DISCUSSION

INFORMATION AND STRATEGY IN ITERATED PRISONER'S DILEMMA*

The prisoner's dilemma (PD) – a two-person, non-zero-sum game – highlights a situation where individual rationality leads to a collectively sub-optimal equilibrium. This has disturbing implications for the social sciences; and the problem has received considerable attention from philosophers, psychologists and economists. It has been used to interpret Rousseau's concept of the 'general will' and has also cropped up in discussions on ethics and morality[1]. It demonstrates with disconcerting simplicity how atomistic action may lead to collective ill, unless there are binding 'social contracts' to guide individual action.

Respite was sought by pointing out that in reality this game was played more than once; and that could result in the equilibrium from not being the socially inferior one. But it was supposedly demonstrated by Luce and Raiffa[2] (LR) that if the game was repeated a finite number of times, then the equilibrium strategy continued to be the sub-optimal one. It will be argued here that this is not in general so and for their result to hold, it is necessary for certain additional assumptions – which are not required in the non-repeated PD – to be satisfied. These assumptions are highly demanding and proliferate at a remarkable pace as the number of repetitions is increased. This is an optimistic result. It shows that in reality, where this game situation is likely to arise more than once, individual rationality may lead to collective good.

I

The following table is the pay-off matrix for individuals A and B. a_1, a_2 and b_1, b_2 are possible strategies for A and B respectively, and α_{ij} (resp. β_{ij}) is the pay-off to A (resp. B) for (a_i, b_j). For simplicity we consider the pay-offs to be monetary payments.

Theory and Decision 8 (1977) 293–298. *All Rights Reserved*

		B	
		b_1	b_2
A	a_1	α_{11}, β_{11}	α_{12}, β_{12}
	a_2	α_{21}, β_{21}	α_{22}, β_{22}

Consider the following assumptions:

(i) $\alpha_{21} > \alpha_{11} > \alpha_{22} > \alpha_{12}$,

(ii) $\beta_{12} > \beta_{11} > \beta_{22} > \beta_{21}$.

Individual A is said to have PD preference if (i) holds and he maximises monetary gains. Similarly for B. We shall assume throughout that both A and B have PD preferences. Consequently, if this game is played once only, then (a_2, b_2) is the equilibrium outcome. This is Pareto inferior to (a_1, b_1). Therefore, if this sort of a preference configuration occurs in reality, as seems quite plausible, then a Pareto inefficient outcome seems inevitable.

However, in reality most games are played more than once. Suppose that the individuals know that the above game will be played more than once. Then A might play a_1 a few times to induce B to play b_1. Once B gets the hint and plays b_1, the equilibrium might settle at (a_1, b_1). A (resp. B) would not break this equilibrium, since though by suddenly playing a_2 (resp. b_2) he can reap some extra gains, B (resp. A) will retaliate with b_2 (resp. a_2) making A worse off eventually.[3] This seems to provide a way out of the Pareto inferior outcome.

However, if the game is played a finite number of times, and the players know how many times, then this 'way out' is allegedly illusory. This paradoxical result was demonstrated by LR using an argument similar to that used in the 'surprise test' paradox:

> Suppose the players are told that game H (i.e. the above game) is to be played exactly twice, and suppose that each player is shrewd enough to see that his second strategy strictly dominates his first one in a single play of the game. Thus, before making their first move, each realizes that in the second game the result is bound to be (a_2, b_2), for, after the first game is played, the second one must be treated as if H is going to be played once and only once. The second play being perfectly determined, the first play of the game can be construed as H being played once and only once. Thus, it appears that (a_2, b_2) must arise on both trials. The argument generalizes: Suppose they know that H is to be played exactly 100 times. Things are clear on the last trial, the

(a_2, b_2) response is assured; hence the penultimate trial, the 99th, is now in strategic reality the last, so it evokes (a_2, b_2); hence the 98th is in strategic reality the last, so it evokes (a_2, b_2) etc. This argument leads to (a_2, b_2) on all hundred trials. Indeed if player 2 is a b_2-conformist on all trials, then player 1 is best off choosing a_2 on all trials, and conversely, i.e., (a_2, b_2) on all trials is an equilibrium pair. (LR, op. cit., p. 98 and 99.)

From the above argument it seems that given the assumption that the two individuals have PD preferences, the equilibrium is (a_2, b_2) whether the game is played once or any finite number of times. I would argue that this is incorrect. The mere assumption that A and B have PD preferences is sufficient to establish (a_2, b_2) as the equilibrium in the single game case. But when the game is repeated N (>1) times, the assumption requirement is much stronger. This can be seen by carefully analysing the case of N games.

Consider the case of $N = 2$. LR argued that in the first game A will move a_2 because A knows that in the final game (a_2, b_2) is bound to occur. But how does A know that? Clearly, to know that, A must know that B has PD preference. Symmetrically, B must know that A has PD preference. Therefore, to prove (a_2, b_2) to be the equilibrium in the first game we need this additional assumption that A (resp. B) knows that B (resp. A) has PD preference.

Now, even if $N = 100$, this assumption is necessary since, without it, we would not be able to prove that (a_2, b_2) would certainly occur in the 99th game. Hence, for all $N > 1$, this assumption is necessary. But is this assumption sufficient for all N? The answer is no. As N becomes larger, the list of assumptions increases.

Consider $N = 3$. We can be sure that A will move a_2 in the first game if he thinks that outcomes of games 2 and 3 are fixed. A will think in this way if (1) A knows that B has PD preference (this will make A sure about the final game's outcome to be (a_2, b_2)) and (2) A knows that B knows that A has PD preference. This will make A sure that B will move b_2 in the second game, and consequently will make him sure that the outcome of the second game will be (a_2, b_2). Remember that because of (1), A will move a_2 in the second game. Given that symmetric conditions hold for B, the outcome of the first game will be (a_2, b_2).

The increasing informational assumption is obvious. We now summarize the above analysis.

When $N = 1$, (a_2, b_2) is the equilibrium if

A and B have PD preferences.

When $N = 2$, (a_2, b_2) would certainly occur in both games if

A and B have PD preferences, and

A (resp. B) knows that B (resp. A) has PD preference.

When $N = 3$, (a_2, b_2) would certainly occur in all three games if

A and B have PD preferences, and

A (resp. B) knows that B (resp. A) has PD preference, and

A (resp. B) knows that B (resp. A) knows that A (resp. B) has PD preference.

What happens in the general case where $N = n$? The result is summarized in the chart below. It is assumed throughout that both A and B have PD preferences.

	A moves a_2 if the following additional assumptions are given	B moves b_2 if the following additional assumptions are given
nth time	No additional asumptions	No additional assumptions
(n-1)th time	(1) A knows B has PD preference	(1*) B knows A has PD preference
(n-2)th time	(1) and (2) A knows that (1*)	(1*) and (2*) B knows that (1)
(n-3)th time	(1), (2) and (3) A knows that (1*) and (2*)	(1*), (2*) and (3*) B knows that (1) and (2)
⋮		
(n-(n-1)th time (i.e. 1st time)	(1), (2) ... ((n-2)) and ((n-1)) A knows that (1*), (2*) ... and ((n-2)*)	(1*), (2*) ... ((n-2)*) and ((n-1)*) B knows that (1), (2) ... and ((n-2))

Hence, while in the case of non-repeated PD the sub-optimal solution is a direct consequence of the individuals having PD preferences, in the n-game case it is, fortunately, not so.

While people are often confronted with PD situations in life, it is not at all clear that they can be expected to fulfil assumptions (1), (2) ... (n-1) and (1*), (2*) ... ((n-1)*). These are extremely strong informational requirements. When these assumptions are not satisfied, the occurrence of (a_1, b_1) in some games becomes a distinct possibility. Under what parametric conditions the 'general will' is likely to be realized is discussed in the case of $N = 2$ in the next section.

However, it is already clear that though a divergence in the 'general will' and the 'will of all' could create serious problems as discussed by Runciman and Sen (op. cit.), the occurrence of such a divergence may not be as widespread as one may have been lead to believe from the LR argument.

II

We assume $N = 2$; A and B have PD preferences and in situations of uncertainty they maximise their individual *expected* earnings; $\alpha_{ij} = \beta_{ji}$, for all i, j (i.e. A and B have symmetric pay-offs); and A's expectation of B's behaviour is the same as B's expectation of A's behaviour. These last two assumptions are not really necessary but are made to allow us restrict our analysis to one individual, since the other will behave symmetrically. Finally, no assumptions like (1), (2) ... ((n-1)) and (1*), (2*) ... ((n-1)*) are granted.

Since A and B have PD preferences, (a_2, b_2) is bound to occur in the final game. Hence, our main interest lies in the first game. Consider the subjective probabilities attached by A to B's moves. The probability of B moving b_2 in the first game is r. If in game 1 A moves a_2, then the probability of B playing b_2 in the second game is p. If in game 1 A moves a_1, then the probability of B playing b_2 in the second game is q. If A moves a_2 in game 1, it is natural for him to expect retaliation from B in the second game. So we may assume $p > q$, though this assumption is not necessary for the subsequent results.

Let A's expected gain from playing a_1 in the first game be $G(1)$.

Hence,

$$G(1) = (1-r)\,\alpha_{11} + r\alpha_{12} + (1-q)\,\alpha_{21} + q\,\alpha_{22}.$$

Similarly, the expected gain from playing a_2 in the first game is $G(2)$.

$$G(2) = (1-r)\,\alpha_{21} + r\alpha_{22} + (1-p)\,\alpha_{21} + p\,\alpha_{22}.$$

In both these expressions, the first two terms are expected earnings from game 1, and the last two terms are expected earnings from game 2. Because of our assumption above, B's subjective probabilities regarding A's behaviour are symmetric. Hence, a similar analysis applies to B.

Let $X = G(1) - G(2)$. Therefore,

$$X = (1-r)\,(\alpha_{11} - \alpha_{21}) + r(\alpha_{12} - \alpha_{22}) + (p-q)\,(\alpha_{21} - \alpha_{22}).$$

If $X > 0$, then A plays a_1 in game 1; and if $X < 0$, then A plays a_2. In case $X = 0$, then A's move is indeterminate in the context of this theory. It is easily checked that suitable values for the parameters involved can be selected such that (i) is satisfied and at the same time $X > 0$. Hence, it is possible for A to have PD preference and at the same time move a_1 in the first game. B's behaviour being symmetric, the occurrence of (a_1, b_1) is a distinct possibility.

Assumption 1 (p. 296) says that $p = q = 1$. If this assumption holds, then $X < 0$; and in line with the analysis in the first section A will move a_2 in both games. But the assumption is an arbitrary one.

As N becomes larger, the assumption structure required to establish (a_2, b_2) as the equilibrium becomes extremely bulky and untenable in any realistic situation. Consequently, the possibility that individual rationality may lead to 'the common good', even when individuals have PD preferences, is no longer as remote as it might have seemed.

University of Reading and KAUSHIK BA3U
The London School of Economics and Political Science

NOTES

* I have benefited from discussions with Prof. Amartya Sen.

[1] For the former see Runciman, W. G. and Sen, A. K.: 1965, 'Games, Justice and the General Will', *Mind* 74; and Smyth, J.: 1972, 'The Prisoners' Dilemma II', *Mind* 81; and for the latter see articles by Sen, A. K. and Watkins, J. W. M. in Korner, S. (ed.), *Practical Reason*, Oxford: Blackwell, 1974.

[2] Luce, R. D. and Raiffa, H.: *Games and Decisions*, Wiley: New York, 1958.

[3] For a more detailed exposition of this see LR, op. cit., p. 98.