IIASA

International Institute for Applied Systems Analysis • A-2361 Laxenburg • Austria
Tel: +43 2236 807 • Fax: +43 2236 71313 • E-mail: info@iiasa.ac.at • Web: www.iiasa.ac.at

INTERIM REPORT IR-98-099 / December

The Social Life of Automata

Karl Sigmund (ksigmund@esi.ac.at)

Approved by
Ulf Dieckmann (dieckman@iiasa.ac.at)
Project Coordinator, *Adaptive Dynamics Network*

Contents

About the Author

Karl Sigmund
Institut für Mathematik
Universität Wien
Strudlhofgasse 4
A-1090 Vienna, Austria
and
Adaptive Dynamics Network
International Institute for Applied Systems Analysis
A-2361 Laxenburg, Austria

The Social Life of Automata

Karl Sigmund

1 Introduction

Some of the major transitions in evolution occured only once, and others several times (cf. Maynard Smith and Szathmary, 1995). Eusocial colonies, for instance, emerged repeatedly among bees, ants, termites, and aphids. This allows one to compare the importance of different factors, for instance ecological opportunities, life history traits, genetical structures etc. and even to predict where similar societies are likely to be found (as Richard Alexander did in anticipating the sterile worker caste of the naked mole-rat, see Sigmund, 1993). In contrast, no parallel to human societies is known in the history of evolution; we seem to be unique in having achieved a social structure distinguished (i) by the levelling of reproductive opportunities, (ii) by the prevalence of division of labour, mutual help and economic exchange between non-related individuals, (iii) by information transfer based on language and (iv) by moral obligations both externally enforced through group sanctions and internaliced through powerful emotions.

We can, of course, learn much about the cultural determinants of human societies by comparing tribes, clans, states and gangs; but all these are manifestations of a universal 'human nature' caused by a major biological transition which, apparently, has occured once only. In order to analyse the mechanisms responsible for it, we have to use thought experiments.

Game theory was devised explicitly as a tool for the social sciences. It was meant to model the independent decision-making process of interacting individuals, each bent upon the 'selfish' goal of maximizing his or her own payoff. Interpreting payoff as reproductive fitness provides a good tool for studying individual selection based on Darwinian competition. But for decades, game theory was handicapped by the fiction of the 'rational player', despite the fact that many economists, and every psychologist, knew better. Furthermore, it was only when biologists started to use game theory that populations of individuals were considered (see Maynard Smith, 1982, Binmore, 1992, Weibull, 1995, Hofbauer and Sigmund, 1998).

The advent of evolutionary game theory has changed all this. Individual players were no longer assumed to be rational, but to follow simple, knee-jerk rules. In the spirit of Richard Dawkins (1976), who claims that we are mere robots, players were therefore modelled by simple automata. Populations of such interacting automata, engaged in the massively parallel kind of problem-solving characteristic of Darwinian evolution, were studied by means of computer simulations or mathematical analysis following their evolution for many generations. New strategic variants (or programs) were introduced either by random processes or by hand, and tested against the current composition of the population. Nonlinear dynamics describing the resulting adaptation or selection processes, were used to analyse the chronicles of these artificial societies (see, e.g., Axelrod, 1997).

This program of evolutionary game theory has been applied to a wide variety of biological and (more recently) economic topics. In this paper, we sketch some recent developments in one particularly active field: the evolution of cooperation. More precisely, we deal with one of the three factors currently recognized as essential, namely with reciprocity. This is not meant to downplay the importance of the other two factors, viz. relatedness (Hamilton, 1963) and group selection (Wilson and Sober, 1994). Doubtlessly close kinship ties, the major basis of cooperation in clones and bee colonies, did prevail in early hominid groups. Furthermore, group selection – more precisely, the individual selective advantage due to belonging to a successful group – was essential, because the major threats to survival were most likely coming from rival groups. This being said, let us turn to the subject of reciprocal altruism, originally introduced in a landmark paper by Trivers (1971) which, to this day, serves as an inspiration to the field.

2 Reciprocal Altruism and the Prisoner's Dilemma

Assume that in an encounter between two players, one is a potential donor and the other a recipient. The donor can give help that the recipient needs. Giving help costs $-c$ to the donor and yields payoff b to the recipient (with the payoff interpreted as Darwinian fitness, i.e. reproductive success, and assuming $0 < c < b$). According to Hamilton's rule, it will pay to help if the degree of relatedness r between donor and recipient (i.e. the probability that a randomly chosen gene from the donor also belongs to the recipient's genome) is larger than the cost-to-benefit ration, i.e.

$$r > c/b.$$

(Since $r \leq 1/2$ under normal circumstances – barring identical twins or high inbreeding – this condition requires $c < b/2$). Does this imply that one should never help an unrelated individual? Not so, according to Trivers, if there is a reasonable chance that the recipient is able to return the help. This is the principle of reciprocal altruism – in Trivers' definition, 'the trading of altruistic acts in which benefit is larger than cost, so that over a period of time both enjoy a net gain' (Trivers, 1985, p. 361). Accordingly, 'reciprocal altruism is expected to evolve when two individuals associate long enough *to exchange roles frequently* as potential altruist and recipient'.

This, however, opens the door to unilateral defection. Suppose that in two consecutive rounds, the players exchange the roles of donor and recipient. If both help each other, both obtain $b - c$. This is higher than the payoff 0 obtained if both refrain from helping. But is one helps and the other does not, then the helper is left with the costs of his act, $-c$, and the recipient gets away with b points. This is just the rank ordering of the payoff values for the Prisoner's Dilemma, which had been studied for many years by game theorists and experimental psychologists. The difference is merely that both players have to decide simultaneously whether to cooperate (play **C**) or to defect (play **D**). If both play **C**, both receive the reward R

for mutual cooperation; if both play **D**, both receive the punishment P; and if one defects unilaterally, he receives the temptation T whereas the other player is left with the sucker's payoff S. For the Prisoner's Dilemma game, it is assumed that

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

The first condition implies that the dominant option is to play **D**: it yields a higher payoff, no matter what the other player is choosing. In the donor-recipient game (with two rounds in alternating roles), one has $T = b, R = b - c, P = 0$ and $S = -c$, so that these inequalities are trivially satisfied.

The conclusion seems inescapable that if the interaction is not repeated, cooperation cannot emerge. It turns out, rather surprisingly, that this conclusion is premature. But we will return to this point only in the concluding discussion – our aim here is rather to follow Trivers in assuming that individuals experience, on average, several interactions.

Let us suppose, then, that the game is repeated with a constant probability w. The number of rounds is a random variable with expected value $(1 - w)^{-1}$. The total payoff is given by $\sum A_n w^n$, with A_n as payoff in the n-th round and w^n the probability for the n-th round to occur. In the limiting case $w = 1$ (the infinitely iterated game) one uses as payoff the limit in the mean, i.e. $(A_1 + ... + A_n)/n$ (provided it exists). If w is sufficiently large, there exists (in contrast to the one-shot game) no strategy which is best against all comers (see Axelrod, 1984). For $w > (T - R)/(T - P)$, for instance, the best reply against *Always*$\mathbf{C}$ is to always defect, whereas against *Grim* (the strategy that cooperates up to the first time that it is been exploited, and from then onwards always defects) it is best to always cooperate.

In a series of round robin tournaments, Axelrod found that the simplest strategy submitted, namely TFT, finished first. Furthermore, Axelrod and Hamilton (1981) explored the emergence of cooperation in evolving populations of players. In particular, they showed that the two strategies *Always*$\mathbf{D}$ and TFT (the Tit For Tat strategy that cooperates in the first round and from then on always repeats the previous move of the co-player) are in bistable equilibrium: none of them can invade the other. But as soon as the frequency of TFT players exceeds a certain threshold (given by $c(1 - w)/w(b - c)$ in the donors-recipient game), keeps growing. For large values of w, this threshold is very small. This means that a small cluster of TFT players can invade a population of defectors: the few interactions with their like more than compensate their loss against the resident majority of defectors.

One may not conclude, however, that a population dominated by TFT can resist invasion by all comers. This becomes particularly obvious if we take into account the possiblity for errors, which must always be present in realistic situations. In fact, the interaction between two TFT players is particularly sensitive to noise. One wrong move causes a whole chain of alternating defections. One further mistake can lead back to mutual cooperation, but just as well to mutual defection. The average payoff decreases drastically. Obviously, the two players should be able to forgive occasionally – not according to a regular pattern, for this could be exploited, but rather on a random basis.

This leads to stochastic strategies. We often do not use hard and fast rules in our everyday interactions, but are guided by factors which are difficult to predict, and which result in a stronger or weaker propensity to opt for this move or that (May, 1987). With TFT, this propensity is 100 percent or 0 percent, depending on whether the co-player cooperated in the previous round or not. With an error rate of 1 percent, TFT cooperates with 99 or with 1 percent probability.

3 Evolutionary Chronicles

Let us consider this in a more general setting. In each round, there are 4 possible outcomes, leading to 4 different payoff values. If we assume that each outcome determines the next move of the player, this yields 16 different strategies (32 if we include the first move). If we allow in addition stochastic strategies with a larger or smaller propensity to cooperate, we obtain a 4-dimensional space of strategies given by quadruples (p_R, p_S, p_T, p_P) where p_i is the probabilty to play $\mathbf{C}$ after outcome i. We can use a computer to find the most successful strategy, by introducing occasionally a small minority of a new, randomly chosen strategy into the population and watching how its frequency develops under the influence of selection. If we run this for a sufficiently long time, we can test a large sample of strategies.

Such mutation-selection chronicles depend on contingencies, and can take very different paths, but they frequently lead to a population dominated by the so-called Pavlov strategy $(1, 0, 0, 1)$ – the strategy that cooperates if and only if, in the previous round, the co-player used the same move as oneself (Nowak and Sigmund, 1993). This strategy embodies a simple win-stay, lose-shift rule: it repeats the former move if the payoff was high (T or R) and switches to the alternative move if the payoff was low (i.e. P or S). Arguably, this is the simplest learning rule. Pavlov seems a hopeless strategy for invading an $Always$**D** population, since it gets suckered every second round. In fact, it needs a retaliatory strategy like TFT or Grim (the strategy that cooperates until suckered, and from then on never cooperates). Once such ‘nice’ strategies (strategies that are not the first to defect) have taken over, Pavlov can invade, because it is tolerant to errors. If two Pavlov players are engaged in a repeated PD game, and one of them commits a mistake and defects, then both players will defect in the next round – the ‘sinner’ because he is happy with his T and repeats the former move, the ‘sucker’ because he shifts to the other option. As a result, both players obtain the low payoff P, switch again and thereby resume mutual cooperation. In addition to being tolerant to errors, Pavlov has also the advantage of being intolerant to $Always$**C** players. After a mistaken defection against such a player, Pavlov keeps defecting. Therefore, indiscriminate altruists cannot spread in a Pavlov population (whereas they could spread by neutral drift in an TFT population). As a result, defectors find no easy victims. They only find Pavlov-players, whom they can exploit in every second round (obtaining $(T + P)/2$ in the mean), whereas Pavlov players obtain R against each other. As long as $2R > T + P$ – or, in the case of the donor-recipient game, as long as $c < b/2$ – a Pavlov population is stable against invasion.

4 Social Norms

How can one formulate this kind of stability? The usual approach would be to test for evolutionarily stability (cf. Maynard Smith, 1982). It can easily be shown, however, that in the context of repeated games, such strategies do not exist. In particular, neither TFT nor *Always***D** are evolutionarily stable, although this has occasionally been claimed. Possibly the most appropriate notion in this context is that of a limit evolutionarily stable strategy. This was originally formulated in terms of extensive games (Selten, 1975), but for repeated games, it is more appropriate to formulate it in terms of strategies implemented by finite automata (Leimar, 1997). In each round, such an automaton can be in one of m internal states. Depending on the state, it plays **C** or **D**; and depending on the outcome of this round, it switches to the next state. In order to test whether such a strategy is a limit ESS, one assumes that it plays against a copy of itself, and looks at what happens at every outcome. If the sequence of moves prescribed by the strategy is better than any alternative, then the strategy is a limit ESS: it is always disadvantageous to deviate from it. In fact, a limit ESS is a social norm: if everybody adheres to it, it wouldn't do not to do.

Let us consider this for TFT. We may view the internal state as defined by the outcome R, S, T or P of the previous round. The action rule is to play **C** after R and T. Fig. 1 shows a graph whose vertices are the four states (the move on top is that by the player, the move below that by the coplayer). The full arrow shows the transition to the next state, if the player uses the move precribed by TFT; the broken arrow shows the transition if the player uses the alternative move. This holds under the assumption that the other player sticks to TFT. We see that in the state S, it would be better to follow the broken arrow: the payoff for the next two moves is $2R$, which is larger than $T + S$. We should note that in principle, the state S should never be reached in a game between two TFT players. But according to the *trembling hand* doctrine of Selten (1975), it can be reached if one player misimplements his move, a mistake which may happen with a small, but positive probability.

If we study the same situation for Pavlov, we obtain the graph of Fig. 2 (again assuming the co-player to use Pavlov). If $2R > T + P$ it is best to follow the full arrows, i.e. Pavlov is a limit ESS. If $2R < T + P$, it is better, when in state P, to deviate from the Pavlov rule and play **D**. In this case, Pavlov is not a limit ESS.

It is easy to find automata that are always norms and lead to cooperation. Let us consider the following example, which has three states and starts in state 1. In state 1 or 2, it plays **C**, and in state 3 it plays **D**. The transition table is given by Table 1. Again, it is easy to check that it is best to always follow the full arrow.

	R	S	T	P
1	1	2	3	1
2	1	2	1	2
3	1	1	3	3

Table 1

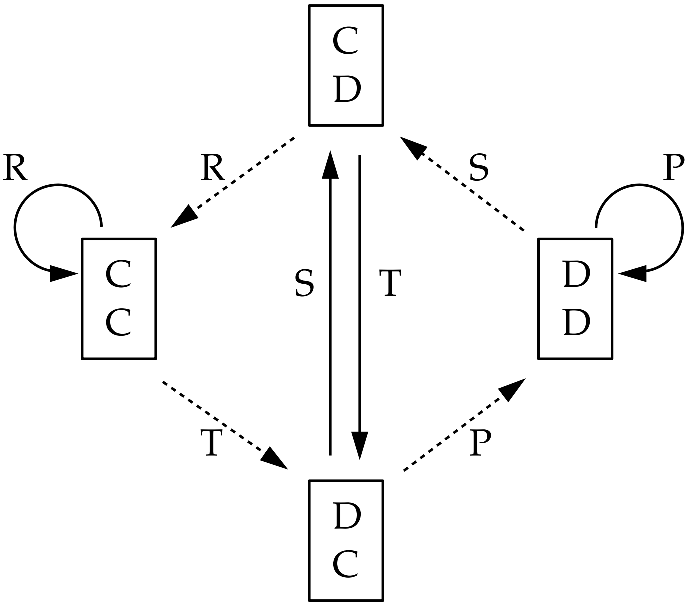

Figure 1: TFT is not a norm

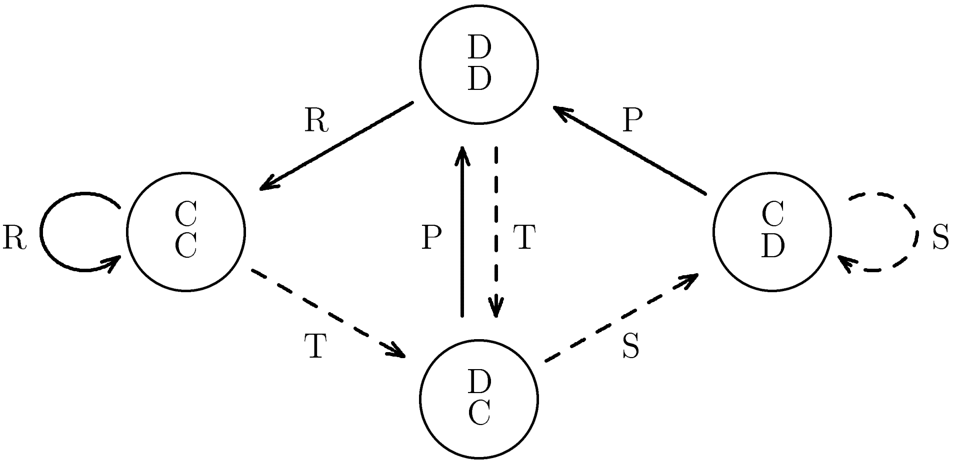

Figure 2: Pavlov is a norm if $2R > T + P$

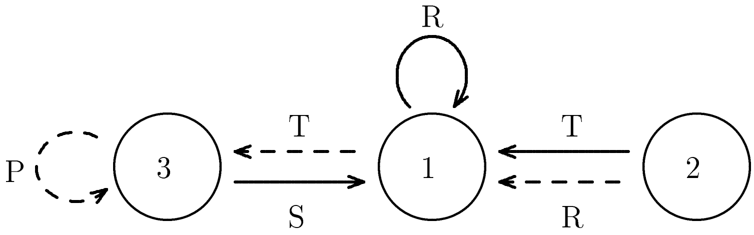

Figure 3: cTFT is a norm

This strategy becomes very transparent if one interprets it as Contrite Tit For Tat (cTFT, originally introduced by Sugden, 1986). It is based on the notion of a *standing* associated to each player, which can be g (good) or b (bad). In each round, the player acts (i.e. opts for **C** or **D**) and obtains a new standing which depends on that action and on the previous standing of both players. The rules for updating the standing are the following: if the co-player has been in good standing, or if both have been in bad standing, one receives a good standing if one cooperates, and a bad standing otherwise. If one has been in good standing and the co-player in bad standing, one receives a good standing no matter what one does.

Thus if one cooperates in a given round, one will always obtain a good standing: but if one defects, one will be in good standing only if the defection has been 'provoked' – i.e. if one has been in good standing and the opponent in bad standing.

cTFT is the strategy which cooperates except if the player is in good standing and the co-player is not. This means that the player defects when provoked, but not otherwise. A player who defects by mistake knows that he lost his good standing, and meekly accepts punishment, i.e. keeps cooperating even if the other player uses **D** on him.

In other words, cTFT begins with a cooperative move, and cooperates except if provoked (or by mistake). If two players using this strategy engage in a repeated Prisoner's Dilemma, and one player defects by mistake, then he loses his good standing. In the next round, he will cooperate, whereas the other player will defect without losing his good standing. From then on both players will be in good standing again and resume their mutual cooperation in the following round.

This is exactly the strategy implemented by the previously described automaton: the state 1 corresponds to both player being in the same standing, the state 2 occurs when the player is in good standing and the co-player not, and state 3 is just the mirror image of 2. Hence cTFT is a norm, and therefore uninvadable. Moreover, it is itself as adept at invading a population of defectors as TFT: it can only be suckered in the first round, and retaliates from then on. Moreover, it is immune against mistakes of implementation.

However, in contrast to Pavlov, cTFT is vulnerable to errors in perception. A player erroneously believing to have been suckered will play **D**. From then on, both cTFT players will remain in state 2 and keep punishing their co-player in good faith. Another weak point of cTFT is that it does not exploit *Always***C**-players, who therefore can spread by neutral drift and thereby open the door to defectors. In Sigmund et al (1998) it is argued that the investigation of strategies implemented by automata should take account of mistakes in perception just as well as of mistakes in implementation. Given that the number of norms is extremely large (see Lindgren, 1991, and Leimar, 1997), it is difficult to predict which strategies will eventually be selected.

5 Indirect Reciprocity

So far, we have described direct reciprocation. Alexander has suggested that another, indirect reciprocation is also operating in human societies, and that it is the basis of all moral systems. In fact, indirect reciprocity was mentioned by Trivers (1971) as 'generalised altruism': the return of an altruistic act is directed towards a third party. 'Individuals ... may respond to an altruistic act that benefits themselves by acting altruistically toward a third individual uninvolved in the initial interaction... In a system of strong multiparty interactions, it is possible that in some situations individuals are selected to demonstrate generalised altruistic tendencies.' This possibility is further stressed in Trivers (1985), who explores the possibility that a sense of fairness may evolve 'in species such as ours in which a system of multi-party altruism may operate such that an individual does not necessarily receive reciprocal benefit from the individual aided but may receive the return from third parties.'

Richard Alexander extended this idea under the heading of 'indirect reciprocity' (see Alexander, 1979 and 1986). With indirect reciprocity, one does not expect a return from the recipient (as with direct reciprocity), but from someone else. Cooperation is thereby channeled towards the cooperative members of the community. A donor provides help if the recipient is likely to help others, or at least if he has not been observed withholding help. According to Richard Alexander (1986), indirect reciprocity, which 'involves reputation and status, and results in everyone in the group continually being assessed and reassessed', plays an essential role in human societies. Alexander argues that systems of indirect reciprocity are the basis of moral systems. (For a dissenting opinion, see the message from the Pope, 1997).

Such scenarios have been considered by game theorists. In Boyd and Richerson (1989) it is assumed that individuals interact in loops such that a cooperative action can be returned, after several steps, to the original donor. According to Boyd and Richerson their model is unlikely to lead to a cooperative outcome, as it requires the loops to be relatively small, closed, and long-lasting. In Binmore's *Fun and Games'* (1992), the principle of indirect reciprocity is pithily resuméd as 'I won't scratch your back if you won't scratch their backs'. Binmore models this by imagining a world in which there are only two people alive at any stage, a mother and her daughter, with the daughter able to provide support to her mother.

The model considered by Nowak and Sigmund (1998a) is more in the tradition of evolutionary game theory. Consider a population of individuals having the options of helping another or not. In each generation, a number of potential donor-recipient pairs are chosen randomly: as before, this implies a cost c to the donor, if the help is actually provided, and a benefit b to the recipient. Furthermore, providing help increases the donor's score by one, whereas the score of a player refusing to help is decreased by one. (Note that the score has nothing to do with the payoff.) Initially all scores are zero. We consider strategies given by integers k; a player with such a strategy helps if and only if the score of the potential recipient is at least k. We can follow the frequencies of the strategies from generation to generation, allowing for occasional mutations.

A remarkably small number of interactions can lead to the emergence of cooperative populations where most members use $k = 0$ or $k = -1$ (for $b = 10$ and $c = 1$, an average of two interactions per lifetime suffices). If the simulation is continued, strategies which are less discriminating spread: players with $k = -5$, for instance, will rarely ever refuse to help, their score will therefore increase faster than average, and hence they will in turn be helped more often. But if the frequency of less discriminating players reaches a certain threshold, then defectors (players with $k = 5$, for instance, who practically never provide help) take over, so that cooperation disappears in the population. Once this happens, the average k-values will drop again, leading eventually back to a cooperative regime of players with maximal discrimination (i.e. $k = 0$).

To summarize, random drift can subvert populations of discriminate altruists by indiscriminate altruists; once their frequency is large, defectors can invade; but as soon as the defectors have reduced the proportion of indiscriminate altruists, the discriminate altruists can fight back and eliminate the defectors. This leads again to a cooperative population which is proof against defectors, but not against indiscriminate altruists, etc.

6 The Good, The Bad, and The Discriminating

In order to obtain an analytic understanding, we can further simplify the model (see Nowak and Sigmund, 1998b), so that only two scores are possible, namely g (for ‘good’) and b (for ‘bad’). Each player has two interactions per round, one as a donor and one as a recipient, against randomly chosen co-players. These two interactions are not with the same co-player. In fact, we may neglect the possiblility that two players are ever paired twice. A player has score g if (and only if) he has provided help in the last round. Let us consider a population with three types of strategies only: type 1, the indiscriminate altruists (with frequency x_1); type 2, the defectors (with frequency x_2) and type 3, the discriminate altruists (with frequency $x_3 = 1 - x_1 - x_2$). Furthermore, we assume that in the first round, discriminators assume that the co-player has score g. It is easy to see that $P_i(1)$, the payoff for type i in the first round, is given by

$$P_1(1) = -c + b(x_1 + x_3),$$

$$P_2(1) = b(x_1 + x_3),$$

and

$$P_3(1) = -c + b(x_1 + x_3).$$

In the n-th round (with $n > 1$) it is

$$P_1(n) = -c + b(x_1 + x_3),$$

$$P_2(n) = bx_1,$$

and

$$P_3(n) = (b-c)(x_1 + x_3^{n-1}x_2)/(x_1+x_2).$$

If there is only one round per generation, then defectors win, obviously. This is no longer the case if there are N rounds, with $N > 1$. The total payoffs $\hat{P}_i := P_i(1) + ... + P_i(N)$ are given by

$$\hat{P}_1 = N[-c + b(x_1 + x_3)],$$

$$\hat{P}_2 = Nbx_1 + bx_3,$$

$$\hat{P}_3 = N(b-c) + x_2[-b + \frac{b-c}{1-x_3}(1 + x_1 + ... + x_3^{N-1} - N)].$$

Let us now assume that the frequencies x_i of the three strategies evolve under the action of selection, with growth rates given by the difference between their payoff $\hat{P}_i$ and the average $\hat{P} = \sum x_i \hat{P}_i$. This yields the replicator equation

$$\dot{x}_i = x_i(\hat{P}_i - \hat{P})$$

on the unit simplex spanned by the three unit vectors $\mathbf{e_i}$ of the standard base. This equation has no fixed point with all $x_i > 0$, hence the three types cannot co-exist in the long run. The fixed points are the point $\mathbf{F_{23}}$ with $x_1 = 0$ and $x_3 + ... + x_3^{N-1} = c/(b-c)$, as well as all the points on the edge $\mathbf{e_1e_3}$. Hence in the absence of defectors, all mixtures of discriminate and indiscriminate altruists are fixed points.

The overall dynamics can be most easily described in the case $N = 2$ (see Fig. 4). The parallel to the edge $\mathbf{e_1e_2}$ through $\mathbf{F_{23}}$ is invariant. It consists of an orbit with ω-limit $\mathbf{F_{23}}$ and α-limit $\mathbf{F_{13}}$. This orbit l acts as a separatrix. All orbits on one side of l converge to $\mathbf{e_2}$. This means that if there are too few discriminating altruists, i.e. if $x_3 < c/(b-c)$, then defectors take over. On the other side of l, all orbits converge to the edge $\mathbf{e_1e_3}$. The limit point lies somewhere between $\mathbf{e_3}$ and $\mathbf{F}$, the point with $x_3 = 2c/b$. In this case, the defectors are eliminated, and a mixture of altruists gets established.

This leads to an interesting behaviour. Suppose that the society consists entirely of altruists. Depending on the frequency x_3 of discriminators, the state is given by a point on the fixed point edge $\mathbf{e_1e_3}$. We may expect that random drift makes the state fluctuate along this edge and that from time to time, mutation introduces a small quantity x_2 of defectors. What happens? If the state is between $\mathbf{F_{13}}$ and $\mathbf{e_1}$, the defectors will take over. If the state is between $\mathbf{e_3}$ and $\mathbf{F}$, they will immediately be selected against, and promptly vanish. But if a minority of defectors invades while the state is between $\mathbf{F}$ and $\mathbf{F_{13}}$, something strange happens. At first, the defectors thrive on the indiscriminate altruists and increase in frequency. But thereby, they deplete their resource, the indiscriminate altruists. After some time, the discriminate altruists take over and eliminate the defectors. The population returns to the edge $\mathbf{e_1e_3}$, but now somewhere between $\mathbf{e_1}$ and $\mathbf{F}$, where the ratio of discriminate to indiscriminate altruists is so large that defectors can no longer invade. The defectors have experienced a Pyrrhic victory. Their only hope is that fluctuations will eventually decrease the frequency of discriminators again. They have to wait until the state is between $\mathbf{F_{13}}$ and $\mathbf{e_1}$. For this, the fluctuations have to

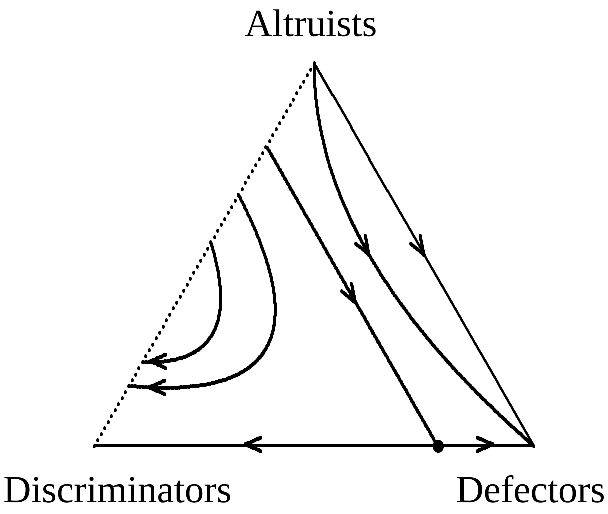

Figure 4: The population dynamics of discriminate and indiscriminate altruists and defectors (see text)

cross the gap between **F** and $\mathbf{F_{13}}$. This takes some time. If defectors try too often to invade, they will never succeed.

This behaviour corresponds precisely to the oscillations observed in the numerical simulations with the more sophisticated model (where the score k can take any integer value). Long periods of cooperation are terminated when the frequency of indiscriminate altruists becomes too large; this is immediately followed by a sharp increase in the number of defectors. Next, these defectors are eliminated by the discriminators, and then another long period of cooperation begins.

7 Know Your Partner

So far we have assumed that all players know each other's score. This is not realistic, of course. Even in small groups, where all members know each other well, there must be many interactions that are not watched by everyone. We must assume that players can have different scores in the eyes of different co-players. If we consider numerical simulations where an interaction between two individuals is observed by a random subset of the other individuals, and assume that only these 'on-lookers' have the possibility to update their perception of the donor's image score, the information is contained in a matrix whose elements s_{ij} denote the image score of player i as seen by player j. In a donor-recipient interaction between j and i, player j will cooperate if $s_{ij} \geq k_j$. If j has no information on i then $s_{ij} = 0$.

The model now depends on the probability that a given individual observes an interaction between two other individuals. We find again that cooperation can easily be established and dominate the population, but a larger number of interactions per generation is needed. For larger groups, it is more difficult to establish cooperation, because the fraction of individuals that obtain information about any particular interaction will be smaller. Therefore, more interactions are required (relative to group size) in order to discriminate against defectors.

We can investigate this analytically by extending the previous two-score model. Let us assume that in each round, a player can be donor or recipient with the same probability 1 / 2, and that q is the probability that a given individual knows the score of a randomly chosen opponent. A discriminator who does not know the score of the co-player will assume with probability 1 that this score is g. If g_n denotes the frequency of g-scorers in the population in round n, and $x_{1g}(n), x_{2g}(n)$ and $x_{3g}(n)$ are the frequencies of g-scoring indiscriminate altruists, unconditional defectors resp. discriminators in round n, then clearly $x_{1g}(n) = x_1$ and $x_{2g}(n) = (1/2)x_{2g}(n-1)$, since a defector is with probability 1/2 in the role of a donor and then unmasks himself. Therefore

$$x_{2,gG}(n) = \frac{x_2}{2^{n-1}} \; .$$

The score of a discriminator remains unchanged if he is a recipient. If he is a potential donor, he will either know the co-player (with probability q) and help if the co-player has score g (as happens with probability g_{n-1}), or else he will not know the co-players score, and help. Since this latter alternative holds with probability $1-q$, this yields

$$x_{3g}(n) = (1/2)x_{3g}(n-1) + (1/2)x_3(1 - q + qg_{n-1}).$$

Since $g_n = x_{1g}(n) + x_{2g}(n) + x_{3g}(n)$, it follows that

$$g_n = sg_{n-1} + \frac{1}{2}(x_1 + (1-q)x_3)$$

with $s = (1 + qx_3)/2$. This recurrence relation implies (together with $g_1 = 1$) that

$$g_n = (\frac{1+qx_3}{2})^{n-1}\frac{x_2}{1-qx_3} + \frac{x_1 + (1-q)x_3}{1-qx_3}.$$

The payoff for the indiscriminate altruists in round n is

$$P_1(n) = -(c/2) + (b/2)(x_1 + x_3).$$

The payoff $P_2(n)$ for the unconditional defectors depends on their score. Those with score b receive $b(x_1 + (1-q)x_3)/2$ and those with score g in addition $qbx_3/2$, so that

$$P_2(n) = (b/2)[x_1 + (1-q)x_3 + x_3q(x_{2g}(n)/x_2)] \; .$$

Finally, a discriminator receives $[-c(qg_n + 1 - q) + bx_1 + (1-q)bx_3)]/2$ if he has score b, and in addition $bqx_3/2$ if he has score g, so that we obtain

$$P_3(n) = -(c/2)(qg_n + 1 - q) + (b/2)(x_1 + x_3) - (b/2)qx_3[1 - (x_{3g}(n)/x_3)].$$

Instead of assuming a fixed number of rounds per generation, let us suppose that it is a random variable. If we assume for instance that $w < 1$ is the probability for a further round, then the total payoff for type i is given by $\hat{P}_i = \sum P_i w^n$. A straightforward computation shows that the phase portrait of the corresponding replicator equation $\dot{x}_i = x_i(\hat{P}_i - P)$ looks as before. The separatix l is now given by $x_3 = c(2-w)/bwq$. In one region, all orbits converge to $\mathbf{e_2}$, i.e. defectors take over.

In the other region, all orbits converge to a point on the edge $\mathbf{e_1 e_3}$ which depends on the initial value. In particular, altruism can only get established if

$$q > c/b,$$

which looks exactly like Hamilton's rule, except that the coefficient of relatedness is replaced by what we may call a coefficient of acquaintanceship, i.e. the probability that a player knows the score of the co-player. If this condition holds, then the average number of rounds per generation, i.e. $1/(1-w)$, must exceed $(bq+c)/(bq-c)$. It is only in this case that in a population consisting of defectors and discriminators, the discriminating strategy is stable.

8 Discussion

It should also be mentioned that this discriminator strategy is related to, but different from the T_1-strategy proposed by Sugden (1986). In Sugden's model, in each round a randomly chosen player needs help, and each of the other players can provide some help (thus the needy player can get as payoff $(m-1)b$, where m is the group size). Sugden's T_1- strategy is based on the concept of *good standing*, just as in the case of direct reciprocation. Players are born with a good standing, and keep it long as they help all needy players with good standing. If not, they lose their good standing. Sugden argues that such a strategy can work as a basis for an insurance principle within the population. We stress that a player can keep his good standing by refusing to help someone of bad standing, whereas in our model, he would lose his g-score whenever he refuses help, even if the potential recipient is a b-scorer. Sugden's T_1 strategy is more sophisticated, but like cTFT, it is vulnerable to errors in perception (see Boerlijst et al, 1997). If information is incomplete, then a player observed while witholding his help may be misunderstood; he may have defected on a player with good standing, or punished someone with bad standing. An eventual error can spread. The discriminator rule is less demanding on the player's mental capabilitites.

Moralistic aggression, social norms and feelings of fairness and solidarity have evolved in hominid groups in order to manage the complex social relations within their familes, bands and tribes (Ridley, 1996). Human decisions and activities are not only guided by economic considerations, but to a large extent by a complex household of emotions. These emotions and passions are themselves the outcome of biological necessities, in particular the need for division of labor and for cooperation. The way of life of the first hominids entering the savannah did not allow for autarchy, but required a highly adapted social life based on reciprocation and an internalized system of norms – a natural law in the fullest sense.

The importance of trading acts of mutual assistance has been recognised for a very long time – we may find it, for instance, in David Hume's book on *Human Nature* (1740): ‘I learn to do service to another, without bearing him any real kindness: because I foresee that he will return my service, in expectation of another of the same kind. I foresee that he will return my service, in expectation of another of the same kind, and in order to maintain the same correspondence of good offices with me *and others* [italics added]. And accordingly, after I have served him and he is in possession of the advantage arising from my action, he is induced to perform his part, as foreseeing the consequences of his refusal.'

An even older formulation of the principle of indirect reciprocation can be found in the biblical injunction 'Give and to you it shall be given'.

It may well be that indirect reciprocation works in other species. A hopeful candidate is the Arabian babbler. Zahavi (1995) mentions that individuals often compete with each other, jostling for the role of a donor. They interfere with the helping of others in feeding the nestlings, in allofeeding between adults, in sentinel activities, in mobbing and in the defence of the common territory. As Zahavi notes, this is difficult to interpret in terms of kin selection, group selection or (direct) reciprocation. If the beneficiary gets help from someone else, so much for the better, it would seem. But if helping is a means for raising one's score, competition for being the one who gives help makes perfect sense.

We cannot expect that direct and indirect reciprocation are always well separated. Alexander (1987) views indirect reciprocity 'as a consequence of direct reciprocity occurring in the presence of interested audiences – groups of individuals who continually evaluate the members of their society as possible future interactants'. The smooth transition from direct to indirect reciprocity is neatly captured in Pollock and Dugatkin (1992), who studied direct reciprocation in the usual context of the repeated Prisoner's Dilemma and allowed the players to occasionally observe a co-player before starting the repeated interaction. They analysed a strategy called Observer Tit For Tat (oTFT). If the future co-player was seen defecting in his last interaction, then oTFT prescribes to defect in the first round. Pollock and Dugatkin were mostly interesting in comparing this strategy with the usual TFT, and showed that it is not always advantageous. This seems surprising at a first glance – why should additional information be a handicap? The reason is that the conclusion drawn by oTFT is not always appropriate. A player last seen defecting can be a defector, but could also be a TFT player punishing a defector, in which case the new interaction starts with the wrong move. But Pollock and Dugatkin also remarked that oTFT may be better able to invade a resident population of defectors (viz. with a smaller initial cluster), and that it can be stable against defectors when TFT is not. They also stressed that the probability of future encounters among specific players engaged in an iterated Prisoner's Dilemma game may in principle decline to zero, i.e. that oTFT can hold its own against defectors when no degree of future interaction with the current partner was presumed. In this case, oTFT reduces precisely to the discriminator strategy. This shows that a gradual transition from direct to indirect reciprocation can lead to the establishment of a discriminating strategy based image scores.

The simple thought experiments performed by means of automata show that self-assembly can lead to norms enforcing cooperation, both through internalisation of fair and tolerant rules by means of inner states and through social pressure based on one's standing within the community.

References

Alexander, R.D. (1979), *Darwinism and Human Affairs*, Univ. Washington Press, Seattle, WA.

Alexander, R.D. (1987) *The Biology of Moral Systems*, Aldine de Gruyter, New York.

Axelrod, R. (1984) *The Evolution of Cooperation*, reprinted 1989 by Penguin, Harmondsworth.

Axelrod, R. (1997), *The Complexity of Evolution*, Princeton Studies in Complexity, Princeton UP.

Axelrod, R. and Hamilton, W.D. (1981), The Evolution of Cooperation, *Science* **211**, 1390-6.

Binmore, K.G. (1992), *Fun and Games: a Text on Game Theory*, Heath and Co, Lexington, Massachussetts.

Boerlijst, M., Nowak, M.A. and Sigmund, K. (1997), The Logic of Contrition, *Journal of Theoretical Biology* **185**, 281-293.

Boyd, R. and Richerson, P.J. (1989), The Evolution of Indirect Reciprocity, *Social Networks* **11**, 213-36.

Dawkins, R. (1976) *The Selfish Gene*, Oxford UP, second ed. 1989.

Hamilton, W.D. (1963), The Evolution of Altruistic Behaviour, *American Naturalist* **97**, 354-6.

Hofbauer, J. and Sigmund, K. (1998), *Evolutionary Games and Population Dynamics*, Cambridge UP.

Hume, D. (1740), *A Treatise of Human Nature*, reprinted in Penguin Classics 1985, Penguin, Harmondsworth.

John Paul II (1997), Message to the Pontifical Academy of Science, *Quarterly Review of Biology* **72**, 381-3.

Leimar, O. (1997), Repeated Games: A State Space Approach, *Journal of Theoretical Biology* **184**, 471-98.

Lindgren, K. (1991) Evolutionary phenomena in simple dynamics, in *Artificial life II* (ed. C.G. Langton et al), Santa Fe Institute for Studies in the Sciences of Complexity, Vol. X, 295-312.

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

Maynard Smith, J. (1982), *The Theory of Games and Evolution*, Cambridge UP.

Maynard Smith, J. and Szathmary, E. (1995), *The Major Transitions in Evolution*, Freeman, Oxford.

Nowak, M.A. and Sigmund, K. (1993), Win-stay, lose-shift outperforms tit-for-tat, *Nature* **364**, 56-8.

Nowak, M.A. and Sigmund, K. (1998a), Evolution of Indirect Reciprocity by Image Scoring, *Nature* **393**, 573-577.

Nowak, M.A. and Sigmund, K. (1998b), The Dynamics of Indirect Reciprocity, *Journal of Theoretical Biology* **194**, 561-574.

Pollock, G.B. and Dugatkin, L.A. (1992), Reciprocity and the Evolution of Reputation, *Journal of Theoretical Biology* **159**, 25-37.

Ridley, M. (1996), *The Origins of Virtue*, Penguin, Harmondsworth.

Selten, R. (1975), Re-examination of the Perfectness Concept for Equilibrium Points in Extensive Games, *International Journal of Game Theory* **4**, 25-55.

Sigmund, K. (1995), *Games of Life*, Penguin, Harmondsworth.

Sugden, R. (1986) *The Evolution of Rights, Co-operation and Welfare.* Blackwell, Oxford.

Trivers, R. (1971), The Evolution of Reciprocal Altruism, *Quarterly Review of Biology* **46**, 35-57.

Trivers, R. (1985), *Social Evolution*, Menlo Park. CA, Benjamin Cummings.

Weibull, J. (1995) *Evolutionary Game Theory*, MIT Press, Cambridge, MA.

Wilson, D.S. and Sober, E. (1994), Re-introducing Group Selection to Human Behavioural Sciences, *Behavioural and Brain Sciences* **17**, 585-654.

Zahavi, A. (1995), Altruism as a Handicap – The Limitations of Kin Selection and Reciprocity, *Journal of Avian Biology* **26**, 1-3.