Robust Mechanisms for Information Elicitation

Aviv Zohar Jeffrey S. Rosenschein
School of Engineering and Computer Science
The Hebrew University of Jerusalem
Jerusalem, Israel
{avivz, jeff}@cs.huji.ac.il

ABSTRACT

We study information elicitation mechanisms in which a principal agent attempts to elicit the private information of other agents using a carefully selected payment scheme based on proper scoring rules. Scoring rules, like many other mechanisms set in a probabilistic environment, assume that all participating agents share some common belief about the underlying probability of events. In real-life situations however, underlying distributions are not known precisely, and small differences in beliefs about these distributions may alter agent behavior under the prescribed mechanism.

We propose designing elicitation mechanisms in a manner that will be robust to small changes in belief. We show how to algorithmically design such mechanisms in polynomial time using tools of stochastic programming and convex programming, and discuss implementation issues for multiagent scenarios.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed AI—*Intelligent Agents, Multiagent Systems*;
H.1.1 [**Models and Principles**]: Systems and Information Theory—*Value of Information*

General Terms

Algorithms, Economics

Keywords

Information Elicitation, Computational Mechanism Design, Robust Mechanisms, Stochastic Programming, Scoring Rules

1. INTRODUCTION

In this paper we examine a scenario in which a principal agent is interested in purchasing information about some event from some other agent (or group of agents) that has private access to that information. The sellers are required to invest some effort in order to learn the information, and may be tempted to guess or report falsely if they expect to benefit from doing so. The buyer of information must therefore design its payments in a way that will induce truthfulness on the part of the sellers. This is ordinarily done using *Proper Scoring Rules* [2]. With a well-designed payment scheme, the expected utility of the sellers is maximized only when they invest the effort to learn the information and reveal it truthfully.

To construct such a mechanism, the designer must take into account the *beliefs* of the sellers about the probabilities of events, since these affect the cost-benefit analysis the sellers make. Unfortunately, these probabilities might not be common knowledge, and may in fact be secret information the agents do not wish to reveal.

We propose to design information elicitation mechanisms to be robust not only against manipulation by the participants, but also against small variations in the beliefs they may hold. The classic approach to dealing with variations in beliefs (or "type") of agents within mechanism design is the use of direct revelation mechanisms. These are mechanisms in which the participants reveal everything to the mechanism, which in turn acts optimally on their behalf — eliminating the need to lie. This approach is not appropriate in scenarios involving information elicitation where information is considered a commodity that is to be sold and not revealed freely.

2. THE SCENARIO

We model the information of agents using discrete random variables. Each seller i is assumed to own a private variable X_i that it can access at a cost of c_i. These variables are not necessarily independent. Once the transaction is complete, the buyer is given access to a random variable denoted Ω. The variable Ω is assumed to be somewhat coupled with the variables X_i, and provides a limited means of verification about their true values. Alternatively, one may think of Ω as some outcome that the buyer is attempting to predict, which eventually becomes known. We denote the governing probability distribution $Pr(\Omega = \omega, X_1 = x_1, \ldots, X_n = x_n)$ by $p_{\omega,x_1 \ldots x_n}$. Payments to the agents are made after the value of Ω is revealed and may thus depend on the outcome, as well as the reports of all the agents (it is impossible to create the incentives for truthfulness without some measure of the correctness of the information provided). We denote the payment to agent i by $u^i_{\omega,x_1,\ldots,x_n}$. When dealing with only one agent, we shall drop the script i from all notations.

We require three things from a proper payment scheme. These are presented below in the case of a single agent:

AAMAS'06 May 8–12 2006, Hakodate, Hokkaido, Japan.

1. **Truth Telling.** Once an agent knows its variable is x, it must have an incentive to reveal it, rather than any lie x'.

$$\forall x, x' \quad s.t. \quad x \neq x' \quad \sum_{\omega} p_{\omega,x} \cdot (u_{\omega,x} - u_{\omega,x'}) > 0 \quad (1)$$

Here $p_{\omega,x}$ is the probability of what actually occurs, while the payment $u_{\omega,x'}$ is based only on what *the agent* reported.

2. **Individual Rationality.** An agent must have a positive expected utility from participating in the game:

$$\sum_{\omega,x} p_{\omega,x} \cdot u_{\omega,x} > c \quad (2)$$

3. **Investment.** The *value of information* for the agent must be greater than its cost. Any guess the agent makes without actually computing its value must be less profitable (in expectation) than revealing the variable:

$$\forall x' \quad \sum_{\omega,x} p_{\omega,x} \cdot (u_{\omega,x} - u_{\omega,x'}) > c \quad (3)$$

When a mechanism is designed for multiple agents, similar requirements should be considered. Their exact nature depends on the level of cooperation possible among the selling agents (transfer of utility, shared information, etc.). They can still be described in the form of linear constraints, but the number of constraints can sometimes be exponentially large in the number of agents.

2.1 Building Non-Robust Mechanisms

The three requirements above can all be characterized using linear constraints and can thus be solved efficiently using linear programming methods. Furthermore, a solution can be found that minimizes some target function such as the expected cost of the mechanism to the buyer.

A great deal of insight into the design problem can be obtained when considering the vectors defined by $\vec{p}_x \triangleq (p_{\omega_1,x} \dots p_{\omega_k,x})$ and $\vec{u}_x \triangleq (u_{\omega_1,x} \dots u_{\omega_k,x})$. Using this notation, the truthfulness constraints can be viewed as the requirement that the probability vectors $\vec{p}_x, \vec{p}_{x'}$ will be linearly separated by the vector of payments $\vec{u}_x - \vec{u}_{x'}$.

The following proposition shows necessary and sufficient conditions for the existence of a proper payment scheme.

PROPOSITION 1. *In the single agent case, a proper payment scheme exists iff the probability vectors $\vec{p}_x$ are pairwise independent. Furthermore, if any proper payment scheme exists then there is one with a mean cost as close to c as desired. Such a scheme is optimal, due to the individual rationality constraint.*

PROOF SKETCH. When two vectors $\vec{p}_x$, $\vec{p}_{x'}$ are linearly dependent, there is no way to satisfy the two truthfulness constraints for x, x'. If they are independent, then the truthfulness constraints can be satisfied (for example) by setting for any $\alpha > 0$ and any $\vec{\beta}$, $\vec{u}_x = \frac{\vec{p}_x}{||\vec{p}_x||} \cdot \alpha + \vec{\beta}$. α and $\vec{\beta}$ can be adjusted independently to satisfy the individual rationality and investment constraints tightly. Notice that this implies that the truthfulness constraints are the only things that might prevent us from finding a proper payment scheme. See [3] for a more detailed proof. $\square$

3. DESIGNING ROBUST MECHANISMS

We shall now turn our attention to the case where the probabilities associated with the random variables Ω, X_i are not common knowledge. From now on, $\hat{p}_{\omega,x}$ shall denote the probabilities the principal agent believes in, and $p_{\omega,x}$ will denote the beliefs of an agent. Since beliefs about probabilities are usually grounded by observations, they are not likely to be very far from the truth. We shall assume that different beliefs are "close" to one another according to some distance metric: $\hat{p}_{\omega,x} = p_{\omega,x} + \epsilon_{\omega,x}$ and $||\vec{\epsilon}|| < \epsilon$. We have opted for the use of the L_∞ norm for measuring distance, since it is easily described using linear constraints, but the methods we present can be easily adapted for any other norm.

3.1 The Robustness of a Payment Scheme

Definition 1. We shall say that a given *payment scheme* $u_{\omega,x}$ is *ϵ-robust* with regard to an elicitation problem with distribution $\hat{p}_{\omega,x}$ if it is a proper solution to every elicitation problem with distribution $\hat{p}_{\omega,x} + \epsilon_{\omega,x}$ such that $\|\vec{\epsilon}\|_\infty < \epsilon$, and is infeasible for at least one problem instance of any larger norm.

The definition above is conservative. An ϵ-robust payment scheme will create the proper incentives for truthfulness for *every* possible belief variation of the agent that is within a distance of ϵ. This represents a "large margins" point of view, where the solution is required to satisfy all constraints with enough redundancy to handle small changes in them.

3.1.1 Determining the Robustness of a Given Scheme

Given a payment scheme $u_{\omega,x}$, and an elicitation problem with probabilities $p_{\omega,x}$, we can determine the robustness level of $u_{\omega,x}$ by finding out how much the probabilities must be perturbed to violate one of the constraints required for a truthful mechanism. We can do this by solving a linear program for every constraint. For example, the following program finds a perturbation that violates the truth-telling constraint for a secret x and a lie x'.

$$\begin{array}{ll}
\min & \epsilon \\
\text{s.t.} & \\
& \sum_{\omega} (\hat{p}_{\omega,x} + \epsilon_{\omega,x})(u_{\omega,x} - u_{\omega,x'}) \le 0 \\
\forall x, \omega & \hat{p}_{\omega,x} + \epsilon_{\omega,x} \ge 0 \\
& \sum_{\omega,x} \epsilon_{\omega,x} = 0 \\
\forall x, \omega & -\epsilon \le \epsilon_{\omega,x} \le \epsilon
\end{array}$$

A solution to this program will be a small perturbation $\epsilon_{\omega,x}$ with a small norm ϵ that manages to violate the constraint. The requirements on $\epsilon_{\omega,x}$ assure us that the perturbed values are still legal probabilities. Once we solve a linear program for every constraint, we simply take the minimal value of ϵ that was found. This is the robustness level of the mechanism.

3.1.2 Efficiently Finding Some ϵ-Robust Mechanism

From a design point of view, we may be interested in finding a solution that is at least ϵ-robust for some ϵ and has a minimal cost. We can do this using tools for *Stochastic Programming*. A stochastic program is simply a mathematical program that contains uncertainty about the exact constraints that need to be satisfied, or the function that is optimized. The following stochastic program describes our problem:

$$
\begin{array}{rl}
\min & \sum_{\omega,x} \hat{p}_{\omega,x} \cdot u_{\omega,x} \\
\text{s.t.} & \\
\forall x \neq x' & \sum_{\omega} p_{\omega,x}(u_{\omega,x} - u_{\omega,x'}) > 0 \\
 & \sum_{\omega,x} p_{\omega,x} \cdot u_{\omega,x} > c \\
\forall x' & \sum_{\omega,x} p_{\omega,x}(u_{\omega,x} - u_{\omega,x'}) > c \\
\text{where:} & \\
\forall x, \omega & p_{\omega,x} = \hat{p}_{\omega,x} + \epsilon_{\omega,x} \\
 & p_{\omega,x} \geq 0 \\
 & \sum_{\omega,x} p_{\omega,x} = 1 \\
 & -\epsilon \leq \epsilon_{\omega,x} \leq \epsilon
\end{array}
$$

Stochastic programs with the conservative solution we require can be solved efficiently. [1] presents methods for doing so by treating the problem as a regular convex programming problem. Naturally, there can be cases where no solution exists for the given robustness level.

3.2 The Robustness Level of a Problem

Definition 2. We define the *robustness level ϵ^* of the problem* $\hat{p}$ as the supremum of all mechanism robustness levels ϵ for which there exists a proper mechanism:

$$\epsilon^* \triangleq \sup_{\vec{u}}\{\epsilon | \exists \vec{u} \text{ that is an } \epsilon\text{-robust mechanism for } \hat{p}\} \quad (4)$$

To find the robustness level of a problem, one can simply perform a binary search. The robustness level is certainly somewhere between 0 and 1. One may test at every desired level in between to see if there exists a mechanism with some specified robustness by solving the stochastic program above. The space between the upper and lower bounds is then narrowed according to the answer that was received.

3.2.1 A Bound for Problem-Robustness

A simple bound for robustness of the problem can be derived from examining the truthfulness conditions. In fact, Proposition 1 for non-robust mechanisms can be viewed as a private case of the following proposition when applied to 0-robust mechanisms:

PROPOSITION 2. *The robustness level ϵ^* of a problem $\hat{p}$ can be bounded by the smallest distance between a vector $\hat{p}_x$ and the optimal hyper-plane that separates it from $\hat{p}_{x'}$:*

$$\epsilon^* \leq \min_{x,x'} ||\hat{p}_x - (\hat{p}_x^{tr} \cdot \vec{\varphi}_{x,x'}) \cdot \vec{\varphi}_{x,x'}||_\infty \quad ; \quad \vec{\varphi}_{x,x'} = \frac{\hat{p}_x + \hat{p}_{x'}}{||\hat{p}_x + \hat{p}_{x'}||_2}$$

The optimal separating hyper-plane is a hyper-plane that separates the points and is of maximal (and equal) distance from both of them.

PROOF SKETCH. If there exist x, x' that give distance ϵ to the hyperplane then the vectors $\hat{p}_x, \hat{p}_{x'}$ can be perturbed towards the hyperplane with a perturbation of norm ϵ, until they are linearly dependent. For this problem instance, according to Proposition 1, there is no possible mechanism. □

In the case where $|\Omega| = 2$, the vectors $\hat{p}_x$ are situated in a two-dimensional plane, and it can be shown that the bound given above is tight — the problem robustness is determined exactly by the closest pair of vectors.

3.3 Mechanisms for Multiple Agents

We assume here that sellers cannot collude, share information, or transfer utility among themselves. Models that allow for collusion are interesting to explore from the viewpoint of robustness, but are beyond the scope of this paper.

When designing mechanisms for multiple agents, the designer has to choose the solution concept it wishes to achieve. It is sometimes possible to design mechanisms that will provide the agents with dominant strategies of truth telling — each agent is better off telling the truth no matter what the other agents do. However, in some cases this is not possible, and the designer must implement a mechanism that provides for good behavior only in equilibrium, when *all* players are behaving well. In these cases the payments made to the players utilize the information given by their peers. Each solution concept translates into a different linear program that may provide a different answer.

In this case, the design of a robust mechanism becomes even more difficult. The designer must now take into account not only the possible beliefs of agents about the probabilities of events, but also their beliefs about the beliefs of other agents. For an agent to believe that some strategy is in equilibrium, it must also be convinced that its counterparts believe that their strategies are in equilibrium, or are otherwise optimal. This will only occur if the agent believes that they believe that it believes that its strategy is in equilibrium — and so on to infinity. Any uncertainty about the beliefs of other agents grows with every step up the belief hierarchy. If agent A knows that all agents have some radius ϵ of uncertainty in beliefs, and its view of the world consists of some probability distribution p it assigns to events, then it is possible that agent B believes the distribution is p' and further believes that agent A believes the distribution is some p'' which is at a distance of up to 2ϵ from p. With an infinite belief hierarchy, it is therefore possible to reach any probability if we go high enough in the hierarchy.

A possible solution to this problem is to use a mixture of solution concepts. The mechanism can often be designed to make each agent's payment depend only on a subset of agents that precedes it. In this case it only needs to take their beliefs into consideration when deciding on a strategy. The necessary belief hierarchy is then finite, which limits the possible range of beliefs about beliefs. The most extreme case of this is to design the mechanism for dominant strategies only. Naturally, a solution constructed in such a way may be less efficient or may not exist at all. Another possibility is to consider only bounded agents that can only consider a finite number of levels in the belief hierarchy.

4. ACKNOWLEDGMENT

This work was partially supported by grant #039-7582 from the Israel Science Foundation.

5. REFERENCES

[1] A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Operations Research Letters*, 25:1–13, 1999.

[2] T. Gneiting and A. E. Raftery. Strictly proper scoring rules, prediction, and estimation. Technical Report 463, Department of Statistics, Univ. of Washington, 2004.

[3] A. Zohar and J. S. Rosenschein. Robust information elicitation mechanisms. Technical Report 2006-3, Leibniz Institute, Hebrew Univ. of Jerusalem, 2006.