Math. Struct. in Comp. Science (2005), *vol.* 15, *pp.* 647–670. © 2005 Cambridge University Press
doi:10.1017/S096012950500472X Printed in the United Kingdom

Semi-pullbacks for stochastic relations over analytic spaces[†]

ERNST-ERICH DOBERKAT[‡]

Chair for Software Technology, University of Dortmund
Email: ernst-erich.doberkat@udo.edu

Received 22 January 2004; revised 6 September 2004

The existence of semi-pullbacks for stochastic relations over analytic spaces is addressed. It is shown by means of a measure extension that the theory of measurable selectors for measurable relations may be employed for its solution. The category of stochastic relations is shown not to have (weak) pullbacks.

1. Introduction

A category is said to have semi-pullbacks if, whenever $f : a \to b, g : c \to b$ is a pair of morphisms with the same target, there exists in the category an object d together with morphisms $r : d \to a, s : d \to c$ such that $f \circ r = g \circ s$. The existence of semi-pullbacks makes sure that bisimulations, which are defined as spans of morphisms, are transitive. Semi-pullbacks are helpful, for example, with stochastic relations when investigating the relationship between modal logics and bisimilar stochastic Kripke models. The argument goes like this: given two stochastic Kripke models $\mathscr{K}_1$ and $\mathscr{K}_2$ that accept exactly the same formulas, we can construct a third Kripke model $\mathscr{L}$ and morphisms Φ_1 and Φ_2 that form a co-span

$$\mathscr{K}_1 \xrightarrow{\Phi_1} \mathscr{L} \xleftarrow{\Phi_2} \mathscr{K}_2.$$

If the category under consideration has semi-pullbacks, the upper left-hand corner of the diagram below may be completed

$$\begin{array}{ccc} \mathscr{M} & \xrightarrow{\Psi_1} & \mathscr{K}_1 \\ {\scriptstyle \Psi_2}\downarrow & & \downarrow{\scriptstyle \Phi_1} \\ \mathscr{K}_2 & \xrightarrow[\Phi_2]{} & \mathscr{L} \end{array}$$

From this a bisimulation is constructed. The crucial point is not the construction of $\mathscr{L}$ (which, nevertheless, is involved in its own right) but rather the existence of the semi-pullback, that is, of the object $\mathscr{M}$ together with the morphisms Ψ_1 and Ψ_2.

[†] Preliminary results were presented at ICALP'03 at Eindhoven, see Doberkat (2003)

[‡] This research was funded in part by *Deutsche Forschungsgemeinschaft*, grant DO 263/8-1, *Algebraische Eigenschaften stochastischer Relationen*

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

This problem has been addressed in a variety of ways, with appropriate choice of the base category:

1 Edalat (Edalat 1999) considers a category of Markov processes such that
   — the state space is an analytic space,
   — the transition probability function is universally measurable.

   The solution to this instance of the problem is essentially based on an explicit construction using regular conditional probabilities, which are available because of the Polish descent of analytic spaces.
   The solution to this problem allows us to tackle the equivalence of bisimilarity and the acceptance of the same formulas for a whole family of modal logics with a countable number of diamonds, all interpreted over analytic spaces with a labelled Markov transition system based on universally measurable transition functions, see Deharnais *et al.* (2002).
2 Markov processes over Polish spaces in which the transition probability function is Borel measurable are considered in Doberkat (2003). The problem was transformed into finding a measurable selector for a set-valued map.
   Based on this solution, it could be shown that bisimilarity and accepting the same set of formulas are equivalent for a simple negation free modal logic with a countable number of diamonds when the labelled Markov transition processes are Borel functions based on Polish spaces, provided one of the processes is small, that is, has a Borel section.

The present paper refines the technique employed in Doberkat (2003) in order to show that semi-pullbacks exist in categories of Markov processes over analytic spaces when the transition probability functions are Borel measurable (rather than universally measurable). For practical purposes, the difference between universal and Borel measurability is negligible, but for structural purposes, it is not. Borel measurability is defined in terms of the inverse image of Borel sets (in exactly the same way as continuity is defined in terms of the inverse image of the open sets, or in the same way as uniform continuity is defined in terms of the inverse image of neighbourhoods), but universal measurability also requires the concept of completing the Borel sets through all finite measures, and thus requires the additional concept of a finite measure. Hence, Borel measurability is conceptually simpler and more basic.

Applications

Once the general availability of semi-pullbacks is ensured, a Hennessy–Milner Theorem on the equivalence of accepting the same formulas and bisimilarity can be established for a general class of modal logics. It becomes easier to investigate the bisimilarity of stochastic relations: for example, it can be shown that two stochastic relations are bisimilar provided they have isomorphic factors; the converse holds under the assumption of compactness as well, so one might speculate whether bisimilarity and having isomorphic factors are equivalent, see (Doberkat 2004b) for a more comprehensive discussion. These investigations require us to study congruences for stochastic relations (Doberkat 2004a).

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

Organisation

The problem and an outline of its solution is discussed in greater detail in Section 2, and some preliminaries are collected in Section 3. At the heart of the solution lies a measure extension, which is provided in Section 4, and Section 5 proposes a solution to the existence of semi-pullbacks for stochastic relations over Polish, and then over analytic spaces. Section 6 wraps it all up and provides pointers to applications.

2. The problem and a roadmap to its solution

This section outlines the problem and indicates a roadmap towards its solution. Terms from topology or measure theory are defined in Section 3; the basic constructions can also be found there. Let $\mathsf{K} = (X, Y, K)$ be a stochastic relation over the Polish spaces X and Y, that is, K is a measurable map from X to the subprobabilities $\mathbf{S}(Y)$ on Y. Assume that K is the target of two morphisms

$$\mathsf{K}_1 \xrightarrow{\mathsf{f}_1} \mathsf{K} \xleftarrow{\mathsf{f}_2} \mathsf{K}_2$$

Thus $\mathsf{f}_i = (\phi_i, \psi_i)$ is for $i = 1, 2$ a pair of Borel measurable and surjective maps $\phi_i : X_i \to X$ and $\psi_i : Y_i \to Y$ such that

$$K(\phi_i(x_i))(B) = \mathbf{S}(\psi_i)(K_i(x_i))(B) \left(= K_i(x_i)(\psi_i^{-1}[B])\right)$$

whenever $x_i \in X_i$ and $B \subseteq Y$ is a Borel set. The spaces involved are Polish. Here $\mathbf{S}(\psi_i) : \mathbf{S}(Y_i) \to \mathbf{S}(Y)$ is the map that is induced on the set of probabilities by ψ_i.

An expansion of the definitions then yields the following commutative diagram:

$$\begin{array}{ccccc}
X_1 & \xrightarrow{\phi_1} & X & \xleftarrow{\phi_2} & X_2 \\
\downarrow{\scriptstyle K_1} & & \downarrow{\scriptstyle K} & & \downarrow{\scriptstyle K_2} \\
\mathbf{S}(Y_1) & \xrightarrow[\mathbf{S}(\psi_1)]{} & \mathbf{S}(Y) & \xleftarrow[\mathbf{S}(\psi_2)]{} & \mathbf{S}(Y_2)
\end{array}$$

We are looking for a stochastic relation L and two morphisms

$$\mathsf{K}_1 \xleftarrow{\mathsf{g}_1} \mathsf{L} \xrightarrow{\mathsf{g}_2} \mathsf{K}_2$$

such that $\mathsf{f}_1 \circ \mathsf{g}_1 = \mathsf{f}_2 \circ \mathsf{g}_2$ holds. Written as diagrams, this means that these diagrams should commute:

$$\begin{array}{ccccc}
X_1 & \xleftarrow{\alpha_1} & A & \xrightarrow{\alpha_2} & X_2 \\
\downarrow{\scriptstyle K_1} & & \downarrow{\scriptstyle L} & & \downarrow{\scriptstyle K_2} \\
\mathbf{S}(Y_1) & \xleftarrow[\mathbf{S}(\beta_1)]{} & \mathbf{S}(B) & \xrightarrow[\mathbf{S}(\beta_2)]{} & \mathbf{S}(Y_2)
\end{array}$$

Then $\mathsf{L} = (A, B, L)$ is the relation involved, and $\mathsf{g}_i = (\alpha_i, \beta_i)$ is the morphism $\mathsf{g}_i : \mathsf{L} \to \mathsf{K}_i$ for $i = 1, 2$.

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

We will define

$$A := \{\langle x_1, x_2\rangle \in X_1 \times X_2 \mid \phi_1(x_1) = \phi_2(x_2)\}$$
$$B := \{\langle y_1, y_2\rangle \in Y_1 \times Y_2 \mid \psi_1(y_1) = \psi_2(y_2)\},$$

and then A and B will be shown to be Polish. Taking α_i, β_i as the projections

$$\alpha_i : A \ni \langle x_1, x_2\rangle \mapsto x_i \in X_I$$
$$\beta_i : B \ni \langle y_1, y_2\rangle \mapsto y_i \in Y_i,$$

we need to find a stochastic relation $L : A \rightsquigarrow B$ that makes $\mathsf{g}_i := (\alpha_i, \beta_i)$ into a morphism $\mathsf{g}_i : \mathsf{L} \to \mathsf{K}_i$ for $i = 1, 2$. Thus L has to satisfy the following constraints for all $\langle x_1, x_2\rangle \in A$:

1 $L(x_1, x_2) \in \mathbf{S}(B)$,
2 $\mathbf{S}(\beta_1)(L(x_1, x_2)) = K_1(x_1)$,
3 $\mathbf{S}(\beta_2)(L(x_1, x_2)) = K_2(x_2)$.

Reformulating again, we put

$$\Gamma(x_1, x_2) := \{\mu \in \mathbf{S}(B) \mid \mathbf{S}(\beta_1)(\mu) = K_1(x_1) \text{ and } \mathbf{S}(\beta_2)(\mu) = K_2(x_2)\},$$

and hence $L(x_1, x_2) \in \Gamma(x_1, x_2)$ for all $\langle x_1, x_2\rangle \in A$. We want $L : A \to \mathbf{S}(B)$ to be a measurable selector for Γ, thus $L : A \rightsquigarrow B$ is a stochastic relation (accounting for the *measurable* in measurable selector) such that $\forall a \in A : L(a) \in \Gamma(a)$ (accounting for the *selector* in measurable selector).

Thus the problem is massaged into finding a measurable selector for the set-valued map Γ. The existence of such a selector can be asserted under the following conditions:

1 $\Gamma(x_1, x_2)$ is a closed subset of $\mathbf{S}(B)$ for each $\langle x_1, x_2\rangle \in A$;
2 the set $\{\langle x_1, x_2\rangle \in A \mid \Gamma(x_1, x_2) \cap C \neq \varnothing\}$ is a measurable subset of A, whenever $C \subseteq \mathbf{S}(B)$ is compact;
3 $\Gamma(x_1, x_2) \neq \varnothing$ for each and any $\langle x_1, x_2\rangle \in A$.

These properties will characterise Γ as a $\mathscr{C}$ – *measurable* relation in Section 3.

It will not be difficult to show that property 1 is satisfied, and property 2 will also easily seen to be fulfilled. This will happen through the use of some helpful properties of Borel sets in Polish spaces. The really crucial property is the third one.

The crucial point

Fix $\langle x_1, x_2\rangle \in A$. We will find without much trouble a measure $\mu_0 \in \mathbf{S}(B)$ such that the equations

$$\mathbf{S}(\beta_1)(\mu_0)(E_1) = K_1(x_1)(E_1)$$
$$\mathbf{S}(\beta_2)(\mu_0)(E_2) = K_2(x_2)(E_2)$$

are true whenever $E_i = \psi_i^{-1}[F_i]$ for some Borel set $F_i \subseteq Y$.

This is equivalent to saying that the measures $\mathbf{S}(\beta_i)(\mu_0)$ and $K_i(x_i)$ coincide on the σ-algebra $\psi_i^{-1}[\mathscr{B}(Y)]$. The latter is usually a proper sub-σ-algebra of $\mathscr{B}(Y_i)$, hence we

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

cannot immediately guarantee equality on the full Borel sets $\mathscr{B}(Y_i)$ with $i = 1, 2$. If, however, we can find a measure, say $\mu_1 \in \mathbf{S}(B)$, such that $\mathbf{S}(\beta_i)(\mu_1)(E_i) = K_i(x_i)(E_i)$ for all E_i ranging over all of $\mathscr{B}(Y_i)$, we may conclude $\mu_1 \in \Gamma(x_1, x_2)$, and then the latter set is in fact shown to be non-empty.

Thus the problem is reduced to finding a single measure on B that has suitable marginal distributions. We will demonstrate that we can extend the marginal distributions of μ_0 to a measure with the desired properties.

This measure extension can be done in two different ways, depending on the nature of the space Y, which otherwise sits quietly in the background and serves patiently as a target space.

1 If Y is Polish, we can and do refer to a special case of Edalat's result, and we are done. This precludes a more general solution that includes analytic spaces.
2 If Y is more generally a separable metric space, an extension can still be found; this assumption will be crucial when the analytic case is targeted. The machinery from Mathematical Analysis for tackling this general case is well established, but somewhat heavy, including such gems as the Riesz Representation Theorem and the Hahn–Banach Theorem.

We propose to take the latter approach here because it opens up the avenue of establishing this result for analytic spaces too.

Discussion

Edalat's solution can be considered constructive: it works with conditional distributions and their properties. These distributions do exist, essentially because of the Radon–Nikodym Theorem, which in turn leans heavily on the way the Lebesgue–Daniell integral is constructed. The basic idea for the more general solution lies in a measure extension process. It works in three steps: the first is to consider the integral of the given measure, which is a linear functional; then to formulate the extension problem in terms of extending this linear functional to a linear functional on a broader domain that corresponds to our needs; and, finally, to represent the extended functional as an integral again. Representing the measure as an integral is easy, since only the standard techniques of integration are involved, but making the extension is complicated and requires the Hahn–Banach Theorem, and converting the linear functional back to a measure is again a bit involved and can be done through the Riesz Representation Theorem. It is this last step that gives us the measure we are looking for.

Through the use of the Hahn–Banach Theorem, we make use of Zorn's Lemma, which in turn is known to be equivalent to the Axiom of Choice. Thus we propose a stronger solution, but we pay more for it. This probably sounds a bit more dramatic and unusual than it really is: when constructing the product of an arbitrary family of sets or when looking into the ultrafilter extension of a Kripke model we silently take the Axiom of Choice for granted.

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

The road map

We will first delve into the problem of extending a measure with given marginal distributions; this is done in Section 4. With the preparations given in Section 3, where a polished array of statements and techniques from topology and probability is collected or provided, we will be in good condition for tackling the problem in Section 5 along the lines sketched here.

3. Polish spaces, measures and all that

This section collects some basic facts from topology and measure theory for convenience and later reference.

A *Polish space* $(X, \mathscr{T})$ is a topological space that is second countable, that is, has a countable dense subset, and is metrisable through a complete metric. A measurable space $(X, \mathscr{A})$ is a set X with a σ-algebra $\mathscr{A}$. The *Borel sets* $\mathscr{B}(X, \mathscr{T})$ for the topology $\mathscr{T}$ is the smallest σ-algebra on X that contains $\mathscr{T}$. A *Standard Borel* space $(X, \mathscr{A})$ is a measurable space such that the σ-algebra $\mathscr{A}$ equals $\mathscr{B}(X, \mathscr{T})$ for some Polish topology $\mathscr{T}$ on X. Although the Borel sets are determined uniquely through the topology, the converse does not hold, as we will see shortly. Given two measurable spaces $(X, \mathscr{A})$ and $(Y, \mathscr{B})$, a map $f : X \to Y$ is *$\mathscr{A} - \mathscr{B}$-measurable* whenever

$$f^{-1}[\mathscr{B}] \subseteq \mathscr{A},$$

where

$$f^{-1}[\mathscr{B}] := \{f^{-1}[B] \mid B \in \mathscr{B}\}$$

is the set of inverse images

$$f^{-1}[B] := \{x \in X \mid f(x) \in B\}$$

of elements of $\mathscr{B}$. Note that $f^{-1}[\mathscr{B}]$ is in any case a σ-algebra; it will be helpful to note that it is the smallest σ-algebra on X that contains all the sets $f^{-1}[B]$ for $B \in \mathscr{B}$. If the σ-algebras are the Borel sets of some topologies on X and Y, respectively, then a measurable map is called *Borel measurable* or simply a *Borel* map. The real numbers $\mathbb{R}$ always carry the Borel structure induced by the usual topology, which will usually not be mentioned explicitly when we talk about Borel maps.

The product $(X_1 \times X_2, \mathscr{A}_1 \otimes \mathscr{A}_2)$ of two measurable spaces $(X_1, \mathscr{A}_1)$ and $(X_2, \mathscr{A}_2)$ is the Cartesian product $X_1 \times X_2$ endowed with the σ-algebra

$$\mathscr{A}_1 \otimes \mathscr{A}_2 := \sigma\left(\{A_1 \times A_2 \mid A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_2\}\right).$$

This is the smallest σ-algebra that contains all the measurable rectangles $A_1 \times A_2$, and is, incidentally, the smallest σ-algebra $\mathscr{E}$ on $X_1 \times X_2$ that makes the projections

$$\pi_i : X_1 \times X_2 \to X_i$$

$\mathscr{E} - \mathscr{A}_i$-measurable for $i = 1, 2$.

An analytic set $X \subseteq Z$ in a Polish space Z is the image $f[Y]$ of a Polish space Y for some Borel measurable map $f : Y \to Z$. Endow X with the trace $\mathscr{A}$ of $\mathscr{B}(Z)$ on X, that

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

is,

$$\mathscr{A} = \mathscr{B}(Z) \cap X := \{B \cap X \mid B \in \mathscr{B}(Z)\}.$$

We also call the elements of $\mathscr{B}(Z) \cap X$ the Borel sets $\mathscr{B}(X)$ of X. A measurable space that is Borel isomorphic to $(X, \mathscr{A})$ is called an *analytic space* (a Borel isomorphism is a Borel measurable and bijective map the inverse of which is also Borel measurable). Note, by the way, that $\mathscr{B}(X) = \{B \subseteq X \mid B \in \mathscr{B}(Z)\}$ if $X \subseteq Z$ is a Borel set itself (this is a slight abuse of notation).

A map $f : X \to Y$ between the topological spaces $(X, \mathscr{T})$ and $(Y, \mathscr{S})$ is $\mathscr{T} - \mathscr{S}$-*continuous* iff the inverse image of an open set from $\mathscr{S}$ is an open set in $\mathscr{T}$. Thus, a continuous map is also measurable with respect to the Borel sets generated by the respective topologies.

When the context is clear, we will write down topological or measurable spaces without their topologies and σ-algebras, respectively, and the Borel sets are always understood with respect to the topology under consideration.

The following statement will be most helpful later in the paper. It implies that, given a measurable map between Polish spaces, we can find a finer Polish topology on the domain, which has the same Borel sets, and renders the map continuous.

Proposition 3.1. Let $(X, \mathscr{T})$ be a Polish space, and $(Y, \mathscr{T}')$ be a second countable metric space. If $f : X \to Y$ is a Borel measurable map, then there exists a Polish topology $\mathscr{T}_0$ on X such that $\mathscr{T}_0$ is finer than $\mathscr{T}$ (hence $\mathscr{T} \subseteq \mathscr{T}_0$), $\mathscr{T}$ and $\mathscr{T}_0$ have the same Borel sets, and f is $\mathscr{T}_0 - \mathscr{T}'$ continuous.

Proof. See Srivastava (1998, Corollaries 3.2.5 and 3.2.6). □

Given two measurable spaces $(X, \mathscr{A})$ and $(Y, \mathscr{B})$, a *stochastic relation* $K : (X, \mathscr{A}) \rightsquigarrow (Y, \mathscr{B})$ is a Borel map from X to the set $\mathbf{S}(Y, \mathscr{B})$, the latter denoting the set of all subprobability measures on $(Y, \mathscr{B})$. The latter set carries the *weak*-σ-algebra*. This is the smallest σ-algebra on $\mathbf{S}(Y, \mathscr{B})$ that renders all maps $\mu \mapsto \mu(D)$ measurable, where $D \in \mathscr{B}$. Hence $K : (X, \mathscr{A}) \rightsquigarrow (Y, \mathscr{B})$ is a stochastic relation iff:

1 $K(x)$ is a subprobability measure on $(Y, \mathscr{B})$ for all $x \in X$,
2 $x \mapsto K(x)(D)$ is a measurable map for each measurable set $D \in \mathscr{B}$.

We will usually deal with stochastic relations between Polish spaces or between analytic spaces. Accordingly, we then call (X, Y, K) a *Polish* or *analytic* object, respectively.

Let Y be a metric space. $\mathbf{S}(Y)$ is usually equipped with the topology of weak convergence. This is the smallest topology on $\mathbf{S}(Y)$ that makes the map $\mu \mapsto \int_Y f \, d\mu$ continuous for each continuous and bounded $f : Y \to \mathbb{R}$. It is well known that for second countable X this topology is also second countable, and that X Polish implies that $\mathbf{S}(X)$ is also Polish (Parthasarathy 1967, Theorems II.6.3 and II.6.5). Moreover, the Borel sets for the topology of weak convergence is just the weak*-σ-algebra (Kechris 1994, Theorem 17.24). If X is a Standard Borel space, $\mathbf{S}(X)$ is one also: select a Polish topology $\mathscr{T}$ on X that induces the measurable structure, then $\mathscr{T}$ will give rise to the Polish topology of weak convergence on $\mathbf{S}(X)$, which in turn has the weak-*-σ-algebra as its Borel sets.

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

An $\mathscr{A} - \mathscr{B}$- measurable map $f : X \to Y$ between the measurable spaces $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ induces a map

$$\mathbf{S}(f) : \mathbf{S}(X, \mathscr{A}) \to \mathbf{S}(Y, \mathscr{B})$$

upon setting $(\mu \in \mathbf{S}(X, \mathscr{A}), D \in \mathscr{B})$

$$\mathbf{S}(f)(\mu)(D) := \mu(f^{-1}[D])$$

It is easy to see that $\mathbf{S}(f)$ is measurable, and we will see in a moment that under some not too restrictive conditions $\mathbf{S}(f) : \mathbf{S}(X, \mathscr{A}) \to \mathbf{S}(Y, \mathscr{B})$ is onto provided $f : X \to Y$ is. If f is continuous, $\mathbf{S}(f)$ is too. We use $\mathbf{P}(X, \mathscr{A})$ to denote the subspace of all probability measures on $(X, \mathscr{A})$.

The category $\mathfrak{Stoch}$ has as objects stochastic relations $\mathsf{K} = (X, Y, K)$ for measurable spaces X, Y and $K : X \rightsquigarrow Y$. A *morphism* $\mathsf{f} : \mathsf{K}_1 \to \mathsf{K}_2$ with $\mathsf{f} = (\phi, \psi)$ between the objects $\mathsf{K}_i := (X_i, Y_i, K_i)$ is comprised of a pair of surjective measurable maps $\phi : X_1 \to X_2$ and $\psi : Y_1 \to Y_2$ such that

$$K_2 \circ \phi = \mathbf{S}(\psi) \circ K_1,$$

that is, the diagram

$$\begin{array}{ccc} X_1 & \xrightarrow{\phi} & X_2 \\ {\scriptstyle K_1}\downarrow & & \downarrow{\scriptstyle K_2} \\ \mathbf{S}(Y_1) & \xrightarrow[\mathbf{S}(\psi)]{} & \mathbf{S}(Y_2) \end{array}$$

is commutative.

We will usually investigate morphisms for stochastic relations based on Polish or analytic spaces, but it does not hurt to define them generally. Recall that an object (X, Y, K) of $\mathfrak{Stoch}$ is *Polish* or *analytic* iff both X and Y are Polish or analytic spaces, respectively. Accordingly, we use $\mathfrak{PStoch}$ and $\mathfrak{AStoch}$ to denote the full subcategories having Polish and analytic objects, respectively, as their objects.

Let $\mathbb{F}(X)$ be the set of all closed and non-empty subsets of the Polish space X. For Polish Y, we say a relation, that is, a set-valued map $F : X \to \mathbb{F}(Y)$, is *$\mathscr{C}$-measurable* iff for a compact set $C \subseteq Y$ the weak inverse

$$\exists F(C) := \{x \in X \mid F(x) \cap C \neq \varnothing\}$$

is measurable. A *selector* s for such a relation F is a single-valued map $s : X \to Y$ such that $s(x) \in F(x)$ holds for each $x \in X$. $\mathscr{C}$-measurable relations have Borel selectors.

Proposition 3.2. Let X and Y be Polish spaces. Then each $\mathscr{C}$-measurable relation F has a measurable selector.

Proof. Since closed subsets of Polish spaces are complete, the assertion follows from Wagner (1977, Theorem 4.2 (e)). □

Postulating measurability for $\exists F(C)$ for open or for closed sets C leads to the general notion of a measurable relation. These relations are a valuable tool in such diverse fields

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

as stochastic dynamic programming (Wagner 1977) and descriptive set theory (Kechris 1994). Overviews are provided in Srivastava (1998, Chapter 5), Himmelberg (1975) and Wagner (1977).

We will need surjectivity of $\mathbf{S}(f)$, when $f : X \to Y$ is measurable and onto. When we apply a selection argument we need to be sure that the set-valued map we are working with does indeed take non-empty values. For these, and other purposes, the concept of universal measurability is needed. Let $\mu \in \mathbf{S}(X, \mathscr{A})$ be a subprobability on the measurable space $(X, \mathscr{A})$. Then $A \subseteq X$ is called *μ-measurable* iff there exist $M_1, M_2 \in \mathscr{A}$ with $M_1 \subseteq A \subseteq M_2$ and $\mu(M_1) = \mu(M_2)$. The μ-measurable subsets of X form a σ-algebra $\mathscr{M}_\mu(\mathscr{A})$. The σ-algebra $\mathscr{U}(\mathscr{A})$ of universally measurable sets is defined by

$$\mathscr{U}(\mathscr{A}) := \bigcap \{\mathscr{M}_\mu(\mathscr{A}) \mid \mu \in \mathbf{S}(X, \mathscr{A})\}$$

(in fact, one usually considers all finite or σ-finite measures, but these definitions lead to the same universally measurable sets). If $f : X_1 \to X_2$ is an $\mathscr{A}_1$-$\mathscr{A}_2$-measurable map between the measurable spaces $(X_1, \mathscr{A}_1)$ and $(X_2, \mathscr{A}_2)$, it is well known that f is also $\mathscr{U}(\mathscr{A}_1)$-$\mathscr{U}(\mathscr{A}_2)$-measurable; the converse does not hold, and one usually cannot conclude that a map $g : X_1 \to X_2$ that is $\mathscr{U}(\mathscr{A}_1)$-$\mathscr{A}_2$-measurable is also $\mathscr{A}_1$-$\mathscr{A}_2$-measurable.

Lemma 3.1. Let X be a Polish and Y a second countable metric space. If $f : X \to Y$ is a surjective Borel map, so is $\mathbf{S}(f) : \mathbf{S}(X) \to \mathbf{S}(Y)$.

Proof.

1 From Arveson (1976, Theorem 3.4.3), we find a map $g : Y \to X$ such that $f \circ g = id_Y$ and g is $\mathscr{U}(\mathscr{B}(Y)) - \mathscr{U}(\mathscr{B}(X))$-measurable.

2 Let $\nu \in \mathbf{S}(Y)$, and define $\mu := \mathbf{S}(g)(\nu)$. Then $\mu \in \mathbf{S}(X, \mathscr{U}(\mathscr{B}(X)))$ by construction. Restrict μ to the Borel sets on X, obtaining $\mu_0 \in \mathbf{S}(X, \mathscr{B}(X))$. Since we have for each set $B \subseteq Y$ the equality $g^{-1}\left[f^{-1}[B]\right] = B$, we see that for each $B \in \mathscr{B}(Y)$

$$\mathbf{S}(f)(\mu_0)(B) = \mu_0(f^{-1}[B]) = \mu(f^{-1}[B]) = \nu(g^{-1}\left[f^{-1}[B]\right]) = \nu(B)$$

holds. □

We say a measurable space $(X, \mathscr{A})$ is *separable* iff the σ-algebra $\mathscr{A}$ has a countable set $(A_n)_{n \in \mathbb{N}}$ of generators that separates points, that is, given $x, x' \in X$ with $x \neq x'$, there exists A_n that contains exactly one of them. A Polish space is separable as a measurable space, as is an analytic space.

Second countable metric spaces are sometimes called separable, and this analogy is justified by the following lemma.

Lemma 3.2. For a separable measurable space $(X, \mathscr{A})$ there exists a second countable metric topology $\mathscr{T}$ on X such that $\mathscr{B}(X, \mathscr{T}) = \mathscr{A}$.

Proof. The proof follows from Srivastava (1998, Proposition 3.3.2 and Remark 3.3.3). □

This innocent looking statement has some remarkable consequences for our context, which we will see in due course. Just for starters, we have the following corollary.

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

Corollary 3.1. Let $(Y, \mathscr{B})$ be a separable measurable space. Then:

1 The diagonal is measurable in the product, that is,

$$\Delta_{Y\times Y} := \{\langle y, y\rangle \mid y \in Y\} \in \mathscr{B} \otimes \mathscr{B}.$$

2 If $f_i : X_i \to Y$ is $\mathscr{A}_i - \mathscr{B}$- measurable, where $(X_i, \mathscr{A}_i)$ is a measurable space $(i = 1, 2)$, then

$$f_1^{-1}[\mathscr{B}] \otimes f_2^{-1}[\mathscr{B}] = (f_1 \times f_2)^{-1}[\mathscr{B} \otimes \mathscr{B}].$$

3 If X is a Polish space and $f : X \to Y$ is $\mathscr{B}(X) - \mathscr{B}$—measurable and onto, then $\mathbf{S}(f) : \mathbf{S}(X, \mathscr{B}(X)) \to \mathbf{S}(Y, \mathscr{B})$ is surjective also.

Proof. The proofs are based on standard arguments; they are given for the reader's convenience. Note that Assertion 3 follows from Lemma 3.1.

1 It is well known (Parthasarathy 1967, Theorem I.3.3) that property 1 holds for second countable metric spaces, so by Lemma 3.2 it holds for separable metric spaces as well.

2 The product σ-algebra $\mathscr{B} \otimes \mathscr{B}$ is generated by the rectangles $B_1 \times B_2$ with B_i taken from some generator $\mathscr{B}_0$ for $\mathscr{B}$ $(i = 1, 2)$. Since

$$(f_1 \times f_2)^{-1}[B_1 \times B_2] = f_1^{-1}[B_1] \times f_2^{-1}[B_2],$$

we see that

$$(f_1 \times f_2)^{-1}[\mathscr{B} \otimes \mathscr{B}] \subseteq f_1^{-1}[\mathscr{B}] \otimes f_2^{-1}[\mathscr{B}].$$

This is true without the assumption of separability. Now let $\mathscr{T}$ be a second countable metric topology on Y with $\mathscr{B} = \mathscr{B}(Y, \mathscr{T})$ and let $\mathscr{T}_0$ be a countable base for the topology. Then

$$\mathscr{T}_p := \{T_1 \times T_2 \mid T_1, T_2 \in \mathscr{T}_0\}$$

is a countable base for the product topology $\mathscr{T} \otimes \mathscr{T}$, and (this is the crucial property)

$$\mathscr{B} \otimes \mathscr{B} = \mathscr{B}(Y \times Y, \mathscr{T} \otimes \mathscr{T})$$

holds (Srivastava 1998, Propositio 3.1.23). Since, it is clear that for $T_1, T_2 \in \mathscr{T}_0$

$$f_1^{-1}[T_1] \times f_2^{-1}[T_2] \in (f_1 \times f_2)^{-1}\left[\mathscr{T}_p\right] \subseteq (f_1 \times f_2)^{-1}[\mathscr{B} \otimes \mathscr{B}],$$

the non-trivial inclusion is inferred from the fact that the smallest σ-algebra containing

$$\{f_1^{-1}[T_1] \times f_2^{-1}[T_2] \mid T_1, T_2 \in \mathscr{T}_0\}$$

equals $f_1^{-1}[\mathscr{B}] \otimes f_2^{-1}[\mathscr{B}]$. □

We will need to make precise statements regarding the measurability of a Borel map; for ease of reference, we record the following technical statement.

Proposition 3.3. Let X be a Polish space and $(Y, \mathscr{B})$ be a separable measurable space. Assume that $g : X \to Y$ is $\mathscr{B}(X) - \mathscr{B}$-measurable and onto. If $f : X \to Y$ is $\mathscr{B}(X) - \mathscr{B}$-measurable such that f is constant on the atoms of $g^{-1}[\mathscr{B}]$, then f is $g^{-1}[\mathscr{B}] - \mathscr{B}$-measurable.

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

Proof. Separability implies that $\{y\} \in \mathscr{B}$ for all $y \in Y$. The atoms of $g^{-1}[\mathscr{B}]$ are just the inverse images $g^{-1}[\{y\}]$ of the points $y \in Y$, because these sets are clearly atomic in that σ-algebra, and form a partition of X. Now let $B \in \mathscr{B}$ be a measurable set. Then, by assumption, $f^{-1}[B]$ is a Borel set in X that is the union of atoms of $g^{-1}[\mathscr{B}]$. Thus the assertion follows from the Blackwell–Mackey Theorem (Srivastava 1998, Theorem 4.5.7). □

4. Extending semi-pullbacks of measures

The main argument in establishing the existence of a semi-pullback in the category of stochastic relations will be a selection argument: we will show that a certain set-valued map will have a (measurable) selector. This will require that this map always takes non-empty values. This section will be devoted to establishing a property of semi-pullbacks for measure spaces, which will, in turn, be crucial in proving non-emptiness. Since it is rather technical in nature, it is convenient to encapsulate this development in a separate section.

We will consider the category $\mathfrak{P}$ of probability spaces that has as objects tuples $(X, \mathscr{A}, \mu)$ with $\mu \in \mathbf{P}(X, \mathscr{A})$ for the measurable space $(X, \mathscr{A})$. We have $\psi : (X, \mathscr{A}, \mu) \to (Y, \mathscr{B}, \nu)$ is a morphism in $\mathfrak{P}$ if $\psi : X \to Y$ is a surjective and $\mathscr{A} - \mathscr{B}$-measurable map that is measure preserving, that is, $\nu = \mathbf{P}(\psi)(\mu)$ holds. For two objects $(X, \mathscr{A}, \mu)$ and $(Y, \mathscr{B}, \nu)$, we have that $\mathfrak{P}$ contains their product $(X \times Y, \mathscr{A} \otimes \mathscr{B}, \mu \otimes \nu)$, with $\mu \otimes \nu$ as the product measure, which is uniquely determined through

$$(\mu \otimes \nu)(A \times B) = \mu(A) \cdot \nu(B).$$

For the following discussion, we fix the Polish spaces X_1 and X_2 with the respective Borel sets as σ-algebras. We assume $(Z, \mathscr{C})$ to be a separable measurable space. For a measurable space $(X, \mathscr{A})$, we use $\mathscr{F}(X, \mathscr{A})$ to denote the linear space of all $\mathscr{A}$-measurable and bounded real-valued maps $g : X \to \mathbb{R}$. Note that $\mathscr{A} \subseteq \mathscr{B}$ implies that $\mathscr{F}(X, \mathscr{A})$ is a linear subspace of $\mathscr{F}(X, \mathscr{B})$.

Now let

$$(X_1, \mathscr{B}(X_1), \mu_1) \xrightarrow{\psi_1} (Z, \mathscr{C}, \nu) \xleftarrow{\psi_2} (X_2, \mathscr{B}(X_2), \mu_2)$$

be a pair of morphisms in $\mathfrak{P}$ with a common target, and assume that

$$\begin{array}{ccc} (S, \mathscr{A}, \theta) & \xrightarrow{\pi_2} & (X_2, \mathscr{B}(X_2), \mu_2) \\ \downarrow{\scriptstyle \pi_1} & & \downarrow{\scriptstyle \psi_2} \\ (X_1, \mathscr{B}(X_1), \mu_1) & \xrightarrow[\psi_1]{} & (Z, \mathscr{C}, \nu) \end{array} \qquad (*)$$

is a semi-pullback diagram in $\mathfrak{P}$ with

$$\begin{aligned} S &:= \{\langle x_1, x_2 \rangle \mid \psi_1(x_1) = \psi_2(x_2)\} \in \psi_1^{-1}[\mathscr{C}] \otimes \psi_2^{-1}[\mathscr{C}] \\ \mathscr{A} &:= S \cap \left(\psi_1^{-1}[\mathscr{C}] \otimes \psi_2^{-1}[\mathscr{C}]\right) \\ &= S \cap (\psi_1 \times \psi_2)^{-1}[\mathscr{C} \otimes \mathscr{C}]. \end{aligned}$$

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

The π_i are again the projections. The equality for $\mathscr{A}$ holds by Corollary 3.1; thus $\mathscr{A}$ is the smallest σ-algebra on S that makes

$$\psi_1 \times \psi_2 : \langle x_1, x_2 \rangle \mapsto \langle \psi_1(x_1), \psi_2(x_2) \rangle$$

measurable.

S is a Borel set, and the crucial step in the technical development will consist of 'lifting' this pullback so that the object $(S, \mathscr{B}(S), \mu)$ for some suitable $\mu \in \mathbf{P}(S, \mathscr{B}(S))$ stands in the upper left-hand corner of the diagram. The essential difference is in the σ-algebras on S: starting with the initial σ-algebra with respect to $\psi_1 \times \psi_2$, we claim that we can find a measure μ on the Borel sets of S so that the properties of a semi-pullback will be preserved. This is clearly a problem of extending the probability measure θ.

Proposition 4.1. The semi-pullback $(*)$ in $\mathfrak{P}$ may be extended to a semi-pullback

$$\begin{array}{ccc} (S, \mathscr{B}(S), \mu) & \xrightarrow{\pi_2} & (X_2, \mathscr{B}(X_2), \mu_2) \\ \Big\downarrow{\scriptstyle \pi_1} & & \Big\downarrow{\scriptstyle \psi_2} \\ (X_1, \mathscr{B}(X_1), \mu_1) & \xrightarrow[\psi_1]{} & (Z, \mathscr{C}, \nu) \end{array}$$

in $\mathfrak{P}$.

Essentially, this entails an extension process, extending $\theta \in \mathbf{P}(S, \mathscr{A})$ to a suitable $\mu \in \mathbf{P}(S, \mathscr{B}(S))$. We establish the existence of this extension in two steps: the first will assume that X_1 and X_2 are compact Polish spaces; and the second will show how to reduce the general case to the compact one.

Proof for the compact case. The line of attack will be as follows: we will construct a linear subspace of $\mathscr{F}(S, \mathscr{B}(S))$ that contains $\mathscr{F}(S, \mathscr{A})$ and some other functions of interest to us, and we will extend the positive linear functional $f \mapsto \int_S f \, d\theta$ linearly to this subspace. A further extension brings us to a positive linear functional Λ on $\mathscr{F}(S, \mathscr{B}(S))$, which can then be represented through a measure $\mu \in \mathbf{P}(S, \mathscr{B}(S))$ so that

$$\Lambda(f) = \int_S f \, d\mu$$

holds. Clearly, μ extends θ and is the measure we are looking for.

By standard arguments from measure theory, the commutativity of the diagram entails that for $i = 1, 2$,

$$\forall f_i \in \mathscr{F}\left(X_i, \psi_i^{-1}[\mathscr{C}]\right) : \int_{X_i} f_i \, d\mu_i = \int_S f_i \circ \pi_i \, d\theta,$$

and, by the same token, it is sufficient to find an extension $\mu \in \mathbf{P}(S, \mathscr{B}(S))$ to $\theta \in \mathbf{P}(S, \mathscr{A})$ such that, for $i = 1, 2$,

$$\forall f_i \in \mathscr{F}(X_i, \mathscr{B}(X_i)) : \int_{X_i} f_i \, d\mu_i = \int_S f_i \circ \pi_i \, d\mu.$$

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

1 Put, for $i = 1, 2$,

$$\mathscr{D}_i := \{f_i \circ \pi_i \mid f_i \in \mathscr{F}(X_i, \mathscr{B}(X_i))\}.$$

Then $\mathscr{D}_i \subseteq \mathscr{F}(S, \mathscr{B}(S))$, and

$$\Lambda_0(f_i \circ \pi_i) := \int_{X_i} f_i \, d\mu_i.$$

So $\Lambda_0 : \mathscr{D}_1 \cup \mathscr{D}_2 \to \mathbb{R}$ is well defined.
In fact, let $g \in \mathscr{D}_1 \cap \mathscr{D}_2$, so that there exist $f_i \in \mathscr{F}(X_i, \mathscr{B}(X_i))$ with

$$g = f_1 \circ \pi_1 = f_2 \circ \pi_2.$$

We claim that f_1 is constant on the atoms of $\psi_1^{-1}[\mathscr{C}]$. Take $x_1, x_1' \in X_1$ with $\psi_1(x_1) = \psi_1(x_1')$. Then there exists $x_2 \in X_2$ such that $\langle x_1, x_2\rangle \in S, \langle x_1', x_2\rangle \in S$. Hence

$$f_1(x_1) = g(x_1, x_2) = f_2(x_2) = g(x_1', x_2) = f_1(x_1').$$

Thus f_1 is $\psi_1^{-1}[\mathscr{C}]$-measurable by Proposition 3.3, and, consequently,

$$\begin{aligned}\int_S g \, d\theta &= \int_S f_1 \circ \pi_1 \, d\theta \\ &= \int_{X_1} f_1 \, d\mu_1.\end{aligned}$$

Similarly,

$$\int_S g \, d\theta = \int_{X_2} f_2 \, d\mu_2$$

is established. This implies that Λ_0 is well defined.

2 Let the linear functional $\Lambda_1 : \mathscr{F}(S, \mathscr{A}) \to \mathbb{R}$ be defined through

$$\Lambda_1(f) := \int_S f \, d\theta.$$

We will look for a joint extension of Λ_0 and Λ_1 to the linear space spanned by $\mathscr{F}(S, \mathscr{A}) \cup \mathscr{D}$, where $\mathscr{D} := \mathscr{D}_1 \cup \mathscr{D}_2$. This requires that both functionals yield the same value on the intersection $\mathscr{F}(S, \mathscr{A}) \cap (\mathscr{D}_1 \cup \mathscr{D}_2)$. Assume first that $g \in \mathscr{F}(S, \mathscr{A}) \cap \mathscr{D}_1$, and thus $g = f_1 \circ \pi_1$ for some $f_1 \in \mathscr{F}(X_1, \mathscr{B}(X_1))$. Since g does not depend on the second component, we may infer from the definition of $\mathscr{A}$ that f_1 is even $\psi_1^{-1}[\mathscr{C}]-$ measurable, and hence

$$\begin{aligned}\Lambda_1(g) &= \int_S g \, d\theta \\ &= \int_S f_1 \circ \pi_1 \, d\theta \\ &= \int_{X_1} f_1 \, d\mu_1 \\ &= \Lambda_0(g).\end{aligned}$$

The argument for $g \in \mathscr{F}(S, \mathscr{A}) \cap \mathscr{D}_2$ is similar.
Let Λ_2 be the joint linear extension of Λ_1 on $\mathscr{F}(S, \mathscr{A})$ and of Λ_0 on $\mathscr{D}$ to the linear space spanned by $\mathscr{F}(S, \mathscr{A})$ and $\mathscr{D}$.
From the construction, it is clear that $\Lambda_2(1) = 1$ holds, and that Λ_2 is monotone.

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

3 The Hahn–Banach Theorem for ordered linear spaces (Jacobs 1978, Lemma IX.1.4) gives a positive linear operator

$$\Lambda : \mathscr{F}(S, \mathscr{B}(S)) \to \mathbb{R}$$

that extends Λ_2. Since each continuous and bounded map $f : X_1 \times X_2 \to \mathbb{R}$ becomes a member of $\mathscr{F}(S, \mathscr{B}(S))$ when restricted to S, we obtain a positive linear operator

$$\Lambda'(f) := \Lambda(f|_S)$$

on the linear space of all continuous maps $X_1 \times X_2 \to \mathbb{R}$. Because $X_1 \times X_2$ is compact, the famous Riesz Representation Theorem yields a probability measure

$$\mu' \in \mathbf{P}(X_1 \times X_2, \mathscr{B}(X_1 \times X_2))$$

with

$$\begin{aligned} \Lambda'(f) &= \int_{X_1 \times X_2} f \, d\mu' \\ &= \int_S f \, d\mu' \end{aligned}$$

for each $f \in \mathscr{F}(X_1 \times X_2, \mathscr{B}(X_1 \times X_2))$. Define, for $B \in \mathscr{B}(S)$, the measure μ by restricting μ' to $\mathscr{B}(S)$, so

$$\mu(B) := \mu'(B \cap S).$$

We will now show that $\mu \in \mathbf{P}(S, \mathscr{B}(S))$ is the measure we are looking for.

4 Let $f \in \mathscr{F}(S, \mathscr{A})$. Then

$$\begin{aligned} \int_S f \, d\theta &= \Lambda_1(f) \\ &= \Lambda_2(f) \\ &= \Lambda'(f) \\ &= \int_S f \, d\mu, \end{aligned}$$

and thus μ extends θ. Let $f_i \in \mathscr{F}(X_1, \mathscr{B}(X_i))$. Then $f_i \circ \pi_i \in \mathscr{D}_i \subseteq \mathscr{D}$, and hence

$$\begin{aligned} \int_{X_i} f_i \, d\mu_i &= \Lambda_0(f_i \circ \pi_i) \\ &= \Lambda_2(f_i \circ \pi_i) \\ &= \Lambda'(f_i \circ \pi_i) \\ &= \int_S f_i \circ \pi_i \, d\mu, \end{aligned}$$

rendering the diagram commutative. □

The compactness assumption was used in the proof only to establish the existence of a measure, given a suitable linear functional on the space of continuous functions. We do not have the Riesz Representation Theorem directly at our disposal in the general case, but we may, nevertheless, still make use of compactness since each Polish space may be embedded into a compact metric space as a measurable subspace.

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

Proof for the general case. The famous characterisation of Polish spaces due to Alexandrov (Srivastava 1998, Remark 2.2.8) states that a topological space is Polish iff it is homeomorphic to a G_δ-subset of $[0,1]^{\mathbb{N}}$. In particular, a Polish space is a measurable and dense subset of a compact metric space. We will capitalise on this: X_1 and X_2 will be embedded into compact metric spaces, and this embedding will take ψ_1, ψ_2 and the measure θ with it. We then apply the extension procedure for the compact case. Restricting what we got from there to the original scenario, we conclude that the assertion also holds for the non-compact case.

1 X_i is a dense measurable subset of a compact metric space $\underline{X_i}$ by Kechris (1994, Theorem 4.14), and $\psi_i : X_i \to Z$ may be extended to a Borel measurable map $\underline{\psi_i} : \underline{X_i} \to Z$ by Srivastava (1998, Proposition 3.3.4).
Define, for $B_i \in \mathscr{B}(\underline{X_i})$,

$$\underline{\mu_i}(B_i) := \mu_i(B_i \cap X_i),$$

and put

$$S_0 := \{\langle x_1, x_2\rangle \in \underline{X_1} \times \underline{X_2} \mid \underline{\psi_1}(x_1) = \underline{\psi_2}(x_2)\}.$$

Then

$$S_0 = \left(\underline{\psi_1} \times \underline{\psi_2}\right)^{-1} \left[\Delta_{\underline{X_1} \times \underline{X_2}}\right].$$

So

$$\begin{aligned} S_0 \;&\in\; \left(\underline{\psi_1} \times \underline{\psi_2}\right)^{-1} [\mathscr{B}(Z \times Z)] \\ &= \underline{\psi_1}^{-1}[\mathscr{C}] \otimes \underline{\psi_2}^{-1}[\mathscr{C}]. \end{aligned}$$

Since $X_i \in \underline{\psi_i}^{-1}[\mathscr{C}]$, and since $S = S_0 \cap (X_1 \times X_2)$, we see that

$$S \in \underline{\psi_1}^{-1}[\mathscr{C}] \otimes \underline{\psi_2}^{-1}[\mathscr{C}].$$

Now, for $E \in \underline{\psi_1}^{-1}[\mathscr{C}] \otimes \underline{\psi_2}^{-1}[\mathscr{C}]$, put

$$\underline{\theta}(E) := \theta(E \cap S).$$

So $\underline{\theta}(S_0 \setminus S) = 0$, and hence $\underline{\theta}$ is concentrated on S.

2 The construction shows that

$$\begin{array}{ccc} (S_0, \mathscr{A}_0, \underline{\theta}) & \xrightarrow{\pi_2} & \left(\underline{X_2}, \mathscr{B}(\underline{X_2}), \underline{\mu_2}\right) \\ \Big\downarrow{\scriptstyle \underline{\pi_1}} & & \Big\downarrow{\scriptstyle \underline{\psi_2}} \\ \left(\underline{X_1}, \mathscr{B}(\underline{X_1}), \underline{\mu_1}\right) & \xrightarrow[\underline{\psi_1}]{} & (Z, \mathscr{C}, \nu) \end{array}$$

commutes, where

$$\mathscr{A}_0 := \left(\underline{\psi_1}^{-1}[\mathscr{C}] \otimes \underline{\psi_2}^{-1}[\mathscr{C}]\right) \cap S_0.$$

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

The compact case applies, hence we can find an extension $\underline{\mu} \in \mathbf{P}(S_0, \mathscr{B}(S_0))$ for $\underline{\theta} \in \mathbf{P}(S_0, \mathscr{A}_0)$, which makes this diagram commute:

$$\begin{array}{ccc} (S_0, \mathscr{B}(S_0), \underline{\mu}) & \xrightarrow{\pi_2} & (\underline{X_2}, \mathscr{B}(\underline{X_2}), \underline{\mu_2}) \\ \downarrow \underline{\pi_1} & & \downarrow \underline{\psi_2} \\ (\underline{X_1}, \mathscr{B}(\underline{X_1}), \underline{\mu_1}) & \xrightarrow[\underline{\psi_1}]{} & (Z, \mathscr{C}, \nu) \end{array}$$

3 We now roll back compactification. For the Borel set $B \subseteq S$, put

$$\mu(B) := \underline{\mu}(B \cap S).$$

So $\mu \in \mathbf{P}(S, \mathscr{B}(S))$, since

$$\underline{\mu}(S_0 \setminus S) = \underline{\theta}(S_0 \setminus S) = 0.$$

The other properties are obvious, so we are done with the general case too. □

The crucial point in this argument has been to prevent any mass from vanishing, that is, to see that $\mu(S) = 1$, which, in turn, could be established from the fact that $\underline{\mu}$ extends $\underline{\theta}$, and for which the incorporation of $\mathscr{F}(S, \mathscr{A})$ into the extension process was responsible.

We can reformulate Proposition 4.1 in terms of subprobability distributions. It then states that there is sometimes a common distribution for two random variables with values in a Polish space with preassigned marginal distributions. This is a cornerstone for the construction leading to the proof of Theorem 5.1, which shows, in particular, where Edalat's work could enter the present discussion.

Proposition 4.2. Let X_1, and X_2 be Polish spaces and $(Z, \mathscr{C})$ be a separable measurable space. Assume also that

$$\psi_i : X_i \to Z \ (i = 1, 2)$$

are measurable and surjective maps. Define

$$S := \{\langle x_1, x_2 \rangle \in X_1 \times X_2 \mid \psi_1(x_1) = \psi_2(x_2)\}$$

and endow S with the trace $\mathscr{B}(S)$ of the product σ-algebra. Assume that subprobability measures $\mu_1 \in \mathbf{S}(X_1), \mu_2 \in \mathbf{S}(X_2), \theta \in \mathbf{S}(S)$ are given such that

$$\forall E_i \in \psi_i^{-1}[\mathscr{C}] : \mathbf{S}(\pi_i)(\theta)(E_i) = \mu_i(E_i) \ (i = 1, 2)$$

where $\pi_1 : S \to X_1, \pi_2 : S \to X_2$ are the projections. Then there exists $\mu \in \mathbf{S}(S)$ such that

$$\forall E_i \in \mathscr{B}(X_i) : \mathbf{S}(\pi_i)(\mu)(E_i) = \mu_i(E_i) \ (i = 1, 2).$$

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

Proof.

1 We want to apply Proposition 4.1, so we need to show how to construct diagram (∗). From the assumption, we see that

$$\begin{aligned}\mu_1(X_1) &= \theta(\pi_1^{-1}[X_1]) \\ &= \theta(S \cap (X_1 \times X_2)) \\ &= \theta(S).\end{aligned}$$

And similarly for μ_2, so $\mu_1(X_1) = \mu_2(X_2) = \theta(S)$. If $\theta(S) = 0$, the assertion is pretty obvious, so we may assume that $\theta(S) > 0$. Hence there is no loss of generality if we assume that all measures are probability measures.

2 Let $C \in \mathscr{C}$. Then

$$\begin{aligned}\mu_1(\psi_1^{-1}[C]) &= \mathbf{P}(\pi_1)(\theta)(\psi_1^{-1}[C]) \\ &= \mathbf{P}(\pi_2)(\theta)(\psi_2^{-1}[C]) \\ &= \mu_2(\psi_2^{-1}[C]),\end{aligned}$$

since $\pi_1 \circ \psi_1 = \pi_2 \circ \psi_2$ holds on S. So, put

$$\nu(C) := \mu_1(\psi_1^{-1}[C]).$$

Then $\nu \in \mathbf{P}(Z, \mathscr{C})$ such that

$$\psi_i : (X_i, \mathscr{B}(X_i), \mu_i) \to (Z, \mathscr{C}, \nu)$$

is a morphism in $\mathfrak{P}$ for $i = 1, 2$. The assumption implies that

$$\pi_i : (S, \mathscr{A}, \theta) \to (X_i, \mathscr{B}(X_i), \mu_i)$$

is a morphism for $i = 1, 2$, where $\mathscr{A}$ is the trace of the σ-algebra $\psi_1^{-1}[\mathscr{C}] \otimes \psi_2^{-1}[\mathscr{C}]$ on S. Consequently, the assertion follows from Proposition 4.1. □

There are other ways of establishing the Proposition in important special cases, as we will discuss briefly.

Remarks

1 If Z is also a Polish space, and if $\psi_i : X_i \to Z$ are bijections, the Blackwell–Mackey Theorem (Srivastava 1998, Theorem 4.5.7) shows that $\psi_i^{-1}[\mathscr{C}] = \mathscr{B}(X_i)$. In this case the given measure $\theta \in \mathbf{S}(S)$ is the desired one. This is because the trace σ-algebra $\mathscr{A}$ equals $\mathscr{B}(S)$:

$$\begin{aligned}\mathscr{A} &= S \cap \psi_1^{-1}[\mathscr{C}] \otimes \psi_2^{-1}[\mathscr{C}] \\ &= S \cap \mathscr{B}(X_1) \otimes \mathscr{B}(X_2) \\ &= S \cap \mathscr{B}(X_1 \times X_2) \\ &= \mathscr{B}(S),\end{aligned}$$

so θ has the desired properties on the proper σ-algebra.

2 The maps $\psi_i : X_i \to Z$ are morphisms in Edalat's category of probability measures on Polish spaces (Edalat 1999), provided Z is a Polish space. The assertion can then

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

be deduced by tracing the development in Edalat (1999, Corollary 5.4). The proof given above applies to Edalat's situation as well, but it should be clear that our proof is independent of Edalat's. The development for the latter depends substantially on the theory of regular conditional probabilities on analytic spaces, so one might gain the impression that the existence of the measure in question also depends on these probabilities. The proof for Proposition 4.1 shows that this is not the case, and that a rather straightforward proof can be given. Hence we are in the lucky position of having two independent proofs *for the Polish case*. Which one is preferred is largely a matter of taste: Edalat's proof working in analytic spaces, or the one proposed here depending on the Hahn–Banach Theorem as a classical tool in analysis (but making use of the sometimes dreaded Axiom of Choice).

We will need an extension theorem for stochastic relations in order to secure the existence of semi-pullbacks for analytic spaces. We begin with a statement on the extension of a probability measure on a sub-σ-algebra. Note that we do not claim the uniqueness of the extension. This is different from the usual measure extensions in Measure Theory.

Lemma 4.1. Let $\mathscr{A}$ be a sub-σ-algebra of the Borel sets of a Polish space X, and assume that θ is a probability measure on $\mathscr{A}$. Then θ can be extended to a probability measure on all of $\mathscr{B}(X)$.

Proof. We will only need to sketch the proof, since the main work has already been done in the proof of Proposition 4.1. Although the assertion is a bit different, the pattern of the argument is very similar to the one presented already.

1 First X is assumed to be compact, then a combination of the Hahn–Banach and Riesz Representation theorems yields the existence of the desired measure.
2 If X is not compact, it is embedded as above as a measurable subset into a compact metric space. There the existence of an extension is established, and exactly the same technique as above moves that measure to the Borel sets of X. □

The application interesting us here is the possibility of establishing an extension to probabilistic relations.

Proposition 4.3. Let X and Y be Polish spaces. Assume that $\mathscr{B} \subseteq \mathscr{B}(Y)$ is a countably generated σ-algebra, and let $K_0 : (X, \mathscr{B}(X)) \rightsquigarrow (Y, \mathscr{B})$ be a stochastic relation. Then K_0 can be extended to a stochastic relation $K : (X, \mathscr{B}(X)) \rightsquigarrow (Y, \mathscr{B}(Y))$.

Proof. We will construct a probability measure on the product $\mathscr{B}(Y) \otimes \mathscr{B}$. We will then extend this measure and obtain the desired extension to the probabilistic relation through disintegration.

1 Let μ be a probability measure on $\mathscr{B}(X)$, and define for $D \in \mathscr{B}(Y) \otimes \mathscr{B}$ the measure

$$\mu_0(D) := \int_X K_0(x)(D_x)\ \mu(dx),$$

where, as usual, $D_x := \{y \in Y \mid \langle x, y\rangle \in D\}$, and, by standard arguments, $D_x \in \mathscr{B}$ for any $D \in \mathscr{B}(Y) \otimes \mathscr{B}$. Let μ_1 be an extension of μ_0 to all of $\mathscr{B}(X \times Y)$. This extension

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

exists by Lemma 4.1. Since μ_1 is a measure on the product of two Polish spaces, there exists a stochastic relation $K_1 : (X, \mathscr{B}(X)) \rightsquigarrow (Y, \mathscr{B}(Y))$ such that

$$\mu_1(D) = \int_X K_1(x)(D_x)\ \mu(dx)$$

for all $D \in \mathscr{B}(X \times Y)$. This follows from the existence of regular conditional distributions on Polish spaces (Parthasarathy 1967, Theorem V.8.1). It is well known that K_1 is unique up to sets of μ-measure zero.

2 Let $\mathscr{B}_0 := \{B_n \mid n \in \mathbb{N}\}$ be a countable generator of the σ-algebra $\mathscr{B}$; we may and do assume that $\mathscr{B}_0$ is closed under finite intersections (otherwise, form all finite intersections of elements of $\mathscr{B}_0$, so that this is still a countable generator that is closed under finite intersections). Now let $E \in \mathscr{B}$, then

$$\mu_0(E \times Y) = \mu_1(E \times Y)$$

by the construction of this extension. Thus, there exists for each $E \in \mathscr{B}$ a set $N(E) \in \mathscr{B}(X)$ with $\mu(N(E)) = 0$ such that

$$\forall x \in X \setminus N(E) : K_0(x)(E) = K_1(x)(E).$$

Now put

$$N := \bigcup_{n \in \mathbb{N}} N(B_n)$$

as the set of all possibly violating x, so $N \in \mathscr{B}(X)$, and $\mu(N) = 0$.

3 We claim that for any $x \in X \setminus N$ the equality $K_0(x)(E) = K_1(x)(E)$ holds for every Borel set $E \in \mathscr{B}$. In fact, put

$$\mathscr{E} := \{E \in \mathscr{B} \mid \forall x \in X \setminus N : K_0(x)(E) = K_1(x)(E)\}.$$

Then $\mathscr{B}_0 \subseteq \mathscr{E}$ by construction, $\mathscr{E}$ contains Y, and $\mathscr{E}$ is closed under complementation and disjoint countable unions. Thus $\mathscr{E} = \mathscr{B}$ can be inferred by the $\pi - \lambda$-Theorem (Kechris 1994, Theorem 10.1.iii). Now let μ_2 be an arbitrary probability measure on $\mathscr{B}(Y)$, and define the stochastic relation K by cases as follows:

$$K(x)(D) := \begin{cases} K_1(x)(D), & x \notin N, D \in \mathscr{B}(Y) \\ K_0(x)(D), & x \in N, D \in \mathscr{B} \\ \mu_2(D), & x \in N, D \notin \mathscr{B}. \end{cases}$$

This relation has the desired properties. □

This proposition applies directly to making a pair of surjective and measurable maps into morphisms under a rather weak condition of separability.

Lemma 4.2. Let $\mathsf{M} := (A, B, M)$ be a stochastic relation between the measurable spaces A and B, and assume that B is separable. If X and Y are Polish spaces with measurable and surjective maps $\phi : X \to A, \psi : Y \to B$, there exists a stochastic relation $\mathsf{K} := (X, Y, K)$ with $K : X \rightsquigarrow Y$ that makes $\mathsf{f} := (\phi, \psi)$ a morphism $\mathsf{f} : \mathsf{K} \to \mathsf{M}$.

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

Proof. Let $\mathscr{B}$ the σ-algebra on B. Then $\psi^{-1}[\mathscr{B}]$ is a countably generated sub-σ-algebra of $\mathscr{B}(Y)$. Define, for $x \in X$ and $D \in \mathscr{B}$,

$$K_0(x)(\psi^{-1}[D]) := M(\phi(x))(D).$$

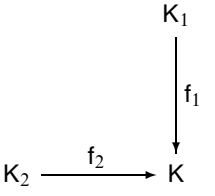
Then $K_0 : (X, \mathscr{B}(X)) \rightsquigarrow (Y, \psi^{-1}[\mathscr{B}])$ is a stochastic relation that can be extended to a stochastic relation $K : (X, \mathscr{B}(X)) \rightsquigarrow (Y, \mathscr{B}(Y))$ by Proposition 4.3. It is plain from the construction that $\mathbf{S}(\psi) \circ K = M \circ \phi$. □

5. Semi-pullbacks

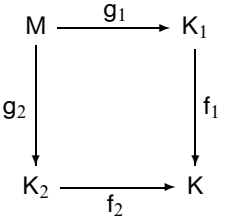
We will show now that semi-pullbacks exist in a rather general setting, generalising the construction in Doberkat (2003) and Edalat (1999). This will ultimately lead to us showing that semi-pullbacks exist for analytic objects, and it will be shown that the object underlying such a semi-pullback is Polish.

The next lemma is our central Lemma.

Lemma 5.1. Let K_i be Polish objects, and assume that $\mathsf{K} = (X, Y, K)$ is a stochastic relation, where X, Y are separable measurable spaces. In $\mathfrak{Stoch}$, each diagram



has a semi-pullback



with a Polish object M.

Proof.

1 Assume $\mathsf{K}_i = (X_i, Y_i, K_i)$ with $\mathsf{f}_i = (\phi_i, \psi_i), i = 1, 2$. In view of Lemma 3.2, we may, and do, assume that the respective σ-algebras on X and Y are the Borel sets of second countable metric spaces. Because of Proposition 3.1, we may assume that the respective σ-algebras on X_1 and X_2 are obtained from Polish topologies that render ϕ_1 and K_1 as well as ϕ_2 and K_2 continuous. These topologies are fixed for the proof. Put

$$\begin{aligned} A &:= \{\langle x_1, x_2\rangle \in X_1 \times X_2 \mid \phi_1(x_1) = \phi_2(x_2)\}, \\ B &:= \{\langle y_1, y_2\rangle \in Y_1 \times Y_2 \mid \psi_1(y_1) = \psi_2(y_2)\}. \end{aligned}$$

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

Then both A and B are closed, and hence Polish. $\alpha_i : A \to X_i$ and $\beta_i : B \to Y_i$ are the projections, $i = 1, 2$. The diagrams

$$\begin{array}{ccccc} X_1 & \xrightarrow{\phi_1} & X & \xleftarrow{\phi_2} & X_2 \\ \downarrow{\scriptstyle K_1} & & \downarrow{\scriptstyle K} & & \downarrow{\scriptstyle K_2} \\ \mathbf{S}(Y_1) & \xrightarrow[\mathbf{S}(\psi_1)]{} & \mathbf{S}(Y) & \xleftarrow[\mathbf{S}(\psi_2)]{} & \mathbf{S}(Y_2) \end{array}$$

are commutative by assumption, so we know that for $x_i \in X_i$

$$\begin{aligned} K(\phi_1(x_1)) &= \mathbf{S}(\psi_1)(K_1(x_1)) \\ K(\phi_2(x_2)) &= \mathbf{S}(\psi_2)(K_2(x_2)) \end{aligned}$$

both hold. The construction implies that

$$(\psi_1 \circ \beta_1)(y_1, y_2) = (\psi_2 \circ \beta_2)(y_1, y_2)$$

is true for $\langle y_1, y_2 \rangle \in B$, and $\psi_1 \circ \beta_1 : B \to Y$ is surjective.

2 Fix $\langle x_1, x_2 \rangle \in A$. We now use separability of the target spaces: Corollary 3.1 shows that the image of a surjective map under $\mathbf{S}$ is onto again, so there exists $\mu_0 \in \mathbf{S}(B)$ with

$$\mathbf{S}(\psi_1 \circ \beta_1)(\mu_0) = K(\phi_1(x_1)),$$

and, consequently,

$$\mathbf{S}(\psi_i \circ \beta_i)(\mu_0) = \mathbf{S}(\psi_i)(K_i(x_i)) \ (i = 1, 2).$$

But this means

$$\forall E_i \in \psi_i^{-1}[\mathscr{B}(Y)] : \mathbf{S}(\beta_i)(\mu_0)(E_i) = K_i(x_i)(E_i) \ (i = 1, 2).$$

Put

$$\Gamma(x_1, x_2) := \{\mu \in \mathbf{S}(B) \mid \mathbf{S}(\beta_1)(\mu) = K_1(x_1) \wedge \mathbf{S}(\beta_2)(\mu) = K_2(x_2)\}.$$

Then Proposition 4.2 shows that $\Gamma(x_1, x_2) \neq \varnothing$.

3 Since K_1 and K_2 are continuous, $\Gamma : A \to \mathbb{F}(\mathbf{S}(B))$ is easily established. The set

$$\exists\Gamma(C) = \{\langle x_1, x_2 \rangle \in A \mid \Gamma(x_1, x_2) \cap C \neq \varnothing\}$$

is closed in A for compact $C \subseteq \mathbf{S}(B)$. In fact, let $(\langle x_1^{(n)}, x_2^{(n)} \rangle)_{n \in \mathbb{N}}$ be a sequence in this set with $x_i^{(n)} \to x_i$, as $n \to \infty$ for $i = 1, 2$, so $\langle x_1, x_2 \rangle \in A$. There exists $\mu_n \in C$ such that $\mathbf{S}(\beta_i)(\mu_n) = K_i(x_i^{(n)})$. Because C is compact, there exists a converging subsequence $\mu_{s(n)}$ and $\mu \in C$ with $\mu = \lim_{n \to \infty} \mu_{s(n)}$ in the topology of weak convergence. Continuity of K_i implies that $\mathbf{S}(\beta_i)(\mu) = K_i(x_i)$, and, consequently, $\langle x_1, x_2 \rangle \in \exists\Gamma(C)$, so this set is closed, and hence measurable.

4 From Proposition 3.2, we can now infer that there exists a measurable map $M : A \to \mathbf{S}(B)$ such that $M(x_1, x_2) \in \Gamma(x_1, x_2)$ holds for every $\langle x_1, x_2 \rangle \in A$. Thus $M : A \rightsquigarrow B$ is a

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

stochastic relation with

$$K_1 \circ \alpha_1 = \mathbf{S}(\beta_1) \circ M,$$
$$K_2 \circ \alpha_2 = \mathbf{S}(\beta_2) \circ M.$$

Hence $\mathsf{M} := (A, B, M)$ is the desired semi-pullback. □

This statement includes the following interesting special cases.

Theorem 5.1. Semi-pullbacks exist in the category $\mathfrak{PStoch}$ of stochastic relations over Polish spaces.

Proof. This follows immediately from Lemma 5.1. □

Corollary 5.1. Suppose that in the diagram of Lemma 5.1 the target object K is an analytic object. Then a Polish semi-pullback of that diagram exists in $\mathfrak{Stoch}$.

Proof. This also follows immediately from Lemma 5.1. □

Now suppose that we have an analytic object M. This means that we can find a Polish object K and a morphism $\mathsf{g} : \mathsf{K} \to \mathsf{M}$ by Lemma 4.2. This is because analytic spaces are surjective images of Polish spaces under measurable maps, and because analytic spaces are, being measurable spaces, separable. This observation has the interesting consequence that semi-pullbacks do exist for analytic spaces.

Corollary 5.2. Let M_i be an analytic object for $i = 1, 2$ and assume that $\mathsf{M} := (A, B, M)$ is a stochastic relation for the separable measurable spaces A, B. For the pair

$$\mathsf{M}_1 \xrightarrow{\mathsf{f}_1} \mathsf{M} \xleftarrow{\mathsf{f}_2} \mathsf{M}_2,$$

there exist both a Polish object K and morphisms

$$\mathsf{M}_1 \xleftarrow{\mathsf{g}_1} \mathsf{K} \xrightarrow{\mathsf{g}_2} \mathsf{M}_2$$

forming a semi-pullback.

Proof. We can find Polish objects K_i and morphisms in $\mathfrak{Stoch}$ extending the diagram

$$\begin{array}{ccccc} \mathsf{K}_1 & & & & \mathsf{K}_2 \\ \downarrow{\scriptstyle \mathsf{h}_1} & & & & \downarrow{\scriptstyle \mathsf{h}_2} \\ \mathsf{M}_1 & \xrightarrow{\mathsf{f}_1} & \mathsf{M} & \xleftarrow{\mathsf{f}_2} & \mathsf{M}_2 \end{array}$$

Now, using Lemma 5.1, find a semi-pullback

$$\mathsf{K}_1 \xleftarrow{\mathsf{t}_1} \mathsf{K} \xrightarrow{\mathsf{t}_2} \mathsf{K}_2$$

for the diagram

$$\mathsf{K}_1 \xrightarrow{\mathsf{h}_1 \circ \mathsf{f}_1} \mathsf{M} \xleftarrow{\mathsf{h}_2 \circ \mathsf{f}_2} \mathsf{K}_2.$$

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

Here K is a Polish object, and $\mathsf{t}_1, \mathsf{t}_2$ are morphisms in $\mathfrak{Stoch}$. Putting $\mathsf{g}_i := \mathsf{h}_i \circ \mathsf{t}_i$ for $i = 1, 2$ now establishes the claim. □

Thus we have, in particular, established the following theorem.

Theorem 5.2. The category $\mathfrak{AStoch}$ of stochastic relations over analytic spaces has semi-pullbacks. They may be chosen as Polish objects.

We close with a negative result, indicating that it is not possible to strengthen the results obtained here towards the existence of weak pullbacks or to pullbacks. Recall that a semi-pullback $r : d \to a, s : d \to c$ for a pair of morphisms $f : a \to b, g : c \to b$ is a weak pullback iff whenever $r' : d' \to a, s' : d' \to c$ forms a commutative diagram with f and g (that is, $f \circ r' = g \circ s'$ holds), there exists a morphism $h : d' \to d$ with $r' = r \circ h, s' = s \circ h$. If morphism h is uniquely determined, d together with r and s is called a pullback.

The category $\mathfrak{Po}$ is the full subcategory of $\mathfrak{P}$ (*cf.* Section 4) that has probability spaces based on Polish spaces as objects.

Proposition 5.1. Let (X, μ) and (Y, ν) be objects in $\mathfrak{Po}$, and assume that $\phi : (X, \mu) \to (Y, \nu)$ is a morphism in $\mathfrak{Po}$ such that $\phi : X \to Y$ is not bijective. Then the kernel pair

$$(X, \mu) \xrightarrow{\phi} (Y, \nu) \xleftarrow{\phi} (X, \mu)$$

does not have a weak pullback in $\mathfrak{Po}$.

Proof. Assume that (P, ρ) is a weak pullback for that kernel pair with morphisms based on the maps $\pi_1 : P \to X$ and $\pi_2 : P \to X$. Because the category of Polish spaces with surjective Borel maps has finite products and equalisers, we conclude that

$$P = \{\langle x_1, x_2 \rangle \in X \times X \mid f(x_1) = f(x_2)\},$$

and that π_1, π_2 are just the projections, see Mac Lane (1997, Exercise III.4.7). Because the identity $id_X : (X, \mu) \to (X, \mu)$ is plainly a morphism with $f \circ id_X = f \circ id_X$, we find a morphism $\chi : (X, \mu) \to (P, \rho)$ such that $\pi_1 \circ \chi = \pi_2 \circ \chi$. Thus each element of P must be of the form $\langle x, x \rangle$, contradicting the fact that ϕ is not injective. □

Because $\mathfrak{Po}$ is a full subcategory of $\mathfrak{PStoch}$, we may conclude with the following corollary.

Corollary 5.3. Neither $\mathfrak{PStoch}$ nor $\mathfrak{AStoch}$ have weak pullbacks.

6. Conclusion

This paper shows that semi-pullbacks do exist in the category of stochastic relations over analytic spaces, using Borel selections for measurable set-valued maps. At the core of the proof lies an extension argument for a measure on a Polish space that has given marginal distributions. It employs the Hahn–Banach Theorem as well as the Riesz Representation Theorem. The solution thus makes use of the Axiom of Choice, in contrast to Edalat's approach (Edalat 1999), which works with the conditional distribution under somewhat stronger assumptions.

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press

The result proposed here find applications, for example, when interpreting modal logic through stochastic relations: it can be shown that probabilistic Kripke models for a variety of modal logics are bisimilar iff they accept exactly the same formulas. This probabilistic version of the Hennessy–Milner Theorem is a generalisation of the well-known characterisation of bisimulation through a simple negation-free modal logic with a countable number of diamonds, which was established in Deharnais *et al.* (2002) and Doberkat (2003). The general nature of the result presented here removes the restrictions that had to be imposed in these results. Analytic spaces have the nice property that they are closed under factorisation of smooth equivalence relations that occur quite naturally for stochastic relations. The existence of semi-pullbacks helps in investigating these congruences, see Doberkat (2004a) for some first results. Based on the result derived here, and on an Isomorphism Theorem for them, it can be shown that two stochastic relations are bisimilar provided they have isomorphic factors; the converse also holds under the assumption of compactness – see Doberkat (2004b).

Acknowledgements

The author thanks Georgios Lajios for his critical comments. Professor Michael Ershov (Essen) was helpful in pointing out a way to simplify the approach to the measure extension in Section 4, leading ultimately to a much more general solution than the author originally had in mind. The referees' comments helped to clarify the representation.

The diagrams in this paper were typeset using Paul Taylor's wonderful `diagrams` package.

References

Arveson, W. (1976) *An Invitation to C*-Algebra*, Graduate Texts in Mathematics, Springer-Verlag **39**.

Desharnais, J., Edalat, A. and Panangaden, P. (2002) Bisimulation of labelled Markov-processes. *Information and Computation* **179** (2) 163–193.

Doberkat, E.-E. (2003) Semi-pullbacks and bisimulations in categories of stochastic relations. In: Proc. ICALP'03. *Springer-Verlag Lecture Notes in Computer Science* **2719** 996–1007.

Doberkat, E.-E. (2004a) Factoring stochastic relations. *Information Processing Letters* **90** (4) 161–166.

Doberkat, E.-E. (2004b) Stochastic relations: congruences, bisimulations and the Hennessy–Milner Theorem. Technical Report 786, Department of Computer Science, University of Dortmund.

Edalat, A. (1999) Semi-pullbacks and bisimulation in categories of Markov processes. *Mathematical Structures in Computer Science* **9** (5) 523–543.

Himmelberg, C. J. (1975) Measurable relations. *Fund. Math.* **87** 53–72.

Jacobs, K. (1978) *Measure and Integral*, Academic Press, New York.

Kechris, A. S. (1994) *Classical Descriptive Set Theory*, Graduate Texts in Mathematics, Springer-Verlag **156**.

Mac Lane, S. (1997) *Categories for the Working Mathematician*, second edition, Graduate Texts in Mathematics, Springer-Verlag **5**.

Parthasarathy, K. R. (1967) *Probability Measures on Metric Spaces*, Academic Press.

Srivastava, S. M. (1998) *A Course on Borel Sets*, Graduate Texts in Mathematics, Springer-Verlag **180**.

Wagner, D. H. (1977) A survey of measurable selection theorems. *SIAM J. Control Optim.* **15** (5) 859–903.

https://doi.org/10.1017/S096012950500472X Published online by Cambridge University Press