

Evolutionary escape from the prisoner's dilemma

Lee Worden^{a,b,*}, Simon A. Levin^{a,c}

^aProgram in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544-1000, USA

^bEnergy and Resources Group, University of California, Berkeley, CA 94720-3050, USA

^cDepartment of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544-1003, USA

Received 23 February 2006; received in revised form 4 October 2006; accepted 9 October 2006

Available online 18 October 2006

Abstract

The classic prisoner's dilemma model of game theory is modified by introducing occasional variations on the options available to players. Mutation and selection of game options reliably change the game matrix, gradually, from a prisoner's dilemma game into a byproduct mutualism one, in which cooperation is stable, and "temptation to defect" is replaced by temptation to cooperate. This result suggests that when there are many different potential ways of interacting, exploring those possibilities may make escape from prisoner's dilemmas a common outcome in the world. A consequence is that persistent prisoner's dilemma structures may be less common than one might otherwise expect.

© 2006 Published by Elsevier Ltd.

Keywords: Prisoner's dilemma; Evolutionary game theory; Adaptive dynamics; Evolution of cooperation; Byproduct cooperation; Byproduct mutualism; Division of labor game; Tragedy of the commons; Cultural evolution; Agent-based models; Agent-based simulation; Complex systems

I'm not very happy with most of the applications of games theory, because it tends to perpetuate the rules of the game as perceived at a given moment by the players ...Nobody knows a thing about changing the rules of the game. —Gregory Bateson (Brand, 1974)

1. Cooperation and altruism

There is a substantial and lively literature on "the evolution of cooperation," and "cooperation" in general, which uses evolutionary game theory to study tensions between individual self-interest and acting for the greater good. This research typically focuses on the famous prisoner's dilemma game, and tends to use the word "cooperation" as a synonym for altruism, which is cooperation that cannot be explained by direct self-interest because one can benefit more by not cooperating (Axelrod, 1984; Sober and Wilson, 1998; Leigh, 2001).

*Corresponding author. Energy and Resources Group, University of California, Berkeley, CA 94720-3050, USA. Tel.: +1 510 642 8779; fax: +1 510 642 1085.

E-mail address: worden@berkeley.edu (L. Worden).

In the study of biological mutualism—cooperation between individuals of different species, such as plants and pollinators, hosts and endosymbionts, and so forth—while attention is certainly given to the problem of "cheating" on the exchange of favors between species, it is also well accepted that helping behavior need not have an associated cost and may also be the behavior that most benefits the helping individual (Connor, 1995b for instance). When this is the case, there is no temptation to cheat and no need to explain how cooperation is maintained.

The coincidence of selfish and cooperative behavior is certainly possible in within-species interactions as well (Connor, 1995b; Dugatkin et al., 1992), and most researchers in this field recognize this possibility and take care to distinguish cooperation from altruism. Even so, the appearance of the word "cooperation" in a biology journal has become a sure sign of a paper on the evolution of altruism, and some writers explicitly conflate the two ("This is a study of the problem of collective action in social contexts, which is the Prisoner's Dilemma writ large" (Hardin, 1982 p. xiii); "We know that any form of mutual aid is threatened by exploitation" (Sigmund and Nowak, 1997)).

The widespread conflation of cooperation and altruism in the study of behavior reinforces a widely shared view of the world in which cooperation is a mysterious anomaly whose existence is difficult to explain because of the universal temptation to defect. Prisoner's dilemma and tragedy of the commons scenarios, which embody these assumptions, appear to the lay reader or student as authoritative scientific statements legitimizing a bleak Hobbesian picture of humans as selfish, greedy individualists whose antisocial tendencies must be kept in check by coercive social forces (Picardi and Seifert, 1976; Evans and Zarate, 1999; Whitfield, 2002 for example). This characterization is not justified by these models, which ignore the range of non-costly cooperation situations, thus obscuring the possibility of easy, stable mutualistic behavior.

This bias has been pointed out and is being addressed by some researchers, making the distinction between byproduct mutualism and altruism (Connor, 1995a; Dugatkin et al., 1992; Hauert et al., 2006; Wright, 1980), and modeling particular means by which stable cooperative interactions can arise among organisms (Clutton-Brock, 2002; Mesterton-Gibbons, 1991). In this paper we model a different route to cooperation, in which by persistently “tinkering” with strategies of interaction, organisms can transform mutual defection into stable mutualistic cooperation.

2. The prisoner's dilemma

The prisoner's dilemma (Axelrod, 1984; Rapoport and Chammah, 1985) is the *Drosophila* of costly cooperation models, the center of a massive research program modeling it in any number of settings and variations in search of techniques for maintaining cooperation in the face of temptation to defect. The story from which it takes its name is of two prisoners who are being questioned separately by police about a crime (Luce and Raiffa, 1957). Each has two options, to confess or to keep quiet. If both prisoners keep quiet (“cooperate”), both will go free since the police have no evidence. If both confess and implicate one another (“defect”), both will receive a reduced sentence for cooperating with the police. If one prisoner agrees to provide incriminating evidence while the other remains silent, the first will be set free and rewarded, while the other will suffer the full wrath of the law. By manipulating rewards and punishments in this way, the police create a situation in which each player benefits more from defecting than from cooperating, whether the other chooses to cooperate or not. If one prisoner confesses, the other will do better to confess as well; while if she does not, the other can go free either way, but can also collect the payoff for collaborating if he turns her in. As a result, when the prisoner's dilemma is made into an evolutionary game theory model, mutual defection is globally stable (a convergence stable ESS) and mutual cooperation is unstable (Boyd and Lorberbaum, 1987; Maynard Smith, 1982). By seeking to maximize their own rewards, both prisoners betray their own interests and condemn themselves to long jail sentences unnecessarily.

The prisoner's dilemma is considered to be a succinct summary of the problem of how humans and other organisms can be induced to help each other in the face of temptations to seek individual benefit at the expense of collective well-being. When expanded into an n -player social dilemma, it takes the form of the tragedy of the commons scenario (Hardin, 1971), which represents the problem of managing a collectively held resource or other public good, such as a fishery or the world's atmosphere, when individuals can profit by overexploiting it at the expense of sustainability. Since human beings and other organisms “cooperate” continually and promiscuously, and sometimes even behave altruistically, the failure of cooperation in the prisoner's dilemma is seen as a problem to be solved, to explain how cooperation—or rather altruism—is sustained (May, 1981).

The most famous solution to this “paradox” is to model the prisoner's dilemma as a repeated game in which players can make their choice to cooperate or defect based on what the other player did previously. In this scenario, the famous Tit-for-Tat strategy, invented by psychologist Anatol Rapoport, which simply returns whatever the other player did previously, and variants are very successful (Axelrod, 1984), suggesting the importance of “reciprocal altruism” (Trivers, 1971) in stabilizing cooperation. Cooperation can also be maintained by spatial locality (Axelrod, 1984; Nowak and May, 1992; Oliphant, 1994), kin selection (Wade and Breden, 1980), group selection (Boyd and Richerson, 1990, 2002), mutual recognition based on arbitrary characteristics (Holland, 1993), or various other possible mechanisms that allow players to cooperate preferentially with others who are likely to have similar strategies to themselves (Cohen et al., 1999; Pepper and Smuts, 2002). Additionally, punishment (Boyd and Richerson, 1992; Henrich and Boyd, 2001), conformity (Boyd and Richerson, 1985), and even simple communication (Miller et al., 1998) can stabilize cooperation in these scenarios.

3. Non-dilemmas and transformations

The two major alternatives to the prisoner's dilemma structure that are currently receiving attention are the snowdrift game, and byproduct mutualism. In the snowdrift, or hawk-dove, game (Maynard Smith and Price, 1973; Sugden, 1986) defection is favorable when cooperation is the norm, and cooperation is favorable when defection is the norm, so that while defection is not expected to replace cooperation, it is always likely to persist in the community. In byproduct mutualism (Connor, 1995b; Dugatkin et al., 1992), cooperation is the only stable strategy, since defection is always an inferior option even for the defector. Byproduct mutualism differs from the other two scenarios in that “temptation to defect” is absent when cooperation is the norm, making cooperation stable and non-paradoxical.

Beyond comparing different cooperation scenarios, the question of how one scenario might transform into another

is much more sparsely investigated. One relevant contribution is a model of reproducing insects presented by Mesterton-Gibbons (1991). These insects, looking for sites to lay their eggs, have to choose whether to lay in sites where an egg is already placed, which is an action that hurts the chances of both eggs but gives, of course, a better chance than not laying an egg. If the insects are capable of distinguishing their own eggs from those laid by others, then the problem has the structure of the prisoner's dilemma, and insects are expected to defect by aggressively laying eggs on sites others have used, unless reciprocity or another such mechanism intervenes to make cooperation stable. However, if insects are not capable of identifying their eggs, then as long as the survival probability of a solitary egg is modestly superior to that of an egg sharing a site with another, defection is no longer a favorable option and evolutionary game theory predicts that the insects will cooperate by sticking to unused sites. Thus evolution on some scale could conceivably remove the dilemma by limiting the organisms' senses, though that question is not addressed directly in the paper.

Turner and Chao (2003) report an in vitro experiment in which an RNA phage can develop a variant that reproduces prolifically in the host, causing a fitness loss for all phages present. This situation has the prisoner's dilemma structure, and the variant (defector) strain replaces the other. However, this is only true when a relatively large number of phages are allowed to infect a single host. When multiple infection is not possible, cooperator phages predominate. Further, the cooperator phages that are evolved in the absence of multiple infection are capable of coexisting with the defector phages evolved in multiple infection circumstances, unlike the cooperators evolved with the defectors. This suggests, as the authors propose, that these objects may be able to evade the dilemma by limiting coinfection.

We approach the problem in a different way, not by demonstrating the possibility of conversion of a game structure in a particular type of organism, but by investigating a general scenario in which there are many different potential ways to interact, and whatever one is currently doing can always be done a little differently. By exploring variants—"tinkering" with the current arrangement—over the long term the dilemma structure is reliably abandoned for stable cooperation.

4. Game theory and dynamics

Mathematically, the two-player prisoner's dilemma is represented as a game with a pay off matrix summarizing the outcomes of the four combinations of choices to cooperate or defect:

$$\begin{array}{c} C \quad D \\ C \begin{pmatrix} R & S \end{pmatrix} \\ D \begin{pmatrix} T & P \end{pmatrix} \end{array}$$

The entries of this matrix represent the payoff to the player whose choice appears in the left-hand column, given the other player's choice in the top row. The prisoner's dilemma is characterized by the fact that the "reward" R exceeds the "punishment" P , while the "temptation to defect" T is greater than the reward, and the "sucker's payoff" for the cooperator who is betrayed, S , is even worse than the punishment: $T > R > P > S$. This matrix encapsulates the paradox of the prisoner's dilemma, since the entries in the D row strictly dominate those above them, making defection always better rewarded, while mutual defection leads to the inferior outcome P .

In this paper we model players who play repeatedly with each other, adjusting their behavior by learning what choices yield higher payoffs at each moment. Individuals in this model do not evolve, die or reproduce. Only their behaviors change, as the same players remain, and continually re-evaluate their strategies. The learning model we use is based on one introduced by biologist Calvin B. Harley (1981), who showed that if there is an evolutionarily stable strategy, that will be the strategy learned by his rule.

Given two or more possible choices, or plays, for instance C and D , each game player's current strategy is a vector of probabilities, adding up to one, describing how likely the player is to choose each play. In each round, each player plays with a randomly chosen partner, and each player takes the payoff received as a reward or punishment for the play that was chosen, increasing or decreasing the future likelihood of making that choice. The details of this process and the rest of the simulation implementation are listed in an Appendix, below.

As this process is iterated, players learn to favor whichever plays are likely to yield the greatest payoff given the other players' strategies. Since the expected payoff associated with each play changes as all the other players change their strategies, the population of players may or may not converge on an optimum.

Typically, a few plays will be in use at a given time, and others will be abandoned by all the players. These plays' probabilities are constrained to remain a little above zero—in the simulations shown, at 0.001 or more. This keeps a "library" of abandoned plays, so that they are occasionally retested and can be re-adopted later, if conditions become favorable.

5. Equilibration of game strategies

All the simulations we present here are begun with a prisoner's dilemma payoff matrix whose payoffs are

$$\begin{array}{c} C \quad D \\ C \begin{pmatrix} 3 & -2 \end{pmatrix} \\ D \begin{pmatrix} 4 & -1 \end{pmatrix} \end{array}$$

Ten players are simulated, with $c = 0.01$. All players begin with 90% probability of choosing the C play.

Fig. 1 is a graph over time of which entries of the payoff matrix are in active use, in a representative simulation run using the above game matrix. The diagonal terms R and P (3 and -1) are shown in black and the off-diagonal terms T and S (4 and -2) are shown in red. Only payoffs involving plays that occur more than 5% of the time are shown. As the game players learn that cooperation does not pay as well as defection, they shift their strategies to pure defection. As the probability that a player will choose the C play declines to zero, the red lines at -2 and 4 and the black line at 3 disappear, since $C-D$ and $C-C$ interactions become rare, so that a single black line is left at -1 , the payoff given to each of a pair of defectors.

Fig. 2 displays the probability of cooperation over time in the same simulation. In this figure, this is how often both players simultaneously choose C . In the general case, we define probability of cooperation in our figures to mean the fraction of games played in which both players receive a positive payoff.

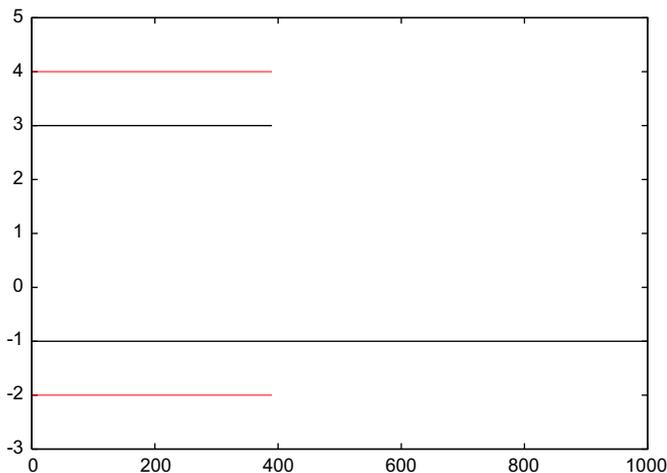


Fig. 1. Changes in the ensemble of payoffs over time in the standard prisoner's dilemma. Diagonal payoff matrix entries appear in black, off-diagonal entries in red.

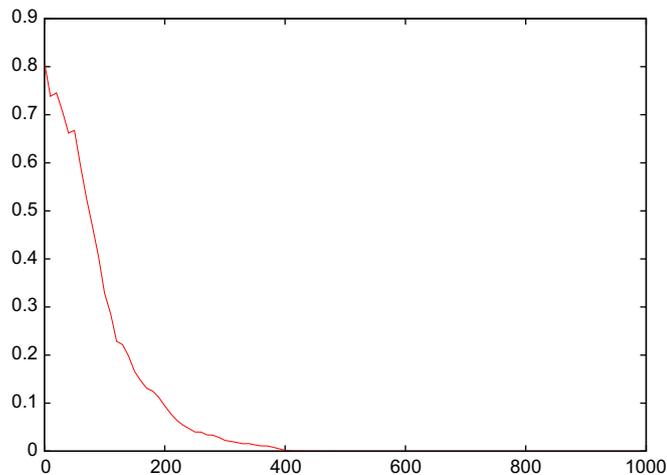


Fig. 2. Probability of cooperation vs. time in the standard prisoner's dilemma.

6. Evolution of behaviors and their interactions

As we have described it so far, this iterated learning model is not especially new or interesting. When it is applied to the prisoner's dilemma, players behave selfishly and converge on universal mutual defection. What we now add to this model is the possibility of long-term change to the game itself, represented by the payoff matrix.

The learning model, in which the frequency of each behavior in the population rises or falls as players learn, settling on an equilibrium or fluctuating attractor, with some behaviors possibly going extinct, can be thought of as the "ecology" of the behaviors. As in an evolutionary ecology project, the ecological model can be paired with an evolutionary model that introduces variant behaviors.

Here, we use an evolutionary model very similar to one we have used elsewhere in connection with ecological models (Worden, 2003). The entries of the payoff matrix, which characterize each play in terms of how it interacts with all the other plays, are perturbed by small numbers to represent the introduction of a new, previously undiscovered play, which is a slight variant of one of the existing ones. For instance, given the standard prisoner's dilemma options C and D , if a variation on the D play is introduced, the resulting payoff matrix is

$$\begin{pmatrix} R & S & S + \Delta p_{13} \\ T & P & P + \Delta p_{23} \\ T + \Delta p_{31} & P + \Delta p_{32} & P + \Delta p_{33} \end{pmatrix},$$

with each Δp value chosen independently from a continuous distribution of small values centered on zero.

This modeling choice is a version of the conventional "trembling hand" assumption. Players follow their chosen strategies almost all of the time, but occasionally make unpredictable errors. Here the "error" is not a random choice from the assigned list, but a behavior a little different from those already known to the players, with slightly different consequences.

Periodically, when the learning dynamics have unfolded for a long time, a variant play is introduced in this way, with a small probability of use by each player, and the players' learning dynamics then determine whether the new play is adopted and whether older plays are abandoned. On a slower time-scale than the learning dynamics, the set of plays currently in use undergoes evolutionary change, and the payoff matrix undergoes gradual qualitative changes.

7. Conversion of defection to cooperation

Fig. 3 presents the results of six simulation runs chosen at random. These runs are not selected to have a particular outcome.

Like Fig. 1, this figure shows only the payoffs for choices that are in use at least 5% of the time. All the choices ever generated are in play in the system, but most are very unpopular at most times. Strictly speaking, the

payoff matrix for the game being played is very large, growing to 101×101 entries by the end of a simulation. At the same time, the matrix for the choices actually in use is much smaller and more descriptive. When we refer to the game matrix at a given time, below, we refer to this reduced matrix, corresponding to the payoffs shown in the figures. Since all other choices are sporadically

re-evaluated by players, it is generally safe to assume that the choices not in use are dominated by those that are in use.

It is clear that essentially all the plays developing in these systems are descended from the *D* play of the original game. Nonetheless, payoffs higher than the original game's *C–C* combination are reached.

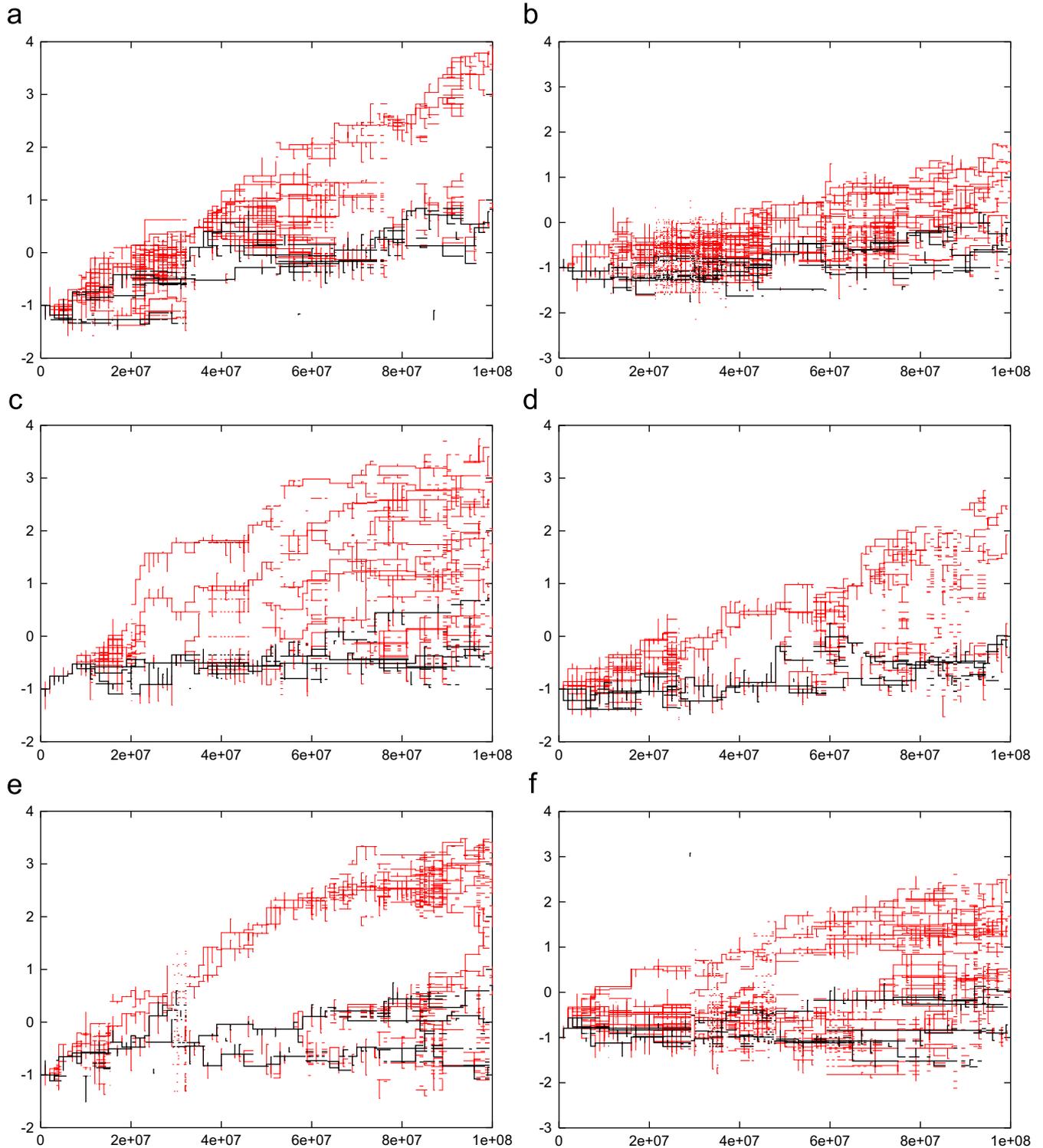


Fig. 3. Evolutionary trees of payoffs vs. time, from six randomly selected simulation runs.

Red values—representing off-diagonal matrix entries—rise steadily in evolutionary time. Black—diagonal—entries, on the other hand, do not show the same tendency, rising slowly if at all.

The number of black lines shown at a moment in time reveals the number of plays in active use at that time. There is a strong tendency for the descendants of the D choice to split quickly into two coexisting lineages, represented by two black branches of the evolutionary tree accompanied by two red branches representing the cross terms of the payoff matrix. A matrix for two plays that coexist at equilibrium is one in which each off-diagonal payoff exceeds the diagonal one in its column, because in this case the more popular one play is, the more favorable it is to choose the other. This configuration appears in these figures as two red lines above two black lines. As this pattern is accentuated through many cycles of mutation and selection, off-diagonal payoffs rise from negative to positive. As an example, the game matrix at time 5×10^7 of Fig. 3e is approximately

$$\begin{pmatrix} -0.83 & 1.85 \\ 1.69 & -0.13 \end{pmatrix}. \quad (1)$$

As positive matrix entries emerge, the probability of cooperation rises from zero, as seen in Fig. 4. As that figure shows, probability of cooperation tends to become positive and remain positive, and generally rises. In some cases it rises to near one and stays at that level until the end of the simulation.¹ Fig. 5 plots the average payoff in the game over time for the same simulation runs, showing that the average payoff rises about linearly as the matrix evolves, though a bit more slowly when the game is more diverse.

The typical two coexisting lineages can diversify further into more than two, as is seen late in the development of Figs. 3c and e and early in Figs. 3b and f. Also, lineages sometimes go extinct, for instance when a row of the payoff matrix comes to be dominated by another one. Extinctions of lineages are visible at several points in Fig. 3a, as the spectrum of game choices becomes more and less diverse several times. (Old plays sometimes come back to life because of a synergy with a new combination of active plays. Also, disused plays are occasionally selected to be the parent of a new mutant play. This is why the tree of descent of payoffs is not always connected, with lineages sometimes appearing out of nowhere. This dynamic does not seem to alter the overall pattern observed.) At other times, a single choice evolves by simple mutation and replacement, without diversifying. As we will discuss further below, evolution without

diversification tends to change the payoff in a positive direction. Two simulation runs illustrating this pattern are shown in Fig. 6.

As these figures suggest, there are two very common general patterns of evolution of the payoff matrix. These two patterns describe the overwhelming majority of simulated trajectories of this model. First, and more common, is the pattern of early diversification followed by increasing off-diagonal payoffs. This leads to a matrix like the one marked (1) above, or a more complex but conceptually similar one. This can be classified as a “division of labor game”, since the best payoffs are achieved when the players take different actions from each other as often as possible. Since we have not made it possible for the simulated players to coordinate with each other, the best they are able to do is settle on a mixed strategy that maximizes the chance of the positive payoffs. If they were able to coordinate, they could easily guarantee positive payoffs for both players on every round. As it is, the mean payoff awarded to a player steadily increases, as the off-diagonal payoffs increase, and becomes positive. In some cases (as in Fig. 3a), the diagonal terms also rise until all payoffs are positive.

The second common pattern, evolutionary increase in the payoff for a play appearing in isolation (seen in Fig. 6), often leads to a game whose payoffs are all positive. This can also take the form of a division of labor game, but from the point of view of cooperation this does not really matter because cooperation is inevitable no matter which choices the players make. This outcome can result either from a series of replacements without diversification, or from diversification followed by extinction of some branches, as in Fig. 6b.

A look into the structure of the payoff matrix can shed some light on these patterns. For one thing, a cycle of diversification and extinction tends to leave larger diagonal payoffs in the matrix than before. As an example, suppose two plays are coexisting with a game matrix of

$$\begin{pmatrix} -0.7 & -0.65 \\ -0.6 & -0.75 \end{pmatrix}.$$

(Coexistence requires that each off-diagonal entry be larger than the diagonal entry in its column.) Now the first play evolves by spawning a variant play that replaces it. Then the matrix will be

$$\begin{pmatrix} -0.7 + \Delta p_{11} & -0.65 + \Delta p_{21} \\ -0.6 + \Delta p_{12} & -0.75 \end{pmatrix}.$$

The first play will dominate the second one and eliminate it from circulation if the upper left value exceeds the lower left one. This can happen at any value of Δp_{11} , but is much more likely when Δp_{11} is positive. Then the remaining entry of the game matrix

$$(-0.7 + \Delta p_{11})$$

¹By the end of one of these figures, after 100 mutations, the old plays that have been abandoned to the “library” at a frequency of 0.001 each are occupying up to 10% of the spectrum of choices. Thus a perfectly cooperative game could arise and have a measured probability of cooperation as little as 0.81 (the square of the frequency of the cooperative plays). This is why the probability of cooperation is less than 1 when it is pinned at the top of the scale.

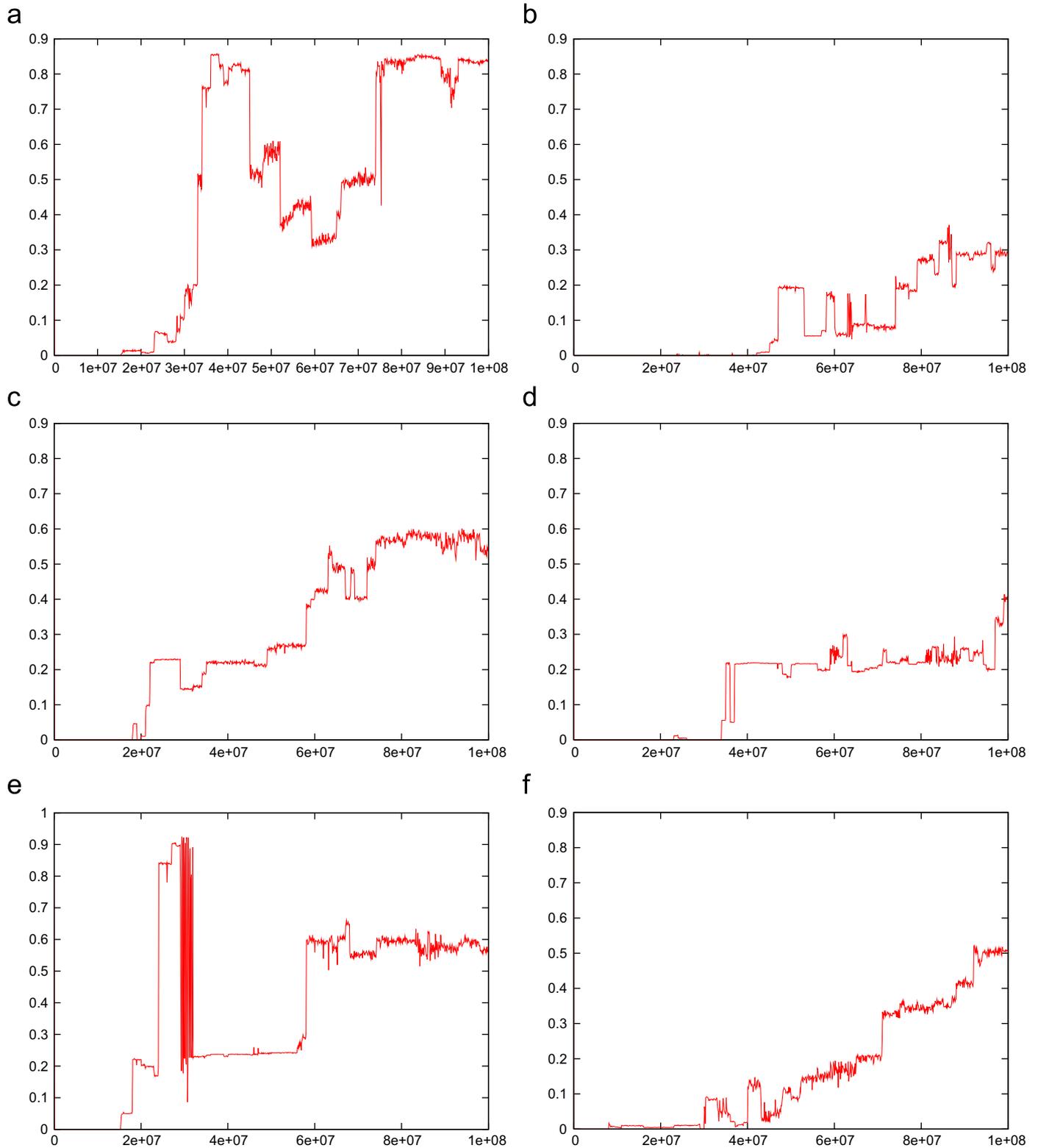


Fig. 4. Probability of cooperation vs. time in the same simulations as in Fig. 3.

exceeds either of the diagonal entries that were present before the mutation.

The same logic applies to direct replacement of a lone matrix entry. Taking the initial matrix (once the C play is abandoned) as an example, when a mutant form of

“defection” arises the matrix becomes

$$\begin{pmatrix} -1.0 & -1.0 + \Delta p_{21} \\ -1.0 + \Delta p_{12} & -1.0 + \Delta p_{22} \end{pmatrix};$$

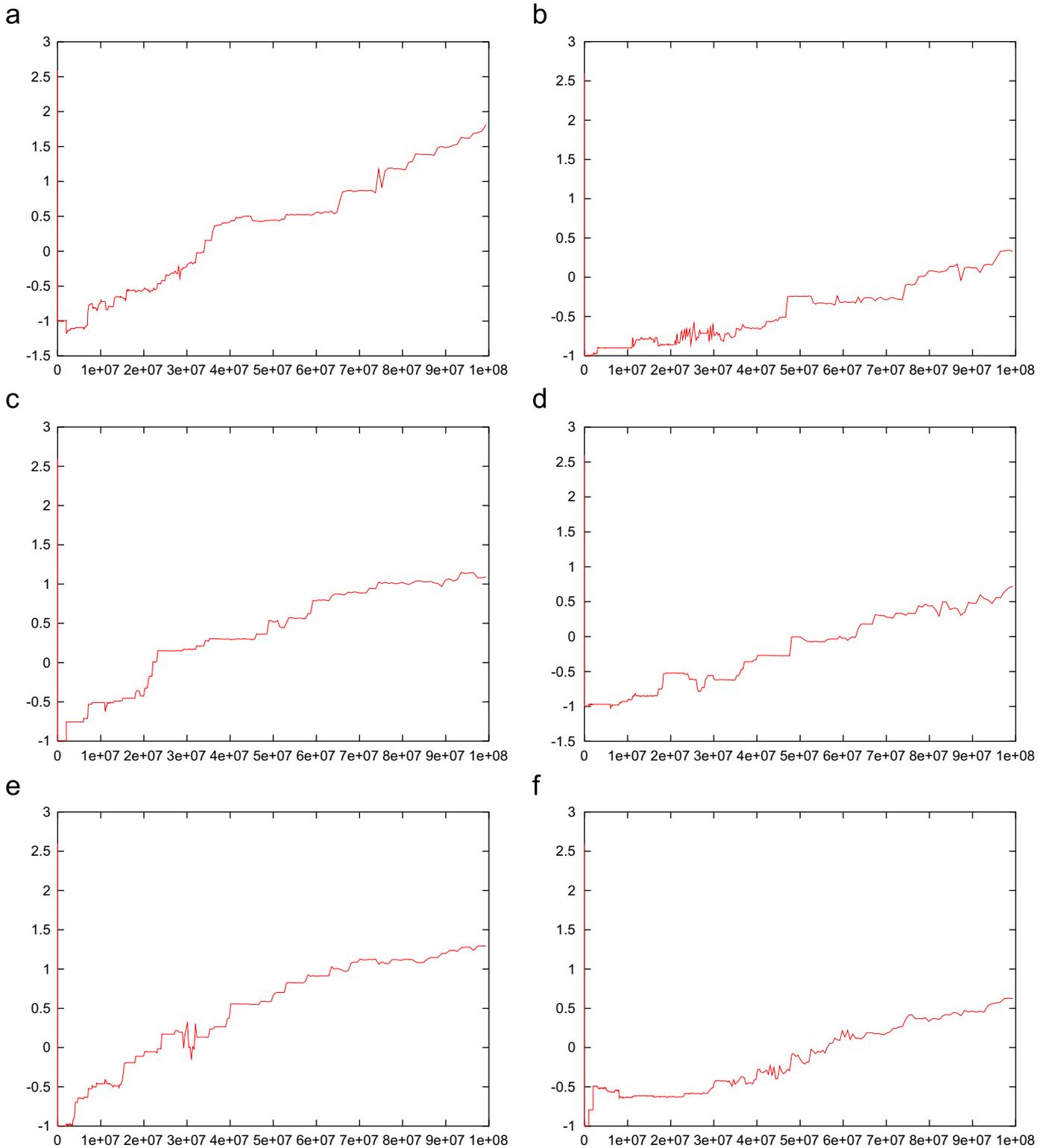


Fig. 5. Mean payoff vs. time in the same simulations as in Fig. 3.

the mutant play is initially adopted if $\Delta p_{12} > 0$; it replaces the parent if $\Delta p_{22} > \Delta p_{21}$, given which Δp_{22} is more likely positive than negative, and the new lone matrix entry is expected (on average) to be larger than the old one. On the other hand, if the mutant play coexists with the parent play, that is, when $\Delta p_{22} < \Delta p_{21}$, the new diagonal entry is

not expected to be larger than the old one. In fact, it seems it is more likely to be smaller. These analytical observations are consistent with the pattern seen in simulations that diagonal entries tend to grow in times of replacement of a single play, but tend to wander with much less of a trend in times of coexisting plays. Diagonal entries may be

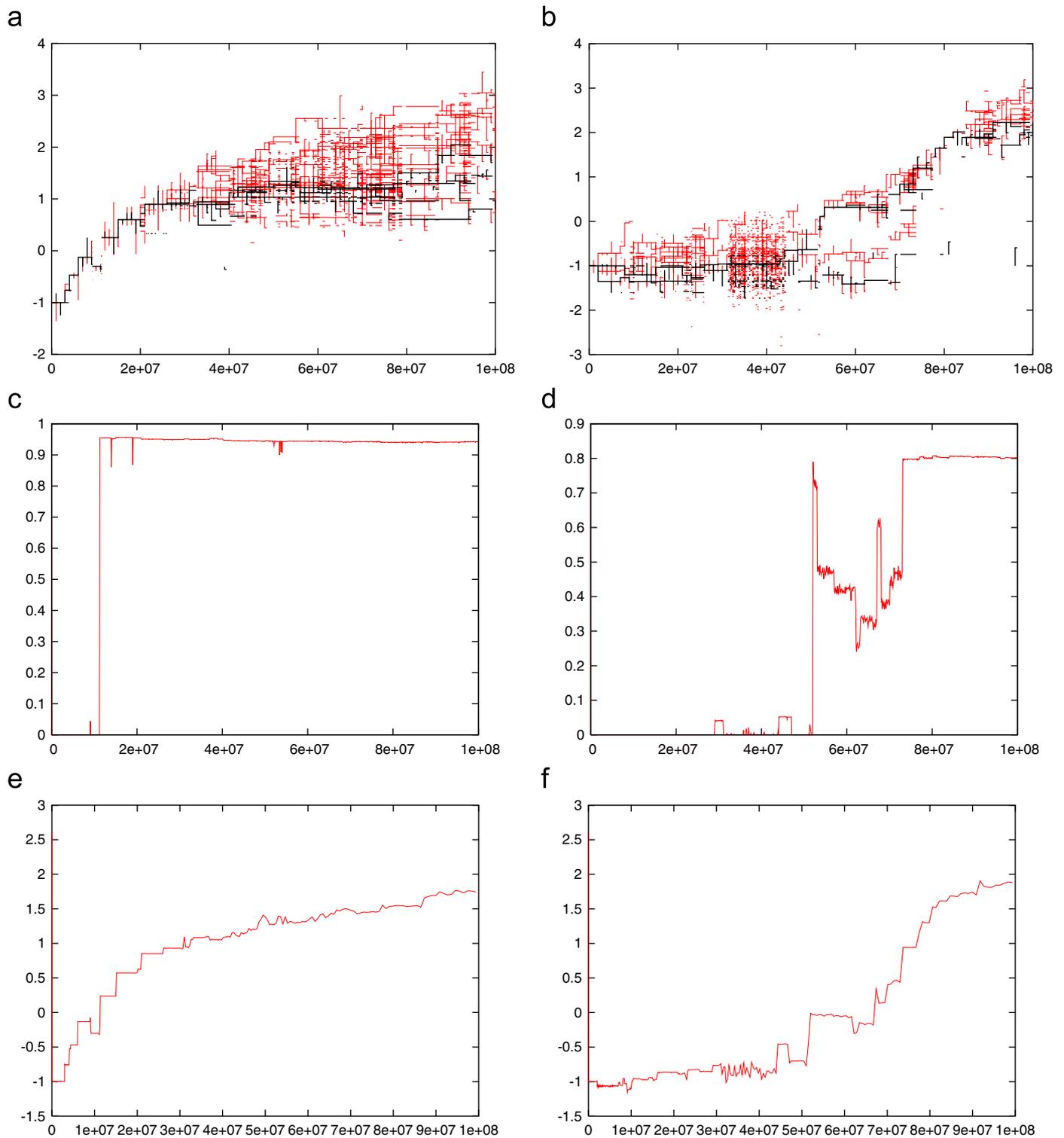


Fig. 6. (From top to bottom) Evolution of payoff matrix, probability of cooperation, and mean payoff vs. time for two simulation runs in which replacement of a single matrix entry drives payoffs upward, culminating in a pure cooperation game in which all payoffs are positive.

influenced by a fine-grained mix of coexistence and replacement events, as suggested by the fact that the number of coexisting plays continually rises and falls on the evolutionary time-scale.

8. No temptation to defect

In the prisoner's dilemma game matrix used as the initial matrix in these simulations, the D play dominates because

each payoff in the D row is the largest in its column:

$$\begin{array}{c} C \quad D \\ C \begin{pmatrix} 3 & -2 \\ 4 & -1 \end{pmatrix} \\ D \end{array}$$

Now consider the matrix for the original D play and the play that dominates just after time 8.2×10^7 in Fig. 6b (let us call this play C^+):

$$\begin{array}{c} C^+ \quad D \\ C^+ \begin{pmatrix} 2.01 & -1.34 \\ -1.09 & -1 \end{pmatrix} \\ D \end{array}$$

Here neither play dominates, and either is stable against adoption of the other. Thus an important difference between this play and the original C is that when C^+ is the norm, players using C^+ not only get the best payoff available, they also are not subject to any temptation to defect. A defector in a population of C^+ players gets a payoff of -1.09 , which compared to 2.01 induces a strong temptation *not* to defect.

Since all discarded options are kept in the system at non-zero frequency, the set of plays active at any time in the simulation is stable against all the ones that have gone before. Consequently, all the cooperative outcomes seen in the figures are stable with regard to both the original C and D plays. Cooperative outcomes tend to diversify into several plays, all combinations of which yield a positive payoff; that cluster of plays is stable against defection in the same way as above, though the analysis is more tedious. The pattern of cooperation without temptation to defect is robust and repeatable.

9. Discussion

In various studies of iterated prisoner's dilemma games, widespread cooperation can be achieved when players respond to others' past actions with strategies such as Tit-for-Tat. In many such models, though, such cooperative communities are vulnerable to invasion by "cheater" strategies that can take advantage of the sucker's payoff in certain cases and escape punishment. In the present model, the model players that we have presented with a dilemma can readily find their way out of the dilemma and achieve positive average payoffs, by mutating the available choices into choices that often produce beneficial outcomes when all players behave selfishly. In some cases, they are able to devise options that always produce positive payoffs. Neither of these outcomes (either positive average payoff, or invariably positive payoff) is undermined by "temptation to defect", that is, an incentive for a player to choose an option that denies positive payoffs to the other player. Cooperation is thus coincident with selfish behavior. In this case, mechanisms such as indirect reciprocity, kin selection, group selection, punishment, communication, and the like

are not necessary since the mutually optimal outcome is easily achieved without them.

The work of Mesterton-Gibbons (1991) and Turner and Chao (2003), mentioned earlier, suggests scenarios in which the transformation from dilemma to stable cooperation may be possible. In harmony with Turner and Chao's results, Mancur Olson, in his classic text *The Logic of Collective Action* (1965, pp. 32–33), points out that small group size can sometimes avert collective action problems. This is evident if one considers a group of size one, in which "defection" is clearly not a favorable option. If the cost of contributing (cooperating) remains constant, while the benefit of cooperating (relative to not cooperating) becomes less and less significant as the group size increases, the balance will gradually shift from favoring cooperation to discouraging it. These examples point to the existence, whether common or rare, of particular "degrees of freedom" along which interactions that have the structure of a collective action problem can be changed into more favorable arrangements.

The model we have presented here addresses the question somewhat differently, suggesting that by "tinkering" with any number of small adjustments, a community can gradually find its way from a dilemma to an arrangement with a more favorable incentive structure. No single solution such as reducing group size or reducing discrimination between individuals need be relied on, though these possibilities are admitted by our scenario.

It seems reasonable that in most biological and social situations, there are generally many ways of varying ones interactions with others. Given that intuition, though such things can be difficult to quantify, our modeling assumptions may be at least as apropos as modeling approaches that allow only two options, or a one-dimensional space of options between two extremes. If so, our results suggest that prisoner's dilemma or other social dilemma scenarios, where they arise, may be more readily removed by communities than is often imagined, and consequently, that they may not occur persistently in actual communities as frequently as imagined.

In fact, since the only difference between the model scenario presented here and a standard prisoner's dilemma model is the introduction of slightly perturbed variant strategies, it follows that familiar results generated by prisoner's dilemma models, such as the emergence of Tit-for-Tat or Pavlov strategies, depend on the absence of sufficient variation to make escape to stable cooperation possible. Therefore, when those results are invoked in discussions about how to achieve cooperation, it seems that the lack of available variation implied in those results requires justification as much as does the assumption of one particular kind of variation.

In the mid-twentieth century, cultural anthropologist Ruth Benedict made a comparative study of a number of different societies that she and her colleagues had studied around the world, investigating differences in how societies are organized and the different qualities they take on.

Focusing on aggression, meaning behavior aimed at hurting, expelling, or humiliating another person, “from all comparative material”, she reports, “the conclusion that arises is that societies where non-aggression is conspicuous have social orders in which the individual by the same act and at the same time serves his own advantage and that of the group” (Maslow and Honigman, 1970). While Benedict is not necessarily distinguishing between advantageous behaviors that require no coordination due to byproduct cooperation and advantageous behaviors that are sustained by, for instance, reciprocal altruism or punishment, her point is at the very least suggestive that instead of relying on reciprocity, social norms, or other means to mitigate temptation to defect in collective action problems, there may also sometimes be an opportunity to construct a much better situation by introducing unconsidered options for behavior that remove the disjuncture between selfishness and generosity.

The prisoner’s dilemma represents the most difficult and intractable of a spectrum of possibilities. It is not always possible to arrange a confluence between what people want to do and what others want them to do. It is important to understand how best to handle such situations, but it is equally crucial to recognize that there are many other ways to achieve cooperative interactions, so that we do not mistake a non-dilemma for a dilemma, and so that we are alert to possibilities of transforming a dilemma into something more promising. More generally, emphasizing the difficulty of achieving cooperation in prisoner’s dilemma, social dilemma and tragedy of the commons situations without acknowledging the existence of other situations in which cooperation is much easier is both incomplete theoretically and dangerous socially. It can lead us to make overly pessimistic and destructive policy decisions, and can cause us to overlook opportunities to construct easy forms of cooperation, as well as making it more difficult to explain the ubiquity of cooperative human behavior.

Acknowledgments

This work was funded by a Department of Energy Computational Science Graduate Fellowship, by the Sloan and Mellon foundations, and by a Ciriacy-Wantrup Postdoctoral Fellowship. Conversations with Junling Ma, Jonathan Dushoff, Joshua Plotkin, Stefan Geritz, Derrick Jensen, Peter Richerson, and Richard Norgaard have been especially helpful. We are grateful for the help of Narin Dickerson, Trevor Lango, Jerry Vigil, Matthew Tom, Stephen Fung, Jesse Edelman, and two anonymous reviewers. SAL thanks Harvard’s Program in Evolutionary Dynamics for its hospitality in 2006.

Appendix A. Model details

In our simulations, each game player is represented by a strategy vector, which assigns a probability to each play. The number of entries in each strategy vector is thus always

the same as the number of rows or columns in the game matrix. Ten players are simulated.

The game matrix is initially equal to the prisoner’s dilemma matrix, as given above. Each player’s initial strategy is 0.9 probability for *C* and 0.1 for *D*. At each time step, each player is paired with a randomly chosen other player (so that every player plays twice per time step on average). Each of the pair chooses a play, and each player’s strategy vector is adjusted according to the payoff received.

If s_i is a player’s probability of choosing play i , when i is chosen and receives payoff p , s_i is adjusted in accordance with the payoff so that play i becomes more likely if p is positive and less likely if p is negative. This adjustment has two steps: first s_i is replaced by $s_i + 0.01p$ or 0.001, whichever is larger; then the strategy vector is normalized, by scaling all the probabilities that are larger than 0.001 to make the total probability equal to one. This algorithm guarantees that all probabilities are kept between 0.001 and 1.0.

The learning dynamics often converges to a near steady state (with small stochastic fluctuations), but sometimes displays persistent oscillations. In either case, one million time steps appears to be sufficient to escape from transient dynamics.

After every million time steps, a new play is added to the matrix, with payoff entries derived from those of an existing play. The “parent” play is chosen randomly from the current set of plays, each play having a probability equal to its current frequency of use in the population. The payoffs in the derived play’s row and column are constructed from the corresponding payoffs of the parent play by adding a perturbation to each, chosen independently from a normal distribution with mean zero and standard deviation 0.2.

An entry for the new play is added to each player’s strategy vector, giving each player an initial probability for that play of 0.05, or half the parent play’s probability, whichever is smaller. The parent play’s probability is decreased by the corresponding amount. This is equivalent to assuming that one player discovers the new play and quickly communicates it to the other players, so that all have the option of trying it.

Each simulation runs until 99 new plays have been introduced.

While all plays remain present in the population, at a frequency of at least 0.001, only those whose overall frequency is 0.05 or more are shown in the plots of payoffs vs. time, in order to illustrate how the players’ strategies change.

References

- Axelrod, R., 1984. The Evolution of Cooperation. Basic Books, New York.
- Boyd, R., Lorberbaum, J.P., 1987. No pure strategy is evolutionarily stable in the repeated prisoner’s dilemma game. *Nature* 327, 58–59.
- Boyd, R., Richerson, P.J., 1985. Culture and the Evolutionary Process. University of Chicago Press, Chicago.

- Boyd, R., Richerson, P.J., 1990. Group selection among alternative evolutionarily stable strategies. *J. Theor. Biol.* 145, 331–342.
- Boyd, R., Richerson, P.J., 1992. Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethol. Sociobiol.* 13, 171–195.
- Boyd, R., Richerson, P.J., 2002. Group beneficial norms can spread rapidly in a structured population. *J. Theor. Biol.* 215, 287–296.
- Brand, S., 1974. *Two Cybernetic Frontiers*. Random House, New York, Bookworks, Berkeley, CA.
- Clutton-Brock, T., 2002. Breeding together: kin selection and mutualism in cooperative vertebrates. *Science* 296, 69–75.
- Cohen, M.D., Riolo, R.L., Axelrod, R., 1999. The emergence of social organization in the prisoners' dilemma: how context-preservation and other factors promote cooperation. Santa Fe Institute Working Paper 99-01-002.
- Connor, R.C., 1995a. Altruism among non-relatives: alternatives to the 'prisoner's dilemma'. *TREE* 10 (2), 84–86.
- Connor, R.C., 1995b. The benefits of mutualism: a conceptual framework. *Biol. Rev.* 70, 427–457.
- Dugatkin, L.A., Mesterton-Gibbons, M., Houston, A.I., 1992. Beyond the prisoner's dilemma: toward models to discriminate among mechanisms of cooperation in nature. *TREE* 7 (6), 202–205.
- Evans, D., Zarate, O., 1999. *Introducing Evolutionary Psychology*. Totem Books, USA, Icon Books, UK.
- Hardin, R., 1971. Collective action as an agreeable *n*-prisoners' dilemma. *Behav. Sci.* 16, 472–481.
- Hardin, R., 1982. *Collective Action*. Johns Hopkins University Press, Baltimore, MD.
- Harley, C.B., 1981. Learning the evolutionarily stable strategy. *J. Theor. Biol.* 89, 611–633.
- Hauert, C., Michor, F., Nowak, M.A., Doebeli, M., 2006. Synergy and discounting of cooperation in social dilemmas. *J. Theor. Biol.* 239, 195–202.
- Henrich, J., Boyd, R., 2001. Why people punish defectors: weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *J. Theor. Biol.* 208, 79–89.
- Holland, J.H., 1993. The effect of labels (tags) on social interactions. Santa Fe Institute Working Paper 93-10-064.
- Leigh, E.G., 2001. Mutualism, evolution of. In: Levin, S.A. (Ed.), *Encyclopedia of Biodiversity*, vol. 4. Academic Press, San Francisco, pp. 281–289.
- Luce, R.D., Raiffa, H., 1957. *Games and Decisions*. Wiley, New York (The story of the two prisoners appears on p. 95).
- Maslow, A.H., Honigman, J.J., 1970. Synergy: some notes of Ruth Benedict. *Am. Anthropol.* 72, 320–333.
- May, R.M., 1981. The evolution of cooperation. *Nature* 292, 291–292.
- Maynard Smith, J., 1982. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.
- Maynard Smith, J., Price, G.R., 1973. The logic of animal conflict. *Nature* 246, 15–18.
- Mesterton-Gibbons, M., 1991. An escape from 'the prisoner's dilemma'. *J. Math. Biol.* 29, 251–269.
- Miller, J., Butts, C., Rode, D., 1998. Communication and cooperation. Santa Fe Institute Working Paper 98-04-037.
- Nowak, M., May, R.M., 1992. Evolutionary games and spatial chaos. *Nature* 359, 826–829.
- Oliphant, M., 1994. Evolving cooperation in the non-iterated prisoner's dilemma: the importance of spatial organization. In: Langton, C.G. (Ed.), *Artificial Life IV*. Addison-Wesley, Reading, MA, pp. 349–352.
- Olson, M.J., 1965. *The Logic of Collective Action*. Harvard University Press, Cambridge, MA.
- Pepper, J.W., Smuts, B.B., 2002. A mechanism for the evolution of altruism among nonkin: positive assortment through environmental feedback. *Am. Nat.* 160 (2), 205–213.
- Picardi, A., Seifert, W., 1976. A tragedy of the commons in the Sahel. *Technol. Rev.* (May), 42–51.
- Rapoport, A., Chammah, A.M., 1985. *Prisoner's Dilemma*. University of Michigan Press, Ann Arbor.
- Sigmund, K., Nowak, M., 1997. The natural history of mutual aid. In: Stadler, F. (Ed.), *Wissenschaft als Kultur. Österreichs Beitrag zur Moderne. Veröffentlichungen des Instituts Wiener Kreis*, vol. 6. Springer, Wien, New York, pp. 259–272.
- Sober, E., Wilson, D.S., 1998. *Unto Others*. Harvard University Press, Cambridge.
- Sugden, R., 1986. *The Economics of Rights, Co-operation and Welfare*. Blackwell, Oxford, New York.
- Trivers, R., 1971. The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57.
- Turner, P.E., Chao, L., 2003. Escape from prisoner's dilemma in RNA phage $\phi 6$. *Am. Nat.* 161 (3), 497–505.
- Wade, M.J., Breden, F., 1980. The evolution of cheating and selfish behavior. *Behav. Ecol. Sociobiol.* 7, 167–172.
- Whitfield, J., 2002. The police state: on the surface, beehives and ant nests seem to be model societies, with each individual striving for the common good. But maintaining this social order sometimes calls for brutal tactics. *Nature* 416, 782–784.
- Worden, L., 2003. Evolution, constraint, cooperation and community structure in simple models. Ph.D. Thesis, Program in Applied and Computational Mathematics, Princeton University, available online at (<http://two.ucdavis.edu/~worden/pub/thesis.pdf>).
- Wright, S., 1980. Genic and organismic selection. *Evolution* 34 (5), 825–843.