<table>
<thead>
<tr>
<th>項目</th>
<th>内容</th>
</tr>
</thead>
<tbody>
<tr>
<td>タイトル</td>
<td>知識構造在決断理論 (代数システム、形式言語と計算)</td>
</tr>
<tr>
<td>著者</td>
<td>Matsuhisa, Takashi</td>
</tr>
<tr>
<td>引用</td>
<td>数理解析研究所講究録 (2000), 1166: 170-177</td>
</tr>
<tr>
<td>発行日</td>
<td>2000-08</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/64342">http://hdl.handle.net/2433/64342</a></td>
</tr>
<tr>
<td>言語</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>版</td>
<td>publisher</td>
</tr>
</tbody>
</table>

京都大学
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
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 $\Omega$ be a non-empty set called a state-space, $N$ a set of two players 1, 2, and let $2^{\Omega}$ be the family of all subsets of $\Omega$. Each member of $2^{\Omega}$ is called an event and each element of $\Omega$ 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 $\Omega$ into $2^{\Omega}$. Given our interpretation, an player $i$ for whom $P_{i}(\omega) \subseteq E$ knows, in the state $\omega$, that some state in the event $E$ has occurred. In this case we say that in the state $\omega$ the player $i$ believes $E$.

An $i$'s belief operator is an operator $B_{i}$ on $2^{\Omega}$ such that $B_{i}E$ is the set of states of $\Omega$ in which $i$ believes that $E$ has occurred; that is,

$$B_{i}E = \{ \omega \in \Omega | P_{i}(\omega) \subseteq E \}. \tag{1}$$

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

- **N:** $B_{i}\emptyset = \emptyset$ and $B_{i}\emptyset = \emptyset$;
- **K:** $B_{i}(E \cap F) = B_{i}E \cap B_{i}F$;

The set $P_{i}(\omega)$ will be interpreted as the set of all the states of nature that $i$ believes to be possible at $\omega$, and $B_{i}E$ 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 $\Omega$ and also will call $P_{i}(\omega)$ the $i$'s possibility set at $\omega$. An event $E$ is said to be an $i$'s truism if $E \subseteq B_{i}E$.
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

<table>
<thead>
<tr>
<th>Player 1</th>
<th>$V_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>1, 1</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0, 3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Player 2</th>
<th>$V_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>3, 0</td>
</tr>
<tr>
<td>$a_2$</td>
<td>2, 2</td>
</tr>
</tbody>
</table>

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 $V = \langle \Omega, (P_i)_{i=1,2}, (A_i)_{i=1,2}, (V_i)_{i=1,2}\rangle$ with the following structures:

- $\Omega$ is a state-space;
- $P_i : \Omega \rightarrow 2^\Omega$ is an $i$'s information function;
- $A_i$ is a set of available actions for player $i$;
- $V_i : A_1 \times A_2 \times \Omega \rightarrow \mathbb{R}$ 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 $V = \langle \Omega, (P_i)_{i=1,2}, (A_i)_{i=1,2}, (V_i)_{i=1,2}\rangle$ given as below is a utility theory with belief:

- $\Omega = \{\omega_1, \omega_2\}$
- $P_1 : \Omega \rightarrow 2^\Omega$ is given by $P_1(\omega_1) := \{\omega_1\}, P_1(\omega_2) := \{\omega_2\}, P_2(\omega_1) := \{\omega_2\}$, and $P_2(\omega_2) := \{\omega_1\}$;
- $A_1 = A_2 = \{a_1, a_2\}$;
- $V_i : A_1 \times A_2 \times \Omega \rightarrow \mathbb{R}$ is defined by
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: $a_1, a_2, \cdots, a_n$ (Actions for players 1, 2)
  - Predicates: $=$ (Equality on the actions) $\mathrm{dom}_1, \mathrm{dom}_2$. (Dominant actions)

- **Terms and Sentences:**
  1. The variables are terms;
  2. If $s$ and $t$ are two terms then $s = t$ and $\mathrm{dom}_1(s), \mathrm{dom}_2(t)$ are atomic sentences;

The sentences of the language form the least set containing all atomic sentences $P_m (m = 0, 1, 2, \ldots)$ closed under the following operations:

- nullary operators for falsity $\perp$ 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 $\varphi$,' $\Box_E \varphi$ as 'everybody believes $\varphi$.' The sentence $\mathrm{dom}_i(a_k)$ is interpreted as '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:

\[(N) \quad \square, T \quad \text{for } * = 1, 2, E\]

\[(\text{RE}_\square) \quad \varphi \leftrightarrow \psi \quad \square \varphi \leftrightarrow \square \psi \quad \text{for } * = 1, 2, E;\]

\[(\text{Def}_E) \quad \square_E \varphi \leftrightarrow \square_1 \varphi \land \square_2 \varphi;\]

\[(\text{RE}_{\text{dom}}) \quad \frac{a_k = a_l}{\text{dom}_i(a_k) \rightarrow \text{dom}_i(a_l)} \quad \text{for } i = 1, 2.\]

A sentence \( \varphi \) 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 \( \langle \mathcal{V}, v, \pi, \models \rangle \) 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) \( A_1 = A_2 = A := \{a_1, a_2, \ldots, a_n\} \);
  (ii) \( V_i(\cdot, \cdot; \omega) \) is injective on \( A \times A \) for each \( \omega \in \Omega \);

• \( v_M : \{a_k \mid k = 1, 2, \ldots, n\} \rightarrow A \) is a valuation of variables into available actions;

• \( \pi : \{P_m \mid m = 0, 1, 2, \ldots\} \times \Omega \rightarrow \{true, false\} \) is a truth assignment such that
  (i) \( \pi(a_k = a_l, \omega) = true \) if and only if \( v_M(a_k) = v_M(a_l) \);
  (ii) \( \pi(\text{dom}_1(a_k), \omega) = true \) if and only if
       \( V_1(v_M(a_k), b; \omega) \geq V_1(c, b; \omega) \) for all \( b, c \in A \);
       \( \pi(\text{dom}_2(a_k), \omega) = true \) if and only if
       \( V_2(b, v_M(a_k); \omega) \geq V_2(a, b; \omega) \) for all \( b, c \in A \).

• Truth \( \models^\mathcal{M}_\omega \varphi \) at \( \omega \) in \( \mathcal{M} \) is inductively defined as follows:
  (i) \( \models^\mathcal{M}_\omega v \) if and only if \( \pi(v, \omega) = true \) for each atomic sentence \( v \);
  (ii) \( \models^\mathcal{M}_\omega T \);
  (iii) \( \models^\mathcal{M}_\omega \varphi \rightarrow \psi \) if and only if \( \models^\mathcal{M}_\omega \varphi \) implies \( \models^\mathcal{M}_\omega \psi \);
  (iv) \( \models^\mathcal{M}_\omega \square \varphi \) if and only if
       \( \emptyset \neq P_i(\omega) \subseteq ||\varphi||^\mathcal{M}_\omega := \{\xi \in \Omega \mid \models^\mathcal{M}_\xi \varphi\} \), for \( i = 1, 2 \);
  (v) \( \models^\mathcal{M}_\omega \square_E \varphi \) if and only if
       \( \emptyset \neq P_E(\omega) := \bigcap_{i=1,2} P_i(\omega) \subseteq ||\varphi||^\mathcal{M}_\omega \).
4. Uniqueness of Dominant Actions

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

\[ |=_{M_L^{sym}} \Box_E (\text{dom}_1(a_k) \land \text{dom}_2(a_l)) \rightarrow a_k = a_l. \]

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

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

\[ d_i(E) = \{ v_\mathcal{M}(a) \in A \mid E \subseteq \|\text{dom}_i(a)\|^{\mathcal{M}} \}. \]

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 $|=_{\omega} \Box_E (\text{dom}_1(a_k) \land \text{dom}_2(a_l))$. Then we obtain

\[ |=_{\omega} \Box_i \text{dom}_i(a_m) \text{ for } m = k, l. \]

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

\[ \{ v_\mathcal{M}(a_k) \} = d_1(P_1(\omega)) \overset{(1)}{=} d_1(P_E(\omega)) \overset{(1),(2)}{=} d_2(P_E(\omega)) \overset{(3)}{=} d_2(P_2(\omega)) \overset{(1)}{=} \{ v_\mathcal{M}(a_l) \}. \]

Thus we obtain that $v_\mathcal{M}(a_k) = v_\mathcal{M}(a_l)$, and so $|=_{\omega} a_k = a_l$. \qed

4.3. Remarks.

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

\[ T: \quad B_i(F) \subseteq F. \]

(See Example 2.)

(ii) There is no role of common-belief in Proposition 1.
The weak statement is not true that
\[\models_{M_{L}^{ym}} \Box_1 \text{dom}_1(a_1) \land \Box_2 \text{dom}_2(a_2) \rightarrow a_1 = a_2.\]
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

\[(UDA) \quad \Box_E(\text{dom}_1(a_k) \land \text{dom}_2(a_l)) \rightarrow a_k = a_l.\]

It immediately follows from Proposition 1 that

**Theorem 1.** The logic \( L^{um} \) is sound with respect to \( M_{L^{um}}^{sym} \): i.e.,
\[\vdash_{L^{um}} \varphi \quad \Rightarrow \quad \models_{M_{L^{um}}^{sym}} \varphi.\]

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 \( M_{L^{um}}^{sym} \): i.e.,
\[\vdash_{L^{um}} \varphi \quad \Leftrightarrow \quad \models_{M_{L^{um}}^{sym}} \varphi.\]

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

**Theorem 3.** The system \( L^{um} \) has finite model property; i.e.,
\[\vdash_{L^{um}} \varphi \quad \Leftrightarrow \quad \models_{M_{L^{um},FIN}^{sym}} \varphi.\]

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