A THEORY OF GAMES WITH TRULY PERFECT INFORMATION

by

Earl A. Thompson

University of California at Los Angeles

and

Roger Faith

University of California at Santa Barbara

UCLA Discussion Paper No. 61

INTRODUCTION*

In conventional game theory, each individual either must consider the strategies of the others as given or must accept artificial constraints on the contingent strategies he may communicate to others. The information which individuals thus possess is less accurate than that frequently implied in discussions of social institutions in that the latter discussions frequently have an individual knowing how other individuals would respond to each of his different possible strategies and freely selecting and communicating contingent strategies. An example occurs when an individual knows how others would respond to his different reactions to attempts at taking his property and commits himself to a reaction to thievery which prevents the theft from occurring, although actually performing the reaction would impose more damages on him than his property is worth. In this paper, we drop the conventional game-theoretic assumption, replacing it with the assumption of perfect information regarding player reactions and unrestricted strategy selection. As our model implies neither inaccurate information regarding player reactions nor the limitation on the communication of contingent strategies implied by the conventional assumption of "perfect information", we label our assumption "truly perfect information".

If we wish the term "game" to include behavior under truly perfect information, then the current "theory of games", given the inaccuracies in each player's information about the reaction of others or the limitations on the admissible strategies, should be relabelled, "a theory of interpersonal gambling and parlor games", where a player is restricted to either guess about how the other players will react to changes in his strategy or to a prespecified order of actions. Our model would then be included in the general theory of games as a theory of rational institutions (i.e., rational reactions), a theory in which each individual determines his rational reactions to the strategies of others with no constraints on the set of feasible strategies and communicates these reactions costlessly in full knowledge of the rationally-determined responses of the other individuals. Thus, in contrast to the received theory, which accepts a given set of institutions, our theory determines the institutions that a group of non-cooperating individuals would rationally adopt in the absence of informational limitations.

There is room for our model in a game theorist's taxonomy in that ours is a model implying"cooperation" or "communication" in which rules are being determined. However, conventional game theory does not contain a general model of <u>rational</u> communication. As a result, existing cooperative game theory contains only conjecture as to the characteristics of solutions to the communication process.

The weakness of conventional game theory's treatment of contingent strategies has also been noted by Schelling (7) and Howard (4). Our central contributions are in constructing a general game and a solution concept appropriate to formally removing the weakness, evaluating the resulting solutions with respect to Pareto optimality, applying the results to the derivation of various economic and political institutions, and comparing the solutions with those resulting from standard game theory.

In Section I, we specify our general game and show that a necessary consequence of rational strategy selection under truly perfect information is a "hierarchy" of players characterized by a recursive set of rationally-determined reaction functions for all but one player, who chooses a simple action and thereby determines the solution set of actions. Such solutions always exist for finite games.

In Section II, we show that the solution set to any finite game with truly perfect information always contains a Pareto optimum. The importance of this optimality result is that while numerous authors have conjectured the Pareto optimality of n-person conflict under sufficiently perfect information, the conjecture has never been derived as the result of the individually rational behavior of the players within the model. However, we also show that the conjecture is not always true, i.e., that the solution set of a game with truly perfect information may also contain Pareto nonoptimal points. No amount of perfecting of information or communication will remove these inefficient solution points. The only way, to our knowledge, of removing these inefficient solution points is to introduce a higher-order game involving competition to determine the order of strategy selection.[1/]

Section III outlines an application of the model to the derivation of alternative economic and political institutions under alternative assumptions on the environment and notes the rough historical accuracy of the derived institutional evolution. It also outlines a biological application of the n-player optimality theorem.

In Section IV, we contrast our derived game to Von Neumann-Morgenstern perfect information games, to Von Neumann-Morgenstern-Nash imperfect information games, to Nash cooperative games and to Friedman's supergame. With respect to the perfect information, zero-sum games of Von Neumann-Morgenstern, our theory yields identical rational reaction functions and solutions. However, this equivalence exists only as a consequence of the zero-sum condition imposed upon the payoffs and vanishes under variable-sum games. In contrast to the Von Neumann-Morgenstern-Nash, "no-regret" equilibria for general-sum games, our solution to the "prisoner's dilemma" is an efficient solution even in the absence of two-way communication between the players. A basic feature that distinguishes our game from the Nash cooperative game and the Friedman "supergame", both of which are designed to allow sufficient communication to prevent Pareto nonoptimal solutions in 2-person games, is that we do not assume Pareto optimality to be a characteristic of solutions or of points on rational reaction functions. The fact that Pareto optimality may not characterize all of the solutions to 2-person interactions no matter how perfect the communication indicates the weakness in these other approaches. We illustrate the non-optimality possibility in a possibly empirically relevant game which we call the "slave master's insensitivity". Another distinguishing feature of our theory is that it is a non-normative theory yielding a determinate distribution of payoffs.

I. THE GENERAL GAME

A. Description

Let G be a n-person, non-cooperative, finite game. A player of G is denoted i, i = 1,...,n, and we shall let N represent the set of n players of G. Each player has an action possibilities set, X_i, consisting of a finite number of possible pure actions, x_i, for all i.

An outcome of G is an n-dimensional vector of pure actions, and is denoted

$$x = (x_1, x_2, \ldots, x_n), \quad x_i \in X_i, \quad \text{for all } i = 1, \ldots, n. \qquad (1)$$

We shall denote a vector of pure actions, one from each player except i, as

$$x_{-i} = (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n). \tag{2}$$

Each player, i, has a payoff function,

$$\pi_i(x) = \pi_i(x_i, x_{-i}), \qquad i = 1, \ldots, n, \tag{3}$$

which describes i's payoff under all possible outcomes of G.

A strategy of a player is a description of what the player will do. One possible strategy is an action, which simply states what pure action the player will choose. Other strategies, chosen prior to certain, specified strategies of others and perhaps requiring a commitment by the player, also exist. These are called contingent strategies. Among these strategies is a reaction strategy, which state the various pure actions the player will choose under the various possible actions of other players. There are also reactions contingent on reactions which are themselves contingent on reactions, and so on. One possible strategy, for example, is expressed by the infinite statement:

"I will play x_1' if you say you will react with x_2'' when I play x_1'' if I say I will react with x_1''' when you play x_2''' if you say you will react with x_2'''' when I play x_1'''' if I say I will react with x_1''''' when you play x_2''''' if you say...."[2]

We assume all players have perfect information regarding the strategies of others and "rationality", i.e., that each player selects the strategy that maximizes his payoff given the responses of the other players to his strategy.

B. <u>The Necessity of Priority in Strategy Selection</u>

If there is to be truly perfect information, there must be one player among the set of n players that determines his strategy prior to the other players, where the response strategies of the other players are known a priori by the prior selector because they are the rational responses to the given strategy of the prior selector.[3/] Given the strategy of the first strategy selector, another player makes a prior strategy selection over the set of the remaining n-2 players in full knowledge of the previous selector's strategy. This strategy selection process continues until all n players have selected a strategy.

A strategy of a prior strategy selector among a set of players describes his response to <u>each</u> of the possible strategies of the other players in the set. Therefore, since each player has an available set of pure actions, each of the possible strategies of a prior strategy selector always contains a reaction strategy.

It follows that the n^{th} strategy selector, who faces only prior strategy selectors can achieve maximum payoffs by simply selecting a pure action from X_n, thereby triggering a chain of reactions up through the first strategy selector in the game. This is because the n^{th} strategy selector, facing the prior strategies of the other players, sees that the eventual set of actions, or outcome, must be consistent with the chosen reaction strategies of each of the n-1 prior selectors. Hence, if player n responds with a pure action, he will have a free choice over all outcomes consistent with the prior reaction strategies. But if n responds with a contingent strategy, thus giving further choices to the prior strategy selectors, he can only reduce his original choice out of the same set of possible outcomes without expanding the set of possible outcomes, as any eventual set of actions must be consistent with the initial n - 1 reaction functions.

Since each prior selector's strategy will be followed by a set of pure actions by the subsequent strategy selectors, each prior selector's strategy set is larger than the set of pure actions contained in the choice set X_i only in that it includes player i's possible reactions to all of the possible actions of the subsequent strategy selectors.

Let us label the n players according to their priority in strategy selection. Then, i's strategy set, S_i, is described by:

$$S_i = \{x_i|(x_{i+1},\ldots,x_n), \text{ all } x_i \in X_i, \text{ and } (x_{i+1},\ldots,x_n) \in X_{i+1}\times\cdots\times X_n\}, \qquad (4)$$

where $x_i|(x_{i+1},\ldots,x_n)$ is i's action contingent upon the vector of actions from players i+1 to n. Of course, $X_i \subset S_i$. Player i's optimal strategy, $i < m$, is summarized by a reaction function.

$$f_i: \mathop{\times}_{j=i+1}^{n} X_j \rightarrow X_i, \qquad (5)$$

which assigns an action to i for each vector of actions from player i+1 to n, given the prior reaction functions of player 1 through i-1.

Any player i who has priority over a non-empty subset of N is called a strategy-maker (or, maker) relative to the players in that subset. Every other player in the subset is called a strategy-taker (or, taker) relative to i. Player 1, who has priority over all others, is called the primary maker. Player n, who is not a maker relative to any other player in N but a taker relative to all others, is called the pure taker. Player i, $i \neq 1$ or n, is a maker relative to players i+1 through n, and is taker relative to players 1 through i-1.

We thus obtain a hierarchy of makers where the primary maker exhibits a reaction function

$$x_1 = f_1(x_2,x_3,\ldots,x_n), \tag{6}$$

the secondary maker exhibits,

$$x_2 = f_2(x_3,\ldots,x_n),$$

the tertiary maker exhibits,

$$x_3 = f_3(x_4,\ldots,x_n),$$

and so on up to the n-1st player's simple reaction function,

$$x_{n-1} = f_{n-1}(x_n).$$

C. Solutions

Solution sets for all n players are easily constructed for a given set of reaction functions. The pure taker selects an action that maximizes his payoffs given the set of n-1 reaction functions. The n-1st player then follows his established reaction function, which gives the two actions necessary for the n-2nd player to determine his action. This process continues until the primary maker's action is determined.

With the assumption of truly perfect information and the condition of rationality imposed upon the takers of a given maker, the maker can solve for his solution reaction function.

This determination can be made since each maker, i, knows: (1) the reaction functions previously selected by his makers, (2) the reaction functions that will be chosen by i's takers given i's choice of reaction function and the previously selected functions of i's makers and (3) player n's rational response to these reaction functions.

Thus, a solution to G is directly obtained from a set of n-1 independent reaction functions, $\{f_i^*, i \neq n\}$ where f_i^* is a solution to

$$\max_{f_i \varepsilon F_i} \quad \pi_i \, [f_i, f_{i+1}(f_i), f_{i+2}(f_i, f_{i+1}(f_i)), \ldots, x_n(f_i), f_1^*, f_2^*, \ldots, f_{i-1}^*], \qquad (7)$$

where F_i is the set of all possible functions relating $\chi_{j=i+1}^{n} X_j$ to X_i. The set of actions, $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$, implied by the set of solution reaction functions is called a solution outcome, or simply a "solution," to G. The recursive and finite structure of the optimization problems of the maker and taker obviously insures that the solution set is non-empty.

D. Determination of Priority and the Role of Commitments

Our game G, with its predetermined priority, is not symmetric in that the solution to G will vary with the order of priority. One may think of the priority in strategy making in the above model as being arbitrarily determined by an umpire of the game, or by some random device allocating the positions in the hierarchy.

The priority of strategy selection can also be determined by playing a "higher-order" game, G'. In such a game, the n players interact to determine the order of strategy selection according to the value (payoff differential) of selecting strategy prior to the other players. Since in order for any player to know this value he must determine his solution reaction function given a particular position in the hierarchy, a solution to G' also implies a solution to G.[4/]

Another reason for introducing a higher order game is to assure that a maker can carry out his chosen strategy. That is, a player may contract, in a higher-order game, with outside parties to pay the outside parties a large sum in case the player does not carry out his announced strategy. For example, a player may adopt a strategy of imposing punishments on other players if they do not adopt certain strategies, punishments which are painful for the player to impose. Without a commitment to an outside party to carry out this strategy or else suffer even greater pain, the announced strategy might be regarded as an empty threat by the other players. This higher order game of contracting with outside parties and the higher order game, G', which determines the order of strategy selection may be combined into a single, higher-order game in which players interact to determine the order in which they make strategy commitments with outside players. (See Faith and Thompson (2).)

II. PARETO OPTIMALITY

In this section, we shall first show that the solution set to any finite game with truly perfect information always contains a Pareto optimal point. The solution set is simply the set of solution outcomes, i.e., the set of all outcomes such that the reaction function for the maker given the rational response of the takers yields maximum payoffs to the maker.

To prove that the solution set must contain at least one Pareto optimal point, suppose it is not the case. Then, for any set of solution actions, say the set $s = (s_1, s_2, \ldots s_n)$, there is a set of non-solution actions, say $p = (p_1, p_2, \ldots, p_n)$, such that

$$\pi_i\ (p) \geq \pi_i\ (s) \quad \text{for all } i, \text{ and}$$
$$\pi_i\ (p) > \pi_i\ (s) \quad \text{for some } i. \tag{8}$$

If player n chooses action p_n, then players 1, ..., n-1 must choose an action set other than p_{-n}; otherwise player n would, in view of (8), be at least indifferent between p and s so that p would be in the solution set. We must now construct a set of reaction functions the same as the original set except in the reactions of some players to subsets of p. In this set of reaction functions, first player 1 exhibits a reaction of p_1 to p_{-1}. In view of (8), he is no worse off than at the solution point, s. This is because the only set of actions which subsequent decision makers will possibly revalue is p_{-1}. All other sets draw the same response from player 1 and therefore have the same payoffs to them. Then, player 2 reacts to p_3, p_4, ..., p_n with p_2 and is similarly not hurt. The same is applied to players 3 to n-1. Thus, under the constructed set of reaction functions, each of the n-1 makers are at least as well off as under the original set of reaction functions. Therefore, since at least one player's solution reaction function yields him no greater payoff than a non-solution reaction function, the supposition that p was not in the solution set has led to a contradiction.

The existence of Pareto optimal points in the solution set does not imply that Pareto nonoptimal points cannot also belong to the solution set.

In fact, Pareto nonoptimal points do belong to the solution sets to certain games. These occur when there are Pareto nonoptimal points for which the only Pareto superior points have less than n parties better-off. Then, assuming that a maker is one of the parties who cannot be made better-off in moving to a Pareto optimum point, and that the Pareto optimum is in the solution set, it is clear that a Pareto nonoptimum yielding the same payoff to the maker but a lower payoff to his takers is also a solution as long as the other choices offered by the maker are even worse for the takers than that Pareto nonoptimum. A quasi-realistic example, a game we call the "slave master's insensitivity", is discussed in Section IVb. (Figure 5 illustrates the Payoff Matrix.)

Another inefficiency, which is discussed in the following section, occurs when the model is completed to determine the positions in the hierarchy and n represents the entire population. In that case, there is no "outside player" to sell-off the rights to make commitments in order to remove the substantial resource drain on society when the primary maker establishes the priority of his commitment.

III. APPLICATIONS

The above nonoptimalities do not have prescriptive relevance when applied to the entire population. The solution will not be changed, even when it is Pareto nonoptimal, because there is no higher order of enforcement to appeal to for allocative correction. No institutions exist before the first reaction function is formed. But once the first commitment is made, the theorem may have prescriptive relevance when applied to the remaining set of n-1 players. That is, the primary maker, once established, may be concerned with the inefficient solutions or the resource drains in the process of becoming secondary maker, etc., and alter his reaction functions accordingly.

To lend substance to this general remark, we shall now outline an application of the general model to the derivation of alternative institutions. First, let us assume that there are punishment actions sufficient to induce actions that result in an overall payoff maximum to the primary maker.[5/] We can then call this individual the dictator and the remainder of the individuals his subjects. The remaining set of reaction functions are degenerate, as the actions of the subjects are exactly those prescribed by the dictator. However, for anything but small, tribal societies, it is implausible to assume that the dictator costlessly knows which set of actions achieves his maximum payoff. The dictator may then gain by switching to the other extreme, permitting a system which is unconstrained, or anarchistic, except for lump sum taxes. Applying our optimality theorem (assuming away the insensitivity solution) to the remaining n-1 individuals, and assuming a sufficiently inexpensive process of collecting lump-sum taxes, this would indeed be an optimal strategy for the dictator. The potential resource drains in allowing the n-1 individuals to compete for priority in the remaining commitment hierarchy can be substantially reduced by the dictator's assigning these priorities himself. However, the dictator may find that the process of making lump-sum tax, or "side payments", between the remaining n-1 individuals is a heavy, socially unnecessary, resource drain on the group. To reduce the resulting deadweight losses, the dictator may impose a restriction on reaction functions, preventing actions which benefit one of the other individuals at the expense of the remaining n-2 other individuals, given a dictatorially determined set of benchmark actions. This restriction on reaction functions defines a system of private property (see Thompson (8)). But the restriction also prevents the application of the above optimality theorem to the n-1 players. While the equilibrium would generally be significantly nonoptimal, imposing further, anti-trust-type restrictions on reaction functions restores the Pareto optimality of the resulting equilibrium (see Faith and Thompson (2)).

While the existence of the punishment actions assumed above is not implausible, the conclusion that individually rational social interaction implies dictatorial allocations may, on the surface, appear to be implausible. We must remember, however, that most private and tribal societies prior to the industrial revolution had private property but dictatorially determined taxation. The industrial revolution introduced technolgies of easy duplication rather than revolutionizing the quality of the goods which a dictator could receive. Hence, the new dictatorial solution gave many more goods to the subjects, assuming some benevolence on the part of the dictator. The losses in utility to the subjects within a private property system under imperfect information (e.g., losses from externalities and avoidable transaction costs) become magnified once the subjects are allowed significant surpluses. Since a voting mechanism often provides alterations in the private property system which are superior to those provided by a dictator armed with the best economic theories (Thompson (9)), it is likely that the rational benevolent dictator, facing the new, complex, significant losses in utility to his subjects, would set up a voting system to alter the private property system. The apparent control that voters have over the benefits of their dictators (military leaders) in the real world, we submit, is an illusion. Military leaders, as a group, have both tenure and an absence of significant institutional constraints on the goods and services they can command. The frequency of military takeovers of democratic governments that generate unsatisfactory results from the standpoint of the military in medium-poor countries is evidence for the dominance of the military. It is also further evidence, given the additional fact that wealthier countries typically have non-military authorities while poorer countries typically have direct military rule, for the above argument on the rationale for popular democracies.

Another apparent difficulty arises from the fact that some countries have gone to authoritarian socialism rather than democracy following the industrial revolution. This fact, however, can be easily explained within our model. The prominent two countries of this variety, USSR and Red China, historically had numerous military centers of power and, as a result, a substantial pyramid of military leaders to satiate with goods. Hence, the industrial revolution of the 18th and 19th centuries may not have generated sufficient surplus for the military to grant significant economic surpluses to the subjects. Since confiscatory lump-sum taxes are imperfect, working in the real world to sharply discourage the private construction of the large scale plants appropriate to modern technology, these countries have reverted to command economics in order to achieve the requisite investments in spite of relatively heavy informational demands of such systems. Our theory also predicts that as the centers of power concentrate and the productive capacities of these countries expand, the political and economic institutions in these countries will fall into line with the rest of the world.

While our optimality results have no prescriptive relevance when applied to the entire population, they may have some descriptive, biological relevance to the entire population. To illustrate this, we add the biological assumption that each player's payoffs are an increasing function of only the survival probabilities of himself and, perhaps, some other of his n-player group. (The possibility of symbiosis prevents us from identifying the n-player group as a single species.) If, by a thorough natural selection of preferences, this function were symmetric and identical to all players there would be no conflict as all players would share the same set of optimal actions, x^*. Each player, i, would, in turn, rationally select x^*_i, and there would be no need for reaction functions (i.e., social institutions). But it is empirically clear that selection has not been so thorough, that players typically place a higher value on their own (or their own family's) survival than on the survival of others.

It follows from our optimality theorem (ruling out the insensitivity solution and exogenously fixing hierarchial positions) that our social equilibrium has the property that there is no alternative set of actions such that the survival probability is greater for at least one player and no lower for all other players. In this sense, our social interaction cannot be biologically dominated by any other social interaction. Our interpretation of this result is that the kind of social intelligence which we find in man and many other mammals, where rational reaction functions are both taught and learned, will not be replaced, through natural selection, by another form of social intelligence. But this conclusion is conditioned by the assumption that the competition for hierarchical position does not absorb signifcant resources. If sufficient resources were devoted to achieving hierarchical position, inflexible, instinctive behavior, even if only instinctive acceptance of hierarchial positions, would dominate rational choice behavior. Similarly, if sufficient resources were required to establish strategy commitments, socially instinctive behavior regarding carrying out one's promise would dominate narrow rationality in player reactions. Also, if preferences were allowed to vary freely between players, the set of identical, symmetric payoffs mentioned above would dominate any other as this set would obviate any competition for hierarchial position. Hence, it is not surprising to see a persistence of benevolence and instinctive social behavior in higher mammals.

IV. CONTRAST TO TRADITIONAL THEORY OF GAMES

In this part, we shall contrast the approach to 2-person games with truly perfect information developed above with the approach of conventional game theory. In particular, we shall contrast our solutions to "perfect information" solutions, no-regret solutions, Nash's cooperative solutions and Friedman's supergame solutions.

A. Perfect Information Solutions

The constructive solution concept pioneered by Von Neumann and Morgenstern (VNM) for two-person, zero sum games is the minimax solution, where each player selects a strategy which maximizes the minimum gain (or minimizes the maximum loss) over all possible outcomes to the game. VNM's theorem on zero-sum "perfect information games" -- i.e., games with a fixed, temporal order of moves and perfect knowledge of the previous mover's move and the payoff matrix -- states that there is always a saddle point, and thus a minimax solution to such games.

To illustrate this theorem and contrast the VNM solution method to our own, consider a two-person, zero-sum game without a saddle point (and thus without a minimax solution) in pure actions (e.g., figure 1). Following VNM, when there is perfect information, there is one player, say player #1, whose choice set must be expanded beyond the set of simple actions to include that player's possible reactions to the various actions of his opponent. Player #2 thus has the "first move." The expanded payoff matrix is shown in figure 2 where $x_1' | x_2''$, for example, means that player #1 adopts his first action if player #2 adopts his second. The matrix has a saddle point at $(x_1' | x_2',\ x_1'' | x_2'')$ for player 1 and x_2'' for player 2. The saddle point corresponds to the solution to a game in extensive form in which player 2 must first pick an action knowing that player 1 will follow with his payoff maximizing action given 1's action. VNM use this to support the generality of the minimax solution.

The extensive form of a game with truly perfect information differs from that of VNM's perfect information game in that the former permits the second mover to commit himself to a reaction function prior to the move of the first mover. This results in a solution concept, outlined in the above section, which differs from that used in the extensive form of VNM's perfect information game. Using our solution concept for the game described in figure 2, first player 1, who is now the maker, rationally adopts his third strategy, i.e., plays action x_1' if player 2 plays x_2' and action x_1'' if player 2 plays x_2'', after scanning all possible strategies and computing his payoffs under 2's rational response. Then, facing this reaction function, player 2's payoff maximizing choice is x_2''. Finally, player 1 performs the action to which he is committed, x_1''. This is the same solution point as the VNM solution point. But this coincidence arises only because of the fact that in zero sum games, rational reaction functions are those which have maximal returns to the maker for each of the possible actions of the taker. In general sum games, rational reaction functions will not generally have this property. In general sum games, it will often pay an individual to commit himself to reactions which, in order to discourage certain actions by the taker, force the maker as well as the taker to be worse off than he could be for the given action of the taker. In general, then, the VNM solution concept is inappropriate with truly perfect information, as it reflects an unacceptable incapacity of the second mover to adopt and communicate strategies which do not simply maximize his payoff for the given action of the first mover.

1 \ 2	x_2'	x_2''
x_1'	4	1
x_1''	2	3

Figure 1. Payoff Matrix A - No Saddle Point

1 \ 2	x_2'	x_2''
x_1'	4	1
x_2''	2	3
$x_1'\|x_2', x_1''\|x_2''$	4	3
$x_1''\|x_2', x_1'\|x_2''$	2	1

Figure 2. Majorant of Payoff Matrix A

The fact that there is a difference between VNM's "perfect information" solution applied to variable sum games and our own solutions to these games is easy to see. If we apply VNM's perfect information solution to the "prisoners' dilemma game" in Fig. 3, we obtain the inefficient solution (x_1'', x_2''): The normal form with player 1 moving second is shown in Fig. 4. Subject to VNM's disallowance of commitment strategies, player 2's optimal strategy, in light of player 1's rational response, is to play x_2''. This leads player 1 to play x_1'' (or $x_1'|x_2', x_1''|x_2''$), resulting in the outcome (x_1'', x_2''). Our solution has player 1, while moving second, able to select a conditional strategy and communicate it to player 2 before he moves. Player 1 rationally selects $(x_1'|x_2', x_1''|x_2'')$ in view of 2's rational responses to 1's various possible strategies(reaction functions). Player 2 then rationally chooses x_2', so our solution is the efficient one, (x_1', x_2').

1 \ 2	don't confess x_2'	confess x_2''
don't confess x_1'	4,4	0,10
confess x_1''	10,0	1,1

Figure 3. Payoff Matrix $[\pi_1, \pi_2]$

1 \ 2	don't confess x_2'	confess x_2''
don't confess x_1'	4,4	0,10
confess x_1''	10,0	1,1
$x_1' \mid x_2'$, $x_1'' \mid x_2''$	4,4	1,1
$x_1'' \mid x_2'$, $x_1' \mid x_2''$	10,0	0,10

Figure 4. Majorant form of prisoners' dilemna.

B. "No-Regret Solutions"

A "no-regret solution" is a pair of strategies such that no player can increase his payoff by changing his strategy, given the strategies of the other players. (See Nash (5)) No-regret solutions are inappropriate to games with either perfect or truly perfect information. Such solutions describe behavior in which each individual, although he may know the other's strategy, assumes that the other's strategy is unaffected by his own strategy -- a straightforward generalization of a Cournot duopoly model. As in the model of Cournot (1), each player is behaving irrationally under truly perfect information because the other player will, by the rule, generally alter his action in response to a change in his opponent's action.

To illustrate the difference in the two concepts of equilibrium, we again consider the "prisoners' dilemma" game. This is depicted in normal form in figure 3 or, if reaction strategies exist, in figure 4.[6] In either case, the Nash equilibrium outcome to the game is the action pair (x_1'', x_2''). Under truly perfect information, this solution is <u>irrational</u>, since either player, if maker, can assure himself a higher payoff (4 rather than 1 in the illustration) by exhibiting a reaction function where he plays his first action if the other player plays his first action and he plays his second action if the other player plays his second, given the payoff-maximizing behavior of the taker as constrained by the maker's rational reaction function. This results in an equilibrium at (x_1', x_2'), which is unattainable under the Nash solution concept. Hence, in spite of the absence of two-way communication, our solution to a prisoners' dilemma game is an efficient solution.

Notice that the maker's reactions are not the payoff maximizing actions of the maker for given actions of the taker, as they were in the zero-sum case.

Finally, we compare our solution to the Nash non-cooperative solution when the payoff matrix takes a form admitting our Pareto non-optimal case. Figure 6 contains an example of such a matrix, which we call the "slave master's insensitivity". The Nash non-cooperative solution has the slave resting while the master insults the slave. This is both non-optimal and empirically unrealistic. Making the master the strategy maker, our solution set contains an optimum (10,0) as the master will beat the slave if and only if he rests in order to induce the maximal output from the slave. But the solution set also contains (10, -4) as the master may also insult the slave, lowering the slave's benefit to -4 without altering either the master's payoff or the slave's optimal taker decision. Our solution set thus contains Pareto non-optimal as well as Pareto optimal points.

master \ slave		work x_2'	rest x_2''
beat the slave	x_1'	5,-10	0,-6
insult the slave	x_1''	10,-4	1,0
leave the slave alone	x_1'''	10,0	0,4

Figure 5. The Slave Master's Insensitivity

1 \ 2		don't confess x_2'	confess x_2''
don't confess	x_1'	4,4	0,10
confess	x_1''	10,0	1,1
beat up the other prisoner	x_1'''	-1,-2	-1,-2

Figure 6. Amended Prisoner's dilemna

The non-optimality disappears when we allow a competitive bidding process to determine which of the two players is to be the strategy maker (see Faith and Thompson (2)).

C. Cooperative Games & Supergames

The Nash cooperative game (6) and Friedman's supergame (3) are two games without truly perfect information which yield Pareto optimal solutions. In Nash's cooperative game the players reach an outcome mutually superior to the Nash non-cooperative outcome by first simultaneously committing themselves to playing a "threat" action if they are not satisfied with the size of their individual payoff and their simultaneously selecting actions among the sets of actions which do not call forth the threat action of the other player. Assuming convexity of the smoothed set of player payoffs and maximizing a symmetric social welfare function, Nash determines an equilibrium set of actions, and derives the form of the optimal threat. The optimal threats, defined as the payoff maximizing reply to the threats of others, are the respective Nash non-cooperative equilibrium actions of each player.

Supergames are infinite sequences of conventional non-cooperative games. In Friedman's model, to achieve a sequence of Pareto outcomes as a solution each player selects a reaction strategy where: one plays his Pareto action in game t, if all other players play their Pareto actions in game t-1, and one plays his Nash non-cooperative action in t and all subsequent games in the sequence, given any other actions of his opponents in game t-1. In the initial game (t=1), it is assumed all players choose their Pareto action. It follows that each player is discouraged from choosing an action in any game which maximizes his payoff given the others' Pareto actions.

To illustrate cooperative, supergame, and truly perfect information game solutions, first consider the simple prisoner's dilemma game (Fig. 3). The three models' solutions all occur at (x_1', x_2') for a payoff of (4, 4). If we now add a new action, "beating up one's fellow prisoner"(Fig. 6), the cooperative and supergame solutions remain at (x_1', x_2'). However, if player 1 is the maker, he can achieve (x_1'', x_2') as a solution by committing himself to play action x_1''' whenever player 2 plays something other than x_2'. It is no longer rational for player 1 to resort to his Nash non-cooperative action, x_1'', if 2 deviates from this Pareto outcome. Thus, with this ability to punish and perfect information about reaction functions, the Pareto optimum occurring at (x_1', x_2') is not a rational solution.

Neither the Nash nor Friedman model, however, is sufficiently descriptive of the player interaction to derive rational player reactions. Pareto optimal solutions to these models are obtained via the imposition of outside conditions which render a description of rational player behavior unnecessary. Nash simply assumes that all possible solutions are Pareto optimal and imposes a symmetric social welfare function generating a unique solution to determine a distribution.[7/] Friedman's Pareto optimal solutions are attained by assigning both players a set of consistent but generally irrational reaction functions. They are, for example, irrational under sufficiently high discount rates (as Friedman points out), finite horizons, sufficiently slow learning, the availability of strategy commitments such as that in the above example (Fig. 6), or a sensitivity of one player's rational actions to the other's various non-Pareto actions.

A resulting weakness of these models is that neither generally solves the positive problem of payoff distribution. That is, the mutual conflict of interests is not resolved. Nash avoids the problem by assuming uniqueness of the equilibrium and a symmetric social welfare function. Friedman, while permitting multiple, non-equivalent, Pareto optimal equilibria, simply assumes that all players somehow agree which Pareto optimum to work for. Our model of truly perfect information, however, besides yielding rational, Pareto optimal solutions (as long as each player's payoff-maximizing outcome is unique), is able to select among multiple Pareto optima.

FOOTNOTES

* The authors are indebted to James Friedman, Louis Makowski, Joe Ostroy and Lloyd Shapley for some very helpful comments.

1. See Faith and Thompson (2).

2. While a general game theory containing contingent strategies has been produced by Howard (4), he does not assume truly perfect information and therefore does not develop a solution concept. Rather, he adopts, without substantive justification, the Von-Neuman-Morgenstern-Nash concept of no-regret solutions. We shall illustrate in Section IV below how our solutions differ from these no-regret solutions to games with unrestricted contingent strategies.

3. In case there is more than one response strategy yielding the maximum payoff, we assume, for the sake of determinancy, that the player gives the choice among his maximal strategies to the prior strategy selector. We call the resulting strategy the player's "rational response".

4. For an example of this simultaneous solution determination in a model of non-competitive interdependence, see Faith and Thompson (2).

5. This assumption does not require that the primary maker be able to simultaneously "punish" the rest of the players. He may simply "punish" the secondary maker if players 3 to n do not perform the actions dictated by the primary maker. The rational reaction function of the secondary maker is then to "punish" the tertiary maker if players 4 to n do not perform the specified actions. "Punishment" here refers to an action a player will select if and only if subsequent strategy selectors do <u>not</u> select a set of actions which maximize the payoff of the first player. Our assumption is that these punishments will result in lower

payoffs to the victims than those which they receive at the set of actions which maximize the payoffs of the primary maker.

6. Howard (4) has shown that it is redundant to consider contingent strategies beyond reaction strategies when applying Nash solutions to a game with unrestricted contingent strategies. However, as we have shown above, the appropriate solution, when contingent strategies are available, is a perfect information (or truly perfect information) solution rather than a Nash solution.

7. Any player symmetry such as that imposed by Nash is generally impossible under truly perfect information given an order of individual strategy selectors. However, in a higher level game where players compete for priority in strategy selection, player symmetry is a formally acceptable, perhaps even reasonable, assumption. For a model of such competition, see Faith and Thompson (2).

REFERENCES

1. Cournot, A.A., Researches into Mathematical Principles of Wealth. MacMillan Co., New York, 1929.
2. Faith, Roger L., and Thompson, Earl A., "A Model of Rational Non-Competitive Interdependence", Working Paper, Department of Economics, UCLA, 1974.
3. Friedman, James W., "A Non-Cooperative Equilibrium for Supergames", The Review of Economic Studies, 38 (1971), pp. 1-12.
4. Howard, Nigel, Paradoxes of Rationality, M.I.T. Press, Cambridge, 1971.
5. Nash, John, "Noncooperative Games", Annals of Mathematics, 54 (1951) pp. 286-295.
6. ______________, "Two Person Cooperative Games", Econometrica 21, (1953), pp. 128-140.
7. Schelling, Thomas C., The Strategy of Conflict, Oxford University Press, New York, 1963.
8. Thompson, Earl A., "A Reformulation of Orthodox Value Theory", UCLA Working Paper, 1971.
9. Thompson, Earl A., "Taxation and National Defense", J.P.E., July/Aug., 1974, pp. 755-783.
10. Von Newmann, J., and Morgenstern, O., Theory of Games and Economic Behavior, Princeton University Press, Princeton, 1944.