Procedural Type Spaces*

Konrad Grabiszewski[†]

This Version: April 2009

Abstract

Type space is of fundamental importance in epistemic game theory. This paper shows how to build type space if players approach the game in a procedural way advocated by rationalizability. If an agent fixes a strategy profile of her opponents and ponders which of their beliefs about her set of strategies make this profile optimal, such an analysis is represented by transition probabilities and yields disintegrable beliefs. Our construction requires that underlying space is separable.

1 Introduction

Fix a game played by Ann and Bob with their strategy sets, S^a and S^b, respectively. Ann's first-order belief is her conjecture over Bob's choices. It is natural to assume that Bob ponders Ann's strategies, as well, and that Ann knows this. Hence, she tries to link Bob's alternatives with his first-order beliefs. Ann fixes Bob's strategy s^b and selects his conjectures that make s^b optimal. Bob conducts the same analysis and, in consequence, we obtain infinite structures representing players' thinking about the game. This way of interactive reasoning lies behind the concept of rationalizability introduced by Bernheim [1] and Pearce [14]. According to the former:

> *Since the state of the world, as perceived by A, is uncertain, he must construct some assessment of B's action and optimize accordingly. (...) A knows that B has an as-*

*This paper would not exist without the enormous support and guidance of Adam Brandenburger. I thank Amanda Friedenberg for our talks about epistemic game theory and her valuable insight and help. I also thank Jean-Pierre Benoît and David Pearce for their comments. Support from the Asociación Mexicana de Cultura A.C. and the Stern School of Business is gratefully acknowledged.

†Address: Department of Business Administration, ITAM, Rio Hondo 1, Col. Progreso Tizapan, C.P. 01080 Mexico D.F., Mexico, konrad.grabiszewski@itam.mx, http://allman.rhon.itam.mx/~konrad.grabiszewski/

sessment of what A will do for which B's strategy is a best response. (...) A must not only have an assessment of what B will do subject to which A's choice is a best response, but for every forecast of B's strategy to which A ascribes positive probability, A must also be able to construct some conjecture of B's assessment of A's action, for which this forecast of B's strategy is a best response. Since conformity with Savage's axioms is common knowledge, this reasoning can be extended indefinitely. If it is possible to justify the choice of a particular strategy by constructing infinite sequence of self-justifying conjectured assessments in this way, then I call the strategy "rationalizable."

The infinite hierarchy of beliefs and type space were introduced by Harsanyi [6] (see Myerson [12] for a non-technical review). Type space is an essential tool of epistemic game theory (see a recent three-article survey by Brandenburger [3], Heifetz [8], and Siniscalchi [16]). In particular, we want to know whether the collection of all of Ann's hierarchies, T^a, and the space of her beliefs over $S^b \times T^b$, $P^a(S^b \times T^b)$ are homeomorphic. Proving this establishes the existence of the universal type space (see Siniscalchi [16] and Friedenberg [5] for discussion of universal, terminal, and complete type spaces). Coherency of agents' conjectures is a minimal condition. That is, we require the higher and lower order beliefs to agree on appropriate spaces. If this is not true, then it is impossible for an infinite hierarchy to induce a unique belief over $S^b \times T^b$. However, coherency is not enough, as it is shown by Heifetz and Samet [9]. Their result is based on violating topological assumptions in the Kolmogorov Extension Theorem. In order to obtain a positive answer, we need to introduce restriction on either underlying space or on agents' hierarchies. Mertens and Zamir [11] and Brandenburger and Dekel [4] focus on topological constraints. The former assumes the space of uncertainty to be compact, while the latter considers a Polish space. In Heifetz [7], that space is Hausdorff separable[1], and agents' beliefs are defined as regular probability measures.

Our construction of the universal type space is based on reconceptualizing the idea of an agent's belief. We want to capture a procedure leading to rationalizability and, for that reason, we call our belief and type procedural. Ann's first-order belief is defined as a probability λ^a over the set of Bob's strategies, S^b. However, instead of defining higher-order beliefs in a standard way directly on product spaces, we use the notion of a transition probability. A second level of Ann's hierarchy consists of λ^a and a family of transition probabilities between S^b and $P^b(S^a)$. For each of Bob's strategies, s^b, a transition probability, v, assigns Ann's conjecture, $v(s^b)$, over the set of Bob's first-order beliefs, $P^b(S^a)$. In the words of Bernheim, "*A must also be able to construct some conjecture of B's assessment of A's action, for which this forecast of B's strategy is a best response.*" Ann's belief, λ^a, and transition probability, v, generate unique belief over $S^b \times P^b(S^a)$. The collection of transition probabilities that constitutes a part of the second-order belief is determined by the notion

[1] Separability of S is not stated directly in the paper. However, it is implied by the regularity of probability measures, as support of such a measure is separable (see Section II.3 in Parthasarathy [13] or Appendix III in Billingsley [2]).

of equivalency. We say that v and $\widetilde{v}$ are equivalent with respect to λ^a, if they are the same except for the set of measure zero.

In our construction, the second-order belief includes the first-order belief. Defining the former solely as a family of transition probabilities might seem more natural. After all, at the second level, Ann only connects Bob's alternatives with his first-order conjectures. This process does not involve Ann's first-order belief and is represented by transition probabilities. However, we show that excluding first-order belief from the definition of the second-order belief is not desirable as it would not allow us to injectively relate beliefs to the probabilities they generate. In consequence, building homeomorphism between T^a and beliefs over $S^b \times T^b$ would be impossible.

We assume that the underlying space of uncertainty is separable topological, and we prove the existence of the universal type space in this setup. It is important to note that we do not require S^a and S^b to be even topological spaces, in order to prove the existence of a bijective map between Ann's types and her beliefs about $S^b \times T^b$. However, we do need separability for the continuity of this map. This is distinct from the previously mentioned literature, where topological assumptions are necessary for the existence of such a map.

In Section 2, we discuss the idea of a procedural belief, a key element of our analysis. We also propose a topology associated with the set of these beliefs. This topology, as expected, is closely related to the standard weak-* topology on the set of measures. In Section 3, we prove the existence of canonical homeomorphism. In Section 4, we compare the standard construction with our construction of type space. In particular, we show that adding disintegrability conditions to the standard notion of coherent type is equivalent to our procedural approach. Appendix includes proofs not discussed in the main text.

2 Procedural Beliefs

Let X and Y be separable topological spaces endowed with Borel σ-algebras $\mathcal{E}$ and $\mathcal{F}$, respectively. E and F are generic elements of $\mathcal{E}$ and $\mathcal{F}$, respectively. We endow $X \times Y$ with the product σ-algebra. B is its generic element and $B(x)$ is a section of B at x. $P(X)$ denotes a collection of probability measures on X with the weak-* topology assigned to it.

$\nu : X \times \mathcal{F} \to \mathbb{R}$ is a transition probability between X and Y, if for each x, $\nu\ (x; .)$ is probability measure on Y, and for each $F \in \mathcal{F}$, $\nu\,(.; F)$ is measurable function. If λ^X is a probability measure on X, then for a given ν, there is the unique probability $\lambda^{X \times Y}$ on $X \times Y$, such that for each B

$$\lambda^{X \times Y}(B) = \int_X \nu(x; B(x)) d\lambda^X. \tag{1}$$

Any measure on $X \times Y$ that can be represented as in (1) will be called disintegrable with respect to (its marginal) λ^X.

Under what assumptions is a measure disintegrable? Using Valadier [17], we can prove the Disintegration Theorem: if X and Y are Polish, then every probability measure on $X \times Y$ is disintegrable.

We say that two transition probabilities, ν and $\widetilde{\nu}$, are equivalent with respect to λ^X, if for any F, $\nu(x; F) = \widetilde{\nu}(x; F)$ λ^X-a.s. Let $[\nu]/_{\lambda^X}$ denote the family of transition probabilities equivalent to ν, with respect to λ^X. Since equivalent ν and $\widetilde{\nu}$ differ only on the set of measure zero, it is easy to show the following result.

Lemma 2.1

Let ν and $\widetilde{\nu}$ be equivalent with respect to λ^X. For each measurable $B \subset X \times Y$,

$$\lambda^{X\times Y}(B) = \int_X \nu(x; B(x))\, d\lambda^X = \int_X \widetilde{\nu}(x; B(x))\, d\lambda^X = \widetilde{\lambda}^{X\times Y}(B).$$

Lemma 2.1 is essential to understanding our definition of a procedural belief, as well as the construction of the topology on the set of procedural beliefs.

Definition 2.1 ***Procedural Belief***

Fix measurable spaces X and Y. For a probability λ^X and a transition probability ν, between X and Y, procedural belief is λ^X together with a family of transition probabilities equivalent to ν, with respect to λ^X.

Let $K(X;Y)$ denote the collection of procedural beliefs on $X \times Y$. Let K_α be a procedural belief for some $(\lambda^X_\alpha, \nu_\alpha)$. Lemma 2.1 says that each procedural belief yields unique disintegrable probability on $X \times Y$. In Lemma 2.2, we show the inverse: for a $\lambda^{X\times Y}$ that is disintegrable with respect to its marginal λ^X, there exists a unique procedural belief that generates it.

Lemma 2.2

Take distinct and disintegrable $\lambda^{X\times Y}$ and $\widetilde{\lambda}^{X\times Y}$. There exist unique and distinct K and $\widetilde{K}$ that generate $\lambda^{X\times Y}$ and $\widetilde{\lambda}^{X\times Y}$, respectively.

The reader may ponder why we require λ^X to be part of a procedural belief. Alternatively, we could define procedural belief as a collection of transition probabilities equivalent to some ν, with respect to some measure. This might seem to be a more natural construction. If the first-level belief is a probability measure on X, then the second-level belief would be some family of transition probabilities between X and Y. However, this definition would make it impossible to bijectively relate procedural beliefs to probabilities on $X \times Y$. The result in Lemma 2.3 serves as our argument.

Lemma 2.3

Fix a transition probability v. There are distinct λ^X and $\widetilde{\lambda}^X$ such that $[\nu]/_{\lambda^X} = [\nu]/_{\widetilde{\lambda}^X}$.

To prove Lemma 2.3, consider two probabilities that are absolutely continuous with respect to each other. Both assign measure zero to exactly the same sets. In consequence, ν and $\widetilde{\nu}$ are equivalent with respect to λ^X and $\widetilde{\lambda}^X$. For a given transition probability, ν, both probabilities generate the same family of equivalent transition probabilities. Thus, collection of such families cannot be injectively mapped to the set of measures on X.

Observe that a procedural belief is a linear bounded operator on the space of real-valued, continuous bounded functions on $X \times Y$, $CB(X \times Y)$. We assign the weak-* topology to $K(X;Y)$. That is, K_α converges to K if and only if, for every $f \in CB(X \times Y)$,

$$\int_X d\lambda_\alpha^X \int_Y f\nu_\alpha(x;dy) \to \int_X d\lambda^X \int_Y f\nu(x;dy). \tag{2}$$

From the (generalized) Fubini Theorem, we know that $\int_X d\lambda_\alpha^X \int_Y f\nu_\alpha(x;dy) = \int_{X\times Y} fd\lambda_\alpha^{X\times Y}$, where $\lambda_\alpha^{X\times Y}$ is the unique measure constructed from K_α. Thus, the convergence of procedural beliefs is a weak-* convergence of the measures these beliefs yield.

For a product of more than two spaces, $\nu^{0,..,n-1;n}$ will denote a transition probability between $X_1 \times ... \times X_{n-1}$ and X_n. Let $K^{0,...,n-1;n}$ be a procedural belief induced by $\lambda^{0,...,n-1}$ and $\nu^{0,..,n-1;n}$. Let $\widetilde{P}(X_0 \times ... \times X_n)$ denote the collection of fully disintegrable measures on $X_1 \times ... \times X_n$. That is, $\lambda^{0,...,n} \in \widetilde{P}(X_0 \times ... \times X_n)$, if there is a probability λ^0 on X_0 and a collection of transition probabilities $\{\nu^{0;1}, \nu^{0,1;2}, ..., \nu^{0,...,n-1;n}\}$ such that for any measurable $B \subset X_0 \times ... \times X_n$, we have:

$$\lambda^{0,...,n}(B) = \int_{X_0} d\lambda^0 \int_{X_1} \nu^{0;1}(x_0;dx_1) ... \int_{X_n} 1_B \nu^{0,...,n-1;n}(x_0,...,x_{n-1};dx_n)$$

where 1_B is an indicator function. We can construct $\widetilde{P}(X_0 \times ... \times X_n)$ inductively. First, we take the elements of $P(X_0)$ and combine these with all transition probabilities $\nu^{0;1}$, as in (1). We obtain $\widetilde{P}(X_0 \times X_1)$. Next, we take these probabilities and combine them with transition probabilities $\nu^{0,1;2}$, again using (1) to obtain $\widetilde{P}(X_0 \times X_1 \times X_2)$, and so on. It is important to note that the construction of $\widetilde{P}(X_0 \times X_1 \times X_2)$ cannot be based on $P(X_0 \times X_1)$. Instead, we need $\widetilde{P}(X_0 \times X_1)$; otherwise, we obtain non-fully disintegrable measures.

Our choice of topology for $K(X;Y)$ turns out to be very useful. First, it implies that $K(X;Y)$ and $\widetilde{P}(X \times Y)$ are topologically equivalent spaces.

Lemma 2.4

There exists a homeomorphism, $\gamma : K(X;Y) \to \widetilde{P}(X \times Y)$.

Second, following Varadarajan [18], we can show that if X and Y are separable (Polish) metric spaces, then $K(X;Y)$ is metrizable as a separable (Polish) metric space. Metrizability is required and expected, as similar results obtain for a set of Borel probabilities.

The existing literature uses the standard construction of beliefs on products. That is, these probabilities are directly defined on $X \times Y$. Our notion of belief is different and less general because not all probabilities on $X \times Y$ can be represented, as in (1). However, topological assumptions in Mertens and Zamir [11] and Brandenburger and Dekel [4] eliminate their advantage of a more general definition of belief. According to the Disintegrability Theorem, their agents' beliefs are disintegrable and, hence, can be represented as procedural beliefs. Heifetz [7] assumes less in terms of topological requirements than the Disintegrability Theorem demands. However, his construction is not more general than ours. Beliefs in Heifetz [7] are defined as regular probability measures and as shown in Leão Jr., et al. [10], they are disintegrable. On the other hand, we can construct a disintegrable measure on $X \times Y$ that is not regular. Suppose that both X and Y are separable metric spaces; however, X is not complete. There is a probability measure, λ^X, on X that is not regular. In consequence, λ^X and a transition probability, ν, generate a non-regular measure, $\lambda^{X\times Y}$, on $X \times Y$ (see Remark 2, Appendix III in Billingsley [2] for details).

3 Procedural Type Spaces

Let S be a topological space endowed with Borel σ-algebra. This is an uncertainty space faced by the players. A procedural type is an infinite collection of procedural beliefs. In order to construct it, we inductively define spaces:

$\Omega_0 := S$

$\Omega_1 := P(S)$

$\Omega_2 := K(S; P(S)) = K(\Omega_0; \Omega_1)$

$\vdots$

$\Omega_n := K(\Omega_0 \times ... \times \Omega_{n-2}; \Omega_{n-1})$.

Let $W_0 := \underset{i=1}{\times} \Omega_i$ be the (canonical) space of procedural types, with generic element, $w := (\lambda^0, K^{0;1}, K^{0,1;2}, ...)$. Each $K^{0,...,n-2;n-1}$ consists of a measure, $\lambda^{0,...,n-2}$, on $\Omega_0 \times ... \times \Omega_{n-2}$ and a family of equivalent transition probabilities between $\Omega_0 \times ... \times \Omega_{n-2}$ and Ω_{n-1}. Equivalency is defined with respect to $\lambda^{0,...,n-2}$. Belief $K^{0,...,n-2;n-1}$ generates a unique disintegrable probability over $\Omega_0 \times ... \times \Omega_{n-1}$.

We say that type w is coherent, if, for each $n \geq 2$, $\lambda^{0,...,n-2}$ is a probability induced by the preceding level. Let W_1 be the set of coherent types. W is the set of types that satisfy both coherency and the common belief of coherency. To believe an event means to assign measure 1 to it.

We want to show that W and $\widetilde{P}(S \times W)$ are homeomorphic. First, we prove that $p : W_1 \to \widetilde{P}(S \times W_0)$ is a homeomorphism.

Lemma 3.1

If S is separable, then there exists a homeomorphism, $p : W_1 \to \widetilde{P}(S \times W_0)$.

To show that p is a function, we take $w \in W_1$. The Ionescu-Tulcea Theorem (see Chapter II.9 in Shiryaev [15]) demonstrates the existence of a unique probability, λ^∞, on $S \times W_0$, which we denote as $p(w)$. To show injectivity, take distinct w and $\widetilde{w}$. There is n such that w and $\widetilde{w}$ generate distinct $\lambda^{0,...,n}$ and $\widetilde{\lambda}^{0,...,n}$. Since w and $p(w)$ agree on cylinders, $p(w) \neq p(\widetilde{w})$. Surjectivity is a consequence of taking $\widetilde{P}(S \times W_0)$, instead of $P(S \times W_0)$, as a range of p.

The standard method of constructing type spaces is based on the Kolmogorov Extension Theorem. In contrast to that approach, we do not need to make any topological assumptions to prove the existence of p. However, we require S to be separable for the continuity of p, as we use the convergence-determining-class technique (see Chapter 1.2 in Billingsley [2]).

In order to prove that $q : W \to \widetilde{P}(S \times W)$ is a homeomorphism, we use the approach based on Proposition 2 in Brandenburger and Dekel [4].

Proposition 3.1

If S is separable, then there exists a homeomorphism, $q : W \to \widetilde{P}(S \times W)$.

First, we prove that W_1 is closed in W_0. Take a sequence of coherent types $\{w_t\}$ that converge to w. Each w_t generates an infinite hierarchy of probabilities $(\lambda_t^0, \lambda_t^{0,1}...)$. That is, we have convergence $(\lambda_t^0, \lambda_t^{0,1}...) \longrightarrow (\lambda^0, \lambda^{0,1}...)$. Note that if $\{\lambda_t^{0,...,n}\}$ converges to $\lambda^{0,...,n}$, then the sequence of marginals of $\{\lambda_t^{0,...,n}\}$ converges to the marginal of $\lambda^{0,...,n}$. Hence, w is coherent. Based on this result, we conclude that each $W_n := \{w \in W_1 : p(w)(S \times W_{n-1}) = 1\}$ is also closed in W_0. Note that $W_0 \supset W_1$ and, by induction, $W_0 \supset W_1 \supset W_2....$ We define W as an intersection, $W = \bigcap_{n=1}^\infty W_n$. Using the continuity of probability measure and the fact that $\{W_n\}$ is a decreasing sequence, we can show that $W = \{w \in W_1 : p(w)\,(S \times W) = 1\}$. From this, we derive that $p(W) = \{\lambda^\infty \in \widetilde{P}(S \times W_0) : \lambda^\infty\,(S \times W) = 1\}$. Since $\{\lambda^\infty \in \widetilde{P}(S \times W_0) : \lambda^\infty\,(S \times W) = 1\}$ and $\widetilde{P}(S \times W)$ are homeomorphic and $p(W)$ and W are homeomorphic, we deduce the desired relation.

4 Procedural and Standard Type Spaces

We inductively define the standard type:

$X_0 := S$

$X_1 := X_0 \times P(X_0)$

$\vdots$

$X_n := X_{n-1} \times P(X_{n-1})$.

$T_0 := \underset{i=0}{\times} P(X_i)$ is the space of standard types with a generic element, $t = (\mu^0, \mu^{0,1}, ...)$. We say that a standard type is disintegrable if each $\mu^{0,...,n}$ is disintegrable, as in (1). However, such a type does not to have to be coherent, as in Brandenburger and Dekel [4]. Hence, we say that a type is coherently disintegrable if disintegration of the n-level conjecture is conducted with respect to the $n-1$-level belief. Let T_1 be the set of standard types that satisfy coherent disintegrability, while T is the set of types that satisfy both coherent disintegrability and the common belief of coherent disintegrability. We prove the existence of homeomorphisms, $f : T_1 \to \widetilde{P}(S \times T_0)$ and $g : T \to \widetilde{P}(S \times T)$.

Lemma 4.1

If S is separable, then there exists a homeomorphism, $f : T_1 \to \widetilde{P}(S \times T_0)$.

Proposition 4.1

If S is separable, then there exists a homeomorphism, $g : T \to \widetilde{P}(S \times T)$.

The existence of f does not require S to be Polish, as in the Kolmogorov Extension Theorem. This is derived from the fact that the beliefs are disintegrable and, hence, representable, as in (1). Once again, we use the Ionescu-Tulcea Theorem. Proof of bijectivity and bicontinuity follows the same technique we employed in the previous section.

Note that, in terms of topological assumption, we require less than Brandenburger and Dekel [4]. However, this not a gain at zero cost. Our requirement of beliefs being coherently disintegrable is stronger than the sole coherency they require.

The next proposition describes the relationship between the procedural and standard types.

Proposition 4.2

Suppose that S is separable; we have the following commutative diagram:

$$\begin{array}{ccc} W & \xrightarrow{\Psi} & T \\ {\scriptstyle q}\big\downarrow & & \big\downarrow{\scriptstyle g} \\ \widetilde{P}\left(S \times W\right) & \xrightarrow[\varphi]{} & \widetilde{P}\left(S \times T\right) \end{array}$$

In Propositions 3.1 and 4.1, we established the existence of homeomorphisms, q and g. Next, we show that two notions of type space – the procedural under coherency and common belief of coherency, as well as the standard under coherent disintegrability and common belief of coherent disintegrability – are topologically equivalent.

Proposition 4.3

There exists a homeomorphism, $\Psi : W \to T$.

In order to show that $\widetilde{P}(S \times W)$ and $\widetilde{P}(S \times T)$ are homeomorphic spaces, we need the following result.

Lemma 4.2

Suppose $\Psi_1 : X_1 \to Y_1$ and $\Psi_2 : X_2 \to Y_2$ are homeomorphisms. Then there exists homeomorphism, $\varphi : \widetilde{P}\left(X_1 \times X_2\right) \to \widetilde{P}\left(Y_1 \times Y_2\right)$.

In order to apply Lemma 4.2, replace both X_1 and Y_1 with S, X_2 with W and Y_2 with T. Let Ψ_1 be an identity function, i, and let Ψ_2 be Ψ from Proposition 4.3. Since both i and Ψ are homeomorphisms, there exists a homeomorphism, $\varphi : \widetilde{P}(S \times W) \to \widetilde{P}(S \times T)$.

Finally, we need to show the commutativity. That is, for any $w \in W$ and measurable $B \subseteq S \times T$, we have $\varphi\left(q(w)\right)(B) = g\left(\Psi(w)\right)(B)$. Note that q is a restriction of p on W. Hence, $q(w)(A) = p(w)(A)$ for any measurable $A \subseteq S \times W$. By the construction of φ in Lemma 4.2, we know that $\varphi\left(q(w)\right)(B) = q(w)(\overline{\Psi}^{-1}(B)) = p(w)(\overline{\Psi}^{-1}(B))$, where $\overline{\Psi} := (i, \Psi)$ is a homeomorphism $\overline{\Psi} : S \times W \to S \times T$. According to Claim 5.5 in the Appendix, $p(w)(\overline{\Psi}^{-1}(B)) = f\left(\Psi(w)\right)(B)$. Thus, $\varphi(q(w))(B) = f\left(\Psi(w)\right)(B)$. Since g and f are related in the same way as p and q, we have $f\left(\Psi(w)\right)(B) = g\left(\Psi(w)\right)(B)$. This concludes our proof of commutativity.

5 Appendix

Proof of Lemma 2.2

Fix disintegrable $\lambda^{X\times Y}$. From (1), we obtain marginal λ^X and transition probability, v. Collecting all transition probabilities equivalent to v with respect to that marginal creates a procedural belief, K. We need to prove that different procedural beliefs generate different probabilities on $X \times Y$. Take distinct K and $\widetilde{K}$. Either marginals on X or families of transition probabilities that make these two differ. If $\lambda^X \neq \widetilde{\lambda}^X$, then $\lambda^{X\times Y} \neq \widetilde{\lambda}^{X\times Y}$ as their marginals disagree. Thus, consider the case of $[\nu]/_{\lambda^X} \neq [\widetilde{\nu}]/_{\widetilde{\lambda}^X}$. This implies that neither ν is a member of $[\widetilde{\nu}]/_{\widetilde{\lambda}^X}$ nor $\widetilde{\nu}$ is a member of $[\nu]/_{\lambda^X}$. We need to prove the following result.

Claim 5.1 *Let ν and $\widetilde{\nu}$ be such that $\forall B$, $\lambda^{X\times Y}(B) = \widetilde{\lambda}^{X\times Y}(B)$. Then, ν and $\widetilde{\nu}$ are equivalent transition probabilities with respect to $\lambda^X = \widetilde{\lambda}^X$.*

> Proof: Fix some measurable subset, F of Y. Let $X = C_1 \cup C_2 \cup C_3$ be a disjoint decomposition of X, such that $C_1 := \{x : \nu(x; F) > \widetilde{\nu}(x; F)\}$, $C_2 := \{x : \nu(x; F) < \widetilde{\nu}(x; F)\}$, and $C_3 := \{x : \nu(x; F) = \widetilde{\nu}(x; F)\}$. We know that these sets are measurable events, since ν and $\widetilde{\nu}$ are measurable functions. By assumption, $\int_{C_k} \nu(x; F)\, d\lambda^X = \int_{C_k} \widetilde{\nu}(x; F)\, d\lambda^X$ for $k = 1, 2, 3$. Take C_1 and suppose that $\lambda^X(C_1) > 0$. Then, $\int_{C_1} \nu(x; F)\, d\lambda^X > \int_{C_1} \widetilde{\nu}(x; F)\, d\lambda^X$. Since this is a contradiction, $\lambda^X(C_1) = 0$. The same holds for C_2. ■

Assume that $\lambda^{X\times Y} = \widetilde{\lambda}^{X\times Y}$. The above claim implies that ν and $\widetilde{\nu}$ are equivalent transition probabilities. Contradiction. ■

Proof of Lemma 2.4

Let $\gamma : K(X;Y) \to \widetilde{P}(X \times Y)$ be a natural relation, where $\gamma(K)$ is a probability measure induced by procedural belief, K. By Lemma 2.1, γ exists as each K induces unique probability, $\lambda^{X\times Y}$. By Lemma 2.2, γ is injective and surjective. To show the continuity of γ, take $\{K_\alpha\}$ that converge to K. The Fubini Theorem implies that $\int_{X\times Y} f d\lambda_\alpha^{X\times Y} \to \int_{X\times Y} f d\lambda^{X\times Y} \quad \forall f \in CB(X\times Y)$, where $\lambda_\alpha^{X\times Y} = \gamma(K_\alpha)$ and $\lambda^{X\times Y} = \gamma(K)$. That is, $\{\gamma(K_\alpha)\}$ weak-* converges to $\gamma(K)$. Continuity of γ^{-1} can be proved in the same way. ■

Proof of Proposition 4.3

Summary of the proof:

1. We define a special space, $\widetilde{T}_0$.
2. We show that W_0 and $\widetilde{T}_0$ are homeomorphic under Ψ (Claim 5.2).
3. We show that $\Psi(W_k) = T_k$ for $k = 1, 2, ...$ (Claim 5.7).
4. Take $T = \cap T_k = \cap \Psi(W_k) = \Psi(\cap W_k) = \Psi(W)$, where we use the fact that Ψ is the inverse of Ψ^{-1}. Thus, Ψ restricted to W carries it homeomorphically to T.

We inductively define a special space, $\widetilde{T}_0$.

$\widetilde{Z}_0 := X_0 = S$

$\widetilde{Z}_1 := P(\widetilde{Z}_0)$

$\widetilde{Z}_2 := \widetilde{P}(\widetilde{Z}_0 \times \widetilde{Z}_1)$

$\vdots$

$\widetilde{Z}_n := \widetilde{P}(\widetilde{Z}_0 \times ... \times \widetilde{Z}_{n-1})$

Let $\widetilde{T}_0 = \widetilde{Z}_1 \times \widetilde{Z}_2 \times \widetilde{Z}_3 \times ....$ Note that $T_1 \subseteq \widetilde{T}_0 \subseteq T_0$. If $t \in \widetilde{T}_0$, then at each level, we have disintegrable belief. In fact, it is commonly believed that players have disintegrable beliefs, but neither coherency nor belief in coherency is sustained.

Let $\mu^{0,...,n-1}$ be a generic element of $\widetilde{Z}_n$. It is a fully disintegrable probability on $\widetilde{Z}_0 \times ... \times \widetilde{Z}_{n-1}$. Let $\lambda^{0,...,n}$ be a fully disintegrable probability on $\Omega_0 \times ... \times \Omega_n$.

Claim 5.2 *There exists a homeomorphism,* $\Psi : W_0 \to \widetilde{T}_0$.

Proof: We construct Ψ explicitly. Take $w \in W_0$ and write $\Psi((w_1, w_2, ...)) := (\psi_1(w_1), \psi_2(w_2), ...)$. We need to inductively define $\psi_1 : \Omega_1 \to \widetilde{Z}_1, \psi_2 : \Omega_2 \to \widetilde{Z}_2, ...$

Since $\Omega_0 = \widetilde{Z}_0 = S$ and $\Omega_1 = \widetilde{Z}_1 = P(S)$, we define $\psi_0 : \Omega_1 \to \widetilde{Z}_1$ and $\psi_1 : \Omega_1 \to \widetilde{Z}_1$ as identity functions. Since $\Omega_2 = K(\Omega_0; \Omega_1) = K(\widetilde{Z}_0; \widetilde{Z}_1)$ and $\widetilde{Z}_2 = \widetilde{P}(\widetilde{Z}_0 \times \widetilde{Z}_1)$, we set $\psi_2 := \gamma_2$, where γ_2 is from Lemma 2.4.

For $n \geq 3$, we conduct a four-step procedure:

1. Let $\gamma_n : \Omega_n \to \widetilde{P}(\Omega_0 \times ... \times \Omega_{n-1})$ be a homeomorphism of Lemma 2.4.

2. Define $\Psi_{0,...,n-1} : \Omega_0 \times ... \times \Omega_{n-1} \to \widetilde{Z}_0 \times ... \times \widetilde{Z}_{n-1}$ as $\Psi_{0,...,n-1}(w_0, ..., w_{n-1}) := (\psi_0(w_0), ..., \psi_{n-1}(w_{n-1}))$. Below, we show that this is a homeomorphism.

3. Let $\varphi_n : \widetilde{P}(\Omega_0 \times ... \times \Omega_{n-1}) \to \widetilde{P}(\widetilde{Z}_0 \times ... \times \widetilde{Z}_{n-1})$ be a homeomorphism of Lemma 4.2, constructed with the help of $\Psi_{0,...,n-1}$.

4. Combine results, $\psi_n := \varphi_n \circ \gamma_n : \Omega_n \to \widetilde{Z}_n$.

We show inductively that $\Psi_{0,...,n-1}$ is a homeomorphism. First, note that $\Psi_{0,1} = (\psi_0, \psi_1)$ is homeomorphic. Thus, φ_n exists and is a homeomorphism, making the ψ_n homeomorphism. Hence, $\Psi_{0,1,2} = (\psi_0, \psi_1, \psi_2)$ is a homeomorphism, and so on. Since every ψ_n is homeomorphic, Ψ is homeomorphic, as well. ■

The next two Claims, 5.3 and 5.4, are auxiliary results that we need in order to prove Claim 5.6. We omit their proofs, as they are easily replicated manipulations.

Claim 5.3 *Take $w_n \in \Omega_n$ and $w_{n+1} \in \Omega_{n+1}$, such that $\gamma_n(w_n) = \lambda^{0,...,n-1}$ is a probability on $\Omega_0 \times ... \times \Omega_{n-1}$, which is a part of w_{n+1}. Then*

(a) $\gamma_{n+1}(w_{n+1}) = \lambda^{0,...,n}$ is disintegrable with respect to $\gamma_n(w_n) = \lambda^{0,...,n-1}$; and

(b) $\psi_{n+1}(w_{n+1}) = \mu^{0,...,n}$ is disintegrable with respect to $\psi_n(w_n) = \mu^{0,...,n-1}$.

Claim 5.4 *Take $\mu^{0,...,n} \in \widetilde{Z}_{n+1}$, which is disintegrable with respect to $\mu^{0,...,n-1} \in \widetilde{Z}_n$. Then*

(a) $\varphi_{n+1}^{-1}(\mu^{0,...,n}) = \lambda^{0,...,n}$ is disintegrable with respect to $\varphi_n^{-1}(\mu^{0,...,n-1}) = \lambda^{0,...,n-1}$; and

(b) $\psi_n^{-1}(\mu^{0,...,n-1}) = w_n$ and $\psi_{n+1}^{-1}(\mu^{0,...,n}) = w_{n+1}$ are such that $\lambda^{0,...,n-1}$ generated by w_n via γ_n is a probability on $\Omega_0 \times ... \times \Omega_{n-1}$, which is a part of w_{n+1}.

The next result relates homeomorphisms, p and f, of Lemmas 3.1 and 4.1, respectively. We will also use this to prove Claim 5.7.

Claim 5.5 *Let $\overline{\Psi} := (i, \Psi)$.*

(a) For any $w \in W_1$, $p(w)(\overline{\Psi}^{-1}(B)) = f(\Psi(w))(B)$, where $B \subseteq S \times \widetilde{T}_0$.

(b) For any $t \in T_1$, $f(t)(\overline{\Psi}(A)) = p(\Psi^{-1}(t))(A)$, where $A \subseteq S \times W_0$.

Proof: First, we show that our claim holds for cylinders. Extending it to all measurable events follows the technique based on the Sierpinski Class Lemma.

(a) Take $w \in W_1$. Let $\lambda^{0,...,n} = \gamma_{n+1}(w_{n+1})$. Take arbitrary n and any measurable $B^{0,...,n} \subset \widetilde{Z}_0 \times ... \times \widetilde{Z}_n$. Let $C^n := B^{0,...,n} \times \widetilde{Z}_{n+1} \times ...$ be a cylinder with base $B^{0,...,n}$.

Observe that $\overline{\Psi}^{-1}(C^n) = \Psi^{-1}_{0,\dots,n}(B^{0,\dots,n}) \times \Omega_{n+1} \times ....$

$$\begin{aligned} p(w)(\overline{\Psi}^{-1}(C^n)) &= p(w)(\Psi^{-1}_{0,\dots,n}(B^{0,\dots,n}) \times \Omega_{n+1} \times ...) \\ &= \lambda^{0,\dots,n}(\Psi^{-1}_{0,\dots,n}(B^{0,\dots,n})) \\ &= \varphi_{n+1}(\lambda^{0,\dots,n})(B^{0,\dots,n}) \\ &= \varphi_{n+1}(\gamma_{n+1}(w_{n+1}))(B^{0,\dots,n}) \\ &= \psi_{n+1}(w_{n+1})(B^{0,\dots,n}) \\ &= f(\Psi(w))(C^n) \end{aligned}$$

The second equality follows the Ionescu-Tulcea Theorem. The third, forth, and fifth use constructions of φ_{n+1} and ψ_{n+1}, which we conducted in Claim 5.2. Coherency implies the last equality.

(b) Take t. There is unique w, such that $\Psi(w) = t$. Take that w. By (a), $p(w)(\overline{\Psi}^{-1}(B)) = f(\Psi(w))(B)$. That is, $p(\Psi^{-1}(t))(\overline{\Psi}^{-1}(B)) = f(t)(B)$ for any B. Take B, such that $B = \overline{\Psi}(A)$. Then, $p(\Psi^{-1}(t))(A) = f(t)(\overline{\Psi}(A))$. ■

Claim 5.6 $\Psi(W_1) = T_1$.

Proof: Take $w \in W_1$. We show that $\psi(w) \in T_1$. Notice that $\Psi(w) := (\psi_1(w_1), \psi_2(w_2), ...) = (\mu^0, \mu^{0,1}, ...)$. According to Claim 5.3, $\mu^{0,\dots,n}$ is disintegrable, with respect to $\mu^{0,\dots,n-1}$. Thus, $(\mu^0, \mu^{0,1}, ...)$ satisfies disintegrability. Take $t \in T_1$. We show that $\Psi^{-1}(t) \in W_1$. Notice that $\Psi^{-1}(t) = (\psi_1^{-1}(\mu^0), \psi_2^{-1}(\mu^{0,1}), ...) = (w_1, w_2, ...)$. According to Claim 5.4, from a couple $(\mu^{0,\dots,n-1}, \mu^{0,\dots,n})$, where $\mu^{0,\dots,n}$ is disintegrable with respect to $\mu^{0,\dots,n-1}$, we obtain a couple (w_n, w_{n+1}), where $\lambda^{0,\dots,n-1}$ generated by w_n is a probability on $\Omega_0 \times ... \times \Omega_{n-1}$ that is a part of w_{n+1}. Thus, $(w_1, w_2, ...)$ satisfies disintegrability. ■

Claim 5.7 $\Psi(W_k) = T_k$ *for* $k = 1, 2, ...$

Proof: The proof is solved by induction. $k = 1$ is proved in Claim 5.6. Suppose that claim is true for k, and we verify it for $k+1$. Take $w \in W_{k+1}$. By assumption, $p(w)(S \times W_k) = 1$, and $W_k = \Psi^{-1}(T_k)$. Thus, $p(w)(S \times \Psi^{-1}(T_k)) = p(w)(\overline{\Psi}^{-1}(S \times T_k)) = 1$. According to Claim 5.5, this means that $f(\Psi(w))(S \times T_k) = 1$. That is, $\Psi(w) \in T_{k+1}$. Take $t \in T_{k+1}$. By assumption, $f(t)(S \times T_k) = 1$, and $W_k = \Psi^{-1}(T_k)$. Thus, $f(t)(S \times \Psi(W_k)) = f(t)(\overline{\Psi}(S \times W_k)) = 1$. According to Claim 5.5, this means that $p(\Psi^{-1}(t))(S \times W_k) = 1$. That is, $\Psi^{-1}(t) \in W_{k+1}$. ■

Proof of Lemma 4.2

Fix topological spaces X_1, X_2, Y_1, and Y_2. Let $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{F}_1$, and $\mathcal{F}_2$ be their Borel σ-algebras, respectively. To show the existence of φ, take $\lambda^{1,2} \in \widetilde{P}(X_1 \times X_2)$. Then, take $F_1 \in \mathcal{F}_1$ and $F_2 \in \mathcal{F}_2$. We construct $\mu^{1,2} \in \widetilde{P}(Y_1 \times Y_2)$. We disintegrate $\lambda^{1,2}$ with respect to its marginal λ^1. Let $\nu^{1;2}$ be a transition probability that satisfies $\lambda^{1,2}(E_1 \times E_2) = \int_{E_1} d\lambda^1 \int_{E_2} \nu^{1;2}(x_1; dx_2)$, where $E_1 \in \mathcal{E}_1$ and $E_2 \in \mathcal{E}_2$.

We construct $\mu^{1,2}$ by producing marginal, μ^1, and transition probability, $\omega^{1;2}$. Assign $\mu^1(F_1) := \lambda^1(\Psi_1^{-1}(F_1))$. This definition makes sense since $\Psi_1^{-1}(F_1) \in \mathcal{E}_1$ as Ψ_1 is homeomorphic. In a standard way, we can verify that this yields a measure.

Assign $\omega^{1;2}(y_1; F_2) := \nu^{1;2}(\Psi_1^{-1}(y_1); \Psi_2^{-1}(F_2))$. This definition makes sense because $\Psi_1^{-1}(y_1)$ is a singleton, as Ψ_1 is injective. We can verify that $\omega^{1;2}(y_1; .)$ is a probability measure, as we did this for μ^1. Measurability of $\omega^{1;2}(.; F_2)$ comes from the fact that it is a composition of two measurable functions, $\nu^{1;2}(.; \Psi_2^{-1}(F_2)) \circ \Psi_1^{-1}$.

Let $\Psi := (\Psi_1, \Psi_2) : X_1 \times X_2 \to Y_1 \times Y_2$. This is a homeomorphism. Then, take $E \in \mathcal{E}_1 \otimes \mathcal{E}_2$. We show that $\lambda^{1,2}(E) = \varphi(\lambda^{1,2})(\Psi(E))$. For simplicity, $\varphi(\lambda^{1,2}) = \mu^{1,2}$. Since probability is uniquely defined on rectangles, we take $E = E_1 \times E_2$. Now, we have $\Psi(E) = \Psi(E_1 \times E_2) = \Psi_1(E_1) \times \Psi_2(E_2) = F_1 \times F_2$. We need the following result:

Claim 5.8 *Let $h : (X, \mathcal{E}) \to (Y, \mathcal{F})$ be measurable and bijective. Let λ be measure on X and let f be real-valued, non-negative function defined on X. Define measure and function on Y by $\mu(F) := \lambda(h^{-1}(F))$ and $g(y) := f(h^{-1}(y))$. These definitions make sense, due to assumptions on h. Then,*

$$\int_Y g(y) d\mu = \int_X f(h^{-1}(y)) d\lambda.$$

We omit the proof, as it is derived in a standard way from the Monotone Convergence Theorem.

Fix F_2. Following notation from Claim 5.8, we take $\Psi_1 \equiv h, \mu^{1,2} \equiv \mu, \lambda \equiv \lambda^{1,2}$, $1_{F_1}\omega^{1;2}(.; F_2) \equiv g$ and $1_{\Psi_1^{-1}(F_1)}\nu^{1;2}(\Psi_1^{-1}(.); \Psi_2^{-1}(F_2)) = f$. We obtain the desired result:

$$\begin{aligned}
\mu^{1,2}(F_1 \times F_2) &= \int_{Y_1} 1_{F_1} \omega^{1;2}(y_1; F_2)\, d\mu^1 \\
&= \int_{X_1} 1_{\Psi_1^{-1}(F_1)} \nu^{1;2}\left(\Psi_1^{-1}(.); \Psi_2^{-1}(F_2)\right) d\lambda^1 \\
&= \lambda^{1,2}(\Psi^{-1}(F_1 \times F_2)).
\end{aligned}$$

In the same way, we show that $\lambda^{1,2}(E_1 \times E_2) = \mu^{1,2}(\Psi(E_1 \times E_2))$. This implies that $\lambda^{1,2}(E) = \mu^{1,2}(\Psi(E))$, and $\mu^{1,2}(F) = \lambda^{1,2}(\Psi^{-1}(F))$.

To show that φ is injective, take distinct $\lambda^{1,2}, \widetilde{\lambda}^{1,2} \in \widetilde{P}(X_1 \times X_2)$. Then, there is $E \in \mathcal{E}_1 \otimes \mathcal{E}_2$, such that $\lambda^{1,2}(E) \neq \widetilde{\lambda}^{1,2}(E)$. Let $F := \Psi(E)$. Thus, $\varphi(\lambda^{1,2})(F) \neq \varphi(\widetilde{\lambda}^{1,2})(F)$. To show that φ is surjective, take $\mu^{1,2} \in \widetilde{P}(Y_1 \times Y_2)$. We define disintegrable measure on $X_1 \times X_2$, using Ψ_1 and Ψ_2 in the same way that we constructed a measure on $Y_1 \times Y_2$, based on $\lambda^{1,2}$, at the beginning.

Continuity of φ and φ^{-1} is based on the Portmanteau Theorem. We prove only the former. Let $\lambda_n^{1,2}$ weak-* converge to $\lambda^{1,2}$. Let $\mu_n^{1,2} := \varphi(\lambda_n^{1,2})$ and $\mu^{1,2} := \varphi\left(\lambda^{1,2}\right)$. Let F be $\mu^{1,2}$-continuity set. That is, $\mu^{1,2}(\partial F) = 0$, and $\lambda^{1,2}(\Psi^{-1}(\partial F)) = 0$. We know that $\Psi^{-1}(\partial F) = \partial\Psi^{-1}(F)$. As such, $\lambda^{1,2}(\partial\Psi^{-1}(F)) = 0$, which implies that $\Psi^{-1}(F)$ is $\lambda^{1,2}$-continuity set. Thus, by the Portmanteau Theorem $\lim \lambda_n^{1,2}(\Psi^{-1}(F)) = \lambda^{1,2}(\Psi^{-1}(F))$, and since $\lambda_n^{1,2}(\Psi^{-1}(F)) = \mu_n^{1,2}(F)$ and $\lambda^{1,2}(\Psi^{-1}(F)) = \mu^{1,2}(F)$, we have $\lim \mu_n^{1,2}(F) = \mu^{1,2}(F)$. Thus, we have convergence on all μ-continuity sets which, according to the Portmanteau Theorem, implies that $\mu_n^{1,2}$ weak-* converge to $\mu^{1,2}$. ■

References

[1] Bernheim, B.D., "Rationalizable Strategic Behavior," *Econometrica*, 52, 1984, 1007-1028.

[2] Billingsley, P., *Convergence of Probability Measures*, John Wiley & Sons, 1968.

[3] Brandenburger, A., "Epistemic game theory: complete information," *The New Palgrave Dictionary of Economics*, Eds. S.N. Durlauf and L.E. Blume, Palgrave Macmillan, 2008.

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

[5] Friedenberg, A., "When Do Type Structures Contain All Hierarchies of Beliefs?," available on-line http://www.public.asu.edu/~afrieden/.

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

[7] Heifetz, A., "The Bayesian Formulation of Incomplete Information – The Non-Compact Case," *International Journal of Game Theory*, 21, 1993, 329-338.

[8] Heifetz, A., "Epistemic game theory: incomplete information," *The New Palgrave Dictionary of Economics*, Eds. S.N. Durlauf and L.E. Blume, Palgrave Macmillan, 2008.

[9] Heifetz, A., and D. Samet, "Coherent beliefs are not always types," *Journal of Mathematical Economics*, 32, 1999, 475-488.

[10] Leão Jr., D., M.D. Fragoso, and P.R.C. Ruffino, "Regular conditional probability, disintegration of probability and Radon spaces," *Proyecciones: Revista de Matemática*, 23, no. 1, 2004, 15-29.

[11] Mertens, J.F., and S. Zamir, "Formulation of Bayesian Analysis for Games with Incomplete Information," *International Journal of Game Theory*, 14, 1985, 1-29.

[12] Myerson, R.B., "Comments on "Games with Incomplete Information Played by "Bayesian" Players, I-III"," *Management Science*, 50, 2004, 1818-1824.

[13] Parthasarathy, K.R., *Probability Measures on Metric Spaces*, Academic Press, 1967.

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

[15] Shiryaev, A.N., *Probability*, Springer, 1996.

[16] Siniscalchi, M., "Epistemic game theory: beliefs and types," *The New Palgrave Dictionary of Economics*, Eds. S.N. Durlauf and L.E. Blume, Palgrave Macmillan, 2008.

[17] Valadier, M., "Désintégration d'une mesure sur un produit," *C.R. Acad. Sc. Paris*, t. 276, 1973, Série A, 33-35.

[18] Varadarajan, V.S., "Weak Convergence of Measures on Separable Metric Spaces," *Sankhya*, 19, 1958, 15-22.