Diploma Thesis

Trace Semantics for Probabilistic Transition Systems

A Coalgebraic Approach

by

Henning Kerstan

Submitted to the

Faculty of Mathematics

of the

University of Duisburg-Essen

In Partial Fulfillment of the Requirements for the Degree of a

"Diplom Mathematiker"

Date of Issue: May 26, 2011

Date of Submission: September 27, 2011

1st Examiner: Prof. Dr. Rüdiger Schultz

2nd Examiner: Prof. Dr. Barbara König

Abstract

Probabilistic transition systems, short *PTS*, are labelled transition systems where each transition depends on a probability. As in the case of finite automata, we are interested in analyzing the *behaviour* of these systems. A method to do this is to define the *trace* of a state. While the concept of *trace semantics* is easy to grasp for finite automata, the introduction of probabilities complicates the definition of a trace. We need measure and integration theory to obtain a mathematically sound definition of trace semantics for PTS with continuous state space and even for discrete PTS without explicit termination. Instead of defining trace semantics directly, we use a coalgebraic approach: We define two endofunctors on the category of measurable spaces and measurable functions and prove that they can be lifted to endofunctors on the Kleisli category of the (sub-)probability monad. Then we model PTS with and without explicit termination as a coalgebra for this lifted functor and prove that a final coalgebra exists. The unique homomorphism into the final coalgebra yields a definition of trace semantics for PTS which is a sub-probability measure on the set of all finite words in the case of a PTS with explicit termination and a probability measure on the set of all infinite words for PTS without explicit termination.

Contents

1 Introduction **1**

1.1 Labelled Transition Systems 1

1.2 Trace Semantics 3

1.3 Transition Systems as Coalgebra 5

1.4 Main Aim of this Thesis 6

1.5 Structure of this Thesis 6

2 Mathematical Basics **9**

2.1 Set Theory 9

2.1.1 Alphabets and Words 11

2.2 Measure and Integration Theory 12

2.2.1 Sigma-Algebras and Measurable Spaces 13

2.2.2 Semi-Rings of Sets 16

2.2.3 Pre-Measures and Measures 17

2.2.4 Measurable Functions 20

2.2.5 The Lebesgue Integral 21

2.2.6 Disjoint Union Sigma-Algebra 24

2.2.7 Product Sigma-Algebra and Product Measure 25

2.3 Probabilistic Transition Systems 26

2.4 Category Theory 28

2.4.1 Categories 28

2.4.2 Functors 30

2.4.3 Coalgebra 31

2.4.4 Monads and Kleisli Categories 32

2.4.5 Coalgebra in a Kleisli Category 34

2.5 Graph Theory 35

3 The Category of Measurable Spaces **37**

3.1 Definition and Conventions 37

3.2 Initial and Terminal Objects . . . 38
3.3 Products . . . 39
3.4 Coproducts . . . 40
3.5 The (Sub-)Probability Monad . . . 41

4 PTS as Coalgebra 43

4.1 Sigma Algebras on Words . . . 43
4.1.1 A Sigma Algebra on Finite Words . . . 43
4.1.2 A Sigma Algebra on Infinite Words . . . 44
4.2 The Endofunctor $F = \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}}$ and PTS w/o Termination . . . 49
4.2.1 The Endofunctor . . . 49
4.2.2 A Distributive Law . . . 49
4.2.3 PTS w/o Termination as Coalgebra . . . 52
4.2.4 A Final Coalgebra and Trace Semantics for PTS w/o Termination 52
4.2.5 Examples . . . 57
4.3 The Endofunctor $F = \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}} + \mathbf{1}$ and PTS with Termination . . . 59
4.3.1 The Endofunctor . . . 59
4.3.2 A Distributive Law . . . 60
4.3.3 PTS with Termination as Coalgebra . . . 63
4.3.4 A Final Coalgebra and Trace Semantics for PTS with Termination 64
4.3.5 Examples . . . 69

5 Conclusions and Future Work 71

5.1 Conclusions and Remarks . . . 71
5.2 Literature for Further Reading . . . 71
5.3 Open Questions and Future Work . . . 72

Bibliography 75

1 Introduction

In theoretical computer science one discipline is the study of (labelled) transition systems and their behavior. While the conventional definition of labelled transition systems does not include initial or final states, we will slightly modify the definition to include a unique initial state and a unique final state. Doing so we can identify labelled transition systems as a generalization of finite automata. However, for the behavioral analysis of these systems we will "ignore" the initial state again.

1.1 Labelled Transition Systems

Definition 1.1 (Labelled Transition System (LTS))**.** *A* labelled transition system (with initial and final state)*, short* LTS*, is a 6-tuple* $M := (X, \mathcal{L}, \perp, \checkmark, \varepsilon, R)$ *where*

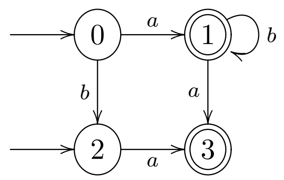
- X *is an arbitrary set of* states*,*
- $\mathcal{L}$ *is a set of* labels*,*
- $\perp \notin X$ *is a unique symbol called* initial state*,*
- $\checkmark \notin X$ *is a unique symbol called* final state*,*
- $\varepsilon \notin \mathcal{A}$ *is a unique label called* empty label*,*
- $R \subseteq (X \cup \{\perp\}) \times (\mathcal{L} \cup \{\varepsilon\}) \times (X \cup \{\checkmark\})$ *is the* transition relation*, a ternary relation linking two states and a symbol from the alphabet.*

For states $x \in X \cup \{\perp\}$ *and* $y \in X \cup \{\checkmark\}$ *and a label* $l \in \mathcal{L} \cup \{\varepsilon\}$ *such that* $(x, l, y) \in R$ *we write* $x \xrightarrow{l} y$ *and say, that the system can make a transition from the state* x *to the state* y*. Depending on the point of view this transition is either induced by the* action l *or the transition generates the* symbol l*. If there is at least one transition of the form* $(\perp, l, y) \in R$ *we say that the LTS has* (explicit) initialization *and analogously we say that the LTS has* (explicit) termination *if there is at least one transition of the form* $(x, l, \checkmark) \in R$*.*

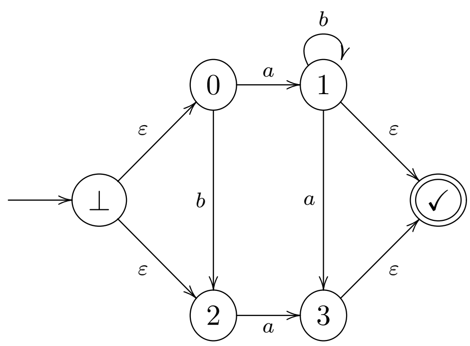
LTS with initial and final state as defined above can be considered to be a generalization of finite automata: Recall that a non-deterministic finite automaton (NFA) is a 5-tuple $(Q, \mathcal{A}, \delta, Q_0, F)$ where Q is a finite set of states, $\mathcal{A}$ is a finite set called alphabet, $\delta\colon Q \times \mathcal{A} \to \mathcal{P}(Q)$ is the transition function and $\emptyset \neq Q_0, F \subseteq X$ are the sets of initial or final states. Without loss of generalization we assume that $\perp, \checkmark \notin Q$ and $\varepsilon \notin \mathcal{A}$ and define

$$R := \{(\perp, \varepsilon, x) \mid x \in X_0\} \cup \{(x, \varepsilon, \checkmark) \mid x \in F\} \cup \{(x, a, y) \mid x \in X, a \in \mathcal{A}, y \in \delta(x, a)\}.$$

Then $(Q, \mathcal{A}, \perp, \checkmark, \varepsilon, R)$ is the corresponding LTS. As example we consider the following NFA:



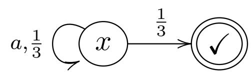
The corresponding LTS is:



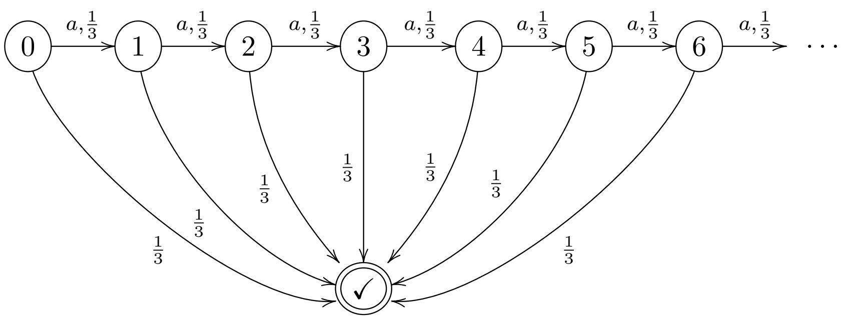
Of course we can also express every deterministic finite automaton (DFA) as LTS in a similar way.

In this thesis we will consider probabilistic versions of LTS, so-called *probabilistic transition systems*. We take a countable set $\mathcal{A}$ and define labels of the form $\mathcal{L} := \mathcal{A} \times [0, 1] \cup [0, 1]$. Thus a label is either a tuple (a, p) or simply a real number p between 0 and 1. The intuitive idea of these systems is, that whenever we are in a state $x \in X$ and there is a transition $x \overset{(a,p)}{\to} y$ or $x \overset{p}{\to} y$ to another state y, this transition occurs with probability p. As we do not restrict the state space X to be a countable set, we must necessarily make use of measure theory to cope with probabilities.

Although we do not yet have a thorough definition of probabilistic transition systems, let us take a look at two "intuitive" introductory examples. The first one is a finite system:



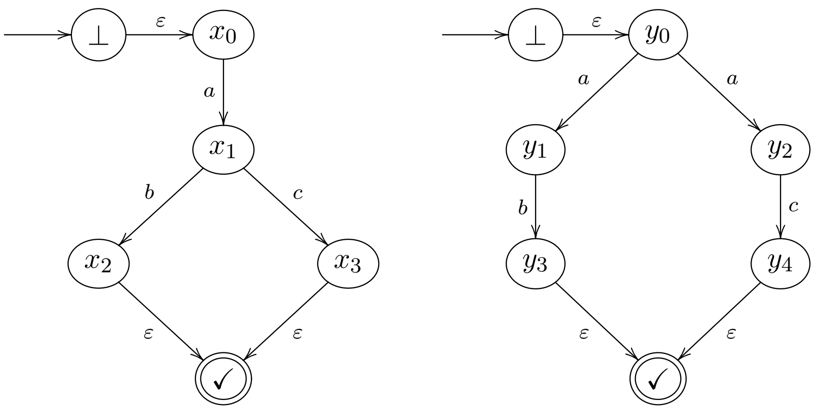
and the second one has a countable state space:



Both systems do not have explicit initialization and in fact we are going to define probabilistic transition systems without initial states. However, although the examples above have explicit termination, we will also consider systems withouth explicit termination in this thesis.

1.2 Trace Semantics

Being able to describe dynamic systems, let us now recall a commonly used method to analyze the behavior of such systems. We take a look at the following two non-deterministic finite automata (represented as LTS):



What is the behavior of these automata? One approach to specifying behavior of any LTS is given by *trace semantics*: For every state of a LTS we define its *trace*. In our example we can obtain a trace of a state x by considering all paths that lead to the final state of the automaton and collecting the labels of the transitions along the path. We get a finite or countable sequence of labels for every such path and call these sequences *words*. The trace of a state x (also called *language of* x in this case) denoted $L(x)$, is the set of all these words. For our example we have:

$$\begin{aligned}
L(x_0) &= \{ab\varepsilon, ac\varepsilon\} = L(y_0),\\
L(x_1) &= \{b\varepsilon, c\varepsilon\},\\
L(x_2) &= L(x_3) = L(y_3) = L(y_4) = \{\varepsilon\},\\
L(y_1) &= \{b\varepsilon\}\\
L(y_2) &= \{c\varepsilon\}
\end{aligned}$$

where ε is treated as a special word, the empty word, with the convention that e.g. $ab\varepsilon = ab$. Now we are able to specify the behavior of a state x simply by its trace. Note that via trace semantics we obtain an equivalence relation on the set of the states, the so-called *trace equivalence*. Two states are related (and thus have the same behavior) if and only if their traces are identical. However, note that trace equivalence is a coarse equivalence as we cannot distinguish the two states x_0 and y_0 in our example although there is a difference in the behavior of the automata: While in the first case the automaton still can make a b- or a c-transition after having made an a-transition, the automaton on the right hand side makes that choice already during the a-transition. There are other, finer notions of behavioral equivalence, e.g. *bisimulation* which can distinguish x_0 and y_0. Yet, trace semantics have proven to be worth studying as well and this is what we are going to do in this thesis.

Although trace semantics can be specified quite easily for finite automata, matters get a lot more complicated in the case of probabilistic transition systems, especially if the state space is uncountably infinite or if we consider systems without explicit termination. For our two previous examples we will see that although the first transition system has a finite state space while the second one has a (countably) infinite state space, the trace of the state x of the first transition system is identical to the trace of any of the states $n \in \mathbb{N}_0$ of the second transition system, both given by the following discrete sub-probability measure:

$$\forall u \in \mathcal{A}^* : \quad P(\{u\}) = \begin{cases} \frac{1}{3^{n+1}} & \text{if } \exists n \in \mathbb{N}_0 : u = a^n \\ 0 & \text{else} \end{cases}$$

Even though it is possible to define the notion of a trace of a probabilistic transition system directly, we will employ a different approach using category theory. In order to understand the subsequent two sections of this introduction a prior knowledge of category theory is advisable. Readers who are unfamiliar with category theory may simply read on in section 1.5. We will provide all the necessary definitions and propositions in chapters 2 and 3 to understand the main results of this thesis.

1.3 Transition Systems as Coalgebra

As for example Bonchi et al. [2011] point out, coalgebras are an appropriate method to model different types of (labelled) transition systems and analyze their behavior. However, note that in a coalgebraic representation of a LTS we usually "neglect" explicit initialization. As a matter of fact this is irrelevant for our considerations on the one hand because our behavioral analysis is restricted to the behavior of singular states and on the other hand because our probabilistic transition systems do not have explicit initialization by definition. Of course, for systems with explicit initialization only the behavior of reachable states is of interest. Let us now present how to model an LTS as coalgebra: Given an endofunctor $F\colon \mathbf{C} \to \mathbf{C}$ on a category $\mathbf{C}$, a coalgebra is a pair (X, α) where X, an object of $\mathbf{C}$, represents the state space of the system and $\alpha\colon X \to FX$, an arrow of $\mathbf{C}$, specifies the transitions of the system. The endofunctor F determines the branching behavior of the system. Depending on the choice of the category and the endofunctor F, a final coalgebra can exist. This final coalgebra is a coalgebra which is unique up to isomorphism and is defined via the universal property that for every other coalgebra there is a unique coalgebra homomorphism κ mapping it into the final coalgebra. In a situation, where a final coalgebra exists and coalgebras are used to model a transition system, the unique homorphism yields some kind of behavior of the system. As an example we can model deterministic finite automata in the category $\mathbf{Set}$ for the endofunctor $\mathbf{2} \times \mathrm{Id}^{\mathcal{A}}\colon \mathbf{Set} \to \mathbf{Set}$ where $\mathcal{A}$ is an alphabet. The final coalgebra is the set of all languages over the alphabet, $\mathcal{P}(\mathcal{A}^*)$, and the unique morphism is a function mapping each state to the language that it accepts (cf. Bonchi et al. [2011]). In this case, the behavioral equivalence induced by the unique map to the final coalgebra is indeed trace equivalence. However, if we model non-deterministic finite automata as coalgebras for the endofunctor $\mathbf{2} \times \mathcal{P}(\mathrm{Id})^{\mathcal{A}}\colon \mathbf{Set} \to \mathbf{Set}$, we obtain bisimularity instead of trace equivalence. Thus the choice of the category and the endofunctor determines what kind of behavioral equivalence the final coalgebra (if it exists) describes. As we are interested in trace equivalence for probabilistic transition systems, we must find an appropriate category and a suitable endofunctor. Our approach to find such a category is based on a paper titled "Generic Trace Semantics via Coinduction" by Hasuo et al. [2007].

1.4 Main Aim of this Thesis

The main aim of this thesis is to show that probabilistic transition systems can be modelled as a coalgebra $(X_T, \alpha^\flat)$ in the Kleisli category of a monad T for the lifting $\overline{F}$ of an endofunctor $F\colon \mathbf{Meas} \to \mathbf{Meas}$ on the category of measurable spaces. For an arbitrary alphabet $\mathcal{A}$ we take:

- $F = \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}} + \mathbf{1}$ and $T = \mathbb{S}$ (the sub-probability monad) for probabilistic transition systems with termination and
- $F = \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}}$ and $T = \mathbb{P}$ (the probability monad) for probabilistic transtion systems without termination.

Moreover, we prove that a final coalgebra exists in both cases if $\mathcal{A}$ is a finite alphabet and use the unique arrow $\mathbf{tr}^\flat$ mapping $(X_T, \alpha^\flat)$ to the final coalgebra to define the trace $\mathbf{tr}(x)$ of a state x:

- For probabilistic transition systems with termination, $\mathbf{tr}(x)$ is a sub-probability measure on the set of all finite words $\mathcal{A}^*$.
- For probabilistic transition systems without termination, $\mathbf{tr}(x)$ is a probability measure on the set of all infinite words $\mathcal{A}^\omega$.

The obtained results coincide with the results for discrete probabilistic transition systems given by Hasuo et al. [2007].

1.5 Structure of this Thesis

This thesis contains a total of five chapters. Below is a short overview of the contents of the chapters.

Chapter 2 In this chapter we present the mathematical basics needed for the subsequent chapters. Most of the presented results are well-known and taken from standard literature of the respective area.

Chapter 3 The measurable spaces form a category together with measurable functions. In preparation for the main part of this thesis we take a closer look at this category and prove that it has finite products and coproducts. For the latter we also develop the necessary measure-theoretic background. Finally, we present the (sub-)probability monad.

Chapter 4 This is the main part of this thesis. We first define σ-algebras on the sets of words allowing us to consider them as measurable spaces. Then we provide distributive laws for the endofunctors $\mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}}$ and $\mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}} + \mathbf{1}$ and the (sub-)probability monad, model probabilistic transition systems as coalgebra in the Kleisli category of the (sub-)probability monads and prove that in certain cases a final coalgebra exists which yields a suitable definition of a trace for probabilistic transition systems. Finally we apply the results to a few examples.

Chapter 5 In the last chapter we will summarize the results we have obtained and give a reference to further literature related to our topic. Additionally, we will present open questions the thesis at hand has not answered.

2 Mathematical Basics

2.1 Set Theory

We assume basic knowledge about sets and common set-theoretic results. However, we will recall some of the definitions and introduce the notation used within this document.

Definition 2.1 (Notation and Special Sets)**.** *We denote by* **1** *the* singleton set*, a set with only one element. Whenever needed we will denote its element by* $\checkmark$*. Furthermore we will use the following notation:*

- *For an arbitrary set X we denote the set of all subsets of X, the* powerset*, by $\mathcal{P}(X)$.*
- *For arbitrary sets A, B the symbol $\subset$ shall denote strict set inclusion, i.e. $A \subset B$ implies that $A \neq B$. For non-strict inclusion we will use the symbol $\subseteq$. The same applies for $\supset$ and $\supseteq$ and for the usual order on the (extended) reals denoted by $<, \leq, >, \geq$.*
- *For arbitrary sets A, B we define their* set theoretic difference *to be $A \setminus B = \{x \in A \mid x \notin B\}$.*
- $\mathbb{N}$ *is the set of natural numbers starting with* 1*,* $\mathbb{N}_0$ *is the set of natural numbers including* 0*.*
- $\mathbb{R}$ *is the set of real numbers,* $\mathbb{R}_+$ *and* $\mathbb{R}_-$ *are the sets of non-negative or non-positive real numbers.*
- $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ *is the set of extended real numbers, equipped with the usual extensions[1] of arithmetic operations on* $\mathbb{R}$*. We recall only the most important convention, namely:* $0 \cdot \pm\infty = \pm\infty \cdot 0 := 0$*.*

[1] For further reference cf e.g. to [Elstrodt, 2007, pp. 104f.].

- *For real intervals we use the following notation*

$$
\begin{aligned}
]a,b[&:= \{x \in \mathbb{R} \mid a < x < b\} \quad \forall a,b \in \overline{\mathbb{R}} \\
[a,b] &:= \{x \in \mathbb{R} \mid a \leq x \leq b\} \quad \forall a,b \in \mathbb{R} \\
]a,b] &:= \{x \in \mathbb{R} \mid a < x \leq b\} \quad \forall a \in \overline{\mathbb{R}}, \forall b \in \mathbb{R} \\
[a,b[&:= \{x \in \mathbb{R} \mid a \leq x < b\} \quad \forall a \in \mathbb{R}, \forall b \in \overline{\mathbb{R}}
\end{aligned}
$$

  and accordingly we define intervals of extended reals.

- *For a total function* $f\colon X \to Y$ *we denote for* $A \in \mathcal{P}(X)$ *and* $B \in \mathcal{P}(Y)$ *the* image *of* A *by* $f(A) := \{f(x) \mid x \in A\}$ *and the* pre-image *of* B *by* $f^{-1}(B) := \{x \in X \mid f(x) \in B\}$. *For* $\mathcal{A} \subseteq \mathcal{P}(X)$ *and* $\mathcal{B} \subseteq \mathcal{P}(Y)$ *we define accordingly* $f(\mathcal{A}) := \{f(A) \mid A \in \mathcal{A}\}$ *and* $f^{-1}(\mathcal{B}) := \{f^{-1}(B) \mid B \in \mathcal{B}\}$.

Definition 2.2 (Characteristic Function)**.** *Let* X *be an arbitrary set,* $A \subseteq X$. *The* characteristic function of A *is the function*

$$\chi_A\colon X \to \{0,1\}, \quad x \mapsto \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

Definition 2.3 (Mutually Disjoint Sets, Disjoint Union, Cartesian Product)**.** *Let* I *be an arbitrary set,* $(X_i)_{i\in I}$ *be a family of arbitrary sets.*

- *We call* $(X_i)_{i\in I}$ *a* disjoint family of sets *iff the* X_i *are* mutually disjoint, *i.e.* $X_i \cap X_j = \emptyset$ *for* $i \neq j$.

- *The* disjoint union *of the* $(X_i)_{i\in I}$ *is the set:*

$$\bigsqcup_{i\in I} X_i := \bigcup_{i\in I} \{(x,i) \mid x \in X_i\}$$

  For every $j \in I$ *there is an injective function* $\iota_j\colon X_j \to \bigsqcup_{i\in I} X_i$ *where* $x \mapsto (x,j)$ *called the* canonical injection. *Note that whenever the family* $(X_i)_{i\in I}$ *is disjoint, the disjoint union coincides with the usual set-theoretic union in an obvious way. For finite* I*, say* $I = \{1,...,n\}$ *we will also write*

$$X_1 + X_2 + ... + X_n = \bigsqcup_{i=1}^{n} X_i$$

  for the (n-th) disjoint union.

- *The* cartesian product *of the X_i is the set*

$$\prod_{i\in I} X_i := \{(x_i)_{i\in I} \mid \forall i \in I \colon x_i \in X_i\}$$

For every $j \in I$ there is a surjective function $\pi_j \colon \prod_{i\in I} X_i \to X_j$ where $(x_i)_{i\in I} \mapsto x_j$ called the canonical projection. *For finite I, say $I = \{1, \ldots, n\}$ we will also write*

$$X_1 \times X_2 \times \ldots \times X_n = \prod_{i=1}^{n} X_i$$

for the (n-th) cartesian product.

2.1.1 Alphabets and Words

Any countable set can be considered to be an *alphabet*. Whenever that is the case, we need some additional definitions and operations which we will introduce here.

Definition 2.4 (Alphabets and Words)**.** *Let $\mathcal{A}$ be a countable set and ε be a symbol such that $\varepsilon \notin \mathcal{A}$. We call $\mathcal{A}$ an* alphabet *and define*

$$\mathcal{A}^0 := \{\varepsilon\}.$$

For every natural number $n \in \mathbb{N}$, we call the elements of the n-th cartesian product words of length n *and denote its elements $u \in \mathcal{A}^n$ as $u = a_1a_2a_3 \ldots a_n$ where $a_1, \ldots, a_n \in \mathcal{A}^n$. The symbol ε is called the* empty word *which has length* 0 *by definition. Furthermore we define for every $n \in \mathbb{N}_0$ the set:*

$$\mathcal{A}^{\leq n} := \bigcup_{i=0}^{n} \mathcal{A}^i$$

and call it the set of words of length of at most n. *The* set of all finite words (over $\mathcal{A}$) *is:*

$$\mathcal{A}^* := \bigcup_{n\in\mathbb{N}_0} \mathcal{A}^n$$

and the set of all infinite words (over $\mathcal{A}$) *is the set:*

$$\mathcal{A}^\omega := \prod_{n\in\mathbb{N}} \mathcal{A} = \{(a_n)_{n\in\mathbb{N}} \mid \forall n \in \mathbb{N} : a_n \in \mathcal{A}\}.$$

Combining the two latter sets we obtain the set of all (finite and infinite) words (over $\mathcal{A}$)*:*

$$\mathcal{A}^\infty := \mathcal{A}^* \cup \mathcal{A}^\omega.$$

Two words can be concatenated by the concatenation function*:*

$$\mathbf{conc}\colon \mathcal{A}^* \times \mathcal{A}^\infty \to \mathcal{A}^\infty$$

where for $x = x_1 \ldots x_m \in \mathcal{A}^m$ *and* $y = y_1 \ldots y_n \in \mathcal{A}^n$ *we define*

$$\mathbf{conc}(x, y) := x_1 \ldots x_m y_1 \ldots y_n \in \mathcal{A}^{m+n}$$

and for x *as above and* $y = (y_n)_{n\in\mathbb{N}} \in \mathcal{A}^\omega$ *we define the sequence* $(z_i)_{i\in\mathbb{N}} := \mathbf{conc}(x, y)$ *where*

$$z_i := \begin{cases} x_i, & 1 \le i \le m \\ y_{i-m}, & i > m \end{cases}$$

The length of a word $u \in \mathcal{A}^\infty$ *is denoted by* $|u|$ *and we define* $|(a_n)_{n\in\mathbb{N}}| := \infty$*.*

Definition 2.5 (Finite Prefixes)**.** *Let* $\mathcal{A}$ *be an alphabet. We define the* prefix-relation $\sqsubseteq \subseteq \mathcal{A}^* \times \mathcal{A}^\infty$ *where two words* $x \in \mathcal{A}^*, y \in \mathcal{A}^\infty$ *are related (notation:* $x \sqsubseteq y$*) iff one of the following conditions is fulfilled:*

(a) $x = \varepsilon$ *and* y *is an arbitrary word.*

(b) There are $m, n \in \mathbb{N}$ *such that* $m \le n$*,* $x = x_1 ... x_m \in \mathcal{A}^m$*,* $y = y_1 ... y_n \in \mathcal{A}^n$ *and* $x_i = y_i$ *holds for all* $i \in \{1, ..., m\}$*.*

(c) There is an $m \in \mathbb{N}$ *such that* $x = x_1 ... x_m \in \mathcal{A}^m$*,* $y = (y_n)_{n\in\mathbb{N}} \in \mathcal{A}^\omega$ *and* $x_i = y_i$ *holds for all* $i \in \{1, ..., m\}$*.*

Whenever $x \sqsubseteq y$ *holds, we call* x *a* (finite) prefix *of* y*. The set of all finite prefixes of a word* $u \in \mathcal{A}^\infty$ *is the (countable) set:*

$$\mathbf{pref}(u) := \{v \in \mathcal{A}^* \mid v \sqsubseteq x\}.$$

2.2 Measure and Integration Theory

The basic idea of measure theory is to assign a *measure* to subsets of a given set X which is nothing else but a non-negative real number (or possibly $+\infty$). Depending on the context, this *measure* might for instance be a length, an area, or a volume. The measures we will be interested in, are *(sub-)probability measures.* In order to define any measure, the first step is to identify what subsets of a set should be measurable. The results in this section are mainly based on or taken from Elstrodt [2007] and Ash [1972].

2.2.1 Sigma-Algebras and Measurable Spaces

Definition 2.6 (Sigma-Algebra, Measurable Space)**.** *Let X be an arbitrary set. We call $\Sigma \subseteq \mathcal{P}(X)$ a* sigma-algebra *(also* σ-algebra*) on X iff the following axioms are fulfilled:*

(1) X itself is an element of Σ, i.e. $X \in \Sigma$.

(2) Σ is closed under complement: $A \in \Sigma \implies X \setminus A \in \Sigma$.

(3) Σ is closed under countable unions: $A_i \in \Sigma\ \forall i \in \mathbb{N} \implies \bigcup_{i\in\mathbb{N}} A_i \in \Sigma$.

The pair (X, Σ) is called a measurable space. *We sometimes denote a measurable space merely by its base set X. When needed, the corresponding σ-algebra is denoted by Σ_X.*

Remark 2.7. *(Additional Properties of σ-Algebras) Let Σ be a σ-algebra on an arbitrary set X. Then the following holds:*

(a) $\emptyset \subseteq \Sigma$.

(b) Σ is closed under countable intersection: $A_i \in \Sigma\ \forall i \in \mathbb{N} \implies \bigcap_{n\in\mathbb{N}} A_i \in \Sigma$.

(c) Σ is closed under finite union.

(d) Σ is closed under finite intersection.

Proof.

(a) $\emptyset = (X \setminus X) \in \Sigma$.

(b) Let $A_i \in \Sigma\ \forall i \in \mathbb{N}$, then:

$$\bigcap_{i\in\mathbb{N}} A_i = \left(X \setminus \left(\bigcup_{i\in\mathbb{N}}(X \setminus A_i)\right)\right) \in \Sigma.$$

(c) Follows immediately from (a) and axiom (3) by taking $S_n = \emptyset$ for all but finitely many $n \in \mathbb{N}$.

(d) Follows immediately from (a) and (b) by taking $S_n = \emptyset$ for all but finitely many $n \in \mathbb{N}$.

□

Let us now consider a few examples of σ-algebras:

Example 2.8. *Let X be an arbitrary set.*

- *The* trivial σ-algebra *on X is the set $\{\emptyset, X\}$. With respect to set inclusion it is the smallest σ-algebra on X.*
- *The powerset $\mathcal{P}(X)$ is a σ-algebra on X. With respect to set inclusion it is the biggest σ-algebra on X.*
- *Let $A \subseteq X$ be an arbitrary subset of X. Then $\{\emptyset, A, X \setminus A, X\}$ is a σ-algebra on X. With respect to set inclusion it is the smallest σ-algebra on X containing A.*

Note that for $X = \emptyset$ and for $X = \mathbf{1}$ the trivial σ-algebra coincides with the powerset and hence in these two cases the choice of the σ-algebra is unique.

A useful property of σ-algebras is that they are closed under intersection in the following manner:

Proposition 2.9. *Let for an arbitrary set I, $(\Sigma_i)_{i \in I}$ be a family of σ-algebras on a common set X. Then their intersection $\bigcap_{i\in I} \Sigma_i$ is a sigma-algebra on X.*

Proof. Trivial checking of the axioms. □

As a result of this property, we can define the notion of a *generated sigma-algebra*. We want certain sets to be measurable, so we take those sets and consider the smallest σ-algebra that contains them.

Definition 2.10 (Generator, Generated Sigma-Algebra)**.** *Let X be a set, $\mathcal{G} \subseteq \mathcal{P}(X)$. We define the* sigma-algebra generated by $\mathcal{G}$ *to be:*

$$\sigma(\mathcal{G}) := \bigcap_{\Sigma \in \Phi(\mathcal{G})} \Sigma$$

*where $\Phi(\mathcal{G}) := \{\Sigma \subseteq \mathcal{P}(X) \mid \mathcal{G} \subseteq \Sigma$ and Σ
mboxisa σ-algebra$\}$. It is the smallest σ-algebra on X (with respect to set inclusion) containing $\mathcal{G}$. The set $\mathcal{G}$ is called the* generator *of $\sigma(\mathcal{G})$. If for all $A, B \in \mathcal{G}$ the intersection $A \cap B$ is itself an element of $\mathcal{G}$, we say that $\mathcal{G}$ is* closed under finite intersection.

With the above definition we can regard σ to be an operator $\sigma\colon \mathcal{P}(X) \to \{\Sigma \subseteq \mathcal{P}(X) \mid \Sigma$ is a σ-algebra on $X\}$. The following properties of this operator are useful for proofs:

Remark 2.11. *(Properties of the σ-operator) Let X be an arbitrary set. The σ-operator is*

- monotone*, i.e. the following holds for every $\mathcal{F}, \mathcal{G} \subseteq \mathcal{P}(X)$:*

$$\mathcal{F} \subseteq \mathcal{G} \quad \Longrightarrow \quad \sigma(\mathcal{F}) \subseteq \sigma(\mathcal{G})$$

- idempotent*, i.e. the following holds for every $\mathcal{G} \subseteq \mathcal{P}(X)$:*

$$\sigma\left(\sigma(\mathcal{G})\right) = \sigma(\mathcal{G})$$

Proof. For monotonicity, let $\mathcal{F} \subseteq \mathcal{G} \subseteq \mathcal{P}(X)$, then $\Phi(\mathcal{F}) \supseteq \Phi(\mathcal{G})$ and hence:

$$\sigma(\mathcal{F}) = \bigcap_{\Sigma \in \Phi(\mathcal{F})} \Sigma \subseteq \bigcap_{\Sigma \in \Phi(\mathcal{G})} \Sigma = \sigma(\mathcal{G}).$$

Idempotence is obvious because $\sigma(\mathcal{G})$ is the smallest σ-algebra that contains itself. □

Let us now take a look at two examples of a generated σ-algebra.

Example 2.12. *Let X be a countable set. We consider the sigma-algebra generated by all singletons and the empty set[2]: $\Sigma := \sigma(\{\emptyset\} \cup \{\{x\} \mid x \in X\})$ and claim that it is identical to the powerset of X. By definition, we have $\Sigma \subseteq \mathcal{P}(X)$. Let now $S \in \mathcal{P}(X)$, then $S = \cup_{x \in S}\{x\}$ is a countable union of singletons. Hence, we have $S \in \Sigma$ and thus $\Sigma = \mathcal{P}(X)$.*

The previous example does not show the necessity of taking a generated σ-algebra. In fact, for countable sets X it is often convenient to take $\mathcal{P}(X)$ as σ-algebra. The most prominent example for taking a smaller σ-algebra than the powerset, is the set of real numbers. The canonical σ-algebra on the reals is the Borel-σ-algebra which is defined below. Explaining why one usually considers this σ-algebra instead of another one is outside the scope of this thesis. For further reference, every standard textbook on measure theory can be consulted. A good summary in German is given in [Elstrodt, 2007, pp. 1-6].

Example 2.13 (The Borel-σ-Algebra on $\mathbb{R}$ and $\overline{\mathbb{R}}$)**.** *The set of real numbers, $\mathbb{R}$, is usually endowed with the* Borel-σ-algebra *$\mathcal{B}$, which can be defined in the following way:*

$$\mathcal{B} := \sigma\left(\{]-\infty, t] \mid t \in \mathbb{R}\}\right).$$

[2] The empty set could be omitted as it is trivially contained in every sigma-algebra. However, later we will see that it useful to include it in the generator.

Analogously, the set

$$\overline{\mathcal{B}} := \sigma\left(\{[-\infty, t] \mid t \in \mathbb{R}\}\right)$$

is the Borel-σ-algebra for $\overline{\mathbb{R}}$.

2.2.2 Semi-Rings of Sets

We know now that the sets we want to consider as "measurable" are defined by a σ-algebra. In order to define how we can "measure" them, it is helpful to define weaker requirements on a subset $\mathcal{S} \subseteq \mathcal{P}(X)$ than those it has to fulfill to be a σ-algebra. In many cases these weaker requirements on $\mathcal{S}$ are sufficient to define a measure on $\sigma(\mathcal{S})$ uniquely.

Definition 2.14 (Semi-Ring of Sets)**.** *Let X be a set. A set $\mathcal{S} \subseteq \mathcal{P}(X)$ is called a* semi-ring (of sets) *iff it satisfies the following axioms:*

(1) The empty set is contained in $\mathcal{S}$: $\emptyset \in \mathcal{S}$.

(2) $\mathcal{S}$ is closed under finite intersection: $A, B \in \mathcal{S} \Rightarrow A \cap B \in \mathcal{S}$.

(3) For every $A, B \in \mathcal{S}$ there are finitely many pairwise disjoint subsets $C_1, ..., C_k \in \mathcal{S}$ such that $A \setminus B = \cup_{i=1}^{k} C_i$.

A simple and yet useful example for a semi-ring of sets is the set of singletons:

Example 2.15. *Let X be a countable set. The set $\mathcal{S} := \{\emptyset\} \cup \{\{x\} \mid x \in X\}$ is a semi-ring of sets.*

Proof. We obviously have $\emptyset \in \mathcal{S}$. Furthermore, we have for every $x, y \in X$: $\{x\} \cap \{y\} = \{x\}$ iff $x = y$ and $\{x\} \cap \{y\} = \emptyset$ iff $x \neq y$. Analogously, we observe that $\{x\} \setminus \{y\} = \{x\}$ iff $x \neq y$ and $\{x\} \setminus \{y\} = \emptyset$ iff $x = y$. $\square$

Remark 2.16. *Every σ-algebra Σ on an arbitrary set X is a semi-ring of sets.*

Proof. We have $\emptyset \in \Sigma$ by definition and we know that σ-algebras are closed under finite intersection. For $A, B \in \Sigma$ we rewrite $A \setminus B = A \cap (X \setminus B) \in \Sigma$. $\square$

Proposition 2.17. *Let X, Y be arbitrary sets and $\mathcal{S}_X \subseteq \mathcal{P}(Y)$ and $\mathcal{S}_Y \subseteq \mathcal{P}(Y)$ be semi-ring of sets. Then the set:*

$$\mathcal{S}_X \oplus \mathcal{S}_Y := \{S_X + S_Y \mid S_X \in \mathcal{S}_X, S_Y \in \mathcal{S}_Y\} \subseteq \mathcal{P}(X + Y)$$

is a semi-ring of sets.

Proof. We check the axioms:

(1) $\emptyset + \emptyset = \emptyset \in \mathcal{S}_1 \oplus \mathcal{S}_2$.

(2) Let $A_1 + A_2, B_1 + B_2 \in \mathcal{S}_1 \oplus \mathcal{S}_2$, then:

$$(A_1 + A_2) \cap (B_1 + B_2) = \underbrace{(A_1 \cap B_1)}_{\in \mathcal{S}_1} + \underbrace{(A_2 \cap B_2)}_{\in \mathcal{S}_2} \in \mathcal{S}_1 \oplus \mathcal{S}_2$$

(3) Let $A_1 + A_2, B_1 + B_2 \in \mathcal{S}_1 \oplus \mathcal{S}_2$, then:

$$(A_1 + A_2) \setminus (B_1 + B_2) = \underbrace{(A_1 \setminus B_1)}_{\in \mathcal{S}_1} + \underbrace{(A_2 \setminus B_2)}_{\in \mathcal{S}_2} \in \mathcal{S}_1 \oplus \mathcal{S}_2$$

□

Proposition 2.18. *Let X, Y be arbitrary sets and $\mathcal{S}_X \subseteq \mathcal{P}(Y)$ and $\mathcal{S}_Y \subseteq \mathcal{P}(Y)$ be semi-ring of sets. Then the set:*

$$\mathcal{S}_X * \mathcal{S}_Y := \{S_X \times S_Y \mid S_X \in \mathcal{S}_X, S_Y \in \mathcal{S}_Y\} \subseteq \mathcal{P}(X \times Y)$$

is a semi-ring of sets.

Proof. Cf. [Elstrodt, 2007, Lemma I.5.4, p. 21] □

Corollary 2.19. *Let X, Y, Z be arbitrary sets and $\mathcal{S}_X \subseteq \mathcal{P}(Y)$, $\mathcal{S}_Y \subseteq \mathcal{P}(Y)$ and $\mathcal{S}_Z \subseteq \mathcal{P}(Z)$ be semi-ring of sets. Then the set:*

$$\mathcal{S}_X * \mathcal{S}_Y \oplus \mathcal{S}_Z := \{(S_X \times S_Y) + S_Z \mid S_X \in \mathcal{S}_X, S_Y \in \mathcal{S}_Y, S_Z \in \mathcal{S}_Z\} \subseteq \mathcal{P}((X \times Y) + Z)$$

is a semi-ring of sets.

Proof. By Proposition 2.18 we know that $\mathcal{S}_X * \mathcal{S}_Y$ is a semi-ring of sets. We then apply Proposition 2.17 to the semi-rings of sets $\mathcal{S}_X * \mathcal{S}_Y$ and $\mathcal{S}_Z$ and obtain that also $(\mathcal{S}_X * \mathcal{S}_Y) \oplus \mathcal{S}_Z = \mathcal{S}_X * \mathcal{S}_Y \oplus \mathcal{S}_Z$ is a semi-ring of sets. □

2.2.3 Pre-Measures and Measures

Definition 2.20 (Pre-Measure, Measure, Measure Space)**.** *Let X be a set and $\mathcal{S} \subseteq \mathcal{P}(X)$ be a semi-ring of sets. A map $\mu\colon \mathcal{S} \to \overline{\mathbb{R}}$ is called a* pre-measure *iff it satisfies the following axioms:*

(1) $\mu(\emptyset) = 0$.

(2) μ *is non-negative.*

(3) μ *is* sigma-additive, *i.e. for every sequence* $(S_n)_{n\in\mathbb{N}}$ *of mutually disjoint sets from* $\mathcal{S}$ *where* $\cup_{n\in\mathbb{N}} S_j \in \mathcal{S}$ *the following holds:*

$$\mu\left(\bigcup_{n\in\mathbb{N}} S_n\right) = \sum_{n=1}^{\infty} \mu(S_n).$$

We call μ *a* sigma-finite *pre-measure iff a sequence* $(S_n)_{n\in\mathbb{N}}$ *of measurable sets* $S_n \in \mathcal{S}$ *exists such that* $\mu(S_n) < \infty$ *for all* $n \in \mathbb{N}$ *and* $\cup_{n\in\mathbb{N}} S_n = X$.
If $\mathcal{S}$ *is a* σ*-algebra on* X *we call* μ *a* measure *and the triple* $(X, \mathcal{S}, \mu)$ *a* measure space. *In this case the measure is called* finite *iff* $\mu(X) < \infty$. *Of course, finiteness of a measure implies* σ*-finiteness.*

Having defined what a pre-measure is, we will now take a look at some of the basic properties such pre-measures fulfill.

Proposition and Definition 2.21 (Properties of Pre-Measures). *Let* X *be a set,* $\mathcal{S} \subseteq \mathcal{P}(X)$ *a semi-ring of sets and* $\mu\colon \mathcal{S} \to \overline{\mathbb{R}}_+$ *be a pre-measure. Then the following holds for all* $A, B \in \mathcal{S}$ *and all sequences* $(S_n)_{n\in\mathbb{N}}$*:*

(a) If $\mu(B) < \infty$ *and* $A \subseteq B$ *then* $\mu(B \setminus A) = \mu(B) - \mu(A)$.

(b) $\mu(A) + \mu(B) = \mu(A \cup B) - \mu(A \cap B)$.

(c) μ *is* subadditive*:*

$$\mu\left(\bigcup_{n=1}^{N} S_n\right) \leq \sum_{n=1}^{N} \mu(S_n)$$

(d) μ *is* σ-subadditive*:*

$$\mu\left(\bigcup_{n\in\mathbb{N}} S_n\right) \leq \sum_{n=1}^{\infty} \mu(S_n).$$

(e) μ *is* continuous from below, *i.e. if* $S_n \subseteq S_{n+1}$ *holds for all* $n \in \mathbb{N}$ *then:*

$$\mu\left(\lim_{n\to\infty} S_n\right) = \mu\left(\bigcup_{n=1}^{\infty} S_n\right) = \lim_{n\to\infty} \mu(S_n).$$

Proof. Cf. [Elstrodt, 2007, Satz II.1.7, pp. 31f.; Satz II.1.10, pp. 32f.]. □

Although the above properties are interesting and useful (especially for calculations), the most useful properties for our interest are of a theoretical nature and given by the following two results which can be summarized in one sentence: A σ-finite measure is already defined by its values on a semi-ring of sets which generates the σ-algebra.

Proposition 2.22. *(Extension Theorem) Let X be an arbitrary set, $\mathcal{S} \subseteq \mathcal{P}(X)$ be a semi-ring of sets. Every σ-finite pre-measure $\mu\colon \mathcal{S} \to \overline{\mathbb{R}}_+$ can uniquely be extended to a measure $\mu\colon \sigma(\mathcal{S}) \to \overline{\mathbb{R}}_+$ on $\sigma(\mathcal{S})$, the σ-algebra generated by $\mathcal{S}$.*

Proof. Cf. [Elstrodt, 2007, Korollar II.5.7, p. 61]. □

Corollary 2.23. *Let X be an arbitrary set, $\mathcal{S} \subseteq \mathcal{P}(X)$ be a semi-ring of sets and $\mu, \nu\colon \sigma(\mathcal{S}) \to \overline{\mathbb{R}}_+$ be σ-finite measures. Then the following holds:*

$$\mu|_{\mathcal{S}} = \nu|_{\mathcal{S}} \quad \Leftrightarrow \quad \mu = \nu$$

Proof. We apply Proposition to the σ-finite pre-measures $\mu|_{\mathcal{S}}$ and $\nu|_{\mathcal{S}}$. For $\mu|_{\mathcal{S}} = \nu|_{\mathcal{S}}$ the uniqueness of the extension yields $\mu = \nu$. The other implication is trivially fulfilled. □

Definition 2.24 (Null Set, Almost Everywhere)**.** *Let (X, Σ, μ) be a measure space. A measurable set $N \in \Sigma$ is called a* null set *iff $\mu(N) = 0$. If a condition defined for all elements $x \in X$ holds outside a null set, it is said to hold* almost everywhere *(denoted "a.e. (μ)" or simply "a.e." if the measure μ is obvious).*

We will be interested in a very prominent class of measures, namely (sub-)probability measures. There is a whole branch of mathematics, called *probability theory*, devoted to the study of these measures and the associated real world problems: random phenomena. However, we will be comfortable with some basic definitions and an "intuitive" understanding of randomness as we will focus our attention on theoretical results covered by these few definitions and the already presented measure-theoretic fundamentals.

Definition 2.25 ((Sub-)Probability Measure, (Sub-)Probability Space)**.** *Let (X, Σ, μ) be a measure space. We call μ a* probability measure *iff $\mu(X) = 1$ and a* sub-probability measure *iff $\mu(X) \leq 1$. In that case (X, Σ, μ) is called* probability space *or* sub-probability space *respectively. We define the sets:*

$$\mathbb{S}(X) := \{P\colon \Sigma \to [0,1] \mid P \textit{ is a sub-probability measure}\} \text{ and}$$
$$\mathbb{P}(X) := \{P\colon \Sigma \to [0,1] \mid P \textit{ is a probability measure}\}.$$

A trivial example of a probability measure is the so-called Dirac measure which we will use a lot in the main part of this thesis:

Definition 2.26 (Dirac Measure)**.** *Let* (X, Σ) *be a measurable space. The* Dirac measure *for a given* $x \in X$ *is the probability-measure:*

$$\delta_x^X : \Sigma \to [0, 1], \quad \delta_x^X(S) = \begin{cases} 1, & x \in S \\ 0, & \textit{else} \end{cases}$$

which assigns the probability 1 *to all measurable sets containing* x *and* 0 *to all other sets.*

2.2.4 Measurable Functions

As a preparation for defining the Lebesgue integral, we must specify a property of functions $f \colon X \to Y$ where (X, Σ_X), (Y, Σ_Y) are measurable spaces:

Definition 2.27 (Measurable Function)**.** *Let* $(X, \Sigma_X), (Y, \Sigma_Y)$ *be measurable spaces. A function* $f \colon X \to Y$ *is called* Σ_X-Σ_Y-measurable *or simply* measurable *iff* $f^{-1}(S) \in \Sigma_X$ *for all* $S \in \Sigma_Y$.

The most simple examples of measurable functions are constant functions $f \colon X \to Y, x \mapsto y_0$ and identity functions $f \colon X \to X, x \mapsto x$. A useful property of measurability of functions is that composition of functions maintains measurability.

Proposition 2.28. *Let* $(X, \Sigma_X), (Y, \Sigma_Y), (Z, \Sigma_Z)$ *be measurable spaces,* $f \colon X \to Y$ *be a* Σ_X*-*Σ_Y*-measurable function and* $g \colon Y \to Z$ *be a* Σ_Y*-*Σ_Z*-measurable function. The composition* $g \circ f \colon X \to Z$ *is* Σ_X*-*Σ_Z*-measurable.*

Proof. Let $S_Z \in \Sigma_Z$. Measurability of g yields $g^{-1}(S_Z) = S_Y \in \Sigma_Y$ and using the measurability of f we obtain:

$$(g \circ f)^{-1}(S_Z) = f^{-1}\left(g^{-1}(S_Z)\right) = f^{-1}(S_Y) \in \Sigma_X$$

which is the desired measurability of $g \circ f$. □

Checking the measurability of a function $f \colon X \to Y$ can be difficult in some instances, especially if Σ_Y is uncountably infinite. However, if Σ_Y is generated by a set $\mathcal{G} \subseteq \mathcal{P}(Y)$ it suffices to show measurability on the generator as the following proposition tells us:

Proposition 2.29. *Let* $(X, \Sigma_X), (Y, \Sigma_Y)$ *be measurable spaces such that* $\Sigma_Y = \sigma(\mathcal{G})$ *for* $\mathcal{G} \subseteq \mathcal{P}(Y)$*. A function* $f \colon X \to Y$ *is* Σ_X*-*Σ_Y*-measurable iff* $f^{-1}(S) \in \Sigma_X$ *for all* $S \in \mathcal{G}$*.*

Proof. Cf. [Elstrodt, 2007, Satz 1.3, p. 86]. □

A prominent example of an uncountable set is the set of real numbers, $\mathbb{R}$. Since the canonical σ-algebra, the Borel-σ-algebra $\mathcal{B}$, is generated by the set of real intervals of the form $]-\infty, t]$ for arbitrary $t \in \mathbb{R}$, the Σ_X-$\mathcal{B}$-measurability (called Borel-measurability) of functions $f\colon X \to \mathbb{R}$ can be checked in the following way:

Example 2.30 (Borel Measurability)**.** *Let (X, Σ) be a measurable space. A function $f\colon X \to \mathbb{R}$ is called* Borel-measurable *iff it is Σ-$\mathcal{B}$-measurable. This is the case iff the following holds:*

$$\{x \mid f(x) \leq t\} = f^{-1}(]-\infty, t]) \in \Sigma$$

for all $t \in \mathbb{R}$. Analogously a function $g\colon X \to \overline{\mathbb{R}}$ is called Borel-measurable *iff it is Σ-$\overline{\mathcal{B}}$-measurable. This is the case iff the following holds:*

$$\{x \mid g(x) \leq t\} = g^{-1}([-\infty, t]) \in \Sigma$$

for all $t \in \mathbb{R}$.

Measurability of a function f is not only useful for integration, but it allows us to specify a new class of measures:

Definition 2.31 (Image Measure)**.** *Let (X, Σ_X, μ) be a measure space, (Y, Σ_Y) be a measurable space and $f\colon X \to Y$ be a Σ_X-Σ_Y-measurable function. We define the* image measure *of μ, f to be:*

$$\mu_f = \mu \circ f^{-1}\colon \Sigma_Y \to [0,1], \quad S \mapsto \mu \circ f^{-1}(S)$$

Note that the composition $\mu \circ f^{-1}$ is only defined for measurable sets $S \in \Sigma_Y$ since only then the pre-image $f^{-1}(S)$ is guaranteed to be an element of Σ_X.

2.2.5 The Lebesgue Integral

Definition 2.32 (Simple Function)**.** *Let (X, Σ) be a measurable space. We call $f\colon X \to \mathbb{R}_+$ a* (non-negative) simple function *iff there are $\alpha_1, ..., \alpha_n \in \mathbb{R}_+$ and $S_1, ..., S_n \in \Sigma_X$ such that*

$$f(x) = \sum_{i=1}^{n} \alpha_i \chi_{S_i}(x).$$

Note that this representation of f is not unique.

Proposition and Definition 2.33. *(The Lebesgue Integral) Let (X, Σ, μ) be a measure space and $f\colon X \to \overline{\mathbb{R}}$ be a Borel-measurable function.*

- *If f is a (non-negative) simple function, say $f(x) = \sum_{i=1}^{n} \alpha_i \chi_{S_i}(x)$, we define the μ-integral of f to be:*
$$\int f \,\mathrm{d}\mu := \sum_{i=1}^{n} \alpha_i \mu(S_i) \in \overline{\mathbb{R}}_+$$
This value is independent of the representation of f.

- *If f is non-negative, we define the μ-integral of f to be:*
$$\int f \,\mathrm{d}\mu := \sup \left\{ \int s \,\mathrm{d}\mu \;\middle|\; s \text{ is a simple function and } 0 \le s \le f \right\} \in \overline{\mathbb{R}}_+$$

- *If f takes on negative values, we consider the* positive part $f^+ := \max\{f, 0\}$ *and the* negative part $f^- := \max\{-h, 0\}$ *of f and notice that $|f| = f^+ + f^-$ and $f = f^+ - f^-$. Both, f^+ and f^-, are non-negative and Borel-measurable functions and thus we can define the μ-integral of f to be:*
$$\int f \,\mathrm{d}\mu := \int f^+ \,\mathrm{d}\mu - \int f^- \,\mathrm{d}\mu \in \overline{\mathbb{R}}$$
if not $\int f^+ \,\mathrm{d}\mu = \int f^- \,\mathrm{d}\mu = +\infty$. In the latter case we say, that the integral does not exist.

*We call f a (μ-)*integrable *or* Lebesgue-integrable *function if the (μ-)integral exists and is finite and write $f \in \mathcal{L}_1(X, \Sigma, \mu)$ in that case. For measurable $S \in \Sigma$ we note that $f \cdot \chi_S\colon X \to \mathbb{R}$ is Borel-measurable and define*
$$\int_S f \,\mathrm{d}\mu := \int f \cdot \chi_S \,\mathrm{d}\mu$$
The following notation might be used synonymously:
$$\int_S f \,\mathrm{d}\mu = \int_S f(x) \,\mathrm{d}\mu(x) = \int_S f(x)\, \mu(\mathrm{d}x).$$

Proof. This can be found in any standard book on integration, e.g. cf. [Ash, 1972, p. 36f.]. □

Now that we know what the Lebesgue integral is, let us recite a few properties of integration in the subsequent propositions:

Proposition 2.34 (Integrability)**.** *Let (X, Σ, μ) be a measurable space. The following statements are equivalent for every function $f\colon X \to \overline{\mathbb{R}}$:*

(i) f is μ-integrable.

(ii) f is measurable and there is a measurable $g\colon X \to \overline{\mathbb{R}}_+$ such that $|f| \leq g$.

Proof. Cf. [Elstrodt, 2007, Satz IV.3.3, p. 130]. □

Proposition 2.35 (Linearity of the Integral)**.** *Let (X, Σ, μ) be a measurable space, $f, g\colon X \to \overline{\mathbb{R}}$ be integrable functions and $\alpha, \beta \in \mathbb{R}$. Then $\alpha f + \beta g$ is integrable and*

$$\int \alpha f + \beta g \, \mathrm{d}\mu = \alpha \cdot \int f \, \mathrm{d}\mu + \beta \cdot \int g \, \mathrm{d}\mu.$$

Proof. Cf. [Elstrodt, 2007, Satz IV.3.6, pp. 131f.]. □

Proposition 2.36 (Monotonicity of the Integral)**.** *Let (X, Σ, μ) be a measurable space and $f, g\colon X \to \overline{\mathbb{R}}$ be μ-integrable functions such that $f \leq g$. Then the following holds:*

$$\int f \, \mathrm{d}\mu \leq \int g \, \mathrm{d}\mu.$$

Proof. Cf. [Elstrodt, 2007, Satz IV.3.7, p. 132]. □

Proposition 2.37. *Let (X, Σ, μ) be a measure space and $f, g\colon X \to \overline{\mathbb{R}}$ such that f is μ-integrable and $f = g$ a.e. (μ). Then gis integrable and the following holds:*

$$\int f \, \mathrm{d}\mu = \int g \, \mathrm{d}\mu.$$

Proof. Cf. [Elstrodt, 2007, Satz IV.4.2 c, p. 141]. □

Corollary 2.38. *Let (X, Σ, μ) be a measure space, $N \in \Sigma$ be a null set and $f\colon X \to \overline{\mathbb{R}}$ be a μ-integrable function. Then the following holds:*

$$\int f \, \mathrm{d}\mu = \int_{X \setminus N} f \, \mathrm{d}\mu.$$

Proof. Apply Proposition 2.37 to the measurable function $g := f \cdot \chi_{X \setminus N}$. □

Given a measure space (X, Σ_X, μ) and a measurable space (Y, Σ_Y) integration with respect to the image measure of μ and a measurable function $f\colon X \to Y$ has the following property, which can sometimes be used to simplify integral calculations:

Proposition 2.39 (Integration and the Image Measure)**.** *Let (X, Σ_X, μ) be a measure space, (Y, Σ_Y) be a measurable space, $f\colon X \to Y$ be a measurable function, $g\colon Y \to \overline{\mathbb{R}}$ be Borel-measurable. Then the following holds:*

$$g \in \mathcal{L}_1(Y, \Sigma_Y, \mu_f) \quad \Leftrightarrow \quad g \circ f \in \mathcal{L}_1(X, \Sigma_X, \mu)$$

Whenever these statements are valid we have:

$$\int_Y g \,\mathrm{d}\mu_f = \int_X g \circ f \,\mathrm{d}\mu$$

and

$$\int_S g \,\mathrm{d}\mu_f = \int_{f^{-1}(S)} g \circ f \,\mathrm{d}\mu \quad \forall S \in \Sigma_Y.$$

Proof. Cf. [Ash, 1972, Theorem 1.6.12, p. 50f]. □

2.2.6 Disjoint Union Sigma-Algebra

Definition 2.40 (Disjoint Union Sigma-Algebra)**.** *Let I be an arbitrary index set and $(X_i, \Sigma_i)_{i\in I}$ be a family of measurable spaces. We consider the disjoint union $\bigsqcup_{i\in I} X_i$ and define the* disjoint union sigma-algebra *to be the finest (i.e. the biggest with respect to set inclusion) σ-algebra on the disjoint union $\bigsqcup_{i\in I} X_i$ for which the canonical injections $\iota_j\colon X_j \to \bigsqcup_{i\in I} X_i, x \mapsto (x, j)$ are measurable (for every $j \in I$). We denote this σ-algebra by $\bigoplus_{i\in I} \Sigma_i$ or for finite I, say $I = \{1, ..., n\}$, by $\Sigma_1 \oplus \Sigma_2 \oplus ... \oplus \Sigma_n$.*

The given definition of the disjoint union sigma-algebra is non-constructive but there is a way of describing the measurable sets explicitly which we show in the following proposition.

Proposition 2.41. *Let I be an arbitrary index set and $(X_i, \Sigma_i)_{i\in I}$ be a family of measurable spaces. We consider the disjoint union $\bigsqcup_{j\in I} X_j$ together with the canonical injections $\iota_i\colon X_i \to \bigsqcup_{j\in I} X_j, x \mapsto (x, i)$. We define:*

$$\mathcal{S}_1 := \left\{ S \in \mathcal{P}\left(\bigsqcup_{i\in I} X_i\right) \;\middle|\; \forall i \in I : \iota_i^{-1}(S) \in \Sigma_i \right\}, \quad \mathcal{S}_2 := \left\{ \bigsqcup_{i\in I} S_i \;\middle|\; \forall i \in I : S_i \in \Sigma_i \right\}$$

Then the following holds: $\bigoplus_{j\in I} \Sigma_j = \mathcal{S}_1 = \mathcal{S}_2$.

Proof. We first show that $\mathcal{S}_2$ is a σ-algebra:

(1) For every $i \in I$ we have $X_i \in \Sigma_j$ and thus $\bigsqcup_{i\in I} X_i \in \mathcal{S}$.

(2) Let $\bigsqcup_{i\in I} A_i \in \mathcal{S}_2$, then the following holds:

$$X \setminus \left(\bigsqcup_{i\in I} A_i\right) = \left(\bigsqcup_{i\in I} X_i\right) \setminus \left(\bigsqcup_{i\in I} A_i\right) = \bigsqcup_{i\in I} \underbrace{(X_i \setminus A_i)}_{\in \Sigma_i} \in \mathcal{S}_2.$$

(3) Let $\left(\bigsqcup_{i\in I} S_n^i\right)_{n\in\mathbb{N}}$ be a countable family of sets from $\mathcal{S}_2$. Then the following holds:

$$\left(\bigcup_{n\in\mathbb{N}} \left(\bigsqcup_{i\in I} S_n^i\right)\right) = \left(\bigsqcup_{i\in I} \underbrace{\left(\bigcup_{n\in\mathbb{N}} S_n^i\right)}_{\in \Sigma_i}\right) \in \mathcal{S}_2.$$

Now we show that $\mathcal{S}_1 = \mathcal{S}_2$ and hence $\mathcal{S}_1$ is a σ-algebra.

"$\subseteq$" Let $S \in \mathcal{S}_1$. We define $S_i := \iota_i^{-1}(S)$ for every $i \in I$, then $S = \bigsqcup_{i\in I} S_i$ and obviously $S_i = \iota_i^{-1}(S) \in \Sigma_i$ for every $i \in I$.

"$\supseteq$" We denote that for every $\bigsqcup_{i\in I} S_i \in \mathcal{S}_2$ we have $\iota_i^{-1}\left(\bigsqcup_{i\in I} S_i\right) = S_i \in \Sigma_i$ for every $i \in I$.

By looking at the definition of $\mathcal{S}_1$ it becomes obvious that $\mathcal{S}_1$ is the finest σ-algebra for which all the injections are measurable and hence we have $\bigoplus_{j\in I} \Sigma_j = \mathcal{S}_1 = \mathcal{S}_2$. □

2.2.7 Product Sigma-Algebra and Product Measure

Definition 2.42 (Initial Sigma-Algebra). *Let I be an arbitrary set, $((X_i, \Sigma_i))_{i\in I}$ be a family of measurable spaces, X be a set and $(f_i)_{i\in I}$ be a family of functions $f_i : X \to X_i$. We define a σ-algebra on X via:*

$$\Sigma := \sigma\left(\bigcup_{i\in I} f_i^{-1}(\Sigma_i)\right).$$

This σ-algebra is called the initial sigma-algebra *for the f_i. By definition this is the smallest σ-algebra on X (with respect to set theoretic inclusion) for which all the f_i are measurable.*

Definition 2.43 (Product Sigma-Algebra). *In Definition 2.42 let $X := \prod_{j\in I} X_j$, for every $i \in I$ let $f_i := \pi_i \colon \prod_{j\in I} X_j \to X_i, (\omega_j)_{j\in I} \mapsto \omega_i$ be the canonical projection. Then the initial σ-algebra is called the* product σ-algebra, *denoted by $\bigotimes_{j\in I} \Sigma_j$. It is the smallest σ-algebra for which the projections are measurable.*

While the definition above is rather theoretic we can also specify generators for the product σ-algebra:

Proposition 2.44 (Generators for the Product σ-Algebra)**.** *Let (X, Σ_X), (Y, Σ_Y) be measurable spaces and $\mathcal{G}_X \subseteq \mathcal{P}(X)$, $\mathcal{G}_Y \subseteq \mathcal{P}(Y)$ such that $X \in \mathcal{G}_X$, $\sigma(\mathcal{G}_X) = \Sigma_X$, $Y \in \mathcal{G}_Y$ and $\sigma(\mathcal{G}_Y) = \Sigma_Y$. Then the set:*

$$\mathcal{G}_X * \mathcal{G}_Y := \{S_X \times S_Y \mid S_X \in \mathcal{G}_X, S_Y \in \mathcal{G}_Y\}$$

is a generator for the product σ-algebra, i.e. the following holds:

$$\sigma(\mathcal{G}_X * \mathcal{G}_Y) = \sigma(\mathcal{G}_X) \otimes \sigma(\mathcal{G}_Y).$$

Particularly the set $\Sigma_X * \Sigma_Y$ is a generator for the product σ-algebra and a semi-ring of sets.

Proof. Cf. [Elstrodt, 2007, Beispiel III.5.3, p. 113] and Proposition 2.18. □

Proposition and Definition 2.45 (Product Measure)**.** *Let (X, Σ_X, μ) and (Y, Σ_Y, ν) be measure spaces, μ, ν be σ-finite. There exists a uniquely determined measure $\mu \otimes \nu \colon \Sigma_X \otimes \Sigma_Y \to \overline{\mathbb{R}}_+$ on $(X \times Y, \Sigma_X \otimes \Sigma_Y)$ such that for all $S_X \times S_Y \in \Sigma_X * \Sigma_Y$ the following holds:*

$$\mu \otimes \nu(S_X \times S_Y) = \mu(S_X) \cdot \nu(S_Y).$$

The measure $\mu \otimes \nu$ is σ-finite and is called the product measure *of μ and ν.*

Proof. Cf. [Elstrodt, 2007, Satz und Definition V.1.5, p. 167]. □

2.3 Probabilistic Transition Systems

Since we now have provided the necessary background from measure theory, we are - at last - ready to give a thorough definition of probabilistic transition systems:

Definition 2.46 (Probabilistic Transition System w/o Termination)**.** *A* probabilistic transition system without (explicit) termination *is a tuple $((X, \Sigma), \mathcal{A}, \alpha)$ where:*

- *(X, Σ) is an arbitrary measurable space, called the* state space*,*
- *$\mathcal{A}$ is an alphabet which we consider as measurable space by endowing it with the sigma-algebra $\mathcal{P}(\mathcal{A})$,*

- $\alpha\colon X \to \mathbb{S}(\mathcal{A} \times X)$ *is the* transition function *which assigns to every state* $x \in X$ *a sub-probability measure on the measurable space* $\big(\mathcal{A} \times X, \mathcal{P}(\mathcal{A}) \otimes \Sigma\big)$*.*

For every $x \in X$ *and every* $a \in \mathcal{A}$ *we define the* probability of making an a-transition from x *to be:*

$$P_{x,a}\colon \Sigma \to [0,1], \quad S \mapsto \alpha(x)(\{a\} \times S).$$

We call the transition system

- discrete *iff the set* X *is at most countable and endowed with the* σ*-algebra* $\mathcal{P}(X)$ *and*
- continuous *iff the set* X *is uncountably infinite.*

As we have already seen in the introduction, a system with explicit termination needs a dedicated state denoting termination:

Definition 2.47 (Probabilistic Transition System with Termination)**.** *A* probabilistic transition system with (explicit) termination *is a tuple* $((X,\Sigma),\mathcal{A},\alpha)$ *where*

- (X,Σ) *is an arbitrary measurable space, called the* state space*,*
- $\mathcal{A}$ *is an alphabet which we consider as measurable space by endowing it with the sigma-algebra* $\mathcal{P}(\mathcal{A})$*,*
- $\alpha\colon X \to \mathbb{S}(\mathcal{A} \times X + \mathbf{1})$ *is the* transition function *which assigns to every state* $x \in X$ *a sub-probability measure on the measurable space* $\big(\mathcal{A} \times X + \mathbf{1}, \sigma\big(\mathcal{P}(\mathcal{A}) * \Sigma \oplus \mathcal{P}(\mathbf{1})\big)\big)$*.*

The unique symbol $\checkmark \in \mathbf{1}$ *denotes* termination *of the system, hence* $\alpha(x)(\mathbf{1})$ *is the probability of terminating at the state* x*. For every* $x \in X$ *and every* $a \in \mathcal{A}$ *we define the* probability of making an a-transition from x *to be:*

$$P_{x,a}\colon \Sigma \to [0,1], \quad S \mapsto \alpha(x)(\{a\} \times S).$$

We call the transition system

- discrete *iff the set* X *is at most countable and endowed with the* σ*-algebra* $\mathcal{P}(X)$ *and*
- continuous *iff the set* X *is uncountably infinite.*

Before we can continue with properties of such transition systems, we need to take a closer look at category theory.

2.4 Category Theory

Category theory is a branch of abstract algebra which allows to unify or generalize concepts from different other disciplines of mathematics. While studying different mathematical structures, one often has maps between those structures which preserve the structure. A few examples are:

- Sets and total functions.
- Groups and group homorphisms.
- Topological spaces and continuous maps.
- Vector spaces over a fixed field F and F-linear transformations.
- Measurable spaces and total, measurable functions.

In category theory, each of these examples form a *category*: the structures are called *objects* and the structure preserving maps are called *arrows*. Many properties of categories are obtained by looking at the arrows instead of the objects. The definitions and results from category theory in this section are taken from or based on Pierce [1991], Mac Lane [1998] and Bonchi et al. [2011].

2.4.1 Categories

Definition 2.48 (Category)**.** *A* category $\mathbf{C}$ *consists of:*

- *A collection of* objects *denoted* $\mathbf{obj}(\mathbf{C})$*.*
- *A collection of* arrows *(also called* morphisms*) denoted* $\mathbf{arr}(\mathbf{C})$*. Associated with each arrow* $f \in \mathbf{arr}(\mathbf{C})$ *are two objects of the category, the* domain *of* f*, denoted by* $\mathrm{dom}(f)$ *and the* codomain *of* f *denoted by* $\mathrm{cod}(f)$*. We denote an arrow* $f \in \mathbf{arr}(\mathbf{C})$ *with domain* A *and codomain* B *by* $f\colon A \to B$*.*
- *A composition operator assigning to a pair of arrows* $f, g \in \mathbf{arr}(\mathbf{C})$ *where* $\mathrm{dom}(g) = \mathrm{cod}(f)$ *a* composite arrow $g \circ f\colon \mathrm{dom}(f) \to \mathrm{cod}(g)$ *satisfying the following* associative law *for all arrows* $f\colon A \to B, g\colon B \to C, h\colon C \to D$*:*

$$h \circ (g \circ f) = (h \circ g) \circ f$$

- *An* identity arrow $\mathrm{id}_A \colon A \to A$ *for every object A satisfying the following* identity law *for every $(f \colon A \to B) \in \mathbf{arr}(\mathbf{C})$:*

$$f \circ \mathrm{id}_A = f \quad \textit{and} \quad \mathrm{id}_B \circ f = f$$

We present a well-known category as our first example:

Example 2.49 (Category of Sets)**.** **Set** *is the category of sets. Its objects are sets and its arrows are total functions. Composition of arrows is set-theoretic function composition and the identity arrows are the identity functions.*

Definition 2.50 (Isomorphism)**.** *Let* $\mathbf{C}$ *be a category. An arrow $f \colon X \to Y$ is called in* isomorphism *iff there is an arrow $f^{-1} \colon Y \to X$, called the* inverse of f, *such that $f \circ f^{-1} = \mathrm{id}_Y$ and $f^{-1} \circ f = \mathrm{id}_X$. Two objects X and Y are called* isomorphic *(denoted $X \cong Y$) iff there is an isomorphism $f \colon X \to Y$.*

Definition 2.51 (Initial and Terminal Objects)**.** *Let* $\mathbf{C}$ *be a category. An object $I \in \mathbf{obj}(\mathbf{C})$ is called an* inital object *iff for all other objects $X \in \mathbf{obj}(\mathbf{C})$ there is a unique arrow $f \colon I \to X$. Analogously, $T \in \mathbf{obj}(\mathbf{C})$ is called a* terminal object *iff for all other objects $X \in \mathbf{obj}(\mathbf{C})$ there is a unique arrow $f \colon X \to T$.*

Example 2.52 (Initial and Terminal Objects in **Set**)**.** *The empty set $\emptyset$ is the only initial object in* **Set** *with the unique arrow being the empty function. Every set with only one element is a terminal object with the unique arrow being the constant function $f \colon X \to \mathbf{1}, x \mapsto \checkmark$ for every object X.*

Definition 2.53 (Product)**.** *Let* $\mathbf{C}$ *be a category and I be an arbitrary index set. For a family of objects $(X_i)_{i \in I}$ we call an object $\prod\limits_{i \in I} X_i$ together with* projection arrows *$\pi_i \colon \prod\limits_{i \in I} X_i \to X_i$ for every $i \in I$ a* product of the X_i *iff for every other object Y with arrows $\tau_i \colon Y \to X_i$ for every $i \in I$ there exists a unique* mediating arrow *denoted by $\langle \tau_i \rangle_{i \in I} \colon Y \to \prod\limits_{i \in I} X_i$ such that the following diagram commutes for every $i \in I$:*

$$\begin{tikzcd}
Y \arrow[d, dotted, "\langle \tau_i \rangle_{i \in I}"'] \arrow[dr, "\tau_i"] & \\
\prod\limits_{i \in I} X_i \arrow[r, "\pi_i"'] & X_i
\end{tikzcd}$$

For finite I, e.g. $I = \{1, \ldots, n\}$, we also denote the product - if it exists - by $X_1 \times X_2 \times \ldots \times X_n$. For $I = \{1, 2\}$ the product is called a binary product.

Definition 2.54 (Coproduct)**.** *Let* $\mathbf{C}$ *be a category and* I *be an arbitrary index set. For a family of objects* $(X_i)_{i \in I}$ *we call an object* $\coprod_{i \in I} X_i$ *together with* injection arrows $\iota_i \colon X_i \to \coprod_{i \in I} X_i$ *for every* $i \in I$ *a* coproduct of the X_i *iff for every other object* Y *with arrows* $\tau_i \colon X_i \to Y$ *for every* $i \in I$ *there exists a unique* mediating arrow *denoted by* $[\tau_i]_{i \in I} \colon \coprod_{i \in I} X_i \to Y$ *such that the following diagram commutes for every* $i \in I$*:*

$$\begin{tikzcd}
\coprod_{i \in I} X_i \arrow[d, dotted, "{[\tau_i]_{i \in I}}"'] & X_i \arrow[l, "\iota_i"'] \arrow[dl, "\tau_i"] \\
Y &
\end{tikzcd}$$

For finite I*, e.g.* $I = \{1, \ldots, n\}$*, we also denote the coproduct - if it exists - by* $X_1 + X_2 + \ldots + X_n$*. For* $I = \{1, 2\}$ *the product is called a* binary coproduct.

2.4.2 Functors

While we have so far considered *arrows* which are structure-preserving maps within a category, we will now focus our attention on *functors* which are structure-preserving maps from one category to another.

Definition 2.55 (Functor)**.** *Let* $\mathbf{C}, \mathbf{D}$ *be categories. A* functor $F \colon \mathbf{C} \to \mathbf{D}$ *is a map that takes*

- *each object* $X \in \mathbf{obj}(\mathbf{C})$ *to an object* $F(X) \in \mathbf{obj}(\mathbf{D})$ *and*
- *each arrow* $(f \colon X \to Y) \in \mathbf{arr}(\mathbf{C})$ *to an arrow* $(F(f) \colon F(X) \to F(Y)) \in \mathbf{arr}(\mathbf{D})$

such that the following holds:

- $F(id_A) = id_{F(A)}$ *for all* $\mathbf{C}$*-objects* A *and*
- $F(f \circ g) = F(f) \circ F(g)$ *for all* $\mathbf{C}$*-arrows* $f \colon B \to C, g \colon A \to B$

If $\mathbf{C}$ *and* $\mathbf{D}$ *are identical, the functor is called an* endofunctor*.*

The following is a trivial example of a functor:

Example 2.56 (Identity Functor)**.** *Let* $\mathbf{C}$ *be an arbitrary category. The map* $\mathrm{Id}_{\mathbf{C}} \colon \mathbf{C} \to \mathbf{C}$ *taking every* $\mathbf{C}$*-object* X *to itself and each* $\mathbf{C}$*-arrow* f *to itself is an endofunctor.*

We will now present two additional popular examples (the first is taken from Ex. 2.1.7 in Pierce [1991] with slight modifications, the second is constructed analogously) of functors for categories with all binary (co-)products, which we will also need in the following chapters of this thesis:

Example 2.57 (Product Functor)**.** *Let* $\mathbf{C}$ *be a category with all binary products. Every* $\mathbf{C}$*-object* A *determines a functor* $(\mathrm{Id}_{\mathbf{C}} \times A)\colon \mathbf{C} \to \mathbf{C}$*, the* right product (with A) functor *that takes each object* X *to* $X \times A$ *and each arrow* $f\colon X \to Y$ *to the arrow* $f \times \mathrm{id}_A\colon X \times A \to Y \times A, (x, a) \mapsto \big(f(x), a\big)$*. The* left product (with A) functor *is defined analogously.*

Example 2.58 (Coproduct Functor)**.** *Let* $\mathbf{C}$ *be a category with all binary coproducts. Every* $\mathbf{C}$*-object* A *determines a functor* $(\mathrm{Id}_{\mathbf{C}} + A)\colon \mathbf{C} \to \mathbf{C}$*, the* right coproduct (with A) functor *that takes each object* X *to* $X + A$ *and each arrow* $f\colon X \to Y$ *to the arrow* $f + \mathrm{id}_A\colon X + A \to Y + A, (x, a) \mapsto f(x) + a$*. The* left coproduct (with A) functor *is defined analogously.*

A structure-preserving map from one functor to another functor is called a *natural transformation*:

Definition 2.59 (Natural Transformation)**.** *Let* $\mathbf{C}, \mathbf{D}$ *be categories and* F *and* G *be functors from* $\mathbf{C}$ *to* $\mathbf{D}$*. A* natural transformation η from F to G *is a function that assigns to every* $\mathbf{C}$*-object* A *a* $\mathbf{D}$*-arrow* $\eta_A\colon F(A) \to G(A)$ *such that for any* $\mathbf{C}$*-arrow* $f\colon A \to B$ *the following diagram commutes in* $\mathbf{D}$*:*

$$\begin{tikzcd}
F(A) \arrow[r, "\eta_A"] \arrow[d, "F(f)"'] & G(A) \arrow[d, "G(f)"] \\
F(B) \arrow[r, "\eta_B"] & G(B)
\end{tikzcd}$$

We denote a natural transformation η *from* F *to* G *by* $\eta\colon F \Rightarrow G$*.*

2.4.3 Coalgebra

Definition 2.60 (Coalgebra)**.** *Let* F *be an endofunctor on a category* $\mathbf{C}$*. A (*F*-)coalgebra is a pair* (X, α) *where* X *is an object of* $\mathbf{C}$ *and* $\alpha\colon X \to FX$ *is a morphism in* $\mathbf{C}$*.*

Definition 2.61 (Coalgebra Homomorphism)**.** *Let* F *be an endofunctor on a category* $\mathbf{C}$*,* $(X, \alpha), (Y, \beta)$ *be* F*-coalgebras. A (*F*-coalgebra) homomorphism* $f\colon (X, \alpha) \to (Y, \beta)$ *is* *a* $\mathbf{C}$*-morphism such that* $F(f) \circ \alpha = \beta \circ f$*, i.e. such that the following diagram commutes:*

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "\alpha"'] & Y \arrow[d, "\beta"] \\ FX \arrow[r, "F(f)"'] & FY \end{tikzcd}$$

Definition 2.62 (Final Coalgebra)**.** *Let* F *be an endofunctor on a category* $\mathbf{C}$*. An* F*-coalgebra* (Ω, ω) *is final iff for any* F*-coalgebra* (X, α) *there exists a unique (*F*-coalgebra) homomorphism* $\phi_X \colon (X, \alpha) \to (\Omega, \omega)$*.*

2.4.4 Monads and Kleisli Categories

Definition 2.63 (Monad)**.** *Let* $\mathbf{C}$ *be a category. A* monad *on* C *is an endofunctor* $T \colon \mathbf{C} \to \mathbf{C}$ *together with*

- *a* unit *natural transformation* $\eta \colon \mathrm{Id} \Rightarrow T$*, i.e. morphisms* $\eta_X \colon X \to TX$ *for every* $\mathbf{C}$*-object* X *and*
- *a* multiplication *natural transformation* $\mu \colon T^2 \Rightarrow T$*, i.e. morphisms* $\mu_X \colon TTX \to TX$ *for every* $\mathbf{C}$*-object* X

such that the following diagrams commute for all $\mathbf{C}$*-objects* X*:*

$$\begin{tikzcd} TX \arrow[r, "\eta_{TX}"] \arrow[dr, "\mathrm{id}_{TX}"'] & T^2X \arrow[d, "\mu_X"] & TX \arrow[l, "T(\eta_X)"'] \arrow[dl, "\mathrm{id}_{TX}"] \\ & TX & \end{tikzcd} \qquad \begin{tikzcd} T^3X \arrow[r, "T(\mu_X)"] \arrow[d, "\mu_{TX}"'] & T^2X \arrow[d, "\mu_X"] \\ T^2X \arrow[r, "\mu_X"'] & TX \end{tikzcd}$$

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[3]. In the following definition we will adopt the notation used by Mac Lane[4], as this notation 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 2.64 (The Kleisli Category of a Monad)**.** *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*

[3] Cf. Bonchi et al. [2011]
[4] Cf. [Mac Lane, 1998, Theorem VI.5.1, p. 147]

$f\colon X \to TY$ 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 follows:*

$$g^\flat \circ f^\flat := (\mu_Z \circ T(g) \circ f)^\flat$$

For every object X_T the identity arrow is $\mathrm{id}_{X_T} = (\eta_X)^\flat$.

Proposition 2.65 (Induced Isomorphism)**.** *Let (T, η, μ) be a monad on a category $\mathbf{C}$. Every isomorphism $\phi\colon X \to Y$ in $\mathbf{C}$ induces an isomorphism $(\eta_Y \circ \phi)^\flat\colon X_T \to Y_T$ in $\mathcal{K}\ell(T)$.*

Proof. Let $\psi\colon Y \to X$ be the inverse arrow of ϕ. We consider the following diagram in the category $\mathbf{C}$:

$$\begin{tikzcd}
Y \arrow[r, "\psi"] \arrow[dr, "\mathrm{id}_Y"'] & X \arrow[r, "\eta_X"] \arrow[d, "\phi"] & TX \arrow[r, "T(\eta_Y\circ\phi)"] \arrow[d, "T(\phi)"'] & TTY \arrow[r, "\mu_Y"] & TY \\
& Y \arrow[r, "\eta_Y"'] & TY \arrow[ur, "T(\eta_Y)"] \arrow[urr, "\mathrm{id}_{TY}"'] & &
\end{tikzcd}$$

and denote that it commutes because:

- ϕ and ψ are inverse isomorphisms: $\phi \circ \psi = \mathrm{id}_Y$,
- $\eta\colon \mathrm{Id}_{\mathbf{C}} \Rightarrow T$ is a natural transformation: $\eta_Y \circ \phi = T(\phi) \circ \eta_X$,
- T is a functor: $T(\eta_Y \circ \phi) = T(\eta_Y) \circ T(\phi)$ and
- μ_Y is the multiplication natural transformation: $\mu_Y \circ T(\eta_Y) = \mathrm{id}_{TY}$.

We conclude that:

$$\mu_Y \circ T(\eta_Y \circ \phi) \circ (\eta_X \circ \psi) = \mathrm{id}_{TY} \circ \eta_Y \circ \mathrm{id}_Y = \eta_Y$$

and hence we have:

$$(\eta_Y \circ \phi)^\flat \circ (\eta_X \circ \psi)^\flat = (\mu_Y \circ T(\eta_Y \circ \phi) \circ (\eta_X \circ \psi))^\flat = \eta_Y^\flat = \mathrm{id}_{Y_T}$$

Analogously we obtain:

$$(\eta_X \circ \psi)^\flat \circ (\eta_Y \circ \phi)^\flat = \mathrm{id}_{X_T}$$

and the proof is complete. □

2.4.5 Coalgebra in a Kleisli Category

Let G be an endofunctor on $\mathcal{K}\ell(T)$. By our previous definition, a G-coalgebra is a tuple $(X_T, \alpha^\flat)$ where X_T is an object of the Kleisli category and $\alpha^\flat\colon X_T \to G(X_T)$ is an arrow in the Kleisli category. Thus the corresponding arrow in the base category $\mathbf{C}$ is $\alpha\colon X \to TGX$.

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 denote 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 2.66 (Distributive Law)**.** *Let (T, η, μ) be a monad on a category $\mathbf{C}$ and $F\colon \mathbf{C} \to \mathbf{C}$ be an endofunctor on $\mathbf{C}$. A natural transformation $\lambda\colon FT \Rightarrow TF$ is called a distributive law iff the following diagrams commute in $\mathbf{C}$:*

$$\begin{tikzcd} FX \arrow[r, "F(\eta_X)"] \arrow[dr, "\eta_{FX}"'] & FTX \arrow[d, "\lambda_X"] \\ & TFX \end{tikzcd} \qquad \begin{tikzcd} FT^2X \arrow[r, "\lambda_{TX}"] \arrow[d, "F(\mu_X)"'] & TFTX \arrow[r, "T(\lambda_X)"] & T^2FX \arrow[d, "\mu_{FX}"] \\ FTX \arrow[rr, "\lambda_X"'] & & TFX \end{tikzcd}$$

Definition 2.67 (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 $\overline{F}(f^\flat) := (\lambda_Y \circ Ff)^\flat$. The corresponding arrow in the base category $\mathbf{C}$ is the arrow $\lambda_Y \circ Ff\colon FX \to TFY$.*

Definition 2.68 (Coalgebra in a Kleisli Category)**.** *Let (T, η, μ) be a monad on a category $\mathbf{C}$ and F be an endofunctor on $\mathbf{C}$ with a distributive law $\lambda\colon FT \Rightarrow TF$. A* (F-)coalgebra *in the Kleisli category $\mathcal{K}\ell(T)$ of the monad T is a $\overline{F}$-coalgebra in $\mathcal{K}\ell(T)$. Thus it is a tuple $(X_T, \alpha^\flat)$ where X_T is an object in $\mathcal{K}\ell(T)$ and $\alpha^\flat\colon X_T \to \overline{F}(X_T)$ is an arrow in $\mathcal{K}\ell(T)$. The corresponding tuple in the base category $\mathbf{C}$ is the tuple (X, α) where $\alpha\colon X \to TFX$.*

2.5 Graph Theory

As we will make use of König's Lemma in a later proof, let us state the few results we need from graph theory:

Definition 2.69 (Undirected Simple Graph)**.** *An* undirected simple graph *is a pair* $G = (V, E)$ *where* V *is an arbitrary set whose elements are called* vertices *or* nodes *and* $E \subseteq \{e \in \mathcal{P}(V) \mid |e| = 2\} = \big\{\{v_1, v_2\} \in \mathcal{P}(V) \mid v_1 \neq v_2\big\}$ *is called the set of* edges*. If both,* V *and* E*, are finite sets the graph is called* finite*, otherwise* infinite*. A* finite/infinite path *is a finite/infinite sequence of vertices where for every vertex in the sequence there is an edge from the vertex to its direct predecessor and an edge to its direct successor (if those exist). Finite and infinite paths have a* start vertex *and a finite path also has an* end vertex*. An arbitrary path is called a* simple path *iff all the vertices in the path are mutually distinct. Two vertices* $v, w \in V$ *are called* connected *if there is a path containing both* v *and* w *and if every pair of vertices* $u, v \in V$ *is connected, the whole graph is called a* connected graph*. For a vertex* $v \in V$*, the degree of* v *is the number:*

$$\deg(v) := |\{e \in E \mid v \in e\}| \in \mathbb{N}_0 \cup \{\infty\}$$

Definition 2.70 (Tree)**.** *An undirected graph* $T = (V, E)$ *is called a* tree *iff any two vertices* $u, v \in V$ *are connected by exactly one path. A tree is called* rooted *iff one vertex* $r \in V$ *has been designated the* root*.*

Proposition 2.71 (König's Lemma)**.** *Let* G *be an infinite, connected graph where every vertex has finite degree. For every vertex* $v \in V$ *there exists an infinite, simple path* $(v_n)_{n\in\mathbb{N}}$ *such that* $v_1 = v$*.*

Proof. Cf. [König, 1936, Satz 3, p. 80]. □

3 The Category of Measurable Spaces

Due to the fact that we want to model *probabilistic* transition systems, it is reasonable to assume that the category in which we want to do that must have something to do with the aspect of measurability. Other authors, e.g. Hasuo et al. [2007], try to avoid this by allowing only discrete probabilistic transition systems, i.e. probabilistic transition systems where the state space is at most countably infinite. Doing so, they are able to work in the well-known category **Set**. We use the category of measurable spaces, **Meas**, instead, which allows us to model probabilistic transition systems with uncountable infinite state spaces. Moreover, it enables us to define trace semantics for probabilistic transition systems without explicit termination which - to our knowledge - is impossible in **Set** even for discrete probabilistic transition systems. In this chapter we are going to provide the necessary definitions and results to work in the category **Meas**.

3.1 Definition and Conventions

Definition 3.1 (The Category of Measurable Spaces)**.** *We define the* category of measurable spaces *denoted* **Meas** *as follows:*

- *The objects of* **Meas** *are the measurable spaces* $X = (X', \Sigma)$ *where* X' *is an arbitrary set endowed with an arbitrary* σ*-algebra* Σ*. For an object* (X', Σ) *we will often use the symbol* X' *to denote the whole tuple* (X, Σ)*. Similarily, given an object* $X = (X', \Sigma)$*, we "forget" the* σ*-algebra and write* X *instead of* X'*.*

- *Let* $X = (X', \Sigma_X), Y = (Y', \Sigma_Y)$ *be measurable spaces. Every* Σ_X*-*Σ_Y *measurable function* $f' \colon X' \to Y'$ *becomes an arrow* $f \colon X \to Y$ *of* **Meas***. Conversely this means that for every arrow* $f \colon X \to Y$ *there is a uniquely determined* Σ_X*-*Σ_Y *measurable function* $f' \colon X' \to Y'$*. Although* f *and* f' *are technically different we will often denote both maps by the same symbol* f*.*

- *Let* $X = (X', \Sigma_X), Y = (Y', \Sigma_Y), Z = (Z', \Sigma_Z)$ *be measurable spaces. For two arrows* $f \colon X \to Y$ *and* $g \colon Y \to Z$ *we consider the associated* Σ_X*-*Σ_Y*-measurable function* $f' \colon X' \to Y'$ *and the* Σ_Y*-*Σ_Z*-measurable function* $g' \colon Y' \to Z'$*. According* *to Proposition 2.28, their composition* $g' \circ f' \colon X' \to Z'$ *is a* Σ_X-Σ_Z*-measurable function. We define* $g \circ f \colon X \to Z$ *to be this function.*

- *Let* $X = (X', \Sigma_X)$ *be a measurable space. The identity arrow* $\mathrm{id}_X \colon X \to X$ *is the measurable identity function* $\mathrm{id}_{X'} \colon X' \to X', x \mapsto x$.

Now that we know what the category of measurable spaces **Meas** is, let us provide some results about it which we will need in the main part of this thesis:

Proposition 3.2 (Isomorphisms)**.** *Let* X, Y *be arbitrary sets,* $\mathcal{S} \subseteq \mathcal{P}(X)$, $f \colon X \to Y$ *a bijective function. Then the following holds:*

(a) $\mathcal{S}$ *is a semi-ring of sets iff* $f(\mathcal{S})$ *is a semi-ring of sets.*

(b) $\mathcal{S}$ *is a* σ*-algebra on* X *iff* $f(\mathcal{S})$ *is a* σ*-algebra on* Y.

(c) The measurable spaces $(X, \sigma(\mathcal{S}))$ *and* $(Y, \sigma(f(\mathcal{S})))$ *are isomorphic in* **Meas**.

Proof. For every set $S \in \mathcal{P}(X)$ the following holds $S \in \mathcal{S} \Leftrightarrow f(S) \in f(\mathcal{S})$. The bijectivity of f yields the following equalities for all $A, B \in \mathcal{P}(X)$ and all sequences $(S_n)_{n\in\mathbb{N}}$ of sets $S_n \in \mathcal{P}(X)$:

$$f(A \setminus B) = f(A) \setminus f(B), \quad f\left(\bigcup_{n\in\mathbb{N}} S_n\right) = \bigcup_{n\in\mathbb{N}} f(S_n) \quad \text{and} \quad \left(\bigcap_{n\in\mathbb{N}} S_n\right) = \bigcap_{n\in\mathbb{N}} f(S_n)$$

With the help of these facts it is easy to prove (a) and (b). We omit this part and prove only (c): Let $g \colon Y \to X$ be the inverse function of f. Then the following holds for all $S_X \in \mathcal{S}$:

$$f^{-1}(f(S_X)) = (g \circ f)(S_X) = S_X \in \sigma(\mathcal{S}) \quad \text{and} \quad g^{-1}(S_X) = f(S_X) \in \sigma(f(\mathcal{S}))$$

and thus by Proposition 2.29 f is a $\sigma(\mathcal{S})$-$\sigma(f(\mathcal{S}))$-measurable and g a $\sigma(f(\mathcal{S}))$-$\sigma(\mathcal{S})$-measurable function. Therefore we have associated arrows $f' \colon (X, \sigma(\mathcal{S})) \to (Y, \sigma(f(\mathcal{S})))$ and $g' \colon (Y, \sigma(f(\mathcal{S}))) \to (X, \sigma(\mathcal{S}))$ in **Meas**. The bijectivity of f yields $f' \circ g' = \mathrm{id}_Y$ and $g' \circ f' = \mathrm{id}_X$ and hence f' is an isomorphism in **Meas**. □

3.2 Initial and Terminal Objects

Proposition 3.3 (Initial Objects in **Meas**)**.** *The measurable space* $(\emptyset, \{\emptyset\})$ *is an initial object in* **Meas**.

Proof. Let (X, Σ) be an arbitrary measurable space. The unique arrow mapping $(\emptyset, \{\emptyset\})$ to (X, Σ) is the empty function $f\colon \emptyset \to X$ which is measurable since for all $S \in \Sigma$ we trivially have $f^{-1}(S) = \emptyset \in \{\emptyset\}$. □

Proposition 3.4 (Terminal Objects in **Meas**)**.** *The measurable space* $(\mathbf{1}, \mathcal{P}(\mathbf{1}))$ *is a terminal object in* **Meas**.

Proof. Let (X, Σ) be an arbitrary measurable space. The unique arrow mapping (X, Σ) to $(\mathbf{1}, \mathcal{P}(\mathbf{1}))$ is the constant function $f\colon X \to \mathbf{1}, x \mapsto \checkmark$ which is measurable since $f^{-1}(\emptyset) = \emptyset \in \Sigma$ and $f^{-1}(\mathbf{1}) = X \in \Sigma$ is fullfilled for every σ-algebra Σ. □

3.3 Products

Proposition 3.5. *(Binary Products in* **Meas***) Let* $(X, \Sigma_X), (Y, \Sigma_Y) \in \mathbf{obj}(\mathbf{Meas})$ *be measurable spaces. The measurable space* $(X \times Y, \Sigma_X \otimes \Sigma_Y)$ *together with the projection arrows* $\pi_1\colon X \times Y \to X, (x, y) \mapsto x$ *and* $\pi_2\colon X \times Y \to Y, (x, y) \mapsto y$ *is a product of* (X, Σ_X) *and* (Y, Σ_Y).

Proof. For an arbitrary object $(C, \Sigma_C) \in \mathbf{obj}(\mathbf{Meas})$ together with arrows $\tau_1\colon C \to X$ and $\tau_2\colon C \to Y$ we consider the following diagram in **Meas**:

$$\begin{tikzcd} & C \arrow[dl, "\tau_1"'] \arrow[d, dashed, "\tau"] \arrow[dr, "\tau_2"] & \\ X & X \times Y \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & Y \end{tikzcd}$$

and define the arrow τ to be:

$$\tau\colon C \to X \times Y, \quad c \mapsto (\tau_1(c), \tau_2(c))$$

Obviously this is the only total function making the diagram commute. Measurability of τ is given as well: For $S_X \times S_Y \in \Sigma_X * \Sigma_Y$ we calculate

$$\tau^{-1}(S_X \times S_Y) = \{c \in C \mid \tau_1(c) \in S_X \text{ and } \tau_2(c) \in S_Y\} = \tau_1^{-1}(S_X) \cap \tau_2^{-1}(S_Y)$$

The latter set is measurable as it is the intersection of two sets which are measurable since both τ_1 and τ_2 are measurable. As the set $\Sigma_X * \Sigma_Y$ is a generator of $\Sigma_X \otimes \Sigma_Y$ this suffices to show measurability of τ. □

We immediately obtain the following result:

Corollary 3.6. *The category of measurable spaces has all finite products.*

Proof. The construction from Proposition 3.5 can inductively be extended to finitely many measurable spaces. An empty product exists as well, since **Meas** has terminal objects. □

3.4 Coproducts

The category **Meas** does not only have all finite products but all finite coproducts as well. Now we are ready to define coproducts in **Meas**:

Proposition 3.7. *(Coproducts in* **Meas***) Let $(X, \Sigma_X), (Y, \Sigma_Y) \in \mathbf{obj}(\mathbf{Meas})$ be measurable spaces. The measurable space $(X + Y, \Sigma_X \oplus \Sigma_Y)$ together with the canonical injection arrows $\iota_1 \colon X \to X + Y, x \mapsto (x, 1)$ and $\iota_2 \colon Y \to X + Y, y \mapsto (y, 2)$ is a coproduct of (X, Σ_X) and (Y, Σ_Y).*

Proof. For an arbitrary object $(C, \Sigma_C) \in \mathbf{obj}(\mathbf{Meas})$ together with arrows $\tau_1 \colon X \to C$ and $\tau_2 \colon Y \to C$ we consider the following diagram in **Meas**:

$$\begin{tikzcd}
X \arrow[r, "\iota_1"] \arrow[dr, "\tau_1"'] & X + Y \arrow[d, dashed, "\tau"] & Y \arrow[l, "\iota_2"'] \arrow[dl, "\tau_2"] \\
& C &
\end{tikzcd}$$

and define the arrow τ to be:

$$\tau \colon X + Y \to C, \quad z \mapsto \begin{cases} \tau_1(x) & \text{if } z = (x, 1), x \in X \\ \tau_2(y) & \text{if } z = (y, 2), y \in Y \end{cases}$$

Obviously this is the only total function making the diagram commute. Measurability of τ is given as well: For $S_C \in \Sigma_C$ we calculate:

$$\begin{aligned}
\tau^{-1}(S_C) &= \{z \in X + Y \mid \tau(z) \in S_C\} \\
&= \{(x, 1) \in X + Y \mid \tau_1(x) \in S_C\} \cup \{(y, 2) \in X + Y \mid \tau_2(y) \in S_C\} \\
&= \underbrace{\tau_1^{-1}(S_C)}_{\in \Sigma_X} + \underbrace{\tau_2^{-2}(S_C)}_{\in \Sigma_Y} \in \Sigma_X \oplus \Sigma_Y
\end{aligned}$$

□

Again we immediately obtain the following result:

Corollary 3.8. *The category of measurable spaces has all finite coproducts.*

Proof. The construction from Proposition 3.7 can inductively be extended to finitely many measurable spaces. An empty product exists as well, since **Meas** has initial objects. □

3.5 The (Sub-)Probability Monad

In the last part of this chapter we introduce the (sub-)probability monad on **Meas**. While Giry [1981] constructs just a probability monad, the very same construction works for sub-probabilities as well as Panangaden [2009] points out. The Kleisli category of this monad can be seen as a "category of stochastic relations". Its objects are measurable spaces and its arrows are conditional probability densities or stochastic kernels (cf. [Panangaden, 2009, Chapter 5, pp. 63-76] for further reference). For our considerations it suffices to know how the functor and the natural transformations of the monad are defined.

Definition 3.9 (Evaluation Map, (Sub-)Probability Functor)**.** *Let (X, Σ) be a measurable space. We define the sets:*

$$\begin{aligned}\mathbb{S}(X) &:= \{P\colon \Sigma \to [0,1] \mid P \textit{ is a sub-probability measure}\} \text{ and}\\ \mathbb{P}(X) &:= \{P\colon \Sigma \to [0,1] \mid P \textit{ is a probability measure}\}.\end{aligned}$$

Let $\mathbb{T} \in \{\mathbb{S}, \mathbb{P}\}$. For every $S \in \Sigma$ we define its evaluation map *to be the map*

$$p_S\colon \mathbb{T}(X) \to [0,1], \quad P \mapsto P(S)$$

that takes a (sub-)probability P from the set of all (sub-)probability measures on X and evaluates it on S. We endow the set $\mathbb{T}(X)$ of all (sub-)probabilities with the initial sigma algebra for the evaluation maps

$$\Sigma_{\mathbb{T}(X)} := \sigma\left(\bigcup_{S\in\Sigma} p_S^{-1}(\mathcal{B})\right)$$

and obtain a measurable space $\big(\mathbb{T}(X), \Sigma_{\mathbb{T}(X)}\big)$. By definition this is the smallest σ-algebra for which all the evaluation maps are Borel-measurable. With these preparations we can define the sub-probability functor $\mathbb{S}\colon \mathbf{Meas} \to \mathbf{Meas}$ *and the* probability functor $\mathbb{P}\colon \mathbf{Meas} \to \mathbf{Meas}$*. For $\mathbb{T} \in \{\mathbb{S}, \mathbb{P}\}$ each measurable space $(X, \Sigma) \in \mathbf{obj}(\mathbf{Meas})$ is mapped to $\mathbb{T}\big((X, \Sigma)\big) := \big(T(X), \Sigma_{T(X)}\big)$ and each arrow $(f\colon X \to Y) \in \mathbf{arr}(\mathbf{Meas})$ is* *mapped to an arrow* $\mathbb{T}(f)$ *where for all* $P \in T(X)$ *we have* $\mathbb{T}(f)(P) := P_f$. *Thus* $\mathbb{T}(f)$ *assigns to every (sub-)probability measure* P *the image measure of* P, f.

Definition and Proposition 3.10 (The (Sub-)Probability Monad)**.** *Let* $\mathbb{T} \in \{\mathbb{S}, \mathbb{P}\}$. *We define the following natural transformations:*

- *For all measurable spaces* (X, Σ) *we define the function:*

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

  which assigns to every $x \in X$ *its Dirac measure on* X. *Taking all the corresponding arrows of* **Meas** *we obtain a natural transformation* $\eta \colon \mathrm{Id}_{\mathbf{Meas}} \to \mathbb{T}$.

- *Again for all measurable spaces* (X, Σ) *we define the function* $\mu_X \colon \mathbb{T}^2(X) \to T(X)$ *where for every (sub-)probability measure* $P \in \mathbb{T}^2(X)$ *a (sub-)probability measure on* X *is defined via*

$$\mu_X(P)(S) := \int p_S \,\mathrm{d}P \quad \forall S \in \Sigma_X$$

  where p_S *is the evaluation map. Taking all the corresponding arrows of* **Meas** *we obtain another natural transformation* $\mu \colon \mathbb{T}^2 \Rightarrow \mathbb{T}$.

Then the triple $(\mathbb{T}, \eta, \mu)$ *is a monad on* **Meas**.

Proof. The proof for $\mathbb{T} = \mathbb{P}$ can be found in [Giry, 1981, Theorem 3.1., pp. 70f.] and for $\mathbb{T} = \mathbb{S}$ a proof is in [Panangaden, 2009, Theorem 5.3, pp. 66f.]. □

4 PTS as Coalgebra

4.1 Sigma Algebras on Words

Due to the fact that we are working in the category of measurable spaces, it is necessary to define suitable σ-algebras on the sets of words, $\mathcal{A}^*$ and $\mathcal{A}^\omega$. In this section we will therefore define sets $\mathcal{S}_* \subseteq \mathcal{P}(\mathcal{A}^*)$ and $\mathcal{S}_\omega \subseteq \mathcal{P}(\mathcal{A}^\omega)$ and show, that they are semi-rings of sets. Moreover, given a non-negative map $\mu\colon \mathcal{S} \to \overline{\mathbb{R}}_+$ for $\mathcal{S} \in \{\mathcal{S}_*, \mathcal{S}_\omega\}$ we will present sufficient criteria to show σ-additivity of μ. Of course, we will then consider $\mathcal{A}^*$ and $\mathcal{A}^\omega$ as measurable spaces by endowing them with the σ-algebras generated by $\mathcal{S}_*$ or $\mathcal{S}_\omega$.

4.1.1 A Sigma Algebra on Finite Words

Definition and Proposition 4.1. *Let $\mathcal{A}$ be an arbitrary alphabet. The set*

$$\mathcal{S}_* := \{\emptyset\} \cup \{\{u\} \mid u \in \mathcal{A}^*\}$$

is a semi-ring of sets. For a family $(S_n)_{n\in\mathbb{N}}$ of mutually disjoint sets from $\mathcal{S}_*$ such that $\bigcup_{n\in\mathbb{N}} S_n \in \mathcal{S}_*$ it necessarily holds, that $S_n \neq \emptyset$ for at most one $n \in \mathbb{N}$. Thus every non-negative map $\mu\colon \mathcal{S}_* \to \overline{\mathbb{R}}_+$ where $\mu(\emptyset) = 0$ is trivially σ-additive and hence a pre-measure. Furthermore the following is valid:

$$\sigma(\mathcal{S}_*) = \mathcal{P}(\mathcal{A}^*)$$

Proof. We have already seen that the singletons form a semi-ring of sets and that the generated σ-algebra is the powerset. The rest of the proposition is obvious. □

Definition and Proposition 4.2. *Let $\mathcal{A}$ be an alphabet. We consider the bijective* *functions*

$$\phi\colon \mathcal{A}^* \to \mathcal{A} \times \mathcal{A}^* + \mathbf{1}, \qquad \phi(u) = \begin{cases} \checkmark, & u = \varepsilon \\ (a, u'), & u = au', a \in \mathcal{A}, u' \in \mathcal{A}^* \end{cases}$$

$$\psi\colon \mathcal{A} \times \mathcal{A}^* + \mathbf{1} \to \mathcal{A}^*, \qquad \psi(u) = \begin{cases} \varepsilon, & u = \checkmark \\ \mathbf{conc}(a, u'), & u = (a, u') \end{cases}$$

and define:

$$\mathcal{S}'_* := \phi(\mathcal{S}_*) = \{\emptyset\} \cup \{\checkmark\} \cup \{\{(a, u)\} \mid a \in \mathcal{A}, u \in \mathcal{A}^*\}.$$

Then $\mathcal{S}'_*$ is a semi-ring of sets and the measurable spaces $(\mathcal{A}^*, \sigma(\mathcal{S}_*))$ and $(\mathcal{A} \times \mathcal{A}^* + \mathbf{1}, \sigma(\mathcal{S}'_*))$ are isomorphic in **Meas***.*

Proof. This follows immediately from Proposition 3.2. □

Proposition 4.3. *Let $\mathcal{A}$ be an alphabet. Then the following holds:*

$$\sigma(\mathcal{S}'_*) = \sigma\big(\mathcal{P}(\mathcal{A}) * \sigma(\mathcal{S}_*) \oplus \mathcal{P}(\mathbf{1})\big)$$

Proof. We obviously have $\mathcal{S}'_* \subseteq \mathcal{P}(\mathcal{A}) * \sigma(\mathcal{S}_*) \oplus \mathcal{P}(\mathbf{1})$ and by the monotonicity of the σ-operator we obtain $\sigma(\mathcal{S}'_*) \subseteq \sigma\big(\mathcal{P}(\mathcal{A}) * \sigma(\mathcal{S}_*) \oplus \mathcal{P}(\mathbf{1})\big)$. The set $\mathcal{A} \times \mathcal{A}^* + \mathbf{1}$ is countable and thus we conclude with Example 2.12 that:

$$\mathcal{P}(\mathcal{A} \times \mathcal{A}^* + \mathbf{1}) = \sigma(\mathcal{S}'_*) \subseteq \sigma\big(\mathcal{P}(\mathcal{A}) * \sigma(\mathcal{S}_*) \oplus \mathcal{P}(\mathbf{1})\big) \subseteq \mathcal{P}(\mathcal{A} \times \mathcal{A}^* + \mathbf{1})$$

and thus equality must hold. □

4.1.2 A Sigma Algebra on Infinite Words

Definition 4.4 (ω-Cones)**.** *Let $\mathcal{A}$ be an arbitrary alphabet. We define for every finite word $a \in \mathcal{A}^*$ the* ω-cone *of a to be:*

$$\mathbf{cone}(a) := \{u \in \mathcal{A}^\omega \mid a \sqsubseteq u\}$$

Definition and Proposition 4.5. *Let $\mathcal{A}$ be a finite alphabet. The* set of all ω-cones of *$\mathcal{A}$ given by:*

$$\mathcal{S}_\omega := \{\emptyset\} \cup \{\mathbf{cone}(a) \mid a \in \mathcal{A}^*\}$$

is a semi-ring of sets. Let $(S_n)_{n\in\mathbb{N}}$ be a family of disjoint sets from $\mathcal{S}_\omega$ such that $\bigcup_{n\in\mathbb{N}} S_n \in \mathcal{S}_\omega$. Then $S_n \neq \emptyset$ holds for only finitely many $n \in \mathbb{N}$. After resorting *the S_n we can hence assume that there exist an $N \in \mathbb{N}$ and finite words $u_n \in \mathcal{A}^*$ for every $n \in \{1, ..., N\}$ such that*

$$S_n = \begin{cases} \mathbf{cone}(u_n), & 1 \leq n \leq N \\ \emptyset, & n > N \end{cases}$$

Moreover there is a finite word $u \in \mathcal{A}^*$ such that

$$\bigcup_{n \in \mathbb{N}} S_n = \bigcup_{n=1}^{N} \mathbf{cone}(u_n) = \mathbf{cone}(u)$$

and $u \sqsubseteq u_n$ for all $n \in \{1, ..., N\}$.

Proof. [1] By definition we have $\emptyset \in \mathcal{S}_\omega$. For $A, B \in \mathcal{S}_\omega$ where $A = \emptyset$ or $B = \emptyset$ obviously $A \cap B = \emptyset \in \mathcal{S}_\omega$ and $A \setminus B \in \mathcal{S}_\omega$. For $a, b \in \mathcal{A}^*$ the following holds:

$$\mathbf{cone}(a) \cap \mathbf{cone}(b) = \begin{cases} \mathbf{cone}(b), & \text{if } a \sqsubseteq b \\ \mathbf{cone}(a), & \text{if } b \sqsubseteq a \\ \emptyset, & \text{else} \end{cases}$$

and hence we also have $\mathbf{cone}(a) \cap \mathbf{cone}(b) \in \mathcal{S}_\omega$. Furthermore:

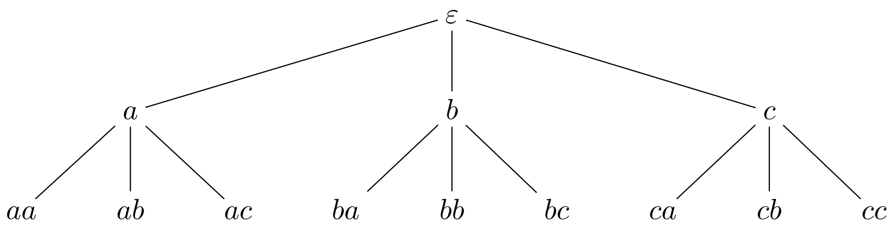
$$\mathbf{cone}(a) \setminus \mathbf{cone}(b) = \begin{cases} \bigcup\limits_{c \in \mathcal{A}^{|b|-|a|} \setminus \{b\}} \mathbf{cone}(ac), & \text{if } a \sqsubseteq b \\ \emptyset, & \text{if } b \sqsubseteq a \\ \mathbf{cone}(a), & \text{else} \end{cases}$$

Obviously, for a finite alphabet, the union in the first case is finite and hence $\mathcal{S}_\omega$ is a semi-ring of sets.

For the second part of the proposition we consider the undirected rooted tree given by $T = (\mathcal{A}^*, E)$ where $E := \{\{u, uv\} \mid u \in \mathcal{A}^*, v \in \mathcal{A}\}$ and $\varepsilon \in \mathcal{A}^*$ is the dedicated root.

[1] Based on ideas (tree-metaphor and the use of Königs Lemma) by Prof. Dr. Barbara König and Mathias Hülsbusch.

For $\mathcal{A} = \{a, b, c\}$ this tree can be depicted as follows:



We can think of the set of all infinite words $\mathcal{A}^\omega$ as being the set of all infinite simple paths starting at the root ε of the tree. For a finite alphabet $\mathcal{A}$ obviously every vertex of this tree has the same, finite degree, namely $|\mathcal{A}| + 1$. Given a finite word $u \in \mathcal{A}^*$, we can identify the cone of u to be the set of all infinite simple paths that begin in ε and contain the vertex u.
Let $(S_n)_{n\in\mathbb{N}}$ be a sequence of mutually disjoint cones such that $\bigcup_{n\in\mathbb{N}} S_n \in \mathcal{S}_\omega$. Then either $\bigcup_{n\in\mathbb{N}} S_n = \emptyset$ and hence $S_n = \emptyset$ for all $n \in \mathbb{N}$ or there must be a finite word $u \in \mathcal{A}$ such that $\bigcup_{n\in\mathbb{N}} S_n = \mathbf{cone}(u)$. We will show by contradiction that this u cannot exist if $S_n \neq \emptyset$ for infinitely many $n \in \mathbb{N}$. Let thus $(S_n)_{n\in\mathbb{N}}$ be a sequence of mutually disjoint cones such that $S_n \neq \emptyset$ for infinitely many $n \in \mathbb{N}$. Without loss of generalization we can assume that for every $n \in \mathbb{N}$ there is an $u_n \in \mathcal{A}^*$ such that $S_n = \mathbf{cone}(u_n) \neq \emptyset$. We assume that there is a suitable $u \in \mathcal{A}^*$ such that $\bigcup_{n\in\mathbb{N}} \mathbf{cone}(u_n) = \mathbf{cone}(u)$. Necessarily we have $u \sqsubseteq u_n$ for every $n \in \mathbb{N}$ or in other words we know that for every $n \in \mathbb{N}$ the vertex u must be contained in the (unique) simple undirected path in T connecting the root ε with u_n. We consider the subgraph $T' := (A, E_A)$ where $A \subseteq \mathcal{A}^*$ is the set of vertices contained in the simple paths connecting u with u_n and $E_A \subseteq E$ is the set of edges contained in them. By our previous observation we conclude that T' is a tree with root u. As the set $\{u_n \mid n \in \mathbb{N}\}$ is infinite, we have thus constructed an infinite, connected graph where every vertex has finite degree. König's Lemma, Proposition 2.71, states that T' contains an infinite, simple path starting at a. This is a contradiction as every path in T' starting at u leads to one of the u_n and is accordingly finite by construction. Thus our assumption must be wrong and there cannot be such a word u as required. $\square$

The previous theorem provides an easy way to check σ-additivity of a given map on $\mathcal{S}_\omega$ as we will see in the next corollary.

Corollary 4.6. *Let $\mathcal{A}$ be a finite alphabet. A non-negative map $\mu\colon \mathcal{S}_\omega \to \overline{\mathbb{R}}_+$ with $\mu(\emptyset) = 0$ is sigma-additive and hence a pre-measure iff for every $u \in \mathcal{A}^*$ the following* *holds:*

$$\mu(\mathbf{cone}(u)) = \mu\left(\bigcup_{a\in\mathcal{A}} \mathbf{cone}(ua)\right) = \sum_{a\in\mathcal{A}} \mu(\mathbf{cone}(ua))$$

Proof. It is obvious, that σ-additivity implies the validity of the given equality. In order to show that the given equality suffices to show σ-additivity, we consider a sequence of cones of the form $(\mathbf{cone}(u_n))_{1\leq n\leq N}$ where $N \in \mathbb{N}$, $N > 1$ ($N = 1$ is trivial!) and $\bigcup_{n=1}^{N} \mathbf{cone}(u_n) = \mathbf{cone}(u)$. We define $U_0 := \{u_1, ..., u_n\}$. Let $v = v_1...v_m \in \{v \in U_0 \mid |v| = \max_{1\leq n\leq N} |\{u_n\}|\}$. Since the cones of the form $\mathbf{cone}(u_n)$ are required to be mutually disjoint, none of the $u_n \in U \setminus \{v\}$ can be a prefix of v. However, all the cones of the form $\mathbf{cone}(v_1...v_{m-1}a)$ where $a \in \mathcal{A}$ are contained in $\mathbf{cone}(u)$ because $u \sqsubset v$ (we must have $u \neq v$ because $N > 1$) . Thus we must have $v_1...v_{m-1}a \in U$ for every $a \in \mathcal{A}$. We know that the following holds:

$$\mu(\mathbf{cone}(v_1...v_{m-1})) = \mu\left(\bigcup_{a\in\mathcal{A}} \mathbf{cone}(v_1...v_{m-1}a)\right) = \sum_{a\in\mathcal{A}} \mu(\mathbf{cone}(v_1...v_{m-1}a))$$

Therefore we may "contract" U_0 and define $U_1 = \{v_1...v_{m-1}\} \cup U_0 \setminus \{v_1...v_{m-1}a \mid a \in \mathcal{A}\}$. By applying the previous steps repeatedly, we obtain the desired σ-additivity of μ. $\square$

While in $(\mathcal{A}^*, \mathcal{P}(\mathcal{A}^*))$ everything is measurable, it is not that obvious which sets - besides the ω-cones themselves - are measurable in the sigma algebra generated by the ω-cones. The following remark will provide some insight into what sets are measurable:

Remark 4.7 (The Sigma Algebra of Cones)**.** *The following sets are measurable sets in $\sigma(\mathcal{S}_\omega)$:*

- *Single infinite words $u \in \mathcal{A}^\omega$: $\{u\} = \bigcap_{v\in\mathbf{pref}(u)} \mathbf{cone}(v)$.*
- *The set of all infinite words which contain a finite word $v \in \mathcal{A}^*$ at a specified position: $\{u \in \mathcal{A}^\omega \mid \exists v_1 \in \mathcal{A}^m : v_1 v \sqsubseteq u\} = \bigcup_{v_1\in\mathcal{A}^m} \mathbf{cone}(v_1 v)$.*
- *The set of all infinite words which contain a finite word $v \in \mathcal{A}^*$: $\{u \in \mathcal{A}^\omega \mid \exists v_1 \in \mathcal{A}^* : v_1 v \sqsubseteq u\} = \bigcup_{v_1\in\mathcal{A}^*} \mathbf{cone}(v_1 v)$.*

Proof. This follows immediately from the fact that the intersection and the unions are countable. $\square$

Proposition 4.8. *Let $\mathcal{A}$ be an alphabet. We consider the bijective functions:*

$$\phi\colon \mathcal{A}^\omega \to \mathcal{A} \times \mathcal{A}^\omega, \quad (u_n)_{n\in\mathbb{N}} \mapsto \big(u_1, (u_{n+1})_{n\in\mathbb{N}}\big)$$
$$\psi\colon \mathcal{A} \times \mathcal{A}^\omega \to \mathcal{A}^\omega, \quad (u,v) \mapsto \mathbf{conc}(u,v)$$

and define:

$$\begin{aligned}\mathcal{S}'_\omega := \phi(\mathcal{S}_\omega) &= \phi\left(\{\emptyset\} \cup \{\mathbf{cone}(u) \mid u \in \mathcal{A}^* \setminus \{\varepsilon\}\} \cup \{\mathbf{cone}(\varepsilon)\}\right)\\ &= \{\emptyset\} \cup \{\{a\} \times \mathbf{cone}(u) \mid a \in \mathcal{A}, u \in \mathcal{A}^*\} \cup \{\mathcal{A} \times \mathbf{cone}(\varepsilon)\}\}\end{aligned}$$

Then $\mathcal{S}'_\omega$ is a semi-ring of sets and the measurable spaces $(\mathcal{A}^\omega, \sigma(\mathcal{S}_\omega))$ and $(\mathcal{A} \times \mathcal{A}^\omega, \sigma(\mathcal{S}'_\omega))$ are isomorphic in **Meas**.

Proof. This follows immediately from Proposition 3.2. □

Proposition 4.9. *Let $\mathcal{A}$ be an alphabet. Then the following holds:*

$$\sigma(S'_\omega) = \mathcal{P}(\mathcal{A}) \otimes \sigma(\mathcal{S}_\omega)$$

Proof. We know from Example 2.12 that $\sigma(\{\emptyset\} \cup \{\{a\} \mid a \in \mathcal{A}\}) = \mathcal{P}(\mathcal{A})$ and conclude that $\mathcal{G} := \{\mathcal{A}\} \cup \{\emptyset\} \cup \{\{a\} \mid a \in \mathcal{A}\}$ fulfills $\mathcal{A} \in \mathcal{G}$ and $\sigma(\mathcal{G}) = \mathcal{P}(\mathcal{A})$. Due to the fact that $\mathcal{A}^\omega = \mathbf{cone}(\varepsilon) \in \mathcal{S}_\omega$ we can apply Proposition 2.44 and conclude that the following holds:

$$\sigma(\mathcal{G} * \mathcal{S}_\omega) = \mathcal{P}(\mathcal{A}) \otimes \sigma(\mathcal{S}_\omega)$$

Finally we show that $\sigma(\mathcal{S}'_\omega) = \sigma(\mathcal{G} * \mathcal{S}_\omega)$:

"$\subseteq$" We note that $\mathcal{S}'_\omega \subseteq \mathcal{G} * \mathcal{S}_\omega$. The monotonicity of the σ-operator yields $\sigma(\mathcal{S}'_\omega) \subseteq \sigma(\mathcal{G} * \mathcal{S}_\omega)$.

"$\supseteq$" Obviously we have $\emptyset \times \mathbf{cone}(u) = \emptyset \in \sigma(\mathcal{S}'_\omega)$ and $\{a\} \times \mathbf{cone}(u) \in \sigma(\mathcal{S}'_\omega)$ for every $a \in \mathcal{A}$ and every $u \in \mathcal{A}^*$. Furthermore the following holds for every $u \in \mathcal{A}^*$:

$$\mathcal{A} \times \mathbf{cone}(u) = \bigcup_{a\in\mathcal{A}} \underbrace{\{a\} \times \mathbf{cone}(u)}_{\in\sigma(\mathcal{S}'_\omega)} \in \sigma(\mathcal{S}'_\omega)$$

and thus we have shown that $\mathcal{G} * \mathcal{S}_\omega \subseteq \sigma(\mathcal{S}'_\omega)$. The monotonicity and the idempotence of the σ-operator yield:

$$\sigma(\mathcal{G} * \mathcal{S}_\omega) \subseteq \sigma\big(\sigma(\mathcal{S}'_\omega)\big) = \sigma(\mathcal{S}'_\omega)$$

$\square$

4.2 The Endofunctor $F = \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}}$ and PTS w/o Termination

4.2.1 The Endofunctor

Definition 4.10 (The Endofunctor $F = \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}}$)**.** *Let $\mathcal{A}$ be an alphabet. We define an endofunctor $F := \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}} + \mathbf{1} \colon \mathbf{Meas} \to \mathbf{Meas}$:*

- *For each measurable space (X, Σ) we define*

$$F\big((X,\Sigma)\big) := (\mathcal{A} \times X + \mathbf{1}, \sigma\big(\mathcal{P}(\mathcal{A}) * \Sigma)\big) = (\mathcal{A} \times X + \mathbf{1}, \mathcal{P}(\mathcal{A}) \otimes \Sigma)$$

- *For each measurable arrow $f \colon (X, \Sigma_X) \to (Y, \Sigma_Y)$, i.e. a Σ_X-Σ_Y measurable total function $f' \colon X \to Y$ we define $F(f)$ to be the arrow corresponding to the $\mathcal{P}(\mathcal{A}) \otimes \Sigma_X$-$\mathcal{P}(\mathcal{A}) \otimes \Sigma_Y$-measurable total function:*

$$\begin{aligned} F'(f) \colon \mathcal{A} \times X &\to \mathcal{A} \times Y \\ (a, x) &\mapsto (a, f'(x)) \end{aligned}$$

Proposition 4.11. *F as defined above is well-defined.*

Proof. We have to verify, that $F(f)'$ is $\mathcal{P}(\mathcal{A}) \otimes \Sigma_X$-$\mathcal{P}(\mathcal{A}) \otimes \Sigma_Y$-measurable. For arbitrary $S_{\mathcal{A}} \times S_Y \in \mathcal{P}(\mathcal{A}) * \Sigma_Y$ we calculate

$$F'(f)^{-1}(S_{\mathcal{A}} \times S_Y) = S_{\mathcal{A}} \times \underbrace{(f')^{-1}(S_Y)}_{\in \Sigma_X} \in \mathcal{P}(\mathcal{A}) * \Sigma_X \subseteq \mathcal{P}(\mathcal{A}) \otimes \Sigma_X$$

Checking that F is a functor is trivial. $\square$

4.2.2 A Distributive Law

Proposition and Definition 4.12 (A Distributive Law)**.** *Let $\mathcal{A}$ be an alphabet, $F = \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}} \colon \mathbf{Meas} \to \mathbf{Meas}$ and let $\mathbb{T} \in \{\mathbb{S}, \mathbb{P}\}$ be either the sub-probability or the probability monad. For every $X \in \mathbf{obj}(\mathbf{Meas})$ we define an arrow:*

$$\begin{aligned} \lambda_X \colon \mathcal{A} \times \mathbb{T}(X) &\to \mathbb{T}(\mathcal{A} \times X) \\ (a, P) &\mapsto \delta_a^{\mathcal{A}} \otimes P. \end{aligned}$$

Then $\lambda \colon F\mathbb{T} \to \mathbb{T}F$ *is a distributive law.*

Proof. We denote that $(\delta_a^{\mathcal{A}} \otimes P)(\mathcal{A} \times X) = P(X)$ holds for all $a \in \mathcal{A}$. Thus $\delta_a^{\mathcal{A}} \otimes P$ is a (sub-)probability measure on $\mathcal{A} \times X$ iff P is a (sub-)probability measure on X. Let us now show that λ is a natural transformation. For an arbitrary arrow $(f \colon X \to Y) \in$ **arr**(**Meas**) we consider the following diagram in **Meas**:

$$\begin{tikzcd} F\mathbb{T}X \arrow[r, "\lambda_X"] \arrow[d, "F\mathbb{T}(f)"'] & \mathbb{T}FX \arrow[d, "\mathbb{T}F(f)"] \\ F\mathbb{T}Y \arrow[r, "\lambda_Y"] & \mathbb{T}FY \end{tikzcd}$$

For arbitrary $(a, P) \in \mathcal{A} \times \mathbb{T}(X)$ and $S := S_{\mathcal{A}} \times S_Y \in \mathcal{P}(\mathcal{A}) * \Sigma_Y$ we calculate:

$$\begin{aligned} (\mathbb{T}F(f) \circ \lambda_X)(a, P)(S) &= (\delta_a^{\mathcal{A}} \otimes P)\left(F(f)^{-1}(S)\right) \\ &= (\delta_a^{\mathcal{A}} \otimes P)(S_{\mathcal{A}} \times f^{-1}(S_Y)) \\ &= \delta_a^{\mathcal{A}}(S_{\mathcal{A}}) \cdot P\left(f^{-1}(S_Y)\right) \\ &= \delta_a(S_{\mathcal{A}}) \cdot P_f(S_Y) \\ &= (\delta_a^{\mathcal{A}} \otimes P_f)(S_{\mathcal{A}} \times S_Y) \\ &= (\lambda_Y \circ F\mathbb{T}(f))(a, P)(S) \end{aligned}$$

Since (sub-)probability measures are finite measures and thus trivially σ-finite, we can use Corollary 2.23 and conclude that the diagram commutes for all arrows $f \in$ **arr**(**Meas**) and hence λ is a natural transformation. Analogously we consider the diagram:

$$\begin{tikzcd} FX \arrow[r, "F(\eta_X)"] \arrow[dr, "\eta_{FX}"'] & F\mathbb{T}X \arrow[d, "\lambda_X"] \\ & \mathbb{T}FX \end{tikzcd}$$

For arbitrary $(a, x) \in F(X)$ and $S := S_{\mathcal{A}} \times S_X \in \mathcal{P}(\mathcal{A}) * \Sigma_X$ we calculate:

$$\begin{aligned} \eta_{F(X)}(a, x)(S) &= \delta_{(a,x)}^{\mathcal{A} \times X}(S_{\mathcal{A}} \times S_X) \\ &= \delta_a^{\mathcal{A}}(S_{\mathcal{A}}) \cdot \delta_x^X(S_X) \\ &= (\delta_a^{\mathcal{A}} \otimes \delta_x^X)(S) \\ &= \lambda_X(a, \delta_x^X) = \lambda_X \circ F(\eta_X)(a, x) \end{aligned}$$

As above, Corollary 2.23 yields commutativity of the diagram. Finally we consider the diagram:

$$\begin{tikzcd}
F\mathbb{T}^2X \arrow[r, "\lambda_{\mathbb{T}(X)}"] \arrow[d, "F(\mu_X)"'] & \mathbb{T}F\mathbb{T}X \arrow[r, "\mathbb{T}(\lambda_X)"] & \mathbb{T}^2FX \arrow[d, "\mu_{F(X)}"] \\
F\mathbb{T}X \arrow[rr, "\lambda_X"'] & & \mathbb{T}FX
\end{tikzcd}$$

For arbitrary $(a, P) \in F\mathbb{T}^2X = \mathcal{A}\times\mathbb{T}(\mathbb{T}(X))$ and $S = S_\mathcal{A}\times S_X \in \mathcal{P}(\mathcal{A})*\Sigma_X$ we calculate:

$$\begin{aligned}
(\lambda_X \circ F(\mu_X))\,(a,P)(S) &= (\lambda_X\,(a, \mu_X(P)))\,(S) \\
&= \left(\delta_a^\mathcal{A} \otimes \mu_X(P)\right)(S) \\
&= \delta_a^\mathcal{A}(S_\mathcal{A}) \cdot (\mu_X(P))\,(S_X) \\
&= \delta_a^\mathcal{A}(S_\mathcal{A}) \cdot \int p_{S_X}\,\mathrm{d}P
\end{aligned}$$

and:

$$\begin{aligned}
\left(\mu_{FX} \circ \mathbb{T}(\lambda_X) \circ \lambda_{\mathbb{T}(X)}\right)(a,P)(S) &= \mu_{FX}\left(\left(\delta_a^\mathcal{A} \otimes P\right)_{\lambda_X}\right)(S) \\
&= \int_{\mathbb{T}FX} p_S\,\mathrm{d}\left(\delta_a^\mathcal{A} \otimes P\right)_{\lambda_X} \\
&\overset{(1)}{=} \int_{\lambda_X^{-1}(\mathbb{T}FX)} p_S \circ \lambda_X\,\mathrm{d}\left(\delta_a^\mathcal{A} \otimes P\right) \\
&\overset{(2)}{=} \int_{\{a\}\times\mathbb{T}(X)} p_S \circ \lambda_X\,\mathrm{d}\left(\delta_a^\mathcal{A} \otimes P\right) \\
&\overset{(3)}{=} \int_{P'\in\mathbb{T}(X)} (\delta_a^\mathcal{A} \otimes P')(S)\,\mathrm{d}P(P') \\
&= \int_{P'\in\mathbb{T}(X)} \delta_a^\mathcal{A}(S_\mathcal{A}) \cdot P'(S_X)\,\mathrm{d}P(P') \\
&= \delta_a^\mathcal{A}(S_\mathcal{A}) \cdot \int p_{S_X}\,\mathrm{d}P
\end{aligned}$$

where we have used:

(1) Proposition 2.39.

(2) The fact that $(\mathcal{A} \setminus \{a\}) \times \mathbb{T}(X)$ is a null set for $\delta_a^\mathcal{A} \otimes P$ in conjunction with Corollary 2.38.

(3) $(p_S \circ \lambda_X)(a, P') = (\delta_a^\mathcal{A} \otimes P')(S)$ and $\forall \mathcal{M} \in \Sigma_{\mathbb{T}(X)} : (\delta_a^\mathcal{A} \otimes P)(\{a\} \times \mathcal{M}) = P(\mathcal{M})$.

As before, commutativity of the diagram is a result of Corollary 2.23. $\square$

4.2.3 PTS w/o Termination as Coalgebra

Although we have defined probabilistic transition systems with the transition function α consisting of sub-probability measures, we are now going to restrict our considerations to systems where α is made up of probability measures as we are only able to prove the final coalgebra theorem for such systems.

Remark 4.13 (PTS without Termination as Coalgebra)**.** *Let $\mathcal{A}$ be an alphabet. We consider the probability monad $(\mathbb{P}, \eta, \mu)$ and the endofunctor $F := \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}}$ on* **Meas** *together with the lifting $\overline{F}\colon \mathcal{K}\ell(\mathbb{P}) \to \mathcal{K}\ell(\mathbb{P})$ induced by the distributive law $\lambda\colon F\mathbb{P} \to \mathbb{P}F$. A probabilistic transition system without explicit termination $((X, \Sigma), \mathcal{A}, \alpha)$, where α comprises merely of probability measures can be expressed as the $\overline{F}$-coalgebra $((X, \Sigma)_{\mathbb{P}}, \alpha^\flat)$.*

Proof. This follows immediately from the definitions of a PTS without termination and a coalgebra in a Kleisli category. □

4.2.4 A Final Coalgebra and Trace Semantics for PTS w/o Termination

Theorem 4.14 (Final Coalgebra)**.** *Let $\mathcal{A}$ be a finite alphabet. We consider the probability monad $(\mathbb{P}, \eta, \mu)$ and the endofunctor $F := \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}}$ on* **Meas** *together with the lifting $\overline{F}\colon \mathcal{K}\ell(\mathbb{P}) \to \mathcal{K}\ell(\mathbb{P})$ induced by the distributive law $\lambda\colon F\mathbb{P} \to \mathbb{P}F$. A final $\overline{F}$-coalgebra is the tuple $\big((\mathcal{A}^w, \sigma(\mathcal{S}_\omega))_{\mathbb{P}}, \kappa^\flat\big)$ where $\kappa^\flat\colon \mathcal{A}^\omega_{\mathbb{P}} \to \overline{F}\left(\mathcal{A}^\omega_{\mathbb{P}}\right)$ is the induced isomorphism associated to:*

$$\kappa := \eta_{F((\mathcal{A}^\omega, \sigma(\mathcal{S}_\omega)))} \circ \phi\colon \mathcal{A}^\omega \to \mathbb{P}(\mathcal{A} \times \mathcal{A}^\omega), \quad u \mapsto \delta^{\mathcal{A}\times\mathcal{A}^\omega}_{\phi(u)}$$

Given an arbitrary $\overline{F}$-coalgebra $\big((X, \Sigma)_{\mathbb{P}}, \alpha^\flat\big)$, the unique arrow mapping this coalgebra to the final coalgebra is $\mathbf{tr}^\flat\colon X_{\mathbb{P}} \to \mathcal{A}^\omega_{\mathbb{P}}$ associated to $\mathbf{tr}\colon X \to \mathbb{P}(\mathcal{A}^\omega)$ where for every $x \in X$ the probability measure $\mathbf{tr}(x)\colon \sigma(\mathcal{S}_\omega) \to [0,1]$ is the uniquely determined extension of the finite (and thus σ-finite) pre-measure $\mathbf{tr}(x)\colon \mathcal{S}_\omega \to [0,1]$ defined by the following equations:

$$\begin{aligned}
\mathbf{tr}(x)(\emptyset) &= 0 \\
\mathbf{tr}(x)(\mathbf{cone}(\varepsilon)) &= 1 \\
\mathbf{tr}(x)(\mathbf{cone}(au)) &= \int \mathbf{tr}_u \,\mathrm{d}P_{x,a} \quad \forall a \in \mathcal{A}, \forall u \in \mathcal{A}^*
\end{aligned}$$

where for every $a \in \mathcal{A}$:

$$P_{x,a}\colon \Sigma_X \to [0,1], \quad P_{x,a}(S) := \alpha(x)(\{a\} \times S)$$

and for every $u \in \mathcal{A}^*$:

$$\mathbf{tr}_u \colon X \to [0,1], \quad x \mapsto \mathbf{tr}(x)(\mathbf{cone}(u)).$$

Proof. First we verify, that the map **tr** defined by the given equations is indeed a finite (and thus σ) finite pre-measure. Using Corollary 4.6 we denote that for **tr** to be a pre-measure, it is sufficient to show σ-additivity of **tr**: Let $u = a_1...a_m \in \mathcal{A}^*$, then:

$$\begin{aligned}
\mathbf{tr}(x)\big(\mathbf{cone}(u)\big) &= \int \mathbf{tr}_{a_2...a_m}\,\mathrm{d}P_{x,a_1} \\
&= \int\limits_{x_1 \in X} \mathbf{tr}(x_1)\big(\mathbf{cone}(a_2...a_m)\big)\,\mathrm{d}P_{x,a_1}(x_1) \\
&= \int\limits_{x_1 \in X} \left(\int \mathbf{tr}_{a_3...a_m}\,\mathrm{d}P_{x_1,a_2}\right) \mathrm{d}P_{x,a_1}(x_1) \\
&= \int\limits_{x_1 \in X} \left(\int\limits_{x_2 \in X} \mathbf{tr}(x_2)\big(\mathbf{cone}(a_3...a_m)\big)\,\mathrm{d}P_{x_1,a_2}(x_2)\right) \mathrm{d}P_{x,a_1}(x_1) \\
&= \int\limits_{x_1 \in X} \dots \int\limits_{x_m \in X} \underbrace{\mathbf{tr}(x_m)\big(\mathbf{cone}(\varepsilon)\big)}_{=1}\,\mathrm{d}P_{x_{m-1},a_m}(x_m)\dots\mathrm{d}P_{x,a_1}(x_1) \\
&= \int\limits_{x_1 \in X} \dots \int\limits_{x_m \in X} 1\,\mathrm{d}P_{x_{m-1},a_m}(x_m)\dots\mathrm{d}P_{x,a_1}(x_1)
\end{aligned}$$

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

$$\begin{aligned}
\mathbf{tr}(x)\big(\mathbf{cone}(ub)\big) &= \int\limits_{x_1 \in X} \dots \int\limits_{x_m \in X} \underbrace{\mathbf{tr}(x_m)\big(\mathbf{cone}(b)\big)}_{=P_{x_m,b}(X)}\,\mathrm{d}P_{x_{m-1},a_m}(x_m)\dots\mathrm{d}P_{x,a_1}(x_1) \\
&= \int\limits_{x_1 \in X} \dots \int\limits_{x_m \in X} P_{x_m,b}(X)\,\mathrm{d}P_{x_{m-1},a_m}(x_m)\dots\mathrm{d}P_{x,a_1}(x_1).
\end{aligned}$$

All integrals exist and are bounded above by 1 and we can use the linearity of the integral to exchange the finite sum and the integrals to obtain:

$$\begin{aligned}
\sum_{b\in\mathcal{A}} \mathbf{tr}(x)\big(\mathbf{cone}(ub)\big) &= \sum_{b\in\mathcal{A}} \int\limits_{x_1\in X} \dots \int\limits_{x_m\in X} P_{x_m,b}(X)\,\mathrm{d}P_{x_{m-1},a_m}(x_m)\dots\mathrm{d}P_{x,a_1}(x_1) \\
&= \int\limits_{x_1\in X} \dots \int\limits_{x_m\in X} \left(\sum_{b\in\mathcal{A}} P_{x_m,b}(X)\right)\mathrm{d}P_{x_{m-1},a_m}(x_m)\dots\mathrm{d}P_{x,a_1}(x_1) \\
&= \int\limits_{x_1\in X} \dots \int\limits_{x_m\in X} \left(\sum_{b\in\mathcal{A}} \alpha(x)(\{b\}\times X)\right)\mathrm{d}P_{x_{m-1},a_m}(x_m)\dots\mathrm{d}P_{x,a_1}(x_1) \\
&= \int\limits_{x_1\in X} \dots \int\limits_{x_m\in X} \underbrace{\big(\alpha(x)(\mathcal{A}\times X)\big)}_{=1}\mathrm{d}P_{x_{m-1},a_m}(x_m)\dots\mathrm{d}P_{x,a_1}(x_1) \\
&= \int\limits_{x_1\in X} \dots \int\limits_{x_m\in X} 1\,\mathrm{d}P_{x_{m-1},a_m}(x_m)\dots\mathrm{d}P_{x,a_1}(x_1) \\
&= \mathbf{tr}(x)\big(\mathbf{cone}(u)\big)
\end{aligned}$$

Hence $\mathbf{tr}(x)$ is σ-additive and since $\mathbf{tr}(x)(\mathcal{A}^\omega) = 1 < \infty$ it is a σ-finite pre-measure which we can extend (uniquely!) to a measure.
In order to show that the given tuple is a final coalgebra, let $\big((X,\Sigma)_{\mathbb{P}}, \alpha^\flat\big)$ be an arbitrary $\overline{F}$-coalgebra. We have to show that $\mathbf{tr}^\flat\colon (X,\Sigma)_{\mathbb{P}} \to (\mathcal{A}^\omega, \sigma(\mathcal{S}_\omega))_{\mathbb{P}}$ is the uniquely determined arrow such that the following diagram commutes:

$$\begin{tikzcd}
(X,\Sigma)_{\mathbb{P}} \arrow[r, "\alpha^\flat"] \arrow[d, "\mathbf{tr}^\flat"'] & (\mathcal{A}\times X, \mathcal{P}(\mathcal{A})\otimes\Sigma)_{\mathbb{P}} \arrow[d, "\overline{F}(\mathbf{tr}^\flat)"] \\
(\mathcal{A}^\omega, \sigma(\mathcal{S}_\omega))_{\mathbb{P}} \arrow[r, "\kappa^\flat", "\cong"'] & (\mathcal{A}\times\mathcal{A}^\omega, \mathcal{P}(\mathcal{A})\otimes\sigma(\mathcal{S}_\omega))_{\mathbb{P}}
\end{tikzcd}$$

We define the following arrows:

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

such that commutativity is equivalent to $g^\flat = h^\flat$ and conclude

$$\begin{aligned}
& g^\flat = h^\flat \\
\Leftrightarrow \quad & g = h \\
\Leftrightarrow \quad & g(x) = h(x) \quad \forall x \in X \\
\Leftrightarrow \quad & g(x)(S) = h(x)(S) \quad \forall x \in X, \forall S \in \mathcal{P}(\mathcal{A}) \otimes \sigma(\mathcal{S}_\omega) \\
\Leftrightarrow \quad & g(x)(S) = h(x)(S) \quad \forall x \in X, \forall S \in \sigma(\mathcal{S}'_\omega) \\
\Leftrightarrow \quad & g(x)(S) = h(x)(S) \quad \forall x \in X, \forall S \in \mathcal{S}'_\omega
\end{aligned}$$

using for the last step that $\mathcal{S}'_\omega$ is a semi-ring of sets, the fact that for all $x \in X$ both $g(x)$ and $h(x)$ are finite and thus σ-finite measures since they are probability measures and Corollary 2.23. For arbitrary $x \in X$ and $S \in \mathcal{S}'_\omega$ we calculate $g(x)(S)$ as follows:

$$\begin{aligned}
g(x)(S) &= (\mu_{\mathcal{A}\times\mathcal{A}^\omega} \circ \mathbb{P}(\kappa) \circ \mathbf{tr})(x)(S) \\
&= \mu_{\mathcal{A}\times\mathcal{A}^\omega} \left(\mathbb{P}(\kappa)(\mathbf{tr}(x))\right)(S) \\
&= \mu_{\mathcal{A}\times\mathcal{A}^\omega} \left(\mathbf{tr}(x)_\kappa\right)(S) \\
&= \int p_S \,\mathrm{d}\mathbf{tr}(x)_\kappa \\
&\stackrel{(1)}{=} \int p_S \circ \kappa \,\mathrm{d}\mathbf{tr}(x) \\
&\stackrel{(2)}{=} \int \chi_{\psi(S)} \,\mathrm{d}\mathbf{tr}(x) = \mathbf{tr}(x)(\psi(S))
\end{aligned}$$

In (1) we used Proposition 2.39 and in (2) the fact that for every S as above and for every $u \in \mathcal{A}^*$ the following holds:

$$p_S \circ \kappa(u) = \delta_{\phi(u)}^{\mathcal{A}\times\mathcal{A}^\omega}(S) = \chi_S(\phi(u)) = \chi_{\psi(S)}(u)$$

Hence we get the following set of equations:

$$\begin{aligned}
g(x)(\emptyset) &= \mathbf{tr}(x)(\emptyset) \\
g(x)(\mathcal{A} \times \mathbf{cone}(\varepsilon)) &= \mathbf{tr}(x)(\mathbf{cone}(\varepsilon)) \\
g(x)(\{a\} \times \mathbf{cone}(u)) &= \mathbf{tr}(x)(\mathbf{cone}(au)) \quad \forall a \in \mathcal{A},\ \forall u \in \mathcal{A}^*
\end{aligned}$$

In order to calculate $h(x)(S)$, let:

$$\rho := \lambda_{\mathcal{A}\times\mathcal{A}^\omega} \circ (\mathrm{id}_{\mathcal{A}} \times \mathbf{tr})\colon \mathcal{A} \times X \to \mathbb{P}(\mathcal{A} \times \mathcal{A}^\omega)$$

and for every $S \in \mathcal{S}'_\omega$ let:

$$\rho_S := p_S \circ \rho \colon \mathcal{A} \times X \to [0,1], \quad (a,x) \mapsto \big(\rho(a,x)\big)(S)$$

Then we obtain the following result:

$$\begin{aligned}
h(x)(S) &= (\mu_{\mathcal{A}\times\mathcal{A}^\omega} \circ \mathbb{P}(\rho) \circ \alpha)(x)(S) \\
&= \mu_{\mathcal{A}\times\mathcal{A}^\omega} \left(\mathbb{P}(\rho)(\alpha(x))\right)(S) \\
&= \mu_{\mathcal{A}\times\mathcal{A}^\omega} \left(\alpha(x)_\rho\right)(S) \\
&= \int p_S \,\mathrm{d}\alpha(x)_\rho \\
&\overset{(1)}{=} \int p_S \circ \rho \,\mathrm{d}\alpha(x) \\
&= \int \rho_S \,\mathrm{d}\alpha(x)
\end{aligned}$$

where we have used Proposition2.39 in (1). Obviously we have $h(x)(\emptyset) = 0$. For $(b,x') \in \mathcal{A} \times X$ the following holds:

$$\begin{aligned}
\rho_{\mathcal{A}\times\mathbf{cone}(\varepsilon)}(b,x') &= \big(\delta_b^{\mathcal{A}} \otimes \mathbf{tr}(x')\big)(\mathcal{A} \times \mathbf{cone}(\varepsilon)) \\
&= \delta_b^{\mathcal{A}}(\mathcal{A}) \cdot \mathbf{tr}(x')\big(\mathbf{cone}(\varepsilon)\big) \\
&= \mathbf{tr}(x')\big(\mathbf{cone}(\varepsilon)\big) \\
&= \mathbf{tr}_\varepsilon(x') \\
&= \left(\mathbf{tr}_\varepsilon \circ \pi_2\right)(b,x')
\end{aligned}$$

where $\pi_2 \colon \mathcal{A} \times X \to X$ is the canonical projection of the product and hence:

$$h(x)(\mathcal{A} \times \mathbf{cone}(\varepsilon)) = \int \mathbf{tr}_\varepsilon \circ \pi_2 \,\mathrm{d}\alpha(x)$$

For $a \in \mathcal{A}, u \in \mathcal{A}^\omega, (b,x') \in \mathcal{A} \times X$ the following holds:

$$\begin{aligned}
\rho_{\{a\}\times\mathbf{cone}(u)}(b,x') &= \big(\delta_b^{\mathcal{A}} \otimes \mathbf{tr}(x')\big)\left(\{a\} \times \mathbf{cone}(u)\right) \\
&= \delta_b^{\mathcal{A}}(\{a\}) \cdot \mathbf{tr}(x')\big(\mathbf{cone}(u)\big) \\
&= \chi_{\{a\}}(b) \cdot \mathbf{tr}_u(x') \\
&= \chi_{\{a\}\times X}(b,x') \cdot (\mathbf{tr}_u \circ \pi_2)(b,x')
\end{aligned}$$

and hence:

$$\begin{aligned} h(x)\left(\{a\} \times \mathbf{cone}(u)\right) &= \int \chi_{\{a\}\times X} \cdot (\mathbf{tr}_u \circ \pi_2)\, \mathrm{d}\alpha(x) \\ &= \int_{\{a\}\times X} \mathbf{tr}_u \circ \pi_2 \, \mathrm{d}\alpha(x) \\ &= \int_X \mathbf{tr}_u \, \mathrm{d}P_{x,a} \end{aligned}$$

Thus we have shown that the commutativity of the diagram is equivalent to the following set of equations:

$$\begin{aligned} \mathbf{tr}(x)(\emptyset) &= 0 \\ \mathbf{tr}(x)(\mathbf{cone}(\varepsilon)) &= \int \mathbf{tr}_\varepsilon \circ \pi_2 \, \mathrm{d}\alpha(x) \quad \forall x \in X \\ \mathbf{tr}(x)(\mathbf{cone}(au)) &= \int \mathbf{tr}_u \, \mathrm{d}P_{x,a} \quad \forall x \in X, \forall a \in \mathcal{A}, \forall u \in \mathcal{A}^* \end{aligned}$$

We denote that $\mathbf{tr}$ as defined in the proposition fulfills these equations because:

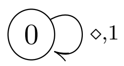
$$\int \mathbf{tr}_\varepsilon \circ \pi_2 \, \mathrm{d}\alpha(x) = \int 1 \, \mathrm{d}\alpha(x) = \alpha(x)(\mathcal{A} \times X) = 1$$

where the last equality follows from the fact that $\alpha(x)$ is a probability measure on $\mathcal{A}\times X$. Furthermore $\mathbf{tr}$ is the is uniquely defined by these equations because $\mathbf{tr}(x)(\mathbf{cone}(\varepsilon)) = \mathbf{tr}(x)(\mathcal{A}^\omega) = 1$ is a necessary requirement for every probability measure. □

Remark 4.15 (Trace Semantics for PTS without Termination)**.** *Let $((X, \Sigma), \mathcal{A}, \alpha)$ be a probabilistic transition system without termination with finite alphabet $\mathcal{A}$ such that α comprises merely of probability measures. For every $x \in X$ the uniquely defined probability measure $\mathbf{tr}(x)$ as defined in Theorem 4.14 is called the* trace *of x.*

4.2.5 Examples

Example 4.16. *Let $X = \{0\}$ and $\mathcal{A}$ be a finite alphabet such that $\diamond \in \mathcal{A}$. We define $\alpha(0) := \delta^{\mathcal{A}\times\{0\}}_{(\diamond,0)}$. The graphical representation of the resulting transition system $((\{0\}, \{\emptyset, \{0\}\}), \alpha)$ is*



We denote that $P_{0,a} = \delta_{\diamond}^{\mathcal{A}}(\{a\})$ *and* $\mathbf{tr}(0)(\mathbf{cone}(\diamond^n)) = \int \mathbf{tr}_{\diamond^{n-1}} \,\mathrm{d}\delta_{(\diamond,0)}^{\mathcal{A}\times\{0\}} = 1$ *and calculate the trace map:*

$$
\begin{aligned}
\mathbf{tr}(0)(\{\diamond^\omega\}) &= \mathbf{tr}(x)\left(\bigcap_{n\in\mathbb{N}_0} \mathbf{cone}(\diamond^n)\right) \\
&= \mathbf{tr}(0)\left(\mathcal{A}^\omega \setminus \left(\bigcup_{n\in\mathbb{N}_0} (\mathcal{A}^\omega \setminus \mathbf{cone}(\diamond^n))\right)\right) \\
&= \mathbf{tr}(0)(\mathcal{A}^\omega) - \mathbf{tr}(0)\left(\bigcup_{n\in\mathbb{N}_0} (\mathcal{A}^\omega \setminus \mathbf{cone}(u))\right) \\
&\geq 1 - \sum_{n\in\mathbb{N}_0} \mathbf{tr}(0)\left(\mathcal{A}^\omega \setminus \mathbf{cone}(\diamond^n)\right) \\
&\geq 1 - \sum_{n\in\mathbb{N}_0} \left(\mathbf{tr}(0)(\mathcal{A}^\omega) - \mathbf{tr}(0)(\mathbf{cone}(\diamond^n))\right) \\
&= 1 - \sum_{n\in\mathbb{N}_0} (1-1) = 1
\end{aligned}
$$

and $\mathbf{tr}(0)(\mathcal{A}^\omega \setminus \{\diamond^\omega\}) = 1 - 1 = 0$. *We obtain:*

$$\mathbf{tr}(x_0) = \delta_{\diamond^\omega}^{\mathcal{A}^\omega}.$$

Example 4.17. *Let* $X := \mathbb{N}_0$ *and* $\mathcal{A}$ *be a finite alphabet such that* $\diamond \in \mathcal{A}$. *We define* $\alpha(n) := \delta_{(\diamond,n+1)}^{\mathcal{A}\times\mathbb{N}}$ *for every* $n \in \mathbb{N}_0$ *and consider the probabilistic transition system* $\big((\mathbb{N}_0, \mathcal{P}(\mathbb{N}_0)), \alpha\big)$ *which has the following graphical representation:*

$$
\begin{tikzcd}
0 \arrow[r, "{\diamond,1}"] & 1 \arrow[r, "{\diamond,1}"] & 2 \arrow[r, "{\diamond,1}"] & 3 \arrow[r, "{\diamond,1}"] & 4 \arrow[r, "{\diamond,1}"] & 5 \arrow[r, "{\diamond,1}"] & 6 \arrow[r, "{\diamond,1}"] & \cdots
\end{tikzcd}
$$

We denote that $P_{n,a} = \delta_{\diamond}^{\mathcal{A}}(\{a\}) \cdot \delta_{n+1}^{\mathbb{N}_0}$ *and claim that* $\mathbf{tr}(m)(\mathbf{cone}(\diamond^n)) = 1$ *for all* $m, n \in \mathbb{N}_0$. *(This can easily be verified by induction.) As in the previous example, we conclude that for all* $m \in \mathbb{N}_0$ *we have* $\mathbf{tr}(m) = \delta_{\diamond^\omega}^{\mathcal{A}^\omega}$.

4.3 The Endofunctor $F = \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}} + \mathbf{1}$ and PTS with Termination

4.3.1 The Endofunctor

Definition 4.18 (The Endofunctor $F = \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}} + \mathbf{1}$)**.** *Let $\mathcal{A}$ be an alphabet. We define an endofunctor $F := \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}} + \mathbf{1} \colon \mathbf{Meas} \to \mathbf{Meas}$:*

- *For each measurable space (X, Σ) we define*

$$F\big((X,\Sigma)\big) := \big(\mathcal{A} \times X + \mathbf{1}, \sigma\big(\mathcal{P}(\mathcal{A}) * \Sigma \oplus \mathcal{P}(\mathbf{1})\big)$$

- *For each measurable arrow $f \colon (X, \Sigma_X) \to (Y, \Sigma_Y)$, i.e. a Σ_X-Σ_Y measurable total function $f' \colon X \to Y$ we define $F(f)$ to be the arrow corresponding to the $\sigma\big(\mathcal{P}(\mathcal{A}) * \Sigma_X \oplus \mathcal{P}(\mathbf{1})\big)$-$\sigma\big(\mathcal{P}(\mathcal{A}) * \Sigma_Y \oplus \mathcal{P}(\mathbf{1})\big)$-measurable total function:*

$$\begin{aligned} F'(f) \colon \mathcal{A} \times X + \mathbf{1} &\to \mathcal{A} \times Y + \mathbf{1} \\ (a, x) &\mapsto (a, f'(x)) \\ \checkmark &\mapsto \checkmark \end{aligned}$$

Proposition 4.19. *F as defined above is well-defined.*

Proof. We have to verify, that $F(f)'$ is $\sigma(\mathcal{P}(\mathcal{A}) * \Sigma_X \oplus \mathcal{P}(\mathbf{1})$-$\sigma(\mathcal{P}(\mathcal{A}) * \Sigma_Y \oplus \mathcal{P}(\mathbf{1})$-measurable. For arbitrary $S_{\mathcal{A}} \times S_Y + S_{\mathbf{1}} \in \mathcal{P}(\mathcal{A}) * \Sigma_Y \oplus \mathcal{P}(\mathbf{1})$ we calculate

$$F'(f)^{-1}(S_{\mathcal{A}} \times S_Y + S_{\mathbf{1}}) = S_{\mathcal{A}} \times \underbrace{(f')^{-1}(S_Y)}_{\in \Sigma_X} + S_{\mathbf{1}} \in \mathcal{P}(\mathcal{A}) * \Sigma_X \oplus \mathcal{P}(\mathbf{1}) \subseteq \sigma\big(\mathcal{P}(\mathcal{A}) * \Sigma_X \oplus \mathcal{P}(\mathbf{1})\big)$$

Checking that F is a functor is trivial. □

Definition 4.20. *Let (X, Σ_X, μ), (Y, Σ_Y, ν) be measure spaces, (Z, Σ_Z) be a measurable space. We define a pre-measure:*

$$\begin{aligned} \mu \odot \nu \colon \Sigma_X * \Sigma_Y \oplus \Sigma_Z &\to \overline{\mathbb{R}}_+ \\ S_X \times S_Y + S_1 &\mapsto \mu(S_X) \cdot \nu(S_Y) \end{aligned}$$

Proposition 4.21. *Let in Definition 4.20 μ and ν be σ-finite measures. Then $\mu \odot \nu$ can uniquely be extended to a σ-finite measure on $\sigma\big(\Sigma_X * \Sigma_Y \oplus \Sigma_Z\big)$.*

Proof. Recall from Corollary 2.2.2 that $\Sigma_X * \Sigma_Y \oplus \Sigma_Z$ is a semi-ring of sets. Due to the fact that μ and ν are σ-finite, we have sequences $(S_X^n)_{n\in\mathbb{N}}$ and $(S_Y^n)_{n\in\mathbb{N}}$ such that for all $n \in \mathbb{N}$ we have $S_X^n \in \Sigma_X$, $S_Y^n \in \Sigma_Y$, $\mu(S_X) < \infty$, $\nu(S_Y) < \infty$, $\bigcup_{n\in\mathbb{N}} S_X^n = X$ and $\bigcup_{n\in\mathbb{N}} S_Y^n = Y$. Moreover, for all $m, n \in \mathbb{N}$ we obviously have $S_X^m \times S_Y^n + Z \in \Sigma_X * \Sigma_Y \oplus \Sigma_Z$, $\mu \odot \nu(S_X^n \times S_Y^n + Z) = \mu(S_X^n) \cdot \nu(S_Y^n) < \infty$ and $\bigcup_{(m,n)\in\mathbb{N}^2} S_X^m \times S_Y^n + Z = X \times Y + Z$. Thus $\mu \odot \nu$ as defined in Definition 4.20 is a σ-finite pre-measure. Hence we can apply the extension theorem, Theorem 2.22, and the proof is complete. □

4.3.2 A Distributive Law

Proposition and Definition 4.22 (A Distributive Law)**.** *Let $\mathcal{A}$ be an alphabet, $F = \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}} + \mathbf{1} \colon \mathbf{Meas} \to \mathbf{Meas}$ and let $\mathbb{T} \in \{\mathbb{S}, \mathbb{P}\}$ be either the sub-probability or the probability monad. For every $X \in \mathbf{obj}(\mathbf{Meas})$ we define an arrow:*

$$\begin{aligned} \lambda_X \colon \mathcal{A} \times \mathbb{T}(X) + \mathbf{1} &\to \mathbb{T}(\mathcal{A} \times X + \mathbf{1}) \\ \checkmark &\mapsto \delta_{\checkmark}^{\mathcal{A}\times X+\mathbf{1}} \\ (a, P) &\mapsto \delta_a^{\mathcal{A}} \odot P. \end{aligned}$$

Then $\lambda \colon F\mathbb{T} \Rightarrow \mathbb{T}F$ is a distributive law.

Although the proof is quite similar to the proof of Proposition 4.12, we will give it for reasons of completeness.

Proof. We denote that $(\delta_a^{\mathcal{A}} \odot P)(\mathcal{A} \times X + \mathbf{1}) = P(X)$ holds for all $a \in \mathcal{A}$. Thus $\delta_a^{\mathcal{A}} \odot P$ is a (sub-)probability measure on $\mathcal{A} \times X + \mathbf{1}$ iff P is a (sub-)probability measure on X. Let us now show that λ is a natural transformation. For an arbitrary arrow $(f \colon X \to Y) \in \mathbf{arr}(\mathbf{Meas})$ we consider the following diagram in $\mathbf{Meas}$:

$$\begin{tikzcd} F\mathbb{T}(X) \arrow[r, "\lambda_X"] \arrow[d, "F\mathbb{T}(f)"'] & \mathbb{T}F(X) \arrow[d, "\mathbb{T}F(f)"] \\ F\mathbb{T}(Y) \arrow[r, "\lambda_Y"] & \mathbb{T}F(Y) \end{tikzcd}$$

For arbitrary $(a, P) \in \mathcal{A} \times \mathbb{T}(X)$ and $S := S_{\mathcal{A}} \times S_Y \times S_{\mathbf{1}} \in \mathcal{P}(\mathcal{A}) * \Sigma_Y \oplus \mathcal{P}(\mathbf{1})$ we calculate:

$$\begin{aligned}
(\mathbb{T}F(f) \circ \lambda_X)(a, P)(S) &= (\delta_a^{\mathcal{A}} \odot P) \left(F(f)^{-1}(S)\right) \\
&= (\delta_a^{\mathcal{A}} \odot P)(S_{\mathcal{A}} \times f^{-1}(S_Y) + S_{\mathbf{1}}) \\
&= \delta_a^{\mathcal{A}}(S_{\mathcal{A}}) \cdot P\left(f^{-1}(S_Y)\right) \\
&= \delta_a^{\mathcal{A}}(S_{\mathcal{A}}) \cdot P_f(S_Y) \\
&= (\delta_a^{\mathcal{A}} \odot P_f)(S_{\mathcal{A}} \times S_Y + S_{\mathbf{1}}) \\
&= (\lambda_Y \circ F\mathbb{T}(f))(a, P)(S)
\end{aligned}$$

and analogously:

$$\begin{aligned}
(\mathbb{T}F(f) \circ \lambda_X)(\checkmark)(S) &= \delta_{\checkmark}^{\mathcal{A}\times X+\mathbf{1}} \left(F(f)^{-1}(S)\right) \\
&= \delta_{\checkmark}^{\mathcal{A}\times X+\mathbf{1}} \left(S_{\mathcal{A}} \times f^{-1}(S_Y) + S_{\mathbf{1}}\right) \\
&= \delta_{\checkmark}^{\mathcal{A}\times Y+\mathbf{1}}(S) \\
&= (\lambda_Y \circ F\mathbb{T}(f))(\checkmark)(S)
\end{aligned}$$

Since (sub-)probability measures are finite measures and thus trivially σ-finite, we can use Corollary 2.23 and conclude that the diagram commutes for all arrows $f \in \mathbf{arr}(\mathbf{Meas})$ and hence λ is a natural transformation. Analogously we consider the diagram:

$$\begin{tikzcd}
FX \arrow[r, "F(\eta_X)"] \arrow[dr, "\eta_{FX}"'] & F\mathbb{T}X \arrow[d, "\lambda_X"] \\
& \mathbb{T}FX
\end{tikzcd}$$

We immediately obtain:

$$\eta_{FX}(\checkmark) = \delta_{\checkmark}^{FX} = \lambda_X(\checkmark) = \lambda_X\left(F(\eta_X)(\checkmark)\right) = \left(\lambda_X \circ F(\eta_X)\right)(\checkmark)$$

and for arbitrary $(a, x) \in F(X)$ and $S := S_{\mathcal{A}} \times S_X \times S_{\mathbf{1}} \in \mathcal{P}(\mathcal{A}) * \Sigma_X \oplus \mathcal{P}(\mathbf{1})$ we calculate:

$$\begin{aligned}
\eta_{F(X)}(a, x)(S) &= \delta_{(a,x)}^{F(X)}(S_{\mathcal{A}} \times S_X \times S_{\mathbf{1}}) \\
&= \delta_a^{\mathcal{A}}(S_{\mathcal{A}}) \cdot \delta_x^X(S_X) \\
&= (\delta_a^{\mathcal{A}} \odot \delta_x^X)(S) \\
&= \lambda_X(a, \delta_x^X)(S) = \left(\lambda_X \circ F(\eta_X)\right)(a, x)
\end{aligned}$$

As above, Corollary 2.23 yields commutativity of the diagram. Finally we consider the diagram:

$$
\begin{tikzcd}
F\mathbb{T}^2X \arrow[r, "\lambda_{\mathbb{T}(X)}"] \arrow[d, "F(\mu_X)"'] & \mathbb{T}F\mathbb{T}X \arrow[r, "\mathbb{T}(\lambda_X)"] & \mathbb{T}^2FX \arrow[d, "\mu_{F(X)}"] \\
F\mathbb{T}X \arrow[rr, "\lambda_X"'] & & \mathbb{T}FX
\end{tikzcd}
$$

For arbitrary $(a, P) \in F\mathbb{T}^2X = \mathcal{A} \times \mathbb{T}(\mathbb{T}(X)) + \mathbf{1}$ and $S = S_{\mathcal{A}} \times S_X + S_{\mathbf{1}}$ we calculate:

$$(\lambda_X \circ F(\mu_X))(\checkmark) = \lambda_X(\checkmark) = \delta_{\checkmark}^{\mathcal{A}\times X+1}$$

and:

$$
\begin{aligned}
(\lambda_X \circ F(\mu_X))(a, P)(S) &= (\lambda_X(a, \mu_X(P)))(S) \\
&= \left(\delta_a^{\mathcal{A}} \odot \mu_X(P)\right)(S) \\
&= \delta_a^{\mathcal{A}}(S_{\mathcal{A}}) \cdot (\mu_X(P))(S_X) \\
&= \delta_a^{\mathcal{A}}(S_{\mathcal{A}}) \cdot \int_X p_{S_X}\,\mathrm{d}P
\end{aligned}
$$

Furthermore we obtain:

$$
\begin{aligned}
\left(\mu_{FX} \circ \mathbb{T}(\lambda_X) \circ \lambda_{\mathbb{T}(X)}\right)(\checkmark)(S) &= \mu_{FX}\left(\left(\delta_{\checkmark}^{\mathcal{A}\times\mathbb{T}(X)+\mathbf{1}}\right)_{\lambda_X}\right)(S) \\
&= \int_{\mathbb{T}FX} p_S\,\mathrm{d}\left(\delta_{\checkmark}^{\mathcal{A}\times\mathbb{T}(X)+\mathbf{1}}\right)_{\lambda_X} \\
&\overset{(1)}{=} \int_{\lambda_X^{-1}(\mathbb{T}FX)} p_S \circ \lambda_X\,\mathrm{d}\delta_{\checkmark}^{\mathcal{A}\times\mathbb{T}(X)+\mathbf{1}} \\
&\overset{(2)}{=} (p_S \circ \lambda_X)(\checkmark) \\
&= \delta_{\checkmark}^{\mathcal{A}\times X+\mathbf{1}}(S)
\end{aligned}
$$

and analogously:

$$\begin{aligned}
\left(\mu_{FX} \circ \mathbb{T}(\lambda_X) \circ \lambda_{\mathbb{T}(X)}\right)(a, P)(S) &= \mu_{FX}\left(\left(\delta_x^{\mathcal{A}} \odot P\right)_{\lambda_X}\right)(S) \\
&= \int_{\mathbb{T}FX} p_S \,\mathrm{d}\left(\delta_a^{\mathcal{A}} \odot P\right)_{\lambda_X} \\
&\overset{(1)}{=} \int_{\lambda_X^{-1}(\mathbb{T}FX)} p_S \circ \lambda_X \,\mathrm{d}\big(\delta_a^{\mathcal{A}} \odot P\big) \\
&\overset{(3)}{=} \int_{\{a\}\times\mathbb{T}(X)} p_S \circ \lambda_X \,\mathrm{d}\big(\delta_a^{\mathcal{A}} \odot P\big) \\
&\overset{(4)}{=} \int_{P' \in \mathbb{T}(X)} (\delta_a^{\mathcal{A}} \otimes P')(S) \,\mathrm{d}P(P') \\
&= \int_{P' \in \mathbb{T}(X)} \delta_a^{\mathcal{A}}(S_{\mathcal{A}}) \cdot P'(S_X) \,\mathrm{d}P(P') \\
&= \delta_a^{\mathcal{A}}(S_{\mathcal{A}}) \cdot \int_{\mathbb{T}(X)} p_{S_X} \,\mathrm{d}P
\end{aligned}$$

where we have used:

(1) Proposition 2.39.

(2) The fact that $\mathcal{A} \times \mathbb{T}(X)$ is a null set for $\delta_{\checkmark}^{\mathcal{A}\times\mathbb{T}(X)+\mathbf{1}}$ in conjunction with Corollary 2.38.

(3) The fact that $(\mathcal{A} \setminus \{a\}) \times \mathbb{T}(X) + \mathbf{1}$ is a null set for $\delta_a^{\mathcal{A}} \odot P$ in conjunction with Corollary 2.38.

(4) $(p_S \circ \lambda_X)(a, P') = (\delta_a^{\mathcal{A}} \odot P')(S)$ and $\forall \mathcal{M} \in \Sigma_{\mathbb{T}(X)} : (\delta_a^{\mathcal{A}} \odot P)(\{a\} \times \mathcal{M}) = P(\mathcal{M})$.

As before, commutativity of the diagram is a result of Corollary 2.23. $\square$

4.3.3 PTS with Termination as Coalgebra

Now we are ready to model probabilistic transition systems with termination as coalgebra:

Remark 4.23 (PTS with Termination as Coalgebra)**.** *Let $\mathcal{A}$ be an alphabet. We consider the sub-probability monad $(\mathbb{S}, \eta, \mu)$ and the endofunctor $F := \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}} + \mathbf{1}$ on* **Meas** *together with the lifting $\overline{F}\colon \mathcal{K}\ell(\mathbb{S}) \to \mathcal{K}\ell(\mathbb{S})$ induced by the distributive law $\lambda\colon F\mathbb{S} \to \mathbb{S}F$. A probabilistic transition system with explicit termination $((X, \Sigma), \mathcal{A}, \alpha)$ can be expressed as the $\overline{F}$-coalgebra $((X, \Sigma)_{\mathbb{S}}, \alpha^{\flat})$.*

Proof. This follows immediately from the definitions of a PTS with termination and a coalgebra in a Kleisli category. $\square$

4.3.4 A Final Coalgebra and Trace Semantics for PTS with Termination

Theorem 4.24 (Final Coalgebra)**.** *Let $\mathcal{A}$ be a finite alphabet. We consider the sub-probability monad $(\mathbb{S}, \eta, \mu)$ and the endofunctor $F := \mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}} + \mathbf{1}$ on* **Meas** *together with the lifting $\overline{F}\colon \mathcal{K}\ell(\mathbb{S}) \to \mathcal{K}\ell(\mathbb{S})$ induced by the distributive law $\lambda\colon F\mathbb{S} \to \mathbb{S}F$. A final $\overline{F}$-coalgebra is the tuple $\left((\mathcal{A}^*, \sigma(\mathcal{S}_*))_{\mathbb{S}}, \kappa^\flat\right)$ where $\kappa^\flat\colon (\mathcal{A}^*, \sigma(\mathcal{S}_*))_{\mathbb{S}} \to \overline{F}\left((\mathcal{A}^*, \sigma(\mathcal{S}_*))_{\mathbb{S}}\right)$ is the induced isomorphism associated to:*

$$\kappa := \eta_{F((\mathcal{A}^*,\sigma(\mathcal{S}_*)))} \circ \phi\colon \mathcal{A}^* \to \mathbb{S}(\mathcal{A} \times \mathcal{A}^* + \mathbf{1}), \quad u \mapsto \delta^{\mathcal{A}\times\mathcal{A}^*+\mathbf{1}}_{\phi(u)}$$

Given an arbitrary $\overline{F}$-coalgebra $\left((X, \Sigma)_{\mathbb{S}}, \alpha^\flat\right)$, the unique arrow mapping this coalgebra to the final coalgebra is $\mathbf{tr}^\flat\colon (X, \Sigma)_{\mathbb{S}} \to (\mathcal{A}^*, \sigma(\mathcal{S}_*))_{\mathbb{S}}$ associated to $\mathbf{tr}\colon X \to \mathbb{S}(\mathcal{A}^*)$ where for every $x \in X$ the sub-probability measure $\mathbf{tr}(x)\colon \sigma(\mathcal{S}_*) \to [0,1]$ is the uniquely determined extension of the finite (and thus σ-finite) pre-measure $\mathbf{tr}(x)\colon \mathcal{S}_* \to [0,1]$ defined by the following equations:

$$\begin{aligned} \mathbf{tr}(x)(\emptyset) &= 0 \\ \mathbf{tr}(x)(\{\varepsilon\}) &= \alpha(x)(\mathbf{1}) \\ \mathbf{tr}(x)(\{au\}) &= \int \mathbf{tr}_u \,\mathrm{d}P_{x,a} \quad \forall a \in \mathcal{A}, \forall u \in \mathcal{A}^* \end{aligned}$$

where for every $a \in \mathcal{A}$:

$$P_{x,a}\colon \Sigma_X \to [0,1], \quad P_{x,a}(S) := \alpha(x)(\{a\} \times S)$$

and for every $u \in \mathcal{A}^*$:

$$\mathbf{tr}_u\colon X \to [0,1], \quad x \mapsto \mathbf{tr}(x)(\{u\})$$

Proof. First we note that the map $\mathbf{tr}$ defined by the given equations is indeed a pre-measure as it is trivially σ-additive according to Proposition 4.1. Let us now show that it is σ-finite: We have $\mathbf{tr}(x)(\{\varepsilon\}) = \alpha(x)(\mathbf{1}) \leq 1$. Let us now assume that $\mathbf{tr}(x)(\{u\}) \leq 1$ for all words $u \in \mathcal{A}^{\leq n}$. For $u \in \mathcal{A}^n$ and arbitrary $a \in \mathcal{A}$ we conclude:

$$\begin{aligned} \mathbf{tr}(x)(\{au\}) &= \int \mathbf{tr}_u \,\mathrm{d}P_{x,a} \\ &= \int_{x' \in X} \underbrace{\mathbf{tr}(x')(\{u\})}_{\leq 1} \,\mathrm{d}P_{x,a}(x') \\ &\leq \int 1 \,\mathrm{d}P_{x,a} = P_{x,a}(X) = \alpha(x)(\{a\} \times X) \leq 1 \end{aligned}$$

and hence we have inductively shown that $\mathbf{tr}(x)(\{u\}) \leq 1$ for all $u \in \mathcal{A}^*$. Together with the trivial decomposition $\mathcal{A}^* = \cup_{u \in \mathcal{A}^*}\{u\}$ this yields that $\mathbf{tr}$ is σ-additive allowing us to apply the extension theorem, Proposition 2.22, to obtain a uniquely defined measure $\mathbf{tr}$. However, we do not know, whether it is a sub-probability measure. In a similar way as above, we will show that it is bounded above by 1 and thus a sub-probability measure. We claim that the following holds for all $n \in \mathbb{N}$:

$$\mathbf{tr}(x)\left(\mathcal{A}^{\leq n}\right) \leq 1$$

For the base case ($n = 0$) we have $\mathbf{tr}(x)\left(\mathcal{A}^{\leq 0}\right) = \mathbf{tr}(x)(\{\varepsilon\}) \leq 1$ and for the inductive step we suppose that for a fixed but arbitrary $n \in \mathbb{N}$ the following holds:

$$\mathbf{tr}(x)\left(\mathcal{A}^{\leq n-1}\right) \leq 1$$

We conclude with the following estimation:

$$\begin{aligned}
\mathbf{tr}(x)\left(\mathcal{A}^{\leq n}\right) &= \mathbf{tr}(x)\left(\bigcup_{u \in \mathcal{A}^{\leq n}}\{u\}\right) \overset{(1)}{=} \sum_{u \in \mathcal{A}^{\leq n}} \mathbf{tr}(x)\left(\{u\}\right) \\
&= \mathbf{tr}(x)(\{\varepsilon\}) + \sum_{a \in \mathcal{A}} \sum_{u \in \mathcal{A}^{\leq n-1}} \mathbf{tr}(x)\left(\{au\}\right) \\
&= \mathbf{tr}(x)(\{\varepsilon\}) + \sum_{a \in \mathcal{A}} \sum_{u \in \mathcal{A}^{\leq n-1}} \int \mathbf{tr}_u \,\mathrm{d}P_{x,a} \\
&\overset{(2)}{=} \mathbf{tr}(x)(\{\varepsilon\}) + \sum_{a \in \mathcal{A}} \int \left(\sum_{u \in \mathcal{A}^{\leq n-1}} \mathbf{tr}_u\right) \mathrm{d}P_{x,a} \\
&= \mathbf{tr}(x)(\{\varepsilon\}) + \sum_{a \in \mathcal{A}} \int_X \left(\sum_{u \in \mathcal{A}^{\leq n-1}} \mathbf{tr}(x')(\{u\})\right) \mathrm{d}P_{x,a}(x') \\
&\overset{(1)}{=} \mathbf{tr}(x)(\{\varepsilon\}) + \sum_{a \in \mathcal{A}} \int_X \underbrace{\left(\mathbf{tr}(x')\left(\mathcal{A}^{\leq n-1}\right)\right)}_{\leq 1} \mathrm{d}P_{x,a}(x') \\
&\overset{(3)}{\leq} \mathbf{tr}(x)(\{\varepsilon\}) + \sum_{a \in \mathcal{A}} \int_X 1 \,\mathrm{d}P_{x,a}(x') \\
&= \mathbf{tr}(x)(\{\varepsilon\}) + \sum_{a \in \mathcal{A}} P_{x,a}(X) \\
&= \alpha(x)(\mathbf{1}) + \sum_{a \in \mathcal{A}} \alpha(x)(\{a\} \times X) \\
&= \alpha(x)(\mathcal{A} \times X + \mathbf{1}) \leq 1
\end{aligned}$$

using:

(1) The fact that the union of singletons of mutually distinct elements is a disjoint union together with the σ-additivity of $\mathbf{tr}(x)$ for every $x \in X$.

(2) The linearity of the integral which can be applied here since $\mathcal{A}$ is finite which in turn implies that $\mathcal{A}^{\leq n-1}$ is finite and all the integrals $\int \mathbf{tr}_u \, \mathrm{d}P_{x,a}$ exist because $\mathbf{tr}_u$ is bounded above by 1.

(3) The induction hypothesis and the monotonicity of the integral.

We have shown:

$$\forall x \in X \ \forall n \in \mathbb{N}_0 : \mathbf{tr}(x)\left(\mathcal{A}^{\leq n}\right) \leq 1$$

which is equivalent to:

$$\forall x \in X \ \sup_{n \in \mathbb{N}_0} \mathbf{tr}(x)\left(\mathcal{A}^{\leq n}\right) \leq 1$$

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. Furthermore, $\mathbf{tr}(x)$ is continuous from below as a measure and we have $\mathcal{A}^{\leq n} \subseteq \mathcal{A}^{\leq n+1}$ for all $n \in \mathbb{N}$ and thus we obtain

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

In order to show that the given tuple is a final coalgebra, let $\left((X,\Sigma)_{\mathbb{S}}, \alpha^\flat\right)$ be an arbitrary $\overline{F}$-coalgebra. We have to show that $\mathbf{tr}^\flat \colon (X,\Sigma)_{\mathbb{S}} \to (\mathcal{A}^*, \sigma(\mathcal{S}_*))_{\mathbb{S}}$ is the uniquely determined arrow such that the following diagram commutes:

$$\begin{tikzcd}
(X,\Sigma)_{\mathbb{S}} \arrow[r, "\alpha^\flat"] \arrow[d, "\mathbf{tr}^\flat"'] & (\mathcal{A} \times X + \mathbf{1}, \sigma(\mathcal{P}(\mathcal{A}) * \Sigma \oplus \mathcal{P}(\mathbf{1})))_{\mathbb{S}} \arrow[d, "\overline{F}(\mathbf{tr}^\flat)"] \\
(\mathcal{A}^*, \sigma(\mathcal{S}_*))_{\mathbb{S}} \arrow[r, "\kappa^\flat", "\cong"'] & (\mathcal{A} \times \mathcal{A}^* + \mathbf{1}, \sigma(\mathcal{P}(\mathcal{A}) * \sigma(\mathcal{S}_*) \oplus \mathcal{P}(\mathbf{1})))_{\mathbb{S}}
\end{tikzcd}$$

We define the following arrows:

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

such that commutativity is equivalent to $g^\flat = h^\flat$ and conclude:

$$\begin{aligned} & g^\flat = h^\flat \\ \Leftrightarrow \quad & g = h \\ \Leftrightarrow \quad & g(x) = h(x) \quad \forall x \in X \\ \Leftrightarrow \quad & g(x)(S) = h(x)(S) \quad \forall x \in X, \forall S \in \sigma(\mathcal{S}'_*) \\ \Leftrightarrow \quad & g(x)(S) = h(x)(S) \quad \forall x \in X, \forall S \in \mathcal{S}'_* \end{aligned}$$

using for the last step that $\mathcal{S}'_*$ is a semi-ring of sets, the fact that for all $x \in X$ both $g(x)$ and $h(x)$ are finite and thus σ-finite measures since they are probability measures and Corollary 2.23. For arbitrary $x \in X$ and $S \in \mathcal{S}'_*$ we calculate $g(x)(S)$ as follows:

$$\begin{aligned} g(x)(S) &= (\mu_{\mathcal{A}\times\mathcal{A}^*+\mathbf{1}} \circ \mathbb{S}(\kappa) \circ \mathbf{tr})(x)(S) \\ &= \mu_{\mathcal{A}\times\mathcal{A}^*+\mathbf{1}} \left(\mathbb{S}(\kappa)(\mathbf{tr}(x))\right)(S) \\ &= \mu_{\mathcal{A}\times\mathcal{A}^*+\mathbf{1}} \left(\mathbf{tr}(x)_\kappa\right)(S) \\ &= \int p_S \,\mathrm{d}\mathbf{tr}(x)_\kappa \\ &\overset{(1)}{=} \int p_S \circ \kappa \,\mathrm{d}\mathbf{tr}(x) \\ &\overset{(2)}{=} \int \chi_{\psi(S)} \,\mathrm{d}\mathbf{tr}(x) = \mathbf{tr}(x)(\psi(S)) \end{aligned}$$

In (1) we used Proposition 2.39 and in (2) the fact that for every S as above and for every $u \in \mathcal{A}^*$ the following holds:

$$p_S \circ \kappa(u) = \delta_{\phi(u)}^{\mathcal{A}\times\mathcal{A}^*+\mathbf{1}}(S) = \chi_S(\phi(u)) = \chi_{\psi(S)}(u)$$

We conclude:

$$\begin{aligned} g(x)(\emptyset) &= \mathbf{tr}(x)(\emptyset) \\ g(x)(\mathbf{1}) &= \mathbf{tr}(x)(\{\varepsilon\}) \\ g(x)(\{(a,u)\}) &= \mathbf{tr}(x)(\{au\}) \quad \forall a \in \mathcal{A}, \forall u \in \mathcal{A}^* \end{aligned}$$

In order to calculate $h(x)(S)$, let:

$$\rho := \lambda_{\mathcal{A}\times\mathcal{A}^*+\mathbf{1}} \circ (\mathrm{id}_{\mathcal{A}} \times \mathbf{tr} + \mathbf{1})\colon \mathcal{A} \times X + \mathbf{1} \to \mathbb{S}(\mathcal{A} \times \mathcal{A}^* + \mathbf{1})$$

and for every $S \in \mathcal{S}'_*$ let:

$$\rho_S := p_S \circ \rho \colon \mathcal{A} \times X + \mathbf{1} \to [0,1], \quad z \mapsto \big(\rho(z)\big)(S)$$

Then we obtain the following result:

$$\begin{aligned}
h(x)(S) &= (\mu_{\mathcal{A}\times\mathcal{A}^*+\mathbf{1}} \circ \mathbb{S}(\rho) \circ \alpha)(x)(S) \\
&= \mu_{\mathcal{A}\times\mathcal{A}^*+\mathbf{1}} \left(\mathbb{S}(\rho)(\alpha(x))\right)(S) \\
&= \mu_{\mathcal{A}\times\mathcal{A}^*+\mathbf{1}} \left(\alpha(x)_\rho\right)(S) \\
&= \int p_S \, \mathrm{d}\alpha(x)_\rho \\
&\overset{(1)}{=} \int p_S \circ \rho \, \mathrm{d}\alpha(x) \\
&= \int \rho_S \, \mathrm{d}\alpha(x)
\end{aligned}$$

where we have used Proposition 2.39 in (1). Obviously we have $h(x)(\emptyset) = 0$. We note that $\rho_{\mathbf{1}} = \chi_{\mathbf{1}}$ and hence:

$$h(x)(\mathbf{1}) = \int \rho_{\mathbf{1}} \, \mathrm{d}\alpha(x) = \int \chi_{\mathbf{1}} \, \mathrm{d}\alpha(x) = \int_{\mathbf{1}} 1 \, \mathrm{d}\alpha(x) = \alpha(x)(\mathbf{1})$$

For $a \in \mathcal{A}, u \in \mathcal{A}^*, (b, x') \in \mathcal{A} \times X$ the following holds:

$$\begin{aligned}
\rho_{\{(a,u)\}}(b, x') &= \big(\delta_b^{\mathcal{A}} \odot \mathbf{tr}(x')\big)(\{(a,u)\}) \\
&= \delta_b^{\mathcal{A}}(\{a\}) \cdot \mathbf{tr}(x')(\{u\}) \\
&= \chi_{\{a\}}(b) \cdot \mathbf{tr}_u(x') \\
&= \chi_{\{a\}\times X}(b, x') \cdot (\mathbf{tr}_u \circ \pi_2)(b, x')
\end{aligned}$$

and:

$$\rho_{\{(a,u)\}}(\checkmark) = \delta_{\checkmark}^{\mathcal{A}\times\mathcal{A}^*+\mathbf{1}}(\{(a,u)\}) = 0$$

and hence:

$$\begin{aligned}
h(x)(\{(a,u)\}) &= \int \rho_{\{(a,u)\}} \, \mathrm{d}\alpha(x) \\
&= \int_{\mathcal{A}\times X+1} \chi_{\mathcal{A}\times X} \cdot \rho_{\{(a,u)\}} \, \mathrm{d}\alpha(x) \\
&= \int_{\mathcal{A}\times X} \rho_{\{(a,u)\}} \, \mathrm{d}\alpha(x) \\
&= \int_{\mathcal{A}\times X} \chi_{\{a\}\times X} \cdot (\mathbf{tr}_u \circ \pi_2) \, \mathrm{d}\alpha(x) \\
&= \int_{\{a\}\times X} \mathbf{tr}_u \circ \pi_2 \, \mathrm{d}\alpha(x) \\
&= \int_X \mathbf{tr}_u \, \mathrm{d}P_{x,a}
\end{aligned}$$

Thus we have shown that the required commutativity of the diagram is equivalent to the following set of equations:

$$\begin{aligned}
\mathbf{tr}(x)(\emptyset) &= 0 \\
\mathbf{tr}(x)(\{\varepsilon\}) &= \alpha(x)(\mathbf{1}) \\
\mathbf{tr}(x)(\{au\}) &= \int \mathbf{tr}_u \, \mathrm{d}P_{x,a} \quad \forall a \in \mathcal{A}, \forall u \in \mathcal{A}^*
\end{aligned}$$

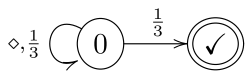
We denote that $\mathbf{tr}$ as defined in the proposition fulfills these equations and that $\mathbf{tr}$ is uniquely defined by these equations. $\square$

4.3.5 Examples

Example 4.25. *Let $X = \{0\}$ and $\mathcal{A}$ be a finite alphabet such that $\diamond \in \mathcal{A}$. We define:*

$$\alpha(0) := \frac{1}{3} \cdot \left(\delta_{(\diamond,0)}^{\mathcal{A}\times\{0\}+\mathbf{1}} + \delta_{\checkmark}^{\mathcal{A}\times\{0\}+\mathbf{1}} \right)$$

The graphical representation of the resulting transition system $((\{0\}, \{\emptyset, \{0\}\}), \alpha)$ is:



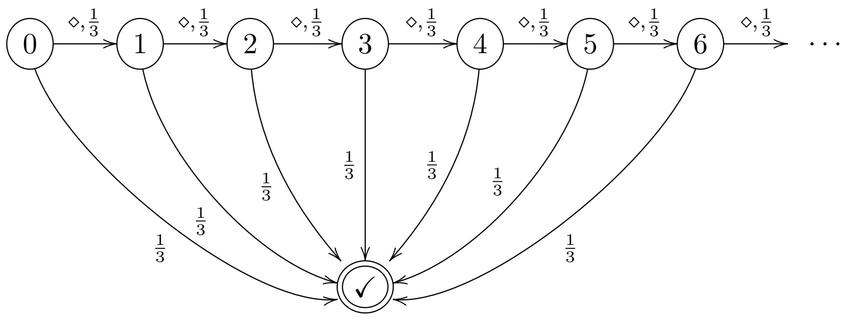
We denote that $P_{0,a} = \frac{1}{3} \cdot \delta_\diamond^{\mathcal{A}}$ and use induction to show that for all $n \in \mathbb{N}$ we have $\mathbf{tr}(0)(\{\diamond^n\}) = \frac{1}{3^{n+1}}$. For the base step we have $\mathbf{tr}(0)(\{\varepsilon\}) = \alpha(0)(\mathbf{1}) = \frac{1}{3}$ and assuming that for an arbitrary but fixed $n \in \mathbb{N}_0$ the induction hypothesis $\mathbf{tr}(0)(\{\diamond^{n-1}\}) = \frac{1}{3^n}$ holds, *we execute the inductive step:*

$$\mathbf{tr}(0)(\{\diamond^n\}) = \int \mathbf{tr}_{\{\diamond^{n-1}\}}\,\mathrm{d}P_{m,\diamond} = \frac{1}{3}\cdot\mathbf{tr}(0)(\{\diamond^{n-1}\}) = \frac{1}{3}\cdot\frac{1}{3^n} = \frac{1}{3^{n+1}}$$

Example 4.26. *Let $X := \mathbb{N}_0$ and $\mathcal{A}$ be a finite alphabet such that $\diamond \in \mathcal{A}$. We define*

$$\alpha(n) := \frac{1}{3}\cdot\left(\delta^{\mathcal{A}\times\mathbb{N}+\mathbf{1}}_{(\diamond,n+1)} + \delta^{\mathcal{A}\times\mathbb{N}+\mathbf{1}}_{\checkmark}\right)$$

for every $n \in \mathbb{N}_0$ and consider the probabilistic transition system $\big((\mathbb{N}_0, \mathcal{P}(\mathbb{N}_0)), \alpha\big)$ which has the following graphical representation:



We denote that $P_{n,a} = \frac{1}{3}\cdot\delta^{\mathcal{A}}_{\diamond}(\{a\})\cdot\delta^{\mathbb{N}_0}_{n+1}$ and use induction to show that for all $m, n \in \mathbb{N}_0$ we have $\mathbf{tr}(m)(\{\diamond^n\}) = \frac{1}{3^{n+1}}$. For the base step note that $\mathbf{tr}(m)(\{\varepsilon\}) = \alpha(m)(\mathbf{1}) = \frac{1}{3}$ and assuming that for all $m \in \mathbb{N}_0$ and an arbitrary but fixed $n \in \mathbb{N}_0$ the induction hypothesis $\mathbf{tr}(m)(\{\diamond^{n-1}\}) = \frac{1}{3^n}$ holds, we execute the inductive step:

$$\mathbf{tr}(m)(\{\diamond^n\}) = \int \mathbf{tr}_{\{\diamond^{n-1}\}}\,\mathrm{d}P_{m,\diamond} = \frac{1}{3}\cdot\mathbf{tr}(m+1)(\{\diamond^{n-1}\}) = \frac{1}{3}\cdot\frac{1}{3^n} = \frac{1}{3^{n+1}}.$$

5 Conclusions and Future Work

5.1 Conclusions and Remarks

Utilizing the category of measurable spaces, we have shown that coalgebras in the Kleisli category of the (sub-)probability monad can be used to model probabilistic transition systems and that the final coalgebra - if it exists - yields a definition of the "trace" of a state and hence a notion of "trace equivalence". This supports the idea that coalgebra captures behavior of dynamic systems. Moreover, as in the category **Set**, the use of a monad to "hide side effects" is a sufficient choice to obtain trace equivalence.
Besides the fact that our approach of working in the category of measurable spaces, **Meas**, nicely generalizes known results from probabilistic transition systems *with explicit termination* and finite state space (cf. Hasuo et al. [2007] who use **Set** and the sub-distribution monad $\mathcal{D}$) to systems with a continuous state space, it also allows us to define trace semantics for PTS *without explicit termination.* To our knowledge, this is impossible without measure theory since for an alphabet $\mathcal{A}$ with at least two distinct elements, the set of infinite words, $\mathcal{A}^\omega$, is uncountably infinite. Thus, even for *discrete* probabilistic transition systems without termination, there is no way to obtain the notion of a trace as a final coalgebra in the Kleisli-category of the sub-distribution monad $\mathcal{D}$ on **Set**.
Last but not least, another fact worth mentioning is, that the existence of the final coalgebra and the uniqueness of the arrow mapping an arbitrary coalgebra into the final coalgebra is so closely connected to the extension theorem for σ-finite measures on semi-rings of sets.

5.2 Literature for Further Reading

This thesis gives only a relatively small insight into the theory of analyzing dynamic systems with transitions depending on randomness. Below is a short list of publications for further reading.

- A good start for the basic ideas of *category theory* can be found in Pierce [1991], for an extensive overview take a look at Mac Lane [1998].

- A general introduction to the theory of *(co)algebras and (co)induction* can be found in Jacobs and Rutten [1997].
- A thorough and extensive introduction to *measure and integration theory* in German is given by Elstrodt [2007].
- The foundation of our approach is based on Hasuo et al. [2007] who characterize trace equivalence as final coalgebra in the Kleisli category of a suitable monad on **Set**.
- While the behavioral equivalance we were interested in is *trace equivalence*, Panangaden [2009] looks at *bisimulation* for probabilistic transition system which he calls "labelled markov processes". He identifies the Kleisli category of the sub-probability monad as a category with *measurable spaces* as objects and *conditional probability densities* or *markov kernels* as arrows.
- The main reference for the *probability monad* and its Kleisli category is of course given by Giry [1981].
- Bonchi et al. [2011] use coalgebra to define an *algorithm for checking behavioral equivalence of dynamic systems* performing both *minimization and determinization* of the given system.

5.3 Open Questions and Future Work

Even though we have provided some new results, we have also generated a couple of new questions which might be interesting for future work:

- Is the finiteness of the alphabet $\mathcal{A}$ necessary for the existence of a final coalgebra or does it exist for countably infinite alphabets as well?
- We have shown that a final coalgebra exists for the combination of $\mathbb{S}$ and $\mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}} + \mathbf{1}$ and for the combination of $\mathbb{P}$ and $\mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}}$. The obvious question is whether there is a final coalgebra for $\mathbb{P}$ and $\mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}} + \mathbf{1}$ or for $\mathbb{S}$ and $\mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}}$? For the first combination, namely $\mathbb{P}$ and $\mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}} + \mathbf{1}$, our current assumption is that the use of an $\infty-$cone σ-algebra similar to that of the ω-cones yields $(\mathcal{A}^\infty, \sigma(\mathcal{S}_\infty))$ as final coalgebra where $\mathcal{S}_\infty \subseteq \mathcal{P}(\mathcal{A}^\infty)$ is again a semi-ring of sets. Our assumption for $\mathbb{S}$ and $\mathcal{A} \times \mathrm{Id}_{\mathbf{Meas}}$ is that the final coalgebra does not exist, but possibly something like a weakly final coalgebra.

- Can the ideas of minimization and determinization as presented in Bonchi et al. [2011] be applied to our setting? What are suitable factorization structures and metrics?
- In our model for probabilistic transition systems without explicit termination we have no specification what a "good" infinite word is. We accept every possible infinite word since there are no acceptance conditions like for instance in a Büchi automaton. An interesting question could thus be, whether there is a way to introduce acceptance conditions for infinite sequences (not only) in probabilistic transition systems.

Bibliography

Robert B. Ash. *Real Analysis and Probability*. Probability and Mathematical Statistics - A Series of Monographs and Textbooks. Academic Press, 111 Fifth Avenue, New York, New York, 1972.

Filippo Bonchi, Mathias Hülsbusch, Barbara König, and Alexandra Silva. A coalgebraic perspective on minimization, determinization and behavioural metrics. Technical Report SEN 11-06, Centrum Wiskunde & Informatica, 2011.

Jürgen Elstrodt. *Maß- und Integrationstheorie*. Springer, 5th edition, 2007.

Michéle Giry. A categorical approach to probability theory. In *Categorical Aspects of Topology and Analysis*, volume 915 of *Lecture Notes in Mathematics*, pages 68–86. Springer, 1981.

Ichiro Hasuo, Bart Jacobs, and Ana Sokolova. Generic trace semantics via coinduction. *Logical Methods in Computer Science*, 3 (4:11):1–36, November 2007.

Bart Jacobs and Jan Rutten. A tutorial on (co)algebras and (co)induction. *Bulletin of the European Association for Theoretical Computer Science*, 62:222–259, 1997.

Dénes König. *Theorie der endlichen und unendlichen Graphen*. Chelsea Publishing Company New York, N.Y., 1936.

Saunders Mac Lane. *Categories for the Working Mathematician*. Springer, 2nd edition, 1998.

Prakash Panangaden. *Labelled Markov Processes*. Imperial College Press, 2009.

Benjamin C. Pierce. *Basic Category Theory for Computer Scientists*. The MIT Press, Cambridge, Massachusetts, 1991.

Hiermit erkläre ich, dass ich diese Diplomarbeit selbstständig verfasst, keine anderen Quellen und Hilfsmittel als die angegebenen verwendet und alle Zitate kenntlich gemacht habe.

I hereby declare that I have written this diploma thesis independently, that I have not used any other sources and aids than the ones stated and that I have marked all citations.

Duisburg, September 27, 2011,

Henning Kerstan