Harsanyi Type Spaces and Final Coalgebras Constructed from Satisfied Theories

Lawrence S. Moss[1]   Ignacio D. Viglizzo[2]

Department of Mathematics
Indiana University
Bloomington, IN 47405 USA

Abstract

This paper connects coalgebra with a long discussion in the foundations of game theory on the modeling of type spaces. We argue that type spaces are coalgebras, that universal type spaces are final coalgebras, and that the modal logics already proposed in the economic theory literature are closely related to those in recent work in coalgebraic modal logic. In the other direction, the categories of interest in this work are usually measurable spaces or compact (Hausdorff) topological spaces. A coalgebraic version of the construction of the universal type space due to Heifetz and Samet [5] is generalized for some functors in those categories. Since the concrete categories of interest have not been explored so deeply in the coalgebra literature, we have some new results. We show that every functor on the category of measurable spaces built from constant functors, products, coproducts, and the probability measure space functor has a final coalgebra. Moreover, we construct this final coalgebra from the relevant version of coalgebraic modal logic. Specifically, we consider the set of theories of points in all coalgebras and endow this set with a measurable and coalgebra structure.

Keywords: Coalgebra, final coalgebra, Harsanyi type space, measurable space, probabilities, beliefs.

1 Introduction: Type Spaces

This paper is a first exploration of the application of ideas and results from coalgebra to the foundational area of game theory concerned with *type spaces*. Type spaces are mathematical structures used in modeling settings where agents are described by their *types*, and these types give us "beliefs about the

---

[1] Email: lsm@cs.indiana.edu
[2] Email: igvigliz@indiana.edu

world", "beliefs about each other's beliefs about the world", "beliefs about each other's beliefs about each other's beliefs about the world", etc. That is, the formal concept of a type space is intended to capture in one structure an unfolding infinite hierarchy related to *interactive belief*.

The 1994 Nobel Prize in Economic Sciences was awarded to John C. Harsanyi, John F. Nash Jr., and Reinhard Selten "for their pioneering analysis of equilibria in the theory of non-cooperative games." In addition to his work on equilibria, Harsanyi will also be remembered for his introduction of type spaces in a three-part paper published in 1967 and 1968 [4]. He showed how to convert a game with incomplete information into one with complete yet imperfect information. This matter is not relevant to our paper, but three related points are noteworthy. First, Harsanyi's notion of types goes further than what we described above: an agent's type gives us their *beliefs about the types of the other agents*. Second, despite this circularity, the informal concept of a *type* (as a "pool" from which all players can be picked) is widespread in areas of non-cooperative game theory and economic theory. And finally, the *formalization* of type spaces was, and to some extent still is, an open area. That is, Harsanyi did not really formalize type spaces in his original paper; this was left to later researchers. Getting back to our very rough informal description above, what exactly are "beliefs"? And how can a structure contain types which give rise to beliefs about other types? What is the relation of this to the infinite hierarchy of beliefs about beliefs about $\cdots$ beliefs about the world? Can we characterize the space of all possible types? So there is a collection of papers on this matter, starting with Böge and Eisele's paper [2] from 1979. Again, we are not much concerned with these conceptual matters in this paper. Most of the important papers for our study are technical contributions dealing with the matter of *universal type spaces*. A universal type space is intended to capture all possible types, so it is an answer to our third question above.

There are some clear conceptual clues that coalgebra could be connected to type spaces. The first is that the notion of "belief" in the game theory literature is typically a probabilistic one. If we replace "belief" with "knowledge" above, then we have a very-well-studied notion, the formalization of knowledge by *possible worlds semantics*. The mathematical structures for possible worlds semantics are sets W of *worlds* with two functions, one giving for each world $w \in W$ some set of "atomic propositions" true at w, and the other giving for each w some set of worlds which are said to be "possible from w". These structures are essentially the coalgebras in the category Set of sets of the functor $F(W) = A \times \mathcal{P}(W)$, where A is the power set of the set of atomic propositions and $\mathcal{P}$ is the power set functor on sets. Perhaps the primary

contribution of coalgebra to this area to date is to show that *modal logic*, the natural logical language for the structures, generalizes to *coalgebraic versions of modal logic.* We'll return to this point shortly.

The second clue has to do with the role of the universal type space in this field. Types are taken to be elements of the universal type space. In the universal type space all possible types are uniquely represented, and the idea is that two types with exactly the same beliefs about the underlying world of "nature" plus the types of the other players are taken to be undistinguishable. This is the same ideology as we find concerning *final coalgebras*, in which (as is well known) points with the same behavior are identified.

Returning to type spaces, we recall that the usual modeling of belief in game theory is via probability. So we would expect that *type spaces should be probabilistic versions of Kripke models.* One should replace the functor $\mathcal{P}$ with something like Δ, where

$$\Delta(W) \quad = \quad \{\mu \,|\, \mu \text{ is a probability measure on } W\}. \tag{1}$$

Indeed, this is the case: most proposals in the literature do end up studying certain mappings from a space X to some variation of the functor Δ applied to X. This is our third clue of the connection. But note that (1) leaves a lot lacking: if W is just a set, how do we know that it has any probability measures? Does it matter which σ-algebra we use? And if W is an object in some other category, say measurable spaces or compact metric spaces, then what structure do we put on $\Delta(W)$?

We intend this paper to be a contribution to this area by connecting it with coalgebra. Here are the main conceptual claims of the paper, as well as the main results:

⋆ The original notion of a Harsanyi type space may be taken to be a coalgebra of a functor F on Meas^I, where Meas is the category of measurable spaces and measurable maps, and I is a discrete category (of agents). But this reformulation is not obvious, because the original notion had an extra condition that the agents "know their own type." We discuss this in Section 2.

⋆ Universal type spaces are final coalgebras.

⋆ The constructions of universal type spaces in the literature are related to constructions found in coalgebra. However, there are differences due primarily to the fact that the work is going on in categories like Meas rather that Set. Putting things differently, the universal type space constructions can be generalized, and this is of interest in coalgebra.

⋆ There are versions of coalgebraic modal logic for relevant functors, and

one can prove the existence of final coalgebras by considering the *satisfied theories* in these logics. This construction parallels the construction of a final coalgebra on Set for $F(W) = A \times \mathcal{P}(W)$ (or rather $A \times \mathcal{P}_{fin}(W)$, using the finite power set functor) using the set of descriptions of all modal theories of all possible worlds. We work a formulation of coalgebraic modal logic based on the one due to Jacobs [7], following work of Kurz [9] and Rößiger [11,12,13]. But we do differ from these papers in that our final coalgebras are not connected to logical theories, but to the theories realized in coalgebras. In the economic literature, this construction is due to Heifetz and Samet [5].

⋆ We formulate a notion of "measure polynomial functor" on Meas. We prove a new result in Section 3: every such polynomial has a final coalgebra with carrier consisting of sets of sentences in a particular language. The method also works for polynomial functors on Set, as we show in Section 5.2.

1.1   Background notions

A measurable space is a pair $M = (M, \Sigma)$, where M is a set and Σ is a σ-algebra of subsets of M. The sets in Σ are called *measurable sets* or *events*. Usually Σ contains all singletons $\{x\}$; we shall almost always assume a weaker condition that for each $x \in M$, $\{x\}$ is the intersection of the measurable subsets of M containing x. A collection $\mathcal{B}$ of subsets of M *generates* a σ-algebra Σ if Σ is the smallest σ-algebra including $\mathcal{B}$. A *measure on* M is a σ-additive function $\mu : \Sigma \to [0, \infty]$. The measure μ is a *probability measure* if $\mu(M) = 1$.

A morphism of measurable spaces $f : (M, \Sigma) \to (N, \Sigma')$ is a function $f : M \to N$ such that for each $A \in \Sigma'$, $f^{-1}(A) \in \Sigma$. This gives a category Meas. Meas has products and coproducts; indeed it has much more structure. There is an endofunctor $\Delta :$ Meas $\to$ Meas defined by: $\Delta(M)$ is the set of probability measures on M with the σ-algebra generated by $\{\beta^p(E) : p \in [0, 1], E \in \Sigma\}$, where

$$\beta^p(E) \quad = \quad \{\mu \in \Delta(M) \,|\, \mu(E) \geq p\}.$$

Here is how Δ acts on morphisms. If $f : M \to N$ is measurable, then for $\mu \in \Delta(M)$ and $A \in \Sigma'$, $(\Delta f)(\mu)(A) = \mu(f^{-1}(A))$. That is, $(\Delta f)(\mu) = \mu \circ f^{-1}$.

We also note some additional structure. First, there is a natural transformation $\delta : Id \to \Delta$ defined by $\delta_M(m)(E) = 1$ if $m \in E$ and 0 if $m \notin E$. We also write δ_m instead of $\delta_M(m)$; this is the Dirac measure supported at m. Second, there is a natural transformation $\gamma : \Delta\Delta \to \Delta$ given by $\gamma_M(\mu)(E) = \int_{\nu \in \Delta(M)} \nu(E) \, d\mu$.

Lemma 1.1 (Giry [3]) (Δ, δ, γ) *is a monad.*

Lemma 1.2 *For each $p \in [0,1]$, β^p may be regarded as a 'predicate lifting'. That is β^p takes measurable subsets of each space M to measurable subsets of $\Delta(M)$, and it is natural in the sense that if $f : M \to N$, then for all measurable $E \subseteq Y$, $\beta^p(f^{-1}(E)) = (\Delta f)^{-1}(\beta^p(E))$.*

Lemma 1.3 *Suppose that the collection $\mathcal{B}$ of sets generates Σ. Then the σ-algebra on $\Delta(M)$ is the generated by the sets of the form $\{\beta^p(E) \mid E \in \mathcal{B}$ and $p \in [0,1]\}$.*

We shall also need one result from measure theory on the extension of probability measures from collections of sets to the σ-algebras they generate. Lemma 1.4 below follows from Dynkin's $\pi - \lambda$ Theorem. For more details, see, for example Billingsley [1], page 36.

Lemma 1.4 *Suppose that μ_1 and μ_2 are probability measures on $\sigma(\mathcal{A})$, where $\mathcal{A}$ is a family of sets which is closed under intersections and $\sigma(\mathcal{A})$ is the σ-algebra generated by $\mathcal{A}$. If μ_1 and μ_2 agree on $\mathcal{A}$, then they agree on $\sigma(\mathcal{A})$.*

We also need at some point to look more closely at products in Meas. Given two measurable spaces, A and B, their product is the cartesian product of the sets A and B, endowed with the σ-algebra generated by the "rectangles" obtained as cartesian product of a measurable subset of A times a measurable subset of B. With this σ-algebra, both projections are measurable. For a subset $E \subseteq A \times B$, the sections of E are the sets: $E_a = \{b : \langle a, b \rangle \in E\}, E^b = \{a : \langle a, b \rangle \in E\}$. Each section of a measurable subset of the product is measurable.

If μ is a probability measure on A and ν a probability measure on B, we can define the probability measure $\mu \times \nu$ on $A \times B$ by $(\mu \times \nu)(E) = \int \mu(E^b) d\nu = \int \nu(E_a) d\mu$.

Going in the other direction, a probability measure μ on $A \times B$ induces via the projections, a measure on each of the factor spaces. These measures are called *marginals*, and denoted by $mar_A\mu = (\Delta \pi_A)\mu = \mu \circ \pi_A^{-1}$; $mar_B\mu = (\Delta \pi_B)\mu = \mu \circ \pi_B^{-1}$.

Lemma 1.5 *Let μ be a probability measure on a product measurable space $A \times B$. If $mar_B\mu = \delta_{b_0}$ for some $b_0 \in B$ then $\mu = mar_A\mu \times \delta_{b_0}$.*

Proof. We only need to prove it for rectangles $G \times F$, where G is a measurable subset of A and F is a measurable subset of B.

We want to prove that $\mu(G \times F) = (mar_A\mu)(G) \times \delta_{b_0}(F)$. We have two cases: if $b_0 \notin F$, this reduces to proving that $\mu(G \times F) = 0$, and if $b_0 \in F$, then we want to show that $\mu(G \times F) = mar_A\mu(G) = \mu(\pi_A^{-1}(G)) = \mu(G \times B)$.

Notice first that for $\mu(G \times B) = \mu(\pi_A^{-1}(G)) = mar_A\mu(G) = mar_A\mu(G) \times \delta_{b_0}(B)$. Also $\mu(A \times F) = mar_B\mu(F) = \delta_{b_0}(F) = (mar_A\mu)(A) \times \delta_{b_0}(F)$.

Now we can prove that if $b_0 \notin F$, then $\mu(G \times F) \leq \mu(A \times F) = 0$, and if $b_0 \in F$, then $\mu(G \times F) = \mu(G \times \{b_0\}) + \mu(G \times (F \setminus \{b_0\}) \leq \mu(G \times \{b_0\}) + \mu(A \times (F \setminus \{b_0\})) = \mu(G \times \{b_0\})$. On the other hand, $\mu(G \times B)$ is also equal to $\mu(G \times \{b_0\}) + 0$.   □

2   Formulation of Type Spaces as Coalgebras

The first constructions of Harsanyi type spaces were based on a hierarchy of beliefs. In this hierarchical approach, as seen for example in [10], types are constructed by detailing the players' beliefs about nature, about the other players' beliefs about nature, etc. Since each type gives a probability measure on the set of all types, there is a function from types to probability measures on states of nature and types. As we noted before, this is almost a coalgebra, except that there are some extra conditions imposed on these functions.

At this point, we need to formulate the category Meas^I. Fix a finite set I of *players*. We assume that $0 \notin I$ and we then define $I_0 = I \cup \{0\}$. This "0" stands for "nature", and so I_0 includes nature as one of the players (but one who won't have beliefs about the other players). We consider I and I_0 as discrete categories.

The objects of Meas^I are families $X = (X_i)_{i \in I}$ of measurable spaces, and the morphisms are also families of measurable maps. Fix a measurable space M to represent the "states of nature", and write X_0 for M. Each player should have beliefs about nature and about the beliefs of the other player. This leads to the following definition.

Let $\mathsf{C} : \mathsf{Meas}^I \to \mathsf{Meas}$ be the functor given by

$$\mathsf{C}X \quad = \quad \prod_{i \in I_0} X_i$$

At first glance, it might look like what we want is to consider for each $x_i \in X_i$, a probability measure on $\mathsf{C}X$. However, this is *not* what we want because it misses an important intuition concerning type spaces. This is that *players should know their own types*. In other words, each player i should only have beliefs about the (joint distribution on) other players' beliefs; i's own beliefs should not even enter in. Thus we define functors $U_i : \mathsf{Meas}^I \to \mathsf{Meas}$ given by

$$U_i(X) \quad = \quad \Pi\{X_j \,|\, j \in I_0, j \neq i\}.$$

U_i acts the obvious way on morphisms. Note that U_i depends on the space X_0 of states of nature, even though our notation does not mention this.

Now $\mathsf{C}X = U_i X \times X_i$ up to isomorphisms. The fact we mentioned above, about player i knowing her own type is modeled in [5] by adding the condition that the corresponding measure on $\mathsf{C}X$ has marginal δ_{x_i} on X_i. Here is where Lemma 1.5 plays a role allowing us to recast type spaces as coalgebras: it is essentially the same to consider measures on $U_i X$ and measures on $\mathsf{C}X$ for which the marginal on X_i is a Dirac measure on a known point.

Finally our main innovation and the one that lets us model type spaces as coalgebras is to work in Meas^I, rather than working with the product of the different spaces considered.

Let $F : \mathsf{Meas}^I \to \mathsf{Meas}^I$ be defined by

$$\mathsf{F}(X) \quad = \quad (\Delta U_i(X))_{i \in I}$$

As before Δ is the probability measure space functor, and once again, our notation elides the underlying space X_0 of states of nature.

This way, instead of having a family of functions in Meas, each one of them with a condition on one of its marginals, any morphism in Meas^I works as a coalgebra structure. The particular functor we use automatically takes care of the condition on marginals.

Definition 2.1 A *Harsanyi type space (over M)* is a coalgebra for the functor F in the category Meas^I. A *universal type space* is a final coalgebra for F in Meas^I.

The points of $\mathsf{C}X$ are called *states of the world*. A point of X_i is called an *i-type*.

Our main result here is that there is a universal Harsanyi type space. Our proof follows that of Heifetz and Samet [5]. (However, they did not formulate type spaces as coalgebras.) In order to make the ideas more transparent, and also because the method is much more general, we shall temporarily forget about all of the machinery involved in the multi-player setting. Instead, we consider functors on Meas built using Δ. We show in Section 3 below that each such functor has a final coalgebra.

3 Coalgebraic Modal Logic for Functors on Measurable Spaces

3.1 Syntax and Semantics

Definition 3.1 The class of *measure polynomial functors* is the smallest class of functors on Meas containing the identity, the constant functor for each measurable space M and closed under products, coproducts, and Δ.

For a measure polynomial functor T, we define $\mathsf{Ing}(T)$, the *ingredients of* T, by the following recursion: For a "constant" space M, $\mathsf{Ing}(M) = \{M, Id\}$, $\mathsf{Ing}(Id) = \{Id\}$ $\mathsf{Ing}(U \times V) = \{U \times V\} \cup \mathsf{Ing}(U) \cup \mathsf{Ing}(V)$, and similarly for $U + V$; $\mathsf{Ing}(\Delta S) = \{\Delta S\} \cup \mathsf{Ing}(S)$. Each T has only finitely many ingredients.

Syntax

We define just below a language $\mathcal{L}(T)$. The language is sorted, and the sorts are the ingredients of T. We write $\varphi : S$ to mean that φ is a formula of sort S; when we need it, we let Form_S denote the set of such formulas.

(i) If $M \in \mathsf{Ing}(T)$ and A is a measurable subset of M, then $A : M$.

(ii) $\mathsf{true}_S : S$.

(iii) If $S \in \mathsf{Ing}(T)$ and $\varphi, \psi : S$, then also $\varphi \wedge \psi : S$.

(iv) If $U \times V \in \mathsf{Ing}(T)$, $\varphi : U$, and $\psi : V$, then $\langle \varphi, \psi \rangle^{U \times V} : U \times V$.

(v) If $U + V \in \mathsf{Ing}(T)$ $(V + U \in \mathsf{Ing}(T))$ and $\varphi : U$, then $\mathsf{inl}^{U+V} \varphi : U + V$ $(\mathsf{inr}^{V+U} \varphi : V + U)$.

(vi) If $\Delta S \in \mathsf{Ing}(T)$ and $\varphi : S$ and $p \in [0, 1]$, then $\beta^p \varphi : \Delta S$.

(vii) If $\varphi : T$, then $[\mathsf{next}]\varphi : Id$.

Semantics

Let $c : X \to TX$ be a coalgebra of T. The semantics assigns to each $S \in \mathsf{Ing}(T)$ and each $\varphi : S$ a subset $[\![\varphi]\!]_S^c \subseteq SX$.

$$
\begin{array}{rcl}
[\![A]\!]_M^c &=& A \\[4pt]
[\![\mathsf{true}_S]\!]_S^c &=& S(X) \\[4pt]
[\![\varphi \wedge \psi]\!]_S^c &=& [\![\varphi]\!]_S^c \cap [\![\psi]\!]_S^c \\[4pt]
[\![\langle \varphi, \psi \rangle^{U \times V}]\!]_{U \times V}^c &=& [\![\varphi]\!]_U^c \times [\![\varphi]\!]_V^c \\[4pt]
[\![\mathsf{inl}^{U+V}\varphi]\!]_{U+V}^c &=& \mathcal{P}\mathsf{inl}_{UX+VX}([\![\varphi]\!]_U^c) \\[4pt]
[\![\mathsf{inr}^{U+V}\varphi]\!]_{U+V}^c &=& \mathcal{P}\mathsf{inr}_{UX+VX}([\![\varphi]\!]_V^c) \\[4pt]
[\![\beta^p\varphi]\!]_{\Delta S}^c &=& \beta^p([\![\varphi]\!]_S^c) \\[4pt]
[\![[\mathsf{next}]\varphi]\!]_{Id}^c &=& c^{-1}([\![\varphi]\!]_T^c)
\end{array}
$$

The notation $\mathcal{P}f(A)$ indicates throughout the paper the image under f of the set A. We check easily that $[\![\varphi]\!]_S^c$ is always a measurable subset of SX. In the sequel, we shall omit the superscripts on the pairing, inl, and inr operators, since they are mostly clear from the context. We also will occasionally omit

the superscript c and the sort subscript when dealing with the semantics of $\varphi : S$ on a particular coalgebra $c : X \to TX$.

Remark 3.2 As we mentioned, if M is a measurable space, then the measurable subsets of M are taken as formulas. This departs from most of the treatments in coalgebraic modal logic, where one would take the *elements* of M as formulas; these formulas are then interpreted by singletons. Our treatment here makes for a more expressive language. We feel that when dealing with a space like $[0,1]$, one might want to have a set denoting a subinterval or a measurable subset of it. Also, there is a technical advantage: at various points, it will be good to know that the set of interpretations of formulas of all sorts coincide with the measurable subsets. To get this, one clearly must start with the measurable subsets of the constants. The only price we pay for this is that we require that all spaces be *separative*: that is, for each $x \in M$, $\{x\}$ is the intersection of all measurable A containing x. (However, this requirement may be lifted: see Section 5.1.)

Lemma 3.3 *Coalgebra morphisms preserve the semantics. That is, if $f : b \to c$ is a morphism of coalgebras $b : X \to TX$ and $c : Y \to TY$, and if $\varphi : S$, then $(Sf)^{-1}(\llbracket \varphi \rrbracket_S^c) = \llbracket \varphi \rrbracket_S^b$.*

3.2   The description operations and the canonical spaces

For each coalgebra $c : X \to TX$ and each $x \in SX$, we define

$$d_S^c(x) \quad = \quad \{\varphi : S \,|\, x \in \llbracket \varphi \rrbracket_S^c\}.$$

In the terminology of this paper's title, each such set $d_S^c(x)$ is a *satisfied theory*.

As an easy consequence of the definition and Lemma 3.3, coalgebra morphisms preserve description maps, i.e. in the conditions of the Lemma, $d_S^c \circ f = d_S^b$.

For each $S \in \mathsf{lng}(T)$, we define S^*, the *canonical domain of sort S*, to be the following measurable space. In each case, the underlying set is

$$S^* \quad = \quad \{d_S^c(x) \,|\, \text{ for some } c : X \to TX, \, x \in SX\}.$$

Note that each S^* is a set; indeed it has cardinality at most $2^{\mathfrak{c}\lambda}$, where $\mathfrak{c} = 2^{\aleph_0}$ is the cardinality of the continuum, and λ is the is maximum of the cardinalities of the sets of measurable subsets of the constant functors in $\mathsf{lng}(T)$. Usually we will use letters like s for elements of S^*. For the σ-algebra, we first take the subsets of S^* of the form

$$|\varphi|_S \quad = \quad \{s \in S^* \,|\, \varphi \in s\}. \tag{2}$$

Then the σ-algebra on each S^* is the one generated by this family of subsets. (Incidentally, each S^* is separative: $s = \bigcap \{|\varphi|_S \,|\, s \in |\varphi|_S\}$.)

Lemma 3.4 *For all $c : X \to TX$, all $S \in \mathsf{Ing}(T)$:*

  (i) *For all $\varphi : S$, $[\![\varphi]\!]_S^c = (d_S^c)^{-1}(|\varphi|)$.*

  (ii) *$d_S^c : SX \to S^*$ is measurable.*

Proof. $x \in [\![\varphi]\!]_S^c$ iff $\varphi \in d_S^c(x)$ iff $d_S^c(x) \in |\varphi|_S$. $\hfill\square$

At this point, we need to have a handle on the σ-algebras on the spaces $S(Id^*)$. We start by defining for each $\varphi : S$ a subset $\widehat{\varphi} \subseteq S(Id^*)$. The definition is by recursion on $\mathsf{Ing}(T)$ (rather than on φ):

For $\varphi : Id$, $\widehat{\varphi} = |\varphi|_{Id}$. For $A : M$, $\widehat{A} = A$. For $\langle \varphi, \psi \rangle : U \times V$, $\widehat{\langle \varphi, \psi \rangle} = \widehat{\varphi} \times \widehat{\psi}$. For $\varphi : U$, $\widehat{\mathsf{inl}\,\varphi} = \mathsf{inl}\,(\widehat{\varphi})$; and $\widehat{\mathsf{inr}\,\varphi} = \mathsf{inr}\,(\widehat{\varphi})$. For $\varphi : S$, $\widehat{\beta^p \varphi} = \beta^p \widehat{\varphi}$. And for all S, $\widehat{\mathsf{true}\,_S} = S(Id^*)$ and $\widehat{\varphi \wedge \psi} = \widehat{\varphi} \cap \widehat{\psi}$.

Lemma 3.5 *The family of sets $\widehat{\varphi}$ for $\varphi :: S$ are a set of generators of the σ-algebra on $S(Id^*)$, and this family is closed under finite intersections.*

3.3   The maps $r_S : S^* \to S(Id^*)$

Lemma 3.6 *There is a family of measurable maps $r_S : S^* \to S(Id^*)$ indexed by the ingredients of T such that the following hold:*

  (i) *For all coalgebras $c : X \to TX$ the diagram below commutes:*

$$
\begin{array}{ccc}
SX & & (3) \\
{\scriptstyle d_S^c}\Big\downarrow & \!\!\!\!\!\raise2pt\hbox{$\searrow$}^{\scriptstyle Sd_{Id}^c} & \\
S^* & \xrightarrow{\;r_S\;} & S(Id^*)
\end{array}
$$

  (ii) *For all $\varphi : S$, $r_S^{-1}(\widehat{\varphi}) = |\varphi|$.*

3.4   The canonical model of $\mathcal{L}(T)$

At this point, we are almost ready to define the canonical model. We need a preliminary concept first. For each $s \in Id^*$, let

$$[\mathsf{next}]^{-1}(s) \quad = \quad \{\varphi : T \,|\, [\mathsf{next}]\varphi \in s\}. \tag{4}$$

Lemma 3.7 *For each $s \in Id^*$, $[\mathsf{next}]^{-1}(s) \in T^*$. Moreover, this defines a measurable injective function $[\mathsf{next}]^{-1} : Id^* \to T^*$.*

Proof. Let $c : X \to TX$ and $x \in X$ be such that $s = d_{Id}^c(x)$. Then $c(x) \in TX$. We claim that $[\mathsf{next}]^{-1}(d_{Id}^c(x)) = d_T^c(c(x))$. For all $\varphi : T$,

$$c(x) \in [\![\varphi]\!]_T^c \quad \text{iff} \quad x \in c^{-1}([\![\varphi]\!]_T^c) = [\![[\mathsf{next}]\varphi]\!]_{Id}^c.$$

This verifies our claim. And from it, we see that $[\mathsf{next}]^{-1}(s)$ indeed belongs to T^*.

For the measurability assertion, note that for $\varphi : T$, the inverse image of $|\varphi|$ under $[\mathsf{next}]^{-1}$ is $|[\mathsf{next}]\varphi|$. (For notational reasons, we prefer not to write out that inverse image as $([\mathsf{next}]^{-1})^{-1}(|\varphi|)$.)

By definition of the σ-algebra on Id^*, this set is measurable.

We conclude by showing that $[\mathsf{next}]^{-1}$ is injective. Assume that $[\mathsf{next}]^{-1}(s) = [\mathsf{next}]^{-1}(t)$. Then s and t agree on all formulas of the form $[\mathsf{next}]\varphi$. So they agree on all formulas of sort Id. Thus $s = t$. □

We define $c^* : Id^* \to T(Id^*)$ to be

$$r_T \circ [\mathsf{next}]^{-1} : \ Id^* \to T^* \to T(Id^*) \tag{5}$$

Note that c^* is injective. We shall show that c^* is a final T-coalgebra. As our title indicates, we build final coalgebras from satisfied theories.

Lemma 3.8 (Truth Lemma) *For all $\varphi : S$, $[\![\varphi]\!]_S^{c^*} = \widehat{\varphi}$.*

3.5   The Final Coalgebra Theorem

Lemma 3.9 $d_{Id}^{c^*} = Id_{Id^*}$.

Proof. If $\varphi : Id$, then by the Truth Lemma, $[\![\varphi]\!]_{Id}^{c^*} = \widehat{\varphi} = |\varphi|$. So for $s \in Id^*$,

$$d_{Id}^{c^*}(s) \quad = \quad \{\varphi : Id \mid s \in [\![\varphi]\!]_{Id}^{c^*}\} \quad = \quad \{\varphi : Id \mid s \in |\varphi|\} \quad = \quad s.$$

□

Lemma 3.10 *For each coalgebra $c : X \to TX$, the diagrams below commute:*

$$
\begin{array}{ccc}
X & \xrightarrow{\ c\ } & TX \\
\ \downarrow{\scriptstyle d_{Id}^c} & & \ \downarrow{\scriptstyle d_T^c} \quad\ \searrow{\scriptstyle Td_{Id}^c} \\
Id^* & \xrightarrow[{[\mathsf{next}]^{-1}}]{} T^* & \xrightarrow[{r_T}]{} T(Id^*)
\end{array}
$$

Hence d_{Id}^c is a morphism of coalgebras.

Proof. The verification of the square is easy, and the triangle comes from Lemma 3.6. □

Theorem 3.11 $c^* : Id^* \to T(Id^*)$ *is a final coalgebra of T.*

Proof. Let $c : X \to TX$ be a T-coalgebra. By Lemma 3.10, d_{Id}^c is a coalgebra morphism. For the uniqueness, suppose that f is any morphism. Since f preserves descriptions, $d_{Id}^{c^*} \circ f = d_{Id}^c$. But by Lemma 3.9, $d_{Id}^{c^*} = Id_{Id^*}$. So $f = d_{Id}^{c^*} \circ f = d_{Id}^c$, just as desired.   □

We conclude with an important corollary of our development. We know of no direct proof of Corollary 3.12 below.

Corollary 3.12 *For each $S \in \mathsf{lng}(T)$, the map $r_S : S^* \to S(Id^*)$ is surjective.*

Proof. Consider the coalgebra $c^* : Id^* \to T(Id^*)$. By Lemma 3.6, $r_S \circ d_S^{c^*} = Sd_{Id}^{c^*}$. And by Lemma 3.9, $d_{Id}^{c^*} = Id_{Id^*}$. Thus $r_S \circ d_S^{c^*} = SId_{Id^*} = Id_{S(Id^*)}$. And this means that r_S is surjective.   □

4   The Universal Harsanyi Type Space

In this section, we show how to adapt our work to the case of Harsanyi type spaces considered as coalgebras on Meas^I. To make the notation more manageable, we assume that $I = \{1, 2, 3\}$. Let M be a fixed separative measurable space. Let Pr_1, Pr_2, and Pr_3 be the obvious projections $Pr_i : \mathsf{Meas}^I \to \mathsf{Meas}$. Let U_1, U_2, and U_3 be the functors $U_i : \mathsf{Meas}^I \to \mathsf{Meas}$ given as follows: $U_1(X_1, X_2, X_3) = M \times Pr_2 \times Pr_3$, $U_2(X_1, X_2, X_3) = M \times Pr_1 \times Pr_3$, $U_3(X_1, X_2, X_3) = M \times Pr_1 \times Pr_2$. Let $T_i = \Delta U_i$. So we are interested in the functor $\mathsf{F} : \mathsf{Meas}^I \to \mathsf{Meas}^I$ given by (T_1, T_2, T_3).

We write $\mathsf{lng}(\mathsf{F})$ for $\{M, U_i, T_i, Pr_i \mid i = 1, 2, 3\}$. Notice that as defined here, all the ingredients of F are functors from Meas^I to Meas.

We formulate our language $\mathcal{L}$ to have formulas of sort S for $S \in \mathsf{lng}(\mathsf{F})$. $\mathcal{L}$ is defined as follows: $\mathsf{true}_{Pr_i} : Pr_i$. If A is a measurable subset of M, then $A : M$. If $\varphi_0 : M$, $\varphi_2 : Pr_2$, $\varphi_3 : Pr_3$, then $\langle \varphi_0, \varphi_2, \varphi_3 \rangle : U_1$. We have similar clauses for U_2 and U_3. We also have clauses for the Δ functors: if $\varphi : U_i$, then $\beta^p \varphi : T_i$. If $\varphi : T_i$, then $[\mathsf{next}_i]\varphi : Pr_i$. (This is not really needed for the M sorts.)

Let $X = (X_1, X_2, X_3)$, and let $c : X \to \mathsf{F}X$ be a coalgebra of F. The semantics assigns to each $S \in \mathsf{lng}(\mathsf{F})$ and each $\varphi : S$ a subset $[\![\varphi]\!]_S^c \subseteq SX$. Here are some representative cases. $[\![\mathsf{true}_{Pr_i}]\!] = X_i$. $[\![[\mathsf{next}_i]\varphi]\!]_{Pr_i}^c = c_i^{-1}([\![\varphi]\!]_{T_i}^c)$, where $c = (c_1, c_2, c_3)$, or in other words, $c_i = Pr_i(c)$. Suppose that $\varphi_0 : M, \varphi_2 : Pr_2$ and $\varphi_3 : Pr_3$, then $\langle \varphi_0, \varphi_2, \varphi_3 \rangle : U_1$. We set $[\![\langle \varphi_0, \varphi_2, \varphi_3 \rangle]\!]_{U_1}^c = [\![\varphi_0]\!]_M^c \times [\![\varphi_2]\!]_{Pr_2}^c \times [\![\varphi_3]\!]_{Pr_3}^c$. Finally, suppose that $\varphi : U_3$ so that $\beta^p \varphi : T_3$. Given $[\![\varphi]\!]_{U_3}^c \subseteq U_3(X)$, we set $[\![\beta^p \varphi]\!]_{T_3}^c = \beta^p([\![\varphi]\!]_{U_3}^c) \subseteq T_3(X)$.

For a F-coalgebra X, and $S \in \mathsf{Ing}(\mathsf{F})$, we define d_S^c so that $d_S^c(x) = \{\varphi : S \mid x \in \llbracket\varphi\rrbracket_S^c\}$. We also define $|\varphi|_S$ and S^* just as before, using the sets $|\varphi|_S$ as the generators of the σ-algebra on S^*. Each d_S^c is again measurable.

Much of the remaining constructions are similar to what we have already seen. The role of Id^* is played by (Pr_1^*, Pr_2^*, Pr_3^*). This turns out to be the carrier of the final coalgebra for F. We define $d_{(Pr_1, Pr_2, Pr_3)}^c : X \to (Pr_1^*, Pr_2^*, Pr_3^*)$ as $(d_{Pr_1}^c, d_{Pr_2}^c, d_{Pr_3}^c)$.

Lemma 4.1 *There is a family of measurable maps $r_S : S^* \to S(Pr_1^*, Pr_2^*, Pr_3^*)$ indexed by $\mathsf{Ing}(\mathsf{F})$ such that the following hold:*

a. *For all coalgebras $c : X \to \mathsf{F}X$, $r_S \circ d_S^c = S d_{(Pr_1, Pr_2, Pr_3)}^c$.*

b. *r_S is injective.*

c. *The σ-algebra on $S(Pr_1^*, Pr_2^*, Pr_3^*)$ is generated by the sets $\overline{\varphi}$ for $\varphi : S$.*

The proof follows closely the work outlined in Lemma 3.6.

Now we have for each $i \in I$ a function $[\mathsf{next}_i]^{-1} : Pr_i^* \to T_i^*$ and $c_i^* = r_{T_i} \circ [\mathsf{next}_i]^{-1}$. The statement corresponding to that of Lemma 3.9 is that $d_{Pr_i}^{c^*} = Id_{Pr_i^*}$, so $d_{(Pr_1, Pr_2, Pr_3)}^{c^*} = Id_{(Pr_1, Pr_2, Pr_3)}$.

Theorem 4.2 *$c^* : (Pr_1^*, Pr_2^*, Pr_3^*) \to \mathsf{F}(Pr_1^*, Pr_2^*, Pr_3^*)$ is a final coalgebra of F.*

The proofs are the same, mutatis mutandis.

As the reader has probably guessed, the work here can be generalized to *polynomials on Meas^I*. These are I-indexed family of functors $T_i : \mathsf{Meas}^I \to \mathsf{Meas}$ built from the projections, constants for separative spaces, products and sums, and Δ. The work in this section generalizes to show that each polynomial on Meas^I has a final coalgebra. The details are not much more than what we have seen.

5 Further Variations and Extensions of the Basic Construction

We have already seen the main construction of final coalgebras for the polynomials on Meas built from separative spaces. We also saw (by example) how to generalize this to systems, thereby building the universal Harsanyi type spaces. The rest of the paper offers variations and extensions of the basic technique.

5.1 Extension: Non-separative Spaces

The main point of working with separative spaces is that the points of such a space may be recovered from the σ-algebra (as the set of all singleton intersections) and from the satisfied theories in our language. This allowed us to have a language for the measurable *subsets* of various spaces. In the absence of separativity, we need to do a little more work. We need to add the singleton sets $\{m\}$ to the formulas of sort M. Now the semantics of some sentences might not be measurable. So to compensate, we insist that in formulas of the form $\beta^p\varphi$, all of the subterms of constant sort are measurable. This side condition makes the semantics well-defined. The rest of the argument is essentially the same as what we have seen.

5.2 Variation: Kripke Polynomial Functors on Set

In this section, we check that the same method gives representations for final coalgebras for functors on Set built from the identity functor, the finite power set functor, product and coproduct, fixed (finite or infinite) sets, and functions from a fixed set. Following Rößiger and Jacobs, we call these *Kripke polynomial functors* (KPF's) on Set.

   The function space construct means that if S is a KPF and E is a set, then S^E is a KPF. As a functor, it is given by $S^E(a) = (S(a))^E$; this is the set of functions $\alpha : E \to Sa$. And if $f : a \to b$ so that $Sf : Sa \to Sb$, then $S^E f$ is given by $S^E f(\alpha) = (Sf) \circ \alpha$.

   We also take $\mathsf{Ing}(S^E) = \{S^E\} \cup \mathsf{Ing}(S)$.

   To avoid double subscripts or confusion with our notation $\mathcal{P}$ for the power set functor, we shall use $\mathcal{Q}$ for the *finite power set* functor on Set. Being a functor, we shall apply $\mathcal{Q}$ to functions as well as sets, writing, e.g., $\mathcal{Q}r(X)$ for the image $r[X]$ of the finite set X under r.

   It is well-known in the coalgebra literature that the finite Kripke polynomial functors all have final coalgebras. One can prove this by checking that all such functors are bounded and then using the much more general fact that bounded functors on Set have final coalgebras. Alternatively, one can use a logical approach, as done in papers such as [7,9,11,13]. This is the approach that we take. However, our work is a bit different than in the cited works since our final coalgebra is based on the satisfied theories rather than the maximal consistent ones in some logical system. This means that our result is actually weaker, since we do not obtain a completeness result. On the other hand, we believe that it is is easier to get the final coalgebra this way. And the general method works even in the absence of a logic, as we have seen in the work on measurable spaces.

Our syntax is constructed so that if for some set A, the associated constant functor $A \in \mathsf{Ing}(T)$, then each *element* $a \in A$ is a formula of sort A. Further, if $\mathfrak{Q}S \in \mathsf{Ing}(T)$ and $\varphi : S$, then $\Box\varphi : \mathfrak{Q}S$, and if $S^E \in \mathsf{Ing}(T)$ for some set E, $\varphi : S$ and $e \in E$, then $(e)\varphi : S^E$. Finally, we enrich our syntax to allow for boolean negation and conjunction in all sorts.

In our semantics, we define $[\![a]\!]^c_A = \{a\}$, and also

$$[\![\Box\varphi]\!]^c_{\mathfrak{Q}S} \quad = \quad \mathfrak{Q}([\![\varphi]\!]^c_S). \tag{6}$$

As usual, we also use the notation $\Diamond\varphi$ for the dual of $\Box\varphi$. So

$$[\![\Diamond\varphi]\!]^c_{\mathfrak{Q}S} \quad = \quad \{w \in \mathfrak{Q}SX \mid w \cap [\![\varphi]\!]^c_S \neq \emptyset\}.$$

The semantics of the $(e)\varphi$ construct is

$$[\![(e)\varphi]\!]^c_{S^E} \quad = \quad \{f \in S(X)^E \mid f(e) \in [\![\varphi]\!]^c_S\}. \tag{7}$$

The semantics of the boolean connectives are classical.

Most of the rest of the results from earlier in the paper go through with only minor changes, dropping the word "measurable" and anything having to do with the measurable space structure. The only differences are in Lemma 3.6 and the Truth Lemma 3.8. Before we turn to the analogs of those results, we need a general result.

Lemma 5.1 *Let C be any category, let $(X_i)_{i \in I}$ be a family of C-objects, assume that $X = \sum X_i$ is their coproduct with $\mathsf{in}_i : X_i \to X$ as the coproduct maps. Let $T : C \to C$ be a functor, and let $(c_i)_{i \in I}$ be a family of T-coalgebra maps, so $c_i : X_i \to TX_i$. Then there is a map $c : X \to TX$ so that for all i, in_i is a T-coalgebra morphism $\mathsf{in}_i : c_i \to c$.*

Proof. For each $i \in I$, $T\mathsf{in}_i \circ c_i : X_i \to TX$. By the universal property of the coproduct, there is some $c : X \to TX$ so that for all i, $T\mathsf{in}_i \circ c_i = c \circ \mathsf{in}_i$. So c is a coalgebra structure map. And as desired, in_i is a T-coalgebra morphism.□

The main difference between our work in this section and the earlier work in Section 3 is that this time we prove that the functions r_S are surjective before we can prove the Truth Lemma. (Recall that in the measurable space setting we derived the surjectivity of r_S in Corollary 3.12.)

Lemma 5.2 *There is a family of bijective maps $r_S : S^* \to S(Id^*)$ indexed by the ingredients of T such that for all coalgebras $c : X \to TX$, $r_S \circ d^c_S = Sd^c_{Id}$.*

The argument for Lemma 5.2 is rather close to what we find in classical work in modal logic.

The same work as in Section 3.5 gives the desired final coalgebra theorem.

Theorem 5.3 *For each Kripke polynomial functor T, $c^* : Id^* \to T(Id^*)$ is a final coalgebra.*

5.3   Other spaces

We are confident that the technique here extends to other kinds of concrete categories, such as compact Hausdorff spaces and the Borel probability measure with the weak ∗-topology. This case had been studied in relation to type spaces beginning with [2]. Indeed, we hope to expand the technique as much as possible in the coming months.

5.4   Conclusion and future directions

This paper has had two overall points. First, we connect work in the economics/game theory area with coalgebras. We feel that most, if not all, of the constructions of universal type spaces and related objects may be obtained by our method. We intend that the final version of our paper will show this. The more common method in the area is to construct an ω^{op}-limit and then follow this by a subspace construction. Though we did not discuss the matter here, we feel that the our method is somewhat simpler. As we have indicated, it generalizes easily to other settings. We also already generalized known results from particular functors to polynomials. Since the theories here are connected to logic, it should be possible to formulate logical systems for all the functors involved and prove completeness theorems. But this we leave to future work.

The construction of using satisfied theories to obtain final coalgebras is quite old in the area of coalgebra and its predecessors. For coalgebras per se, one can find it in Rutten [14]. The technique is perhaps implicit in Kurz, Kupke, and Venema [8]; this paper also contains a note that Jacobs' final coalgebra result in [7] for Kripke polynomials on Set contains an error. Our result does not use maximal consistent sets in a logical system but rather the satisfied theories. This is simpler, though of course one must do more work to get completeness results. From the side of coalgebra, perhaps what is most original here is taking a language for *sets* as opposed to (or, in addition to) a language of points. The interplay of the two languages is well worth studying.

References

[1] Patrick Billingsley, **Probability and Measure**, Second Edition. Wiley Series in Probability and Mathematical Statistics, New York, 1986.

[2] W. Böge and T. Eisele, "On solutions of Bayesian games", International Journal of Game Theory, 8, 1979, 193–215.

[3] Michèle Giry, "A categorical approach to probability theory," in **Categorical Aspects of Topology and Analysis**, Springer Lecture Notes in Mathematics 915, Berlin-New York, 1982, 68–85.

[4] John C. Harsanyi, "Games with incomplete information played by 'Bayesian' players", in *Papers in Game Theory* Theory and Decision Library, Volume 28 D. Reidel Publishing Company.

[5] Aviad Heifetz and Dov Samet, "Topology-free typology of beliefs". Journal of Economic Theory 82 (1998), no. 2, 324–341.

[6] Aviad Heifetz and Dov Samet, "Coherent beliefs are not always types". Journal of Mathematical Economics 32 (1999), 475–488.

[7] Bart Jacobs, "Many-sorted coalgebraic modal logic: a model-theoretic study." Theor. Inform. Appl. 35 (2001), no. 1, 31–59.

[8] Clemens Kupke, Alexander Kurz, and Yde Venema, "Stone coalgebras", in H. P. Gumm (ed.), Proceedings of CMCS 2002.

[9] Alexander Kurz, "Specifying coalgebras with modal logic", in H. P. Gumm (ed.), Proceedings of CMCS 1998.

[10] Jean-François Mertens and Shmuel Zamir, "Formulation of Bayesian analysis for games with incomplete information". International Journal of Game Theory 14 (1985), no. 1, 1–29.

[11] Martin Rößiger, "Languages for coalgebras on datafunctors." Proceedings of CMCS'99 (Amsterdam, 1999), ENTCS 19, Elsevier, Amsterdam, 1999.

[12] Martin Rößiger, "Coalgebras and modal logic." Proceedings of CMCS'2000 (Berlin, 2000), ENTCS 33, Elsevier, Amsterdam, 2000.

[13] Martin Rößiger, "From modal logic to terminal coalgebras." Theoretical Computer Science 260 (2001), no. 1-2, 209–228.

[14] Jan Rutten, "A calculus of transition systems." in A. Ponse et al (eds.), **Modal Logic and Process Algebra, a Bisimulation Perspective**, CSLI Lecture Notes 53, 1995.