

# From Type Spaces to Probability Frames and Back, via Language

Adam Bjorndahl

Carnegie Mellon University  
Department of Philosophy  
Pittsburgh, PA 23185, USA  
adam.bjorndahl@gmail.com

Joseph Y. Halpern

Cornell University  
Ithaca, NY 14853, USA  
halpern@cs.cornell.edu

We investigate the connection between the two major mathematical frameworks for modeling interactive beliefs: *Harsanyi type spaces* and possible-worlds–style *probability frames*. While translating the former into the latter is straightforward, we demonstrate that the reverse translation relies implicitly on a background logical language. Once this “language parameter” is made explicit, it reveals a close relationship between *universal* type spaces and *canonical* models: namely, that they are essentially the same construct. As the nature of a canonical model depends heavily on the background logic used to generate it, this work suggests a new view into a corresponding landscape of universal type spaces.

## 1 Introduction

There are two mathematical frameworks in widespread use for modeling beliefs in multi-agent systems. One approach, popular among computer scientists and logicians, utilizes the *possible worlds* paradigm (see, e.g., [9]). Roughly speaking, a *probability frame* consists of a set of *worlds*, each of which is associated with a set of probability measures (one for each agent), defined on the set of worlds. These probability measures are interpreted as encoding beliefs. Hierarchical beliefs—for example, beliefs about what another agent believes—are naturally captured by the recursive structure of this framework, namely the fact that worlds encode beliefs about worlds. The second approach, more standard in game theory, uses *type spaces*, introduced by Harsanyi [11]. Roughly speaking, types spaces are composed of *states*, encoding “basic” facts about the world (typically including which strategies the players are using), together with *types*, encoding the beliefs of each player in the form of a probability measure defined over the states and the types of her opponents.

What is the relationship between probability frames and type spaces? Aside from a few measure-theoretic technicalities, it is relatively straightforward to transform a type space into a probability frame: essentially, the worlds are state-type pairs. Reversing this transformation is not so straightforward. Given a probability frame, the key question is how to “factor” worlds into states and types. Probability frames encode beliefs about worlds, beliefs about beliefs about worlds, and so on, but this never “bottoms out” in anything like the states in a type space. That is, there is no obvious component of a world that encodes facts such as what strategies the agents are using or the value participants in an auction might assign to an item up for bid. Thus, there seems to be a mismatch between the two approaches.

In this paper, we resolve this mismatch by adding a *language*—a set of basic facts (such as what strategy is used by each agent), represented by primitive propositions—to the picture. In the terminology of modal logic, we pass from frames to models. Given a language, a model is simply a frame together with an *interpretation* that determines for each world  $w$  and primitive proposition  $p$  in the language whether  $p$  is true in world  $w$ . But then we must decide which language to use. We show that the right

choice of language can provide exactly the additional structure needed to “cleanly” factor worlds into states and types. Specifically, we define a transformation on probability models that takes language as a parameter, and show that it produces the familiar type space construction when the language is appropriately expressive.

The value of forging such a connection between the two major mathematical frameworks for modeling belief is obvious: improved communication between researchers working in these respective traditions, and the prospect of importing insights and results from one paradigm to another. And indeed, one immediate application of our language-sensitive translation is a link between two fundamental notions: that of a *canonical* model from the world of modal logic (see, e.g., [2]) and that of a *universal* type space from the theoretical economics literature [17]. Each of these constructions plays a central role in the subfield to which it belongs, and these roles are very similar: each is, in a precise sense, the “largest” structure of its kind—a structure that essentially contains all other such structures. It is perhaps not surprising that they are effectively the *same* structure: roughly speaking, we show in Section 4.3 that canonical models are transformed into universal type spaces.<sup>1</sup> Moreover, since canonical model constructions are highly sensitive to the underlying logical language, this result suggests a new view into a landscape of universal type spaces parametrized by language.

Much of this work was inspired by a beautiful paper of Heifetz and Samet [14]. In it, they construct a measure-theoretic universal type space by a process that closely mimics a standard canonical-model construction (though they do not describe it that way). Our work can be viewed as generalizing their construction to produce a translation from *arbitrary* probability frames to type spaces; our Theorem 9 is then the special case of applying this translation to the canonical model associated with a certain specific logic. In order to emphasize this connection, much of the notation and terminology of this paper duplicates or parallels that used by Heifetz and Samet.

In fact, our “canonical model” construction differs in small but significant ways from the standard construction in modal logic. Typically, worlds in the canonical model are realized as maximal *consistent* sets of formulas from the language, where consistency is, of course, defined relative to some background axiom system. However, the standard finitary axiom system used to reason about probability frames has a problem, namely, it is not *compact*: there exists an infinite set  $F$  of formulas that is not satisfiable such that every finite subset of  $F$  is satisfiable (which means that  $F$  is consistent with the axioms). This renders the corresponding canonical model not a model at all. To avoid this issue, we replace “consistency” with “satisfiability” in our canonical model construction. (Aumann [1] uses an analogous construction.) Meier [16] considers an alternative approach: changing the axiom system. Specifically, he considers an *infinitary* axiom system (with infinitary rules of inference) with respect to which consistency and satisfiability coincide, and constructs a universal type space using a canonical model style construction over this infinitary logic. Although Meier’s logic is infinitary (he allows uncountable conjunctions and disjunctions) and our language is finitary, his canonical model is essentially isomorphic to ours (see Section 4 for further discussion).<sup>2</sup> Conceptually, however, our goals are somewhat different from those of Aumann and Meier. Aumann and Meier focus on the construction of the canonical model. By way of contrast, we approach the issue as a problem first of how to transform an arbitrary probability frame into a type space, and observe afterwards that this translation connects a (suitably defined) notion of canonical model to that of a universal type space.

We are not the first to study the general relationship between type spaces and possible-worlds-style structures. One connection via logic is well known. Sound and complete axiomatizations have been

---

<sup>1</sup>We remark that Meier [16] already observed this connection in the case of an infinitary language.

<sup>2</sup>We thank Martin Meier for pointing this important connection between our work, his work, and that of Aumann.

provided for various logics of probability: Heifetz and Mongin [12] considered a finitary logic where the basic statements have the form  $B_i^\theta \varphi$  (agent  $i$  believes that the probability of  $\varphi$  is at least  $\theta$ )—this is the same logic that we consider—and provided a sound and complete axiomatization in their logic for type systems; Meier [16] did the same for an infinitary logic. Since the axioms are easily seen to be sound in probability frames, and every type structure can be viewed as a probability frame, soundness and completeness of these axiomatizations for probability frames follows. Fagin, Halpern, and Megiddo [5, 4] provided a sound and complete axiomatization of a logic that allowed reasoning about linear combinations of probabilities (i.e., statements such as  $2\ell_i(\varphi) + 3\ell_i(\psi) \geq 1.5$ , which can be read as “twice agent  $i$ ’s probability of  $\varphi$  plus three times agent  $i$ ’s probability of  $\psi$  is at least 1.5”) in probability frames. Since their axioms are easily seen to be sound in type spaces and statements about linear combinations can be expressed in Meier’s infinitary logic, it follows that this axiomatization is also complete for type spaces.

The work on axiomatizations does not produce an explicit translation between type spaces and possible-worlds structures. In more recent work, Galeazzi and Lorini [8] develop a translation between the two and prove a semantic equivalence result. They, too, work at the level of models rather than frames (though they do not explicitly discuss this choice); however, their translations are defined model-theoretically with respect to a single fixed language, rather than taking language as a parameter, making the approach we develop more flexible and more broadly applicable. While the translation they propose from (what we call) probability models into type spaces is not a special case of ours, it is similar in spirit. However, there is one significant difference: in passing through language, our approach effectively identifies worlds that satisfy all the same formulas, while theirs does not (in particular, “duplicate” worlds produce duplicate types under their translation, but not under ours). Semantically speaking, provided we fix an appropriately expressive language, the type spaces we produce are equivalent, once we identify types that satisfy the same formulas. By varying the language, however, our translations take on different characters—they preserve more or less of the type space structure in accordance with what is expressible in the language. Moreover, Galeazzi and Lorini restrict their attention to countable structures, which effectively precludes consideration of structures like universal type spaces or canonical models.

The rest of the paper is organized as follows. Section 2 presents the basic mathematical frameworks within which we work. Section 3 motivates and defines the translations from type spaces to probability frames and vice-versa. Section 4 presents the connection between universal type spaces and canonical models discussed above. Section 5 concludes. Some proofs have been omitted or abridged due to length requirements.

## 2 Preliminaries

The definition of a type space typically includes various topological assumptions that make it easier to prove certain results of interest within that framework [3]. Since our goal is to understand the connection between type spaces and probability frames, we opt instead to work in as minimal a setting as possible, so as not to obscure the translations between the two with additional topological bookkeeping. In particular, following Heifetz and Samet [14], we work with a purely measure-theoretic definition of types spaces.

A **measurable space** is a set  $X$  together with a  $\sigma$ -algebra  $\Sigma_X$  over  $X$ ; elements of  $\Sigma_X$  are called **measurable sets** or **events**. We often drop explicit mention of  $\Sigma_X$  and refer simply to “the measurable space  $X$ ”. We denote by  $\Delta(X)$  the measurable space of all probability measures on  $X$  equipped with the  $\sigma$ -algebra generated by all sets of the form

$$\mathcal{B}^\theta(E) := \{\mu \in \Delta(X) : \mu(E) \geq \theta\},$$

where  $\theta \in [0, 1]$  and  $E \in \Sigma_X$  is an event. Given measurable spaces  $X_1, \dots, X_k$ , the measurable space  $X_1 \times \dots \times X_k$  is just the usual product space equipped with the  $\sigma$ -algebra generated by all sets of the form  $E_1 \times \dots \times E_k$ , where each  $E_i \in \Sigma_{X_i}$ .

Given a probability measure  $\mu$  on  $X$ , the associated **outer measure**, denoted  $\mu^*$ , is defined on arbitrary subsets of  $X$  as follows:

$$\mu^*(A) := \inf\{\mu(E) : E \in \Sigma_X \text{ and } E \supseteq A\}.$$

Obviously, if  $A \in \Sigma_X$ , then  $\mu^*(A) = \mu(A)$ . Otherwise, if  $A$  is not a measurable set, the outer measure of  $A$  can be thought of as a kind of approximation of the measure of  $A$  from above: every event containing  $A$  has probability at least  $\mu^*(A)$ , and for all  $\varepsilon > 0$ , there is an event  $E \supseteq A$  with  $\mu(E) - \mu^*(A) < \varepsilon$ .

Fix a finite set  $I = \{1, \dots, n\}$  of agents. We adopt the usual notational game-theoretic conventions for tuples over  $I$ : Given  $(X_i)_{i \in I}$ , we write

$$X := \prod_{i \in I} X_i \quad \text{and} \quad X_{-i} := \prod_{j \neq i} X_j.$$

We also write  $X'_i \times X_{-i}$  for

$$X_1 \times \dots \times X_{i-1} \times X'_i \times X_{i+1} \times \dots \times X_n$$

and similarly  $(x'_i, x_{-i})$  for

$$(x_1, \dots, x_{i-1}, x'_i, x_{i+1}, \dots, x_n).$$

A **type space (over  $I$ )** is a tuple  $\mathcal{T} = (X, (T_i)_{i \in I}, (\beta_i)_{i \in I})$  where

- $X$  is a measurable space of *states*;
- $T_i$  is a measurable space of  *$i$ -types*;
- $\beta_i : T_i \rightarrow \Delta(X \times T)$  is a measurable function such that the marginal of  $\beta_i(t_i)$  on  $T_i$  is  $\delta_{t_i}$ , the point-mass measure concentrated on  $t_i$ .

Intuitively,  $X$  captures the basic facts about which the agents may be uncertain, while  $i$ -types represent the beliefs of agent  $i$  via the function  $\beta_i$ . These beliefs are not just about the states, but also about the types (and therefore the beliefs) of the agents. In this context, the requirement that  $\beta_i$  be measurable can be thought of as a closure condition on events: for all events  $E \subseteq X \times T$ , the set of points where agent  $i$  assigns  $E$  probability at least  $\theta$ , namely

$$X \times \beta_i^{-1}(\mathcal{B}^\theta(E)) \times T_{-i},$$

is itself an event. The extra condition on  $\beta_i$  is meant to ensure that agent  $i$  is *introspective*: that is, sure of her own beliefs. The point-mass measure  $\delta_{t_i}$  is defined on the measurable subsets of  $T_i$  by

$$\delta_{t_i}(E) = \begin{cases} 1 & \text{if } t_i \in E \\ 0 & \text{if } t_i \notin E. \end{cases}$$

Thus,  $\delta_{t_i}$  assigns probability 1 to all and only the events containing  $t_i$ . Note that in general we cannot simply say that  $\{t_i\}$  has probability 1 according to agent  $i$ , since  $\{t_i\}$  may not be measurable; instead, we can say that every event incompatible with  $t_i$  has probability 0 according to agent  $i$ .<sup>3</sup> Equivalently,  $\delta_{t_i}$  is the unique probability measure on  $T_i$  that assigns  $\{t_i\}$  outer measure 1. A **probability frame (over  $I$ )** is a tuple  $\mathcal{F} = (\Omega, (Pr_i)_{i \in I})$  where

<sup>3</sup>This subtlety does not typically arise in the richer topological setting: provided  $T_i$  is a  $T_1$ -space (see, e.g., [18]; there is an unfortunate clash of notation here),  $\{t_i\}$  is closed and therefore part of the Borel  $\sigma$ -algebra associated with  $T_i$ .

- $\Omega$  is a measurable space of *worlds*;
- $Pr_i : \Omega \rightarrow \Delta(\Omega)$  is a measurable function such that, for each  $\omega \in \Omega$ ,  $Pr_i(\omega)^*(Pr_i^{-1}(Pr_i(\omega))) = 1$ .

Here, all information is encoded in  $\Omega$ , basic facts and beliefs alike. As with type spaces, the measurability of  $Pr_i$  yields a closure condition on events: for all events  $E \subseteq \Omega$ , the set of points where agent  $i$  assigns  $E$  probability at least  $\theta$  is given by  $Pr_i^{-1}(\mathcal{B}^\theta(E))$  and is therefore measurable. And as above, the additional condition on  $Pr_i$  amounts to the stipulation that agent  $i$  is sure of her own beliefs in the sense that at each world  $\omega$ ,  $Pr_i(\omega)$  assigns outer measure 1 to the set

$$Pr_i^{-1}(Pr_i(\omega)) = \{\omega' : Pr_i(\omega') = Pr_i(\omega)\},$$

namely, the set of worlds where her beliefs are given by the measure  $Pr_i(\omega)$ . If this set is measurable, of course, then it is itself assigned probability 1. In much of the literature the measurability of this set is simply assumed. We adopt the slightly more cumbersome definition given above using outer measure because it is more general and because it parallels the introspection condition assumed in type spaces in a way that helps to streamline the translation between the two.

### 3 Translations

Informally, a type space looks like a probability frame where the set of worlds  $\Omega$  has been “factored” into a component representing basic facts—the states—and components representing the beliefs of the agents—the types. As discussed in the introduction, given a probability frame, it is not clear how to perform such a factorization; most of this section is concerned with developing a solution to this problem. The reverse construction, on the other hand, is straightforward, so we begin with it.

**Proposition 1.** *Let  $\mathcal{T} = (X, (T_i)_{i \in I}, (\beta_i)_{i \in I})$  be a type space, and define  $\Omega := X \times T$  and  $Pr_i(x, t) := \beta_i(t_i)$ . Then  $\mathcal{F}_{\mathcal{T}} := (\Omega, (Pr_i)_{i \in I})$  is a probability frame.*

*Proof.* This is the obvious construction; all that needs to be checked is that  $Pr_i$  satisfies the appropriate conditions. Measurability of this function is an easy consequence of the measurability of  $\beta_i$ , since  $Pr_i^{-1}(\mathcal{E}) = X \times \beta_i^{-1}(\mathcal{E}) \times T_{-i}$ . For introspection, observe that

$$Pr_i(x, t)^*(Pr_i^{-1}(Pr_i(x, t))) = \beta_i(t_i)^*(\{(x', t') : \beta_i(t'_i) = \beta_i(t_i)\}) = 1,$$

since every measurable set containing  $\{(x', t') : \beta_i(t'_i) = \beta_i(t_i)\}$  is of the form  $X \times U \times T_{-i}$ , where  $U \subseteq T_i$  is measurable and contains  $t_i$ .  $\square$

In what sense is  $\mathcal{F}_{\mathcal{T}}$  the “right” translation of  $\mathcal{T}$ ? Intuitively, we want to say that the relevant properties of agents and their beliefs that are captured by  $\mathcal{T}$  are also captured by  $\mathcal{F}_{\mathcal{T}}$ , and in some sense preserved by this translation. To make this precise, we formalize the notion of “relevant properties” by identifying them with formulas in a suitably expressive logical language; we then show that the map  $\mathcal{T} \mapsto \mathcal{F}_{\mathcal{T}}$  is truth-preserving with respect to this language (Proposition 2). In addition to providing a formal standard by which to evaluate purported translations between models, making the background language explicit lays the groundwork for the reverse translation, which makes essential use of this structure.

### 3.1 Language

Fix a set  $\Phi$  of *primitive propositions* and a set  $\Theta \subseteq [0, 1]$  of *thresholds*; let  $\mathcal{L}_B^\Theta(\Phi, I)$  be the language recursively generated by the grammar

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \psi \mid B_i^\theta \varphi,$$

where  $p \in \Phi$ ,  $i \in I$ , and  $\theta \in \Theta$ . The parameters  $\Phi$  and  $I$  are omitted when they are clear from context. The other Boolean connectives can be defined in the standard way. We read  $B_i^\theta \varphi$  as “agent  $i$  believes that the probability of  $\varphi$  is at least  $\theta$ ”. Intuitively,  $\Theta$  collects the set of thresholds that the language can express beliefs up to.

There is a standard way of interpreting formulas of  $\mathcal{L}_B^\Theta(\Phi, I)$  in probability frames. A **probability model (over  $(\Phi, I)$ )** is a tuple  $\mathcal{M} = (\mathcal{F}, \pi)$  where  $\mathcal{F}$  is a probability frame (over  $I$ ) and  $\pi : \Phi \rightarrow \Sigma_\Omega$  is an *interpretation*. Recall that  $\Sigma_\Omega$  denotes the  $\sigma$ -algebra associated with the measurable space  $\Omega$ ; the event  $\pi(p) \subseteq \Omega$  is conceptualized as the set of worlds where the primitive proposition  $p$  is true. We can extend this notion of truth to all formulas by defining  $\llbracket \cdot \rrbracket_{\mathcal{M}} : \mathcal{L}_B^\Theta \rightarrow \Sigma_\Omega$  recursively as follows:

$$\begin{aligned} \llbracket p \rrbracket_{\mathcal{M}} &= \pi(p) \\ \llbracket \neg\varphi \rrbracket_{\mathcal{M}} &= \Omega \setminus \llbracket \varphi \rrbracket_{\mathcal{M}} \\ \llbracket \varphi \wedge \psi \rrbracket_{\mathcal{M}} &= \llbracket \varphi \rrbracket_{\mathcal{M}} \cap \llbracket \psi \rrbracket_{\mathcal{M}} \\ \llbracket B_i^\theta \varphi \rrbracket_{\mathcal{M}} &= \{\omega \in \Omega : Pr_i(\omega)(\llbracket \varphi \rrbracket_{\mathcal{M}}) \geq \theta\}. \end{aligned}$$

Of course, the final clause of this definition only makes sense if  $\llbracket \varphi \rrbracket_{\mathcal{M}}$  is measurable, which follows from an easy induction on formulas using the fact that

$$\llbracket B_i^\theta \varphi \rrbracket_{\mathcal{M}} = Pr_i^{-1}(\mathcal{B}^\theta(\llbracket \varphi \rrbracket_{\mathcal{M}})).$$

We say that a formula  $\varphi$  is **true at  $\omega$  (in  $\mathcal{M}$ )** if  $\omega \in \llbracket \varphi \rrbracket_{\mathcal{M}}$ , and that a set  $F$  of formulas is true at  $\omega$  if each  $\varphi \in F$  is true at  $\omega$ . A formula or set of formulas is **valid in  $\mathcal{M}$**  if it is satisfied at all worlds in  $\mathcal{M}$ , and **satisfiable in  $\mathcal{M}$**  if it is true at some world in  $\mathcal{M}$ ; it is **valid** if it is valid in all probability models, and **satisfiable** if it is satisfiable in some probability model.

It is worth noting that the introspection condition on frames, which says that every event containing  $Pr_i^{-1}(Pr_i(\omega))$  has probability 1 according to  $Pr_i(\omega)$ , allows us to deduce the following for all probability models  $\mathcal{M}$  (assuming  $1 \in \Theta$ ):

$$\begin{aligned} \omega \in \llbracket B_i^\theta \varphi \rrbracket_{\mathcal{M}} &\Rightarrow Pr_i^{-1}(Pr_i(\omega)) \subseteq \llbracket B_i^\theta \varphi \rrbracket_{\mathcal{M}} \\ &\Rightarrow Pr_i(\omega)(\llbracket B_i^\theta \varphi \rrbracket_{\mathcal{M}}) = 1 \\ &\Rightarrow \omega \in \llbracket B_i^1 B_i^\theta \varphi \rrbracket_{\mathcal{M}}. \end{aligned}$$

This implies that the formula  $B_i^\theta \varphi \rightarrow B_i^1 B_i^\theta \varphi$  is valid: whenever agent  $i$  believes the probability of  $\varphi$  is at least  $\theta$ , she is sure that she has this belief. A similar argument shows that  $\neg B_i^\theta \varphi \rightarrow B_i^1 \neg B_i^\theta \varphi$  is valid. Of course, this also follows from the stronger assumption that  $Pr_i^{-1}(Pr_i(\omega))$  is itself measurable and has probability 1, but relative to this logical language, such an assumption is overkill.

We can also interpret  $\mathcal{L}_B^\Theta(\Phi, I)$  in type spaces. Although this is not typically done in the literature (though Galeazzi and Lorini [8] do), it allows us to state formally the connection between  $\mathcal{T}$  and  $\mathcal{F}_{\mathcal{T}}$  as defined in Proposition 1, and it highlights the analogies between type spaces and probability frames that we exploit below.

An **interpreted type space** (over  $(\Phi, I)$ ) is a pair  $\mathcal{I} = (\mathcal{T}, \nu)$  where  $\mathcal{T}$  is a type space and  $\nu : \Phi \rightarrow \Sigma_X$  is an *interpretation*; intuitively,  $\nu(p)$  specifies the states of nature where  $p$  is true. As above,  $\nu$  induces a function  $[[\cdot]]_{\mathcal{I}} : \mathcal{L}_B^{\Theta} \rightarrow \Sigma_{X \times T}$  as follows:

$$\begin{aligned} [[p]]_{\mathcal{I}} &= \nu(p) \times T \\ [[\neg\varphi]]_{\mathcal{I}} &= (X \times T) \setminus [[\varphi]]_{\mathcal{I}} \\ [[\varphi \wedge \psi]]_{\mathcal{I}} &= [[\varphi]]_{\mathcal{I}} \cap [[\psi]]_{\mathcal{I}} \\ [[B_i^{\theta} \varphi]]_{\mathcal{I}} &= \{(x, t) \in X \times T : \beta_i(t_i)([[\varphi]]_{\mathcal{I}}) \geq \theta\}. \end{aligned}$$

Now we can formalize the sense in which the map  $\mathcal{T} \mapsto \mathcal{F}_{\mathcal{I}}$  is truth-preserving.

**Proposition 2.** *Let  $\mathcal{I} = (\mathcal{T}, \nu)$  be an interpreted type space, and let  $\mathcal{F}_{\mathcal{I}}$  be the probability frame corresponding to  $\mathcal{I}$  as defined in Proposition 1. Define  $\pi(p) := \nu(p) \times T$ . Then  $\mathcal{M}_{\mathcal{I}} := (\mathcal{F}_{\mathcal{I}}, \pi)$  is a probability model, and for all  $\varphi \in \mathcal{L}_B^{\Theta}$ , we have  $[[\varphi]]_{\mathcal{M}_{\mathcal{I}}} = [[\varphi]]_{\mathcal{I}}$ .*

*Proof.* Proposition 1 tells us that  $\mathcal{F}_{\mathcal{I}}$  is a probability frame, and since  $\nu(p) \in \Sigma_X$ , it is clear that  $\pi(p) \in \Sigma_{X \times T}$ ; it follows that  $\mathcal{M}_{\mathcal{I}}$  is a probability model.

The equality  $[[\varphi]]_{\mathcal{M}_{\mathcal{I}}} = [[\varphi]]_{\mathcal{I}}$  is proved by an easy structural induction on  $\varphi$ . The base cases where  $\varphi \in \Phi$  follows from the definition of  $\pi$ , and the induction steps are all trivial.  $\square$

Proposition 2 is parametrized by the choice of primitive propositions  $\Phi$  and the interpretation  $\nu$ : it says that for any such choice, the correspondence  $\mathcal{T} \mapsto \mathcal{F}_{\mathcal{I}}$  can be extended to a correspondence  $\mathcal{I} \mapsto \mathcal{M}_{\mathcal{I}}$  that is truth preserving with respect to the language  $\mathcal{L}_B^{\Theta}(\Phi)$ . It is worth emphasizing a special case of this result. Given a type space  $\mathcal{T} = (X, (T_i)_{i \in I}, (\beta_i)_{i \in I})$ , recall that the set  $X$  of states is often conceptualized as representing the “basic facts” about the game; for example, the strategy profiles that may be played. As such, when  $X$  is finite (or even just when  $\Sigma_X$  contains all singletons), it is natural to take  $\Phi = X$  and define  $\nu(x) = \{x\}$ ; in this case, intuitively, the primitive propositions simply say what the true state is.

### 3.2 Factoring worlds

We turn now to the reverse translation: the construction of a suitable type space from a given probability frame. As we have observed, the difficulty lies in “factoring” worlds into states and types. Given a probability frame  $\mathcal{F} = (\Omega, (Pr_i)_{i \in I})$ , we might hope to identify types for player  $i$  with probability measures of the form  $Pr_i(\omega)$  for  $\omega \in \Omega$ , but what are the states? This is the crux of the problem: there is nothing in the definition of  $\mathcal{F}$  that allows us to distinguish the “part” of a world  $\omega$  that represents basic facts; indeed, there is no notion of a “basic fact” at all in a probability frame.

A sufficiently rich logical language, however, such as  $\mathcal{L}_B^{\Theta}$ , *does* distinguish “basic” facts from facts about beliefs. For this reason, the construction of a type space naturally operates at the level of probability *models* (which can interpret languages) rather than frames, and depends crucially on the background language.

An  $\mathcal{L}_B^{\Theta}$ -**description** is a set  $D \subseteq \mathcal{L}_B^{\Theta}$  of formulas that is satisfiable and also *maximal* in the sense that, for each  $\varphi \in \mathcal{L}_B^{\Theta}$ , either  $\varphi \in D$  or  $\neg\varphi \in D$ . Given a probability model  $\mathcal{M}$  and a world  $\omega$  in  $\mathcal{M}$ , define the  $\mathcal{L}_B^{\Theta}$ -**description of  $\omega$  in  $\mathcal{M}$**  to be

$$D(\omega) := \{\varphi \in \mathcal{L}_B^{\Theta} : \omega \in [[\varphi]]_{\mathcal{M}}\}.$$

We omit mention of the language and the model when it is safe to do so. It is easy to see that  $D(\omega)$  is an  $\mathcal{L}_B^\Theta$ -description; we call  $D$  the *description map* for  $\mathcal{M}$ . Intuitively,  $D(\omega)$  records all the information about the world  $\omega$  expressible in the language  $\mathcal{L}_B^\Theta$ . Let  $d_0(\omega)$  denote the subset of  $D(\omega)$  consisting of the *purely propositional* formulas: that is, Boolean combinations of the primitive propositions. Let  $d_i(\omega)$  consist of the formulas in  $D(\omega)$  that are Boolean combinations of formulas of the form  $B_i^\theta \varphi$ . Call these the **0-description** and the  **$i$ -description** of  $\omega$ , respectively. We think of the former as recording the basic facts about  $\omega$  (expressible in  $\mathcal{L}_B^\Theta$ ), and the latter as recording the beliefs of agent  $i$  in  $\omega$  (again, expressible in  $\mathcal{L}_B^\Theta$ ).

Fix a probability model  $\mathcal{M} = ((\Omega, (Pr_i)_{i \in I}), \pi)$ . We construct a type space out of  $\mathcal{M}$  by identifying states with 0-descriptions and  $i$ -types with  $i$ -descriptions. Formally, set

$$X := \{d_0(\omega) : \omega \in \Omega\} \text{ and } T_i := \{d_i(\omega) : \omega \in \Omega\}.$$

Intuitively, each state and each type is constituted by a fragment of information about some world  $\omega$  in  $\mathcal{M}$ . We also use this information to define the measure structure: for each  $\varphi \in \mathcal{L}_B^\Theta$ , set

$$E_0(\varphi) := \{x \in X : \varphi \in x\} \text{ and } E_i(\varphi) := \{t_i \in T_i : \varphi \in t_i\};$$

we consider  $X$  and  $T_i$  as measurable spaces equipped with the  $\sigma$ -algebras generated by the collections  $\{E_0(\varphi) : \varphi \in \mathcal{L}_B^\Theta\}$  and  $\{E_i(\varphi) : \varphi \in \mathcal{L}_B^\Theta\}$ , respectively.

The reason we use formulas to pick out events is because, ultimately, we will define each probability measure  $\beta_i(t_i)$  on  $X \times T$  using the information encoded in  $t_i$  about the likelihoods of formulas. For example, if  $B_i^\theta \varphi \in t_i$ , this tells us that  $\beta_i(t_i)$  must assign probability at least  $\theta$  to the subset of  $X \times T$  where  $\varphi$  holds. Of course, in order to make sense of this, we must first define the event in  $X \times T$  that corresponds to  $\varphi$ .

As a first step toward this, we show that given a state-type tuple  $(x, t) \in X \times T$ , the collection of formulas obtained by taking the union of all these partial descriptions, namely  $x \cup \bigcup_i t_i$ , is satisfiable. It is obvious that every 0-description  $x \in X$  and  $i$ -description  $t_i \in T_i$  is individually satisfiable since, by definition, each is satisfied at some world in  $\mathcal{M}$ . On the other hand, there is no guarantee that they are all satisfied at the *same* world in  $\mathcal{M}$  (and in general they may not be), so their joint satisfiability is not so obvious.

**Lemma 3.** *For all  $(x, t) \in X \times T$ , the collection  $x \cup \bigcup_i t_i$  is satisfiable.*

*Proof.* As observed, there are worlds  $\omega_0, \dots, \omega_n$  in  $\mathcal{M}$  such that  $\omega_0$  satisfies  $x$  and  $\omega_i$  satisfies  $t_i$  for  $i = 1, \dots, n$ . We now construct a model  $\mathcal{M}^*$  and world  $\omega^*$  in  $\mathcal{M}^*$  such that  $\mathcal{M}^*$  consists of  $n$  disjoint copies of  $\mathcal{M}$  together with the world  $\omega^*$ ; formally,  $\mathcal{M}^* = ((\Omega^*, (Pr_i^*)_{i \in I}), \pi^*)$ , where

- $\Omega^* = \{(\omega, i) : \omega \in \Omega, i \in \{1, \dots, n\}\} \cup \{\omega^*\};$
- $\pi^*(p) = \begin{cases} \bigcup_{i=1}^n (\pi(p) \times \{i\}) & \text{if } \omega_0 \notin \pi(p) \\ \bigcup_{i=1}^n (\pi(p) \times \{i\}) \cup \{\omega^*\} & \text{if } \omega_0 \in \pi(p) \end{cases}$
- $Pr_i^*(\omega, i)(U \times \{i\}) = Pr_i(\omega)(U)$  for  $\omega \in \Omega$ , and  $Pr_i(\omega^*)(U \times \{i\}) = Pr_i(\omega_i)(U)$  (so the support of  $Pr_i^*(\omega, i)$  and of  $Pr_i(\omega^*)$  is contained  $\Omega \times \{i\}$ ).

It is easy to check that  $\omega^*$  agrees with  $\omega_0$  on propositional formulas and with  $\omega_i$  on  $i$ -descriptions. Thus, the desired result holds.  $\square$

In fact, not only is  $x \cup \bigcup_i t_i$  satisfiable, but it determines a unique  $\mathcal{L}_B^\Theta$ -description.

**Lemma 4.** *There is a unique  $\mathcal{L}_B^\Theta$ -description  $D$  such that  $D \supseteq x \cup \bigcup_i t_i$ .*



*Proof.* By Lemma 3, such a  $D$  exists (take  $D = D(\omega)$  for some  $\omega$  that satisfies  $x \cup \bigcup_i t_i$ ). Uniqueness follows from the following observation, easily proved by structural induction on  $\varphi$ : for all  $\varphi \in \mathcal{L}_B^\Theta$ , either  $x \cup \bigcup_i t_i$  entails  $\varphi$  or  $x \cup \bigcup_i t_i$  entails  $\neg\varphi$ .  $\square$

Let  $D(x, t)$  denote the unique description determined by  $x \cup \bigcup_i t_i$  as in Lemma 4. It is easy to see that  $D(d_0(\omega), d(\omega)) = D(\omega)$ . On the other hand, as mentioned above, the collection of descriptions of the form  $D(x, t)$  may be strictly larger than those of the form  $D(\omega)$ , since some tuples  $(x, t)$  may combine partial descriptions that are not simultaneously satisfied at any world in  $\mathcal{M}$ .

The description  $D(x, t)$  provides a natural way to associate formulas with events in  $X \times T$ . For each  $\varphi \in \mathcal{L}_B^\Theta$ , define

$$[\varphi] := \{(x, t) \in X \times T : \varphi \in D(x, t)\}.$$

**Lemma 5.**  $\Sigma_{X \times T}$  is generated by the collection  $\{[\varphi] : \varphi \in \mathcal{L}_B^\Theta\}$ .

*Proof.* It is easy to see that every  $\varphi \in \mathcal{L}_B^\Theta$  is a Boolean combination of primitive propositions and formulas of the form  $B_i^\theta \psi$ ; it follows that  $\{[\varphi] : \varphi \in \mathcal{L}_B^\Theta\}$  is the algebra generated by all sets of the form  $[p]$  and  $[B_i^\theta \psi]$ . Now observe that  $(x, t) \in [p]$  iff  $p \in x$ , so  $[p] = E_0(p) \times T$ , and similarly,  $(x, t) \in [B_i^\theta \psi]$  iff  $B_i^\theta \psi \in t_i$ , so  $[B_i^\theta \psi] = X \times E_i(B_i^\theta \psi) \times T_{-i}$ . Thus,  $\{[\varphi] : \varphi \in \mathcal{L}_B^\Theta\} \subseteq \Sigma_{X \times T}$ .

To see that  $\Sigma_{X \times T}$  is in fact generated by this collection, it suffices to observe that if each of  $E_0(\varphi_0), E_1(\varphi_1), \dots, E_n(\varphi_n)$  is nonempty, then

$$E_0(\varphi_0) \times E_1(\varphi_1) \times \dots \times E_n(\varphi_n) = [\varphi_0 \wedge \varphi_1 \wedge \dots \wedge \varphi_n].$$

$\square$

We turn now to defining the probability measures  $\beta_i(t_i)$ . Each  $t_i \in T_i$  is a collection of formulas in  $\mathcal{L}_B^\Theta$  that bear on agent  $i$ 's beliefs. We can use these formulas to constrain the space of possible outputs of  $\beta_i(t_i)$ . Moreover, provided  $\mathcal{L}_B^\Theta$  is rich enough, these constraints yield a unique probability measure.

Let  $\mathcal{P}_i$  denote the set of all probability measures  $\mu$  on  $X \times T$  such that, for each  $\varphi \in \mathcal{L}_B^\Theta$  and all  $\theta \in \Theta$ ,

$$\mu([\varphi]) \geq \theta \Leftrightarrow B_i^\theta \varphi \in t_i. \quad (1)$$

**Lemma 6.**  $\mathcal{P}_i \neq \emptyset$ . Moreover, if  $\Theta$  is dense in  $[0, 1]$ , then  $|\mathcal{P}_i| = 1$ .

*Proof.* First we show that  $\mathcal{P}_i$  is nonempty. Let  $\omega$  be a world in  $\mathcal{M}$  such that  $d_i(\omega) = t_i$ . For each  $\varphi \in \mathcal{L}_B^\Theta$ , define

$$\mu_{i, \omega}([\varphi]) = \text{Pr}_i(\omega)([\varphi]_{\mathcal{M}}).$$

One can check that  $\mu_{i, \omega}$  is a pre-measure on the algebra  $\{[\varphi] : \varphi \in \mathcal{L}_B^\Theta\}$  and satisfies (1). By Carathéodory's extension theorem [7, Theorem 1.14], there is a unique extension  $\tilde{\mu}_{i, \omega}$  of  $\mu_{i, \omega}$  to the  $\sigma$ -algebra generated by  $\{[\varphi] : \varphi \in \mathcal{L}_B^\Theta\}$ , which by Lemma 5 is just  $\Sigma_{X \times T}$ . Therefore, by construction,  $\tilde{\mu}_{i, \omega} \in \mathcal{P}_i$ .

If  $\Theta$  is dense in  $[0, 1]$ , then it is easy to see that for all  $\varphi \in \mathcal{L}_B^\Theta$ , if  $\mu \in \mathcal{P}_i$  then

$$\mu([\varphi]) = \sup\{\theta \in \Theta : B_i^\theta \varphi \in t_i\}.$$

It follows that  $\mathcal{P}_i = \{\tilde{\mu}_{i, \omega}\}$ .  $\square$

Let us restrict our attention for the time being to the case where  $\Theta$  is a countable, dense subset of  $[0, 1]$ ; indeed, it is common to assume that  $\Theta = [0, 1] \cap \mathbb{Q}$ . Countability ensures that  $\mathcal{L}_B^\Theta$  contains only countably-many modalities, and by Lemma 6, density allows us to define  $\beta_i(t_i)$  to be the unique element of  $\mathcal{P}_i$ . We then have the following:

**Proposition 7.** *Let  $\mathcal{M}$  be a probability model, and let  $\mathcal{T}_{\mathcal{M}} := (X, (T_i)_{i \in I}, (\beta_i)_{i \in I})$  as defined above. Then  $\mathcal{T}_{\mathcal{M}}$  is a type space. Define  $\nu(p) := E_0(p)$ . Then  $\mathcal{I}_{\mathcal{M}} := (\mathcal{T}_{\mathcal{M}}, \nu)$  is an interpreted type space, and for all  $\varphi \in \mathcal{L}_B^{\ominus}$ , we have*

$$\omega \in \llbracket \varphi \rrbracket_{\mathcal{M}} \Rightarrow (d_0(\omega), d(\omega)) \in \llbracket \varphi \rrbracket_{\mathcal{I}_{\mathcal{M}}}.$$

*Proof.* First we observe that  $\Sigma_{\Delta(X \times T)}$  is generated by all events of the form  $\mathcal{B}^{\theta}([\varphi])$ ; this follows from Lemma 5 together with [14, Lemma 4.5]. Thus, to prove that  $\beta_i$  is measurable it suffices to prove that each set  $\beta_i^{-1}(\mathcal{B}^{\theta}([\varphi]))$  is measurable. By definition, we know that  $\beta_i(t_i)([\varphi]) \geq \theta$  iff  $B_i^{\theta} \varphi \in t_i$ ; it follows that

$$\beta_i^{-1}(\mathcal{B}^{\theta}([\varphi])) = E_i(B_i^{\theta} \varphi),$$

which is measurable by definition. That  $\beta_i(t_i)$  concentrates on  $t_i$  follows from the fact that

$$(x, t) \in [B_i^{\theta} \psi] \Leftrightarrow B_i^{\theta} \psi \in t_i \Leftrightarrow B_i^1 B_i^{\theta} \psi \in t_i \Leftrightarrow \beta_i(t_i)([B_i^{\theta} \psi]) = 1.$$

Finally, the semantic equivalence follows by structural induction on  $\varphi$ . □

## 4 Universal Type Spaces and Canonical Models

### 4.1 Universal type spaces

The existence of a *universal type space* [17] underpins the use of type spaces as a general framework for modeling beliefs: roughly speaking, it guarantees that they do not rule out any possible collection of beliefs. Individual type spaces, of course, can be quite small and omit many configurations of beliefs. The universal type space, by contrast, essentially includes all possible configurations of belief; in particular, this means we need not be concerned with gaps in our representation of games.

Formally, given type spaces  $\mathcal{T} = (X, (T_i)_{i \in I}, (\beta_i)_{i \in I})$  and  $\mathcal{T}' = (X, (T'_i)_{i \in I}, (\beta'_i)_{i \in I})$  (with a common set  $X$  of states), a profile of functions  $f_i : T_i \rightarrow T'_i$  constitutes a **type morphism** from  $\mathcal{T}$  to  $\mathcal{T}'$  provided that, for each  $i \in I$ ,  $t_i \in T_i$ , and each event  $E \subseteq X \times T'$ ,

$$\beta'_i(f_i(t_i))(E) = \beta_i(t_i)(f^{-1}(E)),$$

where  $f : X \times T \rightarrow X \times T'$  is defined by  $f = (id_X, f_1, \dots, f_n)$ . Roughly speaking, this says that each  $f_i$  assigns to each  $t_i \in T_i$  a type  $f_i(t_i) \in T'_i$  that agrees with  $t_i$  on the probabilities of all events, where events in  $\mathcal{T}$  and  $\mathcal{T}'$  are identified via the correspondence given by  $f$ . A type space  $\mathcal{T}^*$  is called **universal (for  $X$ )** if, for every type space  $\mathcal{T} = (X, (T_i)_{i \in I}, (\beta_i)_{i \in I})$ , there exists a unique type morphism from  $\mathcal{T}$  to  $\mathcal{T}^*$ . Thus, each such  $\mathcal{T}$  can be thought of as existing “inside”  $\mathcal{T}^*$  (via the mapping  $f$ ).

Type morphisms are defined so as to preserve the structure of belief. Indeed, given any interpretation  $\nu : \Phi \rightarrow \Sigma_X$ , it is easy to see that if  $(f_1, \dots, f_n)$  is a type morphism from  $\mathcal{T}$  to  $\mathcal{T}'$ , then for any  $(x, t) \in X \times T$  and any  $\varphi \in \mathcal{L}_B^{[0,1]}(\Phi)$ , we have

$$(x, t) \in \llbracket \varphi \rrbracket_{(\mathcal{T}, \nu)} \Leftrightarrow f(x, t) \in \llbracket \varphi \rrbracket_{(\mathcal{T}', \nu)}.$$

As a consequence, the universal type space for  $X$  satisfies all the  $\mathcal{L}_B^{[0,1]}(\Phi)$ -descriptions that are satisfied in *some* type space over  $X$ . It is natural to wonder whether this property characterizes the universal type space; the connection with canonical models we now present essentially amounts to a formalization of this idea.

## 4.2 Canonical models

The classical canonical model construction is used to prove completeness of various modal systems. Given some axiom system AX of interest, a model is constructed wherein each world corresponds to a maximal AX-consistent set of formulas, with additional structure derived from the properties of these sets of formulas.

The construction we present here differs in that we are not concerned with axiomatics—indeed, for logics that fail to be compact (such as, notably, the logic of  $\mathcal{L}_B^{[0,1]}$  as interpreted in probability frames), consistent sets of formulas need not be satisfiable, so the canonical model construction fails. Nonetheless, we can adapt this construction by replacing “consistent” with “satisfiable”; in other words, we can build a model in which the worlds are exactly the  $\mathcal{L}_B^{[0,1]}$ -descriptions.<sup>4</sup> Intuitively, such a model contains a world satisfying every such description; ultimately, we will show that we can obtain a universal type space by constructing such a model and then translating it into a type space as in Section 3.2.

Consider a fixed language  $\mathcal{L}_B^{[0,1]}(\Phi)$  and a class of probability models  $\mathcal{C}$ ; let  $\bar{\Omega}$  denote the set of all  $\mathcal{L}_B^{[0,1]}(\Phi)$ -descriptions satisfiable in some model in  $\mathcal{C}$ . Define  $\hat{\varphi} = \{\bar{\omega} \in \bar{\Omega} : \varphi \in \bar{\omega}\}$ , and let  $\Sigma_{\bar{\Omega}}$  be the  $\sigma$ -algebra generated by the collection  $\mathcal{A} = \{\hat{\varphi} : \varphi \in \mathcal{L}_B^{[0,1]}(\Phi)\}$ . Define  $\mu_{i,\bar{\omega}} : \mathcal{A} \rightarrow [0, 1]$  by

$$\mu_{i,\bar{\omega}}(\hat{\varphi}) = \sup\{\theta \in [0, 1] : B_i^\theta \varphi \in \bar{\omega}\}.$$

It is not hard to check that  $\mu_{i,\bar{\omega}}$  is a pre-measure on the algebra  $\mathcal{A}$ , so, by Carathéodory’s extension theorem, it can be extended to a unique probability measure on  $\Sigma_{\bar{\Omega}}$ ; let  $\bar{P}r_i(\bar{\omega})$  denote this extension. Finally, for each  $p \in \Phi$ , set  $\bar{\pi}(p) = \hat{p}$ .

**Proposition 8.**  *$\bar{\mathcal{M}} = (\bar{\Omega}, (\bar{P}r_i)_{i \in I}, \bar{\pi})$  is a probability model, and for all  $\varphi \in \mathcal{L}_B^{[0,1]}$ , we have  $\llbracket \varphi \rrbracket_{\bar{\mathcal{M}}} = \hat{\varphi}$ . Moreover,  $\bar{\mathcal{M}}$  is universal for  $\mathcal{C}$  in the sense that, for all  $\mathcal{M} \in \mathcal{C}$ , there is a truth-preserving map (namely,  $D$ , the description map for  $\mathcal{M}$ ) from  $\mathcal{M}$  to  $\bar{\mathcal{M}}$ .*

Call  $\bar{\mathcal{M}}$  the **universal probability model for  $\mathcal{C}$  over  $\mathcal{L}_B^{[0,1]}(\Phi)$** . As we mentioned earlier, Meier [16] works with an infinitary version of the language  $\mathcal{L}_B^{[0,1]}(\Phi)$  and constructs a canonical model for that language. Call his language  $\mathcal{L}_B^{\infty,[0,1]}$ . Although  $\mathcal{L}_B^{\infty,[0,1]}$  is infinitary, as observed in [10, Lemma 4.1], every  $\mathcal{L}_B^{[0,1]}$ -description can be uniquely extended to an  $\mathcal{L}_B^{\infty,[0,1]}$ -description. It follows that the canonical model for the language  $\mathcal{L}_B^{[0,1]}$  is isomorphic to the canonical model for  $\mathcal{L}_B^{\infty,[0,1]}$ . Meier shows that the canonical model for  $\mathcal{L}_B^{\infty,[0,1]}$  is universal. Of course, it follows that the canonical model for  $\mathcal{L}_B^{[0,1]}$  is also universal. We give an independent proof of this result here, since it allows us to connect universal type spaces to the language considerations discussed earlier.

## 4.3 Translation

Let  $X$  be a measurable space of states where  $\Sigma_X$  is generated by the singletons  $\{x\}$ .<sup>5</sup> We construct a universal type space for  $X$  by first constructing a universal model as in Section 4.2. Consider the

<sup>4</sup>As we said above, a similar construction appears in [1], though the connection to type spaces is not explored in any depth.

<sup>5</sup>It is possible to weaken this condition to the following: for every  $x, y \in X$ , there exists a “separating event”  $E \in \Sigma_X$  such that  $x \in E$  and  $y \notin E$ . The issue here is that if  $X$  contains points that are not separated in this way, they will not differ on any description and so the universal model construction we employ below will end up identifying them. Notice, however, that this is only a problem because the universal type space for state space  $X$  is required to use  $X$  as the state space, even when  $X$  contains “redundant” states that are not separated by any event. Intuitively, however, this is unnecessary—a slightly relaxed notion of a universal type space would simply require that its state space be rich enough to reflect the measure structure of  $X$ , rather than its set-theoretic structure. And indeed, this is essentially what you get by running the construction below without the separability requirement articulated above.

language  $\mathcal{L}_B^{[0,1]}(X)$  (i.e., where  $\Phi = X$ ) and the class  $\mathcal{C}_X$  of probability models such that  $\{\pi(x) : x \in X\}$  partitions  $\Omega$ . Intuitively, this condition hard-codes the constraint that exactly one state  $x \in X$  is the “true” state of the world.

**Theorem 9.** *Let  $\bar{\mathcal{M}}$  be the universal probability model for  $\mathcal{C}_X$  over  $\mathcal{L}_B^{[0,1]}(X)$ . Then the type space  $\mathcal{T}_{\bar{\mathcal{M}}}$  is universal for  $X$ .*

*Proof.* The state space for  $\mathcal{T}_{\bar{\mathcal{M}}}$  is, by definition, the collection  $\{d_0(\bar{\omega}) : \bar{\omega} \in \bar{\Omega}\}$ ; it is easy to see that each set  $d_0(\bar{\omega})$  contains exactly one element of  $X$ , and this correspondence is a measurable bijection with measurable inverse. So  $\mathcal{T}_{\bar{\mathcal{M}}}$  has the “right” state space.

Next, let  $\mathcal{T} = (X, (T_i)_{i \in I}, (\beta_i)_{i \in I})$  be any type space based on  $X$ . We must produce a (unique) type morphism from  $\mathcal{T}$  to  $\mathcal{T}_{\bar{\mathcal{M}}}$ . To do so, define  $v : \Phi \rightarrow \Sigma_X$  by  $v(x) = \{x\}$ , let  $\mathcal{S} = (\mathcal{T}, v)$  be the corresponding interpreted type space, and consider the model  $\mathcal{M}_{\mathcal{S}}$  obtained from  $\mathcal{S}$  as in Proposition 2. It is easy to see that  $\mathcal{M}_{\mathcal{S}} \in \mathcal{C}_X$ , and because of this, for each  $(x, t) \in X \times T$  and  $i \in I$ , there is a unique  $d_i(\bar{\omega})$  that is satisfied at  $(x, t)$ . In this case, define  $f_i(t_i) = d_i(\bar{\omega})$ .  $\square$

Theorem 9 realizes the intuition that the universal type space for  $X$  is precisely the type space that satisfies all and only the  $\mathcal{L}_B^{[0,1]}(\Phi)$ -descriptions that are satisfied in some type space over  $X$ . Thinking of universal type spaces in this way makes the dependence on language plain, and suggests alternative notions of “universal type spaces” obtained by varying the language over which the universal quantification takes place. That is, given a class of type spaces  $\mathcal{T}$  and a language  $\mathcal{L}$  interpretable in those type spaces in  $\mathcal{T}$ , we can define a type space  $\mathcal{T}^*$  to be **universal for  $\mathcal{T}$  with respect to  $\mathcal{L}$**  provided every  $\mathcal{L}$ -description satisfiable in  $\mathcal{T}$  is (uniquely) satisfied in  $\mathcal{T}^*$ . Naturally, we might hope to construct  $\mathcal{T}^*$  by transforming an appropriate canonical/universal model. The translation defined in Section 3 does the job for languages of the form  $\mathcal{L}_B^\Theta$  when  $\Theta$  is dense in  $[0, 1]$ . Generalizing this result to other languages, both richer and poorer, is the subject of ongoing research.

One natural way to coarsen the language is by dropping the assumption that  $\Theta$  is dense in  $[0, 1]$ . An extreme case of this would be to take  $\Theta = \{1\}$ , corresponding to a standard modal language of qualitative, “probability 1” belief (see, e.g., [9]). In this case, the sets of measures  $\mathcal{P}_{t_i}$  defined in Section 3 encode only information regarding those events that  $t_i$  assigns probability 1 to. Another natural modification to the language is to enrich it with a knowledge modality. Logics of knowledge and belief have been well-studied, and canonical models certainly exist in such settings (see [15] and the references in [9, Chapter 8]). By contrast, *knowledge spaces*, an epistemic analogue to type spaces, have been shown *not* to permit a universal object [6, 13]. What is the source of this mismatch? Does the translation technique we present fundamentally fail to generalize to models of knowledge? Or can the canonical model construction in the modal case inform a new, type-theoretic representation of knowledge that does enjoy a universal model? We leave these questions to future work.

## 5 Conclusion

We have related probability frames and type spaces in a way that makes clear the critical role of language. Our approach allows us to show the deep connections between the canonical models that are standard in the modal logic community and the universal type spaces that play a critical role in epistemic game theory. We believe that further work, considering different choices of language, will further illuminate the connections between these two modeling paradigms.

## References

- [1] R. J. Aumann (1999): *Interactive epistemology II: probability*. *International Journal of Game Theory* 28(3), pp. 301–314, doi:10.1007/s001820050112.
- [2] P. Blackburn, M. de Rijke & Y. Venema (2001): *Modal Logic*. Cambridge Tracts in Theoretical Computer Science, No. 53, Cambridge University Press, Cambridge, U.K., doi:10.1017/CBO9781107050884.
- [3] E. Dekel & M. Siniscalchi (2015): *Epistemic Game Theory*. In H. P. Young & S. Zamir, editors: *Handbook of Game Theory with Economic Applications, Volume 4*, North-Holland, Amsterdam, pp. 619–702.
- [4] R. Fagin & J. Y. Halpern (1994): *Reasoning about knowledge and probability*. *Journal of the ACM* 41(2), pp. 340–367, doi:10.1145/174652.174658.
- [5] R. Fagin, J. Y. Halpern & N. Megiddo (1990): *A logic for reasoning about probabilities*. *Information and Computation* 87(1/2), pp. 78–128, doi:10.1016/0890-5401(90)90060-U.
- [6] R. Fagin, J. Y. Halpern & M. Y. Vardi (1991): *A model-theoretic analysis of knowledge*. *Journal of the ACM* 91(2), pp. 382–428, doi:10.1145/103516.128680. A preliminary version appeared in *Proc. 25th IEEE Symposium on Foundations of Computer Science*, 1984.
- [7] Gerald B. Folland (1999): *Real Analysis: Modern Techniques and Their Applications*. John Wiley & Sons, New York.
- [8] P. Galeazzi & E. Lorini (2016): *Epistemic logic meets epistemic game theory: a comparison between multi-agent Kripke models and type spaces*. *Synthese* 193, pp. 2097–2127, doi:10.1007/s11229-015-0834-x.
- [9] J. Y. Halpern (2003): *Reasoning About Uncertainty*. MIT Press, Cambridge, MA. Second, edition 2017.
- [10] J. Y. Halpern & R. Pass (2009): *A logical characterization of iterated admissibility*. In: *Theoretical Aspects of Rationality and Knowledge: Proc. Twelfth Conference (TARK 2009)*, pp. 146–155, doi:10.1145/1562814.1562836.
- [11] J. Harsanyi (1968): *Games with incomplete information played by ‘Bayesian’ players, Parts I–III*. *Management Science* 14, pp. 159–182, 320–334, 486–502, doi:10.1287/mnsc.14.5.320.
- [12] A. Heifetz & P. Mongin (2001): *Probability logic for type spaces*. *Games and Economic Behavior* 35(1–2), pp. 31–53, doi:10.1006/game.1999.0788.
- [13] A. Heifetz & D. Samet (1998): *Knowledge spaces with arbitrarily high rank*. *Games and Economic Behavior* 22(2), pp. 260–273, doi:10.1006/game.1997.0591.
- [14] A. Heifetz & D. Samet (1998): *Topology-free typology of beliefs*. *Journal of Economic Theory* 82, pp. 324–341, doi:10.1006/jeth.1998.2435.
- [15] W. Lenzen (1978): *Recent work in epistemic logic*. *Acta Philosophica Fennica* 30, pp. 1–219.
- [16] M. Meier (2012): *An infinitary probability logic for type spaces*. *Israel Journal of Mathematics* 192, pp. 1–58, doi:10.1007/s11856-012-0046-8.
- [17] J. F. Mertens & S. Zamir (1985): *Formulation of Bayesian analysis for games of incomplete information*. *International Journal of Game Theory* 14(1), pp. 1–29, doi:10.1007/BF01770224.
- [18] J. Munkres (2000): *Topology*, 2nd edition. Prentice-Hall, Englewood Cliffs, N.J.