Discussion Paper No. 215

Arbitration of Two-Party Disputes

Under Ignorance

by

Ehud Kalai*

and

Robert W. Rosenthal**

April, 1976

* Department of Managerial Economics and Decision Sciences, Graduate School of Management

** Department of Industrial Engineering and Management Sciences, Technological Institute. Research supported by National Science Foundation Grant No. SOC-7401790.

1. Introduction

Two players, about to play a cooperative game, are unable to come to a binding agreement concerning their joint action. They resort to binding arbitration. The arbitrator seeks to assign a fair outcome (in some sense) to the game. There have been several suggestions in the game-theoretic literature as to how the arbitrator should proceed when he has complete information about the game. The problem has been less well-studied for the case of incomplete information on the part of the arbitrator. The major difficulty is that if the arbitrator must rely on the players to reveal information about the game, the players may have incentives to lie, depending on their expectations as to how the information they provide will be used.

In our view, the arbitrator's problem is to design a new noncooperative game for the players and a function from outcomes in this noncooperative game to outcomes in the original game. It would be nice if, based only on his incomplete information, the arbitrator could always construct a noncooperative game with a unique pair of undominated strategies which would give rise to a fair Pareto optimal outcome in the original game. In general, this seems too much to hope for; although our goal here is not to prove impossibility results. We shall show instead that a class of procedures (generalizing Nash's extended bargaining solution [9]), which seems to be a reasonable class for the arbitrator who has complete information about the game, can be recast as a class of schemes for the arbitrator who has very limited information about the game. More specifically, we shall establish certain desirable properties of these schemes when both players have finite pure strategy sets, Von Neumann-Morgenstern utilities, and complete information; all jointly randomized strategies are possible; the arbitrator has complete information about the strategy sets; and the arbitrator has no information about the utilities of the players for any joint strategies.

There is an extensive game-theoretic literature on bargaining and arbitration with complete information (see [3], [5], [6], [8], [9], [11], [12], [15], for example; and [7] and [10] for summaries of other approaches). To our knowledge, the only work on incomplete information is the fixed-threat model of Harsanyi and Selten [4]. In the Harsanyi-Selten model, however, the players have incomplete information; and there is no arbitrator. They propose a generalization of the Zeuthen-Nash solution. In our models, the players are assumed to possess complete information; and the arbitrator's information is either complete or severely restricted.

In Section 2, we shall describe a generalization of the Nash extended bargaining solution and discuss its properties for the case of the arbitrator with complete information. In Section 3, we shall describe how the information-restricted arbitrator may create schemes with desirable properties based on the solutions in Section 2. In Section 4, some possible variations and extensions are discussed.

2. The Complete Information Case

Let (A,B) be a pair of $m \times n$ payoff matrices for a fixed (but arbitrary) 2-player cooperative game in normal form. Let the set of mixed strategies for player 1 be

$$S_1 = \{x = (x_1, \ldots, x_m) \geq 0 : \sum_{i=1}^{m} x_i = 1\}$$

and for player 2 be

$$S_2 = \{y = (y_1, \ldots, y_n) \geq 0 : \sum_{j=1}^{n} y_j = 1\}.$$

The set of correlated strategies is denoted

$$S = \{z = (z_{11}, \ldots, z_{mn}) \geq 0 : \sum_{i=1}^{m} \sum_{j=1}^{n} z_{ij} = 1\}.$$

The pair of mixed strategies (x,y) gives rise to the correlated strategy $z = x^T y$ (i.e., $z_{ij} = x_i y_j$ for $i = 1, \ldots, m$, $j=1, \ldots, n$).

For any correlated strategy z, the expected payoff to player 1 is

$$v_1(z) = \sum_{i=1}^{m} \sum_{j=1}^{n} z_{ij} a_{ij}$$

and to player 2

$$v_2(z) = \sum_{i=1}^{m} \sum_{j=1}^{n} z_{ij} b_{ij}$$

(where a_{ij} and b_{ij} devote the $(i,j)^{th}$ elements of A and B, respectively). A correlated strategy z is <u>efficient</u> for (A,B) if there is no correlated strategy z' such that

$$v_1(z') \geq v_1(z),\ v_2(z') \geq v_2(z), \text{ and } (v_1(z'), v_2(z')) \neq (v_1(z), v_2(z)).$$

The completely-informed arbitrator's problem is to select a correlated strategy $\bar{z}$ with the following properties:

1) $\bar{z}$ is efficient for (A,B);

2) $v_1(\bar{z}) \geq \max_{x \in S_1} \min_{y \in S_2} \sum_i \sum_j x_i y_j a_{ij}$, and $v_2(\bar{z}) \geq \max_{y \in S_2} \min_{x \in S_1} \sum_i \sum_j x_i y_j b_{ij}$;

   and

3) $(v_1(\bar{z}), v_2(\bar{z}))$ is in some sense fair.

Thus, for any bimatrix game, the arbitrator must produce an efficient, individually rational (for otherwise, one of the players has a strong incentive to avoid arbitration) correlated strategy which is, by some standard, equitable.

Let $g: S \rightarrow S$ be any function satisfying

4) $g(s)$ is efficient, for all $s \in S$;

5) $v_k(g(s)) \geq v_k(s)$, for all $s \in S$ and for $k = 1,2$;

6) $v_k \circ g$ is continuous and quasi-concave for $k = 1,2$.

Now consider the noncooperative game H, depending on A, B, and g defined as follows. The pure strategy set for player k is S_k $(k = 1,2)$ and the payoff pair associated with the pure strategy combination $(x,y) \in S_1 \times S_2$ is $(v_1(g(x^T y)), v_2(g(x^T y)))$. It will not be necessary to introduce mixed strategies in the game H. The following is the obvious extension of Nash's observation [8].

Theorem 1: a) There exists at least one Nash equilibrium pair of strategies (x^*,y^*) in H.

b) If (x^*,y^*) is a Nash equilibrium in H, then $g(x^{*T}y^*)$ is efficient for (A,B).

c) All Nash equilibria in H give rise to the same payoffs and are interchangeable.

d) If (x^*,y^*) is a Nash equilibrium in H, then

$$v_1(g(x^{*T}y^*)) \geq \max_{x \in S_1} \min_{y \in S_2} \sum_i \sum_j x_i y_j a_{ij} \text{ and}$$

$$v_2(g(x^{*T}y^*)) \geq \max_{y \in S_2} \min_{x \in S_1} \sum_i \sum_j x_i y_j b_{ij} .$$

Proof: a) The existence argument is a standard one. For each $y \in S_2$, let

$$C_1(y) = \{\bar{x} \in S_1 : \quad v_1(g(\bar{x}^T y)) \geq v_1(g(x^T y)) \text{ all } x \in S_1\};$$

i.e., $C_1(y)$ is the set of best responses to y by player 1. Similarly, define

$C_2(x)$ to be the set of best responses to x by player 2 for all $x \in S_1$. From the continuity of $v_1 \circ g$ and the compactness of S_1, $C_1(y)$ is nonempty for all $y \in S_2$. From the quasi-concavity and continuity of $v_1 \circ g$, $C_1(y)$ is a convex-valued upper semicontinuous correspondence. Similar properties hold for $C_2(x)$. The correspondence which maps (x,y) into $C_1(y) \times C_2(x)$ therefore satisfies the hypotheses of Kakutani's fixed point theorem. Any fixed point is a Nash equilibrium, however.

b) Property 4).

c) Let (x*,y*) and (x',y') be any two Nash equilibria in H. Then

$$v_2(g(x'^T y^*)) \le v_2(g(x'^T y')).$$

With b), however, this implies

$$v_1(g(x'^T y^*)) \ge v_1(g(x'^T y')).$$

Hence, we have

$$v_1(g(x^{*T} y^*)) \ge v_1(g(x'^T y^*)) \ge v_1(g(x'^T y')).$$

By reversing the roles of (x*,y*) and (x',y') we obtain

$$v_1(g(x'^T y')) \ge v_1(g(x^{*T} y^*)).$$

Hence equality everywhere. Similarly for player 2's payoff.

d) Suppose x quaranteed more than $v_1(g(x^{*T} y^*))$ for player 1. Then, in particular, it would yield a higher payoff against y*. But $v_1(x^T y^*) \le v_1(g(x^T y^*)) \le v_1(g(x^{*T} y^*))$, a contradiction. Similarly for player 2. ||

The completely-informed arbitrator may solve his problem by selecting any g which satisfies 4), 5), and 6), computing any Nash equilibrium (x*,y*) for the resulting game H, and producing $z^* = g(x^{*T} y^*)$ as the arbitrated outcome for the original game. The efficiency and individual rationality of z* follow from Theorem 1. The fairness of z* depends largely on the specific form of g. (Note that since all equilibria of H give rise to the same payoffs, it does not matter which equilibrium is selected. Note also that a disequilibrium strategy not only hurts the responder but helps his opponent. Hence any equilibrium strategy guarantees each player the equilibrium payoff, regardless of the action of his opponent.)

We shall describe two particular functions which satisfy 4), 5), and 6). The first example is the function which results in Nash's extended bargaining solution. To construct this function, first consider D(A,B) - the convex hull of $\{(a_{ij}, b_{ij}) : i = 1, \ldots, m, j = 1, \ldots, n\}$. With each nonextreme Pareto optimal point p in D(A,B), associate the set of points in D(A,B) which lie on the line segment having the negative of the slope of the Pareto optimal segment containing p. Associate to each extreme Pareto optimum p in D(A,B), all points in the remaining region touched by p. (See Figure 1.)

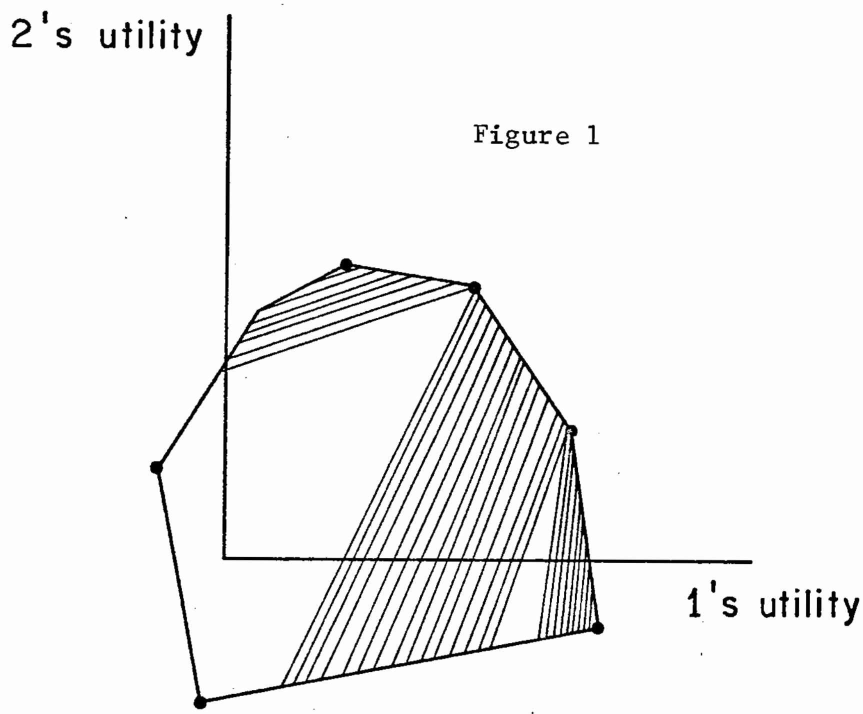

Figure 1

Now for each correlated strategy z in S, first compute $(v_1(z), v_2(z))$, then locate its associated Pareto optimal utility pair in D(A,B) and take g(z) to be any correlated strategy yielding that utility pair. Properties 4), 5), and 6) follow immediately.

The second example is related to a fixed-threat scheme in [6] and [11]. For every point p in D(A,B), consider the line segment joining p to

$(\max_{(i,j)} a_{ij}, \max_{(i,j)} b_{ij})$. Now, associate with p the Pareto point in D(A,B) which lies on its line segment. (See Figure 2.)

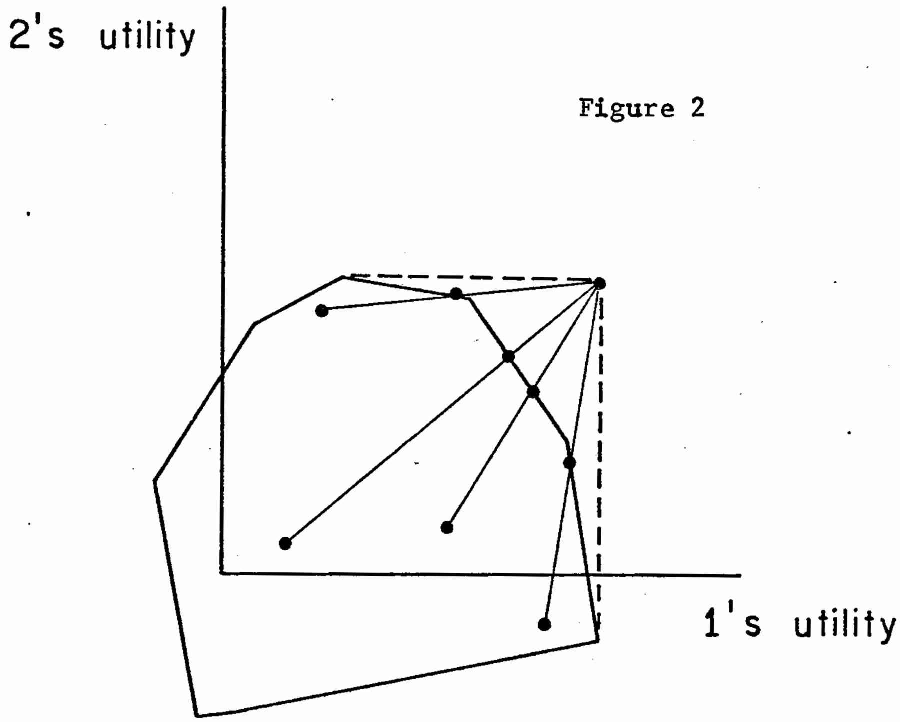

Figure 2

Again, for each z first compute $(v_1(z), v_2(z))$. Then take the associated Pareto optimal point in D(A,B) and choose for g(z) any correlated strategy yielding that utility vector. Again, properties 4), 5), and 6) are easy to verify.

In both instances, arguments for particular notions of fairness represented by the respective g's may be found in the respective references. Note that actual computation of the arbitrated outcome in both of these instances may be difficult in practice.

As demonstrated by these two examples, Theorem 1 can be used to extend many solutions for the fixed-threat bargaining problem to the variable-threat problem.

3. The Incomplete Information Case

The situation in this section is the same as that of the previous section; except that we assume that the arbitrator does not know the pair (A,B), but only that the game is an element of the set Γ of all 2-person, $m \times n$, normal-form games. The players, on the other hand, know exactly which element of Γ is before them. The arbitrator's problem here is to design a non-cooperative game for the players, the outcome of which will become the arbitrated outcome of the original cooperative game. He must do this in such a way that the arbitrated outcome can be expected to be efficient, individually rational, and fair in the game (A,B). This must all be accomplished, of course, in such a way that the arbitrator does not make use of any information concerning the true game (A,B).

Here are some unsatisfactory schemes which the arbitrator might use.

a) Force the players to play (A,B) noncooperatively. Unfortunately, there is no reason to expect a Pareto optimal outcome for this scheme.

b) Pick a player at random. Let this player dictate a correlated strategy for the cooperative game. Unfortunately, even if we could be sure that an efficient strategy would be selected by the dictator, the scheme would not be "ex ante" Pareto optimal in that it produces a random combination of Pareto optima which is not, in general, Pareto optimal. In addition, individual rationality is not guaranteed.

c) Ask each player what his payoff matrix is, then use one of the arbitration schemes from Section 2 on the reported pair of matrices. Unfortunately, for any specified function g, there will usually exist incentives for the players to misreport their utilities. Thus it is not in general likely that this scheme will even result in efficient correlated strategies for the cooperative game.

Consider now the following noncooperative game. First, the arbitrator announces a function g satisfying 4), 5), and 6) for each $(A,B) \in \Gamma$. (For what follows it is simplest to redefine g so that its domain is $(S \times \Gamma)$). For each player k, the set of strategies is $(S_k \times \Gamma)$, i.e., each player selects a mixed strategy for himself (a threat) and a game (interpreted as his report of the true underlying cooperative game). The outcome is determined as follows for the strategy (x, A_1, B_1) of player 1 and (y, A_2, B_2) of player 2. If $A_1 = A_2$ and $B_1 = B_2$, the outcome is $g(x^T y, A_1, B_1)$. If $A_1 \neq A_2$ or if $B_1 \neq B_2$, the outcome is $x^T y$. The utilities for all of these strategy combinations are the utilities for the resulting correlated outcomes in the original cooperative game. The noncooperative game will be denoted G(A,B,g). Note that the game G(A,B,g) can be imposed by the arbitrator without his knowledge of (A,B).

Theorem 2: For fixed (A,B,g), let H be defined as in Section 2. Let (x^*, y^*) be any Nash equilibrium for H. Then $((x^*,A,B), (y^*,A,B))$ is a Nash equilibrium for G(A,B,g) which yields an efficient, individually rational correlated strategy for (A,B).

Proof: Consider 1's best response to (y^*,A,B). No matter what element from S_1 he chooses, he cannot do better than to report (A,B), since g satisfies 5). If he reports (A,B), however, he cannot do better than to select x^*, since (x^*,y^*) is a Nash equilibrium for H. Similarly, (y^*,A,B) is a best response to (x^*,A,B). The efficiency and individual rationality of $g(x^{*T} y^*, A, B)$ follow from Theorem 1. ||

A major difficulty with arbitration schemes of this class is that there generally exist many other Nash equilibria for the game G(A,B,g). Many of these equilibria involve both players reporting a pair $(A',B') \neq (A,B)$. One may think of such equilibria as having one player convincing the other that he is committed to reporting (A',B'). If the choice of (A',B') is judicious, the other player will have no choice but to report it also. See [12] for related considerations. The equilibrium specified in Theorem 2 is appealing, however, in that it is the natural one which is suggested in the absence of communication. This fact alone weakens the ability of either player to convince the other of his commitment to report something other than (A,B). See [13] and [7] for intuitive discussions of this point.

As in the last section, it does not matter to the players which equilibrium they select in H, as long as they both report the true (A,B). The fairness of this particular class of schemes depends on the specific g selected and on our willingness to believe that one of the truthful equilibria of G(A,B,g) will arise. Of course, it would be nice to be able to design a noncooperative game with a unique equilibrium with desirable properties, but that seems to be an impossible task in this setting.

4. Remarks

An obvious direction for further research suggested by this model is a generalization to accomodate incomplete information on the part of the players and different kinds of incomplete information on the part of the arbitrator. The desirable features of the schemes described in Section 3 depend heavily on both players having complete information about the game and the function g selected by the arbitrator, even to the point of both players using the same choice of origin and scale factor for reported utilities.

This may be weakened, of course, by allowing the arbitrator to adjust both reports to a common origin and scale before checking whether or not the reports agree. Such a variation would have no effect on the result as long as the function g was independent of affine utility transformations, which is the case in the two examples of Section 2 (see [6] and [9]). In addition, the arbitrator could tolerate other differences in the reports of the players, as long as $g(x^{*T}y^*, A_1, B_1) = g(x^{*T}y^*, A_2, B_2)$.

Unfortunately, it is not difficult to find examples in which arbitrarily small differences in the reports would invalidate this last equality. This would seem to rule out the use of our schemes except under conditions of essentially complete information on the part of the players, conditions rarely seen in practice.

On the other hand, our schemes make no use of any knowledge of the players' utilities which the arbitrator may possess. A scheme which incorporated such knowledge might enable one to avoid (at least partially) the assumption of complete information on the part of the players or the problem of multiple equilibria.

Another possibility for generalization is to the case of more than two players. In this context, it is interesting to note the similarities between this work and the recent work on incentive mechanisms for economies with public goods and other externalities (e.g., [1], [2], [14]). In particular, the mechanisms described in [2] (and its references) and [14] have properties very similar to the schemes we have proposed in Section 3.

References

1. Groves, T., and Ledyard, J., "Optimal Allocation of Public Goods: A Solution to the 'Free Rider' Problem", Econometrica, forthcoming.

2. Groves, T. and Loeb, M., "Reflections on 'Social Costs and Benefits and the Transfer Pricing Problem'", Journal of Public Economics, forthcoming.

3. Harsanyi, J., "Approaches to the Bargaining Problem Before and After the Theory of Games: A Critical Discussion of Zeuthen's, Hicks', and Nash's Theories", Econometrica, Vol. 24, 1956, pp. 144-157.

4. Harsanyi, J., and Selten, R., "A Generalized Nash Solution for Two-Person Bargaining Games with Incomplete Information", Management Science, Vol. 18, 1972, pp. P80-P106.

5. Kalai, E., "Proportional Solutions to Bargaining Situations", Discussion Paper No. 179, Center for Mathematical Studies in Economics and Management Science, Northwestern University, 1975.

6. Kalai, E. and Smorodinsky M., "Other Solutions to Nash's Bargaining Problem", Econometrica, Vol 43, 1975, pp. 513-518.

7. Luce, D., and Raiffa, H., Games and Decisions, John Wiley and Sons, New York, 1957, Chapter 6.

8. Nash, J., "The Bargaining Problem", Econometrica, Vol. 18, 1950, pp. 155-162.

9. ________, "Two-Person Cooperative Games", Econometrica, Vol. 21, 1953, pp. 128-140.

10. Owen, G., Game Theory, W.B. Saunders Co., Philadelphia, 1968, Chapter 7.

11. Raiffa, H., "Arbitration Schemes for Generalized Two-Person Games", Annals of Mathematics Studies 28 (Kuhn and Tucker, Eds.), Princeton University Press, 1953, pp. 361-387.

12. Rosenthal, R., "An Arbitration Model for Normal-Form Games", Mathematics of Operations Research, Vol. 1, 1976, pp. 83-88.

13. Schelling, T., The Strategy of Conflict, Oxford University Press, London, 1960.

14. Smith, V., "Mechanisms for the Optimal Provision of Public Goods", Technical Report, Department of Economics, University of Arizona.

15. Zeuthen, F., Problems of Monopoly and Economic Warfare, G. Routledge and Sons, London, 1930.