Coalgebraic Trace Semantics for Probabilistic Transition Systems based on Measure Theory*

Henning Kerstan and Barbara König

Universität Duisburg-Essen, Duisburg, Germany
{henning.kerstan,barbara_koenig}@uni-due.de

Abstract. Coalgebras in a Kleisli category yield a generic definition of trace semantics for various types of labelled transition systems. In this paper we apply this generic theory to generative *probabilistic transition systems*, short *PTS*, with *arbitrary (possibly uncountable) state spaces*. We consider the *sub-probability monad* and the *probability monad (Giry monad)* on the category of measurable spaces and measurable functions. Our main contribution is that the existence of a final coalgebra in the Kleisli category of these monads is closely connected to the measure-theoretic extension theorem for sigma-finite pre-measures. In fact, we obtain a practical definition of the trace measure for both *finite* and *infinite traces* of PTS that subsumes a well-known result for discrete probabilistic transition systems.

1 Introduction

Coalgebra [11,16] is a general framework in which several types of transition systems can be studied (deterministic and non-deterministic automata, weighted automata, transition systems with non-deterministic and probabilistic branching, etc.). One of the strong points of coalgebra is that it induces – via the notion of coalgebra homomorphism and final coalgebra – a notion of behavioural equivalence for all these types of systems. The resulting behavioural equivalence is usually some form of bisimilarity. However, [10] has shown that by modifying the category in which the coalgebra lives, one can obtain different notions of behavioural equivalence, such as trace equivalence.

We will shortly describe the basic idea: given a functor F, describing the branching type of the system, a coalgebra in the category **Set** is a function $\alpha\colon X \to FX$, where X is a set. Consider, for instance, the functor $FX = \mathcal{P}_{fin}(\mathcal{A} \times X + \mathbf{1})$, where $\mathcal{P}_{fin}$ is the finite powerset functor and $\mathcal{A}$ is the given alphabet. This setup allows us to specify finitely branching non-deterministic automata where a state $x \in X$ is mapped to a set of tuples of the form (a,y), where $a \in \mathcal{A}, y \in X$, describing transitions. The set contains the symbol ✓ (for termination) – the only element contained in the one-element set $\mathbf{1}$ – whenever x is a final state.

A coalgebra homomorphism maps sets of states of a coalgebra to sets of states of another coalgebra, preserving the branching structure. Furthermore, the final coalgebra – if it exists – is the final object in the category of coalgebras. Every coalgebra has a

* The original publication is available at www.springerlink.com.

unique homomorphism into the final coalgebra and two states are mapped to the same state in the final coalgebra iff they are behaviourally equivalent.

Now, applying this notion to the example above induces bisimilarity, whereas usually the appropriate notion of behavioural equivalence for non-deterministic finite automata is language equivalence. One of the ideas of [10] is to view a coalgebra $X \to \mathcal{P}(\mathcal{A} \times X + \mathbf{1})$ not as an arrow in **Set**, but as an arrow $X \to \mathcal{A} \times X + \mathbf{1}$ in **Rel**, the Kleisli category of the powerset monad. This induces trace equivalence, instead of bisimilarity, with the underlying intuition that non-determinism is a side-effect that is "hidden" within the monad. This side effect is not present in the final coalgebra (which consists of the set $\mathcal{A}^*$ with a suitable coalgebra structure), but in the arrow from a state $x \in X$ to $\mathcal{A}^*$, which is a relation, and relates each state with all words accepted from this state.

In [10] it is also proposed to obtain probabilistic trace semantics for the Kleisli category of the (discrete) subdistribution monad $\mathcal{D}$. Hence coalgebras in this setting are functions of the form $X \to \mathcal{D}(\mathcal{A} \times X + \mathbf{1})$ (modelling probabilistic branching and termination), seen as arrows in the corresponding Kleisli category. From a general result in [10] it again follows that the final coalgebra is carried by $\mathcal{A}^*$, where the mapping into the final coalgebra assigns to each state a probability distribution over its traces. In this way one obtains the finite trace semantics of generative probabilistic systems [17,8].

The contribution in [10] is restricted to discrete probability spaces, where the probability distributions always have at most countable support [18]. This might seem sufficient for practical applications at first glance, but it has two important drawbacks: first, it excludes several interesting systems that involve uncountable state spaces (see for instance the examples in [15]). Second, it excludes the treatment of infinite traces, as detailed in [10], since the set of all infinite traces is uncountable and hence needs measure theory to be treated appropriately. This is an intuitive reason for the choice of the subdistribution monad – instead of the distribution monad – in [10]: for a given state, it might always be the case that a non-zero "probability mass" is associated to the infinite traces leaving this state, which – in the discrete case – can not be specified by a probability distribution over all words.

Hence, we generalize the results concerning probabilistic trace semantics from [10] to the case of uncountable state spaces, by working in the Kleisli category of the (continuous) subprobability monad over **Meas** (the category of measurable spaces). Unlike in [10] we do not derive the final coalgebra via a generic construction (building the initial algebra of the functor), but we define the final coalgebra directly. Furthermore we consider the Kleisli category of the (continuous) probability monad (Giry monad) and treat the case with and without termination. In the former case we obtain a coalgebra over the set $\mathcal{A}^\infty$ (finite and infinite traces over $\mathcal{A}$) and in the letter over the set $\mathcal{A}^\omega$ (infinite traces), which shows the naturality of the approach. For completeness we also consider the case of the subprobability monad without termination, which results in a trivial final coalgebra over the empty set. In all cases we obtain the natural trace measures as instances of the generic coalgebraic theory.

Since, to our knowledge, there is no generic construction of the final coalgebra for these cases, we construct the respective final coalgebras directly and show their correctness by proving that each coalgebra admits a unique homomorphism into the final coalgebra. Here we rely on the measure-theoretic extension theorem for sigma-finite pre-measures.

2 Background Material and Preliminaries

We assume that the reader is familiar with the basic definitions of category theory. However, we will provide a brief introduction to measure theory and integration, coalgebra, coalgebraic trace semantics and Kleisli categories - of course all geared to our needs. For a more detailed analysis of many of the given proofs we refer to [12] which is the primary source for the results presented in this paper. Moreover, there is a long version of this paper ([13]) including all the missing proofs.

2.1 Notation

By **1** we denote a singleton set, its unique element is $\checkmark$. For arbitrary sets X,Y we write $X \times Y$ for the usual cartesian product and the disjoint union $X+Y$ is the set $\{(x,0),(y,1) \mid x \in X, y \in Y\}$. Whenever $X \cap Y = \emptyset$ this coincides with (is isomorphic to) the usual union $X \cup Y$ in an obvious way and we often write $X \uplus Y$. For set inclusion we write $\subset$ for strict inclusion and $\subseteq$ otherwise. The set of extended reals is the set $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ and $\overline{\mathbb{R}}_+$ is the set of non-negative extended reals.

2.2 A Brief Introduction to Measure Theory [2,6]

Measure theory generalizes the idea of length, area or volume. Its most basic definition is that of a *σ-algebra (sigma-algebra)*. Given an arbitrary set X we call a set Σ of subsets of X a *σ-algebra* iff it contains the empty set and is closed under absolute complement and countable union. The tuple (X,Σ) is called a *measurable space*. We will sometimes call the set X itself a measurable space, keeping in mind that there is an associated σ-algebra which we will then denote by Σ_X. For any subset $\mathcal{G} \subseteq \mathcal{P}(X)$ we can always uniquely construct the smallest σ-algebra on X containing $\mathcal{G}$ which is denoted by $\sigma_X(\mathcal{G})$. We call $\mathcal{G}$ the *generator* of $\sigma_X(\mathcal{G})$, which in turn is called *the σ-algebra generated by* $\mathcal{G}$. It is known, that σ_X is a monotone and idempotent operator. The elements of a σ-algebra on X are called the *measurable sets* of X.

Similar to the definition of a σ-algebra we call a subset $\mathcal{S} \subseteq \mathcal{P}(X)$ a *semi-ring of sets* iff it contains the empty set, is closed under pairwise intersection and any relative complement of two sets in $\mathcal{S}$ is the disjoint union of finitely many sets in $\mathcal{S}$. It is easy to see that every σ-algebra is a semi-ring of sets but the reverse is false.

A non-negative function $\mu\colon \mathcal{S} \to \overline{\mathbb{R}}_+$ defined on a semi-ring $\mathcal{S}$ is called a *pre-measure* on X if it assigns 0 to the empty set and is σ-additive, i.e. for a sequence $(S_n)_{n\in\mathbb{N}}$ of mutually disjoint sets in $\mathcal{S}$ where $(\uplus_{n\in\mathbb{N}} S_n) \in \mathcal{S}$ we must have $\mu(\uplus_{n\in\mathbb{N}} S_n) = \sum_{n\in\mathbb{N}} \mu(S_n)$. A pre-measure is called *σ-finite* if there is a sequence $(S_n)_{n\in\mathbb{N}}$ of sets in $\mathcal{S}$ such that their union is X and $\mu(S_n) < \infty$ for all $n \in \mathbb{N}$. Whenever $\mathcal{S}$ is a σ-algebra we call μ a *measure* and the tuple $(X,\mathcal{S},\mu)$ a *measure space*. In that case μ is said to be *finite* iff $\mu(X) < \infty$ and for the special cases $\mu(X) = 1$ (or $\mu(X) \leq 1$) μ is called a *probability measure* (or *sub-probability measure* respectively). The most significant theorem from measure theory which we will use in this paper is the extension theorem for σ-finite pre-measures, for which a proof can be found e.g. in [6].

Proposition 1 (Extension Theorem for σ-finite Pre-Measures). *Let X be an arbitrary set, $\mathcal{S} \subseteq \mathcal{P}(X)$ be a semi-ring of sets and $\mu \colon \mathcal{S} \to \overline{\mathbb{R}}_+$ be a σ-finite pre-measure. Then there exists a uniquely determined measure $\hat{\mu} \colon \sigma_X(\mathcal{S}) \to \overline{\mathbb{R}}_+$ such that $\hat{\mu}|_{\mathcal{S}} = \mu$.*

This theorem can on the one hand be used to construct measures and on the other hand it provides an equality test for σ-finite measures.

Corollary 2 (Equality of σ-finite Measures). *Let X be an arbitrary set, $\mathcal{S} \subseteq \mathcal{P}(X)$ be a semi-ring of sets and $\mu, \nu \colon \sigma_X(\mathcal{S}) \to \overline{\mathbb{R}}$ be σ-finite measures. Then μ and ν are equal iff they agree on all elements of the semi-ring.*

2.3 The Category of Measurable Spaces and Functions

Let X and Y be measurable spaces. A function $f \colon X \to Y$ is called *measurable* iff the pre-image of any measurable set of Y is a measurable set of X. The category **Meas** has measurable spaces as objects and measurable functions as arrows. Composition of arrows is function composition and the identity arrow is the identity function.

The product of two measurable spaces (X, Σ_X) and (Y, Σ_Y) is the set $X \times Y$ endowed with the σ-algebra generated by $\Sigma_X * \Sigma_Y$, the set of so-called "rectangles" which is $\{S_X \times S_Y \mid S_X \in \Sigma_X, S_Y \in \Sigma_Y\}$. It is called the *product σ-algebra* of Σ_X and Σ_Y and is denoted by $\Sigma_X \otimes \Sigma_Y$. Whenever Σ_X and Σ_Y have suitable generators, we can also construct a possibly smaller generator for the product σ-algebra than the set of all rectangles.

Proposition 3 (Generators for the Product σ-Algebra, [6]). *Let X, Y be arbitrary sets and $\mathcal{G}_X \subseteq \mathcal{P}(X), \mathcal{G}_Y \subseteq \mathcal{P}(Y)$ such that $X \in \mathcal{G}_X$ and $Y \in \mathcal{G}_Y$. Then the following holds: $\sigma_{X \times Y}(\mathcal{G}_X * \mathcal{G}_Y) = \sigma_X(\mathcal{G}_X) \otimes \sigma_Y(\mathcal{G}_Y)$.*

We remark that we can construct product endofunctors on the category of measurable spaces and functions.

Definition 4 (Product Functors). *Let Z be a measurable space. The endofunctor $Z \times \mathrm{Id}_{\mathbf{Meas}}$ maps a measurable space X to $(Z \times X, \Sigma_Z \otimes \Sigma_X)$ and a measurable function $f \colon X \to Y$ to the measurable function $F(f) \colon Z \times X \to Z \times Y, (z,x) \mapsto (z, f(x))$. The functor $\mathrm{Id}_{\mathbf{Meas}} \times Z$ is constructed analogously.*

The co-product of two measurable spaces (X, Σ_X) and (Y, Σ_Y) is the set $X + Y$ endowed with $\Sigma_X \oplus \Sigma_Y := \{S_X + S_Y \mid S_X \in \Sigma_X, S_Y \in \Sigma_Y\}$ as σ-algebra, the *disjoint union σ-algebra*. Note that in contrast to the product no σ-operator is needed because $\Sigma_X \oplus \Sigma_Y$ itself is already a σ-algebra whereas $\Sigma_X * \Sigma_Y$ is usually no σ-algebra. For generators of the disjoint union σ-algebra there is a comparable result to the one given above for the product σ-algebra.

Proposition 5 (Generators for the Disjoint Union σ-Algebra). *Let X, Y be arbitrary sets and $\mathcal{G}_X \subseteq \mathcal{P}(X), \mathcal{G}_Y \subseteq \mathcal{P}(Y)$ such that $\emptyset \in \mathcal{G}_X$ and $Y \in \mathcal{G}_Y$. Then the following holds: $\sigma_{X+Y}(\mathcal{G}_X \oplus \mathcal{G}_Y) = \sigma_X(\mathcal{G}_X) \oplus \sigma_Y(\mathcal{G}_Y)$.*

A short proof for this can be found in [13]. As before we can construct endofunctors, the co-product functors.

Definition 6 (Co-Product Functors). *Let Z be a measurable space. The endofunctor* $\mathrm{Id}_{\mathbf{Meas}} + Z$ *maps a measurable space X to $(X+Z, \Sigma_X \oplus \Sigma_Z)$ and a measurable function $f\colon X \to Y$ to the measurable function $F(f)\colon X+Z \to Y+Z$ which acts like f on X and like the identity on Z. The functor* $\mathrm{Id}_{\mathbf{Meas}} + Z$ *is constructed analogously.*

For isomorphisms in **Meas** we provide the following characterization, where again the proof can be found in [13].

Proposition 7 (Isomorphisms in Meas). *Two measurable spaces X and Y are isomorphic in* **Meas** *iff there is a bijective function $\varphi\colon X \to Y$ such that*[1] *$\varphi(\Sigma_X) = \Sigma_Y$. If Σ_X is generated by a set $\mathcal{S} \subseteq \mathcal{P}(X)$ then X and Y are isomorphic iff there is a bijective function $\varphi\colon X \to Y$ such that Σ_Y is generated by $\varphi(\mathcal{S})$. In this case $\mathcal{S}$ is a semi-ring of sets (a σ-algebra) iff $\varphi(\mathcal{S})$ is a semi-ring of sets (a σ-algebra).*

2.4 Kleisli Categories and Liftings of Endofunctors

Given a monad (T, η, μ) on a category $\mathbf{C}$ we can define a new category, the Kleisli category of T, where the objects are the same as in $\mathbf{C}$ but every arrow in the new category corresponds to an arrow $f\colon X \to TY$ in $\mathbf{C}$. Thus, arrows in the Kleisli category incorporate side effects specified by a monad [10,1]. In the following definition we will adopt the notation used by S. Mac Lane [14, Theorem VI.5.1], as it allows us to distinguish between objects and arrows in the base category $\mathbf{C}$ and their associated objects and arrows in the Kleisli category $\mathcal{K}\ell(T)$.

Definition 8 (Kleisli Category). *Let (T, η, μ) be a monad on a category* $\mathbf{C}$. *To each object X of* $\mathbf{C}$ *we associate a new object X_T and to each arrow $f\colon X \to TY$ of* $\mathbf{C}$ *we associate a new arrow $f^\flat\colon X_T \to Y_T$. Together these objects and arrows form a new category $\mathcal{K}\ell(T)$, the* Kleisli category of T, *where composition of arrows $f^\flat\colon X_T \to Y_T$ and $g^\flat\colon Y_T \to Z_T$ is defined as: $g^\flat \circ f^\flat := (\mu_Z \circ T(g) \circ f)^\flat$. For every object X_T the identity arrow is $\mathrm{id}_{X_T} = (\eta_X)^\flat$.*

Given an endofunctor F on $\mathbf{C}$, we now want to construct an endofunctor $\overline{F}$ on $\mathcal{K}\ell(T)$ that "resembles" F: Since objects in $\mathbf{C}$ and objects in $\mathcal{K}\ell(T)$ are basically the same, we want $\overline{F}$ to coincide with F on objects i.e. $\overline{F}(X_T) = (FX)_T$. It remains to define how $\overline{F}$ shall act on arrows $f^\flat\colon X_T \to Y_T$ such that it "resembles" F. We note that for the associated arrow $f\colon X \to TY$ we have $F(f)\colon FX \to FTY$. If we had a map $\lambda_Y\colon FTY \to TFY$ to "swap" the endofunctors F and T, we could simply define $\overline{F}(f^\flat) := (\lambda_Y \circ F(f))^\flat$ which is exactly what we are going to do.

Definition 9 (Distributive Law). *Let (T, η, μ) be a monad on a category* $\mathbf{C}$ *and F be an endofunctor on* $\mathbf{C}$. *A natural transformation $\lambda\colon FT \Rightarrow TF$ is called a distributive law iff for all X we have $\lambda_X \circ F(\eta_X) = \eta_{FX}$ and $\mu_{FX} \circ T(\lambda_X) \circ \lambda_{TX} = \lambda_X \circ F(\mu_X)$.*

Whenever we have a distributive law we can define the lifting of a functor.

[1] For $\mathcal{S} \subseteq \mathcal{P}(X)$ and a function $f\colon X \to Y$ let $\varphi(\mathcal{S}) = \{\varphi(S_X) \mid S_X \in \mathcal{S}\}$.

Definition 10 (Lifting of a Functor). *Let (T,η,μ) be a monad on a category* $\mathbf{C}$ *and F be an endofunctor on* $\mathbf{C}$ *with a distributive law $\lambda\colon FT \Rightarrow TF$. The distributive law induces a* lifting *of F to an endofunctor $\overline{F}\colon \mathcal{K}\ell(T) \to \mathcal{K}\ell(T)$ where for each object X_T of $\mathcal{K}\ell(T)$ we define $\overline{F}(X_T) = (FX)_T$ and for each arrow $f^\flat\colon X_T \to Y_T$ we define $\overline{F}(f^\flat)\colon \overline{F}(X_T) \to \overline{F}(Y_T)$ via $\overline{F}(f^\flat) := (\lambda_Y \circ Ff)^\flat$.*

2.5 Coalgebraic Trace Semantics

We recall that for an endofunctor F on a category $\mathbf{C}$ an (F-)coalgebra is a pair (X,α) where X is an object and $\alpha\colon X \to FX$ is an arrow of $\mathbf{C}$. An F-coalgebra homomorphism between two F-coalgebras $(X,\alpha),(Y,\beta)$ is an arrow $\varphi\colon X \to Y$ in $\mathbf{C}$ such that $\beta\circ\varphi = F(\varphi)\circ\alpha$. We call an F-coalgebra (Ω,κ) final iff for every F-coalgebra (X,α) there is a unique F-coalgebra-homomorphism $\varphi_X\colon X \to \Omega$.

By choosing a suitable category and a suitable endofunctor, many (labelled) transition systems can be modelled as F-coalgebras. The final coalgebra - if it exists - can be seen as the "universe of all possible behaviours" and the unique map into it yields a behavioural equivalence: Two states are equivalent iff they are mapped identically into the final coalgebra. Whenever transition systems incorporate side-effects, these can be "hidden" in a monad. In this case the final coalgebra of an endofunctor in the Kleisli category of this monad yields a notion of trace semantics ([9], [18]). In this case, the side-effects from the original system are not part of the final coalgebra, but are contained in the unique map into the final coalgebra.

2.6 The Lebesgue Integral

Before we can define the probability and the sub-probability monad, we give a crash course in integration loosely based on [2,6]. For that purpose let us fix a measurable space X, a measure μ on X and a Borel-measurable[2] function $f\colon X \to \overline{\mathbb{R}}$. We call f *simple* iff it attains only finitely many values, say $f(X) = \{\alpha_1,...,\alpha_N\}$. The integral of such a simple function f is then defined to be the μ-weighted sum of the α_n, formally $\int_X f\,\mathrm{d}\mu = \sum_{n=1}^{N} \alpha_n\mu(S_n)$ where $S_n = f^{-1}(\alpha_n) \in \Sigma_X$. Whenever f is non-negative we can approximate it from below using non-negative simple functions. In this case we define the integral to be $\int_X f\,\mathrm{d}\mu := \sup\{\int_X s\,\mathrm{d}\mu \mid s \text{ non-negative and simple s.t. } 0 \le s \le f\}$. For arbitrary f we decompose it into its positive part $f^+ = \max\{f,0\}$ and negative part $f^- := \max\{-f,0\}$ which are both non-negative and Borel-measurable. We denote that $f = f^+ - f^-$ and consequently we define the integral of f to be the difference $\int_X f\,\mathrm{d}\mu := \int_X f^+\,\mathrm{d}\mu - \int_X f^-\,\mathrm{d}\mu$ if not both integrals on the right hand side are $+\infty$. In the latter case we say that the integral does not exist. Whenever it exists and is finite we call f a (μ-)integrable function. Instead of $\int_X f\,\mathrm{d}\mu$ we will sometimes write $\int_X f(x)\,\mathrm{d}\mu(x)$ or $\int_{x\in X} f(x)\,\mathrm{d}\mu(x)$ which is useful if we have functions with more than one argument or multiple integrals. Note that this does not imply that singleton sets are measurable.

For every measurable set $S \in \Sigma_X$ its characteristic function $\chi_S\colon X \to \{0,1\}$, which is 1 iff $x \in S$ and 0 otherwise, is integrable and for integrable f the product $\chi_S \cdot f$ is also integrable and we write $\int_S f\,\mathrm{d}\mu$ for $\int_X \chi_S \cdot f\,\mathrm{d}\mu$. Some useful properties of the

[2] A function $f\colon X \to \overline{\mathbb{R}}$ is Borel-measurable iff $\forall t \in \overline{\mathbb{R}} : f^{-1}([-\infty,t]) \in \Sigma_X$.

integral are that it is linear, i.e. for integrable $f,g\colon X \to \overline{\mathbb{R}}$ we have $\int \alpha f + \beta g \,\mathrm{d}\mu = \alpha \int f \,\mathrm{d}\mu + \beta \int g \,\mathrm{d}\mu$ and monotone, i.e. $f \leq g$ implies $\int f \,\mathrm{d}\mu \leq \int g \,\mathrm{d}\mu$. We will state one result explicitly which we will use in our proofs.

Proposition 11 ([2, Theorem 1.6.12]). *Let X,Y be measurable spaces, μ be a measure on X, $f\colon Y \to \overline{\mathbb{R}}$ be a Borel-measurable function and $g\colon X \to Y$ be a measurable function. Then $\mu_g := \mu \circ g^{-1}$ is a measure on Y, the so-called* image-measure *and f is μ_g-integrable iff $f \circ g$ is μ-integrable and in this case we have $\int_S f \,\mathrm{d}\mu_g = \int_{g^{-1}(S)} f \circ g \,\mathrm{d}\mu$ for all $S \in \Sigma_Y$.*

2.7 The Probability and the Sub-Probability Monad

We are now going to present the probability monad (Giry monad) and the sub-probability monad as presented e.g. in [7] and [15]. First, we define the endofunctors of these monads.

Definition 12 (Probability and Sub-Probability Functor). *The* probability-functor $\mathbb{P}\colon \mathbf{Meas} \to \mathbf{Meas}$ *maps a measurable space (X,Σ_X) to the measurable space $\big(\mathbb{P}(X),\Sigma_{\mathbb{P}(X)}\big)$ where $\mathbb{P}(X)$ is the set of all probability measures on Σ_X and $\Sigma_{\mathbb{P}(X)}$ is the smallest σ-algebra such that the* evaluation maps*:*

$$\forall S \in \Sigma_X : \quad p_S\colon \mathbb{P}(X) \to [0,1], P \mapsto P(S) \tag{1}$$

are Borel-measurable. For any measurable function $f\colon X \to Y$ between measurable spaces $(X,\Sigma_X),(Y,\Sigma_Y)$ the arrow $\mathbb{P}(f)$ maps a probability measure P to its image measure:

$$\mathbb{P}(f)\colon \mathbb{P}(X) \to \mathbb{P}(Y), P \mapsto P_f := P \circ f^{-1} \tag{2}$$

If we take sub-probabilities instead of probabilities we can construct the sub-probability functor $\mathbb{S}$ analogously.

Having defined the endofunctors, we continue by constructing the unit and multiplication natural tranformations.

Definition 13 (Unit and Multiplication). *Let $T \in \{\mathbb{S},\mathbb{P}\}$. We obtain two natural transformations $\eta\colon \mathrm{Id}_{\mathbf{Meas}} \Rightarrow T$ and $\mu\colon T^2 \Rightarrow T$ by defining η_X,μ_X for every measurable space (X,Σ_X) as follows:*

$$\eta_X\colon X \to T(X),\, x \mapsto \delta_x^X \tag{3}$$

$$\mu_X\colon T^2(X) \to T(X),\, \mu_X(P)(S) := \int p_S \,\mathrm{d}P \quad \forall S \in \Sigma_X \tag{4}$$

where $\delta_x^X\colon \Sigma_X \to [0,1]$ is the Dirac measure which is 1 on $S \in \Sigma_X$ iff $x \in S$ and 0 otherwise. The map p_S ist the evaluation map (1) *from above.*

If we combine all the ingredients we obtain the following result which also guarantees the soundness of the previous definitions:

Proposition 14 ([7]). *$(\mathbb{S},\eta,\mu)$ and $(\mathbb{P},\eta,\mu)$ are monads on* **Meas***.*

3 Main Results

There is a big variety of probabilistic transition systems [18,8]. We will deal with four slightly different versions of so-called *generative* PTS. The underlying intuition is that, according to a probability measure, an action from the alphabet $\mathcal{A}$ and a set of possible successor states are chosen. We distinguish between probabilistic branching according to sub-probability and probability measures and furthermore we treat systems without and with termination.

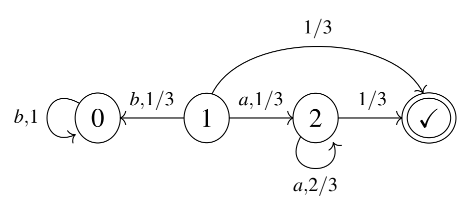
Definition 15 (Probabilistic Transition System (PTS)). *A* probabilistic transition system *is a tuple* $(\mathcal{A}, X, \alpha)$ *where* $\mathcal{A}$ *is a finite alphabet (endowed with* $\mathcal{P}(\mathcal{A})$ *as* σ*-algebra),* X *is the* state space*, an arbitrary measurable space with* σ*-algebra* Σ_X *and* $\alpha \in \{\alpha_0, \alpha_*, \alpha_\omega, \alpha_\infty\}$ *is the* transition function *where:*

$$\alpha_0 \colon X \to \mathbb{S}(\mathcal{A} \times X),\ \alpha_* \colon X \to \mathbb{S}(\mathcal{A} \times X + \mathbf{1})$$
$$\alpha_\omega \colon X \to \mathbb{P}(\mathcal{A} \times X),\ \alpha_\infty \colon X \to \mathbb{P}(\mathcal{A} \times X + \mathbf{1})$$

Depending on the type of the transition function, we call the PTS a $\diamond$-PTS with[3] $\diamond \in \{0, *, \omega, \infty\}$*. For every* $x \in X$ *and every* $a \in \mathcal{A}$ *we define the finite sub-probability measure* $\mathbf{P}_{x,a} \colon \Sigma_X \to [0,1]$ *where* $\mathbf{P}_{x,a}(S) := \alpha(x)(\{a\} \times S)$ *for every* $S \in \Sigma_X$*. Intuitively,* $\mathbf{P}_{x,a}(S)$ *is the probability of making an* a*-transition from the state* $x \in X$ *to any state* $y \in S$*. Whenever* X *is a countable set and* $\Sigma_X = \mathcal{P}(X)$ *we call the PTS* discrete.

We will now take a look at a small example ∞-PTS before we continue to build up our theory.

Example 16 (Discrete PTS with Finite and Infinite Traces). Let $\mathcal{A} = \{a, b\}$, $X = \{0, 1, 2\}$, $\Sigma_X = \mathcal{P}(X)$ and $\alpha := \alpha_\infty \colon X \to \mathbb{P}(\mathcal{A} \times X + \mathbf{1})$ such that we obtain the following system:

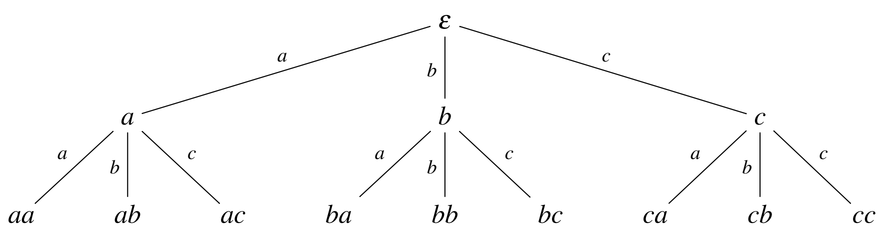


Obviously ✓ is the unique final state which has only incoming transitions bearing probabilities and no labels. This should be interpreted as follows: "From state 1 the system terminates immediately with probability $1/3$".

In order to define a trace measure on these probabilistic transition systems, we need suitable σ-algebras on the sets of words. While the set of all finite words, $\mathcal{A}^*$, is rather simple - we take $\mathcal{P}(\mathcal{A}^*)$ as σ-algebra - the set of all infinite words, $\mathcal{A}^\omega$, and also the set of all finite and infinite words, $\mathcal{A}^\infty$, needs some consideration. For a word $u \in \mathcal{A}^*$ we call the set of all infinite words that have u as a prefix the ω-cone of u, denoted by $\uparrow_\omega \{u\}$, and similarily we call the set of all finite and infinite words having u as a prefix the ∞-cone ([15, p. 23]) of u and denote it with $\uparrow_\infty \{u\}$.

[3] The reason for choosing these symbols as type-identifiers will be revealed later in this paper.

A cone can be visualized in the following way: We consider the undirected, rooted and labelled tree given by $T = (\mathcal{A}^*, E, l)$ with edges $E := \{\{u, uv\} \mid u \in \mathcal{A}^*, v \in \mathcal{A}\}$, edge-labelling function $l \colon E \to \mathcal{A}, \{u, uv\} \mapsto v$ and $\varepsilon \in \mathcal{A}^*$ as the dedicated root. For $\mathcal{A} = \{a, b, c\}$ the first three levels of the tree can be depicted as follows:



Given a finite word $u \in \mathcal{A}^*$, the ω-cone of u is the set of all infinite paths that begin in ε and contain the vertex u and the ∞-cone of u is the set of all finite and infinite simple paths that begin in ε and contain the vertex u (and thus necessarily have a length which is greater or equal to the length of u). Since the sets of cones are no σ-algebras, we will of course take the σ-algebra generated by them. However, the sets of cones can be augmented in such a way that we obtain semi-rings of sets.

Definition 17 (Cones). *Let $\mathcal{A}$ be a finite alphabet and let $\sqsubseteq \subset \mathcal{A}^* \times \mathcal{A}^\infty$ denote the usual prefix relation on words. For $u \in \mathcal{A}^*$ we define its* ω-cone *to be the set $\uparrow_\omega \{u\} := \{v \in \mathcal{A}^\omega \mid u \sqsubseteq v\}$ and analgously we call $\uparrow_\infty \{u\} := \{v \in \mathcal{A}^\infty \mid u \sqsubseteq v\}$ the* ∞-cone *of u. Furthermore we define $\uparrow_0 \{u\} := \emptyset, \uparrow_* \{u\} := \{u\}$.*

With this definition at hand, we can now define the semi-rings we will use to generate σ-algebras on $\mathcal{A}^*, \mathcal{A}^\omega$ and $\mathcal{A}^\infty$.

Definition 18 (Semi-Rings of Sets of Words). *Let $\mathcal{A}$ be a finite alphabet. We define the sets $\mathcal{S}_\diamond := \{\emptyset\} \cup \{\uparrow_\diamond \{u\} \mid u \in \mathcal{A}^\infty\} \subset \mathcal{P}(\mathcal{A}^\diamond)$ for $\diamond \in \{0, *, \omega\}$ and $\mathcal{S}_\infty := \{\uparrow_\infty \{u\} \mid u \in \mathcal{A}^\infty\} \cup \mathcal{S}_* \subset \mathcal{P}(\mathcal{A}^\infty)$.*

Proposition 19. *$\mathcal{S}_0, \mathcal{S}_*, \mathcal{S}_\omega$ and $\mathcal{S}_\infty$ are semi-rings of sets.*

Proving this Proposition is trivial for $\mathcal{S}_0$ and $\mathcal{S}_*$. For $\mathcal{S}_\infty$ we have included a short proof in the long version of this paper, [13], which can easily be adopted to $\mathcal{S}_\omega$.

We remark that many interesting sets will be measurable in the σ-algebra generated by the cones. The singleton-set $\{u\}$ will be measurable for every $u \in \mathcal{A}^\omega$ because $\{u\} = \cap_{v \sqsubseteq u} \uparrow_\omega \{v\} = \cap_{v \sqsubseteq u} \uparrow_\infty \{v\}$ which are countable intersections, or (for ∞-cones only) the set $\mathcal{A}^* = \cup_{u \in \mathcal{A}^*} \{u\}$ and consequently also the set $\mathcal{A}^\omega = \mathcal{A}^\infty \setminus \mathcal{A}^*$ will have to be measurable. The latter will be useful to check to what “extent” a state of a ∞-PTS accepts finite or infinite words/behaviour. One thing about $\mathcal{S}_0$ is worth mentioning: In fact, the above definition yields $\mathcal{S}_0 = \{\emptyset\}$. While this is certainly odd at first sight, it will turn out to be a reasonable specification in our setting.

We will now give a definition of the trace measure which can be understood as the behaviour of a state: it measures the probability of accepting a set of words.

Definition 20 (The Trace Measure). *Let $(\mathcal{A}, X, \alpha)$ be a $\diamond$-PTS. For every state $x \in X$ the trace (sub-)probability measure $\mathbf{tr}_\diamond(x)\colon \sigma_{\mathcal{A}^\diamond}(\mathcal{S}_\diamond) \to [0,1]$ is uniquely defined by the following equations:*

$$\forall a \in \mathcal{A}, \forall u \in \mathcal{A}^* : \quad \mathbf{tr}_\diamond(x)\big(\uparrow_\diamond \{au\}\big) := \int_{x' \in X} \mathbf{tr}_\diamond(x')(\uparrow_\diamond \{u\})\, \mathrm{d}\mathbf{P}_{x,a}(x') \tag{5}$$

and $\mathbf{tr}_\diamond(x)(\emptyset) = 0$, $\mathbf{tr}_*(x)(\uparrow_* \{\varepsilon\}) = \alpha(x)(\mathbf{1})$, $\mathbf{tr}_\omega(x)(\uparrow_\omega \{\varepsilon\}) = 1$, $\mathbf{tr}_\infty(x)(\uparrow_\infty \{\varepsilon\}) = 1$ *and* $\mathbf{tr}_\infty(x)(\{u\}) = \mathbf{tr}_\infty(x)(\uparrow_\infty \{u\}) - \sum_{a \in \mathcal{A}} \mathbf{tr}_\infty(x)(\uparrow_\infty \{au\})$ *where applicable.*

We need to verify that everything is well-defined. In the next proposition we explicitly state what has to be shown.

Proposition 21. *The equations in Definition 20 yield a σ-finite pre-measure $\mathbf{tr}_\diamond(x)\colon \mathcal{S}_\diamond \to [0,1]$ for $\diamond \in \{0, *, \omega, \infty\}$ and every $x \in X$. Moreover, the unique extension of this pre-measure is a (sub-)probability measure.*

Before we prove this proposition, let us try to get a more intuitive understanding of Definition 20 and especially equation (5). First we check how the above definition reduces when we consider discrete systems.

Remark 22. Let $(\mathcal{A}, X, \alpha)$ be a discrete[4] $*$-PTS, i.e. $\alpha\colon X \to \mathbb{S}(\mathcal{A} \times X + \mathbf{1})$. Then $\mathbf{tr}_*(x)(\varepsilon) := \alpha(x)(\checkmark)$ and (5) is equivalent to:

$$\forall a \in \mathcal{A}, \forall u \in \mathcal{A}^* : \quad \mathbf{tr}_*(x)(au) := \sum_{x' \in X} \mathbf{tr}_*(x')(u) \cdot \mathbf{P}_{x,a}(x')$$

which is equivalent to the discrete trace distribution presented in [9] for the sub-distribution monad $\mathcal{D}$ on **Set**.

Having seen this coincidence with known results, we proceed to calculate the trace measure for our example (Ex. 16) which we can only do in our more general setting because this ∞-PTS is a discrete probabilistic transition system which exhibits both finite and infinite behaviour.

Example 23 (Example 16 cont.). We calculate the trace measures for the ∞-PTS from Example 16. We have $\mathbf{tr}_\infty(0) = \delta_{b^\omega}^{\mathcal{A}^\infty}$ because

$$\begin{aligned}
\mathbf{tr}_\infty(0)(\{b^\omega\}) &= \mathbf{tr}_\infty(0)\left(\cap_{k=0}^\infty \uparrow_\infty \{b^k\}\right) = \mathbf{tr}_\infty(0)\left(\mathcal{A}^\infty \setminus \cup_{k=0}^\infty \left(\mathcal{A}^\infty \setminus \uparrow_\infty \{b^k\}\right)\right) \\
&= \mathbf{tr}_\infty(0)\,(\mathcal{A}^\infty) - \mathbf{tr}(0)\left(\cup_{k=0}^\infty \left(\mathcal{A}^\infty \setminus \uparrow_\infty \{b^k\}\right)\right) \\
&\geq 1 - \sum_{k=0}^\infty \mathbf{tr}_\infty(0)\left(\mathcal{A}^\infty \setminus \uparrow_\infty \{b^k\}\right) \\
&= 1 - \sum_{k=0}^\infty \left(1 - \mathbf{tr}_\infty(0)\left(\uparrow_\infty \{b^k\}\right)\right) = 1 - \sum_{k=0}^\infty (1-1) = 1
\end{aligned}$$

Thus we have $\mathbf{tr}_\infty(0)(\mathcal{A}^*) = \mathbf{tr}_\infty(0)\,(\uplus_{u \in \mathcal{A}^*} \{u\}) = 0$ and $\mathbf{tr}_\infty(0)(\mathcal{A}^\omega) = 1$. By induction we can show that $\mathbf{tr}_\infty(2)(\{a^k\}) = (1/3) \cdot (2/3)^k$ and thus $\mathbf{tr}_\infty(2)(\mathcal{A}^*) = 1$ and $\mathbf{tr}_\infty(2)(\mathcal{A}^\omega) = 0$. Furthermore we calculate $\mathbf{tr}_\infty(1)(\{b^\omega\}) = 1/3$, $\mathbf{tr}_\infty(1)(\uparrow_\infty \{a\}) = 1/3$ and $\mathbf{tr}_\infty(1)(\{\varepsilon\}) = 1/3$ yielding $\mathbf{tr}_\infty(1)(\mathcal{A}^*) = 2/3$ and $\mathbf{tr}_\infty(1)(\mathcal{A}^\omega) = 1/3$.

[4] If Z is a countable set and $\mu\colon \mathcal{P}(Z) \to [0,1]$ is a measure, we write $\mu(z)$ for $\mu(\{z\})$.

Recall, that we still have to prove Proposition 21. In order to simplify this proof, we provide a few technical results about the sets $\mathcal{S}_*, \mathcal{S}_\omega, \mathcal{S}_\infty$ for which proofs are given in [13] or in [12].

Lemma 24 (Countable Unions). *Let $(S_n)_{n\in\mathbb{N}}$ be a sequence of mutually disjoint sets in $\mathcal{S}_\omega$ or in $\mathcal{S}_\infty$ such that $\biguplus_{n\in\mathbb{N}} S_n$ is itself an element of $\mathcal{S}_\omega$ or $\mathcal{S}_\infty$. Then $S_n = \emptyset$ for all but finitely many n.*

Lemma 25 (Sigma-Finiteness 1). *A non-negative map $\mu\colon \mathcal{S}_* \to \overline{\mathbb{R}}_+$ where $\mu(\emptyset) = 0$ is always σ-additive and thus a pre-measure.*

Lemma 26 (Sigma-Finiteness 2). *A non-negative map $\mu\colon \mathcal{S}_\omega \to \overline{\mathbb{R}}_+$ where $\mu(\emptyset) = 0$ is σ-additive and thus a pre-measure iff $\mu(\uparrow_\omega \{u\}) = \sum_{a\in\mathcal{A}} \mu(\uparrow_\omega \{ua\})$ for all $u \in \mathcal{A}^*$.*

Lemma 27 (Sigma-Finiteness 3). *A non-negative map $\mu\colon \mathcal{S}_\infty \to \overline{\mathbb{R}}_+$ where $\mu(\emptyset) = 0$ is σ-additive and thus a pre-measure iff $\mu(\uparrow_\infty \{u\}) = \mu(\{u\}) + \sum_{a\in\mathcal{A}} \mu(\uparrow_\infty \{ua\})$ for all $u \in \mathcal{A}^*$.*

Using these results, we can now prove Proposition 21.

Proof (of Proposition 21). For $\diamond = 0$ nothing has to be shown because $\sigma_\emptyset(\{\emptyset\}) = \{\emptyset\}$ and $\mathbf{tr}_0(x)\colon \{\emptyset\} \to [0,1]$ is already uniquely defined by $\mathbf{tr}_0(x)(\emptyset) = 0$. Lemma 25 and Lemma 27 yield immediately that for $\diamond \in \{*, \infty\}$ the equations define a pre-measure. The only difficult case is $\diamond = \omega$ where we will, of course, apply Lemma 26. Let $u = u_1...u_m \in \mathcal{A}^*$ with $u_k \in \mathcal{A}$ for every k, then multiple application of (5) yields:

$$\mathbf{tr}_\omega(x)\big(\uparrow_\omega \{u\}\big) = \int\limits_{x_1\in X} \ldots \int\limits_{x_m\in X} 1 \,\mathrm{d}\mathbf{P}_{x_{m-1},u_m}(x_m) \ldots \mathrm{d}\mathbf{P}_{x,u_1}(x_1)$$

and for arbitrary $a \in \mathcal{A}$ we obtain analogously:

$$\mathbf{tr}_\omega(x)\big(\uparrow_\omega \{ua\}\big) = \int\limits_{x_1\in X} \ldots \int\limits_{x_m\in X} \mathbf{P}_{x_m,a}(X) \,\mathrm{d}\mathbf{P}_{x_{m-1},u_m}(x_m) \ldots \mathrm{d}\mathbf{P}_{x,u_1}(x_1).$$

All integrals exist and are bounded above by 1 so we can use the linearity and monotonicity of the integral to exchange the finite sum and the integrals to obtain that indeed $\sum_{a\in\mathcal{A}} \mathbf{tr}_\omega(x)\big(\uparrow_\omega \{ua\}\big) = \mathbf{tr}_\omega(x)\big(\uparrow_\omega \{u\}\big)$ is valid using the fact that $\sum_{a\in\mathcal{A}} \mathbf{P}_{x_m,a}(X) = \sum_{a\in\mathcal{A}} \alpha(x)(\{a\}\times X) = \alpha(x)(\mathcal{A}\times X) = 1$. Hence also $\mathbf{tr}_\omega(x)\colon \mathcal{S}_\omega \to \overline{\mathbb{R}}_+$ is σ-additive and thus a pre-measure.

Now let us check that the pre-measures are σ-finite. For $\diamond \in \{\omega, \infty\}$ this is obvious and in these cases the unique extension must be a (sub-)probability measure because by definition we have $\mathbf{tr}_\omega(x)(\mathcal{A}^\omega) = 1$ and $\mathbf{tr}_\infty(x)(\mathcal{A}^\infty) = 1$ respectively. For the remaining case ($\diamond = *$) we remark that $\mathcal{A}^* = \biguplus_{u\in\mathcal{A}^*} \{u\}$ which is countable and disjoint. Using induction on the length of $u \in \mathcal{A}^*$ and monotonicity of the integral we can easily verify that $\mathbf{tr}_*(x)(\{u\})$ is always bounded by 1 and hence also in this case $\mathbf{tr}_*(x)$ is σ-finite. Again by induction we can see that for all $n \in \mathbb{N}_0$ we have $\mathbf{tr}_*(x)\left(\mathcal{A}^{\leq n}\right) \leq 1$. Since $\mathbf{tr}_*(x)$ is a measure (and thus non-negative and σ-additive), the sequence given by $\left(\mathbf{tr}_*(x)\left(\mathcal{A}^{\leq n}\right)\right)_{n\in\mathbb{N}_0}$ is a monotonically increasing sequence of real numbers bounded above by 1 and hence has a limit. Furthermore, $\mathbf{tr}_*(x)$ is continuous from below as a measure and we have $\mathcal{A}^{\le n} \subseteq \mathcal{A}^{\le n+1}$ for all $n \in \mathbb{N}_0$ and thus can conclude that

$$\mathbf{tr}_*(x)\left(\mathcal{A}^*\right) = \mathbf{tr}_*(x)\left(\bigcup_{n=1}^{\infty} \mathcal{A}^{\le n}\right) = \lim_{n\to\infty} \mathbf{tr}_*(x)\left(\mathcal{A}^{\le n}\right) = \sup_{n\in\mathbb{N}_0} \mathbf{tr}_*(x)\left(\mathcal{A}^{\le n}\right) \le 1.$$

For more details take a look at [12, Proofs of Theorems 4.14 and 4.24]. $\square$

Now that we know that our definition of a trace measure is mathematically sound, we remember that we wanted to show that it is "natural", meaning that it arises from the final coalgebra in the Kleisli category of the (sub-)probability monad. We now state our main theorem which presents a close connection between the unique existence of the map into the final coalgebra and the unique extension of a family of σ-finite measures.

Theorem 28 (Main Theorem). *Let $T \in \{\mathbb{S}, \mathbb{P}\}$, F be an endofunctor on* **Meas** *with a distributive law $\lambda\colon FT \Rightarrow TF$ and $(\Omega_T, \kappa^\flat)$ be an $\overline{F}$-coalgebra where $\Sigma_{F\Omega} = \sigma_{F\Omega}(\mathcal{S}_{F\Omega})$ for a semi-ring $\mathcal{S}_{F\Omega}$. Then $(\Omega_T, \kappa^\flat)$ is final iff for every $\overline{F}$-coalgebra $(X_T, \alpha^\flat)$ there is a unique (sub-)probability measure $\mathbf{tr}(x)\colon \Sigma_\Omega \to [0,1]$ for every $x \in X$ such that:*

$$\forall S \in \mathcal{S}_{F\Omega}: \quad \int_\Omega p_S \circ \kappa \,\mathrm{d}\mathbf{tr}(x) = \int_{FX} p_S \circ \lambda_\Omega \circ F(\mathbf{tr})\,\mathrm{d}\alpha(x) \tag{6}$$

Proof. We consider the final coalgebra diagram in $\mathcal{K}\ell(T)$:

$$\begin{array}{ccc} X_T & \xrightarrow{\alpha^\flat} & \overline{F}X_T \\ {\scriptstyle \mathbf{tr}^\flat}\downarrow & & \downarrow{\scriptstyle \overline{F}(\mathbf{tr}^\flat) = (\lambda_\Omega \circ F(\mathbf{tr}))^\flat} \\ \Omega_T & \xrightarrow{\kappa^\flat} & \overline{F}\Omega_T \end{array}$$

By definition $(\Omega_T, \kappa^\flat)$ is final iff for every $\overline{F}$-coalgebra $(X_T, \alpha^\flat)$ there is a unique arrow $\mathbf{tr}^\flat\colon X_T \to \Omega_T$ making the diagram commute. We define:

$$g^\flat := \kappa^\flat \circ \mathbf{tr}^\flat \text{ (down, right)} \quad h^\flat := \overline{F}(\mathbf{tr}^\flat) \circ \alpha^\flat \text{ (right, down)}$$

and note that commutativity of this diagram is equivalent to:

$$\forall x \in X, \forall S \in \mathcal{S}_{F\Omega}: \quad g(x)(S) = h(x)(S) \tag{7}$$

because for every $x \in X$ both $g(x)$ and $h(x)$ are (sub-)probability measures and thus σ-finite measures which allows us to apply Corollary 2. We calculate:

$$\begin{aligned} g(x)(S) &= (\mu_{F\Omega} \circ T(\kappa) \circ \mathbf{tr})(x)(S) = \mu_{F\Omega}\left(T(\kappa)(\mathbf{tr}(x))\right)(S) \\ &= \mu_{F\Omega}\left(\mathbf{tr}(x)_\kappa\right)(S) = \int p_S \,\mathrm{d}\mathbf{tr}(x)_\kappa = \int p_S \circ \kappa \,\mathrm{d}\mathbf{tr}(x) \end{aligned}$$

and if we define $\rho := \lambda_\Omega \circ F(\mathbf{tr})\colon FX \to TF\Omega$ we obtain:

$$\begin{aligned} h(x)(S) &= (\mu_{F\Omega} \circ T(\rho) \circ \alpha)(x)(S) = \mu_{F\Omega}\left(T(\rho)(\alpha(x))\right)(S) = \mu_{F\Omega}\left(\alpha(x)_\rho\right)(S) \\ &= \int p_S \,\mathrm{d}\alpha(x)_\rho = \int p_S \circ \rho \,\mathrm{d}\alpha(x) = \int p_S \circ \lambda_\Omega \circ F(\mathbf{tr})\,\mathrm{d}\alpha(x) \end{aligned}$$

and thus (7) is equivalent to (6). $\square$

We immediately obtain the following corollary.

Corollary 29. *Let in Theorem 28* $\kappa = \eta_{F\Omega} \circ \varphi$*, for an isomorphism* $\varphi\colon \Omega \to F\Omega$ *in* **Meas** *and let* $\mathcal{S}_\Omega \subseteq \mathcal{P}(\Omega)$ *be a semi-ring such that* $\Sigma_\Omega = \sigma_\Omega(\mathcal{S}_\Omega)$*. Then equation* (6) *is equivalent to:*

$$\forall S \in \mathcal{S}_\Omega: \quad \mathbf{tr}(x)(S) = \int p_{\varphi(S)} \circ \lambda_\Omega \circ F(\mathbf{tr})\, \mathrm{d}\alpha(x) \tag{8}$$

Proof. Since φ is an isomorphism in **Meas** we know from Proposition 7 that $\Sigma_{F\Omega} = \sigma_\Omega(\varphi(\mathcal{S}_\Omega))$. For every $S \in \Sigma_\Omega$ and every $u \in \Omega$ we calculate:

$$p_{\varphi(S)} \circ \kappa(u) = p_{\varphi(S)} \circ \eta_{F\Omega} \circ \varphi(u) = \delta^{F\Omega}_{\varphi(u)}(\varphi(S)) = \chi_{\varphi(S)}(\varphi(u)) = \chi_S(u)$$

and hence we have $\int p_{\varphi(S)} \circ \kappa\, \mathrm{d}\mathbf{tr}(x) = \int \chi_S\, \mathrm{d}\mathbf{tr}(x) = \mathbf{tr}(x)(S)$. □

Since we want to apply this corollary to sets of words, we now define the necessary isomorphism φ using the characterization given in Proposition 7.

Proposition 30. *Let* $\varphi\colon \mathcal{A}^\infty \to \mathcal{A} \times \mathcal{A}^\infty + \mathbf{1}$*,* $\varepsilon \mapsto \checkmark$*,* $au \mapsto (a,u)$*. Then* φ*,* $\varphi|_{\mathcal{A}^*}$ *and* $\varphi|_{\mathcal{A}^\omega}$ *are bijective functions*[5] *and the following holds:*

$$\sigma_{\mathcal{A}\times\mathcal{A}^\omega}(\varphi(\mathcal{S}_\omega)) = \mathcal{P}(\mathcal{A}) \otimes \sigma_{\mathcal{A}^\omega}(\mathcal{S}_\omega) \tag{9}$$

$$\sigma_{\mathcal{A}\times\mathcal{A}^*+\mathbf{1}}(\varphi(\mathcal{S}_*)) = \mathcal{P}(\mathcal{A}) \otimes \sigma_{\mathcal{A}^*}(\mathcal{S}_*) \oplus \mathcal{P}(\mathbf{1}) \tag{10}$$

$$\sigma_{\mathcal{A}\times\mathcal{A}^\infty+\mathbf{1}}(\varphi(\mathcal{S}_\infty)) = \mathcal{P}(\mathcal{A}) \otimes \sigma_{\mathcal{A}^\infty}(\mathcal{S}_\infty) \oplus \mathcal{P}(\mathbf{1}) \tag{11}$$

We recall that – in order to get a lifting of an endofunctor on **Meas** – we also need a distributive law for the functors we are using to define PTS. A proof for the following proposition is given in [12, Prop. and Def. 4.12 and 4.22].

Proposition 31 (Distributive Laws for the (Sub-)Probability Monad). *Let* $T \in \{\mathbb{S}, \mathbb{P}\}$*. For every measurable space* (X, Σ_X) *we define*

$$\lambda_X\colon \mathcal{A} \times TX \to T(\mathcal{A} \times X),\ (a,P) \mapsto \delta^{\mathcal{A}}_a \otimes P$$

where $\delta^{\mathcal{A}}_a \otimes P$ *denotes the product measure*[6] *of* $\delta^{\mathcal{A}}_a$ *and* P*. Then we obtain a distributive law* $\lambda\colon \mathcal{A} \times T \Rightarrow T(\mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}})$*. In an analogous manner we obtain another distributive law* $\lambda\colon \mathcal{A} \times T + \mathbf{1} \Rightarrow T(\mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}} + \mathbf{1})$ *if we define*

$$\lambda_X\colon \mathcal{A} \times TX + \mathbf{1} \to T(\mathcal{A} \times X + \mathbf{1}),\ (a,P) \mapsto \delta^{\mathcal{A}}_a \odot P, \checkmark \mapsto \delta^{\mathcal{A}\times X+\mathbf{1}}_{\checkmark}$$

for every measurable space (X, Σ_X) *where* $(\delta^{\mathcal{A}}_a \odot P)(S) := (\delta^{\mathcal{A}}_a \otimes P)(S \cap (\mathcal{A} \times X))$ *for every* $S \in \mathcal{P}(\mathcal{A}) \otimes \Sigma_X \oplus \mathcal{P}(\mathbf{1})$*.*

With this result at hand we can finally apply Corollary 29 to the measurable spaces $\emptyset, \mathcal{A}^*, \mathcal{A}^\omega, \mathcal{A}^\infty$, each of which is of course equipped with the σ-algebra generated by the semi-rings $\mathcal{S}_0, \mathcal{S}_*, \mathcal{S}_\omega, \mathcal{S}_\infty$ as defined in Proposition 19, to obtain the final coalgebra and the induced trace semantics for PTS as presented in the following corollary.

---

[5] For a function $f\colon X \to Y$ and $X' \subset X$ we consider $f|_{X'}$ to be $f|_{X'}\colon X' \to f(X')$.

[6] $\delta^{\mathcal{A}}_a \otimes P$ is the unique extension of the measure defined via $\delta^{\mathcal{A}}_a \otimes P(S_{\mathcal{A}} \times S_X) := \delta^{\mathcal{A}}_a(S_{\mathcal{A}}) \cdot P(S_X)$ for all $S_{\mathcal{A}} \times S_X \in \mathcal{P}(\mathcal{A}) * \Sigma_X$.

Corollary 32 (Final Coalgebra and Trace Semantics for PTS). *A PTS* $(\mathcal{A}, X, \alpha)$ *is an* $\overline{F}$*-coalgebra* $(X_T, \alpha^\flat)$ *in* $\mathcal{K}\ell(T)$ *and vice versa. In the following table we present the (carriers of) final* $\overline{F}$*-coalgebras* $\left(\Omega_T, \kappa^\flat\right)$ *in* $\mathcal{K}\ell(T)$ *for all suitable choices of* T *and* F *(depending on the type of the PTS).*

Type	*Monad T*	*Endofunctor F*	*Carrier* Ω_T
0	$\mathbb{S}$	$\mathcal{A} \times X$	$(\emptyset, \{\emptyset\})_T$
$*$	$\mathbb{S}$	$\mathcal{A} \times X + \mathbf{1}$	$(\mathcal{A}^*, \sigma_{\mathcal{A}^*}(\mathcal{S}_*))_T$
ω	$\mathbb{P}$	$\mathcal{A} \times X$	$(\mathcal{A}^\omega, \sigma_{\mathcal{A}^\omega}(\mathcal{S}_\omega))_T$
∞	$\mathbb{P}$	$\mathcal{A} \times X + \mathbf{1}$	$(\mathcal{A}^\infty, \sigma_{\mathcal{A}^\infty}(\mathcal{S}_\infty))_T$

In all cases $\kappa = \eta_{F\Omega} \circ \varphi$ *where* φ *is the isomorphism as defined before. The unique map* $\mathbf{tr}^\flat$ *into the final coalgebra is* $\mathbf{tr}_\diamond(x)$ *as given in Definition 20 for every* $x \in X$.

4 Conclusion, Related and Future Work

We have shown how to obtain coalgebraic trace semantics in a general measure-theoretic setting, thereby allowing uncountable state spaces and infinite trace semantics.

Our work is clearly inspired by [10], generalizing their instantiation to generative probabilistic systems. Probabilistic systems in the general measure-theoretic setting were in detail studied by [21], but note that the author considers bisimilarity and constructs coalgebras in **Meas**, whereas we are working in Kleisli categories based on **Meas**.

In [5] and [15] a very thorough and general overview of properties of labelled Markov processes including the treatment of temporal logics is given. However, the authors do not explicitly cover a coalgebraic notion of trace semantics.

Infinite traces in a general coalgebraic setting have already been studied in [4]. However, this generic theory, once applied to probabilistic systems, is restricted to coalgebras with countable carrier while our setting, which is undoubtedly specific, allows arbitrary carriers for coalgebras of probabilistic systems.

As future work we plan to apply the minimization algorithm introduced in [1] and adapt it to this general setting, by working out the notion of canonical representatives for probabilistic transition system.

Furthermore we plan to define and study a notion of probabilistic trace distance, similar to the distance measure studied in [20,19]. We are also interested in algorithms for calculating this distance, perhaps similar to what has been proposed in [3] for probabilistic bisimilarity.

Acknowledgement. We would like to thank Paolo Baldan, Filippo Bonchi, Mathias Hülsbusch and Alexandra Silva for discussing this topic with us and giving us some valuable hints. Moreover, we are grateful for the detailed feedback from our reviewers.

References

1. Adámek, J., Bonchi, F., Hülsbusch, M., König, B., Milius, S., Silva, A.: A coalgebraic perspective on minimization and determinization. In: Birkedal, L. (ed.) Foundations of Software

Science and Computational Structures. Lecture Notes in Computer Science, vol. 7213, pp. 58–73. Springer (2012)
2. Ash, R.B.: Real Analysis and Probability. Probability and Mathematical Statistics – A Series of Monographs and Textbooks, Academic Press, 111 Fifth Avenue, New York, New York (1972)
3. Chen, D., van Breugel, F., Worrell, J.: On the complexity of computing probabilistic bisimilarity. In: Birkedal, L. (ed.) Foundations of Software Science and Computational Structures. Lecture Notes in Computer Science, vol. 7213, pp. 437–451. Springer (2012)
4. Cîrstea, C.: Generic infinite traces and path-based coalgebraic temporal logics. Electronic Notes in Theoretical Computer Science 264(2), 83–103 (Aug 2010)
5. Doberkat, E.: Stochastic relations: foundations for Markov transition systems. Chapman & Hall/CRC studies in informatics series, Chapman & Hall/CRC (2007)
6. Elstrodt, J.: Maß- und Integrationstheorie. Springer, 5th edn. (2007)
7. Giry, M.: A categorical approach to probability theory. In: Categorical Aspects of Topology and Analysis. Lecture Notes in Mathematics, vol. 915, pp. 68–86. Springer (1981)
8. van Glabbeek, R., Smolka, S.A., Steffen, B., Tofts, C.M.N.: Reactive, generative and stratified models of probabilistic processes. Information and Computation 121, 59–80 (1995)
9. Hasuo, I., Jacobs, B., Sokolova, A.: Generic trace theory. In: International Workshop on Coalgebraic Methods in Computer Science. Electronic Notes in Theoretical Computer Science, vol. 164, pp. 47–65. Elsevier (2006)
10. Hasuo, I., Jacobs, B., Sokolova, A.: Generic trace semantics via coinduction. Logical Methods in Computer Science 3 (4:11), 1–36 (November 2007)
11. Jacobs, B., Rutten, J.: A tutorial on (co)algebras and (co)induction. Bulletin of the European Association for Theoretical Computer Science 62, 222–259 (1997)
12. Kerstan, H.: Trace Semantics for Probabilistic Transition Systems - A Coalgebraic Approach. Diploma thesis, Universität Duisburg-Essen (September 2011), http://jordan.inf.uni-due.de/publications/kerstan/kerstan_diplomathesis.pdf
13. Kerstan, H., König, B.: Coalgebraic trace semantics for probabilistic transition systems based on measure theory. Tech. rep., Abteilung für Informatik und Angewandte Kognitionswissenschaft, Universität Duisburg-Essen (2012), http://jordan.inf.uni-due.de/publications/kerstan/coalgpts_concur12_long.pdf
14. Mac Lane, S.: Categories for the Working Mathematician. Springer, 2nd edn. (1998)
15. Panangaden, P.: Labelled Markov Processes. Imperial College Press (2009)
16. Rutten, J.: Universal coalgebra: a theory of systems. Theoretical Computer Science 249, 3–80 (2000)
17. Sokolova, A.: Coalgebraic Analysis of Probabilistic Systems. Ph.D. thesis, Technische Universiteit Eindhoven (2005)
18. Sokolova, A.: Probabilistic systems coalgebraically: A survey. Theoretical Computer Science 412(38), 5095–5110 (2011), cMCS Tenth Anniversary Meeting
19. van Breugel, F., Worrell, J.: Approximating and computing behavioural distances in probabilistic transition systems. Theoretical Computer Science 360, 373–385 (2005)
20. van Breugel, F., Worrell, J.: A behavioural pseudometric for probabilistic transition systems. Theoretical Computer Science 331, 115–142 (2005)
21. Viglizzo, I.: Final sequences and final coalgebras for measurable spaces. In: Fiadeiro, J., Harman, N., Roggenbach, M., Rutten, J. (eds.) Algebra and Coalgebra in Computer Science, Lecture Notes in Computer Science, vol. 3629, pp. 395–407. Springer (2005)