数理解析研究所講究録
1166 巻 2000 年 86-94

Agreeing theorem in an S-4 logic model *

平瀬 和基 (Kazuki Hirase) ‡

Graduate School of Economics, Keio University, Mita 2-15-45, Minatoku, Tokyo, 108-8345, Japan

This paper introduces an application of the S-4 logic. There are two aims in this paper. Aim 1 is to check the relation between our model and the S-4 logic. We'll see the soundness and completeness of the S-4 logic with respect to the model by using the concept of structure. Aim 2 is to prove Agreeing theorem in the model. Roughly speaking, Agreeing theorem insists that if peoples' information structure satisfy some conditions, then their posteriors are equal.

1. INTRODUCTION

The word "knowledge" and especially "common knowledge" play a very important role in game theory. Intuitively, an event is common knowledge if everyone knows it, everyone knows that everyone knows it, everyone knows that everyone knows that everyone knows it and so on. Then how can we treat (common) knowledge formally?

Aumann (1976) tried to solve this problem. He introduced the formal notion of common knowledge based on partitional information structure and showed a theorem that players who have the common prior can not agree to disagree, that is, if their posteriors for a given event are common knowledge, then these must be equal, even though they are based on different information. In our paper, we call this theorem as Agreeing Theorem.

After Aumann, many papers have studied knowledge. Milgrom (1981), and Monderer and Samet (1989) treated knowledge by different approaches. Milgrom (1981) applied axiomatic approach [1] to model knowledge. Monderer and Samet (1989) used probability approach[2]. They managed to ap-

*I would like to thank Shin-ichi Suda and Takashi Matsuhisa for helpful comments.

‡E-mail address: hirase@pc.highway.ne.jp

[1]This is the approach which defines the set of all states in which a player knows a given event. After Milgrom's paper, many papers have been written by this approach.

[2]Probability approach defines the event in which player n believes E with probability at least p.

proximate (common) knowledge with belief. They also proved that Agreeing Theorem holds when knowledge is replaced by belief.

While various approaches to model knowledge have shown up, Agreeing Theorem has been modified too. Geanakoplos and Polemachakis (1982) explained the process of agreeing. Samet (1990) studies non-partitional information structure. In his paper, a state describes everything, even information structure. Based on this idea, he showed Agreeing Theorem holds in non-partitional case. Matsuhisa and Kamiyama (1997) generalized Samet's result using lattice and filter theory.

This paper also studies non-partitional information structure, or an extension of Aumann's Agreeing Theorem. We would like to prove Agreeing Theorem after showing the relation between the model and the logic. In section 5, we will see the model is one of the S-4 logic. Agreeing Theorem is proved in section 8.

2. S-4 LOGIC

S-4 logic is denoted as $\langle L, S, AR\rangle$.

L means language. Language consists of N, PV, logical connectives, and players' modal operators. N means a set of players 1 and 2. Now, we restrict the number of the players to 2 persons for simplicity. But we can extend the results to n persons case easily. PV is a set of propositional variables, or atomic sentences. Logical connectives are $\wedge, \vee, \rightarrow, (,)$, and $\neg$. Players' modal operators are $\Box_1$ and $\Box_2$.

The second element of S-4 logic is S. S means a set of sentences, or a set of formulae. S is inductively constructed from L.

$(S1) : PV \subseteq S$

$(S2) : \phi, \psi \in S \Rightarrow \neg\phi, \phi \rightarrow \psi, \phi \wedge \psi, \phi \vee \psi, \Box_n\phi \in S \ (n = 1, 2)$

$(S3)$: Every sentence is constructed by a finite number of applications of (S1) and (S2).

The third element of the S-4 logic is AR. AR means axioms and rules. AR consists of PL, inference rules and modal axioms and rules.

PL is propositional logic, or a set of all tautologies, that is, for all $\phi, \psi, \chi \in S$.

$(PL1) : \phi \rightarrow (\psi \rightarrow \phi)$

$(PL2) : (\phi \rightarrow (\psi \rightarrow \chi)) \rightarrow ((\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \chi)$

$(PL3) : (\neg\phi \rightarrow \neg\psi) \rightarrow ((\neg\phi \rightarrow \psi) \rightarrow \phi)$

$$(PL4) : \phi \wedge \psi \to \phi$$

$$(PL5) : \phi \to \phi \vee \psi$$

Inference rules are Modus Ponens (MP), $\wedge$-rule, and $\vee$-rule.

$$(MP) : \frac{\phi \to \psi \quad \phi}{\psi}$$

$$(\wedge - rule) : \frac{\phi \to \psi \quad \phi \to \chi}{\phi \to \psi \wedge \chi}$$

$$(\vee - rule) : \frac{\psi \to \phi \quad \chi \to \phi}{\psi \vee \chi \to \phi}$$

where $\phi, \psi, \chi \in S$.

For modal part, we assume axioms $K, T, 4$ and N. K is, in other words, the Axiom of Distribution. T is the Axiom of Knowledge. 4 is the Positive Introspection. And N is the Necessitation rule.

$$(K) : (\Box_n(\phi \to \psi) \to (\Box_n\phi \to \Box_n\psi))$$

$$(T) : \Box_n\phi \to \phi$$

$$(4) : \Box_n\phi \to \Box_n\Box_n\phi$$

$$(N) : \frac{\phi}{\Box_n\phi}$$

where $\phi, \psi \in S$ and $n = 1, 2$

With these axioms and rules, we can define the provability of a sentence in the logic.

DEFINITION 2.1. A *proof* is a finite tree satisfying (PR1) and (PR2).

$(PR1)$: A sentence is associated with each node, and the sentence associated with every leaf node is an instance of $(PL1)$–$(PL5)$, K, T, or 4.

$(PR2)$: Each adjoining node forms an instance of MP, $(\wedge - rule)$, $(\vee - rule)$, or N.

We say that $\phi (\in S)$ is *provable* in the S-4 logic if and only if there exist a proof which root is associated with ϕ.

3. STRUCTURE

Structure is $\langle \Omega, P_1, P_2 \rangle$. Ω is a nonempty finite state space. So 2^Ω is called a set of events.

Players' *information functions* P_1 and P_2 is a function from the state space Ω to the event set 2^Ω. The set $P_n(\omega)$ means the event which player n recognize when the real state is ω. The set $P_n(\omega)$ is called player n's information set or possibility set at ω.

We assume that each players' information function satisfy the following.

$$(P-1) : \omega \in P_n(\omega)$$

$$(P-2) : \omega' \in P_n(\omega) \Rightarrow P_n(\omega') \subseteq P_n(\omega)$$

$$\text{for } \forall \omega, \omega' \in \Omega \text{ and } n = 1, 2$$

P-1 means the condition that each player never excludes the real state. When the real state is ω, the player n thinks that ω may have occurred. From P-2, we have that if there is a state ξ, so that $\xi \in P_n(\omega')$ and $\xi \notin P_n(\omega)$ then $\omega' \notin P_n(\omega)$. So, P-2 says that player n at ω can make consideration as follows: "The state ξ is excluded. If it were the state ω', I would not exclude ξ. Thus it must be that the state is not ω'."

P-1 and P-2 play very important roles in the relation to the S-4 logic. We call these three tuples $\langle \Omega, P_1, P_2 \rangle$ an information structure.

In Aumann's paper, P-3: $\omega' \in P_n(\omega) \Rightarrow P_n(\omega') \supseteq P_n(\omega)$ $(n = 1, 2)$ for $\omega, \omega' \in \Omega$ was also assumed. So we can say that our model is an extension of Aumann's.

Consider the case that $\omega' \in P(\omega)$ and there is a state $\xi \in P(\omega)$ that is not in $P(\omega')$. Then, P-3 says that a player at ω can conclude, from the fact that he (she) can not exclude ξ, that the state is not ω', a state at which he (she) would be able to exclude ξ.

Note the following proposition holds.

PROPOSITION 3.1. *Player n's information function P_n satisfies P-1, 2, and 3 if and only if there is a partition of Ω such that for any $\omega \in \Omega$ the set $P_n(\omega)$ is the element of the partition that contains ω.*

Proof. Suppose that P_n satisfies P-1, 2, and 3. If $P_n(\omega)$ and $P_n(\omega')$ intersect and $\xi \in P_n(\omega) \cap P_n(\omega')$ then by P-2 and 3, we have $P_n(\omega) = P_n(\omega') = P_n(\xi)$. By P-1 we have $\bigcup_{\omega \in \Omega} P_n(\omega) = \Omega$.

The other direction is obvious. ∎

Thus, Aumann's paper treated a partitional information structure. But we don't assume P-3. We treat a non-partitional information structure.

4. MODEL

The model $\mathcal{M}$ consists of L, S, an information structure, a truth assignment π, and a valuation relation $\models$, i.e., $\mathcal{M} = \langle L, S, \Omega, P_1, P_2, \pi, \models\rangle$.

A truth assignment π is a function from $PV \times \Omega$ to the set $\{\top, \bot\}$.

From this truth assignment, the valuation relation $\models$ decides the validity of the sentences. For any sentence ϕ and ψ, and for $n = 1, 2$, we define the valuation relation as follows.

(VR1): For any $v \in PV$, $\models_\omega \iff \pi(v, \omega) = \top$

(VR2): $\models_\omega \neg\phi \iff \models_\omega \phi$ does not hold.

(VR3): $\models_\omega \phi \rightarrow \psi \iff \models_\omega \neg\phi or \models_\omega \psi$

(VR4): $\models_\omega \phi \wedge \psi \iff \models_\omega \phi$ and $\models_\omega \neg\phi$

(VR5): $\models_\omega \phi \vee \psi \iff \models_\omega \phi$ or $\models_\omega \psi$

(VR6): $\models_\omega \Box_n\phi \iff P_n(\omega) \subseteq \{\xi \in \Omega : \models_\xi \phi\}$ for $n = 1, 2$

5. SOUNDNESS AND COMPLETENESS

With these preparations of logic, structure, and model, we can prove the following theorem, This theorem is well known by logicians as *soundness and completeness* (of the S-4 logic)

THEOREM 5.1. *A sentence ϕ is provable in the S-4 logic $\iff \models_\omega \phi$ for $\forall\omega \in \Omega$ in the model $\mathcal{M}$.*

Proof (sketch[3]). For soundness ($\Rightarrow$), we can verify that each sentence of AR[4] is valid at $\forall\omega \in \Omega$ in the model using the properties P-1 and P-2. For completeness ($\Leftarrow$), we can show that P-1 corresponds to Axiom T, P-2 corresponds to Axiom 4.[5] Our model where P-1 and P-2 assumed is the canonical model of the S-4 logic. ▮

[3]See Chellas (1980), Hughes and Cresswell (1996) for detail.

[4]A rule $\frac{\phi}{\psi}$ must be modified by a sentence $\phi \rightarrow \psi$.

[5]Note that P-3 corresponds to Axiom 5:$\Box_n\neg\phi \rightarrow \Box_n\Box_n\neg\phi$ for $\phi \in S$ $n = 1, 2$. And Aumann's model is one of the S-5 logic.

6. KNOWLEDGE AND COMMON KNOWLEDGE

Since Agreeing Theorem treats an epistemic condition for the agreement of the posteriors, We have to define the concept of knowledge, common knowledge, and posterior. This section defines the knowledge and common knowledge. The definitions here is based on Aumann(1976). Posterior is defined in the next section.

DEFINITION 6.1. Player n *knows* $E(\in 2^\Omega)$ at $\omega \iff P_n(\omega) \subseteq E$. $(n = 1, 2)$.

From the meaning of the information function, the player n knows that some state in $P_n(\omega)$ has occurred. Hence if $P_n(\omega) \subseteq e$, (of course) the player n know the state in E has occurred. With this interpretation, we have defined the player's knowledge.

Before defining common knowledge, we define the self-evident event.

DEFINITION 6.2. $F(\in 2^\Omega)$ is a *self evident* between 1 and 2 $\iff \omega \in F \Rightarrow P_n(\omega) \subseteq F$ for $n = 1, 2$.

An event F is a self-evident event between 1 and 2, if whenever it occurs players 1 and 2 know that it occurs. Now, we define common knowledge.

DEFINITION 6.3. E is *common knowledge* at ω between 1 and 2 $\iff$ there exist a self evident event F between 1 and 2 such that $\omega \in F \subseteq E$.

An event E is common knowledge between 1 and 2, if there is a self-evident event between 1 and 2 containing ω whose occurrence implies E.

7. PRIOR AND POSTERIOR

This section defines player's posterior of E based on a prior. We assume the existence of a prior and it is common for both players (So, a prior is called common prior.). Let the *common prior* be a probability measure μ on Ω. We denote the common prior to E as $\mu[E]$. And we assume $\mu[E] > 0$ for any event $E(\neq \emptyset)$. We consider that each player forms his (her) posterior based on the common prior. We assume *player n 's posterior* to some event at ω is a probability measure on Ω, $Q_n(\cdot;\omega)$. and we define $Q_n(E;\omega) = \dfrac{\mu[E \cap P_n(\omega)]}{\mu[P_n(\omega)]}$ $(n = 1, 2)$ for any event $E(\neq \emptyset)$. This is the conditional probability of E on $P_n(\omega)$. Agreeing Theorem in the next section shows an epistemic condition for the agreement of the posteriors.

8. AGREEING THEOREM

THEOREM 8.1 (Agreeing Theorem). *Suppose that* $\{\Omega, P_1, P_2\}$ *is an information structure,* $E \in 2^{\Omega} \setminus \{\emptyset\}, \omega \in \Omega, q_1 \in [0,1], q_2 \in [0,1]$, *and that* μ *is the common prior. If* $\{\omega \in \Omega : Q_1(E;\omega) = q_1\} \cap \{\omega \in \Omega : Q_2(E;\omega) = q_2\}$ *is common knowledge at* ω *between* 1 *and* 2, *then* $q_1 = q_2$.

To prove the theorem, we have to show some lemmata.

LEMMA 8.1. *For any self-evident event* F *between* 1 *and* 2, *and for* $n = 1, 2$, $F = P_n^1 \cup ... \cup P_n^t$, *where* $P_n^1, ..., P_n^t$ *are* $P_n(\omega_1), ..., P_N(\omega_t)$ *such that* $\omega_1, ..., \omega_t \in F$ *(t is a positive integer)* .

Proof. From P-1, $\omega_i \in P_n(\omega_i)$ for all $\omega_i \in \Omega$. So $F \subseteq P_n^1 \cup ... \cup P_n^t$, where $Pn^1, ..., P_n^t$ are $P_n(\omega_1), ..., P_n(\omega_t)$ such that $\omega_1, ..., \omega_t \in F$. And since F is a self-evident event, $P_n^1 \cup ... \cup P_n^t \subseteq F$, where $P_n^1, ..., P_n^t$ are $P_n(\omega_1), ..., P_n(\omega_t)$ such that $\omega_1, ..., \omega_t \in F$. ▌

We prepare some notations. From here, we abbreviate the index of the player and the subscript of P means the number of a stage. Let $I_m^i \equiv \#\{h : \omega_i \in P_m^h\}$ and $i(m) \in \arg\max_i I_m^i$. For all $k = 1, ..., t$, $P_o^k = P^k$. For h such that $h \in \{h \in \{1, ..., t\} : \omega_{i(m-1)} \in P_{m-1}^h$ and $h \neq i(m-1)\}$, $P_m^h = P_{m-1}^h \setminus P_{m-1}^{i(m-1)}$. For other $h \in \{1, ..., t\}$, $P_m^h = P_{m-1}^h$. Note that, with these notations, if $\max_i I_m^i = 1$ for some m, $\{P_m^1, ..., P_m^t\}$ is a partition of f.

Now we show that there exist some m^* for which $\max_i I_m^i = 1$. To show this, it is enough to prove that lemma 8.2: if P-2 holds till m-stage and $\max_i I_m^i \geq 2$, then $P_m^{i(m)} \neq \emptyset$ and lemma 8.3: P-2 holds until $\max_i I_m^i \geq 2$. Lemma 8.4:if P-2 holds at m-stage, then $P_m^{i(m)} \subseteq P_m^h$ hold for all $\in \{h \in \{1, ..., t\} : \omega^{i(m)} \in P_m^h$ and $h \neq i(m)\}$.

We formally define P-2 holds at m-stage.

DEFINITION 8.1. We say *P-2 holds at m-stage* if and only if $\omega_j \in P_m^{j'} \Rightarrow P_m^j \subseteq P_m^{j'} (j, j' \in \{1, 2, ..., t\})$.

LEMMA 8.2. *If P-2 holds till m-stage and* $\max_i I_m^i \geq 2$, *then* $P_m^{i(m)} \neq \emptyset$.

Proof. Suppose that $\max_i I_m^i \geq 2$ and $P_m^{i(m)} = \emptyset$. Since $\omega_{i(m)} \in P_0^{i(m)}$ from P-1, for some $l(< m)$-stage (1) $\omega_{i(m)} \in P_l^{i(m)}$, (2) $\omega_{i(m)} \in P_l^{i(l)}$, (3) $\omega_{i(l)} \in P_l^{i(m)}$, and $i(m) \neq i(l)$ hold.

From P-2 at $l(< m)$-stage, $\omega_{i(m)} \in P_l^h \Rightarrow P_l^{i(m)} \subseteq P_l^h$. Since (3) $\omega_{i(l)} \in P_l^{i(m)}$, $\omega_{i(l)} \in P_l^h$. Hence, $\omega_{i(m)} \in P_l^h \Rightarrow \omega_{i(l)} \in P_l^h$.

Since $\omega_{i(m)} \in P_l^h \Rightarrow \omega_{i(l)} \in P_l^h$ and (2) $\omega_{i(m)} \in P_l^{i(l)}$, $I_k^{i(m)} = 1$ or 0 for all $k(> l)$-stage.

This contradicts $\omega_{i(m)} \in \max_i I_m^i \geq 2$. Therefore $\max_i I_m^i \geq 2 \Rightarrow P_m^{i(m)} \neq \emptyset$. ∎

LEMMA 8.3. *P-2 holds until* $\max_i I_m^i \geq 2$.

Proof. We show lemma 8.3 by induction. When $m = 0$, P-2 at 0-stage holds from P-2. We show P-2 at s-stage ($s = 1, ..., k$) $\Rightarrow$ P-2 at $k+1$-stage.

Note that $P_{k+1}^h = P_k^h \setminus P_k^{i(k)}$ (if $h \in H_k$) and $P_{k+1}^h = P_k^h$ (if $h \notin H_k$) where $H_k = \{h \in \{1, ..., t\} : \omega_{i(k)} \in P_k^h \text{ and } h \neq i(k)\}$.

CASE1: $j, j' \in H_k$

$P_k^j \setminus P_k^{i(k)} \subseteq P_k^{j'} \setminus P_k^{i(k)}$. Hence $P_{k+1}^j \subseteq P_{k+1}^{j'}$.

CASE2: $j, j' \notin H_k$

From P-2 at k-stage, $P_k^j \subseteq P_k^{j'}$. Hence $P_{k+1}^j \subseteq P_{k+1}^{j'}$.

CASE3: $j \in H_k$, $j' \notin H_k$

From P-2 at k-stage, $P_k^j \subseteq P_k^{j'}$. Hence $P_{k+1}^j \subseteq P_{k+1}^{j'}$.

CASE4-1: $j \notin H_k, j' \in H_k$ and $j \neq i(k)$

From P-2 at k-stage, $P_k^j \subseteq P_k^{j'}$. We have to show $P_k^j \subseteq P_k^{j'} \setminus P_k^{i(k)}$.

When $P_k^j \bigcap P_k^{i(k)} = \emptyset$, $P_k^j \subseteq P_k^{j'} \setminus P_k^{i(k)}$ holds (And lemma 8.3 holds.).

The case $P_k^j \bigcap P_k^{i(k)} \neq \emptyset$ is the matter.

Now suppose that there is some ω_o such that $\omega_o \in P_k^j \bigcap P_k^{i(k)}$. (We shall derive a contradiction from this assumption.) Then $\omega_o \in P_k^j$. Since $j \notin H_k, j' \in H_k$ and $j \neq i(k)$, $\omega_{i(k)} \notin P_k^j$.

Then, from the definition of $i(k)$ ($i(k) \in \arg\max_i I_k^i$), there is some $h \in \{1, 2, ..., t\}$ such that $\omega_{i(k)} \in P_k^h$ and $\omega_o \notin P_k^h$. For this h, $P_k^{i(k)} \subseteq P_k^h$ holds, since P-2 at k-stage and $\omega_{i(k)} \in P_k^i$ hold. But, the assumption $\omega_o \in P_k^j \bigcap P_k^{i(k)}$ means $\omega_o \in P_k^{i(k)}$, that is, $\omega_o \in P_k^h$. This is a contradiction.

CASE4-2: $j \notin H_k, j' \in H_k$ and $j = i(k)$

We have to show $\omega_{i(k)} \in P_{k+1}^{j'} \setminus P_k^{i(k)} \Rightarrow P_k^{i(k)} \subseteq P_k^{j'} \setminus P_k^{i(k)}$.

When $\omega_{i(k)} \in P_k^{i(k)}$, $\omega_{i(k)} \notin P_{k+1}^{j'} \setminus P_k^{i(k)}$. And it is impossible that $\omega_{i(k)} \notin P_k^{i(k)}$ holds. the proof is same as the proof of lemma 8.1.

For all $h \in H_m$, $\omega_{i(m)} \in P_m^h$. P-2 at m-stage means $P_m^{i(m)} \subseteq P_m^h$. Therefore lemma 8.3 holds. ∎

Lemma 8.4. *If P-2 holds at m-stage, then* $P_m^{i(m)} \subseteq P_m^h$ *hold for all* $h \in \{h \in \{1, ..., t\} : \omega^{i(m)} \in P_m^h$ *and* $h \neq i(m)\}$.

Proof. We show that if P-2 holds at m-stage, then $\max_i I_m^i \geq 2 \Rightarrow P_m^{i(m)} \subseteq P_m^h$ for $h \in H_m$, where $H_m = \{h : \omega_{i(m)} \in P_m^h$ and $h \neq i(m)\}$. For all $h \in H_m$, $\omega_{i(m)} \in P_m^h$. From P-2 at m-stage, $P_m^{i(m)} \subseteq P_m^h$. Hence lemma 8.4 holds. ▌

Proof (Agreeing Theorem). From lemmata, we can get a partition, $\{P_{m^*}^1, ..., P_{m^*}^t\}$. Since $\{P_{m^*}^1, ..., P_{m^*}^t\}$ is a partition, the rest of the proof follows Aumann(1976). ▌

REFERENCES

1. Aumann, R.: "Agreeing to Disagree," *The Annals of Statistics*, **4**, (1976) 1236-1239.
2. Brandenburger, A.: "Knowledge and Equilibrium in Games," *Journal of Economic Perspectives*, **6**, (1992) 83-101.
3. Chellas, B.: *Modal Logic: An Introduction*, Cambridge University Press, Cambridge, (1980)
4. Dekel, E. and Gul, F.: "Rationality and knowledge," *Advances in economic theory 7th.*, (1997) 89-72.
5. Geanakoplos, J.: "Common Knowledge," *Journal of Economic Perspectives*, **6**, (1992) 53-82.
6. Geanakoplos, J. and Polemachakis, H.: "We can't disagree forever," *Journal of Economic Theory*, **28**, (1982) 192-200.
7. Hughes, G., and Cresswell, M.: *A New Introduction to Modal Logic*, Routledge, London and New York, (1996)
8. Matsuhisa, T. and Kamiyama, K.: "Lattice structure of knowledge and agreeing to disagree," *Journal of Mathematical Economics*, **27**, (1997) 389-410.
9. Milgrom, P.: "An Axiomatic Characterization of Common Knowledge," *Econometrica*, **49**, (1981) 219-222.
10. Osborne, M., and Rubinstein, A.: *A Course in Game Theory*, MIT Press, Princeton, (1994)
11. Rubinstein, A.: *Modeling Bounded Rationality*, MIT Press, Princeton, (1998)
12. Samet, D.: "Ignoring Ignorance and Agreeing to Disagree," *Journal of Economic Theory*, **52**, (1990) 190-207.