International Journal of Game Theory, Vol. 17, Issue 3, page 177–186

A Note on Reactive Equilibria in the Discounted Prisoner's Dilemma and Associated Games

By E. Kalai[1], D. Samet[2], and W. Stanford[3]

Abstract: Necessary and sufficient conditions are derived under which the Tit-for-Tat strategy pair is a subgame perfect equilibrium in the standard discounted prisoner's dilemma. The sensitivity of the equilibrium to changes in the discount parameters is shown to be a common feature of all such nontrivial "reactive" equilibria. This appears to rule them out as reasonable solutions to the discounted prisoner's dilemma. In remarks, we indicate how these results can be extended to a fairly general class of games.

1 Introduction

The purpose of this note is to investigate the subgame perfection (see Selten [7]) properties of certain pure strategy equilibria in infinitely repeated versions of the classic prisoner's dilemma and related classes of games. Concentrating on the prisoner's dilemma for the moment, the equilibria we consider consist of strategy pairs in which at least one of the players bases current period action only on the prior actions of his opponent, ignoring the history of his own actions in the supergame. A strategy of this type is called *reactive* for obvious reasons. The corresponding equilibria will also be called *reactive*. In the example below, we demonstrate the existence of a perfect reactive equilibrium in which the collusive outcome obtains along the equilibrium path. This equilibrium is seen to be quite special, however; being extremely sensitive to variations in the discount rates of the players. We show that this is a feature common to all such nontrivial reactive equilibria. In particular, the class of perfect reactive equilibria which are robust to changes in the discount rates consists of the single trivial equilibrium in which both players choose their noncooperative action in every

1 Ehud Kalai, Northwestern University, Evanston, Illinois, USA.
2 Dov Samet, Bar-Ilan University, Ramat-Gan, Israel.
3 William Stanford, University of Illinois at Chicago, Chicago, Illinois, USA.

0020–7276/88/3/177–186 $2.50

period, independent of prior history. This result is demonstrated in the Proposition. The inherent sensitivity of nontrivial reactive equilibria to changes in basic parameters of the model would thus appear to rule them out as reasonable solutions to the discounted prisoner's dilemma. In remarks at the end, we indicate how these results can be extended to a fairly general class of games.

Our interest in these results has two main sources. First, we would like to understand better the structural characteristics of collusive pure strategy perfect equilibria in repeated games under discounting. We would contrast this with the question of the possible nature of *outcome paths* to perfect equilibria with discounting. This issue was recently settled in an important paper by Abreu [1]. In contrast, our understanding of the equilibria themselves remains incomplete, largely due to the fact that they may be extremely complex. By singling out reactive strategies and equilibria, we accomplish a reduction in the complexity which strategies are allowed to assume. In the presence of the "robust perfection" requirement alluded to above, this reduction in complexity is, perhaps surprisingly, essentially complete.

Second, reactive strategies and equilibria have appeared in the literature in various repeated games contexts, and so it seems worthwhile to point out their characteristics in complete information games with discounting. For example, the Tit-for-Tat strategy (which is clearly reactive) has received much attention in connection with its success in Axelrod's [3] prisoner's dilemma tournament. It has also found its way into the economics literature. See, for example, Kreps, Milgrom, Roberts and Wilson [4], in which a finite horizon, incomplete information framework is adopted. Also Rubinstein [6], in an infinite horizon game of imperfect information with no discounting, examines strategies in which "society" imposes penalties on individuals based on the "reasonableness" of their records. This is also a reactive strategy. Finally Aumann [2], in a discussion of cooperative behavior in the repeated prisoner's dilemma, considers strategies which are defined in terms of "finite memory." Precisely, each player's mind is assumed to have a finite number of "states." Before each stage, each player may change the state of his mind in a manner which depends only on the prior state and the prior period action of his opponent. Current period action is then allowed to depend only on the (new) state. In the "memory zero" case, the new state may not depend on the old one, but only on the prior period action of the other player. An analysis of this case (with no discounting) distinguishes the Tit-for-Tat equilibrium, and hence the cooperative outcome in every period. Since any "memory zero" strategy is clearly reactive, this result will be seen in contrast to our findings.

2 Formal Definitions and a Lemma

The notation and definitions given below are adapted from Rubinstein [5].

Definition 1: The basic game we consider consists of the following two player, bimatrix game:

		Player Two: N_2	Player Two: C_2
Player One	N_1	0,0	a_1, b_2
Player One	C_1	a_2, b_1	1,1

To make this a prisoner's dilemma, we require $a_1 > 1$, $b_1 > 1$, $a_2 < 0$, $b_2 < 0$. C_i is associated with "cooperation" and N_i with "noncooperation." $G = (S_1, S_2, \pi_1, \pi_2)$ denotes the one-shot simultaneous game, where $S_i = \{C_i, N_i\}$ is the pure strategy set of player i and $\pi_i : S_1 \times S_2 \to R$ is the payoff to player i as given in the above matrix. Elements of S_i are generally denoted by s_i and are referred to as *actions*. $s \equiv (s_1, s_2)$ and $S \equiv S_1 \times S_2$.

Definition 2: $G^\infty(\alpha, \beta)$ denotes the supergame with discounting obtained by repeating G countably many times, where the discount parameter associated with player one is represented by α and for player two by β. f_i denotes a *pure strategy* for player i. It is a set of functions $f_i = \{f_i(t)\}_{t=1}^\infty$, where $f_i(1) \in S_i$ and for $t \geqslant 2, f_i(t) : S^{t-1} \to S_i$.

The set of supergame strategies of i will be denoted by F_i. F is the set of strategy pairs: $F = F_1 \times F_2$.

Given $f = (f_1, f_2) \in F$, the outcome at time t will be denoted by $s(f)(t)$, and defined inductively below:

$$s(f)(1) = f(1) = (f_1(1), f_2(1))$$

$$s(f)(t) = f(t)(s(f)(1), \ldots, s(f)(t-1)).$$

The set $\{s(f)(t)\}_{t=1}^\infty$ will be referred to as the *outcome path* induced by f.

Definition 3: $f \in F$ is a *Nash equilibrium* if for all $f_1' \in F_1$

$$\sum_{t=1}^{\infty} \alpha^{t-1} \pi_1(s(f_1', f_2)(t)) \leqslant \sum_{t=1}^{\infty} \alpha^{t-1} \pi_1(s(f)(t))$$

and for all $f_2' \in F_2$,

$$\sum_{t=1}^{\infty} \beta^{t-1} \pi_2(s(f_1, f_2')(t)) \leqslant \sum_{t=1}^{\infty} \beta^{t-1} \pi_2(s(f)(t)).$$

Note that our convention is to discount to the beginning of period one and we assume that period t payoffs are received at the beginning of period t.

Definition 4: Given $f \in F$ and $r(1), \ldots, r(T) \in S$, the pair of strategies induced by f after a "history" $h(T) = (r(1), \ldots, r(T)) \in S^T$ is denoted $f|_{h(T)} \in F$ and is defined by

$$(f|_{h(T)})(t)(s(1), \ldots, s(t-1)) = f(T+t)(r(1), \ldots, r(T), s(1), \ldots, s(t-1)).$$

where $s(\tau) \in S$, $\tau = 1, 2, \ldots, t-1$. Denote by $s_i(\tau)$ the projection of $s(\tau)$ onto its i-th coordinate.

$f \in F$ is a *subgame perfect equilibrium* if f is a Nash equilibrium and for all $T = 1, 2, \ldots$ and all histories $h(T) \in S^T$, $f|_{h(T)}$ is a Nash equilibrium. We will henceforth refer to subgame perfect equilibria as *perfect equilibria.*

Before proceeding, we state a lemma which is central to the results which follow. It seems worth pointing out that appropriate versions of this lemma also apply in general n-person repeated games where the evaluation relation is other than discounting, say the limit of long run average payoffs or the overtaking criterion for example. We will refer to the result of this lemma as the principle of *subgame parity*.

Lemma (Subgame Parity): Suppose $f \in F$ is a perfect equilibrium and two histories $h(T)$ and $\bar{h}(T)$ induce $f_2|_{h(T)} = f_2|_{\bar{h}(T)}$. Then

$$\sum_{t=1}^{\infty} \alpha^{t-1} \pi_1(s(f|_{h(T)})(t)) = \sum_{t=1}^{\infty} \alpha^{t-1} \pi_1(s(f|_{\bar{h}(T)})(t)).$$

The lemma states that if two histories $h(T)$ and $\bar{h}(T)$ induce the same strategy for player two, then player one must receive equal payoffs along the two induced outcome paths. This is simply because, by perfection, the two induced strategies of player one (which may differ) must both be best responses to the single induced strategy of player two under these conditions. Of course, a symmetric statement can be made with the roles of the players reversed. As we will see, subgame parity seems tailor made for addressing questions concerning reactive equilibria.

3 Example: The Tit-for-Tat Equilibrium

This example motivates the formulation and proof of the main result of this note, which is contained in the Proposition below.

The Tit-for-Tat strategy requires that the player using it play cooperatively in period one and then to cooperate in period t if and only if his opponent cooperated in period $t-1$. Suppose both players adopt the Tit-for-Tat strategy. Note that in this case, both are playing reactive strategies. We derive necessary and sufficient conditions on the payoff matrix entries for this to compose a perfect equilibrium.

With regard to necessary conditions, consider two histories $h(T)$ and $\bar{h}(T)$ such that $r(T) = (C_1, C_2)$ and $\bar{r}(T) = (C_1, N_2)$. Note that:

$$(f_2|_{h(T)})(1) = (f_2|_{\bar{h}(T)})(1) = C_2$$

and

$$(f_2|_{h(T)})(t)(s(1), \ldots, s(t-1)) = (f_2|_{\bar{h}(T)})(t)(s(1), \ldots, s(t-1)) = C_2$$

if and only if $s_1(t-1) = C_1$ for $t = 2, 3, \ldots$. In other words, the two histories $h(T)$ and $\bar{h}(T)$ induce the same strategy for player two; namely, to cooperate in period $T+1$, and then play in accordance with Tit-for-Tat in all subsequent periods. By subgame parity, the payoff player one receives along the induced outcome path is the *same* in both cases. In case $r(T) = (C_1, C_2)$, this payoff is $1/(1-\alpha)$, while $\bar{r}(T) = (C_1, N_2)$ yields an induced payoff of $a_1/(1-\alpha^2) + \alpha a_2/(1-\alpha^2)$, as is easily checked. Thus we have

$$1/(1-\alpha) = a_1/(1-\alpha^2) + \alpha a_2/(1-\alpha^2). \tag{1}$$

Similarly, if $r(T) = (N_1, C_2)$ and $\bar{r}(T) = (N_1, N_2)$ the equation

$$0 = a_2/(1-\alpha^2) + \alpha a_1/(1-\alpha^2) \tag{2}$$

results. Solving the system (1) and (2) yields $a_1 = 1/(1-\alpha)$, $a_2 = -\alpha/(1-\alpha)$. A symmetric argument gives $b_1 = 1/(1-\beta)$, $b_2 = -\beta/(1-\beta)$. Thus if Tit-for-Tat vs Tit-for-Tat is a perfect equilibrium (with discount parameters α and β fixed) then the payoff matrix must take the form:

		Player Two	
		N_2	C_2
Player One	N_1	0,0	$1/(1-\alpha)$, $-\beta/(1-\beta)$
	C_1	$-\alpha/(1-\alpha)$, $1/(1-\beta)$	1,1

Conversely, it is straightforward to check that Tit-for-Tat vs Tit-for-Tat is a perfect equilibrium if the payoff matrix is of this form and the discount parameters of the players are α and β respectively. As it turns out, this is because *any* sequence of actions is a best response to Tit-for-Tat following *any* history, under these conditions.

Thus we have a perfect collusive equilibrium which is, in a sense, accidental: The equilibrium relies on a coincidental relationship between the payoff matrix entries and the discount parameters of the players. That all nontrivial perfect reactive equilibria share this sensitivity to discount parameter changes (for fixed payoff entries) is demonstrated in the next section.

4 On Reactive Equilibria

We consider strategy pairs $f = (f_1, f_2)$ which satisfy the following assumptions:

A1. Player one employs a completely arbitrary supergame strategy: $f_1 \in F_1$. Player two employs a *reactive* strategy: $f_2 \in F_2$. Thus $f_2(1) \in S_2$ and for all $t \geq 2$, $f_2(t)(s(1), \ldots, s(t-1)) = f_2(t)(\bar{s}(1), \ldots, \bar{s}(t-1))$ if $s_1(\tau) = \bar{s}_1(\tau)$ for $\tau = 1, \ldots, t-1$.

A2. There exists a set $I_1 \subset (0, 1)$ with a limit point α in $(0, 1)$ such that for all $\alpha_1 \in I_1$, $f = (f_1, f_2)$ is a perfect equilibrium in $G^\infty(\alpha_1, \beta)$.

A1 is self explanatory. A2 states that the perfectness of f is reasonably robust to changes in player one's discount rate. The fact that I_1 has a limit point $\alpha \in (0, 1)$ means that every neighborhood of α contains an $\alpha_1 \in I_1$, where $\alpha_1 \neq \alpha$. In particular, I_1 must be infinite, but could consist of an arbitrarily small subinterval of $(0, 1)$, for example.

The following Proposition shows that, taken together A1 and A2 are strong assumptions indeed.

Proposition: A strategy pair $f \in F$ satisfies A1 and A2 if and only if

$$f_i(t) \equiv N_i \quad \text{for } i = 1, 2 \text{ and } t = 1, 2, \ldots$$

Proof: Regarding sufficiency, it is clear that the trivial equilibrium satisfies A1 and A2.

To prove necessity, consider two histories $h(T)$ and $\bar{h}(T)$ such that $r_1(\tau) = \bar{r}_1(\tau)$ for $\tau = 1, \ldots, T$. By the definition of induced strategy and A1, the two histories $h(T)$ and $\bar{h}(T)$ induce the same strategy for player two. By A2 and subgame parity applied to all $\alpha_1 \in I_1$:

$$\sum_{t=1}^{\infty} \alpha_1^{t-1} \pi_1(s(f|_{h(T)})(t)) = \sum_{t=1}^{\infty} \alpha_1^{t-1} \pi_1(s(f|_{\bar{h}(T)})(t)).$$

In essence, we have two power series which converge to the same function on the set I_1. A well known result from real analysis then yields:

$$\pi_1(s(f|_{h(T)})(t)) = \pi_1(s(f|_{\bar{h}(T)})(t)) \quad \text{for } t \geqslant 1. \tag{3}$$

But π_1 is one-to-one on S, so that:

$$s(f|_{h(T)})(t) = s(f|_{\bar{h}(T)}) \quad \text{for } t \geqslant 1.$$

Thus

$$\sum_{t=1}^{\infty} \beta^{t-1} \pi_2(s(f|_{h(T)})(t)) = \sum_{t=1}^{\infty} \beta^{t-1} \pi_2(s(f|_{\bar{h}(T)})(t)). \tag{4}$$

This says that the discounted payoffs to player two along the induced outcome paths are equal, and so are *independent* of player two's actions in periods $1, 2, \ldots, T$. Thus in no period can the current or past actions of player two have any effect on his future discounted payoffs along (induced) equilibrium paths. Using perfectness again, this means that player two must choose a stage game best response to the prescribed action of player one in every period. Since N_2 is strongly dominant for player two, $f_2(t) \equiv N_2$ for all t. This easily implies $f_1(t) \equiv N_1$ for all t. Q.E.D.

5 Remarks

There are two key features to the proof of the Proposition. The first consists of showing that player two's future discounted payoffs are independent of his own current and past actions (Eq. (4) in the proof), and thus that player two must play myopic best response in each period. Beyond A1 and A2, this fact was seen to depend on the structure of the payoff matrix to player one: π_1 is one-to-one on $S = S_1 \times S_2$ in the prisoner's dilemma. (As, of course, is π_2-all statements in this discussion have their symmetric counterparts.)

In the definitions below, we single out two fairly general classes of stage games for which Eq. (4) continues to hold given A1 and A2. In what follows, we assume that S is non-empty and compact, and $\pi_i : S \to R$ is continuous in the product topology for $i = 1, 2$.

Definition 5: Let im $(\pi_i) = \{\pi_i(s) | s \in S\}$ for $i = 1, 2$. The stage game G will be called a game with *completely dependent payoffs* if there exists a one-to-one and onto map $g : \text{im}(\pi_1) \to \text{im}(\pi_2)$, satisfying $g(\pi_1(s)) = \pi_2(s)$ for all $s \in S$.

As is easily verified, the prisoner's dilemma is a game with completely dependent payoffs as is, for example, the game derived from the prisoner's dilemma by replicating a row in the payoff matrix, giving player one an extra strategy. Thus

Player Two

	0,0	a_1, b_2
Player One	a_2, b_1	1,1
	0,0	a_1, b_2

also represents a game with completely dependent payoffs. Under A1 and A2, and thus given Eq. (3) in the proof of the Proposition, if G has completely dependent payoffs we conclude directly that Eq. (4) holds.

Definition 6: The stage game G will be called a game with *individually responsive* payoffs if for all $s_2 \in S_2$, π_1 is one-to-one on $S_1 \times \{s_2\}$ and for all $s_1 \in S_1$, π_2 is one-to-one on $\{s_1\} \times S_2$.

Again, the prisoners dilemma is a game with individually responsive payoffs, but the *augmented* prisoner's dilemma defined as above falls outside this class. Conversely, for a game with individually responsive but *not* completely dependent payoffs, consider a battle of the sexes game:

Player Two

Player One		
	1,1	0,0
	0,0	1,2

In this case, given $h(T)$ and $\bar{h}(T)$ as in the proof, Eq. (3) continuous to hold. But we also know (since player two has a reactive strategy) that $(f_2|_{h(T)})(1) = (f_2|_{\bar{h}(T)})(1)$, and thus the fact that payoffs are individually responsive and Eq. (3) force $(f_1|_{h(T)})(1) = (f_1|_{\bar{h}(T)})(1)$. An inductive argument now yields Eq. (4).

The second important feature of the proof is its reliance on the existence of strongly dominant strategies for the players. We can relax this to the assumption that G has Nash equilibria, but at a cost. If we strengthen A1 to the assumption that *both* players adopt reactive strategies and strengthen A2 to a symmetric statement, a symmetric argument allows us to conclude that both must play myopic best response in each period. Thus, some stage game Nash equilibrium must be picked out (generally as a function of history) in each period by the strategy pair f, and so collusive outcome paths are ruled out.

Obviously, in some cases it will be possible to leave A1 unchanged and still prove a relevant result. An example of this is given by the augmented prisoners' dilemma in which player one has two undominated strategies and player two retains a strongly dominant strategy in the stage game.

References

1. Abreu D (1984) Infinitely repeated games with discounting: a general theory. Unpublished manuscript
2. Aumann R (1981) Survey of repeated games. In: Essays in game theory and mathematical economics in honor of Oskar Morgenstern. Bibliographisches Institut, Mannheim Wien Zürich
3. Axelrod R (1980) Effective choice in the prisoner's dilemma. Journal of Conflict Resolution 24:3–25

4. Kreps D, Milgrom P, Roberts J, Wilson R (1982) Rational cooperation in the finitely-repeated prisoners' dilemma. Journal of Economic Theory 27:245–252
5. Rubinstein A (1979) Equilibrium in supergames with the overtaking criterion. Journal of Economic Theory 21:1–9
6. Rubinstein A (1979) An optimal conviction policy for offenses that may have been committed by accident. In: Applied game theory. Physica-Verlag, Würzburg, Germany
7. Selten R (1975) Reexamination of the perfectness concept for equilibrium points in extensive games. International Journal of Game Theory 4:22–55

Received September 1985
Revised version May 1986