Online ISSN 1347-4456 Print ISSN 1347-4464

ABAS: Annals of Business Administrative Science

Vol. 2, No. 4, October 2003

Global Business Research Center

http://www.gbrc.jp

Annals of Business Administrative Science

Vol. 2, No. 4, October 2003

Annals of Business Administrative Science is published quarterly in January, April, July, and October, by Global Business Research Center, Marunouchi, Chiyoda-ku, Tokyo, JAPAN. http://www.gbrc.jp

Note on Complete Proof of Axelrod's Theorem

Takashi SHIMIZU
Graduate School of Arts and Sciences, University of Tokyo
E-mail: tshimizu@waka.c.u-tokyo.ac.jp

Nobuo TAKAHASHI
Graduate School of Economics, University of Tokyo
E-mail: nobuta@e.u-tokyo.ac.jp

Abstract: This note will give a complete proof of Axelrod's theorem that characterizes the advantage of Tit-for-Tat (TFT) strategy in the repeated prisoner's dilemma. Despite of its importance in Axelrod's study, the proof of the theorem is incomplete. First, the fault of the proof is depicted and two approaches for complementation are shown. Then, we provide the complete proof using these two approaches.

Keyword: repeated prisoner's dilemma, Tit-For-Tat, cooperation

In the repeated prisoner's dilemma game, the advantage of Tit-for-Tat (TFT) strategy is characterized by the following theorem (Axelrod, 1981, Theorem 2, see also Axelrod & Hamilton, 1981):

Axelrod's Theorem. For any strategy A,

$$V(TFT \mid TFT) \geq V(A \mid TFT) \Leftrightarrow w \geq \max\left\{\frac{T-R}{R-S}; \frac{T-R}{T-P}\right\} \quad (1)$$

where $V(A|B)$ is the payoff which strategy A can get when playing with strategy B; w is the discount factor; T, R, P, and S are the payoffs of the prisoner's dilemma game and satisfy $T>R>P>S$ and $2R>T+S$ (see Table 1).

This theorem provides the condition under which any strategy cannot get higher payoff than TFT. It is a very important theorem in Axelrod's study and was reprinted in his famous book (Axelrod, 1984, Appendix B). In the last two decades, scholars have been arguing the theorem and Axelrod's

Table 1. Payoff matrix

	Cooperation (C)	Defection (D)
Cooperation (C)	(R, R)	(S, T)
Defection (D)	(T, S)	(P, P)

"collective stability" concept (see Bendor & Swistak, 1997, as a recent example). But the sufficiency of the theorem is not proved completely by Axelrod (1981, 1984) as Taylor (1987) suggested.

In this note, we will give a complete proof of Axelrod's theorem. We start with a brief discussion about the incompleteness of Axelrod's proof, then will show that there are two approaches to complete the proof of sufficiency.

Axelrod's Proof

Axelrod (1981) first shows an alternative formulation of the theorem, that is, $V(TFT|TFT) \geq V(A|TFT)$ if and only if the TFT can get higher or at least the same payoff to ALL D strategy (which defects on every move) and the strategy which alternates defection and cooperation (hereafter DCDC strategy). Then he proves that any strategy cannot get higher payoff than TFT if and only if neither ALL D nor DCDC can get higher payoff than TFT. Since Axelrod's proof is so ambiguous, we paraphrase and demonstrate the proof made by Axelrod (1981).

Proof:

(I) First we prove that the alternative formulation mentioned above is equal to the condition of Axelrod's theorem.

$$V(TFT \mid TFT) = R + wR + w^2R + \cdots = \frac{R}{1-w} \tag{2}$$

$$V(ALLD \mid TFT) = T + wP + w^2P + \cdots = T + \frac{wP}{1-w} \tag{3}$$

$$V(DCDC \mid TFT) = T + wS + w^2T + w^3S + \cdots = \frac{T + wS}{1-w^2} \tag{4}$$

therefore,

$$\begin{aligned} & w \geq \max\left\{\frac{T-R}{R-S}; \frac{T-R}{T-P}\right\} \\ & \Leftrightarrow V(TFT \mid TFT) \geq V(ALLD \mid TFT) \text{and} \\ & V(TFT \mid TFT) \geq V(DCDC \mid TFT). \end{aligned} \tag{5}$$

Thus, these two formulations are equivalent.

(II) From (I), we can paraphrase the theorem as follows:

$$\begin{aligned} & V(TFT \mid TFT) \geq V(A \mid TFT) \\ & \Leftrightarrow V(TFT \mid TFT) \geq V(ALLD \mid TFT) \text{and} \\ & V(TFT \mid TFT) \geq V(DCDC \mid TFT). \end{aligned} \tag{6}$$

The necessity of this is trivial. If any strategy cannot get higher payoff than TFT, then neither ALL D nor DCDC can get higher payoff than TFT.

We therefore turn to sufficiency. TFT has only two states depending on what the other player did on the previous move. We call them state 1 and state 2. When the other player (hereafter player 2) chose C on the previous move, TFT is in state 1 and chooses C on the next move. On the contrary, when player 2 chose D on the previous move, TFT is in state 2 and chooses D. We can consider that TFT is in state 1 on the first move.

If a strategy, A, is interacting with TFT, the best A can do when TFT is in state 1, is to choose C or D. Similarly, the best A can do when TFT is in state 2 is to choose C or D.

Thus we can say that there are only four strategies which can be the best against TFT, if the following lemma holds:

> *Lemma 1. Only the strategies that act dependently on the state of TFT can be the best strategy against TFT.*

We define S_{CC} as the strategy which chooses C in state 1 and chooses C in state 2, S_{DC} as the strategy which chooses D in state 1 and C in state 2, S_{CD} as the strategy which chooses C in state 1 and D in state 2, and S_{DD} as the strategy which chooses D in state 1 and D in state 2.

When S_{CC} interacts with TFT, it chooses C on every move. S_{CD} also chooses C on every move, because TFT chooses C on the first move. S_{DC} repeats the sequence DC, and S_{DD} chooses D on every move.

Therefore, for any strategy A, $V(TFT|TFT) \geq V(A|TFT)$ if $V(TFT|TFT) \geq V(ALL\ D|TFT)$ and $V(TFT|TFT) \geq V(DCDC|TFT)$. This condition is satisfied in Axelrod's theorem as shown in (I). Thus the proof would be completed if we can prove Lemma 1.

Nevertheless, Axelrod had never found out the existence of this embedded lemma. Unfortunately, this Lemma 1 is neither trivial nor self-evident. To prove this lemma, we have to demonstrate that the strategies which act dependently on TFT's state do better than any other strategies, that is, strategies which choose C on one move and choose D on another move in response to state 1 (and/or state 2). Therefore, we can conclude that the proof of sufficiency by Axelrod (1981) is incomplete.

This criticism on Axelrod (1981), however, might not be fair. Looking at Axelrod and Hamilton (1981), the proof which must have been in his mind might be similar to the proof we introduce in the next section: the proof using the concept of subgame.

However, the proof suggested in Axelrod and Hamilton (1981) is also insufficient. In the next section, we will consider this type of proof and provide the complete proof.

Then we will provide another approach to prove the theorem. Here we will prove that any strategy which defects on any move (whether it acts dependently on TFT's state or not) cannot get higher payoff than TFT.

Proof 1: Use of Subgame

Let us now consider the former approach to prove Axelrod's theorem. In order to do so, we use the concept of subgame. This type of proof was already suggested by Axelrod and Hamilton (1981) and Maynard Smith (1982, Appendix K), and then provided by Taylor (1987). But all of these proofs are also insufficient. They only showed section (I) of the following proof and did not examine the case which we will discuss in section (II). We should complement this insufficiency.

In this note, we define *subgame* as a part of the game beginning from the *n*th move for any $n \geq 1$. By definition, the game itself is a subgame. Moreover, any subgame in the repeated prisoner's dilemma is the same as the whole game. Hereafter, the subgame beginning from the *n*th move is denoted by g_n.

We provide a proof of sufficiency by using the concept of subgame.

Proof:

(I) From the player 2's (the player who plays game with TFT player) point of view, all subgames can be divided into two types; (1) subgames beginning with state 1, or (2) those beginning with state 2. Type 1 includes the game itself since TFT is in state 1 on the first move.

Let us now imagine the best strategy for each type of subgame. We make another lemma about the best strategies (instead of lemma 1).

Lemma 2: There are one or more best strategies against TFT for each type of subgame.

Contrary to Lemma 1, this lemma is trivial. Since the choice of TFT is dependent solely on the choice of the strategy which is interacting with TFT, the payoff of this strategy is fixed. Thus there must be the best strategies for each subgame. First, we consider the case where there is one best strategy for each type.

(i) If S_1 , the best strategy for type 1, that is, the best strategy for the game, chooses C on the first move, then the second move of TFT is C and g_2 is also type 1. Thus the second move of S_1 must be C. Similarly, S_1 must be the repeated sequence of C (CCCCCC…).

(ii) On the contrary, if S_1 choose D on the first move, then g_2 is type 2. Here we have to consider two cases; the case in which S_2, the best strategy for type 2, chooses C on its first move, or the case in which S_2 chooses D on its first move.

In the former case, S_1 chooses D at first and then chooses C, since the best strategy for g_2 begins with C. Thus, g_3 is type 1 and the third move must be D. Similarly, S_1 must be the repeated sequence of DC (DCDCDC…).

In the latter case, S_1 chooses D at first and then chooses D. Thus g_2 is also type 2 and the third move must be D. Similarly, S_1 must be the repeated sequence of D (DDDDDD…).

(iii) Therefore, $V(TFT|TFT) \geq V(A|TFT)$ for any strategy A if $V(TFT|TFT) \geq V(ALL\ D|TFT)$ and $V(TFT|TFT) \geq V(DCDC|TFT)$. This condition is satisfied in Axelrod's theorem, thus the theorem is proved if there is only one best strategy for each type of subgame.

(II) Let us now turn to the case where there are several best strategies for each type of subgame. In order to prove the theorem, we have to discuss only two cases; the case in which there are several S_1 and only one S_2, and the case in which there are several S_1 and S_2.

(i) We first focus on the case in which there are several S_1 and only one S_2. From (I), if all of S_1 choose C at first, then they are exactly the same strategy (CCCCCC…). Thus we have to consider S_1

beginning with C and beginning with D.

For the S_1 beginning with C, g_2 is type 1. It follows from this that if S_1 continues choosing C until the nth move, then g_{n+1} is equal to the game as a whole. Hence, in the case above, we can deal with the choice of D on the n+1th move similarly to the choice of D on the first move. It also follows that the repeated sequence of C is included in the set of S_1, similar to (I).

For the S_1 beginning with D, g_2 is type 2. If the S_2 chooses D at first, then g_3 is also type 2, and thus g_n is also type 2 for all $n \geq 3$. It indicates that S_1 beginning with D is the repeated sequence of D. This implies that the payoff of CCCCCC… and that of DDDDDD… are the same. This is satisfied when $w=(T-R)/(T-P)$. It must be noted that any strategies which change its choice from the repetition of C to that of D on *any* move is also included in the set of S_1, because the choice of D on the n+1th move after the repetition of C can be treated similarly to the choice of D on the first move.

On the contrary, if S_2 chooses C at first, then the second move must be C and g_3 returns to type 1. This implies the payoff of DCCCCC… and that of CCCCCC… are the same. This is satisfied when $w=(T-R)/(R-S)$. In this case, any strategy which chooses D on *any* move and chooses C on the next move is also the best strategy (see also (I) of the next section).

(ii) Now, we consider the case in which there are several S_1 and S_2. By using similar methods of (i), we can conclude that all of the sequence CCCCCC…, DCCCCC…, and DDDDDD… are the best strategy and the payoff of them are equal. This condition is satisfied when $w=(T-R)/(T-P)=(T-R)/(R-S)$. It is noteworthy that defection on any move cannot change payoff, so that *any* strategy is the best strategy against TFT (see also (II) of the next section). Thus the proof is completed.

Proof 2: Consideration of All Type Defection

In the previous section we provide the complete proof of Axelrod's theorem by using the concept of subgame. In this section we take another approach and prove that any strategies, which defect on any move cannot get higher payoff than TFT. Though this approach is the same in its essentials as the proof that the pair of TFT is Nash equilibrium, our proof is related to the proof by Axelrod or the proof in the previous section. The former version of this proof is presented in Shimizu (1997).

Proof:

(I) The games of TFT versus TFT have the sequence of moves indicated in Table 2. We consider a strategy $D_1(k)$ of Player 2 which has the sequence of moves in Table 3.

The difference between Table 2 and Table 3 is only the kth and (k+1)th moves within the boxes. Then we define $U_2(k)$ as the total payoff except for the kth and (k+1)th moves. We obtain

$$V(TFT \mid TFT) = w^{k-1}(R + wR) + U_2(k) \tag{7}$$

$$V(D_1(k)\,|\,TFT) = w^{k-1}(T + wS) + U_2(k). \qquad (8)$$

Therefore, for any k, k=1,2,3,...,

$$V(TFT\,|\,TFT) \geq V(D_1(k)\,|\,TFT) \Leftrightarrow w \geq \frac{T-R}{R-S}. \qquad (9)$$

We can see from this that one-time defection on any move cannot change the payoff of player 2 if w=(T-R)/(R-S). This is the case we mentioned in (II) of the previous section.

(II) Now, we consider strategies $D_n(k)$ and $D_{n+1}(k)$ of Player 2 that are respectively defined as Tables 4 and 5. The case of n=1 has been considered in (I), then it is assumed that n≥2.

The difference between Table 4 and Table 5 is only the (k+n)th and (k+n+1)th moves within the boxes. Then we define $U_2(k+n)$ as the total payoff except for the (k+n)th and (k+n+1)th moves. We have

$$V(D_n(k)\,|\,TFT) = w^{k+n-1}(S + wR) + U_2(k+n) \qquad (10)$$

$$V(D_{n+1}(k)\,|\,TFT) = w^{k+n-1}(P + wS) + U_2(k+n). \qquad (11)$$

Therefore, for any k, k=1,2,3,...,

$$V(D_n(k)\,|\,TFT) \geq V(D_{n+1}(k)\,|\,TFT) \Leftrightarrow w \geq \frac{P-S}{R-S} \qquad (12)$$

$$V(D_n(k)\,|\,TFT) \leq V(D_{n+1}(k)\,|\,TFT) \Leftrightarrow w \leq \frac{P-S}{R-S}. \qquad (13)$$

(i) If w≥ (P-S)/(R-S), then

$$V(D_1(k)\,|\,TFT) \geq V(D_2(k)\,|\,TFT) \geq \cdots \geq V(D_n(k)\,|\,TFT) \geq \cdots \qquad (14)$$

for any k.

Therefore, player 2 cannot increase his or her payoff by increasing the number of defection. Thus all we have to do is compare $V(TFT|TFT)$ with $V(D_1(k)|TFT)$.

From (I) we obtain,

$$V(TFT\,|\,TFT) \geq V(D_1(k)\,|\,TFT) \geq V(D_2(k)\,|\,TFT) \geq \cdots \Leftrightarrow w \geq \frac{T-R}{R-S} \text{ and } w \geq \frac{P-S}{R-S}. \qquad (15)$$

Table 2. TFT versus TFT

Player	Strategy	1	....	k-1	k	k+1	k+2	....
1	TFT	C	....	C	C	C	C	....
2	TFT	C	....	C	C	C	C	....
Payoff of Player 2		R	....	R	R	R	R	....

Table 3. TFT versus $D_1(k)$

Player	Strategy	1	....	k-1	k	k+1	k+2	....
1	TFT	C	....	C	C	D	C	....
2	$D_1(k)$	C	....	C	D	C	C	....
Payoff of Player 2		R	....	R	T	S	R	....

It must be noted that

$$V(TFT\mid TFT)\geq V(D_1(k)\mid TFT)\Leftrightarrow w\geq\frac{T-R}{R-S} \quad (16)$$
$$\Leftrightarrow V(TFT\mid TFT)\geq V(DCDC\mid TFT)$$

and thus the comparison between $V(TFT|TFT)$ and $V(D_1(k)|TFT)$ is equivalent to that between $V(TFT|TFT)$ and $V(DCDC|TFT)$.

(ii) If $w\leq (P-S)/(R-S)$, then for any k,

$$\begin{aligned}&V(D_1(k)\mid TFT)\leq V(D_2(k)\mid TFT)\leq\cdots\leq V(D_n(k)\mid TFT)\leq\cdots\\&\leq V(TFT\mid TFT)-w^{k-1}V(TFT\mid TFT)+w^{k-1}V(ALLD\mid TFT)\\&=(1-w^{k-1})V(TFT\mid TFT)+w^{k-1}V(ALLD\mid TFT).\end{aligned} \quad (17)$$

In this case, player 2 can increase his or her payoff by increasing the number of defection. However, he or she cannot do better than TFT or ALL D. Thus, the point is that TFT can get higher payoff than ALL D or not.

Therefore,

$$V(D_1(k)\mid TFT)\leq V(D_2(k)\mid TFT)\leq\cdots\leq V(TFT\mid TFT)$$
$$\Leftrightarrow w\geq\frac{T-R}{T-P}\text{ and } w\leq\frac{P-S}{R-S} \quad (18)$$

since

$$V(TFT\mid TFT)\geq V(ALLD\mid TFT)\Leftrightarrow\frac{R}{1-w}\geq T+\frac{wP}{1-w}$$
$$\Leftrightarrow w\geq\frac{T-R}{T-P}. \quad (19)$$

From (i) and (ii), we obtain, for both cases,

$$V(TFT\mid TFT)\geq V(D_n(k)\mid TFT) \quad (20)$$

for any k, $n(\geq 2)$ if $w\geq (T-R)/(R-S)$ and $w\geq (T-R)/(T-P)$.

If $w=(T-R)/(R-S)=(T-R)/(T-P)$, then all of the payoff of TFT, ALL D, and DCDC are equal. Moreover, if $w= (T-R)/(R-S)$ and $w= (T-R)/(T-P)$,

$$V(D_1(k)\mid TFT)=V(D_2(k)\mid TFT)=\cdots=V(D_n(k)\mid TFT) \quad (21)$$

Table 4. TFT versus $D_n(k)$

Player	Strategy	1	....	k-1	k	k+1	....	k+n-1	k+n	k+n+1	k+2	....
1	TFT	C	....	C	C	D	....	D	D	C	C	....
2	$D_n(k)$	C	....	C	D	D	....	D	C	C	C	....
Payoff of Player 2		R	....	R	T	P	....	P	S	R	R	....

Table 5. TFT versus $D_{n+1}(k)$

Player	Strategy	1	....	k-1	k	k+1	....	k+n-1	k+n	k+n+1	k+2	....
1	TFT	C	....	C	C	D	....	D	D	D	C	....
2	$D_{n+1}(k)$	C	....	C	D	D	....	D	D	C	C	....
Payoff of Player 2		R	....	R	T	P	....	P	P	S	R	....

since $w=(T\text{-}R)/(R\text{-}S)=(T\text{-}R)/(T\text{-}P)$ implies $w=(P\text{-}S)/(R\text{-}S)$. We can conclude that defection of any number of times on any move cannot change the payoff of player2 if the above condition is satisfied. This is what we mentioned in (II) of the previous section.

(III) From (I) and (II), when confronted with TFT, any defection cannot increase its own payoff. Thus, any strategy having more than one D in its sequence of moves cannot get higher payoff than TFT.

Now, we define $D_0'(k)$ as a strategy having more than one D in its sequence of moves except for the $(k\text{-}1)$th move through the $(k+n+1)$th move. And $D_n'(k)$ is defined as a strategy which has additional defections from the kth move to the $(k+n)$th move in comparison with $D_0'(k)$. By using similar methods of proving (I) and (II), we obtain $V(D_0'(k)|TFT) \geq V(D_n'(k)|TFT)$ for any k, n if $w \geq (T\text{-}R)/(R\text{-}S)$ and $w \geq (T\text{-}R)/(T\text{-}P)$. Thus the proof is completed.

References

Axelrod, R. (1981). The emergence of cooperation among egoists. *American Political Science Review*, *75*, 306-318.

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*, 1390-1396.

Bendor, J., & Swistak, P. (1997). The evolutionary stability of cooperation. *American Political Science Review*, *91*, 290-307.

Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge, UK: Cambridge University Press.

Shimizu, T. (1997). Yugen hanpuku shujin no jirenma ni okeru kyocho-kodo no shinka [Evolution of cooperation in the finitely repeated prisoner's dilemma]. *The Japanese Journal of Behaviormetrics*, *24*, 101-111. (in Japanese)

Taylor, M. (1987). *Possibility of cooperation: Studies in rationality and social change*. Cambridge UK: Cambridge University Press.

〔Received October 8, 2003; accepted October 14〕