Revelation principle of direct mechanism : A Primer

Jin-ock Kim*

CONTENTS

Ⅰ. Introduction

Students who understand a classical or traditional microeconomics are not acquainted well with the economic decision problems in the presence of private information held among economic agents. The purpose of this paper is to introduce some concept of mechanism design, especially direct mechanism of revelation principle to the students or some economists who are not well acquainted with this concept. The private information held by the agents is a valuable resource that the agents have even though it is not physical. The economic agent utilizes his valuable own information to maximize his own utility in any economic situations. In these circumstances the principal or the mechanism designer proposes some mechanism or game form in that the agent sends the messages which depend on his private hidden information to the principal. Here we consider some Bayesian game forms with communication channel between the principal and the agents in economic transactions. Given the indirect mechanism or the Bayesian game form, Bayesian Nash Equilibrium (BNE) exists. If we restrict the range of message profile

* Professor, Department of Economics, College of Economics and Commerce, Jeju National University

of BNE to the type spaces of agents, then we get the direct mechanism. In thedirect mechanism the profile of agents'types of true revelation is a BNE.

In section 1, we introduce how to construct the direct mechanism of revelation principle form the indirect mechanism or Bayesian game form with communication. In section 2, several examples of mechanism design problems in the real world are shown. In section 3 we applied revelation principle to the various economic environments including price discrimination, bilateral trading problems, auctions, and the moral hazard. Concluding remarks follow.

II. Some mechanism design problems

There are many examples of mechanism design problems in a real world. In this section we show some examples among those. The first example is, suppose that you are going out somewhere with your friends. The available options are to go bars, clubs, nice restaurants, and theatres. Given the different preferences over the options, it is very difficult to elicit overall agreement among friends. The more difficulty in drawing consensus among friends arises when each individual's preference over available options is not known publicly. There are many mechanisms to solve this problem. One possible mechanism is to select one option in the following way: Firstly you would set two options, say bars and nice restaurants, among which people could select one option. And then you would set two options which consist of one chosen option in the first stage, say, nice restaurants, and the new one, say clubs, and so on. The first problem in this mechanism is who should be the mechanism selector. That is, some bargaining problems in selecting mechanism designer among the players arise. The second problem is that each individual may utilize his private preferences opportunistically in each stage of selection procedures to maximize his utility.

The second example is, consider a situation where n agents must decide whether to build a public good of bridge whose costs should be borne by the agents themselves. The decision problem is easy to take when we knew the people's preferences: The economic efficiency condition for this decision problem is that the total benefits for n agents exceed the total costs borne by the agents. The problems in designing an efficient mechanism occur when preference of each agent is not known publicly. The first thing of coming to mind is to ask people what their preferences are. But as rational utility maximizing citizens, they will report their private information of their individual preferences truthfully when doing so maximizes their utilities. The task the social planner or mechanism designer confronts is to elicit truthful revelation of private information held by the agents.

The third example is, consider an environment where the seller wants to sell an indivisible object to one of the many buyers (participants) in the market. The buyer who gets the object must pay the money. Of course the seller wants to maximize the payment she gets from the participants. In terms of the terminology in economics, we say the seller or the principal allocates the objet efficiently in the sense that the object should be allocated to the buyer whose valuation for the object is the highest among the participants. But the problem the principal faces is that he does not know who values highest. That is, the principal does not know the private type or valuation of each of the participating agents in the market. That information is not publicly known. The problem in this case is also to elicit truthful revelations of types held by the participants. That is exactly the direct mechanism design problem.

The last example can be found in voting scheme in the democratic institution: Majority voting vs. something else, direct vs. representative democracy, etc.

Ⅲ. Revelation principle of direct mechanism

Above examples shows that the principal or a mechanism designer designs an indirect mechanism or a Bayesian game form which is played by agents with private information: the principal offers a mechanism which can be accepted or rejected, once accepted then agents play according to a mechanism. For a formal structure of this mechanism, suppose that there are n individuals, numbered 1 to n. Let θ_i denote the set of possible types for individual i and let $\theta=\theta_1 \times \theta_2 \times \ldots\theta_n$ denote the set of possible combinations of types of all individuals. Given the Principal's chosen mechanism or outcome function y that depends on messages strategically sent by agents. Of course, Given the agents' message profile or a Bayesian strategy profile

$\varphi(\theta) = (\varphi_1(\theta_1), \varphi_2(\theta_2), \ldots, \varphi_n(\theta_n))$ where $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$, the mechanism $y(\varphi(\theta))$ is implemented by the principal. Then the principal has a utility, $v(y, \theta)$, and agent i has a utility, $u_i(y, \theta)$.

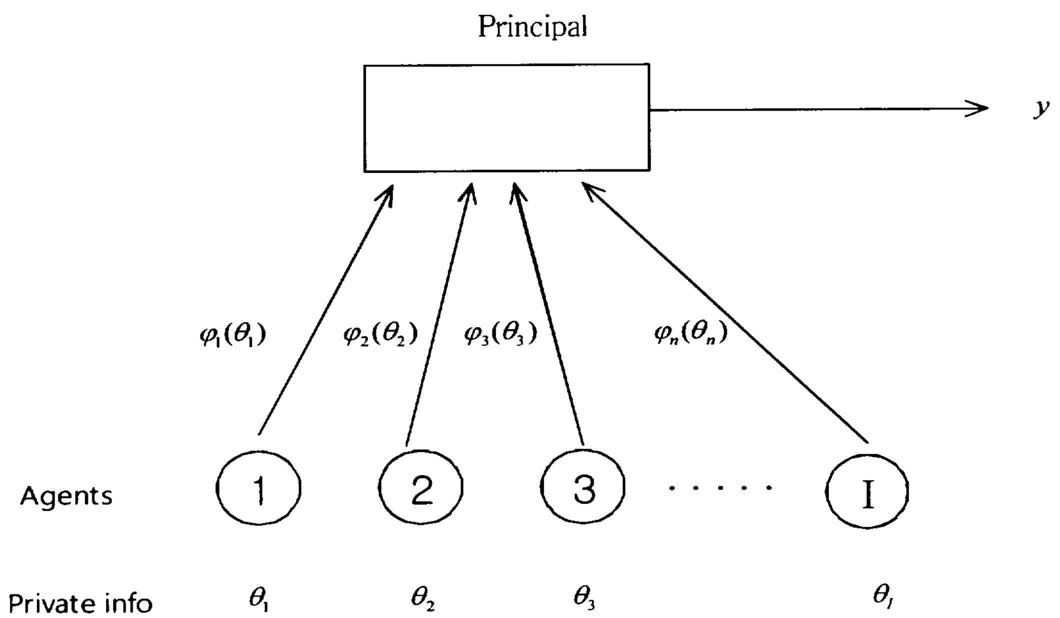
The solution concept of this message game is Bayesian Nash Equilibrium (BNE). The existence theorem of strategic form game implies that there is a BNE in this indirect mechanism. Suppose the BNE strategy profile of this indirect mechanism is $\varphi^*(\theta) = (\varphi_1^*(\theta_1), \ldots \varphi_n^*(\theta_n))$ when type profile $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$ is given. This BNE says that each agent i's conditional expectation of utility is maximized when he uses his Bayesian strategy $\varphi_i^*(\cdot)$ given the other players'Bayesian strategy profile, $\varphi_{-i}^*(\cdot)$. That means that for agent i of type θ_i and every Bayesian strategy φ_i', $E[u_i(y(\varphi^*(\theta)), \theta) \mid \theta_i] \geq E[u_i(y(\varphi_{-i}^*(\cdot), \varphi_i'), \theta) \mid \theta_i]$. Here $E(\cdot \mid \cdot)$ denotes the conditional expectation operator. That is, no agent should have a profitable deviation. We might also impose participation constraint in the mechanism. In this case, the conditional expectation of utility each agent has should excel the agent's outside option.

This indirect mechanism is reduced to the direct mechanism by restricting individual's message space into his original type space. In this direct mechanism each individual is asked to report his type to the mediator or mechanism designer who turns reported types into equilibrium Bayesian strategy profile which results in outcome through outcome function y. Mathematically we define direct mechanism $\mu(\theta) = y(\varphi^*(\theta))$ for every type profile θ which is an element of set of type profiles

θ. Suppose an individual I reports false type θ_i' instead of true type θ_i while the other players report their true types. Then, given true type profile, θ, the direct mechanism $\mu(\theta_{-i},\theta_i') = y(\varphi_{-i}^*(\theta_{-i}),\varphi_i^*(\theta_i'))$ is implemented. The indirect mechanism shows that individual i has no incentive to report his type falsely. That is, the conditional expectation of agent i of any type truthfully revealing his type is greater than when falsely reports his type. This is the well known revelation principle.

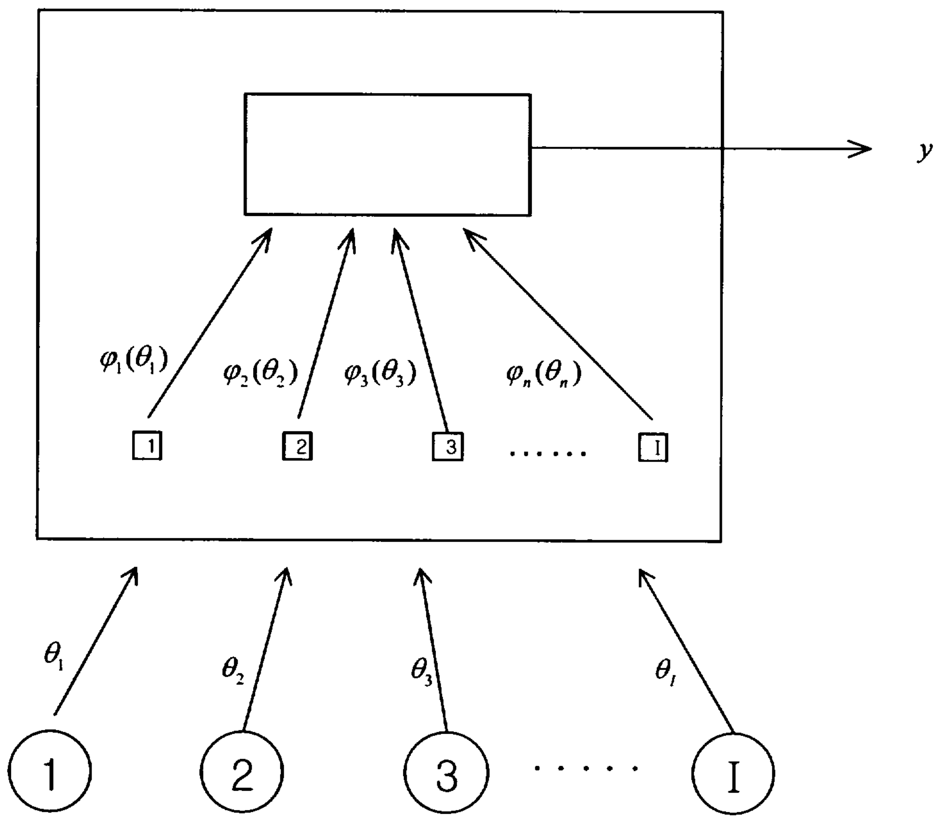
We show an indirect mechanism graphically in <Figure 1>. Each agent in the indirect mechanism sends his message which depends on his private type. The principal commits to a mechanism which transforms these messages into some outcome y. The black box in the figure takes in the messages $\varphi_1(\theta_1)$ through $\varphi_n(\theta_n)$ sent by agents and spits out an outcome y. In this graph, outcome function or indirect mechanism is common knowledge while each agent's own type is private information which is not publicly known.

<Figure 1>An Indirect MECHANISM



A direct mechanism is represented graphically in <Figure 2>. In this graph, Firstly agents report their private types into the bigger box including the black box. Secondly Agents in the bigger box sent the messages into the black box. Lastly the principal in the black box spits out the outcome as a function of messages sent by agents.

<Figure 2>The Corresponding DIRECT REVELATION MECHANISM



Ⅳ. Applications of direct mechanism of revelation principle

In this section we show how the revelation principle is applied in designing a contract for some economic transactions where the players' private information is discrete or continuous random variable in the perspective of uninformed player.

4.1 Two type model of monopoly price discrimination

Suppose the principal (P) or a monopolist produce a quantity q of a good at a constant marginal cost c and no fixed cost. He seeks to have a contract with a buyer (B) whose quasi linear preferences are

$V(q,T,\theta_i)=\theta_i v(q)-t$

Where t is the transfer payment form the buyer to the seller or principal and θ_i is the buyer's type which may represent a parameter affecting the buyer's willingness to pay. The function v(q) is the value function depending on the quantity purchased. We usually assume the first differential of the value function is positive while the second differential of that is negative. This assumption says that interior solutions always applied.

When the seller sells q unit to the buyer and receives a transfer, his profit is:

$\pi = t - cq$

Suppose that there are two types of buyers in the market. Thus, the set of types is $\{\theta_h, \theta_l\}$where $\theta_h > \theta_l$and $\theta_l$occurs with probability λ.

The extensive form of the game is as follows. P designs a contract or mechanism on the basis of taking or leaving it. If A rejects, both sides earn zero. If A accepts, then the contract is executed. If the principal knows buyer's type perfectly, he could extract all of the surpluses in this transaction by designing so called the first best contract. Skipping this case we turn on the case of the principal not knowing buyer's type. That is, there is asymmetric information in the sense that buyer's type is blind to the principal. The revelation principle is applied in dealing with this problem. The seller's objective is to elicit true type form the buyers whom he confront. We specify a direct mechanism or a menu of contracts $(y(\theta) = (t(\theta), q(\theta)))$as a function of buyer's report, θ. Suppose buyer's type is θ_h.

Then the outcome or contract for the high type (θ_h) is, $y(\theta_h) = (t(\theta_h), q(\theta_h))$. The revelation principle says that the high type consumer should prefer the outcome $y(\theta_h)$to the outcome $y(\theta_l)$ designed for the low type buyer. This fact may be represented by the following incentive compatibility constraint for the high type buyer.

$$\theta_h v(q(\theta_h)) - t(\theta_h) \geq \theta_h v(q(\theta_l)) - t(\theta_l) \text{(ICH)}$$

In the same way we represent the incentive compatibility constraint for the low type buyer as follows.

$$\theta_l v(q(\theta_l)) - t(\theta_l) \geq \theta_l v(q(\theta_h)) - t(\theta_h) \text{(ICL)}$$

The participation constraint or individual rationality constraint (IRC)for a buyer is represented depending on the buyer's type as follows.

$$\theta_h v(q(\theta_h)) - t(\theta_h) \geq 0 \qquad \text{(IRH)}$$

$$\theta_l v(q(\theta_l)) - t(\theta_l) \geq 0 \text{(IRL)}$$

Facing these constraints the principal should select the mechanism or optimal contract to maximize his expected utility. That is,

$$\underset{t(\theta), q(\theta)}{\text{argmax}} \; \lambda(t(\theta_l) - cq(\theta_l)) + (1 - \lambda)(t(\theta_h) - cq(\theta_h))$$

Solving this problem, we know that the quantity supplied to the high type is the same as in the first best contract while the low type buyer gets less than the socially optimal contract. On the contrary to the first best case the principal should give the high type so called information rents. We say this is the second best contract in the presence of intrinsic asymmetric information which is blind to the principal. The principal should pay the information rent to the buyer who has the valuable information of his type in this transaction.

4.2 The Bilateral trading problem

There are two agents who are called seller (S) and buyer (B) respectively. The seller has some good to sell, but the cost of producing is not known to the buyer. The value of this object to the buyer is also private information. Seller's cost or value, $V_s$is distributed over a given interval from a_s to b_s.Buyer's valuation, $V_b$is also distributed over a given interval froma_b to b_b. For agent i=b or s, let $f_i(\cdot)$be the probability density function for V_i. We let F_s and F_b be the cumulative distribution function corresponding to f_s and f_b.

We assume that each agent knows his own valuation at the time of bargaining, but considers the other's valuation as a random variable, distributed as above. We also assume that individuals are risk neutral and have additively separable utility for money and the object. These two agents are going to engage in some bargaining game or mechanism to transfer the object from the seller to the buyer. The question we face is to design some efficient mechanism to maximize overall surplus in this transaction. To do that we propose a direct bargaining mechanism in that each agent is asked to report his true type or value to the mediator or coordinator who then determine whether the object is transferred, and how much the buyer should pay. We characterize the direct mechanism by two outcome functions as specified in $y(v) = (p(v),x(v))$ where $v = (v_s, v_b)$. Here p(v) is the probability that the object is transferred to the buyer, and x(v) is the transfer payment to the seller if the v is the reported valuations of the seller and the buyer. Given the reported valuations, v of the agent, the seller's expected gain from trade is represented as follows:

$$u_s(v) = x(v) - v_s p(v)$$

At the beginning of bargaining the seller knows his type of cost, v_s, while he has no information of the buyer's type which he considers as random variable. So the seller's interim expected gain from trade given his true type, v_s is represented as follows:

$u_s(v_s) = \overline{x_s}(v_s) - v_s \overline{p_s}(v_s)$,

Where, $\overline{x_s}(v_s) = \int_{a_b}^{b_b} x(v_s, t)\, f_b(t)$, and $\overline{p_s}(v_s) = \int_{a_b}^{b_b} p(v_s, t) f_b(t)dt$

In the same way we represent buyer's expected gain from trade given the reported valuations, v, as follows:

$u_b(v) = v_b p(v) - x(v)$

The buyer knows his own type v_b at the beginning of bargaining, but he does not know the other agent's type which is considered as random variable to him. So, the buyer's interim expected gain from trade given his type, v_b, is represented as follows:

$u_b(v_b) = v_b \overline{p_b}(v_b) - \overline{x_b}(v_b)$

Where, $\overline{x_b}(v_b) = \int_{a_s}^{b_s} x(t, v_b)\, f_s(t)dt$, and $\overline{p_b}(v_b) = \int_{a_s}^{b_s} p(t, v_b) f_s(t)dt$

The revelation principle says that the direct mechanism or a pair of outcome function (y=(p, x)) is incentive compatible in the Bayesian sense if and only if for every v_s and v_s' in $[a_s, b_s]$,

$u_s(v_s) \geq \overline{x_s}(v_s') - v_s \overline{p_s}(v_s')$,

And for every v_b and v_b' in $[a_b, b_b]$,

$u_b(v_b) \geq v_b \overline{p_b}(v_b') - \overline{x_b}(v_b')$

The direct mechanism is individually rational if and only if

$u_s(v_s) \geq 0$ and $u_b(v) \geq 0$

for every v_s in $[a_s, b_s]$ and for every v_b in $[a_b, b_b]$.

That is, individual rationality is that each individual has non negative expected gains from trade after he knows his own valuation.

Given this direct bargaining mechanism, Myerson and Satterthwaite(1983) show the main theorem as follows:

Theorem: For any incentive compatible mechanism, the sum of interim expected gains from trade for the extremely high type seller and the extremely low type buyer is represented as,

$$u_s(b_s)+u_b(a_b)$$

$$=\int_{a_b}^{b_b}\int_{a_s}^{b_s}([v_b-\frac{1-F_b(v_b)}{f_b(v_b)}]-[v_s+\frac{F_s(v_s)}{f_s(v_s)}])\times p(v_s,v_b)f_s(v_s)f_b(v_b)dv_sdv_b \quad (1)$$

Furthermore, if $p(\cdot,\cdot)$is any function mapped into [0,1], then there exists a function $x(\cdot,\cdot)$ such that (p,x)is incentive compatible and individually rational if and only if $\bar{p}_s(\cdot,\cdot)$ is weakly decreasing and $\overline{p_b}(\cdot,\cdot)$ is weakly increasing, and

$$u_s(b_s)+u_b(a_b)\geq 0.$$

Sketch of proof. Suppose first that we are given an incentive compatible mechanism. Manipulating incentive compatibility for the seller by interchanging the role of true or false type we can show that

$$u_s(v_s)=u_s(b_s)+\int_{v_s}^{b_s}\bar{p}_s(t)dt \quad (2)$$

Similar argument for the buyer shows that

$$u_b(v_b)=u_b(a_b)+\int_{a_b}^{v_b}\overline{p_b}(t)dt \quad (3)$$

Furthermore, we get:

$$\int_{a_b}^{b_b}\int_{a_s}^{b_s}([v_b-v_s])p(v_s,v_b)f_s(v_s)f_b(v_b)dv_sdv_b$$

$$=u_s(b_s)+u_b(a_b)+\int_{a_b}^{b_b}\int_{a_s}^{b_s}(F_s(v_s)f_b(v_b)-(1-F_b(v_b))f_s(v_s))\ p(v_s,v_b)dv_sdv_b.$$

Above equation gives us equation (1) which in turn implies inequality (2). We have proven the first sentence and the "only if"part of the second sentence in the theorem.

To complete the proof of Theorem, suppose that $\overline{p_s}(\cdot,\cdot)$ is weakly decreasing and $\overline{p_b}(\cdot,\cdot)$ is weakly increasing. The next step is to construct the payment function $x(\cdot,\cdot)$ such that the direct mechanism is an individually and incentive compatible. We skip this part. For more details, see Myerson and Satterthwate.

Above type of the direct mechanism design using revelation principle was proposed originally by Myerson(1981). He designed optimal auction market where the seller wants to sell one indivisible objet to the highest bidder among the bidders in the market whose values for the object are not known publicly. In this market there are n bidders and one seller called the principal, whose objective is to maximize his profits. The direct mechanism or a pair of outcome functions (y=p, t) for auctioning the object the bidders in the market is represented as follows:

$y(v) = (p(v,), x(v))$, where $v = (v_1, \ldots, v_n)$, $p(v) = (p_1(v_1), \ldots, p_n(v_n))$, $x(v) = (x_1(v_1), \ldots, x_n(v_n))$.

Here, of course v is a type or valuation profile. For agent i, $p_i(v)$is the probability that agent i

gets the object, and $x_i(v)$is i's expected payment when v is the vector of announcements.

Myerson characterizes the conditions for the direct mechanism (p, x) to be feasible and optimal in the perspective profit maximizing seller. Myerson derived the celebrated revenue equilibrium theorem in which the seller's expected utility from a feasible auction mechanism is completely determined by the probability function p and the numbers of utilities of bidders whose types are the lowest respectively. That is, once we know who get the object in each possible situation as specified by p and how much expected utility each bidder would get if his value estimate were at its lowest possible level, then the seller's expected utility from the auction does not depend on the payment function x.

4.3 Moral hazard problem

In adverse selection problem agent's hidden type does matter in contracting something occurred between the principal who is blind to the agent's type and the agent. Moral hazard is similar to the adverse selection in the sense that it deals with hidden action of the agent instead of agent's hidden type.

Suppose the business man (principal) wants to hire an agent who can help his business. There are two states of the world, a good state(or state 1)and a bad state(or state 2). The income or profit in state 1 is y while that in the state 2 is y−L, which is lower than that in state 2.The income is also affected by the agent's effort level in the way of the probability of state being occurred: if the agent undertakes a high effort level, e_h, then the probability of the good state being realized is p_h if the agent undertakes a low effort level, e_l, then the probability of the bad state being realized is p_l $p_h > p_l$. Let $u(\cdot)$ denote the agent's utility function which depends on eared income. We assume that agent is risk averse while the principal is risk neutral. The first differential of $u(\cdot)$ is positive and the second differential of it is negative. Furthermore we assume that the principal is a monopolist. He faces no competitor in hiring agents in doing his business.

If the principal can monitor agent's effort fully, the first best contract in the perspective of the principal is easy to be designed. Skipping this case we turn on to the so called second best contract. Given that monitoring the agent's effort level is costly, the principal's task is for the agent to exert the high effort level. To do this, the principal designs the direct mechanism or menu of contracts among which the agent selects. The menu of contracts or direct mechanism depending on the state and agent's effort level is given as,

$y(\cdot,\cdot) = \{y_1(e_h), y_2(e_h)\}$.

Where y denotes income level depending on the state and effort level, and the subscripts 1 or 2 attached toy denote states 1 or2,

Here we ignored for the case of agent's exerting low effort level because our main concern is to induce high effort level.

Above mechanism should satisfy the following incentive-compatibility constraints (IC):

$$p_h u(y_1(e_h) + (1-p_h)u(y_2(e_h)) - e_h \geq p_l u(y_1(e_h) + (1-p_l)u(y_2(e_h)) - e_l$$

This incentive constraint is that the agent's utility is higher when exerting high effort level rather than when exerting low effort level. The contract for the high effort level should be self selected by the agent. Taking also into account of individual's rationality constraint (IR) the principal the principal maximizes his profits

$$\max_{\{y_1(e_h),y_2(e_h)}\pi = \{y-(1-(1-p_h)L\} - \{P_h y_1(e_h) + (1-p_h)y_2(e_h))\}$$

Subject to (IC), and

$$p_h u(y_1(e_h) + (1-p_h)u(y_2(e_h)) - e_h \geq \bar{u}(IR)$$

Where $\bar{u}$ denotes the agent's utility from outside option

Solving this problem we can show that the agent's income level in state 1 is greater than that in state 2 when he exerts higher effort level. This is the incentive structure for the principal to extract the higher effort level from the agent. The constant income level which the agent received in this incentive scheme regardless of the state occurred does not induce any higher effort level. That is the reason why we don't care to induce lower effort level from the agent: the agent would not exert higher effort level when there is no incentive to do that.

V. Concluding remarks

We have reviewed the concept of direct mechanism and it applications in economic transactions occurred between players. We derived the revelation principle of direct mechanism for the indirect mechanism. The indirect mechanism proposed by the principal or the mechanism designer is a Bayesian game form with some communication channels between the principal and the agents in the game. By restricting the range of message spaces to the original type spaces of the agents, we show the profile of true revelation of types is the Bayesian equilibrium in the direct mechanism. The next step in the spirit of this paper (some kind of introductory paper for a novice) is to review some extensions to dynamic settings and also refined concept of mechanism design.

References

Roger Myerson and Mark Satterthwaite(1983), Efficient mechanisms for bilateral trading, JOURNAL OF Economic Theory, 29, 102, p. 125-139.

Roger Myerson(1981), optimal auction design, Mathematics of Operation Research, 6, p. 58-73.

■ Various lecture notes:

Botond Koszegi(2004), Mechanism Design.

Susan Athey(1999), Lecture notes for advanced microeconomics.