Asociación Argentina de Mecánica Computacional

Mecánica Computacional Vol XXVI, pp.543-558
Sergio A. Elaskar, Elvio A. Pilotta, Germán A. Torres (Eds.)
Córdoba, Argentina, Octubre 2007

A COALGEBRAIC APPROACH TO TYPE SPACES

Ignacio D. Viglizzo

Departamento de Matemática, Departamento de Ciencias e Ingeniería de la Computación, INMABB-CONICET, Universidad Nacional del Sur, Avda. Alem 1253, 8000 Bahía Blanca, Argentina, viglizzo@criba.edu.ar

Keywords: Harsanyi type spaces, coalgebra, beliefs.

Abstract. When two or more players are engaged in a game with uncertainties, they need to consider what the other players' beliefs may be, which in turn are influenced by what they think the first player's ideas are. Harsanyi defined type spaces simply as a set in which all possible players-as defined by their beliefs- could be found. Later on, more meaningful constructions of this set were performed.

The theory of coalgebra, on the other hand, has been created to deal with circular phenomena, so its application to the problem of type spaces is only natural. We show how to apply it and we use the more general framework of category theory to compare the relative strength of previous solutions to the problem of defining type spaces.

1 MOTIVATION

A player in a game can be optimistic, pessimistic, cautious, daring, suspicious, paranoid, etc. We want to describe players according to their behaviors. Players displaying different kinds of behaviours will be called players of different *type*. Since behaviors are determined by the players' beliefs (assuming they are rational), these are going to be the focus of our attention. To get a mathematical definition, we need to be clear on which kind of games we are talking about, and then we can proceed to see how we can describe the 'type' of a player.

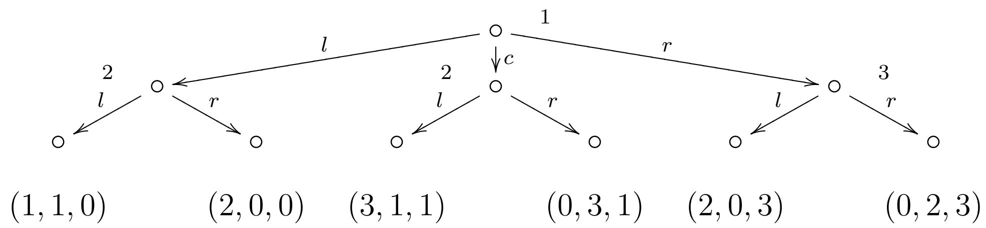
Definition 1.1. (following Osborne and Rubinstein (1994)) An *extensive game with perfect information* $G = (N, H, P, U_n)$ consists of:

- A set N, the set of players.
- A set H of sequences (finite or infinite) that satisfies the following three properties:
  - The empty sequence $\emptyset$ is in H.
  - If $(a_k)_{k=1,\ldots,K} \in H$ (where K may be infinite) and $L < K$ then $(a_k)_{k=1,\ldots,L} \in H$.
  - If an infinite sequence $(a^k)_{k=1}^{\infty}$ satisfies $(a^k)_{k=1,\ldots,L} \in H$ for every positive integer L, then $(a^k)_{k=1}^{\infty} \in H$.

  The members of H are called *histories*. A history $(a^k)_{k=1,\ldots,K} \in H$ is *terminal* if it is infinite or there is no a^{K+1} such that $(a^k)_{k=1,\ldots,K+1} \in H$. The set of terminal histories is denoted with Z.
- A function $P : H \setminus Z \to N$, that indicates for each history in H which one of the players takes an action after the history.
- Functions $U_n : Z \to \mathbb{R}$ for $n \in N$ that give for each terminal history and each player, the *payoff* of that player after that history.

The set H can be seen as a tree with root $\emptyset$, with its nodes labeled by the function P, and the leaves labeled by the functions U_n. We indicate the elements a^k on the edges of the tree so following a particular branch from the root will give the history that names each node.

Example 1.1.



In the diagram above we have a game where $N = \{1, 2, 3\}$; $P(\emptyset) = 1$ meaning that player 1 gets to decide the first move in the game, and has three options available: l, c, r (the letters stand for left, center or right, respectively). If player 1 chooses l or c, then player 2 decides what's the next action, and she has options l and r available. If player 1 chooses r instead, it is player 3 who decides what's the final move. Under each terminal node in the tree, a triple indicates the values of the utility functions U_1, U_2 and U_3. So, for example if the history of the game is (c, l), then player 1 gets a payoff of 3, while players 2 and 3 get a payoff of 1.

Alternatively, extensive games with complete information can be given by indicating (instead of the utility functions) a family of preorders $(\prec_n)_{n\in N}$ that represent the *preferences* of the players. For our purposes, it will be enough to assume that all players prefer to maximize their payoffs and are indifferent to what other players' payoffs are.

Games with incomplete information are games in which the incompleteness of the information arises in three main ways.

1. The players may not know the *physical outcome function* of the game which specifies the physical outcome produced by each strategy available to the players.

2. The players may not know their own or some other players' *utility functions*, which specify the utility payoff that a given player i derives from every physical outcome.

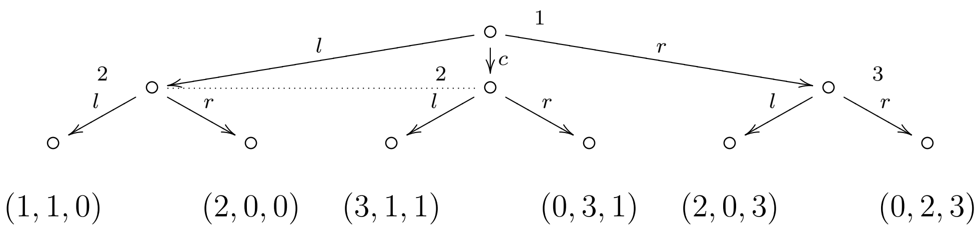
3. The players may not know their own or some other players' *strategy space*, i.e. the set of all strategies available to various players.

'All other causes of incomplete information can be reduced to these three basic cases– indeed sometimes this can be done in two or more different (but essentially equivalent) ways' (Harsanyi, 1967). The challenge is to be able to take the best possible decisions when these uncertainties are present. A breakthrough in this field was made in 1967, when a series of papers by John C. Harsanyi, (Harsanyi, 1967, 1968a,b) saw print. The idea was to tame the uncertainty by transforming the games with incomplete information into games with complete but *imperfect* information.

Definition 1.2. An extensive game with *imperfect information* is a game $G = (N, H, P, U_n, \mathcal{I}_n)$ where N, H, P and U_n are as in Definition 1.1, and for each player $n \in N, \mathcal{I}_n$ is a partition on the set $H_n = \{h \in H \setminus Z : P(h) = n\}$. The equivalence classes in this partitions are called *information sets*.

The idea here is that player n knows in which information set the game currently is, but doesn't know exactly the whole history that has lead the game into that set. Note that the players still have perfect information. They know the payoffs in all the possible outcomes.

Example 1.2.



Now the dotted line indicates that the set $\{l, c\}$ is an information set for player 2. She does not have information about whether player 1 moved to the right or to the center, but she does know what the payoffs will be in each case, and also knows that, since it's her turn, player 1 did not choose r.

If all the information sets contain exactly one node of the tree, we have a game with perfect information. The information sets allow us to represent games in which the players make their moves simultaneously (and thus don't know when making their decision what are the other players' moves), and also to represent situations in which "nature" or "chance" make a move we cannot predict. This feature will be exploited later.

Some further assumptions are made about the games under study. In the first place, it is assumed that the beliefs the players of the game have can be represented through probability measures (this is called the *Bayesian approach*). It is also assumed that the players are aware of the extent of the knowlege or ignorance of the other players, and that they will always act "rationally", that is, they will take the action that gives them the highest possible expected payoff, based on the information available to them. The notion of rationality is quite hard to formulate and still topic of debate in among game theorists.

In Harsanyi's words, (Harsanyi, 1967):

> It seems to me that the basic reason why the theory of games with incomplete information has made so little progress so far lies in the fact that these games give rise, or at least appear to give rise, to an infinite regress in reciprocal expectations on the part of the players.

The argument is the following: suppose the game has incomplete information and just two players. Player 1 has some beliefs about what are the actual values of the missing information. This is represented as a probability measure over the space of all possible values the unknown could take. Player 1 also knows that player 2 cannot know the actual value and hence resorts to using a probability distribution representing her beliefs as well. In order to take a decision, player 1 then must form some mental model of what player 2's beliefs are. Player 2's beliefs include those that, in turn, player 2 has about player 1's beliefs. This kind of reasoning promptly leads to an infinite regression of unfolding beliefs. Harsanyi calls any model of this kind a *sequential-expectations* model for games with incomplete information.

Harsanyi was concerned with finding ways of analising these games with incomplete information. The solution he offered involved the construction of a game with complete but imperfect information based on the given one with incomplete information. In the new game, there are new chance moves that are assumed to occur before the two players choose their strategies. In these random moves, the actual payoff of the two players are determined, but being a game with imperfect information, the players only know they are in some information set, and a probability distribution for the random moves (this probability distribution is assumed to be common knowledge to all the players). Using conditional probabilities, they can then derive the different expected values they need to assess the strategies to be taken in the game.

There is an alternative interpretation of the random moves added to the game. Instead of assuming that they determine important characteristics of the players (in particular, their payoffs), it could be assumed that the players themselves are being chosen at random from 'certain hypothetical populations containing individuals of different "types", each possible "type" of a player i being characterized by a different attribute vector c_i, i.e., by a different combination of production costs, financial resources, and states of information.'(Harsanyi, 1967)

It is these populations that we'll call *type spaces*, and their elements will be of course, *types*. While Harsanyi assumes the type space was given, he already suggested they could be constructed from the considerations about beliefs explained above:

> As we have seen, if we use the Bayesian approach, then the sequential-expectations model for any given [incomplete information] game G will have to be analyzed in terms of infinite sequences of higher and higher-order subjective probability distributions, i.e. subjective probability distributions over subjective probability distributions (Harsanyi, 1967).

Harsanyi was discouraged from this approach by the technical difficulties it presented:

> Probability distributions over some space of payoff functions or of probability distributions, and more generally probability distributions over function spaces, involve certain mathematical difficulties [...]. However, as Aumann has shown (Aumann, 1961) and (Aumann, 1964), these mathematical difficulties can be overcome. But even if we succeed in defining the relevant higher order probability distributions in a mathematically admissible way, the fact remains that the resulting model –like *all* models based on the sequential-expectations approach–will be extremely complicated and cumbersome.

The difficulty pointed out by Aumann in (Aumann, 1961) is that if X and Y are measurable spaces and we denote by Y^X the set of all measurable functions from X to Y, then there is no natural way of endowing Y^X with a σ-algebra that makes the evaluation function $ev : Y^X \times X \to Y$ given by $ev(f, x) = f(x)$ measurable. Aumann proposes in (Aumann, 1964) to choose a single real number that represents a probability distribution. In our approach, the problem is overcome by considering the spaces ΔX of all probability measures over X instead of looking at all the measurable functions in $[0, 1]^X$ that have integral 1 over X.

So, to formalize the notion of types that Harsanyi had in mind, we want a mathematical object, the *type space*, such that each element or *type* will have associated to it, in a natural way, beliefs (represented by probability distributions) over the states of nature and the types of the other players in the game. In a game with N players, each player will assume one of the types $t \in T$, as if they were roles in a play.

A first approach would be to find a correspondence $T \cong \Delta(S \times T)$, where the set T would be the type space and S the *states of nature*. The states of nature are the possible values the unknown variables in the game can take. We want both S and T to be measurable spaces so we can define probability measures on them. Let $m : T \to \Delta(S \times T)$ be the desired isomorphism. Then for each $t \in T$, $m(t)$ represents the beliefs of a player of type t.

To find this isomorphism, we will introduce first the notion of coalgebras, for which we will also need some of the language provided by category theory.

2 COALGEBRAS

The theory of coalgebras was introduced to model certain circular phenomena, like the theory of non wellfounded sets (see Aczel (1988)). It has been found to encompass many different examples, and it has abstracted interesting properties out of them.

Given a category $\mathcal{C}$ and an endofunctor $F : \mathcal{C} \to \mathcal{C}$, a *coalgebra* for the functor F (or F-*coalgebra*) is a pair (X, c) consisting of an object X in the category $\mathcal{C}$ and a morphism $c : X \to F(X)$. Given F-coalgebras (X, c) and (Y, d), a F-*coalgebra morphism* is a $\mathcal{C}$-morphism $f : X \to Y$ such that the following diagram commutes:

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle c}\downarrow & & \downarrow{\scriptstyle d} \\ F(X) & \xrightarrow[F(f)]{} & F(Y) \end{array}$$

A *final object* in a category $\mathcal{C}$ is an object $\mathbf{1}$ such that for any $\mathcal{C}$-object A there exists a unique morphism $!_A : A \to 1$. We will be interested in the final objects of the categories of F-coalgebras (*final coalgebras*), mainly because of the following property:

Theorem 2.1. *(Lambek's Lemma, Lambek (1968)) If (Z, e) is a final F-coalgebra, then e is an isomorphism between Z and $F(Z)$.*

Final coalgebras need not to exist in general. For example, the powerset functor in the category of sets has no final coalgebra (if it had one, Lambek's Lemma would contradict Cantor's theorem).

Since we are representing beliefs as probability measures, we will work within the category Meas of measurable spaces and measurable functions between them.

We will consider the endofunctor Δ in Meas that assigns to each measurable space M, the set ΔM of all the probability measures over M, endowed with the σ-algebra Σ_Δ generated by the sets of the form $\beta^p(E)$ where E is a measurable subset of M and

$$\beta^p(E) = \{\mu \in \Delta M : \mu(E) \geq p\} \tag{1}$$

If $f : M \to N$ is measurable, we define $\Delta f : \Delta M \to \Delta N$ as follows: for $\mu \in \Delta(M)$ and $E \in \Sigma'$,

$$(\Delta f)(\mu)(E) = \mu(f^{-1}(A)).$$

Definition 2.1. The class of *measure polynomial functors* is the smallest class of functors on Meas containing the identity Id, the constant functor M for each measurable space M, and closed under binary products, coproducts and Δ.

Example 2.1. If we consider a fixed measurable space like the real interval $[0, 1]$ with its borel subsets as the measurable ones, we can build the measure polynomial functors $\Delta([0, 1] \times X), (X + [0, 1]) \times \Delta X$, etc.

Theorem 2.2. *(Moss and Viglizzo, 2006; Viglizzo, 2005b) All polynomial measure functors have final coalgebras.*

3 TYPE SPACES AS COALGEBRAS

We now see that the isomorphism $m : T \to \Delta(S \times T)$ would be a byproduct of finding the final coalgebra for the functor taking a measurable space X to $\Delta(S \times X)$. There are some problems with this approach. If the game has N players, then each player i with type t_i should have beliefs about the types of all the other players, so the functor to use could be

$$F(T) = \Delta(S \times T^N). \tag{2}$$

Furthermore, we want each type to know his own type. To model this, we need some definitions and results from measure theory.

Given a probability distribution μ over a product space $X \times Y$, its *marginals* are the distributions μ_X and μ_Y over the spaces X and Y respectively defined by $mar_X\mu(E) = \mu(E \times Y)$ and $mar_Y\mu(F) = \mu(X \times Y)$ for all E measurable subset of X and F measurable subset of Y. Using the functor Δ and the projections, we may write this as $mar_X\mu = (\Delta\pi_X)\mu = \mu \circ \pi_X^{-1}$; $mar_Y\mu = (\Delta\pi_Y)\mu = \mu \circ \pi_Y^{-1}$.

If X is a measurable space and x is a point in X, let δ_x be the probability distribution given by $\delta_x(E) = 1$ if $x \in E$ and 0 otherwise.

So we don't want T to be isomorphic to $\Delta(S \times T^N)$, but to the subset of $\Delta(S \times T^N)$ of probability distributions in which the marginal of each $m(t_i)$ on the i-th copy of T is the distribution δ_{t_i} which has support on the point t_i. Adding this extra condition to the definition would steer us away from the definition of coalgebras on Meas, but we can overcome this difficulty by changing the functor in an appropriate way. The key observation here is that for any product of measurable spaces $A \times B$ and $b_0 \in B$ such that the singleton $\{b_0\}$ is measurable, there is an isomorphism between the spaces $\{\mu \in \Delta(A \times B) : mar_B\mu = \delta_{b_0}\}$ and ΔA.

The following Lemma proves that in the case above, it is enough to know the marginals to determine the measure.

Lemma 3.1. *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 $\mu(G \times \{b_0\}) + 0$. □

So now we can model the introspection condition by considering coalgebras for the functor

$$F(T) = \Delta(S \times T^{N-1}). \tag{3}$$

The problem of finding a *universal type space*, that is, a type space containing all the possible types a player could adopt, could be solved by finding the final coalgebra for the functor $F(X) = \Delta(S \times X^{N-1})$. This can be done using Theorem 2.2. Lambek's Lemma 2.1 provides the isomorphism we are looking for.

But when we look at a single coalgebra for this functor, that is, a measurable map $m : T \to \Delta(S \times T^{N-1})$ we get a somewhat unsatisfactory model. Why should all the players come from the same type space? It would be better to be more general and to assume that there are type spaces $T_1, T_2, \dots, T_N$ and the type of player i is selected from the corresponding T_i. This motivates the following definition:

Definition 3.1. Let Meas^N be the N-fold product of the category Meas. Each object M in Meas^N is a N-tuple of measurable spaces $(M_1, \dots, M_N)$, and the morphisms are N-tuples of measurable functions $f_i : M_i \to M_i'$. Let $Proj_i^N$: Meas$^N \to$ Meas be the i-th projection functor.

Definition 3.2. We define then a *type space* for a game with N players over the measurable space S of states of nature, as a coalgebra for the endofunctor in Meas^N given by $T = (T_1, T_2, \dots, T_N)$ where for $1 \leq i \leq N$,

$$T_i = \Delta(S \times \prod_{j \neq i} Proj_j^N). \tag{4}$$

The diagram for a coalgebra (X, m) of this functor is:

$$\begin{array}{cccc} (X_1, & X_2, & \dots, & X_N) \\ \downarrow m_1 & \downarrow m_2 & & \downarrow m_N \\ (\Delta(S \times \prod_{j \neq 1} X_j), & \Delta(S \times \prod_{j \neq 2} X_j), & \dots, & \Delta(S \times \prod_{j \neq N} X_j)) \end{array}$$

The definition above is a particular case of the more general one that follows.

Definition 3.3. A *measure polynomial functor on many variables* $T : \mathsf{Meas}^N \rightarrow \mathsf{Meas}$ is a functor built from the functors $Proj_1^N, \ldots, Proj_N^N$ and constant functors for measurable spaces, using either products, coproducts and Δ. For any natural number N', we can extend the notion of a measure polynomial functor to functors $T = (T_1, \ldots, T_{N'}) : \mathsf{Meas}^N \rightarrow \mathsf{Meas}^{N'}$ such that each $T_i, 1 \leq i \leq N'$, is a measure polynomial functor on many variables from Meas^N to Meas as defined above.

Example 3.1. For a fixed measurable space M, consider the polynomial functor on three variables $F : \mathsf{Meas}^3 \rightarrow \mathsf{Meas}^2$ given by:

$$F \ = \ (\ \Delta(Proj_1^3 + Proj_2^3)\ ,\ ((\Delta Proj_3^3) \times Proj_2^3) + M\ )$$

We are going to center our attention on measure polynomial functors on many variables that are endofunctors of the category Meas^N, and the coalgebras for those functors.

Theorem 3.1. *(Viglizzo, 2005a) If $T : \mathsf{Meas}^N \rightarrow \mathsf{Meas}^N$ is a measure polynomial functor in many variables, then it has a final coalgebra.*

Going back to the type spaces for a game with N players, application of the Theorem above yields a final type space, also known in the literature as *universal type space*. We also get the following Lemma:

Lemma 3.2. *If $T : \mathsf{Meas}^N \rightarrow \mathsf{Meas}^N$ is the functor given by $(T_i = \Delta(S \times \prod_{j \neq i} Proj_j^N))_{1 \leq i \leq N}$, and $(Z_i)_{1 \leq i \leq N}$ is a final coalgebra for T then for each i, Z_i is isomorphic to $\Delta(S \times \prod_{j \neq i} Z_j)$ and all the spaces $Z_i, 1 \leq i \leq N$ are isomorphic.*

The fact that all the type spaces in the universal type space for a game with N players are isomorphic, together with the fact that all final coalgebras for a given functor are isomorphic justifies naming it *the* universal type space for the game.

4 A BRIEF REVIEW OF THE LITERATURE ON TYPE SPACES

There have been several constructions of type spaces and universal type spaces in the literature, each one trying to capture the intuitive idea behind the definition in a slightly different way. Here we review them, as we compare them with the framework we just exposed.

4.1 Armbruster, Böge and Eisele

In *Bayesian Game Theory* (Armbruster and Böge, 1979), W. Armbruster and W. Böge present their approach to the study of games with unknown utility functions, in which the players "will have at least a subjective probability distribution on [the] alternatives". This is called the *Bayesian assumption*. In order to construct "canonical representations for the players' subjective probability measures", the following notion is introduced, and attributed to Böge, in a lecture on game theory given in 1970.

Definition 4.1. Let $S_1^0, \ldots, S_N^0$ be compact Hausdorff spaces. An N-tuple of compact sets and continuous maps $(S_1, \ldots, S_N, \rho_1, \ldots, \rho_N)$ is called an *oracle system* for $S_1^0, \ldots, S_N^0$ if for all i, $\rho_i : S_i \rightarrow S_i^0 \times \prod_{j \neq i} \Delta_r(S_j)$. Here Δ_r is the functor that assigns to each topological space X the space of all the probability distributions over X with the σ-algebra of its borel sets.

This is the same as saying that (S, ρ) is a coalgebra for the functor $T = (S_i^0 \times \prod_{j \neq i} \Delta_r(Id_j))_{1 \leq i \leq N}$ in the category CHaus^N where CHaus is the category of compact Hausdorff spaces and continuous functions. The underlying assumption here is that each player has a different space of state of nature S_i^0 in which their unknowns lie.

The final coalgebra is constructed by taking the projective limit of the corresponding final sequence. This final coalgebra is called the *canonical oracle system*. Note that not all the components of the functor are the same, so in general the spaces Z_i will not be isomorphic to each other as in Lemma 3.2. This is a reasonable assumption, and using Theorem 3.1, one can extend the definition and existence of canonical oracle systems to the general case of measurable spaces.

It is important to note that here appears for the first time a coalgebra (not necessarily the final one) as a model of the beliefs of a player. This transcends the idea of just looking for the space of all possible types, to give more restricted models that can be useful to describe situations in more manageable terms.

W. Böge and Th. Eisele present a slightly different approach in the paper *On Solutions of Bayesian Games*, (Böge and Eisele, 1979). Here again the topological setting is the category CHaus. The space over which the behavior of the players is selected is similar to the one we proposed in (2), but with certain restrictions.

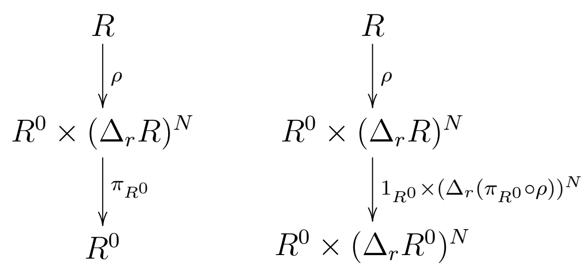
Given a compact space of states of nature R^0, a nonempty subspace $R^1 \subseteq R^0 \times (\Delta_r R^0)^N$ of common a-priori information is selected.

Definition 4.2. A system (R, ρ) with

$$\rho : R \to R^0 \times (\Delta_r R)^N$$

is called a *system of complete reflections over the information set* R^1 if

$$(1_{R^0} \times (\Delta_r(\pi_{R^0} \circ \rho))) \circ \rho \subseteq R^1 \subseteq R^0 \times (\Delta_r R^0)^N. \qquad (5)$$



The space R^1 has to satisfy a couple of conditions, the first one specifying that each player knows what their beliefs are, and the second one saying that each player will try to maximize their utility function. These requirements preclude the systems of complete reflections from being coalgebras. We have seen before how the first condition, of each player knowing their beliefs, can be dealt with by taking a different functor.

The construction of the final object in the category of systems of complete reflections is done by taking the projective limit, and restricting the spaces so that the image of the map ρ for the final object has image contained in R^1.

4.2 Mertens and Zamir

The paper *Formulation of Bayesian Analysis for Games with Incomplete Information* by Jean-François Mertens and Shmuel Zamir, (Mertens and Zamir, 1984), is the most often cited one in the literature about type spaces.

Starting from a compact space S called *parameter-space* or set of states of nature, they seek to define a set Y of the "states of the world" in which every point contains all characteristics, beliefs and mutual beliefs of all players. The equations that summarize their goals are:

$$Y = S \times T^N \tag{6}$$

$$T = \text{ the set of all probability distributions on } (S \times T^{N-1}) \tag{7}$$

These equations are, of course, intended to be solved up to isomorphism. Equation (7) is essentially our (3). Some of the definitions in this work are interesting and we will analyse them here, trying to understand their motivation and how they are accounted for in our model.

Definition 4.3. (Mertens and Zamir, 1984) Let S be a compact space. An *S-based abstract beliefs space (BL-space)* is an $(N + 3)$ tuple $(C, S, f, (t^i)_{i=1}^N)$ where C is a compact set, f is a continuous mapping $f : C \to S$ and $t^i, i = 1, \ldots, N$, are continuous mappings $t^i : C \to \Delta(C)$ (with respect to the weak-* topology) satisfying:

$$\tilde{c} \in C \text{ and } \tilde{c} \in Supp(t^i(c)) \Rightarrow t^i(\tilde{c}) = t^i(c). \tag{8}$$

The condition (8) specifies that "a player assigns positive probability (in the discrete case) only to those points in C in which he has the same beliefs. In other words, he is certain of his own beliefs." It can be rewritten as:

$$\tilde{c} \in C \text{ and } \tilde{c} \in Supp(t^i(c)) \Rightarrow \tilde{c} \in (t^i)^{-1}[t^i(c)].$$

Or the following equivalent equations:

$$Supp(t^i(c)) \subseteq (t^i)^{-1}[t^i(c)]$$
$$t^i(c)[(t^i)^{-1}[t^i(c)]] = 1$$
$$(\Delta t^i)t^i(c) = \delta_{t^i(c)}.$$

Thus, even though the first impression could be that Belief spaces are coalgebras for the functor $FX = S \times \Delta X$, we see immediately that we need the function f to have the specific codomain S, and we need many different functions t^i with codomain ΔC.

However, we can see that an adaptation from our definitions yields spaces with the same properties: If (X, m) is a type space for a game over S with N players, as in Definition 3.2, then let

$$C = S \times \prod_{i=1}^N X_i$$
$$C_{-i} = S \times \prod_{j \neq i} X_j$$

Let π_i and π_{-i} be the projections from C to X_i and C_{-i}, respectively. Now for all $c \in C$, let $t^i : C \to \Delta C$ be defined by

$$t^i(c) = m_i \pi_i(c) \times \delta_{\pi_i(c)}.$$

Thus $t^i(c) \in \Delta C$. Letting $\pi_S : C \to S$ be the projection, we have that

Proposition 4.1. *$(C, S, \pi_S, (t^i)_{i=1}^N)$ is a BL-space.*

Proof. We only need to check that condition (8) is satisfied. Notice that the type spaces of Definition 3.2 are defined for any measurable space S, and the functions m_i need not be continuous, just measurable. Condition (8) is stated in terms of the support of the probability measure $t^i(c)$, which does not necessarily exist in the more general case. We will prove the condition $t^i(c)[(t^i)^{-1}[t^i(c)]] = 1$ which is equivalent to (8) when the support is defined.

$$\begin{aligned} t^i(c)[(t^i)^{-1}(t^i(c))] &= t^i(c)[(t^i)^{-1}(m_i\pi_i(c) \times \delta_{\pi(c)})] \\ &= t^i(c)[((m_i \times \delta) \circ \pi_i)^{-1}(m_i\pi_i(c) \times \delta_{\pi_i(c)})] \\ &= t^i(c)[(\pi_i)^{-1}(m_i \times \delta)^{-1}(m_i\pi_i(c) \times \delta_{\pi_i(c)})] \end{aligned}$$

The set $(m_i \times \delta)^{-1}(m_i\pi_i(c) \times \delta_{\pi_i(c)})$ is not empty, since at least $\pi_i(c)$ is in it. It is also equal to the set $m_i^{-1}m_i\pi_i(c) \cap \delta^{-1}(\delta_{\pi_i(c)}) = m_i^{-1}m_i\pi_i(c) \cap \{\pi_i(c)\}$ so its inverse image under π_i is $C_{-i} \times \{\pi_i(c)\}$. Therefore

$$\begin{aligned} t^i(c)[(t^i)^{-1}(t^i(c))] &= t^i(c)[C_{-i} \times \{\pi_i(c)\}] \\ &= m_i\pi_i(c)(C_{-i}) \times \delta_{\pi_i(c)}(\pi_i(c)) \\ &= 1 \end{aligned}$$

□

Note that in Mertens and Zamir's approach, the universal type spaces are constructed by constructing first the universal BL-space Y and then taking taking $T = t^i(Y)$, while here we have shown how to construct belief spaces from the type spaces.

Definition 4.4. (Mertens and Zamir, 1984) A *coherent beliefs hierarchy [over S] of level K* ($K = 1, 2, \ldots$) is a sequence $(C_0, C_1, \ldots, C_K)$ where:

1. C_0 is a compact subset of S and for $k = 1, \ldots, K, C_k$ is a compact subset of $C_{k-1} \times [\Delta(C_{k-1})]^N$ (as topological spaces). We denote by ρ_{k-1} and t^i the projections of C_k onto C_{k-1} and the i-th copy of $\Delta(C_{k-1})$ respectively.

$$C_0 \xleftarrow{\rho_0} C_1 \xleftarrow{\rho_1} \cdots \xleftarrow{\rho_{K-1}} C_K$$

2. $$\rho_{k-1}(C_k) = C_{k-1}; k = 1, \ldots, K$$

3. For all $c_k \in C_k$, let $c_{k-1} = \rho_{k-1}(c_k)$. Then for all i, and $k = 2, \ldots, K$,

   H1) the marginal distribution of $t^i(c_k)$ on C_{k-2} is $t^i(c_{k-1})$;

   H2) the marginal distribution of $t^i(c_k)$ in the i-th copy of $\Delta(C_{k-2})$ is the unit mass at $t^i(c_{k-1}) = t^i(\rho_{k-1}(c_k))$.

The coherent hierarchies are used to build the universal beliefs space Y. Thay can be seen as the first K steps in the iteration that leads to the final sequence. The additional conditions we see come from different complications introduced in the construction. Part 2 of the definition states that the projections should be surjective. This condition is necessary here because the spaces C_k are compact subspaces of $C_{k-1} \times (\Delta C_{k-1})^N$ and not that whole space.

Conditions **H1**) and **H2**) of part 3 have the following intuitive meaning:

H1) says that player i's k-level beliefs coincide with his $(k-1)$ level beliefs in whatever concerns hierarchies up to level$(k-2)$. Condition **H2**) says that player i knows his own previous order beliefs. (Mertens and Zamir, 1984)

Under a more technical light, **H1**) can be written as

$$(\Delta\rho_{k-2})t^i(c_k) = t^i(c_{k-1}) = t^i(\rho_{k-1}(c_k)) \tag{9}$$

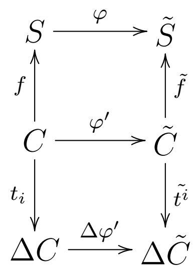
for every $c_k \in C_k$. This condition is saying that c_k is an element of the projective limit of the spaces C_k. The condition **H2**) can be written as: for every $c_k \in C_k$,

$$(\Delta\rho_{k-2})t^i(c_k) = t^i(c_{k-1}) = t^i\rho_{k-1}(c_k). \tag{10}$$

There is some abuse of notation here: for each number $k \geq 1$, functions $t^i : C_k \to \Delta C_{k-1}$ are defined, so there is a different function t^i that is applied to c_k and another one that's applied to c_{k-1}, and it should be clear which one is needed in each ocurrence of t^i. Having (10) is needed in order to obtain (8) in the projective limit.

Morphisms between BL-spaces are defined as follows:

Definition 4.5. (Mertens and Zamir, 1984) A *beliefs morphism* (BL-morphism) from a BL-space $(C, S, f, (t^i)_{i=1}^N)$ to a BL-space $(\tilde{C}, \tilde{S}, \tilde{f}, (\tilde{t}^i)_{i=1}^N)$ is a pair (φ, φ') where φ' is a continuous mapping from C to $\tilde{C}$ and φ is a continuous mapping of S to $\tilde{S}$ such that for each $i; i = 1, 2, \ldots, n$, the following diagram commutes:



Given a fixed space S of states of nature, the *universal BL-space* is the final object in the category of BL-spaces over S and BL-morphisms. The universal BL-space over a fixed space S is built by taking the projective limit Y of a sequence of coherent beliefs hierarchies:

Let $Y_0 = S; Y_1 = S \times \Delta S \times \ldots \times \Delta S$ and for $k \geq 2$, let $Y_k = \{y_k \in Y_{k-1} \times [\Delta(Y_{k-1})]^N :$ **H1)** For all i the marginal distribution of $t^i(y_k)$ on Y_{k-2} is $t^i(y_{k-1})$ and **H2)** the marginal distribution of $t^i(y_k)$ on $\Delta^i(Y_{k-2})$ is the unit mass at $t^i(y_{k-1})\}$.

With this definition, for each value of k, the sequence $(Y_0, \ldots, Y_k)$ is a coherent beliefs hierarchy over S of level k, and it also is the biggest one that can be constructed. All the coherent hierarchies of beliefs can be mapped to the ones constructed above, and all the BL-spaces can be mapped in a unique way to their limit Y.

It is clear from the proof given that the spaces under consideration are assumed to be compact Hausdorff topological spaces. Mertens and Zamir use Riesz's Representation theorem to prove what essentially amounts to Theorem 4.1 below, but one needs to also assume that the probability measures involved are all regular, as (Armbruster and Böge, 1979) and (Heifetz, 1993) point out.

Theorem 4.1. *(see Heifetz (1993), also Mètivier (1963), Theorem III.3.2) Let X_n be a sequence of Hausdorff topological spaces and $f_n : X_{n+1} \to X_n$ a surjective continuous map for $n \geq 0$. If μ_n is a regular Borel probability measure on X_n such that $\mu_{n+1}f_n^{-1} = \mu_n$ for all $n \geq 0$, then there is a unique regular Borel probability measure μ on the projective limit of the X_n such that for all $n \geq 0$, $\mu\pi_n^{-1} = \mu_n$.*

4.3 Heifetz and Samet

Aviad Heifetz and Dov Samet, in their paper *Topology-Free Typology of Beliefs*, (Heifetz and Samet, 1998), are the first to solve the problem of finding the universal type space in the general case of measurable spaces. They present two constructions of the space, much in the spirit of the two constructions of final coalgebras for measure polynomial functors presented in Moss and Viglizzo (2006) and Viglizzo (2005b), respectively.

Their methods have provided us with guiding insight for the constructions of final coalgebras we cited before. In the language-based construction, we have presented new languages $\mathcal{L}(T)$ based on each measure polynomial functor T. We have also introduced two important refinements:

Their operator $B_i^p(e)$ is used to express that a player i believes that an event represented by e has probability bigger than p. In our formulation, this would be expressed as $[\mathsf{next}]_i\beta^p(e)$, a formula of sort Id_i for the functor T from definition 4 that is, the syntactic operator B_i^p has been factored in two parts. This allows us to have more expressive power and describe points of coalgebras that are not of the form $\Delta S(X)$ for some measure polynomial functor S.

4.4 Types and Non-wellfounded sets

4.4.1 Lismont

In the paper (Lismont, 1992), Lismont claims to use the set theory $\mathbf{ZFC}^-\mathbf{AFA}$ to prove the existence of a universal beliefs space in the sense of Mertens and Zamir. He proposes the functor Λ for a *class* C, as the class of all probability measures on sets $c \subseteq C$ (which therefore need to be measurables spaces). Thus, any subset of C needs to be a measurable space. This does not seem to be a very reasonable assumption. Consider a non-measurable subset s of a measurable space c. Since $s \subseteq c \subseteq C$, $\Lambda s \subseteq \Lambda C$, but the inclusion cannot possibly be a measurable function.

With this definition, Λ is a monotone functor, an assumption that simplifies some results, but $\Lambda[0,1]$, for example, is much bigger than $\Delta[0,1]$. It includes all probability measures over all the σ-algebras that are possible over all the subsets of $[0,1]$ (ΔX was defined over a measurable space (X, Σ) as (the measurable space of) all probability measures on the algebra Σ).

So this operator Λ is not so much of an operator in Meas, as one in Set.

Closely following Mertens and Zamir's paper (Mertens and Zamir, 1984), the goal is to "construct" Y and T such that $Y = K \times T^n$ and $T = \Lambda(K \times T^{n-1})$.

If K is a non empty set, $\Lambda_i(C)$ is the class of all $\lambda \in \Lambda(K \times C^n)$ such that a support of λ is a subset of $K \times C^{i-1} \times \{\lambda\} \times C^{n-i}$ (i-coherent probabilities).

Then

$$\Lambda_*(C) = K \times \prod_{1 \le i \le n} \Lambda_i(C)$$

It's easy to verify that the operator Λ_* is set continuous as defined in (Aczel, 1988). Therefore, there a biggest fixed point Ω, and letting $\Theta_i = \Lambda_i(\Omega)$, we get

$$\Omega = K \times \prod_{1 \le i \le n} \Theta_i.$$

Ω is non empty if K is non empty (this result uses the solution Lemma from the theory of non-wellfounded sets).

Notice that now the equalities are actual identities in AFA.

A *beliefs space* is , as in (Mertens and Zamir, 1984), a measurable subset $B \subseteq Y$ so that for each $v = \langle k, t_1, \ldots, t_n \rangle, h(t_i)(B) = 1$.

Then, a language Φ is defined, with propositional variables obtained from the measurable subsets of B, boolean connectives, and a modality $B_{i,r}\varphi$ for every $\varphi \in \Phi$ meaning that player i assigns probability bigger or equal than r to φ.The language induces an equivalence relation $u \equiv v$ iff for all $\varphi \in \Phi, u \in [\![\varphi]\!] \Leftrightarrow v \in [\![\varphi]\!]$. The main theorem is:

Theorem 4.2. *For all belief spaces B there is an application $* : B \to \Omega, u \mapsto u_*$ so that for all $u, v \in B, u_* = v_* \Leftrightarrow u \equiv v$.*

Lismont admits that even though Y can be embedded in Ω, this doesn't mean that Mertens and Zamir construction follows as an special case of the fixed point result in AFA.

4.4.2 Heifetz

In (Heifetz, 1996), Heifetz proposes a similar idea. Every beliefs space B can be mapped (onto) it's non-wellfounded version, constructed again from setting up the corresponding equations. In the non-wellfounded version of the space, the homeomorphisms become equalities.

The setting of the paper is compact metric spaces, so it can use the result from Mertens and Zamir. Let Y be the universal beliefs space from Mertens and Zamir's work. Given any beleifs space B, a map H from B to Y is defined. Then the image $H(B)$ is proved to be in a 1-to-1 correspondence with $\bar{B}$, the nwf-version of B (similar to B_* in Lismont's work). This way, $\bar{B}$ can borrow the topology from Y, instead of having it generated by the formulas in the language as in Lismont. Heifetz would come back to that idea in (Heifetz and Samet, 1998).

The correspondence above is obtained by showing that $\overline{H(\omega)}$ is also a solution for the equation that yields $\bar{\omega}$ for any $\omega \in B$. These solutions are unique under the theory of non-wellfounded sets.

The probability measures are characterized as a family of pairs $\langle E, r \rangle$ such that E is a measurable set and the measure of E is bigger than or equal to r. An example is given where the inadequacy of dealing with equality instead is shown.

4.5 Other related work

Among other work related to type spaces, we'd like to mention some in particular.

Spyros Vassilakis, in (Vassilakis, 1991), identifies the final sequence method as the right one to obtain a solution for $X = \Delta(S \times X)$ in the category of Compact Hausdorff spaces. He also suggests further applications in (Vassilakis, 1990).

Brandenburger and Dekel in (Brandenburger and Dekel, 1993) propose a similar construction to that of (Mertens and Zamir, 1984), and explore the relation of the concept of types with the one of common knowledge.

Probabilistic logic applied to type spaces has been studied by Heifetz and Mongin in (Heifetz and Mongin, 2001), and Meier in (Meier, 2001). Meier also explored the simpler case of type spaces when the probabilities are given by *finitely additive* measures in (Meier, 2002).

This being just a cursory overview of the literature on this topic, it shows the interest in the problem, and also suggest directions for further development in both the applications and the general theory of coalgebras presented in section 2.

REFERENCES

Aczel P. *Non-Well-Founded Sets*, volume 14 of *CSLI Lecture Notes*. CSLI Publications, Stanford, 1988.

Armbruster W. and Böge W. Bayesian game theory. In *Game theory and related topics (Proc. Sem., Bonn and Hagen, 1978)*, pages 17–28. North-Holland, Amsterdam, 1979.

Aumann R.J. Borel structures for function spaces. *Illinois Journal of Mathematics*, 5:614–630, 1961.

Aumann R.J. Mixed and behavior strategies in infinite games. In *Advances in Game Theory*, volume 52 of *Annals of Mathematical Studies*. Princeton University Press, 1964.

Barwise J. and Moss L. *Vicious Circles*, volume 60 of *CSLI Lecture Notes*. CSLI Publications, Stanford, 1996.

Böge W. and Eisele T. On solutions of Bayesian games. *Internat. J. Game Theory*, 8(4):193–215, 1979. ISSN 0020-7276.

Brandenburger A. and Dekel E. Hierarchies of beliefs and common knowledge. *J. Econom. Theory*, 59(1):189–198, 1993. ISSN 0022-0531.

Dudley R.M. *Real analysis and probability*. The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1989. ISBN 0-534-10050-3.

Harsanyi J.C. Games with incomplete information played by “Bayesian” players. I. The basic model. *Management Sci.*, 14:159–182, 1967.

Harsanyi J.C. Games with incomplete information played by “Bayesian” players. II. Bayesian equilibrium points. *Management Sci.*, 14:320–334, 1968a.

Harsanyi J.C. Games with incomplete information played by “Bayesian” players. III. The basic probability distribution of the game. *Management Sci.*, 14:486–502, 1968b.

Heifetz A. The Bayesian formulation of incomplete information- the non-compact case. *International Journal of Game Theory*, 21:329–338, 1993.

Heifetz A. Non-well-founded type spaces. *Games and Economic Behavior*, 16(2):202–217, 1996.

Heifetz A. and Mongin P. Probability logic for type spaces. *Games and Economic Behavior*, 35:31–53, 2001.

Heifetz A. and Samet D. Topology-free typology of beliefs. *Journal of Economic Theory*, 82(2):324–341, 1998.

Lambek J. A fixpoint theorem for complete categories. *Math. Z.*, 103:151–161, 1968. ISSN 0025-5874.

Lismont L. Une classe universelle anti-fondée pour les croyances. *C. R. Acad. Sci. Paris Sr. I Math.*, 315(11):1117–1120, 1992.

Meier M. An infinitary probability logic for type spaces. Technical report, CORE, Universitè Catholique de Louvain, Belgium, 2001.

Meier M. Finitely additive beliefs and universal type spaces. Technical report, Center for operations research and econometrics, Université catholique de Louvain, 2002.

Mertens J.F. and Zamir S. Formulation of Bayesian analysis for games of incomplete information. *International Journal of Game Theory*, 14:1–29, 1984.

Mètivier M. Limites projectives de mesures. martingales. applications. *Ann. Mat. Pura Appl. (4)*, 63:225–352, 1963.

Moss L.S. and Viglizzo I.D. Final coalgebras for functors on measurable spaces. *Information and Computation*, 204, 2006.

Osborne M.J. and Rubinstein A. *A course in game theory*. MIT Press, Cambridge, MA, 1994. ISBN 0-262-65040-1.

Vassilakis S. Rules for changing the rules. Technical report, Stanford - Institute for Thoretical Economics, 1990.

Vassilakis S. Functorial fixed points: a non-technical introduction. Working Paper 266, Department of Economics, University of Pittsburg, 1991.

Viglizzo I.D. *Coalgebras on Measurable Spaces*. Ph.D. thesis, Indiana University, 2005a.

Viglizzo I.D. Final sequences and final colagebras for measurable spaces. In J.L. Fiadeiro, N. Harman, M. Roggenbach, and J. Rutten, editors, *Algebra and Coalgebra in Computer Science: First International Conference, CALCO 2005, Swansea, UK, September 3-6, 2005. Proceedings*, volume 3629, pages 395–407. Springer-Verlag GmbH, 2005b.