Effects of increasing the number of players and memory size in the iterated Prisoner's Dilemma: a numerical approach

CH. HAUERT AND H. G. SCHUSTER

Department of Theoretical Physics, Christian-Albrechts-University of Kiel, Leibnizstrasse 15/242, D-24098 Kiel, Germany (hauert@theo-physik.uni-kiel.de)

SUMMARY

The Prisoner's Dilemma has become a paradigm for the evolution of altruistic behaviour. Here we present results of numerical simulations of the infinitely iterated stochastic simultaneous Prisoner's Dilemma considering players with longer memory, encounters of more than two players as well as different pay-off values. This provides us with a better foundation to compare theoretical results to experimental data. We show that the success of the strategy *Pavlov*, regardless of its simplicity, is far more general by having an outstanding role in the iterated N-player N-memory Prisoner's Dilemma. Besides, we study influences of increased memory sizes in the iterated two-player Prisoner's Dilemma, and present comparisons to results of experiments with first-year students.

1. INTRODUCTION

The Prisoner's Dilemma (PD) in its simultaneous or alternating form has often been used to explain altruistic behaviour. It is a very simple game where any rational player defects when played one round only but where it pays to cooperate if repeated with high probability. In biology such altruistic behaviour has been observed, for example, in sticklebacks (Milinski 1987), baboons (Trivers 1971; Packer 1977) and vampire bats (Wilkinson 1984).

Parallel to such biological observations, numerous theoretical investigations were made concerning various aspects of the PD (see, for example, Binmore & Samuelson 1992; Wu & Axelrod 1996). The effects of different memory sizes in the two-player PD have been studied, for example, by Axelrod (1987) and Lindgren (1991). Deterministic strategies in the N-player PD were studied, for example, by Boyd & Richerson (1988). Axelrod & Dion (1988) presented a review on the N-player PD and Dawes (1980) summarized experimental approaches.

In this paper we present a generalization to N players with M-step memory of the stochastic iterated PD worked out by Nowak & Sigmund (1993) for two players with two-step memory. In particular we investigate whether the strategy, *Pavlov*, remains successful when increasing the memory size and the number of players.

Due to the complexity and the nonlinearity of the system, our investigations are based on extensive numerical simulations. Unfortunately, the dimension of the strategy space increases exponentially with the memory size of the players. This forces us to limit our simulations to $N, M \leqslant 6$.

2. THE GAME

The players in PD have two options. They must either cooperate ('c') or defect ('d'). In the simultaneous PD all players make their decisions or moves at once. The pay-off for each player depends only on the outcome of the previous round, and thus remains unaffected by the memory length. The following rules were used to determine the pay-off matrix for an arbitrary number of players.

1. A cooperating player receives a reward of R points from every cooperating opponent, but only the sucker's pay-off S from defecting ones.
2. A defecting player gets the temptation of T points from every cooperating opponent, but only the punishment P from defecting ones.

Because $T > R > P > S$ and $R > \frac{1}{2}(T+S)$, it always pays for an individual to switch from cooperation to defection, but all other players will profit if a single one switches from defection to cooperation. Thus rational players will end up getting only P instead of R points from every opponent. Hence the dilemma remains intact. Setting $R = 3$, $S = 0$, $T = 5$, $P = 1$, the above rule leads for two players to the familiar pay-off matrix introduced by Axelrod (1984).

Whether a player opts for c or d in the next round depends on the history of the game and thus on the memory size of the player (see figure 1). The strategy adopted by a player is encoded in a vector. The entries specify the conditional probabilities to cooperate for every possible outcome in the history recalled by the player. A player remembering the last M moves plays a strategy consisting of 2^M conditional probabilities. The probabilities are indexed by

Proc. R. Soc. Lond. B (1997) **264**, 513–519
Printed in Great Britain

© 1997 The Royal Society
TEX Paper

(*a*) History of a two player game:

Time	*Player 1*	*Player 2*
⋮	⋮	⋮
t - 2	c	c
t - 1	d	c
t	c	d

(*b*) Memory of the two players (four steps):

Time	*Player 1*	*Player 2*
⋮	⋮	⋮
t - 1	ccdc	cccd
t	dccd	cddc

Figure 1. (*a*) Three sample rounds of a two-player game. (*b*) The respective memories of the two players at times $t-1$ and t. The remembered history of each player is used to index the conditional probabilities to cooperate in this situation. Thus, player 1 will cooperate with probability p_{dccd} at $t+1$ and player 2 with probability p_{cddc}.

$$\left|\begin{array}{llll} p^1_{\mathrm{c\cdots cc}}p^2_{\mathrm{c\cdots cc}}\cdots p^N_{\mathrm{c\cdots cc}} & p^1_{\mathrm{c\cdots cc}}p^2_{\mathrm{c\cdots cc}}\cdots(1-p^N_{\mathrm{c\cdots cc}}) & \cdots & (1-p^1_{\mathrm{c\cdots cc}})\cdots(1-p^N_{\mathrm{c\cdots cc}}) \\ p^1_{\mathrm{c\cdots cd}}p^2_{\mathrm{c\cdots dc}}\cdots p^N_{\mathrm{dc\cdots c}} & p^1_{\mathrm{c\cdots cd}}p^2_{\mathrm{c\cdots dc}}\cdots(1-p^N_{\mathrm{dc\cdots c}}) & \cdots & (1-p^1_{\mathrm{c\cdots cd}})\cdots(1-p^N_{\mathrm{dc\cdots c}}) \\ \vdots & & & \\ p^1_{\mathrm{d\cdots dd}}p^2_{\mathrm{d\cdots dd}}\cdots p^N_{\mathrm{d\cdots dd}} & p^1_{\mathrm{d\cdots dd}}p^2_{\mathrm{d\cdots dd}}\cdots(1-p^N_{\mathrm{d\cdots dd}}) & \cdots & (1-p^1_{\mathrm{d\cdots dd}})\cdots(1-p^N_{\mathrm{d\cdots dd}}) \end{array}\right|. \qquad (1)$$

$$\left|\begin{array}{ccccccc} p^1_{\mathrm{c\cdots cc}}p^2_{\mathrm{c\cdots cc}} & \cdots & (1-p^1_{\mathrm{c\cdots cc}})(1-p^2_{\mathrm{c\cdots cc}}) & 0 & \cdots & & \\ 0 & \cdots & 0 & p^1_{\mathrm{c\cdots cd}}p^2_{\mathrm{c\cdots dc}} & \cdots & 0 & \cdots \\ \vdots & & & & & & \\ p^1_{\mathrm{cdc\cdots c}}p^2_{\mathrm{dc\cdots cc}} & \cdots & (1-p^1_{\mathrm{cdc\cdots c}})(1-p^2_{\mathrm{dc\cdots c}}) & 0 & \cdots & & \\ \vdots & & & & & & \\ 0 & \cdots & 0 & 0 & \cdots & p^1_{\mathrm{d\cdots dd}}p^2_{\mathrm{d\cdots dd}} & \cdots \end{array}\right|. \qquad (2)$$

the respective history. p_{dccd}, for example, specifies the probability to cooperate at time $t+1$ after receiving T at $t-1$ and S at t in a two-player four-step memory game. Note that moves in the past stand to the left, and most recent moves to the right, of the index. The move of the player himself is noted first in each round. The probability to cooperate given a specified but arbitrary history x will be referred to as p_x.

First we look at the iterated PD for N players ($N > 2$) with M-step memory. Strategies, as explained previously, do not incorporate mechanisms to identify opponents. Moreover, the players cannot profit from distinguishing the opponents' moves because the encounters are set-up symmetrically (see equations (3)). Thus, the probability to cooperate depends only on the *number* of cooperating respectively defecting opponents in *each* round. This reduces the strategy space dramatically from $2^M = 2^{mN}$ to $2^m N$ dimensions, where $m \in \mathbb{N}$ corresponds to the number of rounds recalled by the players. Therefore, for example, for three players with three-step memory the following identities must hold: $p_{\mathrm{ccd}} \equiv p_{\mathrm{cdc}}$ and $p_{\mathrm{dcd}} \equiv p_{\mathrm{ddc}}$.

The transition from one round to the next is given by a Markov chain. Each element in the Markov matrix (1) specifies the transition probability of outcome x at time t to x' at $t+1$ for N players remembering one round ($m = 1$). p^i_x specifies the probability to cooperate for player i given the history x (see matrix (1)). The rows of matrix (1) are labelled by all possible histories remembered by the players at time t, and the columns by the history at $t+1$. Note that the history looks different for every player.

As mentioned above the $2^{mN} \times 2^{mN}$ matrix might be reduced to $2^m N \times 2^m N$ dimensions. However, this would complicate the calculation of the transition probabilities, and combinatorial factors would be needed for the calculation of the pay-off (see equations (3)).

Now we turn to the two-player PD with more than one-round memory ($M > 2$). Since all transitions are no longer possible, a considerable number of transition probabilities in the Markov matrix (2) are zero. As previously, the rows in matrix (2) indicate the remembered history at time t and the columns at $t+1$. Note that again the history looks different for each player. The elements of Markov matrix (1) are strictly positive (due to errors in interpreting and implementing a move) and those of matrix (2) are positive. Since both matrices are irreducible there exists a unique left eigenvector to the eigenvalue 1 (Frobenius's theorem). The elements of the eigenvector specify the frequencies of the respective patterns in history. To determine the pay-off of the different players we have to multiply the frequency of each pattern with the corresponding pay-off value. Accumulating these values for each player yields his mean pay-off. In evolutionary game dynamics the pay-off of a player determines his reproductive success or similarly his frequency in the population (Maynard Smith 1982). If the population consists at time t of I different strategies with frequencies f^t_i then their

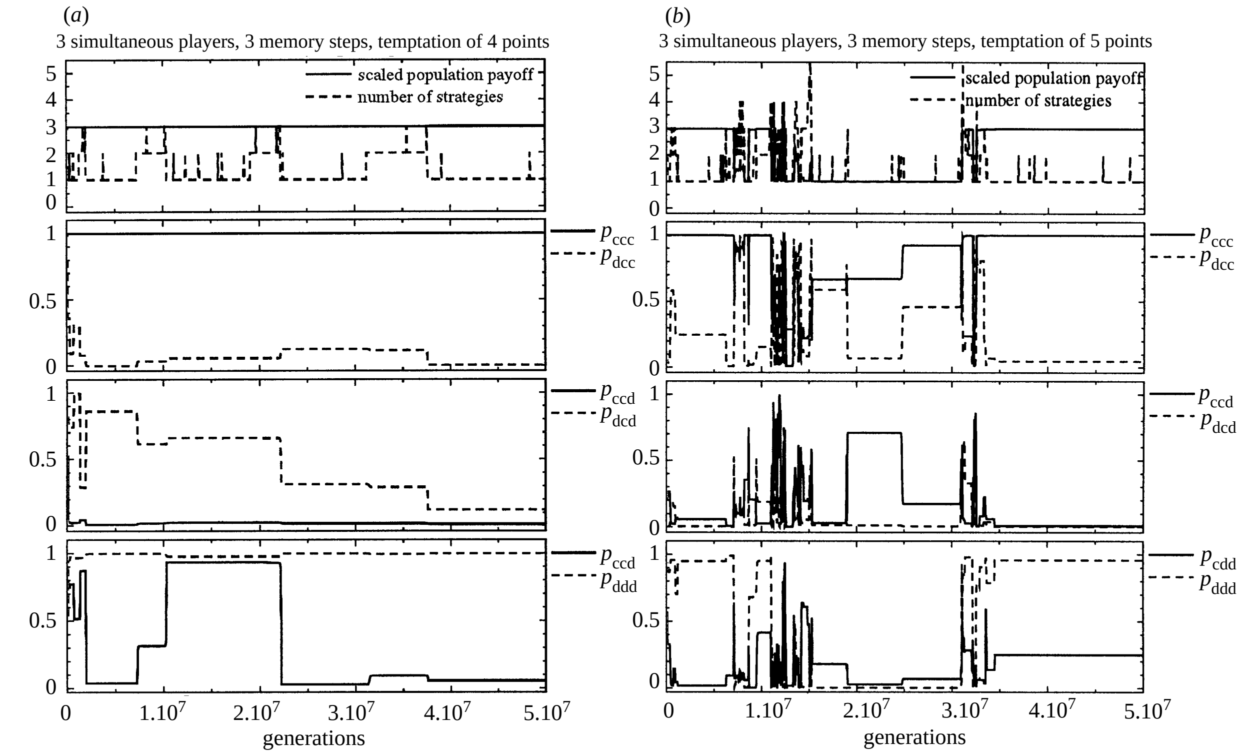

Figure 2. The evolving strategies for three simultaneous players with temptation $T = 4$ (a) and $T = 5$ (b) are very similar. Both strategies are nice ($p_{\rm ccc} > 0.98$) and try to restore cooperation after a round of mutual defection ($p_{\rm ddd} \approx 1$). All other p_x are close to zero. Thus errors are not ignored regardless of the temptation T. The increased temptation in (b) makes it more difficult to establish cooperation, resulting in a longer period of chaotic transitions.

frequencies at $t+1$ are given by

$$f_i^{t+1} = f_i^t \frac{a_i}{a},$$

where

$$\left.\begin{aligned} a_i &= \sum_{n_1,n_2,\cdots,n_{N-1}=1}^{I} f_{n_1}^t f_{n_2}^t \cdots f_{n_{N-1}}^t \\ &\qquad \times a_{i,n_1,n_2,\ldots,n_{N-1}}, \\ a &= \sum_{i=1}^{I} f_i^t a_i \end{aligned}\right\} \tag{3}$$

and

$$\sum_{i=1}^{I} f_i^t = 1, \quad \forall t.$$

The $a_{i,n_1,n_2,\ldots,n_{N-1}}$ stand for the mean pay-off obtained by strategy i when playing against strategies $n_1 \cdots n_{N-1}$ in the N-player PD. A strategy i obtaining a higher total pay-off than the population pay-off ($a_i > a$) will spread.

At $t = 0$ we start with a homogeneous population playing a completely random strategy: $p_x = 0.5$. On average, every 100 generations a mutation occurs and a new strategy is introduced into the population. The initial frequency of a new strategy is set to 0.0011. Whenever the frequency of any strategy drops below a fixed noise level of 0.001 it is removed. New strategies are chosen in a straightforward manner: An appropriate number of random values within the range of $[0.001, 0.999]$ are drawn as probabilities to cooperate for each possible history recalled by the player. In order to explore the corners of the M-dimensional cube, the random numbers were drawn according to the U-shaped density distribution $(\pi x(1-x))^{-1/2}$.

3. RESULTS

The following figures display the evolution of the most successful strategies out of ten simulation runs. Since the search space is far too big to perform reliable statistics, the surviving strategies were merged in a new population. Their frequencies were set proportional to the population pay-off weighted by their frequencies at the end of the respective simulation. The evolution of the strategy achieving the highest pay-off in this population is shown in the figures.

For each arrangement of players and memory lengths we present comparisons between successful strategies for pay-off values of $R = 3$, $S = 0$, $T = 4$, $P = 1$ (in figures labelled (a)) and an increased temptation of $T = 5$ (labelled (b)). The figures show the population pay-off, the number of strategies, as well as the weighted mean probabilities to cooperate calculated over all strategies in the population.

(a) *N-player N-step memory*

Figures 2 and 3 show the results for the simultaneous iterated N-player N-memory PD ($N = 3, 4$). The figures are arranged as follows: the top graph shows the time evolution of the population pay-off as well

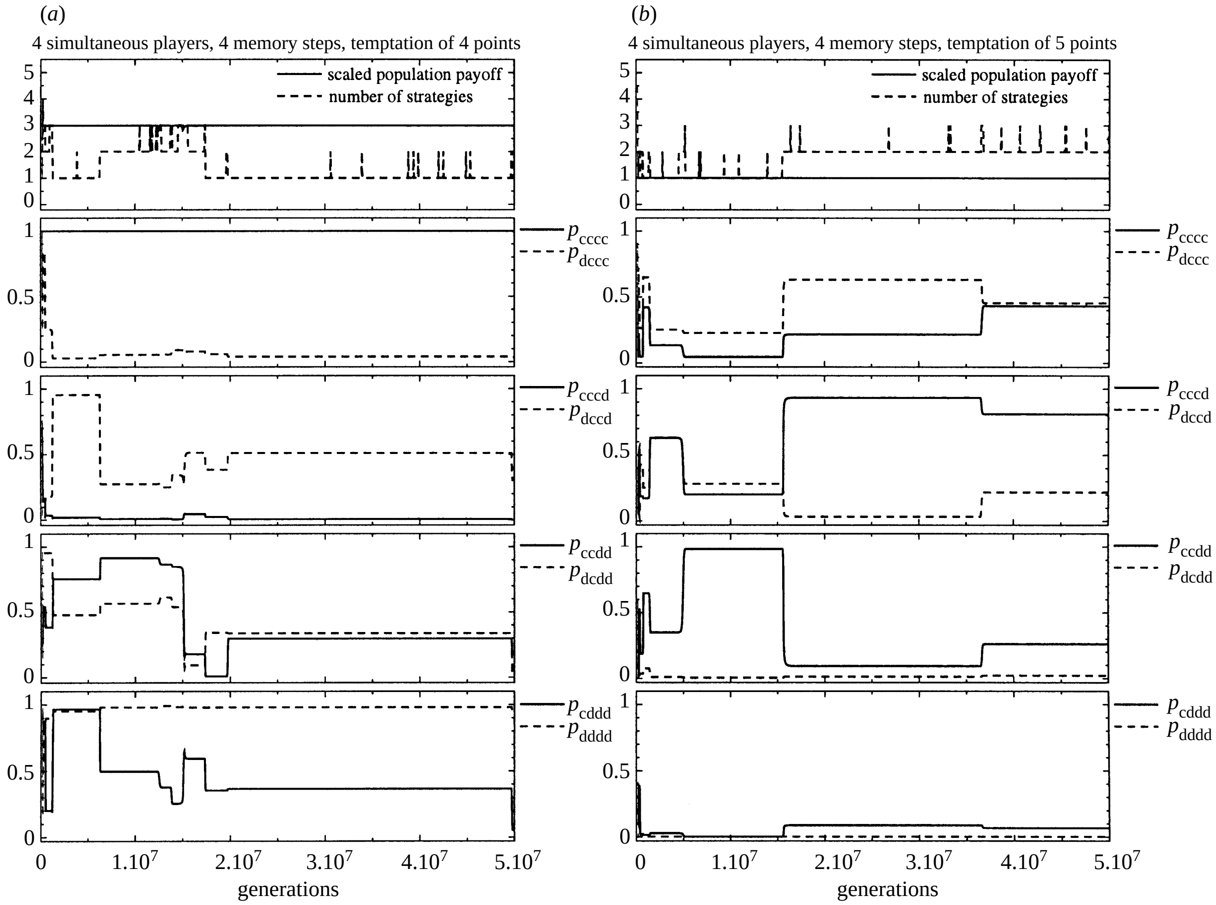

Figure 3. In the simultaneous four-player PD we observed significant differences depending on the temptation $T = 4$ (a) and $T = 5$ (b). In (a) the winning strategy remains nice ($p_{cccc} > 0.98$) and is close to *generic Pavlov*. Strategies in (b) show the characteristics of traitors ($p_{cccc} < 0.5$), leading to a poor population pay-off close to the minimum of 1.

as the number of strategies present in the population. In order to simplify comparisons the population pay-off was scaled such that a round of mutual co-operation pays three points regardless of the number of players. Each graph below shows the conditional probabilities for the two possible moves c and d of the player himself and for an increasing number of defecting opponents.

Figures 2a, b show simulations for the three-player three-memory step PD. Both situations lead to strategies cooperating only after a round of mutual cooperation or defection ($p_{c\cdots c}$, $p_{d\cdots d} \approx 1$, all other $p_x \ll 1$). Such strategies will be called *generic Pavlov* in the following. Intuitively it is clear that for an increasing number of players it gets more and more difficult to establish cooperation. Thus, for successful cooperative strategies it becomes very important to correct errors readily when playing against one's own kind without giving defectors a chance to exploit. In this context the temptation T plays a crucial role too. The strategy *generic Pavlov* implements these features in a straightforward manner.

At the end of ten simulation runs, 14 out of 16 strategies were nice ($p_{ccc} > 0.98$) for $T = 4$, whereas only 11 out of 18 strategies were nice for $T = 5$.

In figures 3a, b simulations of the four-player PD are shown. Only in figure 3a, with the lower temptation $T = 4$, were cooperative strategies close to *generic Pavlov* found. Eleven out of 12 strategies surviving the ten simulation runs were nice, whereas for $T = 5$ only 12 out of 17 were nice ($p_{cccc} > 0.98$). This might be the reason why no nice strategy won in figure 3b. Defectors can intrude in any population when introduced with high enough frequency. In this case the losses of cooperators against defectors are not outweighed by their wins against other cooperators. Nevertheless, for $T = 5$ the second best strategy was nice. It achieved only a slightly lower pay-off of 4.715 compared to 4.775 points for the winning strategy. This strategy is very close to the one shown in figure 3a, and again is similar to *generic Pavlov* ($p_{cccc} = 0.998$, $p_{cccd} = 0.002$, $p_{ccdd} = 0.051$, $p_{cddd} = 0.124$, $p_{dccc} = 0.003$, $p_{dccd} = 0.612$, $p_{dcdd} = 0.093$, $p_{dddd} = 0.834$).

Results of the five-player PD with $T = 4$ and $T = 5$ are not shown. Due to enormous CPU time requirements for each temptation only two simulation runs were made, ending with two nice strategies for $T = 4$, whereas for $T = 5$ only two out of three strategies were nice ($p_{ccccc} > 0.98$). In general, the nice strategies seem to fit well in the concept of *generic Pavlov*, but further simulations are needed in order to obtain reliable results.

To put *generic Pavlov* in the context of win–stay lose–shift strategies we see that $(N-1)R$ and $iT + (N-1-i)P$ with $i = 1, 2, \ldots, N-1$ are con-

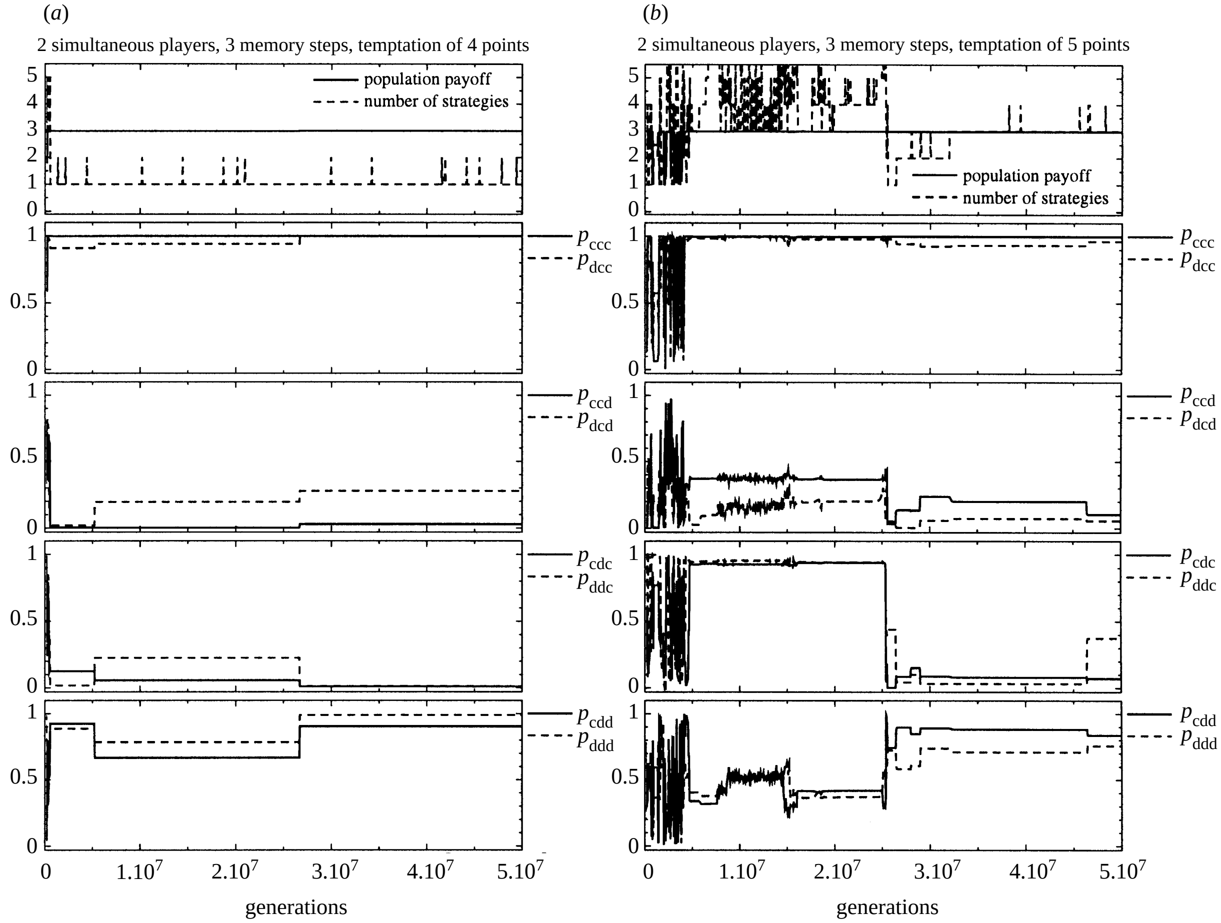

Figure 4. For two players with three-step memory very similar strategies were found regardless of the temptation $T = 4$ (*a*) and $T = 5$ (*b*). The increased memory size does not lead to significant advantages and thus the strategies in (*a*) and (*b*) remain very close to *Pavlov*. The increased temptation in (*b*) leads to a longer period of chaotic transitions before cooperation is established. Interestingly, we see that *generous TFT* seems to trigger the change to cooperation and paves the way for Pavlovian strategies. This catalysing effect of TFT-like strategies was observed by Nowak & Sigmund (1993). Due to the different temptations we also see that the readiness to cooperate after a round of mutual defection is lower in (*b*), for obvious reasons.

sidered as wins. Whereas pay-offs of $(N-1)P$ and $iS + (N-1-i)R$ with $i = 1, 2, \ldots, N-1$ are taken as losses. The number of winning outcomes equals those considered as losses. Note that for $N \geqslant 4$ the top scores considered as losing are higher than the lowest winning scores. Thus the decision strongly depends on the players' own moves.

(*b*) *Two-player M-step memory*

Figures 4 and 5 show simulations of the simultaneous iterated two-player M-memory PD ($M = 3, 4$). Again the top graph shows the population pay-off as well as the number of strategies in the population. The graphs below show the mean conditional probabilities to cooperate grouped according to the four possible outcomes of the last round. This representation simplifies comparisons to strategies with two-step memory such as *Pavlov* or *generous TFT* (see later).

In order to study influences of increasing the memory size we first summarize results of similar simulations done by other researchers. A straightforward reasoning shows that for players without memory a defecting strategy may take over any population regardless of its composition. The situation significantly changes for players remembering the opponent's last move. Players will adopt the well-studied cooperative strategy 'tit-for-tat' (TFT)—do whatever the opponent did in the previous round. Later TFT is outperformed by the more forgiving strategy *generous TFT* (Nowak & Sigmund 1992) with $p_{cc} = 1$, $p_{cd} = \frac{1}{3}$, $p_{dc} = 1$ and $p_{dd} = \frac{1}{3}$. The situation again changes considerably for players with two-step memory leading to the strategy *Pavlov* (Nowak & Sigmund 1993)—implementing the rule win–stay lose–shift ($p_{cc} = 1$, $p_{cd} = 0$, $p_{dc} = 0$ and $p_{dd} \approx 1$).

In figure 4 simulations are shown for strategies with three-step memory, remembering the last round plus the opponent's move in the previous round. The simulations led to similar strategies very close to *Pavlov*. The increased memory size allows refined actions but the strategies show no new features. This is in agreement with simulations of the PD with deterministic but not error-free strategies done by Lindgren (1991). The population is dominated only for a very short period by deterministic memory-three strategies. Thus there is no significant advantage in

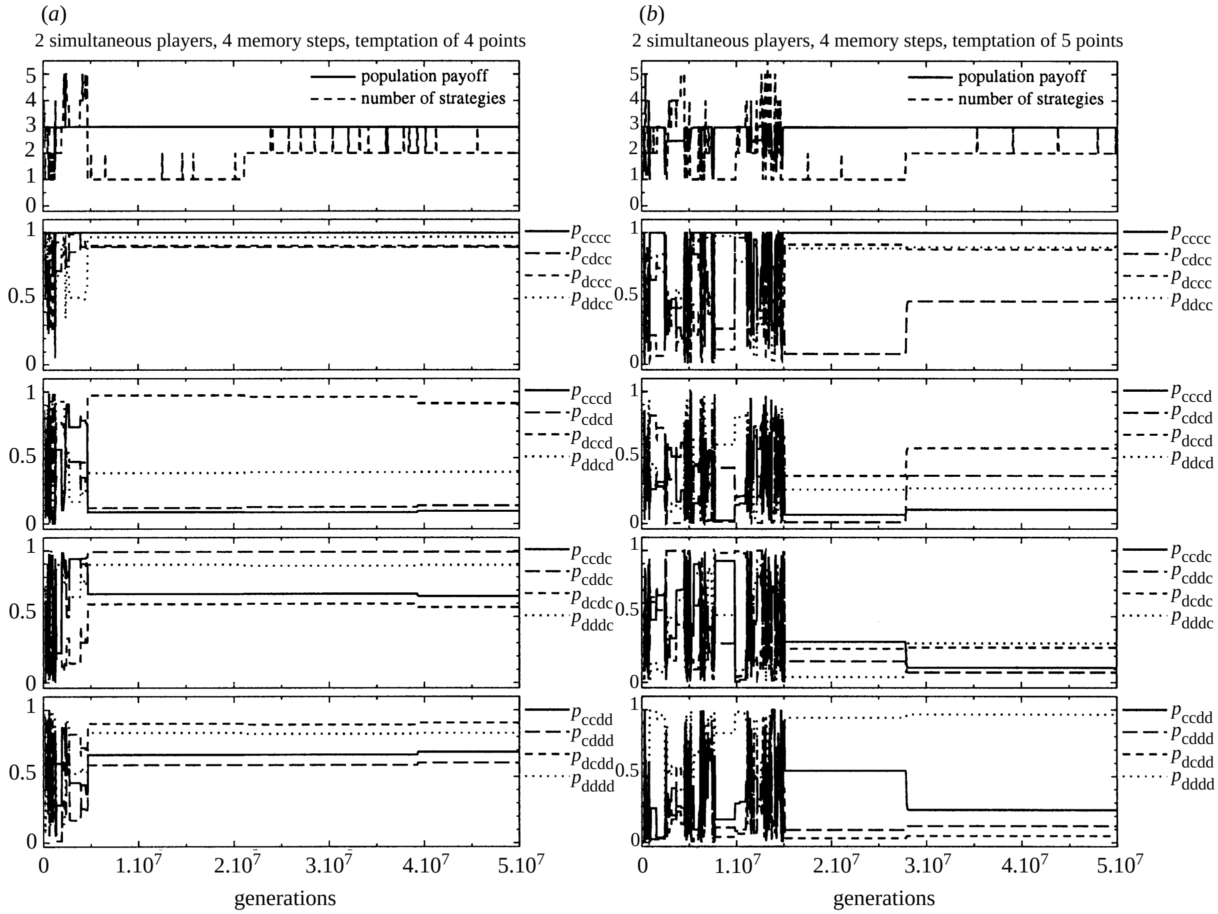

Figure 5. Increasing the memory size of the players to two rounds leads to interesting strategies depending heavily on the temptation $T = 4$ (*a*) and $T = 5$ (*b*). Both situations lead to nice strategies ($p_{\mathrm{cccc}} > 0.98$) but the strategies in (*a*) are close to *generous TFT*, whereas in (*b*) Pavlovian strategies establish. Important differences to these strategies due to the longer memory are discussed in the text.

increasing the memory from two to three steps. In figure 4*b* the increased temptation makes it much harder to establish cooperation. The ten simulation runs ended with 21 nice strategies out of 22 for $T = 4$, and for $T = 5$ only 16 out of 24 were nice ($p_{\mathrm{ccc}} > 0.98$).

Results of simulations for strategies with four-step memory are shown in figure 5. The temptation of $T = 4$ in figure 5*a* leads to strategies close to *generous TFT*, with the important difference that given the history dccd the strategy accepts punishment for defecting and tries to restore cooperation ($p_{\mathrm{dccd}} \approx 1$). Such characteristics are implemented in a different way by a strategy called *contrite TFT* (see, for example, Boerlijst *et al.* 1997). Figure 5*b* shows that the increased temptation again leads to a longer chaotic period until cooperation is established. The resulting strategies are closer to *Pavlov* than *generous TFT* with several modifications due to the longer memory. If a player obtained the sucker's pay-off in the previous round he would not cooperate readily, even when followed by a round of mutual cooperation ($p_{\mathrm{cdcc}} \approx 0.5$). After a round of mutual defection the previous round determines whether cooperation is restored ($p_{\mathrm{dddd}} \approx 1$, but $p_{\mathrm{ccdd}}, p_{\mathrm{cddd}}, p_{\mathrm{dcdd}} \approx 0$). Thus increasing the temptation stresses the importance of the past. The longer memory makes it easier to establish cooperation since defectors can be detected more readily. At the end of the ten simulation runs for $T = 4$ all 22 strategies were nice, and for $T = 5$ all 17 ($p_{\mathrm{cccc}} > 0.98$) were nice.

Results of simulations for five- and six-step memory are not shown. Because of enormous CPU time requirements, we made only five simulation runs for each memory size with $T = 4$. For memory-five, as well as for memory-six, all 11 strategies were nice ($p_{\mathrm{ccccc}} > 0.98$, $p_{\mathrm{cccccc}} > 0.98$). The results of these simulations will be used for comparisons with experimental data (see table 1).

4. CONCLUSIONS

The simulations clearly show that an increasing number of players as well as an increasing temptation T hinder the establishment of cooperation. Nevertheless, we see that cooperative solutions of the dilemma still do exist. The strategy *generic Pavlov*, cooperating with a very high probability only after a round of mutual cooperation or defection ($p_{\mathrm{c\cdots c}}, p_{\mathrm{d\cdots d}} \approx 1$, all other $p_x \ll 1$), seems to be very powerful in the simultaneous N-player N-memory PD. Interestingly, even in the N-player PD it does not seem to be advantageous to ignore errors of an opponent. On

Table 1. *The experimental data listed above are taken from Wedekind & Milinski (1996) showing averages drawn from simultaneous and alternating PDs. The best strategy found for each memory length was used to calculate the average probability to cooperate after experiencing cc, cd, dc or dd in the last round. Both simulations and experiments were made for pay-off values of* $R = 3$, $S = 0$, $T = 4$, $P = 1$. *The data marked with an asterisk '*' represents Pavlov, found by Nowak & Sigmund (1993). Note that with increasing memory size the* $\tilde{p}_x$ *seem to converge rapidly to* 0.5. *Thus, to an observer with only one-round memory, the moves of a strategy with three-round memory are essentially indistinguishable from a random strategy with* $p_x = 0.5$.

experimental data				
type	p_{cc}	p_{cd}	p_{dc}	p_{dd}
Pavlovian	0.65	0.25	0.11	0.17
generous TFT-like	0.88	0.20	0.95	0.17
simulated reduced data				
memory length	$\tilde{p}_{cc}$	$\tilde{p}_{cd}$	$\tilde{p}_{dc}$	$\tilde{p}_{dd}$
2	0.999	0.001	0.001	0.999*
3	0.998	0.152	0.010	0.950
4	0.954	0.311	0.652	0.897
5	0.584	0.452	0.592	0.612
6	0.598	0.617	0.423	0.492

the contrary, an error is followed by a round of mutual defection before cooperation is restored. Literally speaking, this helps to avoid further misunderstandings, and it is the fastest way to restore cooperation when playing against one's own kind without giving defectors a chance to exploit.

Increasing the memory size helps to establish cooperation because traitors can be detected more easily. This is reflected by the vanishing number of traitors present at the end of the simulation runs for longer memories. The resulting strategies remain close to *Pavlov* and *generous TFT* with several though important differences.

Recently, Wedekind & Milinski (1996) did interesting experiments with first-year students in order to find out what strategy humans are adopting in the PD. In these experiments the unknown but finite length of the game and the players memory are likely to have strong influences on the strategies played. To get an idea how the memory affects the readiness to cooperate we calculated the mean probability to cooperate over the remembered history after experiencing cc, cd, dc and dd in the last round. Thus, for a memory-four strategy we would have, for example, $\tilde{p}_{cc} = \frac{1}{4}(p_{cccc} + p_{cdcc} + p_{dccc} + p_{ddcc})$.

Comparing experimental and simulation data in table 1 we see that considerable differences exist regardless of the memory lengths. Thus, either the experimental set-up is not suitable for such comparisons or, more likely, other effects such as the finite length of the games and mechanisms such as contrition must also have strong influences. In fact, it is not a very efficient way to deal with resources when remembering only a certain number of moves, instead of certain patterns triggering the decisions.

We thank K. Sigmund for revising the manuscript and for helpful comments.

REFERENCES

Axelrod, R. 1987 The evolution of strategies in the iterated Prisoner's Dilemma. In *Genetic algorithms and simulated annealing* (ed. D. Lawrence), pp. 32–41. London: Pitman.

Axelrod, R. 1984 *The evolution of cooperation.* New York: Basic Books.

Axelrod, R. & Dion, D. 1988 The further evolution of cooperation. *Science* **242**, 1385–1390.

Binmore, K. G. & Samuelson, L. 1992 Evolutionary stability in repeated games played by finite automata. *J. Econ. Theory* **57**, 278–305.

Boerlijst, M. C., Nowak, M. A. & Sigmund, K. 1997 The logic of contrition. *J. Theor. Biol.* (Submitted.)

Boyd, R. & Richerson, P. J. 1988 The evolution of reciprocity in sizeable groups. *J. Theor. Biol.* **132**, 337–356.

Dawes, M. R. 1980 Social dilemmas. *A. Rev. Psychol.* **31**, 169–193.

Lindgren, K. 1991 Evolutionary phenomena in simple dynamics. In *Artificial life II* (ed. C. G. Langton, J. D. Farmer, S. Rasmussen & C. Taylor), pp. 295–312. Redwood City: Addison-Wesley.

Maynard Smith, J. 1982 *Evolution and the theory of games.* Cambridge University Press.

Milinski, M. 1987 Tit for tat in sticklebacks and the evolution of cooperation. *Nature, Lond.* **325**, 434–435.

Nowak, M. A. & Sigmund, K. 1992 Tit for tat in heterogenous populations. *Nature, Lond.* **355**, 250–253.

Nowak, M. A. & Sigmund, K. 1993 A strategy of win–stay, lose–shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature, Lond.* **364** 56–58.

Packer, C. 1977 Reciprocal altruism in *Papio anubis. Nature, Lond.* **265**, 441–443.

Trivers, R. 1971 The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–57.

Wedekind, C. & Milinski, M. 1996 Human cooperation in the simultaneous and the alternating Prisoner's Dilemma: Pavlov versus generous tit-for-tat. *Proc. Natn. Acad. Sci. USA* **93**, 2686–2689.

Wilkinson, G. S. 1984 Reciprocal food-sharing in the vampire bat. *Nature, Lond.* **308**, 181–184.

Wu, J. & Axelrod, R. 1995 How to cope with noise in the iterated Prisoner's Dilemma. *J. Conflict Resolution* **39**, 1, 183–189.

Received 27 November 1996; accepted 17 December 1996