17

Cooperation in Heterogeneous Populations

Karl Sigmund[1] and Martin Nowak[2]

ABSTRACT This paper deals with stochastic reactive strategies for the Iterated Prisoner's Dilemma. It considers populations of individuals meeting randomly, and noisy interactions. Both the analysis of monomorphic and heteromorphic populations show that the reciprocal strategy Tit For Tat acts like a pivot: it triggers an evolution towards cooperation, but is not the ultimate beneficiary of such an evolution.

17.1 Introduction

Ever since the early days of game theory, the Prisoner's Dilemma (PD) has been viewed as a major paradigm for discussing cooperation between selfish individuals in the absence of constraints. From 1979 onwards, the topic received a further strong impetus by Axelrod's computer tournaments (see Axelrod, 1984). In addition to theoretical work on the mathematical aspects of the game, a wide variety of research on its psychological, biological, political, and economical implications and applications has developed (see Axelrod and Dion, 1988).

The two players of a PD game each have a choice between two strategies, namely, to cooperate (C) or to defect (D). If both players cooperate, each receives as payoff a reward R which is larger than the payoff P (punishment) they receive if both defect. But a player defecting unilaterally obtains a payoff T (the temptation) which is still higher than R, while the opponent receives only S (the sucker's payoff) which is even lower than P. In addition to $T > R > P > S$, it is assumed that $T + S < 2R$, so that mutual cooperation is better than agreeing on a unilateral defection and sharing the payoffs. (In all the computer simulations, Axelrod's payoff values $T = 5, R = 3, P = 1$, and $S = 0$ are used, but other choices lead to similar results.) Since D dominates C (no matter what the co-player does, it is always better to defect), both players ought to opt for D: But then, their payoff P is smaller than the payoff R they would have obtained by both cooperating.

Many people experience difficulties in accepting defection as the "rational" solution (for an entertaining experiment, see chapter 30 of Hofstadter, 1985); but, to quote Luce and Raiffa (1957), "no, there appears no way around this dilemma.

[1]Department of Mathematics, University of Vienna, A-1090 Vienna, Strudlhofgasse 4, Austria
[2]Department of Zoology, University of Oxford, Oxford OX1 3PS, South Parks Road, England

We do not believe there is anything irrational or perverse about the choice of D and we must admit that if we were actually in this position we would make these choices."

One way of persuading oneself of the validity of this conclusion is to omit any appeal to "rationality" and to consider a *population* (biological or artificial) playing this game in randomly selected pairwise encounters. If one assumes that a successful strategy spreads (by inheritance, imitation, or infection, say), then necessarily defectors spread faster and end up by taking over the whole population.

Nevertheless, there is a sizeable amount of cooperation around us, and a tendency towards cooperation within us. How can this be explained? One possible answer would be that the PD is only of marginal relevance for the emergence of cooperation. But this suggestion can quickly be discarded. It seems, in fact, that most business transactions have the same structure as the PD: the players profit by cooperating in a joint enterprise (like trading or bartering), but could profit still more by not delivering their own full share. This remark leads to another possible explanation for cooperation: PD-like transactions are so common because they can be repeated. What we observe is frequently an Iterated Prisoner's Dilemma (or IPD), rather than a single-shot Prisoner's Dilemma.

If one assumes that there exists a constant probability w for repeating the encounter between the two players, one obtains a vastly more complex game. There is now an infinity of different strategies, and it can easily be shown that, if w (the *shadow of the future*, to use Axelrod's expression) is larger than $\frac{T-R}{T-P}$, there is no dominating strategy any more, (i.e., no strategy which is best regardless of what the other player does). In this case, the question of which strategy the other player is likely to choose becomes essential. In 1979, Axelrod put the diverging opinions on this matter to the test by inviting game theorists to submit their strategies (in the form of programs) to a computer tournament. The rest is history: the winner of the round-robin contest was Tit For Tat (or TFT), submitted by A. Rapoport. It consists in playing C in the first round and then doing whatever the other player has done in the previous round: first cooperate, then imitate, to put it in a nutshell.

The tournaments' outcome was surprising, not only because of TFT's utter simplicity, but also because, as is easily seen, a TFT-player does never better than the co-player. Whoever uses TFT can never be ahead in the count (but never lags far behind either). The strong point of TFT is that it elicits cooperation from the co-player. By virtue of its transparency, it allows the co-player to dictate the evolution of the game and thereby manipulates this co-player to opt for cooperation.

Since the IPD is not a zero-sum game, the TFT-program can win the tournament although it achieves at best a draw against any other program. It earns

a large payoff by cooperating with strategies which, against other strategies or their like, engage in fruitless defections. This implies that the full advantage of TFT is only displayed within a population. In further work, Axelrod kept investigating this context: first, by repeating the tournament with a considerably larger number of entries, and then by applying game dynamics, i.e., selection: the population of programs was submitted, generation after generation, to a round-robin contest, and the frequencies of the programs increased at a rate proportional to their payoff in the previous generation. Again, TFT did best.

This led to a heightened interest in the role of reciprocity in human and animal behaviour. However, Axelrod's simulations dealt with a highly selective sample of strategies submitted by sophisticated contestants, running its programs in an error-free world. These two features are rather artificial. In order to approximate more realistic scenarios, May (1987) suggested to "take stochasticities into account and investigate representative samples". The following paper is a step in this direction. It analyses a highly heterogeneous population of players under conditions where noise blurs the perception and implementation of the players' moves. This approach shows, among other things, that TFT is a catalyst rather than a beneficiary of an emerging cooperation.

We start with a short account of game dynamics and a brief sample of biological applications. Then, we introduce stochastic reactive strategies and discuss the evolution both of monomorphic and polymorphic populations. Finally, we compare this with related work and sketch possible extensions.

17.2   Game Dynamics

One can compute the payoff A_n obtained in the n-th round by a player using strategy E against a player using strategy E', and hence the total payoff $A(E, E') = \sum_n A_n w^n$ (or, in the limiting case $w = 1$ of the infinitely iterated PD, $A(E, E') = \lim \frac{1}{n}(A_1 + ... + A_n)$). Consider now a large population where the strategies $E_1, ..., E_n$ exist with frequencies $x_1, ..., x_n$ (with $\sum_i x_i = 1$). The average payoff for strategy E_i in this population is given by $f_i(x) = \sum_j x_j A(E_i, E_j)$. According to the usual discrete game dynamics (see, e.g., Hofbauer and Sigmund, 1988), the frequency x_i' of E_i in the next generation is given by

$$x_i' = x_i \frac{f_i(x)}{\bar{f}},$$

where $\bar{f} = \sum_i x_i f_i(x)$ is the average payoff in the population. Thus the composition of the population changes from generation to generation. This composition in turn affects the average payoff for the strategies. Nowak and Sigmund (1989) show that with just three or four different strategies a rich spectrum of dynamic behaviour can be obtained, including bistability and limit cycles.

If w is sufficiently large, a population playing $AllD$ (i.e., always defecting) can

be invaded by *TFT* players if they enter in a cluster exceeding a certain size, see Axelrod (1984). (*TFT*-players can make up among themselves more than they lose against defectors. They therefore increase in frequency, which in turn increases the frequency of contacts with other *TFT*-players and hence their advantage over *AllD*). Imagine that such a cluster is given by a family, for instance. A *TFT*-population, on the other hand, cannot be invaded by a cluster of defectors. *AllC*-players, however, do as well as the *TFT*-players in a *TFT*-population. Thus, there is no selection against *AllC*, so that it can spread by random drift. But once a sizeable portion of the population consists of individuals who do not retaliate, *AllD* players can invade again.

Stochastic events can threaten a *TFT*-population in a more direct way too: any mistake gives rise to a long sequence of alternating defections. In a heterogeneous population, this effect is less pronounced: when Axelrod repeated his tournament with an error rate of 1%, *TFT* still finished first. But in a homogeneous population of retaliators, even rare errors are very costly, at least if w is large. It is obviously better to be generous and to forgive, at least occasionally, a defection by the co-player. One such generous strategy is Tit For Two Tats (or *TFTT*), which defects only if the co-player defected in the previous *two* rounds. (This strategy would have won the first of Axelrod's tournaments had it been used: but when it was entered in the second tournament, it finished only twenty-first.)

In a mixed population of *TFT* and *TFTT*, all players do equally well if there are no mistakes. In a noisy environment, however, *TFTT* has a considerable advantage. But if a minority of the population acquires the knack to cooperate and defect in turns, then it can systematically exploit *TFTT*, wheras *TFT* holds its own. On the other hand, a small group of players using Suspicious Tit For Tat (or *STFT*), which defects on the first move and then repeats whatever the other did in the previous round, will hurt *TFT* considerably, while *TFTT* escapes almost unharmed.

17.3   Reciprocation in Animal Behaviour

The Prisoner's Dilemma is a favoured game of experimental psychologists (see, e.g., Selten and Stoecker, 1986). But while human communities offer many examples of *IPD*s, these are usually influenced by factors like social pressure, law enforcement, ideologies, etc., which alter the payoffs in a complex way. Biological populations pose less problems in that respect. Some of the leading sociobiologists, like Hamilton (see Axelrod and Hamilton, 1981) and Trivers (1985), have studied the aspects of cooperation and reciprocation in animal behaviour extensively. Here we list a few recent well documented cases:

(a) Predator inspection. If sticklebacks notice a pike in their vicinity, they tend to approach (up to a certain distance) in order to obtain information on this new enemy. The risk of such an approach is considerably reduced if sticklebacks

approach in pairs. As long as the sticklebacks approach together or alternate in taking the lead, they cooperate. In an ingenious experiment, Milinski (1987) had a single stickleback confronted with a dummy pike, and used a mirror to make the stickleback think that it had a companion. Depending on the inclination of the mirror, this fake companion kept either abreast or lagged behind by a few inches consistently. In the former case, the deluded stickleback usually approached the dummy much closer, which suggests that its strategy is based on reciprocity.

(b) Grooming. The so-called cleaner fishes attend to the teeth of much larger fish. This is to their mutual benefit, but the larger fish could be tempted to cap their profit by swallowing the cleaner. The stability of the interaction seems to be due to the probability of a repetition; indeed, such grooming has only been observed among fish with a fixed home-range offering a reliable meeting place (see Axelrod and Hamilton, 1982).

(c) Egg-sharing. Sea-basses are equipped with male and female sex organs. In the course of a mating, they switch their sexual roles up to ten times in a row. Since eggs are more expensive than sperm, sea-basses wait, after producing some eggs, until their partner produces a few on its own (Fischer, 1980). By dividing the interaction into several steps, the sea-basses thus change a single-shot PD into an IPD.

(d) Food-sharing. Vampire bats who found no victim may, upon returning to their cave, be fed by their luckier brethren, who regurgitate some of their stomach's content, in what seems to be a case of reciprocal altruism (Wilkinson, 1984).

(e) Wife-sharing. One young male baboon starts to quarrel with the overlord of an estrous female, while his comrade uses the digression to mount her sneakily. On the next occasion, the two young males exchange their roles (see Packer, 1979).

(f) Mutual assistance. Vervet monkeys living in a group can recognize each other's vocalizations. Their support calls were tape-recorded and later played by a hidden loudspeaker. A monkey is more likely to respond if it has been groomed by the caller a short while ago (see Cheney and Seyfarth, 1982). Interestingly, responsiveness within a family is higher, but remains unaffected by recent grooming – acts of kindness appear to elicit less response among kin.

17.4   Homogeneous Populations

What is common to these real-life examples is that the decision whether or not to cooperate depends on what the other player has been doing, but not in a strictly deterministic way. Individuals do not obey a clear-cut rule, rather they exhibit a stronger or weaker tendency to play C. This appears to be a fuzzy pro-

gram. To model it, we consider stochastic reactive strategies defined by triplets (y, p, q), where y is the probability to play C in the first round, while p and q are the conditional probabilities to play C after an opponent's C and, D respectively, in the previous round. These strategies correspond to points in the unit cube. $AllC$ is the corner point $(1, 1, 1)$, $AllD$ is the point $(0, 0, 0)$, TFT is the point $(1, 1, 0)$, and $STFT$ the point $(0, 1, 0)$. These strategies, however, are not properly stochastic: they correspond to extremal points of our strategy set. As soon as we take noise into account, we are in the interior of the cube.

It is clear that these (y, p, q)-strategies form only a small part of the huge set of all possible strategies for the IPD. For instance, $TFTT$ does not belong to this class. But a drastic restriction is necessary whenever one wants to consider representative samples; otherwise, such samples would have to be far too large and therefore unsuited for analysis. Furthermore, the class of strategies considered is sufficiently broad to contain highly diverse specimens, covering a wide range of cooperative, defective, reciprocal, or fully random behaviours; as a consequence, this set of strategies allows for complex dynamical extensions.

If a player using strategy $E = (y, p, q)$ is matched against a player using strategy $E' = (y', p', q')$, then the probability of the former to cooperate in the n-th round determines the co-player's probability to cooperate in the following round and hence also the own level of cooperation in the $(n+2)$-nd round. This echo-effect determines recursive relations which imply that the probability of the first player choosing C converges (for $n \to \infty$) to

$$c(E, E') = \frac{q + (p - q)q'}{1 - (p - q)(p' - q')}.$$

In particular, the probability to cooperate with a copy of oneself converges to

$$c(E, E) = \frac{q}{1 - p + q}.$$

If (p', q') lies on the line from (p, q) to $(1, 0)$, then $c(E, E) = c(E', E') = c(E, E') = c(E', E)$, and hence E and E' have the same asymptotic level of cooperativity. If (p', q') lies above this line, then E' is more cooperative, both against E and against a copy of itself. Hence this line divides the (p, q)-space into one part where cooperativity is higher (which includes $AllC$) and one part where it is lower (which includes $AllD$).

It is straightforward to compute the payoff $A(E, E')$, see Nowak and Sigmund (1989b). For the sake of simplicity, we shall only consider the limiting case $w = 1$ (but emphasize that our results essentially apply whenever w is large). We then can forget about y, the probability to cooperate in the first move, which only affects the initial phase and plays no role in the long run. We then obtain as payoff for strategy $E = (p, q)$ against $E' = (p', q')$ the value

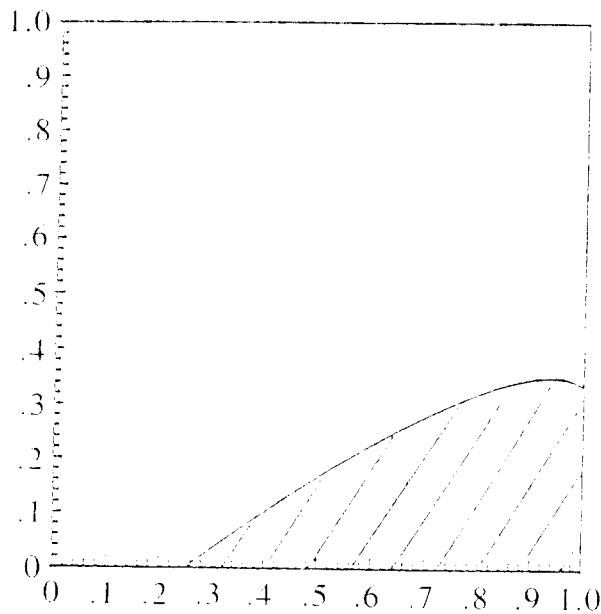$$A(E, E') = (R + P - T - S)c(E, E')c(E'E) + (S - P)c(E, E') + (T - P)c(E', E) + P.$$

FIGURE 17.1. The square of (p, q)-strategies (see text). Point (1.0) corresponds to TFT; in its neighborhood (the cooperation-rewarding zone), a population can be invaded by mutants with a higher p- or q-value.

This allows us to analyse the evolution of populations which are monomorphic in the sense that the great majority of players settles on one fixed strategy E'. We may characterise those strategies E which do better, against E', than E' itself (see Nowak, 1990). A small minority using such a strategy E can invade the E'-population. It turns out that the space of our (p, q)-strategies, i.e., the unit square, is divided into two regions (see Fig.17.1). In the so-called C-region, which contains $TFT = (1, 0)$, an E'-population can be invaded by strategies E having a higher level of cooperation, but not by any strategies with a lower level. In the D-region, which contains $AllC = (1, 1)$ and $AllD = (0, 0)$, the converse holds. If E' is on the boundary line between these two regions, no E does better than E' in an E'-population, and only those E on the line through E' and $(1, 0)$ do equally well, i.e., have a chance to invade by neutral drift. This shows that a tendency towards cooperation exists only in the C-region, i.e., in a neighborhood of TFT. It must be emphasized that this is not a tendency *towards* TFT: it pays most to deviate from the "consensus"-strategy E' played by the homogeneous population, not in the direction towards TFT (such a move would not affect payoff at all), but in an *orthogonal* direction.

In the limiting case of very small noise, the strategy $(1, q)$ with $q = \min[1 - \frac{T-R}{R-S}, \frac{R-P}{T-P}]$ can be shown to be optimal in the sense that it affords the highest payoff for a population adopting it, compared to all reactive strategies immune to invasion by less cooperative strategies (i.e., with lower p-and q-values), see Molander (1985) and Nowak and Sigmund (1990). This strategy, which is called Generous Tit For Tat (or $GTFT$), can be viewed as a mixture of TFT and $AllC$. It is grateful in the sense that it never forgets a good deed, and tolerant in the sense that it sometimes forgives a defection (but not always).
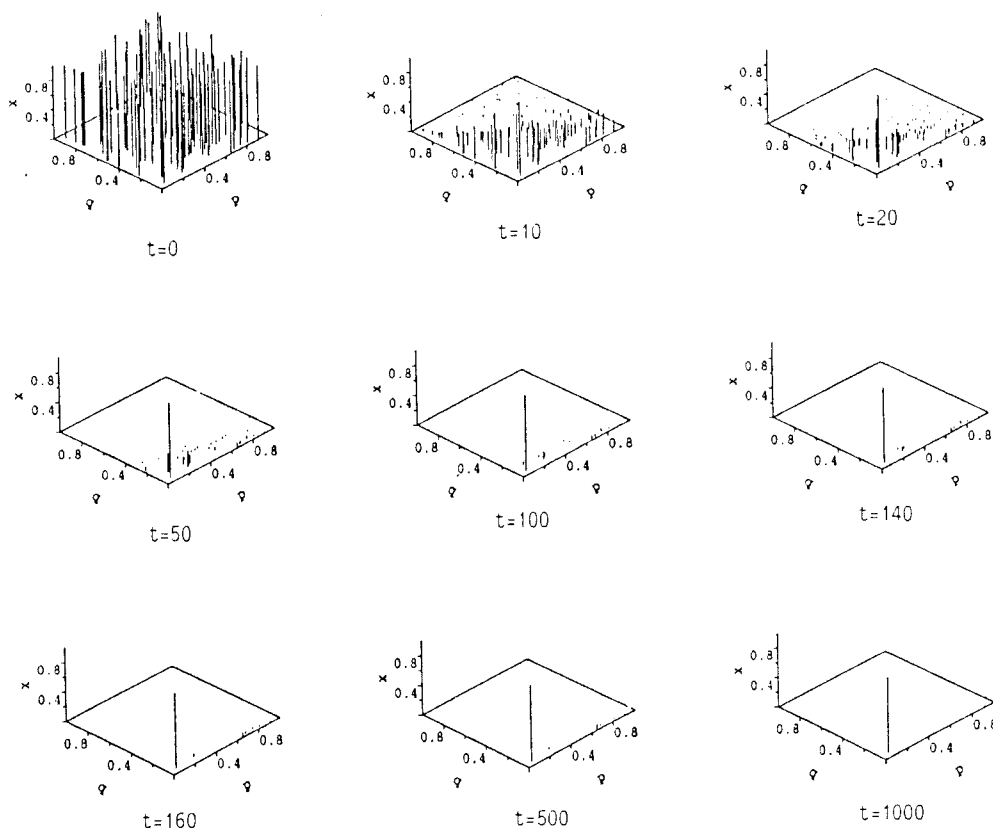
FIGURE 17.2. The evolution towards defection. We start out with 100 randomly chosen strategies, all equally frequent at the beginning. The defectors, close to the $(0,0)$-corner, increase in frequency and ultimately take over.

17.5   Heterogeneous Populations

If the population is heterogeneous, the evolution is considerably more complex. If there are two, three, or four different strategies in the population, the game dynamics can still be studied analytically (see Nowak and Sigmund 1989b). For still more diversified populations, one has to resort to computer simulations.

In Nowak and Sigmund (1992) we studied random samples consisting of hundred different (p,q)-strategies (which means that 10000 interactions have to be computed in each generation). If the strategies are uniformly distributed on the unit square, all initially equally frequent, the evolution tends in almost every case towards $AllD$. This means that the strategies closest to $(0,0)$ increase in frequency from generation to generation, while all others vanish (see Fig.17.2). Such a tendency is due to the fact that a large percentage of the initial random sample has high q-values and does not retaliate against exploiters. It pays to defect against inverterate "suckers". By the time these born victims are eliminated from the population, the defectors are so preponderant that cooperation cannot be established.

If the initial population contains a strategy which is *very* close to TFT, however, the evolution takes a remarkably different turn (see Fig. 17.3). The first
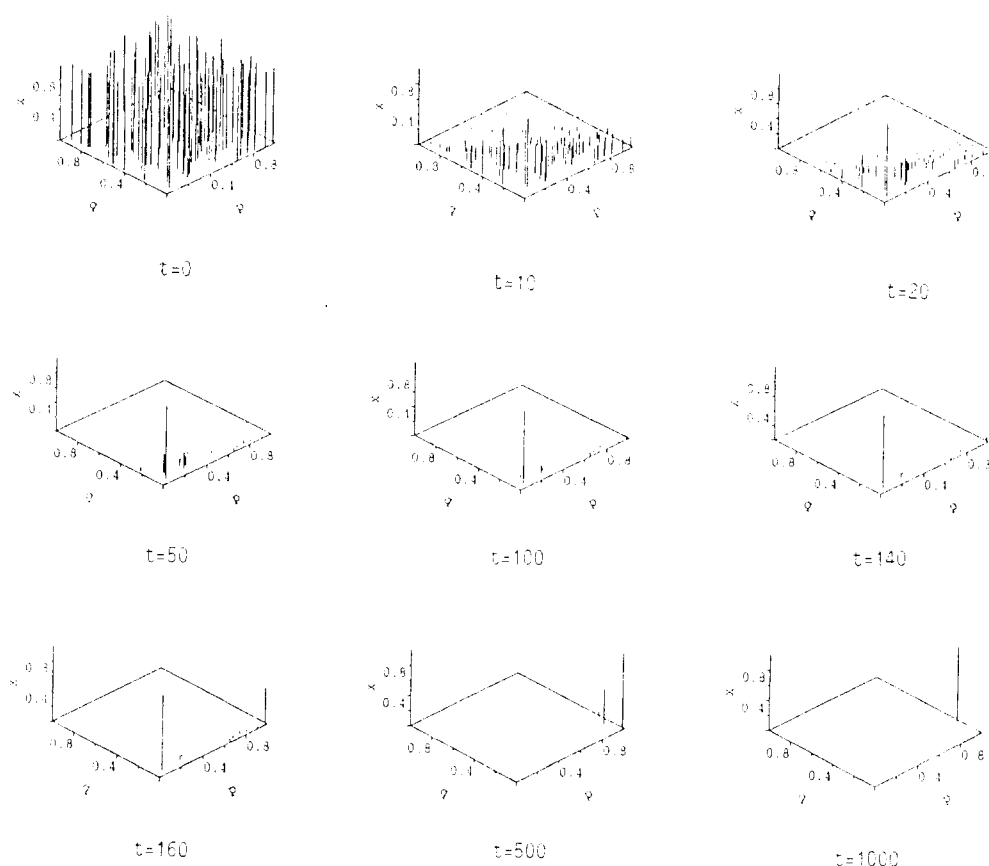
FIGURE 17.3. This time, a strategy close to TFT (with $p = 0.99$ and $q = 0.01$) has been added by hand. At first, defectors thrive, but when the "Suckers" who do not retaliate are driven from the population, the TFT-like strategy takes over. This is in turn superseded by a more generous strategy.

phase is almost indistiguishable from the runs described previously. The strategies close to $AllD$ grow rapidly, feasting on their diet of "suckers". All reciprocating strategies close to $(1,0)$ seem to vanish. But when the victims are so reduced that the defectors can no longer exploit them, the small band of reciprocators returns upon the scene. The fortune of the exploiters takes a decided turn for the worse. Their frequencies dwindle. Now it is the TFT-like strategies who dominate the population. But their hegemony is transient, too. For, when the defectors are exterminated, the reciprocators give way to more generous strategies and vanish also. When the evolution finally grinds to a halt, after several hundred generations, the population consists almost exclusively of strategies close to $GTFT$. Even if we allow for random fluctuations, for instance by introducing occasionally a small percentage of another strategy, this will not alter the final equilibrium.

It should be noted that this turn towards cooperation can only be triggered by a strategy that is extremely near to the TFT-corner $(1,0)$. In general, a random sample of 100 strategies *uniformly* distributed over the unit square will not contain such a strategy. In the simulations, it can either be introduced by hand, or by choosing a probability distribution which places more weight upon

the boundaries of the range $]0, 1[$ of p- and q-values. The most plausible candidate for such a distribution has the familiar U-shaped density function

$$f(p) = \frac{1}{\pi \sqrt{p(1-p)}},$$

which is a particular beta-distribution playing a considerable role in statistics. Such a weighting of the values close to all or nothing (i.e. to 0 and 1) makes sense, since a change in, say, 1% in their vicinity can be much more relevant for the future evolution than a change from 50% to 49%. Whenever random samples according to such a distribution are used, a dramatic turn of the tide in favour of cooperation is obtained: while the exploiters seem to be unstoppable at first, and after ten generations only $AllD$-like strategies appear to be present, a minority of reciprocators survives and acts like a lever to twist the momentum away from defection. After fifty generations, all q-values larger than 10% have gone into hiding, and the battle between $AllD$ and TFT takes a determined shift in the latters' favour. But after a hundred generations, when only p-values larger than 95% remain on stage, higher q-values re-enter and generosity starts to pay. In such an evolution, TFT-like strategies act as catalysers. They need be present only in a trifling amount in order to trigger the emergence of cooperation. During the transient phase of the reaction, their concentrations grow, but then vanish again.

The simulation in polymorphic populations confirms the result obtained by an analysis of the monomorphic scenario: TFT is the pivot rather than the aim of evolution. One way to visualize this pivotal role of TFT is to ask when an $AllD$ population can be invaded by a cluster of E-strategists, for a given size $d \in]0, 1[$ of the cluster. In a dimorphic population where d is the relative frequency of $E = (p, q)$ and $1 - d$ the relative frequency of $AllD$, the amount of E will grow if and only if

$$dA(E, E) + (1 - d)A(E, AllD) > dA(AllD, E) + (1 - d)A(AllD, AllD).$$

For any given value d, this defines a neighborhood U_d of TFT in the strategy square. If d decreases, the neighborhood shrinks to the corner point $(1, 0)$. Thus TFT is the strategy that can invade defectors in a cluster of minimal size.

17.6   Discussion

There are several directions in which these results could be extended. In particular, the range of strategies ought to be enlarged. The (p, q)-strategies depend only on the co-player's move in the previous round. It would be more realistic to admit strategies with a memory of several rounds, taking the own moves into account. In principle, this is straightforward: such a process is simply a Markov chain. With a memory of length 2, for instance, there are 16 relevant histories.

Each strategy is a 16-tuple $p_0, ..., p_{15}$ of probabilities to play C after a given history. This yields a 16×16 transition matrix, which has a unique left eigenvector $\pi_0, ..., \pi_{15}$ corresponding to the eigenvalue 1 which is stochastic (i.e., $\pi_i \geq 0$ and $\sum \pi_i = 1$). This describes the probabilities for cooperation between the two players in the stationary state, and hence the asymptotic payoff. The problem with this approach is less the computation of the eigenvector (the 16×16-matrix has many 0-entries), but the fact that the dimensionality of the strategy space is so huge that samples have to be impossibly large in order to represent all possibilities.

This problem occurs even if we consider only stochastic strategies with memories of lenght 1 (but taking the own move into account). Even if one uses U-shaped distributions for the sample, it may be that the TFT-corner is underrepresented and cannot exert its policing function. Some simulations lead towards $AllD$, while others display the turn towards cooperation as described in the previous section.

A very interesting approach is due to Lindgren (1991), who considered strategies which are less stochastic than ours, but have longer memories (2 or 3). More precisely, the probabilities to play C, given this or that history, are either ϵ or $1 - \epsilon$, where ϵ is a small number corresponding to the noise. This means that the strategies are "intended" to be deterministic, but may suffer with a small probability ϵ from a mistake. Lindgren's simulations exhibit highly complex evolutionary histories, frequently leading to meta-stable states displaying coexistence, and even symbiosis. An interesting result is the emergence of a large class of strategies which (a) cooperate if both players have cooperated in the previous two rounds, defected in the previous two rounds, or simultaneously switched from D to C in the last round; and (b) defect if they, or the co-player, have unilaterally switched from C to D. Such strategies handle mistakes in a surprising, but efficient way: if one player accidentally plays D, both players play two rounds of D and then resume playing C. Such strategies have already been proposed by Micko et al. (1977); they are difficult to exploit systematically, but do not do very well against TFT.

Lindgren's model extends simulations by Axelrod (1987), who used a genetic algorithm to simulate the evolution of a population of strategists whose memory has a length of 3. Such algorithms use random mutations and recombinations to introduce new strategies and submit this diversity to culling by selection. Axelrod also observed evolutionary histories which first lead toward defection and later turn toward cooperation.

It would be interesting to see whether increasing the memory capacity or the ability to randomize is more advantageous in playing the IPD. Preliminary investigations seem to show that the situation is bistable. On one hand, (p, q)-strategies cannot invade a meta-stable population of the type described by Lindgren; on the other hand, strategies of that type cannot invade a population of

GTFT-players. In order to be efficient under conditions of noise and uncertainty, a strategy will probably need both to keep track of the last few rounds and to react in an unpredictable way to isolated defections.

17.7   References

Axelrod R. (1984). *The evolution of cooperation.* New York: Basic Books Inc. (reprinted by Penguin, Harmondsworth)

Axelrod, R. (1987). The. evolution of strategies in the iterated prisoner's dilemma. In D. Davis (Ed.), *Genetic algorithms and simulated annealing* (pp. 32-43). London: Pitman.

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

Axelrod, R., & Hamilton, W.D. (1981). The evolution of cooperation. *Science, 211,* 1390-1396.

Cheney, D.L., & Seyfarth, D. (1982). Recognition of individuals within and between groups of free-ranging vervet monkeys. *American Zoologist, 22,* 519-529.

Fischer, E.A. (1980). The relationship between mating system and simultaneous hermaphroditism in the coral reef fish. *Hypoplectrus nigricans. Animal Behaviour, 28,* 620-633.

Hofbauer, J., & Sigmund, K. (1988). *The theory of evolution and dynamical systems.* Cambridge: Cambridge University Press.

Hofstadter, D. (1985). *Metamagical themas: Questing for the essence of mind and pattern.* New York: Basic Books.

Lindgren, K. (1991) Evolutionary phenomena in simple dynamics. In Langton et al. (Eds.), *Artificial Life II* (pp. 295-312), Proceedings of the Santa Fe Institute Studies, Addison-Wesley.

Luce, R.D., & Raiffa, H. (1957). *Games and decisions.* New York: Wiley.

May, R.M. (1987). More evolution of cooperation. *Nature, 327,* 15-17.

Micko, H.-Ch., Brückner, G., & Ratzke, H. (1977). Theories and strategies for prisoner's dilemma. In W.F. Kempf & B.H. Repp (Eds.), *Mathematical Models for Social Psychology* (pp. 214-276). Bern: Huber.

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

Molander, P. (1985). The optimal level of generosity in a selfish, uncertain environment. *Journal of Conflict Resolution, 29,* 611-618.

Nowak, M. (1990). An evolutionarily stable strategy may be inaccessible, *Theoretical Population in Biology, 142,* 237-241.

Nowak, M., & Sigmund, K. (1989a). Oscillations in the evolution of reciprocity. *Journal of Theoretical Biology, 147,* 21-26.

Nowak, M., & Sigmund, K. (1989b). Game dynamical aspects of the prisoner's dilemma. *Journal of Applied Mathematics and Computation, 30,* 191-213.

Nowak, M. & Sigmund, K. (1990). The evolution of reactive strategies in iterated games. *Acta Applicandae Mathematicae, 20*, 247-265.

Nowak, M., & Sigmund, K. (1992). Tit for tat in heterogeneous populations. *Nature, 353*, 250-253.

Packer, C. (1979). Reciprocal altruism in *Papio anubis. Nature, 265*, 441-443.

Trivers, R. (1985). *Social evolution.* Menlo Park: Benjamin-Cummings.

Selten, R., & Stoecker, R. (1986). End behaviour in sequences of finite prisoner's dilemma supergames. *Journal of Economical Behavior and Organization, 7*, 47-70.

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