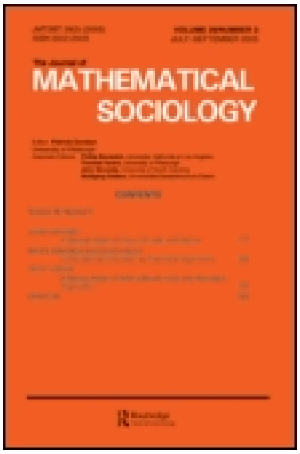

The Journal of Mathematical Sociology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gmas20

How to make cooperation the optimizing strategy in a two-person game

Bernard Grofman [a] & Jonathan Pool [b]
[a] University of California, Irvine
[b] University of Washington
Version of record first published: 26 Aug 2010.

J. Mathematical Sociology
1977, Volume 5, pp. 173-186

HOW TO MAKE COOPERATION THE OPTIMIZING STRATEGY IN A TWO-PERSON GAME

Bernard Grofman

University of California, Irvine

and

Jonathan Pool

University of Washington

This paper demonstrates the existence of a partial tit-for-tat (matching) strategy which, when used by one player in an iterated "Prisoner's Dilemma" game, will induce a response of pure cooperation in the other player if that player behaves optimally. The minimum matching frequency of such a strategy is shown to be monotonically related to the Rapoport-Chammah "Cooperation Index."

A two-person game, shown in Figure 1, is a simple situation in which to study the conditions of cooperative and competitive behavior. One of the most interesting such games, and the one that has been by far most extensively examined (Oskamp, 1971), is the symmetric Prisoner's Dilemma game, defined by three restrictions on the values of the payoffs:[1]

$$a = a', \ b = b', \ c = c', \ d = d' \tag{1}$$

$$c > a > d > b \tag{2}$$

$$2a > b + c > 2d \tag{3}$$

Research has focused both on the theoretical properties of the game (Anatol Rapoport, 1960, 1966; Anatol Rapoport and Chammah, 1965; Amnon Rapoport, 1967; Axelrod, 1967; Howard, 1966a, 1966b, 1970; Shubik, 1970; Grofman, 1975; Grofman and Pool, 1975) and on how people actually play it (see Anatol Rapoport and Chammah, 1965, and review essays and bibliography in Wrightsman, O'Connor and Baker, 1972).

Experimental studies have examined how both personality differences and structural characteristics, particularly variations in the payoff matrix, affect "cooperation," i.e., the choice of alternative 1 by one or both players (see literature

[1] Our defining characteristics for the P.D. game are those given in Rapoport and Orwant (1962) plus the usual stipulation of symmetry. As Oskamp (1971) has pointed out, many games labeled as P.D. in the experimental gaming literature do not in fact satisfy conditions (2) and (3).

$$
\begin{array}{ccc}
 & & \text{Column} \\
 & & \begin{array}{cc} 1 & 2 \end{array} \\
\text{Row} & \begin{array}{c} 1 \\ \\ 2 \end{array} & \begin{bmatrix} (a,\ a') & (b,\ c) \\ E_{11} & E_{12} \\ (c',\ b') & (d,\ d') \\ E_{21} & E_{22} \end{bmatrix}
\end{array}
$$

Figure 1. General two-person game. (Note: This figure employs standard notation, i.e., "Row" and "Column" are the two players, E_{ij} is the outcome when Row chooses alternative i and Column chooses alternative j, and (i, j) shown over an outcome indicates the payoffs to Row and Column, respectively, associated with that outcome.

review in Wrightsman, O'Connor and Baker, 1972, pp. 57-65). In general, however, there has been little theoretical underpinning to these studies; rather they have manipulated fairly arbitrarily selected payoff matrix characteristics. An important exception is those studies which have used as an independent variable the Rapoport and Chammah (1965) "cooperation index." This index, given by

$$\text{Cooperation Index} = \frac{a - d}{c - b}\,, \tag{4}$$

has been asserted by Rapoport and Chammah (1965) to be one of the two basic ratios which (a) can be used to characterize a P.D. game, (b) are invariant with respect to the addition of a constant to all matrix entries, and (c) make use of *all* the available information about the payoff matrix. The utility of this index as a predictor of cooperation in iterated P.D. games has been reasonably well demonstrated (Rapoport and Chammah, 1965; Terhune, 1968; but also see Axelrod, 1967).

There have also been numerous studies dealing with the effect on "cooperation" of variations in the strategy used by the experimenter, who (unbeknownst to the subject) assumes the role of the other player (see the comprehensive review in Oskamp, 1971). While little or no theoretical rationale is customarily given as to why we might expect one strategy to be more successful in eliciting cooperative behavior than another, many data have been collected. Unfortunately, very few of these studies have dealt with the effect of response-contingent strategies. Instead they have usually used preprogrammed strategies according to which the simulated opponent's moves are independent of the subject's.

As far as we are aware, no theoretical model has ever been developed capable of predicting the effects on cooperative behavior of the *interaction* between the reward structure of the matrix and the nature of the strategy used by the other player, nor has there been other than incidental experimental investigation of this interaction. It is this interaction with which this paper will be concerned. We shall show that the Rapoport and Chammah (1965) Cooperation Index can be derived as a special case of a more general model which deals with the interaction between payoff entries and the (response-contingent) strategy of the other player.

In this analysis we shall confine ourselves to iterated Prisoner's Dilemma games in which both players are following what we have called "class 1" decision rules (Grofman and Pool, 1975), i.e., rules which define a player's probability of choosing alternative 1 on trial n of the game as a function only of the $(n-1)^{th}$ trial outcome. A class 1 decision rule is thus the simplest kind of strategy that responds to the other player's prior behavior.

Within this class of rules, we shall focus on a yet simpler subclass that defines the probabilities of choosing 1 and 2 as a function only of the *other* player's move on trial n-1, and does so symmetrically with respect to the other player's two options. Such rules are called partial tit-for-tat (TFT) strategies and are defined each by a probability, p, that the player's n^{th} move will be the same as the other player's $(n-1)^{th}$ move. Although some experiments have investigated the effects on cooperation of a TFT strategy employed by a simulated opponent, they have not generated the information needed to compare rates of cooperation for varying levels of p (Grofman and Pool, 1975). In one study, however, a closely related relationship was discovered (Pool and Grofman, 1975), suggesting that cooperation would vary monotonically and positively with the p of the other player's TFT strategy. We shall now show why this should be the case if the subject is acting so as to maximize his or her own payoff in the long run.

Consider a player facing another player who is using a partial TFT strategy as defined above. It is easy to show under what conditions the former will prefer a pure strategy of 1 on every move to a pure strategy of 2 on every move. The expected long-run (average) payoffs yielded to the user of these two strategies are $ap + (1-p)b$ and $(1-p)c + pd$, respectively (Grofman and Pool, 1975). So if

$$ap + b(1-p) > c(1-p) + pd, \tag{5}$$

then in an iterated P.D. game a strategy of pure cooperation will be preferable to a strategy of pure defection (iterated minimax) against a player known to be using a partial TFT strategy. We may readily show that there always exists a p such that inequality (5) is satisfied, and that any such p must be $> \frac{1}{2}$.

Lemma 1: For a,b,c,d satisfying the constraints in (2) and (3), there always exists a p, $p > \frac{1}{2}$, such that (5) is satisfied.

Proof: We may rewrite (5) as

$$p > \frac{c - b}{a - b + c - d}, \tag{6}$$

because (2) requires $a - b + c - d$ to be positive. The maximum value for p is clearly 1. If

$$1 > \frac{c - b}{a - b + c - d}, \tag{7}$$

we are done with the first half of our proof. But (7) must be true, since (3) requires that $a > d$. The denominator is thus greater than the numerator, which is also positive by (2).

To show that $p > \frac{1}{2}$ is equally simple. We must verify that

$$\frac{c - b}{a - b + c - d} > \frac{1}{2}. \tag{8}$$

Since, by (2), $c > a$ and $d > b$, it follows that $c + d > a + b$, hence $c - b > a - d$. Thus

$$\frac{c - b}{(c-b) + (a-d)} > \frac{c - b}{(c-b) + (c-b)} = \frac{1}{2}. \qquad \text{Q.E.D.}$$

Lemma 1 tells us that there always exists a partial tit-for-tat strategy which can induce a rational opponent to prefer pure cooperation over pure defection (i.e., the choice of alternative 2, as required by an iterated minimax decision-rule) and that it must involve more than a 50% level of reinforcement. (Obviously, a 50% TFT strategy is the same as random choice, and a less than 50% TFT strategy is one involving more "uncopying" than copying of the opponent's last move.) Note that in our proof we have not used the third of the defining characteristics of the symmetric P.D. game.

The above results may be related to the Rapoport and Chammah (1965) Cooperation Index. Expression (5) may be rewritten as

$$\frac{a - d}{c - b} > \frac{1 - p}{p}. \tag{9}$$

The left side of this inequality is the Rapoport-Chammah Cooperation Index given in expression (4). The greater the left side of expression (9), the smaller is the value of p ($p > \frac{1}{2}$) that is needed to produce a situation in which pure cooperation is preferred to pure defection. Similarly, against any given level of p, the greater the value of $(a-d)/(c-b)$, the greater the attractiveness of the strategy of pure cooperation vis-a-vis that of pure defection. If a player sees the other player's responses as contingent on his or her own behavior, then the long-run consequences of defecting may outweigh the short-run gain. The expression $(a-d)/(c-b) - (1-p)/p$ gives and index of the (marginal expected asymptotic) gain to be had from cooperation (as opposed to defection) against another player who is using a partial tit-for-tat strategy. Thus, we would expect that for given p, or perceived p, the higher the value of $(a-d)/(c-b)$, the greater the extent of cooperative behavior.

In order to strengthen the above conclusions it would be useful to show not just that pure cooperation is preferred to pure defection when inequality (5) holds, but also that the pure cooperation strategy is overall optimal, i.e., has the highest (long run expected average) payoff of any class 1 strategy, for p sufficiently large.

The general expression for a row class 1 strategy is given in Figure 2, where x_{ij} represents row's probability of choosing alternative 1 given outcome E_{ij} on the last trial.

$$\begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{cccc} E_{11} & E_{12} & E_{21} & E_{22} \\ \end{array}$$

$$\begin{array}{cc} & \begin{array}{cccc} E_{11} & E_{12} & E_{21} & E_{22} \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} x_{11} & x_{12} & x_{21} & x_{22} \\ 1 - x_{11} & 1 - x_{12} & 1 - x_{21} & 1 - x_{22} \end{bmatrix} \end{array}$$

Figure 2. Row's class 1 decision-rule.

The combination of the general row strategy in Figure 2 and the column TFT strategy of Figure 3 gives rise to the transition matrix below (see Figure 4). The steady state vector

$$\begin{array}{cc} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} E_{11} \\ E_{12} \\ E_{21} \\ E_{22} \end{array} & \begin{bmatrix} p & 1 - p \\ p & 1 - p \\ 1 - p & p \\ 1 - p & p \end{bmatrix} \end{array}$$

Figure 3. Partial tit-for-tat decision rule.

$$\begin{array}{cc} & \begin{array}{cccc} E_{11} & E_{12} & E_{21} & E_{22} \end{array} \\ \begin{array}{c} E_{11} \\ E_{12} \\ E_{21} \\ E_{22} \end{array} & \begin{bmatrix} px_{11} & (1-p)x_{11} & p(1-x_{11}) & (1-p)(1-x_{11}) \\ px_{12} & (1-p)x_{12} & p(1-x_{12}) & (1-p)(1-x_{12}) \\ (1-p)x_{21} & px_{21} & (1-p)(1-x_{21}) & p(1-x_{21}) \\ (1-p)x_{22} & px_{22} & (1-p)(1-x_{22}) & p(1-x_{22}) \end{bmatrix} \end{array}$$

Figure 4. Transition matrix when column plays partial tit-for-tat.

for this transition matrix is found by solving the following set of equations:

$$\alpha_1 = px_{11}\alpha_1 + px_{12}\alpha_2 + (1-p)x_{21}\alpha_3 + (1-p)x_{22}\alpha_4$$

$$\alpha_2 = (1-p)x_{11}\alpha_1 + (1-p)x_{12}\alpha_2 + px_{21}\alpha_3 + px_{22}\alpha_4$$

$$\alpha_3 = p(1-x_{11})\alpha_1 + p(1-x_{12})\alpha_2 + (1-p)(1-x_{21})\alpha_3 + (1-p)(1-x_{22})\alpha_4$$

$$\alpha_4 = (1-p)(1-x_{11})\alpha_1 + (1-p)(1-x_{12})\alpha_2 + p(1-x_{21})\alpha_3 + p(1-x_{22})\alpha_4$$

$$1 = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 \tag{10}$$

After some manipulation we obtain as solution vector

$$\left(\frac{rv}{q}, \frac{sv}{q}, \frac{tw}{q}, \frac{uw}{q}\right) \tag{11}$$

where $r = (1-p) - (1-2p)x_{12}$

$s = p + (1-2p)x_{11}$

$t = (1-p) - (1-2p)x_{22}$

$u = p + (1-2p)x_{21}$

$v = (1-p)x_{21} + px_{22}$

$w = 1 + (p-1)x_{12} - px_{11}$

$q = rv + sv + tw + uw$

We wish thus to find the values of $(x_{11}, x_{12}, x_{21}, x_{22})$ for which

$$\frac{arv + bsv + ctw + duw}{q} \tag{12}$$

is maximized. First let us show that (12) attains its maximum on one or more vertices of the four-dimensional unit hypercube, i.e., when each $x_{ij} = 0$ or 1.

Lemma 2: For any p, $\frac{1}{2} < (c-b)/(a-b+c-d) < p \leq 1$, there is a deterministic class 1 strategy ($x_{ij} \in (0,1)$ for all i,j) that yields at least as high an (asymptotic average) payoff as any probabilistic class 1 strategy ($0 \leq x_{ij} \leq 1$ for all i,j) in an iterated P.D. game against a player using a partial TFT strategy.

Proof: We need to show that for any $x_{i_1j_1}$, $x_{i_2j_2}$, $x_{i_3j_3}$ expression (12) is maximized when $x_{i_4j_4} = 0$ or 1. This is the case if (12) is a monotonically non-increasing or non-decreasing function of $x_{i_4j_4}$ for $0 \leq x_{i_4j_4} \leq 1$. To prove this for x_{11}, we can rewrite (12) as

$$E = \frac{arv + bsv + ctw + duw}{q} = \frac{k_1 + k_2x_{11}}{k_3 + k_4x_{11}}$$

where $k_1 = v(ar+bp) + (ct+du)[1+(p-1)x_{12}]$

$k_2 = bv(1-2p) - p(ct+du)$

$$k_3 = v(r+p) + [1+(p-1)x_{12}](t+u)$$

$$k_4 = v(1-2p) - p(t+u)$$

Then

$$\frac{\partial E}{\partial x_{11}} = \frac{\frac{\partial(k_1+k_2x_{11})}{\partial x_{11}}(k_3+k_4x_{11}) - \frac{\partial(k_3+k_4x_{11})}{\partial x_{11}}(k_1+k_2x_{11})}{(k_3+k_4x_{11})^2} = \frac{k_2k_3-k_1k_4}{(k_3+k_4x_{11})^2}$$

For E to be monotonic, it is sufficient that $\partial E/\partial x_{11}$ be determinate, finite, and of constant sign over the entire range $0 \leq x_{11} \leq 1$. Since $\partial E/\partial x_{11}$ is a constant divided by a function of x_{11} that always is positive or 0, one or more of these conditions can be violated only if, for some x_{11}, $0 = k_3+k_4x_{11} = q$. q, however, is the sum of products of quantities (r,s,t,u,v,w) which can all be shown, by rewriting their formulas in (11), to be ≥ 0. Hence $q = 0$ implies that one or more of the following is true:

$r=s=t=u=0$, which implies $x_{12}=x_{22}=0$, $p=x_{11}=x_{21}=1$

$r=s=w=0$, which implies $x_{12}=0$, $p=x_{11}=1$

$v=t=u=0$, which implies $x_{22}=0$, $p=x_{21}=1$

$v=w=0$, which implies $x_{22}=0$, $p=x_{11}=1$ or $x_{21}=x_{22}=0$, $x_{11}=x_{12}=1$

Removal of the constraints on x_{11} and application of the remaining sets of constraints, one at a time, makes E in each case either indeterminate for all x_{11} or a constant with respect to x_{11}. Thus the assumption that E has a higher value for some x_{11} between 0 and 1 than the greater of its values for $x_{11}=0$, $x_{11}=1$, contradicts itself. Parallel results are easily obtained for x_{12}, x_{21}, and x_{22}. Q.E.D.

Since we now know some deterministic strategy is always optimal against a player using partial TFT, we can examine the deterministic strategies individually to prove our first theorem, which states that pure cooperation will be optimal against some partial TFT strategy.

Theorem 1: There always exists p, $\frac{1}{2} < (c-b)/(a-b+c-d) < p \leq 1$, such that the strategy of pure cooperation yields at least as high an (asymptotic average) payoff as any other class 1 decision-rule in an iterated P.D. game against a player using a partial tit-for-tat strategy.

Proof: We need only establish that such a p must exist. We know from Lemma 1 that if it exists it must satisfy the middle inequality of the theorem, and from Lemma 2 that (12) attains its maximum on one or more of the vertices of the unit hypercube. The values of (12) at all 16 vertices are given in Figure 5. We wish to prove that there exists p, $p \leq 1$, such that the following inequalities will hold (these inequalities are derived from the expressions in Figure 5):

$$\frac{(1-p)a + pb + c + d}{3} < pa + (1-p)b \qquad (13)$$

$(x_{11}, x_{12}, x_{21}, x_{22})$	Asymptotic Payoffs[a]
(0, 0, 0, 0)	$(1-p)c + pd$
(0, 0, 0, 1)	$\frac{(1-p)a}{3} + \frac{pb}{3} + \frac{c}{3} + \frac{d}{3}$
(0, 0, 1, 0)	$\frac{(1-p)a}{3} + \frac{pb}{3} + \frac{c}{3} + \frac{d}{3}$
(0, 0, 1, 1)	$\frac{(1-p)a}{2} + \frac{pb}{2} + \frac{pc}{2} + \frac{(1-p)d}{2}$
(0, 1, 0, 0)	$(1-p)c + pd$
(0, 1, 0, 1)	$\frac{a}{4} + \frac{b}{4} + \frac{c}{4} + \frac{d}{4}$
(0, 1, 1, 0)	$\frac{a}{4} + \frac{b}{4} + \frac{c}{4} + \frac{d}{4}$
(0, 1, 1, 1)	$\frac{a}{3} + \frac{b}{3} + \frac{pc}{3} + \frac{(1-p)d}{3}$
(1, 0, 0, 0)	$(1-p)c + pd$
(1, 0, 0, 1)	$\frac{a}{4} + \frac{b}{4} + \frac{c}{4} + \frac{d}{4}$
(1, 0, 1, 0)	$\frac{a}{4} + \frac{b}{4} + \frac{c}{4} + \frac{d}{4}$
(1, 0, 1, 1)	$\frac{a}{3} + \frac{b}{3} + \frac{pc}{3} + \frac{(1-p)d}{3}$
(1, 1, 0, 0)[b]	$pa + (1-p)b$ *or* $(1-p)c + pd$
(1, 1, 0, 1)	$pa + (1-p)b$
(1, 1, 1, 0)	$pa + (1-p)b$
(1, 1, 1, 1)	$pa + (1-p)b$

[a]Calculated from the expressions in (11).

[b]The transition matrix consists of two absorbing chains. In which chain the process will be absorbed depends upon the initial moves of the two players. If the first move by row is cooperative, the asymptotic payoff will be $pa + (1-p)b$. If the first move by row is noncooperative, then the asymptotic payoff will be $(1-p)c + pd$.

Figure 5. Asymptotic payoffs in a P.D. game where strategies are defined by parameter choices in the transition matrix of Figure 4.

$$\frac{a + b + c + d}{4} < pa + (1-p)b \tag{14}$$

$$\frac{a + b + pc + (1-p)d}{3} < pa + (1-p)b \tag{15}$$

$$\frac{(1-p)a + pb + pc + (1-p)d}{2} < pa + (1-p)b \quad (16)$$

$$(1-p)c + pd < pa + (1-p)b \quad (17)$$

We have already established in Lemma 1 the existence of p for which (17) holds. To establish the other four inequalities we simply let $p = 1$. All four inequalities then follow from the fact that $b+c < 2a$ is a defining characteristic of the P.D. game. Q.E.D.

In proving Theorem 1 we have established that there always exists a class of partial tit-for-tat strategies which can induce cooperation in "rational" opponents, and that such strategies must involve at least a 50% level of reinforcement. If $a+d = b+c$, a condition which many of the P.D. games used in the experimental literature have satisfied, then we may prove a somewhat stronger result.

Theorem 2: If $a+d = b+c$, then for *any* p, $\frac{1}{2} < (c-b)/(a-b+c-d) < p < 1$, the strategy of pure cooperation yields at least as high an (asymptotic average) payoff as any other class 1 strategy in an iterated P.D. game against a player using a partial tit-for-tat strategy.

Proof: We wish to show that, if (6) holds, then $a+d = b+c$ is a sufficient condition for inequalities (13) through (17) to hold for all p: $\frac{1}{2} < (c-b)/(a-b+c-d) < p < 1$. To establish this result for (13) it is sufficient to let $p = (c-b)/(a-b+c-d)$ and $a+d = b+c$. When we do so, we find that the right side of (13) simplifies to $(b+c)/2$, as does the left side. This equality establishes the desired result, since the left side of (13) is monotonically decreasing in p and the right side is monotonically increasing. Analogous results are readily established for expressions (14) through (17). Q.E.D. Note that where $a+d = b+c$ the expression for (5) becomes $p > (c-b)/2(a-b)$.

We have established the pure cooperation strategy as an (asymptotically) optimal response to a partial tit-for-tat strategy in the iterated P.D. game in general, for p sufficiently large, and in particular for $p > (c-b)/(a-b+c-d)$ when $a+d = b+c$. Other strategies may, however, have identical (asymptotic expected average) payoffs. For example, inspection of (10) reveals that if $x_{11} = x_{12} = 1$, then $\alpha_3 = \alpha_4 = 0$. Hence the strategy given in Figure 6 has the same (asymptotic expected average) payoff as the strategy of pure cooperation. Nevertheless, if the inequality of (5) holds (and Lemma 1 establishes the existence of p for which it must hold), the strategy given by Figure 6 will yield a lower (average expected) payoff than pure cooperation against the partial tit-for-tat strategy because, the smaller are x_{21} and x_{22}, the longer it will be (on the average) before the process is absorbed into states α_1 and α_2 if the row player chooses 2 on the first iteration, and payoffs in the transient states will [because of (5)] be sub-optimal, except possibly on the first move. (See the discussion of absorbing chains in Kemeny, Snell, and Thompson, 1957.)

We wish next to present some results analogous to those established above, as to the conditions under which pure

c

$$\begin{array}{c} \\ 1 \\ 2 \end{array}\begin{array}{c} \begin{array}{cccc} E_{11} & E_{12} & E_{21} & E_{22} \end{array} \\ \begin{bmatrix} 1 & 1 & x_{21} & x_{22} \\ 0 & 0 & 1-x_{21} & 1-x_{22} \end{bmatrix} \end{array}$$

Figure 6. Sub-optimal although asymptotically optimal strategy.

defection will be optimal against an opponent in a P.D. game using a (partial) tit-for-tat strategy.

Lemma 3: For a,b,c,d satisfying the constraints given in (2) and (3), there always exists a p, $p \geq \frac{1}{2}$, such that the reverse inequality of (5), i.e.,

$$ap + b(1-p) < c(1-p) + pd, \tag{18}$$

is satisfied.

Proof: Our proof is analogous to that for Lemma 1. We may rewrite (18) as

$$p < \frac{c-b}{a-b+c-d} \, . \tag{19}$$

Let $p = \frac{1}{2}$. Condition (2) specifies $a < c$ and $b < d$. From these inequalities it follows that $\frac{1}{2} < (c-b)/(a-b+c-d)$. Hence a p satisfying (18) must exist. Q.E.D.

Theorem 3: There always exists a p, $(c-b)/(a-b+c-d) > p \geq \frac{1}{2}$, such that the strategy of pure defection yields at least as high an (asymptotic average) payoff as any other class 1 decision rule in an iterated P.D. game against a player using a partial tit-for-tat strategy.

Proof: We wish to prove that there exists a p such that the following inequalities (derived from the expressions in Figure 5) will hold:

$$\frac{(1-p)a + pb + c + d}{3} < (1-p)c + pd \tag{20}$$

$$\frac{a + b + c + d}{4} < (1-p)c + pd \tag{21}$$

$$\frac{a + b + pc + (1-p)d}{3} < (1-p)c + pd \tag{22}$$

$$\frac{(1-p)a + pb + pc + (1-p)d}{2} < (1-p)c + pd \tag{23}$$

$$pa + (1-p)b < (1-p)c + pd \tag{24}$$

We have already shown in Lemma 3 the existence, within the specified range, of a p for which (24) holds. To establish the other four inequalities we again simply let $p = \frac{1}{2}$. All four inequalities then follow from the P.D. defining conditions.

Theorem 4: If $a+d = b+c$, then for any p, $(c-b)/(a-b+c-d) > p > 0$, the strategy of pure defection yields at least as high an (asymptotic average) payoff as any other class 1 strategy in an iterated P.D. game against a player using a partial tit-for-tat strategy.

Proof: We wish to show that if (18) holds, then $a+d = b+c$ is a sufficient condition for the inequalities (20) through (24) to hold for all p: $(c-b)/(a-b+c-d) > p > 0$. To establish this result for (20) it is sufficient to let $p = (c-b)/(a-b+c-d)$ and $a+d = b+c$. When we do so, we find that the right side simplifies to $(b+c)/2$, as does the left side. This establishes the result, since the right side decreases more rapidly in p than does the left side. Analogous results are readily established for expressions (21) through (24).

Theorems 2 and 4 establish that when $a+d = b+c$, a strategy of pure cooperation is optimal for $p > (c-b)/(a-b+c-d)$ and a strategy of pure defection is optimal for $p < (c-b)/(a-b+c-d)$ against a player in an iterated P.D. game known to be using a partial tit-for-tat strategy. Hence a row player in an iterated P.D. game who knows that column is using a strategy of the form given in Figure 3 need only consider two strategies: pure cooperation and pure defection.

As we have noted elsewhere (Grofman and Pool, 1975), however, much empirical research remains to be done on behavior against TFT strategies. In particular, it might be expected that sufficient utility attached to the maximization of relative, as opposed to absolute, gain would rule out a strategy of pure cooperation, since only through defection can one ever gain more than the other player (see Grofman, 1975). It is possible, however, to show that as long as any value at all is attached to absolute gain there will still exist a TFT strategy compelling a "rational" player to prefer pure cooperation over pure defection.

Theorem 5: Let K and $1 - K$ be the relative weights attached by a player in an iterated P.D. game to absolute gain maximization and relative gain maximization, respectively. If $K > 0$, then there exists some p such that pure cooperation is preferred to pure defection against a player using a p% partial tit-for-tat strategy, and

$$p > \frac{(2-K)(c-b)}{K(a-d) + (2-K)(c-b)} > \frac{1}{2} \qquad (25)$$

Proof: The right-hand inequality of (25) follows from condition (2), since $K \leq 2-K$ and $a-d < c-b$. We wish to show that there exists p:

$$pKa+(1-p)[Kb+(1-K)(b-c)] > (1-p)[Kc+(1-K)(c-b)]+pKd. \qquad (26)$$

After some rearranging, we find that expression (26) is equivalent to the left-hand inequality of expression (25). Let $p = 1$. Since $K(a-d) > 0$ for all K: $K \neq 0$, this establishes the left side of the inequality. Q.E.D.[2]

[2]In general, however, for middle-range K values, neither pure cooperation nor pure defection would be the optimal row strategy in response

It is instructive to see how K influences the necessary p value for an actual P.D. matrix. For the P.D. game matrix given in Figure 7, if (as usual) $K = 1$, then the minimum p specified by (25) is 11/16 (.69). If $K = \frac{1}{2}$, a $p \geq 33/38$ (.87) is required. If $K = \frac{1}{4}$, then a $p \geq 77/82$ (.94) is needed. Thus, as expected, as concern for absolute gain maximization is replaced with concern for relative gain maximization, a reinforcement level (p) nearer to 1 is needed to induce cooperation rather than defection in a "rational" player.

$$\begin{array}{cc} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} (5, 5) & (-3, 8) \\ (8, -3) & (0, 0) \end{bmatrix} \end{array}$$

Figure 7. Representative P.D. game.

CONCLUSION

If we look at average expected payoff over some finite number of trials (and discount the payoff of the first move) or if we look at asymptotic average expected payoff, then we have established the existence of partial (p%) tit-for-tat strategies capable of inducing pure cooperation in a "rational" player cognizant of the strategy the opponent is using and unable to change it. Indeed, even when a player's utility function weights both absolute and relative payoff, we have shown that pure cooperation will be preferred to pure defection for sufficiently large p. We have also shown that the lower bound for p (the reinforcement index in such partial tit-for-tat strategies) is a monotonic function of the Rapoport and Chammah "Cooperation Index," and thus is, as expected, an indicator of the extent to which given payoff matrices may be expected to induce cooperation behavior.

In particular, we hope to have shown that even in P.D. games, where the dominant strategy has powerful attractions, iterated minimax need not be the optimal response against certain other strategies (see Fox, 1972; Grofman, 1972). However, whether and to what extent partial tit-for-tat strategies will in fact induce cooperation is a matter for experimental investigation--an investigation which the present authors have recently started.

In general, we believe that many of the most interesting questions in gaming experiments require for their investigation the use of preprogrammed experimenter strategies which are subject-response contingent. Despite the voluminous literature in experimental gaming, we also believe that the investigation of the complex interaction between communicating one's own intentions, influencing others, and maximizing

---

to a column p% tit-for-tat strategy. An expression to be maximized analogous to that given in (12) can, of course, be derived for this mixed relative and absolute gain maximization case, but as yet we have found no quick way to derive the maximum of this expression.

expected gain has only just begun.

ACKNOWLEDGEMENT

The authors wish to acknowledge the support given to this research by the University of Mannheim, the Stiftung Rehabilitation (Heidelberg), the Deutscher Akademischer Austauschdienst, the Council for European Studies, the Alexander von Humboldt-Stiftung, and the SUNY Research Foundation.

REFERENCES

Axelrod, R. (1967) Conflict of interest: an axiomatic approach. *Journal of Conflict Resolution*, 11, 87-99.

Fox, J. (1972) The learning of strategies in a simple two-person zero-sum game without saddlepoint. *Behavioral Science*, 17, 300-308.

Grofman, B. (1972) *Rationality and choice in the context of two-person non-cooperative games.* Unpublished doctoral dissertation, The University of Chicago.

Grofman, B. (1975) Prisoner's dilemma: paradox reconsidered. *Frontiers of Economics*, 101-119.

Grofman, B. & Pool, J. (1975) Bayesian models for iterated Prisoner's Dilemma games. *General Systems*, 20, 185-194.

Howard, N. (1966a) The theory of meta-games. *General Systems*, 11, 167-186.

Howard, N. (1966b) The mathematics of meta-games. *General Systems*, 11, 187-200.

Howard, N. (1970) *Games, metagames, and rationality.* Cambridge, Mass.: M.I.T. Press.

Kemeny, J., Snell, L. & Thompson, J. (1957) *Introduction to finite mathematics.* Englewood Cliffs, N.J.: Prentice-Hall.

Oskamp, S. (1971) Effects of programmed strategies on cooperation in the Prisoner's Dilemma and other mixed-motive games. *Journal of Conflict Resolution*, 15, 225-259.

Pool, J. & Grofman, B. (1975) Computer programs as a means of efficiency and control in cross-cultural experimental games. *Experimental Study of Politics*, 4(2), 27-57.

Rapoport, Amnon (1967) Optimal policies for the Prisoner's Dilemma. *Psychological Review*, 74, 136-148.

Rapoport, Anatol (1960), *Fights, games and debates.* Ann Arbor: University of Michigan Press.

Rapoport, Anatol (1966) *Two person game theory.* Ann Arbor: University of Michigan Press.

Rapoport, Anatol & Chammah, A. (1965) *Prisoner's Dilemma.* Ann Arbor: University of Michigan Press.

Rapoport, Anatol & Orwant, C. (1962) Experimental games: a review. *Behavioral Science*, 7, 1-37.

Shubik, M. (1970) Game theory, behavior, and the paradox of the Prisoner's Dilemma. *Journal of Conflict Resolution*, 14, 181-194.

Terhune, K. W. (1968) Motives, situation and interpersonal conflict within the Prisoner's Dilemma. *Journal of Personality and Social Psychology*, Monograph Supplement, 8, 1-24.

Wrightsman, L. S., Jr., O'Connor, J. & Baker, N. J. (Eds.) (1972) *Cooperation and competition: readings on mixed-motive games*. Belmont, Calif.: Wadsworth Publishing Company.