# The Hierarchical Refinement of Probabilistic Relations

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

November 12, 2001

**Abstract**

Refining a specification means replacing a part of it through a more elaborated one. It is evident how to do this when dealing with ordinary relations: here simply the composition operator is used. We show that the composition operator can also be put to good use when refining a probabilistic relation. In this paper we propose a definition for the refinement of these relations, and we give two representation theorems: the first relates refinements to memoryless probabilistic relations on an infinite product through projection and disintegration. The second one investigates the relation between infinite trees and the refinement operation by showing that each tree can be generated through probabilistic refinement, provided a compactness assumption holds.


**Key words:** Probabilistic relations, relational specifications, refinement, compositionality, concurrency, trees.

# 1 Introduction

Probabilistic relations are the stochastic counterparts to set-based relations. A probabilistic relation is a transition kernel between two measurable spaces, combining measurable maps with subprobability measures (a formal definition is given in Sect. 2).
These kernels are constructed as the probabilistic analogues to set-theoretic relations (see [Pan98, AaPP98] or [Dob01a]). This analogy flows from two sources. First and informally, set-theoretic relations may be used for non-deterministic processes, so that $\{y|\langle x,y\rangle \in R\}$ yields the set of all possible results of computation $R$ after input $x$. Many applications, however, assign probabilities to possible outcomes (because some events are more likely than others), and this leads to the notion of a transition probability which gives us the probability $K(x)(B)$ that upon input $x$ the output will be an element of set $B$. The output of single elements is usually not considered when taking larger spaces like the real numbers into account, where single elements do not carry an individual weight. This intuitive reasoning is supported formally, as e.g. [Pan98] points out: the power set functor forms a monad in the category of sets and has relations as the Kleisli construction; the functor assigning each measurable space its probability measures forms also a monad and has Markov kernels as its Kleisli construction, see [Gir81]. In this way a categorical underpinning of the intuitive reasoning is provided.
The concepts of interactions may be represented through a general form of game sematics by constructing a compact closed category $\mathcal{G}(\mathcal{C})$ over a traced mondial categroy $\mathcal{C}$; Abramsky illustrates this construction through several examples, among others using the category of measurable spaces with probabilistic relations as morphisms [Abr96, § 4.3]. The objects of $\mathcal{C}$ are embedded into $\mathcal{G}(\mathcal{C})$ as boxes which suggest themselves for refinement, replacing boxes through a system of smaller one.
Refinement is an effective technique for the development of programs [GJM91, 4.2.6]. It works iteratively by decomposing a problem into subproblems, and by linking their solution to the solution of the entire problem. In this way the solution space evolves when partial solutions are expanded, so that in each step a closer approximation to the problem's solution is obtained. This paper discusses refinements from the point of view of probabilistic relations: if a process is specified through such a relation, refining an approximate solution means to replace a state that has been reached by some other states that better describe the intended solution. In this way the probabilistic relation describing the entire specification evolves, and the whole process may be pursued by looking at the corresponding sequence of probabilistic relations.
To be specific, suppose that the probabilistic relation $K$ is refined through another relation $L$, then the resulting relation would be $\int L(y)(B)\ K(x)(dy)$, indicating that upon input $x$ the resulting state will be in set $B$ when the initial probability is furnished by $K$, and $L$ serves to describe the probabilities to end up in $B$ via an intermediate state the weight of which depends on $K(x)$.
It is this refinement operation that we study, and it is well to be distinguished from the refinement discussed e.g. in [MMS96, Def. 2.3] in which is a refinement order on the space of all subprobability measures on the discrete space of states is considered.
This paper is organized as follows: we collect some preliminaries in Sect. 2 for the reader's convenience, Sect. 3 defines the refinement operation, illustrates the concept by showing how to do partial refinements, and by relating bisimilarity to refinements; this section gives also a representation through Markov relations on infinite products. Sect. 4 investigates the relationship of refinements and infinite trees and shows under what conditions a tree may be

represented through a sequence of refinements. Conclusions and indications for further work are given in 5.

## 2 Preliminaries

Denote for a set $S$ by $S^*$ as usual the free semigroup on $S$ with $\epsilon$ as the empty word; $S^\infty$ is the set of all infinite sequences based on $S$. If $v \in S^*$, $w \in S^* \cup S^\infty$, then $v \preceq w$ iff $v$ is an initial piece of $w$, in particular $\sigma|k \preceq \sigma$ for all $\sigma \in S^\infty, k \in \mathbb{N}$, where $(s_n)_{n\in\mathbb{N}}|k := s_0 \dots s_k$ is the prefix of $(s_n)_{n\in\mathbb{N}}$ of length $k+1$.
A *tree* $T$ *on* $S$ is a subset of $S^*$ which is closed under the prefix operation, thus $w \in T$ and $v \preceq w$ together imply $v \in T$. The *body* $[T]$ *of* $T$ [Sri98] is the set of all sequences on $S$ each finite prefix is in $T$, thus

$$[T] := \{\sigma \in S^\infty | \forall k \in \mathbb{N} : \sigma|k \in T\}.$$

Apparently a finite tree like a binary search tree or a heap has an empty body.
Denote for a measurable space $X$ by $\mathbb{S}(X)$ the set of a subprobability measures on (the $\sigma$-algebra of) $X$ with the set $\mathbb{P}(X)$ of all probability measures as a subset. We usually omit mentioning the $\sigma$-algebra underlying a measurable space and talk about its members as *measurable subsets*, or as *Borel subsets*, if $X$ is a metric space, see below. $\mathbb{S}(X)$ is endowed with the $* - \sigma$-algebra, i.e. the smallest $\sigma$-algebra that makes for each measurable subset $A$ of $X$ the evaluation map $\mu \mapsto \mu(A)$ measurable.
Define for the measurable map $f : X \to Y$ and for $\mu \in \mathbb{S}(X)$ the *image* of $\mu$ under $f$ by $\mathbb{S}(f)(\mu)(B) := \mu(f^{-1}[B])$, then $\mathbb{S}(f)(\mu) \in \mathbb{S}(Y)$. $\mathbb{P}(f)$ is defined through exactly the same expression, having $\mathbb{P}(X)$ and $\mathbb{P}(Y)$ as domain, and as range, resp. It is easily established that $\mathbb{S}(f) : \mathbb{S}(X) \to \mathbb{S}(Y)$ is measurable. A *probabilistic relation* $K : X \rightsquigarrow Y$ *between* $X$ *and* $Y$ [AaPP98, Pan98, Dob01a, Dob01b] is a measurable map $K : X \to \mathbb{S}(Y)$, consequently it has these properties:

1. for all $x \in X$, $K(x)$ is a subprobability measure on $Y$,
2. for all $B \subseteq Y$ which are measurable, $x \mapsto K(x)(B)$ is a measurable map on $X$, where measurability of real functions always refers to the Borel sets in $\mathbb{R}$.

If each $x \in X$ sends $K(x)$ to $\mathbb{P}(Y)$, then $K$ is called a *Markov relation* or a *transition probability*. When probabilistic relations are used to model computations, Markov relations model terminating computations (so that for a non-Markovian relation $K$ the difference $1 - K(x)(Y)$ may be interpreted as the amount of nontermination on input $x$, since this is the probability for "no state at all" [MMS96]).
Probabilistic relations may be composed similar to set theoretic ones: let $K : X \rightsquigarrow Y$ and $L : Y \rightsquigarrow Z$ be probabilistic relations, then define for $x \in X$ and the measurable subset $C \subseteq Z$ the *(ordinary) product of* $K$ *and* $L$ by

$$(K \odot L)(x)(C) := \int_Y L(y)(C)\ K(x)(dy),$$

thus $K \odot L : X \rightsquigarrow Z$. If $\mu \in \mathbb{S}(X), K : X \rightsquigarrow Y$, define for the measurable subset $A \subseteq X \times Y$

$$(\mu \otimes K)(A) := \int_X K(x)(A_x)\ \mu(dx)$$

with $A_x := \{y \in Y | \langle x, y\rangle \in A\}$. Consequently, $\mu \otimes K \in \mathbb{S}(X \times Y)$.
A Polish space $X$ is a completely metrizable separable topological space; as usual, we take the Borel sets as the $\sigma$-algebra on a Polish space. If $X$ is Polish, so are [Sri98, Par67]

- $\mathbb{S}(X)$ under the topology of weak convergence; the $* - \sigma$-algebra coincides with the Borel sets,
- $X^*$ under the topological sum of $(X_n)_{n\in\mathbb{N}}$,
- $X^\infty$ under the topological product; the Borel sets are the $\sigma$-algebra generated by sets of the form $\prod_{n\in\mathbb{N}} A_n$, where all $A_n \subseteq X$ are Borel sets, and all but a finite number equal $X$.

Assume that $X$ carries a measurable structure, and that $Y$ is a Polish space. A relation $R \subseteq X \times Y$ induces a set-valued map (and again denoted by $R$) upon setting

$$R(x) := \{y \in Y | \langle x, y\rangle \in R\}.$$

If $R(x)$ always takes closed and non-empty values, and if the weak inverse

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

is a measurable set, whenever $G \subseteq Y$ is open, then $R$ is called a *measurable relation on* $X$. Since $Y$ is Polish, $R$ is a measurable relation iff the strong inverse

$$(\forall R)(F) := \{x \in X | F(x) \subseteq F\}$$

is measurable, whenever $F \subseteq Y$ is closed [Him75, Theorem 3.5]. It is well known that a measurable relation $R$ constitutes a measurable subset of $X \times Y$.
Fix for the rest of the paper $X$ as a Polish space.

# 3 The Refinement

**Definition 1** *Let $K, L : X \rightsquigarrow X$ be probabilistic relations on $X$, then $K \odot L$ is the* refinement of $K$ by $L$.

How this construction works is demonstrated best in the discrete case. Suppose $X$ is countable, and $K$ and $L$ are given through stochastic matrices $(p_{x,y})_{x,y\in X}$, and $(q_{x,y})_{x,y\in X}$, hence $p_{x,y}$ is the probability that $K$ produces $y$ on input $x$, and $q_{y,z}$ is similarly the probability that $L$ produces $z$ upon input $y$. Suppose that at a certain time of system specification the specifier arrived at $K$ as the system specification. If state $y$ is to be refined in a subsequent step, we need to replace $y$ by a member from the set $\{z | z \in I_y\}$ of refining states, and we need to replace the transition $x \xrightarrow{p_{x,y}} y$ from $x$ to $y$ by the combined transition $x \xrightarrow{p_{x,y}} y \xrightarrow{q_{y,z}} z$ for a transition from $x$ to $z$ via $y$. Combining $K$ and $L$ yields evidently $\sum_{y\in Y} p_{x,y} \cdot q_{y,z}$ as the probability that the combined system produces $z$ upon input $x$.
Thus if $K(x)(A)$ is the probability that a transition from $x$ will lead into the measurable set $A$, and refining each $y \in A$ by $L(y)$ (so that $L(y)(B)$ is the probability for a transition from $y$ to end up in $B$), then the transition from $x$ to a set $A$ for the refined relation will be $\int L(y)(A)\ K(x)(dy)$, which equals $(K \odot L)(x)(Y)$.

Just to illustrate the concept, suppose that we want to refine only part of the system: we are given a probabilistic relation $K$ on $X$, and a measurable subset $W$, outside which the refined system is not supposed to change. If $x \in W$, however, we are to use the probabilistic relation $L$. Put

$$L_W(y) := \begin{cases} L(y), & y \in W \\ \delta_y, & y \notin W \end{cases}$$

(with $\delta$ as the Dirac kernel), then apparently

$$(K \odot L_W)(x)(A) = \int_W L(y)(A)\ K(x)(dy) + K(x)(X \setminus W).$$

Thus changes are localized to $W$. In particular it is possible building up specification hierarchies through a decreasing sequence

$$X = W_1 \supseteq W_2 \supseteq \dots W_{n-1} \supseteq W_n = \emptyset.$$

To illustrate further, consider bisimilarity: the probabilistic relations $K_1$ and $K_2$ on $X$ are called *bisimilar* iff there exists a measurable subset $A \subseteq X^2$ and a probabilistic relation $T : A \rightsquigarrow X^2$ such that we have for every $x \in A$

$$K_i(\pi_i(x)) = \mathbb{S}\,(\pi_i)\,(T(x))$$

($\pi_i$ is the $i^{th}$ projection, $i = 1, 2$). This is abbreviated by $K_1 \sim_{\langle A,T\rangle} K_2$, and $T : A \rightsquigarrow X_1 \times X_2$ is called the *mediating relation*. From [Dob01b, Prop. 12] it is seen that the refinement operation maintains bisimilarity, provided termination is not affected. This yields:

**Proposition 1** *Let $K_1, K_2, L_1, L_2$ be probabilistic relations on $X$, and assume that both $K_1 \sim_{\langle A,T\rangle} K_2$ and $L_1 \sim_{\langle B,S\rangle} L_2$ holds. Then*

$$K_1 \odot L_1 \sim_{\langle A,(T \odot S)\rangle} K_2 \odot L_2$$

*holds, provided $T : A \rightsquigarrow B$ is a Markov relation.* ■

Thus the computation associated with the mediating relation for $K_1, K_2$ must terminate and produce a transition to an element of the domain for the mediating relation for $L_1, L_2$ with probability 1; under this condition bisimilarity is maintained.
We consider now iterated refinements, and want to find a representation for those relations that are iterated ones. For this purpose we fix a sequence $(K_n)_{n\in\mathbb{N}}$ of probabilistic relations $K_n$ on $X$. Define a sequence $(V_n)_{n\in\mathbb{N}}$ of probabilistic relations $V_n$ on $X$ inductively upon setting

$$V_n := \begin{cases} K_0, & n = 0 \\ V_{n-1} \odot K_n, & n > 0 \end{cases}$$

Then $V_n$ reflects the refinement's state after the $n^{\text{th}}$ step.
The goal is then the characterization of those sequences of probabilistic relations that are the result of a refinement process in the following sense: given $(V_n)_{n\geq 0}$ with $V_n : X \rightsquigarrow X$, under which conditions can we find a sequence $(K_n)_{n\geq 0}$ of probabilistic relations such that

$$V_n = V_{n-1} \odot K_n$$

holds for all $n \in \mathbb{N}$? Because we will deal with infinite products, we will deal with terminating relations (otherwise the infinite products to be constructed will converge to zero).
Denote the resp. projections $(x_n)_{n\geq 0} \mapsto \langle x_0, \dots x_n\rangle$ by $proj_{n+1}$, and $(x_n)_{n\geq 0} \mapsto x_n$ by $\pi_n^\infty$, resp.:

**Definition 2** *A transition probability $L : X \rightsquigarrow X^\infty$ is called* memoryless *iff for each $n \in \mathbb{N}, x \in X$ the projection*

$$proj_{n+1}(L(x))$$

*can be decomposed through disintegration as*

$$proj_n(L(x)) \otimes J_n$$

*with $J_n : X \rightsquigarrow X$, where $J_n$ is independent of $x$.*

Fix $x$, and interpret $\mu := L(x)$ in Definition 2 as the joint distribution of a stochastic process $(\zeta_i)_{i\geq 0}$ with $\zeta_i : \Omega \to X$ over the probability space $(\Omega, \mathcal{A}, \mathsf{P})$. Since $X$ is Polish, there exists for each $n \in \mathbb{N}$ a regular conditional distribution of $\zeta_n$ conditional to $\langle \zeta_0, \dots, \zeta_{n-1}\rangle$ (cp. [Par67, Theorem V.8.1]), hence a Markov kernel $J_n : X^n \rightsquigarrow X$ such that for the Borel sets $B_1 \subseteq X^n, B_2 \subseteq X$

$$\begin{aligned}\mu_n(B_1 \times B_2) &= \mathsf{P}(\langle \zeta_0, \dots, \zeta_{n-1}\rangle \in B_1, \zeta_n \in B_2) \\ &= \int_{B_1} J_n(x_0, \dots x_{n-1})(B_2)\ \mu_{n-1}(d\langle x_0, \dots, x_{n-1}\rangle).\end{aligned}$$

Consequently, a memoryless distribution corresponds to a Markov process, since in this case $J_n(x_0, \dots, x_{n-1})$ only depends on the last state $x_{n-1}$ and not on the whole history $x_0, \dots, x_{n-1}$. Taking things a little further, consider a sequence $L_n : X \rightsquigarrow X$ of Markov relations. Define for $x \in X$ the probabilistic relation $L_\infty : X \rightsquigarrow X^\infty$ upon setting

$$L_\infty(x) := \bigotimes_{n\in\mathbb{N}} L_n(x),$$

then $L_\infty$ is evidently memoryless. This corresponds to the case of an independent stochastic process.
Memoryless transition probabilities will play a leading role in characterizing those relations that arise through refinements.

**Proposition 2** *The following conditions are equivalent for a sequence $(V_n)_{n\geq 0}$ of Markov relations $V_n : X \rightsquigarrow X$:*

1. *There exists a sequence of probabilistic relations $(K_n)_{n\geq 0}$ such that $(V_n)_{n\geq 0}$ is the sequence of refinements for $(K_n)_{n\geq 0}$,*
2. *There exists a memoryless transition probability $L_\infty : X \rightsquigarrow X^\infty$ such that $V_n = \pi_n^\infty(L_\infty)$ holds for each $n \in \mathbb{N}$.*

**Proof:** Denote by $\pi_k^{n+1}$ the $k^{\text{th}}$ projection $X^{n+1} \ni \langle x_0, \dots x_n\rangle \mapsto x_k \in X$, and let for the Borel set $B \subseteq X^{n+1}$, $\langle x_0, \dots, x_n\rangle \in X^n$ the *cut of $B$ at $\langle x_0, \dots, x_n\rangle$* be the set

$$B_{\langle x_0, \dots, x_n\rangle} := \{x \in X | \langle x_0, \dots, x_n, x\rangle \in B\}.$$

*"1. ⇒ 2."*: Assume that $(V_n)_{n\geq 0}$ comes from refining $(K_n)_{n\geq 0}$. Define inductively for $n \in \mathbb{N}$ and the Borel set $B \subseteq X^{n+1}$ with $x \in X$ fixed

$$\begin{aligned} L_0(x) &:= K_0(x), \\ L_{n+1}(x)(B) &:= \int_{X^{n+1}} K_{n+1}(y_n)(B_{\langle y_0,\ldots,y_n\rangle})\ L_n(x)(d\langle y_0,\ldots,y_n\rangle), \end{aligned}$$

then $L_n : X \rightsquigarrow X^{n+1}$ is readily established, and

$$\pi^{n+1}_{n+1}(L_n(x)) = V_n(x)$$

is proved by induction now. The case $n = 0$ is trivial, and the inductive step proceeds as follows:

$$\begin{aligned} V_{n+1}(x)(A) &= \int_X K_{n+1}(y)(A)\ \pi^{n+1}_{n+1}(L_n(x))(dy) \\ &= \int_{X^{n+1}} K_{n+1}(\pi^{n+1}_{n+1}(z))(X^{n+1} \times A)_z)\ L_n(x)(dz) \\ &= L_{n+1}(x)(X^{n+1} \times A) \\ &= \pi^{n+1}_{n+1}(L_{n+1}(x))(A). \end{aligned}$$

For the Borel set $A \subseteq X^{n+1}$ it is immediate that $L_n(x)(A)$ equals $L_{n+1}(x)(A \times X)$. This property enables the unique extension of all these probabilities to $X^\infty$. Let $\mathcal{A}_0$ be the collection of all sets of the form $\prod_{i\geq 0} B_i$ such that $B_i \subseteq X$ is a Borel set, and $B_i = X$ for almost all indices $i$. Then $\mathcal{A}_0$ generates the Borel sets of $X^\infty$, and it is closed under finite intersections. Defining for $x \in X$

$$\hat{\mu}_x(\prod_{i\geq 0} B_i) := L_n(x)(B_1 \times \cdots \times B_n),$$

where $n$ is a arbitrary integer with $B_{n+1} = B_{n+2} = \cdots = X$, then $\hat{\mu}_x$ is well-defined on $\mathcal{A}_0$, and extends for each $x$ uniquely to a probability measure $L_\infty(x) \in \mathbb{P}(X^\infty)$. The function $x \mapsto L_\infty(x)(B)$ is measurable for each Borel set $B \subseteq X^\infty$. This is so since

$$\{B \subseteq X^\infty | x \mapsto L_\infty(x)(B) \text{ is measurable}\}$$

is a $\sigma$-algebra which contains $\mathcal{A}_0$. Consequently,

$$\begin{aligned} V_n(x) &= \pi^{n+1}_{n+1}(L_n(x)) \\ &= \pi^{n+1}_{n+1}(proj_{n+1}(L_\infty(x))) \\ &= \pi^\infty_n(L_\infty). \end{aligned}$$

The construction of $(L_n)_{n\geq 0}$ implies that $L_\infty$ is memoryless.

*"2. ⇒ 1."*: Let $L_\infty$ be memoryless with $V_n = \pi^\infty_n(L_\infty)$. We have to find $K_n : X \rightsquigarrow X$ such that $(V_n)_{n\geq 0}$ is the refinement associated with $(K_n)_{n\geq 0}$. Put

$$L_n(x) := proj_{n+1}(L_\infty(x)),$$

and represent $L_{n+1}$ through disintegration as

$$L_{n+1}(x) = L_n(x) \otimes K_{n+1},$$

such that $K_n : X \rightsquigarrow X$, since $L_\infty$ is memoryless. Hence if $B \subseteq X^{n+2}$ is a Borel set,

$$L_{n+1}(x)(B) = \int_{X^{n+1}} K_{n+1}(\pi_{n+1}^{n+1}(z))(B_z)\ L_n(x)(dz).$$

Thus we get for the Borel set $A \subseteq X$

$$\begin{aligned} V_{n+1}(x)(A) &= L_{n+1}(x)(X^{n+1} \times A) \\ &= \int_X K_{n+1}(y)(A)\ \pi_{n+1}^{n+1}(L_n(x))(dy) \\ &= (V_n \odot K_{n+1})(x)(A). \end{aligned}$$

This was to be shown. ■

This proposition has a somewhat unexpected consequence:

**Corollary 3** *Each sequence $(V_n)_{n\geq 0}$ of probabilistic relations $V_n : X \rightsquigarrow X$ is the sequence of refinements for some $K_n : X \rightsquigarrow X$.* ■

This result indicates that refining programs does not in itself provide sensible solutions: an arbitrary deterministic program can be derived through a sequence of whimsical, albeit goal directed refinement steps; Cor. 3 is the probabilistic counterpart for this observation.

## 4 Trees

A refinement process spawns a tree by opening the possibility of different avenues to exploit. Suppose that in a non-deterministic refinement step $n$ the refining replacement is governed by a relation $R_n \subseteq X \times X$, so that the refinement for $x$ consists in replacing it by all the elements in $R_n(x)$. Put $\mathcal{R} := \{R_n | n \in \mathbb{N}\}$, and

$$|\mathcal{R}| := \{v \in X^* | v_0 \in \mathsf{dom}(R_0), v_j \in R_{j-1}(v_{j-1}) \text{ for } 1 \leq j \leq |v|\} \cup \{\epsilon\},$$

then $|\mathcal{R}|$ is a tree with body

$$[|\mathcal{R}|] = \{\alpha \in X^\infty | \alpha_0 \in X \ \&\ \forall j \geq 1 : \alpha_j \in R_{j-1}(\alpha_{j-1})\}.$$

In fact, each tree $T$ can be represented in this way: Define $S_0 := X \cap T$, and let for $k \geq 0$

$$\langle x_k, x_{k+1}\rangle \in S_{k+1} \Leftrightarrow \exists\ w \in T, |w| = k : w x_k x_{k+1} \in T.$$

Then an inductive argument shows that

$$\begin{aligned} \langle x_k, x_{k+1}\rangle \in S_{k+1} \quad \Leftrightarrow \quad & \exists\ x_0 \in S_0 \exists x_1 \in S_1(x_0) \ldots \exists x_{k-1} \in S_{k-1}(x_{k-2}) : \\ & x_k \in S_k(x_{k-1}) \wedge x_0 x_1 \ldots x_k x_{k+1} \in T, \end{aligned}$$

and from this

$$T = |\{S_n | n \geq 0\}|$$

is deduced easily.

The notion of a measurable tree helps characterizing the situation:

**Definition 3** *The tree $T \subseteq X^*$ is called a* measurable tree *iff the following conditions are satisfied:*

1. *$[T] \neq \emptyset$,*
2. *$T$ is a Borel set in $X^*$,*
3. *$v \mapsto T^\bullet(v) := \{x \in X | vx \in T\}$ constitutes a measurable closed valued map on $T$.*

The last condition implies that $T^\bullet(v)$ is a closed subset of $X^*$ for all $v \in T$. The condition $[T] \neq \emptyset$ makes sure that $\forall v \in T : T^\bullet(v) \neq \emptyset$, so that the tree continues to grow, hence $T$ has the proper range for a measurable relation. This property makes the formulations below a little easier and could be done without, since it is not essential, but since we work in a context in which it will hold anyway, we decided to incorporate it into the definition.
If all relations $R_n$ are measurable relations, then $|\mathcal{R}|$ is a measurable tree:

**Lemma 4** *Construct $|\mathcal{R}|$ from $\{R_n | n \in \mathbb{N}\}$ as above, then this tree has the following properties:*

1. *$|\mathcal{R}| \subseteq X^*$ is a Borel set, provided each $R_n \subseteq X \times X$ is,*
2. *$|\mathcal{R}| = \{\langle v, x\rangle | v \in |\mathcal{R}|, x \in X \text{ so that } vx \in |\mathcal{R}|\}$ is a measurable relation, provided each $R_n$ is.*

**Proof:** 1. Define inductively

$$\begin{aligned} B_0 &:= X, \\ B_1 &:= R_1, \\ B_{k+1} &:= B_k \times X \cap X^k \times R_{k+1}. \end{aligned}$$

Then each $B_k$ is a Borel subset of $X^{k+1}$ under assumption 1. Since $|\mathcal{R}| = \bigcup_{k \geq 0} B_k$, the assertion follows.
2. It is easy to see that

$$\begin{aligned} (\forall|\mathcal{R}|)(F) &= \bigcup_{n \geq 1} \left(B_{n-1} \cap X^{n-1} \times (\forall R_{n-1})(F)\right), \\ (\exists|\mathcal{R}|)(G) &= \bigcup_{n \geq 1} \left(B_{n-1} \cap X^{n-1} \times (\exists R_{n-1})(G)\right), \end{aligned}$$

holds. This implies the second part of the assertion. ■
Define for $\mu \in \mathbb{S}(X)$ the support $\mathsf{supp}(\mu)$ of $\mu$ as the smallest closed subset $F \subseteq X$ such that $\mu(F) = \mu(X) > 0$. Because finite measures on Polish spaces are $\tau$-additive, $\mu(\mathsf{supp}(\mu)) = \mu(X)$, and $x \in \mathsf{supp}(\mu)$ iff $\mu(U) > 0$ for each neighborhood $U$ of $x$. The support for the null measure is defined as the empty set. Probabilistic relations yield measurable relations in a rather natural way [Dob81]: $K : X \rightsquigarrow X$ with $K \neq 0$ induces a relation

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

so that $\mathsf{supp}(K(x))$ takes closed values, and it is measurable, because for an open set $G$ of $X$ the weak inverse $(\exists \mathsf{supp}\ K)(G)$ is measurable. Measurability is also established for the strong inverse $(\forall \mathsf{supp}\ K)(F)$ whenever $F \subseteq X$ is closed, and, somewhat surprisingly, for $(\forall \mathsf{supp}\ K)(A)$ whenever $A \subseteq X$ is an arbitrary Borel set [Dob01b, Cor. 4]. The converse, viz., that a measurable relation may be generated through the support of a probabilistic relation, holds also under compactness conditions, which will be used for establishing the representation below.

**Proposition 5** *Let $(V_n)_{n\geq 0}$ be the sequence of refinements for $(K_n)_{n\geq 0}$, then*

$$|\{\mathsf{supp}\ V_n | n \in \mathbb{N}\}|$$

*constitutes a measurable tree.*

**Proof:** For each $n \in \mathbb{N}$ we have the measurable relation $\mathsf{supp}\ V_n$, thus the assertion follows from Lemma 4. ■

We can show now that under a compactness condition a measurable tree may be generated from some probabilistic refinement. Recall that the set $\mathcal{K}(X)$ of all compact non-void subsets of $X$ is a Polish space when endowed with the Vietoris topology, and that measurability of a compact-valued relation $R \subseteq X \times X$ is equivalent to measurability of the map $R : X \to \mathcal{K}(X)$.

**Proposition 6** *Let $T$ be a tree on $X$, and assume that $T \cap X^k$ is compact for each $k \geq 0$. Then there exists a sequence $(K_n)_{n\geq 0}$ of Markov relations $K_n : X \rightsquigarrow X$ such that*

$$T = |\{\mathsf{supp}\ V_n | n \in \mathbb{N}\}|$$

*for the corresponding sequence $(V_n)_{n\geq 0}$ of refinements.*

**Proof:** 1. Define the sequence $(S_k)_{k\geq 0}$ of relations for $T$ as above, then there exists for each $k \geq 1$ a measurable subset $D_k \subseteq X$ such that

$$S_k : D_k \to \mathcal{K}(X)$$

is a measurable map. This will be shown now. Fix $k \geq 0$, and let $(x_n)_{n\geq 0} \subseteq S_{k+1}(x')$ be a sequence, thus we can find $v_n \in T \cap X^k$ with $v_n x' x_n \in T \cap X^{k+2}$. Since the latter set is compact, we can find a convergent subsequence $(v_{n_\ell} x' x_{n_\ell})_{\ell \geq 0}$ and $vx'x \in T$ with $v_{n_\ell} x' x_{n_\ell} \to vx'x$, as $\ell \to \infty$ . Consequently, $S_{k+1}(x')$ is closed, and sequentially compact, hence compact, since $X$ is Polish. Thus $S_{k+1}(x) \in \mathcal{K}(X)$, provided the former set is not empty. The domain $D_k$ of $S_k$ is

$$\begin{aligned} D_k &= \pi_X \left[\{\langle v, x\rangle \in T \times X | T(vx) \neq \emptyset\}\right] \\ &= \pi_X \left[\{\langle v, x\rangle \in T \times X | T(vx) \cap X \neq \emptyset\}\right]. \end{aligned}$$

If we can show that $(\forall S_{k+1})(F)$ is Borel in $X$ whenever $F \subseteq X$ is closed, then measurability of $D_k$ will follow (among others).

2. In fact, if $F \subseteq X$ is closed, then the compactness assumption for $T$ implies that

$$\{\langle v, x\rangle \in T \times X | T(vx) \cap F \neq \emptyset\}$$

is closed, consequently,

$$H^{(F)} := \{\langle v, x\rangle \in T \times X | T(vx) \subseteq F\}$$

is a $G_\delta$ set, since $F$ is one. Hence $H^{(F)}$ is Borel. Because the section $H^{(F)}_x$ is compact for each $x \in X$, the Novikov Theorem [Sri98, Th. 5.7.1] implies now that

$$\pi_X \left[H^{(F)}\right] = (\forall S_{k+1})(F)$$

is measurable.

3. The map $S_{k+1} : D_k \to \mathcal{K}(X)$ is measurable for each $k \geq 0$, hence there exists an extension $S^*_{k+1} : X \to \mathcal{K}(X)$ such that $S^*_{k+1}$ is Borel [Sri98, Prop. 3.2.3]. Because $S^*_{k+1}$ takes compact and nonempty values in a Polish space we can find by [Dob81, Cor. IV.11] a Markov relation $V_{k+1} : X \rightsquigarrow X$ such that $S^*_{k+1} = \mathsf{supp}\ V_{k+1}$. Hence

$$T = |\{\mathsf{supp}\ V_n | n \in \mathbb{N}\}|,$$

and from Corollary 3 it is apparent that $(V_k)_{k\geq 0}$ is the sequence of refinements for suitably chosen probabilistic relations $K_n : X \rightsquigarrow X$. ■

Concluding the discussion, an automata theoretic representation of a measurable tree $T$ is presented. Suppose that $T^\bullet$ takes always compact values, or that $X$ is $\sigma$-compact. Assume further that $(\exists T^\bullet)(G)$ is open in $T$ whenever $G \subseteq X$ is open, where $T$ inherits its topology from $X^*$. Then it can be shown that there exists a measurable automaton map $\phi : \mathbb{N}^* \to X^*$ (thus $|\phi(\beta)| = |\beta|$, and $\phi(\beta)$ is a prefix of $\phi(\beta n)$ for each $n \in \mathbb{N}$) such that

$$\{\text{last}(\phi(\beta)) | \beta \in \mathbb{N}^{k+1}\} \cap T(x_0 \ldots x_{k-1})$$

is dense in $T(x_0 \ldots x_{k-1})$. The proof constructs a learning system [Men73] from $T$ and utilizes a representation given for these systems in terms of deterministic automata [Dob81, Prop. IV.15]; the technical details are omitted. Consequently the measurable tree can be represented quite concisely as a single deterministic automaton with input $\mathbb{N}$ and output $X$.

# 5 Conclusion and Further Work

The refinement operation was defined for probabilistic relations, and two representation theorems were given. For the Markovian case., i.e., the case that terminating computations are modelled, the refinement operation was related to memoryless relations on an infinite product through projection and disintegration, and the connection between trees and refinements that — so apparent in the deterministic case — could be established for the probabilistic case, too. These representations hinge on topological assumptions about the underlying universe in which specifications are done; without these assumptions the necessary tools from measure theory are not available. For trees, we need even a compactness assumption.

The refinement in the present paper is performed using the ordinary product of probabilistic relations. If termination is also to be considered, the demonic product of these relations will have to be used, and studied systematically. This product has been proposed and investigated in [Dob01b] as a probabilistic counterpart to the demonic product for nondeterministic relations, see e.g. [DMN, BKS97]. The study of the interplay between this demonic refinement and the corresponding refinement for the nondeterministic fringe relations will provide some insight into the structural similarities of these relations, which show already some striking similarities when viewed through the monadic looking glass [Gir81, Pan98].

The refinement studied here is of the hierarchical type. Ghezzi at al. [GJM91, 4.2.6] discuss the well know fact that hierarchical decompositions are not always adequate when larger systems are constructed. Consequently we want to see how probabilistic relations may be used for more general decompositions of systems, contributing to a general theory of compositionality, and in this way to interaction.

## References

[AaPP98] 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.

[Abr96] S. Abramsky. Retracing some paths in process algebra. In *Proc. CONCUR'96*, volume 1119 of *Lecture Notes in Computer Science*, pages 1 – 17, Berlin, 1996. Springer-Verlag.

[BKS97] C. Brink, W. Kahl, and G. Schmidt, editors. *Relational Methods in Computer Science*. Advances in Computing Science. Springer-Verlag, Wien, New York, 1997.

[DMN] J. Desharnais, A. Mili, and T. T. Nguyen. Refinement and demonic semantics. In *[BKS97]*, pages 166 – 184.

[Dob81] E.-E. Doberkat. *Stochastic Automata — Nondeterminism, Stability, and Prediction*, volume 113 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, 1981.

[Dob01a] E.-E. Doberkat. The converse of a probabilistic relation. Technical Report 113, Chair for Software Technology, University of Dortmund, June 2001.

[Dob01b] E.-E. Doberkat. The demonic product of probabilistic relations. Technical Report 116, Chair for Software Technology, University of Dortmund, September 2001.

[Gir81] 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.

[GJM91] C. Ghezzi, M. Jazayeri, and D. Mandrioli. *Fundamentals of Software Engineering*. Prentice-Hall International, Englewood Cliffs, NJ, 1991.

[Him75] C. J. Himmelberg. Measurable relations. *Fund. Math.*, 87:53 – 72, 1975.

[Men73] W. Menzel. An extension of the theory of learning systems. *Acta Informatica*, 2:357 – 381, 1973.

[MMS96] C. Morgan, A. McIver, and K. Seidel. Probabilistic predicate transformers. *ACM Trans. Prog. Lang. Syst.*, 18(3):325 – 353, May 1996.

[Pan98] 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.

[Par67] K. R. Parthasarathy. *Probability Measures on Metric Spaces*. Academic Press, New York, 1967.

[Sri98] S. M. Srivastava. *A Course on Borel Sets*. Number 180 in Graduate Texts in Mathematics. Springer-Verlag, Berlin, 1998.