Cumulative Benefit Games: Achieving Cooperation when Players Discount the Future

DAVID W. STEPHENS*

Department of Ecology, Evolution and Behavior, University of Minnesota, St. Paul, MN 55108-6097, U.S.A.

(Received on 7 June 1999, Accepted in revised form on 6 March 2000)

Experimental studies with captive animals show strong preferences for immediate reward. Several authors have argued that these tendencies to discount delayed reward may severely limit the Iterated Prisoner’s Dilemma game as a model of animal cooperation. This paper explores a simple mechanism, dubbed cumulative games, that can, in principle, promote cooperative action even when there is strong temporal discounting. In the simplest type of cumulative game a pair of players does not receive benefits at the end of each play, as in a conventional repeated game, but must complete a sequence of games before collecting the accumulated benefits. In a preliminary analysis pitting tit-for-tat against all-D, I show that accumulation can promote a conditionally cooperative strategy even when there is strong temporal discounting. However, the delays created by accumulation de-value the pairwise interaction, so although the relative value of cooperation increases, the total value of the interaction decreases. I investigate accumulation further by simulating the evolution of a broader class of strategies. These simulation studies show that accumulation, and small discounting rates (high future value) can both promote cooperative action. The limitations of these results are discussed.

Introduction

The Iterated Prisoner’s Dilemma (IPD) game has been a central part of the theory of non-kin cooperation for nearly two decades (Axelrod & Hamilton, 1981). The IPD is an important and informative thought experiment that has stimulated a large body of ingenious theory. Recently, however, several authors have challenged the pre-eminence of the IPD, questioning its empirical support, and criticizing its relatively narrow realm of applicability (Packer & Ruttan, 1988; Heinrich, 1988, 1989; Pruett-Jones & Pruett-Jones, 1994; Noé, 1990; Dugatkin et al., 1992; Mesterton-Gibbons & Dugatkin, 1992; Grinnell et al., 1995; Heinsohn & Packer, 1995; Connor, 1995a; Clements & Stephens, 1995; Stephens et al., 1995, 1997; Pusey & Packer, 1997). My laboratory has been among those criticizing the IPD on empirical grounds. We have found that pairs of captive bluejays are strongly attracted to the mutual defection equilibrium of IPDs even though they have experienced massive amounts of repetition (Clements & Stephens, 1995). Moreover, although the complete empirical picture is complex, this tendency to favor defection over the long-run has been widely reported elsewhere.
(e.g. Scodel et al., 1959; Colman, 1982; Gardner et al., 1984; Green et al., 1995).

Following the conventional view of the IPD, cooperators forego a large immediate gain (the temptation) in order to achieve a stream of intermediate size gains in the long run. Theory suggests (e.g. Kagel et al., 1986; Benson & Stephens, 1996; Stephens et al., 1997) however, that animals should de-value delayed benefits. Moreover, as my colleagues and I have pointed out (Stephens et al., 1997), experimental measurements of animal discounting show a surprisingly strong effect on preference. For example, my laboratory has shown that a 40 s re-arrangement of the payoff within an IPD matrix can turn cooperative behavior on and off. This tendency to discount future benefits limits cooperativeness in the IPD, because it devalues the stream of future gains that are supposed to make cooperation worthwhile. I have hypothesized, therefore, that the relevance of the IPD to animal cooperation may be severely limited by the strong discounting that I and other experimentalists have observed.

In evaluating this hypothesis, we would like to know a great deal more about discounting. For example, current studies of discounting all focus on food reward, and focus on a relatively narrow range of animal species (typically small birds and mammals). In this paper, however, I assume that strong discounting is the norm and ask whether there is some way to arrange an IPD-like game that can overcome strong animal preferences for immediate reward? I will argue for a simple, and intuitively appealing mechanism that, in theory, can overcome discounting. The idea is that if the gains derived from a sequence of plays accumulate, only becoming available to the players at the end of a given number of interactions, then not only is the “immediacy” of the temptation to cheat effectively removed, but a reciprocating opponent has a chance to punish a cheater before it realizes any gains from its anti-social behavior.

**Review of the Basic Model**

To understand the role of discounting and accumulation in PD-like games, it is helpful to review the basic theory behind cooperation in the IPD. Consider the PD matrix

<table>
<thead>
<tr>
<th>Column player</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Row player</td>
<td>R = 3</td>
<td>S = 0</td>
</tr>
<tr>
<td>D</td>
<td>T = 5</td>
<td>P = 1</td>
</tr>
</tbody>
</table>

(Matrix 1)

where C represents “cooperation” and D “defection”, and the matrix entries show the row player’s benefits; the numerical payoffs are a famous example; in general we require $T > R > P > S$ and $R > (T + S)/2$. Following Axelrod & Hamilton’s (1981) famous analysis, stable mutual cooperation can be achieved if the opponents play repeatedly, and they adopt a conditionally cooperative strategy like tit-for-tat (abbreviated TFT; a tit-for-tat player chooses C on the first play, and copies his opponent’s move on all subsequent plays). Then the game matrix

<table>
<thead>
<tr>
<th>Column player</th>
<th>TFT</th>
<th>All-D</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFT</td>
<td>$R + \frac{w}{1 - w} R$</td>
<td>$S + \frac{w}{1 - w} P$</td>
</tr>
<tr>
<td>All-D</td>
<td>$T + \frac{w}{1 - w} P$</td>
<td>$P + \frac{w}{1 - w} P$</td>
</tr>
</tbody>
</table>

(Matrix 2)

shows the relative merits of TFT and an alternative, consistently defecting, opponent (All-D). Where, $w$ is usually interpreted as the probability that the game will continue from one play to next. Focusing on the term $w/(1 - w)$ (which is the expected repetition of the game), one can see that all entries in the matrix are in the familiar slope–intercept form for the equation of a straight line (cf. the solid lines in Fig. 1). Comparing the “TFT vs. TFT” cell to the “All-D vs. TFT” cell, one can see that the All-D vs. TFT cell has the higher intercept (since $T > R$), but smaller slope (since $R > P$). This means that defecting is a sensible strategy if the game is played only a few times, but for larger amounts of repetition the conditionally cooperative strategy tit-for-tat can be sensible.
FIG. 1. A comparison of the temporal benefit patterns in “typical” and cumulative-benefit games. show the conventional analysis of All-D vs. TFT. The thicker line shows how the benefits increase with time when All-D plays TFT, while the thinner line shows TFT vs. TFT. The two “dashed” staircases show how benefits increase with time in a cumulative benefit game (accumulation period, \( n \)). with triangles at its vertices shows All-D vs. TFT and the thinner (with circles) shows TFT vs. TFT. In conventional games there exists a time, early in the game, where a defector is “ahead” of a TFT player, and small \( w \)'s make it hard to overcome this early advantage. Cumulative benefit games can eliminate this early advantage.

**DISCOUNTING**

To see how discounting enters our analysis, consider the calculation of the TFT vs. TFT cell above. To calculate the expected benefits in this case, we simply write

\[
R + wR + w^2R + w^3R + \cdots = R + \frac{w}{1-w}R,
\]

where we “discount” the reward obtained at the \( i \)-th play by the probability that play will continue until the \( i \)-th play, \( w^i \). If in addition to this termination risk discounting, the value of a benefit obtained at the \( i \)-th play is actually less than the value available in the preceding play, then we must modify this calculation. In the simplest case, imagine that the \( w \) is the product of two elements \( \beta \) the probability that play will continue from one play to the next, and \( \alpha \) the proportion of a benefit’s value that is retained from one play to next, then we can re-write the expression above as

\[
R + \beta \alpha R + \beta^2 \alpha^2 R + \beta^3 \alpha^3 R + \cdots = R + \frac{\beta \alpha}{1-\beta \alpha} R.
\]

Clearly, this does not change the form of the model in any meaningful way. We still require a relatively large \( w \) to maintain cooperative behavior, but we interpret \( w \) more broadly. My hypothesis, that strong discounting limits cooperativeness, is equivalent to the claim that \( \alpha \) is typically small, say smaller than 0.1, and this limits the magnitude of \( w \), because \( \alpha > \beta \alpha = w \). I remark that the unfortunate tendency to suppose that \( w \) is determined solely by the likelihood of repetition is peculiar to biological modelers (see May, 1981; Stephens *et al.*, 1995; Dugatkin, 1997, for exceptions); for example, Axelrod (1984) called \( w \) “the shadow of the future”, a phrase that allows considerable room for interpretation. In the following, I call \( w \) the future value term, a name chosen to suggest that \( w \) measures the general value of future benefits regardless of whether repetition or discounting determines this “value”.

**Cumulative-benefit Games**

In this section, I accept the premise that the future value term \( w \) is small, possibly because \( \alpha \) is small, and I ask whether one can construct a scenario where cooperation can be maintained in spite of \( w \)’s smallness.

Consider two individuals playing a sequence of games, as one typically imagines in the repeated games literature, except that some mechanism withholds the benefits until the pair has completed \( n \) plays. If, for example, \( n = 2 \), the players get nothing at the end of the first play, but obtain the combined benefits of the first and second play, after they play a second time. I call \( n \) the accumulation period. To see the effect of
accumulation, consider how one calculates the benefits to All-D when All-D plays TFT. In the simplest case, we assume that players lose their accumulated benefits if play terminates before a bout of \( n \) is complete. In the first bout of \( n \) plays, All-D expects

\[
w^{n-1}[T + (n - 1)P],
\]

where \( T + (n - 1)P \) is the gain from the first \( n \) plays (by convention we “start the clock” at the end of the first play, so there are \( n - 1 \) rather than \( n \) intervals). Similarly, expected gains from the next bout of \( n \) plays are

\[
w^{2n-1}nP,
\]

where \( nP \) are the gains from a completed bout of \( n \) plays and \( w^{2n-1} \) is the expected proportion of that value retained at the end of the second bout. The expected gains from a sequence of these plays is

\[
w^{n-1}([T + (n - 1)P] + w^n P + w^{2n}nP + \cdots) = w^{n-1} \left( [T + (n - 1)P] + \frac{w^n}{1 - w^n}nP \right).
\]

This is the first of four cells in a modified version of the classic TFT vs. All-D game matrix. The other three cells can be calculated following the same plan, yielding the matrix

\[
\begin{array}{c|cc}
 & \text{TFT} & \text{All-D} \\
\hline
\text{TFT} & w^{n-1} \left[ R + (n - 1)R + \frac{w^n}{1 - w^n}nR \right] & w^{n-1} \left[ S + (n - 1)P + \frac{w^n}{1 - w^n}nP \right] \\
\text{All-D} & w^{n-1} \left[ T + (n - 1)P + \frac{w^n}{1 - w^n}nP \right] & w^{n-1} \left[ P + (n - 1)P + \frac{w^n}{1 - w^n}nP \right]
\end{array}
\]

Accumulation has two effects. It devalues the interaction, but enhances the relative value of conditionally cooperative strategies. The combined discounting and repetition term \( w^{n-1} \) devalues all entries in the matrix, this happens because accumulation means that one must wait at least \( n - 1 \) intervals to obtain any benefit. If a potential player can choose an alternative activity, then accumulation makes any participation in the game less attractive. If, however, a pair enters into the game, presumably because the available alternatives are not very attractive, then accumulation enhances the relative value of a conditionally cooperative strategy like TFT. If, as I have argued, \( w \) is small, the relative magnitudes in the matrix above are approximately

\[
\begin{array}{c|cc}
 & \text{TFT} & \text{All-D} \\
\hline
\text{TFT} & R + (n - 1)R & S + (n - 1)P \\
\text{All-D} & T + (n - 1)P & P + (n - 1)P
\end{array}
\]

(Matrix 3)

The form of this matrix is exactly the same as Matrix 2, the classical, elementary analysis of TFT vs. All-D. In parallel with the classical analysis, TFT can be stable against All-D, if \( n - 1 \) is large enough; \( n - 1 \) is not, however, the expected repetition of the game, but the period over which benefits accumulate. As in Axelrod and Hamilton’s original analysis some form of conditional cooperation is required: unconditional cooperation (All-C) fares no better against All-D in an accumulated game than it does in a traditional IPD. The key difference between this and Matrix 2’s traditional analysis is that here, if \( n \) is large enough, All-D never realizes any advantage; because by the time payoffs are distributed All-D’s ill-gotten \( T \) has been diluted by \( n - 1 \) \( P \) payoffs, so All-D achieves a smaller payoff than it could have had by playing TFT. Figure 1 shows this graphically. Figure 1 shows the conventional TFT vs. TFT, All-D vs. TFT payoffs as intersecting straight lines. The comparable payoffs for an
accumulated game are shown as dashed-staircases that “sample” the conventional lines at the points of accumulation. If the period of accumulation is large enough, the short-term advantage of All-D is eliminated simply because benefits cannot be collected until after the two (conventional) benefit lines have crossed.

This idea raises many intriguing questions about how such games might arise in nature. I defer this question to the discussion, but comment that the idea that one can encourage “cooperativeness” by controlling the timing of payoffs is part of the everyday logic of human commerce. The plumber is much more likely to explain the merits of one repair technique over another, or to consult about some other minor annoyance, if the homeowner withholds payment until after he or she is satisfied with the plumber’s workmanship and explanations.

Evolved Strategies in Cumulative Games

As a first step, I have pitted tit-for-tat against All-D in order to study how accumulation might effect cooperative behavior; however, the literature of the Prisoner’s Dilemma contains many possible strategies together with studies of strategic interaction [see Dugatkin, (1997) for a thoughtful, up-to-date review]. One would like, therefore, to consider the combined effects of accumulation (n) and future value preservation (w) for a broader class of strategies. I did this by simulating the evolution of a fairly general class of strategies; that is, the class of strategies that can be represented by a 5-element vector of probabilities (t, r, p, s, c), where t is the probability of cooperating after visiting the T cell of the matrix (note that in a cumulative game the benefit of T might not be collected until some later time); similarly, r, p and s are the probabilities of cooperating after the R, P and S cells have been visited; c is the probability of initial cooperation. This scheme, similar to one used by Nowak & Sigmund (1993), and identical to that of Stephens et al. (1995) can be used to represent many of the “classic” strategies of the field [e.g. TFT is \( (t = 1, r = 1, p = 0, s = 0, c = 1) \), Pavlov is \( (t = 0, r = 1, p = 1, s = 0, c = \text{unspecified}) \)].

Consider a pair of strategies, \( S_1 = (t_1, r_1, p_1, s_1, c_1) \) and \( S_2 = (t_2, r_2, p_2, s_2, c_2) \). From the point of view of player 1, we can specify the dynamic transitions within a sequence of plays by constructing a transition matrix \( M \); this matrix specifies the probabilities of transitions from the states \( \{T, R, P, S\} \) on the i-th play to the same states on the next, or \( i + 1 \)st, play. Construction of this matrix is straightforward (details in Stephens et al., 1995). In addition, we need a vector of probabilities, \( c \), that give the probabilities of the four states before play has begun; this vector can be constructed from the \( c_1 \) and \( c_2 \) components of the strategy vectors. From the point of the view of the two players, the transition matrix \( M \) and the initial distribution vector \( c \) represent a complete description of their pair-wise interaction. Appendix A shows that player 1’s payoff is

\[
(T R P S)^w (1 + M + M^2 + \cdots + M^{n-1})
\]

\[
\times (1 - w^n M^n)^{-1} c,
\]

where \( n \) is the period of accumulation. Player 2’s payoff is

\[
(S R P T)^w (1 + M + M^2 + \cdots + M^{n-1})
\]

\[
\times (1 - w^n M^n)^{-1} c
\]

(the payoff vector is simply re-ordered to reflect the fact that whenever player 1 gets \( T \), player 2 gets \( S \)). Note that the terms \( T, R, P \) and \( S \) are used in two ways; to denote the payoffs accruing when the events DC, CC, DD and CD occur—as in the row vector above, and as a shorthand for the events themselves as in the set notation \( \{T, R, P, S\} \) used earlier in this paragraph.

I studied eqns (1) and (2) using an evolutionary algorithm, following the principles outlined by Back (1996). Appendix B shows the structure of my algorithm, and gives many of its key parameters. I restricted my attention to the classical payoff matrix \( (T = 5, R = 3, P = 1, S = 0) \), focusing my attention on investigating a wide range of accumulation periods \( n = 1, 5, 10 \) and 25) and future value parameters \( w = 0.01, 0.25, 0.75 \) and 0.99). I ran 20, 5000 generation runs at each \( (n, w) \) pair.
Results: Simulated Evolution of Strategies

PATTERNS OF COOPERATIVENESS

I considered three similar measures of cooperativeness that vary from short to long term. In each case, I take the mean strategy in the population after 5000 generations, and ask how cooperative it is when playing against itself. The three measures are:

Short term: $P_0(CC)$, the probability that two average players will achieve mutual cooperation on the first play of an interaction. This measure uses only the $c$ element of the strategy vector ($c^2 = P_0(CC)$), and ignores the other four elements ($t, r, p, s$).

Intermediate term: $P_5(CC)$, the expected proportion of mutual cooperation ($CC$ events) in the first five plays in an interaction.

Long term: $P_{stat}(CC)$, the stationary probability of mutual cooperation (the relative frequency of $CC$ after a long sequence of plays). This measure uses only the “reactive” elements of the strategy vector ($t, r, p$ and $s$), and ignores the initial cooperativeness ($c$).

Figure 2 shows the effects of future value ($w$) and accumulation ($n$) on these measures of cooperativeness. All three measures show a similar pattern. I found high levels of cooperation for all values of the future value parameter, $w$, when accumulation periods were large ($n = 10$ and 25). Similarly, I found high levels of cooperation for all values of the accumulation period ($n$) when the future value parameter $w$ was sufficiently high ($w = 0.99$). This similarity supports the idea, discussed above, that accumulation (high $n$) and “repetition” (high $w$) promote cooperative behavior in a logically similar way. One sees this pattern most clearly for the shortest term measure of cooperativeness ($P_0(CC)$). The pattern is less clear, for the longest term measure ($P_{stat}(CC)$); plots of this measure [Fig. 2(c)] show some surprises. First, we see elevated levels of $P_{stat}(CC)$ when there is no accumulation $n = 1$, and a very low future value term ($w = 0.01$). This anomaly arises when $w = 0.01$ because the animal’s time horizon is so short that only the $c$ element of the strategy matters (it evolves toward zero), while the other elements of the strategy vector drift about irrerelevantly leading to $P_{stat}(CC)$ near the random level of 1/4. Intriguingly, the same logic does not seem to apply when there is a small accumulation period ($n = 5$) even when $w$ is equally low. Evidently, $n = 5$ lengthens the time horizon enough to make the other elements of the strategy relevant, but not enough to encourage cooperation. Second, when $w = 0.99$ we see higher levels of $P_{stat}(CC)$ with accumulation levels lower ($n = 1$ and 5); this is the first hint we see of a general pattern in which high accumulation periods (high $n$) favor short-term cooperativeness, and high future value (high $w$) favors long-term cooperativeness. The analysis of strategies in the next section will help to clarify this pattern.

EVOLVED STRATEGIES

Figure 3 shows a somewhat bewildering summary of strategies that evolved in each of the 16 cases I studied. The inverted L-shaped region formed by the top row and left-hand column represent high $n$ and high $w$ cases that are the “most cooperative” strategies. Inspection of this region suggests that cooperative strategies are quite similar to one another, with high $r$ (probability of cooperating after mutual cooperation) and high $c$ (probability of initial cooperation) and smaller values for the other three elements of the strategy. A exploratory analysis showed a striking bimodality in cooperativeness; with strategies being either “cooperators” or “defectors”.

Table 1 shows a comparison of the most and least cooperative strategies, where “most cooperative” means the median of the most cooperative third of the strategies (cooperativeness measured as $P_5(CC)$), and “least cooperative” means the median of the least cooperative third.

For the evolved cooperator, the initial and stationary distributions differ markedly (Table 1); the relative frequency of mutual cooperation erodes dramatically as play continues, while the relative frequency of mutual defection builds from near zero at first, to over one-third (cf. initial and stationary distributions). The transition matrix shows that departures from mutual cooperation (state $R$) are most likely to lead to the weakly attractive mutual defection state ($P$). The “cooperator” achieves a long-run preponderance
FIG. 2. The figures shows median values of each of three measures of cooperatives. I conducted 20 independent runs at each combination of $w$ and $n$, so the plotted points shows medians of these 20 values. Points corresponding to different accumulation periods ($n$ values) are marked with distinct symbols: $(\times)$ for $n = 25$, $(+)$ for $n = 10$, $(\triangle)$ for $n = 5$, and $(\square)$ for $n = 1$.

of mutual cooperation because mutual $C$ persists once established; however, once players leave this state the strategy lacks a clear path back to mutual cooperation. In contrast, the evolved defector is remarkably consistent.

This dynamic pattern shown by the evolved cooperator (i.e. departure from $CC$ leads to $DD$, leads weakly back to $CC$) reminds one of Pavlov (Nowak & Sigmund, 1993), except that the link from $DD$ back to $CC$ is much weaker. The evolved cooperator also resembles retaliator, a strategy that cooperates initially, cooperates with a cooperator, but permanently switches to defection if its opponent defects. Its not reasonable to think of our evolved cooperator as a mixture of Pavlov [in 5-element notation $(0,1,1,0,1)$], and retaliator $(0,1,0,0,1)$; a bit closer to retaliator than to Pavlov.
IS ALL COOPERATION THE SAME?

As outlined above, differences in \( w \) and \( n \) lead to large differences in cooperativeness, (Fig. 2), and these differences arise from a striking bimodality in strategies as shown in Table 1. In the next few paragraphs, I consider a more subtle level of strategic variation. I ask whether there are effects of future value \( w \) and accumulation \( n \), within the class of “cooperative strategies” (defined as strategies with \( P_5(CC) \) values in the upper one-third).

I begin by contrasting two extremes to help frame the analysis. Table 2 compares the conventional “high future value/no accumulation case” \((w = 0.99, n = 1)\), with a “high accumulation/low
### Table 1
Comparison of most and least cooperative strategies

<table>
<thead>
<tr>
<th>Least cooperative</th>
<th>Most cooperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_s(\text{CC}) &lt; 0.03$</td>
<td>$P_s(\text{CC}) &gt; 0.69$</td>
</tr>
<tr>
<td><strong>Strategy</strong></td>
<td><strong>Strategy</strong></td>
</tr>
<tr>
<td>$r$</td>
<td>$t$</td>
</tr>
<tr>
<td>0.12</td>
<td>0.18</td>
</tr>
</tbody>
</table>

#### Transition matrix

<table>
<thead>
<tr>
<th>From</th>
<th>To</th>
<th>$T$</th>
<th>$R$</th>
<th>$P$</th>
<th>$S$</th>
<th>$T$</th>
<th>$R$</th>
<th>$P$</th>
<th>$S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>0.17</td>
<td>0.23</td>
<td>0.06</td>
<td>0.10</td>
<td>0.15</td>
<td>0.04</td>
<td>0.20</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>$R$</td>
<td>0.02</td>
<td>0.14</td>
<td>0.00</td>
<td>0.02</td>
<td>0.03</td>
<td>0.92</td>
<td>0.07</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>$P$</td>
<td>0.71</td>
<td>0.40</td>
<td>0.88</td>
<td>0.71</td>
<td>0.67</td>
<td>0.00</td>
<td>0.53</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>$S$</td>
<td>0.10</td>
<td>0.23</td>
<td>0.06</td>
<td>0.17</td>
<td>0.15</td>
<td>0.04</td>
<td>0.20</td>
<td>0.15</td>
<td></td>
</tr>
</tbody>
</table>

#### Distributions

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Stationary</th>
<th>Initial</th>
<th>Stationary</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>0.08</td>
<td>0.07</td>
<td>0.07</td>
<td>0.12</td>
</tr>
<tr>
<td>$R$</td>
<td>0.01</td>
<td>0.01</td>
<td>0.86</td>
<td>0.42</td>
</tr>
<tr>
<td>$P$</td>
<td>0.83</td>
<td>0.86</td>
<td>0.00</td>
<td>0.34</td>
</tr>
<tr>
<td>$S$</td>
<td>0.08</td>
<td>0.07</td>
<td>0.53</td>
<td>0.67</td>
</tr>
</tbody>
</table>

### Table 2
Comparison of cooperation achieved via accumulation and future value

<table>
<thead>
<tr>
<th>Future value cooperator</th>
<th>Accumulation cooperator</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 1$ and $w = 0.99$</td>
<td>$n = 25$ and $w = 0.01$</td>
</tr>
<tr>
<td><strong>Strategy</strong></td>
<td><strong>Strategy</strong></td>
</tr>
<tr>
<td>$r$</td>
<td>$t$</td>
</tr>
<tr>
<td>0.17</td>
<td>0.14</td>
</tr>
</tbody>
</table>

#### Transition matrix

<table>
<thead>
<tr>
<th>From</th>
<th>To</th>
<th>$T$</th>
<th>$R$</th>
<th>$P$</th>
<th>$S$</th>
<th>$T$</th>
<th>$R$</th>
<th>$P$</th>
<th>$S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>0.14</td>
<td>0.05</td>
<td>0.25</td>
<td>0.14</td>
<td>0.20</td>
<td>0.04</td>
<td>0.20</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>$R$</td>
<td>0.03</td>
<td>0.90</td>
<td>0.23</td>
<td>0.03</td>
<td>0.03</td>
<td>0.92</td>
<td>0.07</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>$P$</td>
<td>0.69</td>
<td>0.00</td>
<td>0.27</td>
<td>0.69</td>
<td>0.66</td>
<td>0.00</td>
<td>0.53</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>$S$</td>
<td>0.14</td>
<td>0.05</td>
<td>0.25</td>
<td>0.14</td>
<td>0.11</td>
<td>0.04</td>
<td>0.20</td>
<td>0.20</td>
<td></td>
</tr>
</tbody>
</table>

#### Distributions

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Stationary</th>
<th>Initial</th>
<th>Stationary</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>0.13</td>
<td>0.11</td>
<td>0.07</td>
<td>0.12</td>
</tr>
<tr>
<td>$R$</td>
<td>0.71</td>
<td>0.57</td>
<td>0.85</td>
<td>0.42</td>
</tr>
<tr>
<td>$P$</td>
<td>0.03</td>
<td>0.21</td>
<td>0.01</td>
<td>0.34</td>
</tr>
<tr>
<td>$S$</td>
<td>0.13</td>
<td>0.11</td>
<td>0.07</td>
<td>0.12</td>
</tr>
</tbody>
</table>

The future value case ($w = 0.01$, $n = 25$). Table 2 shows that the accumulation cooperator is more likely to cooperate initially, and less likely to cooperate after an instance of mutual defection (state $P$). Accumulation emphasizes initial cooperation and de-emphasizes returning to mutual cooperation after a perturbation. In contrast, the conventional, future value cooperator seems to emphasize returning to cooperation after a mutual defection, and de-emphasizes initial cooperation. Crudely speaking, the accumulation cooperator is a more “reliant-like” strategy that appears to arise because success in the first bout is the key to success in the accumulated game (a “whatever gets you through the bout” strategy); the future value cooperator is more
“Pavlov-like” taking a long-term view, apparently at the expense of initial cooperation.

In a more formal exploration of the effects of accumulation and future value, I performed a two-way analysis of variance on each of the five elements of the strategy (i.e. five separate ANOVAs using arcsin transformed probabilities as the dependent measure), still restricting my attention to the one-third most cooperative strategies. In each case, I examined the residuals of the fitted ANOVA model for signs of non-normality and for evidence of systematic changes in variance. These examinations justify my use of ANOVA for these analyses. I found that neither w nor n had any effect on the s, r or t elements of cooperative strategies—partially confirming the impression that cooperative strategies are similar. I did find significant differences in initial cooperativeness (c) and in the tendency to cooperate following mutual defection (p), in agreement with the simpler comparison made earlier.

I found simple main effects of w and n on c, but no interaction (w: F_{3, 98} = 34.0, P < 10^{-7}; n: F_{3, 98} = 5.9, P < 10^{-3}). The probability of initial cooperation, c, is markedly smaller when the future value parameter is highest [Fig. 4(a)]. There is a less marked, but significant, trend toward smaller c values at high levels of accumulation [again see Fig. 4(a)].

Analysis of variance reveals a more complex pattern for p, the probability of cooperation following mutual defection: a significant w by n interaction (F_{3, 98} = 3.7, P < 0.01) and significant main effects of both w and n (w: F_{3, 98} = 16.6, P < 10^{-3}; n: F_{3, 98} = 3.4, P < 0.05). Again, one sees a marked difference between the highest w case (w = 0.99) and all the others, with larger p values occurring when w = 0.99 regardless of n’s magnitude. Superimposed on this we see an intriguing pattern for p, the probability of cooperation following mutual defection in which p decreases with accumulation (n) when future value is highest (w = 0.99), but decreases with accumulation at lower future values [Fig. 4(b)].

Figure 4 shows several intriguing patterns. Firstly, one see a striking difference between the highest w cases (w = 0.99) and all others. When w = 0.99, c is lowest and p highest: in words, high future value emphasizes long-term cooperation— i.e. returning to mutual cooperation from mutual defection—and de-emphasizes initial cooperation. It makes sense to view this difference (i.e. w = 0.99 vs. w < 0.99) as the difference between long-term and short-term cooperation. Notice, for example, that with the low w cases (w < 0.99) strategies shift toward long-term behavior as n increases (i.e. c decreases and p increases with n).

THE c/p TRADEOFF

The pattern seen here suggests a trade-off between c and p; with high c and low p combinations occurring when accumulation favors short-term cooperativeness, and low c-high p combinations occurring with high future value favors long-term cooperativeness. Why should such a trade-off exist? Why not cooperative initially and in the long-run? To answer these questions, consider a situation with low future value and an intermediate level of accumulation. This situation places a premium on extracting the most benefit from a relatively short interaction, the best way to accomplish this is to establish mutual cooperation initially; if however, someone defects then the other player must move quickly to avoid a string of sucker's payoff. Once mutual defection is established there is typically little time left in the bout, and attempts to re-establish mutual cooperation via generosity are a gamble that may lead to mutual cooperation or to the sucker’s payoff. Without the deadline of bout termination the future value cooperator has the luxury of relying on the dynamics of the game to establish mutual cooperation even if this means accepting a few sucker’s payoffs.

Discussion

Elsewhere (Clements & Stephens, 1995; Stephens et al., 1995), my colleagues and I have argued that strong preferences for immediate reward (discounting) may limit the IPD's value as a model of animal cooperation. In this paper, I ask whether we can re-organize games to overcome strong discounting. The simplest way to achieve cooperation in the face of discounting is via by-product mutualism. In mutualism there is no temptation to cheat, so mutually beneficial joint action is in both player's short-term best interests. Clearly, cooperation via mutualism is
fundamental; and as a growing number of authors have suggested, probably the basic form of cooperation in animal behavior (see, for example, Pusey & Packer, 1997; Dugatkin, 1997; Stephens et al., 1997).

There is, however, considerable theoretical interest in how one might achieve cooperation in the face of a temptation to cheat. I suggest, here, a relatively simple mechanism to achieve cooperation even when there is a temptation to cheat and strong discounting. I argue that delaying all the benefits from a sequence of plays via accumulation, can eliminate the immediate temptation to cheat and enhance the relative value of a conditionally cooperative strategy. In broad outline, the cumulative payoff game is similar to the conventional Iterated Prisoner’s Dilemma with the accumulation period \((n)\) playing a role.
similar to the conventional model’s probability of repeated play (cf. e.g. Matrices 2 and 3). Although accumulation increases the relative value of conditionally cooperative strategies, the delay inherent in cumulative games reduces the overall value of the interaction. Cumulative games, therefore, are likely to play their most significant role in promoting cooperation when a player’s options elsewhere are poor. Many readers will find the idea that “hard times” promote cooperativeness to be an intuitively appealing one; while this is an immediate consequence of the accumulated games model, it is less clear how this prediction might arise in either conventional iterated games, or in simple mutualism.

The idea that accumulation enhances cooperativeness agrees with our intuition about human interactions. I encourage my colleagues to cooperate by reminding them of the benefits to be obtained at the end of a successful interaction. If I want to guarantee that my child will be well-behaved, I structure my bribe so that the reward follows his good behavior: “if you’re very good I’ll buy you a toy after grandma’s visit”. The idea presented here is that structuring payoffs in this way, is a mechanism to overcome the temptations of immediacy and the selfishness that immediate gratification implies.

COMPARISON WITH THE CONVENTIONAL IPD

Unilateral cooperation (All-C) is unstable in both the cumulative and conventional iterated games: a cooperator must somehow punish defection if mutual cooperation is to be stable. In the conventional game a defector realizes the large payoff $T$ immediately, it may get less in the long-run because its opponent tries to punish it, but it has $T$ now. In the accumulated game a large payoff $T$ may accrue, but it is not available until the end of the bout; the defector’s opponent has time to make its punishment count! This assumes, of course, that the “defection” occurs before the end of the bout. Why should not a player defect on the last play of a bout of accumulation? After all, when only a single play remains, your opponent cannot punish you before you collect your “winnings”. Following the familiar backward-induction argument (see for example Mesterton-Gibbons, 1992) of the finitely repeated PD, one might reasonably expect All-D to predominate in games with fixed accumulation periods. While this scenario seems logical, All-D did not predominate in my simulated evolution of strategies. This may be because of the threat of punishment in the next bout, or because the form of the strategies studied here did not permit a rule such as “defect only on play $n - 1$”. However, one possible advantage of the retaliator-like strategies that arose under accumulation, is that they reduce a player’s chances of being suckered at the end of a bout. In any case, the fixed bout-length model studied here is a simplification. A more complete model would overcome this objection by adding some uncertainty to bout length.

NATURAL ACCUMULATED GAMES?

The reader might reasonably ask whether accumulated games arise in nature. While I did not intend the situation modeled here to represent any particular natural situation, I believe that there may be many natural situations in which interacting individuals achieve substantial benefits only after engaging in a sequence of interactions. Of course, we do not require that benefits accumulate is some literal “bin”, the key idea is that continued cooperative joint action somehow increases the benefits available at some future time. For example, a pair of birds engaged in cooperatively feeding their nestlings can only achieve the benefit of fledged young after a protracted sequence of feeding events. Consider a pair of carnivores, say liensesses, cooperatively stalking a large prey animal. Stalking is not a single action but a sequence of approach and pause actions; and a successful kill can be achieved only after a sequence of “approaches”. The vocal duetting long-tailed manakins (Chiroxiphia linearis, McDonald & Potts, 1994) provides a spectacular and informative example. In this system, pairs of males attract females via elaborate vocal duetting on small traditional lekking arenas. However, only the $z$-male obtains any matings. McDonald and Pott’s interpretation is that the $\beta$ accepts its secondary role in order to establish itself as the new owner of the lekking arena when the $z$ male dies. In contrast to the model studied here, the long-tailed manakin
eventually inheriting the arena, and (b) the repeatedly interactions increase its likelihood of faces an accumulated game if we suppose (a) that some mechanism of retaliation that reduces the accumulated game for the player. The player faces an accumulated game if we suppose (a) that repeated interactions increase its likelihood of eventually inheriting the arena, and (b) the has some mechanism of retaliation that reduces the player’s chances of inheritance.

EXTENSIONS OF THE MODEL

This discussion brings us naturally to the question of how one might extend and improve upon the current model. The manakin example suggests that asymmetries are important. Notice, however, that in the manakin scenario the players not only face different game matrices, but different “blends” of immediate and delayed consequences. This suggests that it will be important to consider matrices that have some immediate, and some accumulating (or delayed) consequences. Another interesting class of extensions concerns the rules of bout termination (to simplify terminology here, I am using bout to mean bout of accumulation, rather that the duration of play). As mentioned above, the finite-length bouts of accumulation may act against cooperative behavior, so one might explore models of random bout length. There are, however, many other ways in which bouts might end. If, as I have argued, long bouts of accumulation encourage cooperativeness, then this opens the possibility that a player might manipulate his opponent’s cooperativeness by exercising some control over when benefits are distributed. In its most extreme form, this is the bribery scenario mentioned above. Another simple possibility is that the level of accumulated benefits determines the timing of payoff distribution: that is, benefits are realized not after some fixed number of plays, but when a specific payoff condition is reached, e.g. when a kill is made.

If games are organized into bouts of accumulation, then this raises the possibility that strategies might reflect this bout structure in some way. For example, an animal might be more (or less) forgiving between bouts than within bouts, or it might “start fresh” at the beginning of each new bout. Clearly, the idea of bouts of accumulation introduces a new wrinkle into strategies that may be a productive direction for further study.

ALTERNATIVES TO THE IPD

As mentioned earlier, a handful of authors have questioned the IPD’s empirical usefulness (e.g. Clements & Stephens, 1995; Heinishon & Packer, 1995; Stephens et al., 1995). An even larger group, however, view the IPD scenario as too restrictive to justify its historical role as the central paradigm of animal cooperation (Dugatkin et al., 1992; Mesterton-Gibbons & Dugatkin, 1992; Connor, 1995a; Pusey & Packer, 1997). In response, I presume, to these sentiments there is a growing literature of alternative models. For example, Clutton-Brock & Parker (1995) suggest that the simple mechanism of punishment deserves a more careful consideration; similarly, Nowak & Sigmund (1998) have recently reformulated the venerable idea of indirect reciprocity (Alexander, 1987; Boyd & Richerson, 1989; Pollock & Dugatkin, 1992). Taken together this growing family of models represents an important step toward an enlarged view of animal cooperation. While the model advocated here is very much in sympathy with this enterprise, it is somewhat different. Rather than advocating an entirely new scenario for cooperation, the accumulated games model suggests how we might modify the conventional IPD to overcome the specific problem of temporal discounting.

Connor’s (1995a, b) model of reciprocity via parceling is similar to the accumulation model in several respects. To outline these similarities, I briefly explain the parceling model. Connor develops his model in the context of allogrooming in ungulates (Hart & Hart, 1992). In this system, one individual initiates bouts of reciprocal grooming, when it delivers a brief bout of grooming to another. The recipient may then choose to reciprocate with a brief stint of grooming, and so the cycle repeats until one player terminates it. Suppose that the fitness benefits of grooming are strongly nonlinear, approaching a step function: small amounts of grooming fail to satisfy a recipient’s needs for parasite removal, with benefits rising sharply when prolonged grooming achieves a satisfactory state.

Obviously, an initiator who satisfied a recipient’s need in one long, unconditional bout of grooming would invite cheating (why reciprocate when you have everything you need?). By the
converse logic an initiator can reduce the temptation to cheat by stopping its bout of grooming before the recipient has all the grooming it wants. The recipient may still be tempted to cheat if it is easy to find a string of gullible initiators. If, however, finding donors is costly, the initiator may be able to reduce its bout of grooming to a size that effectively eliminates the temptation to cheat.

The step function that relates benefit to grooming received, and the parceling of grooming into brief bouts creates an accumulated game; because these two factors combine to create a situation where a sequence of reciprocated bouts is required to achieve a meaningful “payoff”. I remark that my description of Connor’s model differs somewhat from the original. The idea of step function like benefit function is not explicit in Connor’s formal, but implied by the idea that an individual receiving a small initiation bout of grooming is still far from satisfied. In addition, Connor correctly emphasizes the need for uncertainty in the duration of interactions, which I have de-emphasized in order to simplify the situation, and to stress the similarity with the fixed bout accumulation model presented above.

Conclusions

I have argued that one can promote the relative value of cooperative action by arranging games so that payoffs accumulate over a sequence of actions. This agrees with our intuition: a bribe will be more effective if it is withheld until after the desired action. In addition, cooperation in these cumulative benefit games should be most important in difficult times. The idea of cumulative benefit games has implications for future experiment and observational work. For example, my laboratory is conducting an experiment in which food benefits accumulate, visible but unavailable over a sequence of plays. For the field worker, the possibility of cumulative benefit games suggests that careful observation of the temporal relationship between action and consequence may enrich our understanding of natural cooperative systems, and provide much needed guidance for future models. Finally, cumulative benefit games emphasize a new dimension of strategic interaction: can animals encourage others to cooperate by sub-dividing games (as Conner has suggested), or by exerting some other form of control over the timing of their partner’s benefits?

The National Science Foundation supported the research described here (IBN-9507688 & IBN-9896102). I am grateful for the thoughtful comments of my colleagues at the Universities of Nebraska and Minnesota, especially Jeff Stevens and Craig Packer.

REFERENCES


of nine probabilities, \( u = (t, r, p, s, c, t', r', p', s') \). Where \( t \) represents the probability of cooperating given that the focal player defected and its opponent cooperated in the previous play (i.e. a benefit of \( T \) accrues to the focal player). The \( r, p \) and \( s \) probabilities represent the probabilities of cooperation following benefit accruals of \( R, P \) and \( S \), respectively. The term \( c \) represents the probability of initial cooperation. The parameters \( t, r, p, s \) apply within the bout of \( n \) plays, while the analogous terms denoted with prime symbols \( t', r', p', s' \) apply between bouts (thus allowing for the possibility that a bout-structured game, might have a bout-structured strategy).

If we have two such strategy vectors in hand, say \( S_1 \) and \( S_2 \), we can calculate three entities that characterize the strategy interaction: (i) \( M \) the within-bout transition matrix, (ii) \( B \) the between-bout transition matrix, and (iii) \( c \) the initial distribution of states. One can readily calculate these entities from the definitions of the components of the strategy vector (see, e.g. Stephens et al., 1995). Finally, I define \( v \) to be the column vector of payoffs \((T \ R \ P \ S)\). Note that I define these entities so that they reflect the states experienced by the \( S_1 \) player.

**Bout-by-Bout Construction of Benefits**

The expected benefit from the first bout of \( n \) plays is

\[
v(I + M + M^2 + \cdots + M^{n-1})c,
\]

which must be devalued by \( w^{n-1} \). The expected benefits from the second bout of \( n \) can be calculated in exactly the same way, once we recognize that the “initial state” at the beginning of this bout is \( BM^{n-1}c \) and not simply \( c \), i.e.

\[
v(I + M + M^2 + \cdots + M^{n-1})BM^{n-1}c,
\]

which must be devalued by \( w^{2n-1} \). For the third bout, we have

\[
v(I + M + M^2 + \cdots + M^{n-1})(BM^{n-1})^2c
\]

devalue by \( w^{3n-1} \) and so on. The \( u \) player’s total benefit is

\[
w^{n-1}v(I + M + M^2 + \cdots + M^{n-1})c + w^{2n-1}v(I + M + M^2 + \cdots + M^{n-1})BM^{n-1}c
\]
\[ + w^{3n-1} v(I + M + M^2 + \cdots + M^{n-1})(BM^{n-1})^2 c + \cdots. \quad (A.1) \]

Rearranging, we have
\[ w^{n-1} v(I + M + M^2 + \cdots + M^{n-1}) \]
\[ \times [I + w^n BM^{n-1} + w^{2n}(BM^{n-1})^2 + \cdots] c. \]

A standard result from linear algebra (Burden & Faires, 1985) show that the infinite series
\[ I + w^n BM^{n-1} + w^{2n}(BM^{n-1})^2 + \cdots = (I - w^n BM^{n-1})^{-1}, \]
so we may write the total benefit compactly as
\[ w^{n-1} v(I + M + M^2 + \cdots + M^{n-1}) \]
\[ \times (I - w^n BM^{n-1})^{-1} c. \quad (A.2) \]

I remark that matrix inversion is not used in the numerical calculation reported in the text. Numerical matrix inversion is a notoriously sensitive business, while there are many robust algorithms for the numerical solution of linear systems. My computer programs, therefore, first solve the system
\[ (I - w^n BM^{n-1}) y = c \]
for \( y \), and then calculate the total benefit using
\[ w^{n-1} v(I + M + M^2 + \cdots + M^{n-1}) y. \quad (A.3) \]

I remark, also, that considerable computable effort can be saved by recognizing that there exists a fixed relationship between the payoff received by the focal player and the payoff received by its opponent. So, we can calculate the term
\[ w^{n-1} (I + M + M^2 + \cdots + M^{n-1}) \]
\[ \times (I - w^n BM^{n-1})^{-1} c \]
only once; multiply by row vector \( (T, R, P, S) \) to find the focal player’s total benefit, and by \( (S, R, P, T) \) to find the opponent’s benefit.

**APPENDIX B**

**Evolutionary Algorithm**

Create Initial Population. Create 200 vectors of 5 probabilities, each entry uniformly distributed between 0 and 1.

Repeat for 5000 generations:

1. Zero fitness vector. Set the fitness of each member of the population to zero.
2. Play the game.

   **Repeat for three bouts of play:**
   
   Pair up players Randomly form 100 pairs from the 200 members of the population.
   
   Calculate fitness consequences of pairwise play. Use “chromosomes” to calculate transition matrix \( (M) \) and initial state vector \( (c) \), then apply eqns (1) and (2) to determine fitness increment. Add fitness increment to appropriate element of fitness vector.
   
   **End bout loop.**
3. Breed a new population

   Form parent pool. Find 20 most fit individuals and assign them to the “parent pool”.

   Repeat to form 200 new individuals.

   Choose Parents. Choose two individuals from the pool at random.

   Form Offspring. Combine parental elements in “sexual haploid” fashion: i.e. offspring gets Mom’s \( t \)-value with probability 1/2 and Dad’s \( t \) otherwise, and so on for all 5 elements of the “chromosome”.

   Mutate Offspring Mutation occurs via element-wise addition of a normally distributed random variable (mean 0, S.D. 0.1). If this makes an element become negative it is “chopped” to 0; similarly an element that exceeds 1 is “chopped” back to 1.

   **End breeding loop.**

**End generation loop.**

Report results. Statistical summaries of a “winning” chromosomes.