REVIEW

Taylor, Michael, *Anarchy and Cooperation*, John Wiley and Sons, London, 1976, p. 151, £6.25.

Consider a group whose members could, through joint action, achieve a collective good of considerable value. However, the inaction of any one or any handful of individuals in the group does not significantly decrease the payoff to the group as a whole, and those who fail to act can become "free riders" – benefiting from the efforts of their fellows without incurring costs themselves. Under such circumstances every individual in the group may have a very strong incentive to "let George do it", and thus no one may act and the common welfare may, accordingly, through inaction, be neglected. Analogously, consider the problem of individual *restraint* from action when it is in the individual's own interest that he act, but in the collective interest that he not. Such a situation has been imaginatively sketched by Garrett Hardin (1968). Cf. Hume, 1960, p. 538; cited in Schick, 1977, p. 786) under the title 'Tragedy of the Commons'.

Hardin asks us to imagine a commons, a pasture open to all. The village herdsmen keep animals on the commons. Each herdsman is assumed to seek to maximize his own gain. As long as the total number of animals is below the carrying capacity of the commons, a herdsman can add an animal to his herd without affecting the amount of grazing of any of the animals, including his own. But beyond this point, the 'tragedy of the commons' is set in motion. Asking himself now whether he should add another animal to his herd, he sees that this entails for him a gain and a loss: on the one hand, he obtains the benefit from this animal's yield (milk, meat, or whatever); on the other hand, the yield of each of his animals is reduced because there is now overgrazing. The benefit obtained from the additional animal accrues entirely to the herdsman. The effect of overgrazing, on the other hand, is shared by all the herdsmen; every one of them suffers a slight loss. Thus, says Hardin, the benefit to the herdsman who adds the animal is greater than *his* loss. He therefore adds an animal to the commons. For the same reason, he finds that it pays him to add a second animal, and a third, a fourth, and so on. The same is true for each of the other herdsmen. The result is that the herdsmen collectively bring about a situation in which each of them derives less benefit from his herd than he did before the carrying capacity of the commons was exceeded. The process of adding animals may indeed continue until the ability of the commons to support livestock collapses entirely (p. 2).

Anarchy and Cooperation deals with situations like the two types sketched above. In the works of its author (p. v) it is about "cooperation in the

Theory and Decision 12 (1980) 107–114. 0040–5833/80/0121–0107$00.80.

absence of government". In it Taylor offers a critique of what he (and I) regard as the most persuasive justification of the state, to wit, "the argument that, without the state, people would not voluntarily cooperate to provide themselves with certain public goods: goods, that is to say, which any member of the public may benefit from, whether or not he contributes in any way to their provision" (p. v). This argument, which is quite popular with contemporary academics, suggests that, without the state, people would not be able to "act so as to realize their *common* interests" (p. v, emphasis ours).

Following Hardin (1971) (cf. also Dawes, 1973), Taylor looks at the N-person Prisoner's Dilemma game as an exemplification (for sufficiently large N) of the public goods/free-rider problem and discusses the issue of strategy choice in the Prisoner's Dilemma supergame (the ordinary Prisoner's Dilemma game played an indefinite number of times). In connecting the P.D. game to the literature on the "tragedy of the commons", Taylor shows how formal theory can be applied to important problems of public policy.

> In recent years, an old argument about the necessity of a strong and centralized state has found new supporters. Their common concern is the environmental crisis; the combined effects of rapid population growth, depletion of nonrenewable resources, and environmental deterioration. They believe that only through powerful state action can the crisis be either averted or survived . . . .
>
> At the heart of most of the environmentalist justifications of the scope of state power is the argument that people will not voluntarily *restrain* themselves from doing those things of which the environmental crisis is the aggregate effect. They will not voluntarily refrain from hunting whales and other species threatened with extinction, from having 'too many children', from discharging untreated wastes into rivers and lakes, and so on. (pp. 1–2, emphasis ours.)

Taylor goes on to make original contributions in three areas. First, following a suggestion by Shubik (1958, 1970), he looks at the time sequence of behavior in an iterated P.D. game where players have a discount function which maps future payoffs onto their equivalent present value. Taylor points out that, for many public goods, "the choice of whether to contribute to their provision and of how much to contribute is a recurring choice . . . . This is true of the choice of how much to exploit the 'commons', how many whales to take in each year or other time period, how much to treat industrial waste before discharging it into the lake, and so on. It is also true of the individual's choice of whether or not to behave peaceably, to refrain from robbery and fraud, and so on" (p. 28.) Second, contrary to some claims in the

literature, he shows that not all efforts for the provision of public goods result in prisoners dilemma situations, even when the group involved is very large (see Chapter 2). Third, and most importantly, Taylor shows that in a sequential P.D. game, where some or all players' actions may be *contingent* on the past behavior of some or all of the other players, voluntary cooperation may be rational for each player in the game – and this conclusion follows without requiring either state imposed coercion or sanctions or altruistic utility functions.

Taylor's work on the P.D. supergame (played noncooperatively) is both imaginatively conceived and intriguing in its practical implications. In Chapter 2, for the 2-person iterated P.D. game, Taylor looks at five classes of strategies and determines necessary and sufficient conditions for an equilibrium to exist when each player chooses a strategy from among these strategy sets. The strategy classes he considers are

- C^∞: C is chosen in every ordinary game.
- D^∞: D is chosen in every ordinary game.
- A_k: (k is a strictly positive integer): C is chosen in the first game, and it is chosen in each subsequent game as long as the other player chooses C in the previous game; if the other Defects in any game, D is chosen for the next k games; C is then chosen no matter what the other player's last choice is; it continues to be chosen as long as the other player chooses C in the preceding game; when the other player Defects, D is chosen for $K + 1$ games; and so on; the number of games in which the other player is 'punished' for a Defection increases by one each time; and each time there is a return to C. [The limiting case of A_k is when $K \to \infty$.]
- B: C is chosen in the first game; thereafter the choice in each game is that of the other player in the preceding game. [Tit-for-tat.]
- B': D is chosen in the first game; thereafter the choice in each game is that of the other player in the preceding game. [Tit-for-tat.]

Taylor calculates the (discounted) payoff associated with each strategy pair and specifies the outcomes in a 5×5 table. Within this table, Taylor finds that a strategy vector which contains C^∞ is never in equilibrium; shows that (A_k, B') is never in equilibrium and that a strategy vector containing D^∞ and any one of A_k, B and C^∞ is never in equilibrium; and states conditions for (A_k, A_k), (A_k, B), (B, B), (B, B'), (B, B') and (B', D^∞) to be in equilibrium. (Of course, (D^∞, D^∞) is always in equilibrium.) Thus, under certain conditions, strategy pairs resulting in repeated mutual cooperation will in fact be equilibrium. Taylor also shows that, in addition to mutual defection and mutual cooperation, alternation between (C,D) and (D,C) is possible as an equilibrium outcome when the strategies available to the players are A_k, B, B', C^∞ and D^∞. Since specifying the requisite conditions involves a number of lengthy algebraic statements, we shall not attempt to list these conditions here. (See Taylor, pp. 33–39.)

The strategies defined above $(A_k, B, B', C^\infty, D^\infty)$ are, of course, only five out of an infinite number of possible supergame strategies, and Taylor recognizes his analysis to be incomplete, but asserts that "it does illustrate, at least, how mutual cooperation throughout the supergame can be the outcome under certain conditions" (p. 32.) He then goes on to claim that "the five strategies include those which are most likely to be considered, at least at a conscious level, by real players; *inasmuch as this is the case, the analysis given is complete*" (p. 32, emphasis ours.) About this claim I am, however, rather skeptical. I believe there to be other strategies which "reasonable" players might wish to consider when in a sequential P.D. game. (See Axelrod, 1978; Grofman and Pool, 1975.)

In Chapter 3, Taylor offers a similar analysis for the N-person supergame. Here, he assumes that each player's payoffs in each play depend upon two things only; his own strategy choice (C or D) and the *number* of other players choosing C in that game. In addition to C^∞ and D^∞, Taylor generalizes the three previously considered contingent strategies (A_k, B and B') to the N-person case as follows (p. 44.)

$A_{k,n}$: (k is a strictly positive integer): C is chosen in the first game; it continues to be chosen as long as *at least n other players* $(N > n > 0)$ also choose C (in the preceding game); if the number of other Cooperators falls below n, then D is chosen for the next k games; C is then chosen in the next game no matter what the other players chose in the preceding game; it continues to be chosen as long as at least n other players choose C in the preceding game; when the number of other Cooperating players next falls below n, D is chosen for $k + 1$ games; and so on; the number of games in which the other players are 'punished' for Defection increases by one each time; and each time there is a return to C. [The limiting case of $A_{k,n}$ is $A_{k,\infty}$ when $n \to \infty$.]

B_n: C is chosen in the first game; thereafter, if the number of other players choosing C in the preceding game is at least n, C is chosen; otherwise D is chosen. $(n > 0.)$

B'_n: D is chosen in the first game, thereafter, if the number of other players choosing C in the preceding game is at least n, C is chosen; otherwise D is chosen. $(n > 0.)$

In order to simplify the analysis somewhat, Taylor considers a number of special cases, the most important of which is the 4×4 matrix consisting of outcomes when players are restricted to the strategy classes $A_{k,n}, B_n, C^\infty$ and D^∞. In this case Taylor shows that even when "some players insist on using D^∞, cooperation may still be rational for all the rest" (p. 50), but the possibility of cooperation rests upon the existence of sufficient numbers of players who use a strategy in which their continued cooperation is contingent upon the cooperation of *all* their fellows except for those *committed* to defection. In other words, there must be a sufficient number of espousers of an "I'll cooperate as long as everybody, but the incorrigibles, does likewise" point of view, for cooperation to rationally sustain itself among a subset of the group. (Of course, payoff functions must also satisfy certain algebraic inequalities; see Taylor, pp. 49–50.) This is certainly not an intuitively obvious finding but it appears to me to be an important one. While Taylor offers other findings on equilibria in various special cases of P.D. supergames, including an extended treatment of Hardin's (1974) model, we shall not summarize his rather detailed results here.

In Chapter 4, Taylor examines the shape of individual utility functions and the social ethics they represent, looking at transformations of a P.D. payoff matrix which may be generated when one or both the players are characterized by something other than an ethic of simple egoism, e.g., altruism, malevolence, relative gain maximization, etc. Taylor finds that, in the one-shot P.D. game, pure altruism and pure malevolence transform a P.D. payoff matrix into one which no longer satisfies the P.D. conditions, while relative gain maximization leads to what he calls a "Game of Difference" (cf. Shubik, 1971) which under certain relatively easy to satisfy conditions remains a P.D. game and becomes one in which the temptation to defect is greater than in the original untransformed payoff matrix; Taylor then goes on to extend the Game of Difference to the N-person case, and he shows that this game remains a P.D. game provided some very reasonable conditions (e.g., the greater the number of individuals who cooperate, the greater the payoff to any individual who does not cooperate) are satisfied. This game is of particular importance in that "it is the sort of game which Hobbes assumes people to be playing in the 'state of nature'" (p. 94.) Taylor's analysis of the implications of alternative social ethics for behavior in the P.D. game goes considerably beyond what had been available in the P.D. literature (cf. Grofman, 1976); e.g., of identity) the nature of the game. In looking at altruism and other-regardingness Taylor makes a valuable start in this direction, but missing is any trace of sociological analysis.

On balance *Anarchy and Cooperation* makes an important theoretical contribution and is a work with which all students of the logic of collective action and of social ethics should become acquainted.

School of Social Sciences, BERNARD GROFMAN
University of California, Irvine

REFERENCES

Axelrod, Robert, 'Artifical Intelligence and the Iterated Prisoner's Dilemma', Delivered at The Annual Meeting of the Public Choice Society, March 3–5, 1978, New Orleans.

Braybrooke, David, 'The Insoluble Problem of the Social Contract', *Dialogue* **15** (March 1971), pp. 3–37.

Dawes, Robyn, 'The Commons Dilemma Game', *Oregon Research Bulletin* **13**, No. 2 (September, 1973).

Grofman, Bernard, 'Prisoner's Dilemma: Paradox Reconsidered', in Gordon Tullock (ed.), *Frontiers of Economics, Vol. I*, Center for the Study of Public Choice, Blacksburg, Virginia, 1976.

Grofman, Bernard and Pool, Jonathan, 'Bayesian Models for Iterated Prisoners' Dilemma Games', *General Systems* **20** (1975) pp. 185–194.

Grofman, Bernard and Pool, Jonathan, 'How to Make Cooperation the Optimizing Strategy in a Two-Person Game', *Journal of Mathematical Sociology* 5 (1977), 173–186.

Hardin, Garrett, 'The Tragedy of the Commons', *Science* **162** (1968), 1243–1248.

Hardin, Russell, 'Collective Action as an Agreeable *N*-Prisoners' Dilemma', *Behavioral Science* **16** (1971), 472–481.

Hume, David, *A Treatise of Human Nature*, Clarendon Press, Oxford, 1960.

Macpherson, C. B., *The Political Theory of Possessive Individualism: Hobbes to Locke*, Clarendon Press, Oxford, 1962.

Schelling, Thomas C., 'Game Theory and the Study of Ethical Systems.' *Journal of Conflict Resolution*, Vol. 12 (1968), pp. 34–44.

Schick, Frederic, 'Democracy and Interdependent Preferences', *Theory and Decision* 3 (October 1972), 55–75.

Schick, Frederic, 'Some Notes on Thinking Ahead', *Social Research* **44** (Winter, 1977), 786–800.

Shubik, Martin, *Strategy and Market Structure*, Wiley, New York, 1958.

Shubik, Martin, 'Game Theory, Behavior, and the Paradox of the Prisoner's Dilemma: Three Solutions', *Journal of Conflict Resolution* **14** (1970), 181–193.

Shubik, Martin, 'Games of Status', *Behavioral Science* **16** (1971), 117–129.

Valvanis, Stefan, 'The Resolution of Conflict When Utilities Interact', *Journal of Conflict Resolution* 2 (1958) 156–169.