Critical Comparisons between the Nash Noncooperative Theory and Rationalizability*†

Tai-Wei Hu‡and Mamoru Kaneko§

17 April 2012

Abstract

The theories of Nash noncooperative solutions and of rationalizability intend to describe the same target problem of *ex ante* individual decision making, but they are distinctively different. We consider what their essential difference is by giving a unified approach and parallel derivations of their resulting outcomes. Our results show that the only difference lies in the use of quantifiers for each player's predictions about the other's possible decisions; the universal quantifier for the former and the existential quantifier for the latter. Based on this unified approach, we discuss the statuses of those theories from the three points of views: Johansen's postulates, prediction/decision criteria, and the free-will postulate vs. complete determinism. One conclusion we reach is that the Nash theory is coherent with the free-will postulate, but we would meet various difficulties with the rationalizability theory.

JEL Classification Numbers: B40, C70, C72

Key words: Nash equilibrium, Solvability, Rationalizability, Prediction/Decision Criterion, Infinite Regress, Simultaneous Equations

1. Introduction

We make critical comparisons between the theory of *Nash noncooperative solutions* due to Nash [20] and the theory of *rationalizable strategies* due to Bernheim [3] and Pearce [21]. Each theory is intended to be a theory of *ex ante* individual decision making in

*The authors are partially supported by Grant-in-Aids for Scientific Research No.21243016 and No.2312002, Ministry of Education, Science and Culture.

†The authors thank Geir B. Ascheim and Jean-Jacques Herings for valuable comments and discussions on an earlier draft of this paper.

‡MEDS Department, Northwestern University, Evanston, IL, (t-hu@kellogg.northwestern.edu)

§IPPS, University of Tsukuba, Ibaraki 305-8573, Japan (kaneko@sk.tsukuba.ac.jp)

a game, and thus focuses on the decision-making process before the actual play of the game. The difference in their resulting outcomes has been well analyzed and known. However, their conceptual difference has not been much discussed. In this paper, we evaluate these two theories while considering certain conceptual bases of game theory and addressing the question of logical coherence of these theories with them.

We begin with a brief review of these theories. It is well known that Nash [20] provides the concept of Nash equilibrium and proves its existence in mixed strategies. However, it is less known that the main focus of [20] is on *ex ante* individual decision making. In that paper, various other concepts are developed, including interchangeability, solvability, subsolutions, symmetry, and values; those concepts are ingredients of a theory of *ex ante* individual decision making, though the aim is not explicitly stated in [20]. This view is discussed in Nash's [19] dissertation (p.23) and a few other papers such as Johansen [11] and Kaneko [12][1]. We call the entire argumentation the *Nash noncooperative theory*[2].

On the other hand, in the literature, the theory of rationalizability is typically regarded as a faithful description of *ex ante* individual decision making in games, expressing the common knowledge of "rationality". Mas-Colell *et al.* [15], p.243 wrote "*The set of rationalizable strategies consists precisely of those strategies that may be played in a game where the structure of the game and the player's rationality are common knowledge among the players.*" This view is common in many standard game theory/microeconomics textbooks.

The literature exhibits a puzzling feature: Both theories target *ex ante* individual decision making, and both are widely used by many researchers. However, their formal definitions, predicted outcomes, and explanations differ considerably. This puzzling feature raises the following questions: How should we make comparisons between these theories? Then, what are their main differences? How would the difference be evaluated? What are bases for such an evaluation? This paper attempts to answer these questions.

We formulate the two theories in terms of prediction/decision criteria, which gives a unified framework for comparisons of these theories. For the Nash theory, the criterion is given by the following requirements:

N1^o: player 1 chooses his best strategy against <u>all</u> of his predictions
about player 2's choice based on N2^o;

N2^o: player 2 chooses his best strategy against <u>all</u> of his predictions
about player 1's choice based on N1^o.

We may say that player 1 makes a decision if it satisfies N1^o; however, to determine this

[1] Millham [17] and Jansen [10] study the mathematical structure of the solution and subsolutions, but do not touch the view.

[2] The mathematical definition of Nash equilibrium allows different interpretations such as a steady state in a repeated situation (one variant is the "mass-action" interpretation due to Nash [19], pp.21-22), but we do not touch other interpretaitons. See Kaneko [13] for them.

decision, N1^o requires a prediction about 2's possible decisions, which are determined by N2^o. The symmetric form N2^o determines a decision for 2 if he predicts 1's decisions. In this sense, these requirements are circular. Also, they can be regarded as a system of simultaneous equations with players' decisions/predictions as unknown. In Section 3, we show that the system N1^o-N2^o characterizes the Nash noncooperative solution as the greatest set satisfying them if the game is solvable (the set of Nash equilibria is interchangeable); and for an unsolvable game, a maximal set satisfying them is a subsolution.

The rationalizable strategies are characterized by another prediction/decision criterion R1^o-R2^o:

R1^o: player 1 chooses his best strategy against <u>some</u> of his predictions about player 2's choice based on R2^o;

R2^o: player 2 chooses his best strategy against <u>some</u> of his predictions about player 1's choice based on R1^o.

These are obtained from N1^o-N2^o simply by replacing the quantifier "for all" by "for some" before predictions about the other player's decisions. These requirements are closely related to the BP-property ("best-response property" in Bernheim [3] and Pearce [21]), and the characterization result is given in Section 3.

The above prediction/decision criteria and characterization results unify the Nash noncooperative theory and rationalizability theory, and pinpoint their difference: It is the choice of the universal or existential quantifiers for predictions about the other player's possible decisions. To evaluate this difference, we first review the discussion of *ex ante* decision making in games given in Johansen [11]. In his argument, a theory of *ex ante* decision making in games should describe a player's active inferences based on certain axioms about his own and the other's decision-making. Johansen gives four postulates for the Nash solution, although his argument there is still informal and contains some ambiguities.

Our formulation of N1^o-N2^o may be viewed as an attempt to formalize his postulates in the language of classical game theory. The pinpointed difference between the two theories clarifies the precise requirements in those postulates to obtain the Nash theory. One of Johansen's postulates requires that any possible decision be a best response to the predicted decisions, which is violated by the "for some" requirement in R1^o-R2^o. His postulates help to clarify N1^o-N2^o, and *vice versa.* Nevertheless, his postulates contain some subtle concepts, which go beyond the language of classical game theory.

One such concept is "rationality". In the theory of rationalizability, "rationality" is typically regarded as equivalent to payoff maximization. In Johansen's postulates, however, payoff maximization is separated from "rationality", and is only one component of "rationality". We also take this broader view of "rationality"; in our formulation, it includes, but not limited to, the prediction/decision criterion and logical abilities to understand their implications. This broader view allows further research on decision criterion (such as the additional principles needed for specific classes of unsolvable games) and investigations of how players' logical abilities affect their decisions.

To evaluate the difference further, we go to deeper methodological assumptions: the *free-will postulate* vs. *complete determinism*. The former, stating that each player has free will, is automatically associated with decision making. The quantifier "for all" in N1^o-N2^o is coherent with the application of the free-will postulate between the players. On the other hand, as will be argued in Section 4, the theory of rationalizability is better understood from the perspective of complete determinism. Indeed, the epistemic justification for rationalizability begins with a complete description of players' actions as well as mental states, and characterizes classes of those states by certain assumptions.

As a result, our problem is a choice between two methodological assumptions, the free-will postulate and complete determinism. This choice is discussed in Morgenstern [18] and Heyek [8] in the context of economics and/or social science in general. Based upon their arguments, we will conclude that the free-will postulate is more coherent with large part of social science than complete determinism. From this perspective, the Nash theory is preferable to rationalizability.

The Nash theory might be less preferred in that it does not recommend definite decisions for unsolvable games. However, it may not be a defect from the perspective that it points out that additional principles, other than the decision criteria given above, are needed for decision making in unsolvable games. A general study of such additional principles is beyond the scope of this paper, but we remark that many applied works appeal to principles such as symmetry (which is already discussed in Nash [20]) and the Pareto criterion. As an instance, we will give an argument with the Pareto principle for the class of games of strategic complementarity in Section 3.1. From a theoretical perspective, our approach provides a framework to discuss coherence between basic decision criteria and additional principles.

The paper is written as follows: Section 2 introduces the theories of Nash noncooperative solutions and rationalizable strategies; we restrict ourselves to finite 2-person games for simplicity. Section 3 formulates N1^o-N2^o and R1^o-R2^o, and gives two theorems characterizing the Nash noncooperative theory and rationalizability. In Section 4, we discuss implications from them considering foundational issues. Section 5 gives a summary and states continuation to the companion paper.

2. Preliminary Definitions

In this paper, we restrict our analysis to finite 2-person games with pure strategies. In Section 3.3, we discuss required changes for our formulation to accommodate mixed strategies.

We begin with basic concepts in a finite 2-person game. Let $G = (N, \{S_i\}_{i \in N}, \{h_i\}_{i \in N})$

be a finite 2-person game, where $N = \{1, 2\}$ is the set of players, S_i is the finite set of pure strategies and $h_i : S_1 \times S_2 \to R$ is the payoff function for player $i \in N$. We assume $S_1 \cap S_2 = \emptyset$. When we take one player $i \in N$, the remaining player is denoted by j. Also, we write $h_i(s_i; s_j)$ for $h_i(s_1, s_2)$. The property that s_i is a *best-response* against s_j, i.e.,

$$h_i(s_i; s_j) \geq h_i(s_i'; s_j) \text{ for all } s_i' \in S_i, \tag{2.1}$$

is denoted by $\mathrm{Best}(s_i; s_j)$. Since $S_1 \cap S_2 = \emptyset$, the expression $\mathrm{Best}(s_i; s_j)$ has no ambiguity. A pair of strategies (s_1, s_2) is a *Nash equilibrium* in G iff $\mathrm{Best}(s_i; s_j)$ holds for both $i = 1, 2$. We use $E(G)$ to denote the set of all Nash equilibria in G. The set $E(G)$ may be empty.

Nash Noncooperative Solutions: A subset E of $S_1 \times S_2$ is *interchangeable* iff

$$(s_1, s_2), (s_1', s_2') \in E \text{ imply } (s_1, s_2') \in E. \tag{2.2}$$

It is known that this requirement is equivalent for E to have the product form, as stated in the following lemma.

Lemma 2.1. Let $E \subseteq S_1 \times S_2$ and let $E_i = \{s_i : (s_i; s_j) \in E \text{ for some } s_j \in S_j\}$ for $i = 1, 2$. Then, E satisfies (2.2) if and only if $E = E_1 \times E_2$.

Now, let $\mathbf{E} = \{E : E \subseteq E(G) \text{ and } E \text{ satisfies (2.2)}\}$. We say that E is the *Nash solution* iff E is nonempty and is the greatest set in $\mathbf{E}$, i.e., $E' \subseteq E$ for any $E' \in \mathbf{E}$. We say that E is a *Nash subsolution* iff E is a nonempty maximal set in $\mathbf{E}$, i.e., there is no $E' \in \mathbf{E}$ such that $E \subsetneq E'$.

Table 2.1

	$\mathbf{s}_{21}$	$\mathbf{s}_{22}$
$\mathbf{s}_{11}$	$(2, 2)$	$(1, 1)$
$\mathbf{s}_{12}$	$(1, 1)$	$(0, 0)$

Table 2.2

	$\mathbf{s}_{21}$	$\mathbf{s}_{22}$
$\mathbf{s}_{11}$	$(1, 1)$	$(1, 1)$
$\mathbf{s}_{12}$	$(1, 1)$	$(0, 0)$

When $E(G) \neq \emptyset$, $E(G)$ is the Nash solution if and only if $E(G)$ satisfies (2.2). When the Nash solution exists for game G, G is called *solvable*. The game of Table 2.1 is solvable. On the other hand, a game G may be unsolvable for two reasons: either $E(G) = \emptyset$ or $E(G)$ is nonempty but violates (2.2). For a game G with $E(G) \neq \emptyset$, a subsolution exists always; specifically, for any $(s_1, s_2) \in E(G)$, there is a subsolution E^o containing (s_1, s_2). This E^o may not be unique: The game of Table 2.2 is not solvable and has two subsolutions: $\{(\mathbf{s}_{11}, \mathbf{s}_{21}), (\mathbf{s}_{11}, \mathbf{s}_{22})\}$ and $\{(\mathbf{s}_{11}, \mathbf{s}_{21}), (\mathbf{s}_{12}, \mathbf{s}_{21})\}$, and both include $(\mathbf{s}_{11}, \mathbf{s}_{21})$.

In Section 3, it will be argued that the Nash solution can be regarded as a theory of *ex ante* decision making in games. Here we give two comments about this argument.

First, for a solvable game, the theory recommends the set of possible decisions for each player, i.e., the set of Nash strategies for him; moreover, the recommendation also includes the set of predicted decisions of the other player. This means that from player 1's perspective, $E_1(G) = \{s_1 \in S_1 : (s_1, s_2) \in E(G) \text{ for some } s_2\}$ describes player 1's possible decisions, while $E_2(G) = \{s_2 \in S_2 : (s_1, s_2) \in E(G) \text{ for some } s_1\}$ is player 1's predictions of player 2's decisions. As shown later, predictions about player 2's decisions are crucial to determine player 1's possible decisions from the perspective of *ex ante* decision making in games.

Second, the Nash theory does not provide a definite recommendation for decisions if the game is unsolvable, even if a subsolution exists. Suppose that G has exactly two subsolutions, say, $F^1 = F_1^1 \times F_2^1$ and $F^2 = F_1^2 \times F_2^2$ with $F_i^1 \neq F_i^2$ for $i = 1, 2$. One may think that the Nash theory would recommend the set $E_i = F_i^1 \cup F_i^2$ for player i as the set of possible decisions to play G. However, this is not valid; we cannot find a set E_1' or E_2' such that $E_1' \times (F_2^1 \cup F_2^2)$ or $(F_1^1 \cup F_1^2) \times E_2'$ satisfies interchangeability.

Rationalizable Strategies: Now, we turn to rationalizability. The pure strategy version introduced here is known as *point-rationalizability* due to Bernheim [3]. We begin with the iterative definition of rationalizability. A sequence of sets of strategies, $\{(R_1^\nu(G), R_2^\nu(G))\}_{\nu=0}^\infty$, is inductively defined as follows: for $i = 1, 2$, $R_i^0(G) = S_i$, and

$$R_i^\nu(G) = \{s_i : \text{Best}(s_i; s_j) \text{ holds for some } s_j \in R_j^{\nu-1}(G)\} \text{ for any } \nu \geq 1. \tag{2.3}$$

We obtain rationalizable strategies by taking the intersection of these sets, i.e., $R_i(G) = \bigcap_{\nu=0}^\infty R_i^\nu(G)$ for $i = 1, 2$; a pure strategy $s_i \in S_i$ is *rationalizable* iff $s_i \in R_i(G)$.

It is shown by induction on ν that $R_i^\nu(G)$ is nonempty for all ν and $i = 1, 2$. Also, each sequence $\{R_i^\nu(G)\}_\nu$ is monotonically decreasing. Because each $R_i^\nu(G)$ is finite and nonempty, $R_i^\nu(G)$ becomes constant after some $\overline{\nu}$; as a result, $R_i(G)$ is nonempty. These facts are more or less known, but we give a proof for completeness.

Lemma 2.2. $\{R_i^\nu(G)\}_\nu$ is a decreasing sequence of nonempty sets, i.e., $R_i^\nu(G) \supseteq R_i^{\nu+1}(G) \neq \emptyset$ for all ν.

Proof. We show by induction over ν that the two sequences $\{R_i^\nu(G)\}_\nu$, $i = 1, 2$, are decreasing with respect to the set-inclusion relation. Once this is shown, since S_i is finite, we have $R_i(G) = \bigcap_{\nu=0}^\infty R_i^\nu(G) \neq \emptyset$. For the base case of $\nu = 0$, we have $R_i^0(G) = S_i \supseteq R_i^1(G)$ for $i = 1, 2$. Now, suppose the hypothesis that this inclusion holds up to ν and $i = 1, 2$. Let $s_i \in R_i^{\nu+1}(G)$. By (2.3), $\text{Best}_i(s_i; s_j)$ holds for some $s_j \in R_j^\nu(G)$. Since $R_j^{\nu-1}(G) \supseteq R_j^\nu(G)$ by the supposition, $\text{Best}_i(s_i; s_j)$ holds for some $s_j \in R_j^{\nu-1}(G)$. This means $s_i \in R_i^\nu(G)$. ■

Criterion for Prediction/Decision Making: Our discussion of *ex ante* decision making in games begins with a prediction/decision criterion[3]. While comparison between the Nash theory and rationalizability is our concern, some simpler examples of

[3]A general concept of a prediction/decision criterion is formulated in an epistemic logic of shallow depths in Kaneko-Suzuki [14].

decision criteria may be helpful. First, utility maximization can be regarded as a decision criterion in a non-interactive context, which recommends the set of decisions maximizing a given utility function. In game theory, a classical example of a decision criterion is the *maximin* criterion due to von Neumann-Morgenstern [24]: It recommends a player to choose a strategy maximizing the guarantee level (that is, the minimum payoff for a strategy). In $G = (N, \{S_i\}_{i \in N}, \{h_i\}_{i \in N})$, let E_i be a nonempty subset of $S_i, i = 1, 2$. The set E_i is interpreted as the set of possible decisions for player i based on the maximin criterion. The criterion is formulated as follows:

NM1: for each $s_1 \in E_1$, s_1 maximizes $\min_{s_2 \in S_2} h_1(s_1; s_2)$;

NM2: for each $s_2 \in E_2$, s_2 maximizes $\min_{s_1 \in S_1} h_2(s_2; s_1)$.

These are not interactive, since NMi, $i = 1, 2$, can recommend a decision without considering NMj, and player i needs to know only his own payoff function. Thus, no prediction is involved for decision making with this criterion.

A more sophisticated criterion may allow one player to consider the other's criterion. One possibility is the following:

N1: for each $s_1 \in E_1$, Best(s_1, s_2) holds for all $s_2 \in E_2$;

NM2: for each $s_2 \in E_2$, s_2 maximizes $\min_{s_1 \in S_1} h_2(s_2; s_1)$.

The criterion N1 requires player 1 to predict player 2's decisions and to choose his best decision against that prediction, while player 2 still adopts the maximin criterion. In this sense, their interpersonal thinking stops at the second level. In the Nash theory and rationalizability theory, we would meet some circularity and their interpersonal thought goes beyond the second level.

There may be multiple pairs of (E_1, E_2) that satisfies a given decision criterion. Without other information than the criterion and components of the game, a player (and we) cannot make a further choice of particular strategies among those satisfying the criterion. In the case of NM1-NM2, E_i should consist of all strategies maximizing $\min_{s_2 \in S_2} h_1(s_1; s_2)$; that is, E_i is the greatest set satisfying NMi. In the case of N1-NM2, this should also be applied to player 1's predictions about 2's choice: E_2 in N1 should be the greatest set satisfying NM2. We will impose this greatest-set requirement for E_i in Section 3; this is not a mere mathematical requirement, but is very basic for the consideration of *ex ante* decision making, as it will be discussed later.

3. Parallel Derivations of the Nash Noncooperative Solutions and Rationalizable Strategies

In this section we give two parallel decision criteria, and derive the Nash noncooperative solutions and the rationalizable strategies from those criteria. Our characterization results pinpoint the difference between the two theories. This difference is used as the basis for our evaluation of these two theories, which comes in Section 4. We give remarks on the mixed strategy versions of those derivations in Section 3.3.

3.1. The Nash Noncooperative Solutions

The decision criterion for the Nash solution formalizes the statements $\mathrm{N1}^o$ and $\mathrm{N2}^o$ in Section 1. This criterion, N1-N2, is formulated as follows: Let E_i be a subset of $S_i, i = 1, 2$, interpreted as the set of possible decisions based on N1-N2,

N1: for each $s_1 \in E_1$, $\mathrm{Best}(s_1; s_2)$ holds for all $s_2 \in E_2$;

N2: for each $s_2 \in E_2$, $\mathrm{Best}(s_2; s_1)$ holds for all $s_1 \in E_1$.

These describe how each player makes his decisions; when one player's viewpoint is fixed, one of N1-N2 is interpreted as decision making, and the other is interpreted as prediction making. For example, from player 1's perspective, N1 describes his decision making, and N2 describes his prediction making.

Mathematically, N1 and N2 can be regarded as a system of simultaneous equations with unknown E_1 and E_2. First we give a lemma showing that (E_1, E_2) satisfies N1-N2 if and only if it consists only of Nash equilibria.

Lemma 3.1. Let E_i be a nonempty subset of S_i for $i = 1, 2$. Then, (E_1, E_2) satisfies N1-N2 if and only if any $(s_1, s_2) \in E_1 \times E_2$ is a Nash equilibrium in G.

Proof. (Only-If): Let (s_1, s_2) be any strategy pair in $E_1 \times E_2$. By N1, $h_1(s_1, s_2)$ is the largest payoff over $h_1(s'_1, s_2), s'_1 \in S_1$. By the symmetric argument, $h_2(s_1, s_2)$ is the largest payoff over s'_2's. Thus, (s_1, s_2) is a Nash equilibrium in G.
(If): Let $(s_1, s_2) \in E_1 \times E_2$ be a Nash equilibrium. Since $h_1(s_1, s_2) \geq h_1(s'_1, s_2)$ for all $s'_1 \in S_1$, we have N1. We have N2 similarly. ■

Regarding N1-N2 as a system of simultaneous equations with unknown E_1 and E_2, there may be multiple solutions; indeed, any Nash equilibrium pair as a singleton set is a solution for N1-N2. However, the sets E_1 and E_2 should be based only on the information of the game structure G. This implies that we should look for the pair of greatest sets (E_1, E_2) that satisfies N1-N2[4].

The following theorem dictates that N1-N2 is a characterization of the Nash solution theory.

Theorem 3.2 (The Nash Noncooperative Solutions): (0): G has a Nash equilibrium if and only if there is a nonempty pair (E_1, E_2) satisfying N1-N2.

[4] If any additional information is available, then we extend N1-N2 to include it and should consider the pair of greatest sets satisfying the new requirements.

(1): Suppose that G is solvable. Then the greatest pair (E_1, E_2) satisfying N1-N2 exists and $E = E_1 \times E_2$ is the Nash solution $E(G)$.

(2): Suppose that G has a Nash equilibrium but is unsolvable. Then E is a Nash subsolution if and only if (E_1, E_2) is a nonempty maximal pair satisfying N1-N2.

Proof. **(0)**: If (s_1, s_2) is a Nash equilibrium of G, then $E_1 = \{s_1\}$ and $E_2 = \{s_2\}$ satisfy N1-N2. Conversely, if a nonempty pair (E_1, E_2) satisfies N1-N2, then, by Lemma 3.1, any pair $(s_1, s_2) \in E_1 \times E_2$ is a Nash equilibrium of G.

(1):(If): Let (E_1, E_2) be the greatest pair satisfying N1-N2. It suffices to show $E(G) = E_1 \times E_2$. By Lemma 3.1, any $(s_1, s_2) \in E_1 \times E_2$ is a Nash equilibrium. Conversely, let $(s'_1, s'_2) \in E(G)$ and $E'_i = \{s'_i\}$ for $i = 1, 2$. Since this pair (E'_1, E'_2) satisfies N1-N2, we have $(s'_1, s'_2) \in E'_1 \times E'_2 \subseteq E_1 \times E_2$. Hence, $E(G) = E_1 \times E_2$.

(Only-If): Since E is the Nash solution, it satisfies (2.2). Hence, E is expressed as $E = E_1 \times E_2$ by Lemma 2.1. Since it consists of Nash equilibria, (E_1, E_2) satisfies N1-N2 by Lemma 3.1. Since $E(G) = E = E_1 \times E_2$, (E_1, E_2) is the greatest pair having N1-N2.

(2): (If): Let (E_1, E_2) be a maximal pair satisfying N1-N2, i.e., there is no (E'_1, E'_2) satisfying N1-N2 with $E_1 \times E_2 \subsetneq E'_1 \times E'_2$. By Lemma 3.1, $E_1 \times E_2$ is a set of Nash equilibria. Let E' be a set of Nash equilibria satisfying (2.2) with $E_1 \times E_2 \subseteq E'$. Then, E' is also expressed as $E'_1 \times E'_2$. Since $E'_1 \times E'_2$ satisfies N1-N2 by Lemma 3.1, we have $E'_i \subseteq E_i$ for $i = 1, 2$ by maximality for (E_1, E_2). By the choice of E', we have $E_1 \times E_2 = E'$. Thus, E is a maximal set satisfying interchangeability(2.2).

(Only-If): Since E is a subsolution, it satisfies (2.2). Hence, E is expressed as $E = E_1 \times E_2$. Also, by Lemma 3.1, (E_1, E_2) satisfies N1-N2. Since $E = E_1 \times E_2$ is a subsolution, (E_1, E_2) is a maximal set satisfying N1-N2. ■

When G has a Nash equilibrium but is unsolvable, there are multiple pairs of maximal sets (E_1, E_2) satisfying N1-N2. We do not have those problems in NM1-NM2 in Section 2.3, for which the greatest pair always exists and is nonempty. The reason for this difference may be the interactive nature of N1-N2, which is lacking in NM1-NM2.

For an unsolvable game G with a Nash equilibrium, there is no single definite recommended set of decisions and predictions based on N1-N2, even though the decision criterion and game structure are commonly understood between the players. Each maximal pair (E_1, E_2) satisfying N1-N2 may be a candidate, but it requires further information for the players to choose among them. Thus, N1-N2 alone is not sufficient to provide a definite recommendation for unsolvable games. Theorem 3.2 gives a demarcation line between the games with a definite recommendation and those without it.

One possible way to reach a recommendation for an unsolvable game is to impose an additional criterion, such as the symmetry requirement in Nash [20], so as to select a certain subset of Nash equilibria. The game of Table 2.2 is unsolvable, but it has a unique symmetric equilibrium $(\mathbf{s}_{11}, \mathbf{s}_{21})$. Hence, if we add the symmetry criterion, we convert an unsolvable game to a solvable game.

Another possible criterion is the Pareto-criterion. It may work to choose one subsolution for some class of games. For example, it is known that a finite game of strategic complementarity (or super modularity) has a Nash equilibrium in pure strategies, and under some mild condition, that if it has multiple equilibria, they are Pareto-ranked (see Vives [23] for an extensive survey of this theory and its applications). For those games, when there are multiple equilibria, each equilibrium constitutes a subsolution. However, when we add the Pareto-criterion, the subsolution which Pareto dominates the other subsolutions is chosen. Since a finite game version of this theory is not well known, we give a brief description of this theory in our context.

Assume that the strategy set S_i is linearly ordered so that S_i is expressed as $\{1, ..., \ell_i\}$ for $i = 1, 2$. Here, $S_1 \cap S_2 = \emptyset$ is violated but is recovered by a light change. We say that a game G has the *SC property* iff (1): for $i = 1, 2$, $h_i(s_i; s_j)$ is *concave* with respect to s_i, i.e., for all $s_i = 1, ..., l_i - 2$ and all $s_j \in S_j$

$$h_i(s_i + 1; s_j) - h_i(s_i; s_j) \geq h_i(s_i + 2; s_j) - h_i(s_i + 1; s_j); \tag{3.1}$$

and (2): $h_i(s_i; s_j)$ is *strategically complement*, i.e., for all $s_1 \in S_1 \backslash \{\ell_1\}$ and $s_2 \in S_2 \backslash \{\ell_2\}$,

$$h_i(s_i + 1; s_j) - h_i(s_i; s_j) \leq h_i(s_i + 1; s_j + 1) - h_i(s_i; s_j + 1). \tag{3.2}$$

Then, the following are more or less known results, but we give a proof for self-containedness[5].

Lemma 3.3. Let G be a game with the SC property.

(1): G has a Nash equilibrium in pure strategies.

(2): Suppose u-single peakedness, i.e., for each $i = 1, 2$ and $s_j \in S_j$, $h_i(s_i; s_j)$ has a unique maximum over S_i.Then, when G has multiple equilibria, they are linearly ordered with strict Pareto-dominance.

Proof. **(1)**: We will use Tarski's fixed point theorem: Let $(A, \leq)$ be a complete lattice, i.e., any subset of A has both infimum and supremum with respect to $\leq$. A function $\varphi : A \to A$ is called *increasing* iff $a \leq b$ implies $\varphi(a) \leq \varphi(b)$. Tarski's theorem states that φ is an increasing function on a complete lattice $(A, \leq)$ to itself, then φ has a fixed point. See Vives [23] (the Appendix) and Cousot-Cousot [6] for relevant concepts.

We define the partial order $\leq$ over $S_1 \times S_2$ by: $(s_1, s_2) \leq (s_1', s_2') \iff s_i \leq s_i'$ for $i = 1, 2$. Then, $(S_1 \times S_2, \leq)$ is a complete lattice. We also define the *(least) best-response function* $f : S_1 \times S_2 \to S_1 \times S_2$ as follows: for $i = 1, 2$ and $s_j \in S_j$,

$$f_i(s_j) = \min\{t_i : \text{Best}(t_i; s_j) \text{ holds}\}. \tag{3.3}$$

[5] Intervals of reals are typically adopted for these results. But Tarski's fixed point theorem is applied for the existence result in our case, too. In fact we can construct an algorithm to find a Nash equilibrium.

Now, $f(s_1, s_2) = (f_1(s_2), f_2(s_1))$ for each $(s_1, s_2) \in S_1 \times S_2$. We show that this f is increasing. Then, f has a fixed point (s_1^o, s_2^o), which is a Nash equilibrium.

Now, suppose $s_j < s_j'$. Let $f_i(s_j) = t_i$. By (3.3) and (3.2), we have $0 < h_i(t_i; s_j) - h_i(t_i - 1; s_j) \leq h_i(t_i; s_j') - h_i(t_i - 1; s_j')$. By (3.1), we have $0 < h_i(t_i; s_j') - h_i(t_i - 1; s_j') \leq h_i(k_i; s_j') - h_i(k_i - 1; s_j')$ for all $k_i \leq t_i$. Thus, $h_i(t_i; s_j') \geq h_i(k_i; s_j')$ for all $k_i \leq t_i$. This implies that player i's best response to s_j' is at least as small as t_i, i.e., $f_i(s_j') = t_i' \geq t_i$.

(2): Let $(s_1, s_2), (s_1', s_2')$ be two Nash equilibria with $s_i < s_i'$. By the monotonicity of f shown in (1), $s_j = f_j(s_i) \leq f_j(s_i') = s_j'$. If $s_j = s_j'$, then $h_i(\cdot; s_j)$ takes a maximum at s_i and s_i'. This is not allowed by u-single peakedness. ■.

When an SC game G with u-single peakedness has multiple equilibria, G is unsolvable by (2). However, if we add one criterion for player i's prediction/decision criterion, then we can choose one solution for any SC game with u-single peakedness. It may be better to state the result as a theorem.

Theorem 3.4. Let G be an SC game with u-single peakedness. Suppose that (E_1, E_2) and (E_1', E_2') satisfy N1-N2. Then, if for $i = 1, 2$, $h_i(s) \geq h_i(s')$ for some $s \in E_1 \times E_2$ and $s' \in E_1' \times E_2'$, then $E_1 \times E_2$ consists of the unique NE Pareto-dominating all other NE's.

Thus, one subsolution is chosen by adding the Pareto-criterion to N1-N2.

3.2. Rationalizable Strategies

The decision criterion for rationalizability theory, which formalizes the statements R1^o and R2^o in Section 1, is given as follows: for E_1 and E_2,

R1: for each $s_1 \in E_1$, Best$(s_1; s_2)$ holds for some $s_2 \in E_2$;

R2: for each $s_2 \in E_2$, Best$(s_2; s_1)$ holds for some $s_1 \in E_1$.

This criterion differs from N1-N2 only in that the quantifier "for all" before players' predictions in N1-N2 is replaced by "for some". In fact, R1-R2 is the pure-strategy version of the BP-property given by Bernheim [3] and Pearce [21]. The greatest pair (E_1, E_2) satisfying R1-R2 exists and coincides with the sets of rationalizable strategies $(R_1(G), R_2(G))$. A more general version of the following theorem is reported in Bernheim [3] (Proposition 3.1); we include the proof for self-containment.

Theorem 3.5 (Rationalizability): $(R_1(G), R_2(G))$ is the greatest pair satisfying R1-R2.

Proof. Suppose that (E_1, E_2) satisfies R1-R2. First, we show by induction that $E_1 \times E_2 \subseteq R_1^\nu(G) \times R_2^\nu(G)$ for all $\nu \geq 0$, which implies $E_1 \times E_2 \subseteq R_1(G) \times R_2(G)$. Since $R_i^0(G) = S_i$ for $i = 1, 2$, $E_1 \times E_2 \subseteq R_1^0(G) \times R_2^0(G)$. Now, suppose $E_1 \times E_2 \subseteq$

$R_1^\nu(G) \times R_2^\nu(G)$. Let $s_i \in E_i$. Due to the R1-R2, there is an $s_j \in E_j$ such that $\text{Best}(s_i; s_j)$ holds. Because $E_j \subseteq R_j^\nu(G)$, we have $s_j \in R_j^\nu(G)$. Thus, $s_i \in R_i^{\nu+1}(G)$.

Conversely, we show that $(E_1(G), E_2(G))$ satisfies R1-R2. Let $s_i \in R_i(G) = \bigcap_{\nu=0}^\infty R_i^\nu(G)$. Then, for each $\nu = 0, 1, 2, ...$, there exists $s_j^\nu \in R_j^\nu$ such that $\text{Best}(s_i; s_j^\nu)$ holds. Since S_j is a finite set, we can take a subsequence $\{s_j^{\nu_t}\}_{t=0}^\infty$ in $\{s_j^\nu\}_{\nu=0}^\infty$ such that for some $s_j^* \in S_j$, $s_j^{\nu_t} = s_j^*$ for all ν_t. Then, s_j^* belongs to $R_j(G) = \bigcap_{\nu=0}^\infty R_j^\nu(G)$. Also, $\text{Best}_i(s_i; s_j^*)$ holds. Thus, $(R_1(G), R_2(G))$ satisfies R1-R2. ■

Existence of a Theoretical Prediction: Theorem 3.5 and Lemma 2.2 imply that the greatest pair satisfying R1-R2 exists and consists of the sets of rationalizable strategies. Interchangeability is automatically satisfied by construction. In this respect, the rationalizability theory appears preferable to the Nash theory in that it avoids the issues due to emptiness or multiplicity of subsolutions. We take a different perspective to reverse this preference: Difficulties involved in the Nash theory identify situations where additional requirements other than N1-N2 are required for prediction/decision making. In this sense, the Nash theory is a more precise and potentially richer theory of *ex ante* decision making in an interactive situations.

Set-theoretical Relationship to the Nash Solutions: It follows from Theorem 3.5 that each strategy of a Nash equilibrium is a rationalizable strategy. Hence, the Nash solution, if it exists, is a subset of the set of rationalizable strategy profiles. However, the converse does not necessarily hold. Indeed, consider the game of Table 3.4, where the subgame determined by the $2nd$ and $3rd$ strategies for both players is the "matching pennies".

Table 3.4

	$\mathbf{s}_{21}$	$\mathbf{s}_{22}$	$\mathbf{s}_{23}$
$\mathbf{s}_{11}$	$(5, 5)$	$(-2, -2)$	$(-2, -2)$
$\mathbf{s}_{12}$	$(-2, -2)$	$(1, -1)$	$(-1, 1)$
$\mathbf{s}_{13}$	$(-2, -2)$	$(-1, 1)$	$(1, -1)$

This game has a unique Nash equilibrium, $(\mathbf{s}_{11}, \mathbf{s}_{21})$. Hence, the set consisting of this equilibrium is the Nash solution.

Both $\mathbf{s}_{11}$ and $\mathbf{s}_{21}$ are rationalizable strategies. Moreover, the other four strategies, $\mathbf{s}_{12}, \mathbf{s}_{13}$ and $\mathbf{s}_{22}, \mathbf{s}_{23}$ are also rationalizable: Consider $\mathbf{s}_{12}$. It is a best response to $\mathbf{s}_{22}$, which is a best response to $\mathbf{s}_{13}$, and $\mathbf{s}_{13}$ is a best response to $\mathbf{s}_{23}$, which is a best response to $\mathbf{s}_{12}$. That is, we have the following relations:

$$\text{Best}(\mathbf{s}_{12}; \mathbf{s}_{22}),\ \text{Best}(\mathbf{s}_{22}; \mathbf{s}_{13}), \text{Best}(\mathbf{s}_{13}; \mathbf{s}_{23}),\ \text{and}\ \text{Best}(\mathbf{s}_{23}; \mathbf{s}_{12}).$$

By Theorem 3.5, those four strategies are rationalizable. In sum, all the strategies are rationalizable in this game.

This example shows that even for solvable games, the Nash solution may differ from rationalizable strategies[6]. As we shall see later, the game of Table 3.4 becomes unsolvable if mixed strategies are allowed, while the rationalizable strategies remain the same.

3.3. Mixed Strategy Versions

Theorems 3.2 and 3.5 can be carried out in mixed strategies without much difficulty. The use of mixed strategies may give some merits and demerits to each theory. Here, we give comments on the mixed strategy versions of the two theories.

The mixed strategy versions can be obtained by extending the strategy sets S_1 and S_2 to the mixed strategy sets $\Delta(S_1)$ and $\Delta(S_2)$, where $\Delta(S_i)$ is the set of probability distributions over S_i. The notion of Nash equilibrium is defined in the same manner with the strategy sets $\Delta(S_1)$ and $\Delta(S_2)$: Once the Nash equilibrium is defined, the Nash solution, subsolution, etc. are defined in the same manner. However, the mixed strategy version of rationalizability requires some modification: A sequence of sets of strategies, $\{(\tilde{R}_1^\nu(G), \tilde{R}_2^\nu(G))\}_{\nu=0}^\infty$, is inductively defined as follows: for $i = 1,2$, $\tilde{R}_i^0(G) = S_i$, and for any $\nu \geq 1$,

$$\tilde{R}_i^\nu(G) = \{s_i : \text{Best}(s_i; s_j) \text{ holds for some } m_j \in \Delta(\tilde{R}_j^{\nu-1}(G))\}.$$

A pure strategy $s_i \in S_i$ is *rationalizable* iff $s_i \in \tilde{R}_i(G) = \bigcap_{\nu=0}^\infty \tilde{R}_i^\nu(G)$.

Requirements N1-N2 are modified by replacing S_i by $\Delta(S_i)$, $i = 1,2$;for $E_i \subseteq \Delta(S_i)$, $i = 1,2$,

N1^m: for each $m_1 \in E_1$, $\text{Best}(m_1; m_2)$ holds for all $m_2 \in E_2$,

N2^m: for each $m_2 \in E_2$, $\text{Best}(m_2; m_1)$ holds for all $m_1 \in E_1$.

Notice that N1^m-N2^m is the same as N1-N2 with different strategy sets. Moreover, Theorem 3.2 still holds without any substantive changes.

In a parallel manner, the mixed strategy version of rationalizability can also be obtained: for $E_i \subseteq \Delta(S_i)$, $i = 1,2$,

R1^m: for each $m_1 \in E_1$, $\text{Best}(m_1; m_2)$ holds for some $m_2 \in E_2$,

R2^m: for each $m_2 \in E_2$, $\text{Best}(m_2; m_1)$ holds for some $m_1 \in E_1$.

This is a direct counterpart of R1-R2 in a game with mixed strategies. In this case,

[6] When a 2-person game has no Nash equilibria, each player has at least two rationalizable strategies. If a player has a unique rationalizable strategy, it is a Nash strategy. Moreover, when each player has a unique rationalizable strategy, then the pair of them is a unique Nash equilibrium. Example 3.4 states that the converse does not necessarily hold.

a player is allowed to play mixed strategies. However, in the original version of rationalizability in Bernheim [3] and Pearce [21], the players are allowed to use pure strategies only; indeed, mixed strategies are interpreted as a player's beliefs about the other player's decisions. We can reformulate R1^m-R2^m based on this interpretation of mixed strategies: In R1^m, the first occurrence of m_1 is replaced by a pure strategy in the support of E_1, and R2^m is modified in a parallel manner. This reformulation turns out to be mathematically equivalent to R1^m-R2^m.

With the replacement of R1-R2 by R1^m-R2^m in Theorem 3.5, the following statement holds:

Theorem 3.5′. $(\Delta(\tilde{R}_1(G)), \Delta(\tilde{R}_2(G)))$ is the greatest pair satisfying R1^m-R2^m.

A simple observation is that a rationalizable strategy in the pure strategy version is also a rationalizable strategy in the mixed strategy version. Similarly, since a Nash equilibrium in pure strategies is also a Nash equilibrium in mixed strategies, it may be conjectured that if a game G has the Nash solution E in the pure strategies, it might be a subset of the Nash solution in mixed strategies. In fact, this conjecture is answered negatively.

Consider the game of Table 3.4. This game has seven Nash equilibria in mixed strategies:

$$((1,0,0),(1,0,0)),\ ((0,\tfrac{1}{2},\tfrac{1}{2}),(0,\tfrac{1}{2},\tfrac{1}{2})),\ ((\tfrac{4}{18},\tfrac{7}{18},\tfrac{7}{18}),(\tfrac{4}{18},\tfrac{7}{18},\tfrac{7}{18}))$$
$$((\tfrac{1}{8},\tfrac{7}{8},0),(\tfrac{3}{10},\tfrac{7}{10},0)),\ ((\tfrac{1}{8},0,\tfrac{7}{8}),(\tfrac{3}{10},0,\tfrac{7}{10})),\ ((\tfrac{3}{10},\tfrac{7}{10},0),(\tfrac{1}{8},0,\tfrac{7}{8})),\ ((\tfrac{3}{10},0,\tfrac{7}{10}),(\tfrac{1}{8},\tfrac{7}{8},0)).$$

This set does not satisfy interchangeability (2.2). For example, $((1,0,0),(1,0,0))$ and $((0,\frac{1}{2},\frac{1}{2}),(0,\frac{1}{2},\frac{1}{2}))$ are Nash equilibria, but $((0,\frac{1}{2},\frac{1}{2}),(1,0,0))$ is not a Nash equilibrium. Thus, (2.2) is violated, and the set of all mixed strategy Nash equilibria is not the Nash solution. This result depends upon the choice of payoffs: In Table 3.5, (s_{11}, s_{21}) is a unique Nash equilibrium even in mixed strategies, while all pure strategies are still rationalizable.

Table 3.5

	$\mathbf{s}_{21}$	$\mathbf{s}_{22}$	$\mathbf{s}_{23}$
$\mathbf{s}_{11}$	$(5,5)$	$(\frac{1}{2},\frac{1}{2})$	$(\frac{1}{2},\frac{1}{2})$
$\mathbf{s}_{12}$	$(\frac{1}{2},\frac{1}{2})$	$(1,-1)$	$(-1,1)$
$\mathbf{s}_{13}$	$(\frac{1}{2},\frac{1}{2})$	$(-1,1)$	$(1,-1)$

4. Evaluations of N1-N2 and R1-R2 as Prediction/Decision Criteria

Our unified approach pinpoints the difference between the Nash and rationalizability theories: the choice of quantifier "for all" or "for some" for each player's predictions.

Here we evaluate this difference reflecting upon on the conceptual bases of game theory. We take Johansen's [11] argument on the Nash theory as our starting point. Then, we make comparisons between the two theories by considering two methodological principles: the free-will postulate and complete determinism. We also consider multiplicity in prediction/decision criteria and how we should take it in our research activities.

4.1. Johansen's Argument

Johansen [11] gives the following four postulates for prediction/decision making in games and asserts that the Nash noncooperative solution is derived from those postulates for solvable games. For this, he assumes (p.435) that the game has the unique Nash equilibrium, but notes (p.437) that interchangeability is sufficient for his assertion.

Postulate J1. A player makes his decision $s_i \in S_i$ on the basis of, and only on the basis of information concerning the action possibility sets of two players S_1, S_2 and their payoff functions h_1, h_2.

Postulate J2. In choosing his own decision, a player assumes that the other is rational in the same way as he himself is rational.

Postulate J3. If any[7] decision is a rational decision to make for an individual player, then this decision can be correctly predicted by the other player.

Postulate J4. Being able to predict the actions to be taken by the other player, a player's own decision maximizes his payoff function corresponding to the predicted actions of the other player.

Notice that the term "rational" occurs in J2 and J3, and "payoff maximization" in J4. The term "rational" in Johansen's argumentation is broader than its typical meaning in the game theory literature referring to "payoff maximization." Indeed, He regards these four postulates together as an attempt to define "rationality"; "payoff maximization" is only one component of "rationality". We may further disentangle "rationality" using two concepts: the first is the prediction/decision criterion, which includes J4 as its component, and the second is the logical ability. Here we explain the four postulates with these two concepts.

Postulate J1 is the starting point for his consideration of *ex ante* decision making. Postulate J2 requires the decision criterion be symmetric between the decision maker and the other player in his mind. Postulate J3 requires each player's prediction about the other's decision be correctly made. Postulate J4 corresponds to the payoff maximization requirement. In the following, we first elaborate Postulates J2 and J3, and then use J1-J4 as a reference point for our critical comparisons between N1-N2 and R1-R2.

[7] This "any" was "some" in Johansen's orginal Posutlate 3. According to logic, this should be "any". However, this is expressed as "some" by many scientists (even mathematicians).

Postulate J2 implies that from player 1's perspective, the decision criterion has to be symmetric between the two players. In our context, this is interpreted as applied to the choice of prediction/decision criterion. Both N1-N2 and R1-R2 satisfy this symmetric requirement. The combination N1-NM2 discussed in Section 2 violates symmetry, and so does N1-R2, which will be further discussed in Section 4.3.

Postulate J3 is interpreted in the following manner: First, player 1 thinks about the whole situation, taking player 2's criterion as given, and makes inferences from this thinking. Based on such inferences, player 1 makes a prediction about 2's decisions. This prediction is *correct* in the sense that player 1's prediction criterion is the same as 2' decision criterion and 1 has the same logical ability as player 2's. In this sense, predictability in J3 is a result of a player's contemplation of the whole interactive situation[8]. In this reasoning, "rationality" in J3 emphasizes symmetry in players' interpersonal logical abilities, while that in J2 emphasizes symmetry in his prediction/decision criterion.

Postulates J1-J3 are compatible with N1-N2 and R1-R2. Only Postulate J4 makes a distinction between the Nash theory and rationalizability theory. If we read Postulate J4 in light of his assertion that interchangeability is a sufficient condition for J1-J4 to lead to the Nash solution, we can interpret J4 as adopting "for all" predicted actions of the other player's possible decisions.

Johansen [11] does not give a formal analysis of his postulates. Our N1-N2 may be regarded as a formulation of these postulates in the language of classical game theory. In this sense, Theorem 3.2 formalizes Johansen's assertion that the Nash solution is characterized by J1-J4. If we modify Postulate J4 so that the "for all" requirement is replaced by the "for some" requirement, Theorem 3.5 for R1-R2 would be a result. We still need to discuss what are bases for the choice of "for all" or "for some".

4.2. The Free-will Postulate vs. Complete Determinism

Here, we evaluate the difference between N1-N2 and R1-R2, based on two conflicting meta-theoretical principles: the free-will postulate and complete determinism.

The Free-will Postulate: This states that players have freedom to make choices following their own will. Whenever the social science involves value judgements for individual beings and/or the society, they rely on the free-will postulate as a foundation[9]. In a single person decision problem, utility maximization may effectively void this postulate[10]. However, in an interactive situation, even if both players are very smart, it

[8] Bernheim's [4], p.486, interpretation of J3 in his criticism against these postulates is quite different from our reasoning. In his framework, predictability simply means that the belief about the other player's action, which is exogenously given, coincides with the actual action.

[9] The free-will postulate is needed for deontic concepts such as responsibility for individual choice and also for individual and social efforts for future developments.

[10] This does not imply that utility maximization even for 1-person problem violates the free-will postualte; he has still freedom to ignore his utility.

is still possible that individual decision making, based on utility maximization alone, may not result in a unique decision. This is first argued in Morgenstern [18], using the paradox of Moriarty chasing Holmes. This is still a central problem in game theory; the free-will postulate constitutes an important part of this problem. In this respect, the free-will postulate still remains relevant to game theory.

Consider applications of the postulate at two different layers in terms of interpersonal thinking:

(i): It is applied by the outside observer to the (inside) players;

(ii): It is applied by an inside player to the other player.

In application (i), the outside theorist respects the free will of each player; the theorist can make no further refinement than the inside player. This corresponds to the greatestness requirement for (E_1, E_2) in Theorems 3.2.(1) and Theorem 3.5. In (ii), when one player has multiple predictions about the other's decisions, the free-will postulate, applied to interpersonal decision making, requires the player take all possible predictions into account. N1-N2 is consistent with this requirement in that it requires each player's decision be optimal against all predictions[11].

Criterion R1-R2 involves some subtlety in judging whether it is consistent with application (ii). The main difficulty is related to the interpretation of "for some" before the prediction about the other's decision. This leads us to another view, "complete determinism."

Complete Determinism: The quantifier "for some" in R1-R2 has two different interpretations:

(a): it requires only the mere existence of a rationalizing strategy;

(b): it suggests a specific rationalizing strategy predetermined for some other reason.

Interpretation (a) is more faithful to the mathematical formulation of R1-R2 as a decision criterion. If we accept (a), then arbitrariness of the rationalizing strategy shows no respect to the other player's free will, but we would not find a serious difficulty in R1-R2 with the free-will postulate in that R1-R2 is a prediction/decision criterion adopted by a player. However, this reminds us Aesops' *sour grapes* that the fox finds one convenient reason to persuade himself: For R1-R2, it suffices to find any rationalizing strategy. This interpretation of "rationalization" is at odds with the purpose of a theory of *ex ante* decision-making for games, since such a theory is supposed to provide a rationale for players' decisions as well as predictions. Interpretation (a) requires no rationale for each specific rationalizing strategy.

Interpretation (b) resolves the arbitrariness in (a): According to (b), there are some further components, not explicitly included in the game description G and R1-R2, that

[11] There are many other criteria consistent with the requirement. For example, player 1 uses the maximin criterion to choose his action against E_2. Another possibility is to put equal probability on each action in E_2 and to apply expected utility maximization.

determine a specific rationalizing strategy. However, a specific rationalizing strategy for each step has to be uniquely determined, for otherwise the player would have to arbitrarily choose among different strategies or to look for a further reason to choose some of them. Thus, interpretation (b) violates Johansen's postulate J1.

Interpretation (b) deserves a further analysis, since it is related to complete determinism, which has been regarded as very foundational in natural sciences. To determine a specific rationalizing strategy, one possibility is to refer to a full description of the world including players' mental states; this presumes some form of determinism. We consider only complete determinism for simplicity. Such a full description in a situation with two persons may require an infinite hierarchy of beliefs. Indeed, there is a literature, beginning from Aumann [2][12], to justify the rationalizability theory or alike along this line (see Tan-Werlang [22]).

Complete determinism is incompatible with the free-will postulate in that it contains no room for decision; *ex ante* decision making is an empty concept from this perspective. From this view, R1-R2 is regarded as a partial description of a law of causation.

Except for conflicting against the free-will postulate, complete determinism may not be very fruitful as a methodology for social science in general, which is aptly described by Hayek [8], Section 8.93: *"Even though we may know the general principle by which all human action is causally determined by physical processes, this would not mean that to us a particular human action can ever been recognizable as the necessary result of a particular set of physical circumstances."*

Complete determinism is justified only because of its non-refutability by withdrawing from concrete problems into its own abstract world. In fact, neither complete determinism nor the free-will postulate can be justified by its own basis. Either should be evaluated with coherency of the entire scope and the scientific and/or theoretical discourse.

Our conclusion is that the free-will postulate is needed for the perspective of social sciences, and complete determinism has no such a status in social sciences. The Nash noncooperative theory is constructed coherently with the free-will postulate, but the rationalizability theory meets a great difficulty to reconcile with it.

4.3. Prediction/Decision Criteria

The characterizations of the Nash and rationalizability theories in terms of prediction/decision criteria are helpful to find their differences as well as to understand Jo-

[12] In the problem of common knowledge in the information partition model due to Robert Aumann, the information partitions themselves are assumed to be common knowledge. He wrote in [1], p.1237: *"Included in the full description of a state ω of the world is the manner in which information is imparted to the two persons"*. This can be interpreted as meaning that the primitive state ω includes every information. A person receives some partial information about ω, but behind this, everything is predetermined. This view is shared with Harsanyi [7] and Aumann [2].

hansen's argument and *vice versa*. However, these characterization results also introduce a new problem: Among all possible prediction/decision criteria, why should we focus particularly on the Nash theory or the rationalizability theory? Here we consider a few exmaples of prediction/decision criteria and their resulting outcomes.

Relativistic View: It may be the case that people adopt different prediction/decision criteria. In addition to N1-N2 and R1-R2, as already indicated, NM1-NM2, N1-NM2 and N1-R2 are also possible candidates, among others. Even restricting our focus to N1-N2 and R1-R2, it is natural to ask why we avoid a mixture, such as N1-R2, of those criteria. Moreover, this combination actually generates a different outcome either from N1-N2 or R1-R2. Consider the game which is obtained from Table 3.4 by changing the payoffs in the first row and first column.

Table 4.1

	$\mathbf{s}_{21}$	$\mathbf{s}_{22}$	$\mathbf{s}_{23}$
$\mathbf{s}_{11}$	$(1,1)^{NE}$	$(1,1)^{NE}$	(0,0)
$\mathbf{s}_{12}$	(1,0)	$(1,-1)$	$(-1,1)$
$\mathbf{s}_{13}$	(0,0)	$(-1,1)$	$(1,-1)$

For this game, we can calculate the greatest pairs (E_1, E_2) satisfying N1-N2, R1-R2 and N1-R2 as follows:

N1-N2:	$(\{\mathbf{s}_{11}\}, \{\mathbf{s}_{21}, \mathbf{s}_{22}\})$
R1-R2:	$(\{\mathbf{s}_{11}, \mathbf{s}_{12}, \mathbf{s}_{13}\}, \{\mathbf{s}_{21}, \mathbf{s}_{22}, \mathbf{s}_{23}\})$
N1-R2:	$(\{\mathbf{s}_{11}, \mathbf{s}_{12}\}, \{\mathbf{s}_{21}, \mathbf{s}_{22}\})$

As soon as we start considering different combinations, they could provide actually different recommendations.

This relativistic view may turn our target problem into an empirical study of such criteria in real societies. However, a prediction/decision criterion itself is still an analytic concept that serves as a benchmark to understand the prediction/decision-making process in interactive situations. From this perspective, the focus should rather be a study of the underlying structures and rationales for those criteria; if a criterion is incoherent with other bases, people will eventually avoid it. The goal of such study is then to separate some criteria from others, even if we take the relativistic view that people follow diverse ways of prediction/decision making. For example, Johansen's postulate J2 accepts N1-N2 and R1-R2 but rejects N1-R2 as a legitimate criterion.

In this paper we analyze two specific criteria, N1-N2 and R1-R2, taking Johansen's postulates and the current literature of game theory as given. However, once we enter the relativistic world of prediction/decision criterion, we may require rationales for the postulates such as J2. A full analysis, which would involve broader conceptual bases for prediction/decision criterion and more explicit study of the underlying thought processes for prediction/decision making, is way beyond the current research. Nevertheless, Section 5 mentions a further research possibility on these problems as a continuation of the present paper.

5. Conclusions

5.1. Summary: the Unified Framework and Parallel Derivations

We presented the unified framework and parallel derivations of the Nash noncooperative solutions and rationalizable strategies. The difference between them is pinpointed to be the choice of the quantifier "for all" or "for some" for predictions about the other player's possible decisions. In Section 4, we discussed various conceptual issues by viewing the quantifier "for all" and "for some" from the perspectives of Johansen's postulates, the free-will postulate vs. complete determinism, and prediction/decision criteria.

Comparisons with Johansen's postulates help us well understand our unified framework and derivations. The argument from the perspective of the free-will postulate vs. complete determinism concludes that the Nash theory is more coherent to social sciences as a whole than the rationalizability theory. Nevertheless, as a descriptive concept, it would be possible for some people to use a criterion with "for some" for their decision making. Reflections upon our approach in terms of prediction/decision criteria manifest that vast aspects of prediction/decision making in social context are still hidden.

One such hidden problem is the treatment of the assumption of common knowledge. We started this paper with the quotation to Mas-Colell *et al.* [15] about the standard interpretation of rationalizability theory in terms of common knowledge. It is also a common interpretation that the Nash theory requires the common knowledge of the game structure. In this paper, the notion of common knowledge or even knowledge/beliefs remains interpretational. To study the thought process for prediction/decision making explicitly, we meet new issues and additional framework is necessary. The following section discusses these issues.

5.2. Continuation: Thought Process for Prediction/Decision Making

The present paper employs the standard game theory language. In this language, many essential elements remain informal and hidden, including a player's beliefs or knowledge. Those elements are essential for understanding the thought process for prediction/decision-making. Here we consider only N1-N2, but the argument is also applicable to R1-R2.

Prediction Making (Putting Oneself in the Other's Shoes): N1-N2 is understood as describing both prediction making and decision making: From player 1's perspective, E_1 in N1 is his decision variable, while E_2 in N1 is his prediction variable.

Here, player 1 puts himself into player 2's shoes to make predictions. In fact, this argument could not stop here; by putting himself in 2's shoes, 1 needs to think about 2's predictions about 1's decisions. Continuing this argument *ad infinitum,* we meet the infinite regress described in Diagram 5.1, which is made from the viewpoint of player 1. A symmetric argument from player 2's viewpoint can be constructed.

Double Uses of N1-N2: In the infinite regress, N1 is a decision criterion for 1 and is a prediction criterion for 2, while N2 is a decision criterion for 2 and a prediction criterion for 1. Thus, N1 and N2 are used both as decision and prediction criteria. This double makes the infinite regress in Diagram 4.1 collapse into a system of simultaneous equations described by Diagram 5.2. Theorem 3.2 solves this system of equations.

Diagram 5.1

N1		N1		N1		$\cdots$
$\downarrow$	$\nearrow$	$\downarrow$	$\nearrow$	$\downarrow$	$\nearrow$	
N2		N2		N2		

$\Longrightarrow$

Diagram 5.2

N1
$\downarrow\ \uparrow$
N2

The language of classical game theory is incapable to explicitly distinguish between player 1's and 2's perspectives; as a result, many foundational problems can only be discussed at interpretational levels. One way to formalize those issues is to reformulate the above problem in the epistemic logic framework. Then, we can avoid the collapses from Diagram 5.1 into Diagram 5.2, and explicitly discuss the relationship between the above infinite regress and the common knowledge of N1-N2. In doing so, we will be able to evaluate the standard interpretations, such as the quotation from Mas-Colell *et al.* [15] in Section 1, of the rationalizability theory as well as the Nash theory. Also, we can more explicitly discuss Johansen's [11] argument. The research on these problems will be undertaken in the companion paper [9].

References

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

[2] Aumann, R. J. (1987), Correlated Equilibrium as an Expression of Bayesian Rationality? *Econometrica* 55, 1-18.

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

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

[5] Brandenburger, A., and E. Dekel (1993), Hierarchies of Beliefs and Common Knowledge, *Journal of Economic Theory*, 59, 189-198.

[6] Cousot, P., and R. Cousot, (1979), Constructive Versions of Tarski's fixed Point Theorem, Pacific Journal of Mathematics 82, 43-57.

[7] Harsanyi, J. C. (1967/68), Games with Incomplete Information Played by Bayesian Players, I-III, *Management Science* 14, 159-182, 320-334, 486-502.

[8] Hayek, F. A., (1952), *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology*, The University of Chicago Press, (1952).

[9] Hu, T.-W., and M. Kaneko (2012), Epistemic Evaluations of the Nash Theory and Rationalizability, being prepared.

[10] Jansen, M. J. M., (1981), Maximal Nash Subsets for Bimatrix Games, *Naval Research Logistics Quarterly* 28, 147-152.

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

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

[13] Kaneko, M., (2004), *Game Theory and Mutual Misunderstanding*, Springer, Heidelberg.

[14] Kaneko, M., and N.-Y. Suzuki, (2002), Bounded interpersonal inferences and decision making, *Economic Theory* 19 (2002), 63-103.

[15] Mas-Colell, A., M. D. Whinston, and J. R. Green, (1995), *Mircoeconomic Theory*, Oxford University Press, New York.

[16] Mertens, J., and S. Zamir (1985), Formulation of Bayesian analysis for games with incomplete information, *International Journal of Game Theory* 14, 1-29.

[17] Millham, C. B., (1974), On Nash Subsets of Bimatrix Games, *Naval Research Logistics Quarterly* 21, 307-317.

[18] Morgenstern, O., (1935), Perfect Foresight and Economic Equilibrium. Zeitschrift für Nationalökonomie 6, 337-357.

[19] Nash, J. F., (1950), *Non-cooperative Games,* Ph.D. dissertation, Dept. Mathematics, Princeton University.

[20] Nash, J. F., (1951), Non-cooperative Games, *Annals of Mathematics* 54, 286-295.

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

[22] Tan T. C., and S. R. D. C. Werlang, (1988), The Bayesian Foundations of Solution Concepts of Games, *Journal of Economic Theory* 45, 370-391.

[23] Vives, X., Complementarities and Games: New Developments, (2005), *Journal of Economic Literature* XLIII, 437-479.

[24] von Neumann, J. and O. Morgenstern, (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton.