**No. 724**

# Epistemic Considerations of Decision Making in Games

by

Mamoru KANEKO

April 1997

# Epistemic Considerations of Decision Making in Games

Mamoru KANEKO*
Institute of Policy and Planning Sciences, University of Tsukuba,
Ibaraki 305, Japan (kaneko@shako.sk.tsukuba.ac.jp)

April 1997

**Abstract**

From the *Ex Ante* point of view, an axiomatization of decision making in a game with pure strategies is given with considerations of its epistemic aspects in propositional game (epistemic) logic. Our axiomatization consists of four base axioms for predicted final decisions. One of them is an epistemic requirement, which together with the others leads us to an infinite regress of the knowledge of these axioms. This infinite regress is, in fact, the common knowledge of them. We give meta-theoretical evaluations of the derivation of this infinite regress, and consider its implications in the cases of solvable and unsolvable games (in Nash's sense). For a solvable game, it determines predicted decisions to be the common knowledge of a Nash equilibrium, and for an unsolvable game, it is the common knowledge of a subsolution in Nash's sense. The latter result needs the common knowledge of the additional information of which subsolution would be played. We give also meta-theoretical evaluations of these results.

## 1. Game Logic Approach and Meta-Theoretic Evaluations of Some Game Theoretic Considerations

In this paper, decision making in a game is considered from the *Ex Ante* point of view in an axiomatic manner. For such decision making, players' knowledge and thinking on the game situation are essential. To describe these epistemic aspects as well as the game situation, we will use the propositional fragment of game logic developed in Kaneko-Nagashima [14] and [15]. In the framework of game logic, Kaneko-Nagashima [12], [14] presented base axioms for decision making, and showed that the axioms lead to an infinite regress of the knowledge of the base axioms, which is the common knowledge of them. Then they solved this infinite regress and showed that it determines the

*The research of this paper was partially supported by the Tokyo Center of Economic Research.

final decisions to be the common knowledge of a Nash equilibrium under the common knowledge of the interchangeability of Nash equilibria. The objectives of this paper are to evaluate the derivation of the infinite regress in a meta-theoretic manner; and to give a full consideration of decision making in an unsolvable game as well as meta-theoretic evaluations of it.

In the *Ex Ante* decision making, since each player makes a strategy choice before the actual play of a game, the knowledge of the structure of the game as well as predictions on the other players' strategy choices may be needed. In the literature of game theory, the Bayesian approach to this problem has been dominant since Harsanyi [6]. In the Bayesian approach, the players' knowledge is described by means of subjective probability on possible types of each player, and classical game theory is treated as a trivial case – games with complete information. A game with complete information itself is, however, not trivial in that it has at least the description of the constituents of a game. Although the Bayesian approach has been shown to be quite rich in capturing various economic problems, it is incapable in treating players' logical and mathematical abilities as well as their knowledge of the descriptions of a game in a direct manner. This leads us to the development of the game logic framework.

Game logic is an infinitary predicate logic: infinitary conjunctions are allowed to describe common knowledge explicitly as a conjunctive formula, and it is a predicate logic for the purpose of describing real number theory in its scope, since classical game theory often relies upon real number theory. The objectives of [14] and [15] were to develop the new framework and to show some possible applications. The purpose of this paper is to give fuller discussions on the epistemic axiomatization of the *Ex Ante* decision making in a game.

We meet two kinds of basic problems arising in the *Ex Ante* decision making:

(i): *solution-theoretic* problems;

(ii): *existence-playability* problems.

The first is the problem of how strategy choices are made; and the other is the existence and playability of the solution concept obtained in (i). These problems interact with each other. In this paper, we give the solution-theoretic consideration of the *Ex Ante* decision making problem for a finite game with pure strategies, and will give small results on the playability problem.

When we restrict our attention to games with pure strategies, some games might not playable because of no Nash equilibria in pure strategies. When we allow mixed strategies, existence is obtained, but some form of real number theory is involved and is assumed to be known to players, which is a stringent requirement. If a game has a Nash equilibrium in pure strategies, playability is not so serious as in the case with mixed strategies, though some playability problems may remain for such a game. From the playability point of view, the cases with and without mixed strategies are totally different, but from the solution-theoretic point of view, they do not make much differences. The solution-theoretic considerations given in this paper can be carried over to the case with mixed strategies.

The playability of a game with mixed strategies will be discussed in a separate paper. Restricting our attention to games with pure strategies, the propositional fragment of game logic suffices for our considerations.

A merit of the game logic approach is not only to help us describe the epistemic aspects of decision making in an explicit manner, but also to enables us to evaluate such descriptions and resulting outcomes from them in a meta-theoretic manner. The undecidability result given in Kaneko-Nagashima [14], that the existence of a unique Nash equilibrium is common knowledge (in mixed strategies) but the players cannot know what the Nash equilibrium is specifically, is an example of such meta-theoretical evaluations. The main contributions of this paper are meta-theoretical evaluations of the axiomatization and of the resulting outcomes from it.

Our axiomatization consists of four base axioms for predicted final decisions: D1: *Best Response to Predicted Decisions*; D2: *Identical Predictions*; D3: *Knowledge of Predictions*; and D4: *Interchangeability of Predicted Decisions*. The first two simply induce Nash equilibrium, but they together with the others go much further. The third is an epistemic requirement, and the fourth is a requirement of independent decision making. These two additional requirements differentiate substantially our theory from classical Nash equilibrium theory.

The third axiom together with the second leads us to an infinite regress of the knowledge of those four axioms, which forms the common knowledge of them. We will evaluate this derivation. The evaluation states that to complete the axiomatization, we meet necessarily the infinite regress. The introduction of an epistemic structure also enables us to demarcate playability from the mere knowledge of the existence of a Nash equilibrium, which will be discussed in Subsection 5.2. One result differentiates the case where the players have abstract knowledge of a game from the case where they know the concrete structure of it.

The fourth axiom demarcates between the solvable and unsolvable games in Nash's [19] sense. For a solvable game, a predicted final decision profile is characterized to be the common knowledge of a Nash equilibrium. For an unsolvable game, it is characterized to be the common knowledge of a subsolution, where the players need to share some information of which subsolution would be played. Without such information, the game is not playable.

Our axiomatization can be regarded as a materialization of Johansen's [8] informal argument on Nash equilibrium. The seemingly self-referential nature of his argument corresponds to our infinite regress of the base axioms. His claim that his postulates determine a Nash equilibrium for a game with a unique Nash equilibrium is consistent with our results. The comparison between his and ours will be given in Subsection

3.2. This paper is related also to Aumann-Brandenberger [1] in their objectives and scopes. These authors considered some epistemic conditions for Nash equilibrium in a Bayesian framework. The difference is that they concern necessary conditions for Nash equilibrium, while our concern is the complete characterization of predicted final decisions with meta-theoretical evaluations (the direct comparisons are difficult since the frameworks are different).[1]

As already stated, we use a propositional fragment of game logic of Kaneko-Nagashima [14] and [15]. In fact, it is shown in Kaneko [9] that common knowledge logic developed in Halpern-Moses [5] and Lismont-Mongin [18] can be *faithfully* embedded into our logic (with a slight restriction on our logic). This implies that the results obtained in this paper are all translated into common knowledge logic.

We repeat some results given in [12] and [14], but give proofs to some of them for completeness. We distinguish the results already given by putting * from new ones.

## 2. Preliminaries

### 2.1. Game Theoretical Concepts in the Nonformalized Language

In this subsection, we give basic game theoretical concepts in the nonformalized language, which will be redescribed in the formalized language in Subsection 2.3.

Consider an $n$-person noncooperative game $g$. The players are denoted by $1, ..., n$, and each player $i$ has $\ell_i$ (pure) strategies. We assume that the players do not play mixed strategies. Player $i$'s strategy space is denoted by $\Sigma_i = \{\mathbf{s}_{i1}, ..., \mathbf{s}_{i\ell_i}\}$, and his *payoff function* is a real-valued function $g_i$ on $\Sigma = \{\mathbf{s}_{11}, ..., \mathbf{s}_{1\ell_1}\} \times \cdots \times \{\mathbf{s}_{n1}, ..., \mathbf{s}_{n\ell_n}\}$ for $i = 1, ..., n$. We call a vector in $\Sigma = \Sigma_1 \times \cdots \times \Sigma_n$ a *strategy profile.*

A strategy profile $a = (a_1, ..., a_n)$ is called a Nash *equilibrium* iff for $i = 1, ..., n$,

$$g_i(a) \geq g_i(b_i; a_{-i}) \text{ for all } b_i \in \Sigma_i,$$

where $a_{-i} = (a_1, ..., a_{i-1}, a_{i+1}, ..., a_n)$ and $(b_i; a_{-i}) = (a_1, ..., a_{i-1}, b_i, a_{i+1}, ..., a_n)$. We denote the set of Nash equilibria by $E_g$.

Consider a maximal nonempty subset $E$ of $E_g$ which satisfies the *interchangeability* condition:

$$a, b \in E \text{ and } i = 1, ..., n \text{ imply } (a_i; b_{-i}) \in E. \quad (2.1)$$

This is equivalent to that $a^1, ..., a^n \in E$ implies $(a^1_1, ..., a^n_n) \in E$, which states that if each player $i$ independently chooses his equilibrium strategy $a^i_i$, the resulting profile $(a^1_1, ..., a^n_n)$ is also an equilibrium. We call such a maximal set a *subsolution,* which was

[1] Bacharach [2] is the seminal paper along the line of the research of this paper. Nevertheless, his framework is not sufficient to facilitate the considerations of the epistemic aspects – for example, common knowledge is not formulated in his framework explicitly.

introduced by Nash [19]. Each Nash equilibrium belongs to *at least one* subsolution. We denote the subsolutions of game $g$ by $E_g^1, ..., E_g^\sigma$. We stipulate that when $E_g$ is empty, $\sigma = 0$. When $E_g$ is nonempty, $\sigma \geq 1$. When $\sigma = 1$, game $g$ is said to be *solvable.*[2]

| | $s_{21}$ | $s_{22}$ |
|---|---|---|
| $s_{11}$ | $(2,1)^*$ | $(0,0)$ |
| $s_{12}$ | $(0,0)$ | $(1,2)^*$ |

Table 2.1

| | $s_{21}$ | $s_{22}$ |
|---|---|---|
| $s_{11}$ | $(2,2)^*$ | $(1,2)^*$ |
| $s_{12}$ | $(2,1)^*$ | $(0,0)$ |

Table 2.2.

The game (Battle of the Sexes) of Table 2.1 has $E_g = E_g^1 \cup E_g^2 = \{(s_{11}, s_{21})\} \cup \{(s_{12}, s_{22})\}$, and is not solvable. The game of Table 2.2 is not solvable either and has $E_g = E_g^1 \cup E_g^2 = \{(s_{11}, s_{21}), (s_{11}, s_{22})\} \cup \{(s_{11}, s_{21}), (s_{12}, s_{21})\}$. Here $(s_{11}, s_{21})$ belongs to both subsolutions. The games of Tables 2.3 (Prisoner's Dilemma) and 2.4 have the same equilibrium sets $E_g = \{(s_{12}, s_{22})\}$ and are solvable.[3]

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

Table 2.3

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

Table 2.4

In the next subsection, we give a formal logic in which a game as well as such other epistemic constituents are described.

## 2.2. Game Logic $GL_\omega$

In the following, we use some terms of predicate logic. Since, however, we use neither variables nor quantifiers, the following logic is essentially propositional.

We start with the following list of symbols:

*constant symbols*: $s_{11}, ..., s_{1\ell_1};\ s_{21}, ..., s_{2\ell_2};\ ...;\ s_{n1}, ..., s_{n\ell_n}$ ;

*binary predicate*: $=$ ;

*2n–ary predicates*: $R_1, ..., R_n$ ;

*n–ary predicates*: $D_1, ...., D_n$ ;

*knowledge operators*: $K_1, ..., K_n$ ;

[2] Zero-sum two-person games are solvable if they have Nash equilibria. Other sufficient conditions for solvability are found in Kats-Thisse [17]. A study of subsolutions is found in Jansen [7] (see also its references).

[3] We can modify the games of Table 2.2–2.4 so that they keep the same equilibrium sets but have no dominance structures. See Subsection 7.3.

*logical connectives*: $\neg$ (not), $\supset$ (implies), $\bigwedge$ (and), $\bigvee$ (or) ;
*parentheses*: ( , ).

As introduced in Subsection 2.1, the constants $\mathbf{s}_{11}, ..., \mathbf{s}_{1\ell_1};\ \mathbf{s}_{21}, ..., \mathbf{s}_{2\ell_2};\ ...;\ \mathbf{s}_{n1}, ..., \mathbf{s}_{n\ell_n}$ are the players' (pure) strategies. The binary predicate $=$ is intended to describe identity between strategies for each *single* player. The $2n$–ary predicate $R_i(...;...)$ is player $i$'s preferences over strategy profiles describing his payoff function $g_i$. The $n$-ary predicate $D_i(...)$ describes player $i$'s prediction of the players' strategy choices, that is, $D_i(a_1, ..., a_n)$ means that player $i$ predicts that players $1, ..., n$ could choose $a_1, ..., a_n$ as final decisions in their *ex ante* decision makings. Of course, $a_i$ itself is $i$'s own possible decision. These predicates $D_1, ..., D_n$ will be determined by the specific axioms which will be given in Section 3. By the expression $K_i(A)$, we mean that player $i$ knows that a formula $A$ is provable from his basic knowledge.

First, we develop the space of formulae. For any strategies $a_j, b_j$ in $\Sigma_j$ ($a_j$ and $b_j$ may be identical and $j = 1, ..., n$), $(a_j = b_j)$ is an *atomic formula*, and for strategy profiles $(a_1, ..., a_n), (b_1, ..., b_n)$ in $\Sigma$, the expressions $R_i(a_1, ..., a_n; b_1, ..., b_n)$ and $D_i(a_1, ..., a_n)$ $(i = 1, ..., n)$ are also *atomic formulae.* These atomic formulae correspond to propositional variables in the standard formulation of propositional logic. Since the number of strategies is finite, so is the number of atomic formulae.

Let $\mathcal{P}^0$ be the set of all formulae generated by the standard finitary inductive definition with respect to $\neg, \supset$ and $K_1, ..., K_n$ from the atomic formulae. That is, $\mathcal{P}^0$ is the set of 0-*formulae* defined by the following induction:

(**0-i**): any atomic formula is a 0-formula;

(**0-ii**): if $A$ and $B$ are 0-formulae, so are $(\neg A), (A \supset B)$ and $K_i(A)$.[4]

Suppose that $\mathcal{P}^k$ is already defined $(k = 0, 1, ...)$. Then $\mathcal{P}^{k+1}$ is the set of $(k+1)$-*formulae* defined by the following induction:

($(\mathbf{k+1})$**-i**): any formula in $\mathcal{P}^k \cup \{(\bigwedge \Phi), (\bigvee \Phi) : \Phi$ is a nonempty countable subset of $\mathcal{P}^k\}$ is a $(k+1)$-formula;

($(\mathbf{k+1})$**-ii**): if $A$ and $B$ are $(k+1)$-formulae, so are $(\neg A), (A \supset B)$ and $K_i(A)$.

We denote $\bigcup_{k<\omega} \mathcal{P}^k$ by $\mathcal{P}^\omega$.[5] An expression in $\mathcal{P}^\omega$ is called a *formula.* We abbreviate $\bigwedge\{A, B\}$ and $\bigvee\{A, B\}$ as $A \bigwedge B$ and $A \bigvee B$, and $(A \supset B) \bigwedge (B \supset A)$ as $A \equiv B$. We also

[4] and all 0-formuale are obtained by a finite applications of theose steps. We will not add this qualification in the following inducitve definitions.

[5] Note that $\bigwedge \Phi$ and $\bigvee \Phi$ may not be in $\mathcal{P}^\omega$ for some countable subsets $\Phi$ of $\mathcal{P}^\omega$. However, the space $\mathcal{P}^\omega$ is large enough to discuss common knowledge.

The space $\mathcal{P}^\omega$ is a relatively small fragment of the space of infinitary formulae corresponding to that in Karp [16]. Nevertheless, $\mathcal{P}^\omega$ is already uncountable. Some smaller, countable, space of formulae suffices for our purpose. For example, a countable and constructive space of formulae is provided in Kaneko-Nagashima [12]. We adopt the above space for presentational simplicity.

abbreviate some parentheses in the standard manner. Also, we call $\Phi$ an *allowable set* iff $\Phi$ is a nonempty countable subset of $\mathcal{P}^k$ for some $k < \omega$.

The primary reason for our infinitary language is to express common knowledge explicitly as a conjunctive formula. The common knowledge of a formula $A$ is defined as follows: For any $m \geq 0$, we denote the set $\{K_{i_1}K_{i_2}...K_{i_m}$ : each $K_{i_t}$ is one of $K_1, ..., K_n$ and $i_t \neq i_{t+1}$ for $t = 1, ..., m-1\}$ by $\mathsf{K}(m)$. We assume that $\mathsf{K}(0)$ consists of the null symbol $e$ (i.e., $e(A)$ is $A$ itself for any $A$). We define the *common knowledge formula* of $A$ as

$$\bigwedge\{K(A) : K \in \bigcup_{m<\omega} \mathsf{K}(m)\}, \tag{2.2}$$

which we denote by $C(A)$. If $A$ is in $\mathcal{P}^k$, the set $\{K(A) : K \in \bigcup_{m<\omega} \mathsf{K}(m)\}$ is a countable subset of $\mathcal{P}^k$, and its conjunction, $C(A)$, belongs to $\mathcal{P}^{k+1}$ by (($k+1$)-i). Hence the space $\mathcal{P}^\omega$ is closed with respect to the operation $C(\cdot)$.

Note that $A$ itself is included as a conjunct in $C(A)$, since $\mathsf{K}(0) = \{e\}$. In this sense, according to the literature of epistemic logic, $C(A)$ is "common knowledge" instead of "common belief" which is defined to be the conjunction obtained from (2.2) by excluding $A$.

*Base logic* $\mathrm{GL}_0$ is defined by the following five axiom schemata and three inference rules: for any formulae $A, B, C$ and allowable set $\Phi$,

(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 $A \in \Phi$;

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

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

These axioms and inference rules determine base logic $\mathrm{GL}_0$.

We define game logic $\mathrm{GL}_\omega$ by adding the following axiom schemata and inference rule to $\mathrm{GL}_0$: for any formulae $A, B$ and $i = 1, ..., n$;

($\mathrm{MP}_i$): $K_i(A \supset B) \bigwedge K_i(A) \supset K_i(B)$;

($\bot_i$): $\neg K_i(\neg A \bigwedge A)$;

($\mathrm{PI}_i$): $K_i(A) \supset K_iK_i(A)$;

(C-Barcan): $\bigwedge\{K_iK(A) : K \in \bigcup_{m<\omega} \mathsf{K}(m)\} \supset K_iC(A)$;[6]

and

(Necessitation): $\dfrac{A}{K_i(A)}$.

We will abbreviate Necessitation as Nec, and use $\mathrm{MP}_i$, $\perp_i$, $\mathrm{PI}_i$ as generic names for those for different $i$ .

A *proof* in $\mathrm{GL}_\omega$ is a countable tree with the following properties: (i) every path from the root is finite; (ii) a formula is associated with each node, and the formula associated with each leaf is an instance of the axioms; and (iii) adjoining nodes together with their associated formulae form an instance of the above inferences. We write $\vdash_\omega A$ iff there is a proof $P$ such that $A$ is associated with the root. For any subset $\Gamma$ of $\mathcal{P}^\omega$, we write $\Gamma \vdash_\omega A$ iff $\vdash_\omega \bigwedge\Phi \supset A$ for some nonempty finite subset $\Phi$ of $\Gamma$. When $\Gamma$ is empty, $\Gamma \vdash_\omega A$ is assumed to be $\vdash_\omega A$ itself. We also abbreviate $\Gamma \cup \Theta \vdash_\omega A, \Gamma \cup \{B\} \vdash_\omega A$ as $\Gamma, \Theta \vdash_\omega A, \Gamma, B \vdash_\omega A$, etc.[7]

We will use the following facts without references (see Kaneko-Nagashima [14]).

**Lemma 2.1.** Let $\Gamma, \Theta$ be sets of formulae, and $\Phi$ an allowable set of formulae. Then

(1): If $\Gamma \vdash_\omega A \supset B$ and $\Theta \vdash_\omega B \supset C$, then $\Gamma, \Theta \vdash_\omega A \supset C$;

(2): $\vdash_\omega \bigwedge\Phi$ if and only if $\vdash_\omega A$ for all $A \in \Phi$;

(3): $\vdash_\omega K_i(\bigwedge\Phi) \supset \bigwedge K_i(\Phi)$, and if $\Phi$ is a finite set, then $\vdash_\omega K_i(\bigwedge\Phi) \equiv \bigwedge K_i(\Phi)$, where $K_i(\Phi)$ is the set $\{K_i(A) : A \in \Phi\}$;

(4): $\vdash_\omega \bigvee K_i(\Phi) \supset K_i(\bigvee\Phi)$.

Axiom $\mathrm{MP}_i$ and inference rule Nec in addition to $\mathrm{GL}_0$ give the complete logical ability to each player (see [14]). Axiom $\perp_i$ requires the knowledge of each player to be consistent. This is considerably weaker than the *Veridicality Axiom*: $K_i(A) \supset A$. In the literature of epistemic logic, $K_i$ is sometimes called the *belief* operator in the logic without the veridicality; but we use the terminology "knowledge" since the distinction does not play an important role in this paper. Axiom $\mathrm{PI}_i$, called the *Positive Introspection*, means that if player $i$ knows $A$, he knows that he knows $A$. In fact, these logical

[6] Kaneko [10] imposes a further restriction on C-Barcan that (i) each Barcan formula $A$ contains no infinitary disjunctions and (ii) if it contains an infinitary conjuction, then it is $C(B)$ for some $B$. Then he proved that common knowledge logic, based on KD4-axioms), considered in Halpern-Moses [5] and Lismont-Mongin [18] into game logic $\mathrm{GL}_\omega$ and that $\mathrm{GL}_\omega$ is a conservative extension of the fragment obtained by this embedding. The restriction does not prevent any argument in this paper (here we did not assume it because we do not use it). Hence all the results in this paper can be converted to common knowledge logic.

[7] Since $\mathrm{GL}_\omega$ has Nec, nonlogical axioms should be introduced in this manner, instead of being initial formulae in a proof. For the treatment of nonlogical axioms in a logic with Nec, see Kaneko-Nagashima [13].

and introspective abilities of the players are common knowledge in $GL_\omega$ (see [14]). Since the finitary propositional modal logic defined by $MP_i, \perp_i$, $PI_i$ and Nec in addition to classical propositional logic is called KD4, our logic is an infinitary extension of KD4 with C-Barcan.[8]

Axiom C-Barcan is called the *common knowledge Barcan axiom.* For the development of our framework, C-Barcan will be used to derive the property:

$$\vdash_\omega C(A) \supset K_iC(A) \quad \text{for } i = 1, ..., n. \tag{2.3}$$

That is, if $A$ is common knowledge, then each player $i$ knows that it is common knowledge. This property will play an important role in the epistemic axiomatization of final decisions in later sections. It is proved as follows:

**Proof of (2.3)*:** Let $K$ be an arbitrary element in $\bigcup_{m<\omega} \mathsf{K}(m)$. When $K_i$ is not the outermost symbol of $K$, we have $\vdash_\omega C(A) \supset K_iK(A)$ by L4. When $K_i$ is the outermost symbol of $K$, we have $\vdash_\omega K(A) \supset K_iK(A)$ by $PI_i$. This together with $\vdash_\omega C(A) \supset K(A)$ implies $\vdash_\omega C(A) \supset K_iK(A)$. Thus $\vdash_\omega C(A) \supset K_iK(A)$ for all $K \in \bigcup_{m<\omega} \mathsf{K}(m)$. Hence $\vdash_\omega C(A) \supset \bigwedge\{K_iK(A) : K \in \bigcup_{m<\omega} \mathsf{K}(m)\}$ by $\bigwedge$-Rule. By C-Barcan, we obtain $\vdash_\omega C(A) \supset K_iC(A)$. □

In fact, it is proved in Kaneko [10] that (2.3) is not provable without C-Barcan.

When we restrict the use of axioms and inference rules to those of $GL_0$, we denote the provability relation by $\vdash_0$ . Logic $GL_0$ is an infinitary extension of classical finitary propositional logic; and also, it can be proved that it is sound and complete with respect to the standard two-valued semantics with the straightforward modifications of validity for infinitary conjunctions and disjunctions. We briefly review those results to make subsequent arguments simpler.

An *assignment* $\tau$ is a function to $\{\top, \perp\}$ from the set of all formulae which are atomic or are expressed as $K_i(B)$ for some $B$ and $i$. We first define the truth relation $\models_\tau$ relative to an assignment $\tau$ by the following induction on the structure of a formula in $\mathcal{P}^\omega$ :

(T0): for any formula $A$ which is atomic or is expressed as $K_i(B)$ for some $B$ and $i$, $\models_\tau A$ if and only if $\tau(A) = \top$;

(T1): $\models_\tau \neg A$ if and only if not $\models_\tau A$;

(T2): $\models_\tau A \supset B$ if and only if not $\models_\tau A$ or $\models_\tau B$;

(T3): $\models_\tau \bigwedge \Phi$ if and only if $\models_\tau A$ for all $A \in \Phi$;

(T4): $\models_\tau \bigvee \Phi$ if and only if $\models_\tau A$ for some $A \in \Phi$.

[8]The logic obtained from KD4 by the replacement of $\perp_i$ by Veridicality is called S4. In S4, the truth of knowledge can be defined relative to the investigator, while in KD4, knowledge is required only to be consistent within each player. In this sense, KD4 is a logic of cognitive relativism.

Then the following soundness-completeness holds.

**Theorem 2.1.***(Completeness for $GL_0$): For any formula $A$, $\vdash_0 A$ if and only if $\models_\tau A$ for any assignment $\tau$. [9]

When $\models_\tau A$ for any $\tau$, $A$ is called a *tautology* in $GL_0$. Then we can use tautologies as initial formulae in a proof in $GL_\omega$ instead of instances of L1–L5. This fact may be used in the subsequent analysis.

For subsequent uses, we mention the standard soundness-completeness theorem. We say that a formula $A$ is *finitary* iff it contains no infinitary conjunction and no disjunctions, and that $A$ is *nonepistemic* iff it contains no $K_i$, $i = 1, ..., n$. We denote the set of all nonepistemic finitary formulae by $\mathcal{P}^f$. The space $\mathcal{P}^f$ is closed with respect to $\neg, \supset$ and finitary $\bigwedge, \bigvee$. When we restrict the language of base logic $GL_0$ to $\mathcal{P}^f$, the resulting logic is classical propositional logic, which we denote $GL_0^f$. Base logic $GL_0$ is a conservative extension of $GL_0^f$, i.e., for any $A \in \mathcal{P}^f$,

$$\text{if } \vdash_0 A, \text{ then } A \text{ is provable in } GL_0^f, \tag{2.4}$$

and, of course, the converse holds. Therefore soundness-completeness is modified as follows: for any $A \in \mathcal{P}^f$,

$$\vdash_0 A \text{ if and only if } \models_\tau A \text{ for any assignment } \tau \text{ over the atomic formuale.} \tag{2.5}$$

The equivalent soundness will be used: if there is an assignment $\tau$ over the atomic formulae such that $\models_\tau A$, then $A$ is consistent with respect to $\vdash_0$. We will refer to this as Soundness for $GL_0^f$.

Let $A$ be a formula in $\mathcal{P}^\omega$. Then $\varepsilon A$ is the formula obtained from $A$ by eliminating all the occurrences of $K_1, ..., K_n$ in $A$, and $\varepsilon\Gamma$ is the set $\{\varepsilon A : A \in \Gamma\}$ for any set $\Gamma$ of formulae. Formula $\varepsilon A$ is nonepistemic, and if $A$ is nonepistemic, $\varepsilon A$ is $A$ itself. For example, $\varepsilon(K_i(A) \bigwedge K_i(A \supset B) \supset K_i(B))$ is $\varepsilon A \bigwedge(\varepsilon A \supset \varepsilon B) \supset \varepsilon B$, and $\varepsilon(\neg K_i(A \bigwedge \neg A))$ is $\neg(\varepsilon A \bigwedge \neg\varepsilon A)$. For every instance $A$ of the epistemic axioms, $\varepsilon A$ is a nonepistemic provable formula with respect to $\vdash_0$ and every instance of Nec becomes a trivial inference with an application of $\varepsilon$. This is the reason for Lemma 2.2.(1). We have also the other lemmas (see Kaneko-Nagashima [14]).

**Lemma 2.2.(1)***: If $\Gamma \vdash_\omega A$, then $\varepsilon\Gamma \vdash_0 \varepsilon A$.

**(2)***: If $\Gamma \vdash_0 A$ or $\Gamma \vdash_\omega A$, then $C(\Gamma) \vdash_\omega C(A)$, where $C(\Gamma) = \{C(B) : B \in \Gamma\}$.

**Lemma 2.3.(1)***: $\vdash_\omega C(A \supset B) \supset (C(A) \supset C(B))$;

**(2)***: if $C(\Gamma) \vdash_\omega A$, then $C(\Gamma) \vdash_\omega K_i(A)$ for $i = 1, ..., n$.

[9]This can be proved with a slight modification of the standard proof of the completeness theorem for classical propositional logic (cf., the previous version (1991) of Kaneko-Nagashima [14]).

**Lemma 2.4.** Let $\Phi$ be an allowable set of formulae. Then
(1)*: $\vdash_\omega C(\bigwedge \Phi) \supset \bigwedge C(\Phi)$, and if $\Phi$ is a finite set, then $\vdash_\omega C(\bigwedge \Phi) \equiv \bigwedge C(\Phi)$;
(2)*: $\vdash_\omega \bigvee C(\Phi) \supset C(\bigvee \Phi)$.

### 2.3. Game Theoretical Concepts in the Formal Language $\mathcal{P}^\omega$

Now we describe, in the formalized language $\mathcal{P}^\omega$, the game theoretical concepts given in Subsection 2.1.

First, we make the following axiom: for all distinct $a_i, b_i \in \Sigma_i$ $(i = 1, ..., n)$,

**Axiom (Eq):** $a_i = a_i$ and $\neg(a_i = b_i)$.

We denote the set of instances of this axiom by Eq, which is a finite set of formulae. In the following, we identify the set of instances of an axiom with the axiom itself.

Second, we describe the payoff functions $g_1, ..., g_n$ in terms of predicates $R_1, ..., R_n$ as follows: for strategy profiles $a, b, a', b'$ with $g_i(a) \geq g_i(b)$ and $g_i(a') < g_i(b')$ and $i = 1, ..., n$,

**Axiom (G):** $R_i(a; b)$ and $\neg R_i(a'; b')$.

This axiom describes the payoff functions $g_1, ..., g_n$ as preferences $R_1, ..., R_n$. It holds that for any $a, b \in \Sigma$, either G $\vdash_0 R_i(a; b)$ or G $\vdash_0 \neg R_i(a; b)$.

In the game of Table 2.1, Eq is the set $\{\mathrm{s}_{it} = \mathrm{s}_{it} : i = 1,2 \text{ and } t = 1,2\} \cup \{\neg \mathrm{s}_{it} = \mathrm{s}_{it'} : i = 1,2 \text{ and } t, t' = 1,2 \text{ with } t \neq t'\}$, and G is the union of the preferences of players 1 and 2. Note that G contains both positive and negative preferences.

We define the Nash *equilibrium property* to be the formula:

$$\bigwedge_{i=1}^{n} \bigwedge_{y_i \in \Sigma_i} R_i(a : y_i; a_{-i}), \tag{2.6}$$

which we denote by Nash($a$). Since either G $\vdash_0 R_i(a; b)$ or G $\vdash_0 \neg R_i(a; b)$ for any $a, b \in \Sigma$, Nash($a$) is also decidable for any $a \in \Sigma$ under Axiom G, that is,

$$\text{either G } \vdash_0 \text{Nash}(a) \text{ or G } \vdash_0 \neg\text{Nash}(a). \tag{2.7}$$

Also, it holds that

$$\text{G } \vdash_0 \text{Nash}(a) \text{ if and only if } a \in E_g. \tag{2.8}$$

The right-hand side of (2.8) is a statement in the nonformalized language. Thus (2.8) relates the statement "$a$ is a Nash equilibrium" in the formalized language to the nonformalized *extensional* counterpart. The right-hand side can be described in our formalized language as follows:

$$\bigvee_{y \in E_g} \left( \bigwedge_{i=1}^{n} a_i = y_i \right). \tag{2.9}$$

This is the extensional definition of Nash equilibrium in our language. This formula is denoted by $\text{Nash}^E(a)$. Of course, $\text{Nash}(a)$ is the basic definition of Nash equilibrium, and $\text{Nash}^E(a)$ is a representation of it. Under Eq, this is also decidable for any $a \in \Sigma$, i.e., Eq $\vdash_0 \text{Nash}^E(a)$ or Eq $\vdash_0 \neg\text{Nash}^E(a)$. Without Axiom G, however, $\text{Nash}^E(a)$ does not have the meaning "Nash equilibrium".

If we try to give an intensional definition of a subsolution, it would involve the second-order consideration, which is not allowed in our language.[10] Hence we give the extensional definition of a subsolution in our formal language. We denote the following formula by $\text{Sol}^k(a)$:

$$\bigvee_{y \in E_g^k} \left( \bigwedge_{i=1}^{n} a_i = y_i \right) \tag{2.10}$$

for $k = 1, ..., \sigma$. The formula $\text{Sol}^k(a)$ describes "$a$ belongs to the subsolution $E_g^k$".

**Lemma 2.5.(1):** Eq,G $\vdash_0 \bigwedge_{x \in \Sigma} \left( \text{Nash}(x) \equiv \text{Nash}^E(x) \right)$;
**(2):** Eq,G $\vdash_0 \bigwedge_x \left( \text{Nash}(x) \equiv \bigvee_{k=1}^{\sigma} \text{Sol}^k(x) \right)$.

**Proof.** We prove (1). Let $a$ be an arbitrary profile. Suppose G $\vdash_0 \text{Nash}(a)$. Then $a \in E_g$ by (2.8). Thus Eq $\vdash_0 \bigvee_{y \in E_g} \left( \bigwedge_i a_i = y_i \right)$, i.e., Eq $\vdash_0 \text{Nash}^E(a)$. Since $\text{Nash}(a)$ is decidable under G by (2.7), we have proved Eq,G $\vdash_0 \text{Nash}(a) \supset \text{Nash}^E(a)$. Noting that $\text{Nash}^E(a)$ is decidable under Eq, we can repeat a similar argument to have Eq,G $\vdash_0 \text{Nash}^E(a) \supset \text{Nash}(a)$. Thus Eq,G $\vdash_0 \text{Nash}^E(a) \equiv \text{Nash}(a)$. Since $a$ is an arbitrary profile, we have Eq,G $\vdash_0 \bigwedge_x \left( \text{Nash}(x) \equiv \text{Nash}^E(x) \right)$. □

In Section 6, we need to assume that these equivalences are common knowledge among the players. If we assume that Axioms Eq and G are common knowledge, these equivalences are common knowledge, particularly, (2) becomes

$$C(\text{Eq}), C(\text{G}) \vdash_\omega \bigwedge_x C \left( \text{Nash}(x) \equiv \bigvee_{k=1}^{\sigma} \text{Sol}^k(x) \right). \tag{2.11}$$

This follows from Lemmas 2.5, 2.2.(2), 2.3.(1), and 2.4.(1). Note, however, that the common knowledge of the equivalence of $\text{Nash}^E(a)$ and $\bigvee_{k=1}^{\sigma} \text{Sol}^k(a)$ is provable without these axioms, i.e., $\vdash_\omega \bigwedge_x C \left( \text{Nash}^E(x) \equiv \bigvee_{k=1}^{\sigma} \text{Sol}^k(x) \right)$, since this is simply extensional equivalence.

In Section 4, we argue that if Eq and G are not common knowledge, the common knowledge of the equivalence of (2.11) is not provable.

[10] A logic including variables representing sets. Kaneko-Nagashima [12] formalized a subsolution in the first-order language, but it was not satisfactory.

Note that Axioms Eq and G are consistent with respect to $\vdash_0$, i.e., there is no formula $A$ such that Eq,G $\vdash_0 \neg A \wedge A$. This can be proved by constructing a model for them by the Soundness Theorem for $\mathrm{GL}_0^f$. This consistency implies that $C(\mathrm{Eq})$ and $C(\mathrm{G})$ are also consistent with respect to $\vdash_\omega$ by Lemma 2.2.(1). We will use this consistency in Sections 5 and 6.

## 3. Final Decision Axioms and Johansen's Argument

### 3.1. Final Decision Axioms

In a given game $g$, each player deliberates his and the others' strategy choices and may reach some predictions on their final decisions. Now we describe these "predictions" by $n$-ary predicates $D_1, ..., D_n$, that is, each $D_i(a_1, ..., a_n)$ means that player $i$ predicts that players $1, ..., n$ could choose $(a_1, ..., a_n)$ as their final decisions, where $a_i$ is his own decision. Note that each player's prediction may not be uniquely determined.

The following are base axioms for $D_1(\cdot), ..., D_n(\cdot)$: for each $i = 1, ..., n$,

**Axiom** $\mathbf{D1}_i^0$ : $\bigwedge_x \left(D_i(x) \supset \bigwedge_{y_i} R_i(x_i; x_{-i} : y_i; x_{-i})\right)$;

**Axiom** $\mathbf{D2}_i^0$ : $\bigwedge_x \bigwedge_j (D_i(x) \supset D_j(x))$;

**Axiom** $\mathbf{D3}_i^0$ : $\bigwedge_x (D_i(x) \supset K_i(D_i(x)))$;

**Axiom** $\mathbf{D4}_i^0$ : $\bigwedge_x \bigwedge_y \bigwedge_j (D_i(x) \wedge D_i(y) \supset D_i(x_j; y_{-j}))$.

These are verbally as follows:

$\mathrm{D1}_i^0$: (*Best Response to Predicted Decisions*): If player $i$ predicts final decisions $x_1, ..., x_n$ for the players, then his own final decision $x_i$ maximizes his payoff against his prediction $x_{-i}$, that is, $x_i$ is a best response to $x_{-i}$.

$\mathrm{D2}_i^0$: (*Identical Predictions*): The other players reach the same predictions as player $i$'s.

$\mathrm{D3}_i^0$: (*Knowledge of Predictions*): player $i$ knows his own predictions.

$\mathrm{D4}_i^0$: (*Interchangeability*): Player $i$'s predictions are interchangeable.

In fact, $\mathrm{D2}_i^0$ may be unnecessary for each player's decision making, depending upon a game. In some games such as Prisoner's Dilemma having dominant strategies, each player does not need to predict all the others' decisions for him to make a decision. In this case, he does not infer the others' decisions. However, of course, it is possible to assume that each player predicts all the others' to make a decision. This is a basic assumption for our axiomatization, and we focus on the above system of axioms exclusively in this paper. A "weaker" system of axioms is considered in Kaneko [11], which is remarked in Subsection 7.3.

These axioms are requirements for decision (prediction) making in the mind of each player $i = 1, ..., n$. Hence we assume that these axioms hold and are known to player $i$. These are described by

$$K_i^+(\mathrm{D1}_i^0) ,..., K_i^+(\mathrm{D4}_i^0),$$

where $K_i^+(A)$ denotes $A \bigwedge K_i(A)$. We denote these formulae by $\mathrm{D1}_i$ ,..., $\mathrm{D4}_i$, respectively. Now, the problem is whether these axioms determine "unknown" predicate $D_i(\cdot)$ in terms of the "primitives", $\mathbf{s}_{11}, ..., \mathbf{s}_{1\ell_1}; ...; \mathbf{s}_{n1}, ..., \mathbf{s}_{n\ell_n}, R_1, ..., R_n$ and $K_1, ..., K_n$. In the following, we denote $\mathrm{D1}_i^0 \bigwedge ... \bigwedge \mathrm{D4}_i^0$ and $\mathrm{D1}_i \bigwedge ... \bigwedge \mathrm{D4}_i$ by $\mathrm{D}_i^0$(1–4) and $\mathrm{D}_i$(1–4).

Notice that $D_j(\cdot)$'s occur in Axiom $\mathrm{D2}_i$ for all $j$. These $D_j(\cdot)$'s are determined by other $\mathrm{D}_j$(1–4)'s. If $\mathrm{D}_j$(1–4)'s are not known to player $i$, then $D_j(\cdot)$'s would be just symbols without meaning. Hence $\mathrm{D2}_i$ do not make sense for him without knowing $\mathrm{D}_j$(1–4)'s. In fact, the axiom, $\mathrm{D}_i$(1–4), of player $i$ does not imply $\mathrm{D}_j$(1–4).

**Proposition 3.1.** Let $i, j$ be distinct players. Then

(1): neither $\mathrm{D}_i(1\text{–}4) \vdash_\omega \mathrm{D}_j^0(1\text{–}4)$ nor $\mathrm{D}_i(1\text{–}4) \vdash_\omega \neg\mathrm{D}_j^0(1\text{–}4)$;

(2): neither $\mathrm{D}_i(1\text{–}4) \vdash_\omega \mathrm{D}_j(1\text{–}4)$ nor $\mathrm{D}_i(1\text{–}4) \vdash_\omega \neg\mathrm{D}_j(1\text{–}4)$;

(3): neither $\mathrm{D}_i(1\text{–}4) \vdash_\omega K_i(\mathrm{D}_j(1\text{–}4))$ nor $\mathrm{D}_i(1\text{–}4) \vdash_\omega \neg K_i(\mathrm{D}_j(1\text{–}4))$.

**Proof.** We prove only the first assertion of (3). Suppose $\mathrm{D}_i(1\text{–}4) \vdash_\omega K_i(\mathrm{D}_j(1\text{–}4))$, i.e., $\vdash_\omega \mathrm{D}_i(1\text{–}4) \supset K_i(\mathrm{D}_j(1\text{–}4))$. By Lemma 2.2.(1), $\vdash_0 \varepsilon\mathrm{D}_i(1\text{–}4) \supset \varepsilon K_i(\mathrm{D}_j(1\text{–}4))$, which implies $\vdash_0 \mathrm{D}_i^0(1,2,4) \supset \mathrm{D}_j^0(1,2,4)$, since $\varepsilon\mathrm{D3}_k$ is equivalent to $\bigwedge_x (D_k(x) \supset D_k(x))$. Hence $\models_\tau \mathrm{D}_i^0(1,2,4) \supset \mathrm{D}_j^0(1,2,4)$ for any assignment $\tau$ over the atomic formulae by Soundness for $\mathrm{GL}_0^f$. However, we can construct an assignment $\tau_0$ over the set of atomic formulae such that $\models_{\tau_0} \mathrm{D}_i^0(1,2,4)$ but not $\models_{\tau_0} \mathrm{D}_j^0(1,2,4)$, a contradiction. Therefore it is not the case that $\mathrm{D}_i(1\text{–}4) \vdash_\omega K_i(\mathrm{D}_j(1\text{–}4))$. □

We denote $\bigwedge_j \mathrm{D}_j$(1–4) by D(1–4). Proposition 3.1 implies that for $D_j(\cdot)$'s to be meaningful in Axiom $\mathrm{D2}_i$, we should assume $K_i$(D(1–4)) as well as D(1–4). In fact, these are still insufficient to determine the meanings of all $\mathrm{D}_j(\cdot)$'s in $K_i$(D(1–4)), which will be discussed in Section 4. Here we state only the following undecidability:

$$\text{neither } \mathrm{D}(1\text{–}4) \vdash_\omega K_i(\mathrm{D}(1\text{–}4)) \text{ nor } \mathrm{D}(1\text{–}4) \vdash_\omega \neg K_i(\mathrm{D}(1\text{–}4)). \tag{3.1}$$

That is, the knowledge of D(1–4) is not derived from D(1–4). In Section 4, we will give a general version of this claim, which leads to the infinite regress of the knowledge of D(1–4), i.e., the common knowledge of D(1–4).

Before going to the next subsection, we give some simple observations on the above axioms. In the following, we denote $\bigwedge_j \mathrm{Dt}_j^0$ and $\bigwedge_j \mathrm{Dt}_j$ by $\mathrm{Dt}^0$ and Dt for t = 1,...,4, and $\mathrm{D1}^0 \bigwedge \mathrm{D2}^0$ and $\mathrm{D1} \bigwedge \mathrm{D2} \bigwedge \mathrm{D4}$ by $\mathrm{D}^0$(1,2) and D(1,2,4), etc.

**Lemma 3.2.** $D^0(1,2) \vdash_0 \bigwedge_x (D_i(x) \supset \text{Nash}(x))$.

**Proof.** Let $a$ be an arbitrary profile. Since $D2^0 \vdash_0 D_i(a) \supset D_j(a)$ and $D1^0 \vdash_0 D_j(a) \supset \bigwedge_{y_j} R_j(a_j; a_{-j} : y_j; a_{-j})$ for all $i, j$, we have $D^0(1,2) \vdash_0 D_i(a) \supset \bigwedge_{y_j} R_j(a_j; a_{-j} : y_j; a_{-j})$ for all $j$. Thus $D^0(1,2) \vdash_0 D_i(a) \supset \bigwedge_j \bigwedge_{y_j} R_j(a_j; a_{-j} : y_j; a_{-j})$ by $\bigwedge$-Rule, i.e., $D^0(1,2) \vdash_0 D_i(a) \supset \text{Nash}(a)$. Hence we have the assertion. □

Thus, Nash equilibrium is a necessary condition of Axioms $D^0(1,2)$. In fact, we will show in our full axiomatization that $D_i(a)$ implies $C(\text{Nash}(a))$, i.e., it is the common knowledge that $a$ is a Nash equilibrium. The following lemma is indicative of this common knowledge result.

**Lemma 3.3.** $D(2,3) \vdash_\omega \bigwedge_x (D_j(x) \supset K_i(D_k(x)))$ for any $i, j, k$ ($i, j, k$ may be identical).

**Proof.** Let $a$ be an arbitrary profile. Since $D2 \vdash_\omega K_i(D_i(a) \supset D_k(a))$, we have $D2 \vdash_\omega K_i(D_i(a)) \supset K_i(D_k(a))$ by $MP_i$ and MP. Since $D2 \vdash_\omega D_j(a) \supset D_i(a)$ and $D3 \vdash_\omega D_i(a) \supset K_i(D_i(a))$, we have $D(2,3) \vdash_\omega D_j(a) \supset K_i(D_i(a))$. Hence we have $D(2,3) \vdash_\omega D_j(a) \supset K_i(D_k(a))$. This implies the assertion. □

We prepare one more lemma, which will be used in the later sections.

**Lemma 3.4.(1):** $\varepsilon D(1\text{–}4)$ is consistent with respect to $\vdash_0$ ($\varepsilon D(1\text{–}4), D_i(a)$ is also consistent).

**(2):** $\vdash_\omega D(1\text{–}4)$ does not hold.

**Proof.(1):** $\varepsilon D(1\text{–}4)$ is equivalent to $D^0(1,2,4)$ with respect to $\vdash_0$ . To prove that $D^0(1,2,4)$ is consistent, it suffices to show one assignment $\tau$ on the atomic formulae so that $\models_\tau D^0(1,2,4)$, by Soundness for $GL_0^f$. Define an assignment $\tau$ so that $\tau(D_i(a)) = \bot$ for all profiles $a$ and $i = 1, ..., n$. Then $D1^0$,$D2^0$ and $D4^0$ are true in $\tau$ in the trivial sense. Hence $D^0(1,2,4)$ is consistent with respect to $\vdash_0$.

(2): Suppose $\vdash_\omega D(1\text{–}4)$. Then $\vdash_0 \varepsilon D(1\text{–}4)$, i.e., $\vdash_0 D^0(1,2,4)$, by Lemma 2.2.(1). However, we can construct an assignment $\tau$ in a similar manner as in (1) so that $D^0(1,2,4)$ is false. Hence $\vdash_0 D^0(1,2,4)$ is not the case. Thus $\vdash_\omega D^0(1\text{–}4)$ is not the case, too. □

### 3.2. Comparisons with Johansen's Argument

Now we look at Johansen's [8] argument on Nash equilibrium from the viewpoint of the above axiomatization. Johansen proposed four postulates for *rational* (*individualistic*) *decision making* in a game. We reproduce them with slight modifications in terminology:

(J1): A player makes his decision $a_i \in \Sigma_i$ on the basis of, and only on the basis of information concerning the strategy sets $\Sigma_1, ..., \Sigma_n$, and preferences $R_1, ..., R_n$.

**(J2)**: In choosing his own decision, a player assumes that the other players are rational in the same way that he himself is rational.

**(J3)**: If some decision is the rational decision to make for an individual player, then this decision can be correctly predicted by the other players.

**(J4)**: Being able to predict the actions to be taken by the other players, a player's own decision maximizes his preference relation corresponding to the predicted actions of the other players.

Johansen [8] argues that if a game possesses a unique Nash equilibrium, then each player's Nash strategy is the unique choice satisfying these four postulates. He claims also that the uniqueness of a Nash equilibrium is for simplicity and that interchangeability suffices for his argument ([8], p.424).

To make comparisons between Johansen's postulates and our approach, we should be conscious about one important difference. Johansen's postulates describe the whole situation he is considering. On the other hand, our axioms constitute a part of the description of the game situation – the description of logical abilities is made separately in $GL_\omega$. Therefore we should consider the correspondence between his postulates and our entire approach, and then compare our results with Johansen's claim.

Postulate J1 can be regarded as fulfilled by our approach in that the primitives are available strategies $\Sigma_1, ..., \Sigma_n$ and preferences $R_1, ..., R_n$ but the players' knowledge described by $K_1, ..., K_n$ are implicit in Johansen's postulates.

Postulate J2 may be interpreted in our axiomatization as meaning that each player $i$ follows the final decision axioms assuming that the other players assume these axioms, too. This interpretation may be formulated as assuming D(1–4) $\bigwedge K_i$(D(1–4)) for $i = 1, ..., n$. At the more fundamental level, it is interpreted also as including the assumption that every player has a (complete) logical ability as well as knows (or assume) that the other players have the same logical abilities. In the latter sense, J2 partially corresponds to our $GL_\omega$ (which may be an overinterpretation of Johansen).

Postulate J3 corresponds to the assertion of Lemma 3.3 rather than Axiom D3 itself. This postulate requires that each player be conscious about his and others' "rational decision making". These four postulates require each player to know the postulates themselves for his "rational decision making". In this sense, Johansen's postulates look self-referential. Bernheim [4], p.486 criticized Johansen's postulates as ambiguous and self-referential. The same difficulty is involved in our axiomatization: we need to assume that each player $i$ knows D(1–4), and this leads to an infinite regress of the knowledge of D(1–4), i.e., the common knowledge of it. In our case, the ambiguity is avoided by formulating the whole situation in game logic, and the self-referential becomes the infinite regress, which will be solved later.

Postulate J4 corresponds to Axiom D1. This has often been used, but we should be critical of its use. It is imposed since, in Johansen's argument as well as in ours, we consider predicted final decisions after sufficient deliberation before the actual play. We do not assume it in intermediate steps in decision making. This is a totally different justification of the best response property than the one by the "perfect" competition in economics.[11]

We claim that our entire approach is a materialization of Johansen's [8] argument. This claim will be clearer in Sections 4–5, where we will show that the knowledge of our axioms lead to an infinite regress and show that the infinite regress determines the common knowledge of Nash equilibrium as its solution.

The last comment is on the nonstochastic treatment of our notion of "predictions". Our formulation is a full description of a game situation, and nothing relevant for each player's decision making is hidden, which was stated above. Even if predictions are not uniquely determined, there is no reason to put more weights to some than to others. This does not mean that every prediction is assigned an equal probability. Instead, each prediction is logically possible. In this sense, we are treating "intrinsic" uncertainty.[12]

## 4. Infinite Regress of the Knowledge of the Axioms and its Evaluations

In this section, we show that the process of making Axioms D1 to D4 meaningful leads to an infinite regress of the knowledge of D(1–4), i.e., the common knowledge of D(1–4), and then make proof-theoretical evaluations of this infinite regress.

### 4.1. Infinite Regress of the Knowledge of the Final Decision Axioms

Axioms D1 – D4 seem to need each player to know these axioms. This necessity could be found by looking at Lemma 3.3 carefully. It requires that each player know any other players' predictions on final decisions, but this requirement could not be fulfilled unless the meaning of "predictions on final decisions" is given to the players. In fact, the meaning is determined by the above four axioms themselves. Therefore we should assume in addition to these axioms that each player knows them, i.e., $K_i(\mathrm{D}(1\text{–}4))$ for $i = 1, ..., n$. Recall that D(1–4) is $\mathrm{D1} \wedge ... \wedge \mathrm{D4}$.

Once the players are assumed to know these axioms, each knows the consequences

[11] In the definition of the rationalizability of Bernheim [3] and Pearce [21], the best response property is assumed in the intermediate steps in the hierarchy of "beliefs". There seems to be little reason to assume this best response property in such intermediate steps.

[12] Contary to our theory, the Savage-type subjective probablity theory is a black-box decision theory. The capability of making decision is the central postulate, which is based on something hidden in the black-box – such as past experiences. This differentiates the philosophical background of our approach from Savage's [22] as well as from the Bayesian game theory based on Savage.

from these axioms such as the assertion of Lemma 3.2, $\bigwedge_x K_i(D_i(x) \supset \text{Nash}(x)))$, i.e.,

$$\text{D(1–4)}, \bigwedge_i K_i(\text{D(1–4)}) \vdash_\omega \bigwedge_x K_i(D_i(x) \supset \text{Nash}(x))),$$

The following, however, is also provable: for any $j, k = 1, ..., n$,

$$\text{D(1–4)}, \bigwedge_i K_i(\text{D(1–4)}) \vdash_\omega \bigwedge_x (D_i(x) \supset K_j K_k(D_i(x))). \tag{4.1}$$

In (4.1), "imaginary" player $k$ in the mind of player $j$ knows that $x$ is a profile of predicted decisions. This imaginary player $k$ is not given the meaning, D(1–4), of "final decisions", though "real" $k$ is assumed to know D(1–4). This means that (4.1) does not make sense for the imaginary $k$. Thus we need to assume $K_j K_k(\text{D(1–4)})$, that is, we reach the assumption set:

$$\{L(\text{D(1–4)}): L \in \bigcup_{t \leq 2} \mathsf{K}(t)\}.$$

Once we assume this set of axioms, we would again meet the problem parallel to that arising in (4.1), that is, it holds that for any $K \in \mathsf{K}(3)$,

$$\{L(\text{D(1–4)}): L \in \bigcup_{t \leq 2} \mathsf{K}(t)\} \vdash_\omega \bigwedge_x (D_i(x) \supset K(D_i(x))).$$

That is, if we assume the knowledge of D(1–4) up depth 2, the knowledge of depth 3 is necessarily involved. The imaginary players in the epistemic world of depth 3 should know D(1–4).

In general, we have the following proposition.

**Proposition 4.1*.** For any players $i, j$, finite $m \geq 0$ and $K \in \mathsf{K}(m+1)$,

$$\{L(\text{D(2,3)}) : L \in \bigcup_{t \leq m} \mathsf{K}(t)\} \vdash_\omega \bigwedge_x (D_i(x) \supset K(D_j(x))).$$

**Proof.** Let $a$ be an arbitrary profile. We prove that for any $K \in \mathsf{K}(m+1)$, $\{L(\text{D(2,3)}) : L \in \bigcup_{t \leq m} \mathsf{K}(t)\} \vdash_\omega D_i(a) \supset K(D_j(a))$. For $m = 0$, this is Lemma 3.3. Now we assume the induction hypothesis that the assertion holds for $m$. Let $K = K' K_k \in \mathsf{K}(m+2)$. Then $K' \in \mathsf{K}(m+1)$. By the induction basis and some applications of Nec, MP and $\text{MP}_i$, $K'(\text{D(2,3)}) \vdash_\omega K'(D_i(a) \supset K_k(D_j(a)))$, and then $K'(\text{D(2,3)}) \vdash_\omega K'(D_i(a)) \supset K' K_k(D_j(a))$. This together with the induction hypothesis implies $\{L(\text{D(2,3)}) : L \in \bigcup_{t \leq m+1} \mathsf{K}(t)\} \vdash_\omega D_i(a) \supset K' K_k(D_j(a))$. □

Thus, if we assume the knowledge of Axiom D(1–4) up to depth $m$, the knowledge of depth $m + 1$ is necessarily involved. Hence the knowledge of D(1–4) up to depth $m + 1$

should be added. The following theorem states that this addition is inevitable, which is a general version of (3.1). This will be proved in Subsection 4.2.

**Theorem 4.A.** For any finite $m \geq 0$ and for any $K \in \bigcup_{t>m} \mathsf{K}(t)$,

$$\text{neither } \{L(\mathrm{D}(1\text{–}4)) : L \in \bigcup_{t \leq m} \mathsf{K}(t)\} \vdash_\omega K(\mathrm{D}(1\text{–}4)) \tag{4.2}$$

$$\text{nor } \{L(\mathrm{D}(1\text{–}4)) : L \in \bigcup_{t \leq m} \mathsf{K}(t)\} \vdash_\omega \neg K(\mathrm{D}(1\text{–}4)). \tag{4.3}$$

This theorem states that without assuming $K(\mathrm{D}(1\text{–}4))$ for all $K$ of all depths, some imaginary players living in the mind of the players in some depth could not know the definition (meanings) of $D_i(\cdot)$'s. To avoid this problem, we need to assume

$$\{K(\mathrm{D}(1\text{–}4)) : K \in \bigcup_{t<\omega} \mathsf{K}(t)\}.$$

This is an *infinite regress of the knowledge of axioms* D(1–4). This infinite regress is, in fact, the common knowledge of D(1–4). More explicitly, taking the conjunction of this infinite regress, we obtain the common knowledge of D(1–4), i.e., $C(\mathrm{D}(1\text{–}4))$. This corresponds to the self-reference indicated in Johansen's [8] argument.

We will adopt this infinite regress as an axiom and consider its implications. The following result holds, which was given in Kaneko-Nagashima [12] and [14]. For completeness, we will give a brief proof. In the following, when we write $D_i(\cdot), D_j(\cdot)$ without quantification of $i, j$, they are arbitrary players.

**Proposition 4.2.(1)*** : $C(\mathrm{D}(2,3)) \vdash_\omega \bigwedge_x (D_i(x) \supset C(D_j(x)))$.
**(2)***: $C(\mathrm{D}(1\text{–}3)) \vdash_\omega \bigwedge_x (D_i(x) \supset C(\mathrm{Nash}(x)))$.

**Proof.** (1) follows Proposition 4.1. Consider (2). Let $a$ be an arbitrary profile. First, D(1,2) $\vdash_0 D_i(a) \supset \mathrm{Nash}(a)$ by Lemma 3.2. Hence, by Lemma 2.2.(1), we have $C(\mathrm{D}(1,2)) \vdash_\omega C\,(D_i(a) \supset \mathrm{Nash}(a))$, and $C(\mathrm{D}(1,2)) \vdash_\omega C(D_i(a)) \supset C(\mathrm{Nash}(a))$ by Lemma 2.3.(1). Since $C(\mathrm{D}(2,3)) \vdash_\omega D_i(a) \supset C(D_i(a))$ by (1), we have $C(\mathrm{D}(1\text{–}3)) \vdash_\omega D_i(a) \supset C(\mathrm{Nash}(a))$. □

In fact, $C(\mathrm{Nash}(x))$ will be shown to be the solution of the infinite regress, $C(\mathrm{D}(1\text{–}4))$, for a game satisfying interchangeability, which will be the subject of Section 5. Here, we can ask whether the conclusion of Proposition 4.2.(2) is provable from the knowledge of D(1–4) up to some finite depth. The following theorem states the negative answer. Since this can be proved in the same manner as in the proof of Theorem 4.A, we omit the proof.

**Theorem 4.B.** For any finite $m$,

$$\text{neither } \{L(\mathrm{D}(1\text{–}4)) : L \in \bigcup_{t \leq m} \mathsf{K}(t)\} \vdash_\omega \bigwedge_x (D_i(x) \supset C(\mathrm{Nash}(x)))$$

$$\text{nor } \{L(\mathrm{D}(1\text{–}4)) : L \in \bigcup_{t \leq m} \mathsf{K}(t)\} \vdash_\omega \neg \bigwedge_x (D_i(x) \supset C(\mathrm{Nash}(x))).$$

The above derivation of the infinite regress, $C(\mathrm{D}(1\text{–}4))$ is, so far, heuristic. However, it can be formulated in the following manner.

**Theorem 4.C.** Let $\Gamma$ be a set of formulae. Suppose that $\Gamma \vdash_\omega \bigvee_x K(D_i(x)) \supset K(\mathrm{D}(1\text{-}4))$ for all $K \in \bigcup_{m<\omega} \mathsf{K}(m)$. Then $\Gamma \vdash_\omega \bigvee_x D_i(x) \supset C(\mathrm{D}(1\text{–}4))$.

**Proof.** Since $\mathrm{D}(2,3) \vdash_\omega D_i(a) \supset K_k(D_i(a))$ for $k = 1, ..., n$ by Lemma 3.3, we have, by repeating Nec, MP and $\mathrm{MP}_i$,

$$K(\mathrm{D}(2,3)) \vdash_\omega K(D_i(a)) \supset KK_k(D_i(a)) \text{ for any } K \in \bigcup_{m<\omega} \mathsf{K}(m). \tag{4.4}$$

Now we prove

$$\Gamma \vdash_\omega \bigvee_x D_i(x) \supset K(\mathrm{D}(1\text{–}4)) \tag{4.5}$$

$$\Gamma \vdash_\omega \bigvee_x D_i(x) \supset \bigvee_x K(D_i(x)) \tag{4.6}$$

for any $K \in \mathsf{K}(m), m = 0, 1....$ Then (4.5) implies $\Gamma \vdash_\omega \bigvee_x D_i(x) \supset C(\mathrm{D}(1\text{–}4))$ by $\bigwedge$-Rule.

For $m = 0$, the first is written as $\Gamma \vdash_\omega \bigvee_x D_i(x) \supset \mathrm{D}(1\text{–}4)$, which is the assumption of the theorem for null $K$. The second is a trivial statement.

The induction hypothesis is that (4.5) and (4.6) hold for $m$. Since $K(\mathrm{D}(2,3)) \vdash_\omega K(D_i(a)) \supset KK_k(D_i(a))$ by (4.4), we have

$$K(\mathrm{D}(2,3)) \vdash_\omega \bigvee_x K(D_i(x)) \supset \bigvee_x KK_k(D_i(x))$$

Hence this together with the induction hypothesis implies $\Gamma \vdash_\omega \bigvee_x D_i(x) \supset \bigvee_x KK_k(D_i(x))$. This is (4.6) for $m+1$. Then $\Gamma \vdash_\omega \bigvee_x KK_k(D_i(x)) \supset KK_k(\mathrm{D}(1\text{–}4))$ by the assumption of the theorem, we have (4.5). □

This theorem states that for any $K \in \bigcup_{m<\omega} \mathsf{K}(m)$, if the knowledge of all players to reach final decisions in the sense of $K$ implies the knowledge of D(1–4) in the same sense, then the possibility of final decisions implies the common knowledge of Axioms D(1–4).

Thus this is a formulation of the above heuristic argument for the infinite regress. In the next subsection, we will show that this set $\Gamma$ needs to contain some common knowledge.

The results given in this section are still purely solution-theoretic: they do not require the players to know the game, i.e., neither Axioms Eq nor G. The knowledge of the structure of the game will be needed in Sections 5 and 6.

### 4.2. Evaluations of the Infinite Regress

To prove the undecidability theorems given in the above subsection, we need the depth $\delta(A)$ for a formula $A$. Using this concept, we will evaluate the provability of an epistemic statement.

We define the *depth* $\delta(A)$ by induction on the structure of a formula from the inside:

(0): $\delta(A) = \emptyset$ for any atomic $A$;

(1): $\delta(\neg A) = \delta(A)$;

(2): $\delta(A \supset B) = \delta(A) \cup \delta(B)$;

(3): $\delta(\bigwedge \Phi) = \delta(\bigvee \Phi) = \bigcup_{A \in \Phi} \delta(A)$;

(4): $$\delta(K_j(A)) = \begin{cases} \{(j)\} & \text{if } \delta(A) \text{ is empty} \\ \{(j, i_1, ..., i_m) : (i_1, ..., i_m) \in \delta(A) \text{ and } j \neq i_1\} \cup & \\ \{(i_1, ..., i_m) : (i_1, ..., i_m) \in \delta(A) \text{ and } j = i_1\} & \text{otherwise.} \end{cases}$$

For any set $\Gamma$ of formulae, let $\delta(\Gamma)$ be $\bigcup_{A \in \Gamma} \delta(A)$. Define $\sup \delta(\Gamma) = \sup\{m : (i_1, ..., i_m) \in \delta(\Gamma)\}$. For example, $\delta(\mathrm{D3}_i^o) = \delta(\mathrm{D1}_i) = \{(i)\}$, $\delta(\mathrm{D3}) = \{(i) : i = 1, ..., n\}$ and $\delta(\bigwedge_i K_i(\mathrm{D(1\text{–}4)})) = \{(i, j) : i \neq j\}$.

The following lemma is the key to prove the undecidability results presented, and is proved in the Gentzen-style sequent calculus formulation of $\mathrm{GL}_\omega$ in Kaneko [10] using the (cut,Barcan)-elimination theorem for $\mathrm{GL}_\omega$. [13]

**Lemma 4.3* (Depth Lemma).** Let $K = K_{i_1} ... K_{i_m} \in \mathsf{K}(m)$, $\Gamma$ a set of formulae and $A$ a formula. Assume $\Gamma \vdash_\omega K(A)$. Assume $(i_1, ..., i_m) \notin \delta(\Gamma)$.

(1): Let $\sup \delta(\Gamma \cup \{A\}) < \omega$. Then (a) $\Gamma$ is inconsistent with respect to $\vdash_\omega$; or (b) $\vdash_\omega A$.

(2): (a) $\varepsilon\Gamma$ is inconsistent with respect to $\vdash_0$; or (b) $\vdash_\omega \varepsilon A$.

The first states that if $K_{i_1} ... K_{i_m}(A)$ is derived from the premise $\Gamma$, then, in fact, $\Gamma$ is inconsistent or $A$ is a trivial formula. The second is essentially the same, but needs some modification for some technical difficulty if $\sup \delta(\Gamma \cup \{A\})$ is infinite.

[13] This lemma is an extension of the depth lemma proved in Kaneko-Nagashima [13] for the propositional epistemic logic S4 based the cut-elimination theorem proved by Ohnishi-Matsumoto [20].

We use this lemma to prove Theorems 4.A. Before it, we can give one more result corresponding to Theorem 4.C.

**Theorem 4.D.** Let $\Gamma$ be a set of formulae with $\sup \delta(\Gamma) < \omega$, and assume that $\Gamma, \bigvee_x D_i(x)$ is consistent with respect to $\vdash_\omega$. Then $\Gamma \vdash_\omega \bigvee_x D_i(x) \supset C(\mathrm{D}(1\text{–}4))$ does not hold.

**Proof.** On the contrary, suppose $\Gamma \vdash_\omega \bigvee_x' D_i(x) \supset C(\mathrm{D}(1\text{–}4))$. This is equivalent to that $\vdash_\omega (\bigwedge \Phi) \bigwedge (\bigvee_x D_i(x)) \supset C(\mathrm{D}(1\text{–}4))$ for some finite subset $\Phi$ of $\Gamma$. Let $K \in \mathsf{K}(m)$ for $m > \sup \delta(\Gamma)$. Then $\vdash_\omega (\bigwedge \Phi) \bigwedge (\bigvee_x D_i(x)) \supset K(\mathrm{D}(1\text{–}4))$. The application of Lemma 4.3.(1) to this statement implies that either $(\bigwedge \Phi) \bigwedge (\bigvee_x D_i(x))$ is inconsistent with respect to $\vdash_\omega$ or $\vdash_\omega \mathrm{D}(1\text{–}4)$. The former is impossible by the assumption of the theorem. The latter is also impossible by Lemma 3.4.(2). □

This theorem implies that infinite depth is necessarily involved for $\Gamma$ in Theorem 4.C. In fact, we can prove the knowledge structure of common knowledge is exactly involved in $\Gamma$. However, since it needs more detailed argument, we omit it.

We can see from Lemma 4.3 that the conclusion of (2.11) could not hold without the common knowledge of Axioms Eq and G. That is, for any $L \in \bigcup_{m<\omega} \mathsf{K}(m)$,

$$L(\mathrm{Eq}), L(\mathrm{G}) \vdash_\omega \bigwedge_x C\left(\mathrm{Nash}(x) \equiv \bigvee_k \mathrm{Sol}^k(x)\right) \tag{4.7}$$

is not the case. This is proved as follows: Suppose on the contrary that (4.7) holds. Let $L \in \mathsf{K}(m)$ and $K \in \mathsf{K}(t)$ with $t > m$. Since $L(\mathrm{Eq}), L(\mathrm{G}) \vdash_\omega K\left(\mathrm{Nash}(a) \equiv \bigvee_k \mathrm{Sol}^k(a)\right)$, it follows from Lemma 4.3.(1) that $L(\mathrm{Eq}), L(\mathrm{G}) \vdash_\omega K\left(\mathrm{Nash}(a) \equiv \bigvee_k \mathrm{Sol}^k(a)\right)$ implies

either $\{L(\mathrm{Eq}), L(\mathrm{G})\}$ is inconsistent with respect to $\vdash_\omega$.

or $\vdash_\omega \mathrm{Nash}(a) \equiv \bigvee_k \mathrm{Sol}^k(a)$,

neither of which holds.

We close this section with the proof of Theorem 4.A.

**Proof of Theorem 4.A.** Let us start with the proof of (4.3). Suppose, on the contrary, that $\{L(\mathrm{D}(1\text{–}4)) : L \in \bigcup_{t\leq m} \mathsf{K}(t)\} \vdash_\omega \neg K(\mathrm{D}(1\text{–}4))$. Then $\varepsilon\mathrm{D}(1\text{–}4) \vdash_0 \neg\varepsilon\mathrm{D}(1\text{–}4)$ by Lemma 2.2.(1), which contradicts Lemma 3.4.(1).

Now consider (4.2). Suppose $\{L(\mathrm{D}(1\text{–}4)) : L \in \bigcup_{t\leq m} \mathsf{K}(t)\} \vdash_\omega K(\mathrm{D}(1\text{–}4))$ for some $K = K_{i_1}...K_{i_\ell}$ and $\ell > m$. Then since $K(\mathrm{D}(2,3)) \vdash_\omega K(D_i(a)) \supset KK_j(D_i(a))$ for any $j$ and profile $a$, we have $\{L(\mathrm{D}(1\text{–}4)) : L \in \bigcup_{t\leq m} \mathsf{K}(t)\} \vdash_\omega D_i(a) \supset KK_j(D_i(a))$. Let $j$ be different from the index of the innermost symbol of $K$. This is written as

$$\{L(\mathrm{D}(1\text{–}4)) : L \in \bigcup_{t \leq m} \mathsf{K}(t)\}, D_i(a) \vdash_\omega KK_j(D_i(a)).$$

Since $\sup \delta(\{L(\mathrm{D}(1\text{–}4)) : L \in \bigcup_{t\leq m} \mathsf{K}(t)\} \cup \{D_i(a)\}) = m + 1$, it follows from Lemma 4.3.(1) that either $\{L(\mathrm{D}(1\text{–}4)) : L \in \bigcup_{t\leq m} \mathsf{K}(t)\} \cup \{D_i(a)\}$ is inconsistent or $\vdash_\omega D_i(a)$. In the former case, $\varepsilon\mathrm{D}(1\text{–}4), D_i(a)$ is inconsistent with respect to $\vdash_0$, which is impossible by Lemma 3.4.(1). In the second case, we have $\vdash_0 D_i(a)$, which is also impossible. □

## 5. Solutions for the Infinite Regress: Solvable Games

In Section 4, we derived the infinite regress, $C(\mathrm{D}(1\text{–}4))$, of the knowledge of Axioms D(1–4). We adopt this infinite regress as an axiom and consider its implications. It was shown in Kaneko-Nagashima [12] and [14] that this infinite regress determines the final decision predicate $D_i(a)$ to be $C(\mathrm{Nash}(a))$ under the common knowledge of interchangeability. Here we will evaluate epistemic aspects of this result. We consider also the playability of a game.

### 5.1. Determination of the Final Decision Prediction and its Evaluations

Proposition 4.2.(2) states that under $C(\mathrm{D}(1\text{–}4))$, $D_i(a)$ implies $C(\mathrm{Nash}(a))$. Axiom D4 requires $D_i(\cdot)$ to be interchangeable, while $C(\mathrm{Nash}(\cdot))$ is not necessarily interchangeable. Hence $C(\mathrm{Nash}(a))$ may not capture all the properties of $C(\mathrm{D}(1\text{–}4))$. However, if we assume that the *interchangeability* condition,

$$\bigwedge_x \bigwedge_y \bigwedge_j (\mathrm{Nash}(x) \wedge \mathrm{Nash}(y) \supset \mathrm{Nash}(y_j; x_{-j})), \tag{5.1}$$

is common knowledge, then $C(\mathrm{Nash}(\cdot))$ can be regarded as capturing all the properties of $C(\mathrm{D}(1\text{–}4))$. We denote the formula of (5.1) by Int. The following proposition states that $C(\mathrm{Nash}(\cdot))$ as $D_i(\cdot)$'s satisfies our axioms.

**Proposition 5.1.(1)*:** $\vdash_\omega C\,(\mathrm{D}(1\text{–}3)[C(\mathrm{Nash})])$;

**(2)*:** $C(\mathrm{Int}) \vdash_\omega C\,(\mathrm{D4}[C(\mathrm{Nash})])$,

where D(1–3)$[C(\mathrm{Nash})]$ and D4$[C(\mathrm{Nash})]$ are the formulae obtained from D(1–3) and D4 by substituting $C(\mathrm{Nash}(a))$ for every occurrence of $D_i(a)$ ($a \in \Sigma$ and $i = 1, ..., n$) in D(1–3) and D4.

**Proof.** (1): We prove only $\vdash_\omega C\,(\mathrm{D3}[C(\mathrm{Nash})])$. Since $\vdash_\omega C(\mathrm{Nash}(a)) \supset K_i(C(\mathrm{Nash}(a)))$ for all $i = 1, ..., n$ by (2.3), we have $\vdash_\omega \bigwedge_i \bigwedge_x [C(\mathrm{Nash}(x)) \supset K_i(C(\mathrm{Nash}(x)))]$, i.e., $\vdash_\omega \mathrm{D3}[C(\mathrm{Nash})]$. By Lemma 2.2.(2), $\vdash_\omega C\,(\mathrm{D3}[C(\mathrm{Nash})])$.

(2): Since Int $\vdash_0$ Nash$(a) \bigwedge$Nash$(b) \supset$ Nash$(b_j; a_{-j})$, we have $C$(Int) $\vdash_\omega$ $C$[Nash$(a)$ $\bigwedge$Nash$(b) \supset$ Nash$(b_j; a_{-j})$] by Lemma 2.2.(2). By Lemmas 2.3.(1) and 2.4.(1), $C$(Int) $\vdash_\omega C(\text{Nash}(a)) \bigwedge C(\text{Nash}(b)) \supset C(\text{Nash}(b_j; a_{-j}))$. Since $a, b, j$ are arbitrary,

$$C(\text{Int}) \vdash_\omega \bigwedge_x \bigwedge_y \bigwedge_j [C(\text{Nash}(x)) \bigwedge C(\text{Nash}(y)) \supset C(\text{Nash}(y_j; x_{-j}))].$$

This is $C(\text{Int}) \vdash_\omega$ D4[$C$(Nash)]. By Lemma 2.2.(2), we have $C(\text{Int}) \vdash_\omega C(\text{D4}[C(\text{Nash})])$. □

Thus, $C(\text{Nash}(\cdot))$ is a solution of $C$(D(1–4)) under $C$(Int). Conversely, it can be proved in the same way as Proposition 4.2.(2) that

$$C(\text{D}(1\text{–}3)[\mathcal{A}]) \vdash_\omega \bigwedge_x (A_i(x) \supset C(\text{Nash}(x))) \text{ for } i = 1, ..., n, \tag{5.2}$$

where $\mathcal{A} = \{(A_1(a), ..., A_n(a)) : a \in \Sigma\}$ is any set of profiles of formulae indexed by $a \in \Sigma$. Thus $C(\text{Nash}(\cdot))$ is weaker than any formulae satisfying $C$(D(1–3)). Additionally, Proposition 5.1 states that $C(\text{Nash}(\cdot))$ satisfies $C$(D(1–4)) under $C$(Int). Hence $C(\text{Nash}(\cdot))$ is the deductively weakest formula among those satisfying $C$(D(1–4)). We are looking for the deductively weakest formula for $D_i(\cdot)$, since it contains no additional information other than what we intend to describe by $C$(D(1–4)).

The explicit formulation of the choice of the deductively weakest formula is given as the following axiom schemata:

(WD): $C(\text{D}(1\text{–}4)[\mathcal{A}]) \bigwedge C\left(\bigwedge_i \bigwedge_x [D_i(x) \supset A_i(x)]\right) \supset \bigwedge_i \bigwedge_x [D_i(x) \equiv A_i(x)]$,

where $\mathcal{A} = \{(A_1(a), ..., A_n(a)) : a \in \Sigma\}$ is any set of vectors of formulae indexed by $a \in \Sigma$. Although, in fact, we should, probably, regard $C$(WD) as our axiom, WD suffices for the following results. Therefore we use simply WD instead of $C$(WD).

Since Proposition 4.2.(2) and Proposition 5.1 implies

$$C(\text{D}(1\text{–}4)), C(\text{Int}) \ \vdash_\omega C\left(\text{D}(1\text{–}4)[C(\text{Nash})]\right) \bigwedge C\left(\bigwedge_x (D_i(x) \supset C(\text{Nash}(x)))\right).$$

Since this is the premise of an instance of WD, we obtain the following theorem.

**Theorem 5.A*.** $C(\text{D}(1\text{–}4)), C(\text{Int}), \text{WD} \vdash_\omega \bigwedge_x (D_i(x) \equiv C(\text{Nash}(x)))$.

Ignoring the additional common knowledge operator for Nash$(\cdot)$, Theorem 5.A can be regarded as consistent with Johansen's [8] claim that if the game is solvable (has a unique Nash equilibrium), then only Nash equilibrium satisfies his postulates.

Now we evaluate the above procedure of unique determination. First, we ask whether any formula weaker than $C(\text{Nash}(a))$ satisfies our axioms. In fact, for D(1,2) or D(1,2,4), we can replace $C(\text{Nash}(\cdot))$ by Nash$(\cdot)$, i.e., the following hold:

$$\vdash_\omega C(\text{D}(1,2)[\text{Nash}]); \text{ and } C(\text{Int}) \vdash_\omega C(\text{D4}[\text{Nash}]). \tag{5.3}$$

Thus $C$(D(1,2,4)) does not require the common knowledge operator $C(\cdot)$ for Nash($\cdot$). The common knowledge is required by D3 together with D2, which is stated in the following theorem.

**Theorem 5.B.** Let $\mathcal{A} = \{(A_1(a), ..., A_n(a)) : a \in \Sigma\}$ be a set of vectors of formulae.

**(1):** If $\vdash_\omega$ D(2,3)$[\mathcal{A}]$, then $\vdash_\omega A_i(a) \supset C(A_i(a))$ for any $i$ and $a$.

**(2):** Suppose $\max_{i,a} \sup \delta(A_i(a)) < \omega$. Then $\vdash_\omega$ D(2,3)$[\mathcal{A}]$ if and only if $\vdash_\omega \neg A_i(a)$ or $\vdash_\omega A_i(a)$ for any $i$ and $a$.

**Proof.** **(1):** Suppose $\vdash_\omega$ D(2,3)$[\mathcal{A}]$. It can be proved in the same way as the proof of Lemma 3.3 that $\vdash_\omega$ D(2,3)$[\mathcal{A}] \supset (A_i(a) \supset K_j(A_i(a)))$ for all $i, j$ and $a$. Hence $\vdash_\omega A_i(a) \supset K_j(A_i(a))$ for all $i, j$ and $a$. From this together with Nec, MP and MP$_i$, we have $\vdash_\omega A_i(a) \supset K(A_i(a))$ for all $K \in \bigcup_{t<\omega} \mathsf{K}(t)$. Hence $\vdash_\omega A_i(a) \supset C(A_i(a))$ by $\bigwedge$-Rule.
**(2):** The *if* part is straightforward. The *only-if* part is proved as follows: Consider an arbitrary $A_i(a)$. Since $\sup \delta(A_i(a)) < \omega$, we can choose $K \in \mathsf{K}(m)$ so that $m > \sup \delta(A_i(a))$. Since $\vdash_\omega A_i(a) \supset K(A_i(a))$ by (1), Lemma 4.3.(1) states that $\vdash_\omega \neg A(a)$ or $\vdash_\omega A(a)$. □

Theorem 5.B.(1) states that if some formulae satisfy D2 and D3, then they includes common knowledge. Then (2) implies that if nontrivial formulae satisfy D2 and D3, they have infinite depths.

From the viewpoint of logic, the C–Barcan axiom is indispensable for proving D3$[C(\text{Nash})]$, i.e., $C(\text{Nash}(a)) \supset K_i(C(\text{Nash}(a))$ for $i = 1, ..., n$. As mentioned before, Kaneko [10] proved that there is no (finitary or infinitary) formula $A$ such that $A \supset K_i(A)$ is provable for $i = 1, ..., n$ in GL$_\omega$ without C–Barcan.

The following theorem states that the common knowledge assumption, $C$(Int), is needed to have $C$(D4$[C(\text{Nash})]$).

**Theorem 5.C.** For any $K \in \bigcup_{m<\omega} \mathsf{K}(m)$,

$$\text{neither } K(\text{Int}) \vdash_\omega C(\text{D4}[C(\text{Nash})]) \quad \text{nor } K(\text{Int}) \vdash_\omega \neg C(\text{D4}[C(\text{Nash})])$$

**Proof.** Suppose $K(\text{Int}) \vdash_\omega C(\text{D4}[C(\text{Nash})])$. This implies $\vdash_\omega K(\text{Int}) \supset L(\text{D4}[C(\text{Nash})])$ for any $L \in \bigcup_{m<\omega} \mathsf{K}(m)$. Let the depth of $L$ is larger than that of $K$. By Lemma 4.3.(2), $\vdash_0 \neg$Int or $\vdash_0$ D4[Nash]. The first case is impossible. The second is $\vdash_0$ Int, which is not the case. Either is proved in a semantical way.

Suppose $K(\text{Int}) \vdash_\omega \neg C(\text{D4}[C(\text{Nash})])$. By Lemma 2.2.(1), we have Int $\vdash_0 \neg$D4[Nash], i.e., Int $\vdash_0 \neg$Int, which is impossible. □

## 5.2. Playability of and the Knowledge of a Game

The introduction of an epistemic structure enables us to consider the problem of playability. In our context, the playability of a game is formulated as $\bigvee_x D_i(x)$ – the existence of predicted final decisions. According to Theorem 5.A under the assumption of $C(\text{Int})$, the question is equivalent to whether or not $\bigvee_x C(\text{Nash}(x))$ is obtained from some axioms. This is different from $C(\bigvee_x \text{Nash}(x))$. The former states that there is some strategy profile $x$ such that it is common knowledge that $x$ is a Nash equilibrium, but the second states that the existence of a Nash equilibrium is common knowledge.

The *solvability* of a game is formulated as $\text{Int}\bigwedge(\bigvee_x \text{Nash}(x))$, which we denote by Solv. Since $C(\text{Solv})$ contains the existence of a Nash equilibrium, we have

$$C(\text{Solv}) \vdash_\omega C(\bigvee_x \text{Nash}(x)).$$

Nevertheless, this together with $C(\text{D(1–4)})$,WD does not imply $\bigvee_x D_i(x)$: Instead, we need to have some $\Gamma$ so that

$$C(\text{D(1–4)}), \text{WD}, \Gamma \vdash_\omega \bigvee_x C(\text{Nash}(x)). \tag{5.4}$$

We prove that $C(\text{Solv})$ is not sufficient to have $\bigvee_x D_i(x)$, but $C(\text{G})$ guarantees the playability when $g$ is solvable. Note that if $g$ is solvable, then $C(\text{G}) \vdash_\omega C(\text{Int})$.

**Theorem 5.D.(1)**: Suppose that some player $j$ has at least two strategies, i.e., $\ell_j \geq 2$. Then

neither $C(\text{D(1–4)}),\text{WD}, C(\text{Solv}) \vdash_\omega \bigvee_x D_i(x)$

nor $C(\text{D(1–4)}),\text{WD}, C(\text{Solv}) \vdash_\omega \neg \bigvee_x D(x)$.

**(2)**: Let $g$ be a solvable game. Then $C(\text{D(1–4)}),\text{WD}, C(\text{G}) \vdash_\omega \bigvee_x D(x)$.

Thus $C(\text{G})$ suffices for $\Gamma$ in (5.4) with the condition that $g$ is a solvable game, but $C(\text{Solv})$ is not sufficient. The significance of this theorem is to demarcate the knowledge of abstract conditions on the game from the knowledge of the specific structure of a game. Abstract treatments are convenient for our (investigators') considerations, but the specific knowledge of the game is needed for the players who play the game.

The first assertion of Theorem 5.D corresponds to the undecidability result presented in Kaneko-Nagashima [14] that there is a specific three-person game with a unique Nash equilibrium in mixed strategies such that the playability statement is undecidable, while the common knowledge of the existence of a Nash equilibrium is provable under the common knowledge of real closed field axioms. Their undecidability is caused by the choice of a language the players use. Contrary to theirs, our unplayability is caused by the fact that a game is given abstractly and is not specified. The second assertion states that once a game is fully specified, our undecidability is removed when it is solvable.

In fact, the playability and existence of a Nash equilibrium could not be distinguished without epistemic structures. Ignoring Axiom D3, Theorem 5.A is stated as follows:

$$C(\mathrm{D}(1,2,4)), C(\mathrm{Int}), \mathrm{WD}(1,2,4) \vdash_\omega \bigwedge_x (D_i(x) \equiv \mathrm{Nash}(x)) \tag{5.5}$$

where WD(1,2,4) is the corresponding modification of WD. In this case, $\bigvee_x D_i(x)$ is equivalent to $\bigvee_x \mathrm{Nash}(x)$. Hence

$$C(\mathrm{D}(1,2,4)), C(\mathrm{Solv}), \mathrm{WD}(1,2,4) \ \vdash_\omega \bigvee_x D_i(x).$$

Thus the abstract existence knowledge leads to $\bigvee_x D_i(x)$, contrary to Theorem 5.D.(1).

The first assertion of (1) needs a long proof, but (2) can be proved with what we have already prepared. Therefore we give the proofs of those assertions in the reverse order.

**Proof of (2)**: Since $g$ is a solvable game, it has a particular Nash equilibrium $a$. Then G $\vdash_0$ Nash$(a)$ by (2.7), which implies $C(\mathrm{G}) \vdash_\omega C(\mathrm{Nash}(a))$ by Lemma 2.2.(2). By Theorem 5.A, we have $C(\mathrm{D}(1\text{–}4)), \mathrm{WD}, C(\mathrm{G}) \vdash_\omega D_i(a)$. Hence $C(\mathrm{D}(1\text{–}4)), \mathrm{WD}, C(\mathrm{G}) \vdash_\omega \bigvee_x D_i(x)$. □

**Lemma 5.2.** Let $A$ be a formula including no $D_1, ..., D_n$. If $C(\mathrm{D}(1\text{–}4)), \mathrm{WD}, C(\mathrm{Solv}) \vdash_\omega A$, then $C(\mathrm{Solv}) \vdash_\omega A$.

**Proof.** Suppose $C(\mathrm{D}(1\text{–}4)), \mathrm{WD}, C(\mathrm{Solv}) \vdash_\omega A$. Then

$$\vdash_\omega C(\mathrm{D}(1\text{–}4)) \wedge (\bigwedge \Phi) \wedge C(\mathrm{Solv}) \supset A \text{ for some subset } \Phi \text{ of WD.}$$

Hence there is a proof $P$ of $C(\mathrm{D}(1\text{–}4)) \wedge (\bigwedge \Phi) \wedge C(\mathrm{Solv}) \supset A$. We substitute $C(\mathrm{Nash}(a))$ for each occurrences of $D_i(a)$ ($i = 1, ..., n$ and $a \in \Sigma$) in $P$. Then we have a proof $P'$ of $C(\mathrm{D}(1\text{–}4)[C(\mathrm{Nash})]) \wedge (\bigwedge \Phi[C(\mathrm{Nash})]) \wedge C(\mathrm{Solv}) \supset A$. Note that $C(\mathrm{Solv})$ and $A$ are not affected by these substitutions since they contain no $D_i, i = 1, ..., n$.

Since $C(\mathrm{Solv}) \vdash_\omega C(\mathrm{D}(1\text{–}4)[C(\mathrm{Nash})])$ by Proposition 5.1, we have $\vdash_\omega (\bigwedge \Phi[C(\mathrm{Nash})]) \wedge C(\mathrm{Solv}) \supset A$. Also, $C(\mathrm{Solv}) \vdash_\omega \bigwedge \Phi[C(\mathrm{Nash})]$ by (5.2) since $\Phi$ is a subset of WD. Hence $\vdash_\omega C(\mathrm{Solv}) \supset A$. □

**Proof of the Second Assertion of (1)**: Suppose $C(\mathrm{D}(1\text{–}4)), \mathrm{WD}, C(\mathrm{Solv}) \vdash_\omega \neg \bigvee_x D_i(x)$. By Theorem 5.A, we have $C(\mathrm{D}(1\text{–}4)), \mathrm{WD}, C(\mathrm{Solv}) \vdash_\omega \neg \bigvee_x C(\mathrm{Nash}(x))$. By Lemma 5.2, we have $C(\mathrm{Solv}) \vdash_\omega \neg \bigvee_x C(\mathrm{Nash}(x))$. By Lemma 2.2.(1), we have Solv $\vdash_0 \neg \bigvee_x \mathrm{Nash}(x)$. However, Solv $\vdash_0 \bigvee_x \mathrm{Nash}(x)$, which implies that Solv is inconsistent with respect to $\vdash_0$ . It can be proved that this is not the case. □

For the first assertion of Theorem 5.D.(1), we need one metatheorem, which was proved in Kaneko-Nagashima [15].

**Theorem 5.E.*** (Disjunctive Property): Let $\Gamma$ be a set of nonepistemic formulae, and $A_1, ..., A_k$ nonepistemic formulae. If $C(\Gamma) \vdash_\omega \bigvee_{t=1}^{k} C(A_t)$, then $C(\Gamma) \vdash_\omega C(A_t)$ for some $t = 1, ..., k$.

**Proof of the First Assertion of (1):** Suppose $C(\text{D(1–4)}), \text{WD}, C(\text{Solv}) \vdash_\omega \bigvee_x D(x)$. Then it follows from Theorem 5.A that $C(\text{D(1–4)}), \text{WD}, C(\text{Solv}) \vdash_\omega \bigvee_x C(\text{Nash}(x))$. Then we have, by Lemma 5.2, $C(\text{Solv}) \vdash_\omega \bigvee_x C(\text{Nash}(x))$. Applying Theorem 5.E to this statement, we have $C(\text{Solv}) \vdash_\omega C(\text{Nash}(a))$ by some strategy profile $a$. By Lemma 2.2.(1), we have Solv $\vdash_0$ Nash$(a)$. Then we construct an assignment based on a game $g$ which has a unique Nash equilibrium different from $a$. Here we need the assumption that $\ell_j \geq 2$ for some $j$. By Soundness for $\text{GL}_0^f$, it is not the case that Solv $\vdash_0$ Nash$(a)$. □

## 6. Solutions for the Infinite Regress: Unsolvable Games

When $g$ is an unsolvable game, Proposition 5.1.(2), *a fortiori*, Theorem 5.A, fails to hold: $C(\text{Nash}(\cdot))$ does not satisfy Axiom D4. For an unsolvable game, subsolutions play the role of Nash$(\cdot)$, instead, but we will meet two new difficulties. One, purely game theoretical, is that an unsolvable game has multiple subsolutions and the individual choice of one subsolution may lead to a double cross. Therefore they need to share some information about the choice of a subsolution. The other is that, as already discussed in Subsection 2.3, the subsolution concept needs an extensional description. Therefore we need the common knowledge of G as well as Eq. Once these axioms as well as the choice of a subsolution are assumed to be common knowledge, we would have the result parallel to that obtained in Section 5.

### 6.1. Exchange of Some Information to Choose a Subsolution

Proposition 3.2 becomes the following form.

**Proposition 6.1.(1):** $\text{Eq}, \text{G}, \text{D}^0(1,2,4) \vdash_0 \bigvee_{k=1}^{\sigma} \bigwedge_x \left(D_i(x) \supset \text{Sol}^k(x)\right)$;

(2): $C(\text{Eq}, \text{G}), C(\text{D}(1,2,4)) \vdash_\omega C\left(\bigvee_{k=1}^{\sigma} \bigwedge_x \left(D_i(x) \supset \text{Sol}^k(x)\right)\right)$,

where $C(\text{Eq}, \text{G})$ is $C(\text{Eq} \cup \text{G})$.

**Proof.** (2) follows from (1) and Lemma 2.2.(2). Now we prove (1) in a semantic way.

Let $\tau$ be any truth assignment in which Eq, g, $\text{D}^0(1,2,4)$ are true. It follows from Axioms $\text{D1}^0$ and $\text{D2}^0$ that if $D_i(a)$ is true, then $a$ is a Nash equilibrium. Hence

$\{a : \tau(D_i(a)) = \top\}$ is a subset of $E_g$. Then Axiom D4$^0$ implies that the set $\{a : \tau(D_i(a)) = \top\}$ satisfies the interchangeability condition. Since each subsolution is a maximal set of Nash equilibria satisfying the interchageability condition, $\{a : \tau(D_i(a)) = \top\}$ is a subset of some subsolution, say, $E_g^k$. Thus $\models_\tau \bigwedge_x \left(D_i(x) \supset \mathrm{Sol}^k(x)\right)$. This $k$ may depend upon $\tau$, but $\bigvee_{k=1}^{\sigma} \bigwedge_x \left(D_i(x) \supset \mathrm{Sol}^k(x)\right)$ is true in any truth assignment $\tau$ satisfying Eq, $\mathbf{g}$,D$^0$(1,2,4). By Completeness for GL$_0$, we have Eq, $\mathbf{g}$,D$^0$(1,2,4) $\vdash_0$ $\bigvee_{k=1}^{\sigma} \bigwedge_x \left(D_i(x) \supset \mathrm{Sol}^k(x)\right)$. □

The first states that for a game $g$, if $a$ is a final decision profile, then it belongs to one of the subsolutions, and the second statement is simply a conclusion of the first and Lemma 2.2.(2). In the second, however, which subsolution is implied is unknown, since the disjunction is taken over the subsolutions in the scope of common knowledge. To choose one subsolution, the players need to exchange some information.

We denote the following formula

$$\bigwedge_{k \neq t} \bigvee_{x} \left(D_i(x) \bigwedge \neg\mathrm{Sol}^k(x)\right)$$

by Sub($t$). This states that for any subsolution other than the $t$–th one, some predicted final decision profile does not belong to the subsolution. If this is common knowledge, the players can choose the subsolution Sol$^t$.

**Proposition 6.2.** Let $t$ be one of $1, ..., \sigma$. Then $C$(Eq,G), $C$(D(1–4)), $C$(Sub($t$)) $\vdash_\omega$ $\bigwedge_x (D_i(x) \supset C(\mathrm{Sol}^t(x)))$.

**Proof.** Since $\bigvee_{k=1}^{\sigma} \bigwedge_x \left(D_i(x) \supset \mathrm{Sol}^k(x)\right)$ is equivalent to Sub($t$) $\supset \bigvee_x (D_i(x) \supset \mathrm{Sol}^t(x))$ with respect to $\vdash_0$, we have, by Lemma 2.2.(2), $C$(Eq,G), $C$(D(1,2,4)) $\vdash_\omega$ $C$[Sub($t$) $\supset \bigwedge_x (D_i(x) \supset \mathrm{Sol}^t(x))$]. Hence $C$(Eq,G), $C$(D(1,2,4)) $\vdash_\omega$ $C$(Sub($t$)) $\supset \bigwedge_x [C(D_i(x)) \supset C(\mathrm{Sol}^t(x))]$ by Lemmas 2.3 and 2.4. Since $C$(D(2,3)) $\vdash_\omega \bigwedge_x (D_i(x) \supset C(D_i(x)))$ by Proposition 3.2.(1), we have $C$(Eq,G), $C$(D(1–4)), $C$(Sub($t$)) $\vdash_\omega \bigwedge_x (D_i(x) \supset C(\mathrm{Sol}^t(x)))$. □

Thus the common knowledge of the information, Sub($t$), suffices for $\bigwedge_x [D_i(x) \supset C(\mathrm{Sol}^t(x))]$. In addition to this, if $C(\mathrm{Sol}^t(\cdot))$ satisfies $C$(D(1–4)), then Axiom WD implies that $D_i(a)$ is equivalent to $C(\mathrm{Sol}^t(x))$. The following proposition states that $C(\mathrm{Sol}^t(\cdot))$ satisfies those axioms under the common knowledge of Eq and G, which can be proved in the same manner as in the proof of Proposition 5.1.

**Proposition 6.3.** $C$(Eq,G) $\vdash_\omega C\left(\mathrm{D}h[C(\mathrm{Sol}^k)]\right)$ for $k = 1, ..., \sigma$ and $h = 1, 2, 3, 4$.

By Propositions 6.2 and 6.3 together with Axiom WD, we have following theorem, which is a generalization of Theorem 5.A.

**Theorem 6.A.** Let $t$ be one of $1, \ldots, \sigma$. Then $C(\text{Eq,G}), C(\text{D(1–4)}), \text{WD}, C(\text{Sub}(t)) \vdash_\omega \bigwedge_x (D_i(x) \equiv C(\text{Sol}^t(x)))$.

## 6.2. Necessity of the Exchange of the Information

The next theorem states that without any knowledge of Sub($t$), the choice of one subsolution is impossible.

**Theorem 6.B.** Suppose that game $g$ has at least two subsolutions, i.e., $\sigma \geq 2$. Then for any $k = 1, \ldots, \sigma$ and $i = 1, \ldots, n$,

neither $C(\text{Eq,G}), C(\text{D(1–4)}) \vdash_\omega \bigwedge_x \left(D_i(x) \supset C(\text{Sol}^k(x))\right)$;

nor $C(\text{Eq,G}), C(\text{D(1–4)}) \vdash_\omega \neg \bigwedge_x \left(D_i(x) \supset C(\text{Sol}^k(x))\right)$.[14]

**Proof.** Suppose $C(\text{Eq,G}), C(\text{D(1–4)}) \vdash_\omega \bigwedge_x \left(D_i(x) \supset C(\text{Sol}^k(x))\right)$ for some $k$. Hence we have, by Lemma 2.2.(1),

$$\text{Eq}, \text{G}, \text{D}^0(1,2,4) \vdash_\omega \bigwedge_x \left(D_i(x) \supset \text{Sol}^k(x)\right). \tag{6.1}$$

Now we will show that (6.1) is impossible, which implies that the supposition does not hold. For this, we give an assignment $\tau$ on the atomic formulae so that Eq,G,$\text{D}^0(1,2,4)$ are true in $\tau$ but $\bigwedge_x \left(D_i(x) \supset \text{Sol}^k(x)\right)$ is false. This together with Soundness for $\text{GL}_0^f$ implies that $\bigwedge_x \left(D_i(x) \supset \text{Sol}^k(x)\right)$ is not provable from {Eq,G,$\text{D}^0(1,2,4)$} with respect to $\vdash_0$.

Let $\text{Sol}^{k'}(\cdot)$ be different from $\text{Sol}^k(\cdot)$. Then we define a truth assignment $\tau$ as follows:

$$(1) : \tau(\mathsf{s}_{jt} = \mathsf{s}_{jt'}) = \begin{cases} \top & \text{if } t = t' \\ \bot & \text{otherwise} \end{cases}$$

$$(2) : \tau(R_j(a : b)) = \begin{cases} \top & \text{if } g_j(a) \geq g_j(b) \\ \bot & \text{otherwise} \end{cases}$$

$$(3) : \tau(D_j(a)) = \begin{cases} \top & \text{if } a \in E_g^{k'} \\ \bot & \text{otherwise,} \end{cases}$$

where $j = 1, \ldots, n$. Then every axiom of Eq,G,$\text{D}^0(1,2,4)$ is true, but not $\models_\tau D_i(a) \supset \text{Sol}^k(a)$ for some $a$ in $E_g^{k'}$, since $\tau(D_i(a)) = \top$ but not $\models_\tau \text{Sol}^k(a)$ for any $a \in E_g^{k'} - E_g^k$. Hence $\bigwedge_x \left(D_i(x) \supset \text{Sol}^k(x)\right)$ is false in $\tau$.

[14] In fact, we can prove that the knowledge of Sub($t$) up to a finite depth is not sufficient to derive Theorem 6.A. To prove this, we need to extend $\bigwedge_x (D_i(x) \supset C(\text{Sol}^k(x)))$, but this requires a more development. We will discuss this problem in a separate paper.

Next, suppose $C(\text{Eq,G}), C(\text{D(1–4)}) \vdash_\omega \neg \bigwedge_x (D_i(x) \supset C(\text{Sol}^t(x)))$ for some $k$. By Proposition 2.2.(1), we have $\text{Eq,G,D}^0(1,2,4) \vdash_0 \neg \bigwedge_x (D_i(x) \supset \text{Sol}^t(x))$. Equivalently, $\text{Eq,G,D}^0(1,2,4) \vdash_0 \bigvee_x \left(D_i(x) \wedge \neg\text{Sol}^k(x)\right)$. Let $b \in E_g^{k'} - E_g^k$. We modify the above truth assignment $\tau$ by changing (3) into

$(3') : \tau(D_j(a)) = \bot$ for all $a$ and $j = 1, ..., n$.

In the truth assignment $\tau$ defined, $\bigvee_x \left(D_i(x) \wedge \neg\text{Sol}^k(x)\right)$ is false but $\text{D}^0(1,2,4)$ is true in the trivial way. Hence this is a contradiction. □

The above theorem implies that without sharing some information about a subsolution, a player could not reason which subsolution is played specifically. To choose one subsolution, the players need to communicate with each other to share some information. Hence purely independent decision making, described in Johansen's [8] argument, is not possible for an unsolvable game.

The following theorem states that the sufficient condition given in the previous subsection is the deductively weakest, whose proof is omitted.

**Theorem 6.C.** Let $t$ be one of $1, ..., \sigma$, and $\Gamma$ any set of formulae. If $C(\text{Eq,G}), C(\text{D(1–4)}), \Gamma \vdash_\omega C\left(\bigwedge_x (D_i(x) \equiv C(\text{Sol}^t(x)))\right)$, then $C(\text{Eq,G}), C(\text{D(1–4)}), \Gamma \vdash_\omega C(\text{Sub}(t))$.

The condition $C(\text{Sub}(t))$ eliminates all the possibilities other than the $t$-th subsolution. If the game $g$ has a specific structure, then more specific information may suffice. For example, the following is one possible assumption on the game $g$,

**(Dsub)**: $E_g^1, ..., E_g^\sigma$ are mutually exclusive.

The game of Table 2.1, the battle of the sexes, satisfies this condition. For a game satisfying Dsub, an exchange of an indication of a final decision profile suffices, instead of the elimination of possibilities. That is, for a game satisfying Dsub, the sufficient condition, $C(\text{Sub}(t))$, can be replaced by $C\left(\bigvee_x (D_i(x) \wedge \text{Sol}^t(x))\right)$. This can be proved in the same way as Theorem 6.A.

In the game of Table 2.2, the above condition is not sufficient, since $\bigvee_x (D_i(x) \wedge \text{Sol}^1(x))$ does not eliminate the possibility $C\left(\bigvee_x (D_i(x) \wedge \text{Sol}^2(x))\right)$. Nevertheless, it suffices to exchange the information of some particular pair, for example, the common knowledge that $(\mathbf{s}_{12}, \mathbf{s}_{21})$ is a final decision pair is sufficient, that is,

$$C(\text{Eq, G}), C(\text{D(1–4)}), \text{WD}, C(D_i(\mathbf{s}_{12}, \mathbf{s}_{21})) \vdash_\omega C\left(\bigwedge_x \left(D_i(x) \equiv C(\text{Sol}^1(x))\right)\right). \quad (6.2)$$

Note, however, the exchange of the information of $(\mathbf{s}_{11}, \mathbf{s}_{21})$ does not suffice.

An information exchange of a particular profile may not suffice. The following game has six subsolutions:

| | $s_{21}$ | $s_{22}$ | $s_{23}$ |
|---|---|---|---|
| $s_{11}$ | $(0,0)$ | $(1,1)^*$ | $(1,1)^*$ |
| $s_{12}$ | $(1,1)^*$ | $(1,1)^*$ | $(0,0)$ |
| $s_{13}$ | $(1,1)^*$ | $(0,0)$ | $(1,1)^*$ |

Table 6.1

and each Nash equilibrium belongs to exactly two subsolutions. In this game, the exchange of the information of choosing one subsolution needs to have the original form, $C(\text{Sub}(t))$.

## 7. Remarks on the Final Decision Axioms

### 7.1. Role of Each Axiom

Consider the role of each of Axioms D1–D4 by eliminating it from the others. The elimination of D1 or D2 makes our theory almost meaningless, and without Axiom D3 or D4, the theory would still have some structure but be much poorer.

(1) Axiom D1: Without D1, we would have the same infinite regress. The deductively weakest formula satisfying $C$(D(2–4)) is given as the common knowledge of a strategy profile under the assumption $C$(Eq,G), i.e.,

$$C(\text{Eq}, \text{G}), C(\text{D(2–4)}), \text{WD(2–4)} \vdash_\omega \bigwedge_x \left( D_i(x) \equiv C(\bigvee_y \bigwedge_j (x_j = y_j)) \right),$$

where WD(2–4) is the modification of WD by eliminating D1. This states that the prediction is the common knowledge of a strategy profile. Under these axioms, the players think about the entire situation carefully, but do not make a choice.

(2) Axiom D2: Without D2, we would not meet the common knowledge of the other axioms. Each player has $D_i(\ldots)$ but does not think about the other players' predictions. Hence the others' choices remain arbitrary for him. That is, we have

$$\text{D(1,3,4)}, \text{WD}^-(1,3,4) \vdash_\omega \bigwedge_x \left( D_i(x) \equiv K_i^+(\bigwedge_{x_i} R_i(x_i : x_{-i})) \right) \text{ for } i = 1, ..., n.$$

Here $\text{WD}^-(1,3,4)$ does not require the common knowledge operator for its premise. Thus, there are $n$ independent decision problems without interactions, though the situation itself may be interactive.

(3) Axiom D3: We already discussed the case without D3 in Subsection 5.2.

(4) Axiom D4: Without D4, our theory would be simpler. The infinite regress discussed in Section 5 remains unchanged. The main becomes

$$C(\mathrm{D}(1\text{–}3)), \mathrm{WD}(1\text{–}3) \ \vdash_\omega \bigwedge_x (D_i(x) \equiv C(\mathrm{Nash}(x))).$$

This does not depend upon the interchangeability assumption. In this case, we would loose the demarcation between the solvable and unsolvable games. Thus, the entire consideration of Section 6 would disappear.

The role of D4 could be found by looking at a different formulation of our axiomatization, which is given in the next subsection.

## 7.2. The I-System of Final Decision Axioms

We presented the axioms on final decisions $D_1(\cdot), ..., D_n(\cdot)$ as $n$-ary predicates. A prediction is an attribute of a final decision profile. Therefore each player $i$ makes simultaneously a prediction of a strategy profile for all the players. Kaneko-Nagashima [12],[14] and Kaneko [11] adopted different axiomatic systems so that predictions are separately made for each player. Player $i$ makes a prediction of player $j$'s choice, i.e., it is an attribute of a strategy for each player instead of a strategy profile. We call this system the *I-system*, and the system of the present paper the *D-system*.

To formulate the I-system, we prepare unary predicates $I_{11}(\cdot), ..., I_{1n}(\cdot); ... ; I_{n1}(\cdot), ..., I_{nn}(\cdot)$ and assume the following base axioms: for each $i = 1, ..., n$,

**Axiom** $\mathrm{I1}_i^{\mathrm{o}}$ : $\bigwedge_x \left(\bigwedge_j I_{ij}(x_j) \supset \bigwedge_{y_i} R_i(x_i; x_{-i} : y_i; x_{-i})\right)$;

**Axiom** $\mathrm{I2}_i^{\mathrm{o}}$ : $\bigwedge_x{}' \bigwedge_j \bigwedge_k (I_{ij}(x_j) \supset I_{kj}(x_j))$;

**Axiom** $\mathrm{I3}_i^{\mathrm{o}}$ : $\bigwedge_x \bigwedge_j (I_{ij}(x) \supset K_i(I_{ij}(x)))$;

**Axiom** $\mathrm{I4}_i^{\mathrm{o}}$ : $\bigwedge_i \bigwedge_j \bigwedge_k \left(\bigvee_{x_j} I_{ij}(x_j) \supset \bigvee_{x_k} I_{ik}(x_k)\right)$.

We define $\mathrm{I1}_i, ..., \mathrm{I4}_i$ to be $K_i^+(\mathrm{I1}_i^{\mathrm{o}}), ..., K_i^+(\mathrm{I4}_i^{\mathrm{o}})$. Axioms D4 and I4 connects the D-system and I-system in the following sense:

$$\vdash_\omega \mathrm{I4}_i^{\mathrm{o}} \wedge \left(\bigwedge_x [D_i(x) \equiv \bigwedge_j I_{ij}(x_j)]\right) \equiv \mathrm{D4}_i^{\mathrm{o}} \wedge \left(\bigwedge_j \bigwedge_{x_j} [I_{ij}(x_j) \equiv \bigvee_{y_{-j}} D_i(x_j, y_{-j})]\right).$$

Thus, we can start either with the I-system together the definition of $D_i(a)$ to be $\bigwedge_j \mathrm{I}_{ij}(a)$ or either with the D-system together the definition of $I_{ij}(a_j)$ to be $\bigvee_{y_{-j}} D_i(a_j, y_{-j})$. Furthermore, we can prove the entire equivalence between $\{\mathrm{I}_i(1\text{–}4) \wedge K_i^+[\bigwedge_x (D_i(x) \equiv$

$\bigwedge_j \mathrm{I}_{ij}(x))] : i = 1, ..., n\}$ and $\{\mathrm{D}_i(1\text{–}4) \wedge K_i^+[\bigwedge_j \bigwedge_{x_i}(I_{ij}(x_j) \equiv \bigvee_{y_{-j}} D_i(x_j, y_{-j}))] : i = 1, ..., n\}$.

### 7.3. Fully and Partially Interactive Games

Ignoring the difference between the D-system and I-system, the axiomatization in this paper is a special case of that in Kaneko [11]. In his axiomatization, multiple systems of axioms for decision making may be permitted, depending upon a game. A necessary and sufficient condition on a game for only the system in this paper to be permissible is given in [11], which states that each player needs to predict all the others' decisions for his decision to maximize payoffs. A game is called *fully interactive* iff it satisfies this condition, and otherwise, it is called (properly) *partially interactive.* The system in this paper is permitted for any games, though partially interactive games may have multiple "weaker" systems of axioms.

Here we give some simple examples of fully and partially interactive games. In Prisoner's Dilemma (Table 2.3), since each player has a dominant strategy, $\mathbf{s}_{i2}$, it is possible for each player $i$ to make a decision satisfying utility maximization only with the knowledge of his own payoff function but without predicting the other's decision. This argument is formulated as one system of axioms in [11], but the solution for this system does not give the knowledge of the other player's decision. This argument is not applied to the game of Table 2.4, since only player 1 has a dominant strategy, $\mathbf{s}_{12}$. In this game, 1 can ignore 2's choice, but 2 needs to predict 1's choice. This situation is also formulated as a system of axioms. In these games, if we require each player to infer the other's decision, we would have the system in this paper.

| | $\mathbf{s}_{21}$ | $\mathbf{s}_{22}$ | $\mathbf{s}_{43}$ |
|---|---|---|---|
| $\mathbf{s}_{11}$ | $(5,5)$ | $(1,6)$ | $(3,0)$ |
| $\mathbf{s}_{12}$ | $(6,1)$ | $(3,3)^*$ | $(0,2)$ |
| $\mathbf{s}_{13}$ | $(0,3)$ | $(2,0)$ | $(2,2)$ |

Table 7.1 (and 3, 4's payoffs if $(\mathbf{s}_{12}, \mathbf{s}_{22})$ is played)

| | $\mathbf{s}_{41}$ | $\mathbf{s}_{42}$ | $\mathbf{s}_{43}$ |
|---|---|---|---|
| $\mathbf{s}_{31}$ | $(3,3)$ | $(6,1)$ | $(0,2)$ |
| $\mathbf{s}_{32}$ | $(1,6)$ | $(5,5)$ | $(3,0)$ |
| $\mathbf{s}_{33}$ | $(2,0)$ | $(0,3)$ | $(2,2)$ |

Table 7.2 (otherwise)

The game of Table 7.1 obtained by adding one strategy to each player to Prisoner's Dilemma is fully interactive: this has the same Nash equilibrium $(\mathbf{s}_{12}, \mathbf{s}_{22})$, but without predicting the other's decision, each player cannot make a decision to maximize his payoff. This game differs from the games of Tables 2.3 and 2.4 in that the system in this paper is only permissible.

In fact, there is a great spectrum of games from those with dominant strategies to fully interactive games. One typical (4-person) example is as follows: each player $i$

($i = 1, ..., 4$) has three pure strategies $s_{i1}, s_{i2}, s_{i3}$. The payoffs for players 1 and 2 are determined by their own strategies, which are given by Table 7.1. Those for 3 and 4 depend upon their strategies as well as the choices of 1 and 2 : if 1 and 2 choose $(s_{12}, s_{22})$, then the payoffs of 3 and 4 are given also as Table 7.1 with the replacements of players 1, 2 with 3, 4; and if 1 and 2 choose a strategy pair other than $(s_{12}, s_{22})$, then their payoff matrix is given by Table 7.2 (obtained from 7.1 by permuting the roles of $s_{i1}$ and $s_{i2}$ for $i = 3, 4$). In this game, if 1 and 2 ignore 3 and 4, then 1 and 2 are facing the game of Table 7.1, which is fully interactive. Players 3 and 4 still need to infer the choices of 1, 2, and their part is fully also regarded as fully interactive.

We can find a similar structure in the game of Table 2.4. Player 1 can ignore 2 but 2 needs to predict 1's decision. However, each interactive part consists of a single player.

When a partially interactive game permits a "weaker" system of axioms, it has a part which can be regarded as fully interactive in that each player needs to predict the others' decisions in the part, though sometimes it is trivial in the sense that it consists of a single player. The system of axioms restricted to such a fully interactive part of the game is regarded as the same as the system in this paper. Therefore our analysis is applied also to the fully interactive parts of partially interactive games.

## References


[1] Aumann, R. J. and A. Brandenberger, (1995), Epistemic Conditions for Nash equilibrium, *Econometrica* 63, 1161-1180.

[2] Bacharach, M., (1987), A Theory of Rational Decision in Games, *Erkenntnis* 27, 17-55.

[3] Bernheim, B. D., (1984), Rationalizable Strategic Behavior, *Econometrica* 52, 1007-1028.

[4] Bernheim, B. D., (1986), Axiomatic Characterizations of Rational Choice in Strategic Environments, *Scand. J. Economics* 88, 473-488.

[5] 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.

[6] Harsanyi, J., (1967-68), Games with Incomplete Information Played by 'Bayesian' Players, *Management Sciences* 14, 159-182, 320-334, 486-502.

[7] Jansen, M. J. M., (1981), Maximal Nash Subsets for Bimatrix Game, *Naval Research Logistics Quarterly* 28, 147–152.

[8] Johansen, L., (1982), On the Status of the Nash Type of Noncooperative Equilibrium in Economic Theory, *Scand. J. Economics* 84, 421-441.

[9] Kaneko, M., (1996a), Common Knowledge Logic and Game Logic, ISEP.DP.694, 1996, August.

[10] Kaneko, M., (1996b), Depth of Knowledge and the Barcan Inferences in Game Logic, forthcoming.

[11] Kaneko, M., (1997), Decision Making in Partially Interactive Games, forthcoming.

[12] Kaneko, M. and T. Nagashima, (1991), Final Decisions, Nash Equilibrium and Solvability in Games with the Common Knowledge of Logical Abilities, *Mathematical Social Sciences* 22, 229-255

[13] Kaneko, M., and T. Nagashima, (1994), Axiomatic Indefinability of Common Knowledge in Finitary Logics, To appear in *Epistemic Logic and the Theory of Games and Decision,* eds. M. Bacharach, L. A. Gerard-Varet, P. Mongin and H. Shin.

[14] Kaneko, M., and T. Nagashima, (1996), Game Logic and its Applications I. *Studia Logica* 57, 325-354.

[15] Kaneko, M. and T. Nagashima, (1997), Game Logic and its Applications II. To appear in *Studia Logica.*

[16] Karp, C., (1964), *Languages with Expressions of Infinite Length*, North-Holland.

[17] Kats, A., and J.-F. Thisse, (1992), Unilaterally Competitive Games, *International Journal of Game Theory* 21, 291-299.

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

[19] Nash, J. F., (1951), Noncooperative Games, *Annals of Mathematics* 54, 286-295.

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

[21] Pearce, D. G., (1984), Rationalizable Strategic Behavior and the Problem of Perfection, *Econometrica* 52, 1029-1050.

[22] Savage, L. J., (1954, 1st ed.), *The Foundations of Statistics*, Wiley, New York, (1972).