No. 874

Introduction to Epistemic Logics and their Game Theoretic Applications

by

Mamoru Kaneko

July 2000

Introduction to Epistemic Logics and their Game Theoretic Applications (Int0009.tex) *†

Mamoru Kaneko‡

July 13, 2000

Abstract

This paper is written as an introduction to epistemic logics and their game theoretical applications. It starts with both semantics and syntax of classical logic, and goes to the Hilbert-style proof-theory and Kripke-style model theory of epistemic logics. We give also a brief explanation of the Gentzen-style proof theory. In these theories, we discuss individual decision making in some simple game examples. In particular, we will discuss the distinction between beliefs and knowledge, and how false beliefs play roles in game theoretical decision making. Finally, we discuss extensions of epistemic logics to incorporate common knowledge. In the extension, we discuss also false beliefs on common knowledge.

1. Introduction

1.1. Aim and Some Basic Notions in Logic

This paper is written for economists/game theorists as an introduction to epistemic logics and their game theoretical applications. We believe that epistemic intra-interpersonal introspections and interactions play key roles in social behavior of people, and also that in such interactions, symbolic expressions and manipulations are crucial. For these beliefs, we adopt the research strategy to take and use basic concepts and results developed in mathematical logic.

The declaration of our ultimate aim may help readers to understand our research attitude. Since mathematical logic may be regarded as pursuing foundations of mathematics, the logic approach to game theory may be regarded also as pursuing foundations for extant game theories. Contrary to this view, we have little intentions of doing so. Although our investigations target also social interactions consisting of people and their decision making, we take the attitude of regarding problems of finite

*I thank Jeff Kline for many detailed comments on earlier drafts, and Nobu-Yuki Suzuki for helpful discussions on subjects covered in this paper. I am also grateful to the members, A. Kajii, A. Matsui, K. Tatamitani, S. Turbull, T. Ui, of the logic study group here for helpful discussions and comments.

†I am partially supported by Grant-in-Aids for Scientific Research No.106330003, Ministry of Education, Science and Culture.

‡Institute of Policy and Planning Sciences, University of Tsukuba, Ibaraki 305-8573, Japan (kaneko@shako.sk.tsukuba.ac.jp)

nature as central in such investigations. This research attitude leads us to deviate from extant game theories having strong tendencies of transcendentalities. For our research program, we should keep in mind the warning by N. Whitehead: "It is characteristic of a science in its earlier stages ... to be both ambitiously profound in its aims and trivial in its handlings of details".

Because of methodological differences in economics and logic, there are many entrance barriers which economists/game theorists may encounter. These barriers may well inhibit future economic and game theoretic research, but some constituents of them may be important for future research itself. Despite of our ambition, practically we need to follow the standard development of mathematical logic. Now, it is timely to give a systematic introduction to mathematical logic, specifically, epistemic logics, with some illustrations of game theoretical applications. We hope that this introduction will lead further developments of the logical approach to game theory.

Economics and game theory have the tradition that their mathematical methodology is based primarily on analysis such as topology, functional analysis and probability theory. Also in logic, we can treat these mathematical fields, but what we emphasize by the logical approach is the basic constructions of logic rather than direct applications of extant results in logic. There are various unique constituents in logic that are not found in other mathematical fields. One purpose of this paper is to introduce such unique constituents to economists/game theorists, and the other is to discuss, in particular, epistemic logics together with their game theoretical applications. In this introduction, we mention several pairs of basic concepts unique to logic, and to epistemic logic in particular, which would help the reader understand the subjects better.

The very basic starting point of logic is the separation between symbolic expressions and their intended meanings. When we target human thinking seriously, this separation is unavoidable. It is stated in the terminology of logic as:

A1: Syntax VS Semantics.

As a syntactical notion, we define a *formula* as a symbolic expression based on given primitive symbols. As a semantical notion, we define a truth valuation of such a formula. This separation leads to two different theories:

A2: Proof Theory VS Model Theory.

In the former, mathematical reasoning is captured as grammatical symbol manipulations from given axioms, while in the latter, mathematical models satisfying given axioms are considered. These theories are connected by the, so-called, *soundness-completeness* theorem.

The above connection is important conceptually as well as technically. Model theory talks about models, each of which is assumed to be a *complete* description of a targeted world. On the other hand, a proof theory talks about a given set of nonlogical axioms, which may describe only partially the targeted world. In model theory, a set of candidate models corresponds to this description, that is, all the possibilities for a description are considered. Therefore, the soundness-completeness theorem will take the form between syntactical provability and truthfulness for all the candidate models. It is important to notice that this treatment differs from the recent game theoretical literature of epistemic models since Aumann [1], where a *particular* model is assumed to a description of the targeted world.[1]

Another relevant distinction here is:

A3: Object Theorems and Meta-Theorems.

A theorem whose provability and/or validity is discussed *inside* a logical system belongs to the former, and a theorem on a logical system belongs to the latter. Since our purpose is to investigate the players' inferences, meta-theorems on their inferences are our central concerns. This distinction will be clear when some examples are given.

It may be helpful for the reader who have never read a logic textbook to mention that we use a standard mathematical method for handling a logic system. In other word, mathematical logic is a mathematical theory of a mathematical theory. The mathematical method to handle a logical system is called *meta-mathematics*. This note will be pointed out when it is relevant.[2]

In this paper, we will discuss epistemic logics, which are variants of modal logics, which originally targeted the investigation of "necessity" and "possibility". We can borrow a lot from this literature.[3] In the case of epistemic logics, the above distinction A2 becomes.

A4: Hilbert-Style Proof Theory VS Kripke Semantics.

We will discuss mainly these two theories in this paper. The Hilbert-style proof theory is convenient in its concise presentation. One important strength of the logical approach is to show unprovability (or provability in the negative direction). For this purpose, the Hilbert-style proof theory is difficult to handle, and the Kripke semantics is unavoidable. There is another proof theory called *Gentzen-style*, which enables us to evaluate unprovability as well as provability. In this paper, however, we give only a brief explanation of it in Section 4.4.

Since the above two theories are regarded as deductively equivalent, which is stated by the soundness-completeness theorem, the reader may wonder why we adopt both theories. The answer is double-fold. First, technically speaking, since each theory has some merit and some demerit, it would be more powerful to have both theories. Second, conceptually speaking, both syntactical manipulations and semantical considerations are important in investigations of human-thoughts in social contexts. Therefore, we keep the dualistic attitude.

Another relevant distinction is

A5: Propositional Logics VS Predicate Logics.

Predicate logic is an extension of propositional logic in that it allows for quantifications, $\forall$ (for all) and $\exists$ (exists). When the domain of a problem has an infinite number of objects such as a game with mixed strategies, a predicate logic is unavoidable. If the problem involves only a finite number of objects such as a finite

[1] See Bacharach-Mongin [2] and Bacharach, et al [3] for the recent game theoretical literature of epistemic models.

[2] See Kleene [19] for the distinction between object-mathematics and meta-mathematics.

[3] Hughes-Cresswell [11] and Chellas [5] are standard textbooks of modal logic. The modern literature of epistemic logic was started by Hintikka [9]. Fagin-Halpern-Moses-Verdi [6] and Meyer-van der Hoek [23] are found as textbooks on epistemic logics and other subjects.

game with pure strategies, the quantifications, $\forall$ and $\exists$, can be effectively expressed by conjunction $\bigwedge$ (and) and $\bigvee$ (or), respectively. Therefore, as far as we confine ourselves to finite games, propositional logics suffice.[4] In this introduction paper, we discuss only propositional epistemic logics.

1.2. Logical Approach to Game Theoretical Problems

We now turn our attention from broad distinctions in logic to ones particular to epistemic logics and their game theoretical applications. The first distinction is:

B1: Classical Logic VS Epistemic Logics.

Classical logic is the one which is used in the standard mathematical practices. We adopt classical logic as the base logic for our epistemic logics. This means that epistemic logics are constructed as superstructures of classical logic. In other word, the investigator's reasoning ability is described by classical logic, and the players' are described in epistemic logics. The reasoning ability of each player consists of that by classical logic and some additional inference abilities such as introspection.

The second distinction specific to our subjects is particularly important for game theoretical applications:

B2: Logical Axioms VS Nonlogical (Mathematical) Axioms.

This distinction can be, more or less, arbitrary in classical logic, but it is crucial in epistemic logics (modal logics in general). We take the research strategy of separating game theoretical problems from logical problems as much as possible. To achieve this separation, we will describe game theoretical axioms as nonlogical axioms.

The third distinction is:

B3: Beliefs VS Knowledge.

In this paper, "knowledge" will be defined as "true beliefs", where truth is referred to the outside thinker.[5] Beliefs are also divided into basic and derived ones. A justification of a belief for a player is a derivation (proof) from basic beliefs by himself. However, we do not discuss "justifications" of basic beliefs by a player, which is a limit of the logic approach. We will treat the truthfulness axiom, which makes beliefs true, as a possible axiom rather than a basic axiom. We would like to allow *false beliefs*, rather than to discuss whether a player can or cannot obtain true beliefs. To allow false beliefs enables us to consider the emergence of beliefs from other sources such as individual and social experiences. We will discuss the distinction B3 from the logical point of view in Section 6.

The following distinction may help understand our game theoretical applications:

B4: Solution Theory VS Performance-Playability Theory.

[4] Predicate logics may be relevant to some problems such as complexities of expressions (which are relevant for some situations, e.g., communication within language) even if the problems have only finite objects. Kleene [19] and Mendelson [22] are classical textbooks treating basics of predicate logic.

[5] Since we are talking about foundational issues, objectivity can be defined only relative to the thinker. In our case, the ultimate reference of objectivity is the investigator.

Solution theory addresses what criteria are adopted for decision making, while performance-playability theory takes a solution theory as given and addresses how the theory performs and whether the theory allows the player to make a final decision.

A related distinction is:

B5: Decision VS Prediction.

In a game, each individual player makes a decision under predictions of other players' decisions. A decision is ultimately important for the player, while predictions on other players' decisions are auxiliary. For some games, a decision may be made without predictions. Also, we do not need to assume the same decision criteria for decision and prediction. Rather, these are different by nature. Traditional game theory has not distinguished between them. In this paper, we make this distinction, but will have enough space to examine it fully. The distinctions B5 as well as B4 will be clearer in the second paper of this issue.

The organization of this paper is as follows: Section 2 describes games to be used for illustrations and some decision criteria. Section 3 describes semantics and syntax of classical logic CL, and states the soundness-completeness theorem. Its proof will be given in the Appendix. Section 4 gives various epistemic logics and Kripke semantics, which are connected, again, by the soundness-completeness theorem. Section 6 discusses the relationship between "beliefs" and "knowledge". Section 7 discusses decision criteria, and concludes that for some games, an exact solution for decision making involves common knowledge, but the epistemic logics introduced in Section 4 are incapable to treat such a problem. Therefore, we will discuss an extension of an epistemic logic to incorporate common knowledge in Section 8.

2. Basic Game Theoretical Concepts

To exemplify logical constructs, we will refer to a 2-person finite noncooperative game $g = (g_1, g_2)$ in strategic form. Each player $i = 1, 2$ has ℓ_i pure strategies ($\ell_i \geq 2$). We assume throughout the paper that the players do not play mixed strategies. Player i's strategy space is denoted by $S_i := \{\mathrm{s}_{i1}, ..., \mathrm{s}_{i\ell_i}\}$ for $i = 1, 2$. His *payoff function* is a real-valued function g_i on $S := S_1 \times S_2$.

Let $(s_1, s_2) \in S$. We say that s_1 is a *best response to* s_2 iff $g_1(s_1, s_2) \geq g_1(t_1, s_2)$ for all $t_1 \in S_1$. We say that s_1 is a *dominant strategy* iff s_1 is a best response to s_2 for any $s_2 \in S_2$. Player 2's dominant strategy is defined in the parallel manner. A strategy pair $s = (s_1, s_2)$ is called a *Nash equilibrium* iff s_i is a best response to s_j for $i, j = 1, 2$ ($i \neq j$). We say that s_i is a *Nash strategy* for player i iff (s_1, s_2) is a Nash equilibrium for some s_j, where $i, j = 1, 2$ ($i \neq j$).

In the game $g^1 = (g_1^1, g_2^1)$ of Table 2.1 (Prisoner's Dilemma), the second strategy s_{i2} for each i is a dominant strategy. In the game $g^2 = (g_1^2, g_2^2)$ of Table 2.2 which is obtained from g^1 by changing the payoff 6 in the northeast corner to 2, only player 1 has a dominant strategy, s_{12}. Either game has a unique Nash equilibrium, $(\mathrm{s}_{12}, \mathrm{s}_{22})$.

	s_{21}	s_{22}
s_{11}	$(5,5)$	$(1,6)$
s_{12}	$(6,1)$	$(3,3)^*$

Table 2.1: $g^1 = (g_1^1, g_2^1)$

	s_{21}	s_{22}
s_{11}	$(5,5)$	$(1,\tilde{2})$
s_{12}	$(6,1)$	$(3,3)^*$

Table 2.2: $g^2 = (g_1^2, g_2^2)$

As far as the payoff maximization is the criterion for decision making (and it is, of course, assumed that each player can control only his own strategy), these games need different epistemic assumptions, which we will discuss using epistemic logics. Here, we briefly describe in the standard game theory language what decision criteria are candidates for these games, and later, we will see how such decision criteria are accurately described in epistemic logics.

The following is a simple decision criterion:

DC1: Player i chooses a dominant strategy.

In game $g^1 = (g_1^1, g_2^1)$, this criterion recommends a decision to either player. However, it recommends a strategy only to player 1 in game $g^2 = (g_1^2, g_2^2)$, since 2 has no dominant strategies in g^2. One way out for 2 is to predict 1's decision, assuming that 1 adopts DC1 for 1's choice. We write this criterion as follows:

DC2: Player i, predicting that the other j would choose a strategy following DC1, chooses a best strategy to his predicted strategy for player j.

This criterion differs from DC1 in that it involves a prediction on the other player' decision making. The application of DC2 to player 2 in game g^2 states that he predicts that 1 would choose s_{12} as the dominant strategy and then 2 chooses s_{22} as the best strategy to s_{12}. This argument is a special case of the procedure so called the *iterated elimination of dominated strategies* (cf., Moulin [26] and Myerson [27]). The main concern in literature is the consideration of a resulting outcome of such a procedure, but our concern is the considerations of required epistemic aspects for such a decision criterion and of its performance-playability relative to given beliefs. We will make this point clear later.

To differentiate prediction-making by thinking about the other player's thought from that without it, we consider one primitive criterion:

DC3^0: Player i chooses a Nash strategy.

Here, player i is required to think about the other 's payoff function, but not to think about the other 's thought on his decision. This will be compared with the other extreme criterion, DC3, mentioned below.

The above, DC1, DC2 and DC3^0 are criteria among many possible ones (including the Default Decision Criterion discussed in Kaneko-Suzuki [18]). It may be possible that player 2 adopts the prediction criterion for 1 different from 1's actual decision criterion. We combine DC2 and DC3^0 as follows:

DC2&3^0: (1) Player 1 follows DC3^0, and (2) player 2 follows DC2.

This involves subtle beliefs which appear to be false or inconsistent. It would be difficult to discuss the falsity and/or consistency of DC2&3^0 without formulating

DC2&3^0 in a formal language. Now, the reader may find a necessity of our logical approach. Here, we emphasize that we have already met the problem of false beliefs. Such true and false beliefs are important to discuss the players's views on the game (society) as well as their decision making. We will discuss DC2&3^0 in Section 7.1.

The game, $g^3 = (g^3, g^3)$, of Table 2.3 is obtained from g^2 by adding one strategy to player 2. In this game, neither player has a dominant strategy. Hence, neither DC1 nor DC2 gives a decision to a player. DC3^0 makes a recommendation, but DC3^0 does not guarantees player i to believe that his prediction is believed by the other player. Therefore, we consider the other extreme prediction-decision criterion. It is obtained as a modification of DC2 by requiring predictions for player j in the symmetric manner as for player i:

DC3: (1) Player 1 chooses a best strategy to the prediction made by (2);

(2): player 2 chooses a best strategy to the prediction made by (1).

One problem is whether or not DC3 leads to a Nash equilibrium, but our main concern is to consider the epistemic structure involved in DC3. In fact, if we assume that the same decision criterion is adopted for the player in the mind of the other player, then we would meet the infinite regress:

$$(1) \leftarrow\!\!-- \ ((2) \leftarrow\!\!-- \ ((1) \leftarrow\!\!-- \ ((2) \leftarrow\!\!-- \ \ldots$$

$$(2) \leftarrow\!\!-- \ ((1) \leftarrow\!\!-- \ ((2) \leftarrow\!\!-- \ ((1) \leftarrow\!\!-- \ \ldots$$

This infinite regress is closely related to the common knowledge of these criteria. When an individual player adopts this criterion, the infinite regress appears in his mind, and it takes only the form of an *individual belief* of common knowledge. This difference will be discussed together with the choice of some axiom for an epistemic logic.

Criteria DC1 and DC2 can be discussed in an (finitary) epistemic logic, but to capture the infinite regress in DC3, we need an extension of an epistemic logic to incorporate common knowledge. Even if common knowledge is involved, it may be possible to allow false beliefs. For example, player 1 believes the common knowledge of playing game g^3, while 2 does the common knowledge of playing the game g^4 of Table 2.4. To discuss those problems in meaningful manners, we need carefully to develop (and evaluate) epistemic logics. These problems will be discussed in Sections 7 and 8.

	s_{21}	s_{22}	s_{23}
s_{11}	(5,5)	(1,2)	(4,3)
s_{12}	(6,1)	(3,3)*	(0,2)

Table 2.3: $g^3 = (g_1^3, g_2^3)$

	s_{21}	s_{22}	s_{23}
s_{11}	(6,3)	(1,2)	(0,5)
s_{12}	(5,2)	(3,3)*	(4,1)

Table 2.4: $g^4 = (g_1^4, g_2^4)$

As stated in Section 1, our general attitude is to focus on problems of finite natures rather than of transcendental ones. Sometimes, however, we need to consider the limit case of finite ones in order to understand finite problems. We regard common knowledge as such notions.

3. Classical Logic CL

In this section, we review classical logic CL and its semantics. The reader may feel that this is a detour to the logic approach to game theory. However, we will define epistemic logics as superstructures of classical logic CL, and will use a lot of concepts and results from CL in the development of epistemic logics as well as in applications to game theory. Before going to epistemic logics, the reader should become a bit familiar to CL.

In Section 3.1, we define two sets of formulae. In Section 3.2, we give the classical semantics. In Section 3.3, we give one axiomatic presentation of classical logic CL, and state the soundness and completeness of CL.

3.1. The Sets of Formulae: $\mathcal{P}$ and $\mathcal{P}^n$

We start with the following list of symbols:

propositional variable symbols: $\mathrm{p}_0, \mathrm{p}_1, ...$;

logical connective symbols: $\neg$ (not), $\supset$ (implies), $\bigwedge$ (and), $\bigvee$ (or) ;

unary belief operator symbols: $\mathrm{B}_1, \mathrm{B}_2, ..., \mathrm{B}_n$;

parentheses: (,) ; *set-theoretical brackets*: { , }; and *comma*: , .

We stress that these are pure symbols and are used to be elements of more complex expressions, called formulae. We will associate the *intended* meanings, "not", "implies", "and", "or", with connective symbols, $\neg$, $\supset$, $\bigwedge$, $\bigvee$, respectively. The implication symbol $\supset$ should be distinguished from the set-theoretical inclusion $\subseteq$. These intended meanings will be defined operationally by logical axioms and inference rules. Unary belief operator symbol B_i is applied to each formula. The subscripts of these belief operators are the names of players, i.e., B_i is the belief operator of player i. We denote the set of players by $N = \{1, ..., n\}$. The set-theoretical brackets $\{,\}$ are used to express a finite set of formulae. It is assumed that the number of propositional variables is at least one and at most countable. The set of propositional variables is denoted by PV.

Based on the above list symbols, we define *formulae* inductively as follows:

F1: any $p \in PV$ is a formula;

F2: if A and B are formulae, so are $(\neg A), (A \supset B)$ and $\mathrm{B}_i(A)$ $(i \in N)$;

F3: if $\{A_0, A_1, ..., A_m\}$ is a finite set of formulae with $m \geq 0$, then $(\bigwedge\{A_0, A_1, ..., A_m\})$ and $(\bigvee\{A_0, A_1, ..., A_m\})$ are also formulae;[6]

F4: each formula is obtained by a finite number of applications of F1, F2 and F3.

---

[6]This definition deviates from the standard textbook definition of formulae in that conjunctive and disjunctive connectives $\bigwedge$ and $\bigvee$ are applied to a finite nonempty set of formulae, e.g., $\bigwedge\{A_0, A_1, ..., A_m\}$, rather than to an ordered pair of formulae. We take this deviation to facilitate game theoretical applications. However, the resulting logical systems are equivalent (with respective to provabilities or validities defined in the systems). This formulation does fit to the Gödel numbering. Therefore, if one wants to take the Gödel numbering, then he should return to the standard formulation.

For example, $(\mathrm{p}_0 \supset (\mathrm{p}_1 \supset \mathrm{p}_0))$ is obtained by applications of F1 and F2 three times and twice, respectively, and is intended to mean that if p_0 holds, then p_1 implies p_0. The construction of two formulae is described by the following trees:

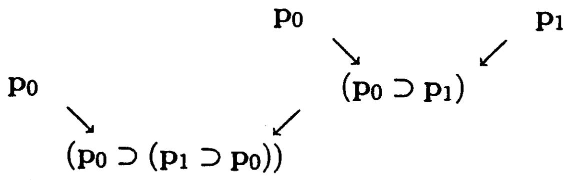

Figure 3.1

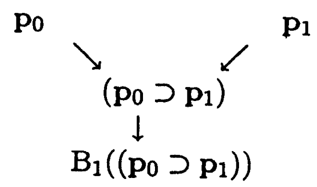

Figure 3.2

In general, a formula has a *finite* tree structure where each terminal node corresponds to a propositional variable and each nonterminal node corresponds to a logical connective or a belief operator. This tree structure will be used to construct inductive proofs. We denote the *set of all formulae* by $\mathcal{P}$.

By a formula $\mathrm{B}_i(A)$ constructed by F2, we intend to mean that player i believes formula A. The behavior of the belief operators is of our central interests. In classical logic CL, we first ignore these formulae, but later, will give a remark on a somewhat nominal treatment of belief operators. We say that a formula A is *nonepistemic* iff A contains no $\mathrm{B}_1, ..., \mathrm{B}_n$. We denote the set of all nonepistemic formulae by $\mathcal{P}^{\mathrm{n}}$. Of course, $\mathcal{P}^{\mathrm{n}} \subseteq \mathcal{P}$.

In this paper, we do not fix exact rules of abbreviations of parentheses (,), but it suffices to follow standard practices of abbreviations so that we could recover the original expressions when necessary. For example, $(\mathrm{p}_0 \supset (\mathrm{p}_1 \supset \mathrm{p}_0))$, $\mathrm{B}_1((\mathrm{p}_1 \supset \mathrm{p}_0))$ and $(\bigwedge \Phi)$ are abbreviated as $\mathrm{p}_0 \supset (\mathrm{p}_1 \supset \mathrm{p}_0)$, $\mathrm{B}_1(\mathrm{p}_1 \supset \mathrm{p}_0)$ and $\bigwedge \Phi$, respectively. We will also abbreviate $\bigwedge\{A, B\}, \bigvee\{A, B, C\}$ as $A \bigwedge B, A \bigvee B \bigvee C$, etc.[7] We denote $(A \supset B) \bigwedge (B \supset A)$ by $A \equiv B$.[8]

To discuss the game theoretical problems of Section 2, we adopt the economics practice to represent a payoff function in terms of preference relations. We start with:

strategy *symbols*: $\mathrm{s}_{11}, ..., \mathrm{s}_{1\ell_1}; \mathrm{s}_{21}, ..., \mathrm{s}_{2\ell_2}$;

4-ary *symbols*: P_1, P_2.

Strategy symbols are identical to those given in Section 2. By a 4-ary symbol P_i, we mean that the expression $P_i(s_1, s_2 : t_1, t_2)$ is allowed for $(s_1, s_2), (t_1, t_2) \in S$. These 4-ary expressions are called *atomic formulae,* and the set of them is denoted by AF. When $\ell_1 = \ell_2 = 2$, AF consists of 32 atomic formulae. When we discuss game theoretical problems, we regard always AF as PV. Also, the sets of formulae $\mathcal{P}$ and $\mathcal{P}^{\mathrm{n}}$ are defined based on AF as the replacement of PV. We see presently how some game theoretical concepts are described. It should be noted that atomic formula $P_i(s_1, s_2 : t_1, t_2)$ is intended to be a *weak preference* for (s_1, s_2) over (t_1, t_2).

[7] In the definition of formulae, we presume the identity of a finite set. Hence, $(\bigwedge\{A_1, A_2\})$ is identical to $(\bigwedge\{A_2, A_1\})$ as a formula.

[8] We introduce four logical connectives, $\neg$, $\supset$, $\bigwedge$ and $\bigvee$. Some of them are sufficient and the others can be defined as abbreviations. However, this abbreviation may be convenient for some presentation purposes, but not necessarily so for other purposes. This is rather a matter of taste.

The statement that s_1 is a best response to player 2's s_2 is described as the formula $\bigwedge\{P_1(s_1, s_2, : t_1, s_2) : t_1 \in S_1\}$, which we denote by $\text{Best}_1(s_1 \mid s_2)$. The statement that s_1 is a dominant strategy for player 1 is expressed as $\bigwedge\{\text{Best}_1(s_1 \mid s_2) : s_2 \in S_2\}$. This means that s_1 is the most preferable whatever player 2 chooses. The formula $\bigwedge\{P_1(s_1, t_2, : t_1, t_2) : t_1 \in S_1 \text{ and } t_2 \in S_2\}$ is equivalent to $\bigwedge\{\text{Best}_1(s_1 \mid s_2) : s_2 \in S_2\}$ in the logic we will define. We denote the former formula by $\text{Dom}_1(s_1)$. In the parallel manner, we define the formulae, $\text{Best}_2(s_2 \mid s_1)$ and $\text{Dom}_2(s_2)$. The statement that (s_1, s_2) is a Nash equilibrium is described as $\text{Best}_1(s_1 \mid s_2) \bigwedge \text{Best}_2(s_2 \mid s_1)$, which is denoted by $\text{Nash}(s_1, s_2)$. The statement that s_1 is a Nash strategy for player 1 is described as $\bigvee\{\text{Nash}(s_1, s_2) : s_2 \in S_2\}$. This will be abbreviated as $\bigvee_{s_2}\text{Nash}(s_1, s_2)$.

Let us recall that a payoff function g_1 of player 1 was given as a real-valued function in Section 2. The purpose of our logic approach is not to formalize the extant game theory, but to investigate game theoretical phenomena which are difficult to investigate without logic. Therefore, we adopt a simple assumption if it is not really our concern. Although it is a practice in economics to require completeness and transitivity for the preference relation $P_1(s : t)$, we adopt a much simpler manner. Specifically, we express a payoff function g_1 by the following set of preferences:

$$\{P_1(s : t) : g_1(s) \geq g_1(t)\} \cup \{\neg P_1(s : t) : g_1(s) < g_1(t)\}. \tag{3.1}$$

This is a set of symbolic expressions and is denoted by $\hat{g}_1$. We can take the conjunction, $\bigwedge \hat{g}_1$, of this set. In the parallel manner, $\hat{g}_2$ and $\bigwedge \hat{g}_2$ are defined. Thus, the payoff functions for both players are described as the set $\hat{g} = \hat{g}_1 \cup \hat{g}_2$ or as the formula $\bigwedge(\hat{g}_1 \cup \hat{g}_2)$. Throughout the paper, we use the convention that a given payoff function h_i is expressed as $\hat{h}_i$ or $\bigwedge \hat{h}_i$ in our formalized language.

3.2. Classical Semantics

So far, we have defined formulae expressing logical or game theoretical ideas, but we have not considered a way of evaluating them. In this subsection, we define semantical notions, "truth" and "falsity". From these, we define another semantical notion, "validity", which will be connected to a syntactical notion, "provability".

First, we give the definition of a semantical valuation of each formula in $\mathcal{P}^n$. We will give a remark on the modification of this definition for $\mathcal{P}$ in the end of Section 3.3.

A function $\kappa : PV \to \{\top, \bot\}$ is called an (classical) *assignment*, where $\top$ and $\bot$ are the symbols designating "true" and "false", respectively. We extend each assignment κ to the function $V_\kappa : \mathcal{P}^n \to \{\top, \bot\}$ by the following induction on the length (tree structure) of a formula:

C0: for any $p \in PV$, $V_\kappa(p) = \top$ if and only if $\kappa(p) = \top$;

C1: $V_\kappa(\neg A) = \top$ if and only if $V_\kappa(A) = \bot$;

C2: $V_\kappa(A \supset B) = \top$ if and only if $V_\kappa(A) = \bot$ or $V_\kappa(B) = \top$;

C3: $V_\kappa(\bigwedge \Phi) = \top$ if and only if $V_\kappa(A) = \top$ for all $A \in \Phi$;

C4: $V_\kappa(\bigvee \Phi) = \top$ if and only if $V_\kappa(A) = \top$ for some $A \in \Phi$.

By induction, the value $V_\kappa(A)$ is defined for every $A \in \mathcal{P}^n$. That is, each step defines its left-hand side by the right-hand side. For example, when $PV = \{p_0, p_1\}$ and $\kappa(p_0) = \top$ and $\kappa(p_1) = \bot$, we calculate $V_\kappa(p_0 \supset (p_1 \supset p_0))$ from the leaves of Figure 3.1, and obtain $V_\kappa(p_0 \supset (p_1 \supset p_0)) = \top$. In fact, the valuation of this formula is always $\top$ independent of κ. Such a formula is said to be a classical tautology, or to be classically valid. More precisely, we say that $A \in \mathcal{P}^n$ is a *classical tautology* iff $V_\kappa(A) = \top$ for all assignments κ. In fact, $A \supset (B \supset A)$ is a tautology for any formulae $A, B \in \mathcal{P}^n$: Indeed, for any κ, if $V_\kappa(A) = \bot$, then $V_\kappa(A \supset (B \supset A)) = \top$ by C3, and if $V_\kappa(A) = \top$, then $V_\kappa(B \supset A) = \top$ by C3 and *a fortiori*, $V_\kappa(A \supset (B \supset A)) = \top$.

To describe game theoretical assumptions, we will use *nonlogical* axioms. Let Γ be a subset of $\mathcal{P}^n$, which is intended to be a set of nonlogical axioms. We say that a classical assignment κ is a *model of* Γ iff $V_\kappa(C) = \top$ for all $C \in \Gamma$. For any $A \in \mathcal{P}^n$, we say that A is a *semantical consequence* of Γ iff $V_\kappa(A) = \top$ for all models κ of Γ, in which case we write $\Gamma \models A$. When Γ is empty, we write $\models A$. In this case A is a tautology. Also, we write $\Gamma \not\models A$ iff not $\Gamma \models A$. Note that this differs from $\Gamma \models \neg A$.

Let us return to game theoretical problems. Consider the game $g^1 = (g_1^1, g_2^1)$ of Table 2.1. Recall that we adopt the set of atomic formulae AF as PV. In the case of a 2×2-game, AF consists of 32 elements. The payoff functions g_1^1 and g_2^1 are described as $\hat{g}_1^1$ and $\hat{g}_2^1$ defined by (3.1). Since $\hat{g}_i^1$ contains either $P_i(s:t)$ or $\neg P_i(s:t)$ for each $P_i(s,t) \in AF$, the value of a model κ of $\hat{g}_i^1$ on $P_i(s:t)$ is uniquely determined, that is, $\hat{g}_1^1 \cup \hat{g}_2^1$ has a unique model κ, in which sense $\hat{g}_1^1 \cup \hat{g}_2^1$ is a *complete* description of game g^1.[9]

Consider the models of $\hat{g}_1^1$. In this case, since a model κ is arbitrary on $P_2(s:t)$, $\hat{g}_1^1$ allows 2^{16} models. For any model κ of the set $\hat{g}_1^1$, we have $V_\kappa(\mathrm{Dom}_1(s_{12})) = \top$. Thus, we have

$$\hat{g}_1^1 \models \mathrm{Dom}_1(s_{12}) \text{ and } \hat{g}_1^1 \models \neg\mathrm{Dom}_1(s_{11}). \tag{3.2}$$

This means that if $\hat{g}_1^1$ is assumed, $\mathrm{Dom}_1(s_{12})$ and $\neg\mathrm{Dom}_1(s_{11})$ are derived as consequences. Also, it holds that $\hat{g}_2^1 \models \mathrm{Dom}_2(s_{21})$ and $\hat{g}_2^1 \models \neg\mathrm{Dom}_2(s_{22})$. When $\hat{g}_1^1 \cup \hat{g}_2^1$ is assumed, it holds that

$$\hat{g}_1^1 \cup \hat{g}_2^1 \models \mathrm{Nash}(s_{12}, s_{22}).$$

However, unless $\hat{g}_2^1$ is assumed, $V_\kappa(\mathrm{Best}_2(s_{22} \mid s_{12})) = \bot$ for some models κ of $\hat{g}_1^1$, and thus

$$\hat{g}_1^1 \not\models \mathrm{Nash}(s_{12}, s_{22}). \tag{3.3}$$

Thus, if enough information is not assumed, we could not conclude that (s_{12}, s_{22}) is a Nash equilibrium.

The above is, more or less, the standard game theoretical argument. Many such standard arguments can be made with the logical structure discussed so far. However, the current structure is still limited in that it cannot address arguments like DC2 of Section 2. Such arguments involve beliefs or predictions about the behavior of the other players. To adequately describe these arguments, we need to introduce epistemic conditions on $\mathrm{B}_i(\cdot)$ which come in Section 4.

[9] up to the orderings determined by the payoff functions g_1 and g_2.

3.3. Classical Logic CL and its Provability $\vdash_0$

As yet we have discussed basic notions like formulae and semantical consequences. In this subsection, we provide a proof-theoretic system of classical logic CL and introduce the notion of a proof. The proof-theoretic system formulates mathematical inferences as pure symbol manipulations, while the consequence relation $\vDash$ is formulated by considering meanings or possibilities of symbolic expressions. We denote this proof-theoretic system also by CL. We will soon find out that the provability and consequence relation are intimately related.

Classical logic CL consists of five axiom schemata and three inference rules, which describe the possible ways of manipulating formulae of $\mathcal{P}^n$. The notion of a proof will be defined by means of such components. More concretely, those five *axiom schemata* and three *inference rules* are as follows: for any formulae A, B, C and finite nonempty set Φ of formulae in $\mathcal{P}^n$,

L1: $A \supset (B \supset A)$;

L2: $(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$;

L3: $(\neg A \supset \neg B) \supset ((\neg A \supset B) \supset A)$;

L4: $\bigwedge \Phi \supset A$, where $A \in \Phi$;

L5: $A \supset \bigvee \Phi$, where $A \in \Phi$;

$$\frac{A \supset B \quad A}{B} \text{ (Modus Ponens)}$$

$$\frac{\{A \supset B : B \in \Phi\}}{A \supset \bigwedge \Phi} \text{ ($\bigwedge$-Rule)} \qquad \frac{\{A \supset B : A \in \Phi\}}{\bigvee \Phi \supset B} \text{ ($\bigvee$-Rule).}$$

Modus Ponens is abbreviated as MP. These axioms and rules are schemata in the sense that formulae, A, B, C and the set Φ can be arbitrary. A particular formula or inference rule of them is called an *instance*, for example, $p_0 \supset (p_1 \supset p_0)$ is an instance of L1.[10]

Let A be a formula in $\mathcal{P}^n$ and Γ a subset of $\mathcal{P}^n$. A *proof* of A from Γ in CL is a finite tree with the following properties:

(1): a formula in $\mathcal{P}^n$ is associated with each node;

(2): the formula associated with each leaf is an instance of the above axioms or belongs to Γ;

(3): adjoining nodes together with their associated formulae form an instance of the above three inference rules;

(4): A is associated with the root node.[11]

---

[10] There are many other formulations of classical propositional logic (see Mendelson [22], pp.37-38). The present axiomatization is given in Kaneko-Nagashima [15].

[11] More explicitly, a proof is given as a triple $(X, \prec: \varphi)$, where $(X, \prec)$ is a finite tree in the sense of graph theory and φ is a function from X to $\mathcal{P}^n$ associating formula $\varphi(x)$ with a each node $x \in X$. An ending node x is called a *leaf*, and the initial node is called the *root*.

We say that A is said to be *provable from* Γ in CL, denoted by $\Gamma \vdash_0 A$, iff there is a proof of A from Γ. When Γ is empty, we write simply $\vdash_0 A$. We write $\Gamma \nvdash_0 A$ iff not $\Gamma \vdash_0 A$.

The following lemma gives simple examples of provable formulae.

Lemma 3.1.(1): $\vdash_0 A \supset A$;

(2): $\{A \supset B, B \supset C\} \vdash_0 A \supset C$.

Proof.(1): The following is a proof tree of $A \supset A$.

$$\dfrac{\dfrac{\overset{\text{L2}}{[A \supset ((A \supset A) \supset A)] \supset [(A \supset (A \supset A)) \supset (A \supset A)]} \quad \overset{\text{L1}}{A \supset ((A \supset A) \supset A)}}{(A \supset (A \supset A)) \supset (A \supset A)}\,\mid \text{MP} \qquad \overset{\text{L1}}{A \supset (A \supset A)}}{A \supset A}\,\mid \text{MP}$$

(2): The following is a proof of $A \supset C$ from $\{A \supset B, B \supset C\}$.

$$\dfrac{\overset{\text{Ass.}}{A \supset B} \qquad \dfrac{\dfrac{\overset{\text{Ass.}}{B \supset C} \quad \overset{\text{L1}}{(B \supset C) \supset (A \supset (B \supset C))}}{A \supset (B \supset C)}\,\mid \text{MP} \qquad \overset{\text{L2}}{[A \supset (B \supset C)] \supset [(A \supset B) \supset (A \supset C)]}}{(A \supset B) \supset (A \supset C)}\,\mid \text{MP}}{A \supset C}\,\mid \text{MP}$$

■

The next lemma will be used without mentioning.

Lemma 3.2.(1): $\Gamma \subseteq \Gamma'$ and $\Gamma \vdash_0 A$ imply $\Gamma' \vdash_0 A$;

(2): if $\Gamma' \vdash_0 A$ and $\Gamma \vdash_0 B$ for all $B \in \Gamma'$, then $\Gamma \vdash_0 A$.

Here, we find the distinction A3 stated in Section 1: Object Theorems vs Meta-Theorems. The claims of Lemma 3.1 are object theorems, while those of Lemma 3.2 are meta-theorems on object theorems, which are not formulated in CL but in metamathematics.

We would like to have a bridge between the classical semantics given in Section 3.2 and the syntactical system CL. To have this connection, we need a key word: We say that a set Γ of formulae in $\mathcal{P}^n$ is *inconsistent* in CL iff $\Gamma \vdash_0 \neg C \bigwedge C$ for some C, and that Γ is *consistent* in CL iff it is not inconsistent in CL.

Theorem 3.3 (Soundness-Completeness). Let Γ be a set of formulae and A a formula. Then

(1): $\Gamma \vdash_0 A$ if and only if $\Gamma \models A$;

(2): there is a model κ of Γ if and only if Γ is consistent with respect to $\vdash_0$.

Assertions (1) and (2) (with the quantifications of all Γ and A for each) are actually equivalent. The *only-if part* of each is called *soundness*, and the *if* part is *completeness*. The term "soundness" means that the syntactical formulation of symbolic inferences suffers from nothing other than the semantical validity, and "completeness" means that the former captures the latter. Thus, the two approaches by considering the meanings or possibilities of symbolic expressions and by grammatical symbol manipulations now become equivalent. In this sense, our description of "logic" is complete. This gives the bridge between model theory and proof theory stated by A2.

The above completeness differs from the "completeness" of $\hat{g}_1^1 \cup \hat{g}_2^1$. The latter is the completeness of nonlogical axioms (the description of the game).[12]

The above theorem should be regarded as quite standard, and the above syntactical system turns to be equivalent to the other formulations of classical logic (cf., Mendelson [22]). Nevertheless, the axiomatization is modified to facilitate game theoretical arguments, and we find no textbook for exactly the same theorem. The soundness part can be proved without much difficulty, following a proof in a textbook, but the completeness part is quite complicated, where complication is caused by a sophistication of the axiomatization. We will give only proofs of completeness and the equivalence of (1) and (2) in the Appendix.

By the above theorem, we can use any tautologies as provable formulae in CL. For example, $\neg A \bigvee A$, $\neg(\neg A \bigwedge A)$, $\neg \bigvee \Phi \equiv \bigwedge\{\neg A : A \in \Phi\}$, $\neg \bigwedge \Phi \equiv \bigvee\{\neg A : A \in \Phi\}$ and $(A \supset B) \equiv (\neg B \supset \neg A)$ are provable in CL. In the subsequent sections, we will use such tautologies without mentioning.

It follows from soundness that CL is contradiction-free.

Corollary 3.4. There is no formula A such that $\vdash_0 \neg A \bigwedge A$.

Let us return to the game theoretical example. Using Theorem 3.3, (3.2) and (3.3) can be written as

$$\hat{g}_1^1 \vdash_0 \mathrm{Dom}_1(\mathrm{s}_{12}) \text{ and } \hat{g}_1^1 \vdash_0 \neg\mathrm{Dom}_1(\mathrm{s}_{11}). \tag{3.4}$$

$$\hat{g}_1^1 \nvdash_0 \mathrm{Nash}(\mathrm{s}_{12}, \mathrm{s}_{22}). \tag{3.5}$$

In fact, (3.4) can be proved in CL without much difficulty. It is easier to prove (3.4) than (3.2). On the other hand, it is difficult to obtain (3.5) directly. For a direct proof of (3.5), we should show that there is no proof P of $\mathrm{Nash}(\mathrm{s}_{12}, \mathrm{s}_{22})$ from $\hat{g}_1^1$. Since there are an infinite number of proof trees, we should narrow down the candidate proofs using the given information $\hat{g}_1^1$ and $\mathrm{Nash}(\mathrm{s}_{12}, \mathrm{s}_{22})$. However, since we allow MP in a proof, some formulae may disappear at some places in a proof. We cannot trace such disappearances in P from the given information. Semantics and it's connection to provability via Theorem 3.3 enables us to show unprovability simply by constructing one countermodel.

Some reader may wonder why we use both model theory and proof theory. The basic research strategy we adopt is to keep the plurality of logical systems for the logical approach to game theory, since plurality gives different views and different opportunities of obtaining new results. Our concern is not the "axiomatization" such as soundness-completeness, but is to obtain new results using whatever is convenient for each purpose.

[12] The latter is the completeness of a theory in the sense of logic literature.

Remark 3.5. When we adopt $\mathcal{P}$ rather than $\mathcal{P}^n$, syntactical system CL is not modified at all; just replace $\mathcal{P}^n$ by $\mathcal{P}$. On the other hand, the classical semantics should be slightly modified: the domain of each assignment κ becomes $PV \cup \{B_i(A) : A \in \mathcal{P}$ and $i \in N\}$ with the same image $\{\top, \bot\}$. Accordingly, C0 is replaced by

C0*: for any $C \in PV \cup \{B_i(A) : A \in \mathcal{P}$ and $i \in N\}$, $V_\kappa(C) = \top$ if and only if $\kappa(C) = \top$.

Then all the other definitions are the same. Here, $B_i(A)$ is treated in the same manner as a propositional variable. Theorem 3.3 holds with these modifications. In the subsequent sections, we use this modified CL with the set of formulae $\mathcal{P}$. The following tautologies will be used, often without mentioning: for any $A, B, C \in \mathcal{P}$ and any finite nonempty subset Φ of $\mathcal{P}$,

(1): $\vdash_0 (A \bigwedge B \supset C) \equiv (A \supset (B \supset C))$;

(2): $\vdash_0 (A \supset B) \bigwedge (B \supset C) \supset (A \supset C)$;

(3): $\vdash_0 \neg \bigwedge \Phi \equiv \bigvee\{\neg A : A \in \Phi\}$ and $\vdash_0 \neg \bigvee \Phi \equiv \bigwedge\{\neg A : A \in \Phi\}$;

(4): $\vdash_0 (A \supset B) \equiv (\neg B \supset \neg A)$.

In the classical logic CL of Remark 3.5, it turns out that $\nvdash_0 B_i(A)$ for any $A \in \mathcal{P}$. Thus, we cannot discuss how a player arrives at his beliefs or his reasoning ability. On the other hand, game theory is particularly interested in the construction of beliefs and the reasoning ability of a player, which is the subject of the next section.

4. Various Epistemic Logics: Proof-Theoretic Approach

This sections present various epistemic logics from the proof-theoretic point of view. There are many possible logical systems and their different mathematical formulations. We discuss some of them, following the standard logic literature. However, the basic principles for epistemic logics may not be clearly seen in the standard formulations. This is caused for the reason that epistemic logics have been developed borrowing technology in the modal logic literature. For game theoretical purposes, we should have clear basic principles for epistemic logics. Therefore, we start this section with describing general principles we adopt. Then we give the standard formulation with some illustrations of game theoretical applications. In our considerations, the logical system, $KD4^n$, emerges at the central position in various systems. In addition, we present the sequent formulation $KD4^n$ in Gentzen-style, which may be skipped for reading this introduction paper but may be important for future purposes.

4.1. Basic Principles for Beliefs

The general idea for the notion of "a belief on A" is stated as:

G: player i believes A if and only if he has an argument for A (from his basic beliefs).

From the proof-theoretical point of view, we formulate "having an argument for A" as "having a proof of A". There are still various options for a formulation of "having a proof". Since our investigator is assumed to have the reasoning ability described by classical logic CL, and since CL is a stable reference point, we take the following formulation:

G1: Player i has the reasoning ability described by classical logic CL.

This means that player i has at least the same reasoning *ability* as the investigator's. That is, *if* player i has the same beliefs as the investigator's, then i can deduce what the investigator can. Note that this does not imply that the players and investigator share the same beliefs in one logical system. Since players' beliefs and reasoning abilities are described inside the investigator's logical system, some descriptions are made purely from the investigator's viewpoint and cannot be shared with players, which will be seen below.

Besides G1, we add other three components to players' beliefs and reasoning abilities: (1) Basic beliefs taken as given; (2) Intrapersonal introspective abilities; and (3) Interpersonal introspective abilities. We emphasize that basic beliefs of player i are taken as given in this paper. Formally, they are given as nonlogical axioms in the terminology in Section 3, and their development or emergence is not discussed in this paper.[13] Regarding (2) and (3), intrapersonal introspection may be still regarded as a problem of individual inferences, but interpersonal one is rather a problem of choosing assumptions on targeted situations. For these problems, there are a great spectrum of options. Here, we adopt simple and clear-cut options for these.

For the intrapersonal introspective ability, we assume:

G2: Player i has the introspective ability on his own abilities described by G1 and G2.

This requires player i only to be conscious of "having a proof". This consciousness of "having a proof" is now regarded as "having a proof of "having a proof"". Therefore, we regard G2 as forming part of basic principle G. However, some subtlety is involved in G2: G2 itself is involved in G2. To see the feasibility of G1–G2, we need an explicit mathematical formulation of them.

The nature of beliefs on other players and/or their beliefs differs considerably from that of beliefs on himself and/or his own. To a large extent, beliefs on other players are based on projecting and/or hypothesizing. In this sense, there are many options for this problem. Here we consider only the following one.

G3: When player i thinks about other players' beliefs, player i assume G1, G2 as well as G3 for the other players in a symmetric manner.

Again, G3 appears in G3 itself. This means that a player imagined in the mind of a player (imagined in the mind of another player ...) follows G3. Thus, the situation is very complicated, but, we obtain the symmetric interpersonal beliefs by assuming the limit case of complications.

To materialize our basic principle G, we need specific assumptions on the above three steps. We should keep the *remark* in mind: Principle G1 is the most basic, G2 is still basic, but G3 is one possible and convenient choice for our research strategy. Weaker possibilities of G3 are discussed together with game theoretical applications in Kaneko-Suzuki [18] of this issue.

[13] The literature of belief revisions is related to the development of basic beliefs. Some papers in this issue discuss belief revisions. Also, the development of basic beliefs is considered from the viewpoint of experiences in Kaneko-Matsui [14].

To formulate G1, it suffices to assume the beliefs on the instances of Logical Axioms L1–L5, and the axioms expressing the inference ability of player i corresponding to inference rules MP, $\bigwedge$-Rule and $\bigvee$-Rule:

BL: $\mathrm{B}_i(A)$, where A is an instance of L1–L5;

BMP: $\mathrm{B}_i(A \supset C) \bigwedge \mathrm{B}_i(A) \supset \mathrm{B}_i(C)$;

B$\bigwedge$: $\bigwedge\{\mathrm{B}_i(A \supset C) : C \in \Phi\} \supset \mathrm{B}_i(A \supset \bigwedge \Phi)$;

B$\bigvee$: $\bigwedge\{\mathrm{B}_i(C \supset A) : C \in \Phi\} \supset \mathrm{B}_i(\bigvee \Phi \supset A)$,

where A, C in BMP, B$\bigwedge$ and B$\bigvee$ are any formulae and Φ is any finite nonempty set of formulae in $\mathcal{P}$. We denote the set of all instances of BL – B$\bigvee$ by Δ_i^0. The following holds:

Theorem 4.1 (Classical Reasoning Ability). Let Γ be a set of formulae in $\mathcal{P}$ and A a formula in $\mathcal{P}$. Then $\Gamma \vdash_0 A$ if and only if $\mathrm{B}_i(\Gamma) \cup \Delta_i^0 \vdash_0 \mathrm{B}_i(A)$.

Sketch of a Proof. The *only-if* part can be proved by induction on a proof P of A from Γ. It suffices to show that for any initial formula C in P, $\mathrm{B}_i(\Gamma) \cup \Delta_i^0 \vdash_0 \mathrm{B}_i(C)$, and that the probability relation $\mathrm{B}_i(\Gamma) \cup \Delta_i^0 \vdash_0$ goes down from the upper formulae of each inference rule into the lower formula.

The *if* part needs some new concept: We define the eraser ε_i of $\mathrm{B}_i(\cdot)$: This goes into a formula and erases the outer $\mathrm{B}_i(\cdot)$ *only once* when ε_i meets $\mathrm{B}_i(\cdot)$. Let P be a proof of $\mathrm{B}_i(A)$ from $\mathrm{B}_i(\Gamma) \cup \Delta_i^0$. Then it suffices to show by induction on the tree structure of P from its leaves that for any formula C in P, $\Gamma \vdash_0 \varepsilon_i C$. ■

Thus, A is provable from Γ in classical logic CL if and only if player i can derive A from his basic beliefs Γ with his reasoning ability described by Δ_i^0. Exactly speaking, the investigator has the description Δ_i^0 of the reasoning ability of player i and deduces what i can deduce from Γ with Δ_i^0. In fact, the investigator conducts deductions in both sides of Theorem 4.1. Theorem 4.1 states that player i has given the same (potential) reasoning ability as the investigator's. Thus, we succeed in formulating G1 explicitly by this method.

For the connection of player i's inner logic to the investigator's, we add

BD: $\neg\mathrm{B}_i(\neg A \bigwedge A)$.

This states that player i has no contradictory beliefs. With this addition, but it holds with BD that Γ is consistent if and only if $\mathrm{B}_i(\Gamma) \cup \Delta_i^0 \cup \mathrm{BD}$ is consistent, where BD is now regarded as the set of instances of BD. In this sense, Axiom BD connects player i's inner logic with the investigator's logic up to their consistencies. This axiom also enables us to show that $\mathrm{B}_i(\Gamma) \cup \Delta_i^0 \cup \mathrm{BD} \vdash_0 \mathrm{B}_i(\neg A)$ implies $\mathrm{B}_i(\Gamma) \cup \Delta_i^0 \cup \mathrm{BD} \vdash_0 \neg\mathrm{B}_i(A)$. The latter is a statement from the investigator's viewpoint, but not a statement from player i's viewpoint. In general, there are many statements purely from the investigator's viewpoint.

We formulate G2 by the following two axioms:

B4: $\mathrm{B}_i(A) \supset \mathrm{B}_i\mathrm{B}_i(A)$;

BI: $\mathrm{B}_i(A)$, where A is an instance of BL - B4.

Axiom B-4 states that if player i believes A, then he believes that he believes A. This requires player i only to be conscious of having an argument for A (from his basic beliefs). Axiom BI states that player i believes that he has the logical and introspective abilities described by BL–B4. We denote, by Δ_i, the set obtained from $\Delta_i^0 \cup \mathrm{BD}$ by adding all the instances of the above axioms. The assumption set Δ_i is the formulation of principles G1 and G2.

When $n \geq 2$, the set of axioms $\Delta_1 \cup ... \cup \Delta_n$ does not yet describe the interpersonal beliefs of players described in G3. Principle G3 requires interpersonal assumptions on logical and introspective abilities, e.g., player i believes that player j has the abilities described by Δ_j, etc. The entire assumption set is given as

$$\Delta^* = \{\mathrm{B}_{i_m}...\mathrm{B}_{i_1}(A) : A \in \Delta_1 \cup ... \cup \Delta_n,\ i_m, ..., i_1 \in N \text{ and } m \geq 0\}, \tag{4.1}$$

where $\mathrm{B}_{i_m}...\mathrm{B}_{i_1}(A)$ is A itself if $m = 0$. This set states that the reasoning and introspection abilities of players described by $\Delta_1 \cup ... \cup \Delta_n$ are common knowledge.

4.2. Standard Formulation of Epistemic Logics

In the literature of modal logic, it is more standard to use one inference rule, called the *Necessitation Rule.* If we adopt the Necessitation Rule, the entire axiomatization of our epistemic logic becomes quite different from the one already given. In this subsection, we follow the standard axiomatization. Here, we add some other axioms, i.e., Axioms T and 5, to the previous axioms. We obtain various logical systems determined by combinations of these axioms, which we treat somewhat in a parallel manner.

We consider the following list of axiom schemata and inference rule, for whose names we follow the literature of modal logic: for any $i \in N$, any $A, C \in \mathcal{P}$ and any finite nonempty subset Φ of $\mathcal{P}$:

K: $\mathrm{B}_i(A \supset C) \supset (\mathrm{B}_i(A) \supset \mathrm{B}_i(C))$;[14]

D: $\neg\mathrm{B}_i(\neg A \bigwedge A)$;

T: $\mathrm{B}_i(A) \supset A$ − − − − − − − − − − truthfulness;

4: $\mathrm{B}_i(A) \supset \mathrm{B}_i\mathrm{B}_i(A)$ − − − − − − positive introspection;

5: $\neg\mathrm{B}_i(A) \supset \mathrm{B}_i(\neg\mathrm{B}_i(A))$ − − − negative introspection;

and

$$\frac{A}{\mathrm{B}_i(A)} \text{ (Necessitation).}$$

By Remark 3.5.(1), Axiom K is equivalent to BMP given above. We abbreviate the Necessitation Rule as Nec. This rule states that if A is provable, so is $\mathrm{B}_i(A)$. In the terminology of Section 4.1, this states that the investigator has a "proof of A", then player i has "a proof of A, too". It is important to notice that Nec may be repeatedly applied even with different players, e.g.,

$$\cfrac{\cfrac{A}{\mathrm{B}_i(A)}}{\cfrac{\mathrm{B}_j\mathrm{B}_i(A)}{\cdots}}$$

[14]K and D come from "*K*ripke" and "*d*eontic logic". See Chellas [5].

Thus, once A is provable, then it becomes *practically* common knowledge.

Note that Nec differs from assuming $A \supset \mathrm{B}_i(A)$ as an axiom. This difference will be clearer when we introduce nonlogical axioms. The treatment of nonlogical axioms differs from that in classical logic CL, which will be explained presently. For a moment, the reader should ignore nonlogical axioms.

The most basic system is defined as

K^n: CL + K + Nec (within $\mathcal{P}$).

A *proof* P in epistemic logic K^n is defined in the same manner as in CL except: (1) instances of K and Nec are allowed and (2) nonlogical axioms are not allowed.

Consider player 1's inference in the game g^1 of Table 2.1. Diagram 4.1 gives a proof of $\mathrm{B}_1(\bigwedge \hat{g}_1^1) \supset \mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12}))$ in K^n, where the uppermost formulae are instances of L4, the uppermost inference is $\bigwedge$-Rule, the second is Nec, the right-hand formula of the third line is an instance of Axiom K, and the last inference is MP. Note that $\bigwedge$-Rule has $|S| = 4$ upper formulae.

$$\cfrac{\cfrac{\cfrac{\{\bigwedge \hat{g}_1^1 \supset P_1(\mathrm{s}_{12}, s_2 : s_1, s_2) : (s_1, s_2) \in S\}}{\bigwedge \hat{g}_1^1 \supset \mathrm{Dom}_1(\mathrm{s}_{12})}}{\mathrm{B}_1(\bigwedge \hat{g}_1^1 \supset \mathrm{Dom}_1(\mathrm{s}_{12}))} \qquad \mathrm{B}_1(\bigwedge \hat{g}_1^1 \supset \mathrm{Dom}_1(\mathrm{s}_{12})) \supset (\mathrm{B}_1(\bigwedge \hat{g}_1^2) \supset \mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12})))}{\mathrm{B}_1(\bigwedge \hat{g}_1^1) \supset \mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12})).}$$

Diagram 4.1

That is, if 1 believes that his payoff function is $\bigwedge \hat{g}_1^1$, then he infers the belief that s_{12} is a dominant strategy. There is also a proof of $\mathrm{B}_1(\bigwedge \hat{g}_1^1) \supset \mathrm{B}_1(\neg\mathrm{Dom}_1(\mathrm{s}_{11}))$, which is derived from $\bigwedge \hat{g}_1^1 \supset \neg\mathrm{Dom}_1(\mathrm{s}_{11})$ in the same manner, but $\bigwedge \hat{g}_1^1 \supset \neg\mathrm{Dom}_1(\mathrm{s}_{11})$ needs a bit long proof.

Various epistemic logics can be defined based on K^n by choices of some of the above axiom schemata. In this paper, we consider the following list of logics:

K: CL + Nec + K;	**KD**^{n}: K^n + D;	**KT**^{n}: KD^n + T;
	KD4^n: KD^n + 4;	**S4**^{n}: $\mathrm{KD4}^n$ + T;
	KD45^n : $\mathrm{KD4}^n$ + 5;	**S5**^{n} : $\mathrm{S4}^n$ + 5.

Diagram 4.2

Let $\mathcal{S}$ be a logic in the above list. A proof in $\mathcal{S}$ is defined in the same manner as in K^n. Let A be a formula in $\mathcal{P}$. We write $\vdash_{\mathcal{S}} A$ iff there is a proof of A in $\mathcal{S}$. The following is a simple observation, which will be used without mentioning: for any $A \in \mathcal{P}$,

$$\vdash_0 A \text{ implies } \vdash_{\mathcal{S}} A \tag{4.2}$$

Note that the provability of $A \in \mathcal{P}$ is mentioned in Remark 3.5.

In the above definition of a proof, we do not allow nonlogical axioms (i.e., Γ in Section 3.3). To describe game theoretical assumptions, we introduce nonlogical axioms in a way different from in Section 3.3. Let Γ be a subset of $\mathcal{P}$ and $A \in \mathcal{P}$.

We define $\Gamma \vdash_{\mathcal{S}} A$ iff $\vdash_{\mathcal{S}} A$ or $\vdash_{\mathcal{S}} \bigwedge \Phi \supset A$ for some finite nonempty subset Φ of Γ. For the reason to avoid a nonlogical axiom in a proof, see the remark below on the Necessitation.

In classical logic CL, for a nonempty finite set Γ of formulae, $\Gamma \models A$ is equivalent to $\models \bigwedge \Gamma \supset A$, which together with the soundness-completeness for CL implies $\Gamma \vdash_0 A$ is equivalent to $\vdash_0 \bigwedge \Gamma \supset A$. Hence it follows from (4.2) that

$$\Gamma \vdash_0 A \text{ implies } \Gamma \vdash_{\mathcal{S}} A. \tag{4.3}$$

The strengths of the provabilities of the above logics are described as follows:

$$\begin{array}{ccccc}
\mathrm{K}^n & \rightarrow & \mathrm{KD}^n & \rightarrow & \mathrm{KT}^n \\
 & & \downarrow & & \downarrow \\
 & & \mathrm{KD4}^n & \rightarrow & \mathrm{S4}^n \\
 & & \downarrow & & \downarrow \\
 & & \mathrm{KD45}^n & \rightarrow & \mathrm{S5}^n
\end{array}$$

Diagram 4.3

where the expression, $\mathcal{S} \rightarrow \mathcal{S}'$, means that the provability of $\mathcal{S}'$ is stronger than that of $\mathcal{S}$, for example, $\vdash_{\mathrm{KD4}^n} A$ implies $\vdash_{\mathrm{S4}^n} A$. Diagram 4.1 is a legitimate proof in all $\mathcal{S}$'s.

The following are basic properties of the epistemic logics of the above list.

Lemma 4.2. For any $A, C \in \mathcal{P}$ and a nonempty finite subset Φ of $\mathcal{P}$,

(1): $\vdash_{\mathcal{S}} \mathrm{B}_i(A \supset C) \bigwedge \mathrm{B}_i(A) \supset \mathrm{B}_i(C)$;

(2): $\vdash_{\mathcal{S}} \bigvee \mathrm{B}_i(\Phi) \supset \mathrm{B}_i(\bigvee \Phi)$, where $\mathrm{B}_i(\Phi) = \{\mathrm{B}_i(A) : A \in \Phi\}$;

(3): $\vdash_{\mathcal{S}} \mathrm{B}_i(\bigwedge \Phi) \equiv \bigwedge \mathrm{B}_i(\Phi)$;

(4): $\vdash_{\mathcal{S}} \mathrm{B}_i(\neg A) \supset \neg \mathrm{B}_i(A)$, where $\mathcal{S}$ includes Axiom D;

(5): if $\Gamma \vdash_{\mathcal{S}} A$, then $\mathrm{B}_i(\Gamma) \vdash_{\mathcal{S}} \mathrm{B}_i(A)$.

Proof. We prove (1) – (4).

(1): This follows from Axiom K and Remark 3.5.(1).

(2): Let A be an arbitrary formula in Φ. Since $\vdash_{\mathcal{S}} A \supset \bigvee \Phi$ by L5, we have $\vdash_{\mathcal{S}} \mathrm{B}_i(A \supset \bigvee \Phi)$ by Nec. By K, $\vdash_{\mathcal{S}} \mathrm{B}_i(A) \supset \mathrm{B}_i(\bigvee \Phi)$. Since this holds for any $A \in \Phi$, we have $\vdash_{\mathcal{S}} \bigvee \mathrm{B}_i(\Phi) \supset \mathrm{B}_i(\bigvee \Phi)$ by $\bigvee$-Rule

(3): $\vdash_{\mathcal{S}} \mathrm{B}_i(\bigwedge \Phi) \supset \bigwedge \mathrm{B}_i(\Phi)$ is the dual of (2). We prove the converse only for $\Phi = \{A, C\}$. Since $\vdash_{\mathcal{S}} A \supset (C \supset A \bigwedge C)$, we have $\vdash_{\mathcal{S}} \mathrm{B}_i(A) \supset (\mathrm{B}_i(C) \supset \mathrm{B}_i(A \bigwedge C))$, using Nec, K and MP a few times. This is equivalent to $\vdash_{\mathcal{S}} \mathrm{B}_i(A) \bigwedge \mathrm{B}_i(C) \supset \mathrm{B}_i(A \bigwedge C)$. When Φ has more than two formulae, we should prove $\vdash_{\mathcal{S}} \bigwedge \mathrm{B}_i(\Phi) \supset \mathrm{B}_i(\bigwedge \Phi)$ by induction on the number of formulae in Φ.

(4): By (3), $\vdash_{\mathcal{S}} \neg \mathrm{B}_i(\neg A \bigwedge A) \equiv \neg(\mathrm{B}_i(\neg A) \bigwedge \mathrm{B}_i(A))$. By Axiom D, we have $\vdash_{\mathcal{S}} \neg(\mathrm{B}_i(\neg A) \bigwedge \mathrm{B}_i(A))$. This is equivalent to $\vdash_{\mathcal{S}} \mathrm{B}_i(\neg A) \supset \neg \mathrm{B}_i(A)$. ∎

We will use the following facts extensively without mentioning.

Lemma 4.3. For any $A, B, C \in \mathcal{P}$ and a finite nonempty subset Φ of $\mathcal{P}$, (1): $\vdash_S A \supset B$ and $\vdash_S B \supset C$ imply $\vdash_S A \supset C$;

(2): $\Gamma \vdash_S A$ for all $A \in \Phi$ if and only if $\Gamma \vdash_S \bigwedge \Phi$.

Proof. (1) follows from Remark 3.5.(2). The *if* part of (2) follows from L4 and MP. The converse is proved by using $\bigwedge$-Rule and MP (L1 in the case of $\Gamma = \emptyset$). ■

Now, let us compare the present formulation of epistemic logics with the formulation given in Section 4.1. The following theorem is due to Kaneko-Nagashima [15]. We omit the proof of it.

Theorem 4.4.(1)(Reasoning Ability described G1 and G2): Let $n = 1$. Then for any $A \in \mathcal{P}$, $\Delta_1 \vdash_0 A$ if and only if $\vdash_{KD4^1} A$.

(2)(Reasoning Abilities described G1, G2 and G3): Let $n \geq 2$. Then for any $A \in \mathcal{P}$, $\Delta^* \vdash_0 A$ if and only if $\vdash_{KD4^n} A$.

In the uni-modal case, $KD4^1$ corresponds to the logic describing basic principles G1 and G2.[15] In the multi-modal case, the common knowledge of $\Delta_1 \cup ... \cup \Delta_n$ describing G1 and G2 for all players is the basic assumptions for $KD4^n$.

Theorem 4.4 helps us understand the basic principle for our epistemic logics, especially, $KD4^n$, but technically speaking, it does not help us evaluate the axioms listed. On the one hand, Theorem 4.4 would hold even if we add each or both of Axioms T and 5 to both sides. On the other hand, if we delete Axiom 4 from both sides, then the left-hand side needs to be modified for the above equivalence. In this case, G2 has a slightly different content. To evaluate various axioms, we should return to the original principle G.

Our purpose is to consider epistemic aspects of decision making in game situations. For this, it does not suffice to consider only mathematical properties of such logics and applications. We need to discriminate some logics as more appropriate than others. We adopt the Inference Rule Nec, Axioms K and D as very basic. We reject Axiom 5 as inappropriate: this rejection is made by recalling our basic principle G. We avoid Axiom T so as to allow false beliefs, but can treat Axiom 4 inside our logics without Axiom 5.

Here, we give remarks on Nec and Axioms T, 4, 5.

Necessitation: We note that when nonlogical axioms are involved, Nec is applied to the whole formula $\bigwedge \Phi \supset A$ and yields $B_i(\bigwedge \Phi \supset A)$. Since Nec can be applied arbitrarily many times, $\bigwedge \Phi \supset A$ becomes common knowledge effectively in the sense that $B_{i_m}...B_{i_1}(\bigwedge \Phi \supset A)$ are all derived.[16] Nevertheless, this does not means that the assumption $\bigwedge \Phi$ becomes common knowledge, but that only the implication $\bigwedge \Phi \supset A$ becomes common knowledge. This note is related to the reason for the introduction of nonlogical axioms Φ in the present form. If $\bigwedge \Phi$ is assumed as an initial formula in a proof, then $\bigwedge \Phi$ becomes common knowledge. To avoid this, we have introduced nonlogical axioms as the premise of $\bigwedge \Phi \supset A$.

Axiom K changes $B_i(\bigwedge \Phi \supset A)$ into $B_i(\bigwedge \Phi) \supset B_i(A)$. The former states that i believes $\bigwedge \Phi \supset A$, while the latter states that if i believes $\bigwedge \Phi$, he believes A,

[15]From Theorem 4.1 and Theorem 4.4, we can regard $KD4^1$ as a logic of provability and introspection. However, it is slightly different from the logic so called *provability logic* in the literature (see Boolos [4]).

[16]We can give a restriction on such applications. See the second paper, [18], of this issue.

too. Thus, Axiom K transforms a statement from i's viewpoint into one from the investigator's.

The difference between Nec and $A \supset \mathrm{B}_i(A)$ (as a logical axiom) may be now clear. If $A \supset \mathrm{B}_i(A)$ is assumed as a logical axiom, then from $\vdash_S \bigwedge \Phi \supset A$, we obtain $\vdash_S \bigwedge \Phi \supset \mathrm{B}_i(A)$, $\vdash_S \bigwedge \Phi \supset \mathrm{B}_{i_m}...\mathrm{B}_{i_1}(A)$. Without $A \supset \mathrm{B}_i(A)$, however, we can obtain only $\vdash_S \mathrm{B}_{i_m}...\mathrm{B}_{i_1}(\bigwedge \Phi) \supset \mathrm{B}_{i_m}...\mathrm{B}_{i_1}(A)$.

Axiom T: This distinguishes knowledge from beliefs. With this axiom, beliefs are always true relative to the thinker (ultimately to the investigator), and without it, beliefs may be false. It is important to discuss the truth and/or falsity of beliefs in the future studies of economics and game theory. Axiom T prohibits the possibility of talking about the falsity of beliefs.

Axiom 4: Although we have adopted this axiom to describe part of G2, we do not think that this is so basic as Nec, K and D. It will be used once in game theoretical arguments in Section 7, but is avoidable with a slightly longer argument. The reasons for this reservation are: First, if we want to examine the role of self-consciousness, we should do it in a logic without Axiom 4. Also, a logic without Axiom 4 would be easier to handle in meta-theoretical respects. So far, we do not have enough developments in theory and applications to give clear distinctions between logics with and without Axiom 4.[17]

Axiom 5: This axiom has been used in game theoretical literature of epistemic models. However, we do not take this as a basic axiom. The reason for this judgement is its negative premise, $\neg \mathrm{B}_i(A)$. According to basic principle G, this means that

(*): the investigator thinks that it is not the case that player i has a proof of A (from i' basic beliefs).

For (*), the investigator needs to examine all the possible (infinitely many) candidates for a proof of A. Notice that this situation differs considerably from the situation suggested by Axiom 4. In this situation, Axiom 5 asserts that the investigator's position is taken by player i himself. Hence, when $\neg \mathrm{B}_i(A)$ holds, Axiom 5 does not allow the possibility:

(**): player i does not have a proof of "not having a proof of A".

Thus, Axiom 5 excludes the possibility that both $\neg \mathrm{B}_i(A)$ and $\neg \mathrm{B}_i(\neg \mathrm{B}_i(A))$ hold. We would like to keep both (*) and (**) as separating the investigator's viewpoint from player i's viewpoint. That is, we do not adopt Axiom 5 in our logical approach. Finally, we note that the contrapositive of Axiom 5 is $\neg \mathrm{B}_i(\neg \mathrm{B}_i(A)) \supset \mathrm{B}_i(A)$ (if Axiom T is additionally assumed, these become equivalent).

4.3. Some Game Theoretical Problems

Let us see how the decision criteria DC1 and DC2 of Section 2 are formulated in epistemic logic $\mathcal{S}$.

[17] We treated intrapersonal and interpersonal introspections separately in Section 4.1. In game theoretical practices, assumptions of interpersonal beliefs play significant roles but not much intrapersonal introspection. In the human history, however, self-consciousness might be evolved as the ability to derive interpersonal beliefs (Mithen [24], pp.217-219). This may give a hint to reconsider the role of Axiom 4 and/or principle G2.

Since $\vdash_{\mathcal{S}} \mathrm{B}_1(\bigwedge \hat{g}_1^1) \supset \mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12}))$ by Diagram 4.1, we have, by Lemma 4.2.(3), $\vdash_{\mathcal{S}} \bigwedge \mathrm{B}_1(\hat{g}_1^1) \supset \mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12}))$. Following our convention of nonlogical axioms in epistemic logic $\mathcal{S}$, we have

$$\mathrm{B}_1(\hat{g}_1^1) \vdash_{\mathcal{S}} \mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12})). \tag{4.4}$$

Similarly, $\mathrm{B}_1(\hat{g}_1^2) \vdash_{\mathcal{S}} \mathrm{B}_1(\neg\mathrm{Dom}_1(\mathrm{s}_{12}))$. In logic $\mathcal{S}$ including Axiom D, we have, by Lemma 4.2.(4),

$$\mathrm{B}_1(\hat{g}_1^1) \vdash_{\mathcal{S}} \neg\mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{11})). \tag{4.5}$$

Hence 1's belief on his own payoff function is enough to decide s_{12} to be the unique dominant strategy.

The counterpart of (3.5) in $\mathcal{S}$ is expected to be:

$$\mathrm{B}_1(\hat{g}_1^1) \nvdash_{\mathcal{S}} \mathrm{B}_1(\mathrm{Nash}(\mathrm{s}_{12}, \mathrm{s}_{22})). \tag{4.6}$$

That is, 1's beliefs $\hat{g}_1^1$ are not enough to answer the question of whether $(\mathrm{s}_{12}, \mathrm{s}_{22})$ is believed to be a Nash equilibrium. In fact, it would be difficult to prove this kind of unprovability purely within epistemic logic $\mathcal{S}$. One way of considering such an unprovability is to use semantics, which will be discussed in Section 5.

Next, we see the formulation of DC2 in $\mathcal{S}$. Consider decision making of player 2 in the game g^2 of Table 2.2. Since $g_1^2 = g_1^1$, it follows from (4.4) and Lemma 4.2.(5) that $\mathrm{B}_2\mathrm{B}_1(\hat{g}_1^2) \vdash_{\mathcal{S}} \mathrm{B}_2\mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12}))$. Also, we have $\mathrm{B}_2(\hat{g}_2^2) \vdash_{\mathcal{S}} \mathrm{B}_2(\mathrm{Best}_2(\mathrm{s}_{22} \mid \mathrm{s}_{12}))$. We combine these two into

$$\mathrm{B}_2\mathrm{B}_1(\hat{g}_1^2), \mathrm{B}_2(\hat{g}_2^2) \vdash_{\mathcal{S}} \mathrm{B}_2\mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12})) \bigwedge \mathrm{B}_2(\mathrm{Best}_2(\mathrm{s}_{22} \mid \mathrm{s}_{12})), \tag{4.7}$$

where we abbreviate $\mathrm{B}_2\mathrm{B}_1(\hat{g}_1^2) \cup \mathrm{B}_2(\hat{g}_2^2)$ as $\mathrm{B}_2\mathrm{B}_1(\hat{g}_1^2)$, $\mathrm{B}_2(\hat{g}_2^2)$. That is, player 2, predicting 1's decision to be s_{12}, would choose s_{22} as a best strategy to s_{12}. This precise formulation of DC2 is possible in epistemic logic $\mathcal{S}$, but not in classical logic CL.

4.4. Gentzen-Style Formulation of $\mathrm{KD4}^n$

The reader can skip this subsection in order to read the remaining of this paper. This subsection is written for the reader who wants to go further to advances in the logic approach to game theory. We give the sequent calculus formulation of $\mathrm{KD4}^n$ in the Gentzen-style. Some other papers in this issue adopt this style. Although it is deductively equivalent to the Hilbert-style formulation, it adds another sense of "logical reality". The following is a very brief introduction. The interested reader may consult some textbooks such as Kleene [19] and Takeuti [30]. Nevertheless, the best introduction may be still Gentzen's [10] original article.

We introduce the concept of a sequent. Let Γ, Θ be finite subsets of $\mathcal{P}$. Using auxiliary symbols [,], and $\rightarrow$, we introduce a new expression $\Gamma \rightarrow \Theta$, which we call a *sequent*. We abbreviate $\Gamma \cup \Delta \rightarrow \Lambda \cup \Theta$ and $\{A\} \cup \Gamma \rightarrow \Theta \cup \{C\}$, respectively, as $\Gamma, \Delta \rightarrow \Lambda, \Theta$ and $A, \Gamma \rightarrow \Theta, C$, etc.

A sequent $\Gamma \rightarrow \Theta$ is associated with each node in a proof in the Gentzen-style. Here, Γ is a set of nonlogical axioms. Thus, a set of nonlogical axioms appears in every step in a proof. The counterpart of $\Gamma \rightarrow \Theta$ in the Hilbert-style formulation is $\bigwedge \Gamma \supset \bigvee \Theta$, where $\bigwedge \emptyset$ and $\bigvee \emptyset$ are meant to be $\neg p \bigvee p$ and $\neg p \bigwedge p$, respectively.

We will explain the relationship of the present formulation with the Hilbert-style epistemic logic KD4^n.

In the following, $\Gamma, \Theta, \Delta, \Lambda, \Phi$ are finite sets of formulae, A, B formulae and Φ is also assumed to be nonempty.

Axiom (Initial Sequent): $A \rightarrow A$,

Structural Rules:

$$\frac{\Gamma \rightarrow \Theta}{\Delta, \Gamma \rightarrow \Theta, \Lambda}\ (\mathrm{Th}) \qquad \frac{\Gamma \rightarrow \Theta, A] \quad A, \Delta \rightarrow \Lambda}{\Gamma, \Delta \rightarrow \Theta, \Lambda}\ (\mathrm{Cut})$$

Operational Rules:

$$\frac{\Gamma \rightarrow \Theta, A}{\neg A, \Gamma \rightarrow \Theta}\ (\neg \rightarrow) \qquad \frac{A, \Gamma \rightarrow \Theta}{\Gamma \rightarrow \Theta, \neg A}\ (\rightarrow \neg)$$

$$\frac{\Gamma \rightarrow \Theta, A \quad B, \Delta \rightarrow \Lambda}{A \supset B, \Gamma, \Delta \rightarrow \Theta, \Lambda}\ (\supset\rightarrow) \qquad \frac{A, \Gamma \rightarrow \Theta, B}{\Gamma \rightarrow \Theta, A \supset B}\ (\rightarrow\supset)$$

$$\frac{A, \Gamma \rightarrow \Theta}{\bigwedge \Phi, \Gamma \rightarrow \Theta}\ (\bigwedge \rightarrow) \text{ where } A \in \Phi \qquad \frac{\{\Gamma \rightarrow \Theta, A : A \in \Phi\}}{\Gamma \rightarrow \Theta, \bigwedge \Phi}\ (\rightarrow \bigwedge)$$

$$\frac{\{A, \Gamma \rightarrow \Theta : A \in \Phi\}}{\bigvee \Phi, \Gamma \rightarrow \Theta}\ (\bigvee \rightarrow) \qquad \frac{\Gamma \rightarrow \Theta, A}{\Gamma \rightarrow \Theta, \bigvee \Phi}\ (\rightarrow \bigvee) \text{ where } A \in \Phi.$$

Epistemic Rule (Necessitation Rule):

$$\frac{\Gamma, \mathrm{B}_i(\Delta) \rightarrow \Theta}{\mathrm{B}_i(\Gamma), \mathrm{B}_i(\Delta) \rightarrow \mathrm{B}_i(\Theta)}\ (\mathrm{B} \rightarrow \mathrm{B}), \text{ where } |\Theta| \leq 1 \text{ and } i \in N,$$

where $|\Theta|$ is the cardinality of Θ.

A *proof* P of $\Gamma \rightarrow \Theta$ in the present system is defined to be a tree in the same manner as in the previous Hilbert-style. That is, a sequent is associated with each node of P, the sequent associated with each leaf of P is an initial sequent, some instances of the inference rules connect nodes of P, and $\Gamma \rightarrow \Theta$ is associated with the root of P. We say that $\Gamma \rightarrow \Theta$ is *provable in* KD4^n, denoted by $\vdash_{\mathrm{KD4}^n} \Gamma \rightarrow \Theta$, iff there is a proof P of $\Gamma \rightarrow \Theta$.

We may regard sequent $\mathrm{B}_1(\hat{g}_1^1) \rightarrow \mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12}))$ as a counterpart of $\mathrm{B}_1(\bigwedge \hat{g}_1^1) \supset \mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12}))$ of (4.5), which is proved as follows:

$$\cfrac{\cfrac{\left\{ \cfrac{P_1(\mathrm{s}_{12}, s_2 : s_1, s_2) \rightarrow P_1(\mathrm{s}_{12}, s_2 : s_1, s_2)}{\hat{g}_1^1 \rightarrow P_1(\mathrm{s}_{12}, s_2 : s_1, s_2)}\ (\mathrm{Th}) \right\}_{(s_1, s_2)}}{\hat{g}_1^1 \rightarrow \mathrm{Dom}_1(\mathrm{s}_{12})}\ (\rightarrow \bigwedge)}{\mathrm{B}_1(\hat{g}_1^1) \rightarrow \mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12}))}\ (\mathrm{B} \rightarrow \mathrm{B})$$

Another counterpart of (4.5) is a sequent $\rightarrow \mathrm{B}_1(\bigwedge \hat{g}_1^1) \supset \mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12}))$, which can be proved similarly to the above.

The relation between the Gentzen-style of and Hilbert-style of KD4^n is as follows.

Theorem 4.5 (Relation to $KD4^n$ in the Hilbert-style). Let Γ and Θ be finite sets of formulae. Then $\vdash_{KD4^n} \Gamma \to \Theta$ if and only if $\vdash_{KD4^n} \bigwedge\Gamma \supset \bigvee\Theta$. In the latter, $\bigwedge\emptyset$ and $\bigvee\emptyset$ are understood as $\neg p \bigvee p$ and $\neg p \bigwedge p$.

The following *cut-elimination theorem* is the main theorem for the Gentzen-style formulation $KD4^n$. It makes the system essentially different from the Hilbert-style formulation of $KD4^n$.

Theorem 4.6 (Cut-Elimination). If $\vdash_{KD4^n} \Gamma \to \Theta$, then there is a cut-free proof P of $\Gamma \to \Theta$.

The cut-elimination theorem states that if a sequent is provable, then we can find a proof of the same endsequent without using (Cut). All the inference rules given above except (Cut) add new symbols from the upper sequent(s) to the lower sequent. Therefore, in a cut-free proof P, any formula occurring at some point in P also occurs as a subformula in the endsequent. This is called the *subformula property*. On the other hand, the Hilbert-style formulation has inference rule Modus Pones, at which one formula is eliminated. In this case, we cannot trace back from a given provable formula what have happened in a proof.

The cut-elimination theorem was first proved for the classical logic by Gentzen [10], and then it was proved for many other systems. The above one is a variant of the cut-elimination theorem for $S4^1$ and some others given by Ohnishi-Matsumoto [28]. One remark is that it is not successful to have the cut-elimination theorem for $S5^n$.

The following is a proof $A, A \supset B, B \supset C \to C$ with a (Cut).

$$\dfrac{\dfrac{A \to A \quad B \to B}{A, A \supset B \to B}(\supset\to) \qquad \dfrac{B \to B \quad C \to C}{B, B \supset C \to C}(\supset\to)}{A, A \supset B, B \supset C \to C}(\text{Cut})$$

This can be proved without using (Cut).

$$\dfrac{\dfrac{A \to A \quad \dfrac{B \to B}{B \to C, B}(\text{Th})}{A, A \supset B \to C, B}(\supset\to) \qquad \dfrac{A \to A \quad \dfrac{C \to C}{B, C \to C}(\text{Th})}{A, A \supset B, C \to C}(\supset\to)}{A, A \supset B, B \supset C \to C}(\supset\to)$$

5. Kripke Semantics: Model-Theoretic Approach

In this section, the basic principle G: "having an argument for A" for the belief on A is formulated in a model-theoretic manner, that is, as "A is a consequence in all the possible models of player i's basic beliefs". Mathematically, this is formulated in the Kripke semantics. The basic principle G1 of Section 4.1 is clear-cut, but the other G2, G3 are less clear-cut. Nevertheless, this model-theoretic approach has some technical advantages over the proof-theoretic approach given in Sections 4.1 and 4.2. As already stated, it is practically difficult to prove the unprovability assertion, (4.6), directly in epistemic logic $\mathcal{S}$. The more complex a formula is, the more difficult to evaluate provability is in $\mathcal{S}$. The Kripke semantics gives a model theory to each epistemic logic $\mathcal{S}$ in a systematic manner, and enables us to evaluate unprovabilities such as (4.6).[18].

5.1. Kripke Models and Completeness

We say that an $n+1$ tuple $(W : R_1, ..., R_n)$ is a *Kripke frame* iff W is an arbitrary nonempty set and each R_i is a binary relation on W. Each element w in W is called a *possible world*, and each R_i an *accessibility relation*. In each world w, we assume the classical truth valuation, i.e., the logical connectives $\neg, \supset, \bigwedge, \bigvee$ are valuated in the manner CV0–CV4. This means that each player has the reasoning ability described by classical logic in each possible world, which expresses the basic principle G1 of Section 4.1. The other principles G2 and G3 are expressed by the relationships to other possible worlds defined by $R_1, ..., R_n$. Each R_i describes the possibilities to valuate $\mathrm{B}_i(A)$ by referring to the truthfulness of A in the worlds accessible from w.

An *assignment in* a Kripke frame $\mathcal{K} = (W : R_1, ..., R_n)$ is a function σ from $W \times PV$ to $\{\top, \bot\}$. A pair $(\mathcal{K}, \sigma)$ of a Kripke frame $\mathcal{K}$ and an assignment σ is called a *Kripke model*. We define the *valuation relation* $(\mathcal{K}, \sigma, w) \models$ and its negation $(\mathcal{K}, \sigma, w) \not\models$ for each $w \in W$ by induction on the length of a formula:

K0: for each $p \in PV$, $(\mathcal{K}, \sigma, w) \models p$ iff $\sigma(w,p) = \top$;

K1: $(\mathcal{K}, \sigma, w) \models \neg A$ iff $(\mathcal{K}, \sigma, w) \not\models A$;

K2: $(\mathcal{K}, \sigma, w) \models A \supset B$ iff $(\mathcal{K}, \sigma, w) \not\models A$ or $(\mathcal{K}, \sigma, w) \models B$;

K3: $(\mathcal{K}, \sigma, w) \models \bigwedge \Phi$ iff $(\mathcal{K}, \sigma, w) \models A$ for all $A \in \Phi$;

K4: $(\mathcal{K}, \sigma, w) \models \bigvee \Phi$ iff $(\mathcal{K}, \sigma, w) \models A$ for some $A \in \Phi$;

K5: $(\mathcal{K}, \sigma, w) \models \mathrm{B}_i(A)$ iff $(\mathcal{K}, \sigma, u) \models A$ for all u with wR_iu.

This inductive definition works simultaneously over the possible worlds. We say that A is *true at world* w in $(\mathcal{K}, \sigma)$ iff $(\mathcal{K}, \sigma, w) \models A$. This valuation is complete in the sense that for any $w \in W$ and $A \in \mathcal{P}$,

$$\text{either } (\mathcal{K}, \sigma, w) \models A \text{ or } (\mathcal{K}, \sigma, w) \models \neg A. \tag{5.1}$$

Step K5 expresses the idea that the truth of $\mathrm{B}_i(A)$ in world w is defined by referring to the truth of A in the accessible worlds from w. Specifically, we note that when $(\mathcal{K}, \sigma, u) \models \neg A$ for some u with wR_iu, we have $(\mathcal{K}, \sigma, u) \not\models \mathrm{B}_i(A)$ by K5, and then $(\mathcal{K}, \sigma, u) \models \neg\mathrm{B}_i(A)$ by K1.

Since a Kripke model is a complete description of the target world, we need to consider a set of Kripke models for connection to the proof theory. We consider a set of Kripke models defined by conditions on the accessibility relations $R_1, ..., R_n$. Then we will have an explicit connection to each proof-theory given in Section 4 to the Kripke semantics.

Specifically, we consider the following conditions on $R_1, ..., R_n$, each of which corresponds to an epistemic axiom:

K – – no condition;

[18]Chellas [5] and Hughes-Cresswell [11] are good textbooks on Kripke semantics, which treat unimodal logics. For multi-modal epistemic logics, see Fagin, et al [6] and Meyer-van der Hoek [23].

D -- *seriality*: for any $w \in W$, there is some $u \in W$ with wR_iu;

T -- *reflexivity*: wR_iw for all $w \in W$;

4 -- *transitivity*: for all $w, u, v \in W$, wR_iu and uR_iv imply wR_iv;

5 -- *euclidean*: for all $w, u, v \in W$, wR_iu and wR_iv imply uR_iv.

We postpone seeing the reasons for these correspondences after the main result. Let $\mathcal{S}$ be an epistemic logic in Diagram 4.2. Then $\mathcal{S}^*$ is the set of all Kripke frames satisfying the conditions on the accessibility relations corresponding to $\mathcal{S}$. For example, if $\mathcal{S}$ is KD4^n, $\mathcal{S}^*$ is the set of Kripke frames satisfying seriality and transitivity.[19] Using this notation, we can state the following soundness-completeness theorem.

Theorem 5.1 (Soundness-Completeness). Let $\mathcal{S}$ be an epistemic logic in Diagram 4.2, and A a formula in $\mathcal{P}$.

(1): $\vdash_{\mathcal{S}} A$ if and only if $(\mathcal{K}, \sigma, w) \models A$ for all Kripke frames $\mathcal{K}$ in $\mathcal{S}^*$, all assignments σ and all $w \in W$.

(2): There is a Kripke frame $\mathcal{K}$ in $\mathcal{S}^*$, an assignment σ and a world $w \in W$ satisfying $(\mathcal{K}, \sigma, w) \models A$ if and only if A is consistent in logic $\mathcal{S}$.

Let Γ be a finite nonempty set. Since $\Gamma \vdash_{\mathcal{S}} A$ is defined by $\vdash_{\mathcal{S}} \bigwedge \Gamma \supset A$, Theorem 5.(1) implies

(1^A): $\Gamma \vdash_{\mathcal{S}} A$ if and only if for all $\mathcal{K}$ in $\mathcal{S}^*, \sigma$ and $w \in W$, $(\mathcal{K}, \sigma, w) \models C$ for all $C \in \Gamma$ imply $(\mathcal{K}, \sigma, w) \models A$.

In the present context, a model of Γ is $(\mathcal{K}, \sigma)$ making all assumptions in Γ true at some world w. Using this terminology, (2) is written: there is a Kripke model of A if and only if A is consistent in logic $\mathcal{S}$.

By Theorem 5.1, we have the equivalence between the provability $\vdash_{\mathcal{S}}$ defined by means of symbol manipulations and the consequence relation $\models$ defined in terms of a set of possible models (complete descriptions). This equivalence is important and useful not only in understanding player i's belief on A as "having an argument for A" but also in investigating the properties of $\vdash_{\mathcal{S}}$ and $\models$. In fact, without the Kripke semantics, we would not go much further than the results in Section 4. The usefulness of the Kripke semantics will be shown by some examples and some other results below.

Let us see the reasons for the correspondences between the epistemic axioms and the conditions on the accessibility relation R_i. In the Kripke semantics, the basic idea for $\mathrm{B}_i(A)$ is that the truthfulness of $\mathrm{B}_i(A)$ at w is determined by looking at the truthfulness of A in all possible worlds accessible from w. Also, in each world, the classical valuation is assumed. Axiom K: $(\mathcal{K}, \sigma, w) \models \mathrm{B}_i(A \supset C) \bigwedge \mathrm{B}_i(A) \supset \mathrm{B}_i(C)$ is derived only by this fact, that is, in any referred world, the truthfulness is

[19] We find the reason for the popularity of S5^n among game theorists. Reflexivity and euclidean imply symmetry, i.e., for all $w, u \in W$, wR_iu implies uR_iw. If $\mathcal{S}$ is S5^n, the conditions for accessibility relations are reflexivity, transitivity, symmetry. Hence, each R_i becomes an equivalence relation. Hence the quotient space $W/R_i = \{\{w \in W : wR_iu\} : u \in W\}$ is a partition of W for $i = 1, ..., n$. The $n+1$-tuple $(W : W/R_1, ..., W/R_n)$ may be regarded as an information partition model (of Aumann [1]). However, a Kripke model describes the possibilities perceived by players but does not describe information processing.

closed under Modus Ponens. Axiom D: $(\mathcal{K},\sigma,w) \models \neg \mathrm{B}_i(\neg A \bigwedge A)$ is derived by the fact that a contradictory formula is not true in *some* accessible world from w, which needs seriality. Axiom T: $(\mathcal{K},\sigma,w) \models \mathrm{B}_i(A) \supset A$ requires that the accessible worlds from w include w itself: otherwise, the truthfulness of $\mathrm{B}_i(A)$ should be independent of that of A. Axiom 4: $(\mathcal{K},\sigma,w) \models \mathrm{B}_i(A) \supset \mathrm{B}_i\mathrm{B}_i(A)$ requires that the accessible worlds be closed with R_i in the sense that if u is accessible from w by finite steps of R_i, u is already accessible directly by R_i. In the same manner, the corresponding conditions to Axiom 5 is understood.

Now we exemplify the above theorem by proving (4.6) with Kripke models. The game (g_1^5, g_2^5) of Table 5.1 has no Nash equilibrium, and the game (g_1^1, g_2^5) of Table 5.2 has the unique Nash equilibrium $(\mathrm{s}_{12}, \mathrm{s}_{21})$. Consider the Kripke model $(\mathcal{K},\sigma)$ described as Diagram 5.3. It is read as follows: Each arrow indexed by i connects the possible worlds with R_i, i.e., $R_i = \{(w_0,w_1),(w_1,w_1)\}$ for $i = 1,2$, and the assignment σ is determined by the set associated with each world, i.e., for any atomic formula $p \in AF$, $\sigma(w_0,p) = \top$ iff $p \in \hat{g}_1^5 \cup \hat{g}_2^5$ and $\sigma(w_1,p) = \top$ iff $p \in \hat{g}_1^1 \cup \hat{g}_2^5$.

	s_{21}	s_{22}
s_{11}	(6,0)	(3,1)
s_{12}	(5,1)	(1,0)

Table 5.1: (g_1^5, g_2^5)

	s_{21}	s_{22}
s_{11}	(5,0)	(1,1)
s_{12}	(6,1)*	(3,0)

Table 5.2: (g_1^1, g_2^5)

$\circlearrowleft_{1,2}\ w_1 : \hat{g}_1^1 \cup \hat{g}_2^5$
$\uparrow_{1,2}$
$w_0 : \hat{g}_1^5 \cup \hat{g}_2^5$

Diagram 5.3: $(\mathcal{K},\sigma)$

This $\mathcal{K}$ is serial and transitive. Since $(\mathcal{K},\sigma,w_1) \models \neg\mathrm{Nash}(\mathrm{s}_{12},\mathrm{s}_{22})$, we have $(\mathcal{K},\sigma,w_1) \models \neg\mathrm{B}_1(\mathrm{Nash}(\mathrm{s}_{12},\mathrm{s}_{22}))$. Also, $(\mathcal{K},\sigma,w_1) \models \mathrm{B}_1(\hat{g}_1^1)$. Hence, $\mathrm{B}_1(\hat{g}_1^1) \nvdash_{\mathrm{KD4}^n} \mathrm{B}_1(\mathrm{Nash}(\mathrm{s}_{12},\mathrm{s}_{22}))$ by (1^A).

The two claims of Theorem 5.1 are equivalent. The *if* part of either claim is called *completeness*, which will be proved in Section 9. The *only-if* part, called *soundness*, is proved as follows:

Lemma 5.2. Let $\mathcal{K} = (W : R_1, ..., R_n)$ in $\mathcal{S}^*$ and σ any assignment in $\mathcal{K}$.

(1): Let A be an instance of L1–L5 or an instance of an epistemic axiom, e.g., $\mathrm{B}_i(A \supset C) \supset (\mathrm{B}_i(A) \supset \mathrm{B}_i(C))$, in $\mathcal{S}$. Then $(\mathcal{K},\sigma,w) \models A$ for any world $w \in W$.

(2): For any $w \in W$, $(\mathcal{K},\sigma,w) \models$ satisfies MP, $\bigwedge$-Rule, and $\bigvee$-Rule, e.g., if $(\mathcal{K},\sigma,w) \models A \supset B$ and $(\mathcal{K},\sigma,w) \models A$, then $(\mathcal{K},\sigma,w) \models B$.

(3): If $(\mathcal{K},\sigma,w) \models A$ for all $w \in W$, then $(\mathcal{K},\sigma,w) \models \mathrm{B}_i(A)$ for all $w \in W$.

Proof. We prove only (3). Let $(\mathcal{K},\sigma,v) \models A$ for all $v \in W$. Let w be any world in W. Since $(\mathcal{K},\sigma,u) \models A$ for all u with wR_iu, we have $(\mathcal{K},\sigma,w) \models \mathrm{B}_i(A)$. ■

Proof of the Only-If Part of (1) of Theorem 5.1. Let P be a proof of A. Then we show, by induction on the tree structure of P from its leaves, the assertion that for any C occurring in P, $(\mathcal{K},\sigma,w) \models C$ for any $\mathcal{K}$ in $\mathcal{S}^*$, assignment σ and world w in $\mathcal{K}$. Each of the inductive steps is verified by Lemma 5.2. ■

The contradiction-freeness of $\mathcal{S}$ follows from the soundness part of Theorem 5.1.(1).

Theorem 5.3 (Contradiction-Freeness). Each $\mathcal{S}$ in Diagram 4.2 is contradiction-free, i.e., there is no formula A in $\mathcal{P}$ such that $\vdash_{\mathcal{S}} \neg A \bigwedge A$.

Proof. Suppose $\vdash_{\mathcal{S}} \neg A \bigwedge A$ for some A. By Theorem 5.1.(1), we have $(\mathcal{K}, \sigma, w) \models \neg A \bigwedge A$ for all $w \in W$ in $\mathcal{K}$, all σ in $\mathcal{K}$ and all $\mathcal{K}$ in $\mathcal{S}^*$. However, this is impossible. ■

As already stated, if $\Gamma \vdash_0 A$, then $\Gamma \vdash_{\mathcal{S}} A$. When Γ and A are nonepistemic, the converse follows from Theorem 5.1, in which case $\mathcal{S}$ is said to be a *conservative extension* of CL.

Theorem 5.4 (Conservativity of $\mathcal{S}$ upon CL). Let $\Gamma \subseteq \mathcal{P}^n$ and $A \in \mathcal{P}^n$. Then $\Gamma \vdash_{\mathcal{S}} A$ if and only if $\Gamma \vdash_0 A$.

5.2. Decision Criterion DC1 with False Beliefs

Here, let us return to a game theoretical problem, specifically, consider DC1 of Section 2 from the viewpoint of false beliefs. In DC1, neither player predicts the other's decision, and each player's own decision making is relevant for $i = 1, 2$. Here we adopt the dominant strategy criterion, $\mathrm{B}_i(\mathrm{Dom}_i(s_i))$, which we denote by $\hat{D}_{ii}(s_i)$. Let us apply $\hat{D}_{ii}(s_i)$ to the game g^1 of Table 2.1. By combining (4.4) and (4.5), we have

$$\mathrm{B}_1(g_1^1), \mathrm{B}_2^1(g_2^1) \vdash_{\mathcal{S}} \bigwedge_i \left(\neg \hat{D}_{ii}(\mathrm{s}_{i1}) \bigwedge \hat{D}_{ii}(\mathrm{s}_{i2})\right). \tag{5.2}$$

That is, each player i infers from the belief on his payoff function that his decision is his second strategy.

The objectivity of the payoff functions is not included in (5.2). If epistemic logic $\mathcal{S}$ includes Axiom T, the assumption set automatically implies that the game to be played is $g^1 = (g_1^1, g_2^1)$. Now, let us adopt KD4^n as $\mathcal{S}$. Then it is possible to add any payoff functions g_i on $\{\mathrm{s}_{11}, \mathrm{s}_{12}\} \times \{\mathrm{s}_{12}, \mathrm{s}_{22}\}$ to (5.2) as an objective one. In other words, the beliefs $\mathrm{B}_1(g_1^1)$, $\mathrm{B}_2^1(g_2^1)$ are false in most cases, for example, we can assume that the true game is $g^2 = (g_1^2, g_2^2)$ rather than $g^1 = (g_1^1, g_2^1)$:

$$\hat{g}^2, \mathrm{B}_1(\hat{g}_1^1), \mathrm{B}_2^1(\hat{g}_2^1) \vdash_{\mathcal{S}} \bigwedge_i \left(\neg \hat{D}_{ii}(\mathrm{s}_{i1}) \bigwedge \hat{D}_{ii}(\mathrm{s}_{i2})\right). \tag{5.3}$$

In this case, 1's belief is true but 2's is false.

Assertion (5.3) is meaningful only if $\hat{g}^2 \cup \mathrm{B}_1(\hat{g}_1^1) \cup \mathrm{B}_2^1(\hat{g}_2^1)$ is consistent in KD4^n. This consistency can be proved by constructing a model for $\hat{g}^2 \cup \mathrm{B}_1(\hat{g}_1^1) \cup \mathrm{B}_2^1(\hat{g}_2^1)$:

$$\begin{array}{c} \circlearrowleft_{1,2}\ w_1 : \hat{g}^1 \\ \uparrow_{1,2} \\ w_0 : \hat{g}^2 \end{array}$$

Diagram 5.4: $(\mathcal{K}, \sigma)$

where σ assigns $\top$ to each atomic formula included in $\hat{g}^1$ at w_1 and to one in $\hat{g}^2$ at w_0. This $\mathcal{K} = (W : R_1, R_2)$ is serial and transitive frame. Then $(\mathcal{K}, \sigma, w_0) \models \hat{g}^2 \cup \mathrm{B}_1(\hat{g}_1^1) \cup \mathrm{B}_2^1(\hat{g}_2^1)$. Hence $\hat{g}^2 \cup \mathrm{B}_1(\hat{g}_1^1) \cup \mathrm{B}_2^1(\hat{g}_2^1)$ is consistent by (1^A) after Theorem 5.1. Since this frame $\mathcal{K}$ does not satisfy reflexivity, $(\mathcal{K}, \sigma)$ is not a model in S4^n. Hence the above consistency proof can be converted to S4^n.

6. Beliefs VS Knowledge

As mentioned already, we adopt the distinction between *beliefs and knowledge* that knowledge is a true belief, while a belief may be false. Here, the truth is referred to the outside thinker, ultimately, the investigator. For example, in $\mathrm{B}_i(A)$, the thinker i *only* believes the truth of A, and the investigator determines the truth of A. In $\mathrm{B}_j\mathrm{B}_i(A)$, thinker j can determine the truth of player i's belief A by referring to j's belief on A. In this case, the investigator may determine the truths of A, $\mathrm{B}_i(A)$ and $\mathrm{B}_j\mathrm{B}_i(A)$.[20] Once Axiom T is assumed, beliefs are true to any outside thinkers. To take human interpersonal epistemic interactions seriously, we would like to allow false beliefs. The situation with Axiom T must be a special case. Therefore, we treat Axiom T in a relativistic manner.

In fact, Axiom T can be treated inside an epistemic logic without it. Let $\mathcal{S}$ be an epistemic logic without Axioms T and 5, and $\mathcal{S}'$ the logic obtained from $\mathcal{S}$ by adding Axiom T. Then we show that $\mathcal{S}'$ is embedded into $\mathcal{S}$. This result guarantees that we capture the distinction between beliefs and knowledge in $\mathcal{S}$ by the procedure for the embedding. In this section, we show the embedding theorem for logics $\mathrm{S4}^n$ and $\mathrm{KD4}^n$. This embedding would not hold with the presence of Axiom 5.

In epistemic logic $\mathcal{S}$, we denote formula $\mathrm{B}_i(A)\bigwedge A$ by $\mathrm{B}_i^+(A)$. It means that player i believes A and A is true to the outside thinker. For example, in $\mathrm{B}_j\mathrm{B}_i^+(A) = \mathrm{B}_j(\mathrm{B}_i(A)\bigwedge A)$, player j thinks about the truth of i's belief comparing with his belief on A. The formula $\mathrm{B}_i^+(A)$ satisfies

$\mathbf{T}^+$: $\vdash_{\mathcal{S}} \mathrm{B}_i^+(A) \supset A$;

$\mathbf{K}^+$: $\vdash_{\mathcal{S}} \mathrm{B}_i^+(A \supset C) \supset (\mathrm{B}_i^+(A) \supset \mathrm{B}_i^+(C))$;

$\mathbf{4}^+$: $\vdash_{\mathcal{S}} \mathrm{B}_i^+(A) \supset \mathrm{B}_i^+\mathrm{B}_i^+(A)$, where Axiom 4 is included in $\mathcal{S}$;

$\mathbf{Nec}^+$: if $\vdash_{\mathcal{S}} A$, then $\vdash_{\mathcal{S}} \mathrm{B}_i^+(A)$.

Thus, if $\mathcal{S}$ includes Axiom 4, the operator $\mathrm{B}_i^+(\cdot)$ behaves like an operator in $\mathrm{S4}^n$. In this definition, however, only the outermost $\mathrm{B}_i(A)$ is replaced by $\mathrm{B}_i^+(A)$, but A may include other $\mathrm{B}_j(\cdot)$. Hence, we cannot yet regard $\mathrm{B}_i^+(\cdot)$ exactly as the operator in $\mathrm{S4}^n$. To have the exact relationship, we need to have a more accurate translation. Now, we focus on the case between $\mathrm{S4}^n$ and $\mathrm{KD4}^n$ (between KT^n and KD^n).

To avoid confusions, we differentiate the formulae for $\mathrm{S4}^n$ from those for $\mathrm{KD4}^n$. We denote, by $\mathcal{P}_{\mathrm{K}}$, the set of all formulae generated by the same list of symbols of Section 3 except for the substitutions of new operator symbols $\mathrm{K}_1,...,\mathrm{K}_n$ for $\mathrm{B}_1,...,\mathrm{B}_n$. Here $\mathrm{K}_i(A)$ is intended to mean that i knows A.

Now we define a translator $\psi : \mathcal{P}_{\mathrm{K}} \to \mathcal{P}$ by the following induction:

T0: for any $p \in PV$, $\psi(p) = p$;

[20] According to philosophical literature, knowledge is defined as "justified true beliefs". In this, justification needs some different sources of authority such as experiences or community (see Moser [25] for debates on justifications of beliefs). Objects targeted by epistemic logics are beliefs inferred from basic beliefs (assumptions or axioms) which are abstracted from experiences. Justifications for inferred beliefs are traced back to those on the basic beliefs. However, when we consider justifications for basic beliefs, we cannot go further to any other sources. In sum, justifications are relativistic notions. To have such justifications, we need a general framework including experiences.

T1: $\psi(\neg A) = \neg\psi A$;

T2: $\psi(A \supset C) = \psi A \supset \psi C$;

T3: $\psi(\bigwedge \Phi) = \bigwedge\{\psi A : A \in \Phi\}$; and $\psi(\bigvee \Phi) = \bigvee\{\psi A : A \in \Phi\}$;

T4: $\psi(\mathrm{K}_i(A)) = \mathrm{B}_i^+(\psi A)$ for $i = 1, ..., n$.

That is, any formula A in $\mathcal{P}_{\mathrm{K}}$ is translated into the corresponding formula in $\mathcal{P}$ which is obtained from A by replacing all occurrences of subformulae of the form $\mathrm{K}_i(C)$ by $\mathrm{B}_i^+(C^*)$, where C^* is obtained from C by the same principle. For example, $\psi(\mathrm{K}_2\mathrm{K}_1(\mathrm{Dom}_1(\mathrm{s}_{12})) = \mathrm{B}_2^+\mathrm{B}_1^+(\mathrm{Dom}_1(\mathrm{s}_{12}))$, which is equivalent to $\mathrm{B}_2\mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12})) \bigwedge \mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12})) \bigwedge \mathrm{B}_2(\mathrm{Dom}_1(\mathrm{s}_{12})) \bigwedge \mathrm{Dom}_1(\mathrm{s}_{12})$.

We have the following theorem, which will be proved in the end of this section.[21]

Theorem 6.1 (Faithful Embedding of S4^n into KD4^n). For any $A \in \mathcal{P}_{\mathrm{K}}$, $\vdash_{\mathrm{S4}^n} A$ if and only if $\vdash_{\mathrm{KD4}^n} \psi A$.

Thus, we can discuss logic S4^n inside KD4^n. When we can forget false beliefs for some game theoretical problems, discussions in S4^n are simpler than in KD4^n in that Axiom T can be used in S4^n. Such results in S4^n can be translated into more general discussions with the presence of false beliefs in KD4^n by the above embedding theorem. Conversely, the theorem may be used to translate some meta-theorems obtained for KD4^n into S4^n: KD4^n is easier than S4^n from the meta-theoretical point of view. One example will be mentioned in Section 7.2. The same embedding assertion holds between KD^n and KT^n. Over all, S4^n and KT^n can be considered inside KD4^n and KD^n, respectively.[22]

Consider the above embedding theorem taking (4.7) as an example: (4.7) holds in $\mathcal{S} = \mathrm{S4}^n$ and is translated by ψ into the following:

$$\mathrm{B}_2^+\mathrm{B}_1^+(\hat{g}_1^2), \mathrm{B}_2^+(\hat{g}_2^2) \vdash_{\mathrm{KD4}^n} \mathrm{B}_2^+\mathrm{B}_1^+(\mathrm{Dom}_1(\mathrm{s}_{12})) \bigwedge \mathrm{B}_2^+(\mathrm{Best}_2(\mathrm{s}_{22} \mid \mathrm{s}_{12})). \tag{6.1}$$

Of course, (6.1) differs from (4.7) in $\mathcal{S} = \mathrm{KD4}^n$. In (6.1), the truth of 1's belief on $\hat{g}_1^2$ is, first, referred to the belief of 2, and then, the truth of 2's belief on $\hat{g}_1^2$ is referred to the investigator's.

In S4^n, all $\mathrm{K}_i(\cdot)$'s are effectively replaced by $\mathrm{B}_i^+(\cdot)$, and it is impossible to distinguish between knowledge and beliefs. However, we can keep this distinction in KD4^n. For example, the following holds:

$$\Gamma \vdash_{\mathrm{KD4}^n} \mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12})) \bigwedge \mathrm{B}_2\mathrm{B}_1^+(\mathrm{Dom}_1(\mathrm{s}_{12})) \bigwedge \mathrm{B}_2(\mathrm{Best}_2(\mathrm{s}_{22} \mid \mathrm{s}_{12})).$$

where $\Gamma = \hat{g}^5 \cup \mathrm{B}_1(\hat{g}_1^2) \cup \mathrm{B}_2\mathrm{B}_1^+(\hat{g}_1^2) \cup \mathrm{B}_2(\hat{g}_2^2)$. The consistency of Γ can be proved in the same manner as in Section 5.2. In Γ, player 2 believes that player 1 has true beliefs on 1's payoff function. Accordingly, player 2 believes that 1's inferred belief, $\mathrm{B}_1(\mathrm{Dom}_1(\mathrm{s}_{12}))$, is also true. Nevertheless, the basic beliefs of both players are all false.

[21] This embedding theorem is a sort of folk theorem among nonclassical logicians.

[22] This embedding theorem fails with the presence of Axiom 5. For example, S5^n cannot be embedded into KD45^n. A counterexample is $\nvdash_{\mathrm{KD45}^n} \neg\mathrm{B}_1^+(\neg p) \supset \mathrm{B}_1^+(\neg\mathrm{B}_1^+(\neg p)))$, where $p \in PV$. This is proved by constructing a Kripke model.

Proof of Theorem 6.1. The *only-if* part of this theorem can be proved by induction on a proof in $S4^n$ from its leaves, using the above T^+, K^+, 4^+ and Nec^+.

The *if* part needs two steps. We define another translator $\varphi : \mathcal{P} \to \mathcal{P}_K$ by T0–T3 and T4$'$: $\varphi(B_i(C)) = K_i(\varphi C)$ for $i = 1, ..., n$. That is, φ is the operator which simply substitutes K_i for all occurrences of B_i in A. Then we can prove that for any $C \in \mathcal{P}$,

$$\vdash_{KD4^n} C \text{ implies } \vdash_{S4^n} \varphi(C). \tag{6.2}$$

This can be proved by induction on a proof in $KD4^n$.

The second step is the following assertion: for any $A \in \mathcal{P}_K$,

$$\vdash_{S4^n} \varphi \cdot \psi(A) \equiv A. \tag{6.3}$$

This is proved by induction on the structure of a formula. For any $p \in PV$, $\varphi \cdot \psi(p)$ is p itself, which implies $\vdash_{S4^n} \varphi \cdot \psi(p) \equiv p$.

Suppose the induction hypothesis that (6.3) holds for any immediate subformulae of A. We should consider the cases: $\neg, \supset, \bigwedge, \bigvee$ and K_i. Here we consider only $\neg$ and K_i.

($\neg$): Let $A = \neg C$. By the induction hypothesis, we have $\vdash_{S4^n} \varphi \cdot \psi(C) \equiv C$. Then $\vdash_{S4^n} \neg\varphi \cdot \psi(C) \equiv \neg C$. Since $\neg\varphi \cdot \psi(C)$ is $\varphi \cdot \psi(\neg C)$ by the definitions of φ and ψ, we have $\vdash_{S4^n} \varphi \cdot \psi(\neg C) \equiv \neg C$.

(K_i): Let $A = K_i(C)$. By the induction hypothesis, we have $\vdash_{S4^n} \varphi \cdot \psi(C) \equiv C$. Hence $\vdash_{S4^n} K_i(\varphi \cdot \psi(C)) \equiv K_i(C)$ by Nec, MP and Axiom K. By the definition of φ and ψ, we have $\varphi \cdot \psi(K_i(C)) = \varphi(B_i(\psi(C)) \bigwedge \psi(C)) = K_i(\varphi \cdot \psi(C)) \bigwedge \varphi \cdot \psi(C)$. Since $\vdash_{S4^n} K_i(\varphi \cdot \psi(C)) \bigwedge \varphi \cdot \psi(C) \equiv K_i(\varphi \cdot \psi(C))$, we have $\vdash_{S4^n} \varphi \cdot \psi(K_i(C)) \equiv K_i(C)$.

It follows from (6.2) that for any $A \in \mathcal{P}_K$, $\vdash_{KD4^n} \psi A$ implies $\vdash_{S4^n} \varphi \cdot \psi(A)$, and the latter is equivalent to $\vdash_{S4^n} A$ by (6.3). ■

7. Solution Theories for DC2 and DC3

As yet our game theoretical consideration was about performance-playability relative to a player's beliefs, taking the dominant strategy criterion as given. Here, we discuss solution theories for DC2 and DC3. In the recent game theoretic terminology, these are axiomatic considerations of decision making. From the descriptive point of view, there must be a lot of possible decision-prediction criteria in that a lot of arbitrariness are involved in behavioral criteria, particularly, in predictions on other players' behavior. In this section, we consider only mathematical formulations of DC2 and DC3. Some other criteria are discussed in Kaneko-Suzuki [18].

For DC2, we consider solution theory and also talk briefly about its performance-playability. Although we give somewhat tedious proofs of an axiomatic characterization of DC2, the point is neither in the characterization nor in the proof, but is in the comparisons with DC3. For DC3, we meet a difficulty caused by an infinite regress of beliefs. This difficulty lead us to an extension of epistemic logic $\mathcal{S}$ to incorporate common knowledge, which is the subject of Section 8.

Throughout this section, we assume $\mathcal{S} = KD4^n$. Let $\{D_{ij}(s_j) : s_j \in S_j$ and $i, j = 1, 2\}$ be a set of formulae indexed by $s_j \in S_j$ and $i, j = 1, 2$. These describe the decision and prediction of player i, that is, it is intended that for $j = i$, $D_{ii}(s_i)$

holds if and only if s_i is player i's decision, and for $j \neq i$, $D_{ij}(s_j)$ holds if and only if i predicts that s_j would be a decision of player j.

7.1. Decision Criterion DC2

For the consideration of decision criterion DC2, we let $i = 2$ and $j = 1$. In this case, only player 2 predicts 1's decision. We assume that 1's decision criterion is given as $\hat{D}_{11}(s_1) = \mathrm{B}_1(\mathrm{Dom}_1(s_1))$. We require 2's decision and prediction to satisfy:

$$\mathrm{DC21}_2\text{: } \bigwedge_{s_1,s_2} (D_{22}(s_2) \wedge D_{21}(s_1) \supset \mathrm{B}_2(\mathrm{Best}_2(s_2 \mid s_1)))\,;$$

$$\mathrm{DC22}_2\text{: } \bigwedge_{s_1} (D_{21}(s_1) \supset \mathrm{B}_2(D_{21}(s_1))) \wedge \bigwedge_{s_2} (D_{22}(s_2) \supset \mathrm{B}_2(D_{22}(s_2)))\,;$$

$$\mathrm{DC23}_2\text{: } \left(\bigvee_{s_2} D_{22}(s_2) \supset \bigvee_{s_1} D_{21}(s_1)\right) \wedge \left(\bigvee_{s_1} D_{21}(s_1) \supset \bigvee_{s_2} D_{22}(s_2)\right);$$

$$\mathrm{DC24}_2\text{: } \bigwedge_{s_1} \left(D_{21}(s_1) \supset \mathrm{B}_2(\hat{D}_{11}(s_1))\right).$$

The first states that if player 2 predicts that 1 would choose s_1, then 2 believes that his decision s_2 is a best strategy against s_1. The second requires player 2 to be conscious of his own decision and prediction. The third states that 2's decision is possible if and only if so is 2's prediction. The last states that 2's prediction is based on $\hat{D}_{11}(s_1) = \mathrm{B}_1(\mathrm{Dom}_1(s_1))$.

Our problem is to find formulae $D_{21}(s_1)$ and $D_{22}(s_2)$ ($s_1 \in S_1$ and $s_2 \in S_2$) satisfying the above four requirements. In fact, contradictory formulae, i.e., the *deductively strongest formulae*, satisfy these requirements in the sense that if we substitute $\neg A \wedge A$ for $D_{22}(s_2)$ and $D_{21}(s_1)$, these would be provable in $\mathcal{S} = \mathrm{KD4}^n$. However, we would like to find the *deductively weakest formulae* satisfying these requirements, since they have no additional properties other than what the requirements describe.[23] One set of candidates are

$$\hat{D}_{21}(s_1) := \mathrm{B}_2\mathrm{B}_1(\mathrm{Dom}_1(s_1));$$

$$\hat{D}_{22}(s_2) := \bigvee_{t_1} \mathrm{B}_2\mathrm{B}_1(\mathrm{Dom}_1(t_1)) \wedge \bigwedge_{t_1} \mathrm{B}_2 \left(\mathrm{B}_1(\mathrm{Dom}_1(t_1)) \supset \mathrm{Best}_2(s_2 \mid t_1)\right).$$

The formula $\hat{D}_{22}(s_2)$ states that 2 has a prediction t_1, and that whatever his prediction t_1 is, 2's decision s_2 is a best response to t_1. These formulae satisfies the above requirements $\mathrm{DC21}_2$,..., $\mathrm{DC24}_2$ in the following sense. First, we denoted, by $\mathrm{DC21}_2(\hat{D})$,..., $\mathrm{DC24}_2(\hat{D})$, the formulae obtained from $\mathrm{DC21}_2$,..., $\mathrm{DC24}_2$ by plugging $\hat{D}_{21}(s_1)$ and $\hat{D}_{22}(s_2)$ into $D_{21}(s_1)$ and $D_{22}(s_2)$. Then we have the following.

Theorem 7.1.(1): $\vdash_{\mathcal{S}} \mathrm{DC21}_2(\hat{D}) \wedge \mathrm{DC22}_2(\hat{D}) \wedge \mathrm{DC24}_2(\hat{D})$.

(2): Let $g = (g_1, g_2)$ be any 2-person game having a unique dominant strategy for player 1. Then $\mathrm{B}_2\mathrm{B}_1(\hat{g}_1), \mathrm{B}_2(\hat{g}_2) \vdash_{\mathcal{S}} \mathrm{DC23}_2(\hat{D})$.

[23] We take the method of looking for the deductively weakest formulae satisfying Axioms $\mathrm{DC21}_2$–$\mathrm{DC24}_2$. This process itself is not described in epistemic logic $\mathrm{KD4}^n$. If we introduce new symbols $D_{ij}(s_j), i,j = 1,2, s_j \in S_j$, then we can treat these axioms inside $\mathrm{KD4}^n$. See Kaneko [13] for this method.

Proof.(1): We prove only $\vdash_S \mathrm{DC21}_2(\hat{D})$. Let (s_1, s_2) be any strategy pair. Then $\vdash_S \hat{D}_{22}(s_2) \supset \mathrm{B}_2(\mathrm{B}_1(\mathrm{Dom}_1(s_1)) \supset \mathrm{Best}_2(s_2 \mid s_1))$ by the definition of $\hat{D}_{22}(s_2)$. Hence $\vdash_S \hat{D}_{22}(s_2) \supset (\mathrm{B}_2\mathrm{B}_1(\mathrm{Dom}_1(s_1)) \supset \mathrm{B}_2(\mathrm{Best}_2(s_2 \mid s_1)))$. This is equivalent to $\vdash_S \hat{D}_{22}(s_2) \bigwedge \mathrm{B}_2\mathrm{B}_1(\mathrm{Dom}_1(s_1)) \supset \mathrm{B}_2(\mathrm{Best}_2(s_2 \mid s_1))$. Thus, $\vdash_S \mathrm{DC21}_2(\hat{D})$.

(2): It is easy to see $\vdash_S \bigvee_{s_2} D_{22}(s_2) \supset \bigvee_{s_1} D_{21}(s_1)$. We prove $\mathrm{B}_2\mathrm{B}_1(\hat{g}_1), \mathrm{B}_2(\hat{g}_2) \vdash_S \bigvee_{s_1} D_{21}(s_1) \supset \bigvee_{s_2} D_{22}(s_2)$. Let t_1^* be the unique dominant strategy for player 1 in g, and t_2 a best response to t_1^*. Since $\mathrm{B}_1(\hat{g}_1) \vdash_S \mathrm{B}_1(\mathrm{Dom}_1(t_1^*))$ and $\mathrm{B}_1(\hat{g}_1) \vdash_S \neg\mathrm{B}_1(\mathrm{Dom}_1(t_1))$ for any $t_1 \neq t_1^*$, we have $\mathrm{B}_1(\hat{g}_1), \hat{g}_2 \vdash_S \mathrm{B}_1(\mathrm{Dom}_1(t_1)) \supset \mathrm{Best}_2(t_2 \mid t_1)$ for all t_1. Thus, $\mathrm{B}_2\mathrm{B}_1(\hat{g}_1)$, $\mathrm{B}_2(\hat{g}_2) \vdash_S \mathrm{B}_2(\mathrm{B}_1(\mathrm{Dom}_1(t_1)) \supset \mathrm{Best}_2(t_2 \mid t_1))$ for all t_1. Hence, $\mathrm{B}_2\mathrm{B}_1(\hat{g}_1)$, $\mathrm{B}_2(\hat{g}_2) \vdash_S \bigwedge_{t_1}\mathrm{B}_2(\mathrm{B}_1(\mathrm{Dom}_1(t_1)) \supset \mathrm{Best}_2(t_2 \mid t_1))$. Hence $\mathrm{B}_2\mathrm{B}_1(\hat{g}_1), \mathrm{B}_2(\hat{g}_2) \vdash_S \bigvee_{s_1} D_{21}(s_1) \supset (\bigvee_{s_1} D_{21}(s_1)) \bigwedge \bigwedge_{t_1}\mathrm{B}_2(\mathrm{B}_1(\mathrm{Dom}_1(t_1)) \supset \mathrm{Best}_2(t_2 \mid t_1))$, i.e., $\mathrm{B}_2\mathrm{B}_1(\hat{g}_1), \mathrm{B}_2(\hat{g}_2) \vdash_S \bigvee_{s_1} D_{21}(s_1) \supset D_{22}(t_2)$. Hence $\mathrm{B}_2\mathrm{B}_1(\hat{g}_1), \mathrm{B}_2(\hat{g}_2) \vdash_S \bigvee_{s_1} D_{21}(s_1) \supset \bigvee_{s_2} D_{22}(s_2)$. ■

Thus, $\hat{D}_{21}(s_1)$ and $\hat{D}_{22}(s_2)$ are candidate formulae for the axioms $\mathrm{DC21}_2, \ldots, \mathrm{DC24}_2$. Conversely, $\hat{D}_{21}(s_1)$ and $\hat{D}_{22}(s_2)$ are the deductively weakest formulae satisfying $\mathrm{DC21}_2, \ldots, \mathrm{DC24}_2$. Hence, $\hat{D}_{21}(s_1)$ and $\hat{D}_{22}(s_2)$ are what we look for.

Theorem 7.2 (Personalized Characterization of DC2). Let $g = (g_1, g_2)$ be any 2-person game having a unique dominant strategy for player 1. If $\mathrm{B}_2(\hat{g}_2), \mathrm{B}_2\mathrm{B}_1(\hat{g}_1) \vdash_S \mathrm{DC21}_2 \bigwedge \ldots \bigwedge \mathrm{DC24}_2$, then

$$\mathrm{B}_2(\hat{g}_2), \mathrm{B}_2\mathrm{B}_1(\hat{g}_1) \vdash_S \bigwedge_{s_1} \left(D_{21}(s_1) \supset \hat{D}_{21}(s_1)\right) \bigwedge \bigwedge_{s_2} \left(D_{22}(s_2) \supset \hat{D}_{22}(s_2)\right). \qquad (7.1)$$

Proof. The former half follows from $\mathrm{DC24}_2$. Let $\Gamma = \mathrm{B}_2(\hat{g}_2) \cup \mathrm{B}_2\mathrm{B}_1(\hat{g}_1)$. For the latter, it suffices to show $\Gamma \vdash_S D_{22}(s_2) \supset \hat{D}_{22}(s_2)$, where s_2 is any strategy for 2. By the former half, $\Gamma \vdash_S D_{22}(s_2) \supset \bigvee_{s_2} \hat{D}_{21}(s_1)$. Hence, it remains to show $\Gamma \vdash_S D_{22}(s_2) \supset \bigwedge_{t_1}\mathrm{B}_2(\mathrm{B}_1(\mathrm{Dom}_1(t_1)) \supset \mathrm{Best}_2(s_2 \mid t_1))$. Let t_1^* be the dominant strategy for 1.

Since $\hat{g}_2, \mathrm{B}_1(\hat{g}_1) \vdash_S \mathrm{Best}_2(t_2 \mid t_1^*) \supset \bigwedge_{t_1}(\mathrm{B}_1(\mathrm{Dom}_1(t_1)) \supset \mathrm{Best}_2(t_2 \mid t_1))$, we have $\mathrm{B}_2(\hat{g}_2), \mathrm{B}_2\mathrm{B}_1(\hat{g}_1) \vdash_S \mathrm{B}_2(\mathrm{Best}_2(t_2 \mid t_1^*)) \supset \bigwedge_{t_1}\mathrm{B}_2(\mathrm{B}_1(\mathrm{Dom}_1(t_1)) \supset \mathrm{Best}_2(t_2 \mid t_1))$. Then $\Gamma \vdash_S \mathrm{B}_2(\mathrm{Best}_2(t_2 \mid s_1)) \bigwedge \mathrm{B}_2\mathrm{B}_1(\mathrm{Dom}_1(s_1)) \supset \bigwedge_{t_1}\mathrm{B}_2(\mathrm{B}_1(\mathrm{Dom}_1(t_1)) \supset \mathrm{Best}_2(t_2 \mid t_1))$. Since $\Gamma \vdash_S D_{22}(s_2) \bigwedge D_{21}(s_1) \supset \mathrm{B}_2(\mathrm{Best}_2(s_2 \mid s_1)) \bigwedge \mathrm{B}_2\mathrm{B}_1(\mathrm{Dom}_1(s_1))$, we have $\Gamma \vdash_S D_{22}(s_2) \bigwedge D_{21}(s_1) \supset \bigwedge_{t_1}\mathrm{B}_2(\mathrm{B}_1(\mathrm{Dom}_1(t_1)) \supset \mathrm{Best}_2(s_2 \mid t_1))$. This is written as $\Gamma \vdash_S D_{21}(s_1) \supset (D_{22}(s_2) \supset \bigwedge_{t_1}\mathrm{B}_2(\mathrm{B}_1(\mathrm{Dom}_1(t_1)) \supset \mathrm{Best}_2(s_2 \mid t_1))$, and then $\Gamma \vdash_S \bigvee_{s_2} D_{21}(s_1) \supset (D_{22}(s_2) \supset \bigwedge_{t_1}\mathrm{B}_2(\mathrm{B}_1(\mathrm{Dom}_1(t_1)) \supset \mathrm{Best}_2(s_2 \mid t_1))$. Since $\Gamma \vdash_S D_{22}(s_2) \supset \bigvee_{s_2} D_{22}(s_2)$, we have $\Gamma \vdash_S D_{22}(s_2) \supset \bigwedge_{t_1}\mathrm{B}_2(\mathrm{B}_1(\mathrm{Dom}_1(t_1)) \supset \mathrm{Best}_2(s_2 \mid t_1))$. ■

It follows from Theorem 7.2 that the deductively weakest formulae $\{D_{2j}(s_j) : s_j \in S_j \text{ and } j = 1, 2\}$ satisfying $\mathrm{DC21}_2 \bigwedge \ldots \bigwedge \mathrm{DC24}_2$ are *uniquely* determined. That is, if $\{D_{2j}^*(s_j) : s_j \in S_j \text{ and } j = 1, 2\}$ satisfy $\mathrm{DC21}_2 \bigwedge \ldots \bigwedge \mathrm{DC24}_2$, and if

$$\mathrm{B}_2(\hat{g}_2), \mathrm{B}_2\mathrm{B}_1(\hat{g}_1) \vdash_S \bigwedge_{s_1} (D_{21}(s_1) \supset D_{21}^*(s_1)) \bigwedge \bigwedge_{s_2} (D_{22}(s_2) \supset D_{22}^*(s_2)), \qquad (7.2)$$

then $\mathrm{B}_2(\hat{g}_2), \mathrm{B}_2\mathrm{B}_1(\hat{g}_1) \vdash_S \bigwedge_{s_j} \left(\hat{D}_{ij}(s_j) \equiv D_{ij}^*(s_j)\right)$.

The assumption that a game g has a unique dominant strategy for player 1 can be relaxed as follows:

$$\hat{g} \vdash_0 \bigwedge_{t_1, t_1'} \bigwedge_{t_2} (\mathrm{Dom}_1(t_1) \bigwedge \mathrm{Dom}_1(t_1') \bigwedge \mathrm{Best}_2(t_2 \mid t_1) \supset \mathrm{Best}_2(t_2 \mid t_1')). \qquad (7.3)$$

That is, though player 2 predicts multiple dominant strategies for player 1, this multiplicity causes no problem for player 2 in game $g = (g_1, g_2)$. If (7.3) is violated, then we should modify player 2' decision criterion or add more information to guarantee him to have a decision.[24]

Let us have a brief look at performance-playability of $\hat{D}_{11}(s_1)$ for 1 and $\hat{D}_{21}(s_1)$, $\hat{D}_{22}(s_2)$ for 2. Suppose that the game played is g^2 of Table 2.2, player 1 believes his payoff function is g_1^5 of Table 5.1, player 2 believes that the game played is g^2 and 1 truly believes his payoff function g_2^2. In this case, both players reach decisions:

$$\hat{g}^2, \mathrm{B}_1(\hat{g}_1^5), \mathrm{B}_1(\hat{g}_2^2), \mathrm{B}_2\mathrm{B}_1^+(\hat{g}_2^2) \vdash_S \left(\bigvee_{s_1} \hat{D}_{11}(s_1)\right) \bigwedge \left(\bigvee_{s_1} \hat{D}_{22}(s_2)\right). \tag{7.4}$$

However, 2 predicts that 1 would choose s_{12}, which differs from 1's decision, i.e.,

$$\hat{g}^2, \mathrm{B}_1(\hat{g}_1^5), \mathrm{B}_2(\hat{g}_2^2), \mathrm{B}_2\mathrm{B}_1^+(\hat{g}_1^2) \vdash_S \hat{D}_{11}(s_{11}) \bigwedge \neg\hat{D}_{21}(s_{11}) \bigwedge \hat{D}_{21}(s_{12}). \tag{7.5}$$

After the play of the game, 2 notices that his prediction was not played by player 1. Consider what 2 may infer from this observation. In this case, player 1 doubts his beliefs on player 1's payoff function and his prediction criterion $\hat{D}_{21}(s_1)$. Suppose that 2 still adopts $\hat{D}_{21}(s_1)$. Then 2 can consider three possibilities: (a) 2' belief $\mathrm{B}_1(\hat{g}_1^2)$ is correct but 2' belief $\hat{g}_1^2$ on 1' payoff function is not, (b): the former is incorrect but the latter is correct, and (c) both are incorrect. Player 2 cannot conclude by his observation which is the case.

The above example shows that our logical approach is powerful in the consideration of such problems. However, the above example suggests an intrinsic difficulty of making beliefs on other players' beliefs from experiential sources.

Finally, consider the case where 1 and 2 adopt decision criteria DC3^0 and DC2, i.e., DC2&3^0 of Section 2. Player 1's DC3^0 is written as $\bigvee_{s_2} \mathrm{B}_1(\mathrm{Nash}(s_1, s_2))$. In this case, if player 1 has beliefs $\mathrm{B}_1(\hat{g}_1^2 \cup \hat{g}_2^2)$, then

$$\mathrm{B}_1(\hat{g}_1^2 \cup \hat{g}_2^2) \vdash_S \bigvee_{s_2} \mathrm{B}_1(\mathrm{Nash}(s_{12}, s_2)).$$

We have, also,

$$\mathrm{B}_1(\hat{g}_1^2 \cup \hat{g}_2^2), \mathrm{B}_2(\hat{g}_2^2), \mathrm{B}_2\mathrm{B}_1(\hat{g}_1^2) \vdash_S \bigvee_{s_2} \mathrm{B}_1(\mathrm{Nash}(s_{12}, s_2)) \bigwedge \hat{D}_{22}(s_{22})$$

Thus, the two players can make decisions without yielding a contradiction. This consistency is proved by constructing a Kripke model of $\mathrm{B}_1(\hat{g}_1^2 \cup \hat{g}_2^2)$,$\mathrm{B}_2(\hat{g}_2^2)$,$\mathrm{B}_2\mathrm{B}_1(\hat{g}_1^2)$. In game g^2, no behavioral problems occur, since the action taken by 1 is "correctly" predicted by player 2. When we replace g^2 by g^3, no logical problems occur, but behaviorally, player 1 can make a decision, though 2 cannot make prediction. This situation may raise a new question.

7.2. Decision Criterion DC3 for player 2

Suppose that player 2 starts believing that his decision-prediction criterion DC2 is inadequate. Among many possible candidates, he adopts DC3 for his decision-prediction criterion. In the following discussion, we consider the decision-prediction

[24]See Kaneko [12] for such modifications in the case of the decision criterion of a Nash strategy.

criterion DC3 from the viewpoint of player 2. Though we consider also 1's prediction-decision making, it occurs in the mind of player 2.

In DC3, we assume that prediction-decision making is reciprocal between 1 and 2, and thus, all of $D_{11}(s_1), D_{12}(s_1), D_{21}(s_1), D_{22}(s_2)$ are relevant. We can keep the requirements for player i corresponding to $\mathrm{DC21}_2$ – $\mathrm{DC23}_2$, which are denoted by $\mathrm{DC31}_i$ – $\mathrm{DC33}_i$, but we modify the fourth one into

$$\mathrm{DC34}_1\text{: } \bigwedge_{s_2} (D_{12}(s_2) \supset \mathrm{B}_1(D_{22}(s_2))) \bigwedge \bigwedge_{s_1} (D_{11}(s_1) \supset \mathrm{B}_1(D_{21}(s_1)))\,;$$

$$\mathrm{DC34}_2\text{: } \bigwedge_{s_1} (D_{21}(s_1) \supset \mathrm{B}_2(D_{11}(s_1))) \bigwedge \bigwedge_{s_2} (D_{22}(s_2) \supset \mathrm{B}_2(D_{12}(s_2)))\,.$$

The former of each means that player i's prediction is based on his belief on the decision criterion for player j, and the latter that player i believes that his decision is predicted by player j. The reciprocity of DC3 stated in Section 2 is involved here.

Recall that we focus on the decision-prediction criterion for player 2. For his own decision-prediction criterion, we assume $\mathrm{DC31}_2, \ldots, \mathrm{DC34}_2$. Since player 2 makes also his prediction on 1's prediction-decision, we assume $\mathrm{B}_2(\mathrm{DC31}_1), \ldots, \mathrm{B}_2(\mathrm{DC34}_1)$. Our problem is to find formulae $\{D_{ij}(s_j) : s_j \in S_j \text{ and } i,j = 1,2\}$ satisfying the requirements $\mathrm{DC31}_2, \ldots, \mathrm{DC34}_2$ and $\mathrm{B}_2(\mathrm{DC31}_1), \ldots, \mathrm{B}_2(\mathrm{DC34}_1)$. To be more specific, the problem is to find formulae which make $\mathrm{DC31}_2 \bigwedge \cdots \bigwedge \mathrm{DC34}_2$ and $\mathrm{B}_2(\mathrm{DC31}_1 \bigwedge \ldots \bigwedge \mathrm{DC34}_1)$ provable under some belief assumptions. In fact, we would meet some difficulty. The following theorem states that only trivial formulae would be candidates for $\mathrm{DC34}_2 \bigwedge \mathrm{B}_2(\mathrm{DC34}_1)$, which is proved in Section 7.3.

Theorem 7.3 (Reciprocal Failure). Suppose $\vdash_{\mathcal{S}} \mathrm{DC34}_2 \bigwedge \mathrm{B}_2(\mathrm{DC34}_1)$. Then for each $(s_1, s_2) \in S$, $\vdash_{\mathcal{S}} \neg(D_{21}(s_1) \bigwedge D_{22}(s_2))$ or $\vdash_{\mathcal{S}} D_{21}(s_1) \bigwedge D_{22}(s_2)$.

Thus, as soon as we assume $\mathrm{DC34}_2 \bigwedge \mathrm{B}_2(\mathrm{DC34}_1)$, $D_{21}(s_1) \bigwedge D_{22}(s_2)$ becomes a trivial formula. Hence, the above axiomatic system is a failure as a description of prediction-decision making. This failure is not caused for a game theoretic reason but for a reason in our logical system. It is the lack of the capability of $\mathcal{S}$ capturing the reciprocity described by $\mathrm{DC34}_2 \bigwedge \mathrm{B}_2(\mathrm{DC34}_1)$. In fact, we could avoid this difficulty if we extend epistemic logic $\mathcal{S}$ to allow the description of common knowledge, which is the subject of Section 8.

Notice that Theorem 7.3 is a meta-theorem evaluating the axiomatic system $\mathrm{DC34}_2 \bigwedge \mathrm{B}_2(\mathrm{DC34}_1)$ in logic $\mathcal{S} = \mathrm{KD4}^n$, while Theorem 7.1 is an object theorem in $\mathcal{S}$.

Before going to the next subsection, we mention one lemma (cf., Chellas [5], p.99). The *if* part is straightforward by Nec. The *only-if* part is essential. This lemma can be proved showing its contrapositive in a semantical manner based on Theorem 5.1. This *only-if* part is straightforward in the Gentzen-style formulation of $\mathrm{KD4}^n$ together with Theorem 4.6 (Cut-Elimination).

Lemma 7.4. For any A, $\vdash_{\mathcal{S}} \mathrm{B}_i(A)$ if and only if $\vdash_{\mathcal{S}} A$.

It is an consequence that $\vdash_{\mathcal{S}} \mathrm{B}_2(\mathrm{DC34}_1)$ is equivalent to $\vdash_{\mathcal{S}} \mathrm{DC34}_1$. After all, the assumption of Theorem 7.3 is $\vdash_{\mathcal{S}} \mathrm{DC34}_2 \bigwedge \mathrm{DC34}_1$.

7.3. Epistemic Depths and the Depth Lemma

To prove Theorem 7.3, we consider the epistemic depth of a formula. First, let $N^{<\omega>} := \{(i_m, ..., i_1) : i_m, ..., i_1 \in N \text{ and } i_{k+1} \neq i_k \text{ for } k = 1, ..., m-1\}$, where we stipulate that $N^{<\omega>}$ includes the null sequence ϵ, i.e., the sequence of length 0. For $e = (i_m, ..., i_1) \in N^{<\omega>}$, $\mathrm{B}_{i_m}...\mathrm{B}_{i_1}(A)$ is denoted by $\mathrm{B}_e(A)$, and $\mathrm{B}_\epsilon(A)$ is stipulated to be A. We define the following concatenation: for $i \in N$ and $e = (i_m, ..., i_1) \in N^{<\omega>}$, let $i * e = (i, i_m, ..., i_1)$ if $i \neq i_m$ and $i * e = (i_m, ..., i_1)$ if $i = i_m$. Also, we let $i * \epsilon = (i)$.

Let $A \in \mathcal{P}$. We define the (*epistemic*) *depth* $\delta(A)$ of A by induction on the length of a formula:

D0: $\delta(p) = \{\epsilon\}$ for any $p \in PV$;

D1: $\delta(\neg C) = \delta(C)$;

D2: $\delta(C \supset D) = \delta(C) \cup \delta(D)$;

D3: $\delta(\bigwedge \Phi) = \delta(\bigvee \Phi) = \bigcup_{C \in \Phi} \delta(C)$;

D4: $\delta(\mathrm{B}_i(C)) = \{i * e : e \in \delta(A)\}$.

For example, $\delta(\mathrm{p}_0 \supset \mathrm{B}_2\mathrm{B}_3(\mathrm{p}_1)) = \delta(\mathrm{p}_0) \cup \delta(\mathrm{B}_2\mathrm{B}_3(\mathrm{p}_1)) = \{\epsilon, (2,3)\}$. We define $\delta(\Gamma) = \bigcup_{C \in \Gamma} \delta(C)$. The following theorem in a weaker form was is given by Kaneko-Nagashima [16], and the present form is due to Kaneko-Suzuki [17].

Theorem 7.5 (Depth Lemma for KD4^n). Let Γ be a subset of $\mathcal{P}$ and $\mathrm{B}_e(A) = \mathrm{B}_{i_m}...\mathrm{B}_{i_1}(A)$ a formula in $\mathcal{P}$. Suppose $e \in N^{<\omega>}$ and $e \notin \delta(\Gamma)$. Then $\Gamma \vdash_{\mathcal{S}} \mathrm{B}_e(A)$ if and only if Γ is inconsistent in $\mathcal{S}$ or $\vdash_{\mathcal{S}} A$.

When $\Gamma = \{C\}$, the assertion is written as: $\vdash_{\mathcal{S}} C \supset \mathrm{B}_e(A)$ if and only if $\vdash_{\mathcal{S}} \neg C$ or $\vdash_{\mathcal{S}} A$.

The reader may find that Theorem 7.5 (hence, Theorem 7.3, too) is translated into S4^n with Theorem 6.1.

Now, let us return to the proof Theorem 7.3.

Lemma 7.6. Suppose the assumption of Theorem 7.3. Then, for any odd m and $(i_m, ..., i_1) \in N^{<\omega>}$ with $i_m = 2$, $\vdash_{\mathcal{S}} D_{21}(s_1) \bigwedge D_{22}(s_2) \supset \mathrm{B}_{i_m}...\mathrm{B}_{i_1}(D_{21}(s_1) \bigwedge D_{22}(s_2))$.

Proof. The claim for $m = 1$ follows from DC34$_2$. Suppose the claim for odd m. Then $\vdash_{\mathcal{S}} \mathrm{B}_1(D_{21}(s_1) \bigwedge D_{22}(s_2)) \supset \mathrm{B}_1\mathrm{B}_{i_m}... \mathrm{B}_{i_1}(D_{21}(s_1) \bigwedge D_{22}(s_2))$. Hence, $\vdash_{\mathcal{S}} \mathrm{B}_2\mathrm{B}_1(D_{21}(s_1) \bigwedge D_{22}(s_2)) \supset \mathrm{B}_2\mathrm{B}_1\mathrm{B}_{i_m}...\mathrm{B}_{i_1}(D_{21}(s_1) \bigwedge D_{22}(s_2))$. It remains to show $\vdash_{\mathcal{S}} D_{21}(s_1) \bigwedge D_{22}(s_2) \supset \mathrm{B}_2\mathrm{B}_1(D_{21}(s_1) \bigwedge D_{22}(s_2))$. Since $\vdash_{\mathcal{S}} D_{11}(s_1) \bigwedge D_{12}(s_2)) \supset \mathrm{B}_1(D_{21}(s_1) \bigwedge D_{22}(s_2))$ by $\vdash_{\mathcal{S}}$ DC34$_1$ by Lemma 7.4, we have $\vdash_{\mathcal{S}} \mathrm{B}_2(D_{11}(s_1) \bigwedge D_{12}(s_2)) \supset \mathrm{B}_2\mathrm{B}_1(D_{21}(s_1) \bigwedge D_{22}(s_2))$. Since $\vdash_{\mathcal{S}} D_{21}(s_1) \bigwedge D_{22}(s_2) \supset \mathrm{B}_2(D_{11}(s_1) \bigwedge D_{12}(s_2))$ by DC34$_2$, we have $\vdash_{\mathcal{S}} D_{21}(s_1) \bigwedge D_{22}(s_2) \supset \mathrm{B}_2\mathrm{B}_1(D_{21}(s_1) \bigwedge D_{22}(s_2))$. ■

Proof of Theorem 7.3. Take an odd m large enough so that $e = (i_m, ..., i_1) \notin \delta(D_{21}(s_1) \bigwedge D_{22}(s_2))$ with $i_m = 2$. Applying Theorem 7.5 to the assertion of Lemma 7.6, we have $\vdash_{\mathcal{S}} \neg(D_{21}(s_1) \bigwedge D_{22}(s_2))$ or $\vdash_{\mathcal{S}} D_{21}(s_1) \bigwedge D_{22}(s_2)$. ■

8. Common Knowledge Logic $\mathcal{CS}$

The difficulty we met in DC3 is caused by the limited capability of epistemic logic $\mathcal{S}$ to express common knowledge. In $\mathcal{S}$, the common knowledge of formula A is expressed by as the set $\Gamma(A) = \{\mathrm{B}_e(A) : e \in N^{<\omega>}\}$, but when this is used as nonlogical axioms, e.g., $\Gamma(A) \vdash_S B$, only a finite subset of $\Gamma(A)$ is used in a proof of B from $\Gamma(A)$. The entirety of $\Gamma(A)$ is never captured in $\mathcal{S}$, which is the reason for the difficulty. In this section, we consider an extension $\mathcal{S}$ to capture the entirety of common knowledge. We find two approaches in literature: the *fixed-point extension* and *infinitary extension*. In this section, we follow the former due to Halpern-Moses [7]. We also discuss decision making criterion DC3 in the extension.[25]

8.1. Common Knowledge Logic

We add one more unary operator symbol, C, to the list of primitive symbols given in Section 3.1. In F2 of the definition of formulae, we allow $\mathrm{C}(A)$ to be a formula. We denote this extended set of formulae by $\mathcal{P}_{\mathrm{C}}$.

Let $\mathcal{S}$ be an epistemic logic in Diagram 4.2. We define the extension $\mathcal{CS}$ as follows:

$\mathcal{CS} : \mathcal{S}$ + CA + CI within the extended set of formulae $\mathcal{P}_{\mathrm{C}}$,

where CA and CI are the following axiom schema and inference rule:

CA: $\mathrm{C}(A) \supset A \bigwedge \mathrm{B}_1\mathrm{C}(A) \bigwedge ... \bigwedge \mathrm{B}_n\mathrm{C}(A)$;

CI: $\dfrac{D \supset A \bigwedge \mathrm{B}_1(D) \bigwedge ... \bigwedge \mathrm{B}_n(D)}{D \supset \mathrm{C}(A)}$.

The provability of $\mathcal{CS}$ is denoted by $\vdash_{\mathcal{CS}}$. Lemmas 4.2 and 4.3 hold also for $\mathcal{CS}$.

Axiom CA is the *fixed-point property* that if A is common knowledge, then A holds and each player believes the common knowledge of A. Using CA, Nec and MP a finitely many times, we have

$$\vdash_{\mathcal{CS}} \mathrm{C}(A) \supset \mathrm{B}_e(A) \text{ for all } e = (i_m, ..., i_1) \in N^{<\omega>}. \tag{8.1}$$

Thus, $\mathrm{C}(A)$ contains, at least, the common knowledge of A. On the other hand, CI states that if formula D has the fixed-point property of the same form as CA, then D contains the common knowledge of A, in other words, $\mathrm{C}(A)$ is the deductively weakest formula having the fixed-point property.

The following fact may help understand the term "fixed-point": By CA and CI, we have, reading both D and A as $\mathrm{C}(A)$ in CI,

$$\vdash_{\mathcal{CS}} \mathrm{C}(A) \supset \mathrm{CC}(A), \tag{8.2}$$

and the converse is provable, too.

To see that $\mathrm{C}(A)$ captures really the entirety of $\{\mathrm{B}_e(A) : e \in N^{<\omega>}\}$ in $\mathcal{CS}$ with no superfluous properties, we will prove the following lemma using semantics:

Lemma 8.1.(1): If $\vdash_{\mathcal{CS}} D \supset \mathrm{B}_e(A)$ for all $e \in N^{<\omega>}$, then $\vdash_{\mathcal{CS}} D \supset \mathrm{C}(A)$.

(2): If $\vdash_{\mathcal{CS}} D \supset \mathrm{B}_i\mathrm{B}_e(A)$ for all $e \in N^{<\omega>}$, then $\vdash_{\mathcal{CS}} D \supset \mathrm{B}_i\mathrm{C}(A)$.

[25]For the fixed-point approach, see also Fagin-Halpern-Moses-Vardi [6], Lismont-Mongin [20],[21] and Meyer-van der Hoek [23]. For the infinitary approach, see Kaneko-Nagashima [15] (including the predicate case) and also Heifetz [8]. Kaneko [13] proved in the propositional case that these approaches can be regarded as equivalent as far as the definition of common knowledge is concerned. However, Wolter [32] proved that this equivalnece does not hold in the predicate case.

We would like to regard C(A) effectively as equivalent to the conjunction of $\Gamma(A) = \{B_e(A) : e \in N^{<\omega>}\}$, though the infinitary conjunction is not allowed in $\mathcal{CS}$. This effective equivalence can be seen by comparing $\mathcal{CS}$ with classical logic CL. In CL, $\bigwedge \Phi$ is fixed by Axiom L4 and $\bigwedge$-Rule. Fact (8.1) corresponds to L4, and Lemma 8.1.(1) does to $\bigwedge$-Rule. Thus, $\mathcal{CS}$ succeeds in having the parallel structure to that of CL. However, we need to extend the Kripke semantics and to prove the soundness-completeness theorem, in order to prove Lemma 8.1. The present author has no direct syntactical proof of Lemma 8.1.

To guarantee that the above axiomatization captures the concept of common knowledge, we need the semantical counterpart of $\mathcal{CS}$. The semantics also facilitates the use of $\mathcal{CS}$ and makes $\mathcal{CS}$ directly comparable with $\mathcal{S}$. Indeed, we can use the same Kripke semantics for $\mathcal{CS}$, and need only to extend the valuation relation $(\mathcal{K}, \sigma, w) \models$ to $\mathcal{P}_C$ from $\mathcal{P}$.

Let $(\mathcal{K}, \sigma) = ((W : R_1, ..., R_n), \sigma)$ be a Kripke model. We extend the valuation relation $(\mathcal{K}, \sigma, w) \models$ from $\mathcal{P}$ to $\mathcal{P}_C$ by K0–K5 and

K6: $(\mathcal{K}, \sigma, w) \models \mathrm{C}(A)$ if and only if $(\mathcal{F}, \sigma, u) \models A$ for all u reachable from w,

where u is *reachable from* w iff there is a sequence $w_0 = w, w_1, ..., w_m = u$ such that for each $k = 0, 1, ..., m-1$, $w_k R_j w_{k+1}$for some $j \in N$. This K6 is equivalent to

K6*: $(\mathcal{K}, \sigma, w) \models \mathrm{C}(A)$ if and only if $(\mathcal{F}, \sigma, w) \models \mathrm{B}_e(A)$ for all $e \in N^{<\omega>}$.

This equivalence can be proved by induction on the length of e.

We state the soundness-completeness of $\mathcal{CS}$, which is proved in Section 9.

Theorem 8.2 (Soundness-Completeness of $\mathcal{CS}$). Let $\mathcal{S}$ be an epistemic logic in Diagram 4.2, and $\mathcal{S}^*$ the set of the Kripke frames satisfying the corresponding conditions on the accessibility relations. Let A be a formula in $\mathcal{P}_C$.

(1): $\vdash_{\mathcal{CS}} A$ if and only if $(\mathcal{K}, \sigma, w) \models A$ for all Kripke frames $\mathcal{K}$ in $\mathcal{S}^*$, all assignments σ and all $w \in W$.

(2): There is a Kripke frame $\mathcal{K}$ in $\mathcal{S}^*$, an assignment σ and a world $w \in W$ satisfying $(\mathcal{K}, \sigma, w) \models A$ if and only if A is consistent in $\mathcal{CS}$.

Claims (1) and (2) are equivalent as for the Soundness-Completeness Theorem (Theorem 5.1.) The *only-if* part of (1) is proved by modifying the corresponding proof for Theorem 5.1 adding the following steps to Lemma 5.2.

Lemma 8.3.(1): $(\mathcal{K}, \sigma, w) \models \mathrm{C}(A) \supset A \bigwedge \mathrm{B}_1\mathrm{C}(A) \bigwedge ... \bigwedge \mathrm{B}_n\mathrm{C}(A)$ for any $\mathcal{K} = (W : R_1, ..., R_n)$ in $\mathcal{S}^*$, assignment σ in $\mathcal{K}$ and $w \in W$.

(2): Let $\mathcal{K}$ be any frame in $\mathcal{S}^*$ and σ any assignment in $\mathcal{K}$. If $(\mathcal{K}, \sigma, w) \models D \supset A \bigwedge \mathrm{B}_1(D) \bigwedge ... \bigwedge \mathrm{B}_n(D)$ for any $w \in W$, then $(\mathcal{K}, \sigma, w) \models D \supset \mathrm{C}(A)$ for any $w \in W$.

Proof. We prove only (2). Let u be any world. Suppose $(\mathcal{K}, \sigma, u) \models D$. Then $(\mathcal{K}, \sigma, u) \models A$ and $(\mathcal{K}, \sigma, u) \models \mathrm{B}_i(D)$ for $i = 1, ..., n$. Let u_m be any world so that it is reachable by m steps ($m \geq 0$). We assume the inductive hypothesis that $(\mathcal{K}, \sigma, u_m) \models A$ and $(\mathcal{K}, \sigma, u_m) \models \mathrm{B}_i(D)$ for $i = 1, ..., n$. Now let $u_m R_i u_{m+1}$. Then $(\mathcal{K}, \sigma, u_{m+1}) \models D$. Since $(\mathcal{K}, \sigma, w) \models D \supset A \bigwedge \mathrm{B}_1(D) \bigwedge ... \bigwedge \mathrm{B}_n(D)$ for any $w \in W$, we have $(\mathcal{K}, \sigma, u_{m+1}) \models A$ and $(\mathcal{K}, \sigma, u_{m+1}) \models \mathrm{B}_i(D)$ for $i = 1, ..., n$. Thus, we have proved $(\mathcal{K}, \sigma, v) \models A$ for any v reachable from u. By K6, $(\mathcal{K}, \sigma, u) \models \mathrm{C}(A)$. ■

One possible use of the above completeness is to prove Lemma 8.1.

Proof of Lemma 8.1. We prove only (1). Suppose $\vdash_{\mathcal{CS}} D \supset \mathrm{B}_i\mathrm{B}_e(A)$ for all $e \in N^{<\omega>}$. Let $(\mathcal{K}, \sigma)$ be any Kripke model. Then $(\mathcal{K}, \sigma, w) \models D \supset \mathrm{B}_i\mathrm{B}_e(A)$ for any world w. Then let u be any world. Suppose $(\mathcal{K}, \sigma, u) \models D$. Then $(\mathcal{K}, \sigma, u) \models \mathrm{B}_i\mathrm{B}_e(A)$. Let u' be any world with uR_iu'. Then $(\mathcal{K}, \sigma, u') \models \mathrm{B}_e(A)$. Since this holds for any $e \in N^{<\omega>}$, we have $(\mathcal{K}, \sigma, u') \models \mathrm{C}(A)$ by K6*. Since this holds for any u' with uR_iu', we have $(\mathcal{K}, \sigma, u) \models \mathrm{B}_i\mathrm{C}(A)$. We proved $(\mathcal{K}, \sigma, u) \models D \supset \mathrm{B}_i\mathrm{C}(A)$. By Theorem 8.2, we have $\vdash_{\mathcal{CS}} D \supset \mathrm{B}_i\mathrm{C}_e(A)$. ∎

The following is an immediate but important consequence from Theorem 8.2.

Theorem 8.4 (Conservativity of $\mathcal{CS}$ upon $\mathcal{S}$). Let A be a formula in $\mathcal{P}$. Then $\vdash_{\mathcal{S}} A$ if and only if $\vdash_{\mathcal{CS}} A$.

This theorem guarantees that for $\mathcal{S} = \mathrm{KD4}^n$, the Reciprocal Failure Theorem and Depth Lemma (Theorems 7.1 and 7.3) can be converted into common knowledge logic $\mathcal{CS}$ with the restrictions of the target formulae to $\mathcal{P}$.

Remark 8.5. In $\mathcal{CS}$, the common knowledge operator C enjoys the properties K, T, 4 and Nec. Hence, $\mathcal{CS}$ may be regarded as $\mathrm{S4}^1$ with respect to C. Therefore, the assertions of Lemma 4.2 hold for C in $\mathcal{CS}$.[26]

8.2. Solution Theory for Decision Criterion DC3

In Section 7.2, we formulated Axioms $\mathrm{DC31}_2$,...,$\mathrm{DC34}_2$,$\mathrm{B}_2(\mathrm{DC31}_1)$,...,$\mathrm{B}_2(\mathrm{DC34}_1)$ for decision criterion DC3 for player 2, and showed that Axioms $\mathrm{DC34}_2$ and $\mathrm{B}_2(\mathrm{DC34}_1)$ lead to the reciprocal failure of allowing only trivial formulae. By the conservativity of the extension $\mathcal{CS}$ upon $\mathcal{S}$, the Reciprocal Failure Theorem (Theorem 7.3) still holds for $\mathcal{P}$ in $\mathcal{CS}$, where $\mathcal{S} = \mathrm{KD4}^n$. Now, however, we can look for candidates in $\mathcal{P}_{\mathrm{C}}$ rather than in $\mathcal{P}$. Here, we show that DC3 can be a meaningful criterion in the extension $\mathcal{CS}$ of $\mathcal{S}$.

To define the candidate formulae, we first modify $\mathrm{Nash}(s_1, s_2)$ into

$$\mathrm{B}_1(\mathrm{Best}_1(s_1 \mid s_2)) \bigwedge \mathrm{B}_2(\mathrm{Best}_2(s_1 \mid s_2)), \tag{8.3}$$

which we denote by $\mathrm{Nash}^*(s_1, s_2)$. This differs from $\mathrm{Nash}(s_1, s_2)$ in that each individual payoff function is taken up to his belief in $\mathrm{Nash}^*(s_1, s_2)$. Then we define the candidate formulae: for $(s_1, s_2) \in S$,

$$\tilde{D}_{11}(s_1) = \bigvee_{s_2} \mathrm{B}_1\mathrm{C}(\mathrm{Nash}^*(s_1, s_2)) \text{ and } \tilde{D}_{12}(s_2) = \bigvee_{s_1} \mathrm{B}_1\mathrm{C}(\mathrm{Nash}^*(s_1, s_2)). \tag{8.4}$$

$$\tilde{D}_{21}(s_1) = \bigvee_{s_2} \mathrm{B}_2\mathrm{C}(\mathrm{Nash}^*(s_1, s_2)) \text{ and } \tilde{D}_{22}(s_2) = \bigvee_{s_1} \mathrm{B}_2\mathrm{C}(\mathrm{Nash}^*(s_1, s_2)). \tag{8.5}$$

Now, we have to see that these formulae satisfy Axioms $\mathrm{DC34}_2$ and $\mathrm{B}_2(\mathrm{DC34}_1)$. We denote by $\mathrm{DC34}_1(\tilde{D})$ and $\mathrm{DC34}_2(\tilde{D})$, the formulae obtained by plugging $\tilde{D}_{ij}(s_j)$ into $D_{ij}(s_j)$ in $\mathrm{DC34}_1$ and $\mathrm{DC34}_2$.

[26] In $\mathcal{CS}$, we are treating common knowledge in the sense of $\vdash_{\mathcal{CS}} \mathrm{C}(A) \supset A$. The common beliefs of A is expressed as $\bigwedge_i \mathrm{B}_i\mathrm{C}(A)$ in $\mathcal{CS}$. However, it will be seen in the following that the individual belief of common knowledge plays an important role in our game theoretical application rather than common beliefs.

Lemma 8.6. $\vdash_{CS} \mathrm{DC34}_2(\tilde{D}) \bigwedge \mathrm{B}_2(\mathrm{DC34}_1(\tilde{D}))$.

Proof. We prove $\vdash_{CS} \mathrm{DC34}_1(\tilde{D})$, which implies $\vdash_{CS} \mathrm{B}_2(\mathrm{DC34}_1(\tilde{D}))$. Recall $\mathrm{DC34}_1(\tilde{D})$ is $\bigwedge_{s_1} \left(\tilde{D}_{11}(s_1) \supset \mathrm{B}_1(\tilde{D}_{21}(s_1))\right) \bigwedge \bigwedge_{s_2} \left(\tilde{D}_{12}(s_2) \supset \mathrm{B}_1(\tilde{D}_{22}(s_2))\right)$. We prove only the first half. Since $\vdash_{CS} \mathrm{C}(\mathrm{Nash}^*(s_1,s_2)) \supset \mathrm{B}_1\mathrm{C}(\mathrm{Nash}^*(s_1,s_2))$ by CA, we have $\vdash_{CS} \mathrm{C}(\mathrm{Nash}^*(s_1,s_2)) \supset \bigvee_{s_2}\mathrm{B}_1\mathrm{C}(\mathrm{Nash}^*(s_1,s_2))$ by L5, i.e., $\vdash_{CS} \mathrm{C}(\mathrm{Nash}^*(s_1,s_2)) \supset \tilde{D}_{11}(s_1)$. Then $\vdash_{CS} \mathrm{B}_2\mathrm{C}(\mathrm{Nash}^*(s_1,s_2)) \supset \mathrm{B}_2(\tilde{D}_{11}(s_1))$. Since this holds for any s_2, we have $\vdash_{CS} \bigvee_{s_2}\mathrm{B}_2\mathrm{C}(\mathrm{Nash}^*(s_1,s_2)) \supset \mathrm{B}_2(\tilde{D}_{11}(s_1))$ by $\bigvee$-Rule, that is, $\vdash_{CS} \tilde{D}_{21}(s_1) \supset \mathrm{B}_2(\tilde{D}_{11}(s_1))$. ■

In fact, requirements $\bigwedge\{\mathrm{DC31}_2,...,\mathrm{DC34}_2\}$ and $\mathrm{B}_2(\bigwedge\{\mathrm{DC31}_1,...,\mathrm{DC34}_1\})$ hold for the above candidates $\tilde{D}_{ij}(s_j) : s_j \in S_j, i,j = 1,2$. Only $\mathrm{DC31}_2$ and $\mathrm{B}_2(\mathrm{DC31}_1)$ need a game theoretical assumption, which is proved in the end of this subsection.

Theorem 8.7. Let $g = (g_1, g_2)$ be a game with a unique Nash equilibrium. Then

(a): $\mathrm{B}_2\mathrm{C}(\hat{g}) \vdash_{CS} \mathrm{DC31}_2(\tilde{D}) \bigwedge \mathrm{B}_2(\mathrm{DC31}_1(\tilde{D}))$;

(b): $\vdash_{CS} \bigwedge\{\mathrm{DC32}_2(\tilde{D}),...,\mathrm{DC34}_2(\tilde{D})\} \bigwedge \mathrm{B}_2(\bigwedge\{\mathrm{DC32}_1(\tilde{D}),...,\mathrm{DC34}_1(\tilde{D})\})$.

The next theorem states that formulae $\tilde{D}_{ij}(s_j), i,j = 1,2$ are the deductively weakest formulae satisfying our requirements. The proof is given in the end of this subsection.

Theorem 8.8 (Personalized Characterization of DC3). Let $g = (g_1, g_2)$ be a 2-person game with a unique Nash equilibrium. Then $D_{ij}(s_j), i,j = 1,2$ satisfy our requirements in the sense of (a) and (b) of Theorem 8.7. Then $\mathrm{B}_2\mathrm{C}(\hat{g}) \vdash_{CS} \bigwedge_{s_1} \left(D_{21}(s_1) \supset \tilde{D}_{21}(s_1)\right) \bigwedge \bigwedge_{s_2} \left(D_{22}(s_2) \supset \tilde{D}_{22}(s_2)\right)$.

These theorems correspond to Theorems 7.1 and 7.2 for DC2. As in the case of these theorems, Theorems 87 and 8.8 imply that the deductively weakest formulae $\tilde{D}_{21}(s_1), \tilde{D}_{22}(s_2), (s_1,s_2) \in S_1 \times S_2$ are uniquely determined.

Nevertheless, there is some unparalleled structure between the theorems for DC2 and for DC3. For example, $\mathrm{DC23}_2(\hat{D})$ requires a game theoretical assumption in Theorem 7.1, while $\mathrm{DC31}_2(\tilde{D}) \bigwedge \mathrm{B}_2(\mathrm{DC31}_1(\tilde{D}))$ does in Theorem 8.7. This non-parallelism is caused by the difference in the formulations of $\hat{D}_{22}(s_2)$ and $\tilde{D}_{ij}(s_j), i,j = 1,2$. If we formulate $\hat{D}_{22}(s_2)$ as

$$\bigvee_{s_1} \mathrm{B}_2(\mathrm{B}_1(\mathrm{Dom}_1(s_1)) \bigwedge \mathrm{Best}_2(s_2 \mid s_1)), \tag{8.6}$$

then this non-parallelism disappears. However, this formulation has some other weak point. The present author does not know what formulae for DC3 correspond to the formula $\hat{D}_{22}(s_2)$ of Section 7.2.

The above characterization is made from the viewpoint of player 2. This personalized characterization makes sense in $\mathcal{CS}$ with $\mathcal{S} = \mathrm{KD4}^n$. If we adopt $\mathcal{S} = \mathrm{S4}^n$, then $\mathrm{B}_i\mathrm{C}(\mathrm{Nash}^*(s_1,s_2))$ is equivalent to $\mathrm{C}(\mathrm{Nash}(s_1,s_2))$ in $\mathcal{CS}$. Furthermore,

$$\vdash_{CS} \tilde{D}_{ij}(s_j) \equiv \bigvee_{s_t} \mathrm{C}(\mathrm{Nash}(s_1,s_2)), \text{ where } t \neq j. \tag{8.7}$$

This is what is described informally by game theorists.[27] In this case, subjectivity disappears, *a fortiori*, false beliefs cannot be discussed. The false beliefs on common knowledge will be discussed in a game theoretical example in Section 8.3.

The assumption on the game having a unique Nash equilibrium can be relaxed to a game with the interchangeable set of Nash equilibria. If interchangeability is violated, we need different assumptions on the individual beliefs on the common knowledge. For example, some strategy pair is assumed to satisfy the criterion, which is assumed to be the common knowledge. This is considered in Kaneko [12] partially. These problems are not yet fully investigated.

We have succeeded in avoiding the reciprocal failure by incorporating common knowledge into epistemic logic $\mathcal{S} = \mathrm{KD4}^n$. Nevertheless, we should recognize that epistemic logic $\mathcal{CS}$ involves two levels of infinities: first, $\mathcal{S}$ allows formulae of any epistemic (finite) depths, e.g., $\mathrm{B}_e(A) = \mathrm{B}_{i_m}...\mathrm{B}_{i_1}(A)$ for any m, and second, $\mathcal{CS}$ allows $\mathrm{C}(A)$ to capture the entirety of $\{\mathrm{B}_e(A) : e \in N^{<\omega>}\}$. After all, common knowledge is an infinitary concept of an idealization, though CA and CI avoid infinitary treatments in an ingenious way.

For the analysis of human decision making, it would be more natural to avoid such an infinitary concept. Nevertheless, this depends upon a situation. If the rules of the game including payoff functions are visible for the players and if they are standing face-to-face, then these constituents may be regarded as common knowledge between these players. This is caused by the special property of vision: Since the speed of light can be regarded as infinity, the mutual verification can be made almost instantaneously (this is reminiscent of the *Analogy of the Sun* of Plato [29], Book VI). On the other hand, payoff functions for games, except well defined parlor games, are usually not visible but belong to individual subjectivity. Therefore, visions do not help to obtain the common knowledge of payoff functions. In sum, the common knowledge assumption may be adequate in a very specific situation, but not in other situations.

Proof of Theorem 8.7. We prove $\mathrm{B}_2\mathrm{C}(\hat{g}) \vdash_{\mathcal{CS}} \mathrm{DC31}_2(\tilde{D})$. Since $\mathrm{B}_1(\hat{g}_1),\mathrm{B}_2(\hat{g}_2) \vdash_{\mathcal{CS}} \mathrm{Nash}^*(s_1,s_2) \bigwedge \mathrm{Nash}^*(t_1,t_2) \supset \mathrm{Nash}^*(t_1,s_2)$ using the uniqueness of a Nash equilibrium, we have $\mathrm{CB}_1(\hat{g}_1), \mathrm{CB}_2(\hat{g}_2) \vdash_{\mathcal{CS}} \mathrm{C}(\mathrm{Nash}^*(s_1,s_2)) \bigwedge \mathrm{C}(\mathrm{Nash}^*(t_1,t_2)) \supset \mathrm{C}(\mathrm{Nash}^*(t_1,s_2))$ by Remark 8.5. Since $\mathrm{C}(\hat{g}) \vdash_{\mathcal{CS}} \mathrm{CB}_i(A)$ for all $A \in \hat{g}_i$ and $i = 1,2$ by K6* and Theorem 8.2, we have $\mathrm{C}(\hat{g}) \vdash_{\mathcal{CS}} \mathrm{C}(\mathrm{Nash}^*(s_1,s_2)) \bigwedge\mathrm{C}(\mathrm{Nash}^*(t_1,t_2)) \supset \mathrm{C}(\mathrm{Nash}^*(t_1,s_2))$. Hence $\mathrm{C}(\hat{g}) \vdash_{\mathcal{CS}} \mathrm{C}(\mathrm{Nash}^*(s_1,s_2)) \bigwedge\mathrm{C}(\mathrm{Nash}^*(t_1,t_2)) \supset \mathrm{B}_2(\mathrm{Best}_2(s_2 \mid t_1))$. Since $\mathrm{C}(\hat{g}) \vdash_{\mathcal{CS}} \mathrm{C}(\mathrm{Nash}^*(r_1,r_2))$ if and only if $\hat{g} \vdash_{\mathcal{CS}} \mathrm{Nash}(r_1,r_2)$ for any (r_1,r_2), we have $\mathrm{C}(\hat{g}) \vdash_{\mathcal{CS}} \mathrm{C}(\mathrm{Nash}^*(s_1,s_2)) \bigwedge\mathrm{C}(\mathrm{Nash}^*(t_1,t_2)) \supset \mathrm{Best}_2(s_2 \mid t_1)$. Thus, $\mathrm{B}_2\mathrm{C}(\hat{g}) \vdash_{\mathcal{CS}} \mathrm{B}_2\mathrm{C}(\mathrm{Nash}^*(s_1,s_2) \bigwedge\mathrm{B}_2\mathrm{C}(\mathrm{Nash}^*(t_1,t_2)) \supset \mathrm{B}_2(\mathrm{Best}_2(s_2 \mid t_1))$. We can introduce $\bigvee$ to the premise: indeed, the following is equivalent to the last formula,

$$\vdash_{\mathcal{CS}} \mathrm{B}_2\mathrm{C}(\mathrm{Nash}^*(s_1,s_2)) \supset (\bigwedge \mathrm{B}_2\mathrm{C}(\hat{g}) \bigwedge \mathrm{B}_2\mathrm{C}(\mathrm{Nash}^*(t_1,t_2)) \supset \mathrm{B}_2(\mathrm{Best}_2(s_2 \mid t_1)))\,.$$

Since s_1 is arbitrary, we have, applying $\bigvee$-Rule,

$$\vdash_{\mathcal{CS}} \bigvee_{s_1} \mathrm{B}_2\mathrm{C}(\mathrm{Nash}^*(s_1,s_2) \supset (\bigwedge \mathrm{B}_2\mathrm{C}(\hat{g}) \bigwedge \mathrm{B}_2\mathrm{C}(\mathrm{Nash}^*(t_1,t_2)) \supset \mathrm{B}_2(\mathrm{Best}_2(s_2 \mid t_1)))\,.$$

[27]This set of decision criteria is considered in Kaneko-Nagashima [15] with the mixed strategies, and is considered in Kaneko [12] with pure strategies.

Using a similar argument, we have $B_2C(\hat{g}) \vdash_{CS} \bigvee_{s_1} B_2C(\mathrm{Nash}^*(s_1, s_2)) \bigwedge \bigvee_{t_2} B_2C(\mathrm{Nash}^*(t_1, t_2)) \supset B_2(\mathrm{Best}_2(s_2 \mid t_1))$, i.e., $B_2C(\hat{g}) \vdash_{CS} \tilde{D}_{22}(s_2) \bigwedge \tilde{D}_{21}(t_1) \supset B_2(\mathrm{Best}_2(s_2 \mid t_1))$. Thus, $B_2C(\hat{g}) \vdash_{CS} \mathrm{DC31}_2(\tilde{D})$.

Similarly, $B_1C(\hat{g}) \vdash_{CS} \tilde{D}_{11}(s_1) \bigwedge \tilde{D}_{12}(t_2) \supset B_1(\mathrm{Best}_1(s_1 \mid t_2))$. By CA, we have $C(\hat{g}) \vdash_{CS} \tilde{D}_{11}(s_1) \bigwedge \tilde{D}_{12}(t_2) \supset B_1(\mathrm{Best}_1(s_1 \mid t_2))$. Thus, $C(\hat{g}) \vdash_{CS} \mathrm{DC31}_1(\tilde{D})$. Hence $B_2C(\hat{g}) \vdash_{CS} B_2(\mathrm{DC31}_1(\tilde{D}))$. ■

For the proof of Theorem 8.8, we first prove the following lemma.

Lemma 8.9. $B_2C(\hat{g}) \vdash_{CS} D_{21}(s_1) \bigwedge D_{22}(s_2) \supset B_2C(\mathrm{Nash}^*(s_1, s_2))$ for $i = 1, 2$.

Proof. Denote $D_{i1}(s_1) \bigwedge D_{i2}(s_2)$ by $D_i(s_1, s_2)$ for $i = 1, 2$. First, by $\mathrm{DC31}_2$, $B_2C(\hat{g}) \vdash_{CS} D_2(s_1, s_2) \supset B_2(\mathrm{Best}_2(s_2 \mid s_1))$. Hence $B_2C(\hat{g}) \vdash_{CS} D_2(s_1, s_2) \supset B_2B_2(\mathrm{Best}_2(s_2 \mid s_1))$. Second, by $B_2(\mathrm{DC31}_1)$, $B_2C(\hat{g}) \vdash_{CS} B_2(D_1(s_1, s_2)) \supset B_2B_1(\mathrm{Best}_1(s_1 \mid s_2))$. Since $\vdash_{CS} D_2(s_1, s_2) \supset B_2(D_1(s_1, s_2))$ by $\mathrm{DC34}_2$, we have $B_2C(\hat{g}) \vdash_{CS} D_2(s_1, s_2) \supset B_2B_1(\mathrm{Best}_1(s_1 \mid s_2))$. Thus, $B_2C(\hat{g}) \vdash_{CS} D_2(s_1, s_2) \supset B_2(\mathrm{Nash}^*(s_1, s_2))$.

In the same manner, we have $B_2C(\hat{g}) \vdash_{CS} B_2(D_1(s_1, s_2)) \supset B_2B_1(\mathrm{Nash}^*(s_1, s_2))$. For this, we use $B_2C(\hat{g}) \vdash_{CS} B_2(D_1(s_1, s_2)) \supset B_2B_1(D_2(s_1, s_2))$.

Suppose the induction hypothesis that $B_2C(\hat{g}) \vdash_{CS} D_2(s_1, s_2) \supset B_2B_e(\mathrm{Nash}^*(s_1, s_2))$ for all $e = (i_m, ..., i_1)$ and $B_2C(\hat{g}) \vdash_{CS} B_2(D_1(s_1, s_2)) \supset B_2B_1B_e(\mathrm{Nash}^*(s_1, s_2))$. We prove these for $(i_{m+1}, i_m, ..., i_1)$. Since $B_2C(\hat{g}) \vdash_{CS} D_2(s_1, s_2) \supset B_2(D_1(s_1, s_2))$, we have $B_2C(\hat{g}) \vdash_{CS} D_2(s_1, s_2) \supset B_2B_1B_e(\mathrm{Nash}^*(s_1, s_2))$ as well as $B_2C(\hat{g}) \vdash_{CS} D_2(s_1, s_2) \supset B_2B_2B_e(\mathrm{Nash}^*(s_1, s_2))$. Finally, since $B_2C(\hat{g}) \vdash_{CS} B_2(D_1(s_1, s_2)) \supset B_2B_1(D_2(s_1, s_2))$, and $B_2C(\hat{g}) \vdash_{CS} B_2B_1(D_2(s_1, s_2)) \supset B_2B_1B_2B_e(\mathrm{Nash}^*(s_1, s_2))$, we have $B_2C(\hat{g}) \vdash_{CS} B_2(D_1(s_1, s_2)) \supset B_2B_1B_2B_e(\mathrm{Nash}^*(s_1, s_2))$. ■

Proof of Theorem 8.8. Since Lemma 8.9 is equivalent to $B_2C(\hat{g}) \vdash_{CS} D_{21}(s_1) \supset (D_{22}(s_2) \supset B_2C(\mathrm{Nash}^*(s_1, s_2)))$. Thus $B_2C(\hat{g}) \vdash_{CS} D_{21}(s_1) \supset (D_{22}(s_2) \supset \bigvee_{t_1} B_2C(\mathrm{Nash}^*(t_1, s_2)))$, and then $B_2C(\hat{g}) \vdash_{CS} \bigvee_{s_1} D_{21}(s_1) \supset (D_{22}(s_2) \supset \bigvee_{t_1} B_2C(\mathrm{Nash}^*(t_1, s_2)))$. By $\mathrm{DC33}_2$, $B_2C(\hat{g}) \vdash_{CS} D_{22}(s_2) \supset \bigvee_{t_1} B_2C(\mathrm{Nash}^*(t_1, s_2))$. ■

8.3. Mutual Misunderstanding of Common Understanding in DC3

Consider the assumption set $B_1C(\bigwedge \hat{g}^3) \cup B_2C(\bigwedge \hat{g}^4)$, where g^3 and g^4 are the games of Tables 2.3 and 2.3. That is, player 1 believes that it is common knowledge that game g^3 is played, while 2 believes that it is common knowledge that g^4 is played. We can prove that $B_1C(\bigwedge \hat{g}^3) \cup B_2C(\bigwedge \hat{g}^4)$ is consistent in $\mathcal{CS}$ with $\mathcal{S} = \mathrm{KD4}^n$, and that

$$B_1C(\hat{g}^3), B_2C(\hat{g}^4) \vdash_{CS} \tilde{D}_{11}(\mathrm{s}_{12}) \bigwedge \tilde{D}_{12}(\mathrm{s}_{22}) \bigwedge \tilde{D}_{21}(\mathrm{s}_{12}) \bigwedge \tilde{D}_{22}(\mathrm{s}_{22}).$$

This states that both players' predictions are behaviorally correct. Nevertheless, decisions and predictions are based on the mutual misunderstanding of common understanding. Neither player would find this misunderstanding by seeing the resulting choice of the other player, since the predictions are behaviorally correct. This argument cannot be done in logic $\mathcal{CS}$ with $\mathcal{S} = \mathrm{S4}^n$, since $B_1C(\bigwedge \hat{g}^3) \cup B_2C(\bigwedge \hat{g}^4)$ becomes inconsistent.

The consistency of $B_1C(\bigwedge \hat{g}^3) \cup B_2C(\bigwedge \hat{g}^4)$ in $\mathcal{CS}$ with $\mathcal{S} = \mathrm{KD4}^n$ is verified by constructing the following model:

$$\circlearrowleft_{12}\ w_1 : \hat{g}^3 \qquad\qquad\qquad \circlearrowleft_{12}\ w_2 : \hat{g}^4$$
$$\nwarrow_1 \quad \nearrow_2$$
$$w_0$$

Diagram 8.1

Then, since $(\mathcal{K}, \sigma, w_1) \models \mathrm{C}(\bigwedge \hat{g}^3)$ and $(\mathcal{K}, \sigma, w_2) \models \mathrm{C}(\bigwedge \hat{g}^4)$, we have $(\mathcal{K}, \sigma, w_0) \models \mathrm{B}_1\mathrm{C}(\bigwedge \hat{g}^3) \bigwedge \mathrm{B}_2\mathrm{C}(\bigwedge \hat{g}^4)$.

The above mutual misunderstanding of common understanding may be observed in our life. The point here is the possibility that each player develops a false and different belief of the common knowledge of the situation. We close our discussions with the following example on such phenomena.

Konnyaku Mondô (Jelly Dialogue): A Japanese traditional rakugo (comic story), "Konnyaku Mondô" (Devil's Tongue Jelly Dialogue), describes exactly such a situation: A (devil's tongue) jelly maker lived in a Buddhist temple pretending to be a monk. A real Buddhist monk came to visit this temple to have a dialogue on Buddhism thoughts. The jelly maker first refused but eventually agreed to have a dialogue. Since the jelly maker did not know how he could communicate with the Monk on Buddhism, he answered the Monk's questions in gestures. The Monk took this as a style of dialogue, and responded in gestures. After several exchanges of gestures, both thought that the jelly maker defeated the Monk. After the dialogue, a witness asked the Monk about the dialogue. The Monk said that the jelly maker had a great Buddhism thought shown by his gestures and should be respected. The jelly maker was asked by another witness and answered: the Monk started talking badly about jelly products with his gestures, made the jelly maker angry, and thus the jelly maker beat the Monk. Thus each of them believed that they had perfectly meaningful dialogue and that it was common knowledge that the jelly maker defeated the monk in the dialogue. However, the Monk believed that they had a Buddhism dialogue, while the jelly maker believed that they had discussed about jelly products (pp.61–70 in [31]).[28]

9. Proof of the Completeness of $\mathcal{CS}$

Here we prove the completeness part, assuming $\mathcal{S} = \mathrm{KD4}^n$. In other cases, we need some modifications (see Halpern-Moses [7]). Let A be a formula, which is consistent in $\mathcal{CS}$. If $\mathrm{C}(\cdot)$ does not occur in A, then the following proof becomes also a proof of the completeness of $\mathcal{S}$, i.e., the *if* part of Theorem 5.2.(2), in which case we can ignore the step (5) of the induction proof of (9.1). We suggest that the reader who is not familiar with proofs in logic should read, before this subsection, the proof of the completeness for CL given in the Appendix.

We denote $\{D : D \text{ is a subformula of } A\} \cup \bigcup_i \{\mathrm{B}_i(D), \mathrm{B}_i\mathrm{C}(D) : \mathrm{C}(D) \text{ is a subformula of } A\}$ by $\mathrm{Sub}(A)$. Then we define $\mathrm{Sub}^+(A) = \mathrm{Sub}(A) \cup \{\neg D : D \in$

[28]The story is slightly modified from the original (pp.61–70 in [31]) to have a shorter consistent story, but the essential part is not changed.

Sub(A)}. We say that a subset Γ of $\text{Sub}^+(A)$ is *maximally consistent* iff it is consistent in $\mathcal{CS}$ and $\Gamma \cup \{D\}$ is inconsistent for any $D \in \text{Sub}^+(A) - \Gamma$. We denote the set $\{\Gamma : \Gamma$ is a maximally consistent subset of $\text{Sub}^+(A)$ in $\mathcal{CS}\}$ by $\text{Con}(A)$.

Each maximally consistent set Γ of $\text{Sub}^+(A)$ has the following properties.

Lemma 9.1. Let $\Gamma \in \text{Con}(A)$. Then

(1): for any $\neg B$ in $\text{Sub}^+(A)$, $B \in \Gamma$ or $\neg B \in \Gamma$;

(2): for any B in $\text{Sub}^+(A)$, $\neg B \in \Gamma \Leftrightarrow B \notin \Gamma$;

(3): for any $B \supset C$ in $\text{Sub}^+(A)$, $B \supset C \in \Gamma \Leftrightarrow \neg B \in \Gamma$ or $C \in \Gamma$;

(4): for any $\bigwedge \Phi$ in $\text{Sub}^+(A)$, $\bigwedge \Phi \in \Gamma \Leftrightarrow B \in \Gamma$ for all $B \in \Phi$;

(5): for any $\bigvee \Phi$ in $\text{Sub}^+(A)$, $\bigvee \Phi \in \Gamma \Leftrightarrow B \in \Gamma$ for some $B \in \Phi$.

For a set u of formulae, we write $u^{\text{B}_i} := \{\text{B}_i(C) : \text{B}_i(C) \in u\}$ and $u^{\backslash \text{B}_i} := \{C : \text{B}_i(C) \in u\}$. We define a Kripke model $\mathcal{K} = (W : R_1, ..., R_n)$ and an assignment σ by

M1: $W = \text{Con}(A)$;

M2: $R_i = \{(u,v) \in W \times W : u^{\backslash \text{B}_i} \cup u^{\text{B}_i} \subseteq v\}$ for all $i \in N$;

M3: for any $(w,p) \in W \times PV$, $\sigma(w,p) = \top$ iff $p \in w$.

First, we verify that each R_i is transitive and serial.

Lemma 9.2.(1): R_i is transitive; and **(2)**: R_i is serial.

Proof. (1): Let $(u,v) \in R_i$ and $(v,w) \in R_i$. Take $\text{B}_i(C)$ from u. Then $\text{B}_i(C) \in v$. This implies $\text{B}_i(C) \in w$ and $C \in w$.

(2): Let $u \in S$. Consider the set $u^{\backslash \text{B}_i} \cup u^{\text{B}_i}$. We prove that this is a consistent set. On the contrary, there is a finite subset $\{C_1, ..., C_\ell, \text{B}_i(C_{\ell+1}), ..., \text{B}_i(C_k)\}$ of $u^{\backslash \text{B}_i} \cup u^{\text{B}_i}$ such that $\vdash_{\mathcal{CS}} C_1 \bigwedge ... \bigwedge C_\ell \bigwedge \text{B}_i(C_{\ell+1}) \bigwedge ... \bigwedge \text{B}_i(C_k) \supset \neg D \bigwedge D$. Then $\vdash_{\mathcal{CS}} \text{B}_i(C_1 \bigwedge ... \bigwedge C_\ell) \bigwedge \text{B}_i(C_{\ell+1}) \bigwedge ... \bigwedge \text{B}_i(C_k) \supset \text{B}_i(\neg D \bigwedge D)$. However, by Axiom D, $\vdash_{\mathcal{CS}} \neg \text{B}_i(\neg D \bigwedge D)$. This means that u itself is inconsistent. Hence $u^{-\text{B}_i} \cup u^{\text{B}_i}$ is consistent. Then we have a maximally consistent subset v of $\text{Sub}^+(A)$ including $u^{-\text{B}_i} \cup u^{\text{B}_i}$. Then $(u,v) \in R_i$. ■

Now we prove by induction on the structure of A that for any $C \in \text{Sub}^+(A)$ and any $v \in W$,

$$C \in v \text{ if and only if } (\mathcal{K}, \sigma, v) \models C. \tag{9.1}$$

Suppose that (9.1) is proved. Since A is consistent, there is a $w \in W$ such that w includes A itself. Thus $(\mathcal{K}, \sigma, w) \models A$ by (9.1).

The 0-th step for (9.1) is the basis of the induction proof.

(0): Let C be a propositional variable in $\text{Sub}^+(A)$. Then $C \in v$ if and only if $\sigma(v, C) = \top$. This is further equivalent to $(\mathcal{F}, \sigma, v) \models C$ by E0. This is (9.1).

Consider a formula in $\text{Sub}^+(A)$ which is not a propositional variable. Now we assume the induction hypothesis that for any immediate subformula D of the formula in question, $D \in v$ if and only if $(\mathcal{K}, \sigma, v) \models D$. Then we prove (9.1) for the formula in question. In each step of (1)-(3), we use Lemma 9.1.

(1): Let the formula in question be expressed as $\neg C$. Suppose $\neg C \in v$. Then $C \notin v$. By the induction hypothesis, $(\mathcal{K}, \sigma, v) \not\models C$. Hence $(\mathcal{K}, \sigma, v) \models \neg C$. The converse is similar.

(2): Let the formula in question be $B \supset C$. Suppose $B \supset C \in v$. Then $\neg B \in v$ or $C \in v$. By the induction hypothesis, $(\mathcal{K}, \sigma, v) \not\models B$ or $(\mathcal{K}, \sigma, v) \models C$. Hence $(\mathcal{K}, \sigma, v) \models B \supset C$. The converse is similar.
(3): Similarly, we can prove (9.1) in the case where the formula in question is expressed as $\bigwedge \Phi$ or $\bigvee \Phi$.
(4): Consider the case of $\mathrm{B}_i(C)$. First, we prove that $(\mathcal{K}, \sigma, v) \models \mathrm{B}_i(C)$ implies $\mathrm{B}_i(C) \in v$. Suppose $(\mathcal{K}, \sigma, v) \models \mathrm{B}_i(C)$. We show the inconsistency of $v^{\setminus \mathrm{B}_i} \cup v^{\mathrm{B}_i} \cup \{\neg C\}$. Suppose that this is consistent. There is a maximally consistent set u in $\mathrm{Sub}^+(A)$ including this set. Thus $(v, u) \in R_i$. However, $(\mathcal{F}, \sigma, v) \models \mathrm{B}_i(C)$ implies $(\mathcal{K}, \sigma, u) \models C$. By the induction hypothesis, we have $C \in u$, a contradiction to $\neg C \in u$. By the inconsistency of $v^{\setminus \mathrm{B}_i} \cup v^{\mathrm{B}_i} \cup \{\neg C\}$, there is a finite subset $\{D_1, ..., D_\ell, \mathrm{B}_i(D_{\ell+1}), ..., \mathrm{B}_i(D_k)\}$ of $v^{\setminus \mathrm{B}_i} \cup v^{\mathrm{B}_i}$ such that $\vdash_{CS} D_1 \bigwedge ... \bigwedge D_\ell \bigwedge \mathrm{B}_i(D_{\ell+1}) \bigwedge ... \bigwedge \mathrm{B}_i(D_k) \supset C$. Then $\vdash_{CS} \mathrm{B}_i(D_1 \bigwedge ... \bigwedge D_\ell) \bigwedge \mathrm{B}_i(D_{\ell+1}) \bigwedge ... \bigwedge \mathrm{B}_i(D_k) \supset \mathrm{B}_i(C)$. Hence $\mathrm{B}_i(C) \in v$.

Conversely, suppose $\mathrm{B}_i(C) \in v$. Then $C \in w$ for all w with $(v, w) \in R_i$ by the definition of R_i. Hence $(\mathcal{F}, \sigma, w) \models C$ for all w with $(v, w) \in R_i$ by the induction hypothesis. This implies $(\mathcal{K}, \sigma, v) \models \mathrm{B}_i(C)$.

(5): Now we prove that $\mathrm{C}(D) \in v$ if and only if $(\mathcal{K}, \sigma, v) \models \mathrm{C}(D)$.

Suppose $\mathrm{C}(D) \in v$. We show by induction on k that if w is reachable from v in k steps, then D and $\mathrm{C}(D)$ are in w. Let $k = 1$. Observe that CA implies $\mathrm{B}_i(D) \in v$ and $\mathrm{B}_i\mathrm{C}(D) \in v$. If w is reachable from v in one step, i.e., $(v, w) \in R_i$ for some i, we have $D \in w$ and $\mathrm{C}(D) \in w$. Now we assume the claim for k. Suppose that w is reachable from v in $k+1$ steps. Then there is a u such that is reachable from v in k steps and $(u, w) \in R_i$. By the induction hypothesis, D and $\mathrm{C}(D)$ are in u. By CA, $\mathrm{B}_i(D) \in u$ and $\mathrm{B}_i\mathrm{C}(D) \in u$. Since $(u, w) \in R_i$, we have $D \in w$ and $\mathrm{C}(D) \in w$. In sum, $D \in w$ for all w reachable from v. By our main induction hypothesis, $(\mathcal{K}, \sigma, w) \models D$ for all w reachable from v. Thus $(\mathcal{K}, \sigma, v) \models \mathrm{C}(D)$ by K6.

Conversely, suppose $(\mathcal{K}, \sigma, v) \models \mathrm{C}(D)$. We define $\mathcal{W} := \{w : (\mathcal{F}, \sigma, w) \models \mathrm{C}(D)\}$. Since $\mathcal{W}$ is a set of subsets of $\mathrm{Sub}^+(A)$, this is a finite set. Let φ_w be $\bigwedge w$, i.e., the conjunction of w, and let $\varphi_{\mathcal{W}} := \bigvee_{w \in \mathcal{W}} \varphi_w$. We are going to prove

$$\vdash_{CS} \varphi_{\mathcal{W}} \supset D \bigwedge \mathrm{B}_1(\varphi_{\mathcal{W}}) \bigwedge ... \bigwedge \mathrm{B}_n(\varphi_{\mathcal{W}}). \tag{9.2}$$

Once this is done, we have $\vdash_{CS} \varphi_{\mathcal{W}} \supset \mathrm{C}(D)$ by CI. Since $v \in \mathcal{W}$, we have $\vdash_{CS} \varphi_v \supset \varphi_{\mathcal{W}}$. Hence $\vdash_{CS} \varphi_v \supset \mathrm{C}(D)$. Thus $\mathrm{C}(D) \in v$.

In the remaining, we prove (9.2). First, $\vdash_{CS} \varphi_{\mathcal{W}} \supset D$ is proved as follows. Let w be an arbitrary world in $\mathcal{W}$. Since $(\mathcal{K}, \sigma, w) \models \mathrm{C}(D)$ implies $(\mathcal{K}, \sigma, w) \models D$ by K6. This implies $D \in w$ by the induction hypothesis. Thus $\vdash_{CS} \varphi_{\mathcal{W}} \supset D$. It remains to prove that $\vdash_{CS} \varphi_{\mathcal{W}} \supset \mathrm{B}_i(\varphi_{\mathcal{W}})$. Let w be arbitrary in $\mathcal{W}$ and $i = 1, ..., n$. It suffices to prove $\vdash_{CS} \varphi_w \supset \mathrm{B}_i(\varphi_{\mathcal{W}})$. Since $\vdash_{CS} \varphi_{\mathcal{W}} \equiv \neg \left(\bigvee_{w' \in W - \mathcal{W}} \varphi_{w'}\right)$, $\vdash_{CS} \varphi_w \supset \mathrm{B}_i(\varphi_{\mathcal{W}})$ is equivalent to $\vdash_{CS} \varphi_w \supset \mathrm{B}_i(\bigwedge_{w' \in W - \mathcal{W}} \neg\varphi_{w'})$. The latter follows if we prove that for each $w' \in W - \mathcal{W}$,

$$\vdash_{CS} \varphi_w \supset \mathrm{B}_i(\neg\varphi_{w'}). \tag{9.3}$$

Suppose that (9.3) does not hold for some $w' \in W - \mathcal{W}$. Then $\varphi_w \bigwedge \neg\mathrm{B}_i(\neg\varphi_{w'})$ is consistent. We will show that this implies $w^{\setminus \mathrm{B}_i} \subseteq w'$ and $w^{\setminus \mathrm{B}_i} \subseteq w'^{\setminus \mathrm{B}_i}$. Once this is proved, we have $(w, w') \in R_i$. By the definition of w, w', we have $(\mathcal{K}, \sigma, w) \models \mathrm{C}(D)$

but $(\mathcal{K}, \sigma, w') \not\models \mathrm{C}(D)$. The latter implies that some $e \in N^{<\omega>}$, $(\mathcal{K}, \sigma, w') \not\models \mathrm{B}_e(D)$. However, $(\mathcal{F}, \sigma, w) \models \mathrm{B}_i\mathrm{B}_e(D)$. This together with $(w, w') \in R_i$ implies $(\mathcal{K}, \sigma, w') \models \mathrm{B}_e(D)$. a contradiction. Overall, we have (9.3).

Finally, we show that $w^{\backslash \mathrm{B}_i} \subseteq w'$ and $w^{\backslash \mathrm{B}_i} \subseteq w'^{\backslash \mathrm{B}_i}$ follows from the consistency of $\varphi_w \bigwedge \neg \mathrm{B}_i(\neg \varphi_{w'})$. There are the two cases to be considered. Consider case (1): $E \notin w'$ for some $\mathrm{B}_i(E) \in w$. Then $\neg E \in w'$. Thus, $\vdash_{CS} E \supset \neg \varphi_{w'}$. Then $\vdash_{CS} \mathrm{B}_i(E) \supset \mathrm{B}_i(\neg \varphi_{w'})$, which contradicts that $\varphi_w \bigwedge \neg \mathrm{B}_i(\neg \varphi_{w'})$ is consistent, since $\mathrm{B}_i(E) \in w$. Next, consider case (2): $\mathrm{B}_i(E) \notin w'$ for some $\mathrm{B}_i(E) \in w$. Then $\neg \mathrm{B}_i(E) \in w'$. Then $\vdash_C \mathrm{B}_i(E) \supset \neg \varphi_{w'}$. This implies $\vdash_{CS} \mathrm{B}_i\mathrm{B}_i(E) \supset \mathrm{B}_i(\neg \varphi_{w'})$, which implies $\vdash_{CS} \mathrm{B}_i(E) \supset \mathrm{B}_i(\neg \varphi_{w'})$. Again, we have a contradiction to the consistency of $\varphi_w \bigwedge \neg \mathrm{B}_i(\neg \varphi_{w'})$. ■

10. Appendix

Our formulation of CL is quite efficient as an axiomatization, but the cost for an efficient axiomatization is practically difficult to prove some necessary steps. For example, the following three claims are needed for developments of our analysis, but they need a lot of tedious steps, which is given in the end of this appendix.

Lemma 10.1.(a): $\vdash_0 (\neg A \supset \neg B) \supset (B \supset A)$; **(b):** $A \supset (B \supset C) \vdash_0 A \bigwedge B \supset C$; **(c):** $A \bigwedge B \supset C \vdash_0 A \supset (B \supset C)$.

Lemma 10.2 (Deduction Theorem): $\Gamma \cup \{A\} \vdash_0 B$ implies $\Gamma \vdash_0 A \supset B$.

Proof. Let P be a proof of B from $\Gamma \cup \{A\}$. We prove by induction on the tree structure of P from its leaves that $\Gamma \vdash_0 A \supset C$ for any C in P. Let C be a formula associated with a leaf of P. Then C is an instance of L1–L5 or a formula in Γ. In either case, $\Gamma \vdash_0 C$. Since $\vdash_0 C \supset (A \supset C)$ by L1, we have $\Gamma \vdash_0 A \supset C$.

Now, let C be a formula associated with a non-leaf node in P. We assume the induction hypothesis that the induction assertion holds for the upper formulae of C in P. We should consider three cases: MP, $\bigwedge$-Rule and $\bigvee$-Rule. We consider MP and $\bigwedge$-Rule.

Suppose that C is inferred from D and $D \supset C$ by MP. The induction hypothesis is that $\Gamma \vdash_0 A \supset D$ and $\Gamma \vdash_0 A \supset (D \supset C)$. By Lemma 3.1.(1) and $\bigwedge$-Rule, $\Gamma \vdash_0 A \supset A \bigwedge D$, and by Lemma 10.1.(b), $\Gamma \vdash_0 A \bigwedge D \supset C$. By Lemma 3.1.(2), we have $\Gamma \vdash_0 A \supset C$.

Suppose that $D \supset \bigwedge \Phi$ is inferred from $\{D \supset E : E \in \Phi\}$. The induction hypothesis is that $\Gamma \vdash_0 A \supset (D \supset E)$ for all $E \in \Phi$. By Lemma 10.1.(b), $\Gamma \vdash_0 A \bigwedge D \supset E$ for all $E \in \Phi$. By $\bigwedge$-Rule, $\Gamma \vdash_0 A \bigwedge D \supset \bigwedge \Phi$. By Lemma 10.1.(c), $\Gamma \vdash_0 A \supset (D \supset \bigwedge \Phi)$. ■

Then we have the following lemma.

Lemma 10.3. $\Gamma \nvdash_0 A$ if and only if $\Gamma \cup \{\neg A\}$ is consistent in CL.

Proof. We denote $\Gamma \cup \{\neg A\}$ by Γ'. Suppose $\Gamma \vdash_0 A$. Then $\Gamma' \vdash_0 A$ and $\Gamma' \vdash_0 \neg A$. Hence $\Gamma' \vdash_0 (A \supset A) \supset A$ and $\Gamma' \vdash_0 (A \supset A) \supset \neg A$ by L1. Thus, $\Gamma' \vdash_0 (A \supset A) \supset \neg A \bigwedge A$. By Lemma 3.1.(1), we have $\Gamma' \vdash_0 \neg A \bigwedge A$, i.e., Γ' is inconsistent.

Suppose that Γ' is inconsistent. Then $\Gamma' \vdash_0 \neg C \bigwedge C$. By Lemma 10.2 and L4, we have $\Gamma \vdash_0 \neg A \supset \neg C$ and $\Gamma \vdash_0 \neg A \supset C$. These together with L3 imply $\Gamma \vdash_0 A$. ■

Proof of the Equivalence of the if Parts of Theorem 3.3: Suppose the *if* part of (1). We prove the negative form that of (2). Suppose that there is no model κ of Γ. Then $\Gamma \models \neg C \bigwedge C$ for some C. By (1), we have $\Gamma \vdash_0 \neg C \bigwedge C$.

Suppose the *if* part of (2). Let $\Gamma \nvdash_0 A$. Then $\Gamma \cup \{\neg A\}$ is consistent with respect to $\vdash_0$ by Lemma 10.3. By (2), there is a model κ of $\Gamma \cup \{\neg A\}$. This means $V_\kappa(A) = \bot$. Hence $\Gamma \not\models A$. ■

Proof of the Completeness of CL: It suffices to prove that if Γ is consistent, there is an assignment κ such that $V_\kappa(B) = \top$ for all $B \in \Gamma$. In the following, we suppose that Γ is consistent in CL.

We order the set $\mathcal{P}$ as follows: $A_1, A_2, \ldots$ We construct a sequence $\Gamma_0, \Gamma_1, \ldots$ by induction on $A_1, A_2, \ldots$ as follows:

G-0: $\Gamma_0 = \Gamma$;

and for any $m > 0$,

$$\text{G-1: } \Gamma_m = \begin{cases} \Gamma_{m-1} \cup \{A_m\} & \text{if } \Gamma_{m-1} \cup \{A_m\} \text{ is consistent} \\ \Gamma_{m-1} & \text{if } \Gamma_{m-1} \cup \{A_m\} \text{ is inconsistent.} \end{cases}$$

We define $\tilde{\Gamma} = \bigcup_m \Gamma_m$.

Lemma 10.4.(1): Each Γ_m is consistent and $\Gamma_0 \subseteq \Gamma_1 \subseteq \ldots$;

(2): $\tilde{\Gamma}$ is consistent.

Proof. (1): This follows the definition of Γ_m.

(2): Suppose that $\tilde{\Gamma}$ is inconsistent. Then there is a proof of $\neg A \bigwedge A$ from a finite subset Γ' of $\tilde{\Gamma}$. This implies $\neg A \bigwedge A \in \Gamma_m$ for some m, a contradiction to (1). ■

Lemma 10.5. The set $\tilde{\Gamma}$ defined above is maximally consistent, i.e., there is no other consistent set strictly including $\tilde{\Gamma}$, and satisfies the following properties:

(0): $\tilde{\Gamma} \vdash_0 A \Leftrightarrow A \in \tilde{\Gamma}$;

(1): either $A \in \tilde{\Gamma}$ or $\neg A \in \tilde{\Gamma}$;

(2): $A \supset B \in \tilde{\Gamma} \Leftrightarrow A \notin \tilde{\Gamma}$ or $B \in \tilde{\Gamma}$;

(3): $\bigwedge \Phi \in \tilde{\Gamma} \Leftrightarrow A \in \tilde{\Gamma}$ for all $A \in \Phi$;

(4): $\bigvee \Phi \in \tilde{\Gamma} \Leftrightarrow A \in \tilde{\Gamma}$ for some $A \in \Phi$.

Proof. The maximality of $\tilde{\Gamma}$ follows from the definition of each Γ_m.

(0): The left direction is immediate. The other direction follows from the definition of $\tilde{\Gamma}$.

(1): It is impossible that both A and $\neg A$ are in $\tilde{\Gamma}$ by Lemma 10.4.(2). If $A = A_m \notin \tilde{\Gamma}$, then $\Gamma_{m-1} \cup \{A_m\}$ is inconsistent. By Lemma 10.3, we have $\Gamma_{m-1} \vdash_0 \neg A_m$. By (0), we have $\neg A \in \tilde{\Gamma}$. The other case is symmetric.

(2): Let $A \supset B \in \hat{\Gamma}$. If $A \in \tilde{\Gamma}$, then $B \in \tilde{\Gamma}$ by MP. Conversely, let $A \notin \tilde{\Gamma}$ or $B \in \tilde{\Gamma}$. First, consider the case: $B \in \tilde{\Gamma}$. Then by L1, we have $A \supset B \in \tilde{\Gamma}$. Second, let $A \notin \tilde{\Gamma}$. Then $\neg A \in \tilde{\Gamma}$ by (1). Since $\neg A \bigwedge A \notin \tilde{\Gamma}$, we have $\neg\neg A \bigwedge A \in \tilde{\Gamma}$. Hence $\neg B \supset \neg\neg A \bigwedge A \in \tilde{\Gamma}$ by L1. By Lemma 10.1.(a), we have $\neg A \bigwedge A \supset B \in \tilde{\Gamma}$. This is equivalent to $\neg A \supset (A \supset B) \in \tilde{\Gamma}$. Hence $(A \supset B) \in \tilde{\Gamma}$.

(3): Suppose $\bigwedge \Phi \in \tilde{\Gamma}$. Since $\vdash_0 \bigwedge \Phi \supset A$ for all $A \in \Phi$, we have $\tilde{\Gamma} \vdash_0 A$ for all $A \in \Phi$.

Conversely, suppose that $A \in \tilde{\Gamma}$ for all $A \in \Phi$. By (0), $\tilde{\Gamma} \vdash_0 A$ for all $A \in \Phi$. By L1, we have $\tilde{\Gamma} \vdash_0 (C \supset C) \supset A$ for all $A \in \Phi$. Hence $\tilde{\Gamma} \vdash_0 (C \supset C) \supset \bigwedge \Phi$.
(4): Suppose $A \notin \tilde{\Gamma}$ for all $A \in \Phi$. Let C be any formula. By (2), $A \supset C \in \tilde{\Gamma}$ and $A \supset \neg C \in \tilde{\Gamma}$ for all $A \in \Phi$. By (0) and $\bigvee$-Rule, we have $\bigvee \Phi \supset C \in \tilde{\Gamma}$ and $\bigvee \Phi \supset \neg C \in \tilde{\Gamma}$. By L3, we have $\neg \bigvee \Phi \in \tilde{\Gamma}$. Conversely, suppose that $A \in \tilde{\Gamma}$ for some $A \in \Phi$. By L5, we have $\bigvee \Phi \in \tilde{\Gamma}$. ■

Now we can define an assignment κ by $\kappa(p) = \top$ iff $p \in \tilde{\Gamma}$. Using this κ, we have V_κ by E0–E4 of Section 4. It remains to prove $V_\kappa(A) = \top$ for all $A \in \Gamma$. For this purpose, it suffices to show that $V_\kappa(A) = \top$ if and only if $A \in \tilde{\Gamma}$. This is proved by induction on the structure of a formula: Lemma 10.5 is used for the following steps.

(1): for any $p \in PV$, $V_\kappa(p) = \top \Leftrightarrow \kappa(p) = \top \Leftrightarrow A \in \tilde{\Gamma}$;
(2): $V_\kappa(\neg A) = \top \Leftrightarrow V_\kappa(A) = \perp \Leftrightarrow A \notin \tilde{\Gamma} \Leftrightarrow \neg A \in \hat{\Gamma}$;
(3): $V_\kappa(A \supset B) = \top \Leftrightarrow V_\kappa(A) = \perp$ or $V_\kappa(B) = \top \Leftrightarrow A \notin \tilde{\Gamma}$ or $B \in \hat{\Gamma} \Leftrightarrow A \supset B \in \hat{\Gamma}$;
(4): $V_\kappa(\bigwedge \Phi) = \top \Leftrightarrow V_\kappa(A) = \top$ for all $A \in \Phi \Leftrightarrow A \notin \tilde{\Gamma}$ for all $A \in \Phi \Leftrightarrow \bigwedge \Phi \in \hat{\Gamma}$;
(5): $V_\kappa(\bigvee \Phi) = \top \Leftrightarrow V_\kappa(A) = \top$ for some $A \in \Phi \Leftrightarrow A \in \tilde{\Gamma}$ for some $A \in \Phi \Leftrightarrow \bigvee \Phi \in \hat{\Gamma}$. ■

Proof of Lemma 10.1.[29] (a),(b) and (c) are proved as (7), (17) and (18). In the following, we use Lemma 3.1.(2) without references.

(1): $\vdash_0 (B \supset C) \supset ((A \supset B) \supset (A \supset C))$.
*): Since $(B \supset C) \supset (A \supset (B \supset C))$ and $(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$ are instances of L1 and L2, we have, by Lemma 3.1.(2), $\vdash_0 (B \supset C) \supset ((A \supset B) \supset (A \supset B))$.

(2): $A \supset (B \supset C) \vdash_0 B \supset (A \supset C)$.
*): By L2, we have $A \supset (B \supset C) \vdash_0 (A \supset B) \supset (A \supset C)$. This together with $B \supset (A \supset B)(-\text{L1})$ implies $A \supset (B \supset C) \vdash_0 B \supset (A \supset C)$ by Lemma 3.1.(2).

(3): $\vdash_0 (A \supset B) \supset ((B \supset C) \supset (A \supset C))$.
*): Regarding $B \supset C$, $A \supset B$ and $A \supset C$ as A, B and C of (2), we have, by (1), $\vdash_0 (A \supset B) \supset ((B \supset C) \supset (A \supset C))$.

(4): $\vdash_0 A \supset ((A \supset B) \supset B)$.
*): Regarding $A \supset B$ and A as A and B of (2), we have, using Lemma 3.1.(1), $\vdash_0 A \supset ((A \supset B) \supset B))$.

(5): $\vdash_0 (A \supset (A \supset B)) \supset (A \supset B)$.
*): Since $(A \supset ((A \supset B) \supset B)) \supset ((A \supset (A \supset B) \supset (A \supset B))$ is an instance of L2 and $\vdash_0 A \supset ((A \supset B) \supset B)$ by (4), we have $\vdash_0 (A \supset (A \supset B)) \supset (A \supset B)$.

(6): $\vdash_0 \neg\neg A \supset A$ (the law of double negation).
*): Since $\neg\neg A \supset (\neg A \supset \neg\neg A)$ and $(\neg A \supset \neg\neg A) \supset ((\neg A \supset \neg A) \supset A)$ are instance of L1 and L3, respectively, we have, by Lemma 3.1.(2), $\vdash_0 \neg\neg A \supset ((\neg A \supset \neg A) \supset A)$. Hence we have, by (2), $\vdash_0 (\neg A \supset \neg A) \supset (\neg\neg A \supset A)$. By Lemma 3.1.(1), we have $\vdash_0 \neg\neg A \supset A$.

(7): $\vdash_0 (\neg A \supset \neg B) \supset (B \supset A)$.

[29] The following proof is due to T. Nagashima.

*) Since $\vdash_0 B \supset (\neg A \supset B)$ by L1 and $\vdash_0 [B \supset (\neg A \supset B)] \supset [((\neg A \supset B) \supset A) \supset (B \supset A)$ by (3), we have $\vdash_0 ((\neg A \supset B) \supset A) \supset (B \supset A)$. This and L3 imply $\vdash_0 (\neg A \supset \neg B) \supset (B \supset A)$.

(8): $\vdash_0 A \supset \neg\neg A$ (the converse of (6)).

*): Since $\vdash_0 \neg\neg\neg A \supset \neg A$ by (6) and $\vdash_0 (\neg\neg\neg A \supset \neg A) \supset (A \supset \neg\neg A)$ by (7), we have $\vdash_0 A \supset \neg\neg A$.

(9): $\vdash_0 (\neg A \supset B) \supset (\neg B \supset A)$.

*) Since $\vdash_0 (B \supset \neg\neg B) \supset ((\neg A \supset B) \supset (\neg A \supset \neg\neg B))$ by (1) and $\vdash_0 B \supset \neg\neg B$ by (8), we have $\vdash_0 (\neg A \supset B) \supset (\neg A \supset \neg\neg B)$. This together with $\vdash_0 (\neg A \supset \neg\neg B) \supset (\neg B \supset A)$ by (7) implies $\vdash_0 (\neg A \supset B) \supset (\neg B \supset A)$.

(10): $\vdash_0 (A \supset B) \supset (\neg B \supset \neg A)$.

*): Since $\vdash_0 (B \supset \neg\neg B) \supset ((A \supset B) \supset (A \supset \neg\neg B))$ by (1) and $\vdash_0 B \supset \neg\neg B$ by (8), we have $\vdash_0 (A \supset B) \supset (A \supset \neg\neg B)$. This together with $\vdash_0 (A \supset \neg\neg B) \supset (\neg B \supset \neg A)$ by (7) implies $\vdash_0 (A \supset B) \supset (\neg B \supset \neg A)$.

(11) $\vdash_0 A \supset (B \supset \neg(A \supset \neg B))$.

*) Since $\vdash_0 A \supset ((A \supset \neg B) \supset \neg B)$ by (6) and $\vdash_0 ((A \supset \neg B) \supset \neg B) \supset (B \supset \neg(A \supset \neg B))$ by (10), we have $\vdash_0 A \supset (B \supset \neg(A \supset \neg B))$.

(12): $\vdash_0 \neg A \supset (A \supset B)$.

*) Since $\neg A \supset (\neg B \supset \neg A)$ is an instance of L1, we have, by (7), $\vdash_0 \neg A \supset (A \supset B)$.

(13): $\vdash_0 \neg(A \supset \neg B) \supset A$.

*): Since $\vdash_0 \neg A \supset (A \supset \neg B)$ by (12) and $\vdash_0 (\neg A \supset (A \supset \neg B)) \supset (\neg(A \supset \neg B) \supset A)$ by (10), we have $\vdash_0 \neg(A \supset \neg B) \supset A$.

(14): $\vdash_0 \neg(A \supset \neg B) \supset B$.

*): Since $\vdash_0 \neg B \supset (A \supset \neg B)$ by L1 and $\vdash_0 (\neg B \supset (A \supset \neg B)) \supset (\neg(A \supset \neg B) \supset B)$ by (9), we have $\vdash_0 \neg(A \supset \neg B) \supset B$.

(15): $\vdash_0 \neg(A \supset \neg B) \supset A \bigwedge B$.

*): Using $\bigwedge$-Rule, it follows from (13) and (14) that $\vdash_0 \neg(A \supset \neg B) \supset A \bigwedge B$.

(16): $\vdash_0 A \supset (B \supset A \bigwedge B)$.

*) Since $\vdash_0 (\neg(A \supset \neg B) \supset A \bigwedge B) \supset [B \supset \neg(A \supset \neg B)) \supset (B \supset A \bigwedge B)]$ by (1), we have $\vdash_0 (B \supset \neg(A \supset \neg B)) \supset (B \supset A \bigwedge B)$ using (15). This and (11) imply $\vdash_0 A \supset (B \supset A \bigwedge B)$.

(17): $A \supset (B \supset C) \vdash_0 A \bigwedge B \supset C$.

*): Since $\vdash_0 A \bigwedge B \supset A$ by L4, we have $A \supset (B \supset C) \vdash_0 A \bigwedge B \supset (B \supset C)$. By (2), $A \supset (B \supset C) \vdash_0 B \supset (A \bigwedge B \supset C)$. This together with $\vdash_0 A \bigwedge B \supset B$ by L4 implies $A \supset (B \supset C) \vdash_0 A \bigwedge B \supset (A \bigwedge B \supset C)$, so $A \supset (B \supset C) \vdash_0 A \bigwedge B \supset C$ by (5).

(18): $A \bigwedge B \supset C \vdash_0 A \supset (B \supset C)$.

*) Since $\vdash_0 (A \bigwedge B \supset C) \supset ((B \supset A \bigwedge B)) \supset (B \supset C))$ by (1) and $A \bigwedge B \supset C$ is an assumption, we have $A \bigwedge B \supset C \vdash_0 (B \supset A \bigwedge B) \supset (B \supset C)$. This and (16) imply $A \bigwedge B \supset C \vdash_0 A \supset (B \supset C)$. ■

References

[1] Aumann, R. J., (1976), Agreeing to Disagree, *Annals of Statistics* 4, 1236–1239.

[2] Bacharach, M., and P. Mongin, Epistemic Logic and the Foundations of Game Theory, *Theory and Decision* 37, 1–6.

[3] Bachrach, M., L. A. Gérard-Varet, P. Mongin, H. S. Shin, (1997), Epistemic Logic and the Theory of Games and Decisions, Kluwwer Academic Publisher, Boston.

[4] Boolos, G., (1979), *The Unprovability of Consistency*, Cambridge University Press.

[5] Chellas, B., (1980), *Modal Logic*, Cambridge University Press, Cambridge.

[6] Fagin, R., J.Y. Halpern, Y. Moses and M. Y. Verdi, (1995), *Reasoning about Knowledge*, The MIT Press, Cambridge.

[7] Halpern, J. H. and Y. Moses, (1992), A Guide to Completeness and Complexity for Modal Logics of Knowledge and Beliefs, *Artificial Intelligence* 54, 319-379.

[8] Heifetz, A., (1999), Iterative and Fixed Point Common Belief, *Journal of Philosophical Logic* 28, 61-79.

[9] Hintikka, J., (1962), *Knowledge and Belief*, Cornell University Press.

[10] Gentzen, G., (1935), Untersuchungen über das logische Schliessen, *Mathematische Zeitschrift* 39, 176-210, 405-431. English translation, Investigations into Logical Deduction, *The Collected Papers of Gerhard Gentzen,* (1969).

[11] Hughes, G. E., and M. J. Cresswell, *A Companion to Modal Logic*, Methuen &Co. London.

[12] Kaneko, M., (1999a), Epistemic Considerations of Decision Making in Games. *Mathematical Social Sciences* 38, 105–137.

[13] Kaneko, M., (1999b), Common Knowledge Logic and Game Logic, *Journal of Symbolic Logic* 64, 685–700.

[14] Kaneko, M., and A. Matsui, (1999), Inductive Game Theory: Discrimination and Prejudices, *Journal of Public Economic Theory* 1, 101–137.

[15] Kaneko, M., and T. Nagashima, Game Logic and its Applications, Part I, (1996), *Studia Logica* 57, 325–354. Part II, (1997), *Studia Logica* 58, 273–303.

[16] Kaneko, M., and T. Nagashima, (1997b), Axiomatic Indefinability of Common Knowledge in Finitary Logics. *Epistemic Logic and the Theory of Games and Decision,* eds. M. Bacharach, *et al.* Kluwer Academic Press, 69–93.

[17] Kaneko, M., and N-Y. Suzuki, (1999), Epistemic Models of Shallow Depths and Decision Making in Games, IPPS. DP. No.828, University of Tsukuba

[18] Kaneko, M., and N-Y. Suzuki, (2000), Logico-Game Analysis of Decision Making and Epistemic Logics of Shallow Depths, this issue.

[19] Kleene, S. C., (1967), *Mathematical Logic*, John Wiley & Sons, Inc. New York.

[20] Lismont, L. and P. Mongin, (1994), On the Logic of Common Belief and Common Knowledge, *Theory and Decision* 37, 75-106.

[21] Limont. L., and P. Mongin (1997), On the Logic of Common Belief and Common Knowledge, *Epistemic Logic and the Theory of Games and Decision,* eds. M. Bacharach, *et al.* Kluwer Academic Press, 3–34.

[22] Mendelson, E., (1987), *Introduction to Mathematical Logic,* Wadsworth.

[23] Meyer, J.-J. Ch., and W. van der Hoek, *Epistemic Logic for AI and Computer Science*, Cambridge University Press, (1995).

[24] Mithen, S., (1996), *The Prehistory of the Mind*, Thames and Hudoson Ltd, London.

[25] Moser, P. K., (1986), Empirical Knowledge – Readings in Contemporary Epistemology, Rowman & Littlefield Publishers, Inc. Chicago.

[26] Moulin, H., (1982), *Game Theory for the Social Sciences*, New York University Press, New York.

[27] Myerson, R. B., (1991), *Game Theory*, Harvard University Press, Cambridge.

[28] Ohnishi, M., and K. Matsumoto, (1957), Gentzen Method in Modal Calculi. I, *Osaka Math. J.* 9, 113-130.

[29] Plato, *The Republic of Plato*, Translated by F. M. Cornford, Oxford University Press, (1941), London.

[30] Takeuti, G., (1987), Proof Theory, North-Holland, Studies in Logic and The Foundations of Mathematics Vol.81, Amsterdam.

[31] Terutoshi, Y. Okitsu K. and S. Enomoto (eds.), (1980), Collections of Meiji-Taisho Comic Stories (in Japanese), Vol. 3. Kodansha, Tokyo.

[32] Wolter, F. (1999), First Order Common Knowledge Logics. To appear in *Studia Logica.*