THE REVELATION PRINCIPLE FOR ARBITRATION GAMES

By

Somdeb Lahiri

W P No. 855
March 1990

The main objective of the working paper series of the IIMA is to help faculty members to test out their research findings at the pre-publication stage.

INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD-380 056
INDIA

Abstract

In this paper we prove the validity of the revelation principle for arbitration games and also establish that a large class of game forms can be represented as a Bayesian Statistician. This adds force and appeal to the concept of an arbitration game.

1. <u>Introduction</u> : The revelation principle originally due to Gibbard (2), has been modified by Myerson (5), in order to be applicable to Bayesian equilibria of games. It has been used widely to explain diverse economic behaviour as equilibrium outcomes of games (see Ledyard (4)). The conventional context of analysis is the Bayesian Collective-Choice problem, which is an incomplete information game in which outcomes are jointly feasible for all players together. A mechanism is used to provide the collective choice, given information provided by the agents. However, no assumption of Bayesian rationality is imposed on the mechanism in the conventional framework.

In Lahiri (3), the framework of an arbitration game has been formally outlined. In essence, an arbitration game is a 'special case' of a Bayesian Collective Choice problem where the mechanism used to provide collective choice is controlled by a <u>Bayesian Statistician</u>, i.e. the mechanism itself is rational. The question that immediately arises is that, to what extent does such an assumption restrict the scope of analysis ? Would the 'revelation principle' still continue to be valid ?

The purpose of this paper is to answer this latter question in the affirmative. This is a key step to further issues regarding the applicability of arbitration games.

2. The Model : Suppose there are n agents numbered $i = 1, \ldots, n$. We let T^i denote the set of possible types of agent i, and let $T = T^1 \times \ldots \times T^n$. We assume for simplicity of exposition, that each T^i is a finite set. D is the set of possible outcomes or group choices. Once again for simplicity, let us assume that D is a finite set of outcomes. Each agent (or player) has a Von Neumann – Morgenstern utility function , $u^i(d,t)$, which denotes the payoff to i if d is the group choice and if $t = (t_1, \ldots, t_n)$ is the vector of players' types. Each agent also has a probability distribution $p^i(t_{-i}, t_i)$ where $t_{-i} \in T_{-i} = T^1 \times \ldots \times T^{i-1} \times T^{i+1} \times \ldots \times T^n$, which denotes the subjective probability that player i would assign to the event t_{-i} if i's actual type were t_i.

The tuple $\alpha = [T, D, u^1, p^1, \ldots, u^n, p^n]$ is called a Bayesian Collective Choice problem by Myerson (5). We assume that each player i knows his own type.

The method by which outcomes are selected, as a function of players' types, is a mechanism or a game-form. A mechanism is a pair $\langle M, g \rangle$ where $M = M^1 \times \ldots \times M^n$. The set M^i is the set of possible messages agent i can use and M is called the language. The outcome function $g(m^1, \ldots, m^n)$ maps M into probability distributions on D. We let $\Delta(D)$ be the set of all such distributions and assume that the structures of α and $\langle M, g \rangle$ are common – knowledge.

Given a mechanism $\langle M, g \rangle$, agents choose messages m^i as a function of their types and their common knowledge. We call a mapping $\beta^i : T^i \rightarrow M^i$ a strategy for i.

Let $\mu(d,m)$ be the probability assigned to d by the distribution $g(m)\in\Delta(D)$. Then

$$\sum_{t_{-i}\in T_{-i}} \sum_{d\in D} u^i(d,t)p^i(t_{-i},t_i)\mu(d,\beta^1(t_1),\dots,\beta^n(t_n))$$

represents the expected utility payoff to i if i's type is t^i and if $\beta=\langle\beta^1,\dots,\beta^n\rangle$ is the vector of strategies used by each player. A Bayesian equilibrium of g in α is a strategy n - tuple $\beta^*=\langle\beta^{*1},\dots,\beta^{*n}\rangle$ such that for each player i, β^{*i} is a function from T^i to M^i satisfying

$$\sum_{t_{-i}\in T_{-i}} \sum_{d\in D} u^i(d,t)p^i(t_{-i},t_i)\mu(d,\beta^*(t))$$

$$= \max_{m^i\in M^i} \sum_{t_{-i}\in T_{-i}} \sum_{d\in D} u^i(d,t)p^i(t_{-i},t_i)\mu(d,\beta^*(t)/m^i)$$

where

$$\langle\beta^*(t)/m^i\rangle=\langle\beta^{*1}(t_1),\dots,\beta^{*(i-1)}(t_{i-1}),m^i,\beta^{*(i+1)}(t_{i+1}),\dots,\beta^{*n}(t_n)\rangle$$

A key-result in Bayesian social choice analysis is the revelation principle. We call any mechanism $\langle M,g\rangle$ a <u>direct-revelation mechanism</u> if $M^i=T^i$ for all agents i. That is, each agent announces a, possibly false, type which is used by $g(t^1,\dots,t^n)$ to pick $\mu\in\Delta(D)$. Of particular interest are direct revelation mechanisms for which truth is a reasonably strategy. We call a mechanism $\langle M,g\rangle$ an <u>incentive-compatible direct-revelation mechanism</u> for (an icdr) if an only if $M^i=T^i$ for all agents and $\beta^{*i}(t_i)=t_i$, for all $t_i\in T^i$ and all i, is a Bayes equilibrium for g in α.

<u>Theorem 1</u> : (Revelation principle, Gibbard (2)). Given $\langle M,g\rangle$, α and $\beta^i : T^i\to M^i$ for each i, such that $\beta^i(T^i) = M^i$:

A. β is a Bayes equilibrium for $\langle M,g\rangle$ in α if and only if

Remark : The condition that $\beta^i(T^i) = M^i$ is needed to show that $B \rightarrow A$. The statement that $A \rightarrow B$ is valid even if $\beta^i(T^i) \subset M^i$.

In this paper we show that the revelation principle is valid in the context of a possibly smaller (sub) class of Bayesian social choice problems which define an arbitration game.

In Lahiri (3) we introduce the concept of an arbitration game. A social value function is a function

$W : D \times T \rightarrow R$

which for each realization $t \in T$ of the true characteristics of the players and for each public decision $d \in D$, gives the value accruing to society.

Let $m \in M$, be the message communicated by the players to an arbitrator. Based on the message $m \in M$, the arbitrator (in the fashion of a true Bayesian) forms posteior beliefs about $t \in T$, which is assumed to be summarized by a conditional probability mass function $f(t|m)$ available to the arbitrator. In this framework the arbitrator solves the following problem :

$$\max_{d \in D} \sum_{t \in T} W(d,t) f(t|m).$$

In response to $m \in M$ communicated by the players, the arbitrator announces $g(m) \in D$ as the solution to the above problem. The mechanism $\langle M, g \rangle$ is now represented by a Bayesian statistician $\langle W, f \rangle$ and we call the triplet $\langle W, f, \alpha \rangle$ an arbitration game. The definition of a Bayes equilibrium for $\langle W, f, \alpha \rangle$ is the same as the definition of a Bayes equilibrium for $\langle M, g \rangle$ in α where $\langle M, g \rangle$ is itself defined by $\langle W, f \rangle$. Further, $\langle W, f \rangle$ (a Bayesian statistician) will be called an icdr in α if the associated mechanism $\langle T, \delta \rangle$ is an icdr in α.

3. Main Theorems : The main theorem we establish in this section is that the revelation principle is valid for arbitration games. The following elementary algebra will be found useful in the sequal :

$$\Pr.\ (t|t')\ \Pr.\ (t') = \Pr.\ (t'|t)\ \Pr.\ (t)$$

$$\therefore \Pr.\ (t|t') = \frac{\Pr.\ (t'|t)\ \Pr.\ (t)}{\Pr.\ (t')}$$

where t is the type of the players and t' is the announced type in a revelation game.

Now $\Pr.\ (t'|t) = \Pr.\ (\beta(t')|t)$

and $\Pr.\ (t') = \Pr.\ (\beta(t'))$, by the change of variable theorem with the notations having the obvious interpretations.

$$\therefore \Pr.\ (t|t') = \frac{\Pr.\ (\beta(t')|t)\ \Pr.\ (t)}{\Pr.\ (t')}$$

$$= \frac{\Pr.\ (\beta(t')|t)\ \Pr.\ (t)}{\sum_{t\in T} \Pr.(\beta(t')|t)\ \Pr.(t)} = \Pr.\ (t|\beta(t'))$$

Thus if f (t|m) is the conditional probability of the true type being 't',if the announced message is 'm', and if β be the strategy for the players, then the conditional probability h(t|t') of the true type being 't' when the announced type is t' defined by

$$h\ (t|t') = f(t|\beta(t'))$$

gives rise to a decision rule $g(\beta(t))\ \forall\ t$ for the arbitrator assuming that $\langle M,g\rangle$ is the mechanism associated with $\langle W,f\rangle$. We thus have the following result :

Theorem 2 : Given $\langle W,f,\alpha\rangle$, and $\beta^i : T^i \to M^i$ for each i, such that $\beta^i(T^i) = M^i$:

A. β is a Bayes equilibrium for $\langle W,f,\alpha\rangle$ if and only if

B. $\langle W,h\rangle$ where $h(t|t') = f(t|\beta(t'))\ \forall\ t,t' \in T$, is an icdr in α.

<u>Proof</u> : β is a Bayes equilibrium for $\langle W,f,\alpha\rangle$

$$\Rightarrow \sum_{t_{-i}\in T_{-i}} u^i(g(\beta(t)),t)p^i(t_{-i},t_i) \geq \sum_{t_{-i}\in T_{-i}} u^i(g(\beta(t)/m^i),t)p^i(t_{-i},t_i)$$

$\forall\ m^i \in M^i$ and $\forall\ t_i \in T^i$

where $g(\beta(t)) = \max\limits_{d\in D} \sum\limits_{t'\in T} W(d,t')f(t'|\beta(t))$

$$\Rightarrow \sum_{t_{-i}\in T_{-i}} u^i(g(\beta(t)),t)p^i(t_{-i},t_i) \geq \sum_{t_{-i}\in T_{-i}} u^i(g(\beta(t/t_i')),t)p^i(t_{-i},t$$

$\forall\ t_i' \in T^i$ and $\forall\ t_i \in T^i$

where $g(\beta(t)) = \max\limits_{d\in D} \sum\limits_{t'\in T} W(d,t')h(t'/t)$

i.e. $\langle W,h\rangle$ is an i.c.d.r. in α.

Conversely suppose that $\langle W,h\rangle$ is an icdr in α.

$$\therefore \sum_{t_{-i}\in T^{-i}} u^i(g'(t),t)p^i(t_{-i},t_i) \geq \sum_{t_{-i}\in T_{-i}} u^i(g'(t/t_i'),t)p^i(t_{-i},t_i)$$

$\forall\ t_i', t_i \in T_i$,

where $g'(t) = \max\limits_{d\in D} \sum\limits_{t\in T} W(d,t')h(t'/t)$

$= \max\limits_{d\in D} \sum\limits_{t\in T} W(d,t')f(t'|\beta(t)) = g(\beta(t))$

$$\therefore \sum_{t_{-i}\in T^{-i}} u^i(g(\beta(t)),t)p^i(t_{-i},t_i) \geq \sum_{t_{-i}\in T_{-i}} u^i(g(\beta(t)|m^i),t)p^i(t_{-i},t_i)$$

$\forall\ t_i \in T^i$ and for all $m^i \in M^i$,

since $\beta(T^i) = M^i$.

Hence β is a Bayes equilibrium for $\langle W,f,\alpha\rangle$.

Q.E.D.

The above result has been restricted to mechanisms which pick out degenerate probability distributions on the space of actions. Such mechanisms are called certainty mechanisms and the underlying rationale for such a drastic oversimplication is that for a finite action space a randomized decision rule is always welfare equivalent to a decision rule which picks out degenerate probability distributions (see Ferguson (1) pp. 34-38). Hence in the context of such problems the revelation principle continues to be valid.

Our next observation is that any mechanism can be rationalized.

<u>Proposition 1</u> : Given $\langle M,g\rangle,\alpha$, with $|M^i|\leq|T^i|$ for all i, there exists $\langle W,f\rangle$ such that $\langle W,f,\alpha\rangle$ is the associated arbitration game.

<u>Proof</u> : Let β be any strategy for $\langle M,g\rangle$ in α such that $\beta^i(T^i) = M^i$ for all i, and let $\langle T,\sigma\rangle$ be the direct revelation mechanism, where $\sigma(t) = g(\beta(t)), \forall t$.

Define, $h(t|t') = 1$ if $t = t'$

$= 0$ if $t \neq t'$

for all $t' \in T$ and let $W : D\times T \to \mathbb{R}$ be a function such that

$W(\sigma(t),t) > W(d,t)$ for all $d \in D$, $d \neq \sigma(t)$

$$\therefore \max_{d\in D} \sum_{t\in T} W(d,t)h(t|t') = W(\sigma(t'),t') \; \forall t' \in T$$

Hence $\langle W,h,\alpha\rangle$ is the arbitration game associated to $[\langle T,\sigma\rangle,\alpha]$.

Defining $f(t|m) = h(t|t')$ where $m = \beta(t')$, completes the proof.

Q.E.D.

The hypothesis that $|M^i|\leq|T^i|$ was required in order to establish the existence of a strategy β, with $\beta^i(T^i) = M^i$ for all i. This latter condition was in turn required to define $f(t|m)$.

4. Conclusion : In this paper we have established that the revelation principle is valid for arbitration games and we have also established in Proposition 1 that a large class of Bayesian Social Choice problems are representable as arbitration games. Hence the concept of an arbitration game derives additional force and appeal. In effect what we have shown is that it is perfectly reasonable to assume that the mechanism which implements a Bayesian equilibrium is under the control of a Bayesian statistician. Thus for instance various auction and resource allocation mechanisms can now be viewed from this stand point.

Reference :

1. T.S. Ferguson : "Mathematical Statistics : A Decision Theoretic Approach", Academic Press, n New York and London, 1973.

2. A. Gibbard : "Manipulation of Voting Schemes : Some General Results", Econometrica 41, 587 - 602 (1973).

3. S. Lahiri : "Arbitration By A Bayesian Statistician And Bounded Rationality", Indian Institute of Management, Ahmedabad, Working Paper No. 849, January, 1990.

4. J.O. Ledyard : "The Scope of the Hypothesis of Bayesian Equilibrium", Journal of Economic Theory 39, 59 - 82 (1986).

5. R. Myerson : "Bayesian equilibrium and incentive Compatibility : An introduction", Discussion Paper No. 548, Center for Mathematical Studies in Economics and Management Science, Northwestern University, Evanston, Ill., February, 1983.