Common assumption of cautious rationality and iterated admissibility*

Emiliano Catonini†

September 2012

Abstract. Iterated admissibility is one of the most appealing solution concepts for complete-information strategic-form games. Still, to understand when it is the appropriate one, conditions under which players want to avoid strategies that are weakly dominated in some reduced game along the procedure (although possibly not in the final set!) must be provided. It is intuitive that these conditions have to incorporate some cautious attitude of the players. Yet, to what extent players are cautious and assume that opponents are must be carefully defined in order to provide a correct motivation for iterated admissibility. Brandenburger, Friedenberg and Keisler (ECMA, 2008) define a notion of rationality, including an "open-mindedness" requirement for lexicographic beliefs, which delivers iterated admissibility when players adopt it, assume (to a defined extent) that opponents adopt it, and so on, up to some finite level. This notion of rationality cannot be commonly assumed by players unless heavy exogenous restrictions to beliefs apply. Here, I provide a weaker notion of cautiousness that can be commonly assumed by players and still captures iterated admissibility. This notion has a very clear and realistic interpretation. Moreover, I carry on the analysis in a type space that encompasses all meaningful lexicographic hierarchies of beliefs, the canonical one, of which I show constructively the existence.

Keywords: epistemic game theory, iterated admissibility, weak dominance, rationality, cautiousness, assumption.

*Preliminary Draft. I want to thank Adam Brandenburger, Amanda Friedenberg and Jerome Keisler for inspiring this work. An additional thank to Adam for the discussion at Stern. A special thank to Pierpaolo Battigalli for precious discussions and suggestions about my work. Thank you also to Burkhard Schipper and to the attendants of the theory informal workshop at Stern. I am grateful to Bocconi university for granting my visit at NYU in spring 2012.

†Bocconi University, PhD Candidate, emiliano.catonini@gmail.com

1 Introduction

In the huge variety of solution concepts for complete-information strategic-form games, iterated admissibility, i.e. iterated deletion of weakly dominated strategies, is surely one of the most appealing. First, it is a decision criterion that does not rely on any pre-existing equilibrium motivation: players can perform it from scratch through nothing else than their strategic reasoning. Second, it reflects an intuitevely reasonable way to behave: to the minimum, it avoids choosing a strategy when there is another one that, when it makes a difference, can only do better.[1] Still, it has to be identified more precisely when iterated admissibility is actually the appropriate solution concept and why, more generally, it is a sound way for players to choose their strategies.

The first step to this end is detecting which kind of conjectures and optimality concept motivate players to avoid strategies that are weakly dominated in some reduced game along the procedure. The following game,[2] where strategy L is eliminated in the first round and strategy B is eliminated in the second, is helpful to follow the next few arguments.

$1\backslash 2$	L	C	R
U	$(4,1)$	$(4,1)$	$(0,1)$
M	$(0,1)$	$(0,1)$	$(4,1)$
D	$(3,1)$	$(2,1)$	$(2,1)$
B	$(9,0)$	$(0,1)$	$(0,1)$

It is known from Pearce [14] that a strategy is not weakly dominated if and only if it is a best reply to some fully mixed conjecture over opponents' strategies. But notice that, differently than the iterated deletion of strongly dominated strategies, iterated admissibility can exclude strategies that are not weakly dominated in the final reduced game (L). To justify this, a player must still consider the possibility that some opponent might play some previously deleted strategy. But to what extent? If previously deleted strategies could be given a positive probability, a player would clearly run the opposite risk of rescuing strategies that are weakly dominated in the final reduced game (B). This tension is solved by lexicographic conjectures and lexicographic best replies.[3] A lexicographic conjecture is a finite list of simple conjectures in a priority ordering. They allow to take into consideration previously deleted strategies and yet, pushing them farther in the list, to deem them as infinitely less likely than strategies that survive more steps of the procedure. A lexicographic best reply is a strategy that, for any other strategy, does not worse than the latter against the conjectures of the list up to the end or up to one against which the former does strictly better.

[1] Moreover, in case the strategic form is derived from an extensive form game without relevant ties among payoffs, iterated admissibility operationalizes extensive form rationalizability ([14] and [2]). Yet, the analysis of the extensive form solution concept is required to understand the epistemic motivations: see [2].

[2] This example (with one strategy added) is due to Pierpaolo Battigalli.

[3] See also Stahl ([15])

Formal definitions will be provided in section 2. Notice that to justify strategy D player 1 must be allowed to hold a lexicographic conjecture with overlapping supports. In the final set, D is a best reply to the simple conjecture that considers C and R equally likely, but it is not a strict best reply. Hence player 1 may wonder about a secondary conjecture to check the desirability of D. If the secondary conjecture were obliged to put probability 1 on the deleted strategy L, strategy D would not be a lexicographic best reply. Instead, considering L and R equally likely as secondary hypothesis makes strategy D a lexicographic best reply.

The second step consists of finding the epistemic hypotheses that identify and conceptually motivate the right lexicographic conjectures, whose lexicographic best replies correspond to the iteratevely admissible strategies. These hypotheses will be defined as notions of cautious rationality, assumption of opponents' cautious rationality, and so on, which characterize players who perform iterated admissibility.

Here comes the contribution of the paper. In section 3, a canonical type space for lexicographic hierarchies of beliefs is constructed. The canonical type space allows players to conceive any meaningful lexicographic hierarchy of beliefs about strategies, so that no exogenous restriction is super-imposed and the states of interest will be entirely identified by the conceptually relevant events. In section 4, compelling notions of assumption, cautiousness and rationality are defined and put at work in this epistemic environment. These notions allow to construct events that not only capture any step of iterated admissibility but can also hold together, defining a cautious rationality and common assumption of cautious rationality non-empty event in which players "share" their being cautious and rational in the sense of this paper.

Brandenburger, Friedenberg and Keisler [6] (henceforth BFK), to whom this work is much indebted, define notions of rationality and assumption that, opportunely combined, deliver the iteratevely admissible strategies. They obtain this result by incorporating in rationality a very strong open-mindedness requirement: players put every state of world in the support of their beliefs, at some level of their lexicographic probability system over the state space. This means that players conceive at the same time every lexicographic hierarchy of beliefs allowed by the type space and consider it possible to some extent. Then, the authors prove the impossibility result that for a rich enough type structure (complete and continuous), players are unable to commonly assume this notion of rationality: the corresponding event is empty. The impossibility ceases to hold for poorer type structures, but this means imposing exogenous restrictions to the hierarchies of beliefs, which could find no justification in the context at hand. Now, suppose that players were able to prove to each other that they are rational in this sense. Then players should assume that everyone is rational; assume that everyone is rational and assumes that everyone is rational; and so on. But if common assumption of rationality is impossible, at some point players must start forming doubts. Why should they? This puzzling result has inspired different papers other than this. Keisler and Lee [12] show that relaxing the continuity hypothesis in the type space, the impossibility may cease to hold. Heifetz, Meier and Schipper [9] take a more radical way out by changing the solution concept.[4] The aim of this paper, instead, is to epistemically characterize precisely iterated admissibility, for its intuitive appeal, but obtaining a non-empty "cautious rationality and common assumption of cautious rationality" event through interpretationally clear innovations. Switching from open-mindedness to a milder cautiousness requirement allows to preserve the characterization and let players commonly believe in their cautiousness and rationality. The idea is simple and realistic: players cannot or are not interested in conceiving and weighing all possible hierarchies of beliefs at the same time.[5] Cautious players just conceive all possible opponents' strategies, the payoff relevant objects. Then, they make a minimal use of higher-order beliefs[6] to put those strategies in a likelihood order, according to hypotheses about opponents' strategic reasoning. For instance, opponents' strategies that are best replies to some cautious conjecture (i.e. cautiously rational ones) are given priority with respect to the ones that are not. Such reduction of the computational burden for players is strictly connected with their ability to commonly assume this notion of cautious rationality.

2 Iterated admissibility and lexicographic beliefs

For all the following player-specific sets X_i, let $X := \prod_{j \in I} X_j$ and $X_{-i} := \prod_{j \neq i} X_j$.

Consider a finite strategic form game $\langle I, (S_i, u_i)_{i \in I} \rangle$, where I is the set of players and for every $i \in I$, S_i is the set of strategies and $u_i : S \to \mathbb{R}$ is the payoff function. For any finite set X, let $\Delta(X)$ be the set of probability measures on it. Define the expected payoff function π_i on $\Delta(S_i) \times \Delta(S_{-i})$ by setting for every $(\sigma_i, \sigma_{-i}) \in \Delta(S_i) \times \Delta(S_{-i})$,

$$\pi_i(\sigma_i, \sigma_{-i}) := \sum_{s_{-i} \in supp \sigma_{-i}} \sum_{s_i \in supp \sigma_i} u_i(s_i, s_{-i}) \sigma_i(s_i) \sigma_{-i}(s_{-i}).$$

A pure strategy or a pure opponents' subprofile of strategies as argument of π_i will indicate the probability distribution putting probability 1 on it.

[4] Also Asheim and Dufwemberg [1] defined a solution concept (fully admissible sets) that captures a form of cautiousness and full belief in rationality and that does not refine, nor is refined, by iterated admissibility.

[5] This is different than impoverishing the type structure: players can still conceive all the meaningful hierarchies of beliefs, simply they will not be obliged to.

[6] I do not rule out in any way that players can put more than necessary or even all hierarchies of beliefs in their conjectures. But the possibility to make a parsimonious use of them is enough to allow common assumption of cautious rationality.

Iterated admissibility is a reduction procedure of the set of strategy profiles that relies on a weak dominance criterion.

Definition 1 *For every player $i \in I$, take a set $\widehat{S}_i \subseteq S_i$. For every strategy $s_i \in \widehat{S}_i$, s_i is weakly dominated over $\widehat{S}$ if there exists $\sigma_i \in \Delta(\widehat{S}_i)$ such that for every $s_{-i} \in \widehat{S}_{-i}$, $\pi_i(s_i, s_{-i}) \leq \pi_i(\sigma_i, s_{-i})$ and there exists $\widehat{s}_{-i} \in \widehat{S}_{-i}$ such that $\pi_i(s_i, \widehat{s}_{-i}) < \pi_i(\sigma_i, \widehat{s}_{-i})$.*

Now iterated admissibility can be defined formally.

Definition 2 *The iterated admissibility procedure is a finite chain of cartesian sets of strategy profiles $S^0 := \prod_{i \in I} S_i^0 \supset ... \supset S^M := \prod_{i \in I} S_i^M$ such that for every $i \in I$ and $s_i \in S_i$:*

1. *$S_i^0 = S_i$;*
2. *for every $n < M$, $s_i \in S_i^{n+1}$ if and only if $s_i \in S_i^n$ and s_i is not weakly dominated over S^n;*
3. *$s_i \in S_i^M$ if and only if s_i is not weakly dominated over S^M.*

Notice that inclusions are strict: then, the chain is finite because the sets of strategies are finite. Moreover, S^M is non-empty because for a player there is always at least one strategy that is not weakly dominated.

When a strategy is not weakly dominated over a set, there exists a fully mixed conjecture over opponents' subprofiles in the set against which the strategy is a best reply.

Proposition 3 *Consider a cartesian set of strategy profiles $\widehat{S} \subseteq S$. For every $i \in I$ and $s_i \in \widehat{S}_i$, if s_i is not weakly dominated over $\widehat{S}$, then there exists $\sigma_{-i} \in \Delta(\widehat{S}_{-i})$ such that for every $s_{-i} \in \widehat{S}_{-i}$, $\sigma_{-i}(s_{-i}) > 0$ and for every $\widehat{s}_i \in \widehat{S}_i$, $\pi_i(s_i, \sigma_{-i}) \geq \pi_i(\widehat{s}_i, \sigma_{-i})$.*

As already argued, looking only at simple fully mixed conjectures may wrongly justify the choice of an iteratevely inadmissible strategy: for a player $i \in I$ there may be strategies that are not weakly dominated over S^M and yet do not belong to S_i^M. The reason is that a player who performs iterated admissibility wants to avoid also strategies that are weakly dominated over some previous set of the chain. Thus, she considers every opponents' subprofile in that set still possible to some extent, but the ones that do not survive the following step are not considered nearly as likely as the ones that do. Therefore, the epistemic characterization will need lists of conjectures that allow to put the states of the world at uncomparable levels of likelihood. These lists are defined here as *lexicographic beliefs*.

Definition 4 *Consider a measurable space X and let $\Delta(X)$ denote the space of probability measures on its Borel field. A lexicographic belief is a finite list $\lambda = (\lambda_1, ..., \lambda_k) \in (\Delta(X))^k$ of such probability measures.*

I will denote by $\Delta^{LEX}(X) := \bigcup_{k \in \mathbb{N}} (\Delta(X))^k$ the set of all lexicographic beliefs over X.

When X is the space of opponents' strategy subprofiles, I will call the lexicographic beliefs *lexicographic conjectures.* As argued in the introduction, I am interested in lexicographic conjectures with possibly overlapping supports, i.e. where there can exist $n \neq m$ such that $supp\lambda_n \bigcap supp\lambda_m \neq \emptyset$. With respect to lexicographic conjectures, I take the standard definition of *lexicographic best reply.*

Definition 5 *Consider a player $i \in I$ and a lexicographic conjecture $\lambda \in \Delta^{LEX}(S_{-i})$. A strategy $s_i \in S_i$ is a lexicographic best reply to $\lambda = (\lambda_1, ..., \lambda_k)$ if for every $s_i' \neq s_i$, there exists $j \leq k$ such that for every $h \leq j$, $\pi_i(s_i, \lambda_h) \geq \pi_i(s_i', \lambda_h)$ and, if $j < k$, $\pi_i(s_i, \lambda_j) > \pi_i(s_i', \lambda_j)$.*

Instead, when X will be the section with respect to opponents of the state space (cross product of the strategy space and of the type space I will construct), I will be interested in lexicographic beliefs with nonoverlapping measures.[7] Such lexicographic beliefs represent a list of mutually exclusive hypotheses about the state of the world: the primary hypothesis, the secondary hypothesis, and so on. This will not prevent marginal lexicographic beliefs on strategies to have overlapping supports; it will just require the belief in the same opponents' strategy subprofile in two different hypotheses to be motivated by two different states of the world. As in BFK, this property is called *mutual singularity* and the lexicographic beliefs that satisfy it are called *lexicographic probability systems.*[8]

Definition 6 *Consider a measurable space X. A lexicographic belief $\lambda = (\lambda_1, ..., \lambda_k) \in \Delta^{LEX}(X)$ is mutually singular if there are measurable sets $E_1, ..., E_k$ in X such that for every $j \leq k$ and $h \neq j$, $\lambda_j(E_j) = 1$ and $\lambda_j(E_h) = 0$. A mutually singular lexicographic belief is called lexicographic probability system.*

I will denote by $\Delta^{LPS}(X) \subset \Delta^{LEX}(X)$ the set of all lexicographic probability systems (henceforth, LPS) over X.

Lexicographic hierachies of beliefs about strategies will be defined in the next section, where they are used to construct the type space that captures them.

[7] This is actually a slightly weaker requirement than nonoverlapping supports when the underlying space is infinite.

[8] The term was coined by Blume, Brandenburger and Dekel [3] with reference also to lists of overlapping measures.

3 A canonical type space for lexicographic hierarchies of beliefs

Here I construct a canonical type space for lexicographic hierachies of beliefs.

I will metrize spaces as follows:

- S_i with the discrete metric;
- $\Delta(X)$, where X is a separable complete metric space (Polish), with the Prohorov metric;
- $\Delta^{LEX}(X)$, where X is a Polish space, by setting the distance between two elements of the same lenght $\lambda = (\lambda_1, ..., \lambda_k)$ and $\widehat{\lambda} = (\widehat{\lambda}_1, ..., \widehat{\lambda}_k)$ as the maximum over $h \leq k$ of the Prohorov distances between λ_h and $\widehat{\lambda}_h$, and the distance between two elements of different lenghts to 1;[9]
- the product of Polish spaces with the product metric.

With these choices, all the spaces are Polish themselves (see [8]).

For every $i \in I$ and $n \in \mathbb{N}$, define inductively the following sets:

$$\begin{aligned} X_i^1 &:= S_{-i}; \\ X_i^{n+1} &:= X_i^n \times \prod_{j \neq i} \Delta^{LEX}(X_{-j}^n). \end{aligned}$$

Moreover, define

$$Z_i^1 := X_i^1,\ Z_i^{n+1} := \prod_{j \neq i} \Delta^{LEX}(X_{-j}^n);$$

then $X_i^n = \prod_{m=1}^{n} Z_i^m$.

Now I can define a *(coherent) lexicographic hierarchy of beliefs about strategies.*

Definition 7 *A (coherent) lexicographic hierarchy of beliefs about strategies is a finite list $\boldsymbol{\delta} = (\delta_1, ..., \delta_k)$ such that for every $h \leq k$, $\delta_h = (\delta^1, \delta^2, ...) \in \prod_{n \in \mathbb{N}} \Delta(X_i^n)$ (and for every $n \in \mathbb{N}$, $\operatorname{marg}_{X_i^n} \delta^{n+1} = \delta^n$)*

Since all the sets previously defined are Polish spaces, the following version of lemma 1 in Brandenburger and Dekel [5] holds.[10]

[9] The Prohorov distance between two elements is at most 1, so triangular inequality is respected.

[10] In [5] the lemma only claims the existence of the homeomorphism because it suffices for the purposes of the paper. However, their proof constructs exactly the homeomorphism specified here through a version of Kolmogorov Existence Theorem (from [8]), which also claims the uniqueness of the images with respect to the marginals requirement (hence the uniqueness of the function with this feature).

Lemma 8 *Let* $D_i := \left\{ \delta = (\delta^1, \delta^2, ...) \in \prod_{n\in\mathbb{N}} \Delta(X_i^n) : \mathrm{marg}_{X_i^n} \delta^{n+1} = \delta^n \right\}$*. There exists a unique function* $f_i : D_i \to \Delta(\prod_{n\in\mathbb{N}} Z_i^n)$ *such that for every* $\delta \in D_i$ *and* $h \in \mathbb{N}$*,* $\mathrm{marg}_{X_i^h} f_i(\delta) = \delta^h$*. Moreover,* f_i *is a homeomorphism.*

Proof. See [5].

Define the set of coherent lexicographic hierarchies of beliefs $C_i := \bigcup_{k\in\mathbb{N}} (D_i)^k$ and metrize it by setting the distance between two elements of the same lenght $\boldsymbol{\delta} = (\delta_1, ..., \delta_k)$ and $\boldsymbol{\delta}' = (\delta'_1, ..., \delta'_k)$ as the maximum over $h \leq k$ of the distances between δ_h and δ'_h, and the distance between two elements of different lenghts to 1. Then the function

$$g_i : C_i \to \Delta^{LEX}(\prod_{n\in\mathbb{N}} Z_i^n) \text{ such that } g_i(\boldsymbol{\delta} = (\delta_1, ..., \delta_k)) := (f_i(\delta_1), ..., f_i(\delta_k)).$$

is a homeomorphism.[11]

Clearly $\prod_{n\in\mathbb{N}} Z_i^n$ is a strict superset of $S_{-i} \times C_{-i}$ because $\prod_{n>1} Z_i^n$ contains also non coherent hierarchies. Moreover I want to achieve mutual singularity in the final type space. The following inductive procedure allows to restrict the sets in the desired way and close the final type space.

Define:

- $\Lambda_i^0 := \left\{ \boldsymbol{\delta} \in C_i : g_i(\boldsymbol{\delta}) \in \Delta^{LPS}(\prod_{n\in\mathbb{N}} Z_i^n) \right\}$;
- $\Lambda_i^n := \left\{ \boldsymbol{\delta} = (\delta_1, ..., \delta_k) \in \Lambda_i^{n-1} : \forall h = 1, ..., k, \ f_i(\delta_h)[S_{-i} \times \Lambda_{-i}^{n-1}] = 1 \right\}$;
- $\Lambda_i := \bigcap_{n\in\mathbb{N}} \Lambda_i^n$.

I have to show that for every $n \in \mathbb{N}$, Λ_i^n is well defined, that is, $S_{-i} \times \Lambda_{-i}^{n-1}$ is measurable.

By corollary C.1 in BFK, for every Polish space X, $\Delta^{LPS}(X)$ is a Borel set in $\Delta^{LEX}(X)$. The function g_i is measurable. Hence, Λ_i^0 is a Borel set in C_i and Λ_{-i}^0 is a Borel set in C_{-i}.

By theorem 17.24 in [11], for every Polish space X, the Borel sigma-algebra on $\Delta(X)$ generated by the Prohorov metric is generated also by the family of maps $\mu \mapsto \mu(A)$ with $\mu \in \Delta(X)$ and Borel set $A \subseteq X$. This requires that for every Borel set $W \subseteq X$, the set $\left\{ \delta \in \Delta(X) : \delta(W^C) > 0 \right\}$ is Borel, hence its complement $\{\delta \in \Delta(X) : \delta(W) = 1\}$ is Borel too. For every lenght $k \in \mathbb{N}$

[11] For any $\boldsymbol{\delta} = (\delta_1, ..., \delta_k) \in C_i$, take a ball around $g_i(\boldsymbol{\delta}) = (f_i(\delta_1), ..., f_i(\delta_k))$ of radius ρ. Since f_i is a homeomorphism, for every $h \leq k$ and for the ball around $f_i(\delta_h)$ of radius ρ, there is a ball around δ_h whose image is contained in the previous ball. Take the smallest radius ε among those balls around δ_h over $h \leq k$. The image of the ball around $\boldsymbol{\delta}$ of radius ε is contained in the ball of $g_i(\boldsymbol{\delta})$ of radius ρ. The same reasoning can be applied inverting g_i.

and $h \leq k$, the projection function $\lambda = (\lambda_1, ..., \lambda_k) \mapsto \lambda_h$ with $\lambda \in (\Delta(X))^k$ is continuous, hence since $\{\delta \in \Delta(X) : \delta(W) = 1\}$ is Borel,

$$L_h^k := \left\{\lambda = (\lambda_1, ..., \lambda_k) \ \in (\Delta(X))^k : \lambda_h(W) = 1\right\}$$

is Borel too.

$$L^k := \left\{\lambda = (\lambda_1, ..., \lambda_k) \in (\Delta(X))^k : \forall h = 1, ..., k, \ \lambda_h(W) = 1\right\} = \bigcap_{h=1,...,k} L_h^k,$$

so it is Borel too.

$$L := \left\{\lambda = (\lambda_1, ..., \lambda_l) \in \Delta^{LEX}(X) : \forall h = 1, ..., l, \ \lambda_h(W) = 1\right\} = \bigcup_{k\in\mathbb{N}} L^k,$$

so it is Borel too. Setting $X := \prod_{n\in\mathbb{N}} Z_i^n$, and $W := S_{-i} \times \Lambda_{-i}^{n-1}$, $\Lambda_i^n = g_i^{-1}(L) \bigcap \Lambda_i^{n-1}$, so it is Borel. Hence, Λ_{-i}^n is a Borel set in C_{-i}.

Now consider that:

- Λ_i is homeomorphic to $g_i(\Lambda_i)$;
- $g_i(\Lambda_i) = \left\{\lambda = (\lambda_1, ..., \lambda_k) \in \Delta^{LPS}(\prod_{n\in\mathbb{N}} Z_i^n) : \forall h = 1, ..., k, \ \lambda_k[S_{-i} \times \Lambda_{-i}] = 1\right\}$, because g_i is onto;
- the latter is homeomorphic to $\Delta^{LPS}(S_{-i} \times \Lambda_{-i})$.

The last homeomorphism is the function that preserves the measures of all sets. Redefine g_i as the composition of itself with this last homeomorphism. So g_i is now a homeomorphism between Λ_i and $\Delta^{LPS}(S_{-i} \times \Lambda_{-i})$ such that for every $\delta \in \Lambda_i$ and for every $h \in \mathbb{N}$, $marg_{X_i^h} g_i(\delta) = \delta^h$. This closes the canonical type space for LPS ($(T_i, g_i)_{i\in I}$ from now on).

All hierarchies in Λ_i are *collectively coherent* (they are coherent, believe that opponents are coherent, and so on); moreover, they display common certainty in mutual singularity. The type space is canonical in the sense that it represents all hierarchies of this kind. Notice that the common certainty in mutual singularity does not mean that the lexicographic hierarchies are composed by mutually singular beliefs of all orders Indeed, beliefs of all first n orders could be even identical at different likelihood levels.

Heifetz, Meier and Schipper [9] construct a canonical type space for LPS with a bottom-up procedure, i.e. building directly only the desired hierarchies and putting them together in the type space. Since they introduce an epistemic hypothesis of mutual singularity of conjectures over opponents' strategies, they obtain mutual singularity in the final type space automatically. The top-down procedure here, instead, allows to throw away only those hierarchies whose representation as lexicographic beliefs over the state space is not mutually singular. Our construction is therefore bigger and the represented hierarchies can be composed by overlapping beliefs for any finite number of orders.

4 Common assumption of cautious rationality and the characterization theorem

In the canonical type space just constructed, the goal is now to identify the conceptually meaningful events that imply iteratevely admissible strategies as behavioral projections. These events will be the result of clear and realistic hypotheses about players' strategic reasoning, which allow to estabilish under which conditions iterated admissibility is the appropriate solution concept.

The first event of interest is the rationality one and it is based on the hypothesis that players play lexicographic best replies to their lexicographic conjectures.

Definition 9 *Rationality is the event* $R := \prod_{i \in I} R_i \subset S \times T$ *such that for every* $i \in I$ *and* $\omega_i = (s_i, t_i) \in R_i$, s_i *is a lexicographic best reply to* $\mathrm{marg}_{S_{-i}} g_i(t_i) = (\lambda_1, ..., \lambda_k)$.

The second event of interest is the cautiousness one and it is based on the hypothesis that players' lexicographic conjectures deem all the opponents' strategy subprofiles as possible to some extent.

Definition 10 *Cautiousness is the event* $C := \prod_{i \in I} C_i \subset S \times T$ *such that for every* $i \in I$ *and* $\omega_i = (s_i, t_i) \in C_i$, $\mathrm{marg}_{S_{-i}} g_i(t_i) = (\lambda_1, ..., \lambda_k)$ *has the following property: for every* $s_{-i} \in S_{-i}$ *there exists* $j \leq k$ *such that* $\lambda_k(s_{-i}) > 0$.

The conjunction of the two is the cautious rationality event. It is the one that translates into the use of a weak dominance criterion.

Definition 11 *Cautious rationality is the event* $R^1 := \prod_{i \in I} (R_i \bigcap C_i) = R \bigcap C$.

The projection on the strategy space of the event cautious rationality will coincide with the first iteration of the iterated admissibility procedure, i.e. with non weakly dominated strategies. To capture the further iterations, I need to identify the events where conjectures give the right priority to the iteratevely admissible strategies, in terms of their likelihood. These events are based on the hypothesis that players hold a kind of belief in opponents' cautious rationality up to some order. This kind of belief in an event (such as cautious rationality) shall not necessarily rule out completely that the event does not occur. This concept is defined here as *assumption*.

Definition 12 *A LPS* $\lambda_i = (\lambda_1, ..., \lambda_k) \in (\Delta(S_{-i} \times T_{-i}))^k$ *assumes* $B \subset S_{-i} \times T_{-i}$ *(at level h) if there exists* $h \leq k$ *such that:*

1. *there are measurable sets* $E_1, ..., E_k$ *in* $S_{-i} \times T_{-i}$ *such that for every* $j \leq h$, $E_j \subseteq B$ *and* $\lambda_j(E_j) = 1$ *and for every* $j > h$, $E_j \bigcap B = \emptyset$ *and* $\lambda_j(E_j) = 1$.

2. *for every* $s_{-i} \in \mathrm{proj}_{S_{-i}} B$, *there exists* $j \leq h$ *such that* $\mathrm{marg}_{S_{-i}} \lambda_j[s_{-i}] > 0$.

The first requirement has the interpretation that players deem the event *infinitely more likely* than its complementary. The definition of this concept in BFK is different because it applies again to open-minded (i.e. full support) LPS only, and has been given a preference-based representation.[12] The second requirement means that players consider every possible behavioral implication of the event infinitely more likely than all other strategy subprofiles. This reflects the view that players, as argued in the introduction, are concerned about conceving all possible moves by the opponents, while using higher-order beliefs only to rank them in a likelihood order. However, point 2 holds also in BFK, although in their model there is no need to specify it in the definition of assumption, because it is already a consequence of open-mindedness.

With this notion of assumption, the corresponding operator that maps subsets of the state space into subsets of the state space can be defined as follows:

$$A_i(B) := \{\omega_i = (s_i, t_i) \in (S_i \times T_i) : g_i(t_i) \text{ } assumes \text{ } B\}.$$

Using the cautious rationality event and the last operator, the right *cautious rationality and m–th order assumption of rationality* events and the *cautious rationality and common assumption of rationality* event can be defined inductively as follows:

$$\forall m \geq 1, \; R^{m+1} := R^m \bigcap (\prod_{i \in I} A_i(R^m_{-i}));$$
$$R^\infty := \bigcap_{m \in \mathbb{N}} R^m.$$

The behavioral implications of the first events correspond step-by-step to the iteratevely admissible strategy profiles. The second event is non-empty too and its behavioral implications coincide with the final set of the iterated admissibility procedure. These facts are summarized in the following characterization theorem.

Theorem 13 *For every $n \geq 0$, $S^n = \text{proj}_S R^n$. Moreover, $S^M = \text{proj}_S R^\infty$.*

Proof.

For every $n \leq M$, $i \in I$ and $s_i \in S_i^n$, take a $\mu_i^n(s_i) \in \Delta(S_{-i}^{n-1})$ such that $supp \mu_i^n(s_i) = S_{-i}^{n-1}$ and for every $s_i' \in S_i$, $\pi_i(s_i, \mu_i^n(s_i)) \geq \pi_i(s_i', \mu_i^n(s_i))$ (it exists by proposition 3). Moreover, for every $i \in I$ and $s_i \in S_i^M$, take a $\mu_i^{M+1}(s_i) \in \Delta(S_{-i}^M)$ such that $supp \mu_i^{M+1}(s_i) = S_{-i}^M$ and for every $s_i' \in S_i$, $\pi_i(s_i, \mu_i^{M+1}(s_i)) \geq \pi_i(s_i', \mu_i^{M+1}(s_i))$.

For every $k \leq M$, define the types $U_i^k := \bigcup_{s_i \in S_i^k} (s_i \times k)$ and set $U_i := \bigcup_{0 \leq k \leq M} U_i^k$.

[12] It would be interesting to check how different an axiomatic treatment of the definition here should be.

For every $i \in I$, define $h_i : U_i \to \Delta^{LPS}(S_{-i} \times U_{-i})$ with the following procedure:

- for every $s_i \in S_i$, take a $\lambda \in \Delta(S_{-i} \times U_{-i})$ such that $supp(marg_{S_{-i}}\lambda) \neq S_{-i}$ and let $h_i((s_i, 0)) := \lambda$;
- for every $0 < k < M$ and $s_i \in S_i^k$, take the $\lambda \in \Delta(S_{-i} \times U_{-i}^{k-1})$ such that for every $s_{-i} \in S_{-i}^{k-1}$, $\lambda[(s_{-i}, (s_{-i}, k-1))] = \mu_i^k(s_i)[s_{-i}]$ and let $h_i((s_i, k)) := (\lambda, h_i((s_i, k-1)))$;
- for every $s_i \in S_i^M$, take the $\lambda_2 \in \Delta(S_{-i} \times U_{-i}^{M-1})$ such that for every $s_{-i} \in S_{-i}^{M-1}$, $\lambda_2[(s_{-i}, (s_{-i}, M-1))] = \mu_i^M(s_i)[s_{-i}]$ and take the $\lambda_1 \in \Delta(S_{-i} \times U_{-i}^M)$ such that for every $s_{-i} \in S_{-i}^M$, $\lambda_1[(s_{-i}, (s_{-i}, M))] = \mu_i^{M+1}(s_i)[s_{-i}]$ and let $h_i((s_i, M)) := (\lambda_1, \lambda_2, h_i((s_i, M-1)))$.

For every $j \in I$ and $u_j \in U_j$, take the lexicographic hierarchy of beliefs $\boldsymbol{\delta}_j(u_j) = (\delta_1, ..., \delta_k)$ induced by $h_j(u_j)$ in the finite type space $(U_i, h_i)_{i \in I}$ and rename $\boldsymbol{\delta}_j(u_j)$ as u_j in C_j (see section 3). Now by the definition of g_j, it must be $g_j(u_j) = h_j(u_j)$ because in such case it is true that for every $l \leq k$ and for every $h \in \mathbb{N}$, $marg_{X_l^h} f_j(\delta_l) = \delta_l^h$ and by lemma 8 there is only one function satisfying this property. Moreover, $g_j(u_j)$ is mutually singular and $\boldsymbol{\delta}_j(u_j)$ is collectively coherent, so it survives all steps of the reduction of the type space and finally $u_j \in T_j$.

Define $m(s_i) := \max\{n \in \mathbb{N} : s_i \in S_i^n\}$. Clearly, for every $i \in I$ and every $s_i \in S_i^1$, $(s_i, (s_i, m(s_i))) \in R_i^1$. By induction, it is immediate to show that $g_i((s_i, m(s_i)))$ assumes $R_{-i}^{m(s_i)-1}, ..., R_{-i}^1$. So it holds that for every $n \leq M$, $S^n \subset proj_S R^n$.

Moreover, notice that for every $i \in I$ and $s_i \in S_i^M$, $g_i((s_i, M))$ assumes also R_{-i}^M, so that $(s_i, (s_i, M)) \in R_i^{M+1}$. But then by induction it is immediate to show that for every $n \in \mathbb{N}$, $g_i((s_i, M))$ assumes R_{-i}^n. So, $S^M \subseteq proj_S R^\infty$.

For the opposite inclusion, take as inductive hypothesis that $S^n \supseteq proj_S R^n$.

Setting $R^0 := S \times T$, it is trivially verified for $n = 0$.

Take any $\omega = (s, t) \in R^{n+1} \subseteq R^n$. Notice that s_i is a lexicographic best reply to the lexicographic conjecture $marg_{S_{-i}} g_i(t_i) = (\nu_1, ..., \nu_k)$, where $g_i(t_i)$ assumes R^n at some level $l \leq k$. By the inductive hypothesis, for every $i \in I$, $s_i \in S_i^n$. Hence it is enough to show that there exists a measure $\sigma_{-i} \in \Delta(S_{-i}^n)$, with $supp\sigma_{-i} = S_{-i}^n$, such that $\pi_i(s_i, \sigma_{-i}) \geq \pi_i(s_i', \sigma_{-i})$ for every $s_i' \in S_i^n$. Any measure $\nu := \alpha_1\nu_1 + ... + \alpha_l\nu_l$ such that the sum-1 weights $\alpha_1, ..., \alpha_l$ satisfy $\alpha_{n+1} \cdot l \cdot (\max_{s \in S} \pi_i(s) - \min_{s \in S} \pi_i(s)) < \alpha_n$ works.

Moreover, since $S^M \supseteq proj_S R^M \supseteq proj_S R^\infty$, it holds $S^M \supseteq proj_S R^\infty$. ■

5 Conclusions

Players are expected to play iteratevely admissible strategies when they are cautiously rational, assume opponents are cautiously rational, and so on, where being cautious, rational and assuming an event like opponents' cautious rationality must be carefully defined. A player is rational when she plays a lexicographic best reply to her lexicographic conjecture about opponents' strategies. A player is cautious when she forms the lexicographic conjecture by taking into consideration every opponents' strategy subprofile as possible to some extent. The definition of assumption allows players to consider in their conjectures also the possibility that the event does not occur, but assigning likelihood priority to the event and to all its possible behavioral implications.

BFK characterize iterated admissibility with rationality and assumption of rationality events, but incorporating in rationality an open-mindedness requirement which is stronger than the cautiousness requirement here: players must form conjectures that assign a priority level and a probability weight to (a neighborhood of) every state of world. As a consequence, in every complete and continuous type space, players cannot commonly assume this notion of open-minded rationality since the corrisponding event is empty. This reflects the computational burden required to players.

This impossibility has been eliminated here by weakening the requirement on players' conjectures. Players are allowed to form parsimonious conjectures, whose supports can also be constituted by a finite number of states of the world. But players always care to order and weigh all possible opponents' moves, the payoff-relevant objects. Assuming opponents' rationality, and so on, and associating to strategies higher-order beliefs allows players to put them in a meaningful likelihood order.

As a result, players are able to form conjectures that commonly assume this notion of cautious rationality, also in a rich type space that does not prevent them from coming up with any meaningful lexicographic hierarchy of beliefs about strategies. The existence of such canonical type space has been shown constructively.

The passage from open-mindedness to cautiousness has the further advantage of reducing the complexity of the analysis. The characterization does not depend on the topology of the type space,[13] which could then be costructed as a simple measure-theoretical object, like in [10] for the non-lexicographic case. The analysis could be further simplified by removing the mutual singularity requirement. It has to be noticed that players' lexicographic beliefs of any order are not required to be mutually singular. Hence, removing mutual singularity over states of the world as a whole would not change the interpretation of the epistemic characterization.

[13] The topological construction of the type space allows to claim the continuity of belief maps and compare the results with BFK's ones. However, the topology is a dispensable object for the characterization result.

Whether the constructed type space is universal or not has not been investigated yet. In the proof of the characterization theorem, a finite type space is mapped into the canonical type space and thus shown to be a belief-closed subset of the latter. If this could be done for any type space, the canonical type space would also be terminal. Hence, a universal type space would exist. On the other hand, it would be interesting to check the existence of a type space for LPS with finite joint support. It is reasonable to think that players will not introduce an infinite dimension to justify the likelihood order they want to give to a finite number of opponents' strategy subprofiles.

However, the epistemic model set up here for the characterization of iterated admissibility can be used for different scopes. For instance, I conjecture that by simply removing the marginal support requirement from the definition of assumption, the same events would characterize the elimination of weakly dominated strategies followed by many rounds of elimination of strongly dominated strategies (which is the appropriate solution concept also under the hypotheses of [7] and [4]).

Finally, the characterization (as in BFK) relies also on strategy-type pairs that are not cautiously rational not because the strategy is not a lexicographic best reply to the conjecture, but because the conjecture does not respect the cautiousness requirement (and the same applies in BFK with the open-mindedness one). This is necessary to assume cautious rationality and form at some level a fully mixed conjecture whenever an opponent has a dominant strategy, which is always rational, and a dominated strategy, which is always irrational. If types without full marginal support on strategies could be put out of the picture, LPS would coincide with Conditional Probability Systems a la Myerson [13] and a unified framework for the epistemic analysis of solution concepts in static and dynamic games could be developed. To avoid the use of such incautious conjectures it is enough to allow that an assumed event, rationality in this case, be given a nonnull probability after the level at which it stops having probability one. This requires to rethink the interpretation of assumption as deeming an event infinitely more likely than the complementary.

References

[1] Asheim, G., Dufwenberg, M., "Admissibility and common belief", *Games and Economic Behavior*, **42**, 2003, 208-234.

[2] Battigalli, P., and M. Siniscalchi, "Strong Belief and Forward Induction Reasoning," *Journal of Economic Theory*, **106**, 2002, 356-391.

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

[4] Borgers, T., "Weak dominance and approximate common knowledge", *Journal of Economic Theory*, **64**, 1994, 265-276.

[5] Brandenburger, A., Dekel, E., "Hierarchies of beliefs and common knowledge", *Journal of Economic Theory*, **59**, 1993, 189-198.

[6] Brandenburger, A., Friedenberg, A., Keisler, J., "Admissibility in games", *Econometrica*, **76**, 2008, 307-352.

[7] Dekel, E. and Fudenberg, D., "Rational behavior with payoff uncertainty", *Journal of Economic Theory*, **52**, 1990, 243-267.

[8] Dellacherie, C., Meyer, P., "Probabilities and potential", *Math. Stud.,* **29**, 1978.

[9] Heifetz, A., Meier, M., Schipper, B., "Comprehensive rationalizability", working paper.

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

[11] Kechris, A., "Classical descriptive set theory", Springer, 1995.

[12] Keisler, J., Lee, B., "Common assumption of rationality", working paper.

[13] Myerson, R., "Multistage games with communication", *Econometrica*, **54**, 1986, 323-358.

[14] Pearce, D., "Rational Strategic Behavior and the Problem of Perfection," *Econometrica*, **52**, 1984, 1029-1050.

[15] Stahl, D., "Lexicographic rationalizability and iterated admissibility", *Economic Letters*, **47**, 1995, 155-159.