Evolutionary and spatial dynamics of the prisoner's dilemma

Martin Nowak

Department of Zoology, University of Oxford

South Parks Road, OX1 3PS, Oxford, U.K.

Tel: (0865) 271197, Fax: (0865) 310447

Email: nowak@uk.ac.ox.vax

Abstract

The prisoner's dilemma (PD) is a well known metaphor for the evolution of cooperation in populations of selfish individuals. In the simple (non-repeated) PD it is always best to defect, but cooperation becomes a promising option if the game is repeated (i.e. if the individual players follow certain strategies, can recognise other players and remember past encounters) or if spatial effects are included (here one can neglect any strategical elaborations or memories).

The paper consists of two parts. The first deals with the evolutionary dynamics of the iterated PD. Very simple strategies can lead to a complex, chaotic dynamics. In heterogeneous populations of probabilistic (erroneous) strategies tit-for-tat (TFT) can catalyse the emergence of cooperation, but is usually replaced by the more forgiving generous-tit-for-tat (GTFT).

In the second part of the paper it is shown how spatial structures can lead to coexistence between cooperators and defectors in a purely non-repeated PD. Such spatial games can give rise to an enormous complexity of chaotically changing, fractal-like patterns. Spatial games lead to a series of new and interesting problems and can defy intuition.

1. Introduction

The Prisoner's Dilemma (PD) is a two-player game: each player can opt for one of the two strategies C (to cooperate) or D (to defect). If both players cooperate, their payoff R is higher than the payoff P for joint defection. But a player defecting unilaterally obtains a payoff T which is larger than R, while the opponent ends up with a payoff S smaller than P. In addition to this rank ordering, one usually also assumes $2R < S+T$. The rational decision, in this game, is to play D, since this yields the higher payoff no matter whether the opponent uses C or D. As a result, both players defect and earn P instead of the larger reward R for joint cooperation.

If the probability that the players repeat the interaction is sufficiently high, there is no longer a single best strategy for this 'iterated prisoner's dilemma' (IPD). But a series of computer tournaments by Axelrod (1984) established the success of a remarkably simple strategy, TFT (Tit For Tat), which consists in playing C in the first round and from then on repeating whatever the adversary did in the previous round. This led Axelrod and Hamilton (1981) to use the IPD for explaining the evolution of cooperation in biological interactions on the basis of reciprocity. This approach has proved to be extremely fruitful. It is not the only paradigm, but certainly the most current in the field (May 1987, Axelrod & Dion 1988, Dugatkin et al 1992).

While reciprocal interactions abound in nature, it is usually difficult to find clear-cut empirical evidence for the implementation of the TFT-strategy (Wilkinson 1984, Milinski 1987, Reboreda & Kacelnik 1991). Furthermore, the uncertainties and mistakes surrounding most biological interactions penalize TFT-populations, since any accidental defection entails a series of alternating retaliations. This vulnerability to errors is not evident under the clinical conditions of computer runs. In applications, however, it should not be overlooked. Moreover, a TFT-population can be subverted, through random drift, by unconditional cooperators who in turn can be invaded by defectors.

Sections 2 and 3 deal with the evolutionary dynamics of the iterated PD and contain summaries of Nowak & Sigmund (1992, 1993). Section 5 gives a brief overview of the spatial PD (Nowak & May 1992, 1993).

2. Tit-for-tat catalyses cooperation

A reactive strategy is given by two probabilities p and q to cooperate after a C (resp. D) of the other player. TFT corresponds to $(1,0)$, $AllD$ to $(0,0)$, $AllC$ to $(1,1)$, and so on. We shall be mostly interested in properly stochastic strategies

where the values of the probabilities are strictly larger than 0 and smaller than 1. For the sake of simplicity, we consider only the infinitely IPD. In this case, the initial move plays no role, since its effect will be 'forgotten' in the long run.

We use the usual game dynamics to study the evolution of representative samples of reactive strategies. If x_i denotes the freqency of strategy i in one generation, then its frequency in the next generation x_i' will be given by

$$x_i' = x_i f_i(x)/\bar{f}. \qquad i = 1, 2, .., n$$

Here $f_i(x)$ denotes the average payoff for strategy i in a population where the frequencies of the strategies are given by $x = (x_1, ..., x_n)$, and $\bar{f} = \sum x_i f_i(x)$ is the average payoff in the population. Initially, all strategies are equally represented.

If we choose $n = 100$ different reactive strategies from a uniform distribution on the unit square, then the evolution proceeds in most cases towards the vicinity of $AllD$: those (p, q)-strategies from the sample which are closest to $(0, 0)$ increase in frequency from generation to generation, while all other strategies vanish. This is an obvious consequence of the fact that a large percentage of the random sample has high q-values and therefore does not relataliate against exploiters. With such a rich diet of 'Suckers', it pays to defect.

The outcome changes drastically if one of the initial strategies is added by hand or by chance TFT, or a close neighbour of it. The first 100 generations are practically indistinguishable from the previous run. The strategies in the vicinity of $AllD$ grow very fast. In our plot, where strategies with very low frequencies are no longer visible (although they are still present in the numerical computations), TFT and all other *reciprocating* strategies (with values close to $(1, 0)$) seem to have disappeared. But an embattled minority is still present, and fights back. The tide turns when (after some 150 generations in our simulations) the 'Suckers' with high q-values are decimated to such a point that exploiters can no longer feed on them. Slowly at first, but with ever increasing momentum, the reciprocators come back. It is the exploiters turn, now, to vanish from the stage. But the TFT-like strategy which has been added, and which caused this reversal of fortune, is not going to profit from it: after having eliminated the exploiters, it is superseded by the strategy closest to $GTFT$, which is defined by $(1, q)$, where $q = min\{1 - (T - R)/(R - S), (R - P)/(T - P)\}$.

These simulations exhibit very clearly that TFT acts as a catalyser for the emergence of cooperation. It is essential for the reaction toward cooperation to get going. It needs to be present, initially, in a very small concentration only; in the intermediate phase of the reaction, its concentration is high; but in the end, only a trace remains.

3. Chaos and cooperation

Next we consider strategies which are entirely specified by the outcome of the previous round. Such strategies can be described by a quadruple of four parameters, (p_1, p_2, p_3, p_4), which denote the probabilities to cooperate after receiving payoff R, S, T, and respectively P, in the previous round. For example $(0,0,0,0)$ is the deterministic rule that always defects $(AllD)$, while $(1,0,1,0)$ plays TFT. The rule $(1,0,0,0)$ plays $GRIM$: after a single D of the adversary, it never reverts to C again. The rule $(1,0,0,1)$ cooperates whenever both players choose the same action in the previous round. It fares poorly against $AllD$, since it reverts each second round to C. For this reason, it has been called 'simpleton' by Rapaport and Chammah (1965). We think that this appelation is not entirely deserved; following Kraines and Kraines (1988), we prefer to call it $PAVLOV$, since it responds to positive and negative conditioning (switching its behaviour whenever one round's payoff is lower than R). There are 16 deterministic rules altogether, which we number from 0 to 15 (the i-th quadruple being the binary expression for i). The strategy corresponding to rule i will be denoted by S_i. Thus S_0 is $AllD$, S_9 is $PAVLOV$. S_{10} is TFT, and S_{15} is $AllC$. The S_i strategies are exactly the 16 corner points of the four dimensional strategy space formed by all (p_1, p_2, p_3, p_4) strategies.

We shall now take uncertainty into account by replacing in the quadruples 1 by $1 - \epsilon$ and 0 by ϵ. The small probability ϵ describes the frequency of errors. If $\epsilon > 0$ the first round no longer matters. The total payoff can be defined as the limit of the mean payoff per round. Among the 16 S_i strategies $AllD$ and $GRIM$ are the only evolutionarily stable strategies (ESS). There are 3 strategies that receive a payoff very close to full cooperation $(R = 3)$ when playing against themselves, these are $AllC$, (1110) and $PAVLOV$.

We now consider a large population of players using the strategies S_0 to S_{15}. By x_i, we denote the frequency of S_i in a given generation. In each generation all the strategies play the infinitely IPD among each other (subject to a small error frequency ϵ). It is easy to compute the average payoff f_i for an S_i-player (which depends on the composition of the population). The evolutionary dynamics map the frequencies x_i after one generation into x_i' according to the following rule: first selection provides each S_i-strategist with a number of offspring proportional to its expected payoff f_i (the higher the payoff, the more offspring). To this is added a tiny number of invaders, u, which may be caused by mutation. This yields a deterministic recurrence equation for the frequencies of the strategies:

$$x_i' = \left(\frac{x_i f_i}{\sum x_i f_i} + u \right) / (1 + nu). \qquad i = 1, .., n$$

Here n denotes the total number of strategies in the population.

This modification of the usual game dynamics allows for recurrent and simultaneous invasion attempts. The resulting dynamics can exhibit complicated periodic and even chaotic orbits. The strategies S_4, S_6, S_7 and S_{12} are almost driven to extinction, but the other strategies and the total payoff for the population display violent oscillations (with large amplitudes for strategies $S_0, S_1, S_8, S_9, S_{10}$ and S_{11}). For small invasion rates, the minima of their frequencies are very close to 0, except for the TFT-like strategy S_{10}, which is best protected against extinction and is in this sense the "safest bet". But whenever the proportion of TFT-players is large, they are superseded by the more generous strategy S_{11} (whose transition rule $(1, 0, 1, 1)$ forgives a defection by the other player if it was matched by an own defection) and $PAVLOV$ (S_9). The S_{11} and $PAVLOV$ population, in turn, is invaded by the parasitic S_1 (which cooperates only if its defection has met with instant chastisment). This paves the way for the strategies close to $AllD$ (S_0) and to $GRIM$ (S_8), which in turn leads to the resurgence of TFT. This is the main cycle in the selective mechanics: but the other strategies introduce the twists leading to chaos. Simple strategies in the iterated prisoner's dilemma can lead to very complicated evolutionary dynamics.

4. Spatial dilemmas

Let us imagine that populations are not completely homogeneous but have certain spatial structures. Let us imagine that animals or molecules are located at certain positions (territories, patches, pixels or cells) and are more likely to interact with their neighbours (= individuals at a closer distance). Therefore the probability that a certain phenotype, A, interacts with another phenotype, B, is no longer just the product of their relative frequencies in the population, but depends in a non- trivial way on the spatial structure of the population.

Let us consider two different strategies C (for cooperation) and D (for defection). If two cooperators interact both receive 1 point. If a defector 'exploits' a cooperator, the defector receives the payoff b and the cooperator 0. The interaction between two defectors also leads to the zero payoff. This game is designed to keep things as simple as possible, without changing the essential properties of the prisoner's dilemma. In fact there is only one parameter, b, the advantage for defectors.

This game is now played on a two dimensional square lattice. Each position is occupied either by a cooperator or a defector. In each generation the payoff of a certain individual is the sum over all interactions with the 8 nearest neighbors (the cells corresponding to the chess king's move) and with its own site. It seems

reasonable to include this self interaction, if one assumes that several animals (a family) or molecules may occupy a single patch. But the general properties of the game do not depend on this assumption, and we have also explored the situation without self interaction.

In the next generation an individual cell is occupied with the strategy that received the highest payoff among all the 8 immediate neighbours of the cell and, of course, the cell itself. Thus whatever happens to a cell depends on the state of the cell, the 8 neighbours and their neighbours. These are altogether 25 different cells. In the terminology of the cellular automata literature our simple game is characterised by a transition matrix with 2^{25} different rules. That is, in cellular automata terms the rules are very complex; our underlying biological 'game', however, enables the rules to be stated very simply.

The rules of our game are now completely defined. The game is deterministic. The outcome depends on the initial configuration and the magnitude of the parameter b.

There are three classes of parameter regions.
(i) If $b < 1.8$ then only C clusters keep growing.
(ii) If $b > 2$ then only D clusters keep growing.
(iii) If $1.8 < b < 2$ then both C and D clusters keep growing.

For $b = 1.15$ almost all cells are occupied by cooperators. Defectors occur either in single isolated cells, which oscillate between $1D$ and $9D$, or in stable unconnected short lines. For $b = 1.35$ the lines of defectors become connected. The basic oscillators are again single defectors , but the end of lines can oscillate, too. These oscillators are generally of period 2. For $b = 1.55$ there are long connected lines and whole lines can oscillate (usually with period 2). Single defectors now oscillate with period 3 (a $1D \rightarrow 9D \rightarrow 5D \rightarrow 1D$ - oscillator). The interaction between large structures can lead to oscillators with very high period. Things are different for $b = 1.79$. Here the pattern is almost completely static. An irregular network of defectors runs over the whole area. For $b = 1.85$ the structure appears to be completely chaotic. There are about 31% cooperators. A large fraction of cells is changing from one generation to the next. The world is covered with defectors , but cooperators exist in many small clusters. These clusters have the tendency to grow. But whenever two such clusters come too close the defectors between them get high payoffs and start to grow. The cooperators win along straight lines, the defectors win along irregular boundaries. The result is a dynamic equilibrium. It is an always changing, but dynamically stable dimorphism. For $b = 2.01$ another static pattern is observed.

An interesting sequence of patterns emerges if a single defector invades a world of cooperators in the parameter region $1.8 < b < 2$. In generation $t = 0$ we start with one defector. This defector first grows to form a 3×3 and then a 5×5 square of defectors. The payoff for the defectors at the 4 corners of this square is $5b$ (which is larger then 9). The payoffs for the defectors along the edges of the square is $3b$ (which is smaller than 6). Therefore the defectors gain at the corners but loose along the lines. The result is a interesting and beautiful growth pattern. We can study these patterns in a finite world (with fixed or cyclic boundary conditions) or in an – effectively– infinite world.

When a single defector invades an infinitely large array of cooperators, we find fractal like structures that repeat themseleves. The whole D-structure takes a square shaped form always at the generations which are the powers of 2. For $t = 0, 1, 2, 4, 8$ and 16 these squares only consist of defectors. For $t = 32$ there are 8 clusters of cooperators left over, but they disappear in the next generation. For higher generations the squares contain many clusters of cooperators, which can persist. The frequency of cooperators within the growing square-like structure of defectors converges to $x \approx 0.318$ which is the same value as in the simulations with random initial conditions. A handwaving, but interseting approximation leads to $x \approx 12 \log 2 - 8 = 0.3178..$ (which is in excellent agreement with the observed value).

Spatial games can be played on regular and random grids, in or 3 dimensions the transition rules can be deterministic (as above) or stochastic (for example by giving a cell in the next generation to one of its neighbours with a probability proportional to this neighbour's payoff). Time can be measured in effectively discrete generations (with an interaction phase followed by a reproduction phase – as above) or continuously (where interaction and reproduction occur simultaneously – see Huberman 1993). In these different approaches we find that cooperators and defectors can coexist under a variety of conditions (Nowak, Bonhoeffer, May 1993).

For many kinds of 'evolutionary games', we find that the outcome of frequency dependent selection can be changed (and reversed) if spatial effects are taken into account. The simplest deterministic models of frequency dependent selection can lead to a rich variety of spatial and temporal dynamics. The success of a given strategy depends on the spatial structure of the population. Selection works on spatial structures.

One of the appealing aesthetic elements in the present approach is the combination between temporal chaos (the unpredictable oscillations of the frequencies of the individual strategies) and highly symmetric fractals. The kaleidoscopes of

spatial evolutionary games have many applications in the 'real world': there is a new industry for tiles, carpets, T-shirts, rose windows and lace doilies.

References

Axelrod R. (1984) *The Evolution of Cooperation*, (Basic Books, New York).

R. Axelrod and W.D. Hamilton (1981) *Science* **211** 1390.

R. Axelrod and D. Dion (1988) *Science* **242**, 1385.

L.A. Dugatkin, M. Mesterton-Gibbons and A.I. Houston (1992) *TREE* **7**, 204-205.

D. Kraines and V. Kraines (1988) *Theory and Decision* **26**, 47.

R.M. May (1987) *Nature* **327**, 15-17.

M. Milinski (1987) *Nature* **325**, 433.

M A Nowak , R M May (1992) Nature 359, 826-829

M A Nowak, R M May (1993) Int J Bifurcation and Chaos, in press

M A Nowak , K Sigmund (1992) Nature 355, 250-253.

M Nowak, K Sigmund (1993) Proc Nat Acad Sciences USA, in press

A. Rapoport and A.M. Chammah (1965) *Prisoner's Dilemma* (Univ. of Michigan Press, Ann Arbor).

J. Reboreda and A. Kacelnik (1991) *Anim. Behav.* 40,1188.

G. Wilkinson (1984) *Nature* **308** 181.