UNIVERSITÄT DORTMUND
Fachbereich Informatik
Lehrstuhl für Software-Technologie

MEMO Nr. 140

Congruences for Stochastic Relations

Ernst-Erich Doberkat

Juli 2003

Internes Memorandum des
Lehrstuhls für Software-Technologie
Prof. Dr. Ernst-Erich Doberkat
Fachbereich Informatik
Universität Dortmund
Baroper Straße 301

D-44227 Dortmund

ISSN 0933-7725

Congruences for Stochastic Relations

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

July 7, 2003

Abstract

We discuss congruences for stochastic relations, stressing the equivalence of smooth equivalence relations and countably generated σ-algebras. Factor spaces are constructed for congruences and for morphisms. Semi-pullbacks are needed when investigating the interplay between congruences and bisimulations, and it is shown that semi-pullbacks exist for stochastic relations over analytic spaces, generalizing a previous result and answering an open question. Equivalent congruences are investigated, and it is shown that stochastic relations that have equivalent congruences are bisimilar. The well-known equivalence relation coming from a Hennessy-Milner logic for labelled Markov transition systems is shown to be a special case in this development.

Keywords: Stochastic relations over Polish and analytic spaces, bisimulation, semi-pullbacks, congruences, factorization.

1 Introduction

The characterization of bisimulation for labelled Markov transition systems through acceptance of the same formulas in a simple negation free Hennessy-Milner logic in [4] and subsequently in [10] used as an essential property the fact that the underlying equivalence relation is countably generated. Take two such transition systems over the state spaces S and S', resp., then the states $s \in S$ and $s' \in S'$ are equivalent iff

$$\forall \varphi \in \Phi : s \models \varphi \Leftrightarrow s' \models \varphi$$

Here Φ is the countable set of all formulas for the logic. Satisfaction of formula φ translates to the Borel set $\{s \in S | s \models \varphi\}$ of states that satisfy φ, and these sets in turn generate the σ-algebra $\mathcal{C}(\Phi, S)$ of measurable sets which are invariant under the equivalence relation just mentioned. $\mathcal{C}(\Phi, S)$ is thus a countably generated sub-σ-algebra of the Borel sets of the state space, the atoms of which are just the equivalence classes for the equivalence relation; it is related to the transition system since equivalent states display the same behavior on each of its members.

Searching a suitable model for characterizing equivalent behavior for stochastic relations, it turns out that those properties of the equivalence relation that are not too closely tied to a labelled Markov transition system may be used to formulate a congruence. An additional property has to be taken into account: while a Markov system operates on a state space, a stochastic relation operates between two spaces: a Markov process may be characterized through a family $(k_a)_{a \in \mathsf{L}}$ of transition probabilities $k_a : S \rightsquigarrow S$, and an equivalence needs only be formulated on S characterizing changes of state, a stochastic relation $K : X \rightsquigarrow Y$ is a transition probability between an input space X and an output space Y, so the counterpart of such an equivalence relation ought to characterize equivalent input behavior reflected by equivalent output behavior. Thus technically two equivalences are needed.

This paper proposes the characterization of congruences for stochastic relations through two equivalence relations. Roughly speaking, equivalent inputs lead to equivalent outputs. While the equivalence of inputs is easily formulated, the equivalence of outputs is more difficult to capture, since it is usually not sensible to assign a probability to single points of the output space. This is so since there may be uncountably many of them. Here the characterization of equivalences through countably generated sub-σ-algebras comes in handy, since equivalent behavior can then be modelled through being assigned the same probability on the corresponding σ-algebra. This leads to the description of a congruence for a stochastic relation $K : X \rightsquigarrow Y$ as a pair $\langle \approx, \mathcal{A} \rangle$ consisting of a countably equivalence relation $\approx$ on the input space X and a countably generated σ-algebra $\mathcal{A}$ on the output space so that

$$x \approx x' \Rightarrow \forall A \in \mathcal{A} : K(x)(A) = K(x')(A)$$

holds. Thus the congruence models identical behavior on a well-defined subset of the universe to be described. We can characterize congruences more uniformly through a pair of countably generated σ-algebras by providing transfer statements between these two forms for the representation of a countably generated σ-algebra. The case of Markov transition systems is thus captured as a special case.

This paper studies these congruences in order to provide some algebraic background to the study of stochastic relations. These relations are the counterpart to their better known cousins, nondeterministic relations, with which the share some properties. It is well known, for example, that nondeterministic relations are closely related to the Kleisli construction for the powerset monad in the category of sets; Giry's work shows that a similar construction can be carried out for the functor that assigns each measurable space the set of all probability measures. In both cases the relational product is the Kleisli product associated with the corresponding monad [12, 17, 8]. Probabilistic relations have a converse [1, 9] in much the same way nondeterministic ones have one; a demonic product can be formulated for them [7] with similar properties as for their nondeterministic counterpart [6], and finally their similarity is so strong that a common monadic formulation can be found for a rather popular software architecture [9].

On the algebraic side, however, their theory is not that elaborate as in the nondeterministic case, as a mere glimpse at the survey [3] on relation algebras indicates. Constructions which can be carried out readily for nondeterministic relations turn out to be rather cumbersome for the stochastic case. This is indicated by the amount of Measure Theory necessary for constructing the converse of a relation, and it becomes evident when it comes to think about semi-pullbacks. Recall that in a category the semi-pullback of a pair of morphisms $f : a \to b, g : c \to b$ with the same target is a pair of morphisms $p : s \to a, q : s \to c$ with a common domain which makes the diagram

$$\begin{array}{ccc} s & \xrightarrow{q} & c \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle g} \\ a & \xrightarrow[f]{} & b \end{array}$$

commutative. It can be shown that a pullback exists in the category of nondeterministic relations (which is a much stronger property than a mere semi-pullback), and it can also be shown that semi-pullbacks exists in the category of stochastic relations, provided the base spaces are Polish, i.e., are second countable topological spaces which can be metrized through a complete metric. But it can also be shown that this does not go much further algebraically: an example shows that weak pullbacks (and hence pullbacks) do not exist in this category [10]. Bisimulations and congruences are closely related also for stochastic relations, as we know from [4, 10]. The links between them is provided through semi-pullbacks, so that a study of congruences suggests also studying this weak form of pullback. Because the factor space of a congruence usually is an analytic space, i.e., the image of a Polish space under a measurable map, is seems worthwhile to study pullbacks also for analytic spaces. Generalizing a previous result we can show that semi-pullbacks also exist in the category of stochastic relations over analytic spaces. This result seems to be new, the technique for obtaining it is quite closely related to the one proposed in [10] which uses the existence of measurable selectors for set valued maps. Given this result, we can show that stochastic relations for which equivalent congruences exist are bisimilar. Congruences are equivalent if they essentially describe the same behavior, essentially meaning that the corresponding σ-algebras are isomorphic.

Overview The paper is organized as follows: we give the necessary background from Measure Theory in Section 2, Section 3 shows how to extend a particular semi-pullback of measures. This is a central technical result for establishing the existence of semi-pullbacks for categories of stochastic relations over a variety of base spaces, most notably over Polish and over analytic spaces in Section 6. Semi-pullbacks require the notion of morphism which is introduced and studied in Section 5; this Section introduces also bisimilarity and studies congruences together with morphisms. Congruences are introduced in Section 4, here also the relations between congruences and countably generated σ-algebras emanating mainly from Blackwell's Theorem are formulated. It is shown in particular that there is an anti-isomorphism between smooth equivalence relations and these σ-algebras. Section 6 also contains construction providing a bisimulation for equivalent congruences.

Acknowledgement The author thanks Georgios Lajios for his critical comments. Michael Ershov (Essen) was helpful in pointing out a way simplifying the approach to the measure extension in Section 3, 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 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 [20, 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 $\langle X, Y, K \rangle$ 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 [18, Theorems II.6.3, II.6.5]. Moreover, the Borel sets for the topology of weak convergence is just the weak*-σ-algebra [16, 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 $(X, \mathcal{A})$.

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 [21, 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 as stochastic dynamic programming [21] and descriptive set theory [16]. Overviews are provided in [20, Chapter 5] and [13, 21].

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 establishing this, 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 [2, 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:

Observation 1 *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 [20, Proposition 3.3.2 and Remark 3.3.3]. □

This innocent 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 Observation 1 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 [20, Thm. 4.5.7]. □

3 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 $\langle X, \mathcal{A}, \mu\rangle$ with $\mu \in \mathbf{P}(X, \mathcal{A})$ for the measurable space $(X, \mathcal{A})$. $\psi : \langle X, \mathcal{A}, \mu\rangle \to \langle Y, \mathcal{B}, \nu\rangle$ 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 $\langle X, \mathcal{A}, \mu\rangle$ and $\langle Y, \mathcal{B}, \nu\rangle$ their product $\langle X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu\rangle$, 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

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

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

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

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 Observation 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 $\langle S, \mathcal{B}(S), \mu\rangle$ 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*

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

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 [14, 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.
□
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 [20, 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 fr the non-compact case.
1. X_i is a dense measurable subset of a compact metric space ${}^{\Box}X_i$ by [16, Theorem 4.14], and $\psi_i : X_i \to Z$ may be extended to a Borel measurable map ${}^{\Box}\psi_i : {}^{\Box}X_i \to Z$ by [20, Proposition 3.3.4].
Define for $B_i \in \mathcal{B}({}^{\Box}X_i)$

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

and put

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

Then

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

thus

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

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

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

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

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

then ${}^\square\theta(S_0 \setminus S) = 0$, hence ${}^\square\theta$ is concentrated on S.
2. The construction shows that

$$\begin{array}{ccc}
\langle S_0, \mathcal{A}_0, {}^\square\theta\rangle & \xrightarrow{{}^\square\pi_2} & \langle {}^\square X_2, \mathcal{B}({}^\square X_2), {}^\square\mu_2\rangle \\
{\scriptstyle {}^\square\pi_1}\downarrow & & \downarrow{\scriptstyle {}^\square\psi_2} \\
\langle {}^\square X_1, \mathcal{B}({}^\square X_1), {}^\square\mu_1\rangle & \xrightarrow[{}^\square\psi_1]{} & \langle Z, \mathcal{C}, \nu\rangle
\end{array}$$

commutes, where

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

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

$$\begin{array}{ccc}
\langle S_0, \mathcal{B}(S_0), {}^\square\mu\rangle & \xrightarrow{{}^\square\pi_2} & \langle {}^\square X_2, \mathcal{B}({}^\square X_2), {}^\square\mu_2\rangle \\
{\scriptstyle {}^\square\pi_1}\downarrow & & \downarrow{\scriptstyle {}^\square\psi_2} \\
\langle {}^\square X_1, \mathcal{B}({}^\square X_1), {}^\square\mu_1\rangle & \xrightarrow[{}^\square\psi_1]{} & \langle Z, \mathcal{C}, \nu\rangle
\end{array}$$

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

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

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

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

The other properties are obvious, so that we are done with the general case, too. □
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 ${}^\square\mu$ extends ${}^\square\theta$, and for which the incorporation of $\mathcal{F}(S, \mathcal{A})$ 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 1, 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 $\nu_1 \in \mathbf{P}(X_1), \nu_2 \in \mathbf{P}(X_2), \nu \in \mathbf{P}(Z, \mathcal{C})$ such that

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

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 $\mu \in \mathbf{P}(S)$ such that

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

holds.

Proof This is a diagram free representation of Proposition 4. □
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 [20, 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 [11], provided Z is a Polish space. The assertion can then be deduced from tracing the development in [11, 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 very heavily 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. 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 2 *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. $\square$
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 $\mathcal{B} \otimes \mathcal{B}(Y)$, 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 $D \in \mathcal{B} \otimes \mathcal{B}(Y)$ 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 $D \in \mathcal{B} \otimes \mathcal{B}(Y)$. Let μ_1 be an extension of μ_0 to all of $\mathcal{B}(X \times Y)$. This extension exists by Lemma 2. 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 [18, 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 [16, 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. □

4 Smooth Equivalence Relations

We will begin in this section a discussion of smooth equivalence relations that will lead us to the definition of congruences and of factor objects. In order to keep a probabilistic grip on the equivalence relations, we require them to be compatible with the Borel structure. The natural way to do this is to have a countable set of Borel measures generate them; we will show that the equivalence relation coming from the Hennessy-Milner logic, in which two states are equivalent iff they satisfy exactly the same formulas has this property. When we want to compare equivalence relations on different spaces, it turns out that the associated σ-algebra is very versatile. This leads to the definition of equivalent congruences, which will be studied in Section 6 from the point of bisimulations, where we show also that the equivalence between bisimulation and satisfaction the same formulas may be treated as a special case.
We fix for this section a Polish space X with its Borel σ-algebra $\mathcal{B}(X)$.

Definition 1 *An equivalence relation $\rho \subseteq X \times X$ is called* smooth *iff one of the following equivalent conditions is satisfied:*

1. *there exists a Polish space Y and a Borel measurable map $f : X \to Y$ such that*

$$x \rho y \Leftrightarrow f(x) = f(y),$$

2. *there exists a sequence $(A_n)_{n\in\mathbb{N}}$ of Borel sets in X such that*

$$x \rho y \Leftrightarrow \forall n \in \mathbb{N} : [x \in A_n \Leftrightarrow y \in A_n]\,.$$

It follows immediately that a smooth equivalence relation is a Borel subset of $X \times X$, see [20, Exercise 5.1.10]. The equivalence classes can be expressed in terms of (A_n):

$$[x]_\rho = \bigcap\{A_n | x \in A_n\} \cap \bigcap\{X \setminus A_n | x \notin A_n\},$$

hence each class is a Borel subset of X.
Smooth relations arise naturally in the context of labelled Markov transition systems and the Hennessy-Milner logic:

Example 1 Let $\langle S, (k_a)_{a\in\mathsf{L}}\rangle$ be a labelled Markov transition system, i.e., for each $k_a : S \rightsquigarrow S$ is a stochastic relation for each $a \in \mathsf{L}$, where L is an at most countable set of actions, and S is a state space. S is assumed to be Polish.

Syntax and semantics of the Hennessy-Milner logic $\mathcal{L}$ are defined as usual (cf. [4, 10]). The syntax is given by

$$\top \mid \phi_1 \wedge \phi_2 \mid \langle a \rangle_q \phi$$

Here $a \in \mathsf{L}$ is an action, and q is a rational number. Satisfaction of a state s for a formula ϕ is defined inductively. This is trivial for $\top$ and for formulas of the form $\phi_1 \wedge \phi_2$. The more complicated case is making an a-move: $s \models \langle a \rangle_q \phi$ holds iff we can find a measurable set $A \subseteq S$ such that $\forall s' \in A : s' \models \phi$ and $k_a(s, A) \geq q$ both hold. Intuitively, we can make an a-move in a state s to a state that satisfies ϕ with probability greater than q.
Denote by Φ the set of all formulas, and put $[\![\varphi]\!] := \{s \in S | s \models \phi\}$ as usual as the set of states that satisfy formula ϕ.
Two states are equivalent iff both satisfy exactly the same formulas. Thus the equivalence $\approx$ is defined through

$$s \approx s' \text{ iff } \forall \phi \in \Phi : \left[s \models \phi \Leftrightarrow s' \models \phi\right].$$

Since $s \models \phi$ iff $s \in [\![\varphi]\!]$, and since Φ is countable, we see that $\approx$ is smooth. —

A set $A \subseteq X$ is called *ρ-invariant* iff $x \in A$ and $x \rho y$ implies $y \in A$, thus A is ρ-invariant iff

$$A = \bigcup \{[x]_\rho \, | x \in A\}$$

holds. The ρ-invariant subsets of X form a σ-algebra $\mathcal{A}_\rho$. If ρ is smooth, then this σ-algebra is generated by $(A_n)_{n \in \mathbb{N}}$, when the latter sequence determines ρ. In fact, it can be said more:

Lemma 3 *Let ρ be a smooth equivalence relation. Then the Borel set $A \subseteq X$ is ρ-invariant iff $A \in \mathcal{A}_\rho$.*

Proof [16, Exercise 14.16]. □
This Lemma implies that a smooth relation does not depend on the specific sequence of generators (as Definition 1 and the representation of the equivalence classes seem to suggest) but rather on the σ-algebra induced by them. It has moreover the pleasant consequence that it permits the identification of smooth equivalence relations and countably generated sub-σ-algebras of $\mathcal{B}(X)$:

Lemma 4 *$\rho \mapsto \mathcal{A}_\rho$ is a order anti-isomorphism between the set $\mathcal{M}_X$ of smooth equivalence relations on X and set $\mathcal{Z}_X$ of all the countably generated sub-σ-algebras of $\mathcal{B}(X)$, where the order is given in each case through inclusion. $\mathcal{M}_X$ has in particular a smallest and a largest element, and is closed under countable intersections.*

Proof Let $\rho, \tau \in \mathcal{M}_X$ with $\rho \subseteq \tau$. Then

$$[x]_\tau = \bigcup \{[y]_\rho \, | y \tau x\},$$

thus each τ-invariant subset is also ρ-invariant, hence $\mathcal{A}_\tau \subseteq \mathcal{A}_\rho$ is inferred. From this observation the assertions follow easily. □
As a consequence of Lemma 4 we see that whenever $(\rho_n)_{n \in \mathbb{N}}$ is a sequence of smooth equivalence relations, then $\bigcap_{n \in \mathbb{N}} \rho_n$ is also smooth (which is not difficult to establish directly), and that

$$\mathcal{A}_{\bigcap_{n \in \mathbb{N}} \rho_n} = \sigma \left(\bigcup_{n \in \mathbb{N}} \mathcal{A}_{\rho_n} \right)$$

holds.
The atoms of $\mathcal{A}_\rho$ for a smooth equivalence relation ρ are the equivalence classes $[x]_\rho$ (recall that an atom A in $\mathcal{A}_\rho$ has the property that $A \neq \emptyset$, and that each subset of A in $\mathcal{A}_\rho$ is either empty or equals A): let $\emptyset \neq B \subseteq [x]_\rho$ and $B \in \mathcal{A}_\rho$. Thus B is ρ-invariant by Lemma 3, and we see that $y \in B$ implies $y \rho x$, hence $[x]_\rho = [y]_\rho \subseteq B$.
We will use Lemma 4 for switching between smooth equivalence relations and countably generated sub-σ-algebras of $\mathcal{B}(X)$. Denote for $\mathcal{D} \in \mathcal{Z}_X$ by $\rho_\mathcal{D}$ the smooth equivalence generated by $\mathcal{D}$.

Definition 2 *Let $K : X \rightsquigarrow Y$ be a stochastic relation, where Y is a Polish space. Then $\langle \mathcal{C}, \mathcal{D}\rangle$ is called a* congruence *for $\langle X, Y, K\rangle$ iff*

1. *$\mathcal{C}$ and $\mathcal{D}$ are countably generated sub-σ-algebras of the Borel sets on X, and Y, resp.*
2. *$x_1 \rho_\mathcal{C} x_2$ implies that $K(x_1)(D) = K(x_2)(D)$ holds for all $D \in \mathcal{D}$.*

Intuitively, a congruence is an equivalence relation on the inputs and a σ-algebra on the outputs such that equivalent inputs spawn the same sub-probability on the σ-algebra. Let $(D_n)_{n\in\mathbb{N}}$ be a sequence of Borel sets generating $\mathcal{D}$, and define

$$F : \begin{cases} X & \to [0,1]^\mathbb{N} \\ x & \mapsto (K(x)(D_n))_{n\in\mathbb{N}}, \end{cases}$$

then $\langle \mathcal{C}, \mathcal{D}\rangle$ is a congruence for $\langle X, Y, K\rangle$ iff $\rho_\mathcal{C} \subseteq \rho_F$, where ρ_F is the smooth equivalence relation induced by F; an equivalent characterization is to say that $\mathcal{D}_F \subseteq \mathcal{C}$ with $\mathcal{D}_F$ as the σ-algebra induced by ρ_F.
It may be noted that for each $D \in \mathcal{D}$ the map $x \mapsto K(x)(D)$ is actually $\mathcal{C} - \mathcal{B}(\mathbb{R})$- measurable. In fact, if $S \in \mathcal{B}(\mathbb{R})$, then

$$A_0 := \{x \in X | K(x)(D) \in S\}$$

is a Borel set in X, and it is $\rho_\mathcal{C}$-invariant by the definition of a congruence. Using Lemma 3 we see now that $A_0 \in \mathcal{C}$.
Fix a Polish object $\langle X, Y, K\rangle$ and a congruence $\langle \mathcal{C}, \mathcal{D}\rangle$ for it. We will show now how to factor $\langle X, Y, K\rangle$ by $\langle \mathcal{C}, \mathcal{D}\rangle$. In constructing this factor objects, we follow essentially the construction outlined in the proof of Proposition 9.4 in [4]. The factor space $X_\mathcal{C}$ is the set of all equivalence classes $[x]_\mathcal{C}$ and is equipped with the largest σ-algebra $\mathcal{A}_\mathcal{C}$ that makes the natural projection $\eta_\mathcal{C} : x \mapsto [x]_\mathcal{C}$ measurable. Then $(X_\mathcal{C}, \mathcal{A}_\mathcal{C})$ is an analytic space [2, Corollary 3.3.5.2]. It turns out that the Borel sets of $X_\mathcal{C}$ are generated by

$$\mathcal{C}_1 := \eta_\mathcal{C}[\mathcal{C}_0],$$

where $\mathcal{C}_0 := \{C_n | n \in \mathbb{N}\}$ is a countable generator for $\mathcal{C}$. First, each $\eta_\mathcal{C}[C_n]$ is an analytic set, since $\eta_\mathcal{C}$ is measurable. On the other hand, $X_\mathcal{C} \setminus \eta_\mathcal{C}[C_n] = \eta_\mathcal{C}[X \setminus C_n]$: Since if $x \notin C_n$, the $\rho_\mathcal{C}$-invariance of C_n implies that $[x]_\mathcal{C} \cap C_n = \emptyset$, thus the non-trivial inclusion is established. Consequently, $\eta_\mathcal{C}[C_n]$ is also coanalytic, hence is a Borel set by Souslin's Theorem [20, Theorem 4.4.3]. By the definition of $\rho_\mathcal{C}$, $\mathcal{C}_1$ separates points, for, if $[x]_\mathcal{C} \neq [x']_\mathcal{C}$ we can find C_n with, say , $x \in C_n, x' \notin C_n$. Consequently, $[x]_\mathcal{C} \in \eta_\mathcal{C}[C_n], [x']_\mathcal{C} \notin \eta_\mathcal{C}[C_n]$. The Unique Structure Theorem [2, Theorem 3.3.5] now implies that $\sigma(\mathcal{C}_1)$ equals the Borel sets of $X_\mathcal{C}$.

We note for later use that

$$\eta_{\mathcal{C}}[C_n \cap C_m] = \eta_{\mathcal{C}}[C_n] \cap \eta_{\mathcal{C}}[C_m]$$

holds. Take $[x]_{\mathcal{C}} \in \eta_{\mathcal{C}}[C_n] \cap \eta_{\mathcal{C}}[C_m]$ thus there exist $x_1 \in C_n, x_2 \in C_m$ such that $x\rho_{\mathcal{C}}x_1, x\rho_{\mathcal{C}}x_2$ The invariance of the elements of $\mathcal{C}$ yields then $[x]_{\mathcal{C}} \subseteq C_n \cap C_m$, which in turn implies the non-trivial inclusion.
The same construction can be carried out for $Y_{\mathcal{D}}$.
Assume that the generator $\mathcal{D}_0 := \{D_n | n \in \mathbb{N}\}$ for $\mathcal{D}$ is closed under finite intersections. Define

$$K_{\mathcal{C},\mathcal{D}}([x]_{\mathcal{C}})(\eta_{\mathcal{D}}[D_n]) := K(x)(D_n),$$

then $K_{\mathcal{C},\mathcal{D}}$ is well-defined on $X_{\mathcal{C}} \times \mathcal{D}_0$. This follows from the assumption that $\langle \mathcal{C}, \mathcal{D}\rangle$ is a congruence for $\langle X, Y, K\rangle$. Since $\eta_{\mathcal{D}}[\mathcal{D}_0]$ generates the Borel sets on $Y_{\mathcal{D}}$, and since this generator is closed under finite intersections, $K_{\mathcal{C},\mathcal{D}}([x]_{\mathcal{C}})$ extends unique to a finite measure on the Borel sets on $Y_{\mathcal{D}}$. Moreover, for each Borel set $D \subseteq Y_{\mathcal{D}}$ the map $q \mapsto K_{\mathcal{C},\mathcal{D}}(q)(D)$ is $\mathcal{A}_{\mathcal{C}} - \mathcal{A}_{\mathcal{D}}$-measurable. This is so since the set of all Borel sets for which this statement is true contains $\mathcal{D}_0$ by construction (and by the remarks following Definition 2) forms a σ-algebra. We can say even a little bit more. Because $\eta_{\mathcal{D}}^{-1}[\eta_{\mathcal{D}}[D]] = D$ holds for each $D \in \mathcal{D}$, we find that

$$K(x)(\eta_{\mathcal{D}}^{-1}[E]) = K_{\mathcal{C},\mathcal{D}}(\eta_{\mathcal{C}}(x))(E)$$

holds for each Borel set $E \subseteq Y_{\mathcal{D}}$, rendering the diagram

$$\begin{array}{ccc} X & \xrightarrow{\eta_{\mathcal{C}}} & X_{\mathcal{C}} \\ {\scriptstyle K}\downarrow & & \downarrow{\scriptstyle K_{\mathcal{C},\mathcal{D}}} \\ \mathbf{S}(Y) & \xrightarrow[\mathbf{S}(\eta_{\mathcal{D}})]{} & \mathbf{S}(Y_{\mathcal{D}}) \end{array}$$

commutative.

Definition 3 *Endow $X_{\mathcal{C}}$ and $Y_{\mathcal{D}}$ with the σ-algebras $\mathcal{A}_{\mathcal{C}}$ and $\mathcal{A}_{\mathcal{D}}$, respectively, and denote these analytic spaces for short by $X_{\mathcal{C}}$ and $Y_{\mathcal{D}}$, resp. The analytic object*

$$\langle X, Y, K\rangle/_{\mathcal{C},\mathcal{D}} := \langle X_{\mathcal{C}}, Y_{\mathcal{D}}, K_{\mathcal{C},\mathcal{D}}\rangle$$

is called the factor object *for $\langle X, Y, K\rangle$ and the congruence $\langle \mathcal{C}, \mathcal{D}\rangle$.*

Another example for congruences and factor spaces of interest is furnished through equivalent congruences.
As a preparation we will have a quick look at how the atoms of a countably generated σ-algebra are characterized through the generators.

Lemma 5 *Let $\mathcal{E} = \sigma(\{E_n | n \in \mathbb{N}\})$ be a countably generated σ-algebra over a set E. Define $A^1 := A, A^0 := E \setminus A$ for $A \subseteq E$. Then*

$$\{\bigcap_{n\in\mathbb{N}} E_n^{\alpha(n)} | \alpha \in \{0,1\}^{\mathbb{N}}\}$$

are exactly the atoms of $\mathcal{E}$.

Proof The proof to [20, 3.1.15] establishes this representation. □

Definition 4 *Let $\mathcal{C}$ and $\mathcal{D}$ be countably generated σ-algebras over set X resp. Y, assume that $\mathcal{C}_0$ is a countable generator for $\mathcal{C}$, and that $\Theta : \mathcal{C}_0 \to \mathcal{D}$ is a map. We say that $\mathcal{C}$* spawns $\mathcal{D}$ via $(\Theta, \mathcal{C}_0)$ *iff*

1. $\{\Theta(C) | C \in \mathcal{C}_0\}$ *generates $\mathcal{D}$,*
2. $[x_1]_{\mathcal{C}} = [x_2]_{\mathcal{C}}$ *implies the equality of*

$$\bigcap\{\Theta(C) | x_1 \in C \in \mathcal{C}_0\} \cap \bigcap\{Y \setminus \Theta(C) | x_1 \notin C \in \mathcal{C}_0\}$$

and

$$\bigcap\{\Theta(C) | x_2 \in C \in \mathcal{C}_0\} \cap \bigcap\{Y \setminus \Theta(C) | x_2 \notin C \in \mathcal{C}_0\}.$$

The first condition makes sure that the information needed to generate $\mathcal{D}$ and hence the equivalence $\rho_{\mathcal{D}}$ is already contained in $\mathcal{C}$ (and consequently in $\rho_{\mathcal{C}}$) and can be transported via Θ. The second condition will permit later on a comparison between the equivalence relations $\rho_{\mathcal{C}}$ and $\rho_{\mathcal{D}}$. It permits mapping equivalence classes, since it makes sure that

$$[x]_{\mathcal{C}} \mapsto \bigcap\{\Theta(C) | x \in C \in \mathcal{C}_0\} \cap \bigcap\{Y \setminus \Theta(C) | x \notin C \in \mathcal{C}_0\}$$

is a well defined map between the atoms of $\mathcal{C}$ and $\mathcal{D}$ (which are exactly the equivalence classes for $\rho_{\mathcal{C}}$ and $\rho_{\mathcal{D}}$). We will reuse the name of the map on generators to denote this map.
This permits the definition of equivalent congruences through proportional ones:

Definition 5 *Let $\langle X, Y, K\rangle$ and $\langle X', Y', K'\rangle$ be Polish objects with congruences $\langle \mathcal{C}, \mathcal{D}\rangle$ and $\langle \mathcal{C}', \mathcal{D}'\rangle$, respectively.*

1. *Call $\langle \mathcal{C}, \mathcal{D}\rangle$* proportional to *$\langle \mathcal{C}', \mathcal{D}'\rangle$ (symbolically $\langle \mathcal{C}, \mathcal{D}\rangle \propto \langle \mathcal{C}', \mathcal{D}'\rangle$) iff there exist countable generators $\mathcal{C}_0$ for $\mathcal{C}$ and $\mathcal{D}_0$ for $\mathcal{D}$ with maps $\Upsilon : \mathcal{C}_0 \to \mathcal{C}'$ and $\Omega : \mathcal{D}_0 \to \mathcal{D}'$ such that*
   (a) $\mathcal{C}$ spawns $\mathcal{C}'$ via $(\Upsilon, \mathcal{C}_0)$, $\mathcal{D}$ spawns $\mathcal{D}'$ via $(\Omega, \mathcal{D}_0)$,
   (b) $\forall x \in X \forall x' \in \Upsilon([x]_{\mathcal{C}}) \forall D_0 \in \mathcal{D}_0 : K(x)(D_0) = K'(x')(\Omega(D_0))$ holds.
2. *Call these congruences* equivalent *iff both $\langle \mathcal{C}, \mathcal{D}\rangle \propto \langle \mathcal{C}', \mathcal{D}'\rangle$ and $\langle \mathcal{C}', \mathcal{D}'\rangle \propto \langle \mathcal{C}, \mathcal{D}\rangle$ holds.*

Thus equivalent congruences behave in exactly the same way since the same behavior is exhibited on each atom, i.e., equivalence class.
We will show now how equivalent congruences on stochastic relations give rise to a factor object built on their sum. This construction will be of use in Section 6 for showing that stochastic relations having equivalent congruences are bisimilar.
Assume that $\langle \mathcal{C}, \mathcal{D}\rangle$ and $\langle \mathcal{C}', \mathcal{D}'\rangle$ are equivalent congruences on the Polish objects $\langle X, Y, K\rangle$, and $\langle X', Y', K'\rangle$, respectively. Construct for $\langle X, Y, K\rangle$ and $\langle X', Y', K'\rangle$ the direct sum

$$\langle X, Y, K\rangle + \langle X', Y', K'\rangle := \langle X + X', Y + Y', K + K'\rangle,$$

where the only non-obvious construction is $K + K'$: put for the Borel set $E \subseteq Y + Y'$

$$(K + K')(z)(E) := \begin{cases} K(z)(E \cap X), & \text{if } z \in X \\ K'(z)(E \cap X'), & \text{if } z \in X', \end{cases}$$

then clearly $K + K' : X + X' \rightsquigarrow Y + Y'$. Define

$$\begin{aligned} \mathcal{C} + \mathcal{C}' &:= \{C + C' | C \in \mathcal{C}, C' \in \mathcal{C}'\} \\ \mathcal{D} + \mathcal{D}' &:= \{D + D' | D \in \mathcal{D}, D' \in \mathcal{D}'\}, \end{aligned}$$

then $\langle \mathcal{C} + \mathcal{C}', \mathcal{D} + \mathcal{D}' \rangle$ is a congruence on the sum $\langle X + X', Y + Y', K + K' \rangle$. Assume $\mathcal{C}$ spawns $\mathcal{C}'$ via $(\Upsilon, \{C_n | n \in \mathbb{N}\})$. One first establishes that

$$\mathcal{C} + \mathcal{C}' = \sigma(\{C_n + \Upsilon(C_n) | n \in \mathbb{N}\}).$$

This is so since $F \subseteq X + X'$ is a member of the sum $\mathcal{C} + \mathcal{C}'$ iff both $F \cap X \in \mathcal{C}$ and $F \cap X' \in \mathcal{C}'$ hold, and since $\mathcal{C}' = \sigma(\{\Upsilon(C_n) | n \in \mathbb{N}\})$ due to the properties of Υ. Similarly, $\mathcal{D} + \mathcal{D}'$ may be represented through $\mathcal{D}$ and Ω, if $\mathcal{D}$ spawns $\mathcal{D}'$ via $(\Omega, \{D_n | n \in \mathbb{N}\})$. Because the σ-algebras in question are countably generated, so is their sum, and because the congruences are equivalent,

$$z \rho_{\mathcal{C}+\mathcal{C}'} z' \Rightarrow \forall F \in \mathcal{D} + \mathcal{D}' : (K + K')(z)(F) = (K + K')(z')(F)$$

holds. To establish this, let $z \in X, z' \in X'$, and consider

$$\mathcal{G} := \{F \in \mathcal{D} + \mathcal{D}' | (K + K')(z)(F) = (K + K')(z')(F)\}.$$

This is a σ-algebra containing the generator $\{D_n + \Omega(D_n) | n \in \mathbb{N}\}$, since the congruences are equivalent. This implies $\mathcal{D} + \mathcal{D}' \subseteq \sigma(\mathcal{G})$.
The factor object

$$\langle X + X', Y + Y', K + K' \rangle /_{\mathcal{C}+\mathcal{C}', \mathcal{D}+\mathcal{D}'}$$

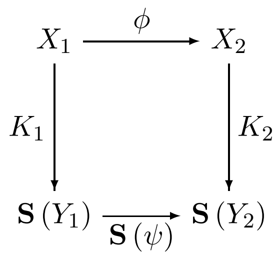
will be used in Proposition 8 for establishing that $\langle X, Y, K \rangle$ and $\langle X', Y', K' \rangle$ are bisimilar, provided they have equivalent congruences.

5 Morphisms

This section introduces morphisms more formally and shows that each morphism gives rise to a congruence in a rather natural way. This is done by imitating the usual construction of obtaining congruences from homomorphisms. It turns out that the factor space for a morphism is itself Polish whenever both source and target are Polish objects, whereas we usually end up with analytic factor spaces. Bisimulations are introduced as spans of morphisms, and we have a brief look at the congruence induced by a bisimulation on the mediating object.
The category $\mathfrak{Stoch}$ has as objects stochastic relations $\langle X, Y, K \rangle$ for measurable spaces X, Y and $K : X \rightsquigarrow Y$. A *morphism* $\langle \phi, \psi \rangle : \langle X_1, Y_1, K_1 \rangle \to \langle X_2, Y_2, K_2 \rangle$ between the objects $\langle X_i, Y_i, K_i \rangle$ is 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



is commutative.
We did encounter morphisms already when constructing the factor object:

Example 2 Let $\langle \mathcal{C}, \mathcal{D}\rangle$ be a congruence on the Polish object $\langle X, Y, K\rangle$. Then

$$\langle \eta_{\mathcal{C}}, \eta_{\mathcal{D}}\rangle : \langle X, Y, K\rangle \to \langle X, Y, K\rangle/_{\mathcal{C},\mathcal{D}}$$

is a morphism. This follows from

$$\mathbf{S}(\eta_{\mathcal{D}}) \circ K = K_{\mathcal{C},\mathcal{D}} \circ \eta_{\mathcal{C}}$$

established just before Definition 3. —

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 $\langle X, Y, K\rangle$ of $\mathfrak{Stoch}$ is dubbed *Polish* or *analytic* iff X and Y are Polish, and analytic spaces, respectively. Accordingly we denote by $\mathfrak{P} - \mathfrak{Stoch}$ and by $\mathfrak{A} - \mathfrak{Stoch}$ the full subcategories having Polish respectively analytic objects as their objects.
Under suitable conditions we can make a pair of surjective and measurable maps into morphisms.

Lemma 6 *Let $M : A \rightsquigarrow B$ be a stochastic relation between the measurable spaces A and B, assume B is separable, and that X and Y are Polish spaces with measurable and surjective maps $\phi : X \to A, \psi : Y \to B$. Then there exists a stochastic relation $K : X \rightsquigarrow Y$ such that*

$$\langle \phi, \psi\rangle : \langle X, Y, K\rangle \to \langle A, B, M\rangle$$

is a morphism.

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. □
This construction will help when it comes to investigate semi-pullbacks for analytic spaces.
Each morphism spawns a congruence in a rather natural way:

Lemma 7 *Let*

$$\langle \phi, \psi\rangle : \langle X_1, Y_1, K_1\rangle \to \langle X_2, Y_2, K_2\rangle$$

be a morphism for the stochastic relations $K_1 : X_1 \rightsquigarrow Y_1$ and $K_2 : X_2 \rightsquigarrow Y_2$, where the spaces involved are all Polish. Then

$$\langle \phi^{-1}[\mathcal{B}(X_2)], \psi^{-1}[\mathcal{B}(Y_2)]\rangle$$

is a congruence for $\langle X_1, Y_1, K_1\rangle$.

Proof Because X_2 and Y_2 are Polish spaces, the σ-algebras $\phi^{-1}[\mathcal{B}(X_2)]$ and $\psi^{-1}[\mathcal{B}(Y_2)]$ are countably generated. Abbreviate $\phi^{-1}[\mathcal{B}(X_2)]$ by $\mathcal{A}$, then

$$x\rho_{\mathcal{A}}x' \Leftrightarrow \phi(x) = \phi(x'),$$

indicating that $\rho_{\mathcal{A}}$ is smooth. This is so since we may choose a countable base $(V_n)_{n\in\mathbb{N}}$ for the topology as a generator for $\mathcal{B}(X_2)$; this base separates points, thus

$$\begin{aligned} x\rho_{\mathcal{A}}x' &\Leftrightarrow \forall n \in \mathbb{N} : \left[\phi(x) \in V_n \Leftrightarrow \phi(x') \in V_n\right] \\ &\Leftrightarrow \phi(x) = \phi(x') \end{aligned}$$

The same argument applies for ψ.
Now let $x\rho_{\mathcal{A}}x'$, and let $D_1 \in \phi^{-1}[\mathcal{B}(X_2)]$ be arbitrary, so that there exists $D_2 \in \mathcal{B}(Y_2)$ with $D_1 = \psi^{-1}[D_2]$. Thus

$$\begin{aligned} K_1(x)(D_1) &= K_1(x)\left(\psi^{-1}[D_2]\right) \\ &= (\mathbf{S}(\psi) \circ K_1)(x)(D_2) \\ &= (K_2 \circ \phi)(x)(D_2) \\ &= K_2(\phi(x))(D_2) \\ &= K_2(\phi(x'))(D_2) \\ &= K_1(x')(D_1). \end{aligned}$$

This establishes the assertion. □
Call a congruence $\langle \mathcal{C}, \mathcal{D} \rangle$ on $\langle X_1, Y_1, K_1 \rangle$ *adapted* to the morphism $\langle \phi, \psi \rangle$ iff

$$\begin{aligned} x\rho_{\mathcal{C}}x' &\Rightarrow \phi(x) = \phi(x') \\ y\rho_{\mathcal{D}}y' &\Rightarrow \psi(y) = \psi(y'), \end{aligned}$$

then the congruence constructed in Lemma 7 is the largest congruence on $\langle X_1, Y_1, K_1 \rangle$ that is adapted to the morphism, as the proof for the Lemma shows.
As expected, morphisms which are compatible with congruences factor through the factor object:

Proposition 7 *Assume that $\langle \mathcal{C}, \mathcal{D} \rangle$ is a congruence on the Polish object $\langle X, Y, K \rangle$, and let for the analytic object $\langle X_1, Y_1, K_1 \rangle$*

$$\langle \phi, \psi \rangle : \langle X, Y, K \rangle \to \langle X_1, Y_1, K_1 \rangle$$

be a morphism which is adapted to $\langle \mathcal{C}, \mathcal{D} \rangle$. Then $\langle \phi, \psi \rangle$ factors uniquely in $\mathfrak{A} - \mathfrak{Stoch}$ through $\langle X, Y, K \rangle /_{\mathcal{C},\mathcal{D}}$.

Proof 1. Because $\langle \phi, \psi \rangle$ are adapted to $\langle \mathcal{C}, \mathcal{D} \rangle$, the maps

$$\begin{aligned} \phi_{\mathcal{C}}([x]_{\mathcal{C}}) &:= \phi(x), \\ \psi_{\mathcal{D}}([y]_{\mathcal{D}}) &:= \psi(y) \end{aligned}$$

are well defined. Since ϕ is $\mathcal{B}(X) - \mathcal{B}(X_1)$-measurable, and since $\mathcal{A}_{\mathcal{C}}$ is the final σ-algebra on $X_{\mathcal{C}}$ with respect to $\eta_{\mathcal{C}}$, $\mathcal{A}_{\mathcal{C}} - \mathcal{B}(X_1)$-measurability of $\phi_{\mathcal{C}}$ is inferred: we can write

$$\mathcal{A}_{\mathcal{C}} = \{A \subseteq X_{\mathcal{C}} | \eta^{-1}[A] \in \mathcal{B}(X)\},$$

thus $\phi_{\mathcal{C}}^{-1}[B_1] \in \mathcal{A}_{\mathcal{C}}$ for $B_1 \in \mathcal{B}(X_1)$, since $\eta_{\mathcal{C}}^{-1}\left[\phi_{\mathcal{C}}^{-1}[B_1]\right] = \phi^{-1}[B_1] \in \mathcal{B}(X)$ due to the measurability of ϕ. A similar argument is used for $\psi_{\mathcal{D}}$. Clearly, these maps are onto.
2. It remains to show that $\langle \phi_{\mathcal{C}}, \psi_{\mathcal{D}} \rangle$ is a morphism. In fact, let $D_1 \subseteq Y_1$ be a Borel set, then

$$\begin{aligned} K_1(\phi_{\mathcal{C}}([x]_{\mathcal{C}}))(D_1) &= K_1(\phi(x))(D_1) \\ &= K(x)(\psi^{-1}[D_1]) \\ &= K_{\mathcal{C},\mathcal{D}}([x]_{\mathcal{C}})(\psi_{\mathcal{D}}^{-1}[D_1]), \end{aligned}$$

because $\psi^{-1}[D_1] = \eta_{\mathcal{D}}^{-1}\left[\psi_{\mathcal{D}}^{-1}[D_1]\right]$, and because $\langle \eta_{\mathcal{C}}, \eta_{\mathcal{D}} \rangle$ is a morphism. Consequently,

$$K_1 \circ \phi_{\mathcal{C}} = \mathbf{S}(\psi_{\mathcal{D}}) \circ K_{\mathcal{C},\mathcal{D}}$$

has been established. Uniqueness of the morphism is obvious. □
It turns out that the factor spaces associated with morphisms are Polish.

Corollary 2 *Denote under the conditions of Lemma 7 by*

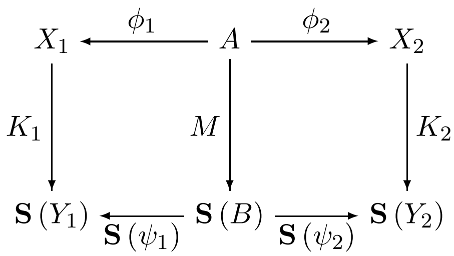
$$\langle X_1, Y_1, K_1 \rangle /_{\langle \phi, \psi \rangle}$$

the factor space associated with the largest congruence $\langle \phi^{-1}[\mathcal{B}(X_2)], \psi^{-1}[\mathcal{B}(Y_2)] \rangle$ *which is adapted to the morphism* $\langle \phi, \psi \rangle$*. Then* $\langle X_1, Y_1, K_1 \rangle /_{\langle \phi, \psi \rangle}$ *is a Polish object and isomorphic to* $\langle X_2, Y_2, K_2 \rangle$*.*

Proof The factor morphism constructed in Proposition 7 is composed of bijections. Since each morphism in $\mathfrak{Stoch}$ is an epi, and since injective maps are underlying exactly the monos, the assertion follows. □
Bisimilarity is introduced as a span of morphisms [15, 19, 10]. For coalgebras based on the category of sets, this definition agrees with the one through relations, originally given by Milner, see [19]. In [5] the authors call a bisimulation what we have introduced as congruence, albeit that paper restricts itself to labelled Markov transition systems, thus technically to stochastic relations $S \rightsquigarrow S$ for some state space S. It seems conceptually to be clearer to distinguish spans of morphisms from equivalence relations. Later we will see that there are some close connections: equivalent congruences induce bisimilarity, as we will establish in Proposition 8.
The stochastic relations $\langle X_1, Y_1, K_1 \rangle$ and $\langle X_2, Y_2, K_2 \rangle$ are called *bisimilar* iff there exists a span of morphisms

$$\langle X_1, Y_1, K_1 \rangle \xleftarrow{\langle \phi_1, \psi_1 \rangle} \langle A, B, M \rangle \xrightarrow{\langle \phi_2, \psi_2 \rangle} \langle X_2, Y_2, K_2 \rangle$$

with a suitable stochastic relation $\langle A, B, M \rangle$; the latter object is said to be *mediating.* If we deal with Polish spaces X_1, X_2, Y_1, Y_2, then we postulate that A and B are Polish spaces, too. A bisimulation yields the familiar commutative diagram



In terms of measures this translates to

$$\begin{aligned} K_1(\phi_1(a))(D_1) &= M(a)(\psi_1^{-1}[D_1]) \\ K_2(\phi_2(a))(D_2) &= M(a)(\psi_2^{-1}[D_2]) \end{aligned}$$

for all $a \in A$ and all Borel sets $D_i \subseteq Y_i$.

Lemma 8 *Suppose that*

$$\langle X_1, Y_1, K_1 \rangle \xleftarrow{\langle \phi_1, \psi_1 \rangle} \langle A, B, M \rangle \xrightarrow{\langle \phi_2, \psi_2 \rangle} \langle X_2, Y_2, K_2 \rangle$$

is a bisimulation where all spaces involved are Polish, and assume that $\langle \mathcal{C}_i, \mathcal{D}_i \rangle$ is a congruence on $\langle X_i, Y_i, K_i \rangle$ which is adapted to $\langle \phi_i, \psi_i \rangle$ for $i = 1, 2$. Then the congruence

$$\langle \sigma(\mathcal{C}_1 \cup \mathcal{C}_2), \sigma(\mathcal{D}_1 \cup \mathcal{D}_2) \rangle$$

is adapted to both $\langle \phi_1, \psi_1 \rangle$ and $\langle \phi_2, \psi_2 \rangle$.

Proof 1. We demonstrate first that

$$\langle \sigma(\mathcal{C}_1 \cup \mathcal{C}_2), \sigma(\mathcal{D}_1 \cup \mathcal{D}_2) \rangle$$

is a congruence for $\langle A, B, M \rangle$. We know from Lemma 7 that

$$\rho_{\sigma(\mathcal{C}_1 \cup \mathcal{C}_2)} = \rho_{\mathcal{C}_1} \cap \rho_{\mathcal{C}_2}.$$

Now let $x\,(\rho_{\mathcal{C}_1} \cap \rho_{\mathcal{C}_2})\,x'$, then we have to show that

$$\forall D \in \sigma(\mathcal{D}_1 \cup \mathcal{D}_2) : M(x)(D) = M(x')(D)$$

holds for each. Let

$$\mathcal{D}_0 := \{D \in \sigma(\mathcal{D}_1 \cup \mathcal{D}_2) \,|M(x)(D) = M(x')(D)\}$$

be the σ-algebra of sets for which the assertion is true. Since in particular $x \rho_{\mathcal{C}_1} x'$ holds, we see that $\mathcal{D}_1 \subseteq \mathcal{D}_0$ because $\langle \mathcal{C}_1, \mathcal{D}_1 \rangle$ is a congruence. Similarly we establish $\mathcal{D}_2 \subseteq \mathcal{D}_0$, thus $\sigma(\mathcal{D}_1 \cup \mathcal{D}_2) \subseteq \mathcal{D}_0$ follows from the fact that $\mathcal{D}_0$ is a σ-algebra. Consequently we have established the properties of a congruence.
2. From Lemma 4 it is immediate that the congruence under consideration is adapted to both $\langle \phi_1, \psi_1 \rangle$ and $\langle \phi_2, \psi_2 \rangle$. □

Corollary 3 *Under the assumptions of Lemma 8, the congruence*

$$\left\langle \sigma\left(\phi_1^{-1}[\mathcal{B}(X_1)] \cup \phi_2^{-1}[\mathcal{B}(X_2)]\right), \sigma\left(\psi_1^{-1}[\mathcal{B}(Y_1)] \cup \psi_2^{-1}[\mathcal{B}(Y_2)]\right) \right\rangle$$

is the largest congruence on $\langle A, B, M \rangle$ which is adapted to both $\langle \phi_1, \psi_1 \rangle$ and $\langle \phi_2, \psi_2 \rangle$.

We are now in a position to introduce semi-pullbacks, to investigate their existence, and to apply this knowledge to congruences.

6 Semi-Pullbacks

We will show now that semi-pullbacks exist in a rather general setting, generalizing the construction in [10] and in [11]. 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. A special case in this discussion is furnished by factor objects generated from the sum of Polish objects having equivalent congruences. Here we use semi-pullbacks to establish the bisimilarity of these objects. The Hennessy-Milner congruence can be subsumed, as we will demonstrate.

Theorem 1 *Let $\langle X_i, Y_i, K_i\rangle$ be Polish objects, and assume that X, Y are separable measurable spaces with a stochastic relation $K : X \rightsquigarrow Y$. In $\mathfrak{Stoch}$ each diagram*

$$\langle X_1, Y_1, K_1\rangle \xrightarrow{\langle\varphi_1, \psi_1\rangle} \langle X, Y, K\rangle \xleftarrow{\langle\varphi_2, \psi_2\rangle} \langle X_2, Y_2, K_2\rangle$$

has a semi-pullback

$$\langle X_1, Y_1, K_1\rangle \xleftarrow{\langle\zeta_1, \xi_1\rangle} \langle A, B, M\rangle \xrightarrow{\langle\zeta_2, \xi_2\rangle} \langle X_2, Y_2, K_2\rangle$$

with a Polish object $\langle A, B, M\rangle$.

The proof is a variant of the one given for [10, Theorem 1]. It requires, however, some variations due to the fact that we deal now with separable measurable spaces as the common target of the morphisms. Hence we give it nearly in full here, in order to render the paper self-contained.

Proof 1. In view of Observation 1 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

$$\begin{array}{ccccc} X_1 & \xrightarrow{\varphi_1} & X & \xleftarrow{\varphi_2} & X_2 \\ \downarrow{\scriptstyle K_1} & & \downarrow{\scriptstyle K} & & \downarrow{\scriptstyle K_2} \\ \mathbf{S}(Y_1) & \xrightarrow[\mathbf{S}(\psi_1)]{} & \mathbf{S}(Y) & \xleftarrow[\mathbf{S}(\psi_2)]{} & \mathbf{S}(Y_2) \end{array}$$

are commutative by assumption, 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. It is shown exactly as in the proof of [10, Theorem 1] that Γ is $\mathcal{C}$-measurable. From Proposition 2 it is now inferred that there exists a measurable map $N : A \to \mathbf{S}(B)$ such that

$$N(x_1, x_2) \in \Gamma(x_1, x_2)$$

holds for every $\langle x_1, x_2 \rangle \in A$. Thus $N : A \rightsquigarrow B$ is a stochastic relation with

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

Hence $\langle A, B, N \rangle$ is the desired semi-pullback. □
This Theorem includes several interesting special cases:

Corollary 4 *Semi-pullbacks exist in the category of stochastic relations over Polish spaces.*

Proof This follows immediately, it was proved first in [10]. □

Corollary 5 *Suppose that in the diagram of Theorem 1 the target object $\langle X, Y, K \rangle$ is an analytic object. Then a Polish semi-pullback of that diagram exists in* $\mathfrak{Stoch}$.

Proof This also follows immediately from Theorem 1. □
Corollary 5 appears to be new. Edalat's Theorem on the existence of semi-pullbacks requires universally measurable transition probability functions, i.e., in the parlance preferred by the present paper, stochastic relations for the semi-pullback, so Corollary 5 seems to be a structural improvement.
This Corollary has an interesting application to the characterization of bisimulations for labelled Markov transition systems. It could be shown that two such systems are bisimilar iff they satisfy the same formulas under the following constraints:

1. the spaces involved are analytic, the stochastic relations are universally measurable. This case was discussed in [4],
2. the spaces involved are Polish, the morphisms used are measurable and onto, and one of the processes is small, i.e. the equivalence induced by satisfaction of formulas has a Borel section. This case was discussed in [10].

Using Corollary 5, the problem can be solved now in full generality, i.e., without imposing any restrictions concerning either the morphisms or the transition systems. We will adopt another angle of view, however, when discussing bisimilarity of Markov transition systems. In Proposition 9 this discussion will be entered into again.
Now suppose that we have an analytic object $\langle A, B, M\rangle$, then we can find a Polish object $\langle X, Y, K\rangle$ and a morphism

$$\langle \gamma, \chi\rangle : \langle X, Y, K\rangle \to \langle A, B, M\rangle$$

by Lemma 6. 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 6 *Let $\langle A_i, B_i, M_i\rangle$ be an analytic object for $i = 1, 2$ and assume that $M : A \rightsquigarrow B$ is a stochastic relation for the separable measurable spaces A, B. For the pair*

$$\langle A_1, B_1, M_1\rangle \xrightarrow{\langle \phi_1, \psi_1\rangle} \langle A, B, M\rangle \xleftarrow{\langle \phi_2, \psi_2\rangle} \langle A_2, B_2, M_2\rangle$$

exist both a Polish object $\langle X, Y, K\rangle$ and surjective Borel measurable maps

$$\langle A_1, B_1, M_1\rangle \xleftarrow{\langle f_1, g_1\rangle} \langle X, Y, K\rangle \xrightarrow{\langle f_2, g_2\rangle} \langle A_2, B_2, M_2\rangle$$

as a semi-pullback.

Proof We can find Polish objects $\langle X_i, Y_i, K_i\rangle$ and morphisms in $\mathfrak{Stoch}$ extending the diagram

$$\begin{array}{ccccc}
\langle X_1, Y_1, K_1\rangle & & & & \langle X_2, Y_2, K_i\rangle \\
{\scriptstyle \langle \gamma_1, \chi_1\rangle}\downarrow & & & & \downarrow{\scriptstyle \langle \gamma_2, \chi_2\rangle} \\
\langle A_1, B_1, M_1\rangle & \xrightarrow{\langle \phi_1, \psi_1\rangle} & \langle A, B, M\rangle & \xleftarrow{\langle \phi_2, \psi_2\rangle} & \langle A_2, B_2, M_2\rangle
\end{array}$$

Now, using Theorem 1, find a semi-pullback

$$\langle X_1, Y_1, K_1\rangle \xleftarrow{\langle \theta_1, \rho_1\rangle} \langle X, Y, K\rangle \xrightarrow{\langle \theta_2, \rho_2\rangle} \langle X_2, Y_2, K_2\rangle$$

for the diagram

$$\langle X_1, Y_1, K_1\rangle \xrightarrow{\langle \phi_1 \circ \gamma_1, \psi_1 \circ \chi_1\rangle} \langle A, B, M\rangle \xleftarrow{\langle \phi_2 \circ \gamma_2, \psi_2 \circ \chi_2\rangle} \langle X_2, Y_2, K_i\rangle$$

Here $\langle X, Y, K\rangle$ is a Polish object, and $\langle \theta_1, \rho_1\rangle, \langle \theta_2, \rho_2\rangle$ are morphisms in $\mathfrak{Stoch}$. Putting

$$\begin{aligned} f_i &:= \gamma_i \circ \theta_i, \\ g_i &:= \chi_i \circ \rho_i \end{aligned}$$

for $i = 1, 2$ will establish the claim. $\square$
Thus we have in particular established:

Theorem 2 *The category $\mathfrak{A} - \mathfrak{Stoch}$ of stochastic relations over analytic spaces has semi-pullbacks.*

This has some consequences:

Corollary 7 *Bisimilarity is transitive both in $\mathfrak{Stoch}$ and $\mathfrak{A} - \mathfrak{Stoch}$.*

Proof The proof proceeds exactly as in [19, Theorem 5.4] □
Suppose two Polish objects have equivalent congruences. Then they are bisimilar, as we will show now as a final application for the existence of semi-pullbacks.

Proposition 8 *If $\langle \mathcal{C}_i, \mathcal{D}_i \rangle$ are equivalent congruences on the Polish objects $\langle X_i, Y_i, K_i \rangle$ for $i = 1, 2$, then $\langle X_1, Y_1, K_1 \rangle$ and $\langle X_2, Y_2, K_2 \rangle$ are bisimilar.*

Proof 1. Construct the sum $\langle X_1 + X_2, Y_1 + Y_2, K_1 + K_2 \rangle$ of the two objects as at the end of Section 4, and let $\langle \kappa_i, \lambda_i \rangle$ be the corresponding injections, which are, however, no morphisms. Let

$$\eta_{\mathcal{C}+\mathcal{C}'}, \eta_{\mathcal{D}+\mathcal{D}'} : \langle X_1 + X_2, Y_1 + Y_2, K_1 + K_2 \rangle \to \langle X_1 + X_2, Y_1 + Y_2, K_1 + K_2 \rangle /_{\mathcal{C}+\mathcal{C}', \mathcal{D}+\mathcal{D}'}$$

be the canonical injection, then

$$\langle \eta_{\mathcal{C}+\mathcal{C}'} \circ \kappa_i, \eta_{\mathcal{D}+\mathcal{D}'} \circ \lambda_i \rangle$$

constitutes a morphism $\langle X_i, Y_i, K_i \rangle \to \langle X_1 + X_2, Y_1 + Y_2, K_1 + K_2 \rangle /_{\mathcal{C}+\mathcal{C}', \mathcal{D}+\mathcal{D}'}$, as will be shown now. The crucial step is establishing surjectivity.
2. We will first give a representation of the equivalence classes for the equivalence relations associated with $\mathcal{C}_1 + \mathcal{C}_2$ and $\mathcal{D}_1 + \mathcal{D}_2$. This will then imply that the injections into the sums composed with the factor maps are indeed onto.
Fix a generator $\{A_n | n \in \mathbb{N}\}$ for $\mathcal{C}_1$, and assume that $\mathcal{C}_1$ spawns $\mathcal{C}_2$ via $(\Upsilon, \{A_n | n \in \mathbb{N}\})$. Then $\{\Upsilon(A_n) | n \in \mathbb{N}\}$ is a generator for $\mathcal{C}_2$, and

$$\mathcal{C}_1 + \mathcal{C}_2 = \sigma(\{A_n + \Upsilon(A_n) | n \in \mathbb{N}\})$$

holds. We claim that each equivalence class $a \in (X_1 + X_2)/_{\mathcal{C}_1 + \mathcal{C}_2}$ can be represented as

$$a = [x_1]_{\mathcal{C}_1} + [x_2]_{\mathcal{C}_2}$$

for some suitably chosen $x_1 \in X_1, x_2 \in X_2$. In fact, suppose $a = [x_1]_{\mathcal{C}_1 + \mathcal{C}_2}$ for some $x_1 \in X_1$. Then

$$[x_1]_{\mathcal{C}_1} = \bigcap \{A_n | x_1 \in A_n\} \cap \bigcap \{X_1 \setminus A_n | x_1 \notin A_n\}$$

holds (see the remark following Definition 1). Consequently,

$$\begin{aligned} \Upsilon([x_1]_{\mathcal{C}_1}) &= \bigcap \{\Upsilon(A_n) | x_1 \in A_n\} \cap \bigcap \{X_2 \setminus \Upsilon(A_n) | x_1 \notin A_n\} \\ &= [x_2]_{\mathcal{C}_2} \end{aligned}$$

for some $x_2 \in X_2$ by Lemma 5. This is so since we know that the atoms of $\mathcal{C}_2$ are exactly the classes $[\cdot]_{\mathcal{C}_2}$. Consequently, we have for each $n \in \mathbb{N}$ that $x_1 \in A_n$ holds iff $x_2 \in \Upsilon(A_n)$ is true. Since $x_1 \in X_1, x_2 \in X_2$, and $X_1 \cap X_2 = \emptyset$, we have established

$$x_1 \in A_n + \Upsilon(A_n) \Leftrightarrow x_2 \in A_n + \Upsilon(A_n)$$

But this means $x_1 \ \rho_{\mathcal{C}_1+\mathcal{C}_2} \ x_2$. Thus we have shown that indeed

$$a = [x_1]_{\mathcal{C}_1} + [x_2]_{\mathcal{C}_2}$$

holds. If $a = [x_2]_{\mathcal{C}_1+\mathcal{C}_2}$ for some $x_2 \in X_2$, the same conclusion would have been reached by interchanging the roles of X_1 and X_2, which is possible because we have equivalent congruences. It is obvious that $[x_1]_{\mathcal{C}_1} + [x_2]_{\mathcal{C}_2}$ for $x_i \in X_i$ forms a $\rho_{\mathcal{C}_1+\mathcal{C}_2}$-class.
In the same way we show that each equivalence class $b \in (Y_1 + Y_2)/_{\mathcal{D}_1+\mathcal{D}_2}$ can be represented as

$$b = [y_1]_{\mathcal{D}_1} + [y_2]_{\mathcal{D}_2}$$

for suitably chosen $y_1 \in Y_1, y_2 \in Y_2$, and that the sum of classes is a class again.
3. The pullback of the pair of morphisms with a joint target constructed in the first step is a Polish object which has the desired properties. $\square$
Let us see how the bisimulation of labelled Markov transition systems from Example 1 fits in. Assume that $\langle S, (k_a)_{a\in \mathsf{L}}\rangle$ and $\langle S', (k'_a)_{a\in \mathsf{L}}\rangle$ are such systems for Polish spaces S and S'. Morphisms (and consequently bisimulations) are defined so that the corresponding properties hold for each action; for example, $\phi : \langle S, (k_a)_{a\in \mathsf{L}}\rangle \to \langle S', (k'_a)_{a\in \mathsf{L}}\rangle$ is a morphism iff $\phi : S \to S'$ is a surjective measurable map such that

$$\forall a \in \mathsf{L} : k'_a \circ \phi = \mathbf{S}(\phi) \circ k_a$$

holds. A bisimulation is then a span of morphisms; a semi-pullback is then a span for a co-span. Congruences are defined in the same manner.
Define as in Example 1 the sets $[\![\varphi]\!]_S := \{s \in S | s \models \varphi\}$ and similarly $[\![\varphi]\!]_{S'}$. Then

$$\mathcal{C}(\Phi, S) := \sigma(\{[\![\varphi]\!]_S \, | \varphi \in \Phi\})$$

is a countably generated σ-algebra on S, similarly, $\mathcal{C}(\Phi, S')$ is defined on S'. Both σ-algebras are countably generated, since the set Φ of all formulas is countable. It is clear that $\mathcal{C}(\Phi, S)$ is a congruence on $\langle S, (k_a)_{a\in \mathsf{L}}\rangle$, since

$$s \approx s' \Rightarrow \forall a \in \mathsf{L} \forall \varphi \in \Phi : k_a(s)([\![\varphi]\!]_S) = k_a(s')([\![\varphi]\!]_S)$$

holds (cf. [4] or [10]).
Now assume that these transition systems are equivalent in the following sense:

$$\langle S, (k_a)_{a\in \mathsf{L}}\rangle \sim \langle S', (k'_a)_{a\in \mathsf{L}}\rangle$$

iff

$$\forall s \in S \exists s' \in S' : s \approx s' \text{ and } \forall s' \in S' \exists s \in S : s' \approx s,$$

where $\approx$ is defined in Example 1. This entails that equivalent labelled Markov transition systems satisfy exactly the same formulas.
Define for $\varphi \in \Phi$

$$\Upsilon([\![\varphi]\!]_S) := [\![\varphi]\!]_{S'} \,.$$

We claim that $\mathcal{C}(\Phi, S)$ spawns $\mathcal{C}(\Phi, S')$ via $(\Upsilon, \{[\![\varphi]\!]_S \, | \phi \in \Phi\})$. From the construction it is clear that Υ maps generators to generators, hence the condition that $\{\Upsilon([\![\varphi]\!]_S) | \phi \in \Phi\}$ generates

$\mathcal{C}(\Phi, S')$ is satisfied. Now assume that $s_1 \rho_{\mathcal{C}(\Phi,S)} s_2$ for $s_1, s_2 \in S$, then we know that we can find $s'_1, s'_2 \in S'$ with $s_1 \approx s'_1, s_2 \approx s'_2$. Since

$$\left[s'_1\right]_{\mathcal{C}(\Phi,S')} = \bigcap\{\Upsilon([\![\varphi]\!]_S) | s_1 \models \phi\} \cap \bigcap\{S' \setminus \Upsilon([\![\varphi]\!]_S) | s_1 \not\models \phi\},$$

and since we have by transitivity of $\approx$

$$\left[s'_1\right]_{\mathcal{C}(\Phi,S')} = \left[s'_2\right]_{\mathcal{C}(\Phi,S')},$$

the second condition in Definition 4 is also satisfied. Consequently, $\mathcal{C}(\Phi, S)$ spawns $\mathcal{C}(\Phi, S')$ via $(\Upsilon, \{[\![\varphi]\!]_S | \phi \in \Phi\})$. We find even that $\mathcal{C}(\Phi, S) \propto \mathcal{C}(\Phi, S')$, because we obtain from $\langle S, (k_a)_{a \in \mathsf{L}}\rangle \sim \langle S', (k'_a)_{a \in \mathsf{L}}\rangle$ that for each $a \in \mathsf{L}$

$$\forall s \in S \forall s' \in \Upsilon([s]_S) \forall \phi \in \Phi : k_a(s)([\![\varphi]\!]_S) = k'_a(s')([\![\varphi]\!]_{S'})$$

holds. Interchanging the roles of S and S', we find $\mathcal{C}(\Phi, S') \propto \mathcal{C}(\Phi, S)$, thus both congruences are equivalent. An application of Proposition 8 yields:

Proposition 9 *The labelled Markov transitions systems $\langle S, (k_a)_{a \in \mathsf{L}}\rangle$ and $\langle S', (k'_a)_{a \in \mathsf{L}}\rangle$ accept the same formulas of the Hennessy-Milner logic iff they are bisimilar.*

Note that this Proposition works, in contrast to the corresponding statements in [4] or [10], without any restrictions either on the transition functions or on the transition system itself.

7 Conclusion and Further Work

We discuss congruences for stochastic relations, capitalizing on the equivalence of smooth equivalence relations and countably generated σ-algebras. This opens up the avenue of investigating factor spaces, and some algebraic legwork has to be done in order to put this development into the right perspective. In particular the question arises under which conditions the semi-pullback for a span of morphisms in the category of stochastic relations does exist. It is known that it exists if the underlying spaces are Polish, and it is shown in this paper that even for analytic spaces the existence of semi-pullbacks can be proven.
At the heart of the proof for this statement lies a measure extension problem: it can be shown that a measure defined on a sub-σ-algebra of the Borel sets of a product of two Polish spaces can be extended to the Borel sets, respecting some prescribed distributions. Varying the sub-σ-algebra within certain not too narrow limits permits some variations of the existence proof for semi-pullbacks which follows an established pattern through the application of some selection theorems from Operations Research. Thus we see an example of proof reuse by establishing the existence of the same construction under different conditions.
The paper's contributions are the following:

1. the existence of semi-pullbacks is established for stochastic relations over analytic spaces,
2. congruences are defined, and it is shown how factor spaces for these congruences can be constructed; the interplay of congruences and smooth equivalence relations is investigated,
3. equivalent congruences are investigated, and it is shown that stochastic relations that have equivalent congruences are bisimilar.

Equivalent congruences arise for example when considering labelled Markov transition systems. Two states are equivalent iff they satisfy the same formulas of a very simple negation free Hennessy-Milner logic: the congruences are equivalent iff each state in one system finds an equivalent state in the other system, and vice versa. It was shown in [4, 10] that such an equivalence exists iff the transition systems are bisimilar. We show that this is actually a special case of the theory developed here.
The congruences investigated in the present paper are all defined over Polish spaces, and the natural question to ask is whether this development can be carried over to stochastic relations over analytic spaces. This would enable us to investigate factor structures arising from congruences much more closely (the present paper remains a bit of a torso in this respect). Another question is whether equivalent congruences can be constructed from bisimilar relations, which would then permit giving a necessary and sufficient conditions on the existence of bisimulations. This would then indicate a nice parallel between labelled Markov transition systems and these relations on one hand, and between stochastic and non-deterministic relations on the other hand.

References

[1] S. Abramsky, R. Blute, and P. Panangaden. Nuclear and trace ideal in tensored *-categories. *Journal of Pure and Applied Algebra*, 143(1 – 3):3 – 47, 1999.

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

[3] C. Brink, W. Kahl, and G. Schmidt, editors. *Relational Methods in Computer Science*. Advances in Computing Science. Springer-Verlag, Wien, New York, 1997.

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

[5] J. Desharnais, R. Jagadeesan, and V. Guptaand P. Panangaden. Approximating labeled Markov processes. In *Proc. 15th Ann. IEEE Symp. on Logic in Computer Science*, pages 95 – 106, June, 2000. IEEE Computer Society.

[6] J. Desharnais, A. Mili, and T. T. Nguyen. Refinement and demonic semantics. In *[3]*, pages 166 – 184.

[7] E.-E. Doberkat. The demonic product of probabilistic relations. In Mogens Nielsen and Uffe Engberg, editors, *Proc. Foundations of Software Science and Computation Structures*, volume 2303 of *Lecture Notes in Computer Science*, pages 113 – 127, Berlin, 2002. Springer-Verlag.

[8] E.-E. Doberkat. The converse of a probabilistic relation. In A. Gordon, editor, *Proc. Foundations of Software Science and Computation Structures'03*, volume 2620 of *Lecture Notes in Computer Science*, pages 233 – 249, Berlin, 2003. Springer-Verlag.

[9] E.-E. Doberkat. Pipes and filters: Modelling a software architecture through relations. *Acta Informatica*, 2003. in print.

[10] 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.

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

[12] M. Giry. A categorical approach to probability theory. In *Categorical Aspects of Topology and Analysis*, volume 915 of *Lecture Notes in Mathematics*, pages 68 – 85, Berlin, 1981. Springer-Verlag.

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

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

[15] A. Joyal, M. Nielsen, and G. Winskel. Bisimulation from open maps. *Information and Computation*, 127(2):164 – 185, 1996.

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

[17] P. Panangaden. Probabilistic relations. In C. Baier, M. Huth, M. Kwiatkowska, and M. Ryan, editors, *Proc. PROBMIV*, pages 59 – 74, 1998. Also available from the School of Computer Science, McGill University, Montreal.

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

[19] J. J. M. M. Rutten. Universal coalgebra: a theory of systems. *Theoretical Computer Science*, 249(1):3 – 80, 2000. Special issue on modern algebra and its applications.

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

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

/99/ T. Bühren, M. Cakir, E. Can, A. Dombrowski, G. Geist, V. Gruhn, M. Gürgrn, S. Handschumacher, M. Heller, C. Lüer, D. Peters, G. Vollmer, U. Wellen, J. von Werne
Endbericht der Projektgruppe eCCo (PG 315)
Electronic Commerce in der Versicherungsbranche
Beispielhafte Unterstützung verteilter Geschäftsprozesse
Februar 1999

/100/ A. Fronk, J. Pleumann,
Der DoDL-Compiler
August 1999

/101/ K. Alfert, E.-E. Doberkat, C. Kopka
Towards Constructing a Flexible Multimedia Environment for Teaching the History of Art
September 1999

/102/ E.-E. Doberkat
An Note on a Categorial Semantics for ER-Models
November 1999

/103/ Christoph Begall, Matthias Dorka, Adil Kassabi, Wilhelm Leibel, Sebastian Linz, Sascha Lüdecke, Andreas Schröder, Jens Schröder, Sebastian Schütte, Thomas Sparenberg, Christian Stücke, Martin Uebing, Klaus Alfert, Alexander Fronk, Ernst-Erich Doberkat
Abschlußbericht der Projektgruppe PG-HEU (326)
Oktober 1999

/104/ Corina Kopka
Ein Vorgehensmodell für die Entwicklung multimedialer Lernsysteme
März 2000

/105/ Stefan Austen, Wahid Bashirazad, Matthais Book, Traugott Dittmann, Bernhard Flechtker, Hassan Ghane, Stefan Göbel, Chris Haase, Christian Leifkes, Martin Mocker, Stefan Puls, Carsten Seidel, Volker Gruhn, Lothar Schöpe, Ursula Wellen
Zwischenbericht der Projektgruppe IPSI
April 2000

/106/ Ernst-Erich Doberkat
Die Hofzwerge — Ein kurzes Tutorium zur objektorientierten Modellierung
September 2000

/107/ Leonid Abelev, Carsten Brockmann, Pedro Calado, Michael Damatow, Michael Heinrichs, Oliver Kowalke, Daniel Link, Holger Lümkemann, Thorsten Niedzwetzki, Martin Otten, Michael Rittinghaus, Gerrit Rothmaier
Volker Gruhn, Ursula Wellen
Zwischenbericht der Projektgruppe Palermo
November 2000

/108/ Stefan Austen, Wahid Bashirazad, Matthais Book, Traugott Dittmann, Bernhard Flechtker, Hassan Ghane, Stefan Göbel, Chris Haase, Christian Leifkes, Martin Mocker, Stefan Puls, Carsten Seidel, Volker Gruhn, Lothar Schöpe, Ursula Wellen
Endbericht der Projektgruppe IPSI
Februar 2001

/109/ Leonid Abelev, Carsten Brockmann, Pedro Calado, Michael Damatow, Michael Heinrichs, Oliver Kowalke, Daniel Link, Holger Lümkemann, Thorsten Niedzwetzki, Martin Otten, Michael Rittinghaus, Gerrit Rothmaier
Volker Gruhn, Ursula Wellen
Zwischenbericht der Projektgruppe Palermo
Februar 2001

/110/ Eugenio G. Omodeo, Ernst-Erich Doberkat
Algebraic semantics of ER-models from the standpoint of map calculus.
Part I: Static view
März 2001

/111/ Ernst-Erich Doberkat
An Architecture for a System of Mobile Agents
März 2001

/112/ Corina Kopka, Ursula Wellen
Development of a Software Production Process Model for Multimedia CAL Systems by Applying Process Landscaping
April 2001

/113/ Ernst-Erich Doberkat
The Converse of a Probabilistic Relation
Oktober 2002

/114/ Ernst-Erich Doberkat, Eugenio G. Omodeo
Algebraic semantics of ER-models in the context of the calculus of relations.
Part II: Dynamic view
Juli 2001

/115/ Volker Gruhn, Lothar Schöpe (Eds.)
Unterstützung von verteilten Softwareentwicklungsprozessen durch integrierte Planungs-, Workflow- und Groupware-Ansätze
September 2001

/116/ Ernst-Erich Doberkat
The Demonic Product of Probabilistic Relations
September 2001

/117/ Klaus Alfert, Alexander Fronk, Frank Engelen
Experiences in 3-Dimensional Visualization of Java Class Relations
September 2001

/118/ Ernst-Erich Doberkat
The Hierarchical Refinement of Probabilistic Relations
November 2001

/119/ Markus Alvermann, Martin Ernst, Tamara Flatt, Urs Helmig, Thorsten Langer, Ingo Röpling, Clemens Schäfer, Nikolai Schreier, Olga Shtern
Ursula Wellen, Dirk Peters, Volker Gruhn
Project Group Chairware Intermediate Report
November 2001

/120/ Volker Gruhn, Ursula Wellen
Autonomies in a Software Process Landscape
Januar 2002

/121/ Ernst-Erich Doberkat, Gregor Engels (Hrsg.)
Ergebnisbericht des Jahres 2001
des Projektes "MuSofT – Multimedia in der SoftwareTechnik"
Februrar 2002

/122/ Ernst-Erich Doberkat, Gregor Engels, Jan Hendrik Hausmann, Mark Lohmann, Christof Veltmann
Anforderungen an eine eLearning-Plattform — Innovation und Integration —
April 2002

/123/ Ernst-Erich Doberkat
Pipes and Filters: Modelling a Software Architecture Through Relations
Juni 2002

/124/ Volker Gruhn, Lothar Schöpe
Integration von Legacy-Systemen mit Eletronic Commerce Anwendungen
Juni 2002

/125/ Ernst-Erich Doberkat
A Remark on A. Edalat's Paper *Semi-Pullbacks and Bisimulations in Categories of Markov-Processes*
Juli 2002

/126/ Alexander Fronk
Towards the algebraic analysis of hyperlink structures
August 2002

/127/ Markus Alvermann, Martin Ernst, Tamara Flatt, Urs Helmig, Thorsten Langer
Ingo Röpling, Clemens Schäfer, Nikolai Schreier, Olga Shtern
Ursula Wellen, Dirk Peters, Volker Gruhn
Project Group Chairware Final Report
August 2002

/128/ Timo Albert, Zahir Amiri, Dino Hasanbegovic, Narcisse Kemogne Kamdem,
Christian Kotthoff, Dennis Müller, Matthias Niggemeier, Andre Pavlenko, Stefan Pinschke,
Alireza Salemi, Bastian Schlich, Alexander Schmitz
Volker Gruhn, Lothar Schöpe, Ursula Wellen
Zwischenbericht der Projektgruppe Com42Bill (PG 411)
September 2002

/129/ Alexander Fronk
An Approach to Algebraic Semantics of Object-Oriented Languages
Oktober 2002

/130/ Ernst-Erich Doberkat
Semi-Pullbacks and Bisimulations in Categories of Stochastic Relations
November 2002

/131 Yalda Ariana, Oliver Effner, Marcel Gleis, Martin Krzysiak,
Jens Lauert, Thomas Louis, Carsten Röttgers, Kai Schwaighofer,
Martin Testrot, Uwe Ulrich, Xingguang Yuan
Prof. Dr. Volker Gruhn, Sami Beydeda
Endbericht der PG nightshift:
Dokumentation der verteilten Geschäftsprozesse im FBI und Umsetzung von Teilen dieser Prozesse im Rahmen eines FBI-Intranets basierend auf WAP- und Java-Technologie
Februar 2003

/132/ Ernst-Erich Doberkat, Eugenio G. Omodeo
ER Modelling from First Relational Principles
Februar 2003

/133/ Klaus Alfert, Ernst-Erich Doberkat, Gregor Engels (Hrsg.)
Ergebnisbericht des Jahres 2002 des Projektes "MuSofT – Multimedia in der SoftwareTechnik"
März 2003

/134/ Ernst-Erich Doberkat
Tracing Relations Probabilistically
März 2003

/135/ Timo Albert, Zahir Amiri, Dino Hasanbegovic, Narcisse Kemogne Kamdem,
Christian Kotthoff, Dennis Müller, Matthias Niggemeier,
Andre Pavlenko, Alireza Salemi,Bastian Schlich, Alexander Schmitz,
Volker Gruhn, Lothar Schöpe, Ursula Wellen
Endbericht der Projektgruppe Com42Bill (PG 411)
März 2003

/136/ Klaus Alfert
Vitruv: Specifying Temporal Aspects of Multimedia Presentations —
A Transformational Approach based on Intervals
April 2003

/137/ Klaus Alfert, Jörg Pleumann, Jens Schröder
A Framework for Lightweight Object-Oriented Design Tools
April 2003

/138/ K. Alfert, A. Fronk , Ch. Veltmann (Hrsg.)
Stefan Borggraefe, Leonore Brinker, Evgenij Golkov, Rafael Hosenberg, Bastian Krol, Daniel Mölle,
Markus Niehammer, Ulf Schellbach, Oliver Szymanski, Tobias Wolf, Yue Zhang
Endbericht der Projektgruppe 415: Konzeption und Implementierung eines digitalen und hypermedialen Automobilcockpits (HyCop)
Mai 2003

/139/ Volker Gruhn, Malte Hülder, Sami Beydeda (Hrsg.)
Endbericht der Projektgruppe 409: Entwicklung von ortsbasierten Diensten für UMTS-Mobilfunkgeräte (mCube)
Mai 2003

/140/ Ernst-Erich Doberkat
Congruences for Stochastic Relations
Juli 2003