Characterizing Permissibility and Proper Rationalizability by Incomplete Information*

EPICENTER Working Paper No.14 (2018)

EPICENTER
Research Center for
Epistemic Game Theory

Shuige Liu†

January 20, 2018

Abstract

We characterize permissibility and proper rationalizability within an incomplete information framework. We define the lexicographic epistemic model for a game with incomplete information, and show that a choice is permissible (properly rationalizable) within a complete information framework if and only if it is optimal for a belief hierarchy within the corresponding incomplete information framework that expresses common full belief in caution, primary belief in the opponent's utilities nearest to the original utilities (the opponent's utilities are centered around the original utilities), and a best (better) choice is supported by utilities nearest (nearer) to the original ones.

JEL Classification: C72

Keywords: *epistemic game theory*, *incomplete information*, *lexicographic belief*, *permissibility*, *proper rationalizability*

1. Introduction

The purpose of noncooperative game theory is to study an individual's decision making in an interactive situation. Since in such a situation one's payoff is not completely determined by her own choice, to make a decision she needs to form a belief on every other participant's choice, on every other participant's belief on every other's choice, and so on. Studying the structure of those belief hierarchies and choices supported by a belief hierarchy satisfying some particular conditions opened up a field called *epistemic game theory.* See Perea [15] for a textbook on this field.

*The author would like to express her gratitude to Andrés Perea for his thorough reading of the early versions of this paper and his numerous helpful comments and suggestions. She also would like to thank János Flesch, Funaki Yukihiko, Christian Bach, Abraham Neyman, Zsombor Z. Méder, and Dmitriy Kvasov for their valuable discussions and encouragements. She thanks all teachers and students in the 4th Epicenter Spring Course on Epistemic Game Theory in Maastricht University for their inspiring teaching and stimulating discussions. She gratefully acknowledge the support of Grant-in-Aids for Young Scientists (B) of JSPS No.17K13707 and Grant for Special Research Project No. 2017K-016 of Waseda University.

†Faculty of Political Science and Economics, Waseda University, 1-6-1 Nishi-Waseda, Shinjuku-Ku, 169-8050, Tokyo, Japan (shuige_liu@aoni.waseda.jp).

In epistemic game theory, various concepts have been developed to describe some specific belief structures. One is *lexicographic belief* (Blume et al. [5], [6]). A lexicographic belief describes a player's subjective conjecture about the opponents' behavior by a sequence of probability distributions over other participants' choices and types, which is different from the adoption of a single probability distribution in a standard probabilistic belief. The interpretation of a lexicographic belief is that every choice-type pair in the sequence is considered to be possible, while a pair occurring ahead in the sequence is deemed *infinitely more likely* than one occurring later. Several concepts have been developed by putting various conditions on lexicographic beliefs intended to capture different types of reasoning about the opponents' behavior. Permissibility and proper rationalizability are two important and interrelated concepts among these.

Permissibility originated from Selten [20]'s perfect equilibrium. It is defined and studied from an epistemic viewpoint by using lexicographic belief in Brandenburger [10][1]. Permissibility is based on two notions: *caution* and *primary belief in the opponents' rationality.* A lexicographic belief is said to be cautious if it does not exclude any choice of the opponents; it is said to primarily believe in the opponents' rationality (Perea [15]) if its first level belief only deems possible those choice-type pairs where the choice is optimal under the belief of the paired type.

Proper rationalizability originated from Myerson [14]'s proper equilibrium which is intended to be a refinement of perfect equilibrium. It is defined and studied in Schuhmacher [19] and Asheim [1] as an epistemic concept. Proper rationalizability shares with permissibility the notion of caution while, instead of primary belief in the opponents' rationality , it is based on a stronger notion called *respecting the opponents' preferences* which means that a "better" choice always occurs in front of a "worse" choice in the lexicographic belief.

We explain these two concepts by an example. Consider a game where player 1 has strategies A and B and player 2 has strategies C, D, and E. Player 2's utility function u_2 is as follows:

u_2	C	D	E
A	3	2	1
B	3	2	1

Consider a lexicographic belief of player 1 on player 2's choices. Caution requires that all three choices of player 2 occur in that belief. Since C is player 2's most preferred choice, primarily believing in player 2's rationality requires that only choice C can be put in the first level of that belief. On the other hand, since C is preferred to D and D is preferred to E for player 2, a lexicographic belief of player 1 respecting 2's preferences should deem C infinitely more likely than D and D infinitely more likely than E, that is, put C before D and D before E in the lexicographic belief.

One motivation for the development of a lexicographic belief is to alleviate the tension between caution and rationality (Blume et al. [5], Brandenburger [10], Börgers [8], Samuelson [18], Börgers and Samuelson [9]). Permissibility and proper rationalizability tried to solve that tension by sacrificing rationality in different senses. That is, though permissibility requires that the first level belief contains only rational choices and proper rationalizability requires that choices should be ordered according to the "level" of rationality, both allow occurences of irrational choices because of caution. This sacrifice of rationality brought some conceptual inconvenience since rationality is a basic assumption in game theory and is reasonable to be adopted as a criterion for each player's belief.

Actually, there is an approach which solves the tension without sacrificing rationality: using an incomplete information framework. That is, instead of considering the uncertainty about opponents' rationality within a complete information framework, we take the uncertainty about the opponents' utility functions and consider types within the incomplete information framework.

[1] An alternative approach without using lexicographic belief is given by Börgers [8].

Then the occurrence of a irrational choice can be explained as that the "real" utility function of an opponent is different from the original one. Both permissibility and proper rationalizability can be characterized within an incomplete information framework. This is the basic idea of this paper.

We use the above example to explain this idea. As mentioned there, though only choice C is rational for player 2, caution requires all three choices C, D, and E to occur in player 1's belief. In a complete information framework, the occurrences of D and E are explained by player 2's irrationality (i.e., "trembling hand"). In contrast, within an incomplete information framework they are explained by the possibility that the "real" utility function of player 2 is not u_2 but v_2 or v_2' as follows:

v_2	C	D	E
A	2	3	1
B	2	3	1

,

v_2'	C	D	E
A	2	1	3
B	2	1	3

Choice D is optimal in v_2 and E is optimal in v_2'. In this way, uncertainty about the opponent's rationality within a complete information framework is transformed into uncertainty about the opponent's real utility function within an incomplete information framework. It can be seen that primary belief in the opponent's rationality in complete information framework is equivalent to the condition that one deems u_2 or a utility function "very similar" to u_2 infinitely more likely to be the real utility function of player 2 than v_2 and v_2', and respecting the opponent's preferences is equivalent to the condition that those alternative utility functions should be ordered by their "similarity" to u_2.

In this paper, we study these equivalences formally for 2-person static form games and provide a characterization of permissibility and proper rationalizability within an incomplete information framework. First, we define the lexicographic epistemic model of a game with incomplete information. Then we show that a choice is permissible (properly rationalizable) within a complete information framework if and only if it is optimal for a belief hierarchy within the corresponding incomplete information framework that expresses common full belief in caution, primary belief in the opponent's utilities nearest to the original utilities (the opponent's utilities are centered around the original utilities), and a best (better) choice is supported by utilities nearest (nearer) to the original ones.

Within the complete information framework, permissibility is weaker than proper rationalizability. This is reflected in our characterization of them within the incomplete information framework: permissibility shares caution with proper rationalizability while the other two conditions of the former are weaker versions of those of the latter.

It should be noted that rationality does not appear in the condition of characterizations. Nevertheless, in our proof we will construct incomplete information models with types satisfying all the conditions as well as rationality. In Section 4.3 we will also give a model with types which satisfies all conditions but does not satisfy rationality. These show that, in contrast to the inconsistency of caution and rationality within the complete information framework, in the incomplete information one the two are logically independent and consistent; we do not need to sacrifice one to save the other. Further, in Section 4.6 we will provide an alternative way to characterize permissibility by using rationality and weak caution.

This paper is not the first one characterizing concepts in epistemic game theory within an incomplete information framework. Perea and Roy [17] characterized ε-proper rationalizability in this approach by using a standard epistemic model without lexicographic beliefs. They showed that a type in a standard epistemic model with complete information expresses common full belief in caution and ε-trembling condition if and only if there is a type in the corresponding model with incomplete information sharing the same belief hierarchy with it which expresses common belief in caution, ε-centered belief around the original utilities u, and belief in rationality under the closest utility function. Since each properly rationalizable choice is the limit of a sequence of ε-proper rationalizable ones, the conditions adopted in their characterizations are very useful for us. Two conditions in our characterization of proper rationalizability, that is, caution and u-centered belief, are faithful translations of their conditions into lexicographic model. However, the most critical condition in their characterization, that is, belief in rationality under the closest utility function, is impossible to be adopted here. The reason is, as will be shown in Section 2.2, that a nearest utility function making a choice optimal does not always exist in lexicographic models. This is a salient difference between standard probabilistic beliefs and lexicographic ones. We define a weaker condition called "a better choice is supported by utilities nearer to the original one" and show that it can be used to characterize proper rationalizability.

Another essential difference between Perea and Roy [17] and this paper is in the way of proof. Equivalence of belief hierarchies generated by types in models with complete and incomplete informations and type morphisms (Böge and Eisele [7], Heifetz and Samet [13], Perea and Kets [16]) play an important role in Perea and Roy [17]'s proof. In contrast, our proofs are based on constructing a specific correspondence between the two models. We show that conditions in a type of one model implies that appropriate conditions are satisfied in the corresponding type in the constructed model. Equivalence of hierarchies follows directly by construction. Our construction can also be used in proving Perea and Roy [17]'s Theorem 6.1. Further, as will be discussed in Section 4.6, our construction shows that rationality is separable from other conditions in characterizing proper rationalizability. This confirms the consistency of caution and rationality within an incomplete information framework.

Our results, as well as Perea and Roy [17]'s, also provide insights in decision theory and general epistemology. They imply that any choice permissible or properly rationalizable within a complete information framework is also optimal for a belief satisfying some reasonable conditions within an incomplete information framework, and vice versa. In other words, by just looking at the outcome, it is impossible to know the accurate epistemic situation behind the choice, that is, whether it is because of players' uncertainty about the opponents' rationality or uncertainty about what are the real utilities of the opponents.

This paper is organized as follows. Section 2 defines permissibility and proper rationalizability in epistemic models with complete information and introduces the lexicographic epistemic model with incomplete information. Section 3 gives the two characterization results and their proofs. Section 4 gives some concluding remarks. Section 5 contains the proofs of all lemmas.

2. Models

2.1. Complete information model

In this subsection, we give a survey of lexicographic epistemic model with complete information. For details, see Perea [15], Chapters 5-6.

Consider a finite 2-person static game $\Gamma = (C_i, u_i)_{i\in I}$ where $I = \{1, 2\}$ is the set of players, C_i is the finite set of choices and $u_i : C_1 \times C_2 \to \mathbb{R}$ is the utility function for player $i \in I$. In the following sometimes we denote $C_1 \times C_2$ by C. We assume that each player has a lexicographic belief on the opponent's choices, a lexicographic belief on the opponent's lexicographic belief on her, and so on. This belief hierarchy is described by a lexicographic epistemic model with types.

Definition 2.1 (Epistemic model with complete information). Consider a finite 2-person static game $\Gamma = (C_i, u_i)_{i\in I}$. A finite *lexicographic epistemic model* for Γ is a tuple $M^{co} = (T_i, b_i)_{i\in I}$ where

(a) T_i is a finite set of types, and

(b) b_i is a mapping that assigns to each $t_i \in T_i$ a lexicographic belief over $\Delta(C_j \times T_j)$, i.e., $b_i(t_i) = (b_{i1}, b_{i2}, ..., b_{iK})$ where $b_{ik} \in \Delta(C_j \times T_j)$ for $k = 1, ..., K$.

Consider $t_i \in T_i$ with $b_i(t_i) = (b_{i1}, b_{i2}, ..., b_{iK})$. Each b_{ik} $(k = 1, ..., K)$ is called t_i's *level-k belief.* For each $(c_j, t_j) \in C_j \times T_j$, we say t_i *deems* (c_j, t_j) *possible* iff $b_{ik}(c_j, t_j) > 0$ for some $k \in \{1, ..., K\}$. We say t_i *deems* $t_j \in T_j$ *possible* iff t_i deems (c_j, t_j) possible for some $c_j \in C_j$. For each $t_i \in T_i$, we denote by $T_j(t_i)$ the set of types in T_j deemed possible by t_i. A type $t_i \in T_i$ is *cautious* iff for each $c_j \in C_j$ and each $t_j \in T_j(t_i)$, t_i deems (c_j, t_j) possible. That is, t_i takes into account each choice of player j for every belief hierarchy of j deemed possible by t_i.

For each $c_i \in C_i$, let $u_i(c_i, t_i) = (u_i(c_i, b_{i1}).,.., u_i(c_i, b_{iK}))$ where for each $k = 1, ..., K$, $u_i(c_i, b_{ik}) := \Sigma_{(c_j,t_j)\in C_j\times T_j} b_{ik}(c_j, t_j) u_i(c_i, c_j)$, that is, each $u_i(c_i, b_{ik})$ is the expected utility for c_i over b_{ik} and $u_i(c_i, t_i)$ is a vector of expected utilities. For each $c_i, c'_i \in C_i$, we say that t_i *prefers* c_i *to* c'_i, denoted by $u_i(c_i, t_i) > u_i(c'_i, t_i)$, iff there is $k \in \{0, ..., K-1\}$ such that the following two conditions are satisfied:

(a) $u_i(c_i, b_{i\ell}) = u_i(c'_i, b_{i\ell})$ for $\ell = 0, ..., k$, and

(b) $u_i(c_i, b_{i,k+1}) > u_i(c'_i, b_{i,k+1})$.

We say that t_i *is indifferent between* c_i *and* c'_i, denoted by $u_i(c_i, t_i) = u_i(c'_i, t_i)$, iff $u_i(c_i, b_{ik}) = u_i(c'_i, b_{ik})$ for each $k = 1, ..., K$. It can be seen that the preference relation on C_i under each type t_i is a linear order. c_i is *rational* (or *optimal*) for t_i iff t_i does not prefer any choice to c_i. A type $t_i \in T_i$ *primarily believes in the opponent's rationality* iff t_i's level-1 belief only assigns positive probability to those (c_j, t_j) where c_j is rational for t_j. That is, at least in the primary belief t_i is convinced that j behaves rationally given her belief.

For $(c_j, t_j), (c'_j, t'_j) \in C_j \times T_j$, we say that t_i *deems* (c_j, t_j) *infinitely more likely than* (c'_j, t'_j) iff there is $k \in \{0, ..., K-1\}$ such that the following two conditions are satisfied:

(a) $b_{i\ell}(c_j, t_j) = b_{i\ell}(c'_j, t'_j) = 0$ for $\ell = 1, ..., k$, and

(b) $b_{i,k+1}(c_j, t_j) > 0$ and $b_{i,k+1}(c'_j, t'_j) = 0$.

A cautious type $t_i \in T_i$ *respects the opponent's preferences* iff for each $t_j \in T_j(t_i)$ and $c_j, c'_j \in C_j$ where t_j prefers c_j to c'_j, t_i deems (c_j, t_j) infinitely more likely than (c'_j, t_j). That is, t_i arranges j's choices from the most to the least preferred for each belief hierarchy of j deemed possible by t_i. It can be seen that respect of the opponent's preferences implies primary belief in the opponent's rationality, since the former requires that each type of the opponent deemed possible in the primary belief should only pair with choices most preferred under that type.

Let P be an arbitrary property of lexicographic beliefs. We define that

(a) $t_i \in T_i$ *expresses* 0*-fold full belief in* P iff t_i satisfies P;

(b) For each $n \in \mathbb{N}$, $t_i \in T_i$ *expresses* $(n+1)$*-fold full belief in* P iff t_i only deems possible j's types that express n-fold full belief in P.

t_i *expresses common full belief in* P iff it expresses n-fold full belief in P for each $n \in \mathbb{N}$.

Definition 2.2 (Permissibility and proper rationalizability). Consider a finite lexicographic epistemic model $M^{co} = (T_i, b_i)_{i\in I}$ for a game $\Gamma = (C_i, u_i)_{i\in I}$. $c_i \in C_i$ is *permissible* iff it is rational for some $t_i \in T_i$ which expresses common full belief in caution and primary belief in rationality. c_i is *properly rationalizable* iff it is rational for some $t_i \in T_i$ which expresses common full belief in caution and respect of preferences.

Since respect of the opponent's preferences implies primary belief in the opponent's rationality, proper rationalizability implies permissibility, while the reverse does not hold.

2.2. Incomplete information model

In this subsection, we define the lexicographic epistemic model with incomplete information which is the counterpart of the probabilistic epistemic model with incomplete information introduced by Battigalli [2] and further developed in Battigalli and Siniscalchi [3], [4], and Dekel and Siniscalchi [12]. We also define some conditions on types in such a model.

Definition 2.2 (Lexicographic epistemic model with incomplete information). Consider a finite 2-person static game form $G = (C_i)_{i\in I}$. For each $i \in I$, let V_i be the set of utility functions $v_i : C_1 \times C_2 \to \mathbb{R}$. A *finite lexicographic epistemic model for G with incomplete information* is a tuple $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ where

(a) Θ_i is a finite set of types,

(b) w_i is a mapping that assigns to each $\theta_i \in \Theta_i$ a utility function $w_i(\theta_i) \in V_i$, and

(c) β_i is a mapping that assigns to each $\theta_i \in \Theta_i$ a lexicographic belief over $\Delta(C_j \times \Theta_j)$, i.e., $\beta_i(\theta_i) = (\beta_{i1}, \beta_{i2}, ..., \beta_{iK})$ where $\beta_{ik} \in \Delta(C_j \times \Theta_j)$ for $k = 1, ..., K$.

Concepts such as "θ_i deems (c_j, θ_j) possible" and "θ_i deems (c_j, θ_j) infinitely more likely than (c'_j, θ'_j)" can be defined in a similar way as in Section 2.1. For each $\theta_i \in \Theta_i$, we use $\Theta_j(\theta_i)$ to denote the set of types in Θ_j deemed possible by θ_i. For each $\theta_i \in \Theta_i$ and $v_i \in V_i$, $\theta_i^{v_i}$ is the auxiliary type satisfying that $\beta_i(\theta_i^{v_i}) = \beta_i(\theta_i)$ and $w_i(\theta_i^{v_i}) = v_i$.

For each $c_i \in C_i, v_i \in V_i$, and $\theta_i \in \Theta_i$ with $\beta_i(\theta_i) = (\beta_{i1}, \beta_{i2}, ..., \beta_{iK})$, let $v_i(c_i, \theta_i) = (v_i(c_i, \beta_{i1}), ..., v_i(c_i, \beta_{iK}))$ where for each $k = 1, ..., K$, $v_i(c_i, \beta_{ik}) := \Sigma_{(c_j,\theta_j)\in C_j\times\Theta_j}\beta_{ik}(c_j, \theta_j)v_i(c_i, c_j)$. For each $c_i, c'_i \in C_i$ and $\theta_i \in \Theta_i$, we say that θ_i *prefers* c_i *to* c'_i iff $w_i(\theta_i)(c_i, \theta_i) > w_i(\theta_i)(c'_i, \theta_i)$. As in Section 2.1, this is also the lexicographic comparison between two vectors. c_i is *rational* (or *optimal*) for θ_i iff θ_i does not prefer any choice to c_i.

Definition 2.3 (Caution). $\theta_i \in \Theta_i$ is *cautious* iff for each $c_j \in C_j$ and each $\theta_j \in \Theta_j(\theta_i)$, there is some utility function $v_j \in V_j$ such that θ_i deems $(c_j, \theta_j^{v_j})$ possible.

This is a faithful translation of Perea and Roy [17]'s definition of caution in probabilistic model (p.312) into lexicographic model. It is the counterpart of caution defined within the complete information framework in Section 2.1; the only difference is that in incomplete information models we allow different utility functions since c_j will be required to be rational for the paired type.

Definition 2.4 (Belief in rationality). $\theta_i \in \Theta_i$ *believes in j's rationality* iff θ_i deems (c_j, θ_j) possible only if c_j is rational for θ_j.

In an incomplete information model, since each type is assigned with a belief on the opponent's choice-type pairs as well as a payoff function, caution and a full belief of rationality can be satisfied simultaneously. The consistency of caution and (full) rationality is the essential difference of models with incomplete information from those with complete information. Rationality does not appear in the conditions for our characterizations. Nevertheless, in the proofs we will construct incomplete information models whose types satisfies all the conditions (including caution) as well as common full belief in rationality. We will discuss more about this consistency between caution and rationality in Sections 4.3 and 4.6.

For each $u_i, v_i \in V_i$, we define the distance $d(u_i, v_i)$ between u_i, v_i by $d(u_i, v_i) = [\Sigma_{c\in C}(u_i(c) - v_i(c))^2]^{1/2}$. This is the Euclidean distance on $\mathbb{R}^C$. We choose it is just out of simplicity. Any distance satisfying the three conditions in Section 3.3 of Perea and Roy [17] also works in our characterization. On the other hand, since the interpretation of $d(v_i, u_i)$ is the similarity between v_i and u_i, and both u_i and v_i are functions representing specific preferences, the Euclidean distance seems cardinal. For example, though multiplying u_i with a positive number leads to the same preferences represented by u_i, its Euclidean distance from u_i may be large. In Section 4.5 we will define an ordinal distance on V_i and show that the characterizations still hold under that distance.

Definition 2.5 (Primary belief in utilities nearest to u and u-centered belief). Consider a static game form $G = (C_i)_{i\in I}$, a lexicographic epistemic model $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ for G with incomplete information, and a pair $u = (u_i)_{i\in I}$ of utility functions.

(5.1) $\theta_i \in \Theta_i$ *primarily believes in utilities nearest to u* iff θ_i's level-1 belief only assigns positive probability to (c_j, θ_j) which satisfies that $d(w_j(\theta_j), u_j) \leq d(w_j(\theta'_j), u_j)$ for all $\theta'_j \in \Theta_j(\theta_i)$ with

$\beta_j(\theta'_j) = \beta_j(\theta_j)$.

(5.2) $\theta_i \in \Theta_i$ has *u-centered belief* iff for any $c_j, c'_j \in C_j$, any $\theta_j \in \Theta_j$, and any $v_j, v'_j \in V_j$ such that $(c_j, \theta_j^{v_j})$ and $(c'_j, \theta_j^{v'_j})$ are deemed possible by θ_i, it holds that θ_i deems $(c_j, \theta_j^{v_j})$ infinitely more likely than $(c'_j, \theta_j^{v'_j})$ whenever $d(v_j, u_j) < d(v'_j, u_j)$.

Definition 2.5 gives restrictions on the order of types in a lexicographic belief. (5.1) requires that θ_i primarily believes in type θ_j only if θ_j's utility function is the nearest to u_j among all types sharing the same belief with θ_j. (5.2) requires that the types of j sharing the same belief deemed possible by θ_i are arranged according to the distance of their assigned utility functions from u_j : the farther a type θ_j's utility function is from u_j, the later θ_j occurs in the lexicographic belief of θ_i. (5.2) is a faithful translation of Perea and Roy [17]'s Definition 3.2 into lexicograhic model and (5.1) is a weaker version of (5.2).

The essential difference between our conditions and Perea and Roy [17]'s for characterization lies in the following definition.

Definition 2.6 (A best (better) choice is supported by utilities nearest (nearer) to u). Consider a static game form $G = (C_i)_{i \in I}$, a lexicographic epistemic model $M^{in} = (\Theta_i, w_i, \beta_i)_{i \in I}$ for G with incomplete information, and a pair $u = (u_i)_{i \in I}$ of utility functions.

(6.1) $\theta_i \in \Theta_i$ *believes in that a best choice of j is supported by utilities nearest to u* iff for any (c_j, θ_j), (c'_j, θ'_j) deemed possible by θ_i with $\beta_j(\theta_j) = \beta_j(\theta'_j)$, if c_j is optimal for $\beta_j(\theta_j)$ in u_j but c'_j is not, then $d(w_j(\theta_j), u_j) < d(w_j(\theta'_j), u_j)$.

(6.2) $\theta_i \in \Theta_i$ *believes in that a better choice of j is supported by utilities nearer to u* iff for any (c_j, θ_j), (c'_j, θ'_j) deemed possible by θ_i with $\beta_j(\theta_j) = \beta_j(\theta'_j)$, if $u_j(c_j, \theta_j) > u_j(c'_j, \theta'_j)$, then $d(w_j(\theta_j), u_j) < d(w_j(\theta'_j), u_j)$.

Definition 2.6 gives restriction on the relation between paired type and choice. (6.1) requires that for each belief of player j, a choice optimal for that belief should be supported by the nearest utility funtion to u_j. (6.2) requires that for each belief of player j, a utility function supporting a "better" choice (i.e., c_j) should be nearer to u_j than one supporting a "worse" choice (i.e., c'_j). It can be seen that (6.2) is stronger than (6.1).

(6.2) is similar to Perea and Roy [17]'s Definition 3.3 which requires that for each (c_j, θ_j) deemed possible by θ_i, $w_j(\theta_j)$ is the nearest utility function in V_j to u_j among those at which c_j is rational under $\beta_j(\theta_j)$. It can be shown by Lemmas 5.4 and 5.5 in Perea and Roy [17] that Definition 2.6 is weaker than Perea and Roy [17]'s Definition 3.3. We adopt it here since a nearest utility function does not in general exist for lexicographic beliefs. That is, given $u_j \in V_j$, $c_j \in C_j$, and a lexicographic belief β_j, there may not exist $v_j \in V_j$ satisfying that (1) c_j is rational at v_j under β_j, and (2) there is no $v'_j \in V_j$ such that c_j is rational at v'_j for β_j and $d(v'_j, u_j) < d(v_j, u_j)$. See the following example.

Example 2.1 (No nearest utility function). Consider a game Γ where player 1 has choices A, B, and C and player 2 has choices D, E, and F. The payoff function u_1 of player 1 is as follows:

u_1	D	E	F
A	1	1	1
B	1	1	0
C	1	0	1

Let $\beta_1 = (D, E, F)$, that is, player 1 deems player 2's choice D infinitely more likely than E and E infinitely more likely than F. In u_1, A is rational for β_1 but B is not. Now we show that there is no nearest utility function to u_1 at which B is rational under β_1. Suppose there is such a function $v_1 \in V_1$. Let $d = d(v_1, u_1)$. It can be seen that $d > 0$. Consider the following v_1' :

v_1'	D	E	F
A	1	1	1
B	$1 + \frac{d}{2}$	1	0
C	1	0	1

B is also rational at v_1' under β_1, while $d(v_1', u_1) = \frac{d}{2} < d = d(v_1, u_1)$, a contradiction. Also, even though B is preferred to C in u_1 under β_1, it can be seen that for each utility function v_1^B in which B is rational under β_1, there is some $v_1^C \in V_1$ satisfying (1) C is optimal in v_1^C under β_1, and (2) $d(v_1^C, u_1) < d(v_1^B, u_1)$. Indeed, this can be done by letting $v_1^C(C, D) = 1 + d(v_1^B, u_1)/2$ and $v_1^C(c_1, c_2) = u_1(c_1, c_2)$ for all other $(c_1, c_2) \in C_1 \times C_2$.

Example 2.1 shows that the relationship between preferences among choices and the distance of utility functions from the original one is more complicated for lexicographic beliefs. That is why we adopt Definition 2.6 here. The following lemma guarantees the existence of utility functions satisfying the condition in Definition 2.6. It shows that, given a utility function u_i and a lexicographic belief β_i, corresponding to the sequence $c_{i1}, ..., c_{iN}$ of i's choices arranged from the most to the least preferred at u_i under β_i, there is a sequence $v_{i1}, ..., v_{iN}$ of utility functions arranged from the nearest to the farthest to u_i such that for each $n = 1, ..., N$, c_{in} is rational at v_{in} under β_i. This lemma plays a similar role in our characterizations as Lemmas 5.4 and 5.5 in Perea and Roy [17].

Lemma 2.1 (Existence of utilities satisfying Definition 2.6). Consider a static game form $G = (C_i)_{i \in I}$, $u_i \in V_i$, and $\beta_i = (\beta_{i1}, \beta_{i2}, ..., \beta_{iK})$ such that $\beta_{ik} \in \Delta(C_j)$ for each $k = 1, ..., K$. Let $\Pi_i(\beta_i) = (C_{i1}, C_{i2}, ..., C_{iL})$ be a partition of C_i satisfying that (1) for each $\ell = 1, ..., L$ and each $c_{i\ell}, c_{i\ell}' \in C_{i\ell}$, $u_i(c_{i\ell}, \beta_i) = u_i(c_{i\ell}', \beta_i)$, and (2) for each $\ell = 1, ..., L-1$, each $c_{i\ell} \in C_{i\ell}$ and $c_{i,\ell+1} \in C_{i,\ell+1}$, $u_i(c_{i\ell}, \beta_i) > u_i(c_{i,\ell+1}, \beta_i)$. That is, $\Pi_i(\beta_i)$ is the sequence of equivalence classes of choices in C_i arranged from the most preferred to the least preferred under β_i.

Then there are $v_{i1}, ..., v_{iL} \in V_i$ satisfying

(a) $v_{i1} = u_i$,

(b) For each $\ell = 1, ..., L$ and each $c_{i\ell} \in C_{i\ell}$, $c_{i\ell}$ is rational at $v_{i\ell}$ under β_i, and

(c) For each $\ell = 1, ..., L-1$, $d(v_{i\ell}, u_i) < d(v_{i,\ell+1}, u_i)$.

3. Characterizations

So far we have introduced two different groups of concepts for static games: one includes permissibility and proper rationalizability within a complete information framework, the other contains various conditions on types within an incomplete information framework. In this section we will show that there are correspondences between them.

3.1. Statements and an example

In this subsection we give two characterization results and an illustrative example.

Theorem 3.1 (Characterization of permissibility). Consider a finite 2-person static game $\Gamma = (C_i, u_i)_{i \in I}$ and the corresponding game form $G = (C_i)$.

Then, $c_i^* \in C_i$ is permissible if and only if there is some finite lexicographic epistemic model $M^{in} = (\Theta_i, w_i, \beta_i)_{i \in I}$ with incomplete information for G and some $\theta_i^* \in \Theta_i$ with $w_i(\theta_i^*) = u_i$ such that

(a) c_i^* is rational for θ_i^*, and,

(b) θ_i^* expresses common full belief in caution, primary belief in utilities nearest to u, and that a best choice is supported by utilities nearest to u.

Theorem 3.2 (Characterization of proper rationalizability). Consider a finite 2-person static game $\Gamma = (C_i, u_i)_{i\in I}$ and the corresponding game form $G = (C_i)_{i\in I}$.

Then, $c_i^* \in C_i$ is properly rationalizable if and only if there is some finite lexicographic epistemic model $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ for G and some $\theta_i^* \in \Theta_i$ with $w_i(\theta_i^*) = u_i$ such that

(a) c_i^* is rational for θ_i^*, and

(b) θ_i^* expresses common full belief in caution, u-centered belief, and that a better choice is supported by utilities nearer to u.

To show these statements, we will construct a correspondence between complete information models and incomplete ones and show that conditions on a type in one model can be transformed into a proper condition on the corresponding type in the constructed model. Before the formal proofs, we use the following example to show the intuition.

Example 3.1. Consider the following game Γ (Perea [15], p.190):

$u_1\backslash u_2$	D	E	F
A	$0,3$	$1,2$	$1,1$
B	$1,3$	$0,2$	$1,1$
C	$1,3$	$1,2$	$0,1$

and the lexicographic model $M^{co} = (T_i, b_i)_{i\in I}$ for Γ where $T_1 = \{t_1\}$, $T_2 = \{t_2\}$, and

$$b_1(t_1) = ((D,t_2),(E,t_2),(F,t_2)),\ b_2(t_2) = ((C,t_1),(B,t_1),(A,t_1)).$$

It can be seen that D is properly rationalizable (and therefore permissible) since it is rational for t_2 which expresses common full belief in caution and respect of preferences. Consider the lexicographic epistemic model $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ with incomplete information for the corresponding game form where $\Theta_1 = \{\theta_{11}, \theta_{12}, \theta_{13}\}$, $\Theta_2 = \{\theta_{21}, \theta_{22}, \theta_{23}\}$, and

$$\begin{aligned}
w_1(\theta_{11}) &= u_1,\ \beta_1(\theta_{11}) = ((D,\theta_{21}),(E,\theta_{22}),(F,\theta_{23})),\\
w_1(\theta_{12}) &= v_1,\ \beta_1(\theta_{12}) = ((D,\theta_{21}),(E,\theta_{22}),(F,\theta_{23})),\\
w_1(\theta_{13}) &= v_1',\ \beta_1(\theta_{13}) = ((D,\theta_{21}),(E,\theta_{22}),(F,\theta_{23})),\\
w_2(\theta_{21}) &= u_2,\ \beta_2(\theta_{21}) = ((C,\theta_{11}),(B,\theta_{12})),(A,\theta_{13})),\\
w_2(\theta_{22}) &= v_2,\ \beta_2(\theta_{22}) = ((C,\theta_{11}),(B,\theta_{12})),(A,\theta_{13})),\\
w_2(\theta_{23}) &= v_2',\ \beta_2(\theta_{23}) = ((C,\theta_{11}),(B,\theta_{12})),(A,\theta_{13})).
\end{aligned}$$

where

v_1	D	E	F
A	0	1	1
B	2	0	1
C	1	1	0

,

v_1'	D	E	F
A	3	1	1
B	2	0	1
C	1	1	0

,

v_2	D	E	F
A	3	2	1
B	3	2	1
C	3	4	1

,

v_2'	D	E	F
A	3	2	1
B	3	2	1
C	3	4	5

.

For each $i \in I$, θ_{i1}, θ_{i2}, and θ_{i3} have the same belief; the only difference lies in their assigned utility functions since each should support some choice. The relation between M^{in} and M^{co} can be seen clearly here: for each $i \in I$, θ_{i1}, θ_{i2}, and θ_{i3} correspond to t_i in the sense that the belief of the former is obtained by replacing every occurence of t_j in the belief of t_i by the type corresponding to t_j in M^{in} at which the paired choice is optimal. It can be seen that θ_{11} expresses common full belief in caution, u-centered belief, and that a better choice is supported by utilities nearer to u (therefore primary belief in utilities nearest to u and that a best choice is supported by utilities nearest to u). Also, since the assigned utility function of θ_{11} is u_1, C is rational for θ_{11}.

This example can be used to show the difference between Theorems 3.1 and 2. Consider the lexicographic epistemic model $(T'_i, b'_i)_{i\in I}$ for Γ where $T'_1 = \{t'_1\}$, $T'_2 = \{t'_2\}$, and

$$b'_1(t'_1) = ((D, t'_2), (F, t'_2), (E, t'_2)),\ b'_2(t'_2) = ((B, t'_1), (C, t'_1), (A, t'_1)).$$

It can be seen that t'_1 is expresses common full belief in caution and primary belief in rationality. We can construct the corresponding lexicographic epistemic model $M^{in} = (\Theta'_i, w'_i, \beta'_i)_{i\in I}$ for the corresponding game form with incomplete information where $\Theta'_1 = \{\theta'_{11}, \theta'_{12}, \theta'_{13}\}$, $\Theta_2 = \{\theta'_{21}, \theta'_{22}, \theta'_{23}\}$, and

$$\begin{aligned}
w'_1(\theta'_{11}) &= u_1,\ \beta'_1(\theta'_{11}) = ((D, \theta'_{21}), (F, \theta'_{22}), (E, \theta'_{23})),\\
w'_1(\theta'_{12}) &= v'_1,\ \beta'_1(\theta'_{12}) = ((D, \theta'_{21}), (F, \theta'_{22}), (E, \theta'_{23})),\\
w'_1(\theta'_{13}) &= v_1,\ \beta'_1(\theta'_{13}) = ((D, \theta'_{21}), (F, \theta'_{22}), (E, \theta'_{23})),\\
w'_2(\theta'_{21}) &= u_2,\ \beta'_2(\theta'_{21}) = ((B, \theta'_{11}), (C, \theta'_{12})), (A, \theta'_{13})),\\
w'_2(\theta'_{22}) &= v'_2,\ \beta'_2(\theta'_{22}) = ((B, \theta'_{11}), (C, \theta'_{12})), (A, \theta'_{13})),\\
w'_2(\theta'_{23}) &= v_2,\ \beta'_2(\theta'_{23}) = ((B, \theta'_{11}), (C, \theta'_{12})), (A, \theta'_{13})).
\end{aligned}$$

It can be seen that θ'_{11} expresses common full belief in caution, primary belief in utilities nearest to u, and that a best choice is supported by utilities nearest to u. On the other hand, it can be seen that t'_1 does not respect player 2's preferences, since E is always preferred to F, while t'_1 deems F infinitely more likely than E. In M^{in}, this can be seen in the violation of u-centered belief in θ'_{11}, that is, though $\beta'_2(\theta'_{22}) = \beta'_1(\theta'_{23})$ and $d(w'_2(\theta'_{22}), u_2) = d(v'_2, u_2) = \sqrt{10} > d(w'_2(\theta'_{23}), u_2) = d(v_2, u_2) = 1$, θ'_{11} deems (F, θ'_{22}) infinitely more likely than (E, θ'_{23}).

3.2. Proof of Theorem 3.1

To show the Only-if part of Theorem 3.1, we construct the following mapping from finite lexicographic epistemic models with complete information to those with incomplete information. Consider $\Gamma = (C_i, u_i)_{i\in I}$ and a finite lexicographic epistemic model $M^{co} = (T_i, b_i)_{i\in I}$ with complete information for Γ. We first define types in a model with incomplete information in the following two steps:

Step 1. For each $i \in I$ and $t_i \in T_i$, let $\Pi_i(t_i) = (C_{i1}, ..., C_{iL})$ be the partition of C_i defined in Lemma 2.1, that is, $\Pi_i(t_i)$ is the sequence of equivalence classes of choices in C_i arranged from the most preferred to the least preferred under t_i. By Lemma 2.1, for each $C_{i\ell}$ there is some $v_{i\ell}(t_i) \in V_i$ such that each choice in $C_{i\ell}$ is rational at $v_{i\ell}(t_i)$ under t_i, and $0 = d(v_{i1}(t_i), u_i) < d(v_{i2}(t_i), u_i) < ... < d(v_{iL}(t_i), u_i)$.

Step 2. We define $\Theta_i(t_i) = \{\theta_{i1}(t_i), ..., \theta_{iL}(t_i)\}$ where for each $\ell = 1, ..., L$, the type $\theta_{i\ell}(t_i)$ satisfies that (1) $w_i(\theta_{i\ell}(t_i)) = v_{i\ell}(t_i)$, and (2) $\beta_i(\theta_{i\ell}(t_i))$ is obtained from $b_i(t_i)$ by replacing every (c_j, t_j) with $c_j \in C_{jr} \in \Pi_j(t_j)$ for some r with (c_j, θ_j) where $\theta_j = \theta_{jr}(t_j)$, that is, $w_j(\theta_j)$ is the utility function among those corresponding to $\Pi_j(t_j)$ in which c_j is the rational for t_i.

For each $i \in I$, let $\Theta_i = \cup_{t_i\in T_i}\Theta_i(t_i)$. Here we have constructed a finite lexicographic epistemic model $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ for the corresponding game form $G = (C_i)_{i\in I}$ with incomplete information. In the following example we show how this construction goes.

Example 3.2. Consider the following game Γ (Perea [15], p.188):

$u_1\backslash u_2$	C	D
A	$1, 0$	$0, 1$
B	$0, 0$	$0, 1$

and the lexicographic epistemic model $M^{co} = (T_i, b_i)_{i\in I}$ Γ where $T_1 = \{t_1\}$, $T_2 = \{t_2\}$, and

$$b_1(t_1) = ((D, t_2), (C, t_2)),\ b_2(t_2) = ((A, t_1), (B, t_1)).$$

We show how to construct a corresponding model $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$. First, by Step 1 it can be seen that $\Pi_1(t_1) = (\{A\}, \{B\})$ and $\Pi_2(t_2) = (\{D\}, \{C\})$. We let $v_{11}(t_1) = u_1$ where A is rational for t_1 and $v_{12}(t_1)$ where B is rational for t_1 as follows. Similarly, we let $v_{21}(t_2) = u_2$ where D is rational under t_2 and $v_{22}(t_2)$ where C is rational under t_2 as follows:

$v_{12}(t_1)$	C	D
A	1	0
B	0	1

,

$v_{22}(t_2)$	C	D
A	2	1
B	0	1

.

Then we go to Step 2. It can be seen that $\Theta_1(t_1) = \{\theta_{11}(t_1), \theta_{12}(t_1)\}$, where

$$\begin{aligned} w_1(\theta_{11}(t_1)) &= v_{11}(t_1),\ \beta_1(\theta_{11}(t_1)) = ((D, \theta_{21}(t_2)), (C, \theta_{22}(t_2))), \\ w_1(\theta_{12}(t_1)) &= v_{12}(t_1),\ \beta_1(\theta_{12}(t_1)) = ((D, \theta_{21}(t_2)), (C, \theta_{22}(t_2))). \end{aligned}$$

Also, $\Theta_2(t_2) = \{\theta_{21}(t_2), \theta_{22}(t_2)\}$, where

$$\begin{aligned} w_2(\theta_{21}(t_2)) &= v_{21}(t_2),\ \beta_2(\theta_{21}(t_2)) = ((A, \theta_{11}(t_1)), (B, \theta_{12}(t_1))), \\ w_2(\theta_{22}(t_2)) &= v_{22}(t_2),\ \beta_2(\theta_{22}(t_2)) = ((A, \theta_{11}(t_1)), (B, \theta_{12}(t_1))). \end{aligned}$$

Let $M^{co} = (T_i, b_i)_{i\in I}$ and $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ be constructed from M^{co} by the two steps above. We have the following observations.

Observation 3.1 (Redundancy). For each $t_i \in T_i$ and each $\theta_i, \theta_i' \in \Theta_i(t_i)$, $\beta_i(\theta_i) = \beta_i(\theta_i')$.

Observation 3.2 (Rationality). Eeach $\theta_i \in \Theta_i(t_i)$ believes in j's rationality.

Observation 3.3 (A better choice is supported by utilities nearer to u). Each $\theta_i \in \Theta_i(t_i)$ believes that a better choice is supported by utilities nearer to u.

We omit their proofs since they hold by construction. Observation 3.1 means that the difference between any two types in a $\Theta_i(t_i)$ is in the utility functions assigned to them. Observation 3.2 means that in an incomplete information model constructed from one with complete information, each type has (full) belief in the opponent's rationality. This is because in the construction, we requires that for each pair (c_j, t_j) occurring in a belief, its counterpart in the incomplete information replaces t_j by the type in $\Theta_j(t_j)$ with the utility function in which c_j is optimal for t_j. It follows from Observation 3.2 that each $\theta_i \in \Theta_i(t_i)$ expresses common full belief in rationality. Observation 3.3 implies that the best choice is supported by utilities nearest to u. It follows that each $\theta_i \in \Theta_i(t_i)$ expresses common full belief in that a best (better) choice is supported by utilities nearest (nearer) to u.

By construction, each t_i shares the same belief about j's choices at each level with each $\theta_i \in \Theta_i(t_i)$; also, for each $t_i \in T_i$, the utility function assigned to $\theta_{i1}(t_i)$ is u_i. It is clear that any c_i rational for t_i is also rational for $\theta_{i1}(t_i)$. Therefore, to show the only-if part of Theorem 3.1, we show that if t_i expresses common full belief in caution and primary belief in rationality, then $\theta_{i1}(t_i)$ expresses common belief in caution, primary belief in utilities nearest to u, and that a best choice is supported by utilities nearest to u.

Lemma 3.1 (Caution^{co} $\to$ Caution^{in}). Let $M^{co} = (T_i, b_i)_{i\in I}$ and $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ be constructed from M^{co} by the two steps above. If $t_i \in T_i$ expresses common full belief in caution, so does each $\theta_i \in \Theta_i(t_i)$.

Lemma 3.2 (Primary belief in rationality $\to$ primary belief in utilities nearest to u). Let $M^{co} = (T_i, b_i)_{i\in I}$ and $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ be constructed from M^{co} by the two steps above. If $t_i \in T_i$ expresses common full belief in primary belief in rationality, then each $\theta_i \in \Theta_i(t_i)$ expresses common full belief in primary belief in utilities nearest to u.

Proof of the Only-if part of Theorem 3.1. Let $M^{co} = (T_i, b_i)_{i\in I}$, $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$

be constructed from M^{co} by the two steps above, $c_i^* \in C_i$ be a permissible choice, and $t_i^* \in T_i$ be a type expressing common full belief in caution and primary belief in rationality such that c_i^* is rational for t_i^*. Let $\theta_i^* = \theta_{i1}(t_i^*)$. By definition, $w_i(\theta_i^*) = u_i$ and $\beta_i(\theta_i^*)$ has the same distribution on j's choices at each level as $b_i(t_i^*)$. Hence c_i^* is rational for θ_i^*. Also, it follows from Observation 3.3, Lemmas 3.1, and 3.2 that θ_i^* expresses common full belief in caution, primary belief in utilities nearest to u, and that a best choice is supported by utilities nearest to u. //

To show the If part, we need a mapping from models with incomplete information to those with complete information. Consider a finite 2-person static game $\Gamma = (C_i, u_i)_{i\in I}$, the corresponding game form $G = (C_i)_{i\in I}$, and a finite epistemic model $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ for G with incomplete information. We construct a model $M^{co} = (T_i, b_i)_{i\in I}$ for Γ with complete information as follows. For each $\theta_i \in \Theta_i$, we define $E_i(\theta_i) = \{\theta_i' \in \Theta_i : \beta_i(\theta_i') = \beta(\theta_i)\}$. In this way Θ_i is partitioned into some equivalence classes $\mathbb{E}_i = \{E_{i1}, ..., E_{iL}\}$ where for each $\ell = 1, .., L$, $E_{i\ell} = E_i(\theta_i)$ for some $\theta_i \in \Theta_i$. To each $E_i \in \mathbb{E}_i$ we use $t_i(E_i)$ to represent a type. We define $b_i(t_i(E_i))$ to be a lexicographic belief which is obtained from $\beta_i(\theta_i)$ by replacing each occurrence of (c_j, θ_j) by $(c_j, t_j(E_j(\theta_j)))$; in other words, $b_i(t_i(E_i))$ has the same distribution on choices at each level as $\beta_i(\theta_i)$ for each $\theta_i \in E_i$, while each $\theta_j \in \Theta_j(\theta_i)$ is replaced by $t_j(E_j(\theta_j))$. For each $i \in I$, let $T_i = \{t_i(E_i)\}_{E_i\in\mathbb{E}_i}$. We have constructed from M^{in} a finite epistemic model $M^{co} = (T_i, b_i)_{i\in I}$ with complete information for Γ.

It can be seen that this is the reversion of the previous construction. That is, let $M^{co} = (T_i, b_i)_{i\in I}$ satisfying that $b_i(t_i) \neq b_i(t_i')$ for each $t_i, t_i' \in T_i$ with $t_i \neq t_i'$, and $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ be constructed from M^{co} by the previous two steps. Then $\mathbb{E}_i = \{\Theta_i(t_i)\}_{t_i\in T_i}$ and $t_i(\Theta_i(t_i)) = t_i$ for each $i \in I$.

In the following example we show how this construction goes.

Example 3.3. Consider the game Γ in Example 3.2 and the model $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ for the corresponding game form where $\Theta_1 = \{\theta_{11}, \theta_{12}\}$, $\Theta_2 = \{\theta_{21}, \theta_{22}\}$, and

$$\begin{aligned}
w_1(\theta_{11}) &= u_1,\ \beta_1(\theta_{11}) = ((D, \theta_{21}), (C, \theta_{22})),\\
w_1(\theta_{12}) &= v_1,\ \beta_1(\theta_{12}) = ((D, \theta_{21}), (C, \theta_{22})),\\
w_2(\theta_{21}) &= u_2,\ \beta_2(\theta_{21}) = ((A, \theta_{11}), (B, \theta_{12})),\\
w_2(\theta_{22}) &= v_2,\ \beta_2(\theta_{22}) = ((A, \theta_{11}), (B, \theta_{12})).
\end{aligned}$$

where $v_1 = v_{12}(t_1)$ and $v_2 = v_{22}(t_2)$ in Example 3.2. It can be seen that $\mathbb{E}_1 = \{\{\theta_{11}, \theta_{12}\}\}$ since $\beta_1(\theta_{11}) = \beta_1(\theta_{12})$ and $\mathbb{E}_2 = \{\{\theta_{21}, \theta_{22}\}\}$ since $\beta_2(\theta_{21}) = \beta_2(\theta_{22})$. Corresponding to those equivalence classes we have $t_1(\{\theta_{11}, \theta_{12}\})$ and $t_2(\{\theta_{21}, \theta_{22}\})$, and

$$\begin{aligned}
b_1(t_1(\{\theta_{11}, \theta_{12}\})) &= ((D, t_2(\{\theta_{21}, \theta_{22}\})), (C, t_2(\{\theta_{21}, \theta_{22}\}))),\\
b_2(t_2(\{\theta_{21}, \theta_{22}\})) &= ((A, t_1(\{\theta_{11}, \theta_{12}\})), (B, t_1(\{\theta_{11}, \theta_{12}\}))).
\end{aligned}$$

To show the If part of Theorem 3.1, we need the following lemmas.

Lemma 3.3 (Caution^{in} → Caution^{co}). Let $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ and $M^{co} = (T_i, b_i)_{i\in I}$ be constructed from M^{in} by the above approach. If $\theta_i \in \Theta_i$ expresses common full belief in caution, so does $t_i(E_i(\theta_i))$.

Lemma 3.4 (Caution^{in} + primary belief in utilities nearest to u + a best choice is supported by utilities nearest to u → Primary belief in rationality). Let $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ and $M^{co} = (T_i, b_i)_{i\in I}$ be constructed from M^{in} by the above approach. If $\theta_i \in \Theta_i$ expresses common full belief in caution, primary belief in utilities nearest to u, and that a best choice is supported by utilities nearest to u, then $t_i(E_i(\theta_i))$ expresses common full belief in primary belief in rationality.

Proof of the If part of Theorem 3.1. Let $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$, $M^{co} = (T_i, b_i)_{i\in I}$ be constructed from M^{in} by the above approach, and $c_i^* \in C_i$ be rational for some θ_i^* with $w_i(\theta_i^*) = u_i$ which expresses common full belief in caution, primary belief in utilities nearest to u, and that a best choice is supported by utilities nearest to u. Consider $t_i(E_i(\theta_i^*))$. Since $w_i(\theta_i^*) = u_i$ and $b_i(t_i(E_i(\theta_i^*)))$ has the same distribution on j's choices at each level as $\beta_i(\theta_i^*)$, c_i^* is rational for $t_i(E_i(\theta_i^*))$. Also, by Lemmas 3.3 and 3.4, $t_i(E_i(\theta_i^*))$ expresses common full belief in caution and primary belief in rationality. Hence c_i^* is permissible in Γ. //

3.3. Proof of Theorem 3.2

To show the Only-if part of Theorem 3.2, we need the following lemmas.

Lemma 3.5 (Respect of preferences → u-centered belief). Let $M^{co} = (T_i, b_i)_{i\in I}$ and $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ be constructed from M^{co} by the two steps in Section 3.2. If $t_i \in T_i$ expresses common full belief in caution and respect of preferences, then each $\theta_i \in \Theta_i(t_i)$ expresses full belief in u-centered belief.

Proof of the Only-if part of Theorem 3.2. Let $M^{co} = (T_i, b_i)_{i\in I}$, $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ be constructed from M^{co} by the two steps in Section 3.2, $c_i^* \in C_i$ be properly rationalizable, and $t_i^* \in T_i$ be a type which expresses common full belief in caution and respect of preferences such that c_i^* is rational for t_i^*. Let $\theta_i^* = \theta_{i1}(t_i^*)$. Since $w_i(\theta_i^*) = u_i$ and $\beta_i(\theta_i^*)$ has the same distribution on j's choices as $b_i(t_i^*)$, c_i is rational for θ_i^*. Also, it follows from Observations 3.3 and Lemmas 3.1 and 3.5 that θ_i^* expresses common belief in caution, u-centered belief, and that a better choice is supported by utilities nearer to u. //

To show the If part, we still use the construction from M^{in} to M^{co} defined in Section 3.2. We need the following lemma.

Lemma 3.6 (Caution^{in} + u-centered belief + a better choice is supported by utilities nearer to u → respect of preferences). Let $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ and $M^{co} = (T_i, b_i)_{i\in I}$ be constructed from M^{in} by the approach in Section 3.2. If $\theta_i \in \Theta_i$ expresses common full belief in caution, u-centered belief, and that a better choice is supported by utilities nearer to u, then $t_i(E_i(\theta_i))$ expresses common full belief in respect of preferences.

Proof of the If part of Theorem 3.2. Let $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$, $M^{co} = (T_i, b_i)_{i\in I}$ be constructed from M^{in} by the approach in Section 3.2, and $c_i^* \in C_i$ be rational for some θ_i^* with $w_i(\theta_i^*) = u_i$ which expresses common belief in caution, rationality, u-centered belief, and that a better choice is supported by utilities nearer to u. Consider $t_i(E_i(\theta_i^*))$. Since $w_i(\theta_i^*) = u_i$ and $t_i(E_i(\theta_i^*))$ and θ_i^* have the same distribution on j's choices in each level, c_i^* is rational for $t_i(E_i(\theta_i^*))$. Also, it follows from Lemmas 3.3 and 3.6 that $t_i(E_i(\theta_i^*))$ expresses common full belief in caution and respect of preferences. Hence c_i^* is properly rationalizable in Γ. //

4. Concluding Remarks

4.1. Faithful parallel to Perea and Roy [17]'s Theorem 6.1

Theorems 3.1 and 3.2 can be rephrased as faithful parallels to Perea and Roy [17]'s Theorem 6.1, focusing on equivalence between belief hierarchies in complete and incomplete information models. We adopt the forms here because the coincidence of belief hierarchies holds by construction, and we think it is unnecessary to mention it independently.

4.2. Extending to n-person cases

Both Perea and Roy [17] and this paper focus on 2-person games. To extend those results to n-person cases, the problem is how to define the distance between utility functions and how to relate the distance with the locations of choice-type pairs. In a 2-person game, a type of i only needs to consider distributions on $\Delta(C_j \times \Theta_j)$. Hence a "cell" in $\beta_i(\theta_i)$ is just a pair (c_j, θ_j), and its location in $\beta_i(\theta_i)$ can be related directly to the distance $d(w_j(\theta_j), u_j)$. In contrast, in an n-person case a "cell" of a lexicographic belief contains $n-1$ pairs such as $\langle (c_1, \theta_1), ..., (c_{i-1}, \theta_{i-1}), (c_{i+1}, \theta_{i+1}), ..., (c_n, \theta_n) \rangle$, and consequently there are $n-1$ distances, i.e., $d(w_1(\theta_1), u_1), ..., d(w_{i-1}(\theta_{i-1}), u_{i-1}), d(w_{i+1}(\theta_{i+1}), u_{i+1}), \ ... \ , d(w_n(\theta_n), u_n)$. Then the problem is how to connect the location of this cell and those distances. We believe that the results of Perea and Roy [17] and this paper can be extended to n-person games with a proper definition of the distances and their relation with locations of "cells" in lexicographic beliefs. Further work is expected in that direction.

4.3. The role of rationality

Rationality has not been used in our characterizations even though in the proofs we construct epistemic models with incomplete information in which each type has a common full belief in rationality. On the other hand, there are also epistemic models with types satisfying all conditions in Theorems 3.1 and 3.2 but not believing in rationality. Here we give an example.

Example 4.1 (Rationality is not needed). Consider the game Γ in Example 3.1 and the lexicographic epistemic model $M^{in} = (\Theta_i, w_i, \beta_i)_{i \in I}$ with incomplete information for the corresponding game form where $\Theta_1 = \{\theta_{11}, \theta_{12}, \theta_{13}\}$, $\Theta_2 = \{\theta_{21}, \theta_{22}, \theta_{23}\}$, and

$$\begin{aligned}
w_1(\theta_{11}) &= u_1, \ \beta_1(\theta_{11}) = ((D, \theta_{21}), (F, \theta_{22}), (E, \theta_{23})), \\
w_1(\theta_{12}) &= v_1, \ \beta_1(\theta_{12}) = ((D, \theta_{21}), (F, \theta_{22}), (E, \theta_{23})), \\
w_1(\theta_{13}) &= v_1', \ \beta_1(\theta_{13}) = ((D, \theta_{21}), (F, \theta_{22}), (E, \theta_{23})), \\
w_2(\theta_{21}) &= v_2, \ \beta_2(\theta_{21}) = ((C, \theta_{11}), (B, \theta_{12})), (A, \theta_{13})), \\
w_2(\theta_{22}) &= v_2', \ \beta_2(\theta_{22}) = ((C, \theta_{11}), (B, \theta_{12})), (A, \theta_{13})), \\
w_2(\theta_{23}) &= v_2'', \ \beta_2(\theta_{23}) = ((C, \theta_{11}), (B, \theta_{12})), (A, \theta_{13})).
\end{aligned}$$

where v_1, v_1', v_2, v_2' are the same as in Example 3.1 and v_2'' are as follows:

v_2''	D	E	F
A	3	2	1
B	3	2	1
C	6	4	5

It can be seen that θ_{11} expresses common full belief in caution, u-centered belief and that a better choice is supported by utilities nearer to u (therefore primary belief in utilities nearest to u and that a best choice is supported by utilities nearest to u are also satisfied) but not rationality, since, for example, D is not rational for θ_{21}. However, consider the model $M^{co} = (T_i, b_i)_{i \in I}$ for Γ constructed from M^{in}. Indeed, since $\mathbb{E}_1 = \{\{\theta_{11}, \theta_{12}, \theta_{13}\}\}$ and $\mathbb{E}_2 = \{\{\theta_{21}, \theta_{22}, \theta_{23}\}\}$, by letting $t_1 = t_1(\{\theta_{11}, \theta_{12}, \theta_{13}\})$ and $t_2 = t_2(\{\theta_{21}, \theta_{22}, \theta_{23}\})$, we obtain $M^{co} = (T_i, b_i)_{i \in I}$ for Γ where $T_1 = \{t_1\}$, $T_2 = \{t_2\}$, and

$$b_1(t_1) = ((D, t_2), (F, t_2), (E, t_2)), \ b_2(t_2) = ((C, t_1), (B, t_1), (A, t_1)).$$

It can be seen that t_1 expresses caution and common full belief in respect of preferences (therefore primary belief in rationality). Further, C is optimal for both θ_{11} and t_1.

On the other hand, rationality can be contained in the characterization. In Section 4.6 we will provide an alternative way to characterize permissibility by using rationality and weak caution.

4.4. Using our construction to show Perea and Roy [17]'s Theorem 6.1

Our proofs are based on constructing a specific correspondence between two models. It can be seen that this correspondence can be translated directly into probabilistic models and be used to show Perea and Roy [17]'s Theorem 6.1. Further, it can be seen that, by using our Lemma 2.1, belief in rationality under closest utility function in Perea and Roy [17] can be replaced by the weaker one (Definition 2.6 (6.2)) here.

4.5. An ordinal distance on V_i

In this note, we use the Euclidean distance to measure similarity between utility functions. As mentioned in Section 2.2, the Euclidean distance is cardinal. We can define an ordinal distance as follows to replace it. Let β_i be a lexicographic belief on $\Delta(C_j \times \Theta_j)$. For each $v_i, u_i \in V_i$, define $d^{\beta_i}(v_i, u_i) = |\{\{c_i, c_i'\} : c_i, c_i' \in C_i$ and the preference between c_i and c_i' under β_i at v_i are different from that at $u_i\}|$. It can be seen that d^{β_i} is a metric on V_i which measures similarity between preferences under β_i represented by v_i and that by u_i, i.e., it measures the ordinal difference between v_i and u_i. This does not belong to the group of distances characterized in Section 3.3 of Perea and Roy [17] since there is no norm on V_i to support d^{β_i}. Lemma 2.1 still holds under d^{β_i} since even if we replace d by d^{β_i} in Lemma 2.1 (c), the constructed utility function sequence in the proof still satisfies it. Hence d in Definition 2.5 can be replaced by d^{β_i} with appropriate β_i and the characterization results still hold. Also, by replacing rationality under closest utility function by our Definition 2.6, Perea and Roy [17]'s Theorem 6.1 still holds under d^{β_i}.

4.6. Characterizing permissibility by rationality and weak caution

In this subsection we provide an alternative characterization of permissibility by using rationality and a condition weaker than caution in Definition 2.3.

Definition 4.1 (Weak caution). Consider a game form $G = (C_i)_{i\in I}$ and a lexicographic epistemic model $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ for G with incomplete information. $\theta_i \in \Theta_i$ is *weakly cautious* iff for each $c_j \in C_j$, there is some $\theta_j \in \Theta_j$ such that θ_i deems (c_j, θ_j) possible.

Definition 4.1 is weaker than Definition 2.3 since it only requires that each choice should appear in the belief of θ_i but does not require that it should be paired with each belief of j deemed possible by θ_i. Nevertheless, we will show in Lemma 4.2 that in with other conditions in this characterization it leads to caution.

Definition 4.2 (Primary belief in u). Consider a static game form $G = (C_i)_{i\in I}$, a lexicographic epistemic model $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ for G with incomplete information, and a pair $u = (u_i)_{i\in I}$ of utility functions. $\theta_i \in \Theta_i$ *primarily believes in* u if θ_i's level-1 belief only assigns positive probability to (c_j, θ_j) with $w_j(\theta_j) = u_j$.

Primary belief in u is stronger than Definition 2.5 (5.2). (5.2) allows the occurance of a type with a utility function which is "very similar" (but not equal) to u_j in the level-1 belief of θ_i, while primary belief in u only allows types with utility function u_j there.

The characterization result is as follows.

Proposition 4.1 (An alternative characterization of permissibility). Consider a finite 2-person static game $\Gamma = (C_i, u_i)_{i\in I}$, the corresponding game form $G = (C_i)_{i\in I}$, and a finite lexicographic epistemic model $M^{co} = (T_i, b_i)_{i\in I}$ for Γ.

Then, $c_i^* \in C_i$ is permissible in M^{co} if and only if there is some finite lexicographic epistemic model $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ with incomplete information for G and some $\theta_i^* \in \Theta_i$ with $w_i(\theta_i^*) = u_i$ such that

(a) c_i^* is rational for θ_i^*, and,

(b) θ_i^* expresses common full belief in caution, rationality, and primary belief in u.

The Only-if part follows directly from Observation 3.2, Lemma 3.1, and the following lemma whose proof can be found in Section 5.

Lemma 4.1 (Primary belief in rationality → Primary belief in u). Let $M^{co} = (T_i, b_i)_{i\in I}$ and $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ be constructed from M^{co} by the two steps above. Then if $t_i \in T_i$ expresses common full belief in primary belief in rationality, then each $\theta_i \in \Theta_i(t_i)$ expresses common full belief in primary belief in u.

To show the If part, we need first to show that weak caution is enough for the characterization. Here, we show that the corresponding concept in complete information model can replace caution and characterize permissibility. Then we can use the mapping between complete and incomplete information models constructed in Section 3.2. Let $M^{co} = (T_i, b_i)_{i\in I}$ be a lexicographic model for $\Gamma = (C_i, t_i)_{i\in I}$ with complete information. $t_i \in T_i$ is *weakly cautious* iff for each $c_j \in C_j$, there is some $t_j \in T_j$ such that t_i deems (c_j, t_j) possible. We have the following lemma.

Lemma 4.2 (Characterizing permissibility by weak caution). Let $M^{co} = (T_i, b_i)_{i\in I}$ be a lexicographic epistemic model for a game $\Gamma = (C_i, u_i)_{i\in I}$. $c_i^* \in C_i$ is permissible if and only if it is rational to some $t_i^* \in T_i$ which expresses common full belief in weak caution and primary belief in rationality.

Also, we need the following lemmas.

Lemma 4.3 (Weak caution^{in} → weak caution^{co}). Let $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ and $M^{co} = (T_i, b_i)_{i\in I}$ be constructed from M^{in} by the above approach. If $\theta_i \in \Theta_i$ expresses common full belief in weak caution, so does $t_i(E_i(\theta_i))$.

We omit the proof of Lemma 4.3 since it can be shown in a similar way as Lemma 3.3.

Lemma 4.4 (Rationality + primary belief in u → Primary belief in rationality). Let $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$ and $M^{co} = (T_i, b_i)_{i\in I}$ be constructed from M^{in} by the above approach. If $\theta_i \in \Theta_i$ expresses common full belief in rationality and primary belief in u, then $t_i(E_i(\theta_i))$ expresses common full belief in primary belief in rationality.

Proof of the If part of Proposition 4.1. Let $M^{in} = (\Theta_i, w_i, \beta_i)_{i\in I}$, $M^{co} = (T_i, b_i)_{i\in I}$ be constructed from M^{in} by the above approach, and $c_i^* \in C_i$ be rational for some θ_i^* with $w_i(\theta_i^*) = u_i$ which expresses common full belief in caution, rationality, and primary belief in u. Consider $t_i(E_i(\theta_i^*))$. Since $w_i(\theta_i^*) = u_i$ and $b_i(t_i(E_i(\theta_i^*)))$ has the same distribution on j's choices at each level as $\beta_i(\theta_i^*)$, c_i^* is rational for $t_i(E_i(\theta_i^*))$. By Lemmas 4.3 and 4.4, $t_i(E_i(\theta_i^*))$ expresses common full belief in weak caution and primary belief in rationality. Also, by Lemma 4.2 $t_i(E_i(\theta_i^*))$ expresses common full belief in caution. Hence c_i^* is permissible in Γ. //

It should be noted that caution cannot be weakened in the characterization of Theorems 3.1 and 3.2. For Theorem 3.1, caution plays an important role in the proof of the If part; without it, primary belief in utilities nearest to u and that a best choice is supported by utilities nearest to u cannot imply primary belief in rationality. For Theorem 3.2, the interpolation method used in the proof of Lemma 4.2 may not work since different types may have different orders there.

An open question is that whether there is a characterization of proper rationalizability by using rationality. More work needs to be done on it.

5. Proofs

Proof of Lemma 2.1. We construct a sequence satisfying (a)-(c) by induction. First, let $v_{i1} = u_i$. Suppose that for some $\ell \in \{1, ..., L-1\}$ we have defined $v_{i1}, ..., v_{i\ell}$ satisfying (a)-(c). Now we show how to define $v_{i,\ell+1}$. It can be seen that there exists $M_{\ell+1} > 0$ such that $v_{i\ell}(c_{i,\ell+1}, \beta_{i1}) + M_{\ell+1} > v_{i\ell}(c_{i\ell}, \beta_{i1})$ for all $c_{i\ell} \in C_{i\ell}$ and $c_{i,\ell+1} \in C_{i,\ell+1}$. We define $v_{i,\ell+1}$ as follows: for each $(c_i, c_j) \in C$,

$$v_{i,\ell+1}(c_i, c_j) = \begin{cases} v_{i\ell}(c_i, c_j) + M_{\ell+1} & \text{if } c_i \in C_{i,\ell+1} \text{ and } c_j \in \text{supp}\beta_{i1} \\ v_{i\ell}(c_i, c_j) & \text{otherwise} \end{cases}$$

It can be seen that each $c_{i,\ell+1} \in C_{i,\ell+1}$ is rational at $v_{i,\ell+1}$ under β_i. Also, since $d(v_{i,\ell+1}, v_{i\ell}) = (M^2_{\ell+1} \times |C_{i,\ell+1}| \times |\text{supp}\beta_{i1}|)^{1/2} > 0$, $d(v_{i,\ell+1}, u_i) = d(v_{i,\ell+1}, v_{i\ell}) + d(v_{in}, u_i) > d(v_{in}, u_i)$. By induction, we can obtain a sequence $v_{i1}, ..., v_{iL} \in V_i$ satisfying (a)-(c). //

It should be noted that, given u_i and β_i, the sequence $v_{i1}, ..., v_{iL}$ satisfying (a)-(c) is not unique. The basic idea behind this inductive construction is depicted as follows. Suppose that $u_i(c_{i1}, \beta_i) > u_i(c_{i2}, \beta_i) > ... > u_i(c_{iN}, \beta_i)$, that is, $\Pi_i(\beta_i) = (\{c_{i1}\}, \{c_{i2}\}, ..., \{c_{iN}\})$, then

$$(c_{i1}, c_{i2}, c_{i3}, ..., c_{iN}) \underset{\longrightarrow}{v_{i2}} (c_{i2}, c_{i1}, c_{i3}, ..., c_{iN},) \underset{\longrightarrow}{v_{i3}} (c_{i3}, c_{i2}, c_{i1} ..., c_{iN}) \; ... \; \underset{\longrightarrow}{v_{iN}} (c_{iN}, c_{i,N-1}, ..., c_{i1})$$

Informally speaking, we take equivalent classes of choices one by one to the foremost location of the sequence according to the order of preference in u_i under β_i. The following example shows how this construction works.

Example 5.1. Consider u_1 in Example 2.1. Under the lexicographic belief $\beta_1 = (D, E, F)$, A is preferred to B and B is preferred to C in u_1, that is, $\Pi_1(\beta_1) = (\{A\}, \{B\}, \{C\})$. We can define v_{11}, v_{12}, v_{13} as follows:

$u_1 = v_{11}$	D	E	F
A	1	1	1
B	1	1	0
C	1	0	1

$\longrightarrow$

v_{12}	D	E	F
A	1	1	1
B	2	1	0
C	1	0	1

$\longrightarrow$

v_{13}	D	E	F
A	1	1	1
B	2	1	0
C	3	0	1

At v_{11}, the order of preferences is (A, B, C) under β_1, at v_{12} it is (B, A, C), and at v_{13} it is (C, B, A).

Proof of Lemma 3.1. We show this statement by induction. First we show that if t_i is cautious, then each $\theta_i \in \Theta_i(t_i)$ is also cautious. Let $c_j \in C_j$ and $\theta_j \in \Theta_j(\theta_i)$. By construction, it can be seen that the type $t_j \in T_j$ satisfying the condition that $\theta_j \in \Theta_j(t_j)$ is in $T_j(t_i)$. Since t_i is cautious, t_i deems (c_j, t_j) possible. Consider the pair (c_j, θ'_j) in $\beta_i(\theta_i)$ corresponding to (c_j, t_j). Since both θ_j and θ'_j are in $\Theta_j(t_j)$, it follows from Observation 3.1 that $\beta_j(\theta_j) = \beta_j(\theta'_j)$. Hence $(c_j, \theta_j^{w_j(\theta'_j)})$ is deemed possible by θ_i. Here we have shown that θ_i is cautious.

Suppose we have shown that, for each $i \in I$, if t_i expresses n-fold full belief in caution then so does each $\theta_i \in \Theta_i(t_i)$. Now suppose that t_i expresses $(n+1)$-fold full belief in caution, i.e., each $t_j \in T_j(t_i)$ expresses n-fold full belief in caution. By construction, for each $\theta_i \in \Theta_i(t_i)$ and each $\theta_j \in \Theta_j(\theta_i)$ there is some $t_j \in T_j(t_i)$ such that $\theta_j \in \Theta_j(t_i)$, and, by inductive assumption, each $\theta_j \in \Theta_j(\theta_i)$ expresses n-fold full belief in caution. Therefore, each $\theta_i \in \Theta_i(t_i)$ expresses $(n+1)$-fold full belief in caution. //

Proof of Lemma 3.2. We show this statement by induction. First we show that if t_i primarily believes in j's rationality, then each $\theta_i \in \Theta_i(t_i)$ primarily believes in utilities nearest to u. Let (c_j, θ_j) be a pair deemed possible in the level-1 belief of θ_i. Consider its correspondence (c_j, t_j) in level-1 belief of t_i. Since t_i primarily believes in j's rationality, c_j is rational for t_j. It follows that $c_j \in C_{j1} \in \Pi_j(t_j)$. By Lemma 2.1 and construction, it follows that $w_j(\theta_j) = u_j$. Since u_j is the nearest function to itself among all utlity functions in V_j, we have shown that θ_i primarily believes in utilities nearest to u.

Suppose we have shown that, for each $i \in I$, if t_i expresses n-fold full belief in primary belief in rationality then each $\theta_i \in \Theta_i(t_i)$ expresses n-fold full belief in primary belief in utilities nearest to u. Now suppose that t_i expresses $(n+1)$-fold full belief in primary belief in rationality, i.e., each $t_j \in T_j(t_i)$ expresses n-fold full belief in primary belief in rationality. Since, by construction, for each $\theta_i \in \Theta_i(t_i)$ and each $\theta_j \in \Theta_j(\theta_i)$ there is some $t_j \in T_j(t_i)$ such that $\theta_j \in \Theta_j(t_j)$, it follows that, by inductive assumption, each $\theta_j \in \Theta_j(\theta_i)$ expresses n-fold full belief in primary belief in utilities nearest to u. Therefore, each $\theta_i \in \Theta_i(t_i)$ expresses $(n+1)$-fold full belief in primary belief in utilities nearest to u. //

Proof of Lemma 3.3. We show this statement by induction. First we show that if θ_i is cautious, then $t_i(E_i(\theta_i))$ is also cautious. Let $c_j \in C_j$ and $t_j \in T_j(t_i(E_i(\theta_i)))$. By construction, $t_j = t_j(E_j)$ for some $E_j \in \mathbb{E}_j$, and there is some $\theta_j \in E_j$ which is deemed possible by θ_i. Since θ_i is cautious, there is some θ'_j with $\beta_j(\theta'_j) = \beta_j(\theta_j)$, i.e., $\theta'_j \in E_j$, such that (c_j, θ'_j) is deemed possible by θ_i. By construction, (c_j, t_j) is deemed possible by $t_i(E_i(\theta_i))$.

Suppose we have shown that, for each $i \in I$, if θ_i expresses n-fold full belief in caution then so does $t_i(E_i(\theta_i))$. Now suppose that θ_i expresses $(n+1)$-fold full belief in caution, i.e., each $\theta_j \in \Theta_j(\theta_i)$ expresses n-fold full belief in caution. Since, by construction, for each $t_j \in T_j(t_i(E_i(\theta_i)))$, there is some $\theta_j \in \Theta_j(\theta_i)$ such that $t_j = t_j(E_j(\theta_j))$, by inductive assumption t_j expresses n-fold full belief in caution. Therefore, $t_i(E_i(\theta_i))$ expresses $(n+1)$-fold full belief in caution. //

Proof of Lemma 3.4. We show this statement by induction. First we show that if θ_i is cautious, primarily believes in utilities nearest to u, and believes in that a best choice is supported by utilities nearest to u, then $t_i(E_i(\theta_i))$ primarily believes in j's rationality. Let (c_j, t_j) be a choice-type pair which is deemed possible in $t_i(E_i(\theta_i))$'s level-1 belief. By construction $t_j = t_j(E_j)$ for some $E_j \in \mathbb{E}_j$, and for some $\theta_j \in E_j$, (c_j, θ_j) is deemed possible in θ_i's level-1 belief. Since θ_i primarily believes in utilities nearest to u, it follows that

$$d(w_j(\theta_j), u_j) \leq d(w_j(\theta'_j), u_j) \text{ for all } \theta'_j \in E_j. \tag{5.1}$$

Suppose that c_j is not optimal for t_j. Let c'_j be a choice optimal to t_j. Since θ_i is cautious, there is some $\theta_j^{v_j} \in E_j$ such that $(c_j, \theta_j^{v_j})$ is deemed possible by θ_i. Then since θ_i believes in that a best choice is supported by utilities nearest to u, it follows that $d(\theta_j^{v_j}, u_j) < d(w_j(\theta_j), u_j)$, which is contradictory to (5.1). Therefore c_j is optimal for t_j. Here we have shown that $t_i(E_i(\theta_i))$ primarily believes in j's rationality.

Suppose we have shown that, for each $i \in I$, if θ_i expresses n-fold full belief in caution, primary belief in utilities nearest to u, and that a best choice is supported by utilities nearest to u, then $t_i(E_i(\theta_i))$ expresses n-fold belief in primary belief in rationality. Now suppose that θ_i expresses $(n+1)$-fold full belief in caution, primary belief in utilities nearest to u, and that a best choice is supported by utilities nearest to u, i.e., each $\theta_j \in \Theta_j(\theta_i)$ expresses n-fold full belief in caution, primary belief in utilities nearest to u, and that a best choice is supported by utilities nearest to u. Since, by construction, for each $t_j \in T_j(t_i(E_i(\theta_i)))$, there is some $\theta_j \in \Theta_j(\theta_i)$ such that $t_j = t_j(E_j(\theta_j))$, by inductive assumption t_j expresses n-fold full belief in primary belief in rationality. Therefore, $t_i(E_i(\theta_i))$ expresses $(n+1)$-fold full belief in primary belief in rationality. //

Proof of Lemma 3.5. We show this statement by induction. First we show that if t_i is caution and respects j's preferences, then each $\theta_i \in \Theta_i(t_i)$ expresses u-centered belief. It can be seen that if t_i is cautious and respects j's preferences, then we can combine all types deemed possible by t_i with the same belief into one type without hurting the caution and respect of j's preference, and every choice optimal for t_i is still optimal for this new type and vice versa. Therefore, without loss of generality we can assume that for each $t_j, t'_j \in T_j$, $b_j(t_j) \neq b_j(t'_j)$. Let $c_j, c'_j \in C_j$, $\theta_j \in \Theta_j$, and $v_j, v'_j \in V_j$ such that $(c_j, \theta_j^{v_j})$ and $(c'_j, \theta_j^{v'_j})$ are deemed possible by θ_i with $d(v_j, u_j) < d(v'_j, u_j)$. Since each type in T_i has a distinct lexicographic belief, it follows that $\theta_j^{v_j}, \theta_j^{v'_j} \in \Theta_j(t_j)$ for some $t_j \in T_j$. By construction it follows that (1) t_i deems both (c_j, t_j) and

(c'_j, t_j) possible, and (2) $u_j(c_j, t_i) > u_j(c'_j, t_i)$. Since t_i respects j's preferences, t_i deems (c_j, t_j) infinitely more likely than (c'_j, t_j), which corresponds to that θ_i deems $(c_j, \theta_j^{v_j})$ infinitely more likely than $(c'_j, \theta_j^{v'_j})$. Here we have shown that θ_i expresses u-centered belief.

Suppose we have shown that, for each $i \in I$, if t_i expresses n-fold full belief in respect of preferences then each $\theta_i \in \Theta_i(t_i)$ expresses n-fold full belief in u-centered belief. Now suppose that t_i expresses $(n+1)$-fold full belief in respect of preferences, i.e., each $t_j \in T_j(t_i)$ expresses n-fold full belief respect of preferences. Since, by construction, for each $\theta_i \in \Theta_i(t_i)$ and each $\theta_j \in \Theta_j(\theta_i)$ there is some $t_j \in T_j(t_i)$ such that $\theta_j \in \Theta_j(t_j)$, by inductive assumption it follows that each $\theta_j \in \Theta_j(\theta_i)$ expresses n-fold full belief in u-centered belief. Therefore, each $\theta_i \in \Theta_i(t_i)$ expresses $(n+1)$-fold full belief in u-centered belief. //

Proof of Lemma 3.6. We show this statement by induction. First we show that if θ_i is cautious, has a u-centered belief, and believes that a better choice is supported by utilities nearer to u, then $t_i(E_i(\theta_i))$ respects j's preferences. First, since θ_i is cautious, By Lemma 3.3, $t_i(E_i(\theta_i))$ is also cautious. Let $c_j, c'_j \in C_j$ and $t_j \in T_j(t_i(E_i(\theta_i)))$ with t_j prefers c_j to c'_j. By construction $t_j = t_j(E_j)$ for some $E_j \in \mathbb{E}_j$, and, since θ_i is cautious, there are $\theta_j, \theta'_j \in E_j$ such that θ_i deems (c_j, θ_j) and (c'_j, θ'_j) possible. Since $\beta_j(\theta_j) = \beta_j(\theta'_j)$ and θ_j has the same probability distribution over C_i at each level as t_j, it follows that $u_j(c_j, \theta_j) > u_j(c'_j, \theta_j)$. Since θ_i believes that a better choice is supported by utilities nearer to u, it follows that $d(w_j(\theta_j), u_j) < d(w_j(\theta'_j), u_j)$. Since θ_i has a u-centered belief, it follows that θ_i deems (c_j, θ_j) infinitely more likely than (c'_j, θ'_j), which implies that $t_i(E_i(\theta_i))$ deems (c_j, t_j) infinitely more likely than (c'_j, t_j). Therefore, $t_i(E_i(\theta_i))$ respects j's preferences.

Suppose we have shown that, for each $i \in I$, if θ_i expresses n-fold full belief in caution, u-centered belief, and that a better choice is supported by utilities nearer to u, then $t_i(E_i(\theta_i))$ expresses n-fold full belief in respect of preferences. Now suppose that θ_i expresses $(n+1)$-fold full belief in caution, u-centered belief, and that a better choice is supported by utilities nearer to u, i.e., each $\theta_j \in \Theta_j(\theta_i)$ expresses n-fold full belief in caution, u-centered belief, and that a better choice is supported by utilities nearer to u. Since, by construction, for each $t_j \in T_j(t_i(E_i(\theta_i)))$, there is some $\theta_j \in \Theta_j(\theta_i)$ such that $t_j = t_j(E_j(\theta_j))$, by inductive assumption t_j expresses n-fold full belief in respect of preferences. Therefore, $t_i(E_i(\theta_i))$ expresses $(n+1)$-fold full belief in respect of preferences. //

Proof of Lemma 4.1. We show this statement by induction. First we show that if t_i primarily believes in j's rationality, then each $\theta_i \in \Theta_i(t_i)$ primarily believes in u. Let (c_j, θ_j) be a pair deemed possible in the level-1 belief of θ_i. Consider its corresponding (c_j, t_j) in level-1 belief of t_i. Since t_i primarily believes in j's rationality, c_j is rational for t_j. It follows that $c_j \in C_{j1} \in \Pi_j(t_j)$. By construction, it follows that $w_j(\theta_j) = u_j$. Here we have shown that θ_i primarily believes in u.

Suppose we have shown that, for each $i \in I$, if t_i expresses n-fold full belief in primary belief in rationality then each $\theta_i \in \Theta_i(t_i)$ expresses n-fold full belief in primary belief in u. Now suppose that t_i expresses $(n+1)$-fold full belief in primary belief in rationality, i.e., each $t_j \in T_j(t_i)$ expresses n-fold full belief in primary belief in rationality. Since, by construction, for each $\theta_i \in \Theta_i(t_i)$ and each $\theta_j \in \Theta_j(\theta_i)$ there is some $t_j \in T_j(t_i)$ such that $\theta_j \in \Theta_j(t_j)$, it follows that, by inductive assumption, each $\theta_j \in \Theta_j(\theta_i)$ expresses n-fold full belief in primary belief in rationality. Therefore, each $\theta_i \in \Theta_i(t_i)$ expresses $(n+1)$-fold full belief in primary belief in u. //

Proof of Lemma 4.2. To show the if part, we need first to show that each weak cautious type can be extended into a cautious one without changing the set of choices rational for it. It is done by an interpolation method as follows. Let t_i be a type satisfying weak caution with $b_i(t_i) = (b_{i1}, ..., b_{iK})$, $c_j \in C_j$, and $t_j \in T_j(t_i)$. Suppose that (c_j, t_j) is not deemed possible by t_i. Since t_i is weakly cautious, there is some $t'_j \in T_j$ such that for some $k \in \{1, ..., K\}$, $b_{ik}(c_j, t'_j) > 0$. Now we extend $(b_{i1}, ..., b_{iK})$ into $(b'_{i1}, ..., b'_{i,K+1})$ by letting (1) $b'_{it} = b_{it}$ for each $t \leq k$, (2) $b'_{it} = b_{i,t-1}$ for each $t > k+1$, and (3) $b'_{i,k+1}$ is obtained by replacing every occurrence of (c_j, t'_j) by (c_j, t_j) in the distribution of b_{ik}. We call $b'_{i,k+1}$ a *doppelganger* of b_{ik}. It can be seen that for each $c_i \in C_i$, and a doppelganger $b'_{i,k+1}$ of b_{ik}, $u_i(c_i, b'_{i,k+1}) = u_i(c_i, b_{ik})$. By repeatedly interpolating doppelgangers into $b_i(t_i)$ for each missed choice-type pairs, finally we obtain a lexicographic belief $(b'_{i1}, ..., b'_{iK'})$ that satisfies caution. We use $\bar{t}_i$ to denote the type with belief $(b'_{i1}, ..., b'_{iK'})$. $\bar{t}_i$ is called a *cautious extension* of t_i. We have the following lemma.

Lemma 4.1 (Extended type preserves rational choices). Let t_i be a weakly cautious type and $\bar{t}_i$ a cautious extension of t_i. Then $c_i \in C_i$ is rational for t_i if and only if it is rational for $\bar{t}_i$.

Proof. (Only-if) Suppose that c_i is not rational for $\bar{t}_i$. Then there is some $c'_i \in C_i$ which is preferred c_i under $b_i(\bar{t}_i) = (b'_{i1}, ..., b'_{iK'})$, that is, there is some $k' \in \{0, ..., K'\}$ such that $u_i(c_i, b'_{i\ell}) = u_i(c'_i, b'_{i\ell})$ for each $\ell \leq k'$ and $u_i(c_i, b_{i,k'+1}) < u_i(c'_i, b_{i,k'+1})$. Let $b_{i,k+1}$ be the entry in $b_i(t_i)$ such that $b'_{i,k'+1}$ is its doppelganger. It follows that in the original $b_i(t_i) = (b_{i1}, ..., b_{iK})$, $u_i(c_i, b_{i\ell}) = u_i(c'_i, b_{i\ell})$ for each $\ell \leq k$ and $u_i(c_i, b_{i,k+1}) < u_i(c'_i, b_{i,k+1})$. Hence c_i is not rational for t_i.

(If) Suppose that c_i is not rational for t_i. Then there is some $c'_i \in C_i$ which is preferred c_i under $b_i(t_i) = (b_{i1}, ..., b_{iK})$, that is, there is some $k \in \{0, ..., K\}$ such that $u_i(c_i, b_{i\ell}) = u_i(c'_i, b_{i\ell})$ for each $\ell \leq k$ and $u_i(c_i, b_{i,k+1}) < u_i(c'_i, b_{i,k+1})$. Let $b'_{i,k'+1}$ be the corresponding doppelganger in $b_i(\bar{t}_i)$ to $b_{i,k+1}$. It follows that in the original $u_i(c_i, b'_{i\ell}) = u_i(c'_i, b'_{i\ell})$ for each $\ell \leq k'$ and $u_i(c_i, b'_{i,k'+1}) < u_i(c'_i, b'_{i,k'+1})$. Hence c_i is not rational for $\bar{t}_i$. //

Proof of Lemma 4.2 (Continued) Since caution implies weak caution, the Only-if part holds automatically. For the If part, suppose that c^*_i is rational for some $t^*_i \in T_i$ which expresses common full belief in weak caution and primary belief in rationality. Consider an epistemic model $(\overline{T}_i, \bar{b}_i)_{i \in I}$ such that for each $i \in I$, $\overline{T}_i = \{\bar{t}_i : t_i \in T_i\}$ and $\bar{b}_i(\bar{t}_i)$ is a cautious extension of $b_i(t_i)$ with replacing each occurrence of t_j by $\bar{t}_j$. By Lemma 4.1, since c^*_i is rational for t^*_i, it is also rational for $\overline{t^*_i}$. Also, it can be seen by construction that $\overline{t^*_i}$ expresses common full belief in caution. Also, since the interpolation always put doppelgangers after the original one, it does not change the level-1 belief, and consequently $\overline{t^*_i}$ expresses common full belief in primary belief in rationality. Therefore, c^*_i is permissible. //

Proof of Lemma 4.4. We show this statement by induction. First we show that if θ_i believes in j's rationality and primarily believes in u, then $t_i(E_i(\theta_i))$ primarily believes in j's rationality. Let (c_j, t_j) be a choice-type pair which is deemed possible in $t_i(E_i(\theta_i))$'s level-1 belief. By construction $t_j = t_j(E_j)$ for some $E_j \in \mathbb{E}_j$, and for some $\theta_j \in E_j$, (c_j, θ_j) is deemed possible in θ_i's level-1 belief. Since θ_i primarily believes in u, it follows that $w_j(\theta_j) = u_j$. Also, since θ_i believes j's rationality, it follows that c_j is rational at u_j under $\beta_j(\theta_j)$, i.e., $b_i(t_j)$. Therefore c_j is rational for t_j. Here we have shown that $t_i(E_i(\theta_i))$ primarily believes in j's rationality.

Suppose we have shown that, for each $i \in I$, if θ_i expresses n-fold full belief in rationality and primary belief in u, then $t_i(E_i(\theta_i))$ expresses n-fold belief in primary belief in rationality. Now suppose that θ_i expresses $(n+1)$-fold full belief in rationality and primary belief in u, i.e., each $\theta_j \in \Theta_j(\theta_i)$ expresses n-fold full belief in rationality and primary belief in u. Since, by construction, for each $t_j \in T_j(t_i(E_i(\theta_i)))$, there is some $\theta_j \in \Theta_j(\theta_i)$ such that $t_j = t_j(E_j(\theta_j))$, by inductive assumption t_j expresses n-fold full belief in primary belief in rationality. Therefore, $t_i(E_i(\theta_i))$ expresses $(n+1)$-fold full belief in primary belief in rationality. //

References

[1] Asheim, G.B. (2001). Proper rationalizability in lexicographic beliefs. *International Journal of Game Theory* **30**, 453-478.

[2] Battigalli, P. (2003). Rationalizability in infinite, dynamic games of incomplete information. *Research in Economics* **57**, 1-38.

[3] Battigalli, P., Siniscalchi, M. (2003). Rationalization and incomplete information. *B.E.Journal of Theoretical Economics* **61**, 165-184.

[4] Battigalli, P., Siniscalchi, M. (2007). Interactive epistemology in games with payoff uncertainty. *Research in Economics* **3**, 1534-5963.

[5] Blume, L., Brandenburger, A., Dekel, E. (1991). Lexicographic probabilities and choice under uncertainty. *Econometrica* 59, 61-79.

[6] Blume, L., Brandenburger, A., Dekel, E. (1991). Lexicographic probabilities and equilibrium refinements. *Econometrica* 59, 81-98.

[7] Böge, W., Eisele, T. (1979). On solutions of Bayesian games. *International Journal of Game Theory* **8**, 193-215.

[8] Börgers, T. (1994). Weak dominance and approximate common knowledge. *Journal of Economic Theory* **64**, 265-276.

[9] Börgers, T. and Samuelson, L. (1992). "Cautious" utility maximization and iterated weak dominance. *International Journal of Game Theory* **21**, 13-25.

[10] Brandenburger, A. (1992). Lexicographic probabilities and iterated admissibility. In *Economic Analysis of Markets and Games*, ed. P. Dasgupta, et al. MIT Press, 282-290.

[11] Dekel, E., Fudenberg, D. (1990). Rational behavior with payoff uncertainty. *Journal of Economic Theory* **52**, 243-267.

[12] Dekel, E., Siniscalchi, M. (2015). Epistemic game theory. In *Handbooks of Game Theory with Economic Applications*, Vol. 4, edited by Young P.H., Zamir, S, 619-702.

[13] Heifetz, A., Samet, D. (1998). Topology-free typology of beliefs. *Journal of Economic Theory* **82**, 324-341.

[14] Myerson, R.B. (1978). Refinements of Nash equilibrium concept. *International Journal of Game Theory* **7**, 73-80.

[15] Perea, A. (2012). *Epistemic Game Theory: Reasoning and Choice*, Cambridge University Press.

[16] Perea, A., Kets, W. (2016). When do types induce the same belief hierarchy? *Games* **7**, 28. http://dx.doi.org/10.3390/g7040028.

[17] Perea, A., Roy, S. (2017) A new epistemic characterization of ε-proper rationalizability. *Games and Economic Behavior* **104**: 309-328.

[18] Samuelson, L. (1992). Dominated strategies and common knowledge. *Games and Economic Behavior* **4**: 284-313.

[19] Schuhmacher, F. (1999). Proper rationalizability and backward induction. *International Journal of Game Theory* **28**, 599-615.

[20] Selten, R. (1978). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory* **4**, 25-55.