strategically complementary, that is, when the payoff matrix A satisfies $a_{i,j} - a_{i,j-1} \leq a_{l,j} - a_{l,j-1},\ j = 1, \ldots, k,\ i, l = 0, \ldots, k,\ l > i$. The PD equidistant payoff matrix A above is supermodular, however, it is a borderline case, since the two dimensions of the matrix are not complementary, but just independent. When we change the payoff matrix of the stage game into a strict supermodular payoff matrix, i.e. $a_{i,j} - a_{i,j-1} < a_{l,j} - a_{l,j-1},\ j = 1, \ldots, k,\ i, l = 0, \ldots, k,\ l > i$, we observe a clear tendency towards more cooperation in the frequency distribution over stable states, when $k > 1$. This can be seen in table 7, where the payoff matrix is generated through Bertrand price competition as specified above this table. Although it is now still the case that there are $k+1$ stable states and that they can all be reached with positive probability, the probability distribution over stable states has more of its mass on the higher labelled than on the lower labelled actions. When we introduce infrequent mutations in the model, a mutant still has considerable effect on the distribution over stable states, as can be seen in tables 8 and 9 where the results are presented for the case $k = 10$ for a single mutation from an all-0 state and an all-10 state respectively. Notice that a mutant in an all-10 state has far less influence on the frequency distribution over stable state than a mutant in an all-0 state does. In the former case the distribution puts more weight on the higher labelled actions than it does on the lower labelled action in the latter case. Again this is caused by the supermodularity of the payoff matrix. The probability that the payoff to a player using a high action is larger than the average payoff of his set of opponents, is larger when the two dimensions of the payoff matrix are complementary than when they are not.

5 Conclusions.

In this paper we have shown that in a local interaction model in which in each round of play a random player from the population is selected to play a stage game with individuals from his neighborhood and in which the stage game is a Prisoner's Dilemma with two possible actions, the probability of ending up in a stable state in which all players play the cooperative action is a priori $\frac{1}{2}$, when the update rule is 'Win Cooperate, Lose Defect', a multi-player version of Tit-for-Tat. Such an update rule is observed in experiments. Therefore this model contributes to the explanation of cooperative behavior in Prisoner's Dilemma-like situations, which is often observed, both in experiments and in real world situations.

Furthermore we have shown that in a multi-action stage game with $k+1$ actions, there are $k+1$ stable state, which can all be reached with positive probability. We ran simulations in order to gain insight into the frequency distribution over stable states after random initialization. It turns out to be symmetric for equidistant payoffs. For strict supermodular payoff matrices, the distribution is puts more weight on the higher actions. This can be seen as a tendency towards cooperative behavior.

References

Binmore, K.G. and L. Samuelson (1992). Evolutionary stability in repetead games played by finite automata. *Journal of Economic Theory 57*.

Binmore, K.G. and L. Samuelson (1997). Muddling through: Noisy equilibrium selection. *Journal of Economic Theory 74*, 235–265.

Camerer, C.F. (1997). Progress in behavioral game theory. *Journal of Economic Perspectives 11*, 167–188.

Ellison, G. (1993). Learning, local interaction, and coordination. *Econometrica 61*, 1047–1071.

Ellison, G. (1995, August). Basins of attraction, long run equilibria, and the speed of step-by-step evolution. Working paper, Massachusetts Institute of Technology.

Eshel, I., L. Samuelson, and A. Shaked (1996). Altruists, egoists and hooligans in a local interaction model. Working paper, University of Wisconsin-Madison, Social Systems Research Institute.

Eshel, I., E. Sansone, and A. Shaked (1996). Evolutionary dynamics of populations with a local interaction structure. Working paper, School of Mathematical Sciences, Tel Aviv University.

Kelley, H.H. and J.W. Thibaut (1978). *Interpersonal Relations: A Theory of Interdependence*. New York: Wiley.

van der Laan, G. and A.F. Tieman (1998). Evolutionary game theory and the modelling of economic behavior. forthcoming in *De Economist*.

van Lange, P.A.M. (1997). Persoonsverschillen in coöperatie, individualisme en competitie. *Nederlands Tijdschrift voor de Psychologie 52*, 101–110. In Dutch.

van Lange, P.A.M., E.M.N. de Bruin, W. Otten, and J.A. Joireman (1997). Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *Journal of Personality and Social Psychology 73*, 733–746.

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

Milgrom, P. and J. Roberts (1990). Rationalizability, learning, and equilibrium in games with strategic complementaries. *Econometrica 58*, 1255–1277.

Milgrom, P. and J. Roberts (1995). Complementarities and fit: Strategy, structure and organizational change in manufacturing. *Journal of Accounting and Economics 19*.

Offerman, T., J. Sonnemans, and A. Schram (1996). Value orientations, expectations, and voluntary contributions in public goods. *Economic Journal 106*, 817–845.

Palomino, F. and F. Vega-Redondo (1996). Convergence of aspirations and (partial) cooperation in the prisoner's dilemma. Discussion paper, Instituto Valenciano de Investigaciones Económicas.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review 83*(5), 1281–1302.

Schlag, K.H. (1996). Why imitate and if so, how? a bounded rational approach to multi-armed bandits. Discussion Paper B-361, Bonn University.

Selten, R. (1991). Evolution, learning, and economic behavior. *Games and Economic Behavior 3*, 3–24.

Thibaut, J.W. and H.H. Kelley (1959). *The Social Psychology of Groups*. New York: Wiley.

Topkis, D.M. (1979). Equilibrium points in nonzero-sum n-person submodular games. *Siam Journal of Control and Optimization 17*, 773–787.

A Simulation Results.

Here we report the results of the some of the most relevant simulations we ran. In each table you will find the label of the action in the first column. In the second column is the number of simulation runs that converged to the action in the first column. The third column contains the frequency distribution over actions. The last column displays the number of rounds of play it took the model to converge to a stable state, averaged over all runs that converged to a particular stable state.

The following 4 tables contain the results of the simulations of the one dimensional model with $N = 38$ that were initialized with a uniform random population state. The payoff matrix A is a PD equidistant payoff matrix, with $a_{0,0} = 500$ and parameter $a = 10$, i.e. $a_{i,j} = 500 - 10i + 20j$.

Action	Nr Cases	Distribution	# Rounds
0	144	0.144	$4,785$
1	698	0.698	$2,726$
2	158	0.158	$4,601$
TOTAL	$1,000$	1	$3,319$

Table 1: Simulation results for $h = 1$, $k = 2$.

Action	Nr Cases	Distribution	# Rounds
0	0	0.000	–
1	70	0.070	$5,241$
2	462	0.462	$3,747$
3	417	0.417	$3,841$
4	51	0.051	$5,098$
5	0	0.000	–
TOTAL	$1,000$	1	$3,960$

Table 2: Simulation results for $h = 1$, $k = 5$.

Action	Nr Cases	Distribution	# Rounds
0	0	0.000	–
1	2	0.002	$6,499$
2	24	0.024	$6,544$
3	224	0.224	$4,501$
4	496	0.496	$3,940$
5	221	0.221	$4,457$
6	33	0.033	$6,118$
7	0	0.000	–
8	0	0.000	–
TOTAL	$1,000$	1	$4,319$

Table 3: Simulation results for $h = 1$, $k = 8$.

Action	Nr Cases	Distribution	# Rounds
0	0	0.000	–
1	0	0.000	–
2	0	0.000	–
3	33	0.033	$5,719$
4	257	0.257	$4,763$
5	458	0.458	$4,307$
6	220	0.220	$4,914$
7	32	0.032	$5,724$
8	0	0.000	–
9	0	0.000	–
10	0	0.000	–
TOTAL	$1,000$	1	$4,650$

Table 4: Simulation results for $h = 1$, $k = 10$.

The following table contains the results of the simulations of the one dimensional model, with $N = 38$ and $k = 10$, that were initialized with a population state in which all but one players play action 0, and the one player plays action 10. The payoff matrix A is the same PD equidistant payoff matrix as before.

Action	Nr Cases	Distribution	# Rounds
0	464	0.464	$3,998$
1	432	0.432	$5,148$
2	98	0.098	$6,514$
3	6	0.006	$6,921$
4	0	0.000	–
5	0	0.000	–
6	0	0.000	–
7	0	0.000	–
8	0	0.000	–
9	0	0.000	–
10	0	0.000	–
TOTAL	$1,000$	1	$4,759$

Table 5: Simulation results for $h = 1$, $k = 10$. All simulations were initialized in an almost all 0 state.

The following table contains the results of the simulations of the one dimensional model, with $N = 38$ and $k = 10$, that were initialized with a population state in which all but one players play action 10, and the one player plays action 0. The payoff matrix A is the same PD equidistant payoff matrix as before.

Action	Nr Cases	Distribution	# Rounds
0	0	0.000	–
1	0	0.000	–
2	0	0.000	–
3	0	0.000	–
4	0	0.000	–
5	0	0.000	–
6	0	0.000	–
7	5	0.005	7,815
8	104	0.104	5,569
9	435	0.435	5,292
10	456	0.456	3,875
TOTAL	1,000	1	4,687

Table 6: Simulation results for $h = 1$, $k = 10$. All simulations were initialized in an almost all 10 state.

The following table contains the result of the simulations of the one dimensional model, with $N = 38$ and $k = 10$, in which the stage game is Bertrand price competition. Actions are now setting prices. And action $l, l = 0, \ldots, k$ is setting a price $p_l = p^{Nash} + \frac{l}{k}\left(p^{cartel} - p^{Nash}\right)$. The demand function is taken to be $D(p_1, p_2) = 20 - p_1 + \frac{1}{2}p_2$ where p_1 is the subject's price and p_2 is the price of the other player. The cost function is $C(q) = q$.

Action	Nr Cases	Distribution	# Rounds
0	0	0.000	–
1	0	0.000	–
2	0	0.000	–
3	2	0.002	5,026
4	42	0.042	5,009
5	268	0.268	4,582
6	424	0.424	4,634
7	221	0.221	5,175
8	39	0.039	5,844
9	4	0.004	4,883
10	0	0.000	–
TOTAL	1,000	1	4,805

Table 7: Simulation results for $h = 1$, $k = 10$. The stage game is Bertrand price competition.

The following table contains the results of the simulations of the one dimensional model, with $N = 38$ and $k = 10$, that were initialized with a population state in which all but one players play action 0, and the one player plays action 10. The stage game is Bertrand price competition, as specified above table 7.

Action	Nr Cases	Distribution	# Rounds
0	150	0.150	$4,374$
1	318	0.318	$5,244$
2	319	0.319	$5,728$
3	154	0.154	$6,149$
4	48	0.048	$6,713$
5	11	0.011	$6,247$
6	0	0.000	–
7	0	0.000	–
8	0	0.000	–
9	0	0.000	–
10	0	0.000	–
TOTAL	$1,000$	1	$5,489$

Table 8: Simulation results for $h = 1$, $k = 10$. The stage game is Bertrand price competition. All simulations were initialized in an almost all 0 state.

The following table contains the results of the simulations of the one dimensional model, with $N = 38$ and $k = 10$, that were initialized with a population state in which all but one players play action 10, and the one player plays action 0. The stage game is Bertrand price competition, as specified above table 7.

Action	Nr Cases	Distribution	# Rounds
0	0	0.000	–
1	0	0.000	–
2	0	0.000	–
3	0	0.000	–
4	0	0.000	–
5	0	0.000	–
6	0	0.000	–
7	0	0.000	–
8	6	0.006	$5,304$
9	305	0.306	$4,837$
10	687	0.688	$2,854$
TOTAL	998	1	$3,475$

Table 9: Simulation results for $h = 1$, $k = 10$. The stage game is Bertrand price competition. All simulations were initialized in an almost all 10 state.

The following 4 tables contain the results of the simulations of the two dimensional model with $N = 4$ that were initialized with a uniform random population state. The payoff matrix A is PD equidistant, with $a_{0,0} = 200$ and parameter $a = 10$, i.e. $a_{i,j} = 200 - 10i + 20j$.

Action	Nr Cases	Distribution	# Rounds
0	243	0.243	$13,982$
1	505	0.505	$13,735$
2	252	0.252	$14,692$
TOTAL	$1,000$	1	$14,036$

Table 10: Simulation results for $h = 2$, $k = 2$.

Action	Nr Cases	Distribution	# Rounds
0	0	0.000	–
1	83	0.083	$15,290$
2	417	0.417	$14,167$
3	420	0.420	$14,848$
4	78	0.078	$16,091$
5	2	0.002	$9,120$
TOTAL	$1,000$	1	$14,686$

Table 11: Simulation results for $h = 2$, $k = 5$.

Action	Nr Cases	Distribution	# Rounds
0	0	0.000	–
1	0	0.000	–
2	39	0.039	$12,104$
3	244	0.244	$14,572$
4	403	0.403	$14,314$
5	265	0.265	$15,883$
6	48	0.048	$18,351$
7	1	0.001	$16,238$
8	0	0.000	–
TOTAL	1000	1	$14,902$

Table 12: Simulation results for $h = 2$, $k = 8$.

Action	Nr Cases	Distribution	# Rounds
0	0	0.000	–
1	0	0.000	–
2	3	0.003	$45,044$
3	41	0.041	$10,708$
4	218	0.218	$4,096$
5	413	0.413	$14,365$
6	257	0.257	$14,751$
7	66	0.066	$13,599$
8	2	0.002	$28,009$
9	0	0.000	–
10	0	0.000	–
TOTAL	$1,000$	1	$14,324$

Table 13: Simulation results for $h = 2$, $k = 10$.

The following table contains the results of the simulations of the two dimensional model, with $N = 4$ and $k = 10$, that were initialized with a population state in which all but one players play action 0, and the one player plays action 10. The payoff matrix A is the PD equidistant matrix specified above table 10.

Action	Nr Cases	Distribution	# Rounds
0	8	0.008	$11,939$
1	152	0.152	$13,353$
2	418	0.418	$13,012$
3	330	0.330	$13,062$
4	82	0.082	$15,166$
5	10	0.010	$9,051$
6	0	0.000	–
7	0	0.000	–
8	0	0.000	–
9	0	0.000	–
10	0	0.000	–
TOTAL	$1,000$	1	$13,209$

Table 14: Simulation results for $h = 2$, $k = 10$. All simulations were initialized in an almost all 0 state.

The following table contains the results of the simulations of the two dimensional model, with $N = 4$ and $k = 10$, that were initialized with a population state in which all but one players play action 10, and the one player plays action 0. The payoff matrix A is the PD equidistant matrix specified above table 10.

Action	Nr Cases	Distribution	# Rounds
0	0	0.000	–
1	0	0.000	–
2	0	0.000	–
3	0	0.000	–
4	0	0.000	–
5	11	0.011	15,115
6	87	0.087	13,922
7	316	0.316	13,138
8	414	0.414	14,485
9	160	0.160	13,869
10	12	0.012	10,723
TOTAL	1,000	1	13,873

Table 15: Simulation results for $h = 2$, $k = 10$. All simulations were initialized in an almost all 10 state.