

## Automata, repeated games and noise

Martin A. Nowak<sup>1</sup>, Karl Sigmund<sup>2</sup>, Esam El-Sedy<sup>3</sup>

<sup>1</sup>Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, U.K.

<sup>2</sup>Institut für Mathematik, Universität Wien Strudlhofgasse 4, A-1090 Vienna, Austria

<sup>3</sup>Department of Mathematics, Ain Shames University, Abassia, Cairo, Egypt

Received 23 April 1993; received in revised form 2 February 1994

**Abstract.** We consider two-state automata playing repeatedly the Prisoner's Dilemma (or any other  $2 \times 2$ -game). The  $16 \times 16$ -payoff matrix is computed for the limiting case of a vanishingly small noise term affecting the interaction. Some results concerning the evolution of populations of automata under the action of selection are obtained. The special role of 'win-stay, lose-shift'-strategies is examined.

### 1 Introduction

It is almost in the nature of a game that it can be repeated a number of times, and a growing part of game theory deals with such iterated encounters. The best-known example is probably the Prisoner's Dilemma, but there are many other iterated games in biology, psychology and economics.

In the Prisoner's Dilemma, two players have the options to cooperate (C) or to defect (D). If both players cooperate, both obtain 3 points, say; if both defect, both receive 1 point; if one player defects and the other cooperates, the defector receives 5 points and the cooperator 0. Clearly strategy D is best, no matter what the other player does; it *dominates* the cooperative option. Hence both players will use D and end up with one point only, instead of three points for mutual cooperation.

This picture greatly changes if the game is repeated with a high probability. There is no longer, in this case, a dominant strategy which yields the best reply against any possible strategy of the co-player, because the short-term advantage obtained by playing D in one round can make one lose the prospect of a future mutual cooperation. This 'shadow of the future' plays a central role in Axelrod's investigations of the Prisoner's Dilemma and in many applications to theoretical biology and the evolution of cooperation (see Axelrod 1984, Axelrod and Hamilton 1981 and Axelrod and Dion 1988). In

particular, it shows that the possibility of further rounds can totally alter the character of a game.

The number of possible strategies for an iterated game grows exponentially with the number of rounds in the game. This makes it almost imperative to restrict the strategy space drastically. One such approach is to consider only strategies defined by finite state automata. We shall follow this approach, but assume, furthermore, that the automata are to some small degree error-prone. Our main motivation is biological. In evolutionary game theory, we have to deal with populations of players following very simple rules (certainly a far cry from fully rational agents) and subject to mistakes in perceiving the action (or identity) of the co-player and in implementing their own behavioural program. We shall only consider very simple interactions, like the Prisoner's Dilemma or Chicken (the Hawk-Dove game of animal behaviourists), but emphasize that our approach is valid in a far more extended context.

Specifically, we consider games with finitely many strategies which are repeated infinitely often. Even the least likely error is bound to occur sometimes. The finite-state automata representing the strategies of the iterated game will always assume all their possible states. In Sect. 2, we shall use probabilistic arguments to derive the payoff matrix for such games if they are affected by errors in implementation. In Sect. 3, we shall deal with game-dynamical aspects: the results, here, are far from complete, due to the high dimension of the problem and the large number of possible cases. In Sect. 4, we deal more systematically with the family of 'win-stay, lose-shift'-strategies, and in Sect. 5, we briefly consider the effect of errors in perception.

This paper is a complement of Nowak and Sigmund (1993a, b). In Nowak and Sigmund (1993a), we have emphasised the dynamical richness of the infinitely iterated Prisoner's Dilemma. If one assumes that strategies are transmitted in the population from one generation to the next in proportion to their success (i.e. to their average payoff in the population), and if this transmission is again error-prone (i.e. subject to a small mutation rate), then one finds a wealth of interesting dynamics including limit cycles and chaotic attractors. This occurs if one considers, for instance, populations composed of *all* conceivable deterministic two-state automata. The wide-spread instability seems to be essentially caused by the prevalence of heteroclinic 'stone-scissors-paper'-cycles on the boundary of the state space. Indeed, we can show that there are no equilibria with all types of automata present: selection drives the population towards the boundary of the state space.

In Nowak and Sigmund (1993b), we have performed extensive evolutionary simulations involving *stochastic* two-state automata (programmed to play one or the other alternative with well-defined probabilities, depending on the outcome of the previous round). In this case, we allowed mutations to create *new* strategies from time to time. The evolutionary chronicles displayed punctuated equilibria: the population was usually either always cooperating or always defecting. There was a distinct trend towards cooperation, and within cooperative populations towards the adoption of 'Pavlov', the only

'win-stay, lose-shift'-strategy which yields cooperation when adopted by the whole population. In particular, tit for tat was outperformed by Pavlov. This result motivates the study of the biologically eminently plausible class of 'win-shift, lose-stay' strategies in the last part of our paper.

## 2 The payoff matrix

We consider a simple  $2 \times 2$ -game with two pure strategies  $C$  and  $D$  and payoff matrix

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} \quad (1)$$

The letters indicate that we have primarily in mind the Prisoner's Dilemma:  $C$  stands for *Cooperate*,  $D$  for *Defect*,  $R$  for *Reward*,  $S$  for the *Sucker's payoff*,  $T$  for *Temptation* and  $P$  for *Punishment*. But we emphasize that the following also applies to all other  $2 \times 2$ -games.

We consider now the iterated game which consists in repeating the simple  $2 \times 2$ -game infinitely often, i.e. with probability 1. It admits a continuum of possible strategies. Following Rubinstein (1986), Abreu and Rubinstein (1988), Miller (1989), Banks and Sundaram (1990), Lindgren (1991), Binmore and Samuelson (1992) and Probst (1993), we restrict our attention to strategies which are given by finite state automata: more precisely, by two-state automata only. Each of the two players is now an automaton which can be, in any given round of the iterated game, in one of two states. These states correspond to the two possible moves  $C$  and  $D$ . The state of the player in the following round depends on the present state and on the other player's move. For instance, an automaton playing tit for tat always assumes the state corresponding to the opponent's previous move.

Hence each such automaton is specified by a graph with two nodes  $C$  and  $D$  (the states of the automaton) and two oriented edges issuing from each node, one labelled  $C$  and the other  $D$ , which specify the transition from the current state to the state in the next round as a reaction to the other player's  $C$  resp.  $D$ . A further arrow points to whichever of the two states is the initial state. (For instance, this is  $C$  for *Tit For Tat* and  $D$  for *Suspicious Tit For Tat*). Fig. 1 displays a few examples. There are  $2^4$  possibilities for the transition arrows, and two possible initial conditions, making altogether for 32 graphs. But some of these graphs describe automata with the same behaviour. For example, if an automaton with initial state  $C$  has the two edges issued from  $C$  doubling back to  $C$ , then it will always remain in the  $C$  mode and hence play *AllC*, no matter where the two edges issued from  $D$  are leading to. The state will never reach  $D$ . Thus there are only 26 different strategies implemented by such automata.

Each round leads one of the four possible outcomes  $(C, C)$ ,  $(C, D)$ ,  $(D, C)$  or  $(D, D)$ , where the first position denotes the option chosen by the player and the second that of the co-player. From the player's point of view, these outcomes

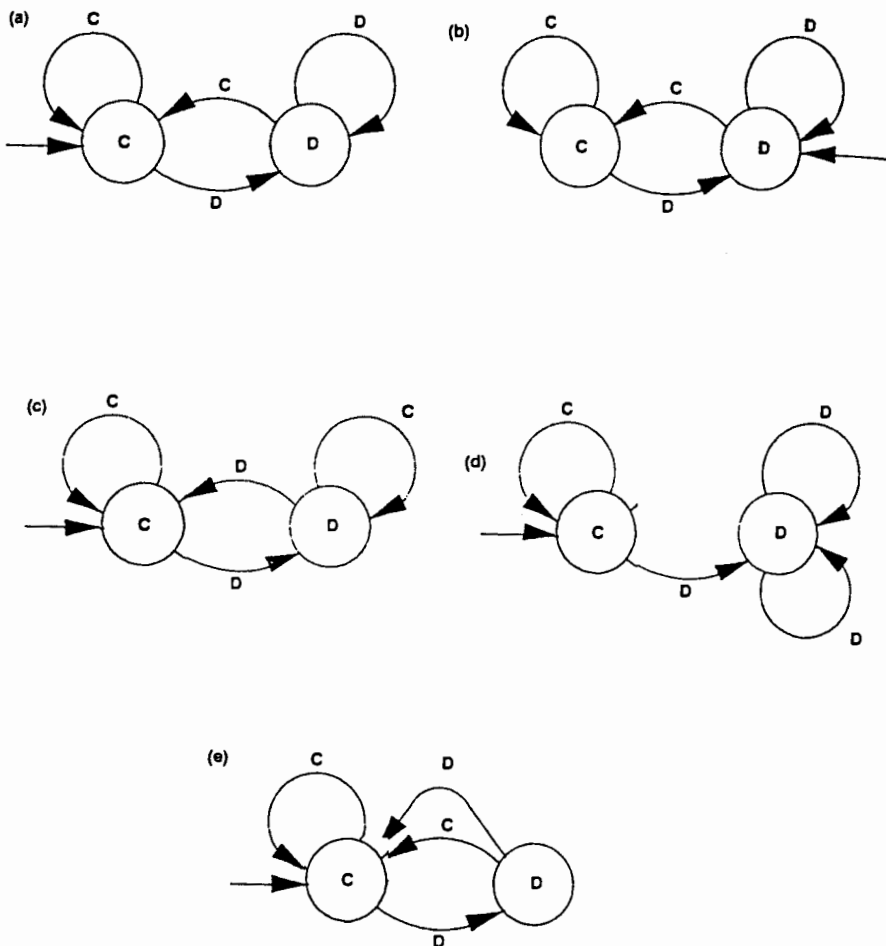


Fig. 1. Some automata: **a** tit for tat; **b** suspicious tit for tat; **c** Pavlov; **d** Grim; **e** Tweedledee (in the terminology of Binmore and Samuelson, 1992). The corresponding transition rules are  $S_{10}$  (for **a** and **b**),  $S_9$  for Pavlov,  $S_8$  for Grim and  $S_{11}$  for Tweedledee

are specified by his payoff  $R, S, T$  or  $P$ , which can be numbered by  $i = 1, 2, 3, 4$ . The 16 possible transition rules can be labelled by quadruples  $(u_1, u_2, u_3, u_4)$  of zeros and ones. Here,  $u_i$  is 1 if the automaton plays  $C$  and 0 if it plays  $D$  after outcome  $i$  ( $i = 1, 2, 3, 4$ ). For instance,  $(1, 0, 1, 0)$  is the transition rule for a tit for tat-player,  $(0, 0, 0, 0)$  that for *AllD*,  $(1, 0, 0, 1)$  that for Pavlov, etc. . . . For convenience, we label these rules by  $S_j$ , where  $j$  ranges from 0 to 15 and is the integer given, in binary notation, by  $u_1u_2u_3u_4$ . Thus  $S_0$  is *AllD*,  $S_9$  is Pavlov and  $S_{10}$  is tit for tat.

How one rule fares against another depends, of course, on the initial condition. Let us consider, for instance, an automaton with rule  $S_8 = (1, 0, 0, 0)$  (a retaliator who never relents after a defection) against an

automaton with rule  $S_{11} = (1, 0, 1, 1)$  (which is more tolerant than tit for tat in the sense that it forgives an opponent's defection if it was matched by a defection of its own).

(a) If both automata start with a  $C$ , they keep playing  $C$  forever. The sequence is:

$$S_8: CCCCCC\dots$$

$$S_{11}: CCCCCC\dots$$

(b) If both automata start with a  $D$ , the  $S_8$ -automaton will keep playing  $D$  forever, while the  $S_{11}$ -automaton will alternate between  $D$  and  $C$ . The sequence looks as follows:

$$S_8: DDDDDDDDDDD\dots$$

$$S_{11}: DCDCDCDCDCD\dots$$

(c) if  $S_8$  starts with  $C$  and  $S_{11}$  with  $D$ , we obtain

$$S_8: CDDDDDDDDDDDD\dots$$

$$S_{11}: DDCDCDCDCDCD\dots$$

(d) if, finally,  $S_8$  starts with  $D$  and  $S_{11}$  with  $C$ , the resulting sequence is

$$S_8: DDDDDDDDD\dots$$

$$S_{11}: CDCDCDCD\dots$$

The payoff in the infinitely repeated game is simply the average payoff per round. In our example, for the player using the transition rule  $S_8$  the payoff is  $\frac{1}{2}(T + P)$  in cases (b), (c) and (d) and  $R$  in case (a).

Let us now introduce a small amount of random noise. More precisely, we shall assume that the *implementation* of a move is subject to errors. (The possibility that the *perception* of a move is erroneous will be dealt with in Sect. 5.) This means that with a small probability  $\epsilon$ , one state is replaced by another. Such events happen only rarely: the average length of a run of unperturbed rounds is  $1/\epsilon$ . The corresponding transition rule is given again by a quadruple like  $S_i$ , but with 0 replaced by  $\epsilon$  and 1 replaced by  $1 - \epsilon$ . These numbers are the probabilities to play  $C$  after  $R, S, T$  or  $P$  in the previous round.

More generally, we can consider stochastic automata with transition rules given by quadruples  $p = (p_1, p_2, p_3, p_4)$  where  $p_i$  is any number between 0 and 1 denoting the probability to play  $C$  after the corresponding outcome of the previous round. The space of all such rules is the four-dimensional unit cube; the corners are just the transition rules  $S_i$ .

If a rule  $p$  is matched against a rule  $q = (q_1, q_2, q_3, q_4)$ , this yields a Markov process where the transitions between the four possible states  $R, S,$

$T$  and  $P$  are given by the matrix

$$\begin{pmatrix} p_1 q_1 & p_1(1 - q_1) & (1 - p_1)q_1 & (1 - p_1)(1 - q_1) \\ p_2 q_3 & p_2(1 - q_3) & (1 - p_2)q_3 & (1 - p_2)(1 - q_3) \\ p_3 q_2 & p_3(1 - q_2) & (1 - p_3)q_2 & (1 - p_3)(1 - q_2) \\ p_4 q_4 & p_4(1 - q_4) & (1 - p_4)q_4 & (1 - p_4)(1 - q_4) \end{pmatrix} \quad (2)$$

(note that  $p_2$  is matched with  $q_3$  and vice versa; one player's  $S$  is the other player's  $T$ ). If  $p$  and  $q$  are in the interior of the strategy cube, then all entries of this stochastic matrix are strictly positive, and hence there exists a unique stationary distribution  $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$  such that  $p_i^{(n)}$ , the probability to be in state  $i$  in the  $n$ -th round, converges to  $\pi_i$  for  $n \rightarrow \infty$  ( $i = 1, 2, 3, 4$ ). The components  $\pi_i$  are strictly positive and sum up to 1. They denote the asymptotic frequencies of  $R, S, T$  and  $P$ . The stochastic vector  $\pi$  is a left eigenvector of the matrix given by (2) for the eigenvalue 1.

It follows that the payoff for a player using  $p$  against an opponent using  $q$  is given by

$$A(p, q) = R\pi_1 + S\pi_2 + T\pi_3 + P\pi_4. \quad (3)$$

We note that the  $\pi_i$  and hence also the payoff are independent of the initial condition, i.e. of the moves of the players in the first round.

For any noise level  $\varepsilon > 0$ , we can therefore compute the payoff obtained by an automaton using transition rule  $S_i$  against an automaton with transition rule  $S_j$ . The initial states of the two automata are irrelevant.

However, if we want to compute the limit value of this payoff for  $\varepsilon \rightarrow 0$ , we cannot simply plug the limiting values into the transition matrix given by (2). If the  $p_i$  are zeros or ones, this stochastic matrix contains many zeros, and is no longer irreducible. Therefore, the stationary distribution  $\pi$  is no longer uniquely defined.

We can instead use a more direct approach. Let us exemplify it for  $S_8$  against  $S_{11}$ . We note that the four possible initial conditions lead (in unperturbed runs) to two possible stationary states  $A$  and  $B$ , where  $A$  denotes the run where both players use  $C$ , while  $B$  is the run where the  $S_8$ -player always plays  $D$  and the  $S_{11}$ -player alternates between  $C$  and  $D$ . Suppose we are in regime  $A$ . A rare perturbation causes one of the two players to play  $D$ : what follows is either scenario (c) or (d), and hence leads after a few steps to regime  $B$ . Suppose now that a perturbation occurs in regime  $B$ . With probability  $\frac{1}{2}$ , it causes the  $S_8$ -player to switch from  $D$  to  $C$ . If this happens while the  $S_{11}$ -player plays  $C$ , we are in scenario (a) and hence back to regime  $A$ ; otherwords, we enter scenario (c), with the result that regime  $B$  is resumed. Suppose now that the perturbation affects the  $S_{11}$ -player. He either plays  $C$  instead of  $D$ , which leads to scenario (c), or else he plays  $D$  instead of  $C$ , which leads to scenario (b): in both cases, the regime  $B$  is reassumed after a few steps.

Thus a perturbation of  $B$  leads with probability  $\frac{1}{4}$  to  $A$ , while a perturbation of  $A$  leads always to  $B$ . The corresponding transition matrix is

$$\begin{pmatrix} 0 & 1 \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}. \quad (4)$$

The corresponding stationary distribution is  $(\frac{1}{5}, \frac{4}{5})$ . This means that asymptotically, an iterated game between an  $S_8$ -player and an  $S_{11}$ -player is in one-fifth of the time in regime  $A$  and in four-fifths of the time in regime  $B$ . The  $S_8$ -player receives an average payoff of

$$\frac{1}{5}R + \frac{4}{5}\left(\frac{T+P}{2}\right)$$

per round. The few rounds of transitions between the two stationary regimes  $A$  and  $B$  occur with a frequency proportional to  $\epsilon$ , and hence do not affect the average payoff in the limiting case.

This argument, repeated for each of the 256 entries, yields the  $16 \times 16$  payoff matrix which is shown in Table 1. More precisely, each entry is a 4-tuple  $(s_1, s_2, s_3, s_4)$ , where the  $s_i$  are the average frequencies of  $R, S, T$  resp.  $P$ . The corresponding payoff is

$$s_1R + s_2S + s_3T + s_4P.$$

A few simple remarks: (1) the entries in the corners of the  $16 \times 16$ -matrix are just the entries of the payoff matrix for the simple  $2 \times 2$ -game; (2) the transpose of the entry  $(s_1, s_2, s_3, s_4)$  is the entry  $(s_1, s_3, s_2, s_4)$ : the quadruples  $(s_1, s_2, s_3, s_4)$  are of the form  $(1, 0, 0, 0)$ ,  $(\frac{2}{3}, \frac{1}{3}, 0, 0)$ ,  $(\frac{1}{2}, \frac{1}{2}, 0, 0)$ ,  $(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}, 0)$ ,  $(\frac{2}{3}, \frac{2}{3}, \frac{1}{3}, 0)$ ,  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0)$ ,  $(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}, 0)$ ,  $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ , or permutations thereof.

Another, less conceptual but more direct approach is used in Esam (1994). Let  $S_i(\epsilon)$  be the strategy obtained from  $S_i$  by replacing 0 by  $\epsilon$  and 1 by  $1 - \epsilon$ . For instance,  $S_{10}(\epsilon)$  is given by  $(1 - \epsilon, \epsilon, 1 - \epsilon, \epsilon)$ . For given  $i$  and  $j$ , let  $P(\epsilon)$  be the transition matrix (2), with  $p = S_i$  and  $q = S_j$ . We may write

$$P(\epsilon) = P + \epsilon Q_1 + \epsilon^2 Q_2,$$

where  $P$  is a stochastic matrix and  $Q_1$  and  $Q_2$  have row sum 0. We may view  $P(\epsilon)$  as a perturbation of the matrix  $P$  and treat the problem of finding the left eigenvector  $s(\epsilon)$  of  $P(\epsilon)$  as a perturbation problem. Thus we set

$$s(\epsilon) = \pi + \epsilon x + \epsilon^2 y + \dots$$

where the stochastic vector  $\pi$  is a solution of the unperturbed problem  $\pi P = \pi$ , whereas the components of the vectors  $x$  and  $y$  must sum up to 0. By expanding  $s(\epsilon)P(\epsilon) = s(\epsilon)$ , and comparing powers of  $\epsilon$ , this yields not only the limiting value  $\pi$  for the payoff (if  $\epsilon$  converges to 0), but also the first order error term  $x$ . (We note that we have, in general, to compare terms up to  $\epsilon^2$  in order to determine  $\pi$  and  $x$ ).

Which strategy is best for the whole population? We shall henceforth consider only the generic situation where payoff values are pairwise distinct and assume that  $R > P$  (this is no restriction of generality, since we can exchange  $C$  and  $D$ ). We have to distinguish two cases: (a) If  $2R > T + S$ , the maximal payoff attainable in a homogeneous population is  $R$ , which is attained if all members play either  $S_9$  or  $S_{14}$  or  $S_{15}$ . (b) If  $2R < T + S$ , the

Table 1. The  $16 \times 16$  payoff matrix. The payoff for strategy  $S_i$  against strategy  $S_j$  is  $s_1R + s_2S + s_3T + s_4R$ , with  $s_i = n_i(n_1 + n_2 + n_3 + n_4)^{-1}$  and  $(n_1, n_2, n_3, n_4)$  is given by the element in the  $i$ -th row and  $j$ -th column of this matrix. The strategy  $S_i$  is specified by the quadruple  $u_1u_2u_3u_4$  which yields the number  $i$  in binary notation ( $i = 0, 1, \dots, 15$ ). For instance,  $S_9$  has as quadruple 1001 and  $S_{13}$  as quadruple 1101. The payoff for  $S_9$  against  $S_{13}$  is  $(2R + T)/3$

	$S_0$	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$	$S_6$	$S_7$	$S_8$	$S_9$	$S_{10}$	$S_{11}$	$S_{12}$	$S_{13}$	$S_{14}$	$S_{15}$
$S_0$	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
$S_1$	0001	0011	0001	0011	0012	0010	0011	0010	0001	0011	0001	0011	0011	0010	0021	0010
$S_2$	0010	0001	0111	1001	0212	1011	0010	1021	0101	1011	0111	1011	0121	0010	0010	0010
$S_3$	0011	0001	0110	0110	0001	1011	0001	1011	0001	0111	0111	0110	1012	1010	2021	1010
$S_4$	0100	1001	0101	1111	0101	1001	1111	1001	0101	1111	0110	0110	1111	1010	1010	1010
$S_5$	0101	0122	0001	0011	0112	0010	0012	0010	0102	0122	0001	0011	0132	0010	0021	0010
$S_6$	0100	1101	1101	1001	0100	1111	1111	1011	0100	1111	1111	1011	0110	0010	0010	0010
$S_7$	0101	0100	0001	1111	0102	1111	0001	1011	0201	0100	1111	1110	1111	2120	2021	1010
$S_8$	0100	1201	1101	1001	0100	1101	1101	1001	0100	0100	1110	1110	1210	2120	1010	1010
$S_9$	0001	0011	0001	0011	0012	0010	0021	0010	0001	1022	0001	1022	1023	1020	1021	1020
$S_{10}$	0101	1101	0111	1111	0212	1111	0010	0010	1202	1000	1111	1000	1111	2010	1020	1010
$S_{11}$	0001	0111	0111	0110	0001	1111	1111	1110	0001	1111	1111	1110	1001	1000	1000	1000
$S_{12}$	0101	1101	0110	0110	0101	1101	1110	1110	1202	1000	1110	2110	2101	1000	1000	1000
$S_{13}$	0101	0211	1102	1111	0312	0110	1111	1120	1203	1111	1001	2011	1111	2130	3021	1010
$S_{14}$	0100	0100	1100	1100	0100	0100	2210	2210	1200	2100	1000	1000	2310	2110	2010	2010
$S_{15}$	0201	0100	2201	1100	0201	0100	2201	1100	1201	1200	1000	1000	3201	2100	1000	1000
	0100	0100	1100	1100	0100	0100	1100	1100	1200	1100	1000	1000	1100	2100	1000	1000



maximal payoff in a homogeneous population is  $\frac{1}{4}(2R + S + T)$ , and it is attained for  $S_{11}$  and  $S_{13}$ .

However, in a biological context, we cannot assume that the agents keep the general welfare in mind. Those members of the population who are able to increase their individual payoff will spread, even if this lowers the average payoff of their group. In order to investigate such an evolution, we have to use the methods of game dynamics.

### 3 Game dynamics for the Prisoner's Dilemma

Let us assume that a population consists of  $n$  different types  $E_1$  to  $E_n$ , and that  $x_i(t)$  is the frequency of type  $E_i$  at time  $t$ , so that the state of population at time  $t$  is given by a vector  $x(t)$  in the unit simplex in  $R^n$ . If  $A$  is the  $n \times n$ -matrix whose element  $a_{ij}$  is the average increase in fitness for an individual of type  $E_i$  encountering an individual of type  $E_j$ , and if individuals meet randomly, then the average increase in fitness for an individual of type  $E_i$  within the population is  $\sum a_{ij}x_j = (Ax)_i$ , and the average increase of fitness within the population is  $\sum x_i(Ax)_i = xAx$ . We interpret each type as a strategy and its payoff as increase in fitness, i.e. as reproductive success. According to the usual replicator dynamics (cf. for instance Hofbauer and Sigmund 1988), we shall assume that the frequencies  $x_i$  evolve according to the ordinary differential equation

$$\dot{x}_i = x_i[(Ax)_i - xAx] . \quad (5)$$

This equation describes the action of selection upon the frequencies of the competing strategies. We note in particular that if some strategies are missing in the population, they are not introduced at a later time. Thus our model is closed, in the sense that it does not admit the emergence of missing types through mutation, migration or recombination. In particular, the corners of the state simplex, i.e. the vectors  $e_i$  of the standard basis, are equilibrium points. They correspond to pure states consisting of type  $E_i$  only.

We shall apply this to the two-state automata playing iterated  $2 \times 2$  games under the effect of a small noise term. Thus the different types in the population correspond to the pure strategies  $S_i$  and the payoff matrix  $A$  is given by Table 1. Needless to say, a complete analysis of the 15-dimensional dynamical system is a hopeless task. We shall consider special cases only, and start with the Prisoner's Dilemma.

#### (a) Competition of two strategies in the Prisoner's Dilemma

Among  $2 \times 2$ -games, the Prisoner's Dilemma is characterised by the inequalities

$$T > R > P > S \quad \text{and} \quad 2R > T + S . \quad (6)$$

**Table 2.** A list of the strategies outcompeting  $S_i$ . Column (a) deals with the general case of the Prisoner's Dilemma, as specified by inequalities (6) in the text for instance,  $S_0$  and  $S_8$  always outcompete  $S_1$  whenever the payoff values  $R, S, T, P$  satisfy (6). Column (b) lists the strategies which outcompete  $S_i$  (in addition to those from the previous column) for the Prisoner's Dilemma under the supplementary condition  $2P < S + T$ . Column (c) lists the strategies which, in addition to the previous ones, outcompete  $S_i$  in the special case  $T = 5, R = 3, P = 1, S = 0$ . Column (d) lists the strategies outcompeting  $S_i$  for the Chicken game with  $T = 1, R = 0, S = -1, P = -10$ . Note the very good performance of the balanced 'win-stay, lose-shift' strategy  $S_9$  in all these cases

	(a)	(b)	(c)	(d)
$S_0$	-	2, 10	-	2, 3, 9, 10, 11, 12, 14
$S_1$	0, 8	-	3, 4, 10	3, 5, 9, 10, 11, 12
$S_2$	-	9, 10, 11	1	5, 9, 10, 11, 14
$S_3$	0, 4, 8, 9	-	11	9, 11
$S_4$	0, 8	-	-	9, 12, 14
$S_5$	0, 1, 4, 8, 9	-	2	9
$S_6$	0, 1, 4, 8, 9	2, 3, 5, 10, 11, 12	-	2, 3, 4, 5, 9, 10, 11, 12, 14
$S_7$	0, 1, 4, 5, 8, 9	-	2, 11, 12	3, 5, 9, 11
$S_8$	-	2, 10	-	2, 3, 4, 9, 10, 11, 12, 14
$S_9$	-	-	0, 1	-
$S_{10}$	9, 14, 15	11	-	9, 11, 13, 14, 15
$S_{11}$	9, 12, 14, 15	-	0, 1, 4, 5, 8	-
$S_{12}$	0, 1, 4, 9	-	2, 8	9, 14
$S_{13}$	0, 1, 3, 4, 5, 9, 12	-	2, 8	9
$S_{14}$	1, 3, 5, 7, 9, 13	-	0, 2, 4, 8	9, 13
$S_{15}$	0, 1, 2, 3, 4, 5, 7, 9, 12, 13	-	8	9, 13

Let us first remark that for a large set of payoff values, no pure strategy is evolutionarily stable: every pure strategy can be invaded, and even outcompeted by another pure strategy. ( $S_i$  can be invaded by  $S_j$  if the equilibrium point  $e_i$  is unstable in the one-dimensional subsystem of (5) obtained by setting  $x_k = 0$  for all  $k \neq i, j$ , i.e. by its restriction to the edge  $e_i e_j$  of the state simplex; this occurs exactly if  $a_{ji} \geq a_{ii}$  and, in case equality holds, if furthermore  $a_{jj} > a_{ij}$ . Strategy  $S_i$  is outcompeted by  $S_j$  if both  $a_{ji} \geq a_{ii}$  and  $a_{jj} \geq a_{ij}$ , with at least one inequality being strict. In this case, even the smallest  $S_j$ -minority introduced into an  $S_i$ -population will eventually drive  $S_i$  to extinction.)

We see from Table 2 that for any choice of payoff values for the Prisoner's Dilemma, all pure strategies except  $S_0, S_2, S_3$ , and  $S_9$  are outcompeted by at least one other strategy. If we assume, in addition to (6), that  $2P < T + S$ , then  $S_0, S_2$  and  $S_8$  will also be always outcompeted, whereas  $S_9$  will be outcompeted (by  $S_0$  and  $S_1$ ) if  $2R < T + P$ . We note that  $S_9$  is particularly good at getting established, in the sense that under general conditions, it outcompetes the greatest number of rival strategies (even more than the *AllD*-strategy  $S_0$ ), and that is as resistant as  $S_0$  and the grim strategy  $S_8$  against being outcompeted. In the next section, we shall discuss this remarkable strategy in more detail. For the moment, we only note that its mirror-image  $S_6$  is the least able at invading or outcompeting a homogeneous population.

We stress the very mediocre showing of the tit for tat strategy  $S_{10}$  in this kind of 'head-on'-confrontations with one single competing strategy. The advantage of tit for tat appears mainly in heterogeneous populations, as noted by Axelrod (1984) already.

Furthermore, the edges  $e_0e_8$ ,  $e_3e_{12}$ ,  $e_5e_{10}$  and  $e_{14}e_{15}$  consist of equilibrium points. It follows that  $S_0$  and  $S_8$  are never evolutionarily stable.  $S_9$  is uninvadable iff  $2R > T + P$ .

(b) *Axelrod's payoff values*

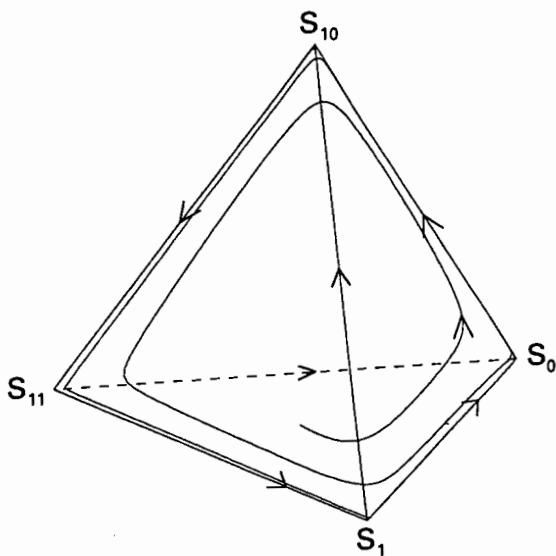
It is very difficult to obtain general results on the higher dimensional game dynamics for the iterated Prisoner's Dilemma, because there are so many different cases to distinguish. We shall only consider one specific example, using the payoff values mentioned in the introduction, which were used by Axelrod in his celebrated computer tournaments:  $T = 5$ ,  $R = 3$ ,  $P = 1$  and  $S = 0$ . (Then  $2P < S + T$  and  $2R = T + P$ .) In this case, no pure strategy is invasion proof, as seen from column (c) of Table 2.

There exist in this case 28 heteroclinic three-cycles: these are triples of strategies  $S_i, S_j$  and  $S_k$  where, like in the Rock-Scissors-Paper game,  $S_i$  is outcompeted by  $S_j$ ,  $S_j$  by  $S_k$  and  $S_k$ , in turn, by  $S_i$  again. The corresponding restriction of the game dynamics (5) to the two-dimensional boundary face of the state simplex which is spanned by the unit vectors  $e_j, e_j$  and  $e_k$  is well understood (see e.g. Hofbauer and Sigmund 1988). It is easy to check that the heteroclinic cycles  $S_0S_{10}S_{11}$ ,  $S_0S_2S_{11}$ ,  $S_0S_2S_9$ ,  $S_0S_{10}S_9$ ,  $S_1S_{10}S_9$ ,  $S_1S_{10}S_{11}$ ,  $S_1S_3S_9$ ,  $S_1S_3S_{11}$  and  $S_2S_{11}S_{12}$  are attractors (within the corresponding face); the cycles  $S_2S_{11}S_5$ ,  $S_7S_{11}S_{14}$ ,  $S_7S_{11}S_{15}$ ,  $S_2S_{11}S_{14}$ ,  $S_2S_{11}S_{15}$ ,  $S_3S_{11}S_{14}$ ,  $S_3S_{11}S_{15}$ ,  $S_2S_{10}S_{14}$  and  $S_2S_{10}S_{15}$  are repellers; the heteroclinic cycles  $S_0S_{10}S_{14}$ ,  $S_0S_2S_1$ ,  $S_1S_{10}S_{14}$ ,  $S_1S_{10}S_{15}$ ,  $S_0S_{10}S_{15}$ ,  $S_1S_8S_2$  are filled with periodic orbits and the cycles  $S_8S_{11}S_2$ ,  $S_8S_{10}S_{11}$ ,  $S_8S_{10}S_{14}$  and  $S_8S_{10}S_{15}$  with homoclinic orbits whose  $\alpha$ - and  $\omega$ -limit is  $e_8$ .

There are many other heteroclinic cycles on the boundary of the state space, involving more than three saddle points, for instance, or saddles which are mixed equilibria. We do not attempt to list them all, but emphasise that they constitute more than a marginal feature of the dynamics.

Indeed, a straightforward computation (best left to a computer) shows that there exists no completely mixed equilibrium in the full 16-strategies system (no equilibrium with all 16 strategies present). It follows (cf. Hofbauer-Sigmund 1988) that every trajectory starting from a completely mixed initial state (i.e. with all  $x_i(0) > 0$ ) converges to the boundary. The behaviour on the boundary seems to be extremely complex, and difficult to attack even by numerical simulations, because some of the  $x_i(t)$  have 0 as  $\liminf$  (for  $t \rightarrow \infty$ ) without however converging to 0.

If this is remedied by adding a small unspecific term to the game dynamical equation, then the numerical simulations exhibit violent oscillations which can be regular or chaotic (see Nowak and Sigmund 1993a, where the



**Fig. 2.** The phase portrait for the subsystem consisting of the strategies  $S_1, S_0, S_{10}$ , and  $S_{11}$ , for the case of the Prisoner's Dilemma with parameter values  $T = 5, R = 3, P = 1, S = 0$ . All orbits in the interior of the four-simplex converge to the attractor consisting of the heteroclinic cycle  $e_1 \rightarrow e_0 \rightarrow e_{10} \rightarrow e_{11} \rightarrow e_1$ . There are two further heteroclinic cycles on the boundary, namely  $e_1 \rightarrow e_{10} \rightarrow e_{11} \rightarrow e_1$  and  $e_0 \rightarrow e_{10} \rightarrow e_{11} \rightarrow e_0$ , each attracting the orbits on the corresponding three-face

numerical simulations were performed for the corresponding difference equation).

In these numerical runs, the strategies  $S_1, S_0, S_{10}, S_{11}$  have the highest average frequencies. It is of interest, therefore, to investigate the evolution of the corresponding subsystem. Again, each of these strategies is dominated by another: in fact, there are no equilibria on the edges, and only two three-strategies equilibria, one on the face  $S_0S_{10}S_{11}$  and the other on the face  $S_1S_{10}S_{11}$ . On both faces, the orbits spiral outward toward the heteroclinic orbit formed by the edges. Again, there is no interior equilibrium in the four-strategy system, and the orbits have to converge to the boundary. The numerical simulations suggest that the heteroclinic cycle formed by the four corners (in the order given above) and the connecting edges is an attractor in this case. (See Fig. 2.) It follows that in the subsystem consisting of these four strategies, each one oscillates wildly: its frequencies have 0 as  $\liminf$  and 1 as  $\limsup$  (as  $t \rightarrow \infty$ ). A random perturbation will eventually cause one of the strategies to reach fixation (the three others will be wiped out), but it is impossible to predict which will be the winner.

It is instructive to follow the asymptotic evolution of this four-dimensional system. The strategy  $S_0$ , which always defects, is outcompeted by the tit for tat strategy  $S_{10}$ . This is a rather weak effect: against  $S_0$ , both  $S_0$  and  $S_{10}$  do equally well, but against  $S_{10}$ ,  $S_{10}$  does better than  $S_0$ . In an  $S_{10}$ -population,

however, mistakes cause long series of retaliation, which lower the average payoff; therefore, the more tolerant strategy  $S_{11}$ , which only defects after having been played for a 'sucker' in the previous round, can invade and take over. This very cooperative strategy can be invaded by the all-defect strategy  $S_0$ , but better still by the parasitic bully  $S_1$ , which cooperates only after its defection has met with defection.  $S_1$  actually takes over for a while, and almost reaches fixation. It can be outcompeted by tit for tat, but better still by the relentlessly defecting  $S_0$ , so that a hegemony of  $S_1$  is followed first by a hegemony of  $S_0$ , while tit for tat takes over much later.

It would be interesting to obtain an analytic proof that in the  $S_0S_1S_{10}S_{11}$ -system, the four-cycle is an attractor. Brannath (1994) has developed criteria which unfortunately fail to cover this precise situation.

### (c) *A Chicken game*

The Chicken game is characterized by the following ranking of the payoff values:

$$T > R > S > P \quad (7)$$

As 'Hawk-Dove'-game, it has been used to explain the prevalence of conventional fighting in innerspecific animal conflicts (Maynard Smith 1982). In this case (C) means to adopt a conventional way to settle a conflict (by display, outstaring the adversary, or engaging in a harmless pushing match) while (D) means to escalate the conflict until the adversary flees or one of the contestants is disabled. In a one-round 'Chicken' game, each pure strategy can be invaded by the other: game dynamics leads to a mixed population in an evolutionary equilibrium. Since contests within animal communities are likely to be frequent, it is plausible to assume that animals often engage in repeated Chicken games, at least as long as the risk of a lasting injury is small.

Rather than considering the general case, we shall restrict our attention to an example, and assume that  $R = 0$ ,  $T = 1$ ,  $P = -10$  and  $S = -1$ . (In this case, the evolutionarily stable equilibrium strategy in the one-round game consists in escalating with a probability of  $\frac{1}{10}$ .) Once more, the corresponding system (6) admits no equilibrium in the interior of the state simplex; hence all orbits have to converge to the boundary.

Again, we see that  $S_9$  does a very good job in any confrontation with one single other strategy. It outcompetes every other strategy, with the exception of  $S_{11}$ . (The edge  $e_{11}e_9$  consists of equilibria.) A Pavlov-population playing the Chicken game usually sticks to conventional display. A one-sided escalation due to some mistake leads to one round of bilateral escalation, after which conventional display is resumed. Again, the strategy  $S_6$  which is the mirror-image of Pavlov does worst in outcompeting other strategies. In contrast to the Prisoner's Dilemma case, there are now many situations where  $S_i$  invades  $S_j$  and  $S_j$  invades  $S_i$ , leading to stable equilibria on the corresponding edges.

#### 4 The 'win-stay, lose-shift' strategies

In Nowak and Sigmund (1993b), we performed extensive game dynamical simulations of the iterated Prisoner's Dilemma in the space of *stochastic* strategies  $(p_1, p_2, p_3, p_4)$  with memory one. In these evolutionary runs, we occasionally introduced small amounts of new strategies, and eliminated those strategies whose frequencies dropped below a certain threshold. Under such a mutation-selection regime, a huge variety of evolutionary chronicles can be observed. The most interesting results were:

(a) a clear display of punctuated equilibrium for the average payoff in the population. For very long periods of stasis, this payoff was confined to a value close to  $P$  (a regime of defection) or to  $R$  (over-all cooperation);

(b) the runs displayed a pronounced tendency towards cooperation;

(c) among the cooperative populations, there was a strong trend to the establishment of  $S_9$ . It outperformed not only tit for tat, but even generous tit for tat, a strategy which tolerates the co-player's defection with a certain probability (see Nowak and Sigmund 1992).

This highlights the interesting properties of the strategy  $S_9$ , which has transition rule  $(1, 0, 0, 1)$ . This strategy was investigated at an early stage in Rapoport and Chamah (1965), who called it a 'simpleton'. It does poorly against *AllD*, since it switches endlessly between  $C$  and  $D$ , and hence is exploited in every second round. Empirical studies on humans have confirmed that  $S_9$  is not particularly suitable to induce cooperation (see Rapoport and Chamah, 1965). In a population of cooperators, however, it fares very well, because it is to a large extent immune against errors, in contrast to *TFT*. If two  $S_9$ -players interact, they cooperate almost always. An accidental mistake leading to a unilateral defection is followed by one round where both players defect, and then by a bilateral resumption of cooperation. Furthermore, an  $S_9$ -player has no qualms about exploiting an *AllC* player. Thereby, the frequency of *AllC* is kept low, which leaves little opportunity for exploiters (whereas in a homogeneous tit for tat-population, *AllC* can spread by neutral drift, which paves the way for an invasion by defectors, as stressed in Axelrod 1984).

In the context of the Prisoner's Dilemma, the strategy  $S_9$  embodies a very simple behavioural principle: win-stay, lose-shift. Indeed, an  $S_9$ -player who experienced a  $T$  or an  $R$  in the previous round will use the same move again; but after a  $P$  or an  $S$ , this player will try the alternative move. It is as if there were a numerical threshold between  $R$  and  $P$ . If a round yields more than this threshold value, it acts as a reinforcer and induces a repetition of the successful move in the following round; if a round yields less, this is experienced as a punishment and cues a shift in behaviour. Because of this almost reflex-like reaction, the strategy  $S_9$  has been christened Pavlov (cf. Kraines and Kraines 1988).

Pavlov's relatives, so to speak, are those 'win-stay, lose-shift' strategies  $S_i$  which act on a similar principle, but with a different threshold. Let us first stick to the Prisoner's Dilemma case. If the threshold is between  $T$  and  $R$ , we

**Table 3.** The family of win-stay, lost-shift strategies. The 12 rows describe the 12 different payoff orderings (Chicken is in the fourth row, the Prisoner's Dilemma occurs in the sixth). They are followed by the corresponding win-stay, lose-shift strategies, ordered according to decreasing aspiration level: (a) the strategy  $S_3$  which shifts behaviour in every round, (b) the ambitious strategy which shifts except if it has achieved the highest payoff, (c) the balanced strategy which shifts only if it has obtained one of the two lowest possible payoff values, (d) the modest strategy which shifts only for the lowest possible payoff, and finally (e) the strategy  $S_{12}$  which sticks to whatever it did in the previous round

$S > T > R > P$	3	7	5	13	12
$T > S > R > P$	3	1	5	13	12
$S > R > T > P$	3	7	15	13	12
$T > R > S > P$	3	1	9	13	12
$S > R > P > T$	3	7	15	14	12
$T > R > P > S$	3	1	9	8	12
$R > S > T > P$	3	11	15	13	12
$R > T > S > P$	3	11	9	13	12
$R > S > P > T$	3	11	15	14	12
$R > T > P > S$	3	11	9	8	12
$R > P > S > T$	3	11	10	14	12
$R > P > T > S$	3	11	10	8	12

obtain the strategy  $S_1$  with transition rule  $(0, 0, 0, 1)$ : this bully-like rule defects if it can exploit a sucker, but switches as soon as it meets a defection. If the threshold lies between  $P$  and  $S$ , we get  $S_8$  with transition rule  $(1, 0, 0, 0)$ . This is a grim strategy which never forgives a defection (and only forgets it by mistake). Finally, there is the strategy  $S_3$  with rule  $(0, 0, 1, 1)$  which always shifts its behaviour from one round to the next, and the strategy  $S_{12}$  with rule  $(1, 1, 0, 0)$  which stubbornly repeats its previous move. These last two examples correspond to the extreme cases of players who are never happy (resp. those who are always content) with the outcome of the previous round.

Obviously, the 'win-stay, lose-shift'-strategies depend on the rank order of the payoff values of the simple  $2 \times 2$ -game. If we assume, as before, that the four values  $R, S, T, P$  are pairwise distinct (which is generically the case) and that  $R > P$  (which is no restriction of generality), then we get 12 different rank orderings. In Table 3, we list these cases, as well as the three 'win-stay, lose-shift'-strategies corresponding to an 'aspiration level' which is (1) *ambitious*, i.e. content only with the highest possible payoff, (2) *balanced*, i.e. content with the two highest payoff values, and (3) *modest*, i.e. content with all but the lowest payoff. The extreme strategies are, of course, always  $S_3$  and  $S_{12}$ , independently of the rank ordering of payoff values.

From Table 3, one can draw a few immediate conclusions. (a) The strategies  $S_0, S_2, S_4$  and  $S_6$  do never occur as 'win-stay, lose-shift'-strategies; (b) the ambitious strategies are  $S_1, S_7$  and  $S_{11}$ , the balanced strategies  $S_5, S_9, S_{10}$  and  $S_{15}$ , the modest strategies  $S_8, S_{13}$  and  $S_{14}$ .

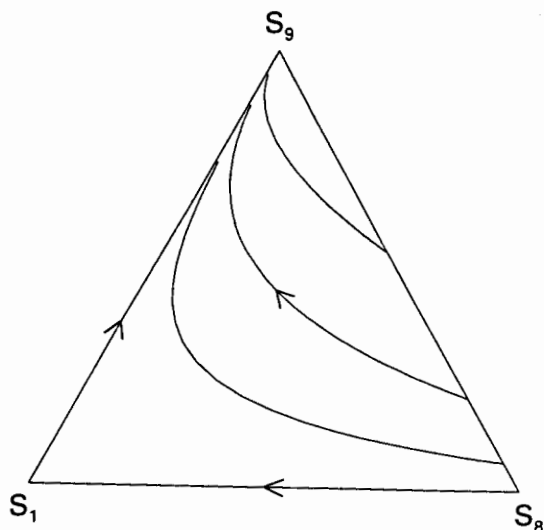


Fig. 3. The phase portrait for the subsystem consisting of the three 'win-stay, lose-shift' strategies  $S_1$ ,  $S_9$  and  $S_8$  for the Prisoner's Dilemma  $T = 5$ ,  $R = 3$ ,  $P = 1$ ,  $S = 0$ . The edge  $e_8e_9$  consists of fixed points. The interior of the simplex is filled with orbits whose  $\alpha$ - and  $\omega$ -limit lies on this edge

If the whole population is ambitious, it can never realise the highest possible level of welfare (which, as we have seen at the end of Sect. 2, is  $R$  or  $\frac{1}{4}(2R + S + T)$ , depending on whether  $2R > T + S$  or not). On the other hand, the balanced strategy, or the modest, or both, can realise the highest possible average payoff in all but the three following cases:

(i) If  $R > P > T > S$ , the highest average payoff  $R$  can never be realised by a 'win-stay, lose-shift'-strategy. In this case, the two players should obviously never deviate from the strategy  $C$ ; but accidental mistakes in a game between two 'win-stay, lose-shift'-players triggers a self-defeating sequence of responses which does not converge to a situation of mutual cooperation.

(ii) If  $S > R > P > T$ , the highest average payoff can only be realised if it is  $R$ , i.e. in the case  $2R > T + S$ .

(iii) The same result holds for  $T > R > P > S$ . We note that this implies that in the Prisoner's Dilemma case, the Pavlov strategy  $S_9$  yields the best outcome if adopted by the whole population. This does not mean that evolution among 'win-stay, lose-shift'-strategies leads necessarily towards this outcome. For  $R = 4$ ,  $S = 0$ ,  $T = 8$ ,  $P = 1$ , for instance, which is not a Prisoner's Dilemma, the modest strategy  $S_8$  can outcompete an  $S_9$ -population, and lead to a society of defectors. But in the case of the Axelrod values  $R = 3$ ,  $S = 0$ ,  $T = 5$  and  $P = 1$ , the edge  $e_8e_9$  consists of equilibrium points, and the  $e_1e_8e_9$ -face is filled with orbits having their  $\alpha$ - and their  $\omega$ -limits on this edge; all orbits in the interior of the system consisting of the 5 'win-stay, lose-shift'-strategies converge to this face (see Fig. 3).



Another interesting situation occurs in the Chicken game for the parameter values  $R = 0$ ,  $P = -10$ ,  $T = 1$  and  $S = -1$ . In the  $S_9, S_1, S_{13}$ -system,  $S_9$  is evolutionarily stable, i.e. neither  $S_1$  nor  $S_{13}$  can invade by themselves; but there exists another stable mixture consisting of 20% of  $S_1$ -players and 80% of  $S_{13}$ -players. This is a bistable system: depending on the initial condition, the outcome is either this  $S_1 - S_{13}$ -mixture or a homogeneous  $S_9$ -population.

## 5 Discussion

One could tentatively divide the history of investigating the Iterated Prisoner's Dilemma into three parts, centering on the following questions: (1) How would a perfectly rational being act? (2) What do human subjects actually do? (3) Which strategies are adopted by simple agents, like animals or automata? Obviously, there is no clear-cut separation between these phases. Nevertheless, one detects a trend from superhuman via human to subhuman. In the 'fifties, game theorists looked mainly for normative prescriptions of the 'right' way to play the Prisoner's Dilemma, whereas the high water mark of psychological experiments was probably reached in the 'sixties (see Rapoport and Chammah (1965), and the bibliography by Guyer and Perkel (1972)). And in the late 'seventies, Axelrod's computer tournaments focussed attention both upon biological applications (see Axelrod and Hamilton, 1981) and upon simple 'programs' to implement strategies.

The idea to use automata for repeated games was explored in Aumann (1981) (in Rapoport and Chammah 1968, we find a forerunner – the notion of a 'simpleton' actually implementing the Pavlov strategy). Rubinstein (1986) used finite automata (each associated with a 'complexity cost' given by its number of states) and lexicographic utilities: first maximize payoff, and then minimize complexity (= number of states). Abreu and Rubinstein (1988) used this (in the case of infinitely repeated games with limit-of-the-mean payoff) to show that two machines in equilibrium must have the same number of states, which must be matched one-to-one and fall into two classes: some states are used exactly once in the beginning of the interaction, and some make up a cycle which is then endlessly repeated. This allows for a very complete analysis in the case of  $2 \times 2$ -games. (Banks and Sundaram (1990) have analysed alternative complexity measures.)

Another approach consists in studying populations of automata evolving under selection. Thus Binmore and Samuelson (1992) show that an evolutionarily viable outcome must be utilitarian, i.e. maximise the sum of the two payoffs (for otherwise, mutant automata recognising each other by some 'secret handshake' could invade). Binmore and Samuelson show that no type of automaton is evolutionarily stable, but that groups of co-existing types of automata can satisfy a modified notion of evolutionary stability, and be proof against any isolated invasion attempted by other strategies. Such outcomes are utilitarian (see also the related work by Probst 1993).

While this body of work sheds interesting light on the evolution of populations playing the iterated Prisoner's Dilemma, it concentrates on the stability of equilibrium situations and does not investigate the dynamics of the game (e.g. whether populations actually converge towards such a modified ESS, or MESS.) Furthermore, the automata are assumed to work with perfect precision. In biological applications, it seems plausible that simple decision rules govern the behaviour of individuals, but that these rules will be frequently affected by errors. In part, the analysis of repeated games played by automata relies essentially on so-called 'trigger states': but in the presence of errors, all states will always be visited. This has drastic effects on the outcome of the game (see Selten and Hammerstein 1984 and Boyd 1989). For the Prisoner's Dilemma, for instance, the MESS consists of three strategies with transition rules (1, 0, 1, 0), (1, 0, 0, 1) and (1, 0, 1, 1), which do equally well against each other (see Samuelson and Binmore 1992); but in the presence of errors, the dynamical system formed by the corresponding strategies  $S_9$ ,  $S_{10}$  and  $S_{11}$  has  $e_9$  as global attractor.

We emphasize that the errors occurring in repeated games can be of varied nature. In our paper, we have assumed that they are errors in implementing a move, somewhat along the line of 'trembling hand' concept (Selten 1975). We can also analyse the effect of errors in perception – in misunderstanding the other's  $C$  for a  $D$ , for example (see, e.g., Axelrod and Dion 1988 and Miller 1989). This type of errors can sometimes lead to quite different results. For instance, the strategy Contribute Tit For Tat (see, e.g. Boyd, 1989) is proof against errors in implementation, but not against errors in perception. This strategy, which does not belong to the class of strategies discussed in this paper, acts in principle like tit for tat, but enters a 'contribute' state if it has, by mistake, played a  $D$  rather than a  $C$ ; in this case, it plays  $C$  after an opponent's  $D$  (i.e. accepts the retaliation), but leaves the contribute state. If such a player erroneously thinks that the other player defected, it will play  $D$  without switching into the contribute state, and will not meekly accept any subsequent retaliation. In contrast to this, Pavlov is proof against both errors in implementation and errors in perception.

Let us denote by  $\varepsilon$  the probability of mistaking the other player's previous move, and by  $\lambda\varepsilon$  the probability of mistaking one's own previous move (usually,  $\lambda$  should be smaller than 1). The perturbation of the tit for tat strategy  $S_{10}$  is  $(1 - \varepsilon, \varepsilon, 1 - \varepsilon, \varepsilon)$ , just as with mistakes in implementation. The perturbation of  $S_9$  is  $(1 - (\lambda + 1)\varepsilon, (\lambda + 1)\varepsilon, (\lambda + 1)\varepsilon, 1 - (\lambda + 1)\varepsilon)$ ; that of  $S_0$  is  $(0, 0, 0, 0)$ , i.e. no perturbation at all. In general, the strategy  $(u_1, u_2, u_3, u_4)$  turns into

$$(1 - (\lambda + 1)\varepsilon)(u_1, u_2, u_3, u_4) + \varepsilon(u_2, u_1, u_4, u_3) \\ + \lambda\varepsilon(u_3, u_4, u_1, u_2) + \lambda\varepsilon^2(v, -v, -v, v),$$

where  $v = u_1 + u_4 - u_2 - u_3$ . Again, one can use the same methods as in Sect. 2 to find the payoff values. For  $S_1$  against  $S_4$ , for instance, one

obtains

$$\frac{S}{2} + \frac{P}{2} + \varepsilon \left( \frac{\lambda}{2} R - \frac{(1+3\lambda)}{4} S + \frac{\lambda}{2} T + \frac{(1-\lambda)}{4} P \right) + O(\varepsilon^2).$$

We note that the error term reduces neither for  $\lambda = 1$  nor for  $\lambda = 0$  to the corresponding perturbation term for mistakes in implementation. Similarly, one can consider the joint effect of errors in implementation and perception; allow for different propensities to mis-implement (or mis-perceive) a  $C$  or a  $D$ ; investigate repeated games where the players move alternately, rather than simultaneously (see Nowak and Sigmund, 1994), etc.

The book by Rapoport and Guyer (1976) contains a full taxonomy of all  $2 \times 2$ -games. We have only considered symmetric games in this paper. We are still very far from fully understanding the effects of selection and mutation on heterogeneous populations of two-state automata engaged in playing repeated games against each other. In a sense, by incorporating noise we simplified the analysis by the factor 4, since the initial states became irrelevant. However, the dynamical systems still have 16 variables; and as our handful of examples shows, the asymptotic behaviour displays a bewildering richness. Axelrod (1987), Miller (1989) and Lindgren (1991) studied evolutionary chronicles of populations of automata playing the iterated Prisoner's Dilemma, using genetic algorithms to generate variety. Lindgren and Miller considered situations where errors occur (in Miller's simulations, they were due to the misperception of an opponents move). A particularly ingenious aspect of Lindgren's approach is to allow for a kind of 'gene duplication' which extends the memory of the player. In many of his runs, the evolution leads towards a class of strategies with memory length 2, which defect after having been played for a sucker, but revert to cooperation after *two* consecutive rounds of mutual defections. Such strategies can be viewed as sophisticated versions of the Pavlov strategy  $S_9$  (see also Sigmund, 1995).

*Acknowledgements* We thank Christian Forst for help with the figures, and the Austrian Forschungsförderungsfonds, Project P8043, for support.

## References

- Aumann, R. (1981) Survey of repeated games, in *Essays in Game Theory and Mathematical Economy in Honor of Oskar Morgenstern*, Wissenschaftsverlag, Bibliographisches Institut, Mannheim
- Axelrod, R. (1984) *The Evolution of Cooperation*, Basic Books, New York (reprinted 1989 in Penguin, Harmondsworth)
- Axelrod, R. (1987) The evolution of strategies in the iterated Prisoner's Dilemma, in Davis, D. (ed.) *Genetic Algorithms and Simulated Annealing*, Pitman, London
- Axelrod, R. and Hamilton, W.D. (1981) The evolution of cooperation, *Science* **211**, 1390-6
- Axelrod, R. and Dion, D. (1988) The further evolution of cooperation, *Science* **242**, 1385-90
- Abreu, D. and Rubinstein, A. (1988) The structure of Nash equilibria in repeated games with finite automata, *Econometrica* **56**, 1259-82

Cell 65, 651-662

13 Marthas, M.L., Sutjipto, S., Higgins, J. *et al.* (1990)

*J. Virol.* 64, 3694-3700

14 Nowak, M.A., May, R.M. and Anderson, R.M. (1990)

*AIDS* 4, 1095-1103

15 Nowak, M.A., Anderson, R.M., McLean, A.R. *et al.*

(1991) *Science* 254, 963-969

16 Nowak, M.A. and May, R.M. (1992) *J. Theor. Biol.* 159, 329-342

17 Nowak, M.A. and May, R.M. (1993) *AIDS* 7 (Suppl. 1), S3-S18

18 Baltimore, D. (1988) *Nature* 335, 395-396

19 Bangham, C.R. and Kirkwood, T.B. (1993) *Trends Microbiol.* 1, 260-264

# Physiological enzymatic cleavage of leukocyte membrane molecules

Vladimír Bažil

*Certain membrane molecules are enzymatically cleaved from the cell surface and then released into the extracellular medium in the form of soluble fragments. This process, commonly initiated by cell stimulation, may regulate the surface expression of such molecules, and may also be responsible for the production of their soluble forms in vivo. Here, Vladimír Bažil provides an overview of the molecules that are cleaved from cells, focusing particularly on leukocyte receptors. In addition, he discusses the mechanisms and putative enzymes involved in this process, as well as the potential physiological significance of such events.*

Several membrane molecules are cleaved from the surface of leukocytes and other cells by endogenous cellular enzymes, thereby releasing their soluble fragments into the extracellular medium (Table 1). This has been demonstrated *in vitro* using a combination of various methods, including: (1) detection of soluble fragments in the culture supernatant of surface-labeled cells; (2) inhibition of the release of the soluble forms by enzyme inhibitors; and (3) comparison of the apparent molecular weight of the membrane form with the released soluble form.

## Two different pathways inducing receptor cleavage

The release of cell-surface molecules following cleavage in the membrane-proximal, extracellular domain is commonly initiated by cell stimulation. Two different pathways inducing this process have been identified. The first of these requires engagement of the relevant membrane protein during the initial stage of cleavage induction. For example, CD14 (Ref. 1), CD43 (Ref. 8), CD44 (Ref. 9) and CD62L (L-selectin)<sup>22,23</sup> are cleaved and released from the surface of leukocytes following incubation *in vitro* with monoclonal antibodies (mAbs) that recognize these individual receptors, possibly simulating the effect of their natural ligands. Crosslinking of the receptors on the cell surface is critical for their effective cleavage. In addition, interleukin 3 (IL-3)<sup>15</sup> and CD27 ligand<sup>24</sup> induce cleavage of their respective receptors, demonstrating that natural ligands are also able to induce this process. The second pathway of receptor cleavage may be initiated either by natural factors which stimulate cells *via* receptors that are differ-

ent from the molecules to be cleaved, or by phorbol 12-myristate 13-acetate (PMA), a potent synthetic activator of protein kinase C. Thus, cytokines and chemotactic peptides induce cleavage of L-selectin<sup>11,25</sup> and tumor necrosis factor receptors (TNFRs)<sup>14</sup>; immunoglobulins interacting with cell-surface Fc receptors induce cleavage of L-selectin<sup>23</sup>; and anti-CD20 mAbs (Ref. 26) or the CD40 ligand<sup>27</sup> stimulate cleavage of CD23. Furthermore, PMA initiates cleavage of CD14 (Ref. 1), CD16 (Ref. 3), TNFRs (Ref. 14), CD43 (Ref. 8), CD44 (Ref. 9), L-selectin<sup>11,25</sup>, IL-6R (Ref. 16), and membrane-anchored precursors of transforming growth factor  $\alpha$  (pro-TGF- $\alpha$ ), *c-kit* ligand 1 (KL-1) and KL-2 (Ref. 21).

The cleavage-triggering event may initiate two distinct activation processes. One mechanism may include a conformational alteration of the membrane molecule to be cleaved, exposing a cleavage site for the enzyme involved. This process may be induced either by direct binding of a particular ligand, or by a signal transduced from a cytoplasmic compartment that results from activation events independent of the molecule to be cleaved. This intracellular signal may be mediated by the modification of the cytoplasmic tail of the receptor, such as by phosphorylation, or by the association/dissociation of the receptor with other molecules or cytoskeletal components. For example, the C-terminal valine residue located in the cytoplasmic tail of pro-TGF- $\alpha$  has been shown to be essential for PMA-induced cleavage of this membrane-bound cytokine<sup>28</sup>. Thus, an 'inside-out' signaling event that emanates from the cytoplasm, and which requires the pro-TGF- $\alpha$