Semi-Pullbacks for Stochastic Relations over Analytic Spaces[1]

Ernst-Erich Doberkat[2]
Chair for Software Technology
University of Dortmund
doberkat@acm.org

January 22, 2004

[1]Preliminary results were presented at ICALP'03 at Eindhoven, see [3]

[2]Research funded in part by *Deutsche Forschungsgemeinschaft*, grant DO 263/8-1, *Algebraische Eigenschaften stochastischer Relationen*

Abstract

The existence of semi-pullbacks for stochastic relations over analytic spaces is addressed. Via a measure extension it is shown 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.

Keywords: Stochastic relations over Polish and analytic spaces, semi-pullbacks, (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. For e.g. stochastic relations, semi-pullbacks are helpful when investigating the relationship between modal logics and bisimilar stochastic Kripke models. The argumentation goes like this: given two stochastic Kripke models $\mathcal{K}_1$ and $\mathcal{K}_2$ that accepts exactly the same formulas, then we can construct a third Kripke model $\mathcal{L}$ and morphisms Φ_1 and Φ_2 that form a co-span

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

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

$$\begin{tikzcd}
\mathcal{M} \arrow[r, "\Psi_1"] \arrow[d, "\Psi_2"'] & \mathcal{K}_1 \arrow[d, "\Phi_1"] \\
\mathcal{K}_2 \arrow[r, "\Phi_2"'] & \mathcal{L}
\end{tikzcd}$$

From this a bisimulation is constructed. The crucial point is not the construction of $\mathcal{L}$ (which nevertheless is involved in its own right) but rather the existence of the semi-pullback, i.e., of the object $\mathcal{M}$ together with the morphisms Ψ_1 and Ψ_2.
This problem has been addressed in a variety of ways, varying the base category suitably:

1. Edalat [6] 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 due to the Polish descent of analytic spaces.

   The solution to this problem makes it possible to tackle the equivalence of bisimilarity and accepting 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 [2].

2. In [3] Markov processes over Polish spaces are considered, where the transition probability function is Borel measurable. The problem was transformed into a 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, i.e., has a Borel section.

The present paper refines the technique employed in [3] 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, 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 as uniform continuity is in terms of the inverse image of neighborhoods), universal measurability requires additionally the concept of completing the Borel sets through all finite measures, thus requires the additional concept of a finite measure. Hence Borel measurability is conceptually simpler and more basic.
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 [4], and it becomes easier to look into congruences for stochastic relations [5].

Organization The problem and an outline of its solution is discussed in greater detail in Section 2, 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.

Acknowledgement The author thanks Georgios Lajios for his critical comments. Professor Michael Ershov (Essen) was helpful in pointing out a way simplifying 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 diagrams in this paper were typeset using Paul Taylor's wonderful `diagrams` package.

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, i.e., 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

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

holds, 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 which is induced on the set of probabilities by ψ_i.

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

$$\begin{tikzcd} X_1 \arrow[r, "\phi_1"] \arrow[d, "K_1"'] & X \arrow[d, "K"'] & X_2 \arrow[l, "\phi_2"'] \arrow[d, "K_2"] \\ \mathbf{S}(Y_1) \arrow[r, "\mathbf{S}(\psi_1)"'] & \mathbf{S}(Y) & \mathbf{S}(Y_2) \arrow[l, "\mathbf{S}(\psi_2)"] \end{tikzcd}$$

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:

$$\begin{tikzcd} X_1 \arrow[d, "K_1"'] & A \arrow[l, "\alpha_1"'] \arrow[r, "\alpha_2"] \arrow[d, "L"'] & X_2 \arrow[d, "K_2"] \\ \mathbf{S}(Y_1) \arrow[r, "\mathbf{S}(\beta_1)"'] & \mathbf{S}(B) & \mathbf{S}(Y_2) \arrow[l, "\mathbf{S}(\beta_2)"] \end{tikzcd}$$

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 $i = 1, 2$.

Well, then. We will define

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

then A and B are Polish. Taking α_i, β_i as the projections

$$\begin{aligned} \alpha_i : A \ni \langle x_1, x_2\rangle &\mapsto x_i \in X_i, \\ \beta_i : B \ni \langle y_1, y_2\rangle &\mapsto y_i \in Y_i, \end{aligned}$$

we need to find a stochastic relation $L : A \rightsquigarrow B$ which 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

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

hence $L(x_1, x_2) \in \Gamma(x_1, x_2)$ for all $\langle x_1, x_2\rangle \in A$. We want $L : A \to \mathbf{S}(B)$ 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 $\{\langle x_1, x_2\rangle \in A \mid \Gamma(x_1, x_2) \cap C \neq \emptyset\}$ is a measurable subset of A, whenever $C \subseteq \mathbf{S}(B)$ is compact,

3. $\Gamma(x_1, x_2) \neq \emptyset$ for each and any $\langle x_1, x_2\rangle \in A$.

These properties will characterize Γ as a $\mathcal{C}-measurable$ 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 real crucial property is the third one.

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

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

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 $\psi_i^{-1}[\mathcal{B}(Y)]$. The latter is a usually a proper sub-σ-algebra of $\mathcal{B}(Y_i)$, hence we cannot guarantee offhand equality on the full Borel sets $\mathcal{B}(Y_i)$ 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 $\mathcal{B}(Y_i)$, 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, then still an extension can 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.

It is the latter way we will propose to go here because it opens up the avenue of establishing this result also for analytic spaces.

Discussion Edalat's solution can be considered constructive: it works with conditional distributions and their properties. These distributions do exist essentially due to the Radon-Nikodym Theorem which in turn leans heavily on the way the Lebesgue-Daniell integral is constructed. The essential observation is that these methods do not involve the Axiom of Choice. This is not true, however, for the more general solution proposed here. We make use of the Hahn-Banach Theorem, which depends directly upon 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 sounds probably 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 take silently the Axiom of Choice for granted.

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 are in a good condition tackling the problem along the lines sketched here in Section 5.

3 Polish Spaces, Measures and All That

This Section collects some basic facts from topology and measure theory for the reader's convenience and for later reference.
A *Polish space* $(X, \mathcal{T})$ is a topological space which is second countable, i.e., which has a countable dense subset, and which is metrizable through a complete metric, a measurable space $(X, \mathcal{A})$ is a set X with a σ-algebra $\mathcal{A}$. The *Borel sets* $\mathcal{B}(X, \mathcal{T})$ for the topology $\mathcal{T}$ is the smallest σ-algebra on X which contains $\mathcal{T}$. A *Standard Borel* space $(X, \mathcal{A})$ is a measurable space such that the σ-algebra $\mathcal{A}$ equals $\mathcal{B}(X, \mathcal{T})$ for some Polish topology $\mathcal{T}$ on X. Although the Borel sets are determined uniquely through the topology, the converse does not hold, as we will see in a short while. Given two measurable spaces $(X, \mathcal{A})$ and $(Y, \mathcal{B})$, a map $f : X \to Y$ is $\mathcal{A} - \mathcal{B}$*-measurable* whenever

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

holds, where

$$f^{-1}[\mathcal{B}] := \{f^{-1}[B] | B \in \mathcal{B}\}$$

is the set of inverse images

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

of elements of $\mathcal{B}$. Note that $f^{-1}[\mathcal{B}]$ is in any case an σ-algebra. If the σ-algebras are the Borel sets of some topologies on X and Y, resp., then a measurable map is called *Borel measurable* or simply a *Borel* map. The real numbers $\mathbb{R}$ carry always the Borel structure induced by the usual topology which will usually not be mentioned explicitly when talking about Borel maps.
The product $(X_1 \times X_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$ of two measurable spaces $(X_1, \mathcal{A}_1)$ and $(X_2, \mathcal{A}_2)$ is the Cartesian product $X_1 \times X_2$ endowed with the σ-algebra

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

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

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

$\mathcal{E} - \mathcal{A}_i$-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 $\mathcal{A}$ of $\mathcal{B}(Z)$ on X, i.e.,

$$\mathcal{A} = \mathcal{B}(X) \cap Z := \{B \cap X | B \in \mathcal{B}(Z)\}.$$

We also call the elements of $\mathcal{B}(X) \cap Z$ the Borel sets of X in a slight misuse of language. A measurable space $(X', \mathcal{A}')$ which is Borel isomorphic to $(X, \mathcal{A})$ is called an *analytic space* (a Borel isomorphism is a Borel measurable and bijective map the inverse of which is also Borel measurable).
A map $f : X \to Y$ between the topological spaces $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ is $\mathcal{T} - \mathcal{S}$*-continuous* iff the inverse image of an open set from $\mathcal{S}$ is an open set in $\mathcal{T}$. 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, resp., and the Borel sets are always understood with respect to the topology under consideration.
The following statement will be most helpful in the sequel. 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 which renders the map continuous.

Proposition 1 *Let $(X, \mathcal{T})$ be a Polish space, $(Y, \mathcal{T}')$ be a second countable metric space. If $f : X \to Y$ is a Borel measurable map, then there exists a Polish topology $\mathcal{T}_0$ on X such that $\mathcal{T}_0$ is finer than $\mathcal{T}$ (hence $\mathcal{T} \subseteq \mathcal{T}_0$), $\mathcal{T}$ and $\mathcal{T}_0$ have the same Borel sets, and f is $\mathcal{T}_0 - \mathcal{T}'$ continuous.*

Proof [12, Cor. 3.2.5, Cor. 3.2.6] □
Given two measurable spaces $(X, \mathcal{A})$ and $(Y, \mathcal{B})$, a *stochastic relation* $K : (X, \mathcal{A}) \rightsquigarrow (Y, \mathcal{B})$ is a Borel map from X to the set $\mathbf{S}(Y, \mathcal{B})$, the latter denoting the set of all subprobability measures on $(Y, \mathcal{B})$. The latter set carries the *weak*-σ-algebra.* This is the smallest σ-algebra on $\mathbf{S}(Y, \mathcal{B})$ which renders all maps $\mu \mapsto \mu(D)$ measurable, where $D \in \mathcal{B}$. Hence $K : (X, \mathcal{A}) \rightsquigarrow (Y, \mathcal{B})$ is a stochastic relation iff

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

We will deal usually with stochastic relations between Polish spaces or between analytic spaces. Accordingly, we call then (X, Y, K) a *Polish* respectively an *analytic object.*
Let Y be a metric space, then $\mathbf{S}(Y)$ is usually equipped with the topology of weak convergence. This is the smallest topology on $\mathbf{S}(Y)$ which 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 [11, Theorems II.6.3, II.6.5]. Moreover, the Borel sets for the topology of weak convergence is just the weak*-σ-algebra [9, Theorem 17.24]. If X is a Standard Borel space, then $\mathbf{S}(X)$ is also one: select a Polish topology $\mathcal{T}$ on X which induces the measurable structure, then $\mathcal{T}$ 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.
An $\mathcal{A} - \mathcal{B}$- measurable map $f : X \to Y$ between the measurable spaces $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ induces a map

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

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

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

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

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

holds, i.e., the diagram

$$\begin{tikzcd} X_1 \arrow[r, "\phi"] \arrow[d, "K_1"'] & X_2 \arrow[d, "K_2"] \\ \mathbf{S}(Y_1) \arrow[r, "\mathbf{S}(\psi)"'] & \mathbf{S}(Y_2) \end{tikzcd}$$

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 an object (X, Y, K) of $\mathfrak{Stoch}$ is dubbed *Polish* or *analytic* iff both X and Y are Polish, and analytic spaces, respectively. Accordingly we denote by $\mathfrak{PStoch}$ and by $\mathfrak{AStoch}$ the full subcategories having Polish respectively analytic objects as their objects.
Let $\mathbb{F}(X)$ be the set of all closed and non-empty subsets of the Polish space X, and call for Polish Y a relation, i.e., a set-valued map $F : X \to \mathbb{F}(Y)$ *$\mathcal{C}$-measurable* iff for a compact set $C \subseteq Y$ the weak inverse

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

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 $x \in X$. $\mathcal{C}$-measurable relations have Borel selectors:

Proposition 2 *Let X and Y be Polish spaces. Then each $\mathcal{C}$-measurable relation F has a measurable selector.*

Proof Since closed subsets of Polish spaces are complete, the assertion follows from [13, Theorem 4.2, (e)]. $\square$
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 as stochastic dynamic programming [13] and descriptive set theory [9]. Overviews are provided in [12, Chapter 5] and [7, 13].
We will need surjectivity of $\mathbf{S}(f)$, when $f : X \to Y$ is measurable and onto. When applying a selection argument we need to be sure that the set-valued map we are working takes indeed non-empty values. For these and other purposes, the concept of universal measurability is needed. Let $\mu \in \mathbf{S}(X, \mathcal{A})$ be a subprobability on the measurable space $(X, \mathcal{A})$, then $A \subseteq X$

is called *μ-measurable* iff there exist $M_1, M_2 \in \mathcal{A}$ 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 $\mathcal{M}_\mu(\mathcal{A})$. The σ-algebra $\mathcal{U}(\mathcal{A})$ of universally measurable sets is defined by

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

(in fact, one considers usually 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 $\mathcal{A}_1$-$\mathcal{A}_2$-measurable map between the measurable spaces $(X_1, \mathcal{A}_1)$ and $(X_2, \mathcal{A}_2)$, then it is well known that f is also $\mathcal{U}(\mathcal{A}_1)$-$\mathcal{U}(\mathcal{A}_2)$-measurable; the converse does not hold, and one usually cannot conclude that a map $g : X_1 \to X_2$ which is $\mathcal{U}(\mathcal{A}_1)$-$\mathcal{A}_2$-measurable is also $\mathcal{A}_1$-$\mathcal{A}_2$-measurable.

Lemma 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 [1, Theorem 3.4.3] we find a map $g : Y \to X$ such that $f \circ g = id_Y$ and g is $\mathcal{U}(\mathcal{B}(Y)) - \mathcal{U}(\mathcal{B}(X))$-measurable.
2. Let $\nu \in \mathbf{S}(Y)$, and define $\mu := \mathbf{S}(g)(\nu)$, then $\mu \in \mathbf{S}(X, \mathcal{U}(\mathcal{B}(X)))$ by construction. Restrict μ to the Borel sets on X, obtaining $\mu_0 \in \mathbf{S}(X, \mathcal{B}(X))$. Since we have for each set $B \subseteq Y$ the equality $g^{-1}\left[f^{-1}[B]\right] = B$, we see that for each $B \in \mathcal{B}(Y)$

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

holds. □

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

Sometimes second countable metric spaces are called separable, and this analogy is justified:

Lemma 2 *For a separable measurable space $(X, \mathcal{A})$ there exists a second countable metric topology $\mathcal{T}$ on X such that $\mathcal{B}(X, \mathcal{T}) = \mathcal{A}$.*

Proof The assertion follows from [12, Proposition 3.3.2 and Remark 3.3.3]. □

This innocently looking statement has some remarkable consequences for our context, as we will see in due course. Just for starters:

Corollary 1 *Let $(Y, \mathcal{B})$ be a separable measurable space. Then*

1. *The diagonal is measurable in the product, i.e.,*

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

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

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

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

Proof Apply Lemma 2 for 1 and 2. Assertion 3 follows from Lemma 1. □
We will need to make precise statements regarding the measurability of a Borel map; for easy reference, the technical statement below is recorded:

Proposition 3 *Let X be a Polish space, $(Y, \mathcal{B})$ be a separable measurable space, and assume that $g : X \to Y$ is $\mathcal{B}(X) - \mathcal{B}$-measurable and onto. If $f : X \to Y$ is $\mathcal{B}(X) - \mathcal{B}$-measurable such that f is constant on the atoms of $g^{-1}[\mathcal{B}]$, then f is $g^{-1}[\mathcal{B}] - \mathcal{B}$ -measurable.*

Proof Separability implies that $\{y\} \in \mathcal{B}$ for all $y \in Y$. The atoms of $g^{-1}[\mathcal{B}]$ 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 since they form a partition of X. Now let $B \in \mathcal{B}$ be a measurable set, then by assumption $f^{-1}[B]$ is a Borel set in X which is the union of atoms of $g^{-1}[\mathcal{B}]$. Thus the assertion follows from the Blackwell-Mackey-Theorem [12, Thm. 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 in turn will be crucial in proving non-emptiness. Since it is rather technical in nature, it is convenient to encapsulate this development into a separate section.
We will consider the category $\mathfrak{P}$ of probability spaces which has as objects tuples $(X, \mathcal{A}, \mu)$ with $\mu \in \mathbf{P}(X, \mathcal{A})$ for the measurable space $(X, \mathcal{A})$. $\psi : (X, \mathcal{A}, \mu) \to (Y, \mathcal{B}, \nu)$ is a morphism in $\mathfrak{P}$ if $\psi : X \to Y$ is a surjective and $\mathcal{A} - \mathcal{B}$-measurable map which is measure preserving, i.e., $\nu = \mathbf{P}(\mu)$ holds. $\mathfrak{P}$ contains for two objects $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ their product $(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu)$, with $\mu \otimes \nu$ as the product measure which is uniquely determined through

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

We fix for the discussion the Polish spaces X_1 and X_2 with the respective Borel sets as σ-algebras. $(Z, \mathcal{C})$ is assumed to be a separable measurable space. Denote for a measurable space $(X, \mathcal{A})$ by $\mathcal{F}(X, \mathcal{A})$ the linear space of all $\mathcal{A}$-measurable and bounded real-valued maps $g : X \to \mathbb{R}$. Note that $\mathcal{A} \subseteq \mathcal{B}$ implies that $\mathcal{F}(X, \mathcal{A})$ is a linear subspace of $\mathcal{F}(X, \mathcal{B})$.
Now let

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

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

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

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

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

The π_i are again the projections. The equality for $\mathcal{A}$ holds by Corollary 1; thus $\mathcal{A}$ is the smallest σ-algebra on S which 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 in “lifting“ this pullback so that the object $(S, \mathcal{B}(S), \mu)$ for some suitable $\mu \in \mathbf{P}(S, \mathcal{B}(S))$ stands in the upper left 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 *The semi-pullback $(*)$ in $\mathfrak{P}$ may be extended to a semi-pullback*

$$\begin{tikzcd} (S, \mathcal{B}(S), \mu) \arrow[r, "\pi_2"] \arrow[d, "\pi_1"'] & (X_2, \mathcal{B}(X_2), \mu_2) \arrow[d, "\psi_2"] \\ (X_1, \mathcal{B}(X_1), \mu_1) \arrow[r, "\psi_1"'] & (Z, \mathcal{C}, \nu) \end{tikzcd}$$

in $\mathfrak{P}$.

This entails essentially an extension process, extending $\theta \in \mathbf{P}(S, \mathcal{A})$ to a suitable $\mu \in \mathbf{P}(S, \mathcal{B}(S))$. We establish the existence of this extension in two steps. The first step 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
0. The line of attack will be as follows: we will construct a linear subspace of $\mathcal{F}(S, \mathcal{B}(S))$ which contains $\mathcal{F}(S, \mathcal{A})$ 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 $\mathcal{F}(S, \mathcal{B}(S))$ which then can be represented through a measure $\mu \in \mathbf{P}(S, \mathcal{B}(S))$, so that

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

holds. Clearly, μ extends θ and is the measure we are looking for.
The commutativity of the diagram entails by standard arguments from measure theory that $(i = 1, 2)$

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

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

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

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

$$\mathcal{D}_i := \{f_i \circ \pi_i | f_i \in \mathcal{F}(X_i, \mathcal{B}(X_i))\},$$

then $\mathcal{D}_i \subseteq \mathcal{F}(S, \mathcal{B}(S))$, and

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

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

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

We claim that f_1 is constant on the atoms of $\psi_1^{-1}[\mathcal{C}]$. 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 $\psi_1^{-1}[\mathcal{C}]$-measurable by Proposition 3, and consequently,

$$\begin{aligned} \int_s g \; d\theta &= \int_S f_1 \circ \pi_i \; d\theta \\ &= \int_{X_1} f_1 \; d\mu_1. \end{aligned}$$

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 $\Lambda_1 : \mathcal{F}(S, \mathcal{A}) \to \mathbb{R}$ 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 $\mathcal{F}(S, \mathcal{A}) \cup \mathcal{D}$, where $\mathcal{D} := \mathcal{D}_1 \cup \mathcal{D}_2$. This requires both functionals yielding the same value on the intersection $\mathcal{F}(S, \mathcal{A}) \cap (\mathcal{D}_1 \cup \mathcal{D}_2)$. Assume first that $g \in \mathcal{F}(S, \mathcal{A}) \cap \mathcal{D}_1$, thus $g = f_1 \circ \pi_1$ for some $f_1 \in \mathcal{F}(X_1, \mathcal{B}(X_1))$. Since g does not depend on the second component, we may infer from the definition of $\mathcal{A}$ that f_1 is even $\psi_1^{-1}[\mathcal{C}]-$ measurable, 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 argumentation for $g \in \mathcal{F}(S, \mathcal{A}) \cap \mathcal{D}_2$ is similar.
Let Λ_2 be the joint linear extension of Λ_1 on $\mathcal{F}(S, \mathcal{A})$ and of Λ_0 on $\mathcal{D}$ to the linear space spanned by $\mathcal{F}(S, \mathcal{A})$ and $\mathcal{D}$.
From the construction it is clear that $\Lambda_2(1) = 1$ holds, and that Λ_2 is monotone.
3. The Hahn-Banach Theorem for ordered linear spaces [8, Lemma IX.1.4] gives a positive linear operator

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

that extends Λ_2. Since each continuous and bounded map $f : X_1 \times X_2 \to \mathbb{R}$ becomes a member of $\mathcal{F}(S, \mathcal{B}(S))$ 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

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

with

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

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

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

then $\mu \in \mathbf{P}(S, \mathcal{B}(S))$ will now be shown the measure we are looking for.
4. Let $f \in \mathcal{F}(S, \mathcal{A})$, then

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

thus μ extends θ. Let $f_i \in \mathcal{F}(X_1, \mathcal{B}(X_i))$, then $f_i \circ \pi_i \in \mathcal{D}_i \subseteq \mathcal{D}$, 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.
$\square$

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. In the general case we do not have the Riesz Representation Theorem directly at our disposal, but compactness may nevertheless be capitalized upon since each Polish space may be embedded into a compact metric space as a measurable subspace.

Proof for the general case

0. The famous characterization of Polish spaces due to Alexandrov [12, 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 capitalize 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 holds also for the non-compact case.

1. X_i is a dense measurable subset of a compact metric space $\underline{X_i}$ by [9, 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 [12, Proposition 3.3.4].

Define for $B_i \in \mathcal{B}(\underline{X_i})$

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

and put

$$S_0 := \{\langle x_1, x_2\rangle \in \underline{X_1} \times \underline{X_2} | \underline{\psi}_1(x_1) = \underline{\psi}_2(x_2)\}.$$

Then

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

thus

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

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

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

Now put for $E \in \underline{\psi_1}^{-1}\left[\mathcal{C}\right] \otimes \underline{\psi_2}^{-1}\left[\mathcal{C}\right]$

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

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

2. The construction shows that

$$\begin{tikzcd} (S_0, \mathcal{A}_0, \underline{\theta}) \arrow[r, "\pi_2"] \arrow[d, "\underline{\pi_1}"'] & \left(\underline{X_2}, \mathcal{B}(\underline{X_2}), \underline{\mu_2}\right) \arrow[d, "\underline{\psi_2}"] \\ \left(\underline{X_1}, \mathcal{B}(\underline{X_1}), \underline{\mu_1}\right) \arrow[r, "\underline{\psi_1}"'] & (Z, \mathcal{C}, \nu) \end{tikzcd}$$

commutes, where

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

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

$$
\begin{tikzcd}
\left(S_0, \mathcal{B}(S_0), \underline{\mu}\right) \arrow[r, "\underline{\pi_2}"] \arrow[d, "\underline{\pi_1}"'] & \left(\underline{X_2}, \mathcal{B}(\underline{X_2}), \underline{\mu_2}\right) \arrow[d, "\underline{\psi_2}"] \\
\left(\underline{X_1}, \mathcal{B}(\underline{X_1}), \underline{\mu_1}\right) \arrow[r, "\underline{\psi_1}"'] & (Z, \mathcal{C}, \nu)
\end{tikzcd}
$$

3. We now roll back compactification. Put for the Borel set $B \subseteq S$

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

then $\mu \in \mathbf{P}\left(S, \mathcal{B}(S)\right)$, since

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

The other properties are obvious, so that we are done with the general case, too. $\square$

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

We reformulate Proposition 4 in terms of probability distributions. It states that there exists 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, it shows in particular where Edalat's work could enter the present discussion.

Proposition 5 *Let X_1, and X_2 be Polish spaces, $(Z, \mathcal{C})$ a separable measurable space, and assume that*

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

are measurable and surjective maps. Define

$$S := \{\langle x_1, x_2\rangle \in X_1 \times X_2 | \zeta_1(x_1) = \zeta_2(x_2)\},$$

and let $\nu_1 \in \mathbf{P}\left(X_1\right), \nu_2 \in \mathbf{P}\left(X_2\right), \nu \in \mathbf{P}\left(Z, \mathcal{C}\right)$ such that

$$\forall E_i \in \zeta_i^{-1}\left[\mathcal{C}\right] : \mathbf{P}\left(\pi_i\right)(\nu)(E_i) = \nu_i(E_i) \ (i = 1, 2)$$

holds, where $\pi_1 : S \to X_1, \pi_2 : S \to X_2$ are the projections; S carries the trace of the product σ-algebra. Then there exists $\mu \in \mathbf{P}\left(S\right)$ such that

$$\forall E_i \in \mathcal{B}(X_i) : \mathbf{P}\left(\pi_i\right)(\mu)(E_i) = \nu_i(E_i) \ (i = 1, 2)$$

holds.

Proof This is a diagram free representation of Proposition 4. $\square$

In important special cases, there are other ways of establishing the Proposition, as will be discussed briefly.

Remark: 1. If Z is also a Polish space, and if $\zeta_i : X_i \to Z$ are bijections, then the Blackwell-Mackey Theorem [12, Thm. 4.5.7] shows that $\zeta_i^{-1}[\mathcal{C}] = \mathcal{B}(X_i)$. In this case the given measure $\nu \in \mathbf{P}(S)$ is the desired one.
2. The maps $\zeta_i : X_i \to Z$ are morphisms in Edalat's category of probability measures on Polish spaces [6], provided Z is a Polish space. The assertion can then be deduced from tracing the development in [6, Cor. 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 one depends substantially on the theory of regular conditional probabilities on analytic spaces, so that the impression might arise that the existence of the measure in question depends on these probabilities, too. The proof for Proposition 4 shows that this is not the case, that rather a 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 3 *Let $\mathcal{A}$ be a sub-σ-algebra of the Borel sets of a Polish space X, and assume that θ is a probability measure on $\mathcal{A}$. Then θ can be extended to a probability measure on all of $\mathcal{B}(X)$.*

Proof 0. We need only sketch the proof, since the main work has already been done in the proof of Proposition 4. Although the assertion is a bit different, the pattern of the argumentation is very similar to the one presented already.
1. First X is assumed to be compact, then a combination of the Hahn-Banach-Theorem and the Riesz Representation Theorem 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 ist 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 to establish an extension to probabilistic relations.

Proposition 6 *Let X and Y be Polish spaces, assume that $\mathcal{B} \subseteq \mathcal{B}(Y)$ is a countably generated σ-algebra, and let $K_0 : (X, \mathcal{B}(X)) \rightsquigarrow (Y, \mathcal{B})$ be a stochastic relation. Then K_0 can be extended to a stochastic relation $K : (X, \mathcal{B}(X)) \rightsquigarrow (Y, \mathcal{B}(Y))$.*

Proof 0. We will construct a probability measure on the product $\mathcal{B}(Y) \otimes \mathcal{B}$, extend this measure and then obtain the desired extension to the probabilistic relation through disintegration.
1. Let μ be a probability measure on $\mathcal{B}(X)$, and define for $D \in \mathcal{B}(Y) \otimes \mathcal{B}$ the measure

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

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

Lemma 3. Since μ_1 is a measure on the product of two Polish spaces, there exists a stochastic relation $K_1 : (X, \mathcal{B}(X)) \rightsquigarrow (Y, \mathcal{B}(Y))$ such that

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

holds for all $D \in \mathcal{B}(X \times Y)$. This follows from the existence of regular conditional distributions on Polish spaces [11, Theorem V.8.1]. It is well known that K_1 is unique up to sets of μ-measure zero.
2. Let $\mathcal{B}_0 := \{B_n | n \in \mathbb{N}\}$ be a countable generator of the σ-algebra $\mathcal{B}$; we may and do assume that $\mathcal{B}_0$ is closed under finite intersections (otherwise form all finite intersections of elements of $\mathcal{B}_0$, then this is still a countable generator which is closed under finite intersections). Now let $E \in \mathcal{B}$, then

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

holds by the construction of this extension, thus there exists for each $E \in \mathcal{B}$ a set $N(E) \in \mathcal{B}(X)$ 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, then $N \in \mathcal{B}(X)$, and $\mu(N) = 0$ holds.
4. 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 $E \in \mathcal{B}$. In fact, put

$$\mathcal{E} := \{E \in \mathcal{B} | \forall x \in X \setminus N : K_0(x)(E) = K_1(x)(E)\},$$

then $\mathcal{B}_0 \subseteq \mathcal{E}$ by construction, $\mathcal{E}$ contains Y, and $\mathcal{E}$ is closed under complementation and disjoint countable unions. Thus $\mathcal{E} = \mathcal{B}$ is inferred by the $\pi - \lambda$-Theorem [9, Theorem 10.1.iii]. Now let μ_2 be an arbitrary probability measure on $\mathcal{B}(Y)$, and define the stochastic relation K by cases as follows:

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

This relation has the desired properties. □
This applies directly to making a pair of surjective and measurable maps into morphisms under a rather weak condition of separability.

Lemma 4 *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* $\phi : X \to A, \psi : Y \to B$, *exists a stochastic relation* $\mathsf{K} := (X, X, K)$ *with* $K : X \rightsquigarrow Y$ *which makes* $\mathsf{f} := (\phi, \psi)$ *a morphism* $\mathsf{f} : \mathsf{M} \to \mathsf{K}$.

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

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

then $K_0 : (X, \mathcal{B}(X)) \rightsquigarrow (Y, \psi^{-1}[\mathcal{B}])$ is a stochastic relation which can be extended to a stochastic relation $K : (X, \mathcal{B}(X)) \rightsquigarrow (Y, \mathcal{B}(Y))$ by Proposition 6. It is plain from the construction that $\mathbf{S}(\psi) \circ K = M \circ \phi$ holds. □

5 Semi-Pullbacks

We will show now that semi-pullbacks exist in a rather general setting, generalizing the construction in [3] and in [6]. This will ultimately lead to showing that semi-pullbacks exist for analytic objects, and it will be shown that the object underlying such a semi-pullback is Polish.
The central Lemma reads as follows:

Lemma 5 *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*

$$\begin{tikzcd} & \mathsf{K}_1 \arrow[d, "\mathsf{f}_1"] \\ \mathsf{K}_2 \arrow[r, "\mathsf{f}_2"] & \mathsf{K} \end{tikzcd}$$

has a semi-pullback

$$\begin{tikzcd} \mathsf{M} \arrow[r, "\mathsf{g}_1"] \arrow[d, "\mathsf{g}_2"'] & \mathsf{K}_1 \arrow[d, "\mathsf{f}_1"] \\ \mathsf{K}_2 \arrow[r, "\mathsf{f}_2"'] & \mathsf{K} \end{tikzcd}$$

with a Polish object M.

Proof 1. Assume $\mathsf{K}_i = (X_i, Y_i, K_i)$. In view of Lemma 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 1 we may assume that the respective σ-algebras on X_1 and X_2 are obtained from Polish topologies which render φ_1 and K_1 as well as φ_2 and K_2 continuous. These topologies are fixed for the proof. Put

$$\begin{aligned} A &:= \{\langle x_1, x_2 \rangle \in X_1 \times X_2 | \varphi_1(x_1) = \varphi_2(x_2)\}, \\ B &:= \{\langle y_1, y_2 \rangle \in Y_1 \times Y_2 | \psi_1(y_1) = \psi_2(y_2)\}, \end{aligned}$$

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

$$\begin{tikzcd} X_1 \arrow[r, "\varphi_1"] \arrow[d, "K_1"'] & X \arrow[d, "K"] & X_2 \arrow[l, "\varphi_2"'] \arrow[d, "K_2"] \\ \mathbf{S}(Y_1) \arrow[r, "\mathbf{S}(\psi_1)"'] & \mathbf{S}(Y) & \mathbf{S}(Y_2) \arrow[l, "\mathbf{S}(\psi_2)"] \end{tikzcd}$$

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

$$\begin{aligned} K(\varphi_1(x_1)) &= \mathbf{S}(\psi_1)(K_1(x_1)) \\ K(\varphi_2(x_2)) &= \mathbf{S}(\psi_2)(K_2(x_2)) \end{aligned}$$

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 $\langle x_1, x_2 \rangle \in A$. Separability of the target spaces now enters: Corollary 1 shows that the image of a surjective map under $\mathbf{S}$ is onto again, so that there exists $\mu \in \mathbf{S}(B)$ with

$$\mathbf{S}(\psi_1 \circ \beta_1)(\mu) = K(\varphi_1(x_1)),$$

consequently,

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

But this means

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

Put

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

then Proposition 5 shows that $\Gamma(x_1, x_2) \neq \emptyset$.
3. Since K_1 and K_2 are continuous, $\Gamma : A \to \mathbb{F}(\mathbf{S}(B))$ is easily established. The set

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

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 $i = 1, 2$, thus $\langle x_1, x_2 \rangle \in A$. 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 $\mathbf{S}(\beta_i)(\mu) = K_i(x_i)$, consequently $\langle x_1, x_2 \rangle \in \exists\Gamma(C)$, thus this set is closed, hence measurable.
4. From Proposition 2 it is now inferred 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 stochastic relation with

$$\begin{aligned} K_1 \circ \alpha_1 &= \mathbf{S}(\beta_1) \circ M, \\ K_2 \circ \alpha_2 &= \mathbf{S}(\beta_2) \circ M \end{aligned}$$

Hence $\mathsf{M} := (A, B, M)$ is the desired semi-pullback. □
This statement includes several interesting special cases:

Theorem 1 *Semi-pullbacks exist in the category $\mathfrak{PStoch}$ of stochastic relations over Polish spaces.*

Proof This follows immediately from Lemma 5. □

Corollary 2 *Suppose that in the diagram of Lemma 5 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. □
Now suppose that we have an analytic object M then we can find a Polish object K and a morphism $\mathsf{g} : \mathsf{K} \to \mathsf{M}$ by Lemma 4. This is so since analytic spaces are surjective images of Polish spaces under measurable maps, and since analytic spaces are — as measurable spaces — separable. This observation has as an interesting consequence that semi-pullbacks do exist for analytic spaces:

Corollary 3 *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*

$$\begin{tikzcd} \mathsf{M}_1 \arrow[r, "\mathsf{f}_1"] & \mathsf{M} & \mathsf{M}_2 \arrow[l, "\mathsf{f}_2"'] \end{tikzcd}$$

there exist both a Polish object K and morphisms

$$\begin{tikzcd} \mathsf{M}_1 & \mathsf{K} \arrow[l, "\mathsf{g}_1"'] \arrow[r, "\mathsf{g}_2"] & \mathsf{M}_2 \end{tikzcd}$$

forming a semi-pullback.

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

$$\begin{tikzcd} \mathsf{K}_1 \arrow[d, "\mathsf{h}_1"'] & & \mathsf{K}_2 \arrow[d, "\mathsf{h}_2"] \\ \mathsf{M}_1 \arrow[r, "\mathsf{f}_1"] & \mathsf{M} & \mathsf{M}_2 \arrow[l, "\mathsf{f}_2"'] \end{tikzcd}$$

Now, using Lemma 5, find a semi-pullback

$$\begin{tikzcd} \mathsf{K}_1 & \mathsf{K} \arrow[l, "\mathsf{t}_1"'] \arrow[r, "\mathsf{t}_2"] & \mathsf{K}_2 \end{tikzcd}$$

for the diagram

$$\begin{tikzcd} \mathsf{K}_1 \arrow[r, "\mathsf{h}_1 \circ \mathsf{f}_1"] & \mathsf{M} & \mathsf{K}_2 \arrow[l, "\mathsf{h}_2 \circ \mathsf{f}_2"'] \end{tikzcd}$$

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:

Theorem 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 the following holds: whenever $r' : d' \to a, s' : d' \to c$ forms a commutative diagram with f and g (i.e., $f \circ r' = g \circ s'$ holds), then there exists a morphism $h : d' \to d$ with $r' = r \circ h, s' = s \circ h$. If morphism h is uniquely determined, then d together with r and s is called a pullback.
The category $\mathfrak{Po}$ is the full subcategory of $\mathfrak{P}$ (cf. Section 4) which has probability spaces based on Polish spaces as objects.

Proposition 7 *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 equalizers, 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 [10, 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. $\square$
Because $\mathfrak{Po}$ is a full subcategory of $\mathfrak{PStoch}$, we may conclude

Corollary 4 *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 [6] that works with the conditional distribution under somewhat stronger assumptions.
The result proposed here find applications e.g. 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 accepts exactly the same formulas. This probabilistic version of the Hennessy-Milner Theorem is a generalization of the well known characterization of bisimulation through a simple negation-free modal logic with a countable number of diamonds, which was established in [2, 3]. 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 factorization of smooth equivalence relations that occur quite naturally for stochastic relations. The existence of semi-pullbacks helps in investigating these congruences, see [5] for first results.

References

[1] W. Arveson. *An Invitation to C*-Algebra*, volume 39 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1976.

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

[3] E.-E. Doberkat. Semi-pullbacks and bisimulations in categories of stochastic relations. In J. C. M. Baeten, J. K. Lenstra, J. Parrow, and G. J. Woeginger, editors, *Proc. ICALP'03*, volume 2719 of *Lecture Notes in Computer Science*, pages 996 – 1007, Berlin, July 2003. Springer-Verlag.

[4] E.-E. Doberkat. Stochastic relations interpreting modal logic. Technical Report 144, Chair for Software-Technology, University of Dortmund, October 2003.

[5] E.-E. Doberkat. Factoring stochastic relations. Technical Report 147, Chair for Software-Technology, University of Dortmund, January 2004.

[6] A. Edalat. Semi-pullbacks and bisimulation in categories of Markov processes. *Math. Struct. in Comp. Science*, 9(5):523 – 543, 1999.

[7] C. J. Himmelberg. Measurable relations. *Fund. Math.*, 87:53 – 72, 1975.

[8] K. Jacobs. *Measure and Integral*. Academic Press, New York, 1978.

[9] A. S. Kechris. *Classical Descriptive Set Theory*. Number 156 in Graduate Texts in Mathematics. Springer-Verlag, Berlin, Heidelberg, New York, 1994.

[10] S. Mac Lane. *Categories for the Working Mathematician*. Number 5 in Graduate Texts in Mathematics. Springer-Verlag, Berlin, 2 edition, 1997.

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

[12] S. M. Srivastava. *A Course on Borel Sets*. Number 180 in Graduate Texts in Mathematics. Springer-Verlag, Berlin, 1998.

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