I. Troch, F. Breitenecker, eds.   ISBN 978-3-901608-35-3

COMMON-KNOWLEDGE RESOLVES MORAL HAZARD

Takashi Matsuhisa
Ibaraki National College of Technology, Ibaraki, Japan

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

Abstract. This article investigates the relationship between common-knowledge and agreement in multi-agent system, and to apply the agreement result by common-knowledge to the principal-agent model under asymmetric information. We treat the two problems: (1) how we capture the fact that the agents agree on an event or they get consensus on it from epistemic point of view, and (2) how the agreement theorem will be able to make progress to settle a moral hazard problem in the principal-agent model under asymmetric information. We shall propose a solution program for the moral hazard in the principalagents model under asymmetric information by common-knowledge. Let us start that the agents have the knowledge structure induced from an equivalence relation associated with the multi-modal logic **S5n**. Each agent obtains the membership value of an event under his/her private information, so he/she considers the event. Specifically consider the situation that the agents commonly know all membership values of the other agents. In this circumstance we shall show the agreement theorem that consensus on the membership values among all agents can still be guaranteed. Furthermore, under certain assumptions we shall show the moral hazard can resolve in the principal-agent model when all the expected marginal costs are common-knowledge among the principal and agents.

1 Introduction

This article considers the relationship between communication 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 flavor. The purposes are first to introduce a knowledge revision system through communication 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 fuzzy structure given by the dual structure of the Kripke semantics for the multi-modal logic **S5n**.

Assume that all agents have a common probability measure. By the membership value of an event under agent i's private information, we mean the conditional probability value of the event under agentĄfs private information. We say that consensus on the set can be guaranteed among all agents (or they agree on it) if all the the membership values are equal.

R.J. 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 Ąeeveryone knows the valuesĄf, and everyone knows that “everyone knows that ‘everyone knows the valuesĄf,Ąh and so on. He showed the famous agreement theorem:

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 organizes as follows. In Section 2 we describe the moral hazard in our principal-agents model. Section 3 and 4 introduce the notion of common-knowledge associated with a partition information structure and the notion of decision function. Section 5 gives the formal statement of Theorem 1 with a sketch of the proof. In the proof, the property ‘Disjoint Union Consistency’ for the function plays crucial role. 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 imagine that there are a president and faculty members in a National College of Technology in Japan. The president wishes increasing the income of his/her college, and he/she frequently encourages them to get any research grants awarded from outside organizations. The research activities of the college is very low, and so the amount of research grants awarded from outside organizations is poor comparing with those of the other National Colleges of Technology in Japan. The president considers this situation comes from the lower stage of the incentives of the faculty members for their research activities. The president proposes to the faculty member the plan to keep their research activities up: The 30 percent of the amount of each research grant awarded from outside organizations goes for overhead costs to the college. The overhead costs will be refunded to each faculty members for improving his/her incentive to research activities. The amount of the refund costs to each member shall be in proportion to his/her educational contribution to the college, where the contribution shall be evaluated by the grades of questionnaires for his/her lessons and by his/her school duties excluding his/her research activities.

Specifically, let us consider the principal-agents model as follows: The principal P is the president in the college and agents $\{1,2,\cdots,k,\cdots,n\}$ are the faculty members. Let k be an agent and e_k the measuring managerial effort for kĄfs research activities. The set of possible efforts for all agents 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 income of the research grant that is obtained from outside organization 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 president 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 president then solves the following problem:

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

Let $W_k(e_k)$ be the total amount of the research grants actually obtained:

$$W_k(e_k) = \left(I_k(e_k) - \frac{3}{10} I_k(e_k)\right) + r_k I_P(e),$$

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

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

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

$$\frac{\partial}{\partial e_k}\mathrm{Exp}[I_k(e_k)] - c'(e_k) = 0$$
$$\left(\frac{7}{10} - r_k\right)\mathrm{Exp}[I_k(e_k)] - c'(e_k) = 0$$

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 i denote an agent. The specification is that $N = \{P,1,2,\cdots,k,\cdots,n\}$ consists of the president 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

A *partition information structure* $\langle \Omega, (\Pi_i)_{i \in N} \rangle$ consists of a state space Ω and a class of agent i's information functions $\Pi : \Omega \to 2^{\Omega}$ satisfying the postulates

1. $\{\Pi_i(\omega) | \omega \in \Omega\}$ is a partition of Ω;
2. $\omega \in \Pi_i(\omega)$.

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:[1]

Definition 1. The *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 on 2^{Ω} 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[2]: For every E, F of 2^{Ω},

N $K_i \Omega = \Omega$, $K_i \emptyset = \emptyset$; **K** $K_i(E \cap F) = K_i E \cap K_i F$;

T $K_i E \subseteq E$ **4** $K_i E \subseteq K_i(K_i E)$;

5 $\Omega \setminus K_i E \subseteq K_i(\Omega \setminus K_i E)$.

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* $K_C : 2^{\Omega} \to \Omega$ is defined by

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

The intended interpretations are as follows: An event E is *common-knowledge* at $\omega \in \Omega$ if $\omega \in K_C E$, and $K_C E$ is the event of common-knowledge of 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.*

(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]. □

[1] C.f.; Fagin et al [2].

[2] 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**.

4 Decision function and Membership values

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 properties of the function f: Let X be an event.

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

Preserving Under Difference: For all events S and T such that $S \subseteq T$, if $f(X;S) = f(X;T) = d$ then $f(X;T \setminus S) = d$.

By the *membership 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 *membership 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 $\omega \in \Omega$.

Example 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 membership value of X is the conditional probability value $d_i(X;\omega) = \mu(X|\Pi_i\omega))$. The pair (X, d_i) can be considered as as a fuzzy set X associated with agent i's membership function d_i. 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 ĄeAgreeing to disagreeĄf theorem of Aumann

We can now state explicitly Theorem 1 known as the agreement theorem of R. J. Aumann [1] as below: Let D be the event of the membership degrees 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 (R.J. Aumann [1]). *Assume the decision function f satisfies the Disjoint Union Consistency. 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 immediately follows 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 Disjoint Union Consistency 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 $i, j \in N$. □

Remark 1. Theorem 3 can be extended to the following cases:

Reflexive and Transitive Information structure: By this we mean $\langle \Omega, (\Pi_i)_{i \in N} \rangle$ in which $\Pi_i : \Omega \to 2^{\Omega}$ satisfies only the two properties: For ach $i \in N$ and for any $\omega \in \Omega$,

Ref $\omega \in \Pi_i(\omega)$;

Tran $\xi \in \Pi_i(\omega)$ implies $\Pi_i(\xi) \subseteq \Pi_i(\omega)$.

D. Samet [6] extends Theorem 3 for the reflexive and transitive Information structure.

Lattice Structure of Knowledge: Matsuhisa and Kamiyama [4] introduces the structure which involves the properties **N** and **K** of knowledge operators, and they shows that Theorem 3 can still valid in thr framework of the lattice structure of knowledge.

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 the information partition $\{\Pi_P(\omega) \mid \omega \in \Omega\}$ of Ω, and each faculty member k has also his/her partition $\{\Pi_k(\omega) \mid \omega \in \Omega\}$:

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 $\sum_{k=1}^{n} I_k(r_k) = 1, 0 \leq r_k \leq 1$.

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

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

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

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

To establish this solution program we have to solve the problem: Construct the information partition 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

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

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

Proof. Follows immediately from Theorem 3. □

Remark 2. To establish Theorem we have to solve the problem: Construct the information partition 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 rk 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 appeared in Parikh and Krasucki [5] for the principal- agents model, where the agents including Principal communicate each other about their expected marginal cost as messages. The recipient of the message revises his/her information structure and recalculates the expected marginal cost under the revised information structure. The agent sends the revised expected marginal cost to another agent according to a communication graph, and so on. In the circumstance does the limiting expected marginal costs actually coincide ? Matsuhisa [3] introduces a fuzzy communication system in the line of Parikh and Krasucki [5], 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.

8 References

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