LARGE SCALE SYSTEMS CONTROL

Optimal Incentive-compatible Mechanisms
in Active Systems

A. K. Enaleev

Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
Received March 24, 2011

Abstract—We consider mechanisms in two-level active systems, where additional conditions of incentive compatibility are imposed on planning procedures and incentive schemes to coordinate preferences of agents and a principal. These conditions ensure plan fulfillment and truth-telling (strategy-proofness). Finally, sufficient conditions of optimal incentive-compatible mechanisms are established.

DOI: 10.1134/S0005117913030120

1. INTRODUCTION

The idea of interests coordination among active elements of hierarchical organizational systems is actually long-standing. Not claiming to cite the earliest work in this field, we emphasize that the idea was discussed in detail as far back as in [10]. In the theory of active systems, interests coordination was given a thorough treatment in [1]. Notably, the author of this book suggested the perfect concordance principle in planning problems and the corresponding incentive-compatible mechanisms (fair play mechanisms) ensuring truth-telling by agents. First research results for optimal incentive-compatible mechanisms in active systems were published in [3, 5] (the case of complete awareness of a principal about models of agents). In the case of incomplete principal’s awareness, optimality of the principle of incentive compatibility was proved in [2] for a given incentive scheme of an agent. Moreover, the paper [4] studied the problem of optimal mechanism design under “strong penalties” for existing deviations of agent’s state from a plan assigned. Finally, the paper [8] showed optimality of incentive-compatible mechanisms in the case when the penalty for such deviations is bounded by a fixed quantity.

In the present paper, we generalize the results derived in [3, 5, 8], as well as provide sufficient conditions of optimal incentive-compatible mechanisms in active systems (both in the case of complete and incomplete awareness of a principal about agents). The incentive-compatibility conditions of mechanisms are formulated by introducing auxiliary constraints on principal’s choice of planning procedures. These conditions guarantee plan fulfillment and truth-telling (revelation of correct information) by agents. The problems of optimal mechanism design are solved for a series of important cases meeting the incentive-compatibility condition.

2. THE MODEL AND PROBLEM STATEMENT

The model of an active system under consideration is described in [1, 8]. Therefore, let us be confined with a brief exposition here. Suppose that an active system comprises a principal and an agent.

Denote by $\Phi(x, y, r)$ the principal’s goal function; assume that $\Phi(y, y, r) \geq \Phi(x, y, r) \geq 0$, where $\Phi(y, y, r)$ appears a continuous function being strictly quasi-concave with respect to $y$ for all $r \in A$. In addition, $x$ designates a plan assigned by the principal, $x \in X$—a set of feasible plans; $y$ is
a state chosen by the agent, \( y \in Y \)—a set of feasible states; \( r \) means a parameter characterizing the system, \( r \in A \)—a set of feasible parameter values. In the sequel, adopt \( X = Y \) for simplicity. Suppose that the sets \( X, Y \) (as well as \( A \)) are bounded and closed. Let \( f(x, y, r) \) be the agent’s goal function. We believe that the plan \( x \) is assigned by the principal according to a certain planning procedure \( x = \pi(\cdot) \), where \( \pi(\cdot) \) maps the set \( A \) into the set \( X \). The principal may assign the agent’s goal function within definite constraints: \( f(x, y, r) \in F \). Here \( F \) specifies a given set of feasible goal functions (incentive schemes).

The combination of the planning procedure \( x = \pi(\cdot) \) and the agent’s goal function \( f(\cdot, \cdot, r) \) forms the mechanism \( \mu = \{ \pi(\cdot), f(\cdot, \cdot, r) \} \). In the theory of active systems (e.g., see [1, 2, 3]), a planning procedure is often called a control law.

To proceed, introduce some assumptions regarding awareness in this active system. The agent knows the parameter \( r \), whereas the principal is aware of the feasible set \( A \) merely. Moreover, under a fixed mechanism \( \mu \), the agent informs the principal of an estimate \( \rho \) of the parameter \( r \), \( \rho \in A \).

Suppose that the mechanism \( \mu \) is given. Consequently, the active system under consideration operates in the following way. The agent reports an estimate \( \rho \) of the parameter \( r \), and the planning procedure \( \pi(\cdot) \) serves for assigning the plan \( x = \pi(\rho) \). Next, the agent chooses his/her state \( y \) by maximizing the goal function \( f(x, y, r) \in y \).

Denote by \( \varphi(x, r) = \max_{y \in Y} f(x, y, r) \) the agent’s preference function. The principal’s preference function is defined by

\[
\Psi(x, r) = \min_{y \in \mathcal{Z}(x, r)} \Phi(x, y, r),
\]

where \( \mathcal{Z}(x, r) \) indicates the set of agent’s rational strategies under the choice of the state \( y \). A rigorous definition of the set of rational strategies \( \mathcal{Z}(x, r) \) could be found below.

Throughout the paper, for simplicity we assume the all necessary operations of maximization and minimization are well-defined.

For the mechanism \( \mu \), consider the efficiency criterion

\[
K(\mu) = \min_{r \in A} \left[ \min_{\rho \in R(r)} \frac{\Psi(\pi(\rho), r)}{\Psi_b(r)} \right],
\]

where \( R(r) \) is the set of agent’s rational strategies under his/her choice of the message \( \rho \) (see the definition of the set \( R(r) \) below), \( \Psi_b(r) \) is a given normalizing function. For instance, the normalizing function can be

\[
\Psi_b(r) = \max_{x \in X} \Psi(x, r), \quad \text{or} \quad \Psi_b(r) = \max_{x \in X} \Phi(x, x, r), \quad \text{or} \quad \Psi_b(r) = \text{const} > 0.
\]

In what follows, we apply the “weak condition of agent’s benevolence” [8]. Accordingly, the set of agent’s rational strategies takes the form

\[
\mathcal{Z}(x, r) = \begin{cases} \{x\}, & \text{if } x \in \text{Argmax}_{y \in Y} f(x, y, r) \\ \text{Argmax}_{y \in Y} f(x, y, r), & \text{otherwise,} \end{cases}
\]

\[
R(r) = \begin{cases} \{r\}, & \text{if } r \in \text{Argmax}_{\rho \in A} \varphi(\pi(\rho), r) \\ \text{Argmax}_{\rho \in A} \varphi(\pi(\rho), r), & \text{otherwise.} \end{cases}
\]