数理解析研究所講究録
1166 巻 2000 年 170-177

Knowledge Structure in Decision Theory

Takashi Matsuhisa

Department of Liberal Arts and Sciences, Ibaraki National College of Technology
866 Nakane, Hitachinaka-shi, Ibaraki-ken 312-8508, Japan
E-mail address mathisa@ge.ibaraki-ct.ac.jp

ABSTRACT. The logic for 'utility maximizers' L^{um} is proposed which is an extension of a system of modal logic for many players. The sound models according to L^{um} are given in terms of game theory. It is shown for the models that two utility maximizing players must take the same actions if they mutually believe that each takes a dominant action, even when they have different informations. We remark that the logic L^{um} have the finite model property.

1. INTRODUCTION

Recently researchers in such diverse fields as Economics, Linguistics, Artificial Intelligence, and Computer Sciences have become interested in reasoning about knowledge. There are pragmatic concerns about the relationship between knowledge and actions, and there are also concerns about the complexity of computing knowledge. Of most interested is the emphasis on considering situation involving the knowledge of a group of players rather than that of a single player.

The purpose of this talk is to develop a theory of decision making among a group of players under uncertainty based on modal logic rather than on probability measures (as in the standard theory.) It is the theory of ' maximizing utility' in which all players are utility maximizers; that is, each player takes the actions being best response to his utility. In the theory there is a 'logic of belief' in which a given proposition is either believed, or disbelieved, or neither believed nor disbelieved. It is

---

Date: July 5, 2000.

1991 *Mathematics Subject Classification.* Primary: 90D35; Secondary: 68Q45, *Journal of Economic Literature Classification*: C62, C78.

Key words and phrases. Logic of belief, dominant action, logic for utility maximizers, finite model property.

This is a discussion paper based on the author's presentation in the R.I.M.S. Symposium on 'Languages, Algebra and Computation,' Kyoto University, March 2000. This paper is preliminary and the final form will be published elsewhere.

noted that there are another specific kinds of theories of decision making: Among others the theory of 'agreeing to disagree' is of most interested, in which all players cannot agree to disagree on their predictions about an event (c.f. Aumann [1], Bacharach [2] and Matsuhisa [6].)

This talk will proceed as follows: In section 2 I review the standard models of belief for a system of modal logic for multi-players and the utility theory with belief. Section 3 proposes a system for 'utility maximizers' that is an extension of a system of modal logic. The sound models according to the system are given in terms of the utility theory with belief. In section 4 I show that two utility maximizers in each sound model must take the same actions if they mutually believe that each takes a dominant action, even when they have different informations. Example (Prisoner Dilemma) demonstrates that they does not always take the same actions in case that each player simply believes that he takes a dominant action. Section 5 presents the logic for 'utility maximizers' L^{um} and remarks that the logic has the finite model property.

2. The Model

Let Ω be a non-empty set called a *state-space*, N a set of two *players* $1, 2$, and let 2^{Ω} be the family of all subsets of Ω. Each member of 2^{Ω} is called an *event* and each element of Ω called a *state*.

2.1. Information and Belief (Binmore [3]).

An *information structure* $(P_i)_{i \in N}$ is a class of mappings P_i of Ω into 2^{Ω}. Given our interpretation, an player i for whom $P_i(\omega) \subseteq E$ knows, in the state ω, that some state in the event E has occurred. In this case we say that in the state ω the player i believes E.

An i's *belief operator* is an operator B_i on 2^{Ω} such that B_iE is the set of states of Ω in which i believes that E has occurred; that is,

$$B_iE = \{\omega \in \Omega | P_i(\omega) \subseteqq E\}. \tag{1}$$

We note that the i's belief operator satisfies the following properties: For every E, F of 2^{Ω},

N: $B_i\Omega = \Omega$ and $B_i\emptyset = \emptyset$;

K: $B_i(E \cap F) = B_iE \cap B_iF$;

The set $P_i(\omega)$ will be interpreted as the set of all the states of nature that i believes to be possible at ω, and B_iE will be interpreted as the set of states of nature for which i believes E to be possible. We will therefore call P_i an i's *possibility operator* on Ω and also will call $P_i(\omega)$ the i's *possibility set* at ω. An event E is said to be an i's *truism* if $E \subseteqq B_iE$

We should note that the information structure P_i is uniquely determined by the belief operator B_i such that $P_i(\omega) = \bigcap_{\omega \in B_i E} E$.

2.2. **Utility and Belief.** By a *utility theory for two player* we mean a triple $\langle N, (A_i)_{i=1,2}, (V_i)_{i=1,2} \rangle$ with the following structure and interpretations: N is a of players $\{1, 2\}$, A_i is a finite set of i's *available actions* (or i's pure strategies) and V_i is an i's *utility-function* of $A_1 \times A_2$ into $\mathbb{R}$. We denote by A_{-i} the set A_j for $j \neq i$.

An action a in A_i is called *dominant* for i if $V_i(a_i, a_{-i}) \geq V_i(b, a_{-i})$ for all $b \in A_i$ and for all $a_{-i} \in A_{-i}$.

Example 1. (Prisoners' dilemma:) Let A be a set of two available actions $\{a_1, a_2\}$ which is common for players $1, 2$. The utility functions (V_1, V_2) are given by

Player 2

Player 1

(V_1, V_2)	a_1	a_2
a_1	1, 1	3, 0
a_2	0, 3	2, 2

In this example we can plainly observe that the action a_1 is dominant for each player $1, 2$. □

Definition 1. By a *utility theory with belief* we mean a tuple $\mathcal{V} = \langle \Omega, (P_i)_{i=1,2}, (A_i)_{i=1,2}, (V_i)_{i=1,2} \rangle$ with the following structures:

- Ω is a state-space;
- $P_i : \Omega \longrightarrow 2^\Omega$ is an i's information function;
- A_i is a set of available actions for player i;
- $V_i : A_1 \times A_2 \times \Omega \longrightarrow \mathbb{R}$ is an i's utility function with the property that $V_i(\cdot, \cdot; \omega)$ is injective on $A_1 \times A_2$ for each $\omega \in \Omega$.

Example 2. A tuple $\mathcal{V} = \langle \Omega, (P_i)_{i=1,2}, (A_i)_{i=1,2}, (V_i)_{i=1,2} \rangle$ given as below is a utility theory with belief:

- $\Omega = \{\omega_1 . \omega_2\}$
- $P_i : \Omega \longrightarrow 2^\Omega$ is given by $P_1(\omega_1) := \{\omega_1\}, P_1(\omega_2) := \{\omega_2\}, P_2(\omega_1) := \{\omega_2\}$, and $P_2(\omega_2) := \{\omega_1\}$;
- $A_1 = A_2 = \{a_1 . a_2\}$;
- $V_i : A_1 \times A_2 \times \Omega \longrightarrow \mathbb{R}$ is defined by

Player 2

Player 1

$(V_1(\cdot,\cdot;\omega_1), V_2(\cdot,\cdot;\omega_1))$

	a_1	a_2
a_1	1, 1	3, 0
a_2	0, 3	2, 2

$(V_1(\cdot,\cdot;\omega_2), V_2(\cdot,\cdot;\omega_2))$

	a_1	a_2
a_1	2, 2	0, 3
a_2	3, 0	1, 1

□

3. System

Let us consider a system of multi-modal logic as follows.

3.1. **Syntax.** The *language* of the system consists of the symbols, the terms and the sentences as follows:

- *Symbols*:

Non-modal operators :	$\neg, \rightarrow, \wedge, \top, \cdots;$	
Modal operators :	$\Box_1, \Box_2, \Box_E;$	
Variables:	$\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$	(Actions for players 1, 2)
Predicates:	$=$	(Equality on the actions)
	$\mathrm{dom}_1, \mathrm{dom}_2.$	(Dominant actions)

- *Terms* and *Sentences*:

  (i) The variables are *terms*;

  (ii) If s and t are two terms then $\mathrm{s} = \mathrm{t}$ and $\mathrm{dom}_1(\mathrm{s})$, $\mathrm{dom}_2(\mathrm{t})$ are *atomic* sentences;

  The *sentences* of the language form the least set containing all *atomic* sentences $\mathbf{P}_m (m = 0, 1, 2, \dots)$ closed under the following operations:

  - nullary operators for *falsity* $\bot$ and for *truth* $\top$;
  - unary and binary syntactic operations for *negation* $\neg$, *conditionality* $\rightarrow$ and *conjunction* $\wedge$, respectively;
  - three unary operations for *modality* $\Box_1, \Box_2, \Box_E$.

Other such operations are defined in terms of those in usual ways.

The intended interpretation of $\Box_i\varphi$ is the sentence that 'player i believes a sentence φ,' $\Box_E\varphi$ as 'everybody believes φ.' The sentence $\mathrm{dom}_i(\mathbf{a}_k)$ is interpreted as '$\mathbf{a}_k$ is a dominant action for i

3.2. **System for utility maximizers.** By this we mean a set of sentences, denoted by L,

- containing a set of all *tautologies* and closed under *substitution* and *modus ponens*;

- has the following *inference rules* and *axioms*:

(N) $\quad \Box_* \top \quad$ for $* = 1, 2, E$

(RE$_\Box$) $\quad \dfrac{\varphi \longleftrightarrow \psi}{\Box_*\varphi \longleftrightarrow \Box_*\psi} \quad$ for $* = 1, 2, E$;

(Def$\Box_E$) $\quad \Box_E\varphi \longleftrightarrow \Box_1\varphi \wedge \Box_2\varphi$;

(RE$_{\mathbf{dom}}$) $\quad \dfrac{\mathbf{a}_k = \mathbf{a}_l}{\mathrm{dom}_i(\mathbf{a}_k) \longleftrightarrow \mathrm{dom}_i(\mathbf{a}_l)} \quad$ for $i = 1, 2$.

A sentence φ is *provable* in L, denoted by $\vdash_L \varphi$, if $\varphi \in L$.

3.3. **Semantics.** A *model* $\mathcal{M}$ for a system L is a tuple $\langle \mathcal{V}, v_{\mathcal{M}}, \pi, \models \rangle$ with the following structures:

- $\mathcal{V} = \langle \Omega, (P_i)_{i=1,2}, (A_i)_{i=1,2}, (V_i)_{i=1,2} \rangle$ is a utility theory with belief such that

  (i) $A_1 = A_2 = A := \{a_1, a_2, \cdots, a_n\}$;
  (ii) $V_i(\cdot, \cdot; \omega)$ is injective on $A \times A$ for each $\omega \in \Omega$;

- $v_{\mathcal{M}} : \{\mathbf{a}_k \mid k = 1, 2, \cdots, n\} \to A$ is a *valuation* of variables into available actions;
- $\pi : \{\mathbf{P}_m \mid m = 0, 1, 2, \dots\} \times \Omega \longrightarrow \{true, false\}$ is a truth assignment such that

  (i) $\pi(\mathbf{a}_k = \mathbf{a}_l, \omega) = true$ if and only if $v_{\mathcal{M}}(\mathbf{a}_k) = v_{\mathcal{M}}(\mathbf{a}_l)$;
  (ii) $\pi(\mathrm{dom}_1(\mathbf{a}_k), \omega) = true$ if and only if
  $V_1(v_{\mathcal{M}}(\mathbf{a}_k), b; \omega) \geq V_1(c, b; \omega)$ for all $b, c \in A$;
  $\pi(\mathrm{dom}_2(\mathbf{a}_l), \omega) = true$ if and only if
  $V_i(b, v_{\mathcal{M}}(\mathbf{a}_l); \omega) \geq V_i(a, b; \omega)$ for all $b, c \in A$.

- *Truth* $\models^{\mathcal{M}}_\omega \varphi$ at ω in $\mathcal{M}$ is inductively defined as follows:
  (i) $\models^{\mathcal{M}}_\omega \mathrm{v}$ if and only if
  $\pi(\mathrm{v}, \omega) = true$ for each atomic sentence v;
  (ii) $\models^{\mathcal{M}}_\omega \top$;
  (iii) $\models^{\mathcal{M}}_\omega \varphi \longrightarrow \psi$ if and only if $\models^{\mathcal{M}}_\omega \varphi$ implies $\models^{\mathcal{M}}_\omega \psi$;
  (iv) $\models^{\mathcal{M}}_\omega \Box_i\varphi$ if and only if
  $\emptyset \neq P_i(\omega) \subseteq ||\varphi||^{\mathcal{M}} \underset{\mathrm{def}}{:=} \{\xi \in \Omega | \models^{\mathcal{M}}_\xi \varphi\}$, for $i = 1, 2$;
  (v) $\models^{\mathcal{M}}_\omega \Box_E\varphi$ if and only if
  $\emptyset \neq P_E(\omega) \underset{\mathrm{def}}{:=} \bigcap_{i=1,2} P_i(\omega) \subseteq ||\varphi||^{\mathcal{M}}$.

4. Uniqueness of Dominant actions

4.1. Let $\boldsymbol{M}_L$ be the class of all models for a system L and $\boldsymbol{M}_L^{sym}$ the subclass of $\mathcal{M} = \langle .., (V_i)_{i=1,2}, ..\rangle$ with $V_1(a,b;\omega) = V_2(b,a;\omega)$ for all $a, b \in A$. We denote $\models_{\boldsymbol{M}_L^{sym}} \varphi$ when $\models_\omega^{\mathcal{M}} \varphi$ for all $\mathcal{M} \in \boldsymbol{M}_L^{sym}$ and for all $\omega \in \mathcal{M}$.

4.2. We will show the uniqueness theorem on dominant actions:

Proposition 1. (Matsuhisa and Hirase [7]:) For a system for utility maximizers L we obtain that

$$\models_{\boldsymbol{M}_L^{sym}} \Box_E(\mathrm{dom}_1(\mathbf{a}_k) \wedge \mathrm{dom}_2(\mathbf{a}_l)) \longrightarrow \mathbf{a}_k = \mathbf{a}_l.$$

That is: If all players believe that each takes his dominant action then they cannot agree to disagree.

Proof. Let $\mathcal{M} \in \mathbf{M}^{sym}$. Set $d_i : 2^\Omega \to 2^A$ by

$$d_i(E) = \{v_{\mathcal{M}}(\mathbf{a}) \in A \mid E \subseteq ||\mathrm{dom}_i(\mathbf{a})||^{\mathcal{M}}\}.$$

We can plainly verify the three properties:

(1) $d_i(E) \subseteq d_i(F)$ if $E \supseteq F$. (by definition of d_i)
(2) $|d_i(E)| \leq 1$ if $E \neq \emptyset$, (because $V_i(\cdot,\cdot;\omega)$ is injective.)
(3) $d_1(E) = d_2(E)$, (because V_1, V_2 are symmetric.)

Suppose $\models_\omega^{\mathcal{M}} \Box_E(\mathrm{dom}_1(\mathbf{a}_k) \wedge \mathrm{dom}_2(\mathbf{a}_l))$. Then we obtain

$$\models_\omega^{\mathcal{M}} \Box_i \mathrm{dom}_i(\mathbf{a}_m) \quad \text{for } m = k, l.$$

It follows from the properties (1), (2), (3) that

$$\begin{aligned}\{v_{\mathcal{M}}(\mathbf{a}_k)\} &\underset{(1)}{=} d_1(P_1(\omega)) \underset{(1),\,(2)}{=} d_1(P_E(\omega)) \\ &\underset{(3)}{=} d_2(P_E(\omega)) \underset{(1),\,(2)}{=} d_2(P_2(\omega)) \underset{(1)}{=} \{v_{\mathcal{M}}(\mathbf{a}_l)\}.\end{aligned}$$

Thus we obtain that $v_{\mathcal{M}}(\mathbf{a}_k) = v_{\mathcal{M}}(\mathbf{a}_l)$, and so $\models_\omega^{\mathcal{M}} \mathbf{a}_k = \mathbf{a}_l$.

□

4.3. **Remarks.**

(i) A model $\mathcal{M}$ is actually a model of belief because it does not satisfy the axiom:

T: $B_i(F) \subseteq F.$

(See Example 2.)

(ii) There is no role of common-belief in Proposition 1.

(iii) The weak statement is not true that

$$\models_{M_L^{sym}} \Box_1 \mathrm{dom}_1(\mathbf{a}_1) \wedge \Box_2 \mathrm{dom}_2(\mathbf{a}_2) \longrightarrow \mathbf{a}_1 = \mathbf{a}_2.$$

In fact, we can plainly observe that Example 2 gives its counter example.

5. Theorems

5.1. **Logic for utility maximizers.** By this we mean the least extension of L, denoted by L^{um}, that contains the axiom

$$\text{(UDA)} \qquad \Box_E(\mathrm{dom}_1(\mathbf{a}_k) \wedge \mathrm{dom}_2(\mathbf{a}_l)) \longrightarrow \mathbf{a}_k = \mathbf{a}_l.$$

It immediately follows from Proposition 1 that

Theorem 1. *The logic L^{um} is sound with respect to $\boldsymbol{M}^{sym}_{L^{um}}$: i.e.,*

$$\vdash_{L^{um}} \varphi \quad \Longrightarrow \quad \models_{\boldsymbol{M}^{sym}_{L^{um}}} \varphi.$$

5.2. **Completeness.** By the similar argument concerning about the 'canonical model' (c.f. Chapter 5 in Chellas [4]) we can prove that:

Theorem 2. *The system L^{um} is complete with respect to $\boldsymbol{M}^{sym}_{L^{um}}$: i.e.,*

$$\vdash_{L^{um}} \varphi \quad \Longleftarrow \quad \models_{\boldsymbol{M}^{sym}_{L^{um}}} \varphi.$$

□

5.3. **Finite model property.** We say a model for L to be *finite* if its state-space is a finite set. Let $\boldsymbol{M}^{sym}_{L,FIN}$ denote the subclass of all finite models in $\boldsymbol{M}^{sym}_{L}$. Furthermore we can prove that:

Theorem 3. *The system L^{um} has finite model property; i.e.,*

$$\vdash_{L^{um}} \varphi \quad \Longleftrightarrow \quad \models_{\boldsymbol{M}^{sym}_{L^{um},FIN}} \varphi.$$

□

We will give the detail proofs of Theorems 2 and 3 in the future paper (Matsuhisa and Hirase [7]) with further discussions.

REFERENCES

[1] R.J. Aumann, *Agreeing to disagree*, Annals of Statistics **4** (1976), 1236–239.
[2] M. Bacharach, *Some extensions of a claim of Aumann in an axiomatic model of knowledge*, Journal of Economic Theory **37** (1985), 167–190.
[3] Ken Binmore, *Fun and Games.* D. C. Heath and Company, Lexington, Massachusetts USA, (1990), xxx+642pp.
[4] B. F. Chellas, *Modal logic: An introduction.* Cambridge University Press, Cambridge, London, New York, New Rochelle, Melbourne, Sydoney, 1980.
[5] R. Fagin, J.Y. Halpern, Y. Moses and M.Y. Vardi, *Reasoning about Knowledge.* The MIT Press, Cambridge, Massachusetts, London, England, 1995.
[6] T. Matsuhisa, *Logic of agreeing to disagree,* Working paper, Ibaraki National College of Technology (1999), in submitting
[7] T. Matsuhisa and K. Hirase, in preparation.