

The Converse of a Stochastic Relation

Ernst-Erich Doberkat

Chair for Software Technology
University of Dortmund
doberkat@acm.org

Abstract. Transition probabilities are proposed as the stochastic counterparts to set-based relations. We propose the construction of the converse of a stochastic relation. It is shown that two of the most useful properties carry over: the converse is idempotent as well as anticommutative. The nondeterminism associated with a stochastic relation is defined and briefly investigated. We define a bisimulation relation, and indicate conditions under which this relation is transitive; moreover it is shown that bisimulation and converse are compatible.

Keywords: Stochastic relations, concurrency, bisimulation, converse, relational calculi, nondeterminism.

1 Introduction

The use of relations is ubiquitous in Mathematics, Logic and Computer Science, their systematic study goes back as far as Schröder's seminal work. Ongoing research with a focus on program specification may be witnessed from the wealth of material collected in [18, 4]. The map calculus [5] shows that these methods determine an active line of research in Logic.

This paper deals with stochastic rather than set-valued relations, it studies the converse of such a relation. It investigates furthermore some similarities between forming the converse for set-theoretic relations and for their stochastic cousins.

For introducing into the problem, let R be a relation, i.e., a set of pairs of, say, states. If $\langle x, y \rangle \in R$, then this is written as $x \rightarrow_R y$ and interpreted as a state transition from x to y . The converse R^\smile shifts attention to the goal of the transition: $y \rightarrow_{R^\smile} x$ is interpreted as y being the goal of a transition from x . Now let $p(x, y)$ be the probability that there is a transition from x to y , and the question arises with which probability state y is the goal of a transition from x . This question cannot be answered unless we know the initial probabilities for the states. Then we can calculate $p_\mu^\smile(y, x)$ as the probability to make a transition from x to y weighted by the probability to start from x conditional to the event to reach y at all, i.e.

$$p_\mu^\smile(y, x) := \frac{\mu(x) \cdot p(x, y)}{\sum_t \mu(t) \cdot p(t, y)}.$$

Consider as an example the simple transition system p on three states given in the left hand side of Fig. 1. The converse p_μ^\smile for the initial probability $\mu := [1/2 \ 1/4 \ 1/4]$ is given on the right hand side.

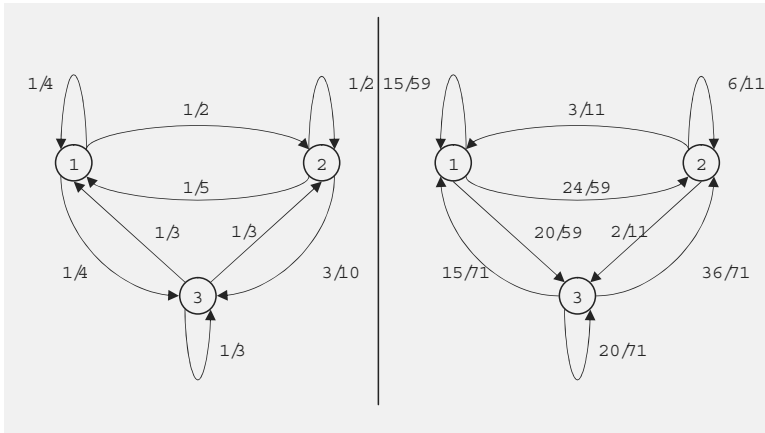


Fig. 1. A Stochastic Relation and Its Converse

The situation is more complicated in the non-finite case, which is considered here; since some measure theoretic constructions do not work in the general case, we assume that the measurable structure comes from Polish, i.e., second countable and completely metrizable topological spaces (like the real line \mathbb{R}). A definition of the converse K_μ^\smile of a stochastic relation K given an initial distribution μ is proposed in terms of disintegration. An interpretation of the converse in terms of random variables is given, and it is shown that the converse behaves with respect to composition like its set-theoretic counterpart, viz., $(K;L)_\mu^\smile = L_{K^\bullet(\mu)}^\smile;K_\mu^\smile$, where $K^\bullet(\mu)$ denotes the image distribution of μ under K , and the composition is the Kleisli composition for the corresponding monad (section 4). This is of course the probabilistic counterpart to the corresponding law for relations R and S , which reads $(R;S)^\smile = S^\smile;R^\smile$.

The set $\{K_\mu^\smile(y) | y \in Y\}$ of all sub-probability measures constituting the converse turns out to have an interesting property: it is topologically rather small, i.e., its closure is compact in the weak topology of sub-probability measures on Y (Cor. 4.2). This indicates that the converse K_μ^\smile does not carry as much information as K or μ do.

A stochastic relation K between X and Y induces a set-theoretic relation R_K (called the *fringe relation*) in the following way: let $\langle x, y \rangle \in R_K$ iff $K(x)(U) > 0$ for each open neighborhood U of y . Relation R_K is considered as K 's nondeterminism, since it indicates the set of all pairs that are possible for the stochastic relation K . The relationship between these relations is briefly investigated in terms of natural transformations between two functors in Sect. 3.

A stochastic relation models the dynamics of a system, which is partly captured through the notion of bisimilarity. Thus the question of stability under bisimilarity arises when constructing the converse. We define in section 5 a suitable notion of bisimilarity and show that this is a transitive relation. It is shown

that the converses K_μ^\smile and L_ν^\smile are bisimilar, provided K and L as well as the initial distributions μ and ν are bisimilar¹.

Acknowledgements. Part of this work could be done while the author was visiting the Dipartimento di Informatica at the University of L'Aquila. The visit was in part supported through grants from the Exchange Programme for Scientists between Italy and Germany from the Italian Ministry of Foreign Affairs/Deutscher Akademischer Austauschdienst and from Progetto speciale I.N.D.A.M./GNIM *Nuovi paradigmi di calcolo: Linguaggi e Modelli*. The author wants to thank Eugenio Omodeo and Gunther Schmidt for getting him interested in relational methods. The referees' comments and suggestions are gratefully acknowledged.

2 Stochastic Relations

Before stochastic relations are introduced, some basic facts from measure theory are recalled. We also introduce some basic operations on these relations.

A *Polish space* (X, \mathcal{T}) is a topological space which has a countable dense subset, and which is metrizable through a complete metric. 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 $f : X \rightarrow Y$ is \mathcal{A} - \mathcal{B} -*measurable* whenever $f^{-1}[\mathcal{B}] \subseteq \mathcal{A}$ holds, where $f^{-1}[\mathcal{B}] := \{f^{-1}[B] \mid B \in \mathcal{B}\}$ is the set of inverse images $f^{-1}[B] := \{x \in X \mid f(x) \in B\}$ of elements of \mathcal{B} . 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 not be mentioned explicitly when talking about Borel maps.

The category \mathfrak{SB} has as objects Standard Borel (SB) spaces, a morphism $f \in \mathfrak{SB}(X, Y)$ between two SB spaces X and Y is a Borel map $f : X \rightarrow Y$.

Recall that a map $f : X \rightarrow Y$ between the topological spaces (X, \mathcal{T}) and (Y, \mathcal{S}) is *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 Polish spaces without their topologies, and the Borel sets are always understood with respect to the topology. $\mathcal{M}(X)$ denotes the vector space of all bounded real-valued Borel maps on the SB-space X .

The set $\mathbf{S}(X)$ denotes the set of all sub-probability measures on the SB space X . The former set carries the weak topology, i.e., the smallest topology which makes the map $\mu \mapsto \int_X f \, d\mu$ for all continuous functions $f : X \rightarrow \mathbb{R}$ continuous as soon as X carries a Polish topology. It is well known that the

¹ The full paper is available as Technische-Berichte/Doberkat_SWT-Memo-113-ps.gz in directory `ftp://ls10-www.cs.uni-dortmund.de/pub`

weak topology on $\mathbf{S}(X)$ is a Polish space [16, Theorem II.6.5], and that its Borel sets are the smallest σ -algebra on $\mathbf{S}(X)$ for which for any Borel set $B \subseteq Y$ the map $\mu \mapsto \mu(B)$ is measurable. This σ -algebra is sometimes called the *weak- \ast - σ -algebra* in stochastic dynamic optimization. Note that the weak- \ast - σ -algebra depends only on the σ -algebra of the underlying SB-space, hence is independent of any specific Polish topology. An argument due to Giry [11] shows that \mathbf{S} is the functorial part of a monad over \mathfrak{SB} , and that the Kleisli morphisms coming with this monad are just the stochastic relations.

Given two Polish spaces X and Y , a *stochastic relation* $K : X \rightsquigarrow Y$ is a Borel map from X to the set $\mathbf{S}(Y)$. Hence $K : X \rightsquigarrow Y$ is a stochastic relation iff

1. $K(x)$ is a sub-probability measure on (the Borel sets of) Y for all $x \in X$,
2. $x \mapsto K(x)(B)$ is a measurable map for each Borel set $B \subseteq Y$.

Composition of stochastic relations is the Kleisli composition: let $K : X \rightsquigarrow Y$ and $L : Y \rightsquigarrow Z$, then define for $x \in X, C \in \mathcal{B}_Z$:

$$(K;L)(x)(C) := \int_Y L(y)(C) K(x)(dy).$$

Standard arguments show that $K;L : X \rightsquigarrow Z$.

In terms of input/output systems, $K(x)(dy)$ may be interpreted that dy is the output of the system modelled by K after input x ; the system does not need to be strictly probabilistic in the sense that each input produces an output with probability 1, i.e., $K(x)(Y) = 1$ does not hold necessarily. $K(x)(Y) < 1$ may occur when K models a non-terminating computation, so that $1 - K(x)(Y)$ is the probability for the event *no output at all*. Note that the Markov processes investigated in [6, 10] are special cases.

Example 1. In the discrete case a stochastic relation p between $\{1, \dots, n\}$ and $\{1, \dots, m\}$ is represented through a non-negative substochastic matrix

$$(p(i, j))_{1 \leq i \leq n, 1 \leq j \leq m}.$$

The composition of two stochastic relations p and q is expressed through matrix multiplication, which is the discrete analogue to the Kleisli product above. \diamond

We collect some constructions and indicate some well known properties which will be helpful in the sequel. It shows how a measurable map and a measure induce a measure on the range of that map, and how a measure and a stochastic relation define a measure on the relation's target space, and on the product space, resp.

Definition 1. *Let X and Y be SB-spaces.*

1. $f^\flat(\mu)(B) := \mu(f^{-1}[B])$ defines a map $\mathfrak{SB}(X, Y) \times \mathbf{S}(X) \rightarrow \mathbf{S}(Y)$ such that

$$\int_Y g \, df^\flat(\mu) = \int_X g \circ f \, d\mu$$

holds for each $g \in \mathcal{M}(Y)$.

2. $K^\bullet(\mu)(B) := \int_X K(x)(B) \mu(dx)$ defines a map $\mathfrak{S}\mathfrak{B}(X, \mathfrak{S}(Y)) \times \mathfrak{S}(X) \rightarrow \mathfrak{S}(Y)$ such that

$$\int_Y g \, dK^\bullet(\mu) = \int_X \int_Y g(y) K(x)(dy) \mu(dx)$$

holds for each $g \in \mathcal{M}(Y)$.

3. $(\mu \otimes K)(D) := \int_X K(x)(D_x) \mu(dx)$ defines a map $\mathfrak{S}(X) \times \mathfrak{S}\mathfrak{B}(X, \mathfrak{S}(Y)) \rightarrow \mathfrak{S}(X \times Y)$ such that

$$\int_{X \times Y} g \, d(\mu \otimes K) = \int_Y \int_X g(x, y) K(x)(dy) \mu(dx)$$

is true whenever $g \in \mathcal{M}(X \times Y)$.

Since the integral in property 1 changes variables, it is sometimes referred to as the *Change of Variables formula*. Property 3 uses $D_x := \{y \in Y \mid \langle x, y \rangle \in D\}$ for the measurable set $D \subseteq X \times Y$; it gives the integral over a product as repeated integrals and contains the Fubini Theorem as special case.

Note that $f^\flat(\mu)$ is $\mathfrak{S}(f)(\mu)$, the former notation being somewhat more light-handed; K^\bullet is just the forgetful functor from the Kleisli category of the Giry monad, and the tensor construction in the third part arises from the tensorial strength of the monad.

Example 2. Illustrating these constructions through the discrete case, assume that $p : \{1, \dots, n\} \rightsquigarrow \{1, \dots, m\}$ is a stochastic relation, and let $\mu \in \mathfrak{S}(\{1, \dots, n\})$ be an initial distribution. Then

1. $f^\flat(\mu)(j) = \sum_{f(i)=j} \mu(i)$ is the probability that $f : \{1, \dots, n\} \rightarrow \{1, \dots, m\}$ hits the value j .
2. $p^\bullet(\mu)(j) = \sum_{i=1}^n \mu(i) \cdot p(i, j)$ is the probability that response j is produced, given the initial probability μ .
3. $(\mu \otimes p)(\langle i, j \rangle) = \mu(i) \cdot p(i, j)$ gives the probability for the input/output pair $\langle i, j \rangle$ to occur, given the initial probability μ (which is responsible for input i), and the probability $p(i, j)$ for output j after input i .

These properties are easily established using elementary computations. \diamond

Some properties of the general constructions are collected for the reader's convenience:

1. $(K;L);M = K;(L;M)$,
2. $(K;L)^\bullet = K^\bullet \circ L^\bullet$ (where \circ denotes the usual composition of maps),
3. for $f \in \mathcal{M}(Z)$ and for $x \in X$ the equality

$$\int_Z f \, d(K;L)(x) = \int_Y \int_Z f(z) L(y)(dz) K(x)(dy)$$

holds.

4. $K; \mathbb{I}_Y = K$ and $\mathbb{I}_X; K = K$, where $\mathbb{I}_X : X \rightsquigarrow X$ is the unit kernel on X which is defined by

$$\mathbb{I}_X(x)(A) := \delta_x(A) := \text{if } x \in A \text{ then } 1 \text{ else } 0 \text{ fi.}$$

It is remarkable that the construction in Def. 1. 3 can be reversed, and this is in fact the cornerstone for constructing the converse of a stochastic relation, as will be seen in Sect. 4. Reversing the construction means that each measure on the product of two SB-spaces can be represented as a product from the measure of a measure and a stochastic relation (or, putting it in terms of Def.1: the map $\langle \mu, K \rangle \mapsto \mu \otimes K$ defined in part 3 between $\mathbf{S}(X) \times \mathfrak{S}\mathfrak{B}(X, \mathbf{S}(Y))$ and $\mathbf{S}(X \times Y)$ is onto).

Proposition 1. *Let X and Y be SB-spaces, and $\zeta \in \mathbf{S}(X \times Y)$. Then there exists a stochastic relation $K : X \rightsquigarrow Y$ such that $\zeta = \pi_{X \times Y, X}^b(\zeta) \otimes K$, $\pi_{X \times Y, X}$ denoting the projection from $X \times Y$ to X .*

Proof. [16, Theorem V.8.1]. \square

The stochastic relation K is uniquely determined up to sets of μ -measure zero; it is known as the *regular conditional distribution of π_Y given π_X* , cf. [16, Ch. V.8]. We will call K a *version of the disintegration of ζ w.r.t. $\pi_{X \times Y, X}^b(\zeta)$* .

Example 3. Let $\zeta \in \mathbf{S}(\{1, \dots, n\} \times \{1, \dots, m\})$, then the probability $p(i, j)$ for input i generating output j is the probability $\zeta(\langle i, j \rangle)$ for the pair $\langle i, j \rangle$ to occur conditioned on the probability $\sum_{t=1}^m \zeta(\langle i, t \rangle)$ that input i is produced at all. Thus relation p satisfies the equation

$$\zeta(\langle i, j \rangle) = \left(\sum_{t=1}^m \zeta(\langle i, t \rangle) \right) \cdot p(i, j).$$

This is the discrete version of Prop. 1. In contrast to the discrete case, however, the version of the disintegration of ζ with respect to its projection usually cannot be computed explicitly in the general case. \diamond

There is a rather helpful interplay between the projection of $\mu \otimes K$ to the second component and $K^\bullet(\mu)$ which will be exploited later on:

Observation 1 *If $\mu \in \mathbf{S}(X)$ is a sub-probability measure, and $K : X \rightsquigarrow Y$ is a stochastic relation, the equality $\pi_{X \times Y, Y}^b(\mu \otimes K) = K^\bullet(\mu)$ holds.*

3 Nondeterminism: The Fringe Relation

Probabilistic modelling is a special case of nondeterministic modelling: we do not only indicate possible outcomes but also assign a weight to them. Thus it comes as a natural construction that each stochastic relation defines a set-valued relation, at least on Polish spaces. This relation is defined now, and we will have

a look at the correspondence between both types of relations. We will assume in this section that the SB-spaces are endowed with a fixed Polish topology.

The support $\text{supp}(\mu)$ of a probability measure $0 \neq \mu \in \mathbf{S}(X)$ is the set of all points $x \in X$ such that each open neighborhood U of x has positive measure. This set is the smallest closed set F with $\mu(F) = \mu(X)$, it is denoted by $\text{supp}(\mu)$; for completeness, $\text{supp}(0) := \emptyset$ is defined for the zero measure $0 \in \mathbf{S}(X)$.

We investigate the set valued map $x \mapsto \text{supp}(K(x))$, when K is a transition probability from the Polish space X to the Polish space Y . This map is the relational counterpart to a stochastic relation, as we will see. It is clear that the map takes values in the set of all closed nonempty subsets of a Polish space, and that for any open subset U of Y the set

$$\text{supp}(K(\cdot))^{-1}[U] = \{x \in X \mid K(x)(U) > 0\}$$

is a measurable subset of X .

As an aside, we look at **supp** from an algebraic point of view. Polish spaces with continuous maps form the category \mathfrak{Pol} . For the Polish space X the space of all its probabilities $\mathbf{S}(X)$ is also a Polish space. We denote by \mathbf{SubPol} the subcategory whose objects are all spaces $\mathbf{S}(X)$ when X ranges over Polish spaces. A morphism $K : \mathbf{S}(X) \rightarrow \mathbf{S}(Y)$ is a continuous map between $\mathbf{S}(X)$ and $\mathbf{S}(Y)$ when both spaces carry their weak topologies. Following a result due to Giry [11, Theorem 1], the functor \mathbf{G} which assigns each Polish space its space of sub-probability measures is the functorial part of a monad in \mathfrak{Pol} . Denote by \mathbf{G}_f the composition of \mathbf{G} with the forgetful functor $\mathbf{SubPol} \rightarrow \mathbf{Set}$, the latter denoting the category of sets with maps as morphisms.

Let $\mathbf{F}(X)$ be the space of all nonempty closed subsets for a Polish space X , endowed with the Vietoris topology. This topology has as a subbase the sets $\{F \mid F \subseteq U_1\} \cap \{F \mid F \cap U_2 \neq \emptyset\}$ for the open sets $U_1, U_2 \subseteq X$.

Here things are a bit more complicated than in the probabilistic setting: if X is a compact metric space, so is $\mathbf{F}(X)$ [13, 4.9.12, 4.9.13]; if X is a Polish space, then the compacta in $\mathbf{F}(X)$ form a Polish space under the Vietoris topology. From [13, 4.9.7] it may be deduced that X is a compact metric space provided $\mathbf{F}(X)$ is a Polish space. Anyway, denote by \mathcal{CL} the category which has $\mathbf{F}(X)$ for Polish X as objects. A morphism $\mathbf{F}(f) := f^\# : \mathbf{F}(X) \rightarrow \mathbf{F}(Y)$ is induced by the continuous map $f : X \rightarrow Y$ through the topological closure of the images under closed sets, hence $f^\#(A) := (f[A])^{\text{cl}}$ is defined. Clearly, $f^\#$ is continuous in the Vietoris topology, since f is under the metric topology, and since $(g \circ f)^\# = g^\# \circ f^\#$, we see that $\mathbf{F} : \mathfrak{Pol} \rightarrow \mathcal{CL}$ is a functor. The discussion above indicates that \mathbf{F} is in general no monad in \mathfrak{Pol} (it is, however, when \mathfrak{Pol} is replaced by the category of all compact metric spaces). Consequently, it is not possible to relate both monads directly. A weaker result may be obtained, however.

Compose this functor with the forgetful functor $\mathcal{CL} \rightarrow \mathbf{Set}$ to obtain the functor \mathbf{F}_f .

Proposition 2. $\text{supp} : \mathbf{G}_f \xrightarrow{\bullet} \mathbf{F}_f$ is a natural transformation.

Hence the map $x \mapsto \text{supp}(K(x))$ relating a transition probability to a set of elements with positive probability is given by a natural transformation.

Leaving this side track, we define the fringe relation:

Definition 2. Let $K : X \rightsquigarrow Y$ be a stochastic relation between the Polish spaces X and Y . The fringe relation R_K associated with K is defined as

$$R_K := \{ \langle x, y \rangle \in X \times Y \mid y \in \text{supp}(K(x)) \}.$$

Conversely, let $R \subseteq X \times Y$ be a set-theoretic relation, then a stochastic relation $K : X \rightsquigarrow Y$ is said to satisfy R (abbreviated by $R \models K$) iff $R_K = R$ holds, hence iff R is just the fringe of K .

Example 4. Let $f : X \rightarrow Y$ be a measurable map between the Polish spaces X and Y , and put $K(x) := \delta_{f(x)}$, δ_y denoting as usual the Dirac measure on y . Then $K : X \rightsquigarrow Y$ is a stochastic relation for which $R_K = \text{Graph}(f)$ holds. \diamond

Investigating the relationship between the stochastic relation K and its fringe R_K , we find that composition carries over as follows:

Observation 2 Let $K : X \rightsquigarrow Y$ and $L : Y \rightsquigarrow Z$ be stochastic relations, then

1. $R_L \circ R_K \subseteq R_{K;L}$,
2. suppose that for each $x \in X$ the probability $K(x)(G)$ is positive for each open set $G \subseteq X$, then also $R_{K;L} \subseteq R_L \circ R_K$.

$R \models K$ indicates that, if R is the nondeterministic specification of a system, stochastic relation K is its probabilistic refinement. Define for $K, K' : X \rightsquigarrow Y$, and for $0 \leq p \leq 1$ the stochastic relation $K \oplus_p K'$ upon defining

$$(K \oplus_p K')(x)(B) := p \cdot K(x)(B) + (1 - p) \cdot K'(x)(B),$$

(thus $(K \oplus_p K')(x)$ is just the convex combination of the measures $K(x)$ and $K'(x)$). The operator \oplus_p is interpreted as a weighted choice operator. It is easy to see that the following holds:

$$\frac{R \models K \quad R \models K' \quad 0 \leq p \leq 1}{R \models (K \oplus_p K')}$$

Consequently, the set of all stochastic relations satisfying a given nondeterministic specification is convex, hence closed under weighted choice. Convexity models the observation that nondeterministic systems are underspecified, as compared to stochastic ones (cf. the discussion in [15]).

Each stochastic relation has a fringe, and the inverse correspondence can be established under suitable topological assumptions: Given a set-valued relation R , a stochastic relation K that satisfies R can be found. For this, R has to take closed values, and a measurability condition is imposed:

Proposition 3. *Let $R \subseteq X \times Y$ a relation (X, Y Polish) such that*

1. $\forall x \in X : R(x) := \{y \in Y \mid \langle x, y \rangle \in R\} \in \mathbf{F}(Y)$,
2. *whenever $U \subseteq Y$ is open, $\{x \in X \mid R(x) \cap U \neq \emptyset\}$ is a measurable subset of X .*

If Y is σ -compact, or if $R(x)$ assumes compact values for each $x \in X$, then there exists a stochastic relation $K : X \rightsquigarrow Y$ with $R \models K$.

Thus each set-valued relation can be represented by a stochastic one under the conditions stated above, so that each nondeterministic specification can be satisfied by a stochastic relation.

4 Converse Relations

Given a sub-stochastic matrix $(p(i, j))_{1 \leq i \leq n, 1 \leq j \leq m}$ representing a stochastic relation $\{1, \dots, n\} \rightsquigarrow \{1, \dots, m\}$ and an initial distribution, the Introduction shows that the probability $p_\mu^\sim(j)(i)$ of responding with $j \in \{1, \dots, m\}$ on a stimulus $i \in \{1, \dots, n\}$ is calculated as

$$p_\mu^\sim(j)(i) = \frac{\mu(i) \cdot p(i, j)}{\sum_t \mu(t) \cdot p(t, j)}.$$

The probability p_μ^\sim under consideration reverses p given an initial distribution, so is regarded as the converse of p (*inverse* might at first sight be considered a better name, but this seems to suggest invertibility of the matrix associated with p).

In view of Examples 3 and 2, this amounts to the disintegration of $\mu \otimes p$ with respect to the distribution $p^\bullet(\mu) = \pi_{X \times Y, Y}^b(\mu \otimes p)$.

This observation guides the way for the definition of the converse for a general stochastic relation. Fix a stochastic relation $K : X \rightsquigarrow Y$, and a sub-probability measure $\mu \in \mathbf{S}(X)$. Then $\mu \otimes K \in \mathbf{S}(X \times Y)$ has a kind of natural converse: define $\tau := r^b(\mu \otimes K)$, where $r : X \times Y \rightarrow Y \times X$ switches components. Thus $r[R] = R^\sim := \{\langle y, x \rangle \mid \langle x, y \rangle \in R\}$, whenever $R \subseteq X \times Y$ is a relation, so r produces the converse. Because $\tau \in \mathbf{S}(Y \times X)$, this measure is — according to Prop. 1 — representable through a stochastic relation $K_\mu^\sim : Y \rightsquigarrow X$ by writing $\tau = \pi_{Y \times X, X}^b(\tau) \otimes K_\mu^\sim$. Since $\pi_{Y \times X, Y}^b(\tau) = K^\bullet(\mu)$ by Obs. 1, the definition of the converse of a stochastic relation now reads as follows.

Definition 3. *The μ -converse K_μ^\sim of the stochastic relation K with respect to the input probability μ is defined by the equation $r^b(\mu \otimes K) = K^\bullet(\mu) \otimes K_\mu^\sim$.*

It is remarked that the converse K_μ^\sim always exists, and that it is unique μ -almost everywhere. Since

$$\mu(A) = (\mu \otimes K)(A \times Y) = (K^\bullet(\mu) \otimes K_\mu^\sim)((Y \times A)^\sim)$$

is true for the Borel set $A \subseteq X$,

$$\mu(A) = \int_X \int_Y K_\mu^\sim(A) K(x)(dy) \mu(dx) = \int_Y K_\mu^\sim(A) K^\bullet(\mu)(dy),$$

we infer that $\mu = (K_\mu^\sim)^\bullet(K^\bullet(\mu)) = (K; K_\mu^\sim)^\bullet(\mu)$ holds. Hence the converse K_μ^\sim solves the equation $\mu = (K; T)^\bullet(\mu)$ for T . This equation does, however, not determine the converse uniquely. This is so because it is an equation in terms of the Borel sets of X , hence may only be carried over to the “strip” $\{A \times Y \mid A \subseteq X \text{ Borel}\}$ on the product $X \times Y$. This is not enough to determine a measure on the entire product.

A *probabilistic interpretation* using regular conditional distributions may be given as follows: Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $\zeta_i : \Omega \rightarrow X_i$ random variables with values in the Polish spaces X_i ($i = 1, 2$). Let μ be the joint distribution of $\langle \zeta_1, \zeta_2 \rangle$, and let μ_i be the marginal distribution of ζ_i . If $\pi_i : X_1 \times X_2 \rightarrow X_i$ are the projections, then clearly $\mu_i = \pi_i(\mu)$. K denotes the regular conditional distribution of ζ_2 given ζ_1 , thus we have for the Borel sets $A_i \subseteq X_i$

$$\begin{aligned} \mathbb{P}(\{\omega \in \Omega \mid \zeta_1(\omega) \in A_1, \zeta_2(\omega) \in A_2\}) &= \mu(A_1 \times A_2) \\ &= \int_{A_1} K(x_1)(A_2) \mu_1(dx_1). \end{aligned}$$

We will show now that $K_{\mu_1}^\sim$ is the regular conditional distribution of ζ_1 given ζ_2 . In fact, let L be the latter distribution, then the definitions of K and L , resp., imply $K^\bullet(\mu_1) = \mu_2$ and $L^\bullet(\mu_2) = \mu_1$. Let $A_i \subseteq X_i$ be Borel sets, then $(K^\bullet(\mu_1) \otimes L)(A_2 \times A_1) = (\mu_1 \otimes K)(A_1 \times A_2)$.

Interpreting a stochastic relation as a regular conditional distribution of a random variable ζ_1 given ζ_2 , its converse may be interpreted as the conditional distribution of ζ_2 given ζ_1 . The start probability μ in the definition of K_μ^\sim is then interpreted as a marginal distribution. This is essentially the probabilistic setting for the definition of the converse in [1].

Returning to the general case, the defining equation for the converse is spelled out in terms of an integral (where $D^x := \{y \in Y \mid \langle y, x \rangle \in D\}$ for $D \subseteq Y \times X$, the cut D_y is defined above):

$$\int_X K(x)(D^x) \mu(dx) = \int_Y K_\mu^\sim(y)(D_y) K^\bullet(\mu)(dy).$$

This will be generalized and made use of later:

Observation 3 *Let $f \in \mathcal{M}(X \times Y)$, then this identity holds:*

$$\int_X \int_Y f(x, y) K(x)(dy) \mu(dx) = \int_Y \int_X f(x, y) K_\mu^\sim(y)(dx) K^\bullet(\mu)(dy).$$

Thus the order of integration of f may be interchanged, as in Fubini’s Theorem, but, unlike that Theorem, we have to adjust the measures used for integration.

Some properties of forming the converse will be investigated now. We begin with an analogue of the property $R^{\smile} = R$ which holds for the set theoretic converse. Taking the initial distribution into account, this property is very similar for the probabilistic case.

Proposition 4. *If $K : X \rightsquigarrow Y$, and if $\mu \in \mathbf{S}(X)$, then $(K_{\mu}^{\smile})_{K^{\bullet}(\mu)}^{\smile} = K$. holds everywhere except possibly on a set of μ -measure zero.*

The question under what condition a stochastic relation may be represented as the converse of another relation is a little more difficult to answer than for the set-valued case. In view of the probabilistic interpretation using conditional distributions, however, the following solution arises naturally.

Corollary 1. *Let $L : Y \rightsquigarrow X$ be a stochastic relation, and $\mu \in \mathbf{S}(X)$. Then these conditions are equivalent:*

1. $\mu = L^{\bullet}(\nu)$ for some $\nu \in \mathbf{S}(Y)$,
2. $L = K_{\mu}^{\smile}$ for some $K : X \rightsquigarrow Y$. \square

Thus $L : Y \rightsquigarrow X$ may be written in a variety of ways as the converse of a stochastic relations, viz., $L = (K_{\nu})_{L^{\bullet}(\nu)}^{\smile}$ for an arbitrary $\nu \in \mathbf{S}(Y)$ (where the relation $X \rightsquigarrow Y$ depends on ν). This is in marked contrast to the set-theoretic case, where the converse of the converse of a relation is the relation itself, hence unique.

Compatibility of composition and forming the converse is an important property in the world of set-theoretic relations. In that case it is well known that $(R;S)^{\smile} = S^{\smile};R^{\smile}$ always holds. The corresponding property for stochastic relations reads

Proposition 5. *Let $K : X \rightsquigarrow Y, L : Y \rightsquigarrow T$ be stochastic relations, and let $\mu \in \mathbf{S}(X)$ be an initial distribution. Then $(K;L)_{\mu}^{\smile} = L_{K^{\bullet}(\mu)}^{\smile};K_{\mu}^{\smile}$ holds.*

We see that there are some algebraic similarities between set-theoretic and stochastic relations. There are exceptions, though. Take e.g. *Schröder's Cycle Rule* $Q \circ R \subseteq S \Leftrightarrow Q^{\smile} \circ \bar{S} \subseteq \bar{R} \Leftrightarrow \bar{S} \circ R^{\smile} \subseteq \bar{Q}$, the bar denoting complementation ([18, 3.2 (xii)] or [4, Def. 3.1.1]). This rule is very helpful in practical applications, but it does not enjoy a direct counterpart for stochastic relations, since the respective notions of negation, and of containment do not carry over. —

If $\mu(A) = 0$ for some Borel set $A \subseteq X$, then $K_{\mu}^{\smile}(y)(A) = 0$ holds $K^{\bullet}(\mu)$ -almost everywhere on Y (i.e., for all $y \in Y$ outside a set of $K^{\bullet}(\mu)$ -measure zero). In fact, we can say more by scrutinizing the relationship between K_{μ}^{\smile}, K and μ . This leads to a rather surprising compactness result of the set of measures comprising the converse.

Recall that for $\mu, \nu \in \mathbf{S}(X)$ the measure ν is called *absolutely continuous* w. r. t. μ iff for every measurable set $A \subseteq X$ the implication $\mu(A) = 0 \Rightarrow \nu(A) = 0$ holds; this is indicated by $\nu \ll \mu$. It is well known [3, Sect. 32] that $\nu \ll \mu$ is equivalent to

$$\forall \varepsilon > 0 \exists \delta > 0 : [\mu(A) < \delta \Rightarrow \nu(A) < \varepsilon].$$

Absolute continuity is used for defining morphisms between probability spaces based on Polish spaces in [1, Def. 7.8] which in turn serves for defining the converse of a stochastic relation; we use it here for characterizing the measures comprising the converse. A subset $M \subseteq \mathbf{S}(X)$ is accordingly called *uniformly absolutely continuous* w.r.t. μ (indicated by $M \ll \mu$) iff given $\varepsilon > 0$ there exists $\delta > 0$ such that $\sup_{\nu \in M} \nu(A) < \varepsilon$ whenever $\mu(A) < \delta$ holds. It will be shown now that the set of measures constituting the converse is uniformly absolutely continuous except on a very small set:

Proposition 6. *Let $K : X \rightsquigarrow Y$ be a stochastic relation, and $\mu \in \mathbf{S}(X)$. Then for each version K_μ^\smile of the converse of K with respect to μ there exists a Borel set $A \subseteq Y$ for which $K(x)(A) = 0$ is true for μ -almost all $x \in X$ so that $\{K_\mu^\smile(y) | y \notin A\} \ll \mu$ holds.*

This implies that the set $\{K_\mu^\smile(y) | y \notin A\}$ is topologically not too large. Since we deal with a specific topology on the set of all sub-probability measures, we fix a Polish topology on the input space which in turn induces the topology of weak convergence on $\mathbf{S}(X)$.

Corollary 2. *Let X be a Polish space, endow $\mathbf{S}(X)$ with the topology of weak convergence, and let Y be an SB-space. Given $K : X \rightsquigarrow Y$ and $\mu \in \mathbf{S}(X)$, there exists a Borel set $A \subseteq Y$ with $K(x)(A) = 0$ for μ -almost all $x \in X$ so that the set $\{K_\mu^\smile(y) | y \notin A\}$ is a relatively compact subset of $\mathbf{S}(X)$.*

Finally, let us have a look at the fringe relation: it turns out that $(R_K)^\smile$ does not necessarily coincide with $R_{K_\mu^\smile}$.

Example 5. Define K as in Example 4, then $\mu \otimes K = id_X \times f^b(\mu)$ holds for $\mu \in \mathbf{P}(X)$, so that $\pi_Y^b(\mu \otimes K) = f^b(\mu)$ is inferred. For the Borel sets $A \subseteq X, B \subseteq Y$ the equalities

$$(\mu \otimes K)(A \times B) = \mu(A \cap f^{-1}[B]) = (f^b(\mu) \otimes K_\mu^\smile)(B \times A).$$

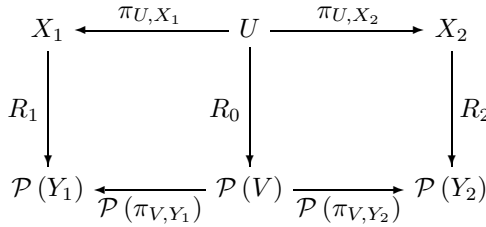
hold. Now put $\mu = \delta_{x'}$ for some $x' \in X$, then the constant relation $K_\mu^\smile(y) = \delta_{x'}$ is a version of the converse, hence

$$R_{K_\mu^\smile} = \{\langle y, x' \rangle | y \in Y\} \neq Graph(f)^\smile = (R_K)^\smile.$$

Thus building the fringe relation and forming the converse does not commute. \diamond

5 Bisimulations

Call the relations $R_1 \subseteq X_1 \times Y_1$ and $R_2 \subseteq X_2 \times Y_2$ bisimilar iff there exists $U \subseteq X_1 \times X_2$ and $V \subseteq Y_1 \times Y_2$ and a relation $R_0 \subseteq U \times V$ (U, V and R_0 are called *mediating*) such that this diagram is commutative:



Here \mathcal{P} is the powerset functor, and relations are interpreted as set valued maps. This is the definition of bisimilarity for coalgebras [17, 2] adapted to the situation at hand.

Defining for a relation R the *yield* relation

$$x \vdash_R y \iff \langle x, y \rangle \in R$$

in analogy to the transition relation \rightarrow_R^a investigated for coalgebras, it is easy to see that R_1 and R_2 are bisimilar iff

1. for all $\langle x_1, x_2 \rangle \in U$, if $x_1 \vdash_{R_1} y_1$, then there exists $y_2 \in Y_2$ such that $\langle y_1, y_2 \rangle \in V$ and $x_2 \vdash_{R_2} y_2$,
2. for all $\langle x_1, x_2 \rangle \in U$, if $x_2 \vdash_{R_2} y_2$, then there exists $y_1 \in Y_2$ such that $\langle y_1, y_2 \rangle \in V$ and $x_1 \vdash_{R_1} y_1$.

In fact, Rutten’s proof [17, Ex. 2.1] carries over. Bisimulations will be studied now for stochastic relations, and the goal is to show that bisimilar relations give rise to bisimilar converses. We first define bisimilarity for stochastic relations and show that under a mild condition bisimilarity is transitive. Then we establish that the operations we are working with, i.e., forming products of measures and relations, and transporting measures through relations, maintain bisimilarity. This holds also for disintegration, and having established this, a small step will be necessary to show that converses will respect bisimilarity.

Bisimulations are usually defined through spans of morphisms in a suitable category. In fact, a stochastic relation $K : X \rightsquigarrow Y$ can be considered as an object $\langle X, Y, K \rangle$ in the comma category $\mathbf{1}_{SB} \downarrow \mathbf{S}$, where $\langle \alpha, \beta \rangle : \langle X, Y, K \rangle \rightarrow \langle X', Y', K' \rangle$ is a morphism iff $\alpha : X \rightarrow X'$ and $\beta : Y \rightarrow Y'$ are measurable such that $K \circ \alpha = \beta \circ K'$ holds. A *1-bisimulation* $\langle O, \Gamma_1, \Gamma_2 \rangle$ for objects O_1 and O_2 is then an object O together with two morphisms $\Gamma_1 : O \rightarrow O_1$ and $\Gamma_2 : O \rightarrow O_2$. This notion of bisimilarity was discussed and investigated in [8] and specialized there to the present notion of bisimulation (called *2-bisimulation* in [8]), which is similar in spirit to the one given above for set valued relations:

Definition 4. Let $K_1 : X_1 \rightsquigarrow Y_1$ and $K_2 : X_2 \rightsquigarrow Y_2$ be stochastic relations, where all participating spaces are SB-spaces. Then $N : U \rightsquigarrow V$ is called a bisimulation for K_1 and K_2 iff these conditions are satisfied:

1. $U \subseteq X_1 \times X_2$ and $V \subseteq Y_1 \times Y_2$ are SB-spaces,
2. $K_1 \circ \pi_{A,X_1} = \pi_{B,Y_1}^b \circ N$ and $K_2 \circ \pi_{A,X_2} = \pi_{B,Y_2}^b \circ N$ hold.

Standard arguments show that N is a bisimulation for K_1 and K_2 iff (in the notation of Def. 4)

$$\int_{Y_i} f_i \, dK_i(x_i) = \int_V f_i \circ \pi_{V,Y_i} \, dN(x_1, x_2)$$

hold for each pair $\langle x_1, x_2 \rangle \in U$, and for each $f_i \in \mathcal{M}(Y_i)$ ($i = 1, 2$). This condition is sometimes easier to handle.

Bisimulation turns out to be transitive under a rather mild condition of surjectivity. This property can be established using the existence of semi-pullbacks for stochastic relations (recall that a *semi-pullback* for a pair of morphisms $f_1 : a_1 \rightarrow c, f_2 : a_2 \rightarrow c$ in a category is a pair of morphisms $g_1 : b \rightarrow a_1, g_2 : b \rightarrow a_2$ with $f_1 \circ g_1 = f_2 \circ g_2$). The plan of attack is as follows: 1-bisimilarity is a transitive relation under the assumption of surjectivity [9, Theorem 2], and the comparison between 1-bisimilarity and bisimilarity from [8, Prop. 5] shows that both notions are equivalent under a condition of measurability. This technical condition which will be established here.

Proposition 7. *Let $K_i : X_i \rightsquigarrow Y_i$ ($i = 1, 2, 3$) be stochastic relations, and assume that $N_1 : U_1 \rightsquigarrow V_1$ and $N_2 : U_2 \rightsquigarrow V_2$ are bisimulations for K_1, K_2 and K_2, K_3 , resp. Assume that all projections are onto. Then there exists a bisimulation $N_3 : U_3 \rightsquigarrow V_3$ for K_1, K_3 .*

In order to show that bisimilar relations give rise to bisimilar converses, it is practical to introduce the notion of bisimilarity for sub-probability measures, too; it is easy to see that the same notion of bisimilarity arises when one restricts oneself to constant stochastic relations.

Definition 5. *Let X_1, X_2 be SB-spaces with $\mu_i \in \mathbf{S}(X_i)$ ($i = 1, 2$). Then $\langle X_1, \mu_1 \rangle$ is said to be bisimilar to $\langle X_2, \mu_2 \rangle$ iff there exists a subset $Z \subseteq X_1 \times X_2$ and $\zeta \in \mathbf{S}(Z)$ such that*

1. Z is a SB-space,
2. $\mu_1 = \pi_{Z, X_1}^b(\zeta)$ and $\mu_2 = \pi_{Z, X_2}^b(\zeta)$.

$\langle Z, \zeta \rangle$ is said to mediate between $\langle X_1, \mu_1 \rangle$ and $\langle X_2, \mu_2 \rangle$.

Bisimulations are maintained by forming products, and by transporting a measure through a stochastic relation, as we will see now:

Proposition 8. *Let $K_i : X_i \rightsquigarrow Y_i$ be bisimilar stochastic relations over the SB-spaces X_i, Y_i for $i = 1, 2$ such that $N : U \rightsquigarrow V$ mediates between them, and assume that $\mu_i \in \mathbf{S}(X_i)$ such that $\langle X_1, \mu_1 \rangle$ and $\langle X_2, \mu_2 \rangle$ are bisimilar with mediating $\langle Z, \zeta \rangle$. Assume that $Z \subseteq U$ holds. then*

1. $\langle Y_1, K_1^\bullet(\mu_1) \rangle$ is bisimilar to $\langle Y_2, K_2^\bullet(\mu_2) \rangle$ with mediating $\langle V, N^\bullet(\zeta) \rangle$,
2. $\langle X_1 \times Y_1, \mu_1 \otimes K_1 \rangle$ is bisimilar to $\langle X_2 \times Y_2, \mu_2 \otimes K_2 \rangle$ with mediating $\langle t[E], t^\flat(\zeta \otimes N) \rangle$, where $E := Z \times V$ and $t(x_1, x_2, y_1, y_2) := \langle x_1, y_1, x_2, y_2 \rangle$.

The argumentation above shows that bisimilar relations and bisimilar initial distributions lead to bisimilar measures on the product. The process can be reversed: the idea is that disintegrating bisimilar measures on a product leads to bisimilar stochastic relations.

Proposition 9. *Let X_i, Y_i be SB-spaces, $\mu_i \in \mathbf{S}(X_i \times Y_i)$ for $i = 1, 2$. Assume that $\langle X_1 \times Y_1, \mu_1 \rangle$ is bisimilar to $\langle X_2 \times Y_2, \mu_2 \rangle$. Define for $i = 1, 2$ the stochastic relations $K_i : X_i \rightsquigarrow Y_i$ as the disintegrations of μ_i w.r.t $\pi_{X_i \times Y_i, X_i}^b(\mu_i)$. Then K_1 is bisimilar to K_2 .*

Showing that bisimilarity is maintained when forming the converse is now an easy consequence:

Corollary 3. *Under the assumptions of Prop. 8, K_{1, μ_1}^\sim is bisimilar to K_{2, μ_2}^\sim .*

6 Related Work

The generalization of set-based relations to probabilistic ones appears straightforward: replace the nondeterminism inherent in these relations by randomness. Panangaden [14] carries out a very elegant construction, arguing as follows: the powerset functor is a monad which has relations as morphisms in its Kleisli category [12], the functor that assigns each measurable space the set of all (sub-) probability measures is also a monad having transition probabilities as morphisms in its Kleisli category [11]. This parallel justifies their characterization as probabilistic relations. The category **SRel** of measurable spaces with transition sub-probabilities is scrutinized closer in [14], and an application to Kozen's semantics of probabilistic programs is given. Stochastic relations are underlying stochastic automata; they were introduced and investigated in [7] as a generalization of finite stochastic machines. Abramsky, Blute, and Panangaden [1] investigate the category **PRel** of probability spaces, hereby introducing the converse of a probabilistic relation as we do through the product measure (Cor. 7.7). The process by which they arrive at this construction (Theorem 7.6) is quite similar to disintegration, as proposed here but makes heavier use of absolute continuity (in fact, morphisms in **PRel** use absolute continuity in a crucial way). The argumentation in the present paper seems to be closer to the set-theoretic case by looking at what happens when we compute the probability for a converse relation. Further investigations of the converse do not include the anti-commutative law. This is probably due to the fact that integration techniques are directly used in the present paper (while [1] prefers arguing with absolute continuity, and consequently, with the Radon-Nikodym Theorem).

The notion of bisimilarity is — as in [8] — adapted from [6, 17] to the situation at hand. Transitivity of bisimulation is demonstrated in [10] for universally measurable stochastic relations case, but left open for the general case of Borel measurable transition probabilities; [9] gives a full solution to this problem.

The observation that each transition probability on a Polish space spawns a measurable set-valued function through the support function, hence a relation,

was used in [7] for investigating the relationship between nondeterministic and stochastic automata. It could be shown that each nondeterministic automaton can be represented through a stochastic one, and that this representation is preserved through the sequential work of the automata. Measurable selections play a major role, but the results are not formulated in terms of monads or categories.

7 Conclusion

Stochastic relations are generalizations of Markov processes. The converse of a stochastic relation is investigated, in particular it is shown that it satisfies some of the algebraic laws which rule their set-theoretic counterparts. Those relations that arise as converses are characterized, and it is shown that the set of all sub-probabilities comprising the converse is topologically quite small, i.e., is relatively compact in the weak topology. It is demonstrated that forming the converse does respect bisimulations — if the models one starts with are bisimilar, the converses will be, too. For a special case which includes the reals it is shown that bisimilarity of stochastic relations is a transitive relation; the proof makes use of the fact that semi-pullbacks exist in the corresponding category.

The nondeterminism inherent in a stochastic relation is identified, and it could be shown that nondeterministic and stochastic relations are related via a natural transformation that is induced by the support of finite measures. It is shown that the stochastic relations satisfying a nondeterministic one is convex, so that a nondeterministic specification provides a large degree of freedom for probabilistic satisfaction.

Further work will address the characterization of bisimilarity more closely in order to find necessary and sufficient conditions indicating under which two probabilistically related components are bisimilar.

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] P. Aczel and N. Mendler. A final coalgebra theorem. In H. H. Pitt, A. Poigne, and D. E. Rydeheard, editors, *Category Theory and Computer Science*, volume 389 of *Lecture Notes in Computer Science*, pages 357 – 365, 1989.
- [3] P. Billingsley. *Probability and Measure*. John Wiley and Sons, New York, 3 edition, 1995.
- [4] C. Brink, W. Kahl, and G. Schmidt, editors. *Relational Methods in Computer Science*. Advances in Computing Science. Springer-Verlag, Wien, New York, 1997.
- [5] D. Cantone, E. G. Omodeo, and A. Policriti. *Set Theory for Computing*. Springer-Verlag, 2001. In print.
- [6] J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation of labelled Markov-processes. Technical report, School of Computer Science, McGill University, Montreal, 1998.

- [7] E.-E. Doberkat. *Stochastic Automata — Nondeterminism, Stability, and Prediction*, volume 113 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, 1981.
- [8] 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.
- [9] E.-E. Doberkat. Semi-pullbacks and bisimulations in categories of stochastic relations. Technical Report 130, Chair for Software Technology, University of Dortmund, November 2002.
- [10] A. Edalat. Semi-pullbacks and bisimulation in categories of Markov processes. *Math. Struct. in Comp. Science*, 9(5):523 – 543, 1999.
- [11] 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.
- [12] S. Mac Lane. *Categories for the Working Mathematician*. Number 5 in Graduate Texts in Mathematics. Springer-Verlag, Berlin, 2 edition, 1997.
- [13] E. Michael. Topologies on spaces of subsets. *Trans. Am. Math. Soc.*, 71(2):152 – 182, 1951.
- [14] 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.
- [15] P. Panangaden. Does combining nondeterminism and probability make sense? *Bulletin of the EATCS*, (75):182 – 189, Oct. 2001.
- [16] K. R. Parthasarathy. *Probability Measures on Metric Spaces*. Academic Press, New York, 1967.
- [17] 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.
- [18] A. Tarski and S. Givant. *A Formalization of Set-Theory Without Variables*. Number 42 in Colloquium Publications. American Mathematical Society, Providence, R. I., 1987.