The Pennsylvania State University

The Graduate School

College of the Liberal Arts

ESSAYS IN GAME THEORY

A Dissertation in
Economics
by
Byung Soo Lee

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2011

The dissertation of Byung Soo Lee was reviewed and approved* by the following:

Edward Green
Professor of Economics
Dissertation Adviser, Chair of Committee

Kalyan Chatterjee
Professor of Economics

Tymofiy Mylovanov
Assistant Professor of Economics

Stephen Simpson
Professor of Mathematics

Vijay Krishna
Distinguished Professor of Economics
Director of Graduate Studies in Economics

*Signatures are on file in the Graduate School.

Abstract

This dissertation consists of two chapters.

Chapter 1: Rationalizing Payoff-Dominant Outcomes

I modify two-player simultaneous-move games with a unique payoff-dominant strategy profile by allowing each player to publicly discard any of her original strategic options in turn before play begins. In this setting, I show that extensive-form rationalizable (EFR) profiles, as defined in Pearce (1984), have payoff-dominant outcomes. Furthermore, a strategy profile in which no player makes any commitments is EFR. Thus, the model is interpreted as one of payoff-dominant focal-point formation via forward induction, which is captured by EFR.

The result is analogous to that in Ben-Porath and Dekel's (1992) model of money-burning games. That the order of beliefs required to obtain payoff-dominant outcomes is uniformly bounded at three across all games is an advantage of the model in this paper over the money-burning approach, which may require the use of arbitrarily high orders of beliefs to obtain the same result. A limitation of this paper, which is also shared by the money-burning model, is that the prediction of payoff-dominant play is not robust to the addition of players.

Chapter 2: Common Assumption of Rationality

In this paper, I provide an epistemic characterization of iterated elimination of inadmissible—that is, weakly dominated—strategies. Along the way, I show that for each finite game there exists a large class of complete lexicographic type structures that admit rationality and common assumption of rationality (RCAR), and provide a versatile methodology for constructing them. This answers a question posed in Brandenburger, Friedenberg, and Keisler (2008) (BFK).

Solving a long-standing puzzle of Samuelson (1992), BFK gave an epistemic characterization of finite rounds of elimination of inadmissible strategies. However, they also gave a negative result showing that RCAR is impossible in a complete lexicographic type structure with continuous mappings. This seemed to cast doubt on the possibility of RCAR as a natural candidate for the desired epistemic condition, but our results revive this possibility. I examine the differences between our positive result and their negative result in isolation, and arrive at the surprising conclusion that a lexicographic type structure contains more information than previously thought.

Contents

List of Figures

Acknowledgments

I am grateful for the generous support that I have received from my adviser, Professor Edward J. Green. Ed freely gave his time to discuss my ideas, read my work, and improve my understanding of economics. Furthermore, he managed to inject a measure of intellectual discipline into my approach to scholarship. However, his support was not limited to the realm of academics. He also took a deep interest in my personal problems and provided both moral and practical assistance, without which I would not have finished my dissertation.

I am also grateful for the suggestions that I received from Professors Kalyan Chatterjee, Tymofiy Mylovanov, and Stephen Simpson while I was conducting my dissertation research.

I am deeply indebted to Professors Adam Brandenburger and Amanda Friedenberg for their helpful and detailed comments on the second chapter. I would also like to thank Professor H. Jerome Keisler, with whom a joint research project extending the results of the second chapter is in progress.

Finally, I would like to thank my parents for their unwavering belief in me. Their love is the foundation of everything that I do.

Chapter 1

Rationalizing Payoff-Dominant Outcomes

1.1 Introduction

An outcome s of a game G is said to payoff-dominate another s' if all players strictly prefer s to s'. In a complete information game, outcomes correspond to strategy profiles. If a strategy profile s^* payoff-dominates all other strategy profiles, it is the game's unique payoff-dominant strategy profile. If such s^* exists in a simultaneous-play game G, then intuition strongly suggests that it should be played. Unfortunately, an interpretation of the normal-form representation as an accurate and exhaustive description of all strategically relevant aspects of the situation may preclude any *logical* justification for unique predictions. Without restrictions on their beliefs about others' play, Bayesian rational players[1] may play any undominated strategy.

[1] i.e., those who maximize subjective expected utility.

Assuming common knowledge of rationality[2] further reduces the set of logically possible predictions to the set of rationalizable strategies (Bernheim, 1984; Pearce, 1984). However, rationalizability does not always yield unique predictions in games with uniquely payoff-dominant strategy profiles. If our intuition is to be believed, it must be that our model of the strategic situation is inaccurate or incomplete.

In this paper, we study a game with a unique payoff-dominant strategy profile from the perspective of extensive-form rationalizability (Pearce, 1984). Extensive-form rationalizability (henceforth EFR) can be characterized by what is called the "best rationalization principle" in Battigalli and Siniscalchi (2002). The best rationalization principle is the assumption that players hold the highest level of common belief of rationality that is logically tenable based on evidence observed. We show that in two-player games with uniquely payoff-dominant strategy profiles, if each player is given a turn to publicly discard any of his original strategic options, then EFR makes unique predictions about payoffs.

Furthermore, the set of EFR predictions includes the strategy profile in which no player makes any commitments. Thus the players may reach a subgame that is strategically identical to the original game. The descriptions of the larger game and the previous history allow the players to infer each others' play in that subgame. We take the view that, in many situations of interest, the original game is an accurate but incomplete description of a reality in which the players have neglected to take advantage of their commitment options. Thus, predictions of payoff-dominant outcomes can be supported even by a non-equilibrium theory if the reality we wish to model allows players to make public commitments.

Schelling (1960) made an influential argument that, in the absence of communi-

[2]The relationship between common knowledge of rationality and rationalizability was established by Aumann (1987); Brandenburger and Dekel (1987a); Tan and Werlang (1988).

cation, players are drawn toward playing an equilibrium that is the “focal point” of their collective expectations. The model in this paper fits the context of Schelling’s argument since, in the subgame that is identical to the original game, no communication has been made. However, the potential for communication in the form of commitments creates a collective expectation of payoff-dominant play via logical deduction. This type of communication of intent via non-communication has been discussed under the umbrella of the term “forward induction” in literature (Kohlberg and Mertens, 1986). While Schelling gave mostly psychological reasons for the formation of focal points, he also stresses the importance of common knowledge of the pervasiveness of those psychological reasons. Therefore, in some sense, common knowledge of previously rejected options justifying payoff-dominant play may be considered a form of focal point formation.

My model of the larger game is closely related to that of Ben-Porath and Dekel (1992), in which a class of two-player games are extended in such a way that one of the players has the option to burn various quantities of utility, including zero, before play begins.[3] As in this paper, Ben-Porath and Dekel (1992) obtain payoff-dominant play in the subgame that is identical to the original game.[4] Ben-Porath and Dekel use iterated admissibility, a decision-theoretic solution concept that coincides with EFR in generic games (see Brandenburger and Friedenberg, 2007).

In Section 1.2, the motivating example is discussed. Section 1.3 sets up the model and a formalism for EFR. Section 1.4 discusses the main results. Section 1.5

[3]Van Damme (1989) independently introduced “money-burning”.

[4]Van Damme and Hurkens (1996) construct an augmentation of the original game that allows simultaneous commitments by either player that is observed before play begins. Again, if neither player makes a commitment, then a subgame identical to the original game is reached. The augmentation in van Damme and Hurkens (1996) does not guarantee payoff-dominant play, which illustrates that considerations about the strategic environment from which the normal-form representation abstracts may be crucial to whether or not...

concludes. All proofs are in Appendix A.1.

1.2 Motivation

Ann, Bob	TGR	TBR
TGR	3,3	0,0
TBR	0,0	1,1

Figure 1.1: Dinner Game

Consider the following scenario. There are exactly two restaurants in town: "The Good Restaurant" (TGR) and "The Bad Restaurant" (TBR). Ann and Bob prefer eating together to eating separately, but if they could eat together, they would prefer to eat at The Good Restaurant instead of The Bad Restaurant (see Figure 1.1). One must take the Green Train to get to TGR and the Blue Train to get to TBR. The train station sells three types of tickets: a green ticket, a blue ticket, and a blue-green ticket. All tickets have the same price. Ann and Bob meet at the train station at lunch time and line up to buy their tickets. Ann goes first, then Bob. For the sake of clarity, let us assume that the ticket counter closes after lunch. Each person can observe the ticket color of the other. They will get off at work at different times in the evening and will then head to one of the two restaurants for dinner.

If Ann believes that Bob is rational, she also believes that if she purchases a green ticket (committing to TGR), Bob will go to TGR as well. If Ann purchases a blue-green ticket and Bob believes that Ann is rational, then Bob also believes that if he purchases a green ticket (committing to TGR), Ann will go to TGR as well since her blue-green ticket allows her to board the Green Train. If both Ann and Bob purchase blue-green tickets and Ann believes that Bob believes that Ann is rational, then Ann must believe that Bob plans to board the Green Train and thus will board the Green Train herself. If both Ann and Bob purchase blue-green tickets and Bob believes that Ann believes that Bob believes that Ann is rational, then Bob must believe that Ann plans to board the Green Train and thus will board the Green Train as well. Thus, if both players purchase blue-green tickets, both players will head to The Good Restaurant.

Note that the ticket-purchasing process, even if it results in no commitments by either side, can cause the formation of a focal point of higher order beliefs for Bob and Ann. Moreover, the focal point gives Ann and Bob certain coordination at The Good Restaurant, which is a very desirable property. Our paper attempts to generalize this logic to a larger class of games that have payoff-dominant outcomes.

1.3 The Setup

Consider an n-player normal-form game $G = \langle (D_i, u_i)_{i=1}^n \rangle$, where D_i is player i's strategy space and $u_i : \prod_{i=1}^n D_i \to \mathbb{R}$ is player i's utility mapping. We use D to denote $\prod_{i=1}^n D_i$. We assume throughout that the game G has a payoff-dominant cylinder.

Definition 1.3.1. *A game $G = \langle (D_i, u_i)_{i=1}^n \rangle$ has a payoff-dominant cylinder if there exists a subset B of D that satisfies the following:*

$$\left[B = \prod_{i=1}^n B_i \right] \wedge \\ \forall i \left[(B_i \in \mathscr{C}_i) \wedge \left(\min_{d \in B} u_i(d) > \max_{d \in D \setminus B} u_i(d) \right) \right], \text{ where } \mathscr{C}_i = 2^{D_i} \setminus \{\varnothing\}.$$

The set B is called a payoff-dominant cylinder.

Now we construct an extensive-form game $\Gamma(G) = \langle (S_i, u_i)_{i=1}^n \rangle$ that augments G.

We begin by defining the sets of possible k-length public histories and the strategies of the augmented game.

Definition 1.3.2 (public histories of length k)**.**

$$\begin{aligned} H(0) &= \{\varnothing\} \\ H(k) &= \left\{ C = \prod_{i=1}^{k} C_i \,\middle|\, \forall i \, [C_i \in \mathscr{C}_i] \right\}, \textit{ for } k = 1, \ldots, n \\ H(n+1) &= \{ (C, d) | \, d \in C \in H(n) \} \\ \mathscr{H} &= \bigcup_{k=0}^{n} H(k). \end{aligned}$$

To elaborate, $H(n+1)$ *is the set of all terminal histories and* $\mathscr{H}$ *is the set of all non-terminal histories.*

Definition 1.3.3 (strategy of the augmented game)**.** *For any player* i*, his strategy in* $\Gamma(G)$ *is a mapping* $s_i : H(i-1) \cup H(n) \to \mathscr{C}_i \cup D_i$ *that satisfies the following:*

$$[s_i(H(i-1)) \in \mathscr{C}_i] \wedge \left[C = \prod_{j=1}^{n} C_j \in H(n) \implies s_i(C) \in C_i \right].$$

This definition says that each player i takes a turn to make a public and enforceable commitment to never play $d_i \notin C_i$. Once every player has taken a turn, the normal-form game $G(C) = \langle (C_i, u_i)_{i=1}^n \rangle$, which modifies G to reflect the commitments made, is played. The definition of strategies in the augmented game $\Gamma(G)$ satisfies the requirements for $\Gamma(G)$ to be a multi-stage game with observed actions, as defined in Battigalli and Siniscalchi (2002).

Given a strategy profile $s = (s_i)_{i=1}^n$ of the game $\Gamma(G)$, we may define the corresponding outcome of the game $d(s) = (d_i(s))_{i=1}^n$ by using $\hat{h}(s, k)$, which is the

k-length history induced by profile s:

$$\begin{aligned}
\hat{h}(s,1) &= s_1(\varnothing) \\
\hat{h}(s,k) &= \hat{h}(s,k-1) \times s_k(\hat{h}(s,k-1)) \text{ where } k = 1,\dots n \\
\hat{h}(s,n+1) &= \left(\hat{h}(s,n), s(\hat{h}(s,n))\right) \\
d_i(s) &= s_i(\hat{h}(s,n)) \text{ for } i = 1,\dots n.
\end{aligned}$$

Since $d(s) \in D$, there is a natural way to extend the domain of $u = (u_i)_{i=1}^n$ to S. For $s \in S$, let

$$u_i(s) = u_i(d(s)) \text{ for } i = 1,\dots n.$$

We provide the definition of k-rationalizable strategies and sequentially rational strategies as given in Battigalli and Siniscalchi (2002) below.

Definition 1.3.4 (conditional probability system)**.** *A conditional probability system (or CPS) for player i is a collection of conditional beliefs $\mu^i = (\mu^i(\cdot|h))_{h\in\mathscr{H}} \in \prod_{h\in H} \Delta(S_{-i}(h))$ such that for all $\bar{s}_{-i} \in S_{-i}$, $h', h'' \in \mathscr{H}$, if h' is a prefix of h'' (i.e., $S_{-i}(h'') \subseteq S_{-i}(h')$) then*

$$\mu^i(\bar{s}_{-i}|h') = \mu^i(\bar{s}_{-i}|h'') \left(\sum_{s_{-i}\in S_{-i}(h'')} \mu^i(s_{-i}|h') \right)$$

The set of CPSs for player i is denoted by CPS_i.

Definition 1.3.5 (k-rationalizability and EFR)**.** *The k-rationalizable strategies for* *player i, Σ_i^k, is defined by the following recursive equivalence relation.*

$$\left(\tilde{s}_i \in \Sigma_i^k\right) \equiv \left(\tilde{s}_i \in \Sigma_i^{k-1}\right) \wedge \exists \mu^i \Bigg[\left(\mu^i \in \mathrm{CPS}_i\right) \wedge \left(\tilde{s}_i \in r_i\left(\mu^i\right)\right) \wedge \forall h \Big[\left((h \in \mathscr{H}) \wedge \left(\Sigma_{-i}^{k-1} \cap S_{-i}(h) \neq \varnothing\right)\right) \implies \left(\mu^i(\Sigma_{-i}^{k-1}|h) = 1\right) \Big] \Bigg]$$

A strategy s_i is extensive form rationalizable if it is in Σ_i^k for all k. If so, we write $s_i \in \Sigma_i^\infty$.

Definition 1.3.6 (sequential rationality)**.** *The set, $r_i(\mu^i)$, of player i's strategies that satisfy sequential rationality with respect to the conditional probability system μ^i:*

$$\tilde{s}_i \in r_i\left(\mu^i\right) \equiv \forall h \left[\left((h \in \mathscr{H}) \wedge (\tilde{s}_i \in S_i(h))\right) \implies \tilde{s}_i \in \operatorname*{argmax}_{s_i \in S_i(h)} U_i(s_i, \mu^i(\cdot|h)) \right]$$

where

$$U_i(s_i, \mu^i(\cdot|h)) = \sum_{s_{-i} \in S_{-i}(h)} u_i(s_i, s_{-i}) \mu^i(s_{-i}|h)$$

1.4 Results

1.4.1 Two-player Games

Theorem 1.4.1. *If G is a two-player game and the assumptions of the general setup are satisfied then $\forall s\,[s \in \Sigma^\infty \implies d(s) \in B]$.*

That is, if all players play EFR strategies then they will never do worse than the worst outcome in B. Theorem 1.4.2 tells us that the subgame where the players play the original normal form game can be reached when all players play EFR strategies of the augmented game. Theorem 1.4.2 is necessary to validate our interpretation that a simultaneous game is played after all players have publicly restricted their own strategic options.

Theorem 1.4.2. *Suppose G is a two-player game and that the assumptions of the general setup are satisfied. Furthermore, let $s \in \Sigma^\infty$ and $B = \{(b_1, b_2)\}$. Let $C = C_1 \times C_2$, $W = W_1 \times W_2$, $s' \in S$ and*

$$\begin{array}{cccc} s_1(\varnothing) = C_1 & s_2(C_1) = C_2 & s_1(C) = d_1 & s_2(C) = d_2 \\ s'_1(\varnothing) = W_1 & s'_2(C_1) = W_2 & s'_1(W) = d_1 & s'_2(W) = d_2 \\ B_1 \subsetneq W_1 & B_2 \subsetneq W_2. & & \end{array}$$

Then $s' \in \Sigma^\infty$.

1.4.2 n-player Games ($n \geq 3$)

In games with 3 or more players, EFR fails to guarantee that the payoff-dominant outcome is reached. The intuition may be illustrated by the game in Figure 1.2 played by Bob, Ann, and Carl. Payoffs for Ann and Carl are suppressed because the players have identical utility functions in this game. After in-L or in-R, Bob is the row player and Ann is the column player. Suppose that Bob plays "in". Ann then infers that Bob expects a payoff greater than or equal to 2. However, there are two paths to payoffs greater than 2 in this game after "in". Furthermore, Carl's choice between L and R is not informative about the actions of Ann and Bob. Therefore Ann cannot deduce whether Bob expects a payoff of 2 after L, R, or both.

The game without Carl, shown in Figure 1.3, does not suffer from such ambiguities

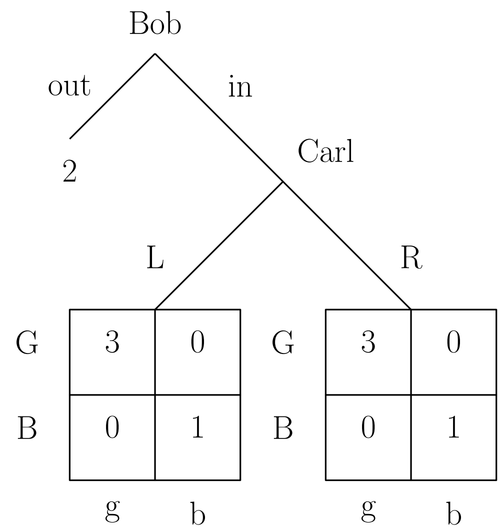

Figure 1.2: Carl makes a move that is irrelevant to him.

because, after “in”, there is exactly one path to a payoff greater than or equal to 2. In our augmented game, players may take many paths to the payoff-dominant outcome since one may commit to playing an option of the original game that is the projection of a payoff-dominant strategy profile of the original game or one may commit to choosing from any subset that contains such a strategy. Therefore, ambiguities about the path expected by other players abound in our augmented game.

For an example of an EFR profile that does not guarantee a payoff-dominant outcome in our augmentation scheme, we will consider the 3-player pure coordination game shown in Figure 1.4. Let $B = \{b\}$ and $W = \{w\}$.

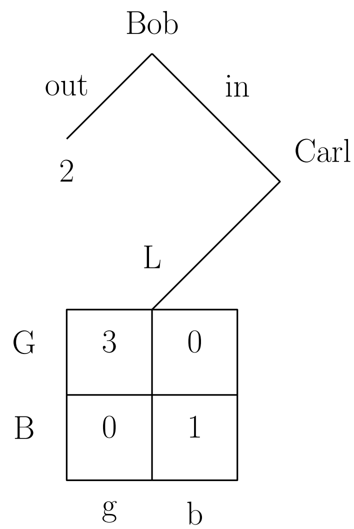

Figure 1.3: Carl is removed.

	b_1		w_1	
	b_2	w_2	b_2	w_2
b_3	3	0	0	0
w_3	0	0	0	1

Figure 1.4: 3-player example

Consider the following strategies:

$$s_1(h) = \begin{cases} b_1 & \text{if } h \subseteq D \wedge (h \succeq B_1 \vee h \succeq D_1 \times B_2) \\ w_1 & \text{if } h \subseteq D \wedge (h \succeq W_1 \vee h \succeq D_1 \times W_2 \vee h \succeq D_{\leq 2} \times W_3) \\ w_1 & \text{if } h = D \\ b_1 & \text{if } h = D_{<3} \times B_3 \\ D_1 & \text{if } h = \varnothing \end{cases}$$

$$s_1'(h) = \begin{cases} b_1 & \text{if } h \subseteq D \wedge (h = D \vee h \succeq B_1 \vee h \succeq D_1 \times B_2) \\ w_1 & \text{if } h \subseteq D \wedge (h \succeq W_1 \vee h \succeq D_1 \times W_2 \vee h \succeq D_{\leq 2} \times W_3) \\ w_1 & \text{if } h = D_{<3} \times B_3 \\ D_1 & \text{if } h = \varnothing \end{cases}$$

$$s_2(h) = \begin{cases} b_2 & \text{if } h \subseteq D \wedge (h \succeq B_1 \vee h \succeq D_1 \times B_2) \\ w_2 & \text{if } h \subseteq D \wedge (h \succeq W_1 \vee h \succeq D_1 \times W_2 \vee h \succeq D_{\leq 2} \times W_3) \\ w_2 & \text{if } h = D \\ b_2 & \text{if } h = D_{<3} \times B_3 \\ D_2 & \text{if } h = B_1 \vee h = D_1 \\ W_2 & \text{if } h = W_1 \end{cases}$$

$$s_2'(h) = \begin{cases} b_2 & \text{if } h \subseteq D \wedge (h = D \vee h \succeq B_1 \vee h \succeq D_1 \times B_2) \\ w_2 & \text{if } h \subseteq D \wedge (h \succeq W_1 \vee h \succeq D_1 \times W_2 \vee h \succeq D_{\leq 2} \times W_3) \\ w_2 & \text{if } h = D_{<3} \times B_3 \\ D_2 & \text{if } h = B_1 \vee h = D_1 \\ W_2 & \text{if } h = W_1 \end{cases}$$

$$s_3(h) = \begin{cases} b_3 & \text{if } h \subseteq D \wedge (h \succeq B_1 \vee h \succeq D_1 \times B_2) \\ w_3 & \text{if } h \subseteq D \wedge (h \succeq W_1 \vee h \succeq D_1 \times W_2 \vee h \succeq D_{\leq 2} \times W_3) \\ b_3 & \text{if } h = D_{<3} \times B_3 \\ b_3 & \text{if } h = D \\ B_3 & \text{if } h \subseteq S_{<3} \wedge (h_1 \cap B_1 \neq \varnothing \wedge h_2 \cap B_2 \neq \varnothing) \\ W_3 & \text{if } h \subseteq S_{<3} \wedge (h_1 \cap B_1 = \varnothing \vee h_2 \cap B_2 = \varnothing) \end{cases}$$

$$s'_3(h) = \begin{cases} b_3 & \text{if } h \subseteq D \wedge (h \succeq B_1 \vee h \succeq D_1 \times B_2) \\ w_3 & \text{if } h \subseteq D \wedge (h \succeq W_1 \vee h \succeq D_1 \times W_2 \vee h \succeq D_{\leq 2} \times W_3) \\ b_3 & \text{if } h = D_{<3} \times B_3 \\ b_3 & \text{if } h = D \\ D_3 & \text{if } h \subseteq S_{<3} \wedge (h_1 \cap B_1 \neq \varnothing \wedge h_2 \cap B_2 \neq \varnothing) \\ W_3 & \text{if } h \subseteq S_{<3} \wedge (h_1 \cap B_1 = \varnothing \vee h_2 \cap B_2 = \varnothing) \end{cases}$$

The following facts are now easily verified:

$$\forall i[s_i \in S_i(D_{\leq 2} \times B_3) \wedge s'_i \in S_i(D)]$$

$$d(s') = b \wedge d(s) = b$$

$$\forall i \left[\hat{h}((s'_i, s'_{-i}), n+1) = (D, b)\right] \wedge \forall i \left[\hat{h}((s_i, s_{-i}), n+1) = (D_{\leq 2} \times B_3, b)\right]$$

$$\forall i \left[i = 3 \implies \hat{h}((s'_i, s_{-i}), n+1) = (D, (b_i, w_{-i})) \right]$$
$$\forall i \left[i \neq 3 \implies \hat{h}((s'_i, s_{-i}), n+1) = (D_{\leq 2} \times B_3, (w_i, b_{-i})) \right]$$
$$\forall i \left[i = 3 \implies \hat{h}((s_i, s'_{-i}), n+1) = (D_{\leq 2} \times B_3, (b_i, w_{-i})) \right]$$
$$\forall i \left[i \neq 3 \implies \hat{h}((s_i, s'_{-i}), n+1) = (D, (w_i, b_{-i})) \right].$$

We may choose CPSs $(\mu^i)_i$ and $(\nu^i)_i$ that satisfy the following properties:

$$\forall i \left[i \neq 3 \implies \left[\mu^i(s'_{-i}|\varnothing) = \mu^i(s'_{-i}|D_1) = \mu^i(s'_{-i}|D_{\leq 2}) = 1 \wedge \mu^i(s'_{-i}|D_{\leq 2} \times B_3) = 0 \right] \right]$$

$$\forall i \left[i \neq 3 \implies \left[\nu^i(s_{-i}|\varnothing) = \nu^i(s_{-i}|D_1) = \nu^i(s_{-i}|D_{\leq 2}) = 1 \wedge \nu^i(s_{-i}|D) = 0 \right] \right]$$

$$\left[\mu^3(s'_{-3}|\varnothing) = \mu^3(s'_{-3}|D_1) = \mu^3(s'_{-3}|D_{\leq 2}) = 1 \right]$$
$$\left[\nu^3(s_{-3}|\varnothing) = \nu^3(s_{-3}|D_1) = \nu^3(s_{-3}|D_{\leq 2}) = 1 \right]$$

Again, it is simple to verify that for each i, s_i is sequentially rational with respect to CPS ν^i ($s_i \in r_i(\nu^i)$) and s'_i is sequentially rational with respect to CPS μ^i ($s'_i \in r_i(\mu^i)$). For each i, since s_i is sequentially rational with respect to a CPS μ^i that places probability 1 on s_{-i} for as long as logically tenable, s_i cannot be eliminated in any round unless s_{-i} is eliminated in the immediately preceding round. Similar logic also guarantees that s'_1, s'_2, s'_3 remain EFR strategies. Therefore, profiles like (s'_3, s_{-3}), which yield less than the payoff-dominant outcome, must remain in Σ^∞.

1.4.3 Iterated Admissibility

Restricting ourselves to profiles that survive iterated admissibility (IA profiles) is an appropriate alternative to EFR if we choose to incorporate avoidance of weakly dominated strategies as part of our definition of rationality. Theorem 1.4.7 shows that all IA profiles have payoff-dominant outcomes in our augmented game. Unfortunately, we cannot obtain an analogue of Theorem 1.4.2, thus our desired interpretation of the original game as a subgame that is reached in the augmented game is invalid if we assume that only IA profiles are played. However, the proof of Theorem 1.4.7 suggests that if there exists a payoff-dominant path in the game tree such that there is perfect information at each and every node along the path, then all IA profiles have payoff-dominant outcomes. Such a conjecture remains to be proven, but we imagine that its proof would parallel that of Theorem 1.4.7 to a large extent.

Definition 1.4.3 (iteratively admissible strategies)**.** *Let* $Y = \prod_{i=1}^{n} Y_i$*.*

$$\mathrm{BR}_i(Y_i, s_{-i}) = \left\{ s_i \,\middle|\, s_i \in Y_i \wedge u_i(s_i, s_{-i}) = \max_{s'_i \in Y_i} u_i(s'_i, s_{-i}) \right\}$$

Set of strategies in Y_i *that are weakly dominated against* Y_{-i}*:*

$$\begin{aligned}&\mathrm{WDD}_i(Y_i, Y_{-i})\\ &\quad = \left\{ s_i \,\middle|\, s_i \in Y_i \wedge \exists s'_i \left[s'_i \in Y_i \wedge \{s_i\} \neq \{s'_i\} = \bigcap_{s_{-i} \in Y_{-i}} \mathrm{BR}_i(\{s_i, s'_i\}, s_{-i}) \right] \right\}\end{aligned}$$

$$
\begin{aligned}
A_i(Y_i, Y_{-i}) &= Y_i \setminus WDD_i(Y_i, Y_{-i}) \\
A(Y) &= \prod_{i=1}^{n} A_i(Y) \\
A^{n+1}(S) &= A(A^n(S)) && \textit{Profiles that survive } n+1 \textit{ rounds} \\
A^{\infty}(S) &= \bigcap_{k \in \mathbb{N}} A^k(S) && \textit{IA profiles}
\end{aligned}
$$

Remark 1.4.4. *For all n, $A^{n+1}(S) \subseteq A^n(S)$ and $A^{\infty}(S)$ is nonempty.*

Lemma 1.4.5 (strategic recombination lemma for admissibility)**.** *For any $m \in \mathbb{N}$, any history $\hat{h}$, and any player k, if $S_k(\hat{h}) = S_k$, then the following holds for all $s'_k, s''_k \in A^m_k(Y)$:*

$$
s'_k(\hat{h}) \neq s''_k(\hat{h}) \implies \exists s'''_k \in A^m_k(Y) \text{ s.t. } s'''_k(h) = \begin{cases} s'_k(h) & \text{if } h \succeq \hat{h} \\ s''_k(h) & \text{otherwise} \end{cases}
$$

In all procedures involving iterated elimination, in order to prove that a strategy is eliminated, we must show that another strategy that eliminates it must exist. If we desire to show that a strategy is not eliminated by another, then we must show that there exists a strategy profile against which it does strictly better than the other. Therefore, a result like Lemma 1.4.5, which is used to prove Theorem 1.4.7, is useful because it may provide the existence of sufficient diversity of strategies to either permit or deny the elimination of a strategy. The condition $S_k(\hat{h}) = S_k$ requires that player k makes his first move at history h. Consider the game in Figure 1.5. It is the pure strategy reduced normal form (or PRNF) of the modified outside option game in Figure 1.2:

After the first round, only “inBB” is eliminated. If admissibility implied strategic

If Carl plays L:

Bob/Ann	gg	gb	bg	bb
inGG	3	3	0	0
inGB	3	3	0	0
inBG	0	0	1	1
inBB	0	0	1	1
out	2	2	2	2

If Carl plays R:

Bob/Ann	gg	gb	bg	bb
inGG	3	0	3	0
inGB	0	1	0	1
inBG	3	0	3	0
inBB	0	1	0	1
out	2	2	2	2

Figure 1.5: PRNF of game in Figure 1.2

recombination at histories where a player makes his second move, that "inBG" and "inGB" remain should imply that "inBB" remains as well; this is clearly not the case. By ignoring the cases when recombination is possible at histories where a player makes his second move, the result becomes a simpler one to prove.

Corollary 1.4.6 (corollary of Lemma 1.4.5)**.** *For any* $m \in \mathbb{N}$*, any history* $\hat{h}$*, and any player* k*, if* $S_k(\hat{h}) = S_k$ *and* $s'_k, s''_k \in A^\infty_k(Y)$ *then:*

$$s'_k(\hat{h}) \neq s''_k(\hat{h}) \implies \exists s'''_k \in A^\infty_k(Y) \text{ s.t. } s'''_k(h) = \begin{cases} s'_k(h) & \text{if } h \succeq \hat{h} \\ s''_k(h) & \text{otherwise} \end{cases}$$

Proof of Corollary 1.4.6. Since $\exists m$ such that $A^m(S) = A^{m+1}(S) = A^\infty(S)$, the result follows from Lemma 1.4.5. □

Theorem 1.4.7. *Suppose* G *is an* n*-player game and let* $B = \{(b_i)_{i=1}^n\}$*. If* $s \in A^\infty(S)$ *then* $d(s) \in B$*.*

1.5 Conclusion

However, as will be apparent in the proof of Theorem 1.4.1, in the case of 2 players, we require only 3 rounds of strategy elimination via the EFR procedure to obtain that result. If a money-burning augmentation of the original game in the style of Ben-Porath and Dekel (1992) is used, then an arbitrarily large number of rounds eliminating weakly dominated strategies can be needed to arrive at the same result. If we think of each round of an iterative procedure as an epistemic requirement on higher order beliefs than implied by the previous round, we may conclude that our result has the benefit of involving fewer (and lower) hierarchies of interactive beliefs.

Our result on iterated admissibility in the augmented game, while not amenable to our desired interpretation of payoff-dominant play in games, is nevertheless interesting in that it suggests some useful and possibly interesting properties preserved by the admissibility operator $A(\cdot)$. In that same vein, the extension of Theorem 1.4.7 to more general situations suggested in Section 1.4.3 may be worth pursuing, as would results similar in spirit to Lemma 1.4.5 that relate actions at information sets to admissibility.

Chapter 2

Common Assumption of Rationality

2.1 Introduction

Analysis of games typically begins under the premise that all players are rational. Furthermore, it is often supposed, at least implicitly, that the rationality of the players is *common knowledge* in the sense of Aumann (1976)—that is, all players know it, all players know that all players know it, and so on. It is then natural to ask which strategic choices are consistent with common knowledge of rationality (CKR).

Bernheim (1984) and Pearce (1984) gave an influential response to this question in which they argued that their notion of *rationalizability* exactly captures the implications of CKR on behavior. The rationalizable set is essentially the iteratively undominated (IU) set—that is, the set of strategy profiles surviving iterated elimination of strongly dominated strategies—with the added virtue of being defined in a way that more starkly emphasizes its intuitive connections to CKR.[1]

Bernheim (1984) and Pearce (1984) motivated their analysis as an extension of Savage's (1954) Bayesian decision theory in which rational actors maximize subjective expected utility (SEU) subject to probabilistic beliefs about the states of the world. Therefore, in these and subsequent papers, CKR is often used interchangeably with *rationality and common belief of rationality* (RCBR), an analogous concept that is better suited for use in Bayesian settings.[2]

More formal analyses followed in Brandenburger and Dekel (1987a) and Tan and Werlang (1988), who showed that RCBR is an epistemic condition that characterizes the IU set. In other words, RCBR implies that IU strategies are played and every IU strategy can be played in some state where RCBR holds. A key fact underpinning this relationship is that SEU maximization characterizes avoidance of strongly dominated strategies. However, it is prima facie reasonable that rationality should incorporate an admissibility requirement—that is, avoidance of *weakly* dominated strategies. A long tradition in statistical decision theory, going as far back as Wald (1939), has advocated admissibility as a minimal criterion of rationality.[3]

In light of the preceding facts, an intuitively appealing conjecture is that *iteratively admissible* (IA) strategies—that is, strategies surviving iterated elimination of weakly dominated strategies—are characterized by RCBR if rationality incorpo-

[1]When we refer to rationalizability in this paper, we will mean *correlated* rationalizability, which omits the independence assumptions of the original definition. The correlated rationalizable set is exactly the IU set.

[2]An event is commonly believed if all players are certain of it, all players are certain that all players are certain of it, and so on, where certainty is understood to mean belief with probability 1. In the literature, common belief is also called common certainty, common belief with probability 1, and common 1-belief. See Brandenburger and Dekel (1987b); Monderer and Samet (1989).

[3]Von Neumann and Morgenstern (1944) justify the requirement from a staunchly objectivist point of view on probability while prefacing the development of their theory of two-person zero-sum games. Furthermore, later surveys by Arrow (1951) and Luce and Raiffa (1957) are uniform in their rejection of inadmissible decision rules.

rates admissibility.[4] However, Samuelson (1992) demonstrated that there are some significant obstacles, which are associated with the limitations of SEU theory, to this research program. Admissibility is typically obtained by requiring that players consider all states of the world to be probabilistically possible. However, a player who believes that her opponents are rational would exclude their inadmissible strategies from consideration. Elegant examples in Samuelson (1992) illustrated the frustrating fact that in many games, an inadmissible strategy may maximize her SEU under such beliefs.

Brandenburger, Friedenberg, and Keisler (2008) (BFK) solved this puzzle by adopting a model of Bayesian rationality that permits the expression of a more general set of beliefs than is allowed by SEU theory. In this framework, they defined the notion of *assuming* an event, which is immune to the aforementioned shortcomings of belief with probability 1 that were pointed out by Samuelson (1992). BFK formulated a condition, *rationality and common assumption of rationality* (RCAR), that gives intuitive support for the IA set as a solution concept.

Despite these advances, the message of BFK was decidedly negative. Given that an admissibility requirement partially reflects the view that rational players should rule out nothing, it makes sense to consider RCAR in model environments, such as *complete* lexicographic type structures, that include a sufficiently rich set of states. If the state space is unduly restricted by the type structure that describes it, the requirement of ruling out nothing in the state space feels arbitrary. Unfortunately, BFK found that in type structures that are both complete and continuous, no state of the world can satisfy RCAR. The issues leading to this nonexistence result are

[4]It is well-known that the order of eliminating weakly dominated strategies matters, whereas the order does not matter when eliminating strongly dominated strategies. When we refer to IA strategies in this paper, we will mean the strategies obtained by simultaneously deleting every weakly dominated strategy of every player in each round.

independent of those that were raised in Samuelson (1992). Furthermore, given that continuity was viewed as a technical condition, these results cast into doubt the existence of any complete type structure in which the RCAR set is nonempty.

In this paper, we revive this research program by giving positive answers to the questions left open by BFK. In the process, we also identify some of the conceptual issues that help us to reconcile the positive results herein with the negative conclusions of BFK.

First, we show that given any fixed game, there is a large class of complete lexicographic type structures in which RCAR can be satisfied. Furthermore, in a complete lexicographic type structure in which the RCAR set is nonempty, 1) RCAR implies that IA strategies are played, and 2) for each IA strategy, there is some RCAR state in which it is played. These results lend support to RCAR as an epistemic condition for IA strategies.

Second, we prove that given each discontinuous type structure, there exists a continuous type structure that describes the exact same set of beliefs. Where the two type structures differ, despite being equivalent in the sense just described, is in how they classify what beliefs assume a given event. Given a belief in a discontinuous type structure, the same belief in an equivalent continuous type structure will, in general, assume fewer events. One implication of this difference is that beliefs in the continuous type structure must meet a higher standard in order to “rule out nothing”. We argue that the discussion of these differences can be conveniently subsumed under the umbrella of *topological distinguishability*. As a disclaimer, we warn that despite the similarity in nomenclature, these issues are completely unrelated to those raised by the extensive literature on *strategic topology*.

Finally, we make a technical contribution by introducing a versatile tool—the method of piecewise isomorphisms—for constructing type structures in which certain sets of interest take a desirable and convenient form. We anticipate that modifications of this method may prove useful in other settings.

The continuing interest in foundations of IA strategies has its basis in the attractive properties they possess, which have long been recognized. Luce and Raiffa (1957, pp. 98–101) realized that IA solutions imply backward induction outcomes in games of interest, most strikingly in the case of the finitely repeated prisoner's dilemma. Luce and Raiffa (1957) also gave an early, albeit informal, justification of the iterated admissibility procedure as a consequence of CKR. Thus, it was known that iterated admissibility is a normal form solution concept with extensive form implications. Indeed, the IA set is invariant up to irrelevant transformations of the game tree à la Thompson (1952) and Dalkey (1953).[5] Therefore, it is consistent with the influential argument advanced by von Neumann and Morgenstern (1944) that the normal form captures all strategically relevant information about the game.

More recently, Brandenburger and Friedenberg (2007) showed that iterated admissibility is equivalent to Pearce's (1984) extensive form rationalizability (EFR) in game trees with generic payoffs. The backward and forward induction properties of EFR, which were investigated in Battigalli and Siniscalchi (2002), thus hold for iterated admissibility as well in such games. Even before such results were known, iterated admissibility was employed by Ben-Porath and Dekel (1992) toward obtaining forward induction outcomes in money-burning games. Iterated admissibility, by reflecting forward induction reasoning, can also refine the equilibria of various signaling games.

This paper is organized as follows. Section 2.2 gives an heuristic treatment of the key issues and the methodological approach we use to deal with them. Section 2.3 reviews the formal framework developed by BFK. Section 2.4 states and interprets

[5]For a more detailed discussion, see Kohlberg and Mertens (1986).

our main results. Section 2.5 concludes with a discussion of the finer conceptual issues and puts our results in the context of the literature. All proofs can be found in the appendix.

2.2 An Heuristic Treatment

2.2.1 Lexicographic Type Structures

We begin by informally describing the framework developed in BFK. Consider a finite game G that is played by Ann and Bob. Let S^a and S^b respectively denote Ann's strategy set and Bob's strategy set. Ann is uncertain of Bob's strategy choice and belief. It follows that her belief should be about Bob's strategy choice and belief, which means that the set of Bob's possible beliefs must be defined before the set of Ann's possible beliefs can be articulated. However, Bob's belief should be about Ann's strategy choice and belief. This inherent circularity can be avoided by allowing Ann's belief to be represented by what is called her type, which is an element of some well-defined space T^a.[6] An analogous object T^b can be introduced for Bob. Ann's possible beliefs can then be described, without self-reference, as beliefs about the ordered pair (s^b, t^b), which is composed of Bob's strategy choice and type.

The sets S^a, S^b, T^a, T^b, together with maps that associate with each type of each player a belief about the other player, form a model environment for epistemic analysis, called a *lexicographic type structure for G*. Figure 2.1 gives a conceptual sketch of the basic components that make up such a structure. Notice that some possible beliefs are not associated with any type in the structure pictured in Figure 2.1. In this paper, we limit our analysis to *complete* structures, which are rich in the sense

[6] This now-ubiquitous innovation is due to Harsanyi (1967).

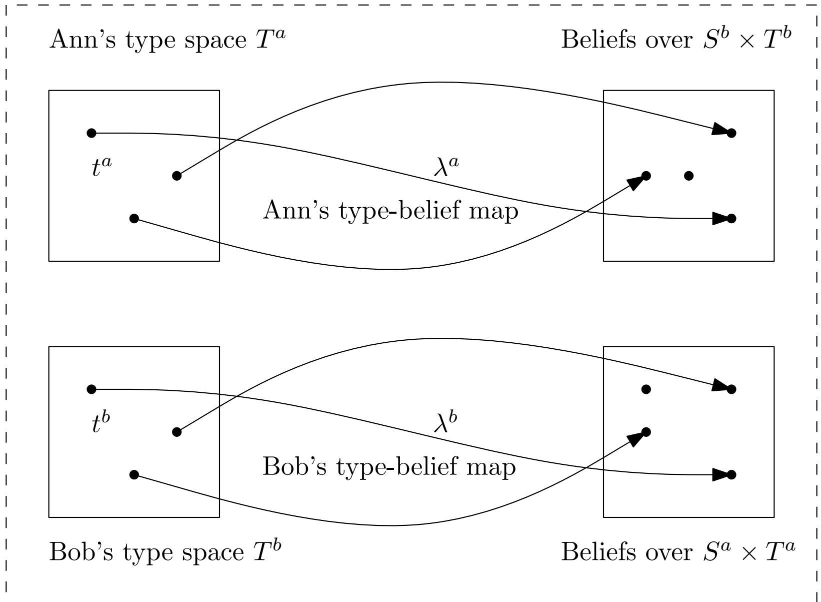

Figure 2.1: A model environment

of associating every possible belief with some type.

The precise notion of a lexicographic type structure in BFK requires that the type spaces be endowed with Polish—that is, separable and completely metrizable—topologies and that the mappings that associate beliefs with types be Borel. The analysis also depends on the precise topological notion of a belief assuming an event. In this section, we omit the technical details.

Lexicographic type structures can be used to analyze the properties of states in which statements regarding rationality, such as those found in the following list, hold.

$a(1)$	Ann is rational;	$b(1)$	Bob is rational;
$a(2)$	Ann is rational and assumes $b(1)$;	$b(2)$	Bob is rational and assumes $a(1)$;
$a(3)$	Ann is rational, assumes $b(1)$, assumes $b(2)$;	$b(3)$	Bob is rational, assumes $a(1)$, assumes $a(2)$;
…		…	

The richness of a complete structure is sufficient to guarantee the existence, for a given m, of a state in which $a(m)$ and $b(m)$ hold. If the statements $a(m)$ and $b(m)$ hold for all m in some state, we say that there is *rationality and common assumption of rationality* (RCAR) at that state. BFK showed that there are complete structures in which no such state exists. They left as an open question whether there are complete structures in which an RCAR state does exist. The main result of our paper, Theorem 2.4.2 answers this question in the affirmative.

A lexicographic type structure allows a proposition like $b(m)$ to be represented as the set of states (or the event) in which it holds; we denote this event by R^b_m. To prove our main theorem, we will develop a methodology for constructing complete lexicographic type structures that allows us to decide in advance which sets represent events about rationality (e.g., $R^b_1, R^b_2, \ldots$). What these sets will be is determined by the mappings λ^a from Ann's types to Ann's beliefs, and λ^b from Bob's types to Bob's beliefs. Once we are given the type spaces T^a and T^b, our task is to construct λ^a and λ^b in such a way that the resulting structure is complete (all beliefs are included) and the RCAR set is nonempty.

2.2.2 Topological Distinguishability

The results in BFK serve as useful guides in our construction. They show that in a complete lexicographic type structure for G, the events $R^b_1, R^b_2, \ldots$ form a strictly decreasing chain. If RCAR holds, Ann must assume each event in the countably infinite list that consists of R^b_1, R^b_2, and so on. It is also shown in BFK that a player can never assume more than a finite number of *topologically distinguishable* events. (Two events are topologically indistinguishable if one approximates the other and vice versa—that is, if they have identical closures.) If, as in Figure 2.2, R^b_{m+1} is

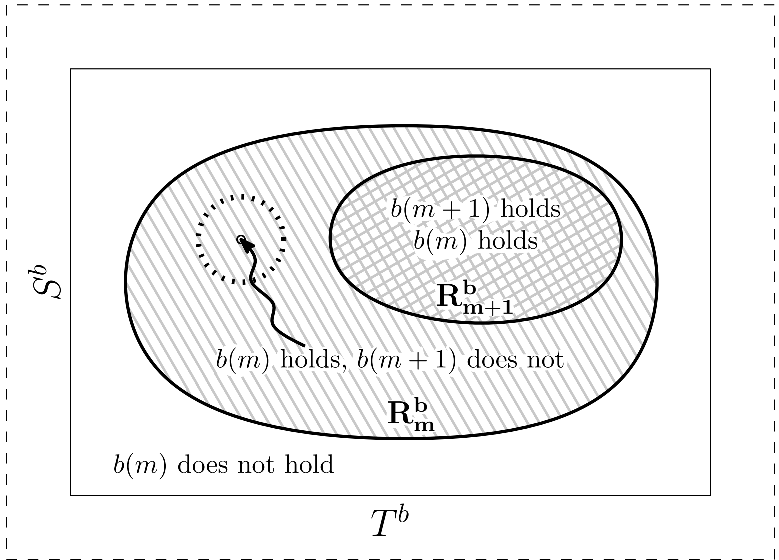

Figure 2.2: Distinguishable $b(m)$ and $b(m+1)$

topologically distinguishable for R^b_m for arbitrarily large m, then RCAR cannot hold in any state since Ann cannot assume an infinite number of distinguishable events. Therefore, in any complete lexicographic type structure where RCAR is nonempty, the events $R^b_m, R^b_{m+1}, \ldots$ must be topologically indistinguishable for all sufficiently large m.

2.2.3 Partition into Optimal Sets

We now explain in heuristic terms the idea behind our construction. As an example, consider the game in Figure 2.3.

		Bob	
		Left (L)	Right (R)
Ann	Up (U)	$1,1$	$0,1$
	Down (D)	$0,1$	$1,0$

Figure 2.3: A game

Iteratively eliminating inadmissible strategies from this game yields more useful guidelines for our construction. Recall that the primary motivation for the formulation of RCAR in BFK is to capture the reasoning implicit in the iterated admissibility procedure. In this game, the strategy R for Bob is ruled out at stage one, and the strategy D for Ann is ruled out at stage two. The strategy sets and the first two levels of admissible sets are as follows:

$$S^a = \{U, D\}\,, \qquad S_1^a = \{U, D\}\,, \qquad S_2^a = \{U\}\,,$$
$$S^b = \{L, R\}\,, \qquad S_1^b = \{L\}\,, \qquad S_2^b = \{L\}\,.$$

The IA set contains the single strategy pair (U, L). Ann's beliefs can be classified according to whether or not they are open-minded. To get admissibility as an output, BFK require that rational players hold open-minded beliefs.

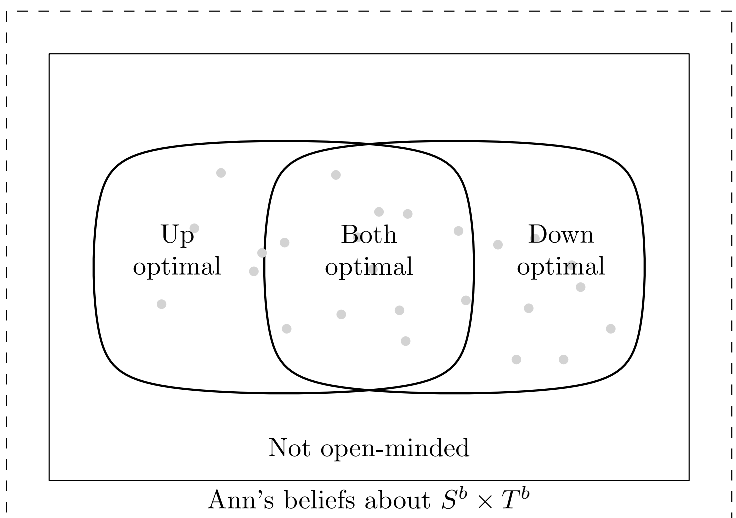

Figure 2.4: Classifying Ann's beliefs

Ann's open-minded beliefs can be further partitioned into three classes, one in which U is the unique optimal strategy, one in which D is the unique optimal strategy, and one in which both U, D are optimal, as in Figure 2.4. (The outer region is the set of beliefs that are not open-minded). The corresponding picture for Bob (not shown) has just two regions, for beliefs that are open-minded and not, because the only admissible strategy is L.

2.2.4 Piecewise Isomorphisms

We will construct a complete lexicographic type structure with RCAR for this game by gluing together countably many Borel isomorphisms between pieces of Ann's type space and corresponding pieces of Ann's belief space. The key fact is the Borel Isomorphism Theorem: *for any pair of Borel sets A, B with the same cardinality in Polish spaces X, Y, there is a one-to-one Borel mapping from A onto B.*[7] For this construction to work, we require that the type spaces T^a, T^b are "large", that is, have the cardinality of the continuum.

At the first stage we *arbitrarily* choose three large disjoint Borel subsets Y_U, Y_D, $Y_{U,D}$ of T^a such that the complement $T^a \setminus (Y_U \cup Y_D \cup Y_{U,D})$ is also large. The mappings λ^a, λ^b are left unspecified, but as we go along we will make a series of promises about them and eventually show that all these promises can be kept.

We start with the promise that λ^a is a Borel isomorphism from T^a onto the set of all Ann's beliefs, and that λ^a maps Y_U, Y_D, and $Y_{U,D}$ onto the sets of open-minded beliefs where U is the unique optimal strategy, where D is the unique optimal strategy, and where both U and D are optimal, respectively. This is illustrated in Figure 2.5. We make the analogous promise for Bob. These promises guarantee that the lexicographic type structure is complete, and that the set of rational states for

[7] This and other useful properties of Polish spaces can be found in Appendix B.3.

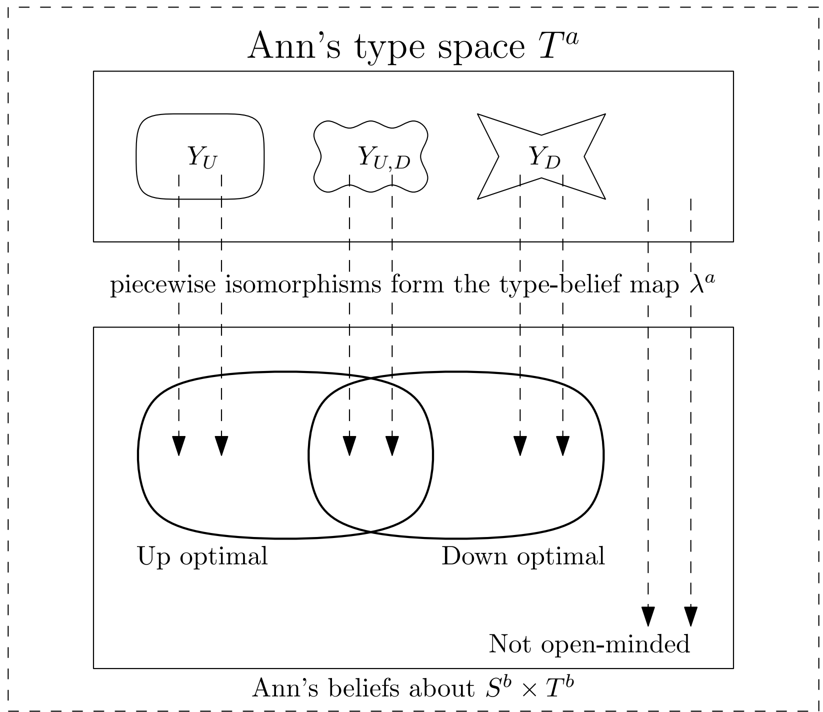

Figure 2.5: Piecewise isomorphisms

Ann will be

$$R_1^a = (\{U\} \times Y_U) \cup (\{D\} \times Y_D) \cup (\{U, D\} \times Y_{U,D}).$$

The set of states where Ann and Bob are rational look like the shaded set in Figure 2.6.

There is still more to be done, because the iterated rationality sets R_m^a continue to shrink as m increases. From stage 2 on, we need only work with the types in Y_U, because Ann's beliefs that are rational at stage 2 and beyond have the unique optimal strategy U. The iterated rationality sets $R_2^a, R_3^a, \ldots$ will be a strictly decreasing chain of subsets of $\{U\} \times Y_U$. We put $Y_1^a = Y_U$ and *arbitrarily* choose a decreasing chain

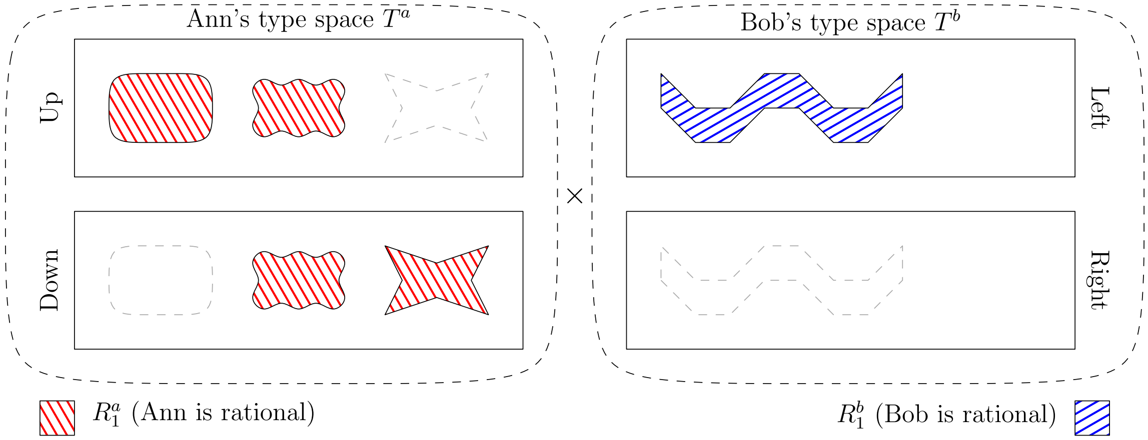

Figure 2.6: Events concerning rationality

of Borel sets $Y_1^a \supseteq Y_2^a \supseteq Y_3^a \supseteq \cdots$ such that the difference sets $Z_m^a \equiv Y_m^a \setminus Y_{m+1}^a$ are large, and the intersection $Y_\infty^a \equiv \bigcap_m Y_m^a$ is large and topologically indistinguishable from each of the sets Y_m^a. It is possible, but not necessary, to do this in such a way that all the sets Y_m^a and Y_∞^a are open. For $m = 1, 2, \ldots$ we make the promise that λ^a maps the set Y_{m+1}^a onto the set of open-minded beliefs with U uniquely optimal that assume $R_1^b, \ldots, R_m^b$. This promise guarantees that $R_{m+1}^a = \{U\} \times Y_{m+1}^a$. This is illustrated in Figure 2.7. As usual, we do the same for Bob.

If all of these promises can be kept, then the intersections $R_\infty^a = \{U\} \times Y_\infty^a$ and $R_\infty^b = \{L\} \times Y_\infty^b$ will be nonempty, and we will have a complete lexicographic type structure for G with nonempty RCAR, as required.

How can our promises be kept? Note that the set of Ann's open-minded beliefs with U uniquely optimal—which we have promised that Y_U will map onto as in

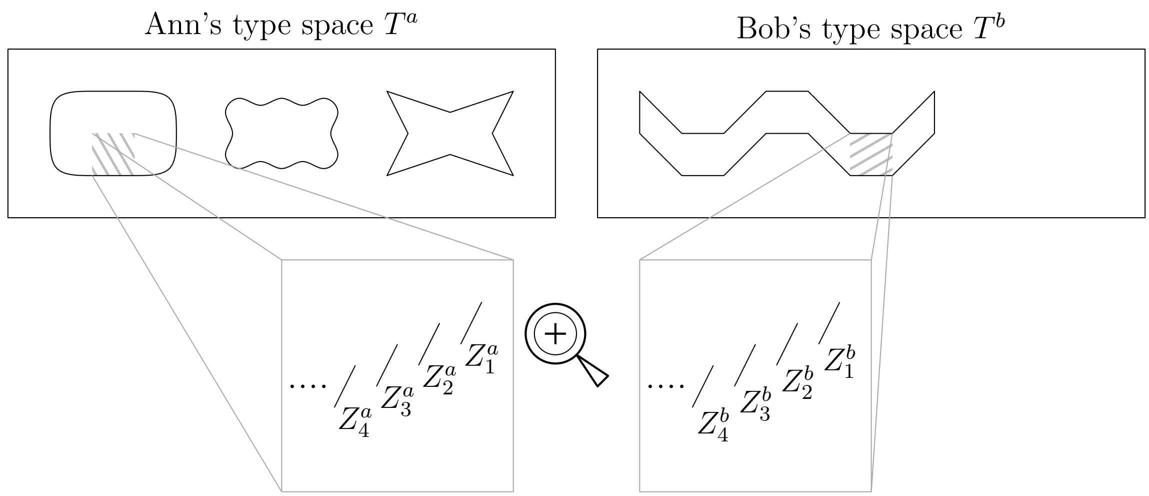

Figure 2.7: Successively removing continuum-large parts.

Figure 2.5—can be partitioned into the following countable list of disjoint classes:

$$A_1 = \text{Beliefs that do not assume } R_1^b;$$
$$A_2 = \text{Beliefs that assume } R_1^b \text{ but not } R_2^b;$$
$$\dots$$
$$A_\infty = \text{Beliefs that assume } R_1^b, R_2^b, \dots.$$

An important step is to show that the set A_∞ is nonempty. This is where we will use our requirement that the set of types Y_∞^a is topologically indistinguishable from $Y_1^a, Y_2^a, \dots$. We then use the Borel Isomorphism Theorem to find Borel mappings that satisfy our initial promise illustrated in Figure 2.5 outside of Y_1^a, a Borel mapping from Z_m onto A_m for each $m > 0$, and a Borel mapping from Y_∞ onto A_∞.[8] Finally, all these mappings can be glued together to get mappings λ^a, λ^b that satisfy all of

[8]Of course, this requires that A_∞ and each A_m is large. This property, as well as many others used in our arguments, follows directly from the myriad of useful technical lemmas included in BFK.

our promises.

2.3 The Underlying Framework

In this section, we briefly review the concepts we will need from BFK. For the remainder of this section, we fix a finite game of complete information

$$G = \left\langle S^a, S^b, \pi^a, \pi^b \right\rangle,$$

where S^a, S^b are strategy spaces and π^a, π^b are payoff functions. The indices a and b stand for Ann and Bob, respectively. *Whenever we state a definition or result involving a and/or b (Ann and/or Bob), it will be understood that we also make the analogous statement with a and b reversed.*

2.3.1 Admissibility

Ann's strategy $s^a \in S^a$ is **admissible** (i.e., not weakly dominated) in the game G if and only if s^a is optimal under some full-support probability measure defined over S^b. Let S_1^a denote the set of Ann's admissible strategies. Given nonempty subsets $X \subseteq S^a$ and $Y \subseteq S^b$, let $G(X, Y)$ denote the reduced game $\left\langle X, Y, \pi^a, \pi^b \right\rangle$. We can then inductively define Ann's m**-admissible** strategy set S_m^a as follows: To get the induction started, we write $S_0^a \equiv S^a$. For each $m \in \mathbb{N}$, let S_{m+1}^a be the set of Ann's admissible strategies in the reduced game $G(S_m^a, S_m^b)$. In other words, S_{m+1}^a is the set of Ann's strategies that are **admissible with respect to** $S_m^a \times S_m^b$.

Note that $S_{m+1}^a \subseteq S_m^a$ for all $m \in \mathbb{N}$. We put $S_\infty^a \equiv \bigcap_{m=0}^\infty S_m^a$. The set $S_\infty^a \times S_\infty^b$ is called the **iteratively admissible** set (henceforth IA set). Since the sets S^a, S^b are finite, we have $S_\infty^a = S_M^a$ and $S_\infty^b = S_M^b$ for some $M \in \mathbb{N}$, and hence the IA set is nonempty.

2.3.2 Lexicographic Probability Systems

Recall that a **Polish space** is a separable topological space that is completely metrizable. Let Ω denote the space of uncertainty faced by the decision maker (e.g., Ann). For now, let us assume only that Ω is Polish and fix a compatible metric. In the conventional Bayesian theory of choice under uncertainty, a decision maker's beliefs are represented by a Borel probability measure on Ω. The set of all Borel probability measures on Ω is denoted by $\mathcal{M}(\Omega)$.

Following an alternative theory developed in Blume, Brandenburger, and Dekel (1991a), BFK adopted the convention that a decision maker's beliefs are represented by a lexicographic probability system. Lexicographic probability systems (henceforth LPSs) are generalizations of probability measures. An **LPS** on Ω is any finite sequence of probability measures on Ω, e.g.,

$$\sigma = (\mu_0, \ldots, \mu_{n-1}) \in \overbrace{\mathcal{M}(\Omega) \times \cdots \times \mathcal{M}(\Omega)}^{n \text{ times}},$$

that satisfies a condition called **mutual singularity**—that is, there exist disjoint Borel sets $U_0, \ldots, U_{n-1}$ in Ω such that $\mu_i(U_i) = 1$ and $\mu_j(U_j) = 0$ for $i \neq j$.[9] The set of all LPSs on Ω is denoted by $\mathcal{L}(\Omega)$. It is immediate that $\mathcal{M}(\Omega) \subsetneq \mathcal{L}(\Omega)$. Additional

[9]The definition of LPSs in Blume, Brandenburger, and Dekel (1991a) did not require mutual singularity. The definition above is from BFK.

notation, which will be convenient later, follows below:

$$\mathcal{N}_n(\Omega) \equiv \overbrace{\mathcal{M}(\Omega) \times \cdots \times \mathcal{M}(\Omega)}^{n \text{ times}};$$

$$\mathcal{N}(\Omega) \equiv \bigcup_{n \in \mathbb{N}} \mathcal{N}_n(\Omega);$$

$$\mathcal{L}_n(\Omega) \equiv \mathcal{L}(\Omega) \cap \mathcal{N}_n(\Omega).$$

We define a Polish topology on $\mathcal{N}(\Omega)$ by following the usual conventions. First, we give $\mathcal{M}(\Omega)$ its weak* topology, which makes it a Polish space. Second, we give $\mathcal{N}_n(\Omega) = \prod_{k=1}^{n} \mathcal{M}(\Omega)$ the product topology. Then, we may view $\mathcal{N}(\Omega)$ as a countable topological union of disjoint Polish spaces $\mathcal{N}_n(\Omega)$. $\mathcal{N}(\Omega)$ with this topology is again a Polish space.

An LPS $\sigma = (\mu_0, \ldots, \mu_{n-1})$ represents a sequence of mutually contradictory hypotheses. We interpret μ_0 as being infinitely more likely than μ_1, which in turn is infinitely more likely than μ_2, and so on. The primary hypothesis μ_0, being more likely than all other hypotheses, can be regarded as the prior belief. The secondary hypothesis μ_1 can be regarded as the conditional belief in the a priori zero-probability (i.e., μ_0-null) event that μ_0 is false. More generally, μ_j is the conditional belief in the event that all a priori more likely hypotheses (i.e., all μ_k such that $k < j$) are false. Such an event would be μ_k-null for all $k < j$.

LPSs generalize the notion of probability measures in a straightforward manner. Not surprisingly, concepts defined with respect to probability measures often have obvious analogs that are defined with respect to LPSs.

The **support** of an n-tuple of measures $\sigma \in \mathcal{N}(\Omega)$ is the union of the supports of the measures that comprise it—that is, the support of $\sigma = (\mu_0, \ldots, \mu_{n-1})$ is simply

$$\operatorname{Supp}\sigma \equiv \bigcup_{j=0}^{n-1} \operatorname{Supp}\mu_j.$$

We say that σ has **full support** if $\operatorname{Supp}\sigma = \Omega$. Equivalently, σ has full support if, for each open U, there exists $j < n$ such that $\mu_j(U) > 0$. The set of all full-support LPSs is denoted by $\mathcal{L}^+(\Omega)$. The set $\mathcal{N}^+(\Omega)$ is defined similarly.

Similarly, Bayesian optimization under belief σ is a straightforward extension of expected utility maximization. Given an act f, for each $j < n$ let u_j be the expected utility of choosing f with respect to μ_j. Then, the vector $u = (u_0, \ldots, u_{n-1})$ is called the **lexicographic expected utility** (henceforth LEU) of f under σ. The order on the hypotheses that comprise σ suggests an obvious way to compare LEU vectors. Given that, for $i < j$, μ_i is infinitely more likely than μ_j, it is natural to assign infinitely more importance to the expected utility of an act under μ_i than one would to its expected utility under μ_j. We write

$$(v_0, \ldots, v_{n-1}) = v >_{\mathsf{LEX}} u = (u_0, \ldots, u_{n-1})$$

and say that v is lexicographically greater than u if there exists $k < n$ such that $v_k > u_k$, and $v_j = u_j$ for all $j < k$. LEU maximization is simply the maximization of LEU with respect to the lexicographic order. Throughout this paper, we use the terms LEU and payoff interchangeably in appropriate contexts.

The idea of certainty (i.e., belief with probability one) admits more than one obvious analog with respect to LPSs. If the decision maker is certain of event $E \subseteq \Omega$, then she considers E to be infinitely more likely than its complement $\Omega \setminus E$. BFK introduced a new epistemic notion, called **assumption**, to capture this property in

LPSs. Intuitively speaking, a decision maker with belief $\sigma = (\mu_0, \ldots, \mu_{n-1})$ **assumes** an event E if she believes **every part** of E to be infinitely more likely than its complement $\Omega \setminus E$. Formally, a Borel set E is assumed under σ at level j if the following three conditions are met (cf. Proposition 5.1 in BFK):

1. $\mu_i(E) = 1$ for each $i \leq j$;
2. $\mu_i(E) = 0$ for each $i > j$; and
3. If U is open with $U \cap E \neq \varnothing$, then $\mu_i(U \cap E) > 0$ for some $i < n$.

Note that, even if E is assumed under σ, it need not be the case that $\Omega \setminus E$ is σ-null. In contrast, if a decision maker is certain of E, then $\Omega \setminus E$ is necessarily a null event.

It is clear that if an event E is assumed under an LPS $\sigma = (\mu_0, \ldots, \mu_{n-1})$, then the level at which E is assumed is unique, is less than n, and is the greatest j such that $\mu_j(E) = 1$.

It is also clear that if $\mu = (\mu_0, \ldots, \mu_{n-1})$ and $\nu = (\nu_0, \ldots, \nu_{n-1})$ are LPSs of the same length n, and μ_j, ν_j have the same null sets for each $j < n$, then μ and ν assume the same events at each level $j < n$. Verbally, the events that an LPS μ assumes depend only on the length of μ and the null sets of the μ_j.

2.3.3 Lexicographic Type Structures

LPSs and associated constructs were used in BFK to build a framework in which the rationale for iterated admissibility can be expressed formally (i.e., in the language of set theory).

Recall the finite game $G = \langle S^a, S^b, \pi^a, \pi^b \rangle$. In the context of the game G, Ann is uncertain of which strategy Bob will choose, what Bob believes about Ann's strategy choice, what Bob believes about what Ann believes about Bob's strategy choice, and so on. To give a parsimonious description of Ann's beliefs about the pair consisting of Bob's strategy and Bob's beliefs while sidestepping the inherent problem of self-reference, BFK followed the convention of implicitly representing beliefs as **types**.[10] Ann's type t^a is an element of a Polish space T^a, called her **type space**. The belief that Ann's type represents is given by a Borel map $\lambda^a : T^a \to \mathcal{L}(S^b \times T^b)$, where T^b denotes Bob's type space. Similarly, Bob's types are interpreted through a Borel map $\lambda^b : T^b \to \mathcal{L}(S^a \times T^a)$. Taken together, these objects form a 6-tuple

$$\mathfrak{T} = \left\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \right\rangle,$$

which is called an (S^a, S^b)**-based lexicographic type structure**, or alternatively, a **lexicographic type structure for** G. Members of $S^a \times T^a \times S^b \times T^b$ are called **states of the world**.

The lexicographic type structure $\mathfrak{T}$ is called **complete** if $\mathcal{L}(S^b \times T^b) = \operatorname{range} \lambda^a$ and $\mathcal{L}(S^a \times T^a) = \operatorname{range} \lambda^b$.[11] A complete lexicographic type structure contains all beliefs about beliefs.

2.3.4 Rationality

The definition of rationality in BFK combines two requirements. The first, that of Bayesian optimization, is captured by LEU maximization. The second, which might be roughly described as a form of agnosticism, is reflected in full-support beliefs. Formally, the LEU of a strategy $s^a \in S^a$ under the LPS $\sigma = (\mu_0, \ldots, \mu_n)$ is the tuple of payoffs $(\pi^a(s^a, \nu_0), \ldots, \pi^a(s^a, \nu_n))$ where $\nu_i = \operatorname{marg}_{S^b} \mu_i$, and s^a is **optimal** under

[10] An innovation due to Harsanyi (1967).

[11] In BFK, a type structure is called complete if $\mathcal{L}^+(S^b \times T^b) \subsetneq \operatorname{range} \lambda^a$ and $\mathcal{L}^+(S^a \times T^a) \subsetneq \operatorname{range} \lambda^b$. This difference is immaterial with respect to both their results and ours.

σ if the LEU of s^a under σ is maximal among all strategies in S^a. A strategy-type pair (s^a, t^a) is **rational** if $\lambda^a(t^a)$ is a full-support LPS, and s^a is optimal under $\lambda^a(t^a)$. The set of all rational pairs (s^a, t^a) is denoted by R_1^a. For each $m > 0$, define R_m^a inductively by

$$R_{m+1}^a \equiv R_m^a \cap \left[S^a \times A^a(R_m^b)\right],$$

where $A^a(R_m^b)$ is the set of Ann's types in T^a such that $\lambda^a(t^a)$ assumes R_m^b. If a state $(s^a, t^a, s^b, t^b) \in R_{m+1}^a \times R_{m+1}^b$, then we say that it satisfies **rationality and m-th order assumption of rationality** (henceforth RmAR).

We write $R_0^b \equiv S^b \times T^b$, and $R_\infty^b \equiv \bigcap_{m \in \mathbb{N}} R_m^b$. Note that R_0^b is trivially assumed under every full-support LPS on $S^b \times T^b$, so every open-minded Ann assumes R_0^b. It is shown in BFK, Lemma C.4, that each of the sets R_m^a, R_m^b is Borel (so the players are able to assume these sets), and that

$$R_m^a = R_1^a \cap \left[S^a \times \bigcap_{i<m} A^a(R_i^b)\right].$$

In words, R_m^a is the set of states for which Ann is rational and assumes that Bob is i-th order rational for each $i \leq m$. If a state (s^a, t^a, s^b, t^b) belongs to $R_\infty^a \times R_\infty^b$, then it satisfies **rationality and common assumption of rationality** (henceforth RCAR). In words, each player is rational and assumes that the other player is m-th order rational for each $m \in \mathbb{N}$. It is shown in BFK that any LPS that assumes each of a countable sequence of events assumes their intersection. It follows that for any RCAR state, each player assumes that the other player is rational at order ∞, that is, Ann assumes R_∞^b and Bob assumes R_∞^a.

2.4 Statements of Results

Section 2.4.1 states our main results, which 1) establish that complete lexicographic type structures with RCAR exist and 2) reconcile that fact with the negative conclusions found in the literature. Section 2.4.2 states some complementary results that relate beliefs about strategies to iterated admissibility. We need these results to prove our main theorems, but they also merit independent consideration because they reveal certain structural commonalities of finite-order reasoning about rationality across complete lexicographic type structures.

In order to easily distinguish new results from previous results from the literature, we will reserve the name "Proposition" for previous results from the literature, and use "Theorem", "Corollary", and "Lemma" for new results.

2.4.1 Main Theorems

Consider the following infinite sequence of statements.

(a1) Ann is rational;

(a2) Ann is rational and assumes (b1);

(a3) Ann is rational, assumes (b1), assumes (b2);

...

(b1) Bob is rational;

(b2) Bob is rational and assumes (a1);

(b3) Bob is rational, assumes (a1), assumes (a2);

...

For each $m > 1$, the statement "a(m+1) and b(m+1)" corresponds to rationality and m-th order assumption of rationality. The conjunction of this infinite sequence of statements corresponds to rationality and common assumption of rationality.

A lexicographic type structure

$$\mathfrak{T} = \left\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \right\rangle$$

for G provides precise interpretations of these statements by implicitly defining the universe of beliefs that each player may hold. Each type t^a of Ann represents a system of beliefs about the infinite sequence of objects (B1), (B2),... below. Similarly, each type t^b of Bob represents a system of beliefs about the sequence (A1), (A2),... below.

(A1) Ann's action s^a;

(A2) Ann's action s^a and her beliefs about (B1);

(A3) Ann's action s^a and her beliefs about (B1),(B2);

...

(B1) Bob's action s^b;

(B2) Bob's action s^b and his beliefs about (A1);

(B3) Bob's action s^b and his beliefs about (A1),(A2);

...

BFK found that if the universe of beliefs implied by $\mathfrak{T}$ is rich enough—that is, $\mathfrak{T}$ is a complete structure—then the set of strategies played when RmAR holds coincides exactly with the set of $(m + 1)$-admissible strategies. Proposition 2.4.1 below is the formal statement of this result.

Proposition 2.4.1 (Theorem 9.1 in BFK)**.** *Fix a finite game G and a complete lexicographic type structure $\mathfrak{T}$ for G. Then, for each $m \in \mathbb{N}$,*

$$\mathrm{proj}_{S^a} R^a_m \times \mathrm{proj}_{S^b} R^b_m = S^a_m \times S^b_m.$$

It is natural to ask whether there is an analogous result that characterizes iterated admissibility using RCAR. Our main results, Theorems 2.4.2 and 2.4.4 below, establish the epistemic foundations of IA along those lines. In particular, Corollary 2.4.3 shows that a complete lexicographic type structure in which the RCAR set is nonempty exists, answering an open question that was asked in BFK.

Theorem 2.4.2 (Existence Theorem)**.** *Fix a finite game* $G = \left\langle S^a, S^b, \pi^a, \pi^b \right\rangle$ *and any uncountable Polish spaces* T^a, T^b*. There exist Borel functions* λ^a, λ^b *such that*

$$\mathfrak{T} = \left\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \right\rangle$$

is a complete lexicographic type structure for G *in which* $R^a_\infty \times R^b_\infty$ *is nonempty.*

Corollary 2.4.3. *Fix a finite game* G*. There exists a complete lexicographic type structure* $\mathfrak{T}$ *for* G *in which* $R^a_\infty \times R^b_\infty$ *is nonempty.*

Theorem 2.4.4. *Fix a finite game* G *and suppose* $\mathfrak{T}$ *is a complete lexicographic type structure for* G *such that* $R^a_\infty \times R^b_\infty$ *is nonempty. Then,*

$$\operatorname{proj}_{S^a} R^a_\infty \times \operatorname{proj}_{S^b} R^b_\infty = S^a_\infty \times S^b_\infty.$$

In words, Corollary 2.4.3 says that there exists a complete lexicographic type structure for G with at least one state that belongs to the RCAR set. Theorem 2.4.4 says that in every complete lexicographic type structure for G in which the RCAR set is nonempty, the set of strategies played when RCAR holds is exactly equal to the IA set.

It is not the case that for every complete lexicographic type structure for G, the set of strategies played when RCAR holds is exactly equal to the IA set. Consider the following two results from BFK.

Proposition 2.4.5 (Theorem 10.1 in BFK)**.** *Fix a finite game G and a complete lexicographic type structure $\mathfrak{T}$ for G such that the maps λ^a, λ^b are continuous. If there exist r^a, s^a, s^b such that $\pi^a(r^a, s^b) \neq \pi^a(s^a, s^b)$, then $R^a_\infty \times R^b_\infty = \varnothing$.*

Proposition 2.4.6 (Proposition 7.2 in BFK)**.** *For each finite game G there exists a complete lexicographic type structure $\mathfrak{T}$ for G such that the maps λ^a, λ^b are continuous.*[12]

These two results together show that there are complete lexicographic type structures for G in which the RCAR set is empty. So, the set of strategies played when RCAR holds is empty, but of course the IA set is nonempty.

How do we reconcile our results with Proposition 2.4.5? In particular, how should we understand the fact that complete lexicographic type structures having nonempty RCAR sets cannot have continuous type-belief maps? To do so, we give an improvement of Proposition 2.4.6. By a **Borel refinement** of a Polish space T we mean a Polish space U such that U has the same set of points and the same Borel σ-algebra as T, and every open set in T is open in U. Thus, U has the same Borel sets but more open sets.

Theorem 2.4.7. *Let*

$$\mathfrak{T} = \left\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \right\rangle$$

be a complete lexicographic type structure for a finite game G. Then, there exist Borel refinements U^a, U^b of T^a, T^b such that the maps

$$\lambda^a : U^a \to \mathcal{L}(S^b \times U^b), \quad \lambda^b : U^b \to \mathcal{L}(S^a \times U^a)$$

[12] Since our definition of complete is slightly different from the definition in BFK, the proof must use Theorem 13.7 instead of Theorem 7.9 in Kechris (1995).

are continuous.

It follows that

$$\mathcal{U} = \left\langle S^a, S^b, U^a, U^b, \lambda^a, \lambda^b \right\rangle$$

is again a complete lexicographic type structure for G, so Theorem 2.4.7 implies Proposition 2.4.6. Since the Borel σ-algebras are unchanged, the sets of LPS are unchanged, that is,

$$\mathcal{L}(S^a \times U^a) = \mathcal{L}(S^a \times T^a), \quad \mathcal{L}(S^b \times U^b) = \mathcal{L}(S^b \times T^b).$$

However, more open sets have been added to the topologies.

No state satisfying RCAR exists in the lexicographic type structure $\mathcal{U}$ by Proposition 2.4.5. Note that while $\mathcal{L}(S^b \times T^b) = \mathcal{L}(S^b \times U^b)$, it is not the case that $\mathcal{L}^+(S^b \times T^b) = \mathcal{L}^+(S^b \times U^b)$. This is because full-support LPSs must assign positive measure to every open set and U^b contains more open sets than T^b does. Effectively, there are fewer full-support types in U^a than there are in T^a. So, the operation of refining the topologies of the type spaces in this fashion shrinks the set of states in which every player is rational.

This is actually the consequence of a more basic change. Recall that, to assume an event, it is necessary to assign positive measure to every "part" of it. These "parts" are topologically distinguishable parts. Refining the topology on a space implies the addition of such "parts" and therefore raises the standards that need to be met in order to assume an event. This operation shrinks the set of full-support types because having a full-support belief is essentially equivalent to assuming the entire state space.

The description of continuous lexicographic type structures in BFK as lexicographic type structures in which neighboring full-support LPSs are associated with neighboring full-support types, while equivalent, does not immediately call this property to attention. Intuitively speaking, a player who is very strongly agnostic cannot assume very much. In the same vein, it might be said that players who are very strongly agnostic cannot assume enough for common assumption of rationality. The above exercise suggests that in complete lexicographic type structures with continuous type-belief maps, all rational types are too strongly agnostic.

2.4.2 Strategic Beliefs and Iterated Admissibility

In this section, we give an alternate definition of the IA set in the style of rationalizability via an iterative refinement of the players' *strategic beliefs*—that is, their marginal beliefs over opponents' strategies.[13] We will show that this refinement process captures some of the structural properties of RmAR sets that are conveniently invariant across all complete lexicographic type structures. In other words, RmAR sets can be described as having the same "shape" in a sense across all such type structures. These properties are incredibly useful in proving the existence theorems of Section 2.4.1.

We suppose throughout that $G = \langle S^a, S^b, \pi^a, \pi^b \rangle$ is a finite game in strategic form. We first introduce some notation. For $r^a, s^a \in S^a$ and a sequence $\nu = (\nu_0, \ldots, \nu_n) \in \mathcal{N}(S^b)$, we say that s^a is **preferred** to r^a under ν, and write $s^a \succ_\nu r^a$, if the LEU of s^a under ν is greater than that of r^a—that is, $\pi^a(s^a, \nu) >_{\mathsf{LEX}} \pi^a(r^a, \nu)$, where $\pi^a(s^a, \nu) \equiv (\pi^a(s^a, \nu_0), \ldots, \pi^a(s^a, \nu_n))$. Note that the leftmost term $\pi^a(s^a, \nu_0)$ has the highest priority. Intuitively speaking, under belief $\nu = (\nu_0, \ldots, \nu_n)$, for each

[13]The refinement process may succinctly be described as *set-valued* lexicographic rationalizability, given its similarity to Stahl's (1995) notion of *lexicographic rationalizability*.

$k \leq n$ the strategies in the support of $(\nu_0, \ldots, \nu_k)$ are infinitely more likely than the strategies outside the support of $(\nu_0, \ldots, \nu_k)$.

Given $\mu, \nu \in \mathcal{N}(S^b)$, we write $\mu \sim \nu$ if for all $r^a, s^a \in S^a$, $s^a \succ_\mu r^a$ if and only if $s^a \succ_\nu r^a$. In other words, $\mu \sim \nu$ if and only if they induce the same preference ordering over pure strategies. It is easy to see that $\sim$ is an equivalence relation. If $\mu = (\mu_0, \ldots, \mu_m)$ and $\nu = (\nu_0, \ldots, \nu_n)$, then the **concatenation** of μ and ν is defined as the sequence

$$\mu\nu \equiv ((\mu\nu)_0, \ldots, (\mu\nu)_{m+n+1}) = (\mu_0, \ldots, \mu_m, \nu_0, \ldots, \nu_n).$$

We note that if $\mu \sim \nu$ and $\mu' \sim \nu'$, then $\mu\nu \sim \mu'\nu'$.

Now let

$$P_1^a \equiv \mathcal{N}^+(S^b) = \left\{\nu : \nu \in \mathcal{N}(S^b) \wedge \operatorname{Supp} \nu = S^b\right\}$$

and define the following for each $m > 0$:

$$P_{m+1}^a \equiv \left\{\nu\nu' : \nu \in \mathcal{N}(S^b) \wedge \operatorname{Supp} \nu = S_m^b \wedge \nu' \in P_m^a\right\}.$$

The set P_1^a can be interpreted as the set of strategic beliefs held by rational types. Fix a complete lexicographic type structure $\left\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \right\rangle$. If Ann is rational then she has a full-support belief $\mu \in \mathcal{L}^+(S^b \times T^b)$ and her strategic belief is $\operatorname{marg}_{S^b} \mu$. Then the set of strategic beliefs that may be held by rational Anns is

$$\left\{\operatorname{marg}_{S^b} \mu : \mu \in \mathcal{L}^+(S^b \times T^b)\right\} = \mathcal{N}^+(S^b) = P_1^a.$$

We will call P_1^a the set of Ann's *rational* strategic beliefs. It readily follows that S_1^a

is the set of strategies played by rational Anns.

Following the intuitive description given above, we can say that if Ann holds a strategic belief $\nu \in P_2^a$ then she considers the event that Bob is rational to be infinitely more likely than the event that he is not. Furthermore, we can say that $\nu = \mathrm{marg}_{S^b} \mu$ for some full-support belief μ of Ann that assumes that Bob is rational. An inductive argument shows that if $\nu \in P_{m+1}^a$ then ν is the marginal on S^b of some full-support belief of Ann that m-th order assumes rationality.

For each $\nu \in \mathcal{N}(S^b)$, let $\mathbb{O}(\nu)$ denote the set of all $s^a \in S^a$ such that s^a is optimal under $\succ_\nu$ (i.e., s^a maximizes LEU under ν). Note that if $\mu \sim \nu$ then $\mathbb{O}(\mu) = \mathbb{O}(\nu)$. For each $m > 0$, define

$$\mathbb{X}_m^a \equiv \{\mathbb{O}(\nu) : \nu \in P_m^a\}.$$

We have the following characterization of m-admissible strategies as strategies that are optimal under strategic beliefs in P_m^a.

Theorem 2.4.8. *For each $m > 0$, $\bigcup \mathbb{X}_m^a = S_m^a$. Thus, each $s^a \in S_m^a$ belongs to some $X^a \in \mathbb{X}_m^a$, and $\mathbb{X}_m^a$ is a set of subsets of S_m^a.*

Note that Theorem 2.4.8 allows us to rewrite the definition of P_{m+1}^a without reference to the m-admissible set S_m^b. In fact, all results in this section would continue to hold even if we had started with the following definition of P_{m+1}^a.

$$P_{m+1}^a \equiv \left\{\nu\nu' : \nu \in \mathcal{N}(S^b) \wedge \mathrm{Supp}\,\nu = \mathbb{O}(P_m^b) \wedge \nu' \in P_m^a\right\},$$

where $\mathbb{O}(P_m^b) = \bigcup\{\mathbb{O}(\mu) : \mu \in P_m^b\}$.

We illustrate Theorem 2.4.8 with a simple example.

Example 2.4.9. *Consider the game from our heuristic treatment in Section 2.2. In this game,* $S^a = \{U, D\}, S^b = \{L, R\}$. *The IA set is the single pair of strategies* $\{(U, L)\}$, *which is reached at stage two. For Ann, we have*

$$S_1^a = \{U, D\}, \quad \mathbb{X}_1^a = \{\{U\}, \{D\}, \{U, D\}\},$$

and for each $m \geq 2$,

$$S_m^a = \{U\}, \quad \mathbb{X}_m^a = \{\{U\}\}.$$

For Bob, for every $m \geq 1$ *we have*

$$S_m^b = \{L\}, \quad \mathbb{X}_m^b = \{\{L\}\}.$$

We are also able to show that for every full-support belief μ of Ann who m-th order assumes rationality, there is some $\nu \in P_{m+1}^a$ such that $\operatorname{marg}_{S^b} \mu$ and ν induce the same preferences over Ann's strategies—that is, the strategic beliefs in P_{m+1}^a are representative of Ann's preferences over her strategies in states of the world satisfying RmAR in a complete lexicographic type structure. Theorem 2.4.10 below gives a precise statement of this relationship.

Theorem 2.4.10. *In a complete lexicographic type structure for a finite game, for each* $m > 0$ *we have:*

1. *If* $(s^a, t^a) \in R_m^a$ *then* $\exists \nu \in P_m^a$ *such that* $\operatorname{marg}_{S^b} \lambda^a(t^a) \sim \nu$; *and*
2. *If* $\nu \in P_m^a$ *then* $\exists (s^a, t^a) \in R_m^a$ *such that* $\operatorname{marg}_{S^b} \lambda^a(t^a) = \nu$.

By definition, a strategy s^a is optimal under $\lambda^a(t^a)$ if and only if s^a is optimal under $\operatorname{marg}_{S^b} \lambda^a(t^a)$. Thus we have the following corollary.

Corollary 2.4.11. *In a complete lexicographic type structure for a finite game G, for each $m > 0$, a set $X^a \subseteq S^a$ belongs to $\mathbb{X}^a_m$ if and only if there is a state $(s^a, t^a) \in R^a_m$ such that X^a is the set of all optimal strategies under $\lambda^a(t^a)$. Moreover, for each m we have $\mathbb{X}^a_m \supseteq \mathbb{X}^a_{m+1}$.*

Theorems 2.4.8 and 2.4.10 tell us something about the "shape" of the RmAR sets in a complete lexicographic type structure for G. To see this, we consider an *arbitrary* relation $Q \subseteq S^a \times T^a$ and subset $X^a \subseteq S^a$, and define

$$\Gamma^a(X^a, Q) \equiv \{t^a \in T^a : X^a = \{s^a : (s^a, t^a) \in Q\}\}.$$

In words, $\Gamma^a(X^a, Q)$ is the set of all $t^a \in T^a$ such that the section of Q at t^a is exactly X^a. It is clear that for each set $Q \subseteq S^a \times T^a$, the family of sets $\{\Gamma^a(X^a, Q) : X^a \subseteq S^a\}$ is pairwise disjoint, and the union of the family is T^a. Thus, the nonempty sets in this family form a finite partition of T^a.

It follows that in any lexicographic type structure for G, and for each nonempty set $X^a \subseteq S^a$ and $m > 0$, $\Gamma^a(X^a, R^a_m)$ is the set of all types t^a for Ann such that X^a is the set of optimal strategies for $\lambda^a(t^a)$, and Ann is open-minded and assumes k-th order rationality for Bob for all $k < m$.[14]

The next corollary shows that the RmAR sets have similar "shapes" in all complete lexicographic type structures for a given finite game G.

Corollary 2.4.12. *In a complete lexicographic type structure for a finite game, for each nonempty $X^a \subseteq S^a$ we have:*

1. *The sequence $\{\Gamma^a(X^a, R^a_m) : m > 0\}$ is a decreasing chain of Borel sets of T^a;*
2. *For each $m > 0$, $\Gamma^a(X^a, R^a_m)$ is nonempty if and only if $X^a \in \mathbb{X}^a_m$;*

[14] Bob is k-th order rational if he is rational and $(k-1)$-th order assumes rationality.

3. *The sequence* $\{\Gamma^a(\varnothing, R^a_m) : m > 0\}$ *is an increasing chain of nonempty Borel sets of* T^a*.*

Corollary 2.4.12 gives us the following useful formula for the RmAR sets in a complete lexicographic type structure for a finite game:

$$R^a_m = \bigcup \{X^a \times \Gamma^a(X^a, R^a_m) : X^a \in \mathbb{X}^a_m\}\,.$$

2.5 Discussion

To understand the results of this paper, let us give them further context by considering the related problem of finding epistemic conditions for IU strategies. The results of Brandenburger and Dekel (1987a) essentially give us the following fact:

Fix a game.

1. For each type structure, the set of strategies consistent with RCBR is a subset of the IU strategies.
2. There exists a finite type structure such that the set of strategies consistent with RCBR is exactly the IU set.

This says that if we look at RCBR across all type structures, we get the IU set. Taken together, our Theorems 2.4.2 and 2.4.4 imply the following analogous fact.

Fix a game.

1. For each complete lexicographic type structure, the set of strategies consistent with RCAR is a subset of the IA strategies.
2. There exists a complete lexicographic type structure such that the set of strategies consistent with RCAR is exactly the IA set.

This says that if we look at RCAR across all complete lexicographic type structures, we get the IA set.[15]

It may appear that this is the end of the matter. However, under the epistemic game theory (EGT) theory approach, the beliefs that the players deem possible—and therefore the type structure that generates them—are part of the description of the strategic situation. From the perspective of the players, the type structures other than the one that describes their strategic situation are simply irrelevant. Such extraneous type structures may exclude types that the players consider possible or include types that the players consider impossible. Thus, while an analyst can find a justification for each IU strategy by looking across all type structures, the player, whose perception is confined to the boundaries defined by the type structure that describes her situation, is not assured of being able to the same.

This raises a question: Can the players themselves see all the IU strategies as the result of a certain thought process? This is equivalent to asking whether there exists a type structure that is "rich" enough in the sense that, for each IU strategy, the type structure includes the description of at least one thought process that justifies it. Indeed, many type structures fail to satisfy such a richness condition.

Brandenburger and Dekel (1987a) showed that, given a fixed game G, we can tailor the type structure so that the IU strategies of G are the output of RCBR.[16] Since this construction depends on the game G, it may be the case that this type structure is not "rich enough" to give us the IU strategies of another game G' as the output of RCBR.

[15]BFK has shown that, if we look at RCAR across all lexicographic type structures, we do not get the IA set.

[16]Brandenburger and Dekel (1987a) use finite partitions, not finite type structures. However, a finite partition structure is essentially equivalent to some finite type structure.

However, this result has no direct analog with respect to RCAR and the IA set. BFK showed that if a lexicographic type structure is complete and continuous then it contains no state that satisfies RCAR.[17] The pessimism in BFK with respect to finding an epistemic condition for IA strategies was due principally to this negative result.

Our Theorem 2.4.2 revives the research program by showing that, given a game G, we can construct complete lexicographic type structures with discontinuous belief maps in which some state satisfies RCAR. This suggested that continuity, which had appeared to be a technical condition ex ante, changes the players' reasoning. Our Theorem 2.4.7 gives a striking way to isolate the effects of continuous belief maps by showing that, given *any* complete lexicographic type structure, there is a corresponding complete and continuous structure that describes exactly the same beliefs—that is, the two type structures are equally rich in at least one sense.

We identify the difference between a complete structure $\mathfrak{T}$ and its continuous counterpart $\mathfrak{U}$ given by Theorem 2.4.7 as one of topological distinguishability, which affects the classification of beliefs rather than changing them. Thus, it turns out that a type structure in the BFK framework captures more information than just what states of the world are considered possible by the players.

$\mathfrak{U}$ gives a finer topologization of the state space than does $\mathfrak{T}$. How should we interpret this difference? The topology on a state space, say Ω, is essentially the set of events that open-minded Bayesians must consider in their decision making. However, we find it more convenient to start with the interpretation that a topology separates and distinguishes hypotheses about the true state of world. Consider two hypotheses, which are respectively represented by events E and E'. If their closures are equal, as determined by a topology $\mathcal{T}$ on Ω, then it may be said that E and E'

[17] In fact, a complete, compact, and continuous lexicographic type structure does not even exist.

are indistinguishable in $\mathcal{T}$ because E approximates E' in an arbitrarily fine way and vice versa.

Whether an event is assumed by a given belief is sensitive to the topology on the state space. As this topology is successively refined, a given belief will be classified as assuming fewer and fewer events. We might then informally describe the difference between $\mathfrak{T}$ and $\mathfrak{U}$ as follows: The players in the environment described by the latter are more careful about saying that they assume something than the players in the environment described by the former. This relationship gives an intuitively appealing reconciliation of BFK's negative result with our positive result. Players described by continuous type structures are just too careful to commonly assume rationality.

Furthermore, it is easily argued that a decision maker who assigns a non-zero probability to each open set is simply giving proper consideration to all distinguishable hypotheses. Therefore, it is even the case that the rational players described by $\mathfrak{U}$ must be more agnostic or open-minded than the rational players described by $\mathfrak{T}$.

Alternate routes to an epistemic condition may also exist. We might consider one of two complementary approaches. The most direct path of attack would be to ask whether there exists a single type structure—perhaps an analog of the universal type structure—in which the IA set of every game is the output of RCAR. The second option is to weaken the criteria for assumption so that they are not sensitive to variations in topological distinguishability. However, we do not want to weaken assumption so much that we no longer get IA as an output of RCAR.

Analyzing the complete lexicographic type structures constructed in this paper may provide some clues about how to achieve these goals. Roughly speaking, beliefs that manifest the so-called *Best Rationalization Principle*, which was articulated in Battigalli (1996), also satisfy common assumption of rationality in our constructions. In other words, if Ann attributes each admissible choice s^b of Bob to a rational decision based on the highest order mutual assumption of rationality that is consistent with it, then her beliefs satisfy common assumption of rationality. Ann, if her beliefs reflect the Best Rationalization Principle, can be viewed as assigning ex ante explanations of all possible actions of Bob—explanations that preserve as much higher order assumption of rationality as possible. Yang (2010)'s notion of *weak assumption* seems well-suited to expressing this particular idea in lexicographic type structures.

Finally, we note that an epistemic characterization of IA may be achieved by eschewing open-mindedness as the generator of admissible choice altogether. Barelli and Galanis (2010) give an alternative epistemic condition for IA that is built on *extended event-rationality*, which, like LEU, is an extension of the standard model of Bayesian rationality. Tie-breaking behavior, not agnosticism, delivers admissible choice in Barelli and Galanis (2010). As might be expected, the above interpretation of topological properties does not carry over to their framework.

Appendix A

Rationalizing Payoff-Dominant Outcomes

A.1 Preliminaries

The proof of Theorem 1.4.1 is given in several steps, each concerning an extra iteration in the construction of Σ^{∞}. The sets of strategies consistent with any non-terminal history h are defined below for the case when we have only two players.

$$S_1(h) = \begin{cases} S_1 & \text{if } h = \varnothing \\ \{\, s_1 |\, s_1 \in S_1 \wedge s_1(\varnothing) = C_1\} & \text{if } h = C_1 \subseteq D_1 \text{ and } h \neq \varnothing \\ S_1(C_1) & \text{if } h = C_1 \times C_2 \subseteq D \text{ and } h \neq \varnothing \end{cases}$$

$$S_2(h) = \begin{cases} S_2 & \text{if } h = \varnothing \text{ or } h \subseteq D_1 \\ \{\, s_2 |\, s_2 \in S_2 \wedge s_2(C_1) = C_2\} & \text{if } h = C_1 \times C_2 \subseteq D \text{ and } h \neq \varnothing \end{cases}$$

Given any $C_1 \times C_2 = C \in \mathscr{H} \cap D$ and any $(s_1, s_2) = s \in S$, the following is true by the definition of strategies.

$$s_1(C) \in C_1 \wedge s_2(C) \in C_2$$

One consequence of Definition 1.3.5 is the following:

$$(\tilde{s}_i \in \Sigma_i^\infty) \implies (\tilde{s}_i \in \Sigma_i^\infty) \wedge \exists \mu^i \Bigg[\left(\mu^i \in \mathrm{CPS}_i\right) \wedge \left(\tilde{s}_i \in r_i\left(\mu^i\right)\right)$$
$$\wedge \forall h \Big[\Big(\Big(\tilde{h} \in \mathscr{H}\Big) \wedge \left(\Sigma_{-i}^\infty \cap S_{-i}(h) \neq \varnothing\right) \Big) \implies \left(\mu^i(\Sigma_{-i}^\infty | h) = 1\right) \Big] \Bigg]$$

A.2 Concerning Σ_1^1

We begin with the following facts, which are consequences of earlier definitions.

$$(\tilde{s}_1 \in \Sigma_1^\infty) \implies \exists \mu^1 \left[\left(\mu^1 \in \mathrm{CPS}_1\right) \wedge \left(\tilde{s}_1 \in r_1\left(\mu^1\right)\right) \right]$$
$$\tilde{s}_1 \in r_1\left(\mu^1\right) \equiv \forall h \left[((h \in \mathscr{H}) \wedge (\tilde{s}_1 \in S_1(h))) \implies \tilde{s}_1 \in \underset{s_1 \in S_1(h)}{\operatorname{argmax}}\, U_1(s_1, \mu^1(\cdot|h)) \right]$$

Consider a non-terminal history $\tilde{h}$ satisfying the following formulas.

$$\Big(\tilde{h} \in \mathscr{H}\Big) \wedge \Big(\tilde{h} = C_1 \times C_2\Big) \wedge (C_1 \cap B_1 \neq \varnothing) \wedge (C_2 \subseteq B_2) \wedge \Big(\tilde{s}_1 \in \Sigma_1^\infty \cap S_1(\tilde{h})\Big)$$

Since $\tilde{h}$ is in $\mathscr{H}$, the following statement must hold.

$$\left(\tilde{s}_1 \in r_1\left(\mu^1\right)\right) \implies \left[\Big(\tilde{s}_1 \in S_1(\tilde{h})\Big) \implies \tilde{s}_1 \in \underset{s_1 \in S_1(\tilde{h})}{\operatorname{argmax}}\, U_1(s_1, \mu^1(\cdot|\tilde{h})) \right]$$

$$\left[(\tilde{s}_1 \in \Sigma_1^\infty) \wedge \left(\tilde{s}_1 \in S_1(\tilde{h})\right)\right] \implies \left[\tilde{s}_1 \in \operatorname*{argmax}_{s_1 \in S_1(\tilde{h})} U_1(s_1, \mu^1(\cdot|\tilde{h}))\right]$$

Given that we know $s_1 \in S_1(\tilde{h})$, we may expand the utility formula in the following way since $\tilde{h} \subseteq D$ (i.e., only $G(\tilde{h})$ remains to be played).

$$U_1(s_1, \mu^1(\cdot|\tilde{h})) = \sum_{s_2 \in S_2(\tilde{h})} u_1(s_1(\tilde{h}), s_2(\tilde{h}))\mu^1(s_2|\tilde{h})$$

The structure of our extensive form game requires that $s_2(\tilde{h}) \in C_2$ and that $s_1 \in S_1(\tilde{h})$ implies $s_1(\tilde{h}) \in C_1$. Therefore, we may assume the first and second inequalities below. The third inequality follows from the Pareto properties of B vis-à-vis $D \setminus B$ that were assumed in the general setup. The last inequality follows from the relation between the max and min operators.

$$\begin{aligned} U_1(\tilde{s}_1, \mu^1(\cdot|\tilde{h})) &= \sum_{s_2 \in S_2(\tilde{h})} u_1(\tilde{s}_1(\tilde{h}), s_2(\tilde{h}))\mu^1(s_2|\tilde{h}) \\ &\geq \max_{s_1 \in S_1(\tilde{h})} \min_{d_2 \in C_2} u_1(s_1(\tilde{h}), d_2) \\ &\geq \max_{d_1 \in C_1} \min_{d_2 \in B_2} u_1(d_1, d_2) \\ &\geq \max_{d_1 \in B_1} \min_{d_2 \in B_2} u_1(d_1, d_2) \\ &\geq \min_{d_1 \in B_1} \min_{d_2 \in B_2} u_1(d_1, d_2) \end{aligned}$$

The Pareto properties of B vis-à-vis $D \setminus B$ that were assumed in the general setup imply that $\tilde{s}_1(\tilde{h}) \in B_1$. Otherwise, we obtain the following contradiction.

$$\begin{aligned} U_1(\tilde{s}_1, \mu^1(\cdot|\tilde{h})) &= \sum_{s_2 \in S_2(\tilde{h})} u_1(\tilde{s}_1(\tilde{h}), s_2(\tilde{h}))\mu^1(s_2|\tilde{h}) \\ &\leq \max_{d_1 \in D_1 \setminus B_1} \max_{d_2 \in B_2} u_1(d_1, d_2) < \min_{d_1 \in B_1} \min_{d_2 \in B_2} u_1(d_1, d_2) \leq U_1(\tilde{s}_1, \mu^1(\cdot|\tilde{h})) \end{aligned}$$

A.2.1 Summary

Therefore, we have proved the following.

$$\begin{aligned} &\left[\left(\tilde{h} \in \mathscr{H}\right) \wedge \left(\tilde{h} = C_1 \times C_2\right) \wedge (C_1 \cap B_1 \neq \varnothing) \wedge (C_2 \subseteq B_2) \wedge (\tilde{s}_1 \in \Sigma_1^\infty \cap S_1(C_1))\right] \\ &\qquad \implies \left[\tilde{s}_1(\tilde{h}) \in B_1\right] \end{aligned}$$

A.3 Concerning Σ_2^1

Again, we begin with the following facts, which are consequences of earlier definitions.

$$\begin{aligned} (\tilde{s}_2 \in \Sigma_2^\infty) &\implies \exists \mu^2 \left[\left(\mu^2 \in \mathrm{CPS}_2\right) \wedge \left(\tilde{s}_2 \in r_2\left(\mu^2\right)\right)\right] \\ \tilde{s}_2 \in r_2\left(\mu^2\right) &\equiv \forall h \left[((h \in \mathscr{H}) \wedge (\tilde{s}_2 \in S_2(h))) \implies \tilde{s}_2 \in \underset{s_2 \in S_2(h)}{\operatorname{argmax}}\, U_2(s_2, \mu^2(\cdot|h))\right] \end{aligned}$$

Consider a non-terminal history $\tilde{h}$ satisfying the following formulas.

$$\left(\tilde{h} \in \mathscr{H}\right) \wedge \left(\tilde{h} = C_1\right) \wedge (C_1 \subseteq B_1)$$

Since $\tilde{h}$ is in $\mathscr{H}$ and $S_2(\tilde{h}) = S_2(C_1) = S_2$, the following statement must hold.

$$\left(\tilde{s}_2 \in r_2\left(\mu^2\right)\right) \implies \left[\tilde{s}_2 \in \operatorname*{argmax}_{s_2 \in S_2(\tilde{h})} U_2(s_2, \mu^2(\cdot|\tilde{h}))\right]$$

Given that we know $s_2 \in S_2(\tilde{h})$, we may expand the utility formula in the following way.

$$U_2(s_2, \mu^2(\cdot|\tilde{h})) = \sum_{s_1 \in S_1(\tilde{h})} u_2(s_1(\tilde{h} \times s_2(\tilde{h})), s_2(\tilde{h} \times s_2(\tilde{h})))\mu^2(s_1|\tilde{h})$$

$$\begin{aligned}
U_2(\tilde{s}_2, \mu^2(\cdot|\tilde{h})) &= \sum_{s_1 \in S_1(\tilde{h})} u_2(s_1(\tilde{h} \times s_2(\tilde{h})), \tilde{s}_2(\tilde{h} \times s_2(\tilde{h})))\mu^2(s_1|\tilde{h}) \\
&\geq \max_{s_2 \in S_2(\tilde{h})} \min_{d_1 \in C_1} u_2(d_1, s_2(\tilde{h} \times s_2(\tilde{h}))) \\
&\geq \max_{s_2 \in S_2} \min_{d_1 \in B_1} u_2(d_1, s_2(\tilde{h} \times s_2(\tilde{h}))) \\
&\geq \max_{d_2 \in D_2} \min_{d_1 \in B_1} u_1(d_1, d_2) \\
&\geq \min_{d_2 \in B_2} \min_{d_2 \in B_2} u_1(d_1, d_2)
\end{aligned}$$

The Pareto properties of B vis-à-vis $D \setminus B$ that were assumed in the general setup imply that $\tilde{s}_2(\tilde{h} \times \tilde{s}_2(\tilde{h})) \in B_2$. Otherwise, we obtain have the following contradiction:

$$\begin{aligned}
U_2(\tilde{s}_2, \mu^2(\cdot|\tilde{h})) &= \sum_{s_1 \in S_1(\tilde{h})} u_2(s_1(\tilde{h} \times s_2(\tilde{h})), \tilde{s}_2(\tilde{h} \times s_2(\tilde{h})))\mu^2(s_1|\tilde{h}) \\
&\leq \max_{d_2 \in D_2 \setminus B_2} \max_{d_1 \in B_1} u_1(d_1, d_2) < \min_{d_2 \in B_2} \min_{d_1 \in B_1} u_2(d_1, d_2) \leq U_2(\tilde{s}_2, \mu^2(\cdot|\tilde{h}))
\end{aligned}$$

A.3.1 Summary

Therefore, we have proved the following:

$$\left[\left(\tilde{h} \in \mathscr{H}\right) \wedge \left(\tilde{h} = C_1\right) \wedge (C_1 \subseteq B_1) \wedge (\tilde{s}_2 \in \Sigma_2^{\infty})\right] \implies \left[\tilde{s}_2(\tilde{h} \times \tilde{s}_2(\tilde{h})) \in B_2\right]$$

A.4 Concerning Σ_2^2

The following is a consequence of earlier definitions.

$$(\tilde{s}_2 \in \Sigma_2^{\infty}) \implies \exists \mu^2 \Bigg[\left(\mu^2 \in \mathrm{CPS}_2\right) \wedge \left(\tilde{s}_2 \in r_2\left(\mu^2\right)\right) \\ \wedge \forall h \Big[((h \in \mathscr{H}) \wedge (\Sigma_1^{\infty} \cap S_1(h) \neq \varnothing)) \implies \left(\mu^2(\Sigma_1^{\infty}|h) = 1\right)\Big]\Bigg]$$

Here, we repeat the definition of sequential rationality of Player 2.

$$\tilde{s}_2 \in r_2\left(\mu^2\right) \equiv \forall h \left[((h \in \mathscr{H}) \wedge (\tilde{s}_2 \in S_2(h))) \implies \tilde{s}_2 \in \operatorname*{argmax}_{s_2 \in S_2(h)} U_2(s_2, \mu^2(\cdot|h))\right]$$

Consider a non-terminal history $\tilde{h}$ satisfying the following formulas.

$$\left(\tilde{h} \in \mathscr{H}\right) \wedge \left(\tilde{h} = C_1\right) \wedge (C_1 \cap B_1 \neq \varnothing) \wedge \left(\Sigma_1^{\infty} \cap S_1(\tilde{h}) \neq \varnothing\right)$$

Since $\tilde{h}$ is in $\mathscr{H}$, the following statement must hold.

$$\left(\tilde{s}_2 \in r_2\left(\mu^2\right)\right) \implies \left[\left(\tilde{s}_2 \in S_2(\tilde{h})\right) \implies \tilde{s}_2 \in \operatorname*{argmax}_{s_2 \in S_2(\tilde{h})} U_2(s_2, \mu^2(\cdot|\tilde{h}))\right]$$

$$\left[(\tilde{s}_2 \in \Sigma_2^\infty) \wedge \left(\tilde{s}_2 \in S_2(\tilde{h})\right)\right] \implies \left[\tilde{s}_2 \in \operatorname*{argmax}_{s_2 \in S_2(\tilde{h})} U_2(s_2, \mu^2(\cdot|\tilde{h}))\right]$$

Given that we know $s_2 \in S_2(C_1)$, we may expand the utility formula in the following way.

$$U_2(s_2, \mu^2(\cdot|\tilde{h})) = \sum_{s_1 \in S_1(\tilde{h})} u_2(s_1(\tilde{h} \times s_2(\tilde{h})), s_2(\tilde{h} \times s_2(\tilde{h})))\mu^2(s_1|\tilde{h})$$

We may rewrite the utility as

$$U_2(s_2, \mu^2(\cdot|\tilde{h})) = \sum_{s_1 \in \Sigma_1^\infty \cap S_1(\tilde{h})} u_2(s_1(\tilde{h} \times s_2(\tilde{h})), s_2(\tilde{h} \times s_2(\tilde{h})))\mu^2(s_1|\tilde{h})$$

since

$$\left(\Sigma_1^\infty \cap S_1(\tilde{h}) \neq \varnothing\right) \wedge \forall h\left[((h \in \mathscr{H}) \wedge (\Sigma_1^\infty \cap S_1(h) \neq \varnothing)) \implies \left(\mu^2(\Sigma_1^\infty|h) = 1\right)\right].$$

Consider a result that we have previously obtained, which is restated below.

$$\left[\left(\tilde{h}' \in \mathscr{H}\right) \wedge \left(\tilde{h}' = C_1' \times C_2'\right) \wedge (C_1' \cap B_1 \neq \varnothing) \wedge (C_2' \subseteq B_2) \wedge (s_1 \in \Sigma_1^\infty \cap S_1(C_1'))\right] \implies \left[s_1(\tilde{h}') \in B_1\right]$$

Applying it to $C_1' = \tilde{h}$ and supposing that $C_2' = s_2(\tilde{h})$, we get the following.

$$\left[\left(\tilde{h} \times s_2(\tilde{h}) \in \mathscr{H}\right) \wedge \left(s_2(\tilde{h}) \subseteq B_2\right) \wedge \left(s_1 \in \Sigma_1^\infty \cap S_1(\tilde{h})\right)\right] \implies \left[s_1(\tilde{h} \times s_2(\tilde{h}) \in \mathscr{H}) \in B_1\right]$$

The optimality of $\tilde{s}_2$ implies that

$$\begin{aligned} &U_2(\tilde{s}_2, \mu^2(\cdot|\tilde{h})) \geq \\ &\qquad U_2(s_2, \mu^2(\cdot|\tilde{h})) = \sum_{s_1 \in \Sigma_1^\infty \cap S_1(\tilde{h})} u_2(s_1(\tilde{h} \times s_2(\tilde{h})), s_2(\tilde{h} \times s_2(\tilde{h})))\mu^2(s_1|\tilde{h}) \\ &\qquad\qquad \geq \min_{d \in B} u_2(d). \end{aligned}$$

The Pareto properties of B vis-à-vis $D \setminus B$ that were assumed in the general setup imply that $\tilde{s}_2(\tilde{h} \times \tilde{s}_2(\tilde{h})) \in B_2$. Otherwise, we have the following contradiction.

$$\begin{aligned} U_2(\tilde{s}_2, \mu^2(\cdot|\tilde{h})) &= \sum_{s_1 \in S_1(\tilde{h})} u_2(s_1(\tilde{h} \times s_2(\tilde{h})), \tilde{s}_2(\tilde{h} \times s_2(\tilde{h})))\mu^2(s_1|\tilde{h}) \\ &\leq \max_{d_2 \in D_2 \setminus B_2} \max_{d_1 \in D_1} u_1(d_1, d_2) < \min_{d \in B} u_2(d) \leq U_2(\tilde{s}_2, \mu^2(\cdot|\tilde{h})) \end{aligned}$$

A.4.1 Summary

Therefore, we have proved the following:

$$\begin{aligned} &\left[\left(\tilde{h} \in \mathscr{H}\right) \wedge \left(\tilde{h} = C_1\right) \wedge (C_1 \cap B_1 \neq \varnothing) \wedge \left(\Sigma_1^\infty \cap S_1(\tilde{h}) \neq \varnothing\right) \wedge (\tilde{s}_2 \in \Sigma_2^\infty)\right] \\ &\qquad\qquad \Longrightarrow \left[\tilde{s}_2(\tilde{h} \times \tilde{s}_2(\tilde{h})) \in B_2\right] \end{aligned}$$

A.5 Concerning Σ_1^3

The following is a consequence of earlier definitions.

$$(\tilde{s}_1 \in \Sigma_1^\infty) \implies \exists \mu^1 \Big[\big(\mu^1 \in \mathrm{CPS}_1\big) \wedge \big(\tilde{s}_1 \in r_1\left(\mu^1\right)\big)$$
$$\wedge \forall h \big[((h \in \mathscr{H}) \wedge (\Sigma_2^\infty \cap S_2(h) \neq \varnothing)) \implies \big(\mu^1(\Sigma_2^\infty | h) = 1\big) \big] \Big]$$

Here, we repeat the definition of sequential rationality of Player 1.

$$\tilde{s}_1 \in r_1\left(\mu^1\right) \equiv \forall h \left[((h \in \mathscr{H}) \wedge (\tilde{s}_1 \in S_1(h))) \implies \tilde{s}_1 \in \operatorname*{argmax}_{s_1 \in S_1(h)} U_1(s_1, \mu^1(\cdot|h)) \right]$$

Consider a non-terminal history $\tilde{h} \in H$ satisfying the following formulas.

$$\left(\tilde{h} = C_1 \times C_2 = C\right) \wedge (C \cap B \neq \varnothing) \wedge \left(\Sigma_2^\infty \cap S_2(\tilde{h}) \neq \varnothing\right) \wedge \left(\tilde{s}_1 \in \Sigma_1^\infty \cap S_1(\tilde{h})\right)$$

Since $S_2(\tilde{h}) \subset S_2(C_1) = S_2$ and $S_1(C_1) = S_1(\tilde{h})$, the results of the previous section imply that

$$(s_2 \in \Sigma_2^\infty) \implies [s_2(C_1 \times s_2(C_1)) \in B_2]\,.$$

Given that we know $s_1 \in S_1(\tilde{h})$, we may expand the utility formula in the following way.

$$U_1(s_1, \mu^1(\cdot|\tilde{h})) = \sum_{s_2 \in S_2(\tilde{h})} u_1(s_1(\tilde{h}), s_2(\tilde{h}))\mu^1(s_2|\tilde{h})$$

We may rewrite the utility as

$$U_1(s_1, \mu^1(\cdot|\tilde{h})) = \sum_{s_2 \in \Sigma_2^\infty \cap S_2(\tilde{h})} u_1(s_1(\tilde{h}), s_2(\tilde{h}))\mu^1(s_1|\tilde{h})$$

since

$$\left(\Sigma_2^\infty \cap S_2(\tilde{h}) \neq \varnothing\right) \wedge \forall h\big[((h \in \mathscr{H}) \wedge (\Sigma_2^\infty \cap S_2(h) \neq \varnothing)) \implies \big(\mu^1(\Sigma_2^\infty|h) = 1\big)\big].$$

Since we know that $(s_2 \in \Sigma_2^\infty) \implies [s_2(C_1 \times s_2(C_1)) \in B_2]$ consider some $s_1 \in S_1(\tilde{h})$ that satisfies $s_1(\tilde{h}) \in C_1 \cap B_1$. Therefore, the optimality of $\tilde{s}_1$ implies

$$\begin{aligned} U_1(\tilde{s}_1, \mu^1(\cdot|\tilde{h})) &\geq \sum_{s_2 \in S_2(\tilde{h})} u_1(s_1(\tilde{h}), s_2(\tilde{h}))\mu^1(s_2|\tilde{h}) \\ &\geq \min_{d \in B} u_1(d) \end{aligned}$$

The Pareto properties of B vis-à-vis $D \setminus B$ that were assumed in the general setup imply that $\tilde{s}_1(\tilde{h}) \in B_1$. Otherwise, we obtain the following contradiction.

$$\begin{aligned} U_1(\tilde{s}_1, \mu^1(\cdot|\tilde{h})) &= \sum_{s_2 \in S_2(\tilde{h})} u_1(\tilde{s}_1(\tilde{h}), s_2(\tilde{h}))\mu^1(s_2|\tilde{h}) \\ &\leq \max_{d_1 \in D_1 \setminus B_1} \max_{d_2 \in B_2} u_1(d_1, d_2) < \min_{d \in B} u_1(d) \leq U_1(\tilde{s}_1, \mu^1(\cdot|\tilde{h})) \end{aligned}$$

A.5.1 Summary

Therefore, we have proved the following.

$$\left(\tilde{h} \in \mathscr{H} \wedge \tilde{h} = C_1 \times C_2 = C \wedge C \cap B \neq \varnothing \wedge \Sigma_2^\infty \cap S_2(\tilde{h}) \neq \varnothing \wedge \tilde{s}_1 \in \Sigma_1^\infty \cap S_1(\tilde{h})\right) \implies \left[\tilde{s}_1(\tilde{h}) \in B_1\right]$$

A.6 Concerning Σ_1^2

From the previous results, we know that the following statement holds.

$$\left[\left(\tilde{h} \in \mathscr{H}\right) \wedge \left(\tilde{h} = C_1\right) \wedge (C_1 \subseteq B_1) \wedge (\tilde{s}_2 \in \Sigma_2^\infty)\right] \implies \left[\tilde{s}_2(\tilde{h} \times \tilde{s}_2(\tilde{h})) \in B_2\right]$$

Suppose that $\tilde{s}_1 \in \Sigma_1^\infty$ and $\tilde{s}_1(\varnothing) \cap B_1 = \varnothing$. Let μ^1 be the CPS that rationalizes $\tilde{s}_1$. Note that $S_2(\tilde{h}) = S_2$. If $s_1(\varnothing) = B_1$ then

$$\begin{aligned}\min_{d \in B} u_1(d) > U_1(\tilde{s}_1, \mu^1(\cdot|\tilde{h})) &\geq \sum_{s_2 \in \Sigma_2^\infty} u_1(s_1(\tilde{h} \times s_2(\tilde{h})), s_2(\tilde{h} \times s_2(\tilde{h})))\mu^1(s_1|\tilde{h}) \\ &\geq \min_{d \in B} u_1(d)\end{aligned}$$

since

$$\left(\Sigma_2^\infty \cap S_2(\tilde{h}) = \Sigma_2^\infty \cap S_2 \neq \varnothing\right) \wedge \forall h \left[(h \in \mathscr{H} \wedge (\Sigma_2^\infty \cap S_2(h) \neq \varnothing)) \implies \mu^1(\Sigma_2^\infty|h) = 1\right].$$

Due to the contradiction, it must be the case that $\tilde{s}_1 \in \Sigma_1^\infty \implies \tilde{s}_1(\varnothing) \cap B_1 \neq \varnothing$.

A.7 About Σ^∞: Putting Things Together

Suppose $s \in \Sigma^\infty$. Let $C = C_1 \times C_2$ and

$$s_1(\varnothing) = C_1 \qquad s_2(C_1) = C_2$$
$$s_1(C) = d_1 \qquad s_2(C) = d_2.$$

Since $\tilde{s}_1 \in \Sigma_1^\infty \implies \tilde{s}_1(\varnothing) \cap B_1 \neq \varnothing$, it is the case that $C_1 \cap B_1 \neq \varnothing$. Therefore,

$$\left[\left(\tilde{h} \in \mathscr{H}\right) \wedge \left(\tilde{h} = C_1\right) \wedge (C_1 \cap B_1 \neq \varnothing) \wedge \left(\Sigma_1^\infty \cap S_1(\tilde{h}) \neq \varnothing\right) \wedge (\tilde{s}_2 \in \Sigma_2^\infty)\right] \implies \left[\tilde{s}_2(\tilde{h} \times \tilde{s}_2(\tilde{h})) \in B_2\right].$$

implies $\tilde{s}_2(C) \in B_2$, which in turn implies that $\tilde{s}_2(C_1) \cap B_2 \neq \varnothing$. Therefore, for all $\tilde{h} \in \mathscr{H}$, we have

$$\left(\tilde{h} = C_1 \times C_2 = C\right) \wedge (C \cap B \neq \varnothing) \wedge \left(\Sigma_2^\infty \cap S_2(\tilde{h}) \neq \varnothing\right) \wedge \left(\tilde{s}_1 \in \Sigma_1^\infty \cap S_1(\tilde{h})\right) \implies \left[\tilde{s}_1(\tilde{h}) \in B_1\right].$$

This implies $\tilde{s}_1(C) \in B_1$. Therefore, we have our first result.

A.8 Proof of Theorem 1.4.2

Proof of Theorem 1.4.2. Since $s \in \Sigma^\infty$, we know the following via the previous result.

$$d_1 = b_1 \qquad d_2 = b_2 \qquad C \cap B \neq \varnothing$$

Consider any number $k \in \mathbb{N}$. Since $\Sigma^0 = S$, we know that $\exists k \left[s' \in \Sigma^{k-1}\right]$. Now, we construct beliefs ν^1 and ν^2 that will allow s' to survive into the next round of elimination. First, note the definition of membership in Σ^k.

$$\left(\tilde{s}_i \in \Sigma_i^k\right) \equiv \left(\tilde{s}_i \in \Sigma_i^{k-1}\right) \wedge \exists \mu^i \Bigg[\left(\mu^i \in \text{CPS}_i\right) \wedge \left(\tilde{s}_i \in r_i\left(\mu^i\right)\right)$$
$$\wedge \forall h \Big[\Big(\Big(\tilde{h} \in \mathscr{H}\Big) \wedge \left(\Sigma_{-i}^{k-1} \cap S_{-i}(h) \neq \varnothing\right) \Big) \implies \left(\mu^i(\Sigma_{-i}^{k-1}|h) = 1\right) \Big] \Bigg]$$

Since $s \in \Sigma^\infty$, we know that there exist CPSs μ^1 and μ^2 that satisfy the above requirement in all rounds (arbitrarily high k) for s_1 and s_2, respectively. Now define ν^1 and ν^2 thusly:

$$\nu^1(s_2'|\varnothing) = \nu^1(s_2'|W_1) = \nu^1(s_2'|W) = 1$$
$$\nu^2(s_1'|\varnothing) = \nu^2(s_1'|W_1) = \nu^2(s_1'|W) = 1$$
$$\forall i \forall Z \left[(Z \neq \varnothing \wedge Z \neq W_1 \wedge Z \neq W) \implies \nu^i(\cdot|Z) = \mu^i(\cdot|Z)\right]$$

Note that the first two lines above guarantee strong belief restriction on CPSs for histories $\varnothing, W_1, W$. The sequential rationality requirements $s_1' \in r_1(\nu^1)$ and $s_2' \in r_2(\nu^2)$ are trivially satisfied at $\varnothing, W_1, W$ since the beliefs at those histories assign probability 1 to the best outcome b.

Borrowing from μ^1 and μ^2 for other histories, as done in the third line, guarantees that the strong belief property is satisfied at those histories. This is by assumption since μ^1 and μ^2 correspond to $s \in \Sigma^\infty$. The sequential rationality requirements $s_1' \in r_1(\nu^1)$ and $s_2' \in r_2(\nu^2)$ are trivially satisfied at those other histories by the same reasoning. □

A.9 Proof of Theorem 1.4.7

Lemma A.9.1. *Let* $Y = \prod_{i=1}^{n} Y_i \subseteq S$. *Assume the following: For any history* $\hat{h}$ *and for any player* k, *if* $S_k(\hat{h}) = S_k$ *then the following holds for all* $s_k^{\#}, s_k^{\#\#} \in Y_k$.

$$s_k^{\#}(\hat{h}) \neq s_k^{\#\#}(\hat{h}) \implies \exists s_k''' \in Y_k \ s.t. \ s_k'''(h) = \begin{cases} s_k^{\#}(h) & if\ h \succeq \hat{h} \\ s_k^{\#\#}(h) & otherwise \end{cases}$$

where $h \succeq \hat{h}$ *is defined as* $\hat{h}$ *is a prefix of* h, *i.e.,* h *is a possible continuation of* $\hat{h}$.

Then, for any history $\hat{h}$ *and for any player* k, *if* $S_k(\hat{h}) = S_k$ *then the following holds for all* $s_k', s_k'' \in A_k(Y)$.

$$s_k'(\hat{h}) \neq s_k''(\hat{h}) \implies \exists s_k''' \in A_k(Y) \ s.t. \ s_k'''(h) = \begin{cases} s_k'(h) & if\ h \succeq \hat{h} \\ s_k''(h) & otherwise \end{cases}$$

Proof of Lemma A.9.1. Since $s_k', s_k'' \in A_k(Y) \subseteq Y_k$, there exist, by assumption, s_k^* such that:

$$s_k^*(h) = \begin{cases} s_k'(h) & \text{if } h \succeq \hat{h} \\ s_k''(h) & \text{otherwise} \end{cases}$$

We need to show that $s_k^*, s_k^{**} \in A_k(Y)$. Suppose, by way of contradiction, that $s_k^* \notin A_k(Y)$. Then, there must exist some $\hat{s}_k \in A_k(Y)$ such that

$$u_k(\hat{s}_k, s_{-k}) \geq u_k(s_k^*, s_{-k}) \qquad \text{for all } s_{-k} \in Y_{-k}$$

$$u_k(\hat{s}_k, \hat{s}_{-k}) > u_k(s_k^*, \hat{s}_{-k}) \qquad \text{for some } \hat{s}_{-k} \in Y_{-k}$$

Now, consider two cases. The first case is when $\hat{s}_{-k} \in S_{-k}(\hat{h})$. Then, $\hat{s}_k$ weakly dominates s'_k *against* $S_{-k}(\hat{h})$ since s^*_k, by definition, does equally well as s'_k against opponents playing strategies consistent with $\hat{h}$. By applying the assumption of the lemma once more, we know that there exists $\hat{s}'_k \in Y_k$ such that

$$\hat{s}'_k(h) = \begin{cases} \hat{s}_k(h) & \text{if } h \succeq \hat{h} \\ s'_k(h) & \text{otherwise} \end{cases}$$

It is evident that $\hat{s}'_k$ weakly dominates s'_k, which yields the desired contradiction of the original assumption that $s'_k \in A_k(Y)$.

Now, consider the second case, when $\hat{s}_{-k} \notin S_{-k}(\hat{h})$. Then, $\hat{s}_k$ weakly dominates s''_k *against* $S_{-k} \setminus S_{-k}(\hat{h})$ since s^*_k, by definition, does equally well as s''_k against opponents playing strategies not consistent with $\hat{h}$. By applying the assumption of the lemma once more, we know that there exists $\hat{s}''_k \in Y_k$ such that

$$\hat{s}''_k(h) = \begin{cases} \hat{s}_k(h) & \text{if } h \not\succeq \hat{h} \\ s''_k(h) & \text{otherwise} \end{cases}$$

It is evident that $\hat{s}''_k$ weakly dominates s''_k, which yields the desired contradiction of the original assumption that $s''_k \in A_k(Y)$.

Since we obtain contradictions for the two possible cases, we have proven the desired result. □

Proof of Lemma 1.4.5. The proof is by induction. The initial condition is satisfied by setting $Y = S$ in Lemma A.9.1. The induction step is proven by Lemma A.9.1. □

Proof of Theorem 1.4.7. Part 1: For a given m, suppose

$$\exists k\left[A^m(S)\cap S(B_{\leq k})\neq\varnothing\wedge\forall s\left[s\in A^m(S)\cap S(B_{\leq k})\implies d(s)=b\right]\right].$$

We want to show that if

$$\exists\ell\left[\left(\ell\leq k\wedge A^{m+1}(S)\cap S(B_{\leq\ell})=\varnothing\right)\right]$$

then

$$\forall s\left[s\in\left[A_\ell^{m+1}(S)\times A_{-\ell}^m(S)\right]\cap S(B_{<\ell})\implies d(s)=b\right].$$

Suppose by way of contradiction that

$$\exists\ell\Big[\left[\left(\ell\leq k\wedge A^{m+1}(S)\cap S(B_{\leq\ell})=\varnothing\right)\right]\\ \wedge\exists s''\left[s''\in\left[A_\ell^{m+1}(S)\times A_{-\ell}^m(S)\right]\cap S(B_{<\ell})\wedge d(s)\neq b\right]\Big].$$

Let ℓ be the smallest number that satisfies the above requirements. Now, we consider the following fact.

$$\exists s'\left[s'\in A^m(S)\cap S(B_{\leq\ell})\wedge d(s')=b\right]$$

and all such s'_ℓ must be weakly dominated by some $s_\ell\in A_\ell^m(S)\setminus S_\ell(B_{\leq\ell})$ against $A_{-\ell}^m(S)$ since $A^m(S)\cap S(B_{\leq\ell})\neq\varnothing$ and $A^{m+1}(S)\cap S(B_{\leq\ell})=\varnothing$. Therefore,

$$u_\ell(s'_\ell,s'_{-\ell})=u_\ell(s_\ell,s'_{-\ell}).$$

However, we can construct $s'''_{-\ell}$ in the following way.

$$s'''_{-\ell}(h) = \begin{cases} s'_{-\ell}(h) & \text{if } h \succeq B_{\leq \ell} \vee h \preceq B_{\leq \ell} \\ s''_{-\ell} & \text{otherwise} \end{cases}$$

Via Lemma 1.4.5, we conclude that $s'''_{-\ell} \in A^m_{-\ell}(S) \cap S_{-\ell}(B_{<\ell})$. Via the weak dominance of s_ℓ, we obtain the following.

$$[u_\ell(s') =] \quad u_\ell(s'_\ell, s'''_{-\ell}) = u_\ell(s_\ell, s'''_{-\ell}) \quad \left[= u_\ell(s_\ell, s''_{-\ell})\right].$$

However,

$$u_\ell(s_\ell, s'''_{-\ell}) = u_\ell(s_\ell, s''_{-\ell}) \neq u_\ell(b) \quad \therefore u_\ell(s_\ell, s''_{-\ell}) < u_\ell(b) = u_\ell(s') = u_\ell(s'_\ell, s'''_{-\ell}),$$

which gives us a contradiction of our assumptions. To recap, we have proven that if, given m,

$$\exists k \left[A^m(S) \cap S(B_{\leq k}) \neq \varnothing \wedge \forall s \left[s \in A^m(S) \cap S(B_{\leq k}) \implies d(s) = b\right]\right],$$

then

$$\begin{aligned} &\exists \ell \left[\left(\ell \leq k \wedge A^{m+1}(S) \cap S(B_{\leq \ell}) = \varnothing\right)\right] \\ &\qquad \implies \forall s \left[s \in \left[A^{m+1}_\ell(S) \times A^m_{-\ell}(S)\right] \cap S(B_{<\ell}) \implies d(s) = b\right] \\ &\qquad\qquad \implies \forall s \left[s \in A^{m+1}(S) \cap S(B_{<\ell}) \implies d(s) = b\right]. \end{aligned}$$

Part 2: Suppose that for some (m, k), $\forall s \left[s \in A^m(S) \cap S(B_{<k}) \implies d(s) = b\right]$. We want to show that $\forall s \left[s \in A^{m+1}(S) \cap S(B_{<k-1}) \implies d(s) = b\right]$. Suppose by way of

contradiction that there exists s such that $s \in A^{m+1}(S) \cap S(B_{<k-1}) \wedge d(s) \neq b$. Then,

$$\left[s \in A^{m+1}(S) \cap S(B_{<k-1}) \wedge d(s) \neq b\right] \implies A^{m+1} \cap S(B_{<k}) = \varnothing$$

since $A^{m+1} \cap S(B_{<k}) \subseteq A^m \cap S(B_{<k})$; otherwise, $s \in A^{m+1}(S) \cap S(B_{<k-1})$ implies $d(s) = b$. The result of Part 1 then implies that there does not exist s such that $s \in A^{m+1}(S) \cap S(B_{<k-1}) \wedge d(s) \neq b$.

$$\therefore \forall s \left[s \in A^m(S) \cap S(B_{<k}) \implies d(s) = b\right] \implies \forall s \left[s \in A^{m+1}(S) \cap S(B_{<k-1}) \implies d(s) = b\right]$$

Part 3: The following initial step is true by definition.

$$\forall s \left[s \in A^0(S) \cap S(B_{<n}) \implies d(s) = b\right]$$

Therefore, the induction step that was proven in Part 2 gives us the following.

$$\exists m \forall (k, s) \left[s \in A^m(S) \cap S(B_{<k}) \implies d(s) = b\right] \quad \therefore \exists m \forall s \left[s \in A^m(S) \cap S(B_{<1}) \implies d(s) = b\right]$$

Our conclusion is then proven by a series of implications.

- Since $S(B_{<1}) = S$, we have $\exists m \forall s \left[s \in A^m(S) \implies d(s) = b\right]$.
- Since $\forall m \left[A^m(S) \subseteq A^\infty(S)\right]$, we have $\forall s \left[s \in A^\infty(S) \implies d(s) = b\right]$.
- Since $\{b\} = B$, we have $\forall s \left[s \in A^\infty(S) \implies d(s) \in B\right]$.

□

Appendix B

Common Assumption of Rationality

B.1 Proofs of Theorems 2.4.4 and 2.4.7

A **Borel refinement** of a Polish space T is a Polish space U such that U has the same points as T, and every open set in T is open in U. To prove Theorem 2.4.7, we need the following results about Borel refinements.

Proposition B.1.1 (13.1 and 13.11 in Kechris (1995)). *Suppose T is a Polish space, Y is a second countable space, and $f : T \to Y$ is a Borel function. Then, there is a Borel refinement U of T such that $f : U \to Y$ is continuous.*

Note that each subspace of a Polish space is second countable.

Proposition B.1.2 (13.3 in Kechris (1995)). *Let T_0 be a Polish space, and for each $n \in \mathbb{N}$ let T_{n+1} be a Borel refinement of T_n. Let T_∞ be the space whose topology is generated by $\bigcup_{n\in\mathbb{N}} T_n$. Then, T_∞ is a Polish space and is a Borel refinement of T_0.*

Proposition B.1.3 (Portmanteau Theorem, 17.20 in Kechris (1995))**.** *Let X be a Polish space, let $\mathcal{M}(X)$ be the space of Borel probability measures on X, and let $\mathcal{O}$ be an open basis for X. A sequence μ_k weakly converges to μ in $\mathcal{M}(X)$ if and only if $\liminf_k \mu_k(O) \geq \mu(O)$ for every $O \in \mathcal{O}$.*

This result is stated for all open sets in Kechris (1995), but the version stated here with an open basis follows from the proof.

Proof of Theorem 2.4.7. Let $T^a_0 = T^a$ and $T^b_0 = T^b$. Using Proposition B.1.1 countably many times, we obtain sequences of Polish spaces T^a_n, T^b_n such that for each n, T^a_{n+1} is a Borel refinement of T^a_n and λ^a is continuous from T^a_{n+1} to $\mathcal{L}(S^b \times T^b_n)$, and similarly with a and b reversed. Let U^a be the space whose topology is generated by $\bigcup_n T^a_n$, and define U^b analogously. By Proposition B.1.2, U^a is a Borel refinement of T^a, and U^b is a Borel refinement of T^b. Since each open set in T^a_{n+1} is open in U^a, λ^a is continuous from U^a to $\mathcal{L}(S^b \times T^b_n)$ for each n.

We show that λ^a is continuous from U^a to $\mathcal{L}(S^b \times U^b)$. To do this, suppose that u^a_k converges to u^a in U^a. Then, for each n, $\lambda^a(u^a_k)$ converges to $\lambda^a(u^a)$ in $\mathcal{L}(S^b \times T^b_n)$. For some ℓ, we have $\lambda^a(u^a) \in \mathcal{N}_\ell(S^b \times U^b)$. Then, $\lambda^a(u^a_k) \in \mathcal{N}_\ell(S^b \times U^b)$ for all but finitely many k, so we may assume this holds for all k. Let $\mu_{k,m}$ be the m-th coordinate of $\lambda^a(u^a_k)$, and let μ_m be the m-th coordinate of $\lambda^a(u^a)$. It follows that for each n and each $m \leq \ell$, $\mu_{k,m}$ weakly converges to μ_m in $\mathcal{M}(S^b \times T^b_n)$. By the Portmanteau Theorem, $\liminf_k \mu_{m,k}(O) \geq \mu_m(O)$ for each m, n and each open set O in $S^b \times T^b_n$. Since the open sets in $S^b \times T^b_n$, $n \in \mathbb{N}$ form an open basis for $S^b \times U^b$, it follows from the other direction of the Portmanteau Theorem B.1.3 that for each $m \leq \ell$, $\mu_{k,m}$ weakly converges to μ_m in $\mathcal{M}(S^b \times U^b)$. Therefore, $\lambda^a(u^a_k)$ converges to $\lambda^a(u^a)$ in $\mathcal{L}(S^b \times U^b)$. This shows that λ^a is continuous from U^a to $\mathcal{L}(S^b \times U^b)$. $\square$

To prove Theorem 2.4.4, we need two results from BFK about assumption.

Proposition B.1.4 (Property 6.2 in BFK)**.** *Let X be a Polish space, E, F be Borel subsets of X, and $\sigma = (\mu_0, \ldots, \mu_{m-1})$ a full-support LPS on X. If σ assumes both E and F the same level, then $\overline{E} = \overline{F}$.*

Proposition B.1.5 (Property 6.3 in BFK)**.** *Let X be a Polish space, $k \in \mathbb{N}$, and $\sigma \in \mathcal{L}^+(X)$. Suppose $E_n, n \in \mathbb{N}$ are Borel sets in X, and E_n is assumed under σ at level k for each $n \in \mathbb{N}$. Then, $\bigcap_{n\in\mathbb{N}} E_n$ is assumed under σ at level k.*[1]

Proof of Theorem 2.4.4. Since $R^a_\infty \times R^b_\infty$ is nonempty, there exists a state of the world (s^a, t^a, s^b, t^b) that satisfies RCAR. Furthermore, the LPS $\lambda^a(t^a)$ has full support and assumes each set in the infinite sequence $(R^b_1, R^b_2, \ldots)$. Let $\lambda^a(t^a) = \sigma = (\mu_0, \ldots, \mu_{n-1})$, where $n \in \mathbb{N}$ is the length of σ. It follows that there is a level $k \leq n$ and an infinite subset $J \subseteq \mathbb{N}$ such that σ assumes R^b_m at level k for each $m \in J$. Pick an element $M \in J$. We have $R^b_\infty = \bigcap_{m\in J} R^b_m$, so by Propositions B.1.5 and B.1.4,

- σ assumes R^b_∞ at level k;
- $\overline{R^b_\infty} = \overline{R^b_M}$.

We may take M large enough so that $S^b_M = S^b_\infty$, the IA set of player b. Since $\{s^b\} \times T^b$ is open for all $s^b \in S^b$, $\overline{R^b_\infty} = \overline{R^b_M}$ implies that $\operatorname{proj}_{S^b} R^b_\infty = \operatorname{proj}_{S^b} R^b_M$. By Proposition 2.4.1, we have $\operatorname{proj}_{S^b} R^b_M = S^b_M$, and it follows that $\operatorname{proj}_{S^b} R^b_\infty = S^b_\infty$. By analogous arguments, $\operatorname{proj}_{S^a} R^a_\infty = S^a_\infty$. □

B.2 Proofs of Theorems 2.4.8 and 2.4.10

For convenience, we let $S^a_0 \equiv S^a$ and $R^a_0 \equiv S^a \times T^a$, and similarly for b.

[1] The proof in BFK establishes the result as stated here, but the statement in BFK did not mention the level.

Lemma B.2.1. *For each $m > 0$,*

$$P^a_m = \left\{\nu^{m-1} \dots \nu^0 : (\forall k < m)[\nu^k \in \mathcal{N}(S^b) \wedge \operatorname{Supp} \nu^k = S^b_k]\right\}.$$

Proof of Lemma B.2.1. Note that the ν's are indexed in the reverse order. The proof is by induction. The base case ($m = 1$) holds trivially. Assume the result for m. Then, by the definition of P^a_{m+1}, the following are equivalent.

- $\mu \in P^a_{m+1}$;
- $\mu = \nu\nu'$ for some $\nu, \nu' \in \mathcal{N}(S^b)$ such that $\operatorname{Supp} \nu = S^b_m$ and $\nu' \in P^a_m$;
- $\mu = \nu^m \nu^{m-1} \dots \dots \nu^0$ for some $\nu^0, \dots, \nu^{m-1}, \nu^m \in \mathcal{N}(S^b)$ such that $\operatorname{Supp} \nu^k = S^b_k$ for all $k \leq m$.

This completes the induction. □

Lemma B.2.2. *For each $m > 0$ we have $P^a_{m+1} \subseteq P^a_m$.*

Proof. Suppose that $\mu \in P^a_{m+1}$. By Lemma B.2.1, μ can be written in the form $\mu = \nu^m \nu^{m-1} \dots \nu^0$ where $\nu^k \in \mathcal{N}(S^b)$ and $\operatorname{Supp} \nu^k = S^b_k$ for all $k \leq m$. Then, $\operatorname{Supp} \nu^m \nu^{m-1} = S^b_{m-1}$, and by Lemma B.2.1 again we have $\mu \in P^a_m$. □

We will need the following result, which is Proposition 1 in Blume, Brandenburger, and Dekel (1991b).

Proposition B.2.3. *For each $\nu \in \mathcal{N}(S^b)$ there is a probability measure $\rho \in \mathcal{M}(S^b)$ such that*

1. $\operatorname{Supp} \rho = \operatorname{Supp} \nu$; *and*
2. $(\rho) \sim \nu$.

Lemma B.2.4. *For each $m \in \mathbb{N}$ and $\sigma \in P^a_{m+1}$ there exists $\nu = (\nu_0, \ldots, \nu_m) \in \mathcal{N}_{m+1}(S^b)$ such that $\nu \sim \sigma$ and $\operatorname{Supp} \nu_{m-k} = S^b_k$ for each $k \leq m$ (so $\nu \in P^a_{m+1}$ by Lemma B.2.1).*

Proof of Lemma B.2.4. We argue by induction on m. The result for $m = 0$ follows from Proposition B.2.3. Suppose the result holds for m, and let $\sigma \in P^a_{m+2}$. Then, $\sigma = \sigma'\sigma''$ where $\sigma' \in \mathcal{N}(S^b)$, $\operatorname{Supp} \sigma' = S^b_{m+1}$, and $\sigma'' \in P^a_{m+1}$. By inductive hypothesis, there exists $\nu = (\nu_1, \ldots, \nu_{m+1}) \in \mathcal{N}_{m+1}(S^b)$ such that $\nu \sim \sigma''$ and $\operatorname{Supp} \nu_{m+1-k} = S^b_k$ for each $k \leq m$. By Proposition B.2.3, there exists $\nu_0 \in \mathcal{M}(S^b)$ such that $\operatorname{Supp} \nu_0 = \operatorname{Supp} \sigma' = S^b_{m+1}$ and $\nu_0 \sim \sigma'$. Then

$$\nu_0\nu = (\nu_0, \nu_1, \ldots, \nu_{m+1}) \sim \sigma,$$

so the result holds for $m + 1$. □

Lemma B.2.5. *If $\nu, \nu' \in \mathcal{N}(S^b)$ then $\mathbb{O}(\nu\nu') \subseteq \mathbb{O}(\nu)$.*

Proof of Lemma B.2.5. It is easily seen that if $r^a \succ_\nu s^a$ then $r^a \succ_{\nu\nu'} s^a$. If $s^a \in \mathbb{O}(\nu\nu')$ then there is no r^a such that $r^a \succ_{\nu\nu'} s^a$, so there is no r^a such that $r^a \succ_\nu s^a$, and thus $s^a \in \mathbb{O}(\nu)$. □

Proof of Theorem 2.4.8. The proof is by induction on m. The base case: Since $P^a_1 = \mathcal{N}^+(S^b)$ and S^a_1 is the set of Ann's admissible strategies, we have

$$\bigcup \mathbb{X}^a_1 \equiv \bigcup \{\mathbb{O}(\nu) : \nu \in P^a_1\} = S^a_1.$$

Now, fix an $m > 1$ and assume the induction hypothesis that $S^a_m = \bigcup \mathbb{X}^a_m$. We will show that $S^a_{m+1} = \bigcup \mathbb{X}^a_{m+1}$ in two steps.

Step 1: We want to show that $S^a_{m+1} \supseteq \bigcup \mathbb{X}^a_{m+1}$. Equivalently, we want to show that $\mathbb{O}(\mu) \subseteq S^a_{m+1}$ for any $\mu \in P^a_{m+1}$. By Lemma B.2.4, there exists $\nu = (\nu_0, \ldots, \nu_m) \in \mathcal{N}_{m+1}(S^b)$ such that $\nu \sim \mu$ and $\operatorname{Supp} \nu_{m-k} = S^b_k$ for each $k \leq m$. Then, $\mathbb{O}(\nu) = \mathbb{O}(\mu)$. By Lemma B.2.5, $\mathbb{O}(\mu) = \mathbb{O}(\nu) \subseteq \mathbb{O}(\nu_0)$. By Lemma B.2.2, $\mu \in P^a_m$, so $\mathbb{O}(\mu) \in \mathbb{X}^a_m$. Then, by the induction hypothesis $S^a_m = \bigcup \mathbb{X}^a_m$ we have $\mathbb{O}(\mu) \subseteq S^a_m$. We note that $\operatorname{Supp} \nu_0 = S^b_m$, so by the definition of S^a_{m+1}, $\mathbb{O}(\nu_0) \cap S^a_m \subseteq S^a_{m+1}$. But, $\mathbb{O}(\mu) \subseteq \mathbb{O}(\nu_0) \cap S^a_m$, so $\mathbb{O}(\mu) \subseteq S^a_{m+1}$.

Step 2: We want to show that $S^a_{m+1} \subseteq \bigcup \mathbb{X}^a_{m+1}$. Equivalently, we want to show that, for each $s^a \in S^a_{m+1}$, there exists a $\mu \in P^a_{m+1}$ such that $s^a \in \mathbb{O}(\mu)$. If $s^a \in S^a_{m+1}$, then for each $k \leq m$ we have $s^a \in S^a_{k+1}$, so there exists $\nu^k \in \mathcal{N}_1(S^b)$ such that $\operatorname{Supp} \nu^k = S^b_k$ and $s^a \in \mathbb{O}(\nu^k)$. By Lemma B.2.5, $s^a \in \mathbb{O}(\mu)$ where $\mu = \nu^m \nu^{m-1} \ldots \nu^0$. By Lemma B.2.1, $\mu \in P^a_{m+1}$. $\square$

For the proof of Theorem 2.4.10, we will need the following two results. The result below is an immediate consequence of Lemma E.2 in BFK.

Proposition B.2.6. *For each LPS $\sigma \in \mathcal{L}^+(S^b \times T^b)$ there are continuum many $\hat{\sigma} \in \mathcal{L}^+(S^b \times T^b)$ such that*

- *For each $s^a \in S^a$, s^a is optimal under σ if and only if s^a is optimal under $\hat{\sigma}$ (i.e., $\sigma \sim \hat{\sigma}$);*
- *For each Borel set $E \subseteq S^b \times T^b$ and each $k \in \mathbb{N}$, E is assumed under σ at level k if and only if E is assumed under $\hat{\sigma}$ at level k.*

Proposition B.2.7 (Lemma E.3 in BFK)**.** *In a complete type structure for G, for each $m \in \mathbb{N}$ we have* $\operatorname{proj}_{S^b} R^b_m = \operatorname{proj}_{S^b}(R^b_m \setminus R^b_{m+1})$.

Proof of Theorem 2.4.10. Part 1: We want to show that if $(s^a, t^a) \in R^a_{m+1}$, then there is a $\nu \in P^a_{m+1}$ such that $\operatorname{marg}_{S^b} \lambda^a(t^a) \sim \nu$. Let $\sigma = (\mu_0, \ldots, \mu_n) = \lambda^a(t^a)$.

For each $k \leq m$, let $[k]$ denote the level at which σ assumes R^b_k. Then, $n = [0] \geq \cdots \geq [m]$. For *Part 1* only, let $\nu^k = \operatorname{marg}_{S^b}(\mu_0, \ldots, \mu_{[k]})$ for each $k \leq m$. Then, $\nu^0 = \operatorname{marg}_{S^b} \sigma$. Since σ assumes R^b_k at level $[k]$, we see from Proposition 2.4.1 that $\operatorname{Supp} \nu^k = \operatorname{proj}_{S^b} R^b_k = S^b_k$. Note that for each $k \leq m$, ν^{k+1} is an initial segment of ν^k. It is readily verified that if ν is an initial segment of ν' then $\nu\nu' \sim \nu'$. It follows by induction that $\nu^m\nu^{m-1} \ldots \nu^k \sim \nu^k$ for each $k < m$, and hence $\nu^m\nu^{m-1} \ldots \nu^0 \sim \nu^0 = \operatorname{marg}_{S^b} \sigma$. By Lemma B.2.1, $\nu^m\nu^{m-1} \ldots \nu^0 \in P^a_{m+1}$.

Part 2: Let $\nu \in P^a_{m+1}$. We want to show that there is some $(s^a, t^a) \in R^a_{m+1}$ such that $\operatorname{marg}_{S^b} \lambda^a(t^a) = \nu$. By Lemma B.2.1, we can write ν as $\nu^m\nu^{m-1} \ldots \nu^0$, where $\operatorname{Supp} \nu^k = S^b_k$ for all $k \leq m$.

We first consider ν^m. We can write $\nu^m = (\nu^m_0, \ldots, \nu^m_n)$ where each ν^m_i is a probability measure on S^b. By Proposition 2.4.1, $\operatorname{Supp} \nu^m = S^b_m = \operatorname{proj}_{S^b} R^b_m$. By Proposition B.2.6, for each $s^b \in S^b_m$, the set $(\{s^b\} \times T^b) \cap R^b_m$ is uncountable. Since S^b is finite and T^b is separable, R^b_m has a countable dense subset. We may therefore pick sets $Y_0, \ldots, Y_n$ such that

- The sets Y_i are countable and pairwise disjoint;
- The union $Y_0 \cup \cdots \cup Y_n$ is a dense subset of R^b_m; and
- For each $i \leq n$, $\operatorname{proj}_{S^b} Y^m_i = \operatorname{Supp} \nu^m_i$.

For each $s^b \in S^b$ and $i \leq n$, we can assign positive measures to the points of $(\{s^b\} \times T^b) \cap Y_i$ that add up to $\nu^m_i(\{s^b\})$. This gives probability measures $\mu^m_i \in \mathcal{M}(S^b \times T^b)$ such that $\operatorname{Supp} \mu^m_i \supseteq Y_i$, $\mu^m_i(Y_i) = 1$, and $\operatorname{marg}_{S^b} \mu^m_i = \nu^m_i$. Then, the measures $\mu^m_i, i \leq n$ are mutually singular, so the $n+1$-tuple $\mu^m \equiv (\mu^m_0, \ldots, \mu^m_n)$ is an LPS such that

- $\operatorname{Supp} \mu^m \supseteq R^b_m$;

- $\mu^m(R^b_m) = (1, 1, \ldots, 1)$; and
- $\operatorname{marg}_{S^b} \mu^m = \nu^m$.

We now consider ν^k for $k < m$. By Proposition B.2.7, $\operatorname{Supp} \nu^k = S^b_k = \operatorname{proj}_{S^b}(R^b_k \setminus R^b_{k+1})$. By the above construction with $R^b_k \setminus R^b_{k+1}$ in place of R^b_m, we obtain an LPS μ^k such that

- $\operatorname{Supp} \mu^k \supseteq R^b_k \setminus R^b_{k+1}$;
- $\mu^k(R^b_k \setminus R^b_{k+1}) = (1, 1, \ldots, 1)$; and
- $\operatorname{marg}_{S^b} \mu^k = \nu^k$.

Now, let μ be the concatenation $\mu \equiv \mu^m \mu^{m-1} \ldots \mu^0$. Then, $\operatorname{marg}_{S^b} \mu = \nu$, and μ is a full-support LPS that assumes R^b_k for all $k \leq m$. By completeness, there exists a type $t^a \in T^a$ such that $\lambda^a(t^a) = \mu$. Since μ has full support, there exists an $s^a \in S^a$ such that (s^a, t^a) is a rational pair. Then, (s^a, t^a) satisfies rationality and m-th order assumption of rationality, so $(s^a, t^a) \in R^a_{m+1}$. □

For each LPS $\sigma \in \mathcal{N}(S^b \times T^b)$, let $\mathbb{O}(\sigma)$ be the set of all strategies $s^a \in S^a$ that are optimal under $\operatorname{marg}_{S^b} \sigma$.

Proof of Corollary 2.4.11. Note that for each LPS $\sigma \in \mathcal{N}(S^b \times T^b)$, we have $\mathbb{O}(\sigma) = \mathbb{O}(\operatorname{marg}_{S^b} \sigma)$. Then, by Theorem 2.4.10,

$$\mathbb{X}^a_m = \{\mathbb{O}(\nu) : \nu \in P^a_m\} = \{\mathbb{O}(\lambda^a(t^a)) : (s^a, t^a) \in R^a_m\}.$$

□

Proof of Corollary 2.4.12. Proof of 1: By Lemma C.4 in BFK, each of the sets R^a_m is Borel. Since S^a is finite, it follows that each of the sets $\Gamma^a(X^a, R^a_m)$ is Borel. Let U^a_m

be the set of types for Ann that are open-minded and assume k-th order rationality for Bob for all $k < m$, that is,

$$U_m^a = \{t^a \in T^a : (\exists s^a)(s^a, t^a) \in R_m^a\}.$$

Then, for each nonempty $X^a \subseteq S^a$,

$$\Gamma^a(X^a, R_m^a) = \{t \in U_m^a : \mathbb{O}(\lambda^a(t^a)) = X^a\}.$$

Since $R_{m+1}^a \subseteq R_m^a$, $U_{m+1}^a \subseteq U_m^a$, and therefore $\Gamma^a(X^a, R_{m+1}^a) \subseteq \Gamma^a(X^a, R_m^a)$. This proves 1.

Proof of 2: By Corollary 2.4.11, the following are equivalent

- $X^a \in \mathbb{X}_m^a$;
- $(\exists (s^a, t^a) \in R_m^a)\, [X^a = \mathbb{O}(\lambda^a(t^a))]$;
- $(\exists t^a \in U_m^a)\, [X^a = \mathbb{O}(\lambda^a(t^a))]$;
- $(\exists t^a)\, [t^a \in \Gamma^a(X^a, R_m^a)]$.

Proof of 3: Since $\mathfrak{T}$ is complete, there exists $t^a \in T^a$ such that Ann is not open-minded, so there is no s^a such that $(s^a, t^a) \in R_1^a$ and hence $t^a \in \Gamma(\varnothing, R_1^a)$. $\Gamma^a(\varnothing, R_m^a)$ is the complement of the union of the sets $\Gamma^a(X^a, R_m^a)$ with X^a nonempty. Thus, by 1, $\{\Gamma^a(\varnothing, R_m^a) : m > 0\}$ is an increasing sequence of Borel sets. □

B.3 Polish Spaces and Assumption

In this section we establish some useful properties of Polish spaces and assumption. A topological space $(X, \mathscr{T})$ is called **Polish** if it is a separable completely metrizable space. It is well known that all uncountable Polish spaces have cardinality equal to $2^{\aleph_0}$ (i.e., the cardinality of the continuum). By Proposition B.3.1 below, every uncountable Borel subset of a Polish space also has cardinality $2^{\aleph_0}$.

The **Cantor space** $\mathcal{C}$ is the set $\{0,1\}^{\mathbb{N}}$ endowed with the product topology. It is a Polish space of cardinality $2^{\aleph_0}$. A **Cantor set** C in a topological space X is a homeomorphic copy of $\mathcal{C}$ in X—that is, $(C, \mathscr{T}|C)$ is homeomorphic to $\mathcal{C}$, where $\mathscr{T}|C = \{U \cap C : U \in \mathscr{T}\}$ is the subspace topology on C. A subset of a topological space is **perfect** if it is closed and has no isolated points.

Proposition B.3.1 (The Perfect Set Theorem for Borel Sets, 13.6 in Kechris, 1995)**.** *Let X be a Polish space and $A \subseteq X$ be Borel. Then, either A is countable, or else A contains a Cantor set and has cardinality $2^{\aleph_0}$.*

Proposition B.3.2 (Cantor-Bendixson, 6.4 in Kechris, 1995)**.** *Let X be a Polish space. Then, X can be uniquely written as a disjoint union $X = P \uplus C$, with P a perfect subset of X and C countably open. Furthermore, every open neighborhood of every $x \in P$ is uncountable.*

Lemma B.3.3. *Let X be an uncountable Polish space and $n \in \mathbb{N}$. Then, there exist disjoint open sets $U_1, \ldots, U_n$ in X such that*

1. *U_i is uncountable for all i;*
2. *$X \setminus \biguplus \{U_1, \ldots, U_n\}$ is uncountable.*

Proof of Lemma B.3.3. By Proposition B.3.2, $X = P \uplus C$ such that P is perfect and C countably open. We can choose $n+1$ distinct points $x_1, \ldots, x_{n+1} \in P$. Since X is metrizable, it is a normal—that is, any two disjoint closed sets in X have disjoint open neighborhoods. It follows that there exist disjoint open sets $U_1, \ldots, U_{n+1}$ such that

$x_j \in U_j$ for all j. By Proposition B.3.2, $U_1, \ldots, U_{n+1}$ are uncountable. Finally, $X \setminus \biguplus \{U_1, \ldots, U_n\}$ is uncountable since it contains U_{n+1}, which is itself uncountable. □

Lemma B.3.4. $\mathcal{C} = \biguplus_{n \in \mathbb{N}} K_n$, *where* $(K_0, K_1, \ldots)$ *is a sequence of disjoint uncountable compact sets.*

Proof of Lemma B.3.4. Let

$$K_0 = \{0\}^{\mathbb{N}} \cup \{c \in \mathcal{C} : c_0 = 1\}.$$

For each $n > 0$, let

$$K_n = \{c \in \mathcal{C} : (\forall k < n) c_k = 0 \wedge c_n = 1\}.$$

For each $n > 0$, K_n is a Cantor set, and therefore it is uncountable and compact. K_0 is the union of a Cantor set and a single point, and therefore it is also uncountable and compact. By construction, $\mathcal{C} = \biguplus_{n \in \mathbb{N}} K_n$, and $(K_0, K_1, \ldots)$ is a sequence of disjoint sets. □

Given a Polish space $(X, \mathcal{O}(X))$, a **Borel subspace** of X is a topological space $(A, \mathcal{O}(A))$ where A is a nonempty Borel subset of X endowed with the subspace topology $\mathcal{O}(A) = \{U \cap A : U \in \mathcal{O}(X)\}$.

Proposition B.3.5 (Borel Isomorphism Theorem, Theorem 15.6 in (Kechris, 1995))**.** *Let* A, B *be Borel subspaces of Polish spaces. If* $\mathrm{card}(A) = \mathrm{card}(B)$*, then there is a one-to-one Borel mapping from* A *onto* B*.*[2]

[2]In Kechris, 1995, this result is stated in terms of standard Borel spaces, which are the measure spaces associated with Borel subspaces of Polish spaces.

Lemma B.3.6. *Let X, Y be Polish spaces, and let*

$$X = \biguplus_{n\in\mathbb{N}} X_n, \quad Y = \biguplus_{n\in\mathbb{N}} Y_n$$

be partitions of X, Y into countably many disjoint Borel sets such that

$$\mathrm{card}(X_n) = \mathrm{card}(Y_n) \text{ for each } n \in \mathbb{N}.$$

Then, there is a one-to-one Borel mapping from X onto Y that maps X_n onto Y_n for each $n \in \mathbb{N}$.

Proof of Lemma B.3.6. Each of the sets X_n, Y_n with its subspace topology is a Borel subspace of a Polish space. By Proposition B.3.5, for each $n \in \mathbb{N}$ there is a one-to-one Borel mapping λ_n from X_n onto Y_n. Then, the union $\lambda = \bigcup_{n\in\mathbb{N}} \lambda_n$ is a one-to-one Borel mapping from X onto Y that sends X_n onto Y_n for each $n \in \mathbb{N}$, as required. □

We will need the following facts from BFK about assumption.

Proposition B.3.7 (Lemma C.3 in BFK)**.** *For each Polish space X and Borel set E in X, the set of $\sigma \in \mathcal{L}^+(X)$ such that E is assumed under σ is Borel.*

Proposition B.3.8 (Lemma B.1 in BFK[3])**.** *Let X be a Polish space, E be a Borel subset of X, $\sigma = (\mu_0, \ldots, \mu_{n-1})$ be a full-support LPS on X, and $k < n$. Then, σ assumes E at level k if and only if the following conditions are met.*

1. $\mu_i(E) = 1$ *for each* $i \leq k$*;*
2. $\mu_i(E) = 0$ *for each* $i > k$*; and*
3. $E \subseteq \bigcup_{i\leq k} \mathrm{Supp}\, \mu_i$*.*

[3]The proof in BFK establishes this fact, but the statement of Lemma B.1 was garbled in BFK.

In a topological space, a set D is said to be **dense in** a set E if $D \subseteq E$ and $\overline{D} = \overline{E}$. Note that if D_1 is dense in E_1 and D_2 is dense in E_2, then $D_1 \cup D_2$ is dense in $E_1 \cup E_2$. Also, if D is dense in E and $D \subseteq F \subseteq E$, then D is dense in F and F is dense in E.

Lemma B.3.9. *Let X be a Polish space, and E an uncountable Borel set in X. Then, there exists a Cantor set $C \subseteq E$ such that $E \setminus C$ is uncountable and $E \setminus C$ is dense in E.*

Proof of Lemma B.3.9. The Cantor space $\mathcal{C}$ contains the Cantor set $\{(0,0),(1,1)\}^{\mathbb{N}}$, and that the complement of this set is uncountable and dense in $\mathcal{C}$. By Proposition B.3.1, E contains a Cantor set D. It follows that D contains a Cantor set C such that $D \setminus C$ is uncountable and dense in D. But $D \subseteq E$, so $C \subseteq E$, and $E \setminus C$ is uncountable and dense in E. $\square$

Lemma B.3.10. *Let X be a Polish space, and U_0 an uncountable open set in X. Then, there exists a decreasing sequence of open sets $(U_0, U_1, U_2, \dots)$ such that*

1. *For all $n \in \mathbb{N}$, $U_n \setminus U_{n+1}$ is uncountable;*
2. *$U_\infty \equiv \bigcap_{n \in \mathbb{N}} U_n$ is an uncountable open set;*
3. *U_∞ is dense in U_0.*

Proof of Lemma B.3.10. By Lemma B.3.9, there exists a Cantor set $C \subseteq U_0$ such that $U_0 \setminus C$ is uncountable and dense in U_0. By Lemma B.3.4, there exists a sequence $(K_0, K_1, \dots)$ of disjoint uncountable compact sets such that $\biguplus_{n \in \mathbb{N}} K_n = C$. For each $n > 0$, define $U_n \equiv U_0 \setminus \biguplus_{j<n} K_j$. Then, U_n is open, and $U_n \setminus U_{n+1} = K_n$, which is uncountable. Moreover, $U_\infty = U_0 \setminus C$, so U_∞ is uncountable, open and dense in U_0. $\square$

Lemma B.3.11. *Let X, Y be Polish spaces with X finite, and let $Z_0 = X \times Y$. Let $\nu = (\nu_0, \ldots, \nu_m) \in \mathcal{N}_{m+1}(X)$, and let $(Z_1, Z_2, \ldots, Z_{m+1})$ be a decreasing sequence of nonempty Borel subsets of Z_0, such that*

- *For all $k \leq m$, $\operatorname{proj}_X Z_k = \operatorname{proj}_X (Z_k \setminus Z_{k+1})$;*
- *For all $k \leq m$, $\operatorname{Supp} \nu_{m-k} = \operatorname{proj}_X Z_k$.*

Then, there exists $\mu = (\mu_0, \ldots, \mu_m) \in \mathcal{L}^+_{m+1}(Z)$ such that

- *$\operatorname{marg}_X \mu = \nu$;*
- *For all $k \leq m$, μ assumes Z_k at level $m - k$;*
- *μ does not assume Z_{m+1};*
- *For each $x \in \operatorname{proj}_X Z_{m+1}$, $\mu_0(Z_{m+1} \cap (\{x\} \times Y)) > 0$.*

Proof of Lemma B.3.11. Using the fact that Polish spaces are separable, there is a countable subset U of Z_0 such that $U \cap Z_{m+1}$ is dense in Z_{m+1}, and $U \cap (Z_k \setminus Z_{k+1})$ is dense in $Z_k \setminus Z_{k+1}$ for each $k \leq m$. It follows that $U \cap Z_k$ is dense in Z_k for each $k \leq m$.

Choose any $\rho \in \mathcal{M}(Z_0)$ such that $\rho(U) = 1$ and $\rho(\{u\}) > 0$ for each $u \in U$. Since U is dense in Z_0, $\rho \in \mathcal{M}^+(Z_0)$. For all $k \in \mathbb{N}$, let $X_k \equiv \operatorname{proj}_X Z_k$. For all $x \in X$ and $k \in \mathbb{N}$, let $Z_k(x) = Z_k \cap (\{x\} \times Y)$. This set is clearly Borel. Since X is finite, it readily follows that

- For all $x \in X_k$ and $k \geq 0$, $Z_k(x) \cap U$ is nonempty and dense in $Z_k(x)$;
- For all $x \in X_k$ and $k \leq m$, $(Z_k(x) \setminus Z_{k+1}(x)) \cap U$ is nonempty and dense in $Z_k(x) \setminus Z_{k+1}(x)$.

Note that for every Borel set V of Z_0 such that $V \cap U$ is nonempty, the conditional measure $\rho(\cdot|V)$ is well-defined. We define $\mu_1, \ldots, \mu_m$ in the following way. For each $k < m$, let

$$\mu_{m-k}(E) \equiv \sum_{x \in X_k} \nu_{m-k}(x)\rho(E|Z_k(x) \setminus Z_{k+1}(x)).$$

Define μ_0 in the following way.

$$\mu_0(E) \equiv \sum_{x \in X_m} \nu_0(x)\rho(E|Z_m(x)).$$

It is clear from these definitions that $\sum_{k=0}^{m} \mu_k$ and ρ are mutually absolutely continuous. Therefore, $\mu = (\mu_0, \ldots, \mu_m)$ is a full-support LPS on Z_0. It is also clear that $\mu_0(Z_{m+1}(x)) > 0$ for each $x \in X_{m+1}$, and that $\text{marg}_X \, \mu_k = \nu_k$ for all $k \leq m$.

For each $k \leq m$, $Z_k \subseteq \text{Supp}(\mu_0, \ldots, \mu_{m-l})$, because $Z_k \cap S$ is dense in Z_k. Using Proposition B.3.8, we can easily verify that, for all $k \leq m$, μ assumes Z_k at level $m-k$. $Z_m \setminus Z_{m+1}$ has a nonempty intersection with U, so μ_0 gives the set positive probability. However, since $\mu_0(Z_m) = 1$, it follows that $\mu_0(Z_{m+1}) < 1$. Proposition B.3.8 makes it clear that μ does not assume Z_k when $k > m$. □

Lemma B.3.12. *Let X, Y be Polish spaces with X being finite, and let $Z_0 = X \times Y$. Let $\nu = (\nu_0, \ldots, \nu_m) \in \mathcal{N}_{m+1}(X)$, and let $(Z_1, Z_2, \ldots)$ be a strictly decreasing sequence of nonempty Borel subsets of Z_0, such that*

- *For all $k \geq 0$, $\text{proj}_X Z_k = \text{proj}_X(Z_k \setminus Z_{k+1})$;*
- *For all $k \leq m$, $\text{Supp}\, \nu_{m-k} = \text{proj}_X Z_k$;*
- *$Z_\infty \equiv \bigcap_{k \in \mathbb{N}} Z_k$ is dense in Z_m.*

Then, there exists $\mu = (\mu_0, \ldots, \mu_m) \in \mathcal{L}^+_{m+1}(Z)$ such that

- *For all $k \leq m$, μ assumes Z_k at level $m - k$;*
- *For all $k > m$, μ assumes Z_k at level 0;*
- $\operatorname{marg}_X \mu = \nu$.

Proof of Lemma B.3.12. For each $k \in \mathbb{N} \cup \{\infty\}$, let $X_k \equiv \operatorname{proj}_X Z_k$, and for each $x \in X$, let $Z_k(x) = Z_k \cap (\{x\} \times Y)$. Since Z_∞ is dense in Z_m and X is finite, we have $X_\infty = X_m$, and for each $x \in X_m$, $Z_\infty(x)$ is dense in $Z_m(x)$. By Lemma B.3.11, there exists $\phi = (\phi_0, \ldots, \phi_m) \in \mathcal{L}^+_{m+1}(Z)$ such that

- $\operatorname{marg}_X \phi = \nu$;
- For each $k \leq m$, ϕ assumes Z_k at level $m - k$;
- For each $x \in X_m$, $\mu_0(Z_\infty(x)) > 0$.

Then, the conditional probability $\phi_0(\cdot|Z_\infty(x))$ is well-defined for each $x \in X_m$. For each $0 < k \leq m$, let $\mu_k = \phi_k$. Define μ_0 in the following way.

$$\mu_0(E) \equiv \sum_{x \in X_m} \nu_0(x)\phi_0(E|Z_\infty(x)).$$

By construction,

$$\operatorname{Supp} \mu_0 = \operatorname{Supp} \phi_0(\cdot|Z_\infty) = \overline{Z_\infty} = \overline{Z_m} = \operatorname{Supp} \phi_0.$$

It is readily apparent, then, that

$$\operatorname{Supp}(\mu_0, \ldots, \mu_m) = \operatorname{Supp}(\phi_0, \ldots, \phi_m) = Z_0.$$

We have $\operatorname{Supp} \nu_0 = X_m$, so

$$\mu_0(Z_\infty) = \mu_0(Z_m) = \nu_0(X_m) = 1.$$

Using Proposition B.3.8, we can easily verify that for all $k \geq m$, μ assumes Z_k at level 0. □

B.4 Proof of Theorem 2.4.2

Recall that RCAR is the intersection R_∞ of an infinite sequence of nested sets. Intuitively, it would be easier to show that RCAR is nonempty if each RmAR set was large in some sense. If RmAR is large in a probabilistic sense (i.e, given probability one by some full-support measure that is fixed across all values of m), then the difference between RmAR and R$(m+1)$AR is small (i.e., given probability zero by the same measure). Furthermore, since probability measures are countably additive, it would also imply that the set of states not satisfying RCAR is small in the same sense. Non-emptiness of RCAR would be an immediate consequence of this fact.

Unfortunately, this idea must be modified before it can be applied to the method of constructing RmAR sets. Since the strategy space is finite, if some strategy s^a is inadmissible then the RmAR sets will have an empty intersection with the open set $\{s^a\} \times T^a$. Therefore, the RmAR sets cannot be given probability one by a full-support measure if some strategies are inadmissible. Instead, we replace "given probability one by a full-support measure" with "given probability one by a measure having full-support on some open set U" as the appropriate notion of large size.

Our general strategy will be to first construct some sequence of sets that we would like to be the RmAR sets, then show the existence of a complete type structure in which the candidates are indeed the RmAR sets. For each m, the candidate for RmAR will need to be shaped like RmAR as described by Corollary 2.4.12 and given probability one by a measure having full-support on the same open set U for all m.

We fix the underlying game $\langle S^a, S^b, \pi^a, \pi^b \rangle$ as we have done throughout the paper. We also fix T^a and T^b and assume that they are uncountable Polish spaces.

Lemma B.4.1. *There exists* $\{\tau_1(X^a) : X^a \in \mathbb{X}_1^a\}$*, a family of pairwise disjoint uncountable open sets such that* $T^a \setminus \bigcup \{\tau_1(X^a) : X^a \in \mathbb{X}_1^a\}$ *is uncountable and closed. Analogous objects exist for Bob.*

Proof of Lemma B.4.1. Since $\mathbb{X}_1^a$ is a finite set, the desired result follows immediately from Lemma B.3.3. □

For the remainder, fix $\{\tau_1(X^a) : X^a \in \mathbb{X}_1^a\}$, which exists by Lemma B.4.1. Furthermore, fix some full-support Borel probability measures $\phi^a \in \mathcal{M}^+(T^b)$ and $\phi^b \in \mathcal{M}^+(T^a)$.

Lemma B.4.2. *Fix* $X^a \in \mathbb{X}_1^a$*. There exists a family* $\{\tau_m(X^a) : m > 1\}$ *of open sets in* T^a *such that, for all* $m \in \mathbb{N}$*,*

1. $\tau_m(X^a) \supseteq \tau_{m+1}(X^a)$;
2. $\tau_m(X^a) \setminus \tau_{m+1}(X^a)$ *is an uncountable* ϕ^b*-null set;*
3. $\tau_\infty(X^a) \equiv \bigcap_{m=1}^{\infty} \tau_m(X^a)$ *is an uncountable open set; and*
4. $\phi^b(\tau_\infty(X^a)) = \phi^b(\tau_m(X^a))$.

Analogous results hold for Bob.

Proof of Lemma B.4.2. The result follows immediately from Lemma B.3.12. □

For each m, fix the family of sets $\{\tau_m(X^a) : X^a \in \mathbb{X}_1^a\}$ that exists by Lemma B.4.1 and Lemma B.4.2.

Definition B.4.3. *For all $m \in \mathbb{N}$, define*

$$\begin{aligned}
\widehat{R}_0^a &\equiv S^a \times T^a \\
\widehat{R}_m^a &\equiv \bigcup \{X^a \times \tau_m(X^a) : X^a \in \mathbb{X}_m^a\} \text{ and} \\
\widehat{R}_\infty^a &\equiv \bigcap_{m=1}^{\infty} \widehat{R}_m^a.
\end{aligned}$$

Analogous objects are defined for Bob.

Lemma B.4.4. *For all $m \in \mathbb{N}_0$,*

1. $\widehat{R}_m^a \supseteq \widehat{R}_{m+1}^a$;
2. $\widehat{R}_m^a$ *is an uncountable open set;*
3. $\widehat{R}_\infty^a$ *is an uncountable open set; and*
4. $\widehat{R}_m^a \setminus \widehat{R}_{m+1}^a$ *is uncountable.*

Analogous results hold for Bob.

Proof of Lemma B.4.4. The desired results follow from Lemma B.4.2 and the finiteness of S^a. □

Lemma B.4.5. *For any $\nu \in \mathcal{M}^+(S^a)$, $\nu \otimes \phi^b \in \mathcal{L}^+(S^a \times T^a)$ denotes the product measure on $S^a \times T^a$. For all $m \in \mathbb{N}$, if $\mathbb{X}_m^a = \mathbb{X}_{m+1}^a$, then $\widehat{R}_m^a \setminus \widehat{R}_{m+1}^a$ is $(\nu \otimes \phi^b)$-null. Analogous results hold for Bob.*

Proof of Lemma B.4.5. Fix $m \in \mathbb{N}$. If $\mathbb{X}^a_m = \mathbb{X}^a_{m+1}$, then

$$\widehat{R}^a_m \setminus \widehat{R}^a_{m+1} = \bigcup \{X^a \times [\tau_m(X^a) \setminus \tau_{m+1}(X^a)] : X^a \in \mathbb{X}^a_m\}.$$

By Lemma B.4.2, $\tau_m(X^a) \setminus \tau_{m+1}(X^a)$ is ϕ^b-null. It follows that $X^a \times (\tau_m(X^a) \setminus \tau_{m+1}(X^a))$ is $(\nu \otimes \phi^b)$-null. Then, $\widehat{R}^a_m \setminus \widehat{R}^a_{m+1}$ is $(\nu \otimes \phi^b)$-null since it is a finite union of such sets. □

Definition B.4.6. *Given any $X^a \in \mathbb{X}^a_1$, let $\Lambda_m(X^a)$ denote the set of all $\mu \in \mathcal{L}^+(S^b \times T^b)$ such that*

1. *$X^a = \{s^a \in S^a : s^a$ is optimal under $\mu\}$; and*
2. *μ assumes $\widehat{R}^b_1, \ldots, \widehat{R}^b_{m-1}$.*

Let $\Lambda_\infty(X^a) \equiv \bigcap_{m=1}^\infty \Lambda_m(X^a)$. Analogous objects are defined for Bob.

Lemma B.4.7. *The following holds for all $m \in \mathbb{N}$.*

1. *For all $X^a \in \mathbb{X}^a_{m+1}$, $\Lambda_m(X^a) \setminus \Lambda_{m+1}(X^a)$ is uncountable; and*
2. *For all $X^a \in \mathbb{X}^a_\infty$, $\Lambda_\infty(X^a)$ is uncountable; and*
3. *For all $X^a \in \mathbb{X}^a_m$, $\Lambda_m(X^a)$ is Borel.*

Analogous results hold for Bob.

Proof. Part 1. Since $X^a \in \mathbb{X}^a_{m+1}$, it follows that $X^a \in \mathbb{X}^a_m$. By the definition of $\mathbb{X}^a_m$, there must exist $\nu = (\nu_0, \ldots, \nu_{m-1}) \in \mathcal{N}^+(S^b)$ such that X^a is exactly the optimal set of strategies under ν; and $\operatorname{Supp} \nu_k = S^b_{m-1-k}$ for all $k = 0, \ldots, m-1$. By Lemma B.3.11, there exists $\mu = (\mu_0, \ldots, \mu_{m-1}) \in \mathcal{L}^+(S^b \times T^b)$ such that $\operatorname{marg}_{S^b} \mu = \nu$; μ assumes $\widehat{R}^b_k$ at level $m-1-k$ for all $k = 0, \ldots, m-1$; and μ does not assume

$\widehat{R}^b_m$. It follows that $\Lambda_m(X^a) \setminus \Lambda_{m+1}(X^a)$ is nonempty. Lemma E.2 in BFK implies that for any μ, there exists a continuum of $\widehat{\mu}$ such that $\operatorname{marg}_{S^b} \widehat{\mu} = \operatorname{marg}_{S^b} \mu$ and $\widehat{\mu}$ assumes the same sets as μ.

Part 2. There exists an $M \in \mathbb{N}$ such that $\mathbb{X}^b_\infty = \mathbb{X}^b_M$ since $\left\{\mathbb{X}^b_n : n \in \mathbb{N}\right\}$ is a decreasing sequence of nonempty finite sets. By the definition of $\mathbb{X}^a_{M+1}$, there must exist $\nu = (\nu_0, \ldots, \nu_M) \in \mathcal{N}^+(S^b)$ such that X^a is exactly the optimal set of strategies under ν; and $\operatorname{Supp} \nu_k = S^b_{M-k}$ for all $k = 0, \ldots, M$. Now, take any $\nu'_0 \in \mathcal{M}^+(S^b)$. By Lemma B.4.5, for all $m \geq M$, $\widehat{R}^b_m \setminus \widehat{R}^b_{m+1}$ is $(\nu \otimes \phi^a)$-null and $\nu \otimes \phi^a \in \mathcal{M}^+(S^b \times T^b)$. It follows, by Lemma B.3.12, that there exists $\mu \in \mathcal{L}^+(S^b \times T^b)$ such that $\operatorname{marg}_{S^b} \mu = \nu$ and μ assumes $\widehat{R}^b_m$ for all $m \geq 0$. Therefore, $\Lambda_\infty(X^a)$ is nonempty. Lemma E.2 in BFK implies that for any μ, there exists a continuum of $\widehat{\mu}$ such that $\operatorname{marg}_{S^b} \widehat{\mu} = \operatorname{marg}_{S^b} \mu$ and $\widehat{\mu}$ assumes the same sets as μ.

Part 3. First, the set of all LPSs μ that assume a Borel set is Borel (See Lemma C.3 in BFK, p. 340). Second, the set of all LPSs μ under which a strategy is optimal is Borel since, by the finiteness of the game, optimality is defined by a finite set of inequalities between two continuous real functions of μ (i.e., expected utility under μ). $\Lambda_m(X^a)$ is a set of LPSs μ for which a finite combination of these conditions are satisfied. Therefore, it is a finite intersection of Borel sets, which is also a Borel set. □

Lemma B.4.8. *There exists a Borel isomorphism $\lambda^a : T^a \to \overline{\mathcal{L}}(S^b \times T^b)$ such that for all $m \in \mathbb{N}$,*

1. *For all $X^a \in \mathbb{X}^a_{m+1}$, $\lambda^a \left(\Lambda_m(X^a) \setminus \Lambda_{m+1}(X^a)\right) = \tau_m(X^a) \setminus \tau_{m+1}(X^a)$; and*

2. *For all $X^a \in \mathbb{X}^a_m \setminus \mathbb{X}^a_{m+1}$, $\lambda^a \left(\Lambda_m(X^a)\right) = \tau_m(X^a)$.*

Analogous results hold for Bob.

Proof of Lemma B.4.8. Fix $X^a \in \mathbb{X}_1^a$. For each $m \in \mathbb{N}$, if $X^a \in \mathbb{X}_{m+1}^a$, then $\Lambda_m(X^a) \setminus \Lambda_{m+1}(X^a)$ is uncountable. Consider two cases.

Case 1: There exists a largest M such that $X^a \in \mathbb{X}_M^a$. Then, $\Xi(X^a)$ is a partition of $\Lambda_1(X^a)$. In the case that $M = 1$, $\Xi(X^a) = \{\Lambda_1(X^a)\}$. Each member of $\Xi(X^a)$ is uncountable.

$$\Xi(X^a) \equiv \{\Lambda_1(X^a) \setminus \Lambda_2(X^a), \ldots, \Lambda_{M-1}(X^a) \setminus \Lambda_M(X^a), \Lambda_M(X^a)\}$$

The corresponding partition $\Pi(X^a)$ of $\tau_1(X^a)$ is defined below. Each member of $\Pi(X^a)$ is uncountable.

$$\Pi(X^a) \equiv \{\tau_1(X^a) \setminus \tau_2(X^a), \ldots, \tau_{M-1}(X^a) \setminus \tau_M(X^a), \tau_M(X^a)\}$$

$\Xi(X^a)$ and $\Pi(X^a)$ are equinumerous (in particular, have size M) and each member of each partition is an uncountable Borel set. Therefore, by the Borel Schröder-Bernstein Theorem (See 15.7, 15.8 in Kechris, 1995, pp. 90-91), there exists a Borel isomorphism $f : \tau_1(X^a) \to \Lambda_1(X^a)$ such that

$$\begin{aligned}
f(\tau_1(X^a) \setminus \tau_2(X^a)) &= \Lambda_1(X^a) \setminus \Lambda_2(X^a) \\
f(\tau_2(X^a) \setminus \tau_3(X^a)) &= \Lambda_2(X^a) \setminus \Lambda_3(X^a) \\
\ldots &= \ldots \\
f(\tau_{M-1}(X^a) \setminus \tau_M(X^a)) &= \Lambda_{M-1}(X^a) \setminus \Lambda_M(X^a) \\
f(\tau_M(X^a)) &= \Lambda_M(X^a).
\end{aligned}$$

Case 2: For all $m \in \mathbb{N}$, $X^a \in \mathbb{X}_m^a$. We define the partitions $\Xi(X^a)$ of $\Lambda_1(X^a)$ and

$\Pi(X^a)$ of $\tau_1(X^a)$ as follows.

$$\Xi(X^a) \equiv \{\Lambda_1(X^a) \setminus \Lambda_2(X^a), \Lambda_2(X^a) \setminus \Lambda_3(X^a), \ldots, \Lambda_\infty(X^a)\}$$
$$\Pi(X^a) \equiv \{\tau_1(X^a) \setminus \tau_2(X^a), \tau_2(X^a) \setminus \tau_3(X^a), \ldots, \tau_\infty(X^a)\}$$

$\Xi(X^a)$ and $\Pi(X^a)$ are equinumerous (in particular, countably infinite) and each member of each partition is an uncountable Borel set. Therefore, by the Borel Schröder-Bernstein Theorem (See 15.7, 15.8 in Kechris, 1995, pp. 90-91), there exists a Borel isomorphism $f : \tau_1(X^a) \to \Lambda_1(X^a)$ such that

$$f(\tau_1(X^a) \setminus \tau_2(X^a)) = \Lambda_1(X^a) \setminus \Lambda_2(X^a)$$
$$f(\tau_2(X^a) \setminus \tau_3(X^a)) = \Lambda_2(X^a) \setminus \Lambda_3(X^a)$$
$$\ldots = \ldots$$
$$f(\tau_\infty(X^a)) = \Lambda_\infty(X^a).$$

Since $T^a \setminus \bigcup\{\tau_1(X^a) : X^a \in \mathbb{X}_1^a\}$ and $\overline{\mathcal{L}}(S^b \times T^b) \setminus \bigcup\{\Lambda_1(X^a) : X^a \in \mathbb{X}_1^a\}$ are uncountable, we conclude that there exists a Borel isomorphism $f : T^a \to \overline{\mathcal{L}}(S^b \times T^b)$ such that f satisfies the properties described in the cases above. $\square$

Lemma B.4.9. *Fix a type structure $\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \rangle$ in which λ^a and λ^b are given by Lemma B.4.8. Then, for all $m \in \mathbb{N}$, $\widehat{R}_m^a = R_m^a$ and $\widehat{R}_m^b = R_m^b$.*

Proof of Lemma B.4.9. First, note that for all $m \in \mathbb{N}$ and $X^a \in \mathbb{X}_m^a$, $\lambda^a(\tau_m(X^a)) = \Lambda_m(X^a)$. For all $X^a \in \mathbb{X}_1^a$, $\Lambda_1(X^a)$ is the set of all full-support beliefs under which X^a is exactly the set of optimal strategies. By Corollary 2.4.12, it follows that

$$R_1^a = \{X^a \times \tau_1(X^a) : X^a \in \mathbb{X}_1^a\} = \widehat{R}_1^a.$$

An analogous result holds for Bob.

It follows by induction that for each $m \in \mathbb{N}$, $\tau_{m+1}(X^a)$ is precisely the set of all full-support types such that

1. Each type in $\tau_{m+1}(X^a)$ assumes $R_1^b, \ldots, R_m^b$; and

2. X^a is exactly the set of optimal strategies for each type in $\tau_{m+1}(X^a)$.

It follows that $R_{m+1}^a = \left\{X^a \times \tau_{m+1}(X^a) : X^a \in \mathbb{X}_{m+1}^a\right\} = \widehat{R}_{m+1}^a$. □

Proof of Theorem 2.4.2. Since $R_\infty = \bigcap_{m=1}^{\infty}(\widehat{R}_m^a \times \widehat{R}_m^b) \neq \varnothing$, there exists an RCAR state in $\left\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \right\rangle$. λ^a and λ^b are isomorphisms, and therefore the lexicographic type structure $\left\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \right\rangle$ is complete. □

Bibliography

Kenneth J. Arrow. Alternative approaches to the theory of choice in risk-taking situations. *Econometrica*, 19(4):404–437, 1951.

Robert Aumann. Agreeing to disagree. *Annals of Statistics*, 4(6):1236–1239, November 1976.

Robert Aumann. Correlated equilibrium as an expression of Bayesian rationality. *Econometrica*, 55(1):1–18, January 1987.

Paulo Barelli and Spyros Galanis. Admissibility and event-rationality. mimeo, May 2010.

Pierpaolo Battigalli. Strategic rationality orderings and the best rationalization principle. *Games and Economic Behavior*, 13(2):178–200, April 1996.

Pierpaolo Battigalli and Marciano Siniscalchi. Strong belief and forward induction reasoning. *Journal of Economic Theory*, 106(2):356–391, October 2002.

Elchanan Ben-Porath and Eddie Dekel. Signaling future actions and the potential for sacrifice. *Journal of Economic Theory*, 57(1):36–51, June 1992.

B. Douglas Bernheim. Rationalizable strategic behavior. *Econometrica*, 52(4):1007–1028, July 1984.

Lawrence Blume, Adam Brandenburger, and Eddie Dekel. Lexicographic probabilities and choice under uncertainty. *Econometrica*, 59(1):61–79, January 1991a.

Lawrence Blume, Adam Brandenburger, and Eddie Dekel. Lexicographic probabilities and equilibrium refinements. *Econometrica*, 59(1):81–98, January 1991b.

Adam Brandenburger and Eddie Dekel. Rationalizability and correlated equilibria. *Econometrica*, 55(6):1391–1402, November 1987a.

Adam Brandenburger and Eddie Dekel. Common knowledge with probability 1. *Journal of Mathematical Economics*, 16(3):237–245, 1987b.

Adam Brandenburger and Amanda Friedenberg. The relationship between rationality on the matrix and the tree. `http://pages.stern.nyu.edu/~abranden/papers.html`, 2007.

Adam Brandenburger, Amanda Friedenberg, and H. Jerome Keisler. Admissibility in games. *Econometrica*, 76(2):307–352, March 2008.

N. Dalkey. Equivalence of information patterns and essentially determinate games. In *Contributions to the Theory of Games*, volume 2, pages 217–245. Princeton University Press, Princeton, NJ, 1953.

John C. Harsanyi. Games with incomplete information played by “Bayesian” players, I-III. part I. the basic model. *Management Science*, 14(3):159–182, November 1967.

Alexander S. Kechris. *Classical Descriptive Set Theory*. Springer-Verlag, New York, 1995.

Elon Kohlberg and Jean-François Mertens. On the strategic stability of equilibria. *Econometrica*, 54(5):1003–1037, September 1986.

R. Duncan Luce and Howard Raiffa. *Games and Decisions*. Wiley, New York, NY, 1957.

Dov Monderer and Dov Samet. Approximating common knowledge with common beliefs. *Games and Economic Behavior*, 1(2):170–190, June 1989.

David G. Pearce. Rationalizable strategic behavior and the problem of perfection. *Econometrica*, 52(4):1029–1050, July 1984.

Larry Samuelson. Dominated strategies and common knowledge. *Games and Economic Behavior*, 4(2):284–313, April 1992.

Leonard J. Savage. *The Foundations of Statistics*. Dover, New York, NY, 1954.

Thomas C. Schelling. *The Strategy of Conflict*. Harvard University Press, Cambridge, MA, 1960.

Dale O. Stahl. Lexicographic rationalizability and iterated admissibility. *Economic Letters*, 47(2):155–159, February 1995.

Tommy Chin-Chiu Tan and Sérgio Werlang. The Bayesian foundations of solution concepts of games. *Journal of Economic Theory*, 45(2):370–391, August 1988.

F. B. Thompson. Equivalence of games in extensive form. RAND Memo RM RM-759, The RAND Corporation, Santa Monica, CA, January 1952.

Eric van Damme. Stable equilibria and forward induction. *Journal of Economic Theory*, 48(2):476–496, August 1989.

Eric van Damme and Sjaak Hurkens. Commitment robust equilibria and endogenous timing. *Games and Economic Behavior*, 15(2):290–311, 1996.

John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ, 1944.

Abraham Wald. Contributions to the theory of statistical estimation and testing hypotheses. *The Annals of Mathematical Statistics*, 10(4):239–326, December 1939.

Chih-Chun Yang. Epistemic foundations of admissibility. mimeo, January 2010.

Vita
Byung Soo Lee

Education

The Pennsylvania State University State College, Pennsylvania 2006–2010
Ph.D. in Economics, expected in August 2011.
Requirements completed in July 2010.

Duke University Durham, North Carolina 2004–2006
M.A. in Economics.

Duke University Durham, North Carolina 1999–2003
B.S. in Economics and Computer Science.

Research Interests

Game theory, interactive epistemology, nonstandard probability.