The Converse of a Probabilistic Relation

Ernst-Erich Doberkat
Chair for Software Technology
Universität Dortmund, Germany
doberkat@acm.org

10th June 2001

Abstract

Transition probabilities are proposed as the stochastic counterparts to set-based relations. We propose the construction of the converse of a probabilistic relation. It is shown that two of the most useful properties carry over: the converse is idempotent, and anticommutative. The last property is shown to hold relative to some initial probability measure. This paper investigates the relation between stochastic and set-based relations through the support function of probabilities.

1 Introduction

Relational methods are ubiquitous in Mathematics, Logic and Computer Science, they go back as far as Schröder's work [13]. Ongoing work with a focus on program specification may be witnessed from the wealth of material collected in [2]. The map calculus [3] shows that these methods determine an active line of research in Logic.

Specifying system properties using relations determine usually some kind of input-output-behavior. This is not always entirely adequate because it is in some sense too nondeterministic. Hence it may be more adequate to say that on input x an output y_1 is produced with probability 0.7, and output y_2 is produced with probability 0.02 rather than saying that on input x the outputs y_1, y_2 may be produced without giving a qualitative assessment.

This paper studies probabilistic relations. It is shown that these relations have converses, and it is demonstrated that the converse behaves nearly like its nondeterministic counterpart. The corresponding law for relations R and S reads

$$(R;S)^{\smile} = S^{\smile};R^{\smile}.$$

Working in a probabilistic setting, we require an initial probability distribution in order to get going. If K and L are probabilistic relations, and if the initial probability is μ, then the converse $K^{\smile}_{\mu}$ depends on μ. After K is done, the initial probability is transformed by K into $K^{\bullet}(\mu)$ for the next step which in turn is modelled by L, so its converse is $L^{\smile}_{K^{\bullet}(\mu)}$. On the other hand, the composed system starts with μ and has the converse $(K;L)^{\smile}_{\mu}$. Thus the simple relational equation translates into

$$(K;L)^{\smile}_{\mu} = L^{\smile}_{K^{\bullet}(\mu)};K^{\smile}_{\mu}$$

A large part of this paper is devoted to a proof of this identity which essentially depends on the possibility of decomposing probabilities on a product space into a start probability and a probabilistic relation.

This is only possible under suitable topological assumptions. We work in this paper under the assumptions that all probabilities are defined on Polish ($\equiv$ topologically complete and separable metric) spaces; Polish spaces have the disintegration property mentioned above.

We will also have a look at the relation between a probabilistic relation K between X and Y, and the non-deterministic relation R_K induced by K. It is described as

$$R_K := \{\langle x, y\rangle \mid x \in X, y \in \mathsf{supp}(K(x))\},$$

where $\mathsf{supp}(K(x))$ is the support of the probability measure $K(x)$, hence the set of all y such that $K(x)(U) > 0$ holds for each open neighborhood of y. $\mathsf{supp}(K(x))$ being closed, R_K is the graph of a measurable relation between X and Y. Thus we get a nondeterministic relation for free, when we define a probabilistic one! The relationship between these relations is investigated, we show that supp is a natural transformation between the functor which assigns to each Polish space its probability measures and the one which assigns it its closed and nonempty subsets. We show that under suitable topological assumptions one gets a probabilistic relation for free from a nondeterministic one.

The rest of the paper is organized as follows: in Sect. 2 we collect some notions and basic techniques from measure theory, the support function in investigated in sect. 3. Probabilistic relations are introduced formally in sect. 4, and some relationships between them and their nondeterministic sidekicks are investigated (for example, the well known Peirce product results through a construction). The converse relation is defined and investigated in sect. 5, Proposition 8 giving the result indicated above. Section 6 discusses related work, in particular the approach proposed by Panangaden and his coworkers, and sect. 7 suggests a conclusion and indicates further work to be done.

Acknowledgements. This work was done while the author was visiting the Computer Science Group of the Dipartimento di Matematica Pura ed Applicata at the University of L'Aquila. This research was in part supported through grants from the *Exchange Programme for Scientists between Italy and Germany* from the Italian Minstry 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.

2 Preliminaries

A topological space is called a *Polish space* iff it is a complete and separable metric space. These spaces are most convenient from a measure theoretic point of view.

Let X be a Polish space, then we equip X always with the σ-algebra $\mathcal{B}_X$ of all Borel sets, hence the σ-algebra generated from the open sets. Probability measures on X are always defined on $\mathcal{B}_X$. The set $\mathsf{Prob}(X)$ of all probability measures is itself a Polish space, when endowed with the topology of weak convergence, which in turn is metrizable by Prohorov's metric.

δ denotes as usual the Dirac kernel

$$\delta_x(A) := (x \in A \ ? \ \mathtt{1} \ : \mathtt{0}),$$

hence δ provides a continuous map $P \to \mathsf{Prob}(P)$.

The Borel sets for the weak topology is also the initial σ-algebra for all evaluations

$$\mathsf{ev}_B : \mu \mapsto \mu(B)$$

for all Borel sets B in X (or for some generator of $\mathcal{B}_X$).

For a second Polish space Y, a *transition probability* (sometimes called a *Markov kernel* or just a *kernel*) K from X to Y may be described as a $\mathcal{B}_X - \mathcal{B}_{\mathsf{Prob}(Y)}$ measurable map $K : X \to \mathsf{Prob}(Y)$, or, equivalently, as a map

$$K : X \times \mathcal{B}_Y \to [0, 1]$$

with the following properties:

- $\forall x \in X : K(x) \in Prob(Y)$,
- $\forall B \in \mathcal{B}_Y : x \mapsto K(x)(B)$ constitutes a measurable map to the unit-interval (which carries always the Borel sets).

We denote a transition probability K from X to Y by $K : X \rightsquigarrow Y$.

For the probability $\mu \in \mathsf{Prob}(X)$ and the measurable map $\varphi : X \to Y$ (where Y is another measurable space) the measure μ may be transported through φ upon setting for the measurable subset B of Y

$$\varphi(\mu)(B) := \mu(\varphi^{-1}[B]).$$

Then $\varphi(\mu)$ is a probability on Y, and the integral w.r.t. $\varphi(\mu)$ is easily calculated through the integral for μ: whenever $g : Y \to \mathbb{R}$ is measurable and bounded, we have

$$\int_Y g \; d\varphi(\mu) = \int_X g \circ \varphi \; d\mu.$$

This will be referred to as the *integral transform*. The proof of this equality is somewhat typical for the establishment of statements dealing with this kind of maps: the equation in question is true, if $g = 1_A$ holds, when $A \subseteq Y$ is measurable with 1_A as the indicator function of A. The set F of all functions g for which the above property holds is a vector space which contains the indicator functions of measurable sets, hence their linear combinations and monotone limits (by the Dominated Convergence Theorem). Consequently, F contains all bounded measurable maps. We will often make use of this argumentation in what follows, without explicitly repeating it.

3 The Support Function

The support $\mathsf{supp}(\mu)$ of a probability measure $\mu \in \mathsf{Prob}(X)$ 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 $\mu(F) = 1$, it is denoted by $\mathsf{supp}(\mu)$. We want to investigate the set valued map $x \mapsto \mathsf{supp}(K(x))$, when K is a transition probability from the Polish space X to the Polish space Y. This map is the relational counterpart to a probabilistic relation, as we will see. It is apparent that the map takes values in the set of all closed nonempty subsets of a Polish space, and that the set $\mathsf{supp}(K(\cdot))^{-1}[U] = \{x \in X | \; K(x)(U) > 0\}$ is a measurable subset of X, whenever $U \subseteq Y$ is open.

We need the characterization of the topology of weak convergence through the well known Portmanteau Theorem [12]:

Proposition 1 *The following conditions are equivalent for a sequence $(\mu_n)_{n\geq 0}$ of probability measures on the Polish space X:*

1. *μ_n converges weakly to μ_0 (in signs: $\mu_n \rightharpoonup_w \mu_0$),*
2. *$\int_X f d\mu_n \to \int_X f d\mu_0$ for each bounded and continuous $f : X \to \mathbb{R}$,*
3. *$\liminf_{n\to\infty} \mu_n(F) \leq \mu_0(F)$ for each closed subset $F \subseteq X$.*

Polish spaces with continuous maps form a category $\mathcal{P}$. For a Polish space X the space of all its probabilities $\mathsf{Prob}(X)$ is also a Polish space. We denote by $\mathcal{PR}$ the category which has $\mathsf{Prob}(X)$ for an object X in $\mathcal{P}$ as its objects. A morphism $K : \mathsf{Prob}(X) \to \mathsf{Prob}(Y)$ is a continuous map between $\mathsf{Prob}(X)$ and $\mathsf{Prob}(Y)$ when both spaces carry their weak topologies. According to a result due to Giry [7, Theorem 1] (cf. [11]), the functor $\mathbf{G}$ which assigns each

Polish space its space of probability measures is actually a monad in $\mathcal{P}$, the unit being the Dirac kernel, multiplication being given by

$$\mathsf{Prob}(\mathsf{Prob}(X)) \ni M \mapsto \left(\lambda B \in \mathcal{B}_X. \int_{\mathsf{Prob}(X)} \tau(B)\ M(d\tau)\right) \in \mathsf{Prob}(X).$$

Denote by $\mathbf{G}_f$ the composition of $\mathbf{G}$ with the forgetful functor $\mathcal{PR} \to \mathcal{SET}$, the latter being 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

$$\{F|\ F \subseteq \bigcup_{1 \leq i \leq n} U_i\} \cap \bigcap_{1 \leq i \leq n} \{F|\ F \cap U_i \neq \emptyset\}$$

for the open sets $U_1, \ldots, U_n \subseteq X$.

Here things are a bit more complicated than in the probabilistic setting: if X is a compact metric space, so is $\mathbf{F}(X)$ [10, 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 [4, Cor. II-9]. From [10, 4.9.7] it may be deduced that X is a compact metric space provided $\mathcal{F}(X)$ is a Polish space. Anyway, denote by $\mathcal{CL}$ the category which has $\mathbf{F}(X)$ for Polish spaces as objects. A morphism

$$\mathbf{F}(f) := f^\sharp : \mathcal{F}(X) \to \mathcal{F}(Y)$$

is induced by the continuous map $f : X \to Y$ through the topological closure of the images under closed sets:

$$f^\sharp(A) := (f[A])^{\mathsf{cl}}.$$

Apparently, $f^\sharp$ is continuous in the Vietoris topology, since f is under the metric topology, hence $\mathbf{F} : \mathcal{P} \to \mathcal{CL}$ is a functor. The discussion above indicates that $\mathbf{F}$ is in general no monad in $\mathcal{P}$ (it is, when $\mathcal{P}$ is replaced by the category of all compact metric spaces). Hence it is not possible to relate two Kleisli categories directly. A weaker result may be obtained, however.

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

Proposition 2 $\mathsf{supp} : \mathbf{G}_f \stackrel{\bullet}{\to} \mathbf{F}_f$ *is a natural correspondence.*

Proof: This follows from the fact that for each continuous map f between metric spaces X and Y and for each $\mu \in \mathsf{Prob}(X)$ the equality

$$(f[\mathsf{supp}(\mu)])^{\mathsf{cl}} = \mathsf{supp}(f(\mu))$$

holds. $\square$

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

This relationship may be analyzed under an other aspect: Given a set-valued relation R, find a transition probability K such that

$$\langle x, y \rangle \in R \Longleftrightarrow x \in \mathsf{supp}(K(x))$$

holds. Apparently, relation R has to satisfy some constraints:

Proposition 3 *Let for the Polish spaces X and Y be R a relation such that*

1. $\forall x \in X : R(x) := \{y \in Y |\ \langle x, y\rangle \in R\} \in \mathbf{F}(Y)$,
2. *whenever $U \subseteq Y$ is open, $\{x \in X |\ R(x) \cap U \neq \emptyset\}$ 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 transition probability K from X to Y such that $R(x) = \mathsf{supp}(K(x))$ holds for all $x \in X$.

Proof: This follows from [6, Cor. IV.11]. □
Let us turn to probabilistic relations.

4 Probabilistic Relations

Modelling a stochastic input-output system through a transition probability K, intuitively $K(x)(dy)$ gives the probability that an input $x \in X$ will produce output $y \in Y$. The system will thus produce on input of x an element of the Borel set $B \subseteq Y$ with probability $K(x)(B)$, so $\mathsf{supp}(K(x))$ may be considered the set of all possible outputs on input x.

Definition 1 *A* probabilistic relation *between the Polish spaces X and Y is a transition probability $K : X \rightsquigarrow Y$.*

Consequently, each probabilistic relation $K : X \rightsquigarrow Y$ induces a (set-theoretic) relation

$$R_K := \{\langle x, y\rangle |\ x \in X, y \in \mathsf{supp}(K(x))\}.$$

It is not difficult to see that R_K is a measurable subset of $X \times Y$.

The relation R_K then relates each input with all possible outputs. In the absence of the possibility of assigning probabilities to single elements, possible outputs are characterized by the positive probability for each open circle with positive radius around it.

The composition between relations is defined as to be expected: 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 indeed $K;L$ is a probabilistic relation between X and Z. The properties that followed are collected for the reader's convenience:

- $(K;L);M = K;(L;M)$,
- $(K;L)^\bullet = K^\bullet \circ L^\bullet$ (where the latter composition denotes compositions of maps between sets of probability measures),
- for the measurable and bounded map $f : Z \to \mathbb{R}$ and for $x \in X$ the integral $\int_Z f\ d(K;L)(x)$ is calculated as $\int_Y \int_Z f(z)\ L(y)(dz)\ K(x)(dy)$,
- $K;\mathbb{I}_Y = K$ and $\mathbb{I}_X;K = K$, where $\mathbb{I}_X : X \rightsquigarrow X$ is the unit kernel on X which is defined by $\mathbb{I}_X(x)(A) := \delta_x(A)$.

There is an interplay between probabilistic relations $X \rightsquigarrow Y$ and probability measures on $X \times Y$ that is of vital interest here. Let $K : X \rightsquigarrow Y$ be a probabilistic relation, and $\mu \in \mathsf{Prob}(X)$ be a probability measure. Introduce for a subset $C \subseteq X \times Y$ the sections

$$\begin{aligned} C_x &:= \{y \in Y \mid \langle x, y\rangle \in C\}, \\ C^y &:= \{x \in X \mid \langle x, y\rangle \in C\}, \end{aligned}$$

then define

$$(\mu \otimes K)(C) := \int_X K(x)(C_x)\ \mu(dx)$$

as a measure on $X \times Y$. It should be noted that

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

holds, whenever $f : X \times Y \to \mathbb{R}$ is measurable and bounded.

The following crucial *disintegration property* shows that each measure on a product space can indeed be written as the product of a measure and a kernel. To be more specific: suppose $\mu \in \mathsf{Prob}(X \times Y)$ is a probability on the Cartesian product $X \times Y$ of X and Y. Then there exists a probabilistic relation $K : X \rightsquigarrow Y$ such that

$$\mu = \pi_X(\mu) \otimes K$$

holds. Here $\pi_X : X \times Y \to X$ is the projection.

Finally, $K : X \rightsquigarrow Y$ induces a map from $\mathsf{Prob}(X)$ to $\mathsf{Prob}(Y)$ upon setting:

$$K^\bullet(\mu)(B) := \int_X K(x)(B)\ \mu(dx).$$

Hence $\int_Y g\ dK^\bullet(\mu) = \int_X \int_Y g(y)\ K(x)(dy)\ \mu(dx)$ holds for every measurable and bounded function $g : X \to \mathbb{R}$.

The relationship between the probabilistic relation K and its set-theoretic sidekick R_K may be interesting to observe. Composition carries over to the sidekick as follows:

Observation 4 *Let $K : X \rightsquigarrow Y$ and $L : Y \rightsquigarrow Z$ be probabilistic relations, and assume that L is continuous. Then*

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

Proof: Since L is continuous, the set $U_L := \{y \in Y \mid L(y)(U) > 0\}$ is open in Y, whenever $U \subseteq Z$ is open.

Now let $\langle x, z\rangle \in R_K \circ R_L$ such that for some y both $\langle x, y\rangle \in R_K$ and $\langle y, z\rangle \in R_L$ hold. If U is an open neighborhood of z, U_L is an open neighborhood of y, thus

$$(K;L)(x)(U) \geq \int_{U_L} L(y)(U)\ K(x)(dy) > 0.$$

This proves (1). If, on the other hand, $(K;L)(x)(U) > 0$ for some open set U containing z, and if $K(x)(U_L) > 0$, then $\langle x, z\rangle \in R_{K;L}$ implies $\langle x, z\rangle \in R_K \circ R_L$. This establishes (2). $\Box$

The condition in the second part of Obs. 4 is e.g. satisfied for $Y = \mathbb{R}$ and the case that

$$K(x)([a,b]) = \int_a^b f(x,y)\ dy,$$

with a strictly increasing and differentiable density $f(x,\cdot)$ for each x.

Recall that for a relation $R \subseteq X \times Y$ and a set $P \subseteq X$ the (left) *Peirce product* $P \odot R$ [2, Ch. 1] is defined as

$$P \odot R := \{\langle x,y\rangle \in R|\ x \in P\}.$$

Observation 5 *If $K : X \rightsquigarrow Y$ is a probabilistic relation, and μ is a probability on X, then*

1. $\mathsf{supp}(\mu \otimes K) \subseteq (\mathsf{supp}(\mu) \odot R_K)^{\mathsf{cl}}$,
2. *if K is continuous,* $\mathsf{supp}(\mu) \odot R_K \subseteq \mathsf{supp}(\mu \otimes K)$ *holds.*

Proof: Because

$$\mathsf{supp}(\mu) \odot R_K = (\mathsf{supp}(\mu) \times Y) \cap R_K,$$

and because R_K is measurable, $\mathsf{supp}(\mu) \odot R_K$ is. From

$$\begin{aligned}(\mu \otimes K)(\mathsf{supp}(\mu) \odot R_K) &= \int_{\mathsf{supp}(\mu)} K(x)(\mathsf{supp}(K(x))\ \mu(dx) \\ &= 1,\end{aligned}$$

we may infer

$$\mathsf{supp}(\mu \otimes K) \subseteq (\mathsf{supp}(\mu) \odot R_K)^{\mathsf{cl}}.$$

On the other hand, continuity of K implies that $(\mu \otimes K)(U \times V) > 0$ for each open neighborhood $U \times V$ of $\langle x,y\rangle \in \mathsf{supp}(\mu) \odot R_K$. $\Box$

5 Converse Relations

Fix a probabilistic relation $K : X \rightsquigarrow Y$, and a probability measure μ on X. Then $\tau := \mu \otimes K$ is a probability on $X \times Y$ which has a kind of natural converse: put for $D \in \mathcal{B}_{X \times Y}$

$$\tau^{\smile}(D) := \tau(D^{\smile})$$

with

$$D^{\smile} := \{\langle x,y\rangle|\ \langle y,x\rangle \in D\}$$

as the set theoretic converse.

Hence $\tau^{\smile}$ is a probability measure on the Polish space $X \times Y$ and is representable by a probability measure γ on Y and a probabilistic relation $K_\mu^{\smile} : Y \rightsquigarrow X$:

$$\tau^{\smile} = \gamma \otimes K_\mu^{\smile}.$$

Note that reverting a stochastic matrix in the finite or denumerable case does not yield necessarily a stochastic matrix again: if $p(i,j)$ is the probability that on input i system p reacts with output j, then a converse interpretation that starts from j is not easy to devise. Thus it is more difficult to transpose a stochastic matrix so that another matrix of this type emerges. The proposal made here is to use a kind of helper probability.

We are ready for the definition of the converse of a probabilistic relation.

Definition 2 *The* μ-converse $K_\mu^\smile$ *of the probabilistic relation* K *with respect to the input probability* μ *is defined by the equation*

$$(\mu \otimes K)^\smile = K^\bullet(\mu) \otimes K_\mu^\smile.$$

It is remarked that the converse $K_\mu^\smile$ always exists, and that it is unique μ-almost everywhere.

Since

$$\begin{aligned}\mu(A) &= (\mu \otimes K)(A \times Y) \\ &= (K^\bullet(\mu) \otimes K_\mu^\smile)((Y \times A)^\smile)\end{aligned}$$

is true for $A \in \mathcal{B}_X$,

$$\begin{aligned}\mu(A) &= \int_X \int_Y K_\mu^\smile(A)\ K(x)(dy)\ \mu(dx) \\ &= \int_Y K_\mu^\smile(A)\ K^\bullet(\mu)(dy),\end{aligned}$$

we infer that

$$\begin{aligned}\mu &= \left(K_\mu^\smile\right)^\bullet(K^\bullet(\mu)) \\ &= (K;K_\mu^\smile)^\bullet(\mu)\end{aligned}$$

holds. Hence the converse $K_\mu^\smile$ 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" $\{A \times Y |\ A \in \mathcal{B}_X\}$ on the product $X \times Y$. This is not enough to determine a measure on the entire product.

The construction implies that for the Borel set $D \subseteq Y \times X$ the equation

$$\int_X K(x)(D^x)\ \mu(dx) = \int_Y K_\mu^\smile(y)(D_y)\ \gamma(dy)(*)$$

holds, which implies

$$\begin{aligned}\gamma(B) &= \int_X K(x)(B)\ \mu(dx) \\ &= K^\bullet(\mu)(B).\end{aligned}$$

Now equation (*) reads a little more symmetric:

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

This will be generalized and made use of later:

Observation 6 *Let* $f : X \times Y \to \mathbb{R}$ *be measurable and bounded. Then this identity holds:*

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

□

Thus we may interchange the order of integration of f 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. Let us have a look at an analogue of the property $R^{\smile\smile} = R$ which holds for the set theoretic converse. We need again a helper probability μ and will see that we have a similar property.

Proposition 7 *If $K : X \rightsquigarrow Y$, and if $\mu \in \mathsf{Prob}(X)$, then*

$$(K_\mu^\smile)_{K^\bullet(\mu)}^\smile =_\mu K.$$

holds.

Proof: The probabilistic relation $(K_\mu^\smile)_{K^\bullet(\mu)}^\smile$ is determined by the equation

$$(K^\bullet(\mu) \otimes K_\mu^\smile)^\smile = \eta \otimes (K_\mu^\smile)_{K^\bullet(\mu)}^\smile$$

with $\eta := K_\mu^\smile(K^\bullet(\mu))$. Equation $(*)$ implies now that $\eta = \mu$, consequently,

$$\mu \otimes K = \mu \otimes (K_\mu^\smile)_{K^\bullet(\mu)}^\smile,$$

as expected. □

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 probabilistic relations reads

Proposition 8 *Let $K : X \rightsquigarrow Y, L : Y \rightsquigarrow T$ be probabilistic relations, and let $\mu \in Prob(X)$ be a probability. Then*

$$(K;L)_\mu^\smile = L_{K^\bullet(\mu)}^\smile;K_\mu^\smile$$

holds.

Proof: We will make use of observation 6 by showing that both relations have the same properties on measurable and bounded functions.

Allora: Let $f : X \times Z \to \mathbb{R}$ be bounded, then

$$
\begin{aligned}
\int_{X\times Z} f\ d(\mu \otimes (K;L)) &= \int_Z \int_X f(x,z)\ (K;L)_\mu^\smile(z)(dx)\ L^\bullet\left(K^\bullet(\mu)\right)(dz) && (1)\\
&= \int_X \int_Z f(x,z)\ (K;L)\,(x)(dz)\mu(dx) && (2)\\
&= \int_X \int_Y \int_Z f(x,z)\ L(y)(dz)\ K(x)(dy)\ \mu(dx) && (3)\\
&= \int_Y \int_X \int_Z f(x,z)\ L(y)(dz)\ K_\mu^\smile(dx)\ K^\bullet(\mu)(dy) && (4)\\
&= \int_Y \int_Z \int_X f(x,z)\ K_\mu^\smile(dx)\ L(y)(dz)\ K^\bullet(\mu)(dy) && (5)\\
&= \int_Z \int_Y \int_X f(x,z)\ K_\mu^\smile(dx)\ L_{K^\bullet\mu}^\smile(z)(dy)\ L^\bullet\left(K^\bullet(\mu)\right)(dz) && (6)\\
&= \int_Z \int_X \int_Y f(x,z)\ K_\mu^\smile(dx)\ L_{K^\bullet\mu}^\smile(z)(dy)\ L^\bullet\left(K^\bullet(\mu)\right)(dz) && (7)\\
&= \int_Z \int_X f(x,z)\ \left(L_{K^\bullet(\mu)}^\smile;K_\mu^\smile\right)(z)(dx)\ L^\bullet\left(K^\bullet(\mu)\right)(dz). && (8)
\end{aligned}
$$

Eq. (1) is a first application of Observation 6, equation (2) applies the definition of $\mu \otimes (K;L)$, to the first integral. In eq. (3) the definition of $K;L$ is expanded, and in eq. (4) Observation 6 is applied to the two outermost integrals, similarly for eq. (6). Fubini's Theorem is used for the interchanges of integrals in eqs. (5) and (7). Finally, equation (8) applies the definition of the composition of kernels to $L^{\smile}_{K^{\bullet}(\mu)}$ and $K^{\smile}_{\mu}$. Comparing the first and the last equalities established the claim.

6 Related Work

The generalization of set-based relations to probabilistic ones appears straightforward: replace the nondeterminism inherent in these relations by randomness. Panangaden [11] carries out a very elegant construction, arguing as follows: the powerset functor is a monad which has relations as morphisms in its Kleisly category [9], 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 [7]. This parallel justifies their introduction as probabilistic relations. The category **SRel** of measurable spaces with transition sub-probabilities is scrutinized closer in [11], and an application to the semantics Kozen's probabilistic programs [8] is given. Abramsky, Blute, and Panangaden [1] investigate that category (now called **Stoch**) in the context of Hilbert spaces and their adjoints, 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 what is called here the image measure. 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 technique are directly used in the present paper (while [1] prefers arguing with absolute continuity, and consequently, with the Radon-Nikodym Theorem).

The observation that each transition probability on a Polish space spawns through the support function a measurable set-valued function, hence a relation, was used in [6] 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

This paper proposes the notion of the converse of a probabilistic relation through disintegrating a probability that measures the converse of a set-theoretic relation. The resulting converse depends on an initial probability. The techniques applied are based on the theory of measures on Polish spaces.

The main contribution of this paper are:

- Some fundamental laws for dealing with converse relations could be established,
- The support function of a probability measure was investigated in terms of monads and shown to be a natural transformation;

- The support function was used to clarify the interplay between set-theoretic and probabilistic relations.

Some questions remain open. Desharnais, Edalat and Panangaden [5] formulate some of their work on stochastic bisimulation using analytic spaces. This gives rise to the question which topological requirements are basic for work on probabilistic relations. What can be carried over to these spaces, that appear as quite natural candidates? The question arises, too, whether some of the rather stringent assumptions can be bypassed (e.g., compactness is needed to ensure that the functor **F** forms a monad in sect. 3 because the Vietoris topology forces innocent assumptions like separability of a hyperspace into compactness). The T-algebras arising from the functor **G** in sect. 3 require identification and interpretation in terms of probabilistic relations.

All this may help in establish a theory of probabilistic relations, comparable in breadth and scope to their set-valued cousins.

References

[1] S. Abramsky and R. Blute amd P. Panangaden. Nuclear and trace ideal in tensored *-categories. Technical report, School of Computer Science, McGill University, Montreal, June 1998.

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

[3] D. Cantone, E. G. Omodeo, and A. Policriti. *Set Theory for Computing.* Springer-Verlag, 2001. In print.

[4] C. Castaing and M. Valadier. *Convex Analysis and Measurable Multifunctions*, volume 580 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1977.

[5] J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation of labelled markov-processes. Technical report, School of Computer Science, McGill University, Montreal, 1998.

[6] E.-E. Doberkat. *Stochastic Automata — Nondeterminism, Stability, and Prediction*, volume 113 of *Lecture Notes in Computer Science.* Springer-Verlag, Berlin, 1981.

[7] 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.

[8] D. E. Kozen. Semantics of probabilistic programs. *Journal of Computer and Systems Sciences*, 30(2):328 – 350, 1985.

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

[10] E. Michael. Topologies on spaces of subsets. *Trans. Am. Math. Soc.*, 71(2):152 – 182, 1951.

[11] 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.

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

[13] E. Schröder. *Vorlesungen über die Algebra der Logik (exakte Logik)*, volume 3, Algebra und Logic der Relative, part 1. B. Teubner, Leipzig, 1895. Reprinted by Chelsea Publishing Co., New York, 1966.