Moral Hazard Resolved by Common-Knowledge in S5n-Logic *

Takashi Matsuhisa

Department of Natural Sciences, Ibaraki National College of Technology
Nakane 866, Hitachinaka-shi, Ibaraki 312-8508, Japan.
E-mail: mathisa@ge.ibaraki-ct.ac.jp

Abstract. This article investigates the role of common-knowledge in the principal-agent model under asymmetric information. We treat the problem: How the common-knowledge condition will be able to settle a moral hazard problem in the principal-agents model under asymmetric information. We shall propose a solution program for the moral hazard in the principal-agents model under asymmetric information by common-knowledge. Let us start that the agents have the knowledge structure induced from a partition relation associated with the multi-modal logic **S5n**. In particular we consider the situation that the agents commonly know all decision values of the other agents. Under certain assumptions we shall show the moral hazard can be resolved in the principal-agents model when all the expected marginal costs are common-knowledge among the principal and agents.
Keywords: Agreeing to Disagree, Common-Knowledge, Moral hazard, Principal-agents model under uncertainty, **S5n**-logic.
AMS 2000 Mathematics Subject Classification: Primary 91A35, Secondary 03B45.
Journal of Economic Literature Classification: C62, C78.

1 Introduction

This article considers the relationship between common-knowledge and agreement in multi-agent system. How we capture the fact that the agents agree on an event or they get consensus on it? We treat the problem from Fuzzy set theoretical flavour. The purposes are first to introduce common-knowledge structure on multi-agent system, and by which we show that all agents can agree on an event, and second to apply the result to solving the moral hazard in a principal-agents model under asymmetric information. Let us consider there are agents more than two and the agents have the structure of the Kripke semantics for the multi-modal logic **S5n**.

Assume that all agents have a common probability measure. By i's decision value of an event under agent i's private information, we mean the conditional probability value of the event under agents' private information. We say that

* The abstract was presented in CSM 2009, Laxenburg (Austria)

consensus on the set can be guaranteed among all agents (or they agree on it) if all the the membership values are equal.

Aumann [1] considered the situation that the agents have common-knowledge of the membership values; that is, simultaneously everyone knows the membership values, and everyone knows that 'everyone knows the values' and everyone knows that "everyone knows that 'everyone knows the values" and so on. He showed the famous agreement theorem for a partition information structure equivalent to the multi-modal logic **S5n**, and Samet [6] and Matsuhisa & Kamiyama [4] extend the theorem to models corresponding to weaker logics (e.g. **S4n** etc.)

Theorem 1 (Aumann [1]). *The agents can agree on an event if all membership values of the event under private information are common-knowledge among them.*

We shift our attention to the principal-agents model as follows: an owner (principal) of a firm hires managers (agents), and the owner cannot observe how much effort the managers put into their jobs. In this setting, the problem known as the moral hazard can arise: There is no optimal contract generating the same effort choices for the manager and the agents. We apply Theorem 1 to solve the problem. The aim is to establish that

Theorem 2. *The owner and the managers can reach consensus on their expected marginal costs for their jobs if their expected marginal costs are common-knowledge.*

This article organises as follows. In Section 2 we describe the moral hazard in our principal-agents model. Sections 3 and 4 introduce the notion of common-knowledge associated with partitional (reflexive, transitive and symmetric) information structure and the notion of decision function. Section 5 gives the formal statement of Theorem 1. In Section 6 we introduce the formal description of a principal-agents model under asymmetric information. We will propose the program to solve the moral hazard in the model: First the formal statement of Theorem 2 is given, and secondly what further assumptions are investigated under which Theorem 2 is true. In the final section we conclude with remarks.

2 Moral Hazard

Let us consider the principal-agents model as follows: There are the principal P and n agents $\{1, 2, \cdots, k, \cdots, n\}$ $(n \geq 1)$ in a firm. The principal makes a profit by selling the productions made by the agents. He/she makes a contact with each agent k that the total amount of all profits is refunded each agent k in proportion to the agent's contribution to the firm.

Let e_k denote the measuring managerial effort for k's productive activities. The set of possible efforts for k is denoted by E_k with $E_k \subseteq \mathbb{R}$. Let $I_k(\cdot)$ be a real valued continuously differentiable function on E_k. It is interpreted as the profit by selling the productions made by the agent k with the cost $c(e_k)$. Here we assume $I'_k(\cdot) \geq 0$ and the cost function $c(\cdot)$ is a real valued continuously differentiable function on $E = \cup_{k=1}^n E_k$. Let I_P be the total amount of all the research grants awarded:

$$I_P = \sum_{k=1}^n I_k(e_k).$$

The principal P cannot observe these efforts e_k, and shall view it as a random variable on a probability space (Ω, μ). The optimal plan for the principal then solves the following problem:

$$\text{Max}_{e=(e_1,e_2,\cdots,e_k,\cdots,e_n)}\{\text{Exp}[I_P(e)] - \sum_{k=1}^n I_k(e_k)\}.$$

Let $W_k(e_k)$ be the total amount of the refund to agent k:

$$W_k(e_k) = r_k I_P(e),$$

with $\sum_{k=1}^n r_k = 1, 0 \leq r_k \leq 1$, where r_k denotes the proportional rate representing k's contribution to the firm. The optimal plan for each agent also solves the problem: For every $k = 1, 2, \cdots, n$,

$$\begin{aligned} &\text{Max}_{e_k}\{\text{Exp}[W_k(e_k)] - c(e_k)\} \\ &\text{subject to } \sum_{k=1}^n r_k = 1, 0 \leq r_k \leq 1. \end{aligned}$$

We assume that r_k is independent of e_k, and the necessity conditions for critical points are as follows: For each agent $k = 1, 2, \cdots, n$, we obtain

$$\begin{aligned} &\frac{\partial}{\partial e_k}\text{Exp}[I_k(e_k)] - c'(e_k) = 0 \\ &r_k \frac{\partial}{\partial e_k}\text{Exp}[I_k(e_k)] - c'(e_k) = 0 \end{aligned}$$

in contraction. This contradictory situation is called the **moral hazard** in the principal-agents model.

3 Common-Knowledge

Let N be a set of finitely many agents and k denote an agent. The specification is that $N = \{P, 1, 2, \cdots, k, \cdots, n\}$ consists of the principal P and the faculty members $\{1, 2, \cdots, k, \cdots, n\}$ in the college. A state-space Ω is a non-empty set, whose members are called *states*. An *event* is a subset of the state-space. If Ω is a state-space, we denote by 2^Ω the field of all subsets of it. An event E is said to occur in a state ω if $\omega \in E$.

3.1 Information and Knowledge

By *partition information structure* we mean $\langle \Omega, (\Pi_i)_{i\in N}\rangle$ in which $\Pi_i : \Omega \to 2^\Omega$ satisfies the three postulates: For each $i \in N$ and for any $\omega \in \Omega$,

Ref $\omega \in \Pi_i(\omega)$;
Trn $\xi \in \Pi_i(\omega)$ implies $\Pi_i(\xi) \subseteq \Pi_i(\omega)$.
Sym If $\xi \in \Pi_i(\omega)$ then $\omega \in \Pi_i(\xi)$.

This structure is equivalent to a Kripke semantics for the multi-modal logic **S5n**. The set $\Pi_i(\omega)$ will be interpreted as the set of all the states of nature that i knows to be possible at ω, or as the set of the states that i cannot distinguish from ω. We call $\Pi_i(\omega)$ i's *information set* at ω.

We will give the formal model of knowledge as follows (C.f.; Fagin et al [2].)

Definition 1. *The* **S5**-knowledge structure *is a tuple* $\langle \Omega, (\Pi_i)_{i\in N}, (K_i)_{i\in N}\rangle$ *that consists of a partition information structure* $\langle \Omega, (\Pi_i)_{i\in N}\rangle$ *and a class of i's* knowledge operator $K_i : 2^\Omega \to 2^\Omega$ *defined by*

$$K_i E = \{\omega \mid \Pi_i(\omega) \subseteq E\}$$

The event $K_i E$ will be interpreted as the set of states of nature for which i knows E to be possible.

We record the properties of i's knowledge operator: For every E, F of 2^Ω,

N $K_i\Omega = \Omega$;
K $K_i(E \cap F) = K_iE \cap K_iF$;
T $K_iE \subseteq E$
4 $K_iE \subseteq K_i(K_iE)$.
5 $\Omega \setminus K_iE \subseteq K_i(\Omega \setminus K_iE)$.

According to these properties we can say the structure $\langle \Omega, (K_i)_{i\in N}\rangle$ is a model for the multi-modal logic **S5n**.

3.2 Common-Knowledge and Communal information

The *mutual knowledge operator* $K_E : 2^\Omega \to 2^\Omega$ is the intersection of all individual knowledge operators:

$$K_E F = \cap_{i\in N} K_i F,$$

which interpretation is that everyone knows E.

Definition 2. *The* common-knowledge operator $K_C : 2^\Omega \to 2^\Omega$ *is defined by*

$$K_C F = \cap_{n\in\mathbb{N}} (K_E)^n F.$$

The intended interpretations are as follows: $K_C E$ is the event that 'everyone knows E' and "everyone knows that 'everyone knows E'," and "'everybody knows that "everyone knows that 'everyone knows E'," ,"' and $\cdots$ An event E is *common-knowledge* at $\omega \in \Omega$ if $\omega \in K_C E$.

Let $M : 2^\Omega \to 2^\Omega$ be the dual of the common- knowledge operator K_C:

$$ME := \Omega \setminus K_C(\Omega \setminus E).$$

By the *communal* information function we mean the function $M : \Omega \to 2^\Omega$ defined by $M(\omega) = M(\{\omega\})$. It can be plainly observed that the communal information function has the following properties:

Proposition 1. *Notations are the same as above. Then*

(i) $\omega \in K_C E$ *if and only if* $M(\omega) \subseteq E$
(ii) *For every* $i \in N$, $M(\omega)$ *can be decomposed into the disjoint union of the components* $\Pi_i(\xi)$ *for* $\xi \in M(\omega)$*: i.e.,* $M(\omega) = \sqcup_{\xi \in M(\omega)} \Pi_i(\xi)$.

Proof. See, Fagin et al [2].

4 Decision Function and Consensus

Let Z be a set of decisions, which set is common for all agents. By a *decision function* we mean a mapping f of $2^\Omega \times 2^\Omega$ into the set of decisions Z. We refer the following property of the function f: Let X be an event.

Disjoint Union Consistency (DUC): For every pair of disjoint events S and T, if $f(X; S) = f(X; T) = d$ then $f(X; S \cup T) = d$;

By the *i's decision function* associated with f under agent i's private information we mean the function d_i from $2^\Omega \times \Omega$ into Z defined by $d_i(X; \omega) = f(X; \Pi_i(\omega))$, and we call $d_i(X; \omega)$ the *i's decision value* of X associated with f under agent i's private information at ω.

Definition 3. *We say that* consensus *on X can be guaranteed among all agents (or they* agree on *it) if* $d_i(X; \omega) = d_j(X; \omega)$ *for any agent* $i, j \in N$ *and in all* $\omega \in \Omega$.

Remark 1. If f is intended to be a posterior probability, we assume given a probability measure μ on a state-space Ω which is common for all agents; precisely, for some event X of Ω, $f(X; \cdot)$ is given by $f(X; \cdot) = \mu(X|\cdot)$. Then the i's decision value of X is the conditional probability value $d_i(X; \omega) = \mu(X|\Pi_i\omega))$. Consensus on X guaranteed among all agents can be interpreted as that the fuzzy sets (X, d_i) and (X, d_j) are equal for any $i, j \in N$

5 Agreeing to Disagree Theorem

We can now state explicitly Theorem 1 as below: Let D be the event of the i's decision values of an event X for all agents at ω, which is defined by

$$D = \cap_{i \in N}\{\xi \in \Omega \mid d_i(X;\xi) = d_i(X;\omega)\}.$$

Theorem 3. *Assume that the agents have the* **S5n***-knowledge structure and the decision function f with satisfying the condition (DUC). If $\omega \in K_C D$ then $d_i(X;\omega) = d_j(X;\omega)$ for any agents $i, j \in N$ and in all $\omega \in \Omega$.*

Proof. By Proposition 1 it is plainly observed that

$$M(\omega) = \sqcup_{\xi \in M(\omega)} \Pi(\xi) \subseteq D \subseteq \{\xi \in \Omega \mid d_i(X;\xi) = d_i(X;\omega)\}.$$

On noting that $d_i(X;\xi) = d_i(X;\omega)$ for any $\xi \in M(\omega)$, it can be observed by (DUC) that $f(X;M(\omega)) = d_i(X;\omega)$ for every $i \in N$, and thus $d_i(X;\xi) = d_j(X;\omega)$ for any $i, j \in N$. □

6 Moral Hazard Revisited

This section investigates the moral hazard problem from the common-knowledge view point. Let us reconsider the principal-agents model and let notations and assumptions be the same in Section 2. We show the evidence of Theorem 2 under additional assumptions **A1-2** below. This will give a possible solution of our moral hazard problem.

A1 The principal P has a partition information $\{\Pi_P(\omega) \mid \omega \in \Omega\}$ of Ω, and each agent k has also his/her a partition information $\{\Pi_k(\omega) \mid \omega \in \Omega\}$:
A2 For each $\omega, \xi \in \Omega$ there exists the decision function $f : 2^\Omega \times 2^\Omega \to \mathbb{R}$ satisfying the Disjoint Union Consistency together with
- **(a)** $f(\{\xi\};\Pi_P(\omega)) = \frac{\partial}{\partial e_0(\xi)} \mathrm{Exp}[I_P(e)|\Pi_P(\omega)]$;
- **(b)** $f(\{\xi\};\Pi_k(\omega)) = \frac{\partial}{\partial e_k(\xi)} \mathrm{Exp}[W_k(e)|\Pi_k(\omega)]$

We have now set up the principal-agents model under asymmetric information. The optimal plans for principal P and agent k are then to solve

PE $\mathrm{Max}_{e=(e_1,e_2,\cdots,e_k,\cdots,e_n)}\{\mathrm{Exp}[I_P(e)|\Pi_P(\omega)] - \sum_{k=1}^n I_k(e_k)\}$;
AE $\mathrm{Max}_{e_k}\{\mathrm{Exp}[W_k(e_k)|\Pi_k(\omega)] - c(e_k)\}$ subject to $\sum_{k=1}^n r_k = 1, 0 \leq r_k \leq 1$.

From the necessity condition for critical points together with **A2** it can been seen that the principal's marginal expected costs for agent k is given by

$$c'_P(e_k(\xi);\omega) = f(\xi;\Pi_P(\omega)),$$

and agent k's expected marginal costs is also given by

$$c'_k(e_k(\xi);\omega) = f(\xi;\Pi_P(\omega)).$$

To establish this solution program we have to solve the problem: Construct the information structure together with decision function such that the above conditions **A1** and **A2** are true. Under these circumstances, a resolution of the moral hazard given by Theorem 2 will be restate as follows: We denote

$$[c'(e(\xi);\omega)] = \cap_{i\in N}\{\zeta \in \Omega | f(\xi; \Pi_i(\zeta)) = f(\xi; \Pi_i(\omega))\}.$$

Theorem 4. *Under the conditions* **A1** *and* **A2** *we obtain that for each* $\xi \in \Omega$*, if* $\omega \in K_C([c'(e(\xi);\omega)])$ *then* $c'_P(e_k(\xi);\omega) = c'_k(e_k(\xi);\omega)$ *for any* $k = 1, 2, \cdots, n$.

Proof. Follows immediately from Theorem 3. □

Remark 2. To establish Theorem 4 we have to solve the problem: Construct the information structure $(\Pi_i)_{i\in N}$ together with decision function f such that the above conditions **A1** and **A2** are true.

7 Concluding Remarks

It ends well this article to pose additional problems for making further progresses:

1. If the proportional rate r_k representing k's contribution to the college depends only on his/her effort for research activities in the principal-agents model, what solution can we have for the moral hazard problem?

2. Can we construct a communication system for the principal- agents model in the line of Parik &Krasucki [5]? Where each agent and the principal communicate privately each other about their expected marginal cost as messages: The principal sends his expected marginal cost as messages. The agent as the recipient of the message revises his/her information structure and recalculates the expected marginal cost under the revised information structure, and he/she sends the revised expected marginal cost to the principal. The principal as the recipient of the message revises his/her information structure and recalculates the expected marginal cost under the revised information structure, and he/she sends the revised expected marginal cost to the agent, and so on. In the circumstance does the limiting expected marginal costs actually coincide ? Matsuhisa (2008) introduces a fuzzy communication system and extends Theorem 3 in the communication model. By using this model Theorem 4 can be extended in the communication framework, and the detail will be reported in near future.

Acknowledges

The paper was financially supported by Ibaraki Kosen Kenkyu Suisin Seido (2009 Promotion for Advancing Scientific Researches at Ibaraki National College of Technology).

References

1. Aumann, R. J. (1976). Agreeing to disagree. *Annals of Statistics* 4: 1236–1239.
2. Fagin, R., Halpern, J. Y. , Moses , Y. and Vardi, M. Y. (1995). *Reasoning about Knowledge*, MIT Press, Cambridge, Massachusetts, London, England.
3. Matsuhisa,T. (2008). Fuzzy communication reaching consensus under acyclic condition. T.-B. Ho and Z.-H. Zhou (Eds.): *PRICAI2008*. Lecture Notes in Computer Science 5351, Springer: 760-767.
4. Matsuhisa, T. and Kamiyama, K. (1997). Lattice structure of knowledge and agreeing to disagree. *Journal of Mathematical Economics* 27: 389-410.
5. Parikh, R. and Krasucki, P. (1990). Communication, consensus, and knowledge. *Journal of Economic Theory* 52: 178–189.
6. Samet, D. (1990). Ignoring ignorance and agreeing to disagree. *Journal of Economic Theory* 52: 190-207.