Incomplete Information and Generalized Iterated Strict Dominance*

CHRISTIAN W. BACH** AND ANDRÉS PEREA***

c.w.bach@liverpool.ac.uk   a.perea@maastrichtuniversity.nl

Abstract. In games with incomplete information, players face uncertainty about the opponents' utility functions. We follow Harsanyi's [*Management Science* **14**, 3 & 5 & 7 (1967-68)] one-person perspective approach to modelling incomplete information. Moreover, our formal framework is kept as basic and parsimonious as possible, to render the theory of incomplete information accessible to a broad spectrum of potential applications. In particular, we formalize common belief in rationality and provide an algorithmic characterization of it in terms of decision problems, which gives rise to the non-equilibrium solution concept of generalized iterated strict dominance.

Keywords: algorithms, common belief in rationality, epistemic game theory, generalized iterated strict dominance, incomplete information, interactive epistemology, interim correlated rationalizability, one-person perspective, solution concepts, static games.

---

* We are grateful to audiences at the National University of Singapore, Maastricht University, University of Liverpool, and the University of Amsterdam for useful and constructive comments.

** Department for Economics, Finance and Accounting, University of Liverpool Management School, Chatham Street, Liverpool, UK-L69 7ZH; EPICENTER, School of Business and Economics, Maastricht University, NL-6200 MD Maastricht

*** Department of Quantitative Economics, School of Business and Economics, Maastricht University, NL-6200 MD Maastricht; EPICENTER, School of Business and Economics, Maastricht University, NL-6200 MD Maastricht

1 Introduction

In games with incomplete information, players do not only face uncertainty about their opponents' choices but also about their utility functions. The analysis of this class of games has been pioneered by Harsanyi (1967-68). In particular, his framework is constructed on the basis of a one-person perspective. Accordingly, the strategic situation is analyzed entirely from the viewpoint of a single player. For instance, as Harsanyi (1967-68, p. 170) writes it is some

> [...] *player j (from whose point of view we are analyzing the game)* [...],

and Harsanyi (1967-68, p. 175) states that

> [...] *we are interested only in the decision rules that player j himself will follow* [...].

Conceptually, a one-person perspective approach treats game theory as an interactive extension of decision theory.

Here, we also take a one-person, hence strictly decision-theoretic, approach to game theory, and model common belief in rationality within the mind of a single player as well as define a corresponding non-equilibrium solution concept – generalized iterated strict dominance – in terms of decision problems. The formal framework is kept as simple and parsimonious as possible, to render the theory of incomplete information games accessible to a potentially vast field of applications beyond economics.

The standard solution concept for static games with incomplete information has been Harsanyi's (1967-68) Bayesian equilibrium. [1] Recently, the idea of rationalizability – due to Bernheim (1984) and Pearce (1984) – has been generalized to incomplete information games. In particular, the solution concepts of weak and strong Δ-rationalizability have been introduced by Battigalli (2003) for dynamic games, and further analyzed by Battigalli and Siniscalchi (2003a)

[1] The solution concept of Bayesian equilibrium is analyzed epistemically by Bach and Perea (2017).

and (2007), Battigalli et al. (2011), Battigalli and Prestipino (2013), as well as Dekel and Siniscalchi (2015). Intuitively, Δ-rationalizability concepts iteratively delete strategy utility pairs by some best response requirement, and allow for exogenous restrictions on the first-order beliefs. Δ-rationalizability has been applied to auctions by Battigalli and Siniscalchi (2003b), to signaling games by Battigalli (2006), as well as to static implementation by Ollar and Penta (2017). Furthermore, a backward inductive variant of rationalizability for dynamic games with incomplete information has been proposed by Penta (2017) and applied to dynamic implementation by Penta (2015). A different incomplete information generalization of rationalizability has been proposed by Ely and Pęski (2006)'s interim rationalizability as well as by Dekel et al. (2007)'s interim correlated rationalizability, respectively. The essential difference to Δ-rationalizability lies in fixing the belief hierarchies on utilities. The robustness of interim correlated rationalizability with regards to perturbations of the belief hierarchies on utilities is studied by Weinstein and Yildiz (2007) as well as by Penta (2013) for static games and by Penta (2012) for dynamic games. Besides, note that all incomplete information rationalizability concepts employ a modeller – and not a one-person – perspective in their formal frameworks.

Generalized iterated strict dominance based on a one-person perspective augments the class of solution concepts for incomplete information games. Intuitively, the algorithm iteratively reduces decision problems by some strict dominance requirement. In contrast to the Δ-rationalizability concepts in the literature, generalized iterated strict dominance is formulated in a one-person perspective by means of decision problems and uses strict dominance arguments instead of best-response arguments. Moreover, it is attempted to keep the formalization as simple as possible. In fact, Battigalli and Siniscalchi (1999) as well as Battigalli (2003) already indicate that Δ-rationalitzability concepts are equivalent to iterated strict dominance procedures for the class of static games. Also, Battigalli et al. (2011) point out that their belief-free rationalizability concept can be characterized by an iterated strict dominance procedure. Besides, we show that generalized iterated strict dominance is behaviourally equivalent to iterated strict dominance once complete information is imposed. Hence, our algorithm can be viewed as a direct generalization of iterated strict dominance from complete to incomplete information games.

In epistemic game theory the central concept is common belief in rationality. For static games with complete information common belief in rationality has been extensively studied and is well understood. In the case of complete information, rationalizability concepts occured first and were introduced by Bernheim (1984) and Pearce (1984). Only later, common belief in rationality was spelled out and connected to rationalizability by Brandenburger and Dekel (1987) as well as by Tan and Werlang (1988). Similarly, for the more general class of static games with incomplete information common belief in rationality has only appeared after the generalized rationalizability concepts. Common belief in rationality has been formalized and employed in different forms for epistemic foundations of the Δ-rationalizability variants by Battigalli and Siniscalchi (1999), (2002), and (2007), Battigalli et al. (2011), as well as Battigalli and Prestipino (2013). Besides, Battigalli et al. (2011) also give an epistemic foundation of interim correlated rationalizability. Also, the literature on common belief in rationality for incomplete information games so far share the use of a modeller perspective.

We propose a formalization of common belief in rationality based on Harsanyi's (1967-68) one-person perspective approach and which is algorithmically characterized by generalized iterated strict dominance. If the belief hierarchies on utilities are kept fixed, then common belief in rationality is behaviourally equivalent to interim correlated rationalizability. In our model the restrictions only concern the belief hierarchies of a single player – the reasoner. In particular, a one-person perspective epistemic framework does not need to introduce states as modeller perspective approaches do. Also, in line with Harsanyi (1967-68) we treat strategic uncertainty and payoff uncertainty symmetrically. Furthermore, rational choice under common belief in rationality is lean in the sense that it does not fix an epistemic model, but rather uses different epistemic models as a way to encode different belief hierarchies. Besides, the epistemic model is kept as basic and parsimonious as possible, in order to maximize accessibility for potential applications.

From a conceptual point of view, we treat the reasoning as foundational and hence prior to the corresponding algorithm which gives rise to a solution concept in the classical sense. Accordingly, the reasoning concept of common belief in rationality within the framework of an epistemic model is constructed first. Only thereafter an incomplete information generalization of iterated strict dominance in terms of decision problems is conducted and shown to characterize reasoning in line with common belief in rationality.

The two notions for incomplete information considered here, i.e. common belief in rationality and its algorithmic analogue generalized iterated strict dominance, can be relevant for numerous applications. In particular, the formal framework is kept as basic and lean as possible, to facilitate and stimulate the use of the concepts for concrete economic problems. In particular, the illustration of the concepts in our examples underlines the accessibility of our framework for applied work. For instance, in pricing games firms may have no information about their competitors' characteristics such as their cost structures. Furthermore, in auctions participants can be uncertain about each others' valuations, which is indeed typically assumed in public auctions or internet auctions. More generally, incomplete information settings of mechanism design or implementation could be considered with the non-equilibrium concept generalized iterated strict dominance. Beyond applications in economics, a basic framework of analysis for incomplete information could also be of use in other fields of strategic enquiry such as management or political theory.

We proceed as follows. In Section 2, the epistemic framework for games with incomplete information and a one-person perspective is formally defined as well as some basic notation fixed. Section 3 then formalizes the reasoning concept of common belief in rationality in this more general setting that admits payoff uncertainty. In Section 4, a solution concept for incomplete information games called generalized iterated strict dominance is constructed as a procedure on decision problems using strict dominance arguments. Section 5 gives a characterization of common belief in rationality by generalized iterated strict dominance as well as in terms of best-response sets. Section 6 relates common belief in rationality to interim correlated rationalizability. It turns out that, if the belief hierarchies on utilities are fixed, then the two concepts are behaviourally equivalent. Section 7 identifies epistemic conditions that characterize complete information from a one person-perspective. Finally, Section 8 offers some concluding remarks.

2 Preliminaries

It is standard in game theory to model a static game by specifying the players, their respective choices, as well as their respective utilities for every choice combination. If these ingredients are assumed to be commonly known among the players, the corresponding games are said to exhibit complete information. The more general class of incomplete information games admits uncertainty about the players' utilities. Accordingly, a game with incomplete information can be formally represented by a tuple

$$\Gamma = \big(I, (C_i)_{i \in I}, (U_i)_{i \in I}\big)$$

where I denotes a finite set of players, C_i denotes player i's finite choice set, and U_i denotes the finite set of player i's utility functions.[2] Every utility function $u_i \in U_i$ is of the form $u_i : \times_{j \in I} C_j \to \mathbb{R}$. The decisive difference between a static game with incomplete and complete information lies in the consideration of a set of utility functions instead of a unique utility function for every player.

In order to formally express beliefs and interactive beliefs about choices and utility functions an epistemic structure needs to be added to the game. The

[2] For simplicity sake, attention is restricted to finite games and finite epistemic models.

following epistemic model enables a compact representation of epistemic mental states of players with regards to choices, utility functions, and higher-order beliefs.

Definition 1. *Let* $\Gamma = \big(I, (C_i)_{i \in I}, U_i)_{i \in I}\big)$ *be a game with incomplete information. An* epistemic model *of* Γ *is a tuple* $\mathcal{M}^{\Gamma} = \big((T_i)_{i \in I}, (b_i)_{i \in I}\big)$*, where for every player* $i \in I$

- T_i *is a finite set of types,*
- $b_i : T_i \to \Delta(C_{-i} \times T_{-i} \times U_{-i})$ *assigns to every type* $t_i \in T_i$ *a probability measure* $b_i[t_i]$ *on the set of opponents' choice type utility function combinations.*

Note that for every type an infinite belief hierarchy about the respective opponents' choices and utility functions can be derived. Also, marginal beliefs can be inferred from a type. For instance, every type $t_i \in T_i$ induces a belief on the opponents' choice combinations by marginalizing the probability measure $b_i[t_i]$ on the space C_{-i}. For simplicity sake, no additional notation is introduced for marginal beliefs. In the sequel, it should always be clear from the context which belief $b_i[t_i]$ refers to. Similarly, marginal belief hierarchies can be derived from a type. For instance, a type's marginal belief hierarchy on choices specifies a belief about the opponents' choice combinations, a belief about the opponents' beliefs about their respective opponents' choice combinations, etc, where all beliefs are obtained by marginalization of the type's full belief hierarchy. For every type $t_i \in T_i$ the marginal belief hierarchy on choices is denoted by t_i^C and the marginal belief hierarchy on utilities is denoted by t_i^U.

Here, payoff uncertainty is treated symmetrically to strategic uncertainty. As the latter concerns the respective opponents' choices, the former is also defined with respect to the respective opponents' utility functions only. This treatment is in line with Harsanyi (1967-68), who also assumes that each player knows his own utility function, and more generally, that the uncertainty concerns the opponents of the player from whose point of view the game is analyzed.[3] However, the special case of players being uncertain about their *own* payoffs could be

[3] Cf. Harsanyi (1967-68), p. 163 and p. 170.

accommodated in Definition 1 by extending the space of uncertainty for every player $i \in I$ from $C_{-i} \times T_{-i} \times U_{-i}$ to $C_{-i} \times T_{-i} \times (\times_{j \in I} U_j)$. Alternatively, a reasoner's actual utility function could be defined as the expectation over the set U_i. This modelling choice does not affect the subsequent results.

Note that in our treatment, a type only specifies the epistemic mental state of a player, not his utility function. In this sense we follow Harsanyi's (1967-68) approach, which separates the utility component from the epistemic component.[4]

Moreover, due to the symmetric treatment of uncertainty about choices and payoffs, types are – analogous to complete information epistemic structures – simply compact ways of representing belief hierarchies. In general, a type holds a belief about the basic space of uncertainty and the opponents' types. In the case of complete information games the basic space of uncertainty consists of the players' choice combinations, while in the more general case of incomplete information games the basic space of uncertainty is extended to the players' choice utility function combinations. Alternatively, for every type $t_i \in T_i$ the probability measure $b_i[t_i]$ could be defined exactly as in the case of complete information, i.e. on the space $C_{-i} \times T_{-i}$, and payoff uncertainty be injected into the epistemic model by assigning a utility function to every type. Again, the subsequent results are essentially independent of this modelling choice.

Note that our epistemic model follows Harsanyi's (1967-68) one-person perspective approach. Accordingly, game theory can be conveived of as an interactive extension of decision theory. Indeed, all epistemic concepts – including iterated ones – are understood and defined as mental states inside the mind of a single person. A one-person perspective approach seems natural in the sense that reasoning is formally represented by epistemic concepts and any reasoning process prior to choice does indeed take place entirely *within* the reasoner's mind. Formally, this approach is parsimonious in the sense that states, describing the beliefs of all players, do not have to be introduced.

[4] Cf. Harsanyi (1967-68), pp. 169-171.

Since the epistemic model according to Definition 1 treats the sources of uncertainty – choices and utilities – symmetrically, our approach is more general than Ely and Pęski (2006) as well as Dekel et al. (2007). Indeed, the latter models formalize incomplete information by fixing the belief hierarchies on the utilities before reasoning about choice is considered.

Some further notions and notation are now introduced. For that purpose consider a game Γ, an epistemic model $\mathcal{M}^{\Gamma}$ of it, and fix two players $i, j \in I$ such that $i \neq j$. A type $t_i \in T_i$ of i is said to *deem possible* some choice type utility function combination (c_{-i}, t_{-i}, u_{-i}) of his opponents, if $b_i[t_i]$ assigns positive probability to (c_{-i}, t_{-i}, u_{-i}). Analogously, t_i deems possible some type t_j of his opponent, if $b_i[t_i]$ assigns positive probability to t_j. For each choice-type-utility function combination (c_i, t_i, u_i), the *expected utility* is given by

$$v_i(c_i, t_i, u_i) = \sum_{c_{-i} \in C_{-i}} \big(b_i[t_i](c_{-i}) \cdot u_i(c_i, c_{-i}) \big).$$

Optimality can now be formally defined.

Definition 2. *Let* $\Gamma = \big(I, (C_i)_{i \in I}, U_i)_{i \in I}\big)$ *be a game with incomplete information,* $\mathcal{M}^{\Gamma}$ *some epistemic model of it,* $i \in I$ *some player,* $u_i \in U_i$ *some utility function for player* i*, and* $t_i \in T_i$ *some type of player* i*. A choice* $c_i \in C_i$ *is* optimal *for the type utility function pair* (t_i, u_i)*, if* $v_i(c_i, t_i, u_i) \geq v_i(c'_i, t_i, u_i)$ *for all* $c'_i \in C_i$*.*

In contrast to standard epistemic models for static games, optimality of a choice is not defined relative to a type, but to a type-utility function pair here. This is due to the existence of payoff uncertainty in addition to strategic uncertainty, as optimality of a choice depends on the respective player's utility function as well as on his first-order belief about his opponents' choices induced by his type.

3 Common Belief in Rationality

In the usual way, interactive reasoning can be constructed based on epistemic models. In fact, conditions are inductively imposed on the different layers of a belief hierarchy. Intuitively, a player believes his opponents to be rational, if – for each of his opponents – he only assigns positive probability to choice type utility function combinations such that the choice is optimal for the respective type utility function pair. Formally, belief in rationality can be defined as follows.

Definition 3. *Let* $\Gamma = \big(I, (C_i)_{i \in I}, (U_i)_{i \in I}\big)$ *be a game with incomplete information,* $\mathcal{M}^{\Gamma}$ *some epistemic model of it, and* $i \in I$ *some player. A type* $t_i \in T_i$ believes in the opponents' rationality*, if* t_i *only deems possible choice type utility function combinations* (c_{-i}, t_{-i}, u_{-i}) *such that* c_j *is optimal for* (t_j, u_j) *for every opponent* $j \in I \setminus \{i\}$*.*

As in the special case of complete information, belief in the opponents' rationality puts a restriction on a type's induced beliefs. However, with incomplete information the opponents' utility functions are part of the uncertainty space of the induced belief of a player's type.

Interactive reasoning about rationality can then be defined by iterating belief in rationality.

Definition 4. *Let* $\Gamma = \big(I, (C_i)_{i \in I}, (U_i)_{i \in I}\big)$ *be a game with incomplete information,* $\mathcal{M}^{\Gamma}$ *some epistemic model of it, and* $i \in I$ *some player.*

- *A type* $t_i \in T_i$ *expresses* 1-fold belief in rationality*, if* t_i *believes in the opponents' rationality.*
- *A type* $t_i \in T_i$ *expresses* k-fold belief in rationality *for some* $k > 1$*, if* t_i *only assigns positive probability to types* $t_j \in T_j$ *for all* $j \in I \setminus \{i\}$ *such that* t_j *expresses* $k-1$*-fold belief in rationality.*
- *A type* $t_i \in T_i$ *expresses* common belief in rationality*, if* t_i *expresses* k*-fold belief in rationality for all* $k \geq 1$*.*

Intuitively, if a player expresses common belief in rationality, then there exists no layer in his belief hierarchy in which the rationality of any player is questioned. Note that the only difference to the complete information case is the generalization of belief in the opponents' rationality. Yet, the way that interactive beliefs are constructed is identical with and without payoff uncertainty. Besides, belief in the opponents' rationality and its iterations purely concern an agent's reasoning and are thus properties of the agent's epistemic set-up – formally, his type – only. Thus, Definition 4 provides a one-person perspective formalization of common belief in rationality.

Finally, the decision rule of rational choice under common belief in rationality can be defined with incomplete information as well.

Definition 5. *Let* $\Gamma = \big(I, (C_i)_{i \in I}, (U_i)_{i \in I}\big)$ *be a game with incomplete information,* $i \in I$ *some player, and* $u_i \in U_i$ *some utility function of player* i*. A choice* $c_i \in C_i$ *of player* i *is* rational for utility function u_i under common belief in rationality*, if there exists an epistemic model* $\mathcal{M}^{\Gamma}$ *of* Γ *with a type* $t_i \in T_i$ *of player* i *such that* c_i *is optimal for* (t_i, u_i) *and* t_i *expresses common belief in rationality.*

Note that in our incomplete information framework rational choice under common belief in rationality does not fix a particular epistemic model. Thus no exogenous restrictions are put on the belief hierarchies and no belief hierarchies are excluded a priori. Yet only the existence of some epistemic model is needed to construct a belief hierarchy expressing common belief in rationality that supports the given choice. This frugality is enabled by construction of the formal framework using the one-person perspective approach.

An illustration of the concept of common belief in rationality is provided by the following example.

Example 1. Consider a two player game with incomplete information between *Alice* and *Bob*, where the choices sets are $C_{Alice} = \{a, b, c\}$ as well as $C_{Bob} = \{d, e, f\}$, respectively, and the sets of utility functions are $U_{Alice} = \{u_A, u'_A\}$ as well as $U_{Bob} = \{u_B, u'_B\}$, respectively. In Figure 1, the utility functions are spelled out in detail.

An interactive – more classical – representation of the game is provided in Figure 2.

u_A

	d	e	f
a	3	2	1
b	2	1	3
c	0	0	0

u'_A

	d	e	f
a	1	3	1
b	2	1	1
c	0	0	0

u_B

	a	b	c
d	3	2	1
e	2	1	3
f	0	0	0

u'_B

	a	b	c
d	1	3	1
e	2	1	1
f	0	0	0

Fig. 1. Utility functions of *Alice* and *Bob*.

Alice \ Bob	d	e	f
a	3, 3	2, 2	1, 0
b	2, 2	1, 1	3, 0
c	0, 1	0, 3	0, 0

Alice \ Bob	d	e	f
a	3, 1	2, 2	1, 0
b	2, 3	1, 1	3, 0
c	0, 1	0, 1	0, 0

Alice \ Bob	d	e	f
a	1, 3	3, 2	1, 0
b	2, 2	1, 1	1, 0
c	0, 1	0, 3	0, 0

Alice \ Bob	d	e	f
a	1, 1	3, 2	1, 0
b	2, 3	1, 1	1, 0
c	0, 1	0, 1	0, 0

Fig. 2. Interactive representation of the two-player game with incomplete information and utility functions as specified in Figure 1.

Suppose the epistemic model $\mathcal{M}^\Gamma$ of Γ given by the sets of types $T_{Alice} = \{t_A, t'_A\}$, $T_{Bob} = \{t_B, t'_B\}$, and the following induced belief functions

- $b_{Alice}[t_A] = (e, t_B, u_B)$,
- $b_{Alice}[t'_A] = (d, t'_B, u'_B)$,
- $b_{Bob}[t_B] = (a, t_A, u_A)$,
- $b_{Bob}[t'_B] = \frac{1}{2}(a, t'_A, u_A) + \frac{1}{2}(b, t'_A, u'_A)$.

Accordingly, type t_A assigns probability 1 to the choice type utility function combination (e, t_B, u_B). Analogeously, the induced beliefs of types t'_A and t_B are obtained. *Bob*'s type t'_B assigns probability $\frac{1}{2}$ to the choice type utility function combination (a, t'_A, u_A) and probability $\frac{1}{2}$ to the choice type utility function combination (b, t'_A, u'_A). Note that *Alice*'s type t_A does not believe in *Bob*'s rationality, as e is not optimal for the type utility function pair (t_B, u_B) she believes him to be characterized by. In particular, it follows that t_A does not express common belief in rationality. However, *Alice*'s type t'_A expresses common belief in rationality. Indeed, t'_A believes in *Bob*'s rationality, as d is optimal for *Bob*'s type utility function pair (t'_B, u'_B). Also, t'_B believes in *Alice*'s rationality, since a is optimal for *Alice*'s type utility function pair (t'_A, u_A) and b is optimal for *Alice*'s type utility function pair (t'_A, u'_A). As t'_A only deems possible *Bob*'s type t'_B, and t'_B only deems possible *Alice*'s type t'_A, it follows inductively that t'_A expresses common belief in rationality. Hence, a is rational for u_A under common belief in rationality, b is rational for u'_A under common belief in rationality, and d is rational for u'_B under common belief in rationality. ♣

A special case that could be of relevance in some applications ensues if the reasoner's beliefs about his opponents' types and about his opponents' utilities are assumed to be independent. Intuitively, a person is made up of two components: doxastic mental states and preferences. Given such a modular notion of a person, it can be of interest to consider beliefs that treat the two components as independent. This issue of independence is also discussed by Dekel et al. (2007). In fact, the following example shows that such an independence condition can refine the set of optimal choices under common belief in rationality.

Example 2. Consider a two player game with incomplete information between *Alice* and *Bob*, where the choices sets are $C_{Alice} = \{a, b, c\}$ as well as $C_{Bob} = \{d, e, f\}$, respectively, and the utility functions are $U_{Alice} = \{u_A\}$ as well as $U_{Bob} = \{u_B, u'_B\}$, respectively. In Figure 1, the utility functions are spelled out in detail.

u_A	d	e	f
a	2	0	2
b	0	2	2
c	1	1	0

u_B	a	b	c
d	1	0	0
e	0	0	0
f	1	1	1

u'_B	a	b	c
d	0	0	0
e	0	1	0
f	1	1	1

Fig. 3. Utility functions of *Alice* and *Bob*.

An interactive representation of the game is provided in Figure 4.

Consider the epistemic model $\mathcal{M}^{\Gamma}$ of Γ given by the sets of types $T_{Alice} = \{t_A, t'_A, t''_A\}$, $T_{Bob} = \{t_B, t'_B\}$, and the following induced belief functions

Alice \ *Bob*	d	e	f
a	$2,1$	$0,0$	$2,1$
b	$0,0$	$2,0$	$2,1$
c	$1,0$	$1,0$	$0,1$

Alice \ *Bob*	d	e	f
a	$2,0$	$0,0$	$2,1$
b	$0,0$	$2,1$	$2,1$
c	$1,0$	$1,0$	$0,1$

Fig. 4. Interactive representation of the two-player game with incomplete information and utility functions as specified in Figure 3.

– $b_{Alice}[t_A] = \frac{1}{2}(d, t_B, u_B) + \frac{1}{2}(e, t'_B, u'_B)$,

– $b_{Alice}[t'_A] = (d, t_B, u_B)$,

– $b_{Alice}[t''_A] = (e, t'_B, u'_B)$,

– $b_{Bob}[t_B] = (a, t'_A, u_A)$,

– $b_{Bob}[t'_B] = (b, t''_A, u_A)$.

Observe that all types in this epistemic model believe in the opponents' rationality. In particular, type t_A thus expresses common belief in rationality. As choice c is optimal for type t_A, *Alice* can rationally choose c under common belief in rationality given her utility function u_A. However, the belief $\frac{1}{2}d + \frac{1}{2}e$ is the unique first-order belief on choices supporting choice c. Since d is only optimal for *Bob* if his utility function is u_B and he assigns probability 1 to *Alice*'s choice a, and e is only optimal for him if his utility function is u'_B and he assigns probability 1 to *Alice*'s choice b, it follows that c can only be optimal for *Alice* under common belief in rationality, if she assigns probability $\frac{1}{2}$ to *Bob* being equipped with utility function u_B and to *Bob* assigning probability 1 to her choosing a as well as probability $\frac{1}{2}$ to *Bob* being equipped with utility function u'_B and to *Bob* assigning probability 1 to her choosing b. Since this belief violates the independence of beliefs on types and utilities, c can be concluded to be ruled out under common belief in rationality with the independence assumption. ♣

4 Generalized Iterated Strict Dominance

An algorithm is now introduced as a solution concept, which extends iterated strict dominance to games with incomplete information with a one-person perspective. The algorithm is built on the notion of a decision problem. Given a game $\Gamma = \big(I, (C_i)_{i\in I}, (U_i)_{i\in I}\big)$, a player $i \in I$, and a utility function $u_i \in U_i$, a decision problem

$$\Gamma_i(u_i) = (D_i, D_{-i}, u_i)$$

for player i consists of choices $D_i \subseteq C_i$ for i, choice combinations $D_{-i} \subseteq C_{-i}$ for i's opponents, as well as the utility function u_i restricted to $D_i \times D_{-i}$. A decision problem describes a game-theoretic choice problem from a one-person perspective, namely the perspective of the reasoner. In a decision problem, choice rules such as strict dominance can be formally defined. Indeed, given a utility function $u_i \in U_i$ for player i and his corresponding decision problem $\Gamma_i(u_i) = (D_i, D_{-i}, u_i)$, a choice $c_i \in D_i$ is called strictly dominated, if there exists a probability measure $r_i \in \Delta(D_i)$ such that $u_i(c_i, c_{-i}) < \sum_{c_i' \in D_i} r_i(c_i') \cdot u_i(c_i', c_{-i})$ for all $c_{-i} \in D_{-i}$.

With the notions of decision problem and strict dominance on decision problems the standard solution concept iterated strict dominance for complete information games can be extended to payoff uncertainty with a one-person perspective. Indeed, the algorithm *generalized iterated strict dominance* is defined as follows.

Definition 6. *Let $\Gamma = \big(I, (C_i)_{i\in I}, (U_i)_{i\in I}\big)$ be a game with incomplete information.*

Round 1. *For every player $i \in I$ and for every utility function $u_i \in U_i$ consider the initial decision problem $\Gamma_i^0(u_i) := \big(C_i^0(u_i), C_{-i}^0(u_i), u_i\big)$, where $C_i^0(u_i) := C_i$ and $C_{-i}^0(u_i) := C_{-i}$.*

- Step 1.1 *Set $C_{-i}^1(u_i) := C_{-i}^0(u_i)$.*
- Step 1.2 *Form $\Gamma_i^1(u_i) := \big(C_i^1(u_i), C_{-i}^1(u_i), u_i\big)$, where $C_i^1(u_i) \subseteq C_i^0(u_i)$ only contains choices $c_i \in C_i$ for player i that are not strictly dominated in the decision problem $\big(C_i^0(u_i), C_{-i}^1(u_i), u_i\big)$.*

Round $k > 1$. *For every player $i \in I$ and for every utility function $u_i \in U_i$ consider the reduced decision problem $\Gamma_i^{k-1}(u_i) := \big(C_i^{k-1}(u_i), C_{-i}^{k-1}(u_i), u_i\big)$.*

Step $k.1$ *Form $C_{-i}^{k}(u_i) \subseteq C_{-i}^{k-1}(u_i)$ by eliminating from $C_{-i}^{k-1}(u_i)$ every opponents' choice combination $c_{-i} \in C_{-i}^{k-1}(u_i)$ that contains for some opponent $j \in I \setminus \{i\}$ a choice $c_j \in C_j$ for which there exists no utility function $u_j \in U_j$ such that $c_j \in C_j^{k-1}(u_j)$.*

Step $k.2$ *Form $\Gamma_i^{k}(u_i) := \big(C_i^{k}(u_i), C_{-i}^{k}(u_i), u_i\big)$, where $C_i^{k}(u_i) \subseteq C_i^{k-1}(u_i)$ only contains choices $c_i \in C_i^{k-1}(u_i)$ for player i that are not strictly dominated in the decision problem $\big(C_i^{k-1}(u_i), C_{-i}^{k}(u_i), u_i\big)$.*

The set $GISD := \times_{i \in I} GISD_i \subseteq \times_{i \in I}(C_i \times U_i)$ is the output of generalized iterated strict dominance*, where for every player $i \in I$ the set $GISD_i \subseteq C_i \times U_i$ only contains choice utility function pairs $(c_i, u_i) \in C_i \times U_i$ such that $c_i \in C_i^{k}(u_i)$ for all $k \geq 0$.*

The algorithm iteratively eliminates strictly dominated choices from decision problems for all players. In every round a decision problem for a player is formed by first eliminating all opponents' choices that are strictly dominated in every decision problem for that opponent in the previous round, and subsequently eliminating the player's choices that are strictly dominated. In fact, for every player the algorithm yields a set of choice utility function pairs as output. Due to the presence of incomplete information the algorithm thus identifies choices relative to payoffs. With generalized iterated strict dominance a non-equilibrium and one-person perspective solution concept is thus added to the theory of games with incomplete information.

Note that generalized iterated strict dominance can be viewed as a direct generalization of iterated strict dominance to incomplete information with a one person-perspective approach. It also closely corresponds to the iterative strict dominance procedures for static games considered by Battigalli (2003, Proposition 3.8) and by Battigalli and Siniscalchi (1999, p. 215), as well as to the interim iterated dominance procedure by Battigalli et al. (2011, p. 14). Some differences to Δ-rationalizability concepts are the use of decision problems in our algorithm as well as of strict dominance – instead of best response – arguments.

The following remark draws attention to some useful properties of the generalized iterated strict dominance algorithm, that directly follow from its definition.

Remark 1. Let $\Gamma = \big(I, (C_i)_{i\in I}, (U_i)_{i\in I}\big)$ be a game with incomplete information. The algorithm generalized iterated strict dominance is non-empty, i.e. $GISD \neq \emptyset$, finite, i.e. there exists $n \in \mathbb{N}$ such that $\Gamma_i^k(u_i) = \Gamma_i^n(u_i)$ for all $k \geq n$, for all utility functions $u_i \in U_i$, and for all players $i \in I$, as well as order-independent, i.e. the final output of generalized iterated strict dominance does not depend on the specific order of elimination.

The non-emptiness of the algorithm follows from the fact that at no round it is possible to delete all choices for a given player by definition of strict dominance. As there are only finitely many choices for every player, the algorithm stops after finitely many rounds. As a choice remains strictly dominated if a decision problem is reduced, the order of elimination does not affect the eventual output of the algorithm.

Finally, generalized iterated strict dominance is illustrated by applying the algorithm to the two player game introduced in Example 1.

Example 3. Consider again the two player game with incomplete information from Example 1. In order to apply $GISD$ to this game decision problems for the two players for each of their respective utility functions need to be formed as in Figure 5, where the choices of the respective decision making player are represented as rows and the opponent's choices as columns.

In both $\Gamma_A^0(u_A)$ and $\Gamma_A^0(u'_A)$ the choice c is strictly dominated by b. For Bob the choice f is strictly dominated by e in his decision problems $\Gamma_B^0(u_B)$ and $\Gamma_B^0(u'_B)$. There are no further choices that can be ruled out for Alice or Bob with strict dominance given either of their utility functions. The 1-fold reduced decision problems Γ_A^1 and Γ_B^1 result as in Figure 6.

$\Gamma_A^0(u_A)$

	d	e	f
a	3	2	1
b	2	1	3
c	0	0	0

$\Gamma_A^0(u'_A)$

	d	e	f
a	1	3	1
b	2	1	1
c	0	0	0

$\Gamma_B^0(u_B)$

	a	b	c
d	3	2	1
e	2	1	3
f	0	0	0

$\Gamma_B^0(u'_B)$

	a	b	c
d	1	3	1
e	2	1	1
f	0	0	0

Fig. 5. Initial decision problems for *Alice* and *Bob*.

$\Gamma_A^1(u_A)$

	d	e	f
a	3	2	1
b	2	1	3

$\Gamma_A^1(u'_A)$

	d	e	f
a	1	3	1
b	2	1	1

$\Gamma_B^1(u_B)$

	a	b	c
d	3	2	1
e	2	1	3

$\Gamma_B^1(u'_B)$

	a	b	c
d	1	3	1
e	2	1	1

Fig. 6. 1-fold reduced decision problems for *Alice* and *Bob*.

In both $\Gamma_A^1(u_A)$ and $\Gamma_A^1(u'_A)$ those choices of Bob are eliminated that are strictly dominated in all initial decision problems Γ_B^0 for Bob, i.e. choice f. Then, the choice b can be deleted for Alice given u_A as it is strictly dominated by a in $(\{a,b\},\{d,e\},u_A)$, but not given u'_A as it is not strictly dominated in $(\{a,b\},\{d,e\},u'_A)$. Moreover, in both $\Gamma_B^1(u_B)$ and $\Gamma_B^1(u'_B)$ those choices of Alice are eliminated that are strictly dominated in all initial decision problems Γ_A^0 for Alice, i.e. choice c. Then, the choice e can be deleted for Bob given u_B as it is strictly dominated by d in $(\{d,e\},\{a,b\},u_B)$, but not given u'_B as it is not strictly dominated in $(\{d,e\},\{a,b\},u'_B)$. The 2-fold reduced decision problems Γ_A^2 and Γ_B^2 result as in Figure 7.

$\Gamma_A^2(u_A)$

	d	e
a	3	2

$\Gamma_A^2(u'_A)$

	d	e
a	1	3
b	2	1

$\Gamma_B^2(u_B)$

	a	b
d	3	2

$\Gamma_B^2(u'_B)$

	a	b
d	1	3
e	2	1

Fig. 7. 2-fold reduced decision problems for *Alice* and *Bob*.

Since there are no strict dominance relations in any of the 2-fold reduced decision problems Γ_A^2 and Γ_B^2, the algorithm stops and returns the set $GISD = GISD_{Alice} \times GISD_{Bob} = \{(a, u_A), (a, u'_A), (b, u'_A)\} \times \{(d, u_B), (d, u'_B), (e, u'_B)\}$ as a solution to this two player game with incomplete information. ♣

5 Characterization

Next it is shown that for the class of incomplete information games, common belief in rationality can be characterized by generalized iterated strict dominance. A fundamental result in game theory – so-called Pearce's Lemma – due to Pearce (1984) connects strict dominance and optimality of choice. Accordingly, a choice in a two-player game with complete information is strictly dominated, if and only if, it is not optimal for any belief about the opponent's choices.[5] Note that a choice $c_i \in C_i$ of some player $i \in I$ is called optimal for a belief $p \in \Delta(C_{-i})$ about the opponents' choices, if $\sum_{c_{-i} \in C_{-i}} p(c_{-i}) \cdot u_i(c_i, c_{-i}) \geq \sum_{c_{-i} \in C_{-i}} p(c_{-i}) \cdot u_i(c'_i, c_{-i})$ for all $c'_i \in C_i$. Similarly, in a game with incomplete information, a choice $c_i \in C_i$ is said to be optimal for a belief utility function pair (p_i, u_i), where $p_i \in \Delta(C_{-i})$ and $u_i \in U_i$, if $\sum_{c_{-i} \in C_{-i}} p(c_{-i}) \cdot u_i(c_i, c_{-i}) \geq \sum_{c_{-i} \in C_{-i}} p(c_{-i}) \cdot u_i(c'_i, c_{-i})$ for all $c'_i \in C_i$.

A slight generalization of Pearce's Lemma to finite incomplete information games is given by the following result.

Lemma 1. *Let $\Gamma = \big(I, (C_i)_{i \in I}, (U_i)_{i \in I}\big)$ be a game with incomplete information, $i \in I$ some player, $u_i \in U_i$ some utility function of player i, and $\Gamma_i(u_i) = \big(D_i, D_{-i}, u_i\big)$ some decision problem of player i. A choice $c_i \in D_i$ is strictly dominated in $\Gamma_i(u_i)$, if and only if, there exists no probability measure $p \in \Delta(D_{-i})$ such that c_i is optimal for (p, u_i) in $\Gamma_i(u_i)$.*

Proof. Consider the two player game $\Gamma' = \big(\{i, j\}, \{D'_i, D'_j\}, \{u'_i, u'_j\}\big)$, where $D'_i = D_i$, $D'_j = \{d_j^{d_{-i}} : d_{-i} \in D_{-i}\}$, $u'_i(d_i, d_j^{d_{-i}}) = u_i(d_i, d_{-i})$ for all $d_i \in D'_i$ and

[5] Besides the original proof in Pearce (1984) a more elementary proof of Pearce's Lemma is provided by Perea (2012).

for all $d_j^{d_{-i}} \in D_j'$, as well as $u_j'(d_i, d_j^{d_{-i}}) = 0$ for all $d_i \in D_i'$ and for all $d_j^{d_{-i}} \in D_j'$. Note that a choice $c_i \in D_i$ is strictly dominated in the decision problem $\Gamma_i(u_i)$, if and only if, it is strictly dominated in the two person game Γ'. By Pearce's Lemma applied to Γ', it then follows that c_i is strictly dominated in $\Gamma_i(u_i)$, if and only if, there exists no probability measure $p \in \Delta(D_{-i})$ such that c_i is optimal for (p, u_i) in $\Gamma_i(u_i)$. ■

Note that optimality in epistemic models according to Definition 2 is defined relative to a type utility function pair, while in the algorithmic setting optimality is defined relative to a pair consisting of a belief about the opponents' choices and a utility function. Of course these two notions of optimality are semantically equivalent, as the relevant belief by the type in an epistemic model is its marginal belief about the opponents' choices.

Equipped with a generalized version of Pearce's Lemma an algorithmic characterization of the epistemic concept of common belief in rationality can be established for games with incomplete information by generalized iterated strict dominance.

Theorem 1. *Let $\Gamma = \big(I, (C_i)_{i \in I}, (U_i)_{i \in I}\big)$ be a game with incomplete information, $i \in I$ some player, $c_i \in C_i$ some choice for player i, and $u_i \in U_i$ some utility function of player i. The choice c_i is rational for u_i under common belief in rationality, if and only if, $(c_i, u_i) \in GISD_i$.*

Proof. For the *only if* direction of the theorem define a set $(C_i \times U_i)^{CBR} := \{(c_i, u_i) \in C_i \times U_i : c_i \text{ is rational for } u_i \text{ under common belief in rationality}\}$ for every player $i \in I$. It is shown, by induction on $k \geq 0$, that for every player $i \in I$ and for every choice utility function pair $(c_i, u_i) \in (C_i \times U_i)^{CBR}$, it is the case that $c_i \in C_i^k(u_i)$. Note that $c_i \in C_i^0(u_i)$ directly holds for all $(c_i, u_i) \in (C_i \times U_i)^{CBR}$ and for all $i \in I$, as $C_i^0(u_i) = C_i$ for all $u_i \in U_i$ and for all $i \in I$. Now consider some $k \geq 0$ and suppose that $c_i \in C_i^k(u_i)$ holds for every player $i \in I$ and for every choice utility function pair $(c_i, u_i) \in (C_i \times U_i)^{CBR}$. Let $i \in I$ be some player, and take some $(c_i, u_i) \in (C_i \times U_i)^{CBR}$. Then, there exists an epistemic model $\mathcal{M}^{\Gamma}$ of Γ with a type $t_i \in T_i$ that expresses common belief in rationality such that c_i is optimal for (t_i, u_i). Take some $(c_j, t_j, u_j) \in C_j \times T_j \times U_j$ such that $b_i[t_i](c_j, t_j, u_j) > 0$. As t_i expresses common belief in rationality, t_j expresses common belief in rationality too, and c_j is optimal for (t_j, u_j). Thus, $(c_j, u_j) \in (C_j \times U_j)^{CBR}$, and, by the inductive assumption, $c_j \in C_j^k(u_j)$. Hence, for every choice $c_j \in \text{supp}(b_i[t_i])$ it is the case that $c_j \in C_j^k(u_j)$ for some utility function $u_j \in U_j$, and thus t_i only assigns positive probability to choices c_j contained in a decision problem $\Gamma_j^k(u_j)$ for some $u_j \in U_j$ for every opponent $j \in I \setminus \{i\}$. Consequently, t_i only assigns positive probability to choice combinations in $C_{-i}^{k+1}(u_i)$. Since c_i is optimal for (t_i, u_i), it follows from Lemma 1 that $c_i \in C_i^{k+1}(u_i)$. Therefore, by induction, $(c_i, u_i) \in GISD_i$ obtains.

For the *if* direction of the theorem, suppose that the algorithm stops after $k \geq 0$ rounds. Then, for every $(c_i, u_i) \in GISD_i$ it is the case that $c_i \in C_i^k(u_i)$. By Lemma 1, c_i is optimal for (p_i, u_i), where $p_i \in \Delta\big(C_{-i}^k(u_i)\big)$. Observe that every $c_{-i} \in C_{-i}^k(u_i)$ only contains, for every player $j \in I \setminus \{i\}$, choices $c_j \in C_j$ such that $(c_j, u_j^{c_j}) \in GISD_j$ for some $u_j^{c_j} \in U_j$. Define a probability measure $p_i^{(c_i,u_i)} \in \Delta(GISD_{-i})$ by

$$p_i^{(c_i,u_i)}(c_{-i}, u_{-i}) = \begin{cases} p_i(c_{-i}), & \text{if } c_{-i} \in C_{-i}^k(u_i) \text{ and } u_{-i} = u_{-i}^{c_{-i}} \\ 0, & \text{otherwise} \end{cases}$$

for all $(c_{-i}, u_{-i}) \in C_{-i} \times U_{-i}$. Construct an epistemic model $\mathcal{M}^{\Gamma} = \{(T_i)_{i \in I}, (b_i)_{i \in I}\}$ of Γ, where

$$T_i := \{t_i^{(c_i,u_i)} : (c_i, u_i) \in GISD_i\}$$

for all $i \in I$, and

$$b_i[t_i^{(c_i,u_i)}](c_{-i}, t_{-i}, u_{-i}) =$$

$$\begin{cases} p_i^{(c_i,u_i)}(c_{-i}, u_{-i}), & \text{if } (c_{-i}, u_{-i}) \in GISD_{-i} \text{ and } t_j = t_j^{(c_j,u_j)} \text{ for all } j \in I \setminus \{i\} \\ 0, & \text{otherwise} \end{cases}$$

for all $(c_{-i}, t_{-i}, u_{-i}) \in C_{-i} \times T_{-i} \times U_{-i}$, for all $t_i^{(c_i,u_i)} \in T_i$ and for all $i \in I$. Observe that, by construction, for every player $i \in I$ and for every $(c_i, u_i) \in$

$GISD_i$, the choice c_i is optimal for $(t_i^{(c_i,u_i)}, u_i)$. Hence, every type $t_i^{(c_i,u_i)}$ believes in the opponents' rationality. It then directly follows inductively that every such type $t_i^{(c_i,u_i)}$ also expresses common belief in rationality. Therefore, for every choice utility function pair $(c_i, u_i) \in GISD_i$, there exists a type $t_i^{(c_i,u_i)}$ within $\mathcal{M}^\Gamma$ such that $t_i^{(c_i,u_i)}$ expresses common belief in rationality and c_i is optimal for $\left(t_i^{(c_i,u_i)}, u_i\right)$. Hence, c_i is rational for u_i under common belief in rationality. ■

Similar algorithmic characterizations of common belief in rationality in incomplete information games can be found in Battigalli and Siniscalchi (1999, Proposition 4), Battigalli (2003, Proposition 3.8) and Battigalli et al. (2011, Section 3.1).

Besides the algorithmic characterization of common belief in rationality, the resulting choice utility function pairs can also be characterized by means of best-response sets. For the case of complete information the notion of best-response set is analyzed by Pearce (1984), and can be formulated in the context of incomplete information as follows.

Definition 7. *Let $\Gamma = \left(I, (C_i)_{i\in I}, (U_i)_{i\in I}\right)$ be a game with incomplete information, and $D_i \subseteq C_i \times U_i$ a set of choice utility function pairs for every player $i \in I$. A tuple $(D_i)_{i\in I}$ is called* best-response-set-tuple, if there exists, for every player $i \in I$ and for every choice utility function pair $(c_i, u_i) \in D_i$, a probability measure $\mu_i \in \Delta(D_{-i})$ such that c_i is optimal for (μ_i, u_i).

In fact, the best-response property enables a characterization of the choice utility function pairs selected by common belief in rationality.

Theorem 2. *Let $\Gamma = \left(I, (C_i)_{i\in I}, (U_i)_{i\in I}\right)$ be a game with incomplete information, $i \in I$ some player, $c_i \in C_i$ some choice of player i, and $u_i \in U_i$ some utility function of player i. There exists a best-response-set-tuple $(D_i)_{i\in I}$ such that $(c_i, u_i) \in D_i$, if and only if, c_i is rational for u_i under common belief in rationality.*

Proof. For the *only if* direction of the theorem it is shown, by induction on $k \geq 0$, that $c_i \in C_i^k(u_i)$ for all $(c_i, u_i) \in D_i$, for all $k \geq 0$, and for all $i \in I$. Let $i \in I$ be some player and $(c_i, u_i) \in D_i$. It then holds that $c_i \in C_i^0(u_i) = C_i$. Now, consider some $(c_i, u_i) \in D_i$ and assume that $k \geq 0$ is such that $c_j \in C_j^k(u_j)$ for every $j \in I$ and for every $(c_j, u_j) \in D_j$. Fix some $(c_i, u_i) \in D_i$, and note that c_i is optimal for (μ_i, u_i), where $\mu_i \in \Delta(D_{-i})$ is some probability measure. By the inductive assumption, $c_j \in C_j^k(u_j)$ for every $(c_j, u_j) \in D_j$ and for every $j \in I \setminus \{i\}$. Hence, μ_i only assigns positive probability to opponents' choices $c_j \in C_j$ which are contained in $C_j^k(u_j)$ for some $u_j \in U_j$. Therefore, μ_i only assigns positive probability to opponents' choice combinations $c_{-i} \in C_{-i}^{k+1}(u_i)$. It follows, by Lemma 1, that c_i is not strictly dominated in the decision problem $\big(C_i^k(u_i), C_{-i}^{k+1}(u_i), u_i\big)$. Thus, $c_i \in C_i^{k+1}(u_i)$, and, by induction on $k \geq 0$, it holds that $(c_i, u_i) \in GISD_i$. Hence, by Theorem 1, c_i is rational for u_i under common belief in rationality.

For the *if* direction of the theorem, it is shown that $(GISD_i)_{i \in I}$ is a best-response-set-tuple. For every $u_j \in U_j$, let $C_j^*(u_j) := \{c_j \in C_j : (c_j, u_j) \in GISD_j\}$ and $C_j^* := \{c_j \in C_j : (c_j, u_j) \in GISD_j \text{ for some } u_j \in U_j\}$. Fix $(c_i, u_i) \in GISD_i$. Consequently, c_i is not strictly dominated in the decision problem $\big(C_i^*(u_i), C_{-i}^*, u_i\big)$. By Lemma 1, c_i is optimal for (p_i, u_i) for some $p_i \in \Delta(C_{-i}^*)$. Hence, c_i is optimal for (μ_i, u_i) for some $\mu_i \in \Delta(GISD_{-i})$. Therefore $(GISD_i)_{i \in I}$ is a best-response-set-tuple. Now, take some $(c_i, u_i) \in C_i \times U_i$ such that c_i is rational for u_i under common belief in rationality. Then, by Theorem 1, it is the case that $(c_i, u_i) \in GISD_i$. ■

Besides, it is actually the case that the algorithm generalized iterated strict dominance always yields the largest best-response-set-tuple as output.

Corollary 1. *Let* $\Gamma = \big(I, (C_i)_{i \in I}, (U_i)_{i \in I}\big)$ *be a game with incomplete information. The set* $GISD \subseteq \times_{i \in I}(C_i \times U_i)$ *is the largest best-response-set-tuple.*

Proof. Let $i \in I$ be some player. By the proof of the *if*-direction of Theorem 2, $(GISD_j)_{j \in I}$ is a best-response-set-tuple. Consider some element $(c_i, u_i) \in D_i$ of a best-response-set-tuple $(D_j)_{j \in I}$ for player i. By Theorems 1 and 2, it follows that $(c_i, u_i) \in GISD_i$. Hence, $GISD_i$ is the largest best-response-set-tuple for player i. ■

6 Interim Rationalizability

Rather recently, interim rationalizability has been proposed in the literature by Ely and Pęski (2006) as well as by Dekel et al. (2007) as a non-equilibrium solution concept for static games with incomplete information. Intuitively, the belief hierarchies on utilities are first fixed and then non-optimal choices are iteratively deleted. In contrast, common belief in rationality does not put any restrictions on the belief hierarchies on utilities. In the specific case of fixed belief hierarchies on utilities, it turns out that the optimal choices under common belief in rationality and Dekel et al.'s (2007) interim correlated rationalizability coincide.

In order to relate common belief in rationality and the associated algorithm generalized iterated strict dominance to interim correlated rationalizability the latter needs to be formally defined. First of all, the necessary framework is introduced.

Definition 8. *Let* $\Gamma = \big(I, (C_i)_{i \in I}, (U_i)_{i \in I}\big)$ *be a game with incomplete information. A* Dekel-Fudenberg-Morris model *of* Γ *is a tuple* $\mathcal{R}^\Gamma = \big((R_i)_{i \in I}, (\tau_i)_{i \in I}\big)$, *where for every player* $i \in I$

- R_i *is a finite set of Dekel-Fudenberg-Morris types,*
- $\tau_i : R_i \to \Delta(R_{-i} \times U_{-i})$ *assigns to every Dekel-Fudenberg-Morris type* $r_i \in R_i$ *a probability measure on the set of opponents' Dekel-Fudenberg-Morris type utility function combinations.*

Note that a Dekel-Fudenberg-Morris model significantly differs from standard epistemic models, as strategic uncertainty is not formally represented via belief hierarchies in the former. Originally, Dekel et al. (2007) also admit own payoff uncertainty, i.e. the induced belief function assigns to every Dekel-Fudenberg-Morris type a probability measure on combinations of opponents' Dekel-Fudenberg-Morris types and all players' utility functions. In order to enable comparability with our model, Definition 8 only considers uncertainty about the opponents' utility functions.[6]

Within the framework of a Dekel-Fudenberg-Morris model the non-equilibrium solution concept of interim correlated rationalizability can be defined next.

Definition 9. *Let* $\Gamma = \big(I, (C_i)_{i \in I}, (U_i)_{i \in I}\big)$ *be a game with incomplete information,* $\mathcal{R}^\Gamma$ *some Dekel-Fudenberg-Morris model of it,* $i \in I$ *some player,* $r_i \in R_i$ *some Dekel-Fudenberg-Morris type of player* i*, and* $u_i \in U_i$ *some utility function of player* i*. The set of player* i*'s* interim correlated rationalizable choices $ICR_i(r_i, u_i)$ *given the Dekel-Fudenberg-Morris type* r_i *and the utility function* u_i *is inductively defined as follows.*

– $ICR_i^0(r_i, u_i) := C_i$,

–

$$ICR_i^k(r_i, u_i) := \{c_i \in C_i : \textit{ there exists } \nu_i \in \Delta(C_{-i} \times R_{-i} \times U_{-i})$$

$$\textit{such that (1), (2), and (3) are satisfied.}\},$$

where

(1) $marg_{R_{-i} \times U_{-i}} \nu_i = \tau_i[r_i]$,

(2) c_i *is optimal for* $\big(marg_{C_{-i}} \nu_i, u_i\big)$,

(3) $\nu_i(c_{-i}, r_{-i}, u_{-i}) > 0$ *implies* $c_j \in ICR_j^{k-1}(r_j, u_j)$ *for all* $j \in I \setminus \{i\}$,

for every $k > 0$,

– $ICR_i(r_i, u_i) := \bigcap_{k \geq 0} ICR_i^k(r_i, u_i)$.

[6] Alternatively, our model could be adapted to admit own payoff uncertainty. A type's expected utility function could then be defined as a convex combination of the respective player's payoffs from the underlying game weighted with the type's marginal beliefs on his own payoffs. However, we intend to model epistemic structures for incomplete information games as close as possible to Harsanyi's original (1967-68) model, and therefore do not admit own payoff uncertainty.

In fact, similarly to Battigalli et al. (2011, Theorem 1), it is shown that interim correlated rationalizability can be epistemically characterized by common belief in rationality for a fixed marginal belief hierarchy on utilities.

Theorem 3. *Let* $\Gamma = \big(I, (C_i)_{i\in I}, (U_i)_{i\in I}\big)$ *be a game with incomplete information,* $\mathcal{R}^\Gamma$ *some Dekel-Fudenberg-Morris model of* Γ*,* $i \in I$ *some player,* $c_i \in C_i$ *some choice of player* i*,* $r_i \in R_i$ *some Dekel-Fudenberg-Morris type of player* i *with marginal belief hierarchy* r_i^U *on utilities, and* $u_i \in U_i$ *some utility function of player* i*. It is the case that* $c_i \in ICR_i(r_i, u_i)$*, if and only if, there exists an epistemic model* $\mathcal{M}^\Gamma$ *of* Γ *with some type* $t_i \in T_i$ *of player* i *and belief hierarchy* t_i^U *such that* t_i *expresses common belief in rationality,* c_i *is optimal for* (t_i, u_i)*, and* $t_i^U = r_i^U$*.*

Proof. For the *only if* direction of the theorem, consider $c_i \in ICR_i(r_i, u_i)$. Then, there exists a probability measure $\nu_i^{c_i,r_i,u_i} \in \Delta(C_{-i} \times R_{-i} \times U_{-i})$ such that $\mathrm{marg}_{R_{-i}\times U_{-i}} \nu_i^{c_i,r_i,u_i} = \tau_i[r_i]$, c_i is optimal for $\big(\mathrm{marg}_{C_{-i}} \nu_i^{c_i,r_i,u_i}, u_i\big)$, and $\nu_i^{c_i,r_i,u_i}(c_{-i}, r_{-i}, u_{-i}) > 0$ implies that $c_j \in ICR_j(r_j, u_j)$ for all $j \in I \setminus \{i\}$. Define the epistemic model $\mathcal{M}^\Gamma = \big((T_i)_{i\in I}, (b_i)_{i\in I}\big)$ with

$$T_i := \{t_i^{c_i,r_i,u_i} : r_i \in R_i, u_i \in U_i, c_i \in ICR_i(r_i, u_i)\}$$

and

$$b_i[t_i^{c_i,r_i,u_i}](c_{-i}, t_{-i}^{c_{-i},r_{-i},u_{-i}}, u_{-i}) := \nu_i^{c_i,r_i,u_i}(c_{-i}, r_{-i}, u_{-i})$$

for all $t_{-i}^{c_{-i},r_{-i},u_{-i}} \in T_{-i}$ and for all $t_i^{c_i,r_i,u_i} \in T_i$. Note that any other choice Dekel-Fudenberg-Morris type utility function tuple receives zero probability. As c_i is optimal for $(\mathrm{marg}_{C_{-i}} \nu_i^{c_i,r_i,u_i}, u_i)$, it follows directly by construction of $\mathcal{M}^\Gamma$ that c_i is optimal for $(t_i^{c_i,r_i,u_i}, u_i)$.

It is now shown that every $t_i^{c_i,r_i,u_i} \in T_i$ believes in the opponents' rationality. Let $t_i^{c_i,r_i,u_i} \in T_i$ and $(c_j, t_j^{c_j,r_j,u_j}, u_j) \in C_j \times T_j \times U_j$ for some player $j \in I \setminus \{i\}$ such that $b_i[t_i^{c_i,r_i,u_i}](c_j, t_j^{c_j,r_j,u_j}, u_j) > 0$. From the preceeding paragraph, it follows that c_j is optimal for $(t_j^{c_j,r_j,u_j}, u_j)$. Hence, $t_i^{c_i,r_i,u_i}$ believes in the opponents' rationality. Since all types in the epistemic model $\mathcal{M}^\Gamma$ believe in the respective opponents' rationality, every type $t_i^{c_i,r_i,u_i} \in T_i$ expresses common belief in rationality.

In order to show that $t_i^U = r_i^U$ for every type $t_i \in T_i$ and for every player $i \in I$, we construct a type morphism $(\psi_i)_{i\in I}$, where for every $i \in I$ the function $\psi_i : T_i \to R_i$ satisfies

$$\tau_i[\psi_i(t_i)](r_{-i}, u_{-i}) = b_i[t_i]\big(\psi_{-i}^{-1}(r_{-i}) \times \{u_{-i}\}\big)$$

for all $(r_{-i}, u_{-i}) \in R_{-i} \times U_{-i}$ and for all $t_i \in T_i$. Towards this end define $\psi_i(t_i^{c_i,r_i,u_i}) := r_i$ for all $t_i^{c_i,r_i,u_i} \in T_i$ and for all $i \in I$. Observe that

$$b_i[t_i^{c_i,r_i,u_i}]\big(\psi_{-i}^{-1}(r_{-i})\times\{u_{-i}\}\big) = b_i[t_i^{c_i,r_i,u_i}]\big(\times_{j\in I\setminus\{i\}}\{t_j^{c_j,r_j,u_j} : c_j \in ICR_j(r_j,u_j)\}\big)$$

$$= \nu_i^{c_i,r_i,u_i}(C_{-i} \times \{r_{-i}\} \times \{u_{-i}\}) = \tau_i[r_i](r_{-i}, u_{-i}) = \tau_i[\psi_i(t_i^{c_i,r_i,u_i})](r_{-i}, u_{-i}).$$

By Heifetz and Samet (1998), Proposition 5.1, it follows that t_i and $\psi_i(t_i)$ induce the same belief hierarchies on utilities for every type $t_i \in T_i$ and for every player $i \in I$, and thus $(t_i^{c_i,r_i,u_i})^U = \big(\psi_i(t_i^{c_i,r_i,u_i})\big)^U = r_i^U$ holds.

Now, take some player $i \in I$ and some choice $c_i \in ICR_i(r_i, u_i)$. Then, it has been shown that c_i is optimal for $(t_i^{c_i,r_i,u_i}, u_i)$, as well as that $t_i^{c_i,r_i,u_i}$ expresses common belief in rationality, and $(t_i^{c_i,r_i,u_i})^U = r_i^U$.

The *if* direction of the theorem is addressed next. For every player $j \in I$, for every Dekel-Fudenberg-Morris type $r_j \in R_j$, for every utility function $u_j \in U_j$, and for every $k \geq 0$ define the set

$$C_j^k(r_j, u_j) := \{c_j \in C_j : c_j \text{ is optimal for } (t_j, u_j)$$

$$\text{for some } t_j \in T_j \text{ that expresses up to } k\text{-fold belief in rationality and } t_j^U = r_j^U\}.$$

It is now shown by induction that $C_j^k(r_j, u_j) \subseteq ICR_j^k(r_j, u_j)$ holds for all $k \geq 0$, for every Dekel-Fudenberg-Morris type $r_j \in R_j$, for every utility function $u_j \in U_j$, and for all $j \in I$. Consider some player $i \in I$. Note that $C_i^0(r_i, u_i) \subseteq ICR_i^0(r_i, u_i)$ obtains directly, as $ICR_i^0(r_i, u_i) = C_i$. Let $k > 0$ and suppose that $C_j^{k-1}(r_j, u_j) \subseteq ICR_j^{k-1}(r_j, u_j)$ for every every Dekel-Fudenberg-Morris type $r_j \in R_j$, for all utility functions $u_j \in U_j$, and for all $j \in I$. Take $r_i^* \in R_i$,

$u_i \in U_i$, and $c_i \in C_i^k(r_i^*, u_i)$. Then, c_i is optimal for (t_i^*, u_i), where t_i^* expresses up to k-fold belief in rationality, and $(t_i^*)^U = (r_i^*)^U$. By Perea (2014), Theorem 4, there exists a set-valued type morphism $\mathcal{F} = (F_i)_{i \in I}$ between $\mathcal{M}^\Gamma$ and $\mathcal{R}^\Gamma$, where $F_j : T_j \twoheadrightarrow R_j$, for all $j \in I$ with $r_i^* \in F_i(t_i^*)$. Hence, for all $t_j \in T_j$ it is the case that

$$F_j(t_j) = \{r_j \in R_j : b_j[t_j]\Big(C_{-j} \times F_{-j}^{-1}\big(F_{-j}(t_{-j})\big) \times \{u_{-j}\}\Big)$$

$$= \tau_j[r_j]\big(F_{-j}(t_{-j}) \times \{u_{-j}\}\big) \text{ for all } t_{-j} \in T_{-j} \text{ and for all } u_{-j} \in U_{-j}\}.$$

Define $\nu_j^{c_j, r_j, u_j} \in \Delta(C_{-j} \times R_{-j} \times U_{-j})$ by $\nu_j^{c_j, r_j, u_j}(c_{-j}, r_{-j}, u_{-j}) := b_j[t_j]\big(\{c_{-j}\} \times F_{-j}^{-1}(r_{-j}) \times \{u_{-j}\}\big)$ whenever $r_j \in F_j(t_j)$. Without loss of generality assume that R_j does not contain two different types inducing the same belief hierarchy on utilities, which ensures that $\mid F_j(t_j) \mid = 1$ for all $t_j \in T_j$. Consequently,

$$\nu_j^{c_j, r_j, u_j}(C_{-j} \times \{r_{-j}\} \times \{u_{-j}\}) = b_j[t_j]\big(C_{-j} \times F_{-j}^{-1}(r_{-j}) \times \{u_{-j}\}\big) = \tau_j[r_j](r_{-j}, u_{-j})$$

whenever $r_j \in F_j(t_j)$. Besides, since c_i is optimal for (t_i^*, u_i), and $b_i[t_i^*]$ has the same marginal belief hierarchy on choices as $\nu_i^{c_i, r_i^*, u_i}$, it follows that c_i is optimal for $(\nu_i^{c_i, r_i^*, u_i}, u_i)$.

Moreover, assume that $\nu_i^{c_i, r_i^*, u_i}(c_{-i}, r_{-i}, u_{-i}) > 0$ and let $j \in I \setminus \{i\}$ be some opponent of player i. Then, $b_i[t_i^*]\big(\{c_j\} \times F_j^{-1}(r_j) \times \{u_j\}\big) > 0$, as

$$b_i[t_i^*]\big(\{c_{-i}\} \times F_{-i}^{-1}(r_{-i}) \times \{u_{-i}\}\big) = \nu_i^{c_i, r_i^*, u_i}(c_{-i}, r_{-i}, u_{-i}) > 0.$$

Consider some $t_j \in F_j^{-1}(r_j)$ such that $b_i[t_i^*](c_j, t_j, u_j) > 0$. Since t_i^* expresses up to k-fold belief in rationality, c_j is optimal for (t_j, u_j), where t_j expresses up to $(k-1)$-fold belief in rationality, and by construction of F as well as by Perea (2014), Theorem 4, it is the case that $t_j^U = r_j^U$. Hence, $c_j \in C_j^{k-1}(r_j, u_j)$, and by the inductive assumption it follows that $c_j \in ICR_j^{k-1}(r_j, u_j)$. Therefore, it holds that $\text{marg}_{R_{-i} \times U_{-i}} \nu_i^{c_i, r_i^*, u_i} = \tau_i[r_i^*]$, the choice c_i is optimal for $(\text{marg}_{C_{-i}} \nu_i^{c_i, r_i^*, u_i}, u_i)$, and that $\nu_i^{c_i, r_i^*, u_i}(c_{-i}, r_{-i}, u_{-i}) > 0$ implies $c_j \in ICR_j^{k-1}(r_j, u_j)$ for all $j \in I \setminus \{i\}$. Consequently, $c_i \in ICR_i^k(r_i^*, u_i)$. It follows by induction that $\bigcap_{k \geq 0} C_j^k(r_j, u_j) \subseteq ICR_j(r_j, u_j)$ for all $j \in I$, for all $r_j \in R_j$, and for all

$u_j \in U_j$. Now, take some type $t_i \in T_i$ that expresses common belief in rationality such that $t_i^U = r_i^U$, and some chocie $c_i \in C_i$ that is optimal for (t_i, u_i). Then, $c_i \in \bigcap_{k \geq 0} C_i^k(r_i, u_i)$ and hence $c_i \in ICR_i(r_i, u_i)$. ■

This section is concluded with an illustration of interim correlated rationalizability by applying the concept to the incomplete information game of Example 1.

Example 4. Consider again the two player game with incomplete information as described in Figure 1.

Suppose the Dekel-Fudenberg-Morris model $\mathcal{R}^\Gamma$ of Γ given by the sets of Dekel-Fudenberg-Morris types $R_{Alice} = \{r_A, r'_A\}$, $R_{Bob} = \{r_B, r'_B\}$, and the following probability measures

- $\tau_{Alice}[r_A] = \frac{1}{2}(r_B, u_B) + \frac{1}{2}(r'_B, u'_B)$,
- $\tau_{Alice}[r'_A] = (r_B, u_B)$,
- $\tau_{Bob}[r_B] = \frac{1}{2}(r_A, u_A) + \frac{1}{2}(r'_A, u'_A)$,
- $\tau_{Bob}[r'_B] = (r_A, u_A)$.

Observe that

- $ICR^1_{Alice}(r_A, u_A) = ICR^1_{Alice}(r_A, u'_A) = ICR^1_{Alice}(r'_A, u_A) = ICR^1_{Alice}(r'_A, u'_A) = \{a, b\}$,
- $ICR^1_{Bob}(r_B, u_B) = ICR^1_{Bob}(r_B, u'_B) = ICR^1_{Bob}(r'_B, u_B) = ICR^1_{Bob}(r'_B, u'_B) = \{d, e\}$,
- $ICR^2_{Alice}(r_A, u_A) = ICR^2_{Alice}(r'_A, u_A) = \{a\}$ and $ICR^2_{Alice}(r_A, u'_A) = ICR^2_{Alice}(r'_A, u'_A) = \{a, b\}$,
- $ICR^2_{Bob}(r_B, u_B) = ICR^2_{Bob}(r'_B, u_B) = \{d\}$ and $ICR^2_{Bob}(r_B, u'_B) = ICR^2_{Bob}(r'_B, u'_B) = \{d, e\}$,
- $ICR^3_{Alice}(r_A, u_A) = ICR^3_{Alice}(r'_A, u_A) = \{a\}$, $ICR^3_{Alice}(r_A, u'_A) = \{a, b\}$, and $ICR^3_{Alice}(r'_A, u'_A) = \{b\}$,
- $ICR^3_{Bob}(r_B, u_B) = ICR^3_{Bob}(r'_B, u_B) = \{d\}$, $ICR^3_{Bob}(r_B, u'_B) = \{d, e\}$, and $ICR^3_{Bob}(r'_B, u'_B) = \{e\}$.

– $ICR^4_{Alice}(r_A, u_A) = ICR^4_{Alice}(r'_A, u_A) = ICR^4_{Alice}(r_A, u'_A) = \{a\}$, and $ICR^4_{Alice}(r'_A, u'_A) = \{b\}$,

– $ICR^4_{Bob}(r_B, u_B) = ICR^4_{Bob}(r'_B, u_B) = ICR^4_{Bob}(r_B, u'_B) = \{d\}$, and $ICR^4_{Bob}(r'_B, u'_B) = \{e\}$.

The procedure of interim correlated rationalizability thus stops after 4 rounds and the output is $ICR_{Alice}(r_A, u_A) = ICR_{Alice}(r'_A, u_A) = ICR_{Alice}(r_A, u'_A) = \{a\}$, and $ICR_{Alice}(r'_A, u'_A) = \{b\}$ for *Alice* as well as $ICR_{Bob}(r_B, u_B) = ICR_{Bob}(r'_B, u_B) = ICR_{Bob}(r_B, u'_B) = \{d\}$, and $ICR_{Bob}(r'_B, u'_B) = \{e\}$ for *Bob*. Note that the choice Dekel-Fudenberg-Morris type utility function tuples selected by interim correlated rationalizability induce the choice utility function pairs (a, u_A) ,(a, u'_A), and (b, u'_A) for *Alice* as well as (d, u_B) , (d, u'_B), and (e, u'_B) for *Bob*. These are exactly the choice utility function pairs selected by generalized iterated strict dominance. Hence, the optimal choices under interim correlated rationalizability and common belief in rationality are the same in this example. ♣

7 Complete Information

So far games with incomplete information have been considered. In particular, a basic non-equilibrium way of strategic reasoning has been spelled out in the face of payoff uncertainty. The construction has been conducted epistemically, i.e. with common belief in rationality, as well as algorithmically, i.e. with generalized iterated strict dominance. Now, the question could be posed what conditions on the interactive reasoning of players in incomplete information games actually dissolve payoff uncertainty. In particular, such conditions would need to restrict the marginal belief hierarchies with respect to the players' utility functions. Before this question can be tackled, the notion of complete information needs to be formally defined in epistemic structures.

Intuitively, complete information means that there is no uncertainty about any player's utility function at any level of interactive reasoning. Given some player $i \in I$, a type utility function pair $(t_i, u_i) \in T_i \times U_i$ can then be said to express complete information, if there exists for every opponent $j \in I \setminus \{i\}$ a utility function $u_j \in U_j$ such that t_i's marginal belief hierarchy t_i^U on utilities is generated by $\big(u_i, (u_j)_{j \in I \setminus \{i\}}\big)$, i.e. $b_i[t_i]\big((u_j)_{j \in I \setminus \{i\}}\big) = 1$, for every opponent $j \in I \setminus \{i\}$ player i only deems possible types $t_j \in T_j$ such that $b_j[t_j]\big((u_k)_{k \in I \setminus \{j\}}\big) = 1$, and for every opponent $j \in I \setminus \{i\}$ player i only deems possible types $t_j \in T_j$ that for every opponent $k \in I \setminus \{j\}$ only deem possible types $t_k \in T_k$ such that $b_k[t_k]\big((u_l)_{l \in I \setminus \{k\}}\big) = 1$, etc. Note that complete information is not defined simply for a type but for a type utility function tuple with the reasoner's actual utility function.

Also, the notion of correct beliefs needs to be invoked in the context of the players' utility functions. A type utility function tuple (t_i, u_i) is said to believe some opponent j to be correct about his utility function and marginal belief hierarchy t_i^U on utilities, if t_i only deems possible types t_j such that $b_j[t_j](u_i) = 1$ and $b_j[t_j]$ assigns probability 1 to t_i^U. Compared to complete information correct beliefs are defined for a type utility function tuple instead of merely for a type, since correct beliefs in the context of payoff uncertainty also concern the reasoner's utility function. With complete information and correct beliefs formally defined, the following theorem characterizes complete information with three doxastic correctness conditions.

Theorem 4. *Let $\Gamma = \big(I, (C_i)_{i \in I}, (U_i)_{i \in I}\big)$ be a game with incomplete information, $\mathcal{M}^\Gamma$ some epistemic model of it, and $i \in I$ some player. A type utility function tuple $(t_i, u_i) \in T_i \times U_i$ of player i expresses complete information, if and only if,*

- *for every opponent $j \in I \setminus \{i\}$, type t_i only deems possible types $t_j \in T_j$ that are correct about i's utility function u_i and marginal belief hierarchy on utilities,*
- *for every opponent $j \in I \setminus \{i\}$, type t_i only deems possible type utility function pairs $(t_j, u_j) \in T_j \times U_j$ that only deem possible types $t_i' \in T_i$ that are correct about j's utilities and j's marginal belief hierarchy on utilities,*

– *for all opponents $j \in I \setminus \{i\}$ and $k \in \setminus\{i,j\}$, type t_i only deems possible types $t_j \in T_j$ that have the same marginal belief on k's utilities and on k's marginal belief hierarchies on utilities as t_i has.*

Proof. Since only t_i's marginal belief hierarchy on utilities is affected by incomplete information and the three doxastic conditions, attention can be restricted to the induced marginal type t_i^U.

For the *if* direction of the theorem suppose that i's utility function is $u_i \in U_i$ and that t_i satisfies the three correctness of beliefs conditions. It is first shown that t_i's marginal type t_i^U only deems possible a unique marginal type t_j^U and a unique utility function $u_j \in U_j$ for every opponent $j \in I \setminus \{i\}$. Towards a contradiction assume that t_i^U assigns positive probability to at least two marginal type utility function pairs (t_j^U, u_j) and $(t_j^{U\prime}, u_j')$ for some opponent $j \in I \setminus \{i\}$. Since t_i believes that j is correct about his utility function and marginal belief hierarchy on utilities, t_i believes that j only deems possible (t_i^U, u_i). Consequently, the marginal type utility function pairs (t_j^U, u_j) and $(t_j^{U\prime}, u_j')$ both only deem possible (t_i^U, u_i). Consider marginal type t_j^U and note that (t_j^U, u_j) believes that i deems it possible that j is characterized by the marginal type utility function tuple $(t_j^{U\prime}, u_j')$. Hence, (t_j^U, u_j) does not believe that i is correct about his utility function and marginal belief hierarchy on utilities. It follows that t_i deems it possible that j does not believe that i is correct about his utility function and marginal belief hierarchy on utilities, a contradiction. For every opponent $j \in I \setminus \{i\}$, type t_i's marginal type t_i^U thus assigns probability 1 to a single marginal type utility function tuple (t_j^U, u_j) and the corresponding type t_j assigns probability 1 to (t_i^U, u_i). By the third condition in Theorem 4 it is ensured that for each opponent the respective other opponents share the same marginal belief on utilities, and thus it follows, by induction, that t_i's marginal belief hierarchy on utilities is generated by $(u_j)_{j \in I}$ and therefore (t_i, u_i) expresses complete information.

For the *only if* direction of the theorem, suppose that (t_i, u_i) expresses complete information and let $(u_j)_{j \in I} \in \times_{j \in I} U_j$ be the tuple of utility functions generating t_i's marginal belief hierarchy on utilities. Then, it directly follows that the three doxastic conditions hold. ■

From a conceptual point of view complete information can thus be modelled entirely within the mind of the reasoner satsfying the three conditions of Theorem 4 instead of restricting the game specification. Accordingly, the specific case of payoff certainty can be obtained subjectively or objectively.

The epistemic and algorithmic concepts of common belief in rationality according to Definition 4 and generalized iterated strict dominance according to Definition 6, respectively, can be considered in the special case of complete information. Indeed, both concepts are then equivalent to their natural complete information analogues.

In epistemic models for complete information games the induced belief functions assign to every type a probability measure on the set of opponents' choice type combinations and not choice type utility function combinations. Interactive uncertainty about payoffs is not modelled, as it is absent from the underlying game. However, common belief in rationality is defined in exactly the same way as in Definition 4 with the only immediate difference that Γ is a game with complete information. In the case of complete information, optimality and belief in rationality are not defined with respect to type utility function pairs, but only with respect to types. Common belief in rationality for incomplete information games with a single utility function for every player is thus equivalent to the standard definition of common belief in rationality for complete information games.

Generalized iterated strict dominance joins the class of solution concepts for incomplete information games. For complete information games the algorithm is equivalent to iterated strict dominance. To recall the definition of iterated strict dominance, let $\Gamma = \big(I, (C_i)_{i \in I}, (u_i)_{i \in I}\big)$ be a complete information game, and consider the sets $C_i^0 := C_i$ and

$$C_i^k := C_i^{k-1} \setminus \{c_i \in C_i : \text{ there exists } r_i \in \Delta(C_i^{k-1})$$

$$\text{such that } u_i(c_i, c_{-i}) < \sum_{c_i' \in C_i} r_i(c_i') \cdot u_i(c_i', c_{-i}) \text{ for all } c_{-i} \in C_{-i}^{k-1}\}$$

for all $k > 0$ and for all $i \in I$. The output of iterated strict dominance is then defined as $ISD := \times_{i \in I} ISD_i \subseteq \times_{i \in I} C_i$, where $ISD_i := \bigcap_{k \geq 0} C_i^k$ for every player $i \in I$. With complete information there is for every player i and for every round k a unique decision problem $\Gamma_i^k(u_i) = \big(C_i^k(u_i), C_{-i}^k(u_i), u_i\big)$, as payoff uncertainty vanishes. Thus, $C_{-i}^k(u_i) = \times_{j \in I \setminus \{i\}} C_j^k$, $C_i^k(u_i) = C_i^k$, and Definition 6 then becomes a formulation of iterated strict dominance in terms of decision problems. Consequently, generalized iterated strict dominance for incomplete information games with a single utility function for every player is equivalent to iterated strict dominance for complete information games.

8 Conclusion

The basic epistemic notion of common belief in rationality has been considered within a one-person perspective model of incomplete information static games that is kept as parsimonious and simple as possible. The algorithmic characterization of this concept in terms of decision problems and strict dominance arguments has led to the non-equilibrium solution concept of generalized iterated strict dominance, which can be seen as a direct incomplete information analogue to iterated strict dominance. This rather natural and basic algorithm provides a tool for economists when analyzing situations involving payoff uncertainty. Due to its simplicity, generalized strict dominance seems suitable for a broad spectrum of potential applications, including management and political theory.

References

Bach, C. W. and Perea, A. (2017): Incomplete Information and Equilibrium. Mimeo.

Battigalli, P. (2003): Rationalizability in Infinite, Dynamic Games of Incomplete Information. *Research in Economics* 57, 1–38.

Battigalli, P. (2006): Rationalization in Signaling Games: Theory and Applications. *International Game Theory Review* 8, 1–27.

Battigalli, P., Di Tillio, E., Grillo, E. and Penta, A. (2011): Interactive Epistemology and Solution Concepts for Games with Asymmetric Information. *B. E. Journal of Theoretical Economics* 11, 1935–1704.

Battigalli, P. and Prestipino, A. (2013): Transparent Restrictions on Beliefs and Forward-Induction Reasoning in Games with Asymmetric Information. *B. E. Journal of Theoretical Economics* 13, 79–130.

Battigalli, P. and Siniscalchi, M. (1999): Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games. *Journal of Economic Theory* 88, 188–230.

Battigalli, P. and Siniscalchi, M. (2002): Strong Belief and Forward Induction Reasoning. *Journal of Economic Theory* 106, 356–391.

Battigalli, P. and Siniscalchi, M. (2003a): Rationalization and Incomplete Information. *B. E. Journal of Theoretical Economics* 3, 1534–5963.

Battigalli, P. and Siniscalchi, M. (2003b): Rationalizable Bidding in First Price Auctions. *Games and Economic Behavior* 45, 38–72.

Battigalli, P. and Siniscalchi, M. (2007): Interactive Epistemology in Games with Payoff Uncertainty. *Research in Economics* 61, 165–184.

Bernheim, B. D. (1984): Rationalizable Strategic Behavior. *Econometrica*, 52, 1007–1028.

Brandenburger, A. and Dekel, E. (1987): Rationalizability and Correlated Equilibria. *Econometrica* 55, 1391–1402.

Dekel, E., Fudenberg, D. and Morris, S. (2007): Interim Correlated Rationalizability. *Theoretical Economics* 1, 15–40.

Dekel, E. and Siniscalchi, M. (2015): Epistemic Game Theory. In *Handbook of Game Theory with Economic Applications*, Vol. 4, 619–702.

Ely, J. C. and Pęski, M. (2006): Hierarchies of Belief and Interim Rationalizability. *Theoretical Economics* 1, 19–65.

Harsanyi, J. C. (1967-68): Games of Incomplete Information played by "Bayesian Players". Part I, II, III. *Management Science* 14, 159–182, 320–334, 486–502.

Heifetz, A. and Samet, D. (1998): Topology-Free Typology of Beliefs. *Journal of Economic Theory* 82, 324–341.

Ollar, M. and Penta, A. (2017): Full Implementation and Belief Restrictions. Mimeo.

Pearce, D. (1984): Rationalizable Strategic Behavior and the Problem of Perfection. *Econometrica*, 52, 1029–1050.

Penta, A. (2012): Higher Order Uncertainty and Information: Static and Dynamic Games. *Econometrica*, 80, 631–660.

Penta, A. (2013): On the Structure of Rationalizability on Arbitrary Spaces of Uncertainty. *Theoretical Economics*, 8, 405–430.

Penta, A. (2015): Robust Dynamic Implementation. *Journal of Economic Theory*, 160, 280–316.

Penta, A. (2017): Backward Induction Reasoning in Incomplete Information Games. Mimeo.

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

Perea, A. (2014): When Do Types Induce the Same Belief Hierarchy? The Case of Finitely Many Types. *EPICENTER Working Paper* 1.

Tan, T. and Werlang, S. R. C. (1988): The Bayesian Foundations of Solution Concepts of Games. *Journal of Economic Theory* 45, 370–391.

Weinstein, J. and Yildiz, M. (2007): A Structure Theorem for Rationalizability with Application to Robust Predictions of Refinements. *Econometrica* 75, 365–400.