Università degli Studi di Udine

Dipartimento di Matematica e Informatica

Dottorato di Ricerca in Informatica

Ph.D. Thesis

Generalized labelled Markov processes, coalgebraically

Candidate:

Giorgio Bacci

Supervisor:

Prof. Marino Miculan

Academic Year 2011/12
Version of May 9, 2013

Author's e-mail: giorgio.bacci@uniud.it

Author's address:

Dipartimento di Matematica e Informatica
Università degli Studi di Udine
Via delle Scienze, 206
33100 Udine
Italia

Contents

1
Introduction

This thesis is concerned with the analysis of generalized labelled Markov processes, that is, dynamical systems with continuos state space, interacting with the environment by means of input labels and producing measurable events by means of transitions to a measurable set of successor states. The term "generalized" is used to stress the fact that transition events can be measured by formal measures on a generic measurable space, without assuming a priori that these are of a certain type, e.g., (sub)probability measures, finite measures, or σ-finite measures. The adjective "Markovian" is usually employed in the probabilistic setting; here it just indicates that the transitions depend entirely on the present state and not on the past history of the system.

We will model Markov processes coalgebraically, in the category of measurable spaces and measurable functions, following the lines of Desharnais et al. [36, 40, 18] and recent books of Panangaden [69] and Doberkat [42, 43] that contain most of the research on probabilistic systems.

The general goal is to make a step forward in the analysis of Markov processes. Unlike many results in the literature, in this thesis we have made great efforts in order to develop the whole theory of Markov processes without assuming particular properties of the measurable state space. In particular, we will never assume that the state space is either Polish or analytic, but we forced ourselves to work only with generic measurable spaces.

Why continuous states? In recent years continuos data have become very important in computer science, especially when one consider real physical models, and dynamical systems evolving in a continuous state, by involving continuous parameters such as concentrations, temperature, pressure, distances, etc. These systems arise in biology, engineering, security (e.g., of wireless networks, telecommunications, etc.) and, of course, in many other fields. Sometimes the use of continuous data cannot be avoided without affecting the behavior of the model: there are situations in which the discretization of continuous parameters may totally change the response of the model. Concrete interesting examples can be found in [18].

Due to the complexity of such models, computer scientists tried to to provide techniques to help the analysis and the reasoning on continuous state systems. This is best done with the help of formal methods, that is, mathematically based languages, techniques and tools for specifying, describe and verify systems. When designing a system, the ultimate goal is to make it operate reliably, despite its complexity. Formal methods are used more and more in industry, not only for verification but also for the preliminary specification of the models. Both these techniques have been proven to greatly improve the product quality.

1.1 Bisimulation for Labelled Markov Processes

The notion of bisimulation is central in the study of concurrent systems. In the case of nondeterministic labelled transition systems (strong) bisimilarity of Milner and Park [65, 70] is *the* basic equivalence equating systems exhibiting the same behavior. Intuitively, two systems are bisimilar if they match each other's moves, in this sense that each of the systems cannot be distinguished from the other by any external observer. For discrete probabilistic transition systems the basic

process equivalence is probabilistic bisimilarity of Larsen and Skou [63]. The main difference regards transition probabilities which have to be taken into account in the behavior of the system. Intuitively, two states are bisimilar if we get the same probability after we have added up the transition probabilities to all the states in an equivalence class of bisimilar states. The adding up is crucial, since the probabilities are not just another labels.

When one moves to probabilistic systems over continuous state spaces, the notion of bisimilarity becomes surprisingly difficult, and many technical problems suddenly arise due to the continuous nature of the state space. The first notion of bisimulation for labelled probabilistic Markov processes, that is, probabilistic systems with generic measurable space of states, has been given categorically by Blute et al. [18] as a span of zig-zag morphisms, that is, measurable surjective maps respecting the transition structure of the Markov process. Since from the beginning, it turned out to be very difficult to prove that the induced notion of bisimilarity is an equivalence relation. This problem was solved by a very involved construction due to Edalat [45] which, although, requires that one works with a Polish or, more generally, an analytic space structure. Under these assumptions they were able to prove that bisimilarity is an equivalence and moreover, they gave a neat logical characterization of it [36], resulting in a very simple logic. In subsequent works [37, 38] the definition of bisimulation has been characterized in more plain mathematical terms, without the need of notions from category theory. This characterization mimics the definition of Larsen and Skou for discrete systems, but few measure-theoretic conditions have been imposed to deal with the fact that not all sets need to be measurable. However, this characterization was given assuming that the bisimulation relation is already an equivalence, hence they do not cover all possible cases.

In [31], Danos et al. introduced a notion alternative to that of bisimulation, the so called *event bisimulation*. This definition is dual to that of bisimulation as it has been given in [18] in the sense that spans are replaced by cospans. With this definition, equivalence for event bisimilarity is always guaranteed and, moreover, they were able to characterize event bisimilarity by the logic without any assumption of analyticity of the state space. This notion have been proven to be equivalent to standard (or state) bisimilarity in the case of analytic spaces.

In this thesis we prove that bisimilarity for Markov processes over generic measurable spaces is an equivalence. Our proof does not assume any Polish or analytic structure on the state space, hence solves the problem posed in [18, 36]. The proof of equivalence is given in terms of a characterization of bisimulation that generalizes that given in [37, 38] to generic binary relations. This characterization is proven to be in one-to-one correspondence with the abstract coalgebraic notion of bisimulation of Aczel and Mendler, hence all the results extend to the coalgebraic setting. In virtue of the proof of equivalence, it is reasonable to ask if the concepts of bisimilarity and event bisimilarity coincide in general, without assuming analyticity on the state space. Unfortunately, as it has been proven by Terraf in [80], this is not the case. Nevertheless, we will see that bisimilarity is contained in event bisimilarity, thus that one of the two inclusion still holds, even without assuming analyticity. The proof of this result is shown coalgebraically, establishing a formal adjunction between the category of bisimulations and that of cocongruences (actually, only a subcategory of the latter). To the best of our knowledge, also this result is new and, together with the counterexample given in [80], concludes the comparison between these two notions of equivalence between Markov processes over generic measurable spaces.

1.2   Structural Operational Semantics for Markov Processes

The *operational semantics* of a programming language accounts for a formal description of the *behavior* of programs, specifying the way they should be executed and the kind of behavior which should be observed. To a programming language can also be given a more abstract mathematical representation by means of a *denotational semantics*. Concretely, to each expression of the language is assigned a *denotation*, i.e., an object in a mathematical domain. In this respect, each program is represented by a function over denotations that maps each input into the corresponding output. An important property of denotational semantics is that it should be compositional, i.e., the denotation of a program expression can be constructed by the denotation of its sub-expression. This allows

inductive reasoning on the structure of programs, and provides a general way to prove properties of these.

Every semantics gives rise to a notion of equivalence between programs, i.e., *semantical equivalence*, which equates programs having the same semantics. It is always a good practice to give to programming languages both an operational and denotational semantics, and this should be done ensuring that the respective semantical equivalences coincide. This property is usually denoted as *full abstraction*. A denotational semantics is fully abstract with respect to a certain operational semantics whenever it holds that two expressions have the same denotation if and only if they are behavioral equivalent. This means that they cannot be distinguished by an external observer that looks at their executions in all possible environments. Full abstraction, usually guarantees that the operational semantics is compositional in the sense that behavioral equivalent subprograms can be substituted without affecting the overall behavior of the system containing them.

Compositionality is usually met when the semantics of the program language terms depends only on the semantics of its subcomponents. Such property is the mantra of Plotkin's *structural operational semantics* (SOS) [71], which is one of the most applied tools for giving operational semantics to recursively defined process description languages when the operational semantics is given in terms of labelled transition systems (deterministic or not). Labelled transition systems are defined by means of a set of derivation rules that allows for a simple description of the transitions of a labelled transition system following the syntactic structure of the terms of the programming language. The great success of the SOS paradigm is mainly due to the fact that many important semantic properties, such as congruence for bisimilarity, can be established simply by inspecting the syntactic format of the rules. Depending on the format of the derivation rules, operational semantics can be more expressive than others. The most popular rule formats for labelled transition systems are the so called GSOS format [17] and the tyft/tyxt rule format [51], but there are many other in the literature each with their own specific features (see [2] for a survey).

In recent years, SOS specification systems have been also developed for *stochastic* and *probabilistic* systems, due to their important applications to performance evaluation, systems biology, etc [55, 21, 54, 41]. For example, Bartels [15] have investigated rule formats both for simple discrete probabilistic systems and Segala systems, and Klin and Sassone [61, 60] proposed rule formats for stochastic systems with discrete state space and, more generally, for weighted transition systems. However, these formats still do not cover the case of *continuous-state* probabilistic and stochastic systems, like calculi with spatial/geometric features introduced in last years [24, 12]. In these models, the behaviour of the system may be influenced by continuous data, which therefore is part of the state of the system. Typical examples are quantitive informations such as density, volumes, concentrations, and spatial informations, such as the position of processes and where transitions take place; e.g., in wireless networks distance may affect data access, or in biological models diffusion alters the signaling pathways, etc.

Working with continuous data is not simple in general, and even very simple process algebras may become extremely difficult to be described in terms of probabilistic transition systems. Consider, for example this simple yet paradigmatic calculus of agents, where CCS-like synchronizations are affected by the concentrations of the agents in the system

$$P, Q ::= \mathbf{0} \mid \alpha.P \mid P \parallel Q \mid c \text{ } \mathbf{of} \text{ } P \qquad \text{where } c \in \mathbb{R}_{\geq 0} \text{ and } \alpha \in A \cup \overline{A} \cup \{\tau\}$$

where $\mathbf{0}$ denotes the null process, $\alpha.P$ denotes the action prefix, $P \parallel Q$ denotes the parallel composition, and c $\mathbf{of}$ P denotes the system with continuous number $c \in \mathbb{R}_{\geq 0}$ of occurrences of P. The idea we aim to model is that the rate of execution of an action $a \in L$ must depend on the availability of the agents that may perform this action. Of course, we want also to be faithful with the intuitive idea that after an action has been performed the occurrence of that action must be removed from the system. So, the problem is how to specify the semantics of a process like c $\mathbf{of}$ P. Any discrete semantic would force us to decide a priory which is the quantity of P to be consumed in c $\mathbf{of}$ P, with a rule of the form

$$\frac{x \xrightarrow{\alpha[r]} x'}{c \text{ } \mathbf{of} \text{ } x \xrightarrow{\alpha[c' \cdot r]} c' \text{ } \mathbf{of} \text{ } x' \parallel (c - c') \text{ } \mathbf{of} \text{ } x}$$

where r denotes the execution rate of the stochastic α-transition in the premise, and c' denotes the concentration of the agent consumed by the transition. Any fixed choice of $c' \leq c$ would be unreasonable in a continuos state semantics, since the uniform probability of choosing the exact value of c' in the interval $[0, c]$ would be always zero. The only satisfactory choice is to change the format of the transitions, in order to give an actual continuous state operational semantics to the calculus.

The operational models we are interested in are, therefore, Markov processes, so that, the notion of interest is no longer a measure on a discrete space, but a measure over a generic measurable space. This leads to transitions of the form $t \xrightarrow{\alpha} \mu$, where t is the current state of the system, $\alpha \in L$ is an action label representing the interactions with an external environment, and μ is an *actual measure* over a measurable space of process terms, measuring the the possible outcomes of P. Regarding our example above the semantics can be given by the following rule

$$\frac{x \xrightarrow{\alpha} \mu}{c \text{ of } x \xrightarrow{\alpha} c \cdot (U[0, c] \times \mu \times \delta_x) \circ (\lambda(c', x', x). \, c' \text{ of } x' \parallel (c - c') \text{ of } x)^{-1}}$$

where $U[0, c](E) = \int_{[0,c] \cap E} \frac{1}{c} \, dx$, for any measurable set E in $\mathbb{R}_{\geq 0}$, denotes the uniform probability distribution over the interval $[0, c] \cap E$, δ_x is the Dirac distribution at x, $\mu \times \nu$ the product measure, and the lambda term (on the right) in the conclusion denotes a (measurable) function taking three arguments an returning a process term.

Semantics with a similar transition format have been considered already by Cardelli and Mardare in [26, 10] for dealing with specific equational stochastic systems. However, differently from the case of discrete processes, the SOS specification given in [26, 10] are rather ad hoc, and they are not based on any general framework for operational descriptions.

In traditional GSOS format, the target of a transition is a term built from the components of the source process, and their corresponding semantics. In our settings, the target of a transition is not a term, but a measure over a generic measurable space, hence the derivations of rules becomes more complicated. We cope with this problem proposing transitions of the form $t \xrightarrow{\alpha} \mu$ where μ is no more a measure but a syntactic expression intended to denote a measure, which we call *measure term*. The syntax of measure terms, and their interpretation as actual measures, is part of the operational specification: a specification is given by a set of rules together with a description of how measures must be combined. Has one may aspect, not all measure interpretations guarantees that bisimilarity is a congruence. Sufficient conditions for ensuring well-behaved interpretation can be established rather easily working at the algebraic and coalgebraic level.

Bialgebraic framework. An abstract formulation of well-behaved SOS specification formats has been proposed by Turi and Plotkin [83, 82], who built a strong bridge between this approach and denotational semantics: the so called *bialgebraic framework*. The key intuition is that rule specification systems can be formulated in terms of certain natural transformations, called *distributive laws*. The models for these distributive laws are *bialgebras*, that is, a pair consisting of a T-algebra $\alpha \colon TX \to X$ and a D-coalgebra $\beta \colon X \to DX$ on the same carrier and such that they are related by a distributive law $\lambda \colon TD \Rightarrow DT$ of a monad T over a comonad D as follows:

$$
\begin{array}{ccccc}
TX & \xrightarrow{\ \alpha\ } & X & \xrightarrow{\ \beta\ } & DX \\
{\scriptstyle T\beta} \downarrow & & & & \uparrow {\scriptstyle D\alpha} \\
TDX & & \xrightarrow{\ \lambda_X\ } & & DTX
\end{array}
$$

Intuitively, the monad T represents the syntax of the programming language and the comonad D models the shape of computations. The algebra $\alpha \colon TX \to X$ and coalgebra $\beta \colon X \to DX$, respectively, denote the denotational and operational models of the system, and the distributive law $\lambda \colon TD \Rightarrow DT$ explains how the syntax distributes over the computations, that is to say, how the computation of syntactic operator depends on the executions of its arguments. Bialgebras form

a category, where any unique morphism from the initial object represent a denotational semantics, and any unique morphism to the final one an operational semantics. Hence, one can always find a canonical fully-abstract semantics: the universal morphism from the initial to the final bialgebras.

The main advantages of this abstract categorical formulation are that it allows for a deeper understanding of the semantical theory leaving out all the technical details due to the specific model at hand, and, more importantly, it permits to instantiate the general framework to different kinds of systems behaviors.

We will see that our *Measure GSOS specification format* induces a distributive law that fits the bialgebraic framework of Turi and Plotkin, so that we will be able to give a universal fully-abstract semantics to Markov processes. Since all morphisms respect the operations of the term language, the universal semantics is compositional and as a trivial consequence, both behavioral equivalence and (state) bisimilarity over Markov processes are congruences with respect to our rule format.

1.3 Behavioral Pseudometrics and Algorithms

When one focuses on quantitative behaviors it becomes obvious that any notion of equivalence is too strict, even that of bisimilarity or behavioral equivalence. Indeed, in many situations it is still of interest knowing whether two systems that may differ by a small perturbation in the continuous parameters have "sufficiently" similar behaviors. This motivated the development of the metric theory for Markov process, initiated by Desharnais et al. [39] and greatly developed and explored by van Breugel, Worrell, and others [88, 87]. It consists in proposing a pseudometric which measures the behavioral similarity of the systems. This pseudometric, of course, must be consistent with behavioral equivalence, that is, two systems must be at distance zero if and only if they exhibit the same behavior. Moreover, working with distances rather than equivalence relations, allows one also to adapt the notion of similarity between systems according to the problems we have to deal with. For example, the pseudometric proposed by Desharnais et al. is parametric in a *discount factor* $\lambda \in (0, 1]$ that controls the significance of the future in the measurement. Having a discount factor $\lambda < 1$ amounts to make the future behavior of the system less significant; if $\lambda = 1$ the future is not discounted and any transition in the present or in the future have the same relevance.

Since van Breugel et al. have presented a fixed point characterization of the bisimilarity pseudometric, several iterative algorithms have been developed in order to compute approximations of the pseudometric up to any degree of accuracy [46, 88, 87]. Recently, Chen et al. [28] proved that, for finite Markov chains the bisimilarity pseudometrics can be computed exactly in polynomial time. The proof consists in describing the pseudometric as the solution of a linear program that can be solved using the *ellipsoid method*. Although the ellipsoid method is theoretically efficient, *"computational experiments with the method are very discouraging and it is in practice by no means a competitor of the, theoretically inefficient, simplex method"*, as stated in [75]. Unfortunately, in this case the simplex method cannot be used to speed up performances in practice, since the linear program to be solved may have an exponential number of constraints.

In this thesis we propose an efficient on-the-fly algorithm for computing exactly the pseudometric of Desharnais et al. [39]. This algorithm is inspired by a characterization of the undiscounted pseudometric given in [28] based on the notion of *coupling* of Markov chains, which we extend to generic discount factors. The advantage of using an on-the-fly approach consists in the fact that we do not need to exhaustively explore the state space nor to construct and store the data structure entirely, we will only need those fragments that are really demanded by the *local* computation.

The efficiency of our algorithm has been evaluated empirically on a consistent set of randomly generated MCs. The results show that our algorithm performs better than the iterative algorithms proposed, for instance in [46, 28].

1.4   Structure of the Thesis and Contributions

We summarize below the content and main contributions of each chapter.

Chapter 2.   It introduces the basic preliminaries on category theory and measure theory, and it is mainly aimed at fixing the notation and the terminology that will be used in the rest of the thesis. All the material in this chapter are not original and can be found in any (good) textbook on category theory and measure theory. The only originality is in the exposition of the results, which are summarized to be used as a short (nevertheless, complete) reference for non-expert readers.

The last section summarizes the definitions and the main properties that are specific to the category of measurable spaces and measurable maps. This section will be often referred to along the thesis, since many definitions will serve in many chapters.

Chapter 3.   It contains material already existing in the literature, and is aimed to collect the main results on the theory of universal algebras and coalgebras.

In particular, it recalls the definitions of algebra and coalgebra for a functor, and the categorical generalizations of the concepts of congruence, bisimulation and cocongruence. Moreover, we also recall the abstract definitions of induction and coinduction in relation with the notion of initial algebra and final coalgebra for a functor. These concepts are then also related to the notions of free and cofree constructions provided by the universal properties of adjoint functors.

Chapter 4.   We provide general techniques for proving the existence and also to characterize initial and final objects in the category of algebra and coalgebras, respectively.

The main contributions of this chapter are (i) a proof for the existence of initial algebras for the class of polynomial functors in the category of measurable spaces, (ii) the proof of existence of final coalgebras for the class of measure functors which are specific to the category of measurable spaces, and (iii) an alternative and general construction for initial algebras and final coalgebras that uses the axiomatic properties of factorization systems in relation to initial and final sequences for an endofunctor. As a side result, we slightly generalize a well-known theorem of stabilization for the final sequence due to Worrell [92, Theorem 4.6].

All the categorical constructions provided in this chapter are given attempting at never assume specific properties of the category of sets, hence many definitions and results may be found a bit counterintuitive for people not used to categorical abstraction. These efforts, however, return in terms of the generality of the constructions, that apply in categories notoriously difficult to handle, such as **Top** (the category of topological spaces and continuous functors), **UMet** (the category of ultrmetric spaces), **PMet** (the category of pseudometric spaces), and **Meas** (the category of measurable spaces and measurable maps).

Chapter 5.   We present a theory of generalized labelled Markov processes which brings together under a unique framework probabilistic and stochastic Markov processes of [36] and [26]. The main contributions are (i) an exact and faithful characterization of the coalgebraic bisimulation for Markov processes in "plain" mathematical terms (ii) the proof that bisimilarity on Markov processes over generic measurable spaces is an equivalence, and (iii) a formal coalgebraic analysis on the relations between the bisimulation and cocongruence on labelled Markov processes, done establishing a formal adjunction between the category of bisimulations and (a subcategory) cocongruences. This adjunction is then proved to induce an equivalence between two suitable subcategories of bisimulations and cocongruences. A consequence of this equivalence is that bisimilarity (i.e., the final object in the category of bisimulations) is well-behaved with respect to behavioral equivalence (i.e. the final cocongruence). Moreover, this establishes sufficient conditions for a bisimulation to "coincide" with a cocongruence. Remarkably, all is proven without assuming that the state space of Markov processes is analytic. These results together with the counterexample due to Terraf [80], that proves that state bisimilarity does not coincides to event bisimilarity for Markov processes over generic measurable spaces, conclude the comparison between these two notions of equivalence.

Chapter 6. In this chapter we consider the problem of modeling syntax and semantics of both probabilistic and stochastic processes with continuous states, i.e. generalized Markov processes.

The main contributions are (i) the definition of a *syntactic* rule format that allows for an easy description of well-behaved semantics for continuos state probabilistic and stochastic Markov processes, (ii) a proof that this rule format induces an abstract GSOS distributive law of a monad over a copointed functor in **Meas** that adheres the bialgebraic framework of Turi and Plotkin [83], and (iii) a technique for the definition of *measures terms interpretations*, that is, natural transformations in **Meas** aimed at giving denotation to measure terms, i.e., expressions specifically designed for describing measures over generic measurable spaces which are employed by the syntactic rule format.

As an example application, we model a CCS-like calculus of processes placed in an Euclidean space. The approach we follow in this case can be readily adapted to other quantitative aspects, e.g. Quality of Service, physical and chemical parameter in biological systems, etc.

Chapter 7. This chapter deals with the problem of exactly computing bisimilarity distances between discrete-time Markov chains introduced by Desharnais et al. [39].

The main contribution consists in the definition and implementation of an efficient on-the-fly algorithm which, unlike other existing solutions, computes exactly the distances between given states and avoids the exhaustive state space exploration. Our technique successively refines over-approximations of the target distances using a greedy strategy which ensures that the state space is further explored only when the current approximations are improved. The efficiency of our algorithm is supported by experimental results, showed in the last section of the chapter, which prove that our algorithm improves, in average, the the execution time of the approximated iterative algorithms.

These results are the fruit of a collaboration with Giovanni Bacci, Radu Mardare and Kim G. Larsen, and have been supported by Sapere Aude: DFF-Young Researchers Grant 10-085054 of the Danish Council for Independent Research, by the VKR Center of Excellence MT-LAB and by the Sino-Danish Basic Research Center IDEA4CPS.

<div align="right">

2

</div>

Preliminaries

2.1 Category theory: definitions and notation

In this section, we recall the basic definitions from category theory that will be used in the thesis. In the following we will assume some familiarity with the notions of category, functors, natural transformations, and commutative diagrams. As a reference we recommend to consult the textbooks by Mac Lane [64] or Borceux [19], but shorter introductions are good as well.

Limits and Colimits. Category theory is the study of universal properties. The most primitive universal property is initiality. An object X in a category $\mathbf{C}$ is *initial* if for every object Y in $\mathbf{C}$ there exists a unique arrow $i \colon X \to Y$. Every notion in category theory can be dualized reversing the direction of the arrows. For example, the notion dual to that of initial object is *final object*, that is, an object X in $\mathbf{C}$ such that for every object Y in $\mathbf{C}$ there exists a unique arrow $f \colon Y \to X$.

Other universal objects are products and coproducts. The *binary product* between two objects X and Y is $\mathbf{C}$ is a triple $(X \times Y, \pi_X, \pi_Y)$ consisting of an object $X \times Y$ in $\mathbf{C}$ and a pair of arrows $\pi_X \colon X \times Y \to X$ and $\pi_Y \colon X \times Y \to Y$, called projections, such that, for every other pair of arrows $f \colon Z \to X$ and $g \colon Z \to Y$, there exists a unique arrow $\langle f, g \rangle \colon Z \to X \times Y$ such that, $f = \pi_X \circ \langle f, g \rangle$ and $g = \pi_Y \circ \langle f, g \rangle$. A *binary coproduct* between two objects X and Y in $\mathbf{C}$ is a triple $(X + Y, in_X, in_Y)$ consisting of an object $X + Y$ in $\mathbf{C}$ and a pair of arrows $in_X \colon X \to X + Y$ and $in_Y \colon Y \to X + Y$, called injections, such that, for every other pair of arrows $f \colon X \to Z$ and $g \colon Y \to Z$, there exists a unique arrow $[f, g] \colon X + Y \to Z$ such that, $f = [f, g] \circ in_X$ and $g = [f, g] \circ in_Y$. These properties are represented diagrammatically as follows:

where dashed lines denote unique arrows.

Other kind of universal objects are pullback and pushouts. The *pullback* of a pair of arrows $f \colon X \to C$ and $g \colon Y \to C$ is a triple (P, p_X, p_Y) consisting of an object P and arrows $p_X \colon P \to X$ and $p_Y \colon P \to Y$ such that $f \circ p_X = g \circ p_Y$ and, for any pair of arrows $q_X \colon Q \to X$ and $q_Y \colon Q \to Y$ such that $f \circ q_X = g \circ q_Y$, there exists a unique morphism $h \colon Q \to P$ such that $q_X = p_X \circ h$ and $q_Y = p_Y \circ h$. The *pushout* of a pair of arrows $f \colon C \to X$ and $g \colon C \to Y$ is a triple (K, k_X, k_Y), with an object K and arrows $k_X \colon X \to K$ and $k_Y \colon Y \to K$ such that $k_X \circ f = k_Y \circ f$ and, for any pair of arrows $q_X \colon X \to Q$ and $q_Y \colon Y \to Q$ such that $q_X \circ f = q_Y \circ f$, there exists a unique

morphism $h\colon K \to Q$ such that $q_X = h \circ k_X$ and $q_Y = h \circ k_Y$. Diagrammatically:
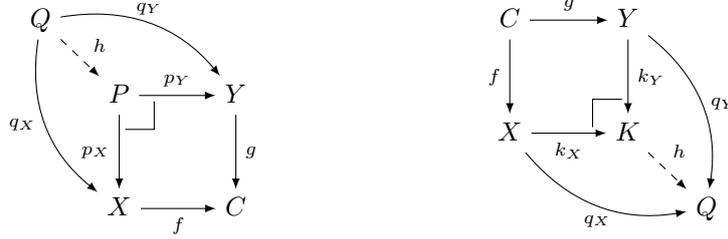
All the above universal objects are just particular cases of the more general notions of *limit* and *colimit over a diagram*. Formally, a *diagram* of type $\mathbf{J}$ in $\mathbf{C}$ is a functor $D\colon \mathbf{J} \to \mathbf{C}$. Intuitively, $\mathbf{J}$ can be thought of as an index category and D as a mapping to morphisms in $\mathbf{C}$ patterned on $\mathbf{J}$.

Definition 2.1.1 (Limit cone) *Let $D\colon \mathbf{J} \to \mathbf{C}$ be a diagram. A* cone *over D is a collection $(h_X\colon U \to DX)_{X\in\mathbf{J}}$ of morphisms in $\mathbf{C}$, such that, for every arrow $f\colon X \to Y$ in $\mathbf{J}$, it holds $Df \circ h_X = h_Y$. A cone $(h_X\colon U \to DX)_{X\in\mathbf{J}}$ is a* limit, *if for any cone $(k_X\colon V \to DX)_{X\in\mathbf{J}}$ over D there exists a unique arrow $u\colon V \to U$, such that $h_X \circ u = k_X$, for all objects X in $\mathbf{J}$.*

Definition 2.1.2 (Colimit cocone) *Let $D\colon \mathbf{J} \to \mathbf{C}$ be a diagram. A* cocone *over D, is a collection $(h_X\colon DX \to U)_{X\in\mathbf{J}}$ of morphisms in $\mathbf{C}$, such that, for every arrow $f\colon X \to Y$ in $\mathbf{J}$, it holds $h_X = h_Y \circ Df$. A cocone $(h_X\colon DX \to U)_{X\in\mathbf{J}}$ is a* colimit, *if for any cocone $(k_X\colon DX \to V)_{X\in\mathbf{J}}$ over D there exists a unique arrow $u\colon U \to V$, such that $u \circ h_X = k_X$, for all objects X in $\mathbf{J}$.*

Examples of limits are final objects, products, and pullbacks, and of colimits initial objects, coproducts, and pushouts. Limits and colimits over $D\colon \mathbf{J} \to \mathbf{C}$ are said *small* if the index category $\mathbf{J}$ has a proper set of objects. A category $\mathbf{C}$ is *complete* if it has all small limits, *cocomplete* if is has all small colimits.

Adjoint functors. There are various definitions for adjoint functors. Their equivalence is elementary but not trivial at all. We recall some of them: via unit and counit laws, via universal morphisms, and via isomorphism of homsets.

Definition 2.1.3 (Adjunction) *An* adjunction *between two functors $F\colon \mathbf{C} \to \mathbf{D}$ and $G\colon \mathbf{D} \to \mathbf{C}$ consists of two natural transformations $\eta\colon Id_{\mathbf{C}} \Rightarrow GF$ and $\epsilon\colon FG \Rightarrow Id_{\mathbf{D}}$, respectively called* unit *and* counit *of the adjunction, satisfying the following composition laws:*

We write $(\eta,\epsilon)\colon F \dashv G$, or simply $F \dashv G$, when there is an adjunction between F and G. This is also described by saying that F is the left adjoint *of G, or G is the* right adjoint *of F.*

An alternative characterization can be given in terms of universal morphisms. In particular, it turns out that each component of η, the unit of the adjunction, is a G-initial arrow and, dually, each component of the counit ϵ is an F-final arrow. The existence of unique initial and final morphisms is used in the so called *free* and *cofree universal construction* of morphisms, respectively.

Theorem 2.1.4 (Universal morphisms) *Let $F\colon \mathbf{C} \to \mathbf{D}$ and $G\colon \mathbf{D} \to \mathbf{C}$ be two functors. Then, $(\eta,\epsilon)\colon F \dashv G$ is equivalent to the following statements:*

i. Free construction: for any pair of objects X in $\mathbf{C}$, Y in $\mathbf{D}$, and any arrow $f\colon X \to GY$ in $\mathbf{C}$, there exists a unique arrow $f^{\#}\colon FX \to Y$ in $\mathbf{D}$ such that the following diagram commutes

$$
\begin{array}{ccc}
\mathbf{C} & \underset{G}{\overset{F}{\rightleftarrows}} & \mathbf{D} \\
\end{array}
$$

$$
\begin{array}{ccc}
X \xrightarrow{\eta_X} GFX & & FX \\
\quad\searrow{\scriptstyle f} \quad \downarrow{\scriptstyle Gf^{\#}} & & \downarrow{\scriptstyle f^{\#}} \\
GY & & Y
\end{array}
$$

ii. Cofree construction: for any pair of objects X in $\mathbf{C}$, Y in $\mathbf{D}$, and any arrow $f\colon FX \to Y$ in $\mathbf{D}$, there exists a unique arrow $f^{\flat}\colon X \to GY$ in $\mathbf{C}$ such that the following diagram commutes

$$
\begin{array}{ccc}
\mathbf{D} & \underset{F}{\overset{G}{\rightleftarrows}} & \mathbf{C} \\
\end{array}
$$

$$
\begin{array}{ccc}
FX & & X \\
{\scriptstyle f}\swarrow \quad \downarrow{\scriptstyle Ff^{\flat}} & & \downarrow{\scriptstyle f^{\flat}} \\
Y \xleftarrow{\epsilon_Y} FGY & & GY
\end{array}
$$

Moreover, $F \dashv G$ is equivalent to have a natural isomorphism between arrows of type $FX \to Y$ and $X \to GY$, for all objects X in $\mathbf{C}$ and Y in $\mathbf{D}$. This can be formalized via the *homset (bi)functor* $Hom_{\mathbf{C}}\colon \mathbf{C}^{op} \times \mathbf{C} \to \mathbf{Set}$, defined by

$$Hom_{\mathbf{C}}(X,Y) = \{f \mid f\colon X \to Y \ \mathbf{C}\text{-morphism}\}$$
$$Hom_{\mathbf{C}}(f,g)(h\colon X \to Y) = g \circ h \circ f\colon X' \to Y'\,.$$

for all objects X, Y in $\mathbf{C}$ and arrows $f\colon X' \to X$, $g\colon Y \to Y'$ in $\mathbf{C}$. Indeed, the composites $Hom_{\mathbf{D}}(F, Id) := (F^{op} \times Id_{\mathbf{D}}) \circ Hom_{\mathbf{D}}$ and $Hom_{\mathbf{C}}(Id, G) := (Id_{\mathbf{C}}^{op} \times G) \circ Hom_{\mathbf{C}}$ describe the type of morphisms we are looking for. Formally:

Theorem 2.1.5 (Isomorphism of homsets) *Let $F\colon \mathbf{C} \to \mathbf{D}$ and $G\colon \mathbf{D} \to \mathbf{C}$ be two functors. Then, $F \dashv G$ if and only if there exists a natural isomorphism $\theta\colon Hom_{\mathbf{D}}(F, Id) \Rightarrow Hom_{\mathbf{C}}(Id, G)$. Explicitly, for all $f\colon X' \to X$ in $\mathbf{C}$ and $g\colon Y \to Y'$ in $\mathbf{D}$, the diagram below commutes*

$$
\begin{array}{ccc}
Hom_{\mathbf{D}}(FX,Y) & \xrightarrow[\cong]{\theta_{X,Y}} & Hom_{\mathbf{C}}(X,GY) \\
{\scriptstyle Hom_{\mathbf{D}}(Ff,g)}\downarrow & & \downarrow{\scriptstyle Hom_{\mathbf{C}}(f,Gg)} \\
Hom_{\mathbf{D}}(FX',Y') & \xrightarrow[\theta_{X',Y'}]{\cong} & Hom_{\mathbf{C}}(X',GY')
\end{array}
$$

Monads and Comonads. Monads are one of the most general mathematical tools. For instance, *every* algebraic theory, that is, every set of operations satisfying equational laws, can be seen as a monad; and algebraic theories are only a minor source of monads. In fact, every "canonical" (or universal) construction between two categories give rise to a monad.

Definition 2.1.6 (Monad) *A* monad *in $\mathbf{C}$ is a triple (T, η, μ) of a functor $T\colon \mathbf{C} \to \mathbf{C}$ and two natural transformations $\eta\colon Id \Rightarrow T$ and $\mu\colon TT \Rightarrow T$, called the* unit *and* multiplication *respectively, such that the three diagrams below commute*

$$
\begin{array}{ccc}
T \xrightarrow{\eta T} TT & TT \xleftarrow{T\eta} T & TTT \xrightarrow{T\mu} TT \\
{\scriptstyle id}\searrow \ \downarrow{\scriptstyle \mu} & {\scriptstyle \mu}\downarrow \ \swarrow{\scriptstyle id} & {\scriptstyle \mu T}\downarrow \qquad \downarrow{\scriptstyle \mu} \\
T & T & TT \xrightarrow{\mu} T
\end{array}
$$

The first two diagrams represent the unit laws of the monad, *and the third the* multiplication law of the monad.

Intuitively, a monad (T, η, μ) on **C** can be understood as a monoid in the category of endofunctors on **C**, the "operation" μ being the associative multiplication of the monoid and η its unit.

Definition 2.1.7 (Comonad) *A comonad in* **C** *is a triple* (D, ϵ, ξ) *of a functor* $D \colon \mathbf{C} \to \mathbf{C}$ *and two natural transformations* $\eta \colon D \Rightarrow Id$ *and* $\xi \colon D \Rightarrow DD$, *called the counit and comultiplication respectively, such that the three diagrams below commute*

$$
\begin{array}{ccc}
D & \quad & D \\
\end{array}
$$

The first two diagrams represent the counit laws of the comonad, *and the third the* comultiplication law of the comonad.

Dually to the case of monads, comonads (D, ϵ, ξ) in **C** can intuitively be understood as *comonoids* in the category of endofunctors on **C**, the "operation" ξ is an associative de-constructor, decomposing the structure given in the shape of the functor D into the (de)-composite DD, and ϵ the counit (or destructor). Comonads have been extensively used in the definitions of non-well-funded dynamical data structures, such as directed containers, infinite streams, infinite trees, etc.

An important result, which will be extensively used in the rest of the thesis, is that every adjunction between two categories gives rise to both a monad and a comonad.

Theorem 2.1.8 *Let* **C** *and* **D** *be two categories,* $F \colon \mathbf{C} \to \mathbf{D}$ *and* $G \colon \mathbf{D} \to \mathbf{C}$ *be adjoint functors* $(\eta, \epsilon) \colon F \dashv G$. *Then, the triple* $(GF, \eta, G\epsilon F)$ *is a monad and the triple* $(FG, \epsilon, F\eta G)$ *a comonad.*

Monic spans and Epic cospans. The categorical generalization of relations $R \subseteq X \times Y$ in **Set** in an arbitrary category **C**, are *monic spans* between objects X and Y in **C**, that is, triples (R, f, g) with R an object in **C** and $f \colon R \to X$, $g \colon R \to Y$ a pair of **C**-morphisms, such that they are *jointly-monic* (i.e., given any pair of morphisms $h, k \colon Z \to R$ in **C**, it holds that $f \circ h = f \circ k$ and $g \circ h = g \circ k$ implies $h = k$). Note that, in categories with binary products, (R, f, g) is a mono span if and only if the canonical morphism $\langle f, g \rangle \colon R \to X \times Y$ is monic, i.e., R is a proper sub-object of $X \times Y$.

The dual notion is given by *epic cospans*, that is, triples (K, f, g) with K an object in **C**, and $f \colon X \to K$, $g \colon Y \to K$ a pair of morphisms in **C**, such that they are *jointly-epic* (i.e., given any pair of morphisms $h, k \colon K \to Z$ in **C**, it holds that $h \circ f = k \circ f$ and $h \circ g = k \circ g$ implies $h = k$). In categories with binary coproducts, (K, f, g) is an epic cospan if and only if the canonical morphism $[f, g] \colon X + Y \to K$ is epic, i.e., K is a proper quotient of $X + Y$. In **Set**, epic cospans over the sets X and Y are in one-to-one correspondence with quotients $(X + Y)/E$ and their canonical injections, where E is an equivalence relation on $X + Y$.

Factorization systems. The categorical generalization of the notion of subset is that of *sub-object*, that is a monic arrow $X' \to X$, so that the object X' is said a sub-object of X. Dually, the generalization of set-quotients is given by the notion of quotients, that is epic arrows $X \to X'$, and in this case the object X' is said a *quotient* of X.

In certain situations, the above notions of sub-object and quotient are too strong and sometimes inadequate. These can be further generalized (and relaxed) using factorization systems. The idea behind the definition of a factorization system is to axiomatize the essential properties of sub-objects and quotients, so that they could be found also in categories with "too few" monic or epic arrows.

Definition 2.1.9 (Factorization system) *A pair* $(\mathcal{L}, \mathcal{R})$ *of classes of morphisms in* $\mathbf{C}$ *is a factorization system if it obeys the following axioms:*

(i) $\mathcal{L}$ *and* $\mathcal{R}$ *are closed under composition with isomorphisms;*

(ii) *every morphism* f *in* $\mathbf{C}$ *factors as* $f = \rho \circ \lambda$, *for some* $\rho \in \mathcal{R}$ *and* $\lambda \in \mathcal{L}$;

(iii) *each lifting problem, i.e., a commutative square* $\rho \circ f = g \circ \lambda$, *where* $\rho \in \mathcal{R}$ *and* $\lambda \in \mathcal{L}$, *has a unique solution* d, *that is, an arrow such that the following diagram commutes:*

$$
\begin{array}{ccc}
A & \xrightarrow{\lambda \in \mathcal{L}} & B \\
{\scriptstyle f}\downarrow & \nearrow^{d} & \downarrow{\scriptstyle g} \\
C & \xrightarrow[\rho \in \mathcal{R}]{} & D
\end{array}
\qquad \text{(lifting problem)}
$$

Example 2.1.10 Typical examples of factorizations systems are the following.

(i) ($\mathbf{Epic}, \mathbf{Monic}$) is a factorization system on $\mathbf{Set}$, where $\mathbf{Epic}$ is the class of epic arrows (hence, surjective set-maps) and $\mathbf{Monic}$ is the class of monic arrows (hence, injective set-maps). In particular, in $\mathbf{Set}$ all epic arrows are also extreme and strong epic, and all monic arrows are also extreme and strong monic.

(ii) ($\mathbf{Epic}, \mathbf{Monic}$) forms a factorization system on $\mathbf{Grp}$, the category of groups and homomorphisms between them. The class $\mathbf{Epic}$ is composed by all surjective homomorphisms, and $\mathbf{Monic}$ by all injective homomorphisms. This factorization system, in particular, is a "lift" of the factorization system in $\mathbf{Set}$ along the obvious forgetful functor $U \colon \mathbf{Grp} \to \mathbf{Set}$ that maps a group $(X, \cdot)$ to its underlying set X, forgetting the group operation $\cdot \colon X \times X \to X$.

(iii) The category $\mathbf{Top}$, of topological spaces and continuous maps, has many factorization systems that "lift" the epic-monic factorization system in $\mathbf{Set}$. e.g., quotients and injections form a factorization system on $\mathbf{Top}$ and so do surjections and subspace embeddings.

In $\mathbf{Top}$, there is a further factorization system $(\mathcal{L}, \mathcal{R})$, where $\mathcal{L}$ does not consist of epimorphisms: just take $\mathcal{R}$ as of the class of *closed* subspace embeddings and $\mathcal{L}$ as the class of dense maps. ∎

Remark 2.1.11 In "sufficiently complete" categories, certain factorization systems come for free. In categories with *intersections* and finite limits, extremal epimorphisms and monomorphisms form a factorization system, and do so epimorphisms and extremal monomorphisms in categories with intersections and equalizers. Here the *intersection* of a family of monomorphisms ('subobjects') $m_i \colon B_i \to A$, indexed over a set I, is the limit of the obvious associated diagram. ∎

Remark 2.1.12 In a complete and well-powered category $\mathbf{C}$, take $\mathcal{R}$ as the class of monomorphisms, and $\mathcal{L}$ as the class of strong epimorphisms. Then, the pair $(\mathcal{E}, \mathcal{M})$ defines an orthogonal factorization system on $\mathbf{C}$ (see [19, Prop. 4.4.2 and 4.4.3]). ∎

Orthogonal factorization systems enjoy several properties, which we will recall after having introduced some preliminary (standard) notations.

Given classes of morphisms $\mathcal{L}$ and $\mathcal{R}$, we say that $\mathcal{L}$ is *orthogonal* to $\mathcal{R}$, written $\mathcal{L} \perp \mathcal{R}$, if for all lifting problems $\rho \circ f = g \circ \lambda$, with $\lambda \in \mathcal{L}$ and $\rho \in \mathcal{R}$, there exists a unique solution. Note, that for a factorization system $(\mathcal{L}, \mathcal{R})$ it always holds that $\mathcal{L} \perp \mathcal{R}$. Let also denote by $\mathcal{L}^{\perp}$ and $\mathcal{R}_{\perp}$ the following classes of morphisms

$$
\mathcal{L}^{\perp} = \{ m \in \mathbf{Morph_C} \mid \mathcal{L} \perp m \}, \qquad \mathcal{R}_{\perp} = \{ m \in \mathbf{Morph_C} \mid m \perp \mathcal{R} \}.
$$

Then, the following alternative characterization for factorization systems holds.

Proposition 2.1.13 $(\mathcal{L}, \mathcal{R})$ *is a factorization system iff the following conditions are satisfied:*

(i) *every morphism* f *in* $\mathbf{C}$ *factors as* $f = \rho \circ \lambda$, *for some* $\rho \in \mathcal{R}$ *and* $\lambda \in \mathcal{L}$;

(ii) $\mathcal{L} = \mathcal{R}_{\perp}$ *and* $\mathcal{R} = \mathcal{L}^{\perp}$.

Of course, this can be taken as an alternative definition of factorization system. What is important here, is that $\mathcal{L} = \mathcal{R}_\perp$ means that in order to prove that a morphism l is in $\mathcal{L}$ it suffices to prove that it is orthogonal to any morphism in $\mathcal{R}$. An obvious corollary is that each class of a factorization system determines the other.

Both classes of the morphisms of a factorization system enjoy a cancellation property:

Lemma 2.1.14 *If $(\mathcal{L}, \mathcal{R})$ is a factorization system, then $\mathcal{L}$ has the* right cancellation property *and $\mathcal{R}$ the* left cancellation property*, that is*

$$g \circ f \in \mathcal{L} \wedge f \in \mathcal{L} \implies g \in \mathcal{L} \qquad\qquad g \circ f \in \mathcal{R} \wedge g \in \mathcal{R} \implies f \in \mathcal{R}$$

Lemma 2.1.15 *If $(\mathcal{L}, \mathcal{R})$ is a factorization system, then $\mathcal{L} \cap \mathcal{R}$ is the class of all isomorphisms.*

The classes $\mathcal{L}$ and $\mathcal{R}$ enjoy many closure properties. In the following we recall some of them.

Lemma 2.1.16 *If $(\mathcal{L}, \mathcal{R})$ is a factorization system, then $\mathcal{L}$ and $\mathcal{R}$ are closed under compositions.*

The classes $\mathcal{L}$ and $\mathcal{R}$ of morphisms of a factorization system in $\mathbf{C}$ behaves well with colimits and limits (taken in the category of arrows of $\mathbf{C}$).

Theorem 2.1.17 *If $(\mathcal{L}, \mathcal{R})$ is a factorization system in $\mathbf{C}$, then $\mathcal{L}$ is closed under all colimits and, dually, $\mathcal{R}$ is closed under all limits.*

That is, if $\mathbf{C}$ has coproducts and $f, g \in \mathcal{L}$, then $f + g \in \mathcal{L}$, where $f + g$ is the coproduct of arrows. Dually, if $\mathbf{C}$ has products, $f, g \in \mathcal{R}$ implies that $f \times g \in \mathcal{R}$, where $f \times g$ is the product of arrows.

Lemma 2.1.18 (Closure under pullback and pushouts) *Let $(\mathcal{L}, \mathcal{R})$ be a factorization system in a category $\mathbf{C}$ with pullback and pushouts. Then $\mathcal{L}$ is closed under pushouts and, dually, $\mathcal{R}$ is closed under pullbacks i.e., if the two diagrams below are respectively a pushout and a pullback, then if $\lambda \in \mathcal{L}$ and $\rho \in \mathcal{R}$, then also $\lambda' \in \mathcal{L}$ and $\rho' \in \mathcal{R}$.*

Lemma 2.1.19 (Closure under transfinite compositions & precompositions) *Let α be a non-empty ordinal, regarded as a category, and $F \colon \alpha \to \mathbf{C}$ is a colimit preserving functor and $G \colon \alpha^{op} \to \mathbf{C}$ is a limit preserving functor. Assume $(\mathcal{L}, \mathcal{R})$ is a factorization system in $\mathbf{C}$, then for every limit ordinal $\beta \leq \alpha$ the following implications hold*

$$\{F(\delta \to \gamma) \mid \delta \leq \gamma < \beta\} \subseteq \mathcal{L} \implies F(0 \to \beta), \qquad\qquad \text{(transfinite composition)}$$
$$\{G(\gamma \to \delta) \mid \delta \leq \gamma < \beta\} \subseteq \mathcal{R} \implies G(\beta \to 0). \qquad\qquad \text{(transfinite precomposition)}$$

Remark 2.1.20 Moreover, note that all the closure properties that have been considered so far in this section are all satisfied by monomorphisms and epimorphisms, hence we can say that factorization systems is the right axiomatization that extends the properties of monomorphisms and epimorphisms. ∎

Union and intersections in factorization systems.   We already seen that factorization systems generalize the concepts of subject and quotient for an object. Here we see how the concepts of union and intersection of subobjects and counion and cointersection of quotients are generalized to generic factorization systems. These constructions can be found in [19, Section 4.2].

Given and object Z in a category $\mathbf{C}$ with factorization system $(\mathcal{L}, \mathcal{R})$, the collection of all $\mathcal{L}$-morphisms with domain Z, and the collection of all $\mathcal{R}$-morphisms with codomain Z form two

categories, denoted by $\mathcal{L}(Z)$ and $\mathcal{R}(Z)$, respectively, with morphisms $f\colon (Z \xrightarrow{\lambda_1} X) \to (Z \xrightarrow{\lambda_2} Y)$ in $\mathcal{L}(Z)$ and $g\colon (X \xrightarrow{\rho_1} Z) \to (Y \xrightarrow{\rho_2} Z)$ in $\mathcal{R}(Z)$ such diagrams below commute in $\mathbf{C}$

Note that, by Lemma 2.1.14, $f \in \mathcal{L}$ and $g \in \mathcal{R}$. With an abuse of notions objects $(Z \xrightarrow{\lambda} X)$ in $\mathcal{L}(Z)$ and $(X \xrightarrow{\rho} Z)$ in $\mathcal{R}(Z)$ will be identified simply by X. This natation does not make confusion since it holds that, two objects X and Y are isomorphic iff there exist a pair of morphisms $f\colon X \to Y$ and $g\colon Y \to X$. This follows noticing that $f \circ g$ and id_Y fit as the unique solution of the same lifting problem, and the same thing happen to $g \circ f$ and id_X, so that f and g are inverses.

Definition 2.1.21 (Union and intersection) *Let $(\mathcal{R}, \mathcal{L})$ be a factorization system in $\mathbf{C}$. The* union *of a family of objects in $\mathcal{R}(Z)$, is their coproduct in $\mathcal{R}(Z)$ (if it exists), and the* intersection *of a family of objects in $\mathcal{R}(Z)$, is their product in $\mathcal{R}(Z)$ (if it exists).*

Therefore, unions and intersections are universal objects in $\mathcal{R}(Z)$. The existence of unions and intersections is ensured by the existence of coproducts and (generalized) pullbacks in the underlying category: assume $(\mathcal{L}, \mathcal{R})$ is a factorization system in a category $\mathbf{C}$ with pullbacks and coproducts, then the union and intersection of X and Y in $\mathcal{R}(Z)$ is respectively given by U and I

where in the diagram on the left (I, ρ'_X, ρ'_Y) is the pullback of the pair (ρ_X, ρ_Y), and in the diagram on the right $\rho_U \circ \lambda_U$ is the $(\mathcal{L}, \mathcal{R})$ factorization of the (unique) arrow $[\rho_X, \rho_Y]$ from the coproduct $X + Y$. Indeed, by Lemma 2.1.18, the composites $\lambda_X \circ \lambda'_X = \lambda_Y \circ \lambda'_Y$ are in $\mathcal{R}$, hence I is an object in $\mathcal{R}(Z)$ and it is the intersection of X and Y (with projections given by λ'_X and λ'_Y) by the universal properties of pullbacks. Similarly, U is the union of X and Y (with injections $\lambda_U \circ \in_X$ and $\lambda_U \circ in_Y$) by the universal property of the coproduct.

The above can be easily dualized in the case of $\mathcal{L}(Z)$, where we have the following definitions of counion and cointersection of collections of objects in $\mathcal{L}(Z)$.

Definition 2.1.22 (Counion and cointersection) *Let $(\mathcal{R}, \mathcal{L})$ be a factorization system in $\mathbf{C}$. The* counion *of a family of objects in $\mathcal{L}(Z)$, is their coproduct in $\mathcal{L}(Z)$ (if it exists), and the* cointersection *of a family of objects in $\mathcal{L}(Z)$, is their product in $\mathcal{L}(Z)$ (if it exists).*

Of course, by a dual argument, counions and cointersections exist in $\mathcal{L}(Z)$ if the underlying category has pullbacks an products.

2.2 Measure Theory

In this section, we recall the basic definitions of measurable space, measurable functions, and measure, and we provide all the tools which will be used in the thesis for dealing with measurable spaces and measures on them. This introductory exposition is very far from been exhaustive, henceforth if some notations or definitions are not sufficiently clear, we suggest to consult the

textbooks by Halmos [53] or Dudley [44], but shorter introductions are still adequate, such as the lecture notes by Tao [79] or the introductory chapters in the books by Panangaden [69] and Dokerkat [42].

Definition 2.2.1 (Measurable Space) *A* measurable space *is a pair* (X, Σ)*, where X is a set and Σ is σ-algebra on X, that is, a collection of subsets of X satisfying the following conditions:*

> *(S.1) $\emptyset \in \Sigma$;*
> *(S.2) if $E \in \Sigma$, then $E^c = X \setminus E \in \Sigma$;*
> *(S.3) if $E_0, E_1, E_2, \ldots \in \Sigma$, then $\bigcup_{n \in \mathbb{N}} E_n \in \Sigma$;*

A subset $E \subseteq X$ is measurable *in (X, Σ) if $E \in \Sigma$.*

From these axioms, it follows that the σ-algebra is also closed under countable intersections (by applying De Morgan's laws), moreover, by (S.1) and (S.2), X is always measurable in (X, Σ). The concept of σ-algebra generalizes the more well-known concept of *boolean algebra*, which requires only closure under complements and *finite* unions. Indeed, by padding a finite union into a countable union by using the empty set, we see that every σ-algebra is a boolean algebra.

Every set X can be always endowed with two canonical σ-algebras: the *discrete* σ-algebra $\mathcal{P}(X)$, i.e., the collection all subsets of X, and the *indiscrete* σ-algebra, i.e., the collection of subsets whose only members are $\emptyset$ and X. The discrete and indiscrete σ-algebras are, respectively, the finest and the coarser σ-algebras which any set can be endowed with. Formally, let X be a set, and denote by $\mathcal{S}(X)$ the family of all the σ-algebras on X. The set $\mathcal{S}(X)$ can be naturally partially ordered by subset inclusion (for which the discrete and indiscrete σ-algebras are the top and the bottom elements, respectively). The following proposition shows that $\mathcal{S}(X)$ is closed under arbitrary intersections, so that $\mathcal{S}(X)$ is a complete lattice:

Proposition 2.2.2 (Intersection of σ-algebra) *Let X be a set and $\mathcal{S}$ an arbitrary large family of σ-algebras on X. Then $\bigcap \mathcal{S}$ is a σ-algebras for X.*

Proof. We have to check that $\bigcap \mathcal{S}$ satisfies the conditions (S.1), (S.2), and (S.3) of Definition 2.2.1. But since all σ-algebras in $\mathcal{S}$ satisfy those conditions, and each element in $\bigcap \mathcal{S}$ belongs to all the elements of $\mathcal{S}$, (S.1), (S.2), and (S.3) are obviously satisfied by $\bigcap \mathcal{S}$ as well. ∎

Corollary 2.2.3 *$\mathcal{S}(X)$ is a complete lattice with respect to subset inclusion.*

Proof. For any given family of σ-algebras $\mathcal{S} \subseteq \mathcal{S}(X)$, the greatest lower bound is given by $\bigcap \mathcal{S}$, which is in $\mathcal{S}(X)$ by Proposition 2.2.2. The least upper bound for $\mathcal{S}$, is just the greatest lower bound of the set of all upper bounds, i.e., $\bigcap \{\Sigma \mid \Sigma' \subseteq \Sigma, \text{ for all } \Sigma' \in \mathcal{S}\}$, which is again in $\mathcal{S}(X)$ by Proposition 2.2.2. ∎

Since the intersection plays the rôle of the greatest lower bound operator, one may be tempted to think that the least upper bound is given by the "union of σ-algebras". However, the union is not even a σ-algebra in general, as shown in Example 2.2.4 below:

Example 2.2.4 (Union of σ-algebras) Let $X = \{x_1, x_2, x_3\}$, and $\Sigma_1, \Sigma_2 \in \mathcal{S}(X)$ be given by

$$\Sigma_1 = \{\emptyset, \{x_3\}, \{x_1, x_2\}, X\}, \qquad\qquad \Sigma_2 = \{\emptyset, \{x_1\}, \{x_2, x_3\}, X\}.$$

The union $\Sigma_1 \cup \Sigma_2$ is not a σ-algebra. Indeed we have that $\{x_1\} \in \Sigma_1 \cup \Sigma_2$ and $\{x_3\} \in \Sigma_1 \cup \Sigma_2$ but $\{x_1, x_3\} \notin \Sigma_1 \cup \Sigma_2$, hence (S.3) is not satisfied. ∎

Proposition 2.2.2 allows one to consider the *smallest σ-algebra* satisfying a generic property as the intersection of all the σ-algebras on X satisfying it. In particular, we have the following definition:

Definition 2.2.5 (Generated σ-algebra) *Let $\mathcal{F}$ be a family of sets in X. We define $\sigma(\mathcal{F})$ to be the intersection of all the σ-algebras that contain $\mathcal{F}$, explicitly*

$$\sigma(\mathcal{F}) = \bigcap \{\Sigma \in \mathcal{S}(X) \mid \mathcal{F} \subseteq \Sigma\}. \qquad \text{(generated σ-algebra)}$$

If Σ is such that $\Sigma = \sigma(\mathcal{F})$, then we say that Σ is generated *by $\mathcal{F}$ and that $\mathcal{F}$ is its* generator.

Important examples of generated σ-algebras are the Borel σ-algebras, that is, the σ-algebras generated by the open (or closed) sets of a topological or metric spaces. For example, $\mathbb{R}$ with the Euclidean metric is turned into a measurable space by means of it Borel σ-algebra. This construction is so common that when $\mathbb{R}$ is considered as a measurable space and no particular σ-algebra has been specified for it, it is always meant to be the Borel σ-algebra generated by the open balls $B_\epsilon(r) = \{r' \in \mathbb{R} \mid |r - r'| < \epsilon\}$, for all $r \in \mathbb{R}$ and real radius $\epsilon > 0$.

Remark 2.2.6 The Borel σ-algebra on $\mathbb{R}$ can be generated also by the following collection of sets:
  (i) open/closed balls with real radius;
 (ii) compact subsets;
(iii) left/right open/closed intervals for all real numbers;
 (iv) open/closed balls with rational radius;
  (v) left/right open/closed intervals for all rational numbers.
Remarkably, the last item considers only a countable family of sets. These results do not hold in general for any metric space, but only for those which are *separable* or *compact*. ∎

Definition 2.2.7 (Measurable function) *Let (X, Σ_X) and (Y, Σ_Y) be two measurable spaces. A function $f \colon X \to Y$ is* measurable *if, for any $E \in \Sigma_Y$, $f^{-1}(E) = \{x \in X \mid f(x) \in E\} \in \Sigma_X$.*

When σ-algebras are generated by some family of subsets, checking measurability for a function is made easier by the observation:

Lemma 2.2.8 *Let (X, Σ_X) and (Y, Σ_Y) be measurable spaces, and assume that Σ_Y is generated by $\mathcal{F}$. Then $f \colon X \to Y$ is measurable if and only if $f^{-1}(F) \in \Sigma_X$ holds for all $F \in \mathcal{F}$.*

The following two propositions are useful when one is dealing with measurable functions.

Proposition 2.2.9 *Let X and Y be sets, $\mathcal{F}$ be some collection of subsets of Y, and $f \colon X \to Y$ be a function. Then $f^{-1}(\sigma(\mathcal{F})) = \sigma(f^{-1}(\mathcal{F}))$.*

Proof. A key fact that will be used in the proof is that the pre-image operation commutes with respect to all σ-algebra operations: if $f \colon X \to Y$ and $A_i, A, B \subseteq Y$, for all $i \in I$, then
  (i) $f^{-1}(\bigcup_{i \in I} A_i) = \bigcup_{i \in I} f^{-1}(A_i)$;
 (ii) $f^{-1}(A \setminus B) = f^{-1}(A) \setminus f^{-1}(B)$; in particular $f^{-1}(Y \setminus A) = X \setminus f^{-1}(A)$.
We prove the two inclusions separately.
   ($\supseteq$) Since $\sigma(f^{-1}(\mathcal{F}))$ is, by definition, the intersection of all σ-algebra over X containing $f^{-1}(\mathcal{F})$, the inclusion can be proved just showing that $f^{-1}(\sigma(\mathcal{F}))$ is a σ-algebra and that it contains $f^{-1}(\mathcal{F})$. Since $\mathcal{F} \subseteq \sigma(\mathcal{F})$, we have $f^{-1}(\mathcal{F}) \subseteq f^{-1}(\sigma(\mathcal{F}))$. To prove that $f^{-1}(\sigma(\mathcal{F}))$ is a σ-algebra we have to show that it contains the empty set $\emptyset$, and that is closed by countable unions and complements. $\emptyset = f^{-1}(\emptyset)$, and since $\emptyset \in \sigma(\mathcal{F})$, we have $\emptyset \in f^{-1}(\sigma(\mathcal{F}))$. Assume $F_1, F_2, \ldots \in f^{-1}(\sigma(\mathcal{F}))$, then, for each $n \in \mathbb{N}$, there exists $A_n \in \sigma(\mathcal{F})$, such that $F_n = f^{-1}(A_n)$. Since $\sigma(\mathcal{F})$ is closed by countable unions, $\bigcup_{n \in \mathbb{N}} A_n \in \sigma(\mathcal{F})$, and in particular $f^{-1}(\bigcup_{n \in \mathbb{N}} A_n) \in f^{-1}(\sigma(\mathcal{F}))$. By (i) above, $f^{-1}(\bigcup_{n \in \mathbb{N}} A_n) = \bigcup_{n \in \mathbb{N}} f^{-1}(A_n)$, thus $\bigcup_{n \in \mathbb{N}} F_n \in f^{-1}(\sigma(\mathcal{F}))$. Closure under complements is proved similarly, using (ii).
   ($\subseteq$) To prove the inclusion it suffices to show that $D = \{A \subseteq Y \mid f^{-1}(A) \in \sigma(f^{-1}(\mathcal{F}))\}$ is a σ-algebra containing $\mathcal{F}$. Indeed, if it is so, $\sigma(\mathcal{F}) \subseteq D$, hence, by definition of D, we have $f^{-1}(\sigma(\mathcal{F})) \subseteq \sigma(f^{-1}(\mathcal{F}))$. The inclusion $\mathcal{F} \subseteq D$, follows since $\mathcal{F} \in \sigma(\mathcal{F})$. It remains to prove that D is a σ-algebra. $\emptyset \in D$, since $\emptyset = f^{-1}(\emptyset)$, and $\emptyset \in \sigma(f^{-1}(\mathcal{F}))$. Let $A_1, A_2, \ldots \in D$, then, for all $n \in \mathbb{N}$, $f^{-1}(A_n) \in \sigma(f^{-1}(\mathcal{F}))$. By (i) and the fact that $\sigma(f^{-1}(\mathcal{F}))$ is closed by countable unions,

we have $f^{-1}(\bigcup_{n \in \mathbb{N}} A_n) = \bigcup_{n \in \mathbb{N}} f^{-1}(A_n) \in \sigma(f^{-1}(\mathcal{F}))$, therefore $\bigcup_{n \in \mathbb{N}} A_n \in D$. Assume $A \in D$, then $f^{-1}(A) \in \sigma(f^{-1}(\mathcal{F}))$. By (ii) and the fact that $\sigma(f^{-1}(\mathcal{F}))$ is closed under complements, $f^{-1}(Y \setminus A) = X \setminus f^{-1}(A) \in \sigma(f^{-1}(\mathcal{F}))$, therefore $Y \setminus A \in D$. ∎

Proposition 2.2.10 *Let X and Y be sets, $\mathcal{F}$ be some collection of subsets of X, and $f \colon X \to Y$ be a function. Then $\sigma(\{A \subseteq Y \mid f^{-1}(A) \in \mathcal{F}\}) = \{A \subseteq Y \mid f^{-1}(A) \in \sigma(\mathcal{F})\}$.*

Proof. Let $\mathcal{K} = \{A \subseteq Y \mid f^{-1}(A) \in \mathcal{F}\}$ and $\Sigma_Y = \{A \subseteq Y \mid f^{-1}(A) \in \sigma(\mathcal{F})\}$. We show that $\sigma(\mathcal{K}) = \Sigma_Y$ proving the two inclusions simultaneously. Since $\mathcal{F} = f^{-1}(\mathcal{K})$, we have that $\sigma(\mathcal{F}) = \sigma(f^{-1}(\mathcal{K}))$. By Proposition 2.2.9, $\sigma(f^{-1}(\mathcal{K})) = f^{-1}(\sigma(\mathcal{K}))$, hence $\sigma(\mathcal{F}) = f^{-1}(\sigma(\mathcal{K}))$. The following sequence of equivalences

$$A \in \Sigma_Y \iff f^{-1}(A) \in \sigma(\mathcal{F}) \iff f^{-1}(A) \in f^{-1}(\sigma(\mathcal{K})) \iff A \in \sigma(\mathcal{K}),$$

proves the equality. ∎

Given a measurable space (Y, Σ_Y), it happens frequently that for a set X and a function $f \colon X \to Y$, one wants to equip X with a σ-algebra Σ_X making f measurable. Also the dual case is of interest, that is, when the map has opposite direction. This can always be done in a canonical way considering the initial and final σ-algebras with respect to the function f:

Proposition 2.2.11 (Initial σ-algebra w.r.t. f) *Let $f \colon X \to Y$ be a function and let (Y, Σ_Y) be a measurable space. The collection $\{f^{-1}(E) \mid E \in \Sigma_Y\}$ form a σ-algebra on X. In particular, it is the smallest σ-algebra on X making f measurable.*

Particular cases are *subspace embeddings*: for a measurable space (X, Σ), the inclusion $i \colon X' \hookrightarrow X$ becomes measurable for a subset $X' \subseteq X$, when X' is endowed with the σ-algebra $\{X' \cap E \mid E \in \Sigma\}$, which, in particular, is the the initial σ-algebra w.r.t. i. In this case i is called (subspace) *embedding*.

Proposition 2.2.12 (Final σ-algebra w.r.t. f) *Let $f \colon X \to Y$ be a function and let (X, Σ_X) be a measurable space. The collection $\{A \subseteq Y \mid f^{-1}(A) \in \Sigma_X\}$ form a σ-algebra on Y. In particular, it is the largest σ-algebra on Y making f measurable.*

Examples are the *quotient spaces*: for a measurable space (X, Σ), the function $q \colon X \to X/\sim$ mapping elements in X to the equivalence classes $[x]_\sim$, for $\sim \subseteq X \times X$ an equivalence relation on X, can be rendered measurable if we equip $X/\sim$ with the final σ-algebras w.r.t. q. In this case q will be called *quotient* and $(X/\sim)$ quotient space.

Initial and final σ-algebras generalize in an obvious way to families of maps. Let (Y_i, Σ_{Y_i}) be measurable spaces and $f \colon X \to Y_i$ be maps, for all $i \in I$. The initial σ-algebra on X w.r.t. to $\{f \colon X \to Y_i \mid i \in I\}$, is given by $\sigma(\bigcup_{i \in I}\{f_i^{-1}(E) \mid E \in \Sigma_{Y_i}\})$. Dually, for (X_i, Σ_{X_i}) measurable spaces and $f \colon X \to Y_i$ maps, for all $i \in I$, the final σ-algebra on Y w.r.t. $\{f \colon X_i \to Y \mid i \in I\}$ is defined as $\bigcap_{i \in I}\{A \subseteq Y \mid f_i^{-1}(A) \in \Sigma_{X_i}\}$, i.e., the σ-algebra generated by the sets $A \subseteq Y$ for which $f_i^{-1}(A)$ is a measurable in (X, Σ), for all $i \in I$.

Lemma 2.2.13 *Let (X, Σ_X), (Y, Σ_Y), and (Z, Σ_Z) be measurable spaces and $f \colon X \to Y$ a map.*
(i) if Σ_X is initial w.r.t. f, a map $g \colon Z \to X$ is measurable iff the composite $f \circ g$ is so;
(ii) if Σ_Y is final w.r.t. f, a map $h \colon Y \to Z$ is measurable iff the composite $h \circ f$ is so.

An exhaustive definition of a σ-algebra may be complicated, and it is often advantageous to work with a generating collection. However, the lack of an explicit representation of the elements of a generated σ-algebra requires more efforts in handling such measurable sets. A very powerful tool used in these situations is the *monotone class theorem*.

Definition 2.2.14 (Monotone class) *A collection $\mathcal{M}$ of subsets of X is a* monotone class *if the following conditions hold, for all $A_0, A_1, A_2, \ldots \in \mathcal{M}$,*
(i) if $A_0 \subseteq A_1 \subseteq A_2 \subseteq \ldots$, then $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{M}$;
(ii) if $A_0 \supseteq A_1 \supseteq A_2 \supseteq \ldots$, then $\bigcap_{n \in \mathbb{N}} A_n \in \mathcal{M}$.

Clearly any σ-algebra is a monotone class. The intersection of monotone classes is again a monotone class, thus we can talk about the monotone class generated by a collection of sets $\mathcal{F}$ just as we did in Definition 2.2.5, and we denote it by $m(\mathcal{F})$.

Proposition 2.2.15 *Any σ-algebra is a monotone class and if a monotone class is also a boolean algebra, then it is a σ-algebra.*

Theorem 2.2.16 (Monotone class theorem) *If $\mathcal{A}$ is a boolean algebra, then $m(\mathcal{A}) = \sigma(\mathcal{A})$.*

Measures.   An important concept related to measure theory is that of *measure space*.

Definition 2.2.17 (Measure space) *Let (X, Σ) be a measurable space. A map $\mu\colon \Sigma \to [0, \infty]$ is a* measure *on (X, Σ) if it obeys to the following axioms:*
  (i) $\mu(\emptyset) = 0$;
  (ii) *whenever $E_0, E_1, E_2, \ldots \in \Sigma$ is a countable collection of disjoint measurable sets, then $\mu(\bigcup_{n\in\mathbb{N}} E_n) = \sum_{n\in\mathbb{N}} \mu(E_n)$.*
In this case the triplet (X, Σ, μ) is called measure space.

Note the distinction between a measure space and a measurable space. The latter has the *capability* to be equipped with a measure, but the former is *actually* equipped with a measure.

There are some distinctions on the types of measure one is used to work with. A measure μ on (X, Σ) is said of *probability* provided that $\mu(X) = 1$; of *subprobability* if $\mu(X) \le 1$; *finite* if $\mu(X) < \infty$; σ-*finite* if $X = \bigcup_{n\in\mathbb{N}} A_n$, for some countable collection $\{A_n \subset X \mid n \in \mathbb{N}\}$, and $\mu(A_n) < \infty$, for all $n \in \mathbb{N}$. Note that, σ-finiteness is weaker than simple finiteness, indeed it is not required that the whole space has finite measure, rather, it is just required that there exists a countable cover of it that for which each patch has finite measure. Clearly, a probability measure is also of subprobability, that, in turn, is finite and σ-finite.

Example 2.2.18 We recall some notable measures and operations on them.
  Dirac measure: Let $x \in X$ and Σ be an arbitrary σ-algebra on X. The map $\delta_x\colon \Sigma \to [0, \infty]$ defined by $\delta_x(E) = \chi_E(x)$, where $\chi_A\colon X \to \{0, 1\}$ is the characteristic function of $A \subseteq X$, is a measure on (X, Σ) and is called the *Dirac measure at x*.
  Zero measure: The constant zero map $\mathbf{0}\colon \Sigma \to [0, \infty]$ is a measure for any measurable space (X, Σ), and it is called the *zero measure*.
  Linear combinations of measures: Let $\mu, \nu\colon \Sigma \to [0, \infty]$ be measures on (X, Σ), then the map $\mu + \nu\colon \Sigma \to [0, \infty]$, defined by $(\mu + \nu)(E) = \mu(E) + \nu(E)$ is also a measure, as is $c\mu\colon \Sigma \to [0, \infty]$, defined as $(c\mu)(E) = c \cdot \mu(E)$, for $c \in [0, \infty]$.
  In the same way, any *finite linear combination* of measures $\mu_1, \mu_2, \ldots, \mu_n$ on (X, Σ)
  $$c_1\mu_1 + c_2\mu_2 + \cdots + c_n\mu_n\colon \Sigma \to [0, \infty] \qquad E \mapsto c_1\mu_1(E) + c_2\mu_2(E) + \cdots + c_n\mu_n(E)$$
  is a measure on (X, Σ).
  Countable combinations of measures: Let $\mu_0, \mu_1, \mu_2, \ldots$ be a countable collection of measures on (X, Σ), then the map $\sum_{n\in\mathbb{N}} \mu_n\colon \Sigma \to [0, \infty]$ defined as
  $$\left(\textstyle\sum_{n\in\mathbb{N}} \mu_n\right)(E) = \textstyle\sum_{n\in\mathbb{N}} \mu_n(E)\,,$$
  is a σ-additive map, hence it is a measure on (X, Σ). Note that, when X has at most a countable number of elements, then any measure can be expressed as a countable linear combination of the form $\sum_{x\in X} c_x \delta_x$, for some suitable set $\{c_x \in [0, \infty] \mid x \in X\}$ of coefficients. This characterization does not hold if X is more than countable. $\blacksquare$

Proposition 2.2.19 (Countable subadditivity) *Let (X, Σ) be a measurable space, μ be a measure on it, and $E_0, E_1, E_2, \ldots$ a countable collection of measurable set on (X, Σ), then*
$$\mu(\textstyle\bigcup_{n\in\mathbb{N}} E_n) \le \textstyle\sum_{n\in\mathbb{N}} \mu(E_n)\,.$$

Proposition 2.2.20 (Monotone convergence) *Let (X, Σ) be a measurable space, μ be a measure on it, and $E_0, E_1, E_2, \ldots$ a countable collection of measurable set on (X, Σ), then*

   (i) if $E_0 \subseteq E_1 \subseteq E_2 \subseteq \ldots$, then

$$\mu(\textstyle\bigcup_{n \in \mathbb{N}} E_n) = \lim_{n \to \infty} \mu(E_n) = \sup_{n \in \mathbb{N}} \mu(E_n) \, ;$$

   (ii) if $E_0 \supseteq E_1 \supseteq E_2 \supseteq \ldots$, and $\mu(E_n) < \infty$, for some $n \in \mathbb{N}$, then

$$\mu(\textstyle\bigcap_{n \in \mathbb{N}} E_n) = \lim_{n \to \infty} \mu(E_n) = \inf_{n \in \mathbb{N}} \mu(E_n) \, .$$

The two properties above are respectively called *upward* and *downward convergence*. Note that, the downward convergence fails if the hypothesis that $\mu(E_n) < \infty$ for at least one $n \in \mathbb{N}$ is dropped.

Outer measures, pre-measures, and product measures   So far we have focused on specific properties of countably additive measures. Now we consider the problem of constructing measures satisfying particular properties. One of the most powerful tools is the Charathéodory extension theorem, which allows one to construct measures from any outer measure.

   One can in turn construct outer measures from another concept known as a pre-measure. With these tools, one can start constructing many more measures, such as Lebesgue-Stieltjes measures, product measures, and Hausdorff measures.

Definition 2.2.21 (Outer measure) *An* outer measure *on a set X is a map $\mu^* \colon \mathcal{P}(X) \to [0, \infty]$ defined on all subsets of X, which obeys the following axioms:*

   (i) $\mu(\emptyset) = 0$;
   (ii) if $A \subseteq B$, then $\mu^*(A) \leq \mu^*(B)$;
   (iii) if $A_0, A_1, A_2, \ldots$ is a countable collection of subsets of X, then $\mu(\bigcup_{n \in \mathbb{N}} A_n) \leq \sum_{n \in \mathbb{N}} \mu(A_n)$.

Outer measures are also known as *exterior measures*. Note that, outer measures are weaker than measures in that they are merely countably subadditive, rather than countably additive. On the other hand, they are able to measure all subsets of X, whereas measures can only measure a σ-algebra of measurable sets.

Definition 2.2.22 (Charathéodory measurability) *Let μ^* be an outer measure on a set X. A set $E \subseteq X$ is said to be* Carathéodory measurable *with respect to μ^* if, for every set $A \subseteq X$, the following holds:*

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E) \, .$$

Roughly speaking, Carathéodory measurable sets are those sets which can "split" any other set and still get finite additivity for the outer measure. So, one may say that these are those sets who behave nicely with respect to the outer measure.

Remark 2.2.23 *Null sets* (or *negligible sets*), i.e., sets A such that $\mu^*(A) = 0$, are always Charathéodory measurable.   ∎

   There is a general construction which produces a σ-algebra and a measure defined on it from an outer measure. This is the Carathéodory extension theorem:

Theorem 2.2.24 (Carathéodory extension theorem) *Let X be a set and μ^* be an outer measure defined on X. Denote by Σ the collection of all subsets which are Carathéodory measureble w.r.t. μ^*. For all $E \in \Sigma$, define $\mu(E) = \mu^*(A)$, then (X, Σ, μ) is a measure space.*

Note that the theorem above states also that the collection of Carathéodory measurable sets is a σ-algebra and, moreover, that such a σ-algebra is the one on which the outer measure can be restricted to produces an actual measure.

   A general situation that one often has to cope with, is that the σ-algebra is already given (usually generated by some family of sets) and that the measure must be defined on such a σ-algebra and not on the one produced by the Carathéodory extension theorem. This problem, can be solved having resort to boolean algebras and to the notion of *pre-measure*.

Definition 2.2.25 (Pre-measure) *A pre-measure on a boolean algebra $\mathcal{A}$ is a finitely additive measure $\mu_0 \colon \mathcal{A} \to [0, \infty]$ with the additional property that whenever $A_0, A_1, A_2, \ldots$ is a countable disjoint collection sets in $\mathcal{A}$ such that $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$, then*

$$\mu_0(\textstyle\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} \mu_0(A_n) \,.$$

Roughly speaking, pre-measures are those finitely additive measures on a boolean algebra which may be extended to a σ-additive measure on $\mathcal{A}$ (note that $\bigcup_{n \in \mathbb{N}} A_n$ is not automatically in $\mathcal{A}$, since $\mathcal{A}$ is closed only under finite unions).

Clearly, the condition for a pre-measure to be already σ-additive is a necessary condition in order to extend it to a σ-additive measure. Using the Carathéodory extension theorem, it can be shown that this necessary condition is also sufficient. More precisely, we have

Theorem 2.2.26 (Hahn-Kolmogorov theorem) *Any pre-measure $\mu_0 \colon \mathcal{A} \to [0, \infty]$ on a boolean algebra $\mathcal{A}$ on X can be extended to a measure $\mu \colon \sigma(\mathcal{A}) \to [0, \infty]$ on $(X, \sigma(\mathcal{A}))$.*

Let us call the measure μ constructed in the above theorem the *Hahn-Kolmogorov extension of* μ_0. However, notice that, this extension is not unique in general. Unicity can be ensured requiring that the pre-measure is σ-additive.

Theorem 2.2.27 (Unicity) *If μ_0 is a σ-finite pre-measure on a boolean algebra $\mathcal{A}$. Then any extension of μ_0 on $\sigma(\mathcal{A})$ is unique.*

A well known example of measure defined using the Hahn-Kolmogorov extension theorem is the *product measure*. In the following example we give its definition and sketch its construction.

Example 2.2.28 (Product measure) Given two measurable spaces (X, Σ_X) and (Y, Σ_Y), one can define the product space $(X, \Sigma_X) \times (Y, \Sigma_Y) = (X \times Y, \Sigma_X \otimes \Sigma_Y)$, where $\Sigma_X \otimes \Sigma_Y$ is the σ-algebra generated by the rectangles $E \times F \subseteq \Sigma_X \times \Sigma_Y$. This σ-algebra is the smallest one that makes the canonical projections $\pi_X \colon X \times Y \to X$ and $\pi_Y \colon X \times Y \to Y$ measurable.

Given two measures $\mu \colon \Sigma_X \to [0, \infty]$ and $\nu \colon \Sigma_Y \to [0, \infty]$ on (X, Σ_X) and (Y, Σ_Y), respectively, one can define a measure $\mu \times \nu \colon \Sigma_X \otimes \Sigma_Y \to [0, \infty]$ on $(X \times Y, \Sigma_X \otimes \Sigma_Y)$ such that

$$(\mu \times \nu)(E \times F) = \mu(E)\nu(F) \qquad \text{for all } E \in \Sigma_X \text{ and } F \in \Sigma_Y.$$

The above measure is called *product measure*, and has the property that its left and right marginal are such that $(\mu \times \nu) \circ \pi_X^{-1} = \nu(Y)\mu$ and $(\mu \times \nu) \circ \pi_Y^{-1} = \mu(X)\nu$. The product measure is defined first as a pre-measure on the boolean algebra $\mathcal{A}$ of all finite unions of measurable rectangles $E \times F \subseteq \Sigma_X \times \Sigma_Y$, and then is extended to the σ-algebra generated by $\mathcal{A}$ using the Hahn-Kolmogorov theorem. It easy to see that $\sigma(\mathcal{A})$ is exactly the product σ-algebra $\Sigma_X \otimes \Sigma_Y$, so that, $\mu \times \nu$ is a well-defined measure on $(X \times Y, \Sigma_X \otimes \Sigma_Y)$.

Notice that, in general, its definition is not unique. Uniqueness for the product measure can be ensured requiring that the measures μ and ν are both σ-finite. ∎

2.3 The category of measurable spaces

Measurable spaces and measurable maps forms a category, denoted by **Meas**. This category is complete and cocomplete: limits and colimits are obtained as in **Set** and endowed, respectively, with initial and final σ-algebra with respect to their cone and cocone maps. Explicitly, the product $X \times Y$ has as underlying set the cartesian product of the underlying sets and σ-algebra generated by the rectangles $E \times F$ for $E \in \Sigma_X$, $F \in \Sigma_Y$ (it is exactly the product space!). The coproduct $X + Y$, has as underlying set the disjoint union of the underlying sets and σ-algebra generated by the insertion maps.

There is an obvious forgetful functor $U \colon \mathbf{Meas} \to \mathbf{Set}$ forgetting the σ-algebra structure of the measurable spaces. This functor is faithful, and has a left adjoint $D \colon \mathbf{Set} \to \mathbf{Meas}$ which assigns to each set X the discrete σ-algebra $\mathcal{P}(X)$. Therefore U preserves all limits —this is *why*

the underlying set of the product of measurable spaces is the cartesian product of their underlying sets. The forgetful functor U also has a right adjoint $I \colon \mathbf{Set} \to \mathbf{Meas}$ which assigns to each set the indiscrete σ-algebra. Therefore U preserves also all colimits.

Functors.   For a *set* A and a measurable space X, the *exponential space* X^A has as underlying set the set of all functions from A to X, and it is endowed with the initial σ-algebra with respect to the family of function $\{ev_a \colon X^A \to X \mid a \in A\}$ of the *evaluation maps at* a, defined as $ev_a(f) = f(a)$, for all $f \colon A \to X$ in X^A. This definition can be extended to a functor $Id^A \colon \mathbf{Meas} \to \mathbf{Meas}$, called *exponent functor*, acting on object as $X \mapsto X^A$ and arrows $f \mapsto f \circ -$. Notice that, since X^A is equipped with the initial σ-algebra making all evaluation maps at elements of A measurable, also f^A is measurable.

Since **Meas** is complete and cocomplete there also the products and coproduct endofunctors. Explicitly, given $F, G \colon \mathbf{Meas} \to \mathbf{Meas}$ two endofunctors, the *product functor* $F \times G$ acts on objects by $X \mapsto FX \times GX$ and on arrows by $f \mapsto Ff \times Gf$, while, the *coproduct functor* $F + G$ by $X \mapsto FX + GX$ and $f \mapsto Ff + Gf$, i.e, each square in

$$
\begin{array}{ccccc}
FX & \xleftarrow{\pi_{FX}} & (F \times G)X & \xrightarrow{\pi_{GX}} & GX \\
{\scriptstyle Ff}\downarrow & & \downarrow{\scriptstyle (F \times G)f} & & \downarrow{\scriptstyle Gf} \\
FY & \xleftarrow[\pi_{FY}]{} & (F \times G)Y & \xrightarrow[\pi_{GY}]{} & GY
\end{array}
\qquad
\begin{array}{ccccc}
FX & \xrightarrow{in_{FX}} & (F + G)X & \xleftarrow{in_{GX}} & GX \\
{\scriptstyle Ff}\downarrow & & \downarrow{\scriptstyle (F + G)f} & & \downarrow{\scriptstyle Gf} \\
FY & \xrightarrow[in_{FY}]{} & (F + G)Y & \xleftarrow[in_{GY}]{} & GY
\end{array}
$$

commutes, for all arrows $f \colon X \to Y$.

Completeness and cocompleteness of allows to consider the class of *polynomial endofunctors*, that is, the smallest class of endofunctors containing the identity Id, the constant functor for all measurable spaces X, and closed under binary product and coproduct.

Another important class of functors that will be used in the thesis are the *measure functors*. Let (X, Σ) be a measurable space and $\Delta(X, \Sigma)$ be the set of all measures $\mu \colon \Sigma \to [0, \infty]$ on (X, Σ). For each measurable set $E \in \Sigma$, there is a canonical evaluation function $ev_E \colon \Delta(X, \Sigma) \to [0, \infty]$, defined by $ev_E(\mu) = \mu(E)$, for each measure $\mu \in \Delta(X, \Sigma)$, and called *evaluation at* E. By means of these evaluation maps, $\Delta(X, \Sigma)$ can be organized into a measurable space $(\Delta(X, \Sigma), \Sigma_{\Delta(X, \Sigma)})$, where $\Sigma_{\Delta(X, \Sigma)}$ the initial σ-algebra with respect to $\{ev_E \mid E \in \Sigma\}$, i.e., the smallest σ-algebra making ev_E measurable w.r.t. the Borel σ-algebra on $[0, \infty]$, for all $E \in \Sigma$. This definition can be extended to a functor $\Delta \colon \mathbf{Meas} \to \mathbf{Meas}$ acting measurable spaces (X, Σ_X) and arrows $f \colon (X, \Sigma_X) \to (Y, \Sigma_Y)$ as follows, for $\mu \in \Delta(X, \Sigma)$

$$
\Delta(X, \Sigma) = (\Delta(X, \Sigma), \Sigma_{\Delta(X, \Sigma)})
$$
$$
\Delta(f)(\mu) = \mu \circ f^{-1}
$$

Note that, since f is measurable, $f^{-1}(E) \in \Sigma_X$, for any $E \in \Sigma_Y$, so that $(\mu \circ f^{-1})$ is a well-defined measure on (Y, Σ_Y). Functoriality of the definition can be checked easily.

We identify four subclasses of measures, and each of these can be easily extended to a functor as we did above:

probability mesures:	$\Delta_1(X, \Sigma) = \{\mu \in \Delta(X, \Sigma) \mid \mu(X) = 1\}$,
subprobability mesures:	$\Delta_{\leq 1}(X, \Sigma) = \{\mu \in \Delta(X, \Sigma) \mid \mu(X) \leq 1\}$,
finite mesures:	$\Delta_{< \infty}(X, \Sigma) = \{\mu \in \Delta(X, \Sigma) \mid \mu(X) < \infty\}$,
σ-finite mesures:	$\Delta_\sigma(X, \Sigma) = \{\mu \in \Delta(X, \Sigma) \mid \mu \text{ is } \sigma\text{-finite}\}$.

Obviously, $\Delta_1(X, \Sigma) \subseteq \Delta_{\leq 1}(X, \Sigma) \subseteq \Delta_{< \infty}(X, \Sigma) \subseteq \Delta_\sigma(X, \Sigma) \subseteq \Delta(X, \Sigma)$, and, more importantly, any property that is satisfied by a certain class of measures is preserved in all its subclasses. Thus, if in the exposition we will require some particular property which is not satisfied by all classes of measures, we will conventionally use the biggest subclass for which the property holds and assume that all the results are valid in all its subclasses (e.g., $\Delta(X, \Sigma)$ denotes that no assumptions on the measures are required, Δ_σ denotes that the measures are required to be σ-finite, and so on).

Remark 2.3.1 The functor $\Delta_1 \colon \textbf{Meas} \to \textbf{Meas}$ is the so called Giry functor (actually monad). As noticed in [68, 57], the σ-algebra we have defined for $\Delta_1(X, \Sigma)$ can also be generated by the following collection of sets

$$\{L_r(E) \mid E \in \Sigma, \, r \in [0,1] \cap \mathbb{Q}\},$$

where $L_r(E) = \{\mu \in \Delta_1(X, \Sigma) \mid \mu(E) \geq r\}$. ∎

Remark 2.3.1 can extended to all measures as the following (folklore?) lemma shows.

Lemma 2.3.2 Let $\Delta(X, \Sigma)$ be the set of measures on (X, Σ). Then the following families of sets generate the same σ-algebra

(i) $\mathcal{F}_0 = \{ev_E^{-1}(O) \mid E \in \Sigma, \, O \text{ Borel-open in } [0, \infty]\}$
(ii) $\mathcal{F}_1 = \{L_r(E) \mid r \in [0, \infty) \cap \mathbb{Q}, \, E \in \Sigma\}$;
(iii) $\mathcal{F}_2 = \{B[s]_r(E) \mid s, r \in [0, \infty) \cap \mathbb{Q}, \, E \in \Sigma\}$.
where $L_r(E) = \{\mu \in \Delta(X, \Sigma) \mid \mu(E) \geq r\}$ and $B[s]_r(E) = \{\mu \in \Delta(X, \Sigma) \mid \mu(E) \in (s - r, r + s)\}$.

Proof. The Borel σ-algebra on $[0, \infty]$ can be generated both by the intervals of the form $[r, \infty)$ or of the form $(s - r, r + s)$, for $s, r \in [0, \infty)$. Therefore, the thesis follows by Proposition 2.2.9, noticing that $L_r(E) = ev_E^{-1}([r, \infty))$ and $B[s]_r(E) = ev_E^{-1}((s - r, r + s))$. ∎

Notice that $\sigma(\mathcal{F}_0)$ is exactly the initial σ-algebra with respect to $\{ev_E \mid E \in \Sigma\}$, hence

$$\Sigma_{\Delta(X, \Sigma)} = \sigma(\mathcal{F}_0) = \sigma(\mathcal{F}_1) = \sigma(\mathcal{F}_2).$$

The following result will be very useful in the thesis.

Lemma 2.3.3 Let $\mathcal{A}$ be a boolean algebra and (X, Σ_X) a measurable space with σ-algebra Σ_X generated by $\mathcal{A}$. Then $\sigma(\mathcal{F}) = \sigma(\mathcal{G})$ where, $L_r(E) = \{\mu \in \Delta_{<\infty}(X, \Sigma_X) \mid \mu(E) \geq r\}$, and

$$\mathcal{F} = \{L_r(E) \mid E \in \Sigma_X \text{ and } r \in \mathbb{Q} \cap [0, \infty)\}, \qquad \mathcal{G} = \{L_r(A) \mid A \in \mathcal{A} \text{ and } r \in \mathbb{Q} \cap [0, \infty)\}.$$

Proof. We will prove the two inclusions separately.

$\sigma(\mathcal{G}) \subseteq \sigma(\mathcal{F})$: Since $\sigma(\mathcal{G})$ is the smallest σ-algebra that contains $\mathcal{G}$, to prove the inclusion it suffices to show that $\mathcal{G} \subseteq \sigma(\mathcal{F})$. By definition of generated σ-algebra $\mathcal{F} \subseteq \sigma(\mathcal{F})$, and $\mathcal{A} \subseteq \sigma(\mathcal{A}) = \Sigma_X$. From this it is clear that $\mathcal{G} \subseteq \mathcal{F}$, and therefore that $\mathcal{G} \subseteq \sigma(\mathcal{F})$.

$\sigma(\mathcal{F}) \subseteq \sigma(\mathcal{G})$: This inclusion is less trivial. Let $D = \{E \in \Sigma_X \mid L_r(E) \in \sigma(\mathcal{G})\}$. Notice that $\mathcal{A} \subseteq D$, indeed, for every $r \in \mathbb{Q} \cap [0, \infty)$ and $A \in \mathcal{A}$, $L_r(A) \in \mathcal{G} \subseteq \sigma(\mathcal{G})$. So that, if we were able to show that D is a σ-algebra, $\Sigma_X = \sigma(\mathcal{A}) = D$ and by definition of D this will imply that $\sigma(\mathcal{F}) \subseteq \sigma(\mathcal{G})$. So, let us prove that D is a σ-algebra. Since $\mathcal{A} \subseteq D$ and $\mathcal{A}$ is a boolean algebra, by the monotone class theorem (Theorem 2.2.16) it is enough to show that D is a monotone class. Assume $E_0 \supseteq E_1 \supseteq E_2 \supseteq \ldots$ be a decreasing countable collection of elements in D, i.e., $L_r(E) \in \sigma(\mathcal{G})$. We show $L_r(\bigcap_{n \in \mathbb{N}} E_n) = \bigcap_{n \in N} L_r(E)$, thus that $\bigcap_{n \in \mathbb{N}} E_n \in D$.

($\subseteq$) Since, for all $k \in \mathbb{N}$, $\bigcap_{n \in \mathbb{N}} E_n \subset E_k$, we have that $L_r(\bigcap_{n \in \mathbb{N}} E_n) \subseteq L_r(E_k)$, therefore $L_r(\bigcap_{n \in \mathbb{N}} E_n) \subseteq \bigcap_{n \in N} L_r(E)$.

($\supseteq$) Let $\mu \in \bigcap_{n \in N} L_r(E)$, hence, for all $n \in \mathbb{N}$, $\mu(E_n) \geq r$. This means that r is a lower bound for $\{\mu(E_n) \mid n \in \mathbb{N}\}$. By Lemma 2.2.20 and since all measures are assumed to be finite, $\mu(\bigcap_{n \in \mathbb{N}} E_n) = \inf_{n \in \mathbb{N}} \mu(E_n)$, so that $\mu(\bigcap_{n \in \mathbb{N}} E_n) \geq r$, thus $\mu \in L_r(\bigcap_{n \in \mathbb{N}} E_n)$.

Assume $E_0 \subseteq E_1 \subseteq E_2 \subseteq \ldots$ be an increasing countable collection of elements in D, that is, $L_r(E) \in \sigma(\mathcal{G})$. We show $L_r(\bigcup_{n \in \mathbb{N}} E_n) = \bigcap_{k > 0} \bigcup_{n \in N} L_{r - \frac{1}{k}}(E)$, which implies $\bigcap_{n \in \mathbb{N}} E_n \in D$. This follows from the following sequence of equivalent statements:

$$\mu \in L_r(\textstyle\bigcup_{n \in \mathbb{N}} E_n) \iff \mu(\textstyle\bigcup_{n \in \mathbb{N}} E_n) \geq r \qquad \text{(by def. } L_r(E))$$
$$\iff \lim_{n \in \mathbb{N}} \mu(E_n) \qquad \text{(by Lemma 2.2.20)}$$
$$\iff \forall k > 0 \,.\, \exists n \in \mathbb{N} \,.\, \mu(E_n) \geq r - \tfrac{1}{k} \qquad \text{(by convergence)}$$
$$\iff \forall k > 0 \,.\, \exists n \in \mathbb{N} \,.\, \mu \in L_{r - \frac{1}{k}}(E_n) \qquad \text{(by def. } L_r(E))$$

$$\iff \forall k > 0 \,.\, \mu \in \bigcup_{n \in \mathbb{N}} L_{r - \frac{1}{k}}(E_n) \qquad\qquad \text{(by union)}$$

$$\iff \mu \in \bigcap_{k > 0} \bigcup_{n \in \mathbb{N}} L_{r - \frac{1}{k}}(E_n) \qquad\qquad \text{(by intersection)}$$

∎

Remark 2.3.4 The proof of Lemma 2.3.3 requires the measures to be finite in order to apply Lemma 2.2.20(ii) that otherwise fails to hold.   ∎

Factorization systems.   The category **Meas** "lifts" in several ways the (**Epic**, **Monic**) factorization system in **Set**. Here we consider two of these factorizations systems, which are defined by factorizations through subspace embeddings and measurable quotient maps.

Consider a set function $f\colon X \to Y$ and the two factorizations depicted below

where $f(X)$ is the image of X under f, $i\colon f(X) \hookrightarrow Y$ its canonical inclusion set-map, $X/\sim$ the quotient w.r.t. $\sim\, = \{(x, x') \in X \times X \mid f(x) = f(y)\}$, and $q_\sim\colon X \to X/\sim$ the canonical quotient set-map, mapping the elements in $x \in X$ to their equivalence classes $[x]_\sim$.

In **Set** this two factorizations are the vey same thing, indeed $f(X)$ and $X/\sim$ are isomorphic. Things change when we consider $f\colon X \to Y$ as a measurable map between the spaces (X, Σ_X) and (Y, Σ_Y). Ideally, for any measurable map we want to produces measurable factorizations, i.e., we need to equip the sets $f(X)$ and $X/\sim$ with a σ-algebra that makes the maps i, $q_\sim$, f', and f'' measurable. This can be done adopting the initial and final σ-algebra constructions. Indeed, by Lemma 2.2.13, if we endow $f(X)$ with the initial σ-algebra with respect to i, and $X/\sim$ with the final σ-algebra with respect to $q_\sim$, we have that also f' and f'' are measurable, since f is so. With these σ-algebras, the map $i\colon f(X) \to Y$ is a *subspace embedding*, and the function $q_\sim\colon X \to X/\sim$ is a *measurable quotient*. Note that, now, $f(X)$ and $X/\sim$, considered as measurable spaces, are no more isomorphic which each other, therefore the two factorizations are different.

These constructions actually "lift" the (**Epic**, **Monic**) factorization system in **Set**, in the sense that the maps given as the solution of any lifting problem in **Set** is inherited in **Meas** since its measurability is ensured by initiality and finality of the σ-algebras of the embedding and quotient, respectively, again by Lemma 2.2.13. We will denote these two factorization systems by (**Epic**, **Emb**) and (**Quot**, **Monic**), where **Emb** denotes the class of measurable embeddings, and **Quot** the class of measurable quotients.

3

Algebras and Coalgebras

In this chapter we review the main concepts of the theory of universal algebras and coalgebras, with the aim to show the connection with denotational and operational semantics for recursively defined process description languages. In contrast with standard introductory expositions, where the theory is presented in the setting of **Set**, here we develop the theory in purely categorical terms, that is, in a way that is independent on the category at hand. This is done in view of how we will employ this theory in the rest of the thesis. For a canonical introductory presentation we recommend [56, 74, 72].

3.1 Algebras and congruences

This section recalls the basic definitions and introduces the notation. Along the exposition we provide formal examples relating algebras and denotational semantics for a term language. Then we proceed reviewing all the classical results and categorical constructions which will be used in the rest of the thesis.

Definition 3.1.1 (F-algebra) *Let $F\colon \mathbf{C} \to \mathbf{C}$ be an endofunctor. An F-algebra is a pair (X, α), where X is an object in $\mathbf{C}$, said* carrier*, and $\alpha\colon FX \to X$ is an arrow in $\mathbf{C}$, said* algebra structure.

In computer science, algebras are typically used to give an abstract categorical formalization to denotational semantics. In the following example we show how interpretations of signatures are elegantly modeled as algebras.

Example 3.1.2 (Operator interpretations as algebras) A *signature* is a pair (Σ, ar), where Σ is a set of operator symbols and $ar\colon \Sigma \to \mathbb{N}$ is an arity function. An *interpretation* of this signature on a set X of denotations is a collection of operators $(\llbracket \sigma \rrbracket\colon X^{ar(\sigma)} \to X)_{\sigma \in \Sigma}$.

Any signature (Σ, ar) gives rise to a **Set**-functor $S = \coprod_{\sigma \in \Sigma} Id^{ar(\sigma)}$ acting on objects X and arrows $f\colon X \to Y$, respectively, as

$$SX = \{\langle \sigma, (x_1, \ldots, x_{ar(\sigma)})\rangle \mid \sigma \in \Sigma, \text{ and } x_1, \ldots, x_{ar(\sigma)} \in X\},$$
$$Sf = \big[\langle \sigma, (x_1, \ldots, x_{ar(\sigma)})\rangle \mapsto \langle \sigma, (f(x_1), \ldots, f(x_{ar(\sigma)}))\rangle\big].$$

$$(3.1.1)$$

where $\langle \sigma, (x_1, \ldots, x_{ar(\sigma)})\rangle$ denotes $in_\sigma^X(x_1, \ldots, x_{ar(\sigma)}) \in SX$.

Any S-algebra (X, α) can be turned into an interpretation $(\alpha \circ in_\sigma^X\colon X^{ar(\sigma)} \to X)_{\sigma \in \Sigma}$. Conversely, by the universal property of coproducts, any interpretation $(\llbracket \sigma \rrbracket\colon X^{ar(\sigma)} \to X)_{\sigma \in \Sigma}$ defines an S-algebra structure $\coprod_{\sigma \in \Sigma} \llbracket \sigma \rrbracket\colon SX \to X$. In fact, we have just defined a correspondence between S-algebras and interpretations for the signature (Σ, ar), given by

$$(X, \alpha) \mapsto (\alpha \circ in_\sigma^X)_{\sigma \in \Sigma} \qquad\qquad (\llbracket \sigma \rrbracket\colon X^{ar(\sigma)} \to X)_{\sigma \in \Sigma} \mapsto (X, \coprod_{\sigma \in \Sigma}\llbracket \sigma \rrbracket)$$

which is clearly bijective as shown by the following commutative diagrams

$$
\begin{array}{ccc}
X^{ar(\sigma)} & \xrightarrow{\ in_\sigma^X\ } & SX \\
& & \Big\downarrow{\scriptstyle \alpha = \prod_{\sigma\in\sigma}(\alpha\,\circ\,in_\sigma^X)} \\
{\scriptstyle \alpha\,\circ\,in_\sigma^X} & \searrow & \\
& & X
\end{array}
\qquad\qquad
\begin{array}{ccc}
X^{ar(\sigma)} & \xrightarrow{\ in_\sigma^X\ } & SX \\
{\scriptstyle [\![\sigma]\!]} & \searrow & \Big\downarrow{\scriptstyle \coprod_{\sigma\in\Sigma}[\![\sigma]\!]} \\
& & X
\end{array}
$$

Note that, this construction can be carried over into any category with products and coproducts, so that signatures can be interpreted also in categories different from **Set**. ∎

Definition 3.1.3 (F-homomorphism) *Let $F\colon \mathbf{C} \to \mathbf{C}$ be an endofunctor and (X,α) and (Y,β) be F-algebras. An arrow $f\colon X \to Y$ in $\mathbf{C}$ is a F-homomorphism between (X,α) and (Y,β) if the following diagram in $\mathbf{C}$ commutes:*

$$
\begin{array}{ccc}
FX & \xrightarrow{\ Ff\ } & FY \\
{\scriptstyle \alpha}\Big\downarrow & & \Big\downarrow{\scriptstyle \beta} \\
X & \xrightarrow[\ f\ ]{} & Y
\end{array}
$$

We continue Example 3.1.2 showing that homomorphisms between interpretations correspond to algebra homomorphisms for the **Set**-functor associated with the signature.

Example 3.1.4 (Homomorphism between interpretations) Let (Σ, ar) be a signature and $([\![\sigma]\!]_X\colon X^{ar(\sigma)} \to X)_{\sigma\in\Sigma}$ and $([\![\sigma]\!]_Y\colon Y^{ar(\sigma)} \to Y)_{\sigma\in\Sigma}$ be two interpretations for the operators in Σ. A function $h\colon X \to Y$ is an *homomorphism* between interpretations if, for all $\sigma\in\Sigma$,

$$h\big([\![\sigma]\!]_X(x_1,\ldots,x_{ar(\sigma)})\big) = [\![\sigma]\!]_Y(h(x_1),\ldots,h(x_{ar(\sigma)})) . \tag{3.1.2}$$

Consider the endofunctor $S\colon \mathbf{Set} \to \mathbf{Set}$ and the bijection between interpretations for (Σ, ar) and S-algebras defined in Example 3.1.2. We show that homomorphisms between interpretations correspond to S-homorphisms. To see this, note that Equation (3.1.2) corresponds to say that the following diagram commutes

$$
\begin{array}{ccc}
X^{ar(\sigma)} & \xrightarrow{\ h^{ar(\sigma)}\ } & Y^{ar(\sigma)} \\
{\scriptstyle [\![\sigma]\!]_X}\Big\downarrow & & \Big\downarrow{\scriptstyle [\![\sigma]\!]_Y} \\
X & \xrightarrow[\ h\ ]{} & Y
\end{array}
$$

Therefore, by the universal property of coproducts, it can be shown that h is an S-algebra homomorphism between $(X, \coprod_\sigma [\![\sigma]\!]_X)$ and $(Y, \coprod_\sigma [\![\sigma]\!]_Y)$. Conversely, any S-algebra homomorphism $h\colon (X,\alpha_X) \to (Y,\alpha_Y)$ is an homomorphism between the signatures $(\alpha_X \circ in_\sigma^X)_{\sigma\in\Sigma}$ and $(\alpha_Y \circ in_\sigma^Y)_{\sigma\in\Sigma}$, since

$$h\big((\alpha_X \circ in_\sigma^X)(x_1,\ldots,x_{ar(\sigma)})\big) = (\alpha_Y \circ in_\sigma^Y)(h(x_1),\ldots,h(x_{ar(\sigma)})) ,$$

which corresponds to the condition of being an S-homomorphism. ∎

Definition 3.1.5 (F-congruence) *Let $F\colon \mathbf{C} \to \mathbf{C}$ be an endofunctor. A monic span (R,f,g) in $\mathbf{C}$ between X and Y is a F-congruence between F-algebras (X,α) and (Y,β) if there exists a (unique) algebra structure $\gamma\colon FR \to R$ on R making the following diagram in $\mathbf{C}$ commute*

$$
\begin{array}{ccccc}
FX & \xleftarrow{\ Ff\ } & FR & \xrightarrow{\ Fg\ } & FY \\
{\scriptstyle \alpha}\Big\downarrow & & {\scriptstyle \gamma}\Big\downarrow & & \Big\downarrow{\scriptstyle \beta} \\
X & \xleftarrow[\ f\ ]{} & R & \xrightarrow[\ g\ ]{} & Y
\end{array}
$$

that is, making f and g morphisms of F-algebras.

The following example shows how F-congruences can be indeed considered as the right categorical generalization of congruential relations with respect to a signature.

Example 3.1.6 (Congruence with respect to a signature) Let (Σ, ar) be a signature and $([\![\sigma]\!]_X \colon X^{ar(\sigma)} \to X)_{\sigma \in \Sigma}$, $([\![\sigma]\!]_Y \colon Y^{ar(\sigma)} \to Y)_{\sigma \in \Sigma}$ be two interpretations. A relation $R \subseteq X \times Y$ is a *congruence with respect to* (Σ, ar) if, for all $\sigma \in \Sigma$ and $(x_1, y_1), \ldots, (x_{ar(\sigma)}, y_{ar(\sigma)}) \in R$ it holds

$$\big([\![\sigma]\!]_X(x_1, \ldots, x_{ar(\sigma)}), [\![\sigma]\!]_Y(y_1, \ldots, y_{ar(\sigma)})\big) \in R \,,$$

that is, the relation R respects the operator interpretations. Congruences for a signature (Σ, ar) correspond to S-congruences for the functor $S \colon \mathbf{Set} \to \mathbf{Set}$ defined in Example 3.1.2.

Any congruence $R \subseteq X \times Y$ as above defines an S-algebra structure $\gamma \colon SR \to R$ as follows

$$\gamma\big(\langle \sigma, ((x_1, y_1), \ldots, (x_{ar(\sigma)}, y_{ar(\sigma)}))\rangle\big) = \big([\![\sigma]\!]_X(x_1, \ldots, x_{ar(\sigma)}), [\![\sigma]\!]_Y(y_1, \ldots, y_{ar(\sigma)})\big) \in R \,.$$

By definition of γ and S we have that

$$\pi_X \circ \gamma\big(\langle \sigma, ((x_1, y_1), \ldots, (x_{ar(\sigma)}, y_{ar(\sigma)}))\rangle\big) = [\![\sigma]\!]_X(x_1, \ldots, x_{ar(\sigma)})$$
$$= \coprod_{\sigma \in \Sigma} [\![\sigma]\!]_X\big(\langle \sigma, (x_1, \ldots, x_{ar(\sigma)})\rangle\big)$$
$$\coprod_{\sigma \in \Sigma} [\![\sigma]\!]_X \circ S\pi_X\big(\langle \sigma, ((x_1, y_1), \ldots, (x_{ar(\sigma)}, y_{ar(\sigma)}))\rangle\big) =$$

that is, $\pi_X \colon R \to X$ is an S-algebra homomorphism between (R, γ) and $(X, \coprod_{\sigma \in \Sigma} [\![\sigma]\!]_X)$ and, similarly, also $\pi_Y \colon R \to Y$ is an S-homomorphism between (R, γ) and $(Y, \coprod_{\sigma \in \Sigma} [\![\sigma]\!]_Y)$. Therefore the (monic) span (R, π_X, π_Y) is an S-congruence.

Conversely, given an S-congruence (R, π_X, π_Y) between S-algebras (X, α_X) and (Y, α_Y), we show that R is a congruence between the interpretations $(\alpha \circ in_\sigma^X)_{\sigma \in \Sigma}$ and $(\alpha \circ in_\sigma^Y)_{\sigma \in \Sigma}$. Assume $(x_1, y_1), \ldots, (x_{ar(\sigma)}, y_{ar(\sigma)}) \in R$, then, using the fact that γ is an S-homomorphism, we have

$$\pi_X \circ \gamma\big(\langle \sigma, ((x_1, y_1), \ldots, (x_{ar(\sigma)}, y_{ar(\sigma)}))\rangle\big) = \alpha_X \circ S\pi_X\big(\langle \sigma, ((x_1, y_1), \ldots, (x_{ar(\sigma)}, y_{ar(\sigma)}))\rangle\big)$$
$$= \alpha_X\big(\langle \sigma, (x_1, \ldots, x_{ar(\sigma)})\rangle\big)$$
$$= \alpha_X \circ in_\sigma^X(x_1, \ldots, x_{ar(\sigma)}) \,,$$

Similarly, $\pi_Y \circ \gamma\big(\langle \sigma, ((x_1, y_1), \ldots, (x_{ar(\sigma)}, y_{ar(\sigma)}))\rangle\big) = \alpha_Y \circ in_\sigma^Y(y_1, \ldots, y_{ar(\sigma)})$. Therefore the pair $(\alpha_X \circ in_\sigma^X(x_1, \ldots, x_{ar(\sigma)}), \alpha_Y \circ in_\sigma^Y(y_1, \ldots, y_{ar(\sigma)}))$ belongs to R. ∎

Examples 3.1.2, 3.1.4, and 3.1.6 should have convinced the reader of the usefulness of F-algebras as good theoretical tools to reason about (denotational) semantics. Moreover, the high level of abstraction provided by the categorical language allows for simpler further generalizations of the results in different domain settings.

The category of F-algebras

For any endofunctor $F \colon \mathbf{C} \to \mathbf{C}$, it is easy to check that F-algebras and F-homomorphisms form a category, denoted by F-**alg**. Composition of arrows is inherited from the underling category $\mathbf{C}$, so that, associativity is always guaranteed to hold.

Notably, the category of F-algebra lifts all limits from the underlying category $\mathbf{C}$, so that if $\mathbf{C}$ is complete, so is F-**alg**. This allows to define derived structures such as products, equalizes, and pullbacks. For example, the binary product of the F-algebras (X, α_X) and (Y, α_Y) is given by

$$
\begin{array}{ccccc}
FX & \xleftarrow{F\pi_X} & F(X \times Y) & \xrightarrow{F\pi_Y} & FY \\
\alpha_X \downarrow & & \alpha_{X \times Y} \downarrow & & \downarrow \alpha_Y \\
X & \xleftarrow{\pi_X} & X \times Y & \xrightarrow{\pi_Y} & Y
\end{array}
$$

where the structure map $\alpha_{X \times Y}$ is the uniquely given by the universal property of the product $X \times Y$ in $\mathbf{C}$, and we denote it by $(X, \alpha_X) \times (Y, \alpha_Y) = (X \times Y, \alpha_{X \times Y})$. Notice that, the carrier of the product of F-algebras corresponds exactly to product of the carriers in $\mathbf{C}$. This holds in general, that is, the carrier of a limit in F-**alg** is the limit in $\mathbf{C}$.

The case is different for colimits. Coproducts of F-algebras need not exist, but when they do, their carrier will often have to be different from the coproduct of the carriers in $\mathbf{C}$. One may be lucky, in that the functor F preserves a certain type of colimit. In that case, this type of colimit exists for F-algebras and it is constructed as in $\mathbf{C}$. In general though, the functors usually fail to preserve arbitrary colimits, and often this isx the case for functors needed to model most applications of interest.

3.2   Initial Algebra and Induction

The lack of general colimits in F-**alg** does not represent a problem when algebras are used to give semantics to programming languages. However, one colimit is generally required to exist: the *initial object*. This universal object is so important in the theory of universal algebras that it deserves its own name.

Definition 3.2.1 (Initial F-algebra) *Let $F\colon \mathbf{C} \to \mathbf{C}$ be an endofunctor. An* initial F-algebra *is an initial object in F-**alg**, i.e., an F-algebra (A, ι) such that for any F-algebra (X, α) there exists a unique F-homomorphism from (A, ι) to (X, α).*

The following result is classical and depends only on the universal property of initial objects.

Theorem 3.2.2 (Lambek's lemma) *Initial F-algebras (A, ι) are fixed points for $F\colon \mathbf{C} \to \mathbf{C}$, that is, $\iota\colon FA \to A$ is an isomorphism in $\mathbf{C}$.*

If one sees categories as generalized preorders and endofunctors as monotone functions, algebras for a functor correspond to prefixed points and initial algebras as the least fixed point.

Of course initial algebras do not need to exist in general, but when they exist they give a useful *induction proof principle*. Indeed, every algebra structure $\alpha\colon FX \to X$ of an arbitrary endofunctor F with initial algebra (A, ι), can be *inductively extended* to an arrow $\alpha^{\#}\colon A \to X$ by taking the unique algebra arrow from the initial algebra to the algebra (X, α). Notably, this principle generalizes the standard set-theoretical mathematical induction based on the notion of well-founded relation, which is briefly recalled below in the particular case of natural numbers.

Recursion on natural numbers. Probably, one of the most known induction proof principle is that on natural numbers. Induction can be performed on $\mathbb{N}$ using the well-foundedness of the canonical order relation on it. Formally, we have the following recursion theorem:

Theorem 3.2.3 (Recursion Theorem) *Given a set X, an element $x \in X$ and a function $g\colon X \to X$, there exists a* unique *function $f\colon \mathbb{N} \to X$ such that for all numbers $n \in \mathbb{N}$*

$$f(0) = x \qquad and \qquad f(n+1) = g(f(n)).$$

The value x of the function f at 0 (i.e., the least element w.r.t. the order relation) is the *base* of the induction and g defines the inductive step. The fact that standard mathematical constructions are inductive reflects the common assumption that the axioms of set theory include the *axiom of foundation* which postulates that the set membership relation is well-founded. The axiom of foundation allows an inductive construction of sets starting from the empty set (the base) and recursively applying the powerset operator. Induction on natural numbers is just a particular case of that on ordinal numbers, usually identified as *transfinite induction*.

Notably, the recursion theorem can be taken as the definition of natural numbers. That is, every set N with a distinguished element $z \in N$ and a unary "successor" operation $s\colon N \to N$ such that the recursion theorem holds, is isomorphic to the natural numbers. The existence and uniqueness of the function f, asserts the universal property characterizing natural numbers: *initiality*.

We shall see how initial arrows subsume an induction proof principle. The next example relates recursion on natural numbers with the universal property of initial algebras.

Example 3.2.4 Consider the **Set**-endofunctor $\mathbf{1} + Id$, then, by the universal property of the coproduct, an algebra (X, α) for $\mathbf{1} + Id$ is given by a set X together with a distinguished element $x \in X$ and a function $g \colon X \to X$ (i.e., the structure map is the coproduct pairing $\alpha = [x, g]$, where $x \in X$ is seen as the constant function $x \colon \mathbf{1} \to X$ from the singleton set to X). Assume that there exists an initial algebra $(N, [e, s])$ for $\mathbf{1} + Id$, then for any given algebra $(X, [x, g])$ there exists a unique function $f \colon N \to X$ such that the following diagram commutes

$$
\begin{array}{ccc}
\mathbf{1} + N & \xrightarrow{\;\mathbf{1} + f\;} & \mathbf{1} + X \\
{\scriptstyle [e, s]} \big\downarrow & & \big\downarrow {\scriptstyle [x, g]} \\
N & \dashrightarrow{\;f\;} & X
\end{array}
$$

Spelling out of the above commutative diagram, we have that f must obey to the the following equations:

$$ f(e) = x \qquad \text{and} \qquad f(s(n)) = g(f(n)) \,. $$

This correspond to the recursion theorem on natural natural numbers (see Theorem 3.2.3), therefore existence and uniqueness of the function f implies that the initial algebra $(N, [e, s])$ is isomorphic to $(\mathbb{N}, [0, (1 + \cdot)])$, thus the latter is initial too.

This may convince the reader how the recursion theorem on natural numbers is nothing but initiality on the category of algebras for the functor $\mathbf{1} + Id$. ∎

An other classical induction proof principle is that of "structural induction on terms". The next example shows that this principle is again an instance of initiality in the category of algebras for some functor. Next we give all the details.

Example 3.2.5 (Structural induction on terms) Let (Σ, ar) be a signature and X a set of variables. The set of terms TX (freely) generated over the variables in X and the signature (Σ, ar) is the smallest set satisfying the following axioms and rules, for all $x \in X$ and $\sigma \in \Sigma$

$$ x \in TX \qquad\qquad \text{and} \qquad\qquad \frac{t_1, \ldots, t_{ar(\sigma)} \in TX}{\sigma(t_1 \ldots t_{ar(\sigma)}) \in TX} $$

The *structural recursion theorem* states that, given any interpretation $h \colon X \to Y$ on the set variables and any interpretation $(\llbracket \sigma \rrbracket \colon Y^{ar(\sigma)} \to Y)$ of the signature, there exists a unique function $f \colon TX \to Y$ such that, for all $x \in X$, $\sigma \in \Sigma$, and $t_1, \ldots, t_{ar(\sigma)} \in TX$

$$ f(x) = h(x) \qquad \text{and} \qquad f(\sigma(t_1, \ldots, t_{ar(\sigma)})) = \llbracket \sigma \rrbracket (f(t_1), \ldots, f(t_{ar(\sigma)})) \,. \qquad (3.2.1) $$

We already seen in Example 3.1.2 that interpretations for (Σ, ar) are algebras for the **Set**-functor $S = \coprod_{\sigma \in \Sigma} Id^{ar(\sigma)}$. The above "extended interpretations" with variables in X are just algebras for the functor $X + S$. For instance, an interpretation as above corresponds to the $(X + S)$-algebra $(Y, [h, \coprod_{\sigma \in \Sigma} \llbracket \sigma \rrbracket])$. The set of terms TX can be naturally endowed with an $(X + S)$-algebra structure $[\eta_X, \psi_X] \colon X + STX \to TX$, where $\eta_X \colon X \to TX$ and $\psi_X \colon STX \to TX$ are defined as follows

$$ \eta_X(x) = x \qquad \text{and} \qquad \psi_X(\langle \sigma, (t_1, \ldots, t_{ar(\sigma)}) \rangle) = \sigma(t_1 \ldots t_{ar(\sigma)}) \,, $$

for all $x \in X$, $\sigma \in \Sigma$, and $t_1, \ldots, t_{ar(\sigma)} \in TX$.

The statement of the structural recursion theorem corresponds to say that for any $(X + S)$-algebra of the form $(Y, [h, \coprod_{\sigma \in \Sigma} \llbracket \sigma \rrbracket])$ there exists a unique function $f \colon TX \to Y$ such that the

following diagram commutes (note that, it corresponds to Equation (3.2.1)):

$$X + STX \xrightarrow{\;X + f\;} X + SY$$

$$[\eta_X, \psi_X] \downarrow \qquad\qquad \downarrow [h, \coprod_{\sigma \in \Sigma}[\![\sigma]\!]]$$

$$TX \dashrightarrow_{f} Y$$

that is, f is nothing but an $(X + S)$-homomorphism between $(TX, [\eta_X, \psi_X])$ and $(Y, [h, \coprod_{\sigma \in \Sigma}[\![\sigma]\!]])$. Since any algebra $(X + S)$-algebra can be turned into an isomorphic algebra of the prescribed form (see Example 3.1.2), the existence and uniqueness of the homomorphism f makes $(TX, [\eta_X, \psi_X])$ an initial algebra. ∎

The abstraction into categorical terms allows to find inductive proof principles in many different settings (e.g. changing the category or the functor of reference). Indeed, it is only required the existence of initial algebras, then the proof principle comes for free.

3.2.1   From Initial Algebras to Adjunctions and back

Any universal construction gives rise to an adjunction and, conversely, any adjunction provides useful free universal constructions.

From the category F-**alg** of algebras for some functor $F: \mathbf{C} \to \mathbf{C}$ to the underlying category $\mathbf{C}$, there is a natural *forgetful functor* $U^F: F\text{-}\mathbf{alg} \to \mathbf{C}$ mapping an F-algebra (X, α) to its carrier X and an F-homomorphism $f: (X, \alpha) \to (Y, \beta)$ to the arrow $f: X \to Y$ in $\mathbf{C}$. In case the forgetful functor $U^F: F\text{-}\mathbf{alg} \to \mathbf{C}$ has a left adjoint, namely, the functor $L^F: \mathbf{C} \to F\text{-}\mathbf{alg}$, the category of F-algebras admits a *free construction*: given any object X in $\mathbf{C}$, any F-algebra (Y, β) and morphism $f: X \to Y = U^F(Y, \beta)$ in $\mathbf{C}$, there exists a *unique* F-homomorphism $f^\#$ such that the following diagrams commute

$$\mathbf{C} \underset{U^F}{\overset{L^F}{\rightleftarrows}} F\text{-}\mathbf{alg}$$

$$X \xrightarrow{\eta_X} U^F L^F X \qquad\qquad L^F X$$
$$\quad {}_{f}\searrow \quad \downarrow {}^{U^F f^\#} \qquad\qquad \downarrow {}^{f^\#}$$
$$\qquad U^F(Y, \beta) \qquad\qquad (Y, \beta)$$

where $\eta: Id \Rightarrow U^F L^F$ is the unit of the adjunction.

Next, we show that the existence of the left adjoint L^F and of initial algebras for some (class of) functors are strictly related with each other.

From initial algebras to adjunctions

Let $F: \mathbf{C} \to \mathbf{C}$ be a functor in a category $\mathbf{C}$ with binary coproducts. Assume, moreover, that for any object X in $\mathbf{C}$ the functor $X + F$ has initial algebra (A_X, ι_X). Under these assumptions we can define a functor $L^F: \mathbf{C} \to F\text{-}\mathbf{alg}$, which will be proved to be the left adjoint of the forgetful functor $U^F: F\text{-}\mathbf{alg} \to \mathbf{C}$.

Before proceeding with the definition of L^F, we give the following correspondence lemma.

Lemma 3.2.6 *Let X be an object in $\mathbf{C}$ and $F: \mathbf{C} \to \mathbf{C}$ be an endofunctor. An arrow $f: A \to B$ is an homomorphism between the $(X + F)$-algebras (A, α_A) and (B, α_B) if and only if the following*

diagram in **C** *commutes:*

$$
\begin{array}{ccccc}
X & \xrightarrow{\alpha_A \circ in_X^A} & A & \xleftarrow{\alpha_A \circ in_{FA}^A} & FA \\
& & \downarrow{f} & & \downarrow{Ff} \\
\alpha_B \circ in_X^B & & B & \xleftarrow{\alpha_B \circ in_{FB}^B} & FB
\end{array}
$$

where in_X^A and in_{FA}^A are the left and right injections of the coproduct $X + FA$, respectively, and in_X^B, in_{FB}^B the left and right injections of $X + FB$.

Proof. One direction of the correspondence follows by definition of the coproduct functor $(X + F)$ which makes the left squares of the two following diagrams commute:

$$
\begin{array}{ccccc}
X & \xrightarrow{in_X^A} & X + FA & \xrightarrow{\alpha_A} & A \\
\downarrow{id_X} & & \downarrow{X + Ff} & & \downarrow{f} \\
X & \xrightarrow{in_X^B} & X + FB & \xrightarrow{\alpha_B} & B
\end{array}
\qquad
\begin{array}{ccccc}
FA & \xrightarrow{in_{FA}^A} & X + FA & \xrightarrow{\alpha_A} & A \\
\downarrow{Ff} & & \downarrow{X + Ff} & & \downarrow{f} \\
FB & \xrightarrow{in_{FB}^B} & X + FB & \xrightarrow{\alpha_B} & B
\end{array}
$$

The other direction holds noticing that $\alpha_A = [\alpha_A \circ in_X^A, \alpha_A \circ in_{FA}^A]$ and $\alpha_B = [\alpha_B \circ in_X^B, \alpha_B \circ in_{FB}^B]$, so that the diagram given in the statement of the lemma makes f an $(X + F)$-homomorphism between (A, α_A) and (B, α_B). ∎

Note that, any $(X + F)$-algebra (A, α_A) can be turned into an F-algebra $(A, \alpha_A \circ in_{FA}^A)$, so that by the lemma above any $(X + F)$-homomorphism becomes an F-homomorphism along this translation.

Let us define the functor $L^F \colon \mathbf{C} \to F\text{-}\mathbf{alg}$. For any object X in $\mathbf{C}$, the initial $(X + F)$-algebra (A_X, ι_X) associated with it can be turned into an F-algebra (A_X, ψ_X), with algebra structure $\psi_X \colon FA_X \to A_X$ given by the composite $\iota_X \circ in_{FA_X}^X$. We take this as the definition of L^F on objects X in $\mathbf{C}$:

$$
L^F X = (A_X, \psi_X = \iota_X \circ in_{FA_X}^X)
$$

Let X and Y two objects in $\mathbf{C}$ and $f \colon X \to Y$ be an arrow between them. From the initial $(Y + F)$-algebra (A_Y, ι_Y) we define an $(X + F)$-algebra on A_Y with algebra structure given by the composite $\iota_Y \circ (f + Fid_{A_Y}) \colon X + FA_Y \to A_Y$. By initiality of (A_X, ι_X), there exists a unique $(X + F)$-homomorphism $f^\# \colon A_X \to A_Y$ making the following diagrams commute (by Lemma 3.2.6)

$$
\begin{array}{ccccc}
X & \xrightarrow{\iota_X \circ in_X^X} & A_X & \xleftarrow{\psi_X} & FA_X \\
\downarrow{f} & & \vdots\,{f^\#} & & \downarrow{Ff} \\
Y & \xrightarrow{\iota_Y \circ in_Y^Y} & A_Y & \xleftarrow{\psi_Y} & FA_Y
\end{array}
\qquad (3.2.2)
$$

In particular, $f^\#$ is an F-homomorphism between $L^F X = (A_X, \psi_X)$ and $L^F Y = (A_Y, \psi_Y)$. Thus, the following definition for L^F on morphisms is well given:

$$
L^F(f \colon X \to Y) = f^\# \colon (A_X, \psi_X) \to (A_Y, \psi_Y).
$$

To prove that this definition is functorial, i.e., $L^F id_X = id_L^F X$ and $L^F g \circ L^F f = L^F(g \circ f)$, one just has to exploit the universal property of the initial algebra.

To prove that L^F is indeed the left adjoint of U^F, we provide the unity $\eta \colon Id \Rightarrow U^F L^F$ and counit $\epsilon \colon L^F U^F \Rightarrow Id$ of the adjunction. These are defined component-wise as follows, for X in $\mathbf{C}$ and (X, α) in $F\text{-}\mathbf{alg}$,

$$
\eta_X = \iota_X \circ in_X^X \colon X \to U^F L^F X = A_X, \quad \epsilon_{(X, \alpha)} = [id_X, \alpha]^\# \colon L^F U^F X = (A_X, \psi_X) \to (X, \alpha),
$$

where $[id_X, \alpha]^{\#}$ is the unique arrow given by initiality between the $(X + F)$-algebras (A_X, ι_X) and $(X, [id_X, \alpha])$, making the following diagrams commute (by Lemma 3.2.6):

$$
\begin{array}{ccccc}
X & \xrightarrow{\iota_X \circ in_X^X} & A_X & \xleftarrow{\psi_X} & FA_X \\
 & & \Big\downarrow {[id_X, \alpha]^{\#}} & & \Big\downarrow {F[id_X, \alpha]^{\#}} \\
 & {}_{id_X} \searrow & X & \xleftarrow{\alpha} & FX
\end{array}
$$

Notice that by the commutativity of the left square of the diagram above (seen, alternatively, as a diagram in $\mathbf{C}$ or in F-**alg**) the following equalities hold, for all X in $\mathbf{C}$ and (X, α) in F-**alg**:

$$U^F \epsilon_{(X,\alpha)} \circ \eta_{U^F(X,\alpha)} = id_{U^F X} \qquad\qquad \epsilon_{L^F X} \circ L^F \eta_X = id_{L^F X}$$

It only remains to prove the naturality of the definitions of η and ϵ. As for η, naturality is given by the commutativity of the left square of Diagram (3.2.2). Note, moreover, that the right square of Diagram (3.2.2) states that ψ_X is natural in X as a natural transformation $\psi \colon FU^F L^F \Rightarrow U^F L^F$. Naturality of ϵ follows by initiality and by the fact that ψ is natural.

The above construction can be summarized are follows.

Theorem 3.2.7 *Let $F \colon \mathbf{C} \to \mathbf{C}$ be a functor in a category $\mathbf{C}$ with binary coproducts. If for any object X in $\mathbf{C}$ the functor $X + F$ has initial algebra, then the forgetful functor $U^F \colon F\text{-}\mathbf{alg} \to \mathbf{C}$ has a left adjoint $L^F \colon \mathbf{C} \to F\text{-}\mathbf{alg}$.*

From adjunctions to initial algebras

We have already seen at the beginning of this section that if the forgetful functor $U^F \colon F\text{-}\mathbf{alg} \to \mathbf{C}$ has a left adjoint, namely, the functor

$$L^F \colon \mathbf{C} \to F\text{-}\mathbf{alg} \qquad\qquad X \mapsto (A_X, \psi_X \colon FA_X \to A_X),$$

then, F-**alg** comes with a principle of free construction, which can be restated as:

$$
\mathbf{C} \xrightleftharpoons[U^F]{L^F} F\text{-}\mathbf{alg}
$$

$$
\begin{array}{ccc}
X \xrightarrow{\eta_X} A_X \;\; {}= U^F L^F X & & A_X \xleftarrow{\psi_X} FA_X \\
{}_{f}\searrow \;\; \Big\downarrow {f^{\#} = U^F f^{\#}} & & {}_{f^{\#}}\Big\downarrow \;\;\;\; \Big\downarrow {Ff^{\#}} \\
Y \;\; = U^F(Y, \beta) & & Y \xleftarrow{\beta} FY
\end{array}
\qquad (3.2.3)
$$

where $\eta \colon Id \Rightarrow U^F L^F$ is the unit of the adjunction.

The algebra $L^F X = (A_X, \psi_X)$ is usually called *free F-algebra over X*, and since left adjoints preserves all colimits, the existence of initial F-algebras is guaranteed if the underlying category $\mathbf{C}$ has an initial object $\mathbf{0}$. Formally we can state the following theorem.

Theorem 3.2.8 (Existence of initial F-algebras) *Let $F \colon \mathbf{C} \to \mathbf{C}$ be a functor on a category with initial object $\mathbf{0}$. If the forgetful functor $U^F \colon F\text{-}\mathbf{alg} \to \mathbf{C}$ has a left adjoint $L^F \colon \mathbf{C} \to F\text{-}\mathbf{alg}$, then $L^F \mathbf{0}$, the free F-algebra over $\mathbf{0}$, is initial.*

Combining Theorems 3.2.7 and 3.2.8, we have that an initial algebra $(A_{\mathbf{0}}, \iota_{\mathbf{0}})$ for the functor $\mathbf{0} + F$ gives rise to an initial algebra $(A_{\mathbf{0}}, \psi_{\mathbf{0}})$ for F, with algebra structure $\psi_{\mathbf{0}} = \iota_{\mathbf{0}} \circ in_{FA_{\mathbf{0}}}^{\mathbf{0}}$. Of course, this happens in case the category $\mathbf{C}$ has both initial object $\mathbf{0}$ and binary coproducts. Note that, the existence of an adjunction $L^F \dashv U^F$ does not require the existence of binary coproducts in $\mathbf{C}$, therefore the principle presented in Theorem 3.2.8 is more general.

The adjunction $(\eta, \epsilon): L^F \dashv U^F$ generates a (free) monad over $\mathbf{C}$, defined as

$$T^F = (U^F L^F, \eta, U^F \epsilon L^F)$$

called the *monad freely generated by F*. By a little abuse of notation the composite functor $U^F L^F: \mathbf{C} \to \mathbf{C}$ is usually simply denoted by $T^F: \mathbf{C} \to \mathbf{C}$, the unit η by η^F, and multiplication $U^F \epsilon L^F$ by μ^F. If F is clear from the context the monad will be denoted simply by (T, η, μ).

Notably, any monad freely generated by a functor $F: \mathbf{C} \to \mathbf{C}$ comes with a *structural induction proof principle*, formalized as follows.

Definition 3.2.9 (Structural induction) *Let $F: \mathbf{C} \to \mathbf{C}$ be a functor and (T^F, η^F, μ^F) be the monad freely generated by F. Then, for any object X in $\mathbf{C}$, arrow $f: X \to Y$, and F-algebra (Y, β), there exists a unique arrow $f^\#: T^F X \to Y$ making the following diagrams commute*

$$
\begin{array}{ccccc}
X & \xrightarrow{\eta^F_X} & T^F X & \xleftarrow{\psi_X} & F T^F X \\
& & & & \\
{\scriptstyle f} \searrow & & \downarrow {\scriptstyle f^\#} & & \downarrow {\scriptstyle F f^\#} \\
& & & & \\
& & Y & \xleftarrow{\beta} & F Y
\end{array}
$$

where $\psi_X: F T^F X \to T^F X$ is the free F-algebra structure over X. The arrow $f^\#$ is said the (free) *inductive extension of β along (the valuation) f.*

Note that, the diagram in the above statement is nothing but Diagram (3.2.3). Indeed it is just a restatement of the principle of free construction given by the adjunction $L^F \dashv U^F$ from which the monad is generated.

Moreover, the structural induction proof principle of Definition 3.4.7 extends to a definition principle for natural transformations:

Lemma 3.2.10 *Let $F, G, H: \mathbf{C} \to \mathbf{C}$ be functors and (T^F, η^F, μ^F) be the monad freely generated by F. Then, any two natural transformations $\phi: G \Rightarrow H$ and $\varphi: FG \Rightarrow H$ uniquely define a natural transformation $\rho: T^F G \Rightarrow H$, making the following diagrams commute*

$$
\begin{array}{ccccc}
G & \xrightarrow{\eta^F G} & T^F G & \xleftarrow{\psi G} & F T^F G \\
& & & & \\
{\scriptstyle \phi} \searrow & & \vdots {\scriptstyle \rho} & & \vdots {\scriptstyle F\rho} \\
& & & & \\
& & H & \xleftarrow{\varphi} & F H
\end{array}
$$

where $\psi: F T^F \Rightarrow T^F$ is the natural transformation induced by free F-algebras structures over the objects in $\mathbf{C}$. The natural transformation ρ is said the (free) *inductive extension of φ along ϕ.*

Proof. We define ρ component-wise. For any object X in $\mathbf{C}$, we define ρ_X as the unique (free) inductive extension of φ_X along the valuation ϕ_X. Naturality of ρ can be established with the uniqueness aspect of the structural induction proof principle for the monad (T^F, η^F, μ^F). ∎

Remark 3.2.11 If the category $\mathbf{C}$ has binary coproducts and, for each object X in $\mathbf{C}$, the functor $X + F$ has initial algebra (A_X, ι_X), then (T^F, η^F, μ^F), the monad freely generated by F, can be defined directly without passing first through the adjunction $L^F \dashv U^F$.

Indeed, the mapping $X \mapsto A_X$ is functorial, with action on arrows $f: X \to Y$ given by $f \mapsto f^\#$, where $f^\#: A_X \to A_Y$ is the unique arrow making Diagram 3.2.2 commute. So that one defines $T^F: \mathbf{C} \to \mathbf{C}$ as follows:

$$T^F X = A_X \qquad\qquad\qquad T^F f = f^\#$$

The unit $\eta^F \colon Id \Rightarrow T^F$ is defined component-wise by $\eta_X = \iota_X \circ in_X^X$ (note that, it correspond to the unit of the adjunction $L^F \dashv U^F$) and the multiplication $\mu^F \colon T^F T^F \Rightarrow T^F$ is defined as the inductive extension of $\psi \colon F T^F \Rightarrow T^F$ along the identity natural transformation $id_{T^F} \colon T^F \Rightarrow T^F$:

$$
\begin{array}{ccccc}
T^F & \xrightarrow{\eta T^F} & T^F T^F & \xleftarrow{\psi T^F} & F T^F T^F \\
& & \Big\downarrow \mu^F & & \Big\downarrow F \mu^F \\
id_{T^F} \searrow & & T^F & \xleftarrow{\psi} & F T^F
\end{array}
$$

The commutativity of the triangle on the left, shows that η^F and μ^F satisfy the left unit law. As for the right unit law, exploit the uniqueness of the inductive extension, noticing that both id_{T^F} and the composite $\mu^F \circ T^F \eta^F$ fit as the unique inductive extension of ψ along η^F. Similarly, one can prove the associativity law by showing that both $\mu^F T^F \circ \mu^F$ and $T^F \mu^F \circ \mu^F$ fit as the inductive extension of ψ along μ^F. ∎

We conclude this section showing that the monad freely generated by any syntactic **Set**-functors gives rise to the *monad of freely generated terms*.

Example 3.2.12 (Terms Monad) Let (Σ, ar) be a signature and $S = \coprod_{\sigma \in \Sigma} Id^{ar(\sigma)}$ be the syntactic **Set**-functor associated to the signature. In Example 3.2.5 we showed that for any set X the functor $X + S$ has initial algebra $(TX, [\eta_X, \psi_X])$, where TX is the set of terms freely generated by the signature (Σ, ar) over the variables in X, and the algebra structure is given by

$$
\eta_X(x) = x \qquad \text{and} \qquad \psi_X(\langle \sigma, (t_1, \ldots, t_{ar(\sigma)}) \rangle) = \sigma(t_1 \ldots t_{ar(\sigma)}),
$$

for all $x \in X$, $\sigma \in \Sigma$, and $t_1, \ldots, t_{ar(\sigma)} \in TX$. By Remark 3.4.9, we can define a monad (T^S, η^S, μ^S) with functor $T^S \colon \mathbf{Set} \to \mathbf{Set}$ given by

$$
\begin{aligned}
T^S X &= TX & T^S f(x) &= f(x) \\
T^S(f \colon X \to Y) &= f^\# \colon TX \to TY & T^S f(\sigma(t_1, \ldots t_{ar(\sigma)})) &= \sigma(T^S f(t_1), \ldots, T^S f(t_{ar(\sigma)}))
\end{aligned}
$$

for all $x \in X$, $\sigma \in \Sigma$, and $t_1, \ldots, t_{ar(\sigma)} \in TX$; unit $\eta_X^S = \eta_X \colon X \to TX$ (the insertion-of-variables function); and multiplication $\mu_X^S \colon TTX \to TX$ (the operation which allows one to plug terms into contexts) inductively defined as follows

$$
\begin{aligned}
\mu_X^S(t) &= t \\
\mu_X^S(\sigma(C_1, \ldots C_{ar(\sigma)})) &= \sigma(\mu^S(C_1), \ldots, \mu^S(C_{ar(\sigma)}))
\end{aligned}
$$

for all $t \in TX$, $\sigma \in \Sigma$, and $C_1, \ldots, C_{ar(\sigma)} \in TTX$ (i.e., contexts). ∎

3.2.2 Algebras for a Monad

We saw that when the the forgetful functor $U^F \colon F\text{-}\mathbf{alg} \to \mathbf{C}$ has a left adjoint $L^F \colon \mathbf{C} \to F\text{-}\mathbf{alg}$, the functor $F \colon \mathbf{C} \to \mathbf{C}$ admits a free monad (T^F, η^F, μ^F) generated by it. Moreover, the adjunction gives rise to a construction of free F-algebra $L^F X = (T^F X, \psi_X)$ for any object X in $\mathbf{C}$, whose algebra structures are natural in the sense that $\psi = (\psi_X \colon F T^F X \to T^F X)_{X \in \mathbf{C}}$ is a natural transformation. This gives rise to a natural transformation $\theta \colon F \Rightarrow T^F$ defined as the composite

$$
\theta = \psi \circ F \eta^F \qquad\qquad F \xrightarrow{\ F \eta^F \ } F T^F \xrightarrow{\ \psi \ } T^F
$$

The naturality of θ allows us to transform any T^F-algebra (X, α) to an F-algebra $(X, \theta_X \circ \alpha)$, and any T^F-homomorphism to an F-homomorphism, by naturality of θ, in the following way

$$
\begin{array}{ccc}
\begin{array}{ccc}
T^F X & \xrightarrow{\alpha} & X \\
{\scriptstyle T^F f}\downarrow & & \downarrow{\scriptstyle f} \\
T^F Y & \xrightarrow[\beta]{} & Y
\end{array}
& \mapsto &
\begin{array}{ccccc}
FX & \xrightarrow{\theta_X} & T^F X & \xrightarrow{\alpha} & X \\
{\scriptstyle Ff}\downarrow & & \downarrow{\scriptstyle T^F f} & & \downarrow{\scriptstyle f} \\
FY & \xrightarrow[\theta_Y]{} & T^F Y & \xrightarrow[\beta]{} & Y
\end{array}
\end{array}
\tag{3.2.4}
$$

This transformation is functorial, i.e., it extends to a functor $K \colon T^F\text{-}\mathbf{alg} \to F\text{-}\mathbf{alg}$.

We will see that this functor takes part into an isomorphism between categories, where $T^F\text{-}\mathbf{alg}$ will be the *Eleinberg-Moore category for the monad* (T^F, η^F, μ^F), i.e., the category of algebras for the monad (T^F, η^F, μ^F).

Definition 3.2.13 (Eleinberg-Moore algebra) *Let (T, η, μ) be a monad in $\mathbf{C}$. An algebra of the monad (T, η, μ) is a T-algebra (X, α) such that the two diagrams below commute, which we call the* unit *and* multiplication laws *of the algebra, respectively.*

$$
\begin{array}{ccc}
X & \xrightarrow{\eta_X} & TX \\
& {\scriptstyle id_X}\searrow & \downarrow{\scriptstyle \alpha} \\
& & X
\end{array}
\qquad\qquad
\begin{array}{ccc}
TX & \xleftarrow{\mu_X} & TTX \\
{\scriptstyle \alpha}\downarrow & & \downarrow{\scriptstyle T\alpha} \\
X & \xleftarrow[\alpha]{} & TX
\end{array}
$$

The category of all algebras for the monad (T, η, μ) and T-homomorphisms between them, is said the Eleinberg-Moore category for (T, η, μ)*, and it is denoted by $(T, \eta, \mu)\text{-}\mathbf{alg}$.*

The unit and multiplication laws can intuitively be explained saying that the algebra structure α must respect the structure of the monad. By an abuse of notation we will denote $(T, \eta, \mu)\text{-}\mathbf{alg}$ simply as $T\text{-}\mathbf{alg}$, when it is clear from the contexts that T belongs to a monad.

Lemma 3.2.14 *Let $F \colon \mathbf{C} \to \mathbf{C}$ be a functor, (T^F, η^F, μ^F) be the monad freely generated by F, and $\theta \colon F \Rightarrow T^F$ be the natural transformation induced by $\psi \colon FT^F \Rightarrow T^F$, the free algebra natural transformation. Then $\psi = \mu^F \circ \theta T^F$.*

Proof. The thesis follows by definition of μ^F (see also Remark 3.4.9) and by the unit law for (T^F, η^F, μ^F), which make commute the right square and the left triangle of the following diagram, respectively

$$
\begin{array}{ccccc}
FT^F & \xrightarrow{F\eta^F T^F} & FT^F T^F & \xrightarrow{\psi T^F} & T^F T^F \\
& {\scriptstyle id}\searrow & \downarrow{\scriptstyle F\mu^F} & & \downarrow{\scriptstyle \mu^F} \\
& & FT^F & \xrightarrow[\psi]{} & T^F
\end{array}
$$

Then $\psi = \mu^F \circ \theta T^F$ follows, since, by definition, $\theta = \psi \circ F\eta^F$. $\blacksquare$

Next we show that the category of F-algebras is isomorphic to that of Eilenberg-Moore algebras for the monad (T^F, η^F, μ^F) freely generated by F.

Lemma 3.2.15 ($F\text{-}\mathbf{alg} \cong T^F\text{-}\mathbf{alg}$) *Let $F \colon \mathbf{C} \to \mathbf{C}$ be a functor, (T^F, η^F, μ^F) be the monad freely generated by F. Then, the categories $F\text{-}\mathbf{alg}$ and $T^F\text{-}\mathbf{alg}$ are isomorphic.*

Proof. We define the functors $K\colon T^F\text{-}\mathbf{alg} \to F\text{-}\mathbf{alg}$ and $H\colon F\text{-}\mathbf{alg} \to T^F\text{-}\mathbf{alg}$, then we prove they are inverse of each other. Let (X,α), (Y,β) be (T^F,η^F,μ^F)-algebras and (P,h), (Q,k) be F-algebras, then we define the two functors as follows

$$K\colon T^F\text{-}\mathbf{alg} \to F\text{-}\mathbf{alg} \qquad\qquad H\colon F\text{-}\mathbf{alg} \to T^F\text{-}\mathbf{alg}$$
$$K(X,\alpha) = (X,\theta_X \circ \alpha) \qquad\qquad H(P,h) = (P,h^*)$$
$$K(f\colon (X,\alpha) \to (Y,\beta)) = f \qquad\qquad H(g\colon (P,h) \to (Q,k)) = g$$

where $\theta\colon F \Rightarrow T^F$ is the natural transformation induced by $\psi\colon FT^F \Rightarrow T^F$ and $\eta^F\colon Id \Rightarrow T^F$, and h^* is defined as the inductive extension of h along id_P, as follows

$$
\begin{array}{ccccc}
P & \xrightarrow{\eta_P^F} & T^F P & \xleftarrow{\psi_P} & FT^F P \\
& \searrow^{id_P} & \downarrow^{h^*} & & \downarrow^{Fh^*} \\
& & P & \xleftarrow{h} & FP
\end{array}
$$

Diagram (3.2.4) proves that K is well defined. As for H we need to prove the unit and multiplication law of the algebra. The unit law follows by definition (left triangle in the above diagram). The multiplication law follows since both $h^* \circ \mu_P^F$ and $h^* \circ T^F h^*$ fit as the unique extension of h along η_P^F. The F-homomorphism $g\colon (P,h) \to (Q,k)$ is proved to be a T^F-homomorphism by showing that both $k^* \circ Tg$ and $g \circ h^*$ fit as the unique inductive extension of ψ_Q along g. Functoriality of H follows similarly, exploiting again the universal property of structural induction. It remains to show that K and H are inverses of each other. On arrows is clear. As for objects, we have

$$KH(P,h) = (P,h^* \circ \theta_P), \qquad\qquad HK(X,\alpha) = (X,(\theta_X \circ \alpha)^*),$$

thus, we need to show $h^* \circ \theta_P = h$ and $(\theta_X \circ \alpha)^* = \alpha$. These are proved by the following diagrams:

$$
\begin{array}{cccc}
& & \theta_P & \\
T^F P & \xleftarrow{\psi_P} FT^F P & \xleftarrow{F\eta_P^F} & FP \\
\downarrow^{h^*} & \downarrow^{Fh^*} & & \\
P & \xleftarrow{h} FP & &
\end{array}
\qquad
\begin{array}{ccccc}
X & \xrightarrow{\eta_X^F} & T^F X & \xleftarrow{\psi_X} & FT^F X \\
& & \downarrow^{\alpha} & & \downarrow^{F\alpha} \\
& id_X & & T^F T^F X & \\
& & X & \xleftarrow{\alpha} T^F X \xleftarrow{\theta_X} & FX
\end{array}
$$

The diagram on the left commutes by definition of h^*, the one on the right by Lemma 3.2.14, naturality of θ, and by unit and multiplication laws for α. In particular, this diagram shows that α is the inductive extension of $\theta_X \circ \alpha$ along the identity id_X, so that, by uniqueness it must coincide with $(\theta_X \circ \alpha)^*$. ∎

Remark 3.2.16 (On the Beck's theorem) The category of Eleinberg-Moore algebras is one of the most studied in the theory of universal algebras and, more in general, in the theory of categories. Its importance is due to the Beck's theorem [] which revealed the strict relationship between monads and adjunctions. In fact, not only every adjunction gives rise to a monad, but also, conversely, every monad spits into an adjunction. In general, there are many categories $\mathbf{D}$ such that a monad in $\mathbf{C}$ spits into an adjunction from $\mathbf{C}$ to $\mathbf{D}$, but there are two canonical ones, namely, the initial and the final ones in a suitable sense. The final one is the Eleinberg-Moore category for the monad. ∎

3.3 Coalgebras and bisimulations

We recall the definition of coalgebra for a functor, a notion dual to that of algebra for a functor (Definition 3.1.1), which has been proved to be a very useful tool for modeling dynamic systems abstractly. The advantages of this approach is that once we have found that the systems we are interested in are coalgebras of some functor, several meaningful notions and results immediately become available. For instance, an abstract notion of *bisimulation*.

In this section, coalgebraic bisimulation will be discussed at length since we want to make the point that some of the peculiarities of bisimulation are due to the fact that the structure of the (largest) bisimulation is in general not uniquely determined. This will prepare to the notion of *cocongruences* and *behavioral equivalences*, for which the structure is uniquely determined.

Definition 3.3.1 (F-coalgebra) *Let $F \colon \mathbf{C} \to \mathbf{C}$ be a functor. An F-coalgebra is a pair (X, α), consisting of an object X in $\mathbf{C}$, carrier, and an arrow $\alpha \colon X \to FX$ in $\mathbf{C}$, coalgebra structure.*

We often call the functor F used to define a class of coalgebras a behaviour functor. We do so only to stress the rôle of F, not to restrict the type of functors under consideration.

Definition 3.3.2 (F-homomorphism) *Let $F \colon \mathbf{C} \to \mathbf{C}$ be an endofunctor and (X, α) and (Y, β) be F-coalgebras. An arrow $f \colon X \to Y$ in $\mathbf{C}$ is a F-homomorphism between (X, α) and (Y, β) if the following diagram in $\mathbf{C}$ commutes:*

$$\begin{array}{ccc} X & \xrightarrow{\;Ff\;} & Y \\ {\scriptstyle \alpha}\downarrow & & \downarrow{\scriptstyle \beta} \\ FX & \xrightarrow{\;f\;} & FY \end{array}$$

The following example shows how *labelled transition systems* and homomorphisms between them can be elegantly modeled as coalgebras.

Example 3.3.3 (Labelled transition systems) For a non-empty set of labels L, an *L-labelled transition system* (LTS) is a pair $(X, \{\xrightarrow{a}\}_{a \in L})$ consisting of a set of states X and an L-indexed collection of labelled transition relations $\xrightarrow{a} \subseteq X \times X$. An homomorphism between two L-labelled transition systems $(X, \{\xrightarrow{a}\}_{a \in L})$ and $(Y, \{\xrightarrow{a}\}_{a \in L})$ is a map $h \colon X \to Y$ such that, for all $a \in L$ and $x, x' \in X$

$$x \xrightarrow{a} x' \text{ in } X \qquad \text{implies} \qquad h(x) \xrightarrow{a} h(x') \text{ in } Y$$

that is, h respects the structure of the system.

Notice that, any relation $R \subseteq X \times X$ can be represented as a function $\tilde{R} \colon X \to \mathcal{P}(X)$, where $\mathcal{P}(X)$ is the powerset of X and the correspondence, for all $x, x' \in X$, is given by

$$x \, R \, x' \qquad \Longleftrightarrow \qquad x' \in \tilde{R}(x) \,.$$

Thus, L-labelled transition systems correspond to coalgebras for the **Set**-functor $\mathcal{P}^L \colon \mathbf{Set} \to \mathbf{Set}$, acting on objects X and arrows $f \colon X \to Y$, respectively, as

$$\mathcal{P}^L X = \{g \mid g \colon L \to \mathcal{P}(X)\}$$
$$\mathcal{P}^L f(g)(a) = \{f(y) \mid y \in g(a)\}$$

for all $a \in L$ and $g \colon L \to \mathcal{P}(X)$. Indeed, given an L-labelled transition system $(X, \{\xrightarrow{a}\}_{a \in L})$ and a $\mathcal{P}^L$-coalgebra (X, α), we get the following correspondence, for all $a \in L$

$$x \xrightarrow{a} x' \qquad \Longleftrightarrow \qquad x' \in \alpha(x)(a) \,.$$

It is immediate to check that any homomorphism between labelled transition systems is also a $\mathcal{P}^L$-homomorphism between the corresponding coalgebraic translations, and vice versa. ∎

Definition 3.3.4 (*F*-bisimulation) *Let* $F\colon \mathbf{C} \to \mathbf{C}$ *be an endofunctor. A monic span* (R, f, g) *in* $\mathbf{C}$ *between* X *and* Y *is a* F-bisimulation *between* F-coalgebras (X, α) *and* (Y, β) *if there exists a (not necessarily unique) coalgebra structure* $\gamma\colon R \to FR$ *on* R *making the following diagram in* $\mathbf{C}$ *commute*

$$
\begin{array}{ccccc}
X & \xleftarrow{\;f\;} & R & \xrightarrow{\;g\;} & Y \\
{\scriptstyle\alpha}\downarrow & & {\scriptstyle\gamma}\downarrow & & \downarrow{\scriptstyle\beta} \\
FX & \xleftarrow{Ff} & FR & \xrightarrow{Fg} & FY
\end{array}
$$

that is, making f *and* g *morphisms of* F-coalgebras.

The requirement that (R, f, g) is a mono span generalizes $R \subseteq X \times Y$ in **Set** to arbitrary categories. Notice that, this notion is similar, but not dual to that of a F-congruence (cf. Definition 3.1.5).

Historical note. The abstract coalgebraic definition of bisimulation given above is due to Aczel and Mendler [4], but the notion of bisimulation first appeared in modal logic under the name of p-morphism [76] and zigzag relation [84, 85]. Then, it became notorious after it was successfully applied in concurrency theory by Park [70] and Milner [65, 66]. Bisimulations were used by Aczel [3] to define equality for non-well founded sets and to prove existence of final coalgebras.

In the coalgebraic context, many different generalized notions of bisimulation have been proposed. The fist work trying to relate them in a formal way is [78]. There Staton identified four definition of bisimulations and gave conditions for the behaviour functor and the underlying category under which they coincide.

For labelled transition systems, there is a well-known notion of (strong) bisimulation. In the next example we are going to show that notion of bisimulation on labelled transition systems exactly corresponds to the coalgebraic one, via the translation we have see in Example 3.3.3.

Example 3.3.5 (Bisimulation on LTSs) Let $(X, \{\xrightarrow{a}_X\}_{a \in L})$ and $(Y, \{\xrightarrow{a}_Y\}_{a \in L})$ be LTSs. A relation $R \subseteq X \times Y$ is a *bisimulation* relation if, whenever $(x, y) \in R$, then for all $a \in L$

- if $x \xrightarrow{a}_X x'$, then there exists $y' \in Y$, such that $y \xrightarrow{a}_Y y'$ and $(x, y) \in R$;

- if $y \xrightarrow{a}_Y y'$, then there exists $x' \in X$, such that $x \xrightarrow{a}_X x'$ and $(x, y) \in R$.

This definition is a particular case of the generalized coalgebraic notion of Definition 3.3.4. Consider the "labelled powerset functor" $\mathcal{P}^L\colon \mathbf{Set} \to \mathbf{Set}$ and the $\mathcal{P}^L$-coalgebras (X, α_X) and (Y, α_Y) corresponding to $(X, \{\xrightarrow{a}_X\}_{a \in L})$ and $(Y, \{\xrightarrow{a}_Y\}_{a \in L})$, respectively, given as in Example 3.3.3:

$$
x \xrightarrow{a}_X x' \iff x' \in \alpha_X(x)(a), \qquad\qquad y \xrightarrow{a}_Y y' \iff y' \in \alpha_Y(y)(a).
$$

To prove that the monic span (R, π_X, π_Y) is $\mathcal{P}^L$-bisimulation we have to provide a $\mathcal{P}^L$-coalgebra structure $\gamma\colon R \to \mathcal{P}^L R$ on R such that $\alpha_X \circ \pi_X = \mathcal{P}^L \pi_X \circ \gamma$ and $\alpha_Y \circ \pi_Y = \mathcal{P}^L \pi_Y \circ \gamma$. We do this as follows, for $(x, y) \in R$ and $a \in L$

$$
(x', y') \in \gamma((x, y))(a) \qquad \iff \qquad x \xrightarrow{a}_X x' \text{ and } y \xrightarrow{a}_Y y'
$$

Note that, γ is well-defined since we are guaranteed that $(x', y') \in R$ by the assumption that R is a bisimulation. The equality $\alpha_X \circ \pi_X = \mathcal{P}^L \pi_X \circ \gamma$ is proved by following equivalences

$$
\begin{aligned}
(x', y') \in \mathcal{P}^L \pi_x \circ \gamma(x, y)(a) &\iff (x', y') \in \mathcal{P}^L \pi_X(\gamma((x, y)))(a) \\
&\iff (x', y') \in \{\pi_X(x', y') \mid (x', y') \in \gamma((x, y))(a)\} \\
&\iff (x', y') \in \{\pi_X(x', y') \mid x \xrightarrow{a}_X x' \text{ and } y \xrightarrow{a}_Y y'\}
\end{aligned}
$$

$$\Longleftrightarrow x' \in \{x' \mid x \xrightarrow{a}_X x'\}$$
$$\Longleftrightarrow x' \in \alpha_X(x)(a)$$
$$\Longleftrightarrow (x', y') \in \alpha_X(\pi_X(x, y))(a)$$
$$\Longleftrightarrow (x', y') \in \alpha_X \circ \pi_X(x, y)(a),$$

Similarly, also $\alpha_Y \circ \pi_Y = \mathcal{P}^L \pi_Y \circ \gamma$ holds, thus (R, π_X, π_Y) is a $\mathcal{P}^L$-bisimulation.

Conversely, assume (R, π_X, π_Y) is a $\mathcal{P}^L$-bisimulation with coalgebra structure $\gamma \colon R \to \mathcal{P}^L R$. Assume $(x, y) \in R$, $a \in L$, and $x \xrightarrow{a}_X x'$. Then, by $\alpha_X \circ \pi_X = \mathcal{P}^L \pi_X \circ \gamma$ and definition of α_X it follows that, there exists $y' \in Y$ such that $(x', y') \in R$ and $(x', y') \in \gamma((x, y))(a)$. From $\alpha_Y \circ \pi_Y = \mathcal{P}^L \pi_Y \circ \gamma$, we have that $y' \in \alpha_Y(y)(a)$, therefore $y \xrightarrow{a}_Y y'$. This proves the first condition for R to be a bisimulation. The second one follows similarly. ∎

Note that, the mediating coalgebra structure γ in Definition 3.3.4 is not necessarily uniquely determined. The following example, taken from [72], make this clear.

Example 3.3.6 (Non-uniqueness) Consider the *powerset functor* $\mathcal{P} \colon \mathbf{Set} \to \mathbf{Set}$, acting on objects X and arrows $f \colon X \to Y$, respectively, as follows, for $X' \subseteq X$

$$\mathcal{P}X = \mathcal{P}(X)$$
$$\mathcal{P}f(X') = f(X') = \{f(x') \mid x' \in X'\}.$$

Let (X, α) be a coalgebra for $\mathcal{P} \colon \mathbf{Set} \to \mathbf{Set}$ defined as $X = \{x_0, x_1, x_2\}$ and with coalgebra structure $\alpha \colon X \to \mathcal{P}X$ depicted below

$$\alpha(x_0) = \{x_1, x_2\}$$
$$\alpha(x_1) = \emptyset$$
$$\alpha(x_2) = \emptyset$$

Clearly the largest bisimulation on (X, α) is $R = \{(x_0, x_0), (x_1, x_1), (x_2, x_2), (x_1, x_2), (x_2, x_1)\}$, but there is no unique coalgebra structure on R as shown by the following two examples:

Uniqueness of the coalgebraic structure depends on the behavior functor. Indeed, if the functor (strongly) preserves pullbacks, the structure must be uniquely determined by the universal property. In the example above, this does not work since the powerset functor $\mathcal{P}$ only weakly preserves pullbacks, thus for the structure it is only guaranteed the existence but not its uniqueness. ∎

The category of F-coalgebras

For any endofunctor $F \colon \mathbf{C} \to \mathbf{C}$, it is easy to check that F-coalgebras and F-homomorphisms form a category, denoted by F-$\mathbf{coalg}$. Note that, F-$\mathbf{coalg}$ is not the proper dual of F-$\mathbf{alg}$, but they can be considered dual in a weak sense.

For example, we saw that F-$\mathbf{alg}$ lifts all limits from the underlying category $\mathbf{C}$, and "dually" the category of F-coalgebra lifts all colimits in $\mathbf{C}$, thus, if $\mathbf{C}$ is complete, so is F-$\mathbf{coalg}$. This allows to define derived structures such as coproducts, coequalizes, and pushouts. For example, the coproduct of the F-coalgebras (X, α_X) and (Y, α_Y) is given by

where the structure map α_{X+Y} is the one uniquely determined by the universal property of the coproduct $X + Y$ in $\mathbf{C}$, and we denote it by $(X, \alpha_X) + (Y, \alpha_Y) = (X + Y, \alpha_{X+Y})$. Notice that, the carrier of the coproduct of F-coalgebras corresponds exactly to the coproduct of the carriers in $\mathbf{C}$. This holds in general, that is, the carrier of a colimit in F-**coalg** is the colimit in $\mathbf{C}$. Categorically speaking, this amounts to say that the *forgetful functor* $U_F\colon F\text{-}\mathbf{coalg} \to \mathbf{C}$, forgetting the coalgebraic structure of the objects, creates and preserves colimits.

As for limits, it is different. Products and equalizers of F-coalgebras need not exist, but when they do, their carrier will often have to be different from the product of the carriers in $\mathbf{C}$. If the behavior functor F preserves a certain type of limit, the same type of limit exist for F-coalgebras and it is constructed as in $\mathbf{C}$. In general, though, the behavior functors usually employed in applications, fail to preserve arbitrary limits, an it would be extremely limiting to restrict the attention to such a class of functors.

Historical note. The existence of limits does not only depend on the type of the behaviour functor. For example, considering **Set** as underlying category, Worrell was able to show in [91] that F-**coalg** is complete, that is, products and equalizers exist, provided that behaviour functor $F\colon \mathbf{Set} \to \mathbf{Set}$ weakly preserves pullbacks and is bounded. Worrell's proof uses the theory of monads and some further category theoretic machinery. A shorter and more elementary proof of the same result was proposed by Gumm et al. in [52], which at the same time extended the result by removing the assumption that the functor F should weakly preserve generalized pullbacks. In doing so, they redefined the notion of boundedness for a functor, and showed that terminal coalgebras, more generally, arbitrarily large cofree coalgebras, exist in F-**coalg**, whenever F is bounded in their sense.

A note on largest bisimulations

We discuss two (canonical) ways to obtain largest bisimulations. The first requires that some suitable classes of morphisms in the base category are split, the second that the behavior functor preserves weak pullbacks. Of course, these requirements are not always met, but when they do largest bisimulations aways exist. These constructions are taken from [62].

Union of bisimulations.   Bisimulations on X and Y consist of monic spans (R, f, g). In case base category $\mathbf{C}$ has binary products, monic spans (R, f, g) are in one-to-one correspondence with monomorphisms of type $R \to X \times Y$. The largest bisimulation on X and Y may be defined as the (generalized) union of all bisimulations on X and Y.

There are several ways to describe the notion of union categorically. An axiomatic approach is via factorization systems (see the last part of Section 2.1). To this end one have to assume that:

  (i) $\mathbf{C}$ has a faction system $(\mathcal{L}, \mathbf{Monic})$, for some suitable class of morphisms $\mathcal{L}$;

 (ii) $\mathbf{C}$ is well-powered, i.e., each object in $\mathbf{C}$ has, up to isomorphism, only a *set* (and not a proper class) of subobjects;

(iii) $\mathbf{C}$ has small coproducts and binary products;

(iv) the morphisms in $\mathcal{L}$ are right invertible.

Well-poweredness for $\mathbf{C}$ is required in order to take coproducts over all bisimulations, and right invertible arrows in $\mathcal{L}$ are asked in order endow R with a suitable coalgebra structure.

Proposition 3.3.7 (Union of bisimulations [62]) *Let $\mathbf{C}$ be a well-powered category with small coproducts, binary products, and $(\mathcal{L}, \mathbf{Monic})$-factorizations such that the morphism in $\mathcal{L}$ are right invertible. Then, for all functors $F\colon \mathbf{C} \to \mathbf{C}$, and F-coalgebras (X, α) and (Y, β), the largest bisimulation between them exists and it is given by $(R, \pi_X \circ \rho, \pi_Y \circ \rho)$, where $\rho\colon R \to X \times Y$ is the union of $\mathcal{B} = \{\langle r'_1, r'_2 \rangle\colon R' \to X \times Y \mid (R', r'_1, r'_2)\ F\text{-bisimulation on } (X, \alpha) \text{ and } (Y, \beta)\}$.*

Proof. Let $\mathbf{Monic}(X \times Y)$ be the category of monic morphism with codomain $X \times Y$. Since $\mathbf{C}$ has binary products, this category is equivalent to the category of monic spans between X and Y. The union $\rho \colon R \to X \times Y$ in $\mathbf{Monic}(X \times Y)$ of all monic spans in $\mathcal{B}$ exists since $\mathbf{C}$ has coproducts and a factorization system $(\mathcal{L}, \mathbf{Monic})$, and by the fact that, the collection $\mathcal{B}$, by well-poweredness of $\mathbf{C}$, is is a proper set (up-to-isomorphsim), so that $\coprod_{R' \in \mathcal{B}} R'$ exists in $\mathbf{C}$. This is formally given by the following diagram.

$$
\begin{array}{ccc}
\coprod_{R' \in \mathcal{B}} R' & \xrightarrow{\ \lambda \in \mathcal{L}\ } & R \\
{\scriptstyle in_{R'}} \Big\uparrow & \ \ \ \ \xrightarrow{\ \ h\ \ } & \Big\downarrow {\scriptstyle \rho} \\
R' & \xrightarrow[\langle r'_1, r'_2 \rangle]{} & X \times Y
\end{array}
$$

where $\rho \circ \lambda$ is the $(\mathcal{L}, \mathbf{Monic})$-factorization of the unique arrow h such that $h \circ in'_R = \langle r'_1, r'_2 \rangle$ for all $\langle r'_1, r'_2 \rangle \colon R' \to X \times Y$ in $\mathcal{B}$, given by the universal property of coproducts. It remains to show that R can be endowed with a coalgebra structure $\gamma \colon R \to FR$ that renders it a bisimulation. By assumption λ has right inverse r, that is, $\lambda \circ r = id_R$. So that we define $\gamma = F\lambda \circ \sigma \circ r$, where σ is the coalgebra structure of the coproduct of the coalgebras over $\mathcal{B}$. It is routine to check that both $\pi_X \circ \rho$ and $\pi_Y \circ \rho$ are F-homomorphisms from (R, ρ) to (X, α) and (Y, β), respectively. Hence $(R, \pi_X \circ \rho, \pi_Y \circ \rho)$ is an F-bisimulation between (X, α) and (Y, β). Moreover it is the "largest" one since, for all $R' \in \mathcal{B}$, $\lambda \circ in'_R$ is a morphism between coalgebras. $\blacksquare$

Remark 3.3.8 The above proposition is a slight generalization of a classical result due to Rutten [72, Theorem 5.5], which was proved in $\mathbf{Set}$. In fact, $\mathbf{Set}$ is the bests category in which this approach can the used. Indeed, it has an $(\mathbf{Epic}, \mathbf{Monic})$-factorization system, and in this setting the most requiring assumption, i.e. asking for right invertible arrows in the left-class of morphisms, is met (if the axiom of choice is assumed to be valid). In general, it difficult to find factorization systems satisfying these prerequisites, indeed, to the best of our knowledge, this result has been applied only in $\mathbf{Set}$. $\blacksquare$

Weak pullback preserving functors.   If the category F-$\mathbf{coalg}$ has terminal object $\mathbf{1}$ we expect that the kernel of the unique morphism from a coalgebra (X, α) to the terminal object $\mathbf{1}$ is the largest bisimulation on (X, α). This holds if pullbacks from the base category $\mathbf{C}$ are lifted to the category of coalgebras. Such a lifting is possible if pullbacks in $\mathbf{C}$ can be (canonically) endowed with a coalgebra structure and this happens, for example, if the behaviour functor weakly preserves them.

Proposition 3.3.9 ([62]) *Let $\mathbf{C}$ be a category with pullbacks and $F \colon \mathbf{C} \to \mathbf{C}$ be a functor weakly preserving them. Let (X, α) and (Y, β) be F-coalgebras and $\mathbf{1}$ be a terminal object in F-$\mathbf{coalg}$. Then, (R, π_1, π_2) is the largest F-bisimulation on (X, α) and (Y, β) if and only if the following diagram in $\mathbf{C}$ is a pullback.*

$$
\begin{array}{ccc}
R & \xrightarrow{\ \pi_2\ } & Y \\
{\scriptstyle \pi_1} \Big\downarrow & & \Big\downarrow {\scriptstyle U_F !_Y} \\
X & \xrightarrow[U_F !_X]{} & U_F \mathbf{1}
\end{array}
$$

Actually, the assumption that $\mathbf{C}$ has all pullbacks and that they are weakly preserved by F could be weakened. Indeed, only the pullback in the diagram of the statement is needed. Moreover, the existence of the a terminal object is not really needed: a weakly terminal one would be sufficient.

Remark 3.3.10 In Proposition 3.3.9, requiring that the functor weakly preserves pullbacks serves only to guarantee the existence of a (not uniquely determined) structure map. In [62], Kurz noted that the proof of Proposition 3.3.9 makes hidden use of split epimorphisms: if (P, p_1, p_2) is a pullback of a cospan (X, f, g), and (R, r_1, r_2) is a weak pullback for the same cospan, then the

unique morphism $h\colon R \to P$ given by the universal property of the pullback is epic and its right inverse gives rise to the non-unique morphism into the weak pullback. This reveals the connection between Propositions 3.3.9 and 3.3.7.   ∎

3.4   Final coalgebras and Coinduction

In Section 3.2 we have seen that the existence of initial algebras give rise to a generalized induction proof principle. Similarly, the existence of terminal objects in the category of coalgebras, induces an abstract proof principle, dual the notion of induction: *coinduction*. Historically, this proof principle is less know then induction, and a practical explanation for this could be that it cannot be encoded in the classical Zermelo-Fraenkel set theory, unless the *anti-foundation axiom* is taken in place of the axiom of foundation [3].

Definition 3.4.1 (Final F-coalgebra) *Let $F\colon \mathbf{C} \to \mathbf{C}$ be an endofunctor. A final F-coalgebra is a terminal object in F-**coalg**, i.e., an F-coalgebra (Z,ω) such that for any F-coalgebra (X,α) there exists a unique F-homomorphism from (X,α) to (Z,ω).*

The following is the dual of Theorem 3.2.2.

Theorem 3.4.2 (Lambek's lemma) *Final F-coalgebras (Z,ω) are fixed points for $F\colon \mathbf{C} \to \mathbf{C}$, that is, $\omega\colon Z \to FZ$ is an isomorphism in $\mathbf{C}$.*

Intuitively, final coalgebras can be explained using notions from the theory of preordered sets, indeed, categories can be seen as generalized preorders and endofunctors as monotone functions. In this way coalgebras correspond to postfixed points and final coalgebras as greatest fixed points.

Of course, final coalgebras do not need to exists in general. For example there is no final coalgebra for the powerset functor $\mathcal{P}\colon \mathbf{Set} \to \mathbf{Set}$. This is a consequence of Lambek's lemma and Cantor's theorem, which says that there is no set X such that $X \cong \mathcal{P}(X)$. Things work differently if instead of all subsets one restricts to only the finite subsets, that is, if we consider the *finite powerset functor* $\mathcal{P}_\omega\colon \mathbf{Set} \to \mathbf{Set}$, which takes objects X and arrows $f\colon X \to Y$, respectively, to

$$\mathcal{P}_\omega X = \{X' \subseteq X \mid |X'| < \omega\}$$
$$\mathcal{P}_\omega f(X') = f(X') = \{f(x') \mid x' \in X'\}\,,$$

where $X' \subseteq X$ is such that $|X'| < \omega$. This functor admits a final coalgebra, and more generally, final coalgebras exist for all κ-bounded versions $\mathcal{P}_\kappa\colon \mathbf{Set} \to \mathbf{Set}$ of the powerset functor, that is, when the cardinality of the subsets is limited to a fixed (limit) cardinal κ. These are all examples of $\mathbf{Set}$-*bounded functors*, that is, functors F for which there exists a global bound to the size of the carrier set of any one-generated F-coalgebras (see Rutten [72] for further details).

The coinduction proof principle is particularly useful when one wants to define operations on coalgebraic structures such as, labelled transition systems (see Example 3.3.3), deterministic and non-deterministic automata, infinite streams, etc. In the following example we show a simple application of the coinduction proof principle on infinite streams.

Example 3.4.3 (Coinduction on infinite streams) A *stream system* for an alphabet of symbols A is a pair (X,t), where X is a set of states and $t\colon X \to A \times X$ is a transition function, i.e., it is a coalgebra for the functor $A \times Id\colon \mathbf{Set} \to \mathbf{Set}$.

The set A^ω of infinite streams of symbols in A is the largest set satisfying the following rules

$$\frac{a \in A \qquad s \in A^\omega}{as \in A^\omega}$$

This set can be naturally endowed with an $(A \times Id)$-coalgebra structure $\langle hd, tl \rangle\colon A^\omega \to A \times A^\omega$ where $hd\colon A^\omega \to A$ and $tl\colon A^\omega \to A^\omega$ are, respectively, the *head* and *tail* functions, defined respectively as follows, for all $a \in A$ and $s \in A^\omega$

$$hd(as) = a\,, \qquad\qquad\qquad tl(as) = s\,.$$

The function $\langle hd, tl \rangle$ is an isomorphism and, in particular, $(A^\omega, \langle hd, tl \rangle)$ is the final coalgebra for $A \times Id$ (see [73] for a detailed coalgebraic analysis of stream systems).

Consider a simple alternating composition operation $alt \colon A^\omega \times A^\omega \to A^\omega$, acting on infinite streams as follows:

$$alt(a_1 a_2 a_3 \cdots, b_1 b_2 b_3 \cdots) = a_1 b_1 a_2 b_2 a_3 b_3 \cdots$$

To define alt formally by coinduction, one uses the finality of $(A^\omega, \langle hd, tl \rangle)$, and pick an $(A \times Id)$-coalgebra on the set $A^\omega \times A^\omega$, to be though as the "conductive step"

$$h \colon A^\omega \times A^\omega \to A \times (A^\omega \times A^\omega) \qquad\qquad h(s_1, s_2) = \big(hd(s_1), (s_2, tl(s_1))\big)$$

and define alt as the unique map making the following diagram commute:

$$
\begin{array}{ccc}
A^\omega \times A^\omega & \dashrightarrow^{\ alt\ } & A^\omega \\
{\scriptstyle h}\downarrow & & \downarrow{\scriptstyle \langle hd, tl \rangle} \\
A \times (A^\omega \times A^\omega) & \xrightarrow[A \times alt]{} & A \times A^\omega
\end{array}
$$

Note that, to define the function alt coinduction is essential, indeed simple (or structural) induction cannot be applied to define it, and any other operation on A^ω. ∎

3.4.1 From Final Coalgebras to Adjunctions and back

In this section, we dualize the results we saw in Section 3.2.1 in the case of algebras. Since all proofs are similar and need only to be dualized, we do not provide them. Our aim is just to reveal the duality of the two approaches and to investigate how *cofree constructions*, provided by right adjoints, give rise to a more structured coinduction proof principle over cofree comonads.

In case the forgetful functor $U_F \colon F\text{-}\mathbf{coalg} \to \mathbf{C}$ has a right adjoint, namely, the functor $R_F \colon \mathbf{C} \to F\text{-}\mathbf{coalg}$, the category of F-coalgebras admits a *cofree construction*: given any object Y in $\mathbf{C}$, any F-coalgebra (X, α) and morphism $f \colon U_F(X, \alpha) = X \to Y$ in $\mathbf{C}$, there exists a *unique* F-homomorphism $f^\flat$ such that the following diagrams commute

$$
\begin{array}{ccc}
\mathbf{C} & \underset{\xleftarrow[U_F]{}}{\overset{R_F}{\xrightarrow{\hspace{3cm}}}} & F\text{-}\mathbf{coalg}
\end{array}
$$

$$
\begin{array}{ccc}
U_F(X, \alpha) & & (X, \alpha) \\
{\scriptstyle f}\swarrow \ \ \downarrow{\scriptstyle U_F f^\flat} & & \ \ \downarrow{\scriptstyle f^\flat} \\
Y \xleftarrow[\epsilon_Y]{} U_F R_F Y & & R_F Y
\end{array}
$$

where $\epsilon \colon U_F R_F \Rightarrow Id$ is the counit of the adjunction.

Similar to the case of algebras, the existence of the right adjoint R_F and of final coalgebras for some (class of) functors are strictly related with each other.

From final coalgebras to adjunctions

Let $F \colon \mathbf{C} \to \mathbf{C}$ be a functor in a category $\mathbf{C}$ with binary products. Assume, moreover, that for any object X in $\mathbf{C}$ the functor $X \times F$ has final coalgebra (Z_X, ω_X). Under these assumptions we can define the right adjoint of the forgetful functor $U_F \colon F\text{-}\mathbf{coalg} \to \mathbf{C}$, namely, $R_F \colon \mathbf{C} \to F\text{-}\mathbf{coalg}$.

For the coalgebras for the functor $X \times F$, we have the following correspondence lemma.

Lemma 3.4.4 *Let X be an object in $\mathbf{C}$ and $F\colon \mathbf{C} \to \mathbf{C}$ be an endofunctor. An arrow $f\colon A \to B$ is an homomorphism between the $(X \times F)$-coalgebras (A, α_A) and (B, α_B) if and only if the following diagram in $\mathbf{C}$ commutes:*

$$
\begin{array}{ccc}
 & A & \xrightarrow{\ \pi_{FA}^A \circ \alpha_A\ } FA \\
\pi_X^A \circ \alpha_A \swarrow & \ \downarrow f & \ \downarrow Ff \\
X \xleftarrow{\ \pi_X^B \circ \alpha_B\ } B & \xrightarrow{\ \pi_{FB}^B \circ \alpha_B\ } & FB
\end{array}
$$

where π_X^A and π_{FA}^A are the left and right projections of the product $X \times FA$, respectively, and π_X^B, π_{FB}^B the left and right projections of $X \times FB$.

Any $(X \times F)$-coalgebra (A, α_A) can be turned into an F-coalgebra $(A, \pi_{FA}^A \circ \alpha_A)$, hence, by the lemma above, any $(X \times F)$-homomorphism is an F-homomorphism with respect to this translation.

We define $R_F\colon \mathbf{C} \to F\text{-}\mathbf{coalg}$, for any object X in $\mathbf{C}$ and arrow $f\colon X \to Y$, as follows

$$
R_F X = (Z_X, \delta_X = \pi_{FZ_X}^X \circ \omega_X)
\qquad
\begin{array}{ccc}
X \xleftarrow{\ \pi_X^X \circ \omega_X\ } Z_X & \xrightarrow{\ \delta_X\ } & FZ_X \\
\ \downarrow f & R_F f = f^\flat \ \downarrow & \ \downarrow Ff \\
Y \xleftarrow{\ \pi_Y^Y \circ \omega_Y\ } Z_Y & \xrightarrow{\ \delta_Y\ } & FZ_Y
\end{array}
$$

where (Z_X, ω_X) is the final $(X \times F)$-coalgebra associated with X, and $f^\flat\colon Z_X \to Z_Y$ is the unique final $(X \times F)$-homomorphism making the above diagram commute (cf. Lemma 3.4.4). In particular, $f^\flat$ is an F-homomorphism between $R_F X = (Z_X, \delta_X)$ and $R_F Y = (Z_Y, \delta_Y)$, thus R_F is well defined. Functoriality is readily proved exploiting the universal property of final coalgebras.

The functor R_F right adjoint to U_F, with unity $\eta\colon Id \Rightarrow L_F U_F$ and counit $\epsilon\colon U_F R_F \Rightarrow Id$ defined component-wise as follows, for X in $\mathbf{C}$ and (X, α) in $F\text{-}\mathbf{coalg}$,

$$
\eta_{(X,\alpha)} = \langle id_X, \alpha \rangle^\flat \colon (X, \alpha) \to (Z_X, \delta_X) = L_F U_F(X, \alpha), \quad \epsilon_X = \pi_X^X \circ \omega_X \colon U_F R_F X = Z_X \to X \,,
$$

where $\langle id_X, \alpha \rangle^\flat$ is the unique arrow given by finality, making the following diagram commute (by Lemma 3.4.4):

$$
\begin{array}{ccc}
 & X & \xrightarrow{\ \alpha\ } FX \\
id_X \swarrow & \langle id_X, \alpha \rangle^\flat \downarrow & \ \downarrow F\langle id_X, \alpha \rangle^\flat \\
X \xleftarrow{\ \epsilon_X\ } Z_X & \xrightarrow{\ \delta_X\ } & FZ_X
\end{array}
$$

Theorem 3.4.5 *Let $F\colon \mathbf{C} \to \mathbf{C}$ be a functor in a category $\mathbf{C}$ with binary products. If for any object X in $\mathbf{C}$ the functor $X \times F$ has final algebra, then the forgetful functor $U_F\colon F\text{-}\mathbf{coalg} \to \mathbf{C}$ has a right adjoint $R_F\colon \mathbf{C} \to F\text{-}\mathbf{coalg}$.*

From adjunctions to final coalgebras

The existence of a right adjoint to the forgetful functor $U_F\colon F\text{-}\mathbf{coalg} \to \mathbf{C}$, namely, the functor

$$
R_F\colon \mathbf{C} \to F\text{-}\mathbf{coalg} \qquad\qquad X \mapsto (Z_X, \delta_X\colon Z_X \to FZ_X) \,,
$$

induces a principle of cofree construction, which can be (re)stated as

$$
\mathbf{C} \underset{U_F}{\overset{R_F}{\rightleftarrows}} F\text{-}\mathbf{coalg}
$$

(3.4.1)

where $\epsilon \colon U_F R_F \Rightarrow Id$ is the counit of the adjunction.

The coalgebra $R_F X = (Z_X, \delta_X)$ is usually called *cofree F-coalgebra over X*, and since right adjoints preserves all limits, the existence of final F-coalgebras is guaranteed if the underlying category $\mathbf{C}$ has an terminal object $\mathbf{1}$.

Theorem 3.4.6 (Existence of final F-coalgebras) *Let $F \colon \mathbf{C} \to \mathbf{C}$ be a functor on a category with terminal object $\mathbf{1}$. If the forgetful functor $U_F \colon F\text{-}\mathbf{coalg} \to \mathbf{C}$ has a right adjoint $R_F \colon \mathbf{C} \to F\text{-}\mathbf{coalg}$, then $R_F \mathbf{1}$, the cofree F-coalgebra over $\mathbf{1}$, is final.*

The adjunction $(\eta, \epsilon) \colon U_F \dashv R_F$ generates a (cofree) comonad over $\mathbf{C}$, defined as

$$
D_F = (U_F R_F, \epsilon, U_F \eta R_F)
$$

called the *comonad cofreely generated by F*. By a little abuse of notation the composite functor $U_F R_F \colon \mathbf{C} \to \mathbf{C}$ is tipically denoted by $D_F \colon \mathbf{C} \to \mathbf{C}$, the counit ϵ by ϵ_F, and comultiplication $U_F \eta R_F$ by ξ_F. If F is clear from the context the comonad will be denoted simply by (D, ϵ, ξ).

Definition 3.4.7 (Structural coinduction) *Let $F \colon \mathbf{C} \to \mathbf{C}$ be a functor and (D_F, ϵ_F, ξ_F) be the comonad cofreely generated by F. Then, for any object Y in $\mathbf{C}$, arrow $f \colon X \to Y$, and F-coalgebra (X, α), there exists a unique arrow $f^\flat \colon Y \to D_F X$ making the following diagram commute*

where $\delta_Y \colon D_F Y \to F D^F X$ is the cofree F-coalgebra structure over Y. The arrow $f^\flat$ is said the (cofree) coinductive extension of α along (the co-valuation) f.

Notice that, the diagram in the above statement is nothing but Diagram (3.4.1).

Lemma 3.4.8 *Let $F, G, H \colon \mathbf{C} \to \mathbf{C}$ be functors and (D_F, ϵ_F, ξ_F) be the comonad cofreely generated by F. Then, any two natural transformations $\phi \colon H \Rightarrow G$ and $\varphi \colon H \Rightarrow FH$ uniquely define a natural transformation $\rho \colon H \Rightarrow D_F G$, making the following diagram commute*

where $\delta \colon D_F \Rightarrow F D_F$ is the natural transformation induced by cofree F-coalgebras structures over the objects in $\mathbf{C}$. In this case, ρ is said the (cofree) coinductive extension of φ along ϕ.

Remark 3.4.9 Notably, the comultiplication $\xi^F \colon D_F \Rightarrow D_F D_F$ fits as the coinductive extension of $\delta \colon D_F \Rightarrow F D_F$ along the identity natural transformation $id_{D_F} \colon D_F \Rightarrow D_F$, i.e., it makes the following diagram commute

$$
\begin{array}{ccc}
& D_F & \xrightarrow{\;\delta\;} & F D_F \\
{\scriptstyle id_{D_F}} \swarrow & \Big\downarrow{\scriptstyle \xi} & & \Big\downarrow{\scriptstyle F\xi} \\
D_F \xleftarrow[\;\epsilon D_F\;]{} & D_F D_F & \xrightarrow[\;\delta D_F\;]{} & F D_F D_F
\end{array}
$$

This property is very useful in proofs, since it gives a simple universal characterization of the comultiplication if it belongs to a comonad cofreely generated by some functor. ∎

We conclude this section showing an example of comonad cofreely generated by functor.

Example 3.4.10 (Infinite Streams Comonad) Let A an alphabet of symbols, and X a set of state variables. For any set X, the functor $X \times (A \times Id) \colon \mathbf{Set} \to \mathbf{Set}$ has final coalgebras.

By Theorem 3.4.5, it follows that the forgetful functor $U_{(A \times Id)} \colon (A \times Id)\text{-}\mathbf{coalg} \to \mathbf{Set}$ has a right adjoint $R_{(A \times Id)} \colon \mathbf{Set} \to (A \times Id)\text{-}\mathbf{coalg}$ mapping objects X in $\mathbf{Set}$ to (Z_X, δ_X), the cofree $(A \times Id)$-coalgebra over X, and arrows $f \colon X \to Y$ to the coinductive extensions of δ_X over $f \circ \epsilon_X$, where ϵ is the counit of the adjunction.

The carrier Z_X, of the cofree $(A \times Id)$-coalgebra over X, is the set $(X \times A)^\omega$ of infinite streams over the enriched alphabet $X \times A$, and the coalgebra structure $\delta_X \colon (X \times A)^\omega \to A \times (X \times A)^\omega$ is given by $\delta_X = \langle hd_X, tl_X \rangle$, where $tl_X \colon (X \times A)^\omega \to (X \times A)^\omega$ is the tail function, and $hd_X \colon (X \times A)^\omega \to A$ is the half-head function (cf. Example 3.4.3). In this respect, we can define the comonad (D, ϵ, ξ) cofreely generated by $A \times Id$ explicitly as follows, for all $a \in A$, $x \in X$, and $s \in (X \times A)^\omega$,

$$
\begin{aligned}
& D \colon \mathbf{Set} \to \mathbf{Set} \\
& \quad DX = (X \times A)^\omega \\
& \quad D(f \colon X \to Y)((x, a)s) = (f(x), a)Df(s)
\end{aligned}
\qquad
\begin{aligned}
& \epsilon_X((x, a)s) = x \\
& \xi_X((x, a)s) = ((x, a)s, a)\xi_X(s) \, .
\end{aligned}
$$

Note that, both the action on morphisms of D and the comultiplication are given applying the structural coinduction proof principle of Definition 3.4.7 (cf. also Remark 3.4.9). ∎

3.4.2   Coalgebras for a Comonad

In Section 3.2.2, we showed that when an endofunctor $F \colon \mathbf{C} \to \mathbf{C}$ admits free algebra constructions, then we can define the monad (T^F, η^F, μ^F) freely generated by F. In particular the category of F-algebras and of Eilenberg-Moore algebras for (T^F, η^F, μ^F) are isomorphic. Dually, we can prove that the category of Eilenberg-Moore coalgebras for the comonad cofreely generated by F is isomorphic to F-$\mathbf{coalg}$.

Definition 3.4.11 (Eilenberg-Moore coalgebra) *Let (D, ϵ, ξ) be a comonad in $\mathbf{C}$. A coalgebra of the comonad (D, ϵ, ξ) is a D-coalgebra (X, α) such that the two diagrams below commute, which we call the* counit *and* comultiplication laws *of the coalgebra, respectively.*

$$
\begin{array}{ccc}
& X & & \qquad X \xrightarrow{\;\alpha\;} DX \\
{\scriptstyle id_X} \swarrow & \Big\downarrow{\scriptstyle \alpha} & & \quad {\scriptstyle \alpha}\Big\downarrow \qquad\qquad \Big\downarrow{\scriptstyle D\alpha} \\
X \xleftarrow[\;\epsilon_X\;]{} & DX & & \quad DX \xrightarrow[\;\xi_X\;]{} DDX
\end{array}
$$

The category of all coalgebras for the comonad (D, ϵ, ξ) and D-homomorphisms between them, is said the Eleinberg-Moore category for (D, ϵ, ξ), *and it is denoted by* (D, ϵ, ξ)-$\mathbf{coalg}$.

The counit and comultiplication laws say that the coalgebra structure α respects the structure of the comonad. By an abuse of notation we will denote (D, ϵ, ξ)-**coalg** simply as D-**coalg**, when it is clear from the contexts that D belongs to a monad.

Let (D_F, ϵ_F, ξ_F) be the comonad freely generated by $F\colon \mathbf{C} \to \mathbf{C}$. We define the natural transformation $\nu\colon D_F \Rightarrow F$ as the composite

$$\nu = F\epsilon \circ \delta \qquad D_F \xrightarrow{\ \delta\ } FD_F \xrightarrow{\ F\epsilon\ } F$$

The naturality of ν allows us to transform any D_F-algebra (X, α) to an F-algebra $(X, \alpha \circ \nu_X)$, and any D_F-homomorphism to an F-homomorphism, by naturality of ν, in the following way

$$
\begin{array}{ccc}
\begin{array}{ccc}
X & \xrightarrow{\ \alpha\ } & D_F X \\
{\scriptstyle f}\downarrow & & \downarrow{\scriptstyle D_F f} \\
Y & \xrightarrow{\ \beta\ } & D_F Y
\end{array}
& \mapsto &
\begin{array}{ccccc}
X & \xrightarrow{\ \alpha\ } & D_F X & \xrightarrow{\ \nu_X\ } & FX \\
{\scriptstyle f}\downarrow & & \downarrow{\scriptstyle D_F f} & & \downarrow{\scriptstyle Ff} \\
Y & \xrightarrow{\ \beta\ } & D_F Y & \xrightarrow{\ \nu_Y\ } & FY
\end{array}
\end{array}
$$

This transformation is functorial, i.e., it extends to a functor $K\colon D_F$-**coalg** $\to F$-**coalg**. Moreover we have that K takes part into the following isomorphism of categories:

Lemma 3.4.12 (F-coalg $\cong D_F$-coalg) *Let $F\colon \mathbf{C} \to \mathbf{C}$ be a functor, (D_F, ϵ_F, ξ_F) be the comonad cofreely generated by F. Then, the categories F-**coalg** and D_F-**coalg** are isomorphic.*

3.5 Cocongruences and Behavioral Equivalences

In Section 3.3, we have seen that seeking for the the largest bisimulation is problematic in general, and one has to require either that behavior functor preserves weak pullbacks, or that the base category has factorization systems with right invertible morphisms. (Propositions 3.3.9 and 3.3.7). Aiming at a general development of universal coalgebra without having recourse to such strong assumptions on the base category and on functors, it has been proposed a better behaved alternative to the notion of bisimulation: *cocongruence*. Cocongruences replace monic spans in the definition of bisimulation with the dual notion of epic cospans.

In this section, we recall the definitions of cocongruence and behavioral equivalence, and we give evidence of how these notions can actually take the place of bisimulation. As a consequence we show why the property of a behavior functor to preserve weak pullbacks is convenient but not necessary. All the results and examples recalled in this section are due to Kurz [62].

Definition 3.5.1 (F-Cocongruence) *Let $F\colon \mathbf{C} \to \mathbf{C}$ be a functor. An epic cospan (K, f, g) in $\mathbf{C}$ between X and Y is an F-cocongruence between F-coalgebras (X, α) and (Y, β) if there exists a (unique) coalgebra structure $\kappa\colon K \to FK$ on K making the following diagram in $\mathbf{C}$ commute*

$$
\begin{array}{ccccc}
X & \xrightarrow{\ f\ } & K & \xleftarrow{\ g\ } & Y \\
{\scriptstyle \alpha}\downarrow & & \vdots\,{\scriptstyle \kappa} & & \downarrow{\scriptstyle \beta} \\
FX & \xrightarrow{\ Ff\ } & FK & \xleftarrow{\ Fg\ } & FY
\end{array}
$$

that is, making f and g morphisms of F-coalgebras.

To explain this definition assume the base category is **Set**. Given an F-cocongruence (K, f, g) on (X, α) and (Y, β) we can define the relation $\sim_K \subseteq X \times Y$ as $x \sim_K y$, if and only if, $f(x) = g(y)$. Actually, the set $\sim_K$ together with the canonical projections is the pullback of (f, g). Note that, (K, κ) is the coalgebra which results from identifying the states related by $\sim_K$. These arguments

are valid in any category with pullbacks, but note that the definition of cocongruence is more general since it does not assume the existence of pullbacks in the category.

However, even without the use of pullbacks, we can still have a relational notion of cocongruence via monic spans respecting the structure of the cocongruence.

Definition 3.5.2 (Span associated with F-cocongruence) *Let $F\colon \mathbf{C} \to \mathbf{C}$ be a functor, and (K, f, g) be an F-cocongruence between F-coalgebras (X, α) and (Y, β). A monic cospan (R, s, t) between X and Y is associated with (K, f, g) if the following diagram commutes in $\mathbf{C}$.*

$$
\begin{array}{ccccc}
 & & R & & \\
 & {}^{s}\swarrow & & \searrow{}^{t} & \\
X & \xrightarrow{\ f\ } & K & \xleftarrow{\ g\ } & Y \\
\alpha\downarrow & & \downarrow\kappa & & \downarrow\beta \\
FX & \xrightarrow[\ Ff\ ]{} & FK & \xleftarrow[\ Fg\ ]{} & FY
\end{array}
$$

Note that, for the object R it is not required any coalgebra structure, hence it is merely a relation between the carriers X and Y and not between their coalgebras. Any span derived from a pullback is monic, so that, in case the base category has pullbacks there is a universal (actually, final) monic span associated to the F-cocongruence. In [78], this notion is referred to as *kernel bisimulation*, in order to emphasize its universal property. It should be mentioned that many authors prefer this definition to the more general of cocongruence because it can be compared with the notion of bisimulation, so that one can say that a bisimulation is a "cocongruence" and viceversa.

Definition 3.5.3 (Behavioral Equivalence) *Let $F\colon \mathbf{C} \to \mathbf{C}$ be a functor. An arrow $e\colon X \to E$ between F-coalgebras (X, α) and (E, ϵ) is an F-behavioral equivalence if it is an epimorphism.*

Note that, behavioral equivalences $e\colon X \to E$ are exactly cocongruences of the form (E, e, e). The name "equivalence" is chosen according to the fact that in **Set** the induced relation $\sim_E$ is always and equivalence. This argument works in general in any category with kernel pairs, which can be considered as the categorical generalization of equivalence relations.

Historical note. The name "cocongruence" is due to Kurz, and it was chosen due to the fact that cocongruences are dual to congruences for algebras. Behavioral equivalences are essentially Aczel and Mendler's congruences but the term "behavioral equivalence" gives a better intuition of its meaning and it does not conflict with other uses of the term "congruence".

To the best of our knowledge, Kurz was the fist who recognized the use of cocongruences as an alternative notion to bisimulations. Later, cocongruences have been adopted by various authors and now they are recognized as *the* alternative to bisimulations, at least when largest bisimulations do not exist. Bartels et al. [16] used cocongruences in order to find reflections of bisimilarities in a hierarchy of coalgebras for discrete state probabilistic systems. Their use, however, only eased the proofs and was not really necessary. Later, Danos et al. [31] proposed cocongruences (called event bisimulations or probabilistic cocongruences) as *the right* alternative to bisimulations in the study of labelled Markov processes, that is, probabilistic systems with continuous state spaces. The use of congruences allowed them to give a logical characterization of probabilistic "bisimilarity" (actually, behavioral equivalence) without having recourse to specific properties of polish or analytic spaces. In previous works, the logical characterization was proved only in the case of analytic spaces which allow for the construction of semi-pullbacks [45], used to prove the existence of the largest bisimulation.

Proposition 3.5.4 ([62]) *Let $\mathbf{C}$ be a category with pullbacks and $F\colon \mathbf{C} \to \mathbf{C}$ be a functor preserving weak pullbacks. Then, F-cocongruences give rise to F-bisimulations via pullbacks.*

Proof. Let (X, α) and (Y, β) be F-coalgebras and (K, f, g) be an F-cocongruence between them which coalgebra structure $\kappa \colon K \to FK$. Consider the pullback for the pair (f, g) in $\mathbf{C}$, namely, (R, r_1, r_2), which, by the universal property of pullbacks, is a monic span. To prove it is an F-bisimulation we have to provide an F-coalgebra structure $\gamma \colon R \to FR$. To do so, consider the following diagram

$$
\begin{array}{ccc}
FR & \xrightarrow{\ Fr_2\ } & FY \\
\end{array}
$$

for which the outer square is a weak pullback, since F preserves weak pullbacks, thus γ exists and makes (R, p_1, p_2) be an F-bisimulation between (X, α) and (Y, β). ∎

Hence, the (final) monic span associated with the cocongruence derived via pullback, is also a bisimulation, provided that the behavior functor weakly preserves pullbacks. This suggests the notion of *bisimulation* associated with a cocongruence.

Definition 3.5.5 (Associated Bisimulation) *Let $F \colon \mathbf{C} \to \mathbf{C}$ be a functor and (K, f, g) be an F-cocongruence between (X, α) and (Y, β). An F-bisimulation (R, s, t) between (X, α) and (Y, β) is* associated *with (K, f, g) if its cospan is so.*

Proposition 3.5.4 also holds for behavioral equivalences (indeed, the proof works in the same way imposing $f = g$), so that, we will also talk of bisimulations associated with behavioral equivalences.

To explain the difference between bisimulations and behavioral equivalences, in the following example we consider a functor which does not preserve weak pullback, the Aczel-Mendler functor, called in this way since it first appeared in [4].

Example 3.5.6 Consider the Aczel-Mendler functor $AM \colon \mathbf{Set} \to \mathbf{Set}$ defined on objects X and arrows $f \colon X \to Y$ as follows, for $x, y, z \in X$

$$
AM(X) = \{(x, y, z) \in X^3 \mid |\{x, y, z\}| \le 2\}
$$
$$
AM(f)((x, y, z)) = (f(x), f(y), f(z))
$$

Coalgebras for this functor can be seen as a kind of deterministic automata taking three types of inputs, namely, 1, 2, and 3, and performing a deterministic transitions to a successor state according to the type of input it has been received. As an example, consider the coalgebra (X, α) defined as below

$$
X = \{x_1, x_2\}
$$
$$
\alpha(x_1) = (x_1, x_1, x_2)
$$
$$
\alpha(x_2) = (x_1, x_2, x_2)
$$

The restriction on the cardinality imposed by the behavior functor can be thought as a constraint in the implementation of these kind of automata, i.e., in every state at least two inputs have to give rise to the same successor.

For an example of behavioral equivalence that is not a bisimulation consider the (final) coalgebra (Z, ω) given by $Z = \{*\}$ (the final object in $\mathbf{Set}$) and $\omega(*) = (*, *, *)$. The unique map to the singleton set, $! \colon X \to Z$, is a surjection and, hence, an epimorphism in $\mathbf{Set}$. This means that $! \colon X \to Z$ is a behavioral equivalence on (X, α), and, since (Z, ω) is final, it is the largest one.

Since $!(x_1) = * = !(x_2)$, x_1 and x_2 are behavioral equivalent, but their are not bisimilar. Indeed, if it would exist an AM-bisimulation relating x_1 and x_2 it will have a coalgebra structure that on (x_1, x_2) will be of the following form

$$
\begin{array}{ccc}
 & (x_1, x_2) & \\
{}^{1}\swarrow & {}^{2}\downarrow & {}^{3}\searrow \\
(x_1, x_1) & (x_1, x_2) & (x_2, x_2)
\end{array}
$$
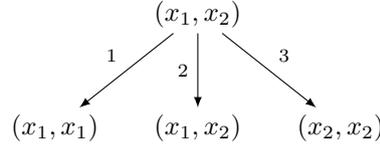
which is not possible due to the cardinality constraint imposed by AM.

This shows that bisimulations may fail to capture behavioral equivalence. This phenomenon is due to the fact that the functor AM imposes constraints which cannot be satisfied by a largest bisimulation. On the other hand, from a behavioral point of view, this constraint is not observable, which is represented by the fact that the largest behavioral equivalence does exist. ∎

The Aczel-Mendler functor can also be used to show that bisimulation may fail to give rise to a smallest bisimulation containing it.

Example 3.5.7 Let $AM\colon \mathbf{Set} \to \mathbf{Set}$ be the Aczel-Mendler functor defined as in Example 3.5.6. Consider the AM-coalgebra (X, α) defined as follows

$$
\begin{array}{lll}
{}^{1/3}\circlearrowright & {}^{1/2/3}\circlearrowright & {}^{1}\circlearrowright \\
x_1 \xrightarrow{\ 2\ } x_2 \xleftarrow{\ 2/3\ } x_3 & &
\end{array}
\qquad
\begin{array}{l}
X = \{x_1, x_2, x_3\} \\
\alpha(x_1) = (x_1, x_2, x_1) \\
\alpha(x_2) = (x_2, x_2, x_2) \\
\alpha(x_3) = (x_3, x_2, x_2)
\end{array}
$$

Note that the relations $R = \{(x_1, x_2)\}$ and $S = \{(x_2, x_3)\}$ are both AM-bisimulations, respectively, with coalgebra structures $\rho\colon R \to AM(R)$ and $\sigma\colon S \to AM(S)$ defined as follows

$$
\rho((x_1, x_2)) = ((x_1, x_2), (x_1, x_2), (x_1, x_2)), \qquad \sigma((x_2, x_3)) = ((x_2, x_3), (x_2, x_3), (x_2, x_3)).
$$

And their union is a bisimulation too, however there is no bisimulation equivalence containing. Indeed, if there were such an equivalence, it should contain the pair (x_1, x_3) at which the coalgebra structure would have been of the following form

$$
\begin{array}{ccc}
 & (x_1, x_3) & \\
{}^{1}\swarrow & {}^{2}\downarrow & {}^{3}\searrow \\
(x_1, x_3) & (x_2, x_2) & (x_1, x_2)
\end{array}
$$
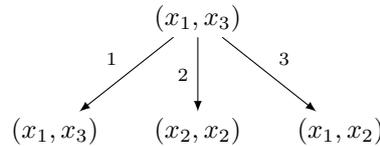
hence, violating the cardinality constraint imposed by AM. ∎

Example 3.5.7 shows also that bisimulations are not closed under composition.

If one considers cocongruences in place of bisimulations the theory works smoothly. Indeed, the existence of largest behavioral equivalences is guaranteed in case the category of coalgebras has final objects, or more generally when the base category has cointersections. Moreover, if the underlying category has pushouts, then cocongruences are also closed by composition. These arguments, together with what we have seen in Examples 3.5.7 and 3.5.7, give evidence of why cocongruences should be preferred instead of bisimulations in case the behavior functor does not preserve weak pullbacks.

4

Initial Algebras and Final Coalgebras via Factorization Systems

Remarks and notation: We assume the reader have elementary knowledge about ordinal numbers and transfinite induction. Ordinal numbers will be ranged over by $\alpha, \beta, \gamma, \kappa, \dots$ and the class of ordinal numbers will be denoted by **Ord**. By a little abuse of notation, **Ord** will also denote the category of ordinal numbers with arrows $\alpha \to \beta$ iff $\alpha \le \beta$, and by α we will also denote the full sub-category of **Ord** of all ordinal numbers less or equal than α. For a functor $F \colon \mathbf{Ord} \to \mathbf{C}$, we define the functor $F{\restriction}\alpha \colon \alpha \to \mathbf{C}$ as the composite $\iota \circ F$, where $\iota \colon \alpha \hookrightarrow \mathbf{Ord}$ is the inclusion functor.

4.1   Initial and Final Sequences

In this section we recall the definition of *initial* and *final sequences* for an endofunctor. These structures, which are dual to each other, were first explicitly given by Barr [13] in order to investigate the relationship between the initial and final coalgebra, and in order to provide sufficient conditions for a functor to be algebraically compact (i.e., when the unique arrow from the initial algebra to the final coalgebra is an isomorphism). These sequences have been successfully used in order to infer properties about the initial algebra and final coalgebra and, moreover, to provide sufficient conditions for a functor to admit such initial and terminal objects (see for example, Barr [14], Adàmek [9, 8, 7, 6], Smyth and Plotkin [77], and Worrell [94, 92]).

4.1.1   Initial Sequences Leads to Initial Algebras

In this section we recall the definition and the main results about initial sequences. The exposition is slightly nonstandard and puts the light on some results which were implicit in previous presentations (e.g., in [13, 9]).

Definition 4.1.1 (Initial Sequence) *Let* **C** *be a category with initial object* **0** *and colimits of ordinal-indexed diagrams, and assume* $T \colon \mathbf{C} \to \mathbf{C}$ *be an endofunctor on* **C***. The* initial sequence *of* T *is a limit-preserving functor* $A \colon \mathbf{Ord} \to \mathbf{C}$ *such that, for all ordinals* $\gamma \le \beta$,

   i. $A(0) = \mathbf{0}$;
  ii. $A(\beta{+}1) = TA(\beta)$;
 iii. $A(\gamma{+}1 \to \beta{+}1) = TA(\gamma \to \beta)$.

The initial sequence is said to stabilize *at some* $\alpha \in \mathbf{Ord}$, *if* $A(\alpha \to \alpha{+}1)$ *is an isomorphism.*

Note that, for all limit ordinals β, $(\gamma \to \beta)_{\gamma < \beta}$ is a colimit in **Ord**, and since A preserves colimits, $(A(\gamma \to \beta))_{\gamma < \beta}$ is a colimit in **C** for the diagram $A{\restriction}\alpha$.

Historical note. In [13], Barr gave a more direct construction of the initial sequence as an ordinal-indexed sequence of objects $(A_\beta)_{\beta \in \mathbf{Ord}}$ with arrows $(f_\beta^\gamma \colon A_\gamma \to A_\beta)_{\gamma \leq \beta}$, uniquely defined by the following conditions, for $\delta \leq \gamma \leq \beta$:

(IS-1) $A_{\beta+1} = TA_\beta$;

(IS-2) $f_{\beta+1}^{\gamma+1} = Tf_\beta^\gamma$;

(IS-3) $f_\beta^\beta = id_{A_\beta}$:

(IS-4) $f_\beta^\gamma \circ f_\gamma^\delta = f_\beta^\delta$;

(IS-5) if β is a limit ordinal, the cocone $(f_\beta^\gamma \colon A_\gamma \to A_\beta)_{\gamma < \beta}$ is a colimit.

The sequence is defined by transfinite induction on $\alpha \in \mathbf{Ord}$, defining A_α and $f_\alpha^\beta \colon A_\beta \to A_\alpha$, for all $\beta \leq \alpha$, and checking at each stage that conditions (IS-1)–(IS-5) hold for the portion of sequence already defined.

First step: Let $\alpha = 0$. The sequence begins with $A_0 = \mathbf{0}$ and $f_0^0 = id_{A_0}$.

Isolated step: Let $\alpha = \alpha' + 1$ and assume by inductive hypothesis that $A_{\alpha'}$ and the arrows $f_{\alpha'}^\gamma$ have been given and satisfy (IS-1)–(IS-5), for all $\gamma < \alpha'$. We define $A_\alpha = TA_{\alpha'}$, and the arrows f_α^β are defined by induction on $\beta \leq \alpha$. We distinguish three cases. If $\beta = \alpha$, then we define $f_\alpha^\beta = id_{A_\alpha}$. If β is a successor ordinal, say $\beta = \beta' + 1$, then we define $f_\alpha^\beta = Tf_{\alpha'}^{\beta'}$. If β is a limit ordinal, $(f_\beta^\gamma \colon A_\gamma \to A_\beta)_{\gamma < \beta}$ is a colimit, by (IS-5). By inductive hypothesis on β, we can consider $(f_\alpha^\gamma \colon A_\gamma \to A_\alpha)_{\gamma < \beta}$, which turns out to be a compatible cocone by (IS-2) and (IS-4). Now, we define f_α^β as the unique map factorizing the cocone.

Limit step: Let α be a limit ordinal. By inductive hypothesis we are given all arrows f_β^γ, for $\gamma \leq \beta < \alpha$, which, indeed, form a chain. We define A_α to be the colimit of this chain and $(f_\alpha^\beta \colon A_\beta \to A_\alpha)_{\beta < \alpha}$ are its injections.

An initial sequence constitutes an ordinal-index diagram $A \colon \mathbf{Ord} \to \mathbf{C}$ in $\mathbf{C}$, and it turns out that for any T-algebra (X, h) one can define a cocone over it.

Lemma 4.1.2 ([13]) *Assume $A \colon \mathbf{Ord} \to \mathbf{C}$ be the initial sequence of T, and (X, h) be a T-algebra. Then, there is a cocone $(h_\alpha \colon A(\alpha) \to X)_{\alpha \in \mathbf{Ord}}$ from A to X, such that, for any ordinal α, the following diagram commutes:*

$$
\begin{array}{ccc}
TX & \xrightarrow{\ h\ } & X \\[1mm]
{\scriptstyle Th_\alpha}\big\uparrow & & \big\uparrow{\scriptstyle h_\alpha} \\[1mm]
A(\alpha{+}1) & \longleftarrow & A(\alpha)
\end{array}
$$

Proof. We define $h_\alpha \colon A(\alpha) \to X$ by transfinite induction on $\alpha \in \mathbf{Ord}$, checking at each step that $h_\alpha = h \circ Th_\alpha \circ A(\alpha \to \alpha{+}1)$ and $h_\beta = h_\alpha \circ A(\beta \to \alpha)$, for all $\beta \leq \alpha$.

We define $h_0 \colon A(0) = \mathbf{0} \to X$ as the unique arrow from the initial object, therefore, by uniqueness, $h_0 = h \circ Th_0 \circ A(0 \to 1)$ holds. Assume by inductive hypothesis that, for all $\beta \leq \alpha$, the arrows h_β are given, and they are such that $h_\beta = h \circ Th_\beta \circ A(\beta \to \beta{+}1)$ and $h_\beta = h_\alpha \circ A(\beta \to \alpha)$. define $h_{\alpha+1} \triangleq h \circ Th_\alpha$. From this it follows that

$$
\begin{aligned}
h_{\alpha+1} &= h \circ Th_\alpha && \text{(by def. } h_{\alpha+1}) \\
&= h \circ T(h \circ Th_\alpha \circ A(\alpha \to \alpha{+}1)) && \text{(by inductive hp.)} \\
&= h \circ T(h \circ Th_\alpha) \circ TA(\alpha \to \alpha{+}1) && \text{(by funct. } T)
\end{aligned}
$$

$$\begin{aligned}
&= h \circ Th_{\alpha+1} \circ TA(\alpha \to \alpha+1) && \text{(by def. } h_{\alpha+1}) \\
&= h \circ Th_{\alpha+1} \circ A(\alpha+1 \to \alpha+2)\,; && \text{(by def. } A)
\end{aligned}$$

$h_{\alpha+1} = h_{\alpha+1} \circ A(\alpha+1 \to \alpha+1)$, since $A(\alpha+1 \to \alpha+1) = id_{A(\alpha+1)}$, and, for all $\beta \leq \alpha$,

$$\begin{aligned}
h_{\beta} &= h_{\alpha} \circ A(\beta \to \alpha) && \text{(by inductive hp.)} \\
&= h \circ Th_{\alpha} \circ A(\alpha \to \alpha+1) \circ A(\beta \to \alpha) && \text{(by inductive hp.)} \\
&= h \circ Th_{\alpha} \circ A(\beta \to \alpha+1) && \text{(by funct. } A) \\
&= h_{\alpha+1} \circ A(\beta \to \alpha+1)\,. && \text{(by def. } h_{\alpha+1})
\end{aligned}$$

Let α be a limit ordinal, and assume by inductive hypothesis that, for all $\beta < \alpha$, the arrows h_{β} are given, and they are such that $h_{\beta} = h \circ Th_{\beta} \circ A(\beta \to \beta+1)$ and $h_{\beta} = h_{\alpha} \circ A(\beta \to \alpha)$. By definition of initial sequence, $(A(\beta \to \alpha))_{\beta < \alpha}$ is a colimit over $A{\upharpoonright}\alpha$, and by (the second part of the) inductive hypothesis we have that $(h_{\beta} \colon A(\beta) \to X)_{\beta < \alpha}$ is a cocone over $A{\upharpoonright}\alpha$. From this, we define h_{α} as the unique arrow such that $h_{\beta} = h_{\alpha} \circ A(\beta \to \alpha)$, for all $\beta < \alpha$. Trivially, $h_{\alpha} = h_{\alpha} \circ A(\alpha \to \alpha)$, since $A(\alpha \to \alpha) = id_{A(\alpha)}$. We have that, for all $\beta < \alpha$,

$$\begin{aligned}
h_{\beta} &= h \circ Th_{\beta} \circ A(\beta \to \beta+1) && \text{(by inductive hp.)} \\
&= h \circ T(h_{\alpha} \circ A(\beta \to \alpha)) \circ A(\beta \to \beta+1) && \text{(by inductive hp.)} \\
&= h \circ Th_{\alpha} \circ TA(\beta \to \alpha) \circ A(\beta \to \beta+1) && \text{(by funct. } T) \\
&= h \circ Th_{\alpha} \circ A(\beta+1 \to \alpha+1) \circ A(\beta \to \beta+1) && \text{(by def. } A) \\
&= h \circ Th_{\alpha} \circ A(\beta \to \alpha+1) && \text{(by funct. } A) \\
&= \big(h \circ Th_{\alpha} \circ A(\alpha \to \alpha+1)\big) \circ A(\beta \to \alpha)\,. && \text{(by funct. } A)
\end{aligned}$$

By uniqueness of the colimiting arrow, the above proves that $h_{\alpha} = h \circ Th_{\alpha} \circ A(\alpha \to \alpha+1)$. $\blacksquare$

Definition 4.1.3 (Initial co-projection) *Let $T \colon \mathbf{C} \to \mathbf{C}$ be a functor, (X, h) be a T-algebra, and $A \colon \mathbf{Ord} \to \mathbf{C}$ be the initial sequence of T. An arrow $k \colon A(\alpha) \to X$ is an* initial co-projection *at α for (X, h), if the following diagram commutes:*

$$
\begin{array}{ccc}
TX & \xrightarrow{\ \ h\ \ } & X \\
{\scriptstyle Tk}\Big\uparrow & & \Big\uparrow{\scriptstyle k} \\
A(\alpha+1) & \xleftarrow{\hspace{1cm}} & A(\alpha)
\end{array}
$$

Given a T-algebra (X, h), Lemma 4.1.2 states that each arrow in the (canonical) cocone $(h_{\alpha} \colon A(\alpha) \to X)_{\alpha \in \mathbf{Ord}}$ for (X, h) over the initial sequence A, is an initial co-projection. It turns out that, for any ordinal α and T-algebra (X, h), initial co-projections at α for (X, h) are unique.

Lemma 4.1.4 (Uniqueness of initial co-projections) *Let $T \colon \mathbf{C} \to \mathbf{C}$ be a functor, (X, h) be a T-algebra, $A \colon \mathbf{Ord} \to \mathbf{C}$ be the initial sequence of T, and $(h_{\alpha} \colon A(\alpha) \to X)_{\alpha \in \mathbf{Ord}}$ be the cocone for (X, h) over A given by Lemma 4.1.9. If $k \colon A(\alpha) \to X$ is an initial co-projection at α for (X, h), then $k = h_{\alpha}$.*

Proof. For $\beta \leq \alpha$, let $k_{\beta} = k \circ A(\beta \to \alpha)$. We will show by transfinite induction on β, that $k_{\beta} = h_{\beta}$, for all $\beta \leq \alpha$. Certainly $k_0 = h_0$ since their domain is the initial object. Assume by inductive hypothesis that $k_{\beta} = h_{\beta}$, then we have

$$\begin{aligned}
k_{\beta+1} &= k \circ A(\beta+1 \to \alpha) && \text{(by def. } k_{\beta+1}) \\
&= h \circ Tk \circ A(\alpha \to \alpha+1) \circ A(\beta+1 \to \alpha) && \text{(by hp.)} \\
&= h \circ Tk \circ A(\beta+1 \to \alpha+1) && \text{(by funct. } A) \\
&= h \circ Tk \circ TA(\beta \to \alpha) && \text{(by def. } A)
\end{aligned}$$

$$
\begin{aligned}
&= h \circ T(k \circ A(\beta \to \alpha)) && \text{(by funct. } T) \\
&= h \circ T(k_\beta) && \text{(by def. } k_\beta) \\
&= h \circ T(h_\beta) && \text{(by inductive hp.)} \\
&= h_{\beta+1}. && \text{(by def. } h_{\beta+1})
\end{aligned}
$$

Assume β is a limit ordinal and, by inductive hypothesis, that $k_\gamma = h_\gamma$, for every $\gamma < \beta$. By definition of k_γ and by the fact that $(h_\gamma \colon A(\gamma) \to X)_{\gamma \leq \alpha}$ is a cocone over $A{\restriction}\alpha$, we have that $k_\gamma = k \circ A(\gamma \to \alpha)$ and $h_\gamma = h_\alpha \circ A(\gamma \to \alpha)$. Therefore, by inductive hypothesis, we obtain $k \circ A(\gamma \to \alpha) = h_\alpha \circ A(\gamma \to \alpha)$. By functoriality of A, this implies that

$$
k \circ A(\beta \to \alpha) \circ A(\gamma \to \beta) = h_\alpha \circ A(\beta \to \alpha) \circ A(\gamma \to \beta).
$$

This holds for all $\gamma < \beta$, and since $(A(\gamma \to \beta))_{\gamma < \beta}$ is a colimit over $A{\restriction}\beta$, by uniqueness of the colimiting arrow, we have $k \circ A(\beta \to \alpha) = h_\alpha \circ A(\beta \to \alpha)$. Thus, by definition of k_β and compatibility of the cocone, we conclude that $k_\beta = h_\beta$. ∎

Remark 4.1.5 Uniqueness of initial co-projections was not explicitly recognized in [13]. Indeed, in [13, Theorem 1.2], uniqueness for the arrows in initial co-projections was only proved assuming their targets are isomorphisms in the initial sequence. Lemma 4.1.4 extends this result and drop the assumption of existence of isomorphisms in the initial sequence. Thank to this lemma, the proof of [13, Theorem 1.2] can be made easier (a restatement of it is given in Theorem 4.1.6). ∎

Initial sequences gives sufficient conditions for the existence of initial algebras. Indeed, if the initial sequence of T stabilizes at some ordinal α, then there exists an initial T-algebra.

Theorem 4.1.6 ([13]) *Let α be an ordinal number, and suppose the initial sequence A of T stabilizes at α, then $(A(\alpha), A(\alpha \to \alpha{+}1)^{-1})$ is an initial T-algebra.*

Proof. Let (X, h) be a T-algebra. By Lemma 4.1.2, there exists a cocone $(h_\beta \colon A(\beta) \to X)_{\beta \in \mathbf{Ord}}$ over A. Moreover, $h_\alpha = h \circ Th_\alpha \circ A(\alpha \to \alpha{+}1)$. Since, by hypothesis, $A(\alpha \to \alpha{+}1)$ is an isomorphism, h_α is a homomorphism of T-algebras from $(A(\alpha), A(\alpha \to \alpha{+}1)^{-1})$ to (X, h). Therefore, $(A(\alpha), A(\alpha \to \alpha{+}1)^{-1})$ is weakly terminal. Uniqueness follows by Lemma 4.1.4. ∎

Remark 4.1.7 One may be tempted to think that if there exists an initial algebra for some endofunctor T, then the initial sequence must lead to it. This is not true in general, even for categories which are complete and cocomplete, as noted by Barr [13].

For example, let $\mathbf{C}$ be the category whose objects are all ordinals, ordered by inclusion, plus one more object $\top$ greater than all the ordinals. Then $\mathbf{C}$ is complete and cocomplete. Let $T \colon \mathbf{C} \to \mathbf{C}$ be the endofunctor defined by $T(\alpha) = \alpha{+}1$, when α is an ordinal, and $T(\top) = \top$. Then, the initial sequence of T consists of all the ordinals and never stabilizes. There is only one T-algebra, namely $(\top, id_\top)$, and it is initial. ∎

4.1.2 Final Sequences leads to Final Coalgebras

In this section, we dualize the definitions and results given in Section 4.1.1 for initial sequences, yielding the notion of final sequence, and providing sufficient conditions for the existence of final coalgebras. Note that, the proofs of all the results are omitted since they are straightforward dualizations of those given in the previous section.

Definition 4.1.8 (Final Sequence) *Let $\mathbf{C}$ be a category with final object $\mathbf{1}$ and limits of ordinal-indexed diagrams, and assume $T \colon \mathbf{C} \to \mathbf{C}$ be an endofunctor on $\mathbf{C}$. The* final sequence *of T is a limit-preserving functor $Z \colon \mathbf{Ord}^{op} \to \mathbf{C}$ such that, for all ordinals $\gamma \leq \beta$,*

i. $Z(0) = \mathbf{1}$;
ii. $Z(\beta{+}1) = TZ(\beta)$;
iii. $Z(\beta{+}1 \to \gamma{+}1) = TZ(\beta \to \gamma)$.

The final sequence is said to stabilize *at some $\alpha \in \mathbf{Ord}$, if $Z(\alpha{+}1 \to \alpha)$ is an isomorphism.*

A final sequence constitutes an ordinal-index diagram $Z \colon \mathbf{Ord}^{op} \to \mathbf{C}$ in $\mathbf{C}$, and it turns out that for any T-coalgebra (X, h) one can define a cone over it.

Lemma 4.1.9 *Assume $Z \colon \mathbf{Ord}^{op} \to \mathbf{C}$ be the final sequence of T, and (X, h) be a T-coalgebra. Then, there is a cone $(h^\alpha \colon X \to Z(\alpha))_{\alpha \in \mathbf{Ord}}$ over Z, such that, for any ordinal α, the following diagram commutes:*

$$
\begin{array}{ccc}
X & \xrightarrow{\ h\ } & TX \\
{\scriptstyle h^\alpha}\downarrow & & \downarrow{\scriptstyle Th^\alpha} \\
Z(\alpha) & \longleftarrow & Z(\alpha{+}1)
\end{array}
$$

The dual notion of initial co-projection is *final projection*.

Definition 4.1.10 (Final projection) *Let $T \colon \mathbf{C} \to \mathbf{C}$ be a functor, (X, h) be a T-coalgebra, and $Z \colon \mathbf{Ord}^{op} \to \mathbf{C}$ be the final sequence of T. An arrow $k \colon X \to Z(\alpha)$ is a* final projection at α *for (X, h), if the following diagram commutes:*

$$
\begin{array}{ccc}
X & \xrightarrow{\ h\ } & TX \\
{\scriptstyle k}\downarrow & & \downarrow{\scriptstyle Tk} \\
Z(\alpha) & \longleftarrow & Z(\alpha{+}1)
\end{array}
$$

For a given T-coalgebra (X, h), Lemma 4.1.9 ensures that each arrow in the (canonical) cone $(h^\alpha \colon X \to Z(\alpha))_{\alpha \in \mathbf{Ord}}$ for (X, h) over the final sequence Z is a final projection. Dualizing Lemma 4.1.2, we have that final projections are unique.

Lemma 4.1.11 (Uniqueness of final projections) *Let $T \colon \mathbf{C} \to \mathbf{C}$ be a functor, (X, h) be a T-coalgebra, $Z \colon \mathbf{Ord}^{op} \to \mathbf{C}$ be the final sequence of T, and $(h^\alpha \colon X \to Z(\alpha))_{\alpha \in \mathbf{Ord}}$ be the cone for (X, h) over Z given by Lemma 4.1.9. If $k \colon X \to Z(\alpha)$ is a final projection at α for (X, h), then $k = h^\alpha$.*

Theorem 4.1.12 ([13]) *Let α be an ordinal number, and suppose the final sequence Z of T stabilizes at α, then $(Z(\alpha), Z(\alpha{+}1 \to \alpha)^{-1})$ is a final T-coalgebra.*

By Theorem 4.1.12, for an endofunctor $T \colon \mathbf{C} \to \mathbf{C}$, a final T-coalgebra may be obtained looking for an isomorphism in the final sequence Z of T. Consequently, this result motivates the search for conditions for the final sequence to stabilize.

Probably one of the most known requirements for stabilization is that the endofunctor T is κ-continuous for some limit ordinal κ, that is, it preserves limits for any diagram $D \colon \kappa^{op} \to \mathbf{C}$.

Lemma 4.1.13 (κ-continuity [8]) *Let κ be a limit ordinal, T be a κ-continuous endofunctor in $\mathbf{C}$, and $Z \colon \mathbf{Ord}^{op} \to \mathbf{C}$ be the final sequence of T. Then, Z stabilizes at κ.*

Proof. We have to prove that $Z(\kappa{+}1 \to \kappa)$ is an isomorphism. By definition of final sequence, $(Z(\kappa \to \gamma))_{\gamma < \kappa}$ is a limit cone over $Z{\upharpoonright}\kappa$. Since T is κ-continuous, we have also that $(TZ(\kappa \to \gamma))_{\gamma < \kappa}$ is a limit cone over $TZ{\upharpoonright}\kappa$. By definition for Z, we have $TZ(\kappa \to \gamma) = Z(\kappa{+}1 \to \gamma{+}1)$, for any $\gamma \leq \kappa$, hence $(Z(\kappa{+}1 \to \gamma{+}1))_{\gamma < \kappa}$ is a limit cone over $TZ{\upharpoonright}\kappa$. However, by the fact that $(Z(\kappa \to \gamma))_{\gamma < \kappa}$ is a limit cone over $Z{\upharpoonright}\kappa$, also $(Z(\kappa \to \gamma{+}1))_{\gamma < \kappa}$ is a limit cone over $TZ{\upharpoonright}\kappa$. Since limits are unique up to isomorphism, any arrow $k \colon Z(\kappa{+}1) \to Z(\kappa)$ such that $Z(\kappa{+}1 \to \gamma{+}1) = Z(\kappa \to \gamma{+}1) \circ k$, for all $\gamma < \kappa$, must be an isomorphism. By functoriality of Z, we have that, for all $\gamma < \kappa$, $Z(\kappa{+}1 \to \gamma{+}1) = Z(\kappa \to \gamma{+}1) \circ Z(\kappa{+}1 \to \kappa)$, therefore $Z(\kappa{+}1 \to \kappa)$ is an isomorphism. ∎

Note that, the proof works even if the functor T just preserves limits over $Z{\upharpoonright}\kappa$, so the lemma can be weakened accordingly.

In many situations, proving κ-continuity for a functor is not an easy task, and often the functors one wants to deal with do not satisfy such a property. However, under mild assumptions on the underlying category $\mathbf{C}$ and on the functor, the existence of the final coalgebra can be proved looking for monic arrows along the final sequence instead of isomorphisms.

Lemma 4.1.14 (Preserve monics) *Let* $T \colon \mathbf{C} \to \mathbf{C}$ *be an endofunctor in a well-powered category* $\mathbf{C}$, *and* $Z \colon \mathbf{Ord}^{op} \to \mathbf{C}$ *be the final sequence of* T. *If* T *preserves monics and, for some ordinal* κ, $Z(\kappa{+}1 \to \kappa)$ *is monic, then* Z *stabilizes.*

Proof. Since $\mathbf{C}$ is well-powered, it suffices to prove that for all ordinals $\alpha \geq \kappa$, the arrows $Z(\alpha \to \kappa)$ are monic. Indeed, there is only a proper set of sub-objects of $Z(\kappa)$ up to isomorphism, thus there must exists an ordinal $\kappa' \geq \kappa$ such that $Z(\kappa'{+}1 \to \kappa')$. We prove the statement by transfinite induction on $\alpha \geq \kappa$. Base case: if $\alpha = \kappa$ then $Z(\alpha \to \kappa) = Z(\kappa \to \kappa) = id_{Z(\kappa)}$ hence is monic. Inductive step: assume, by inductive hypothesis that $Z(\alpha \to \kappa)$ is monic. Then the following holds

$$Z(\alpha{+}1 \to \kappa) = Z(\kappa{+}1 \to \kappa) \circ Z(\alpha{+}1 \to \kappa{+}1) \qquad \text{(by func. } Z\text{)}$$
$$= Z(\kappa{+}1 \to \kappa) \circ TZ(\alpha \to \kappa) \qquad \text{(by def. } Z\text{)}$$

Since $Z(\kappa{+}1 \to \kappa)$ is monic and T preserves monomorphisms, $Z(\alpha{+}1 \to \kappa)$ is a composite of monic arrows, hence is monic. Limit step: assume α is a limit ordinal and, by inductive hypothesis, that for all β such that $\kappa \leq \beta < \alpha$, the arrows $Z(\beta \to \kappa)$ are monic. Since, monomorphisms are closed by transfinite pre-composition, to prove that $Z(\alpha \to \kappa)$ is monic it suffices to prove that, for all ordinals δ such that $\kappa \leq \delta \leq \beta$, $Z(\beta \to \delta)$ is monic. We proceed by transfinite induction on δ. The base case $\delta = \kappa$, follows by the inductive hypothesis on α. For the inductive step, assume that $Z(\beta \to \delta)$ is monic. We have that

$$Z(\delta{+}1 \to \kappa) \circ Z(\beta \to \delta{+}1) = Z(\beta \to \kappa) \circ Z(\beta{+}1 \to \delta{+}1) \qquad \text{(by func. } Z\text{)}$$
$$= Z(\beta \to \kappa) \circ TZ(\beta \to \delta) \qquad \text{(by def. } Z\text{)}$$

By inductive hypothesis on δ, $Z(\beta \to \delta)$ is monic, and since T preserves monic arrows, $TZ(\beta \to \delta)$ is monic. By inductive hypothesis on α, both $Z(\delta{+}1 \to \kappa)$ and $Z(\beta \to \kappa)$ are monic, hence by right cancellability $Z(\beta \to \delta{+}1)$ is monic. Finally, assume δ to be a limit ordinal and, by inductive hypothesis, that for all γ such that $\kappa \leq \gamma < \delta$, the arrows $Z(\beta \to \gamma)$ are monic. By functoriality of Z we have

$$Z(\delta \to \kappa) = Z(\gamma \to \kappa) \circ Z(\delta \to \gamma)$$

By inductive hypothesis on α, both $Z(\delta \to \kappa)$ and $Z(\gamma \to \kappa)$ are monic, hence by right cancellability also $Z(\delta \to \gamma)$ is monic. From this, and by right cancellability again, also $Z(\beta \to \delta)$ is monic, since the following holds

$$Z(\beta \to \gamma) = Z(\delta \to \gamma) \circ Z(\beta \to \delta)$$

and $Z(\beta \to \gamma)$ is monic by inductive hypothesis on δ. $\blacksquare$

Historical note. The proof of Lemma 4.1.14 was implicit in [92, Corollary 3.3], where the existence of a monic arrow in the final sequence and the fact that the category **C** is well-powered was implied assuming **C** to be locally presentable. In the statement of [92, Corollary 3.3], the functor was even required to be accessible (i.e., to preserve limits of κ-filtered diagrams). However, this requirement is not strictly needed for the proof, and it was required only for applying a stronger result, namely [92, Proposition 3.2], which could be avoided.

The assumption of well-poweredness of **C** in Lemma 4.1.14 is very mild, and examples of non-well-powered categories are often quite unnatural. For examples of non-well-powered categories one could consider partially ordered classes: a partially order set (or class) $(X, \leq)$ can be interpreted as a category with objects the elements in X and arrows between two objects if and only if they are related by $\leq$. In this way each arrow is a monomorphism (there is only one arrow between two specified objects), therefore if it has top element, its sub-objects are in bijection with the collection of objects which can be a a proper class. The category $\mathbf{Ord}^{op}$ is such an example.

The other two prerequisites in Lemma 4.1.14 are stronger, in particular, requiring that the final sequence reaches a monic arrow. However, these assumptions can be weakened in the case the underlying category **C** has factorization systems.

Lemma 4.1.15 *Assume $(\mathcal{L}, \mathcal{R})$ be a factorization system for the category $\mathbf{C}$. Let $\mathbf{C}$ be $\mathcal{R}$-well-powered, T be an endofunctor on $\mathbf{C}$, and Z be the final sequence of T. If T preserves $\mathcal{R}$-morphisms and, for some ordinal κ, $Z(\kappa{+}1 \to \kappa)$ is an $\mathcal{R}$-morphism, then Z stabilizes.*

Proof. Since **C** is $\mathcal{R}$-well-powered, it suffices to prove that for all ordinals $\alpha \geq \kappa$, $Z(\alpha \to \kappa) \in \mathcal{R}$. We proceed by transfinite induction on $\alpha \geq \kappa$. Base case: if $\alpha = \kappa$ then $Z(\alpha \to \kappa) = Z(\kappa \to \kappa) = id_{Z(\kappa)} \in \mathcal{R}$. Inductive step: assume, by inductive hypothesis that $Z(\alpha \to \kappa) \in \mathcal{R}$. Then the following holds

$$Z(\alpha{+}1 \to \kappa) = Z(\kappa{+}1 \to \kappa) \circ Z(\alpha{+}1 \to \kappa{+}1) \qquad \text{(by func. } Z\text{)}$$
$$= Z(\kappa{+}1 \to \kappa) \circ TZ(\alpha \to \kappa) \qquad \text{(by def. } Z\text{)}$$

Since $Z(\kappa{+}1 \to \kappa)$ is in $\mathcal{R}$ and T preserves $\mathcal{R}$-morphisms, by Lemma 2.1.16, $Z(\alpha{+}1 \to \kappa) \in \mathcal{R}$. Limit step: assume α is a limit ordinal and, by inductive hypothesis, that for all β such that $\kappa \leq \beta < \alpha$, the arrows $Z(\beta \to \kappa)$ are in $\mathcal{R}$. Since, the class $\mathcal{R}$ is closed under transfinite pre-composition (Lemma 2.1.19), to prove $Z(\alpha \to \kappa) \in \mathcal{R}$ it suffices to show that, for all ordinals δ such that $\kappa \leq \delta \leq \beta$, $Z(\beta \to \delta) \in \mathcal{R}$. We do that by transfinite induction on δ. The base case $\delta = \kappa$, follows by the inductive hypothesis on α. For the inductive step, assume that $Z(\beta \to \delta) \in \mathcal{R}$. We have that

$$Z(\delta{+}1 \to \kappa) \circ Z(\beta \to \delta{+}1) = Z(\beta \to \kappa) \circ Z(\beta{+}1 \to \delta{+}1) \qquad \text{(by func. } Z\text{)}$$
$$= Z(\beta \to \kappa) \circ TZ(\beta \to \delta) \qquad \text{(by def. } Z\text{)}$$

By inductive hypothesis on δ, $Z(\beta \to \delta) \in \mathcal{R}$, and since T preserves $\mathcal{R}$-morphisms, $TZ(\beta \to \delta) \in \mathcal{R}$. By inductive hypothesis on α, both $Z(\delta{+}1 \to \kappa)$ and $Z(\beta \to \kappa)$ are $\mathcal{R}$-morphisms, hence, by Lemma 2.1.14, $Z(\beta \to \delta{+}1) \in \mathcal{R}$. Assume δ be a limit ordinal and, by inductive hypothesis, that for all γ such that $\kappa \leq \gamma < \delta$, the arrows $Z(\beta \to \gamma)$ are in $\mathcal{R}$. By functoriality of Z we have

$$Z(\delta \to \kappa) = Z(\gamma \to \kappa) \circ Z(\delta \to \gamma)$$

By inductive hypothesis on α, both $Z(\delta \to \kappa)$ and $Z(\gamma \to \kappa)$ are $\mathcal{R}$-morphisms, hence, by Lemma 2.1.14, $Z(\delta \to \gamma)$ is in $\mathcal{R}$. From this, and by Lemma 2.1.14 again, also $Z(\beta \to \delta)$ is in $\mathcal{R}$ since, by functoriality, $Z(\beta \to \gamma) = Z(\delta \to \gamma) \circ Z(\beta \to \delta)$, and $Z(\beta \to \gamma) \in \mathcal{R}$ by inductive hypothesis on δ. ∎

Lemma 4.1.15 is not a proper generalization of Lemma 4.1.14. Indeed the hypotheses of the former do not imply that of the latter; and, since any category admits $(\mathbf{Morph}, \mathbf{Iso})$ as a factorization

system and all functors preserve isomorphisms, the lemma can always be applied (vacuously). However, when the factorization system $(\mathcal{L}, \mathcal{R})$, is such that $\mathbf{Monic} \neq \mathcal{R}$, Lemma 4.1.15 can be applied even in the case the functor T do not preserve monic arrows but a different class $\mathcal{R}$ of morphisms. In Section 4.1.3 we show a paradigmatic example of a functor for which Lemma 4.1.14 does not work, but Lemma 4.1.15 do.

Note that, although the combination of Lemma 4.1.15 and Theorem 4.1.12 ensures the existence of a final T-coalgebra, we have no bounds on the cardinal at which the final sequence stabilizes.

Remark 4.1.16 Even though not explicitly stated, all the results we have seen so far can be easily dualized in the case of initial sequences and functors preserving either epic arrows or the left class of a factorization system. ∎

4.1.3   Final coalgebra for $\Delta_{<\infty}\colon \mathbf{Meas} \to \mathbf{Meas}$

In this section, we prove the existence of final coalgebras for the functor $\Delta_{<\infty}\colon \mathbf{Meas} \to \mathbf{Meas}$, that has been introduced in Section 2.3. To this end, we consider the final sequence $Z\colon \mathbf{Ord}^{op} \to \mathbf{Meas}$ for $\Delta_{<\infty}$ and prove stabilization for it using Lemma 4.1.15. Remarkably, Lemma 4.1.14 cannot be applied to $\Delta_{<\infty}$ since it does not preserve monic arrows. Therefore this turns out to be a good example for showing the effectiveness of Lemma 4.1.15 in contrast to Lemma 4.1.14.

Recall from Section 2.3 that, $(\mathbf{Epic}, \mathbf{Emb})$ is factorization system in $\mathbf{Meas}$, the category of measurable spaces and measurable functions, where $\mathbf{Epic}$ denotes the class of epic morphisms (i.e., surjective measurable functions) and $\mathbf{Emb}$ is the class of measurable embeddings, that is injective measurable functions $f\colon (X, \Sigma_X) \to (Y, \Sigma_Y)$ such that $\Sigma_X = \{f^{-1}(E) \mid E \in \Sigma_Y\}$, that is, Σ_X is initial w.r.t. f. The functor $\Delta_{<\infty}$ acts on measurable spaces as $(X, \Sigma_X) \mapsto (\Delta_{<\infty}(X, \Sigma_X), \Sigma_{\Delta_{<\infty}(X,\Sigma_X)})$, where $\Delta_{<\infty}(X, \Sigma)$ is the set of finite measures on (X, Σ_X) and $\Sigma_{\Delta_\sigma(X,\Sigma_X)}$ is the smallest σ-algebra making all evaluation maps $ev_E\colon \Delta_\sigma(X, \Sigma_X) \to [0, \infty)$ measurable, for $E \in \Sigma_X$; and on measurable functions $f\colon (X, \Sigma_X) \to (Y, \Sigma_Y)$ as $\Delta_{<\infty}f(\mu) = \mu \circ f^{-1}$, for all $\mu \in \Delta_{<\infty}(X, \Sigma_X)$.

Since $\mathbf{Meas}$ is complete and cocomplete, the final sequence $Z\colon \mathbf{Ord}^{op} \to \mathbf{Meas}$ for $\Delta_{<\infty}$ is well-defined. Moreover, $\mathbf{Meas}$ is $\mathbf{Emb}$-well-powered, and $\Delta_{<\infty}$ preserves embeddings. Thus, in order to apply Lemma 4.1.15 we "only" need to prove that the final sequence eventually reaches an embedding. Next we show that Z reaches the required embedding after ω steps, that is, $Z(\omega+1 \to \omega) \in \mathbf{Emb}$, where ω denotes the first limit ordinal.

Lemma 4.1.17 (σ-algebra of $Z(\omega+1)$) *Let $Z\colon \mathbf{Ord}^{op} \to \mathbf{Meas}$ be the final sequence for $\Delta_{<\infty}$. Then, the σ-algebra $\Sigma_{Z(\omega+1)}$ on $Z(\omega+1)$ is generated by the following collection of subsets*

$$\mathcal{F} = \{ev_{Z(\omega \to n)^{-1}(E)}^{-1}([r, \infty)) \mid r \in [0, \infty) \cap \mathbb{Q} \text{ and } E \in \Sigma_{Z(n)}\}.$$

Proof. By definition of final sequence, $(Z(\omega \to n))_{n<\omega}$ is a limit cone over $Z{\upharpoonright}\omega$, therefore $Z(\omega)$ has initial σ-algebra w.r.t. its cone, that is $\Sigma_{Z(\omega)} = \sigma(\{Z(\omega \to n)^{-1}(E) \mid E \in \Sigma_{Z(n)}, n < \omega\})$. We show that the family $\mathcal{A} = \{Z(\omega \to n)^{-1}(E) \mid E \in \Sigma_{Z(n)}, n < \omega\}$ is a boolean algebra. The empty set $\emptyset$ is contained in $\mathcal{A}$, since $Z(\omega \to n)^{-1}(\emptyset) = \emptyset$, for all $n < \omega$. Assume $A \in \mathcal{A}$, then there exists $n < \omega$ and $E \in \Sigma_{Z(n)}$ such that $A = Z(\omega \to n)^{-1}(E)$. The following hold by definition of pre-image

$$Z(\omega) \setminus A = Z(\omega) \setminus Z(\omega \to n)^{-1}(E) = Z(\omega \to n)^{-1}(Z(n) \setminus E).$$

Since $Z(n) \setminus E \in \Sigma_{Z(n)}$, then $Z(\omega) \setminus A \in \mathcal{A}$. Assume $A, B \in \mathcal{A}$, then there exits $m, n < \omega$, $E_m \in \Sigma_{Z(m)}$, and $E_n \in \Sigma_{Z(n)}$ such that $A = Z(\omega \to m)^{-1}(E_m)$ and $B = Z(\omega \to n)^{-1}(E_n)$ Without loss of generality assume that $m \leq n$. Then we have

$$
\begin{aligned}
A = Z(\omega \to m)^{-1}(E_m) &= (Z(n \to m) \circ Z(\omega \to n))^{-1}(E_m) && \text{(by func. } Z) \\
&= Z(\omega \to n)^{-1} \circ Z(n \to m)^{-1}(E_m) && \text{(by inverse images)} \\
&= Z(\omega \to n)^{-1}(Z(n \to m)^{-1}(E_m)). && \text{(by composition)}
\end{aligned}
$$

From this we derive the following equality

$$A \cup B = Z(\omega \to n)^{-1}\big(Z(n \to m)^{-1}(E_m)\big) \cup Z(\omega \to n)^{-1}(E_n)$$
$$= Z(\omega \to n)^{-1}\big(Z(n \to m)^{-1}(E_m) \cup E_n\big).$$

Since $Z(n \to m)$ is measurable $Z(n \to m)^{-1}(E_m) \in \Sigma_{Z(n)}$, hence $Z(n \to m)^{-1}(E_m) \cup E_n \in \Sigma_{Z(n)}$. Therefore $A \cup B \in \mathcal{A}$. From these we conclude that $\mathcal{A}$ is a boolean algebra. Now, the thesis follows by Lemma 2.3.3, noticing that by definition of final sequence, $Z(\omega+1) = \Delta_{<\infty}Z(\omega)$, and $L_r(E) = ev_E^{-1}([r,\infty))$. ∎

The first step to prove that $Z(\omega+1 \to \omega)$ is an embedding, is to show that σ-algebra on $Z(\omega+1)$ is initial with respect to $Z(\omega+1 \to \omega)$. This is proved by the following proposition.

Proposition 4.1.18 *Let $Z \colon \mathbf{Ord}^{op} \to \mathbf{Meas}$ be the final sequence for $\Delta_{<\infty}$. Then, the σ-algebra on $Z(\omega+1)$ is initial w.r.t. $Z(\omega+1 \to \omega)$, that is $\Sigma_{Z(\omega+1)} = \{Z(\omega+1 \to \omega)^{-1}(E) \mid E \in \Sigma_{Z(\omega)}\}$.*

Proof. Let $\mathcal{E} = \{Z(\omega+1 \to \omega)^{-1}(E) \mid E \in \Sigma_{Z(\omega)}\}$. The inclusion $\mathcal{E} \subseteq \Sigma_{Z(\omega+1)}$ follows since $Z(\omega+1 \to \omega)$ is measurable. As for the reverse inclusion, we know, by Lemma 4.1.17, that $\Sigma_{Z(\omega+1)}$ is generated by the following collection of subsets

$$\mathcal{F} = \{ev_{Z(\omega \to n)^{-1}(E)}^{-1}([r,\infty)) \mid r \in [0,\infty) \cap \mathbb{Q} \text{ and } E \in \Sigma_{Z(n)}\}.$$

Since $\mathcal{E}$ is already a σ-algebra, to prove $\sigma(\mathcal{F}) \subseteq \mathcal{E}$, we only need to show that $\mathcal{F} \subseteq \mathcal{E}$. This is done noticing that, for all $n < \omega$, $r \in [0,\infty)$ and $E \in \Sigma_{Z(n)}$

$$\begin{aligned}
ev_{Z(\omega \to n)^{-1}(E)}^{-1}([r,\infty)) &= \{\mu \in Z(\omega+1) \mid \mu(Z(\omega \to n)^{-1}(E)) \geq r\} && \text{(by inverse image)} \\
&= \{\mu \in Z(\omega+1) \mid \big(\Delta_{<\infty}Z(\omega \to n)(\mu)\big)(E) \geq r\} && \text{(by def. } \Delta_{<\infty}) \\
&= \{\mu \in Z(\omega+1) \mid \big(Z(\omega+1 \to n+1)(\mu)\big)(E) \geq r\} && \text{(by def. } Z) \\
&= \big(ev_E \circ Z(\omega+1 \to n+1)\big)^{-1}([r,\infty)) && \text{(by def. } ev_E \text{ and inv. image)} \\
&= \big(ev_E \circ Z(\omega \to n+1) \circ Z(\omega+1 \to \omega)\big)^{-1}([r,\infty)) && \text{(by func. } Z) \\
&= Z(\omega+1 \to \omega)^{-1}\big(Z(\omega \to n+1)^{-1} \circ ev_E^{-1}([r,\infty))\big). && \text{(by inverse)}
\end{aligned}$$

Clearly, $ev_E^{-1}([r,\infty)) \in \Sigma_{Z(n+1)}$, so that $Z(\omega \to n+1)^{-1}(ev_E^{-1}([r,\infty))) \in \Sigma_{Z(\omega)}$ follows by measurability of $Z(\omega \to n+1)$. This proves $ev_{Z(\omega \to n)^{-1}(E)}^{-1}([r,\infty)) \in \mathcal{E}$, therefore $\mathcal{F} \subseteq \mathcal{E}$. ∎

The last step is to show that $Z(\omega+1 \to \omega)$ is injective.

Proposition 4.1.19 (Injective) *Let $Z \colon \mathbf{Ord}^{op} \to \mathbf{Meas}$ be the final sequence for $\Delta_{<\infty}$. Then $Z(\omega+1 \to \omega)$ is injective.*

Proof. By definition of final sequence, $(Z(\omega \to n))_{n<\omega}$ is a limit cone over $Z{\restriction}\omega$, therefore $Z(\omega)$ is equipped with the initial σ-algebra w.r.t. its cone, that is $\Sigma_{Z(\omega)}$ is generated by

$$\mathcal{A} = \sigma(\{Z(\omega \to n)^{-1}(E) \mid E \in \Sigma_{Z(n)}, n < \omega\}).$$

By definition of final sequence $Z(\omega+1) = \Delta_{<\infty}Z(\omega)$. Let $\mu, \nu \in \Delta_{<\infty}Z(\omega)$ be to measure on $Z(\omega)$, and assume $Z(\omega+1 \to \omega)(\mu) = Z(\omega+1 \to \omega)(\nu)$. We have to show that $\mu = \nu$. Since μ and ν are clearly σ-finite, and they are also pre-measures on the boolean algebra $\mathcal{A}$ (see the proof of Lemma 4.1.17), by Lemma 2.2.27 it suffices to show that μ and ν agree on all subsets in $\mathcal{A}$. This is shown below, for all $n < \omega$ and $E \in \Sigma_{Z(n)}$

$$\begin{aligned}
\mu(Z(\omega \to n)^{-1}(E)) &= \Delta_{<\infty}Z(\omega \to n)(\mu)(E) && \text{(by def. } \Delta_{<\infty}) \\
&= Z(\omega+1 \to n+1)(\mu)(E) && \text{(by def. } Z) \\
&= Z(\omega \to n+1) \circ Z(\omega+1 \to \omega)(\mu)(E) && \text{(by func. } Z)
\end{aligned}$$

$$
\begin{aligned}
&= Z(\omega \to n{+}1) \circ Z(\omega{+}1 \to \omega)(\nu)(E) &&\text{(by hp.)}\\
&= Z(\omega{+}1 \to n{+}1)(\nu)(E) &&\text{(by func. } Z)\\
&= \Delta_{<\infty} Z(\omega \to n)(\nu)(E) &&\text{(by def. } Z)\\
&= \nu(Z(\omega \to n)^{-1}(E)). &&\text{(by def. } \Delta_{<\infty})
\end{aligned}
$$

Therefore, $Z(\omega{+}1 \to \omega)$ is injective.   ∎

Theorem 4.1.20 *The functor* $\Delta_{<\infty}\colon \mathbf{Meas} \to \mathbf{Meas}$ *has final coalgebra.*

Proof. **Meas** has terminal object and limits for ordinal indexed diagrams, thus the final sequence $Z\colon \mathbf{Ord}^{op} \to \mathbf{Meas}$ for $\Delta_{<\infty}$ is well-defined. $(\mathbf{Epic}, \mathbf{Emb})$ is a factorization system in **Meas**, moreover, by $\mathbf{Emb} \subseteq \mathbf{Monic}$ and the fact that **Meas** is well-powered, it follows that it is also **Emb**-well-powered. By Propositions 4.1.18 and 4.1.19 we have $Z(\omega{+}1 \to \omega) \in \mathbf{Emb}$, so that, by Lemma 4.1.15, the final sequence stabilizes. The existence of $\Delta_{<\infty}$-coalgebra follows by Theorem 4.1.12.   ∎

Remark 4.1.21 The general case $\Delta\colon \mathbf{Meas} \to \mathbf{Meas}$ does not fit the proof we have provided for Proposition 4.1.19 since two measures which are not σ-finite may agree on all the elements of a generator of the σ-algebra but still differ on some measurable set. This case is very subtle and it is always very difficult to prove equality over generic measures over generated σ-algebras, without assuming σ-finiteness. A possible route to a proof of equality could be by transfinite induction on the generation steps that leads to the actual σ-algebra. Indeed, any family of sets can be turned into a σ-algebra by making a transfinite number of closures under complementations and countable unions (at least ω_1 steps are needed). Although this is a plausible strategy, it is always very difficult to treat the limit steps. Another route could be by using the π-λ-theorem, but this requires a different characterization of the σ-algebra on $Z(\omega)$. We already tried these strategies without any success, so we are still wondering if we need to find different ordinals greater then ω for proving the existence of embeddings in the final sequence for Δ.

Moreover, also the case $\Delta_\sigma\colon \mathbf{Meas} \to \mathbf{Meas}$ does not fit the proof of Proposition 4.1.17 since it makes use of Lemma 2.3.3 which holds only assuming the measures are finite.   ∎

Remark 4.1.22 (Final coalgebra for L-labelled Markov kernels) It easy to see that the whole constructions in this section apply also to the composite functor $\Delta_{<\infty}^L\colon \mathbf{Meas} \to \mathbf{Meas}$, for any set L. Hence also for this functor there exits a final coalgebra. The existence of such a terminal object in the category of coalgebras for $\Delta_{<\infty}^L$ will be indispensable in Chapter 6 in order to have a principle of coinduction on $\Delta_{<\infty}^L$-coalgebras, that is, L-labelled Markov kernels (see Section 5.2).   ∎

Historical note. In [86], van Breugel et al. consider the functor $\Delta_{\leq 1}^L\colon \mathbf{Meas} \to \mathbf{Meas}$ of subprobabilities over a measurable space. Given a countable set L, the construction of the final coalgebra of the functor $\Delta_{\leq 1}^L$ is performed using the the connection of the final sequence for this functor with the one for the labelled probabilistic powerdomain $V^L\colon \omega\mathbf{Coh} \to \omega\mathbf{Coh}$ in the category of ω-coherent domains (i.e., topological spaces over ω-continuous dcpo which are compact in their Lawson topology). Furthermore, they proved that the final coalgebra can be regarded as a Polish space, in the category of metric spaces and continuous maps. Their approach gives better insights on the carrier of the final coalgebra, relating it which different structures such as domains with Lawson topologies on them, and Polish metric spaces. However, the entire construction is very complicated and needs one to jump from a category to an other, and moreover it does not give any general proof technique to prove the existence of final coalgebras for different behaviour functors.

Another proof for the existence of final coalgebras for $\Delta_1\colon \mathbf{Meas} \to \mathbf{Meas}$ is due to Viglizzo and Moss in [68] where the existence of final coalgebras was proved for any functor constructed as the finite composition of constant functors, identity functors, binary product and coproduct functors,

and Δ_1 (the so called class of measurable polynomial endofunctor in **Meas**). Their construction uses modal logics: the elements of the final coalgebra are theories (sets of modal formulas) satisfied by states in all possible coalgebras. Viglizzo [90] also provided another construction of a final coalgebra for the same class of functors, avoiding the logic in favor of final sequences. Still the connection their results and ours is very hard to be formalized.

4.1.4 Initial algebras for polynomial endofunctors on Meas

In this section, we prove the existence of initial algebras for polynomial endofunctors on **Meas** (see Section 2.3 for their definition). To this end, we consider initial sequences $A\colon \mathbf{Ord} \to \mathbf{Meas}$ and the dual of Lemma 4.1.15 to prove stabilization of the initial sequence:

Lemma 4.1.23 *Assume* $(\mathcal{L}, \mathcal{R})$ *be a factorization system for the category* **C**. *Let* **C** *be* $\mathcal{L}$-*cowell-powered,* T *be an endofunctor on* **C**, *and* A *be the initial sequence of* T. *If* T *preserves* $\mathcal{L}$-*morphisms and, for some ordinal* κ, $A(\kappa \to \kappa{+}1)$ *is an* $\mathcal{L}$-*morphism, then* A *stabilizes.*

On **Meas**, there are many factorization systems that "lift" the factorization system on **Set**, e.g., quotients and injections (**Quot**, **Monic**), surjections and embeddings (**Epic**, **Emb**), etc. (see Section 2.3). However, in order to apply Lemma 4.1.23 we cannot choose the former factorization system. Indeed, polynomial functors in **Meas** do not preserve quotients (it is well known that binary products fail to do so), but they do preserve surjections. Hence, this is also a good example to show the flexibility of the approach.

Since **Meas** is complete and cocomplete, the initial sequence $A\colon \mathbf{Ord} \to \mathbf{Meas}$ for any polynomial functor $P\colon \mathbf{Meas} \to \mathbf{Meas}$ is well-defined. Moreover, **Meas** is cowell-powered, and P preserves epimorphism (i.e., measurable surjections). Thus, we only need to prove that the initial sequence eventually reaches an epimorphism, and we will show that this happens exactly after ω steps, that is, $A(\omega \to \omega{+}1) \in \mathbf{Epic}$.

Proposition 4.1.24 *Let* $A\colon \mathbf{Ord} \to \mathbf{Meas}$ *be the initial sequence for a polynomial endofunctor* $P\colon \mathbf{Meas} \to \mathbf{Meas}$. *Then* $A(\omega \to \omega{+}1)$ *is an epimorphism.*

Proof. For any polynomial functor in $P\colon \mathbf{Meas} \to \mathbf{Meas}$ there exists an associated a polynomial endofunctor $P'\colon \mathbf{Set} \to \mathbf{Set}$, such that $P'U = UP$, where $U\colon \mathbf{Meas} \to \mathbf{Set}$ is the obvious forgetful functor. We prove that $UA\colon \mathbf{Ord} \to \mathbf{Set}$ is the initial sequence for P'. Clearly, since both A and U preserves colimits, also UA preserves them. Moreover, for all ordinals $\gamma \leq \beta$ the following holds

$$UA(0) = U\mathbf{0} = \mathbf{0}$$
$$UA(\beta{+}1) = UPA(\beta) = P'UA(\beta)$$
$$UA(\gamma{+}1 \to \beta{+}1) = UPA(\gamma \to \beta) = P'UA(\gamma \to \beta)\,.$$

Therefore $UA\colon \mathbf{Ord} \to \mathbf{Set}$ is a well-defined initial sequence for P'. Recall that polynomial functors in **Set** are ω-cocontinuous, that is, preserves colimits of ω-sequences. Therefore the initial sequence UA of P' stabilizes at ω, thus, $UA(\omega \to \omega{+}1)$ is an isomorphism and, in particular, also an epimorphism. Since U reflects epimorphism to **Meas**, $A(\omega \to \omega{+}1)$ is an epic arrow. ∎

The above proof technique works in many situations, for example it can be applied also to categories different from set such as **Top**, **UMet**, **PMet**, etc. The need of reducing problem to consider an associated initial sequence in **Set** is needed because it is not known if all polynomial endofunctors do preserve colimits of ω-sequences in **Meas** (the same holds for **Top**). However the approach is quite elegant and still leads to the the existence of initial algebras for measurable polynomial endofunctors.

Theorem 4.1.25 (Initial algebra) *Any polynomial endofunctor in* **Meas** *has initial algebra.*

4.2   An Alternative Final Coalgebra Construction

In this section, we provide an alternative characterization of final coalgebras out of weakly final ones, which are easier to be found then proper final objects. At first sight, this may not appear a new result, since it is standard that final objects can be obtained taking the coequalizer of all the endomorphisms of a weakly final one. However, adopting this characterization for the final object needs that one already knows all the endomorphisms of the weakly final object —which are often hard to be determined— and that the underlying category has generalized coequalizers. Moreover, even when all this informations are known, this characterization of a final object is highly non-constructive and it may be difficult to work with it.

Our proposal makes use of the final sequence, but rather than trying to determine bounds for stabilization, we use it to provide unique projection for weakly final coalgebras into it (cf. Definition 4.1.10) in order to obtain final coalgebras constructively and without assuming the category has coequalizers. This can be done assuming that the underlying category has a factorization system $(\mathcal{L}, \mathcal{R})$ with $\mathcal{R} \subseteq \mathbf{Monic}$, and that the final sequence has an $\mathcal{R}$-morphisms. Moreover, we propose a general construction that allows to obtain weakly final coalgebras under the (very mild) additional assumption that the underlying category has $\mathcal{R}$-unions. Interestingly, the entire construction can be made functorial, in particular, we will be able to define a quotient functor $Q_{\mathcal{L}} \colon T\text{-}\mathbf{coalg} \to T\text{-}\mathbf{coalg}$, mapping a T-coalgebra to its $\mathcal{L}$-quotient with respect to its canonical projection into the final sequence.

Definition 4.2.1 (Final quotient) *Assume* $\mathbf{C}$ *has factorization system* $(\mathcal{L}, \mathcal{R})$, $T \colon \mathbf{C} \to \mathbf{C}$ *be a functor, and* $Z \colon \mathbf{Ord}^{op} \to \mathbf{C}$ *be the final sequence of* T. *Given a* T-*coalgebra* (X, h) *and a final projection* $k \colon X \to D(\alpha)$ *for it, we say that a* T-*coalgebra* (Q_h, q_h) *is a* final $\mathcal{L}$-*quotient at* α *for* (X, h), *if the following diagrams commute:*

$$
\begin{array}{ccccc}
& & k & & \\
& \overset{\frown}{} & & & \\
X & \xrightarrow{\lambda \in \mathcal{L}} & Q_h & \xrightarrow{\rho \in \mathcal{R}} & Z(\alpha) \\
\downarrow h & & \downarrow q_h & & \downarrow \\
TX & \xrightarrow{T\lambda} & TQ_h & \xrightarrow{T\rho} & Z(\alpha+1) \\
& & Tk & &
\end{array}
$$

where $\rho \circ \lambda$ *is a* $(\mathcal{L}, \mathcal{R})$-*factorization for* k.

Under some assumption on the functor and on its final sequence, each coalgebra has a unique quotient into the final sequence, up to isomorphism.

Lemma 4.2.2 *Assume* $\mathbf{C}$ *has factorization system* $(\mathcal{L}, \mathcal{R})$, $T \colon \mathbf{C} \to \mathbf{C}$ *preserves* $\mathcal{R}$-*morphisms, and* $Z \colon \mathbf{Ord}^{op} \to \mathbf{C}$ *be the final sequence of* T. *If* $Z(\alpha+1 \to \alpha) \in \mathcal{R}$, *for some ordinal* α, *then* (X, h) *has a unique final* $\mathcal{L}$-*quotient at* α.

Proof. By Lemma 4.2.2 and Lemma 4.1.9, there exists a unique final projection at α for (X, h), say $k \colon X \to Z(\alpha)$. Assume $k = \rho \circ \lambda$ be an $(\mathcal{L}, \mathcal{R})$-factorization for it. By hypothesis, T preserves $\mathcal{R}$-morphisms, hence $T\rho \in \mathcal{R}$. Since $Z(\alpha+1 \to \alpha) \in \mathcal{R}$ and $\mathcal{R}$ is closed by composition, we have $Z(\alpha+1 \to \alpha) \circ T\rho \in \mathcal{R}$. By hypothesis $k = Z(\alpha+1 \to \alpha) \circ Tk \circ h$, therefore the diagram below is a lifting problem for $\lambda \in \mathcal{L}$ and $Z(\alpha+1 \to \alpha) \circ T\rho \in \mathcal{R}$ with unique solution q_h.

$$
\begin{array}{ccc}
X & \xrightarrow{\quad \lambda \quad} & Q_h \\
\downarrow{T\lambda \circ h} & \overset{q_h}{\nearrow} & \downarrow{\rho} \\
TQ_h \xrightarrow{T\rho} & Z(\alpha+1) \longrightarrow & Z(\alpha)
\end{array}
$$

This defines a T-coalgebra (Q_h, q_h), and the same diagram proves that it is a final $\mathcal{L}$-quotient for (X, h). Uniqueness (up-to-isomorphism) follows by uniqueness of the $(\mathcal{L}, \mathcal{R})$-factorization. ∎

The above lemma can be made even stronger if we assume that the factorization system $(\mathcal{L}, \mathcal{R})$ is such that $\mathcal{R} \subseteq \mathbf{Monic}$. This will be used in the next theorem which allows us to characterize a final coalgebra as the final $\mathcal{L}$-quotient for a weakly final coalgebra.

Theorem 4.2.3 (Final coalgebra as a final quotient) *Assume the category* $\mathbf{C}$ *has factorization system* $(\mathcal{L}, \mathcal{R})$ *such that* $\mathcal{R} \subseteq \mathbf{Monic}$, $T: \mathbf{C} \to \mathbf{C}$ *preserves* $\mathcal{R}$-*morphisms, and* $Z: \mathbf{Ord}^{op} \to \mathbf{C}$ *is the final sequence of* T. *If* $Z(\alpha{+}1 \to \alpha) \in \mathcal{R}$, *for some ordinal* α, *and* (W, w) *is a weakly final* T-*coalgebra, then the final* $\mathcal{L}$-*quotient at* α *for* (W, w) *is a final* T-*coalgebra.*

Proof. Assume (Q_w, q_w) be the unique final $\mathcal{L}$-quotient at α for (W, w) defined as in Lemma 4.2.2:

$$
\begin{array}{ccccc}
& & w^\alpha & & \\
W & \xrightarrow{\lambda} & Q_w & \xrightarrow{\rho} & Z(\alpha) \\
{\scriptstyle w}\downarrow & & {\scriptstyle q_w}\downarrow & & \downarrow \\
TW & \xrightarrow{T\lambda} & TQ_w & \xrightarrow{T\rho} & Z(\alpha{+}1) \\
& & Tw^\alpha & &
\end{array}
$$

Let (X, h) be a T-coalgebra, then, by weak finality of (W, w), there exists a T-homomorphism $f: (X, h) \to (W, w)$. Since λ is a T-homomorphism between (W, w) and (Q_w, q_w), we have that $\lambda \circ f$ is a T-homomorphism from (X, h) to (Q_w, q_w). This proves weak finalilty. As for uniqueness, let $f, g: (X, h) \to (Q_w, q_w)$ be two morphisms of T-coalgebras. Consider the two composites $\lambda \circ f$ and $\lambda \circ g$. Since both are projections at α for (X, h) into Z, by Lemma 4.1.11, we have that $\lambda \circ f = \lambda \circ g$. Since $\lambda \in \mathcal{R}$ and $\mathcal{R} \subseteq \mathbf{Monic}$, by left cancellabilty of monomorphisms we conclude that $f = g$. ∎

Notably, Theorem 4.2.3 provides a constructive characterization for a final coalgebra. Indeed, our approach gives an actual definition rather than just a characterization: the projection is unique (Lemma 4.2.2) and it is canonically defined for any functor (Lemma 4.1.9); moreover, the carrier of the final coalgebra is uniquely determined by the factorization of the projection (which is unique up to isomorphism) and the structure map is the unique solution of a lifting problem.

4.2.1 A weakly final coalgebra construction

If the category has a factorization system $(\mathcal{L}, \mathcal{R})$ with $\mathcal{R} \subseteq \mathbf{Monic}$, and the final sequence eventually reaches an $\mathcal{R}$-arrow, then, in order to apply Theorem 4.2.3 one needs a weakly final coalgebra. In this section we provide a very simple construction which only additionally requires that the underlying category has coproducts and is $\mathcal{R}$-well-powered. This approach is very general and applies to a wide class of categories and functors.

Proposition 4.2.4 (Weakly final coalgebra) *Assume* $\mathbf{C}$ *has coproducts and a factorization system* $(\mathcal{L}, \mathcal{R})$, $T: \mathbf{C} \to \mathbf{C}$ *preserves* $\mathcal{R}$-*morphisms, and* $Z: \mathbf{Ord}^{op} \to \mathbf{C}$ *be the final sequence of* T *such that* $Z(\alpha{+}1 \to \alpha) \in \mathcal{R}$, *for some ordinal* α. *If* $\mathbf{C}$ *is* $\mathcal{R}$-*well-powered a weakly final coalgebra exists and it is given as the coproduct of all final* $\mathcal{L}$-*quotients at* α.

Proof. Let (X, h) be a T-coalgebra and (Q_h, q_h) be its final $\mathcal{L}$-quotient defined as in Lemma 4.2.2,

so that the following diagram commutes:

$$
\begin{array}{ccccc}
 & & h^\alpha & & \\
X & \xrightarrow{\lambda_h \in \mathcal{L}} & Q_h & \xrightarrow{\rho_h \in \mathcal{R}} & Z(\alpha) \\
\downarrow{h} & & \vdots\, q_h & & \uparrow \\
TX & \xrightarrow[T\lambda_h]{} & TQ_h & \xrightarrow[T\rho_h]{} & Z(\alpha+1) \\
 & & Th^\alpha & &
\end{array}
$$

Notice that Q_h is an $\mathcal{R}$-subobject of $Z(\alpha)$ and this holds for any T-coalgebra (X, h). By hypothesis, **C** is $\mathcal{R}$-well-powered, therefore $\mathcal{Q} = \{(Q_h, q_h) \mid (X, h) \text{ is a } T\text{-coalgebra}\}$ is a proper set of T-coalgebras and we are allowed to take its coproduct $\coprod_{(Q,q) \in \mathcal{Q}} (Q, q)$, which is well defined, since **C** has coproducts and T-**coalg** lift them. $\coprod_{(Q,q) \in \mathcal{Q}} (Q, q)$ is readily seen to be weakly final, since for any T-coalgebra (X, h) the following diagrams commute

$$
\begin{array}{ccccc}
X & \xrightarrow{\lambda_h} & Q_h & \xrightarrow{in_{Q_h}} & \coprod_h Q_h \\
\downarrow{h} & & \downarrow{q_h} & & \downarrow \\
TX & \xrightarrow[T\lambda_h]{} & TQ_h & \xrightarrow[Tin_{Q_h}]{} & T\coprod_h Q_h
\end{array}
$$

hence $in_{Q_h} \circ \lambda_h$ is a T-homomorphism from (X, h) to $\coprod_{(Q,q) \in \mathcal{Q}} (Q, q)$. ∎

Corollary 4.2.5 *Under the assumptions of Proposition 4.2.4 and assuming $\mathcal{R} \subseteq$ **Monic**, the final coalgebra is the $\mathcal{R}$-union of the final $\mathcal{L}$-quotients of all T-coalgebras.*

Proof. Immediate from Proposition 4.2.4, Theorem 4.2.3, and the characterization of $\mathcal{R}$-union as the $\mathcal{L}$-quotient of a coproduct (see Definition 2.1.21 for details about generalized unions in categories with factorization systems). ∎

Historical note. Weakly final coalgebras have always been subject of research since they give "estimates" of the final coalgebra. A canonical final coalgebra can be found in the final sequence as the left inverse of some arrow in it (we will see this in more detail in Section 4.2.3). In **Set** any monic arrow with non empty domain has left inverse, hence to obtain a weakly final coalgebra one only needs to prove that the final sequence eventually reaches a monic arrow at some ordinal cardinal. Moreover, since any **Set**-endofunctor preserves monomorphism the construction given in Theorem 4.2.3 applies very easily in this settings. However, this is a very particular case and does not hold in other categories, such as **CMet** (complete metric spaces), **UMet** (ultra-metric spaces), **POSet** (partially ordered sets), **Meas** (measurable spaces), **Top** (topological spaces), etc. for which our generalization with factorization systems is really useful.

Other sources for weakly final coalgebras are right invertible natural transformations. Indeed, if a final coalgebra (W, w) exists for an endofunctor H, this can be turned into a weakly final one $(W, \theta_W \circ w)$ for the functor T, where $\theta \colon H \Rightarrow T$ has right inverse in each component. A reasonable proof sketch is provided by the following diagram, where (X, h) is any T-coalgebra, r is the right inverse of θ_X and $f \colon (X, r \circ h) \to (W, w)$ is the final H-homomorphism to (W, w):

$$
\begin{array}{ccccccc}
 & & & & id_{TX} & & \\
X & \xrightarrow{h} & TX & \xrightarrow{r} & HX & \xrightarrow{\theta_X} & TX \\
\downarrow{f} & & & & \downarrow{Hf} & & \downarrow{Tf} \\
W & \xrightarrow[w]{\cong} & & & HW & \xrightarrow[\theta_W]{} & TW
\end{array}
$$

This situation is very useful when one is dealing with finitary **Set**-endofunctors F, for which there always be right invertible natural transformations (also known as quotients) $\theta\colon F_\Sigma \Rightarrow F$ from an associated syntactic endofunctor F_Σ [9]. Indeed, syntactic **Set**-endofunctors always have final coalgebras (also with a nice characterization) and right inverses for the components can be found requiring that each component of θ is an epimorphism. However, also in this case, right invertible arrows are always difficult to be found in categories different from **Set**.

4.2.2 The quotient functor

In this section, we show that in categories with factorization system $(\mathcal{L}, \mathcal{R})$ with $\mathcal{R} \subseteq \mathbf{Monic}$, and such that the final sequence $Z\colon \mathbf{Ord}^{op} \to \mathbf{C}$ for an endofunctor $T\colon \mathbf{C} \to \mathbf{C}$ is well defined, to any T-coalgebra can be assigned its final $\mathcal{L}$-quotient at α (cf. Definition 4.2.1), provided that T-preserves $\mathcal{R}$-morphisms and $Z(\alpha{+}1 \to \alpha) \in \mathcal{R}$. Remarkably, this mapping is functorial and yields to the so called final $\mathcal{L}$-quotient functor $Q_\mathcal{L}\colon T\text{-}\mathbf{coalg} \to T\text{-}\mathbf{coalg}$.

Proposition 4.2.6 (Quotient functor) *Let $Z\colon \mathbf{Ord}^{op} \to \mathbf{C}$ be the final sequence of an endofunctor T on a category $\mathbf{C}$ with factorization system $(\mathcal{L}, \mathcal{R})$ with $\mathcal{R} \subseteq \mathbf{Monic}$. If T preserves $\mathcal{R}$-morphisms and $Z(\alpha{+}1 \to \alpha) \in \mathcal{R}$, for some α, then the functor $Q_\mathcal{L}\colon T\text{-}\mathbf{coalg} \to T\text{-}\mathbf{coalg}$ acting on T-coalgebras (X, h) and T-homomorphisms $f\colon (X, h) \to (Y, k)$ as follows*

$$Q_\mathcal{L}(X, h) = (Q_h, q_h) \qquad\qquad Q_\mathcal{L}f = \varphi_h$$

where (Q_h, q_h) is the unique final $\mathcal{L}$-quotient at α for (X, h) given as in Lemma 4.2.2 (cf. definition diagram below on the left), and $\varphi_h\colon Q_\mathcal{L}(X, h) \to Q_\mathcal{L}(Y, k)$ is the unique solution to the lifting problem below on the right

is well defined.

Proof. $Q_\mathcal{L}$ is clearly well-defined on objects. Let $f\colon (X, h) \to (Y, k)$ be an arrow in $T\text{-}\mathbf{coalg}$, $Q_\mathcal{L}(X, h) = (Q_h, q_h)$ and $Q_\mathcal{L}(Y, k) = (Q_k, q_k)$. We have to prove that $Q_\mathcal{L}f = \varphi_f\colon Q_h \to Q_k$ is a T-homomorphism between (Q_h, q_h) and (Q_h, q_k), that is, $T\varphi_f \circ q_h = q_k \circ \varphi_f$. To this end notice that the following holds:

$$
\begin{aligned}
A(\alpha{+}1 \to \alpha) \circ T\rho_k \circ T\varphi_f \circ q_h &= A(\alpha{+}1 \to \alpha) \circ T(\rho_k \circ \varphi_f) \circ q_h && \text{(by func. } T) \\
&= A(\alpha{+}1 \to \alpha) \circ T\rho_h \circ q_h && \text{(by def. } \varphi_f) \\
&= \rho_h && \text{(by def. } q_h) \\
&= \rho_k \circ \varphi_f && \text{(by def. } \varphi_f) \\
&= A(\alpha{+}1 \to \alpha) \circ T\rho_k \circ q_k \circ \varphi_f && \text{(by def. } \rho_k)
\end{aligned}
$$

Since T preserves $\mathcal{R}$-morphisms and $\mathcal{R}$ is closed under composition, we have that the composite $A(\alpha{+}1 \to \alpha) \circ T\rho_k \in \mathcal{R}$. By $\mathcal{R} \subseteq \mathbf{Monic}$ and left cancellability of monomorphisms, we have $T\varphi_f \circ q_h = q_k \circ \varphi_f$. Therefore φ_f is a T-homomorphism.

Fuctoriality is easily proved. Let $f\colon (X,h) \to (Y,k)$ and $g\colon (Y,k) \to (Z,l)$ be morphisms in T-**coalg**. By definition $\rho_h = \rho_l \circ Q_{\mathcal{L}}(g \circ f)$, $\rho_h = \rho_k \circ Q_{\mathcal{L}}(f)$, and $\rho_k = \rho_l \circ Q_{\mathcal{L}}(g)$, therefore $\rho_l \circ Q_{\mathcal{L}}(g \circ f) = \rho_l \circ Q_{\mathcal{L}}(g) \circ Q_{\mathcal{L}}(f)$. Since $\rho_l \in \mathcal{R} \subseteq$ **Monic**, by left cancellability of monomorphisms, we have $Q_{\mathcal{L}}(g \circ f) = Q_{\mathcal{L}}(g) \circ Q_{\mathcal{L}}(f)$. Consider $Q_{\mathcal{L}} id_{(X,h)}$. By definition, $\rho_h \circ Q_{\mathcal{L}} id_{(X,h)} = \rho_h$, therefore $\rho_h \circ Q_{\mathcal{L}} id_{(X,h)} = \rho_h \circ id_X$. Again, by left cancellability of $\rho_h \in \mathcal{R} \subseteq$ **Monic**, we have $Q_{\mathcal{L}} id_{(X,h)} = id_X$, from which follows $Q_{\mathcal{L}} id_{(X,h)} = id_{(X,h)}$. ∎

Remark 4.2.7 $(\mathcal{L}, \mathcal{R})$-factorizations of morphisms are unique only up to isomorphism, hence for a given T-coalgebra (X,h), the associated T-coalgebra (Q_h, q_h) is uniquely determined only up to isomorphism. However, under the hypothesis of Proposition 4.2.6, one can fix any $(\mathcal{L}, \mathcal{R})$-factorization $h_\alpha = \rho_h \circ \lambda_h$, and still the $\mathcal{L}$-quotient functor is well-defined. Moreover, note that functoriality crucially depends on the assumption that all $\mathcal{R}$-morphisms are monic. ∎

Remark 4.2.8 Theorem 4.2.3 may be restated in terms of the $\mathcal{L}$-quotient functor $Q_{\mathcal{L}}$, simply saying that any weakly final coalgebra (W,w) gives rise to a final coalgebra $Q_{\mathcal{L}}(W,w)$. Similarly, Corollary 4.2.4 can be restated saying that the final coalgebra is given as the union of (set) all coalgebras which are images under $Q_{\mathcal{L}}$. ∎

4.2.3   Bound for stabilization of the final sequence

When the final sequence $Z\colon \mathbf{Ord}^{op} \to \mathbf{C}$ of T reaches a left invertible arrow $Z(\alpha{+}1 \to \alpha)$ at some ordinal α, it turns out that its left inverse is the structure map of a weakly final coalgebra. If one, then, additionally assumes that $\mathbf{C}$ has a factorization system $(\mathcal{L}, \mathcal{R})$ with $\mathcal{R} \subseteq$ **Monic**, T preserves $\mathcal{R}$-morphisms, and $Z(\alpha{+}1 \to \alpha) \in \mathcal{R}$, by Theorem 4.2.3, the above result yields a final coalgebra.

Remarkably, under the same assumptions we can prove that the final sequence stabilizes after $\alpha + \alpha$ steps. Our proof is inspired by a stabilization result for final sequences of ω-accessible **Set**-functors due to Worrell [92]. Differently from [92, Theorem 4.6], our proof works out without assuming specific properties of **Set**-arrows, so that its application is no more limited to **Set**.

Lemma 4.2.9 (Left inverse is weakly final) *Let $Z\colon \mathbf{Ord}^{op} \to \mathbf{C}$ be the final sequence of T. If, for some ordinal α, $Z(\alpha{+}1 \to \alpha)$ has left-inverse $l\colon Z(\alpha) \to Z(\alpha{+}1)$, then $(Z(\alpha), l)$ is a weakly final T-coalgebra.*

Proof. Let (X,h) be a T-coalgebra. By Lemma 4.1.9, there exists a cone $(h^\beta\colon X \to Z(\beta))_{\beta \in \mathbf{Ord}}$ such that, for all ordinal β, h^β is a final projection at β, hence $h^\alpha = Z(\alpha{+}1 \to \alpha) \circ Th^\alpha \circ h$. From this we have

$$
\begin{aligned}
l \circ h^\alpha &= l \circ Z(\alpha{+}1 \to \alpha) \circ Th^\alpha \circ h && \text{(by post-composition)} \\
&= id_{Z(\alpha+1)} \circ Th^\alpha \circ h && \text{(by left inverse)} \\
&= Th^\alpha \circ h. && \text{(by identity)}
\end{aligned}
$$

Therefore h^α is a T-homomorphism from (X,h) to $(Z(\alpha), l)$. ∎

Corollary 4.2.10 *Assume the category $\mathbf{C}$ has factorization system $(\mathcal{L}, \mathcal{R})$, such that $\mathcal{R} \subseteq$ **Monic**, $T\colon \mathbf{C} \to \mathbf{C}$ preserves $\mathcal{R}$-morphisms, and $Z\colon \mathbf{Ord}^{op} \to \mathbf{C}$ be the final sequence of T. If $Z(\alpha{+}1 \to \alpha)$ is an $\mathcal{R}$-morphism and has left inverse $l\colon Z(\alpha) \to Z(\alpha{+}1)$, then the final $\mathcal{L}$-quotient at α for $(Z(\alpha), l)$ is a final T-coalgebra.*

Proof. It follows immediately from Lemma 4.2.9 and Theorem 4.2.3. ∎

If the final sequence has a left invertible arrow at some ordinal α, we have that it stabilizes at $\alpha{+}\alpha$ steps. In order to prove this, we need the following proposition.

Proposition 4.2.11 (Worrell [93], Proposition 3.4.6) *Let $T\colon \mathbf{C} \to \mathbf{C}$ be an endofunctor with final sequence $Z\colon \mathbf{Ord}^{op} \to \mathbf{C}$, $\alpha \in \mathbf{Ord}$, and $(Z(\alpha), l)$ be a T-coalgebra such that l is the left inverse of $Z(\alpha{+}1 \to \alpha)$. Then,*

$$l^{\alpha} \circ Z(\alpha{+}\alpha \to \alpha) = Z(\alpha{+}\alpha \to \alpha) \,,$$

where $(l^{\beta}\colon Z(\alpha) \to Z(\beta))_{\beta \in \mathbf{Ord}}$ is the cone over Z for $(Z(\alpha), l)$ defined as in Lemma 4.1.9.

Proof. We show by transfinite induction on $\beta \leq \alpha$ that

$$\forall \beta \leq \alpha\,. \qquad l^{\beta} \circ Z(\alpha{+}\beta \to \alpha) = Z(\alpha{+}\beta \to \beta)\,.$$

By unicity of the final arrow, $l^0 \circ Z(\alpha \to \alpha) = Z(\alpha \to 0)$. Assume, by inductive hypothesis, that $l^{\beta} \circ Z(\alpha{+}\beta \to \alpha) = Z(\alpha{+}\beta \to \beta)$, then we have

$$
\begin{aligned}
l^{\beta+1} \circ Z(\alpha{+}\beta{+}1 \to \alpha) &= Tl^{\beta} \circ l \circ Z(\alpha{+}\beta{+}1 \to \alpha) && \text{(by def. } l^{\beta+1}) \\
&= Tl^{\beta} \circ l \circ Z(\alpha{+}1 \to \alpha) \circ Z(\alpha{+}\beta{+}1 \to \alpha{+}1) && \text{(by func. } Z) \\
&= Tl^{\beta} \circ Z(\alpha{+}\beta{+}1 \to \alpha{+}1) && \text{(by left inverse)} \\
&= Tl^{\beta} \circ TZ(\alpha{+}\beta \to \alpha) && \text{(by def. } Z) \\
&= T(l^{\beta} \circ Z(\alpha{+}\beta \to \alpha)) && \text{(by func. } T) \\
&= TZ(\alpha{+}\beta \to \beta) && \text{(by inductive hp.)} \\
&= Z(\alpha{+}\beta{+}1 \to \beta{+}1)\,. && \text{(by def. } Z)
\end{aligned}
$$

Let β be a limit ordinal, and assume $l^{\gamma} \circ Z(\alpha{+}\gamma \to \alpha) = Z(\alpha{+}\gamma \to \gamma)$, for all $\gamma < \beta$. By definition of final sequence, $(Z(\beta \to \gamma))_{\gamma < \beta}$ is a limit cone over $Z{\restriction}\beta$, and since, for all $\gamma < \beta$, we have that $Z(\beta \to \gamma) \circ l^{\beta} = l^{\gamma}$, the composite $l^{\beta} \circ Z(\alpha{+}\beta \to \alpha)$ is the unique arrow such that

$$Z(\beta \to \gamma) \circ l^{\beta} \circ Z(\alpha{+}\beta \to \alpha) = l^{\gamma} \circ Z(\alpha{+}\beta \to \alpha)\,,$$

for all $\gamma < \beta$. Thus, from the following equality

$$
\begin{aligned}
Z(\beta \to \gamma) \circ Z(\alpha{+}\beta \to \beta) &= Z(\alpha{+}\beta \to \gamma) && \text{(by func. } Z) \\
&= Z(\alpha{+}\gamma \to \gamma) \circ Z(\alpha{+}\beta \to \alpha{+}\gamma) && \text{(by func. } Z) \\
&= l^{\gamma} \circ Z(\alpha{+}\gamma \to \alpha) \circ Z(\alpha{+}\beta \to \alpha{+}\gamma) && \text{(by inductive hp.)} \\
&= l^{\gamma} \circ Z(\alpha{+}\beta \to \alpha)\,, && \text{(by func. } Z)
\end{aligned}
$$

follows that $l^{\beta} \circ Z(\alpha{+}\beta \to \alpha) = Z(\alpha{+}\beta \to \beta)$. ∎

Theorem 4.2.12 *Assume the category $\mathbf{C}$ has factorization system $(\mathcal{L}, \mathcal{R})$, such that $\mathcal{R} \subseteq \mathbf{Monic}$, $T\colon \mathbf{C} \to \mathbf{C}$ preserves $\mathcal{R}$-morphisms, and $Z\colon \mathbf{Ord}^{op} \to \mathbf{C}$ be the final sequence of T. If $Z(\alpha{+}1 \to \alpha)$ is an $\mathcal{R}$-morphism with left inverse $l\colon Z(\alpha) \to Z(\alpha{+}1)$, then Z stabilizes at $\alpha{+}\alpha$.*

Proof. We show that $Z(\alpha{+}\alpha{+}1 \to \alpha)$ is an isomorphism. Assume (Q_l, q_l) be the unique final $\mathcal{L}$-quotient at α for $(Z(\alpha), l)$ defined as in Lemma 4.2.2 by the following commutative diagram, where $\rho \circ \lambda$ is the $(\mathcal{L}, \mathcal{R})$-factorization of l^{α}:

Note that, by Corollary 4.2.10, (Q_l, q_l) is a final T-coalgebra.

By Lemma 4.1.9, (Q_l, q_l) can be extended to a cone $(q_l^\beta : Q_l \to Z(\beta))_{\beta \in \mathbf{Ord}}$ over Z such that the upper square of the following diagram commutes:

$$
\begin{array}{ccc}
Q_l & \xrightarrow{\quad q_l \quad} & TQ_l \\
{\scriptstyle q_l^{\alpha+\alpha}}\downarrow & & \downarrow{\scriptstyle Tq_l^{\alpha+\alpha}} \\
Z(\alpha+\alpha) & \longleftarrow & Z(\alpha+\alpha+1) \\
\downarrow & & \downarrow \\
Z(\alpha) & \longleftarrow & Z(\alpha+1)
\end{array}
$$

The bottom square commutes by functoriality of Z, therefore $Z(\alpha+\alpha \to \alpha) \circ q_l^{\alpha+\alpha}$ is a final projection at α for (Q_l, q_l). By definition, also ρ is a final projection at α for (Q_l, q_l), so that, by Lemma 4.1.11, $Z(\alpha+\alpha \to \alpha) \circ q_l^{\alpha+\alpha} = \rho$. By Proposition 4.2.11, $l^\alpha \circ Z(\alpha+\alpha \to \alpha) = Z(\alpha+\alpha \to \alpha)$, therefore the following equalities hold (recall that $l^\alpha = \rho \circ \lambda$):

$$
Z(\alpha+\alpha \to \alpha) = \rho \circ \lambda \circ Z(\alpha+\alpha \to \alpha) \qquad \text{and} \qquad Z(\alpha+\alpha \to \alpha) \circ q_l^{\alpha+\alpha} = \rho \,.
$$

By hypothesis, $Z(\alpha+1 \to \alpha) \in \mathcal{R}$ and T preserves $\mathcal{R}$-morphisms, therefore $Z(\alpha+\alpha \to \alpha) \in \mathcal{R}$ (any right-class $\mathcal{R}$ of a factorization system is closed by transfinite composition). By definition, also $\rho \in \mathcal{R}$, so that $Z(\alpha+\alpha)$ is isomorphic to Q_l (they are $\mathcal{R}$-subobjects of each other). From this and Lambek lemma it follows that $Z(\alpha+\alpha+1 \to \alpha+\alpha)$ is an isomorphism, with inverse given by $Tq_l^{\alpha+\alpha} \circ q_l \circ \lambda \circ Z(\alpha+\alpha \to \alpha)$. $\blacksquare$

Remark 4.2.13 As already mentioned at the beginning of this section, this result is inspired by the proof of [92, Theorem 4.6], thus a comparison of the two approaches is obligatory.

In [92], Worrell showed that the final sequence of an α-accessible **Set**-functor always stabilizes at $\alpha + \alpha$ steps. In his proof, the isomorphism is found providing an injective and right invertible (hence surjective) map. A detailed look at his proof reveals that his method can be applied only if the underlying category is **Set** or, more generally, if it has a (**StrongMonic**, **Epic**)-factorization system. In the proof of Theorem 4.2.12 we do not assume the existence of right inverses for epimorphisms, so that our proof technique can be extended to categories different from **Set**. $\blacksquare$

This is a remarkably strong result, but it requires an equally strong assumption, i.e., that the final sequence reaches a left invertible arrow. In fact, to verify the existence of a left inverse is almost as difficult as to verify that the arrow is an isomorphism, unless the base category at hand has stronger properties, such as **Set** where any monomorphism with non-empty domain has right inverse.

5

Markov processes coalgebraically

In this chapter, we present a theory of generalized labelled Markov processes, that is, dynamical systems with continuos state space, interacting with the environment by means of input labels and producing measurable events by means of transitions to a measurable set of successor states. The term "generalized" is used to stress the fact that transition events can be measured by generic measures on the state space, without assuming a priori that these are of a certain type (e.g., (sub)probability measures, finite measures, or σ-finite measures). In these terms, probabilistic and stochastic Markov processes of [36] and [26] are particular instances of ours. For these dynamical structures we define a notion of bisimulation which generalizes both Larsen and Skou definition [63], given for probabilistic discrete state systems; Desharnais et al. [36], for subprobabilistic processes over analytic spaces; and Cardelli and Mardare [26, 25], for finite-rate stochastic processes over analytic spaces.

One of the two main technical contributions of the chapter is the proof that bisimilarity on generalized Markov processes is an equivalence. Equivalence for bisimilarity have been already proved by Desharnais et al. in [36], but only assuming that the state space of the Markov processes is analytic. Their proof uses very involved category theoretic constructions in order to prove that the Giry functor (actually monad) [48] on the category of analytic spaces weakly preserves (what they called) semi-pullbacks [45], and hence that these kind of limits can be lifted to the category of coalgebras for that functor. Later in [42, 43], Doberkat proposed a relatively simpler proof of the same result for stochastic relations, and thus for Markov processes, on Polish and analytic spaces. His proof followed after a deeper analysis of the semi-pullback construction of Edalat [45], and uses tools from descriptive set theory that give techniques for inverting measurable functions. None of this proves that bisimilarity of labelled Markov processes on generic measurable spaces is not an equivalence. Our proof does not need to assume that the state space is analytic, hence generalizes the result to generic measurable spaces.

The proof for the equivalence is given in terms of a characterization of the coalgebraic bisimulation in "plain" mathematical terms. This characterization will be proved to be in one-to-one correspondence with the abstract coalgebraic notion hence all the results extend to the coalgebraic setting. The main reason for switching to a different representation is that many standard techniques that are usually employed in the theory of universal coalgebras require the existence of right inverses for epimorphisms, which always exist when one is working in **Set** (assuming the axiom of choice), but are very difficult to be found if one is working in different categories, such as **Meas**, the category of measurable spaces and measurable maps. Remarkably, once one has adopted the alternative characterization, the proof of equivalence becomes extremely easy, and requires very few notions from measure theory. Thanks to this characterization we will also be able to prove that bisimulation is closed under unions.

The other main contribution of the chapter is a formal coalgebraic analysis on the relations between the notion of bisimulation and cocongruence on labelled Markov processes. In [31], Danos et al. introduced a notion alternative to that of bisimulation, the so called event bisimulation, which from the coalgebraic point of view corresponds to what we have called behavioral equivalence (see Definition 3.5.3). They proved that, when the state spaces of the Markov processes are assumed to be analytic, event bisimilarity and (state) bisimilarity coincide. Therefore, in virtue of the proof

of equivalence, it is reasonable to ask if these two concepts coincides in general, without assuming analyticity. Unfortunately, as it has been proved by Terraf in [80], this is not the case. Nevertheless, we will see that bisimilarity is contained in event bisimilarity, thus that one of the two inclusion still holds, even without assuming analyticity. The proof of this result is shown coalgebraically, establishing a formal adjunction between the category of bisimulations and that of cocongruences (actually, only a subcategory of the latter). To the best of our knowledge, also this result is new and, together with the counterexample given in [80], concludes the comparison between these two notions of equivalence between Markov processes over generic measurable spaces.

5.1   Labelled Markov kernels and bisimulation

In this section, we propose the definitions of generalized labelled Markov kernels and processes, using a notation similar to [26], and the definition of bisimulation relation between them. We prove that the induced notion of bisimilarity is indeed a bisimulation and, in particular, that it can be characterized as the union of all bisimulations, so that, it is the largest one. Then, we prove that bisimilarity over a (single) labelled Markov process is an equivalence relation. The combination of these two results allows us to give a more direct characterization of bisimilarity, similar to the one that has been proposed in [26] which was proven to be useful in proofs.

Recall from Section 2.3 the definition of the measurable space of measures: let (X, Σ) be a measurable space and $\Delta(X, \Sigma)$ be the set of all measures $\mu \colon \Sigma \to [0, \infty]$ on (X, Σ). From this set we identify four particular subclasses of measures:

probability mesures:	$\Delta_1(X, \Sigma) = \{\mu \in \Delta(X, \Sigma) \mid \mu(X) = 1\}$,
subprobability mesures:	$\Delta_{\leq 1}(X, \Sigma) = \{\mu \in \Delta(X, \Sigma) \mid \mu(X) \leq 1\}$,
finite mesures:	$\Delta_{< \infty}(X, \Sigma) = \{\mu \in \Delta(X, \Sigma) \mid \mu(X) < \infty\}$,
σ-finite mesures:	$\Delta_\sigma(X, \Sigma) = \{\mu \in \Delta(X, \Sigma) \mid \mu \text{ is } \sigma\text{-finite}\}$.

We denote by $\Delta(X, \Sigma)$, without subscript, each of the above sets (see Section 2.3 for the details about the convention on the subscripts).

For each measurable set $E \in \Sigma$, there is a canonical evaluation function $ev_E \colon \Delta(X, \Sigma) \to [0, \infty]$, defined by $ev_E(\mu) = \mu(E)$, for each measure $\mu \in \Delta(X, \Sigma)$, and called *evaluation at E*. By means of these evaluation maps, $\Delta(X, \Sigma)$ can be organized into a measurable space $(\Delta(X, \Sigma), \Sigma_{\Delta(X, \Sigma)})$, where $\Sigma_{\Delta(X, \Sigma)}$ the initial σ-algebra with respect to $\{ev_E \mid E \in \Sigma\}$, i.e., the smallest σ-algebra making ev_E measurable with respect to the Borel σ-algebra on $[0, \infty]$, for all $E \in \Sigma$.

Definition 5.1.1 (Labelled Markov kernel) *Let (X, Σ) be a measurable space and L a set of action labels. An L-labelled* Markov kernel *is a tuple $\mathcal{M} = (X, \Sigma, \{\theta_a\}_{a \in L})$ where, for all $a \in L$*

$$\theta_a \colon X \to \Delta(X, \Sigma)$$

is a measurable function, called Markov a-transition function*. An L-labelled Markov kernel $\mathcal{M}$ with a distinguished initial state $x \in X$, is said* Markov process*, and it is denoted by $(\mathcal{M}, x)$.*

The adjective "Markovian" is usually employed in the probabilistic setting; here it just indicates that the transitions depend entirely on the present state and not on the past history of the system. Interactions among processes are represented as in process algebras: the labels in L represent all possible interactions of processes with the environment. Depending on the type of measures one is dealing with, there are different interpretations for a labelled Markov transition: let $a \in L$ is a label, $x \in X$ is the current state of the system, and $E \in \Sigma$ is a measurable set of X,

(i) if $\theta_a(x) \in \Delta_1(X, \Sigma)$, then $\theta_a(x)(E)$ represents the probability of taking an a-transition from x to arbitrary elements in E;

(ii) if $\theta_a(x) \in \Delta_{\leq 1}(X, \Sigma)$, then $\theta_a(x)(E)$ is the probability of successfully taking an a-transition from x to arbitrary elements in E, and $1 - \theta_a(x)(E)$ represents the probability of terminating in some state in E from x;

(iii) if $\theta_a(x) \in \Delta_*(X, \Sigma)$, for $* \in \{``<\infty\text{''}, ``\sigma\text{''}, ``\ \text{''}\}$, then $\theta_a(x)(E)$ represents the *rate* of an exponentially distributed random variable characterizing the duration of an a-transition from x to arbitrary elements in E.

Before introducing the definition of bisimulation we need some preliminary notation.

Definition 5.1.2 ($\mathcal{R}$-closed pair) *Let $\mathcal{R} \subseteq X \times Y$ be relation on the sets X and Y, and $E \subseteq X$, $F \subseteq Y$. A pair (E, F) is $\mathcal{R}$-closed if $\mathcal{R} \cap (E \times Y) = \mathcal{R} \cap (X \times F)$.*

Lemma 5.1.3 *Let $\mathcal{R}' \subseteq \mathcal{R} \subseteq X \times Y$. If (E, F) is $\mathcal{R}$-closed, then (E, F) is also $\mathcal{R}'$-closed.*

Proof. Given $\mathcal{R} \cap (E \times Y) = \mathcal{R} \cap (X \times F)$, we prove $\mathcal{R}' \cap (E \times Y) = \mathcal{R}' \cap (X \times F)$.
($\subseteq$) Let $(x, y) \in \mathcal{R}'$ and $x \in E$. By $\mathcal{R}' \subseteq \mathcal{R}$, $(x, y) \in \mathcal{R}$. By (E, F) $\mathcal{R}$-closed, we have $y \in F$.

($\supseteq$) Let $(x, y) \in \mathcal{R}'$ and $y \in F$. By $\mathcal{R}' \subseteq \mathcal{R}$, $(x, y) \in \mathcal{R}$. By (E, F) $\mathcal{R}$-closed, we have $x \in E$. ∎

Lemma 5.1.4 *Let $\mathcal{R} \subseteq X \times X$ be an equivalence relation on X. If (E, F) is $\mathcal{R}$-closed then $E = F$, moreover E is an union of $\mathcal{R}$-equivalence classes.*

Proof. We prove only the inclusion $E \subseteq F$, the reverse is similar. Assume $x \in E$. By reflexivity of $\mathcal{R}$, $(x, x) \in \mathcal{R}$. Since (E, F) is $\mathcal{R}$-closed, we have $x \in F$. To prove that E is an union of $\mathcal{R}$-equivalence classes, it suffices to show that if $x \in E$ and $(x, y) \in \mathcal{R}$, then $y \in E$. This easily follows since $E = F$. ∎

Definition 5.1.5 (Bisimulation) *Let $\mathcal{M} = (X, \Sigma_X, \{\alpha_a\}_{a \in L})$ and $\mathcal{N} = (Y, \Sigma_Y, \{\beta_a\}_{a \in L})$ be two generalized L-labelled Markov kernels. A relation $\mathcal{R} \subseteq X \times Y$ is a* bisimulation *if, for all $(x, y) \in \mathcal{R}$, $a \in L$, and any pair $E \in \Sigma_X$ and $F \in \Sigma_Y$ such that (E, F) is $\mathcal{R}$-closed*

$$\alpha_a(x)(E) = \beta_a(y)(F).$$

Two L-labelled Markov processes $(\mathcal{M}, x)$ and $(\mathcal{N}, y)$ are bisimilar*, written $x \sim y$, if the initial states x and y are related by some bisimulation $\mathcal{R} \subseteq X \times Y$.*

The states that are related by a bisimulation $\mathcal{R}$ must agree on the values which are measured by Markov labelled transitions, for all $\mathcal{R}$-closed pairs of measurable sets of successor states. This amounts to ask that also the reachable states must be related by $\mathcal{R}$, hence that the agreement is preserved by all labelled Markov transitions. Intuitively, we can say that a bisimulation relates only those states that exhibit the same behaviour.

Remark 5.1.6 Definition 5.1.5 generalizes the Larsen and Skou [63] definition of bisimulation on discrete state probabilistic systems and, at the same time, extends both the definitions of *state bisimulation* given by Desharnais et al. in [18] for (sub)probabilistic Markov processes, and that of *rate bisimulation* proposed by Cardelli and Mardare [26] for stochastic Markov processes, which have been given only considering binary relations which are already equivalence relations[1]. Indeed, by Lemma 5.1.4, it is easy to see that their definitions coincides with Definition 5.1.5 above, in the case the relations to be considered are assumed to be equivalences. ∎

Proposition 5.1.7 (Union of bisimulations) *Let $\mathcal{F}$ be a family of bisimulation relations on $\mathcal{M} = (X, \Sigma_X, \{\alpha_a\}_{a \in L})$ and $\mathcal{N} = (Y, \Sigma_Y, \{\beta_a\}_{a \in L})$. Then $\bigcup \mathcal{F}$ is a bisimulation.*

Proof. We have to show that if $(x, y) \in \bigcup \mathcal{F}$, then for all $a \in L$, and any pair $E \in \Sigma_X$ and $F \in \Sigma_Y$ such that (E, F) is $\bigcup \mathcal{F}$-closed

$$\alpha_a(x)(E) = \beta_a(y)(F).$$

Assume $(x, y) \in \bigcup \mathcal{F}$, $a \in L$, and $E \in \Sigma_X$ and $F \in \Sigma_Y$ such that (E, F) is $\bigcup \mathcal{F}$-closed. By $(x, y) \in \bigcup \mathcal{F}$, there exists a bisimulation relation $\mathcal{R} \subseteq X \times Y$ such that $(x, y) \in \mathcal{R}$. Obviously $\mathcal{R} \subseteq \bigcup \mathcal{F}$ thus, by Lemma 5.1.3, (E, F) is $\mathcal{R}$-closed. Since $(x, y) \in \mathcal{R}$ and $\mathcal{R}$ is a bisimulation relation, we have $\alpha_a(x)(E) = \beta_a(y)(F)$. ∎

---

[1] Actually, in [31], the definition of state bisimulation is given without mentioning that the relation must be an equivalence, but without that requirement many subsequent results do not hold (e.g. Lemmas 4.1, 4.6, 4.8, Proposition 4.7, and Corollary 4.9). However, looking at the proofs it looks clear that they were imposing this condition.

Corollary 5.1.8 $\sim$ *is the largest bisimulation relation.*

Proof. By definition $\sim = \bigcup\{\mathcal{R} \subseteq X \times Y \mid \mathcal{R}$ is a bisimulation$\}$, thus, by Lemma 5.1.7 it is a bisimulation and in particular it is the largest one. ∎

Theorem 5.1.9 (Equivalence) *Let $\mathcal{M} = (X, \Sigma, \{\theta_a\}_{a \in L})$ be an L-labelled Markov kernel. Then the bisimilarity relation $\sim \subseteq X \times X$ on $\mathcal{M}$ is an equivalence.*

Proof. Symmetry is trivial: if $\mathcal{R} \subseteq X \times X$ is a bisimulation, then so is $\mathcal{R}^{-1} = \{(y,x) \mid (x,y) \in \mathcal{R}\}$.

For reflexivity, we prove that the identity relation Δ_X is a bisimulation, i.e., for all $x \in X$, $a \in A$, and measurables $E, F \in \Sigma$ such that (E, F) is Δ_X-closed,

$$\theta_a(x)(E) = \theta_a(x)(F). \tag{5.1.1}$$

Since Δ_X is and equivalence, by Lemma 5.1.4, $E = F$, therefore Equation (5.1.1) holds trivially.

It remains to prove transitivity. To this end, it suffices to show that, given $\mathcal{R}_1$ and $\mathcal{R}_2$ bisimulations on $\mathcal{M}$, there exists a bisimulation $\mathcal{R}$ on $\mathcal{M}$ that *contains* the relational composition of $\mathcal{R}_1$ and $\mathcal{R}_2$, denoted by $\mathcal{R}_1; \mathcal{R}_2 = \{(x,z) \mid (x,y) \in \mathcal{R}_1$ and $(y,z) \in \mathcal{R}_2$ for some $y \in X\}$.

Let $\mathcal{R}$ be the (unique) smallest equivalence relation containing $\mathcal{R}_1 \cup \mathcal{R}_2$. $\mathcal{R}$ can be defined as $\mathcal{R} = \Delta_X \cup \bigcup_{n \in \mathbb{N}} \mathcal{S}_n$, where

$$\mathcal{S}_0 \triangleq \mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_1^{-1} \cup \mathcal{R}_2^{-1} \qquad\qquad \mathcal{S}_{n+1} \triangleq \mathcal{S}_n; \mathcal{S}_n.$$

It is easy to see that $\mathcal{R}_1; \mathcal{R}_2 \subseteq \mathcal{R}$. We are left to show that $\mathcal{R}$ is a bisimulation. By Lemma 5.1.4, it suffices to prove that for all $a \in L$, and measurable sets $E, F \in \Sigma$ such that (E, E) is $\mathcal{R}$-closed,

$$\text{for all } (x,y) \in \mathcal{R}: \quad \theta_a(x)(E) = \theta_a(y)(E). \tag{5.1.2}$$

Now, if $(x,y) \in \mathcal{R}$, then $(x,y) \in \Delta_X$ or $(x,y) \in \mathcal{S}_n$ for some $n \geq 0$. If $(x,y) \in \Delta_X$ then $x = y$ hence Equation (5.1.2) trivially holds. We show now, by induction on $n \geq 0$, that for all $(x,y) \in \mathcal{S}_n$, Equation (5.1.2) holds.

Base case ($n = 0$): for all $(x,y) \in \mathcal{R}_j$ ($j = 1, 2$), Equation (5.1.2) holds since, by Lemma 5.1.3 and $\mathcal{R}_j \subseteq \mathcal{R}$, $(E, E) \in \Sigma(\mathcal{R}_j)$, and by the hypothesis that $\mathcal{R}_j$ is a bisimulation. For all $(x,y) \in \mathcal{R}_j^{-1}$ ($j = 1, 2$) we have that $(y, x) \in \mathcal{R}_j$, hence (5.1.2) holds too.

Inductive case ($n + 1$): for $n \geq 0$, the inductive hypothesis is given by

$$\text{for all } (x',y') \in \mathcal{S}_n: \quad \theta_a(x')(E) = \theta_a(y')(E). \tag{5.1.3}$$

Then, it is easy to see that Equation (5.1.2) holds for all $(x,y) \in \mathcal{S}_{n+1}$: by definition, there exists some $z \in X$ such that $(x, z) \in \mathcal{S}_n$ and $(z, y) \in \mathcal{S}_n$. Hence, applying Equation (5.1.3) twice, we get $\theta_a(x)(E) = \theta_a(z)(E) = \theta_a(y)(E)$. ∎

Remark 5.1.10 Remarkably, the proof above does not need any specific property or results from the theory of measures and measurable spaces; it directly follows by the definition of bisimulation (Definition 5.1.1) and $\mathcal{R}$-closed pair (Definition 5.1.2).

As a technical remark, note that, in Theorem 5.1.9, transitivity is verified adopting a strategy that avoids to prove that bisimulation relations are closed under composition. A reason for avoiding it is that this would have required that (semi-)pullbacks of relations in **Meas** are weakly preserved by the behaviour functor $\Delta: \mathbf{Meas} \to \mathbf{Meas}$ (see Section 2.3). Recently, in [80] Terraf showed that this is not the case. The proof of this result is based on the existence of a non-Lebesgue-measurable set V in the open unit interval $(0, 1)$, which is used to define two measures on the σ-algebra extended with V such that they differ in this set[2]. In the light of this, the simplicity of the proof of Theorem 5.1.9 is even more remarkable. ∎

---

[2]It must be noticed that the counterexample given in [80, Theorem 12] requires one to assume that all zig-zag morphisms (hence, Δ_L^L-homomorphisms) are surjective. Such an assumption is not generally considered valid, so that one may argue that this is not yet the definitive counterexample.

From Theorem 5.1.9 and Corollary 5.1.8 we have the following characterization of bisimilarity:

Proposition 5.1.11 *Let $\mathcal{M} = (X, \Sigma, \{\theta_a\}_{a \in L})$ be an L-labelled Markov kernel, then, for $x, y \in X$:*

$$x \sim y \quad \Longleftrightarrow \quad \text{for all } a \in L \text{ and } E \in \Sigma \text{ such that } (E, E) \text{ is } \sim\text{-closed.} \quad \theta_a(x)(E) = \theta_a(y)(E),$$

Proof. The implication from left to right is an immediate consequence of the fact that $\sim$ is an equivalence relation (Theorem 5.1.9) and that $\sim$ is a bisimulation (Corollary 5.1.8). We are left to prove the implication from right to left. To this end, assume $x, y \in X$ have the following property:

$$\text{for all } a \in L \text{ and } E \in \Sigma \text{ such that } (E, E) \text{ is } \sim\text{-closed}, \quad \theta_a(x)(E) = \theta_a(y)(E). \qquad (5.1.4)$$

We prove that $x \sim y$ showing a bisimulation $\mathcal{R}$ such that $(x, y) \in \mathcal{R}$. Let $\mathcal{R}$ be the smallest equivalence relation containing $\{(x, y)\}$ and $\sim$. This can be defined as $\mathcal{R} = \Delta_X \cup \bigcup_{n \in \mathbb{N}} \mathcal{S}_n$, where

$$\mathcal{S}_0 \triangleq \{(x, y), (y, x)\} \cup \sim \qquad\qquad \mathcal{S}_{n+1} \triangleq \mathcal{S}_n ; \mathcal{S}_n.$$

(";" denotes relation composition). By Lemma 5.1.4, it suffices to prove that for all $a \in L$, and $E' \in \Sigma$ such that (E', E') is $\mathcal{R}$-closed

$$\text{for all } (x', y') \in \mathcal{R} : \quad \theta_a(x')(E') = \theta_a(y')(E'). \qquad (5.1.5)$$

If $(x', y') \in \mathcal{R}$, then $(x', y') \in \Delta_X$ or $(x', y') \in \mathcal{S}_n$ for some $n \geq 0$. If $(x', y') \in \Delta_X$ then $x' = y'$, therefore Equation (5.1.5) holds trivially. We show, by induction on $n \geq 0$, that for all $(x', y') \in \mathcal{S}_n$, Equation (5.1.5) holds.

Base case ($n = 0$): assume $(x', y') \in \sim$. Since, $\sim \subseteq \mathcal{R}$, by Lemma 5.1.3, (E', E') is $\sim$-closed. Thus, Equation (5.1.5) holds since, by Corollary 5.1.8, $\sim$ is a bisimulation relation. If $x' = x$ (resp. $x' = y$) and $y' = y$ (resp. $y' = x$), then property (5.1.4) holds. Again, by Lemma 5.1.3 and $\sim \subseteq \mathcal{R}$, we have that (E', E') is $\sim$-closed, thus Equation (5.1.5) holds trivially.

Inductive case ($n + 1$): for $n \geq 0$, the inductive hypothesis is as follows

$$\text{for all } (x'', y'') \in \mathcal{S}_n : \quad \theta_a(x'')(E') = \theta_a(y'')(E'). \qquad (5.1.6)$$

Then, it is easy to see that Equation (5.1.5) holds for all $(x', y') \in \mathcal{S}_{n+1}$. Indeed, by definition, there exists some $z \in X$ such that $(x', z) \in \mathcal{S}_n$ and $(z, y') \in \mathcal{S}_n$, hence, applying Equation (5.1.6) twice, we have $\theta_a(x')(E') = \theta_a(z)(E') = \theta_a(y')(E')$. ∎

5.2 Characterization of the coalgebraic bisimulation

In this section we prove that the definition of bisimilarity we have used so far coincides with the abstract coalgebraic notion of bisimilarity. To avoid confusion between the two notions, we will refer to the one given in Definition 5.1.5 as *state bisimulation* and to that in Definition 3.3.4 as *coalgebraic bisimulation*. We show that any coalgebraic bisimulation induces a state bisimulation and vice versa, hence the two notion of bisimilarity coincide. This correspondence is easy to determine when the state bisimulation is assumed to be an equivalence, but is no more easy when one consider generic relations $R \subseteq X \times Y$ over different sets. A similar result has been already proposed by de Vink and Rutten in [33, 34] for probabilistic transitions system over the category of ultrametric spaces and non-expansive maps, but the correspondence was only found under the assumption that the relation R has a Borel decomposition. The approach we use here do not assume any extra condition on the relation and it can be used also to drop the assumption in [34, Theorem 5.8].

Remark 5.2.1 Unfortunately, we were not able to extend this correspondence to all the types of Markov kernels we have considered so far. Indeed, at a certain point of the proof we have to assume that the measures are finite. Nevertheless, we provide all results at the maximum level of generality we were able to achieve, in the hope that someone can extend the proof and complete the correspondence. ∎

In order to model generalized Markov kernels as coalgebras one needs a suitable category and a suitable functor. The most natural choice for a category is **Meas**, the category of measurable spaces and measurable functions, and for a functor is $\Delta\colon \mathbf{Meas} \to \mathbf{Meas}$ (see Section 2.3).

Proposition 5.2.2 *Generalized L-labelled Markov kernels are exactly the Δ^L-coalgebras on* **Meas**.

Proof. The correspondence is trivial. Given a generalized Markov kernel $\mathcal{M} = (X, \Sigma_X, \{\theta_a\}_{a \in L})$ we define a Δ^L-coalgebra (X, α) as follows, for $a \in L$

$$\alpha\colon X \to \Delta^L X \qquad\qquad\qquad \alpha(a) = \theta_a$$

A map $\alpha\colon X \to \Delta^L X$ is measurable iff $ev_E \circ ev_a \circ \alpha$ is measurable, for all $a \in L$ and $E \in \Sigma_X$. But $ev_a \circ \alpha = \theta_a$ is measurable by definition of Markov kernel, hence α is so. Conversely, given an Δ^L-coalgebra (X, α), we define a Markov kenel $\mathcal{M} = (X, \Sigma_X, \{\theta_a\}a \in L)$ as follows, where for all $a \in L$, $\theta_a = ev_a \circ \alpha = \alpha(\cdot)(a)$. Measurability of θ_a follows since it is the composite of measurable functions. It is immediate to see that the two translations are inverses of each other.   ∎

Due to this correspondence, we will make no distinction between Δ^L-coalgebras and L-labelled Markov kernels, and the translation from one model to the other will be used without reference.

Remark 5.2.3 The result above is very well-known for the case of the Giry functor Δ_1 and its subprobability variant $\Delta_{\leq 1}$. Still we recall the proof to convince the reader that the result can extended to all the other variants of measure functors we have introduced.   ∎

The coalgebraic definition of bisimulation between coalgebras is given in terms of a monic spans between the carriers, such that there exist coalgebra structures that make them actual cospans between coalgebras (cf. Definition 3.3.4). In categories with binary products monic spans (R, f, g) are in one-to-one correspondence with monic arrows $R \to X \times Y$. Thus, without loss of generality, we restrict our attention only to relations $R \subseteq X \times Y$ with measurable canonical projections $\pi_X\colon R \to X$ and $\pi_Y\colon R \to Y$. This will be convenient especially to make a comparison with the results in Section 5.1.

Proposition 5.2.4 *Let (R, π_X, π_Y) be a Δ^L-bisimulation between the coalgebras (X, α) and (Y, β). Then, R is a state bisimulation.*

Proof. We prove that $R \subseteq X \times Y$ is a state bisimulation between the Markov kernels $\mathcal{M} = (X, \Sigma_X, \{\alpha(\cdot)(a)\}_{a \in L})$ and $\mathcal{N} = (Y, \Sigma_Y, \{\beta(\cdot)(a)\}_{a \in L})$. Thus, we have to show that for all $(x, y) \in R$, $a \in L$ and $E \in \Sigma_X$, $F \in \Sigma_Y$ such that (E, F) is R-closed, the following equality holds

$$\alpha(x)(a)(E) = \beta(y)(a)(F)\,.$$

Notice first that, $\pi_X^{-1}(E) = (E \times Y) \cap R$ and $\pi_Y^{-1}(F) = (X \times F) \cap R$, so that, for R-closed pairs (E, F) it holds that $\pi_X^{-1}(E) = \pi_Y^{-1}(F)$. Since (R, π_X, π_Y) is a Δ^L-bisimulation, there exists a coalgebraic structure $\gamma\colon R \to \Delta^L R$ on R making the following diagram commute

$$
\begin{array}{ccccc}
X & \xleftarrow{\;\pi_X\;} & R & \xrightarrow{\;\pi_Y\;} & Y \\
{\scriptstyle\alpha}\downarrow & & {\scriptstyle\gamma}\downarrow & & \downarrow{\scriptstyle\beta} \\
\Delta^L X & \xleftarrow[\Delta^L \pi_X]{} & \Delta^L R & \xrightarrow[\Delta^L \pi_Y]{} & \Delta^L Y
\end{array}
$$

From the commutativity of the diagram above, we have that

$$
\begin{aligned}
\alpha(x)(a)(E) &= \alpha(\pi_X(x, y))(a)(E) & \text{(by def. } \pi_X) \\
&= (\alpha \circ \pi_X)(x, y)(a)(E) & \text{(composition)} \\
&= (\Delta^L \pi_X \circ \gamma)(x, y)(a)(E) & \text{(by } \Delta^L\text{-homomorphism)}
\end{aligned}
$$

$$\begin{aligned}
&= \Delta \pi_X(\gamma(x,y)(a))(E) && \text{(by def. } Id^L) \\
&= \gamma(x,y)(a) \circ \pi_X^{-1}(E) && \text{(by def. } \Delta) \\
&= \gamma(x,y)(a) \circ \pi_Y^{-1}(F) && \text{(by } (E,F) \text{ } R\text{-closed)} \\
&= \Delta \pi_Y(\gamma(x,y)(a))(F) && \text{(by def. } \Delta) \\
&= (\Delta^L \pi_Y \circ \gamma)(x,y)(a)(F) && \text{(by def. } Id^L) \\
&= (\beta \circ \pi_Y)(x,y)(a)(F) && \text{(by } \Delta^L\text{-homomorphism)} \\
&= \beta(y)(a)(F) && \text{(by def. } \pi_Y)
\end{aligned}$$

∎

For the other half of the correspondence, i.e., that any state bisimulation is a Δ^L-bisimulation, we need some preliminary work that involves results and techniques from measure theory.

First, notice that in Definition 5.1.5, no σ-algebra is assigned to the relation $R \subseteq X \times Y$, so that in order to make it an object in **Meas** we have to provide it one. This, moreover, must be defined such that the canonical projections will be rendered measurable. The most natural choice is the initial σ-algebra w.r.t. $\pi_X \colon R \to X$ and $\pi_Y \colon R \to Y$, that is, the smallest one making the two projections measurable. This is generated by the following family of sets

$$\mathcal{F} = \{(E \times F) \cap R \mid E \in \Sigma_X \text{ and } F \in \Sigma_Y\}.$$

Second, we have to provide a measurable Δ^L-coalgebra structure on R such that makes the canonical projections coalgebra morphisms. To do so, we will use the following result.

Proposition 5.2.5 *Let (X, Σ_X) and (Y, Σ_Y) be measurable spaces, $R \subseteq X \times Y$, and Σ_R denote the σ-algebra generated by the collection of all subsets of the form, $(E \times F) \cap R$, for $E \in \Sigma_X$ and $Y \in \Sigma_Y$. Then, for any measure $\mu \colon \Sigma_X \to [0, \infty]$, there exists a measure $\tilde{\mu} \colon \Sigma_R \to [0, \infty]$ such that*

$$\tilde{\mu}((E \times F) \cap R) = \mu(E), \qquad \text{for all } E \in \Sigma_X \text{ and } Y \in \Sigma_Y.$$

Moreover, if μ is σ-finite, $\tilde{\mu}$ is unique.

Proof. We define $\tilde{\mu}$ has the Hahn-Kolmogorov extension (Theorem 2.2.26) of a pre-measure $\tilde{\mu}_0$ defined on a suitable boolean algebra $\mathcal{A}$ such that $\sigma(\mathcal{A}) = \Sigma_R$.

Let $\mathcal{A}$ be the collection of all *finite* unions $\bigcup_{i=0}^{k}(E_i \times F_i) \cap R$ such that $k \in \mathbb{N}$ and, for all $0 \le i \le k$, $E_i \in \Sigma_X$ and $F_i \in \Sigma_Y$. Certainly, $\sigma(\mathcal{A}) = \Sigma_R$. To prove that $\mathcal{A}$ is a boolean algebra on R, is suffices only prove that is closed under complements (it is already closed under finite union). This is immediate by the following equality:

$$R \setminus ((E \times F) \cap R) = ((X \setminus E \times F) \cap R) \cup ((E \times Y \setminus F) \cap R).$$

Now we define $\tilde{\mu}_0 \colon \mathcal{A} \to [0, \infty]$. Note that any set $S \in \mathcal{A}$ can always decomposed into a finite union of pair-wise disjoint sets $S = \bigcup_{i=0}^{k}(E_i \times F_i) \cap R$. We then define the quantity $\tilde{\mu}_0(S)$ associated to such disjoint union S by

$$\tilde{\mu}_0(S) := \textstyle\sum_{i=0}^{k} \mu(E_i) \tag{5.2.1}$$

It is easy to show that this definition does not depend on how S is decomposed into a disjoint union. To see this, note that any two representations $\bigcup_{i=0}^{k}(E_i \times F_i) \cap R$ and $\bigcup_{i=0}^{k'}(E_i' \times F_i') \cap R$ of the same set can be decomposed into a common refinement $\bigcup_{i=0}^{k''}(E_i'' \times F_i'') \cap R$, so that, by the well definition of μ they must agree on it. Moreover, by construction, $\tilde{\mu}_0$ is finitely additive. It remains to show that, if $S \in \mathcal{A}$ is the countable disjoint union of sets $S_0, S_1, S_2, \ldots \in \mathcal{A}$, then

$$\tilde{\mu}_0(S) = \textstyle\sum_{n \in \mathbb{N}} \tilde{\mu}_0(S_n).$$

Slitting up S into disjoint product sets, and restricting S_n to each of these product sets in turn, for all $n \in \mathbb{N}$, by finite additivity of $\tilde{\mu}_0$, we may assume without loss of generality that $S = (E \times F) \cap R$,

for some $E \in \Sigma_X$ and $F \in \Sigma_Y$. In the same way, by breaking up each S_n into a component product sets and using finite additivity of $\tilde{\mu}_0$ again, we may assume without loss of generality that each S_n takes the form $S_n = (E_n \times F_n) \cap R$, for some $E_n \in \Sigma_X$ and $F \in \Sigma_Y$. By definition of $\tilde{\mu}_0$, Equation (5.2.1) is rewritten as

$$\mu(E) = \sum_{n \in \mathbb{N}} \mu(E_n).$$

This holds trivially by σ-additivity of μ.

Now that we have proven that $\tilde{\mu}_0 \colon \mathcal{A} \to [0, \infty]$ is a pre-measure, we can define $\tilde{\mu} \colon \Sigma_R \to [0, \infty]$ as the Hahn-Kolmogorov extension of of $\tilde{\mu}_0$, and since, $((E \times F) \cap R) \in \mathcal{A}$, for all $E \in \Sigma_X$ and $F \in \Sigma_Y$, we have

$$\tilde{\mu}((E \times F) \cap R) = \tilde{\mu}_0((E \times F) \cap R) = \mu(E).$$

Thus the required condition is satisfied.   ∎

Proposition 5.2.6 *Let $R \subseteq X \times Y$ be a state bisimulation between the (finite) L-labelled Markov kernels $(X, \Sigma_X, \{\alpha_a\}_{a \in L})$ and $(Y, \Sigma_Y, \{\beta_a\}_{a \in L})$. Then (R, π_X, π_Y) is a $\Delta_{<\infty}^L$-bisimulation.*

Proof. We want to prove that (R, π_X, π_Y) is a $\Delta_{<\infty}^L$-bisimulation between (X, α) and (Y, β), where $\alpha(a) = \alpha_a$ and $\beta(a) = \beta_a$, for all $a \in L$. To this end we have to equip R with σ-algebra Σ_R such that $\pi_X \colon R \to X$ and $\pi_Y \colon R \to Y$ are measurable, and provide a measurable coalgebra structure $\gamma \colon R \to \Delta_{<\infty}^L R$ making the following diagram commutes

$$
\begin{array}{ccccc}
X & \xleftarrow{\ \pi_X\ } & R & \xrightarrow{\ \pi_Y\ } & Y \\
{\scriptstyle \alpha}\downarrow & & {\scriptstyle \gamma}\downarrow & & \downarrow{\scriptstyle \beta} \\
\Delta_{<\infty}^L X & \xleftarrow[\Delta_{<\infty}^L \pi_X]{} & \Delta_{<\infty}^L R & \xrightarrow[\Delta_{<\infty}^L \pi_Y]{} & \Delta_{<\infty}^L Y
\end{array}
\qquad (5.2.2)
$$

As for Σ_R we take the initial σ-algebra w.r.t. $\pi_X \colon R \to X$ and $\pi_Y \colon R \to Y$, so that both the projections are measurable. Σ_R can also be characterized as the σ-algebra generated by the collection of all finite unions of subsets of the form $(E \times F) \cap R$, for $E \in \Sigma_X$ and $F \in \Sigma_Y$. Let $\gamma \colon R \to \Delta_{<\infty}^L R$ be defined as $\gamma((x, y))(a) = \tilde{\alpha_a}(x)$, for all $a \in L$, $(x, y) \in R$, where $\tilde{\alpha_a}$ is given by Lemma 5.2.5. Therefore, $\gamma((x, y))(a) \colon \Sigma_R \to [0, \infty)$ is the (unique) measure on (R, Σ_R), such that

$$\gamma((x, y))(a)((E \times F) \cap R) = \alpha(x)(a)(E), \qquad \text{for all } E \in \Sigma_X \text{ and } Y \in \Sigma_Y.$$

To prove that γ is measurable, by Lemmas 2.2.8 and 2.3.3 it suffices to show that for any finite union of the form $S = \bigcup_{i=0}^k ((E_i \times F_i) \cap R)$, where $E_i \in \Sigma_X$ and $F_i \in \Sigma_Y$, for $0 \leq i \leq k$, $(\gamma(\cdot)(a))^{-1}(L_r(S)) \in \Sigma_R$. We may assume, without loss of generality, that $S = \bigcup_{i=0}^k ((E_i \times F_i) \cap R)$ is given as a disjoint union (otherwise we may represent it as a disjoint one taking a the disjoint refinements for the sets $(E_i \times F_i) \cap R$).

$$
\begin{aligned}
(\gamma(\cdot)(a))^{-1}(L_r(S)) &= \{(x, y) \in R \mid \gamma((x, y))(a) \in L_r(S)\} & \text{(by inverse image)} \\
&= \{(x, y) \in R \mid \gamma((x, y))(a)(S) \geq r\} & \text{(by def. } L_r(E)) \\
&= \{(x, y) \in R \mid \sum_{i=0}^k \gamma((x, y))(a)((E_i \times F_i) \cap R) \geq r\} & \text{(by finite additivity)} \\
&= \{(x, y) \in R \mid \sum_{i=0}^k \alpha(x)(a)(E_i) \geq r\} & \text{(by def. } \gamma) \\
&= \{(x, y) \in R \mid \alpha(x)(a)(\bigcup_{i=0}^k E_i) \geq r\} & \text{(by finite additivity)} \\
&= \{(x, y) \in R \mid \alpha(x)(a) \in L_r(\bigcup_{i=0}^k E_i)\} & \text{(by def. } L_r(E)) \\
&= ((\alpha(\cdot)(a))^{-1}(\bigcup_{i=0}^k E_i) \times Y) \cap R & \text{(by inverse image)}
\end{aligned}
$$

Since α is measurable, $(\alpha(\cdot)(a))^{-1}(\bigcup_{i=0}^k E_i) \in \Sigma_X$, so that $(\gamma(\cdot)(a))^{-1}(L_r(S)) \in \Sigma_R$.

Now we prove that Diagram (5.2.2) commutes. To this end, consider the following two sets: $R(\cdot, F) = \{x \in X \mid (x, y) \in R \text{ and } y \in F\}$ and $R(E, \cdot) = \{y \in Y \mid (x, y) \in R \text{ and } x \in E\}$, for $E \in \Sigma_X$ and $F \in \Sigma_Y$. $R(\cdot, F)$ and $R(E, \cdot)$ are measurable, since following equalities hold

$$\pi_X^{-1}(E) = (E \times Y) \cap R = (X \times R(E, \cdot)) \cap R = \pi_Y^{-1}(R(E, \cdot)), \qquad (5.2.3)$$

$$\pi_Y^{-1}(F) = (X \times F) \cap R = (R(\cdot, F) \times Y) \cap R = \pi_X^{-1}(R(\cdot, F)). \qquad (5.2.4)$$

In particular, we have also that the pairs $(E, R(E, \cdot))$ and $(R(\cdot, F), F)$ are R-closed, so that, since R is a state bisimulation, for any $(x, y) \in R$, $a \in L$, $E \in \Sigma_X$, and $F \in \Sigma_Y$ the following hold

$$\alpha(x)(a)(E) = \beta(y)(a)(R(E, \cdot)), \qquad \alpha(x)(a)(R(\cdot, F)) = \beta(y)(a)(F). \qquad (5.2.5)$$

With this in mind, we prove the commutativity of the right square of Diagram (5.2.2):

$$
\begin{aligned}
(\Delta_{<\infty}^L \pi_Y \circ \gamma)((x, y))(a)(F) &= \gamma((x, y))(a)(\pi_Y^{-1}(F)) && \text{(by def. } \Delta_{<\infty}^L) \\
&= \gamma((x, y))(a)((R(\cdot, F) \times Y) \cap R) && \text{(by (5.2.4))} \\
&= \alpha(x)(a)(R(\cdot, F)) && \text{(by def. } \gamma) \\
&= \beta(y)(a)(F) && \text{(by (5.2.5))} \\
&= (\beta \circ \pi_Y)((x, y))(a)(F) && \text{(by def. } \pi_Y)
\end{aligned}
$$

for all $(x, y) \in R$, $a \in A$, and $F \in \Sigma_Y$. Commutativity of the left square of Diagram (5.2.2) is trivial and follows even without using Equation (5.2.5). ∎

Remark 5.2.7 In the proof of Proposition 5.2.6 we have to impose that the measures are finite due to the use of Lemma 2.3.3. However, all the other constructions follow without any extra condition. Interestingly, if it were possible to extend Lemma 2.3.3 to generic measures, the coalgebraic structure imposed on R may be not unique. Indeed, if the coalgebra structures α and β, are such that, $\alpha(x)(a)$ and $\beta(y)(a)$ are non-σ-finite, for some $(x, y) \in R$ and $a \in L$, the extension provided by Proposition 5.2.5 is not necessarily unique. However, it must be noticed that is situation is not unusual for coalgebraic bisimulations (see Example 3.3.6). ∎

Remark 5.2.8 (Ultrametric spaces) The coalgebraic treatment of continuous probabilistic systems originated in the work of de Vink and Rutten [33, 34] on the category of ultrametric spaces and nonexpansive maps. They were the first to compare the notion of probabilistic bisimulation of Larsen and Skou [63] with the abstract coalgebraic definition of Aczel and Mendler. The comparison between the two notions was fully achieved only in the discrete case. For the continuous case, they were able to establish the correspondence only under the assumption that the bisimulation relation has a Borel decomposition.

The construction given in Proposition 5.2.5 can be applied also to Borel measures, so that the proof-strategy of Proposition 5.2.6 do not need can be employed to prove [34, Theorem 5.8] with the benefits of dropping the assumption of the existence of Borel decompositions. ∎

Theorem 5.2.9 (Characterization) *State bisimulation and $\Delta_{<\infty}$-bisimulation coincide.*

Proof. It follows directly by Propositions 5.2.4 and 5.2.6. ∎

5.3 Relating Cocongruences and Bisimulations

In this section, we consider the problem of relating bisimulation and cocongruences between L-labelled Markov kernels with finite measures, hence $\Delta_{<\infty}^L$-coalgebras. We will show sufficient conditions under which we get a bijection between these two notions. This is done setting up an adjunction between the category of bisimulations and (a subcategory) cocongruences between two fixed coalgebras. This bijection is based on a standard adjunction occurring between span and cospans in categories with pushouts and pullbacks. Our aim is to lift this adjunction to the

categories of bisimulations and cocongruences for the functor $\Delta^L_{<\infty}$. This lifting is very simple to be implemented if the behaviour functor preserves weak pullbacks. However, our case is not so lucky, since $\Delta^L_{<\infty}$ does not enjoy such a property (see Viglizzo [89]). Nevertheless, we show that under some assumptions (i.e., the restriction to a suitable subcategory of cocongruences) this adjunction can still be lifted.

For sake of clarity we first recall the bijection in the simplified setting of monic span and epic cospan in a generic category $\mathbf{C}$, then, we show how to lift the construction to bisimulations and cocongruences for the functor $\Delta^L_{<\infty}\colon \mathbf{Meas} \to \mathbf{Meas}$.

The category $\mathbf{MSpan}_{\mathbf{C}}(X,Y)$ has as objects monic spans (R,f,g) between X and Y in $\mathbf{C}$, and arrows $f\colon (R,r_1,r_2) \to (S,s_1,s_2)$ which are morphisms $f\colon R \to S$ in $\mathbf{C}$ making the diagrams below commute:

$$
\begin{array}{ccc}
R \xrightarrow{\ f\ } S & \qquad & R \xrightarrow{\ f\ } S \\
{}_{r_1}\searrow \ \ \swarrow_{s_1} & & {}_{r_2}\searrow \ \ \swarrow_{s_2} \\
X & & Y
\end{array}
$$

Since the legs of each span in $\mathbf{MSpan}_{\mathbf{C}}(X,Y)$ are jointly monic, we have that, given any two objects, there is only one morphism between them. As a consequence, if we have that for two monic spans (R,r_1,r_2) and (S,s_1,s_2) there exist morphisms $f\colon (R,r_1,r_2) \to (S,s_1,s_2)$ and $g\colon (S,s_1,s_2) \to (R,r_1,r_2)$, then they are isomorphic, with isomorphism given by f and g. Indeed, by uniqueness, the composites $f \circ g$ and $g \circ f$ must be equal to id_S and id_R, respectively.

The category $\mathbf{ECospan}_{\mathbf{C}}(X,Y)$, of epic cospans between X and Y in $\mathbf{C}$, is defined analogously, and enjoys the same properties we have discussed above in the case of monic spans. In the following, we will often omit the subscript $\mathbf{C}$ when the category of reference is understood.

Assume the category $\mathbf{C}$ has pullbacks and pushouts. Then, given any cospan we can take the pullback over it producing a span and, conversely, any span yields a cospan via its pushouts. Formally, we can define two functors: $Pb_{(X,Y)}\colon \mathbf{ECospan}(X,Y) \to \mathbf{MSpan}(X,Y)$ mapping each epic span in $\mathbf{MSpan}(X,Y)$ to its pullback, and $Po_{(X,Y)}\colon \mathbf{MSpan}(X,Y) \to \mathbf{ECospan}(X,Y)$ mapping each monic span in $\mathbf{MSpan}(X,Y)$ to its pushout. These operations are well defined, indeed, spans deriving from pullbacks are always monic, and cospans deriving from pushouts are always epic. As for morphisms, let $f\colon (R,r_1,r_2) \to (S,s_1,s_2)$ be a morphism in $\mathbf{MSpan}(X,Y)$, then $Po_{(X,Y)}(f)$ is defined as the unique arrow, given by the universal property of pushout, making the following diagram commute

where $Po_{(X,Y)}(R,r_1,r_2) = (R',r_1',r_2')$ and $Po_{(X,Y)}(S,s_1,s_2) = (S',s_1',s_2')$. The action on arrows for $Pb_{(X,Y)}$ is defined similarly, using the universal property of pullbacks. Functoriality of $Pb_{(X,Y)}$ and $Po_{(X,Y)}$ follows by the universal properties of pullbacks and pushouts, respectively. When the domain (X,Y) of the spans and cospans in $\mathbf{MSpan}(X,Y)$ and $\mathbf{ECospan}(X,Y)$, respectively, are clear from the context, the subscript in $Po_{(X,Y)}$ and $Pb_{(X,Y)}$ will be omitted.

For these two functors we have the following properties.

Lemma 5.3.1 *Let $\mathbf{C}$ be a category with pushouts and pullbacks. Then, the following hold*

  (i) $Po\,Pb\,Po \cong Po$;

(ii) $PbPoPb \cong Pb$.

Proof. Both (i) and (ii) follow by the universal properties of pullbacks and pushouts. We show (i), (ii) is similar. We need to show that there are inverse natural transformations $f: Po \Rightarrow PoPbPo$ and $g: PoPbPo \Rightarrow Po$. We define them component-wise as follows. Let (R, r_1, r_2) be an object in **MSpan**(X, Y), and consider the diagram below
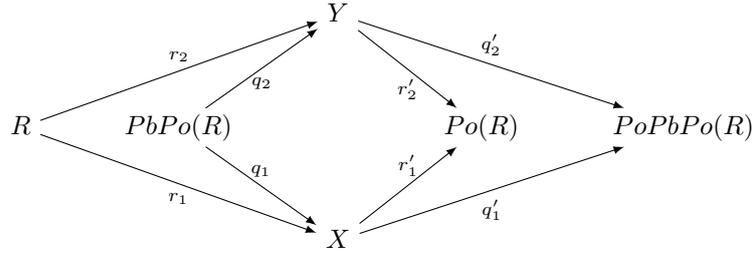
where

$$Po(R, r_1, r_2) = (Po(R), r_1', r_2') \qquad \text{hence} \qquad r_1' \circ r_1 = r_2' \circ r_2 \qquad (5.3.1)$$

$$PbPo(R, r_1, r_2) = (PbPo(R), q_1, q_2) \qquad \text{hence} \qquad r_1' \circ q_1 = r_2' \circ q_2 \qquad (5.3.2)$$

$$PoPbPo(R, r_1, r_2) = (PoPbPo(R), q_1', q_2') \qquad \text{hence} \qquad q_1' \circ q_1 = q_2' \circ q_2 \qquad (5.3.3)$$

By the universal property of pullbacks, from (5.3.1) and (5.3.2) there exists $k: R \to PbPo(R)$ such that $q_1 \circ k = r_1$ and $q_2 \circ k = r_2$. From these two equalities we have that

$$\begin{aligned} q_1' \circ r_1 &= q_1' \circ q_1 \circ k \\ &= q_2' \circ q_2 \circ k \qquad \qquad \text{(by (5.3.3))} \\ &= q_2' \circ r_2 \end{aligned}$$

From this and (5.3.3), by the universal property of pushouts, there exists $f_R: Po(R) \to PoPbPo(R)$ such that $f_R \circ r_1' = q_2'$ and $f_R \circ r_2' = q_2'$. Now, by the universal property of pushouts, from (5.3.2) and (5.3.3), there exists $g_R: PoPbPo(R) \to Po(R)$ such that $g_R \circ q_1' = r_1'$ and $g_R \circ q_2' = r_2'$. Since both f_R and g_R are opposite morphisms of epic cospans, they are necessarily inverses of each other, thus $Po(R, r_1, r_2)$ and $PoPbPo(R, r_1, r_2)$ are isomorphic. Naturality of f and g follows by the universal property of pushouts. ∎

Proposition 5.3.2 *Let* **C** *be a category with pushouts and pullbacks, then* $Po \dashv Pb$.

Proof. Due to the universal properties of pushouts and pullbacks, for any pair of objects (R, r_1, r_2) in **MSpan**(X, Y) and (S, s_1, s_2) in **ECospan**(X, Y), it holds that,

$$Hom(Po(R, r_1, r_2), (S, s_1, s_2)) \cong Hom((R, r_1, r_2), Pb(S, s_1, s_2)),$$

that is, Po is left adjoint to Pb. ∎

The unit $\eta: Id \Rightarrow PbPo$ and counit $\epsilon: PoPb \Rightarrow Id$ of the adjunction $Po \dashv Pb$, are defined component-wise as follows, for all objects (R, r_1, r_2) in **MSpan**(X, Y) and (K, k_1, k_2) in **ECospan**(X, Y).
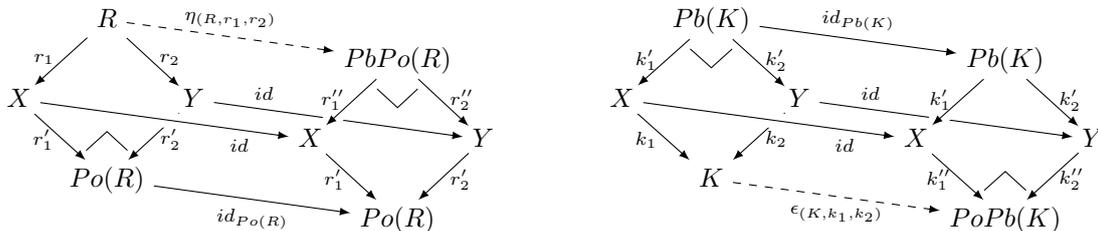
As any adjunction, $Po \dashv Pb$ gives rise to a monad $(PbPo, \eta, Pb\epsilon Po)$ in $\mathbf{MSpan}(X, Y)$ and a comonad $(PoPb, \epsilon, Po\eta Pb)$ in $\mathbf{ECospan}(X, Y)$, which, by Lemma 5.3.1, are idempotent.

If the categories $\mathbf{MSpan}(X, Y)$ and $\mathbf{ECospan}(X, Y)$ are quotiented by isomorphism, they can be considered as partial orders, so that the adjunction is a Galois connection, and the monad $(PbPo, \eta, Pb\epsilon Po)$ and comonad $(PoPb, \epsilon, Po\eta Pb)$, respectively, defines a *closure operator* on monic spans and an *interior operator* on epic cospans.

5.3.1   The comonad $PoPb$ and monad $PbPo$ in Meas

In this section, we give an explicit characterization of the monad $PbPo$ in $\mathbf{MSpan}_{\mathbf{Meas}}(X, Y)$ and a comonad $PoPb$ in $\mathbf{ECospan}_{\mathbf{Meas}}(X, Y)$, for generic objects X and Y in $\mathbf{Meas}$. This will also serve as a preliminary step for determining the adjunction between $\Delta^L_{<\infty}$-bisimulations and (a subclass of) $\Delta^L_{<\infty}$-cocongruences.

Since $\mathbf{Meas}$ has pullbacks and pushouts, the functors $Pb \colon \mathbf{ECospan}(X, Y) \to \mathbf{MSpan}(X, Y)$ and $Po \colon \mathbf{MSpan}(X, Y) \to \mathbf{ECospan}(X, Y)$ are well defined, for any pair of objects X and Y. Moreover, since $\mathbf{Meas}$ has both binary products and coproducts, we can identify the categories $\mathbf{MSpan}(X, Y)$ and $\mathbf{ECospan}(X, Y)$, respectively, as the categories of relations $R \subseteq X \times Y$ (with measurable canonical projections) and quotients $(X+Y)/E$ (with measurable canonical injections), where E is an equivalence relation on $(X + Y)$.

To give an explicit explicit characterization of the monad $(PbPo, \eta, Pb\epsilon Po)$ and comonad $(PoPb, \epsilon, Po\eta Pb)$ we first need to understand how the functors $PbPo$ and $PoPb$ act on objects and arrows. To this end, recall that pullbacks and pushouts in $\mathbf{Meas}$ are defined as in $\mathbf{Set}$ and equipped with initial and final σ-algebras with respect to the cone-projections and cocone-injections, respectively. This allows us to split the characterization in two parts: one regarding the underlying sets of the measurable spaces, and the other dealing with the σ-algebras structures. The last part, is the easies one, since once one has the characterization of the underlying sets it just need to equip them with the initial and final σ-algebras.

In $\mathbf{Set}$, the pushout of a relation $R \subseteq X \times Y$ is given by $(X + Y)/R^*$, where R^* is the smallest equivalence relation on $X + Y$ containing $\{(in_X(x), in_Y(y)) \mid (x, y) \in R\}$; the pullback of a quotient $(X + Y)/E$, is given by $\{(x, y) \in X \times Y \mid (in_X(x), in_Y(y)) \in E\}$.

Consider the comonad $PoPb$, first. Let $(X + Y)/E$ be a quotient in $\mathbf{ECospan}(X, Y)$. It is easy to see that $PoPb((X + Y)/E) = (X + Y)/E$. Indeed

$$PoPb((X + Y)/E) = Po(\{(x, y) \in X \times Y \mid (in_X(x), in_Y(y)) \in E\}) = (X + Y)/R^*$$

where R^* is the smallest equivalence on $X + Y$ containing E. But since E is already an equivalence, $R^* = E$. Thus, the counit $\epsilon \colon PoPb \Rightarrow Id$ and comultiplication $Po\eta Pb \colon PoPb \Rightarrow PoPbPoPb$ are just the "identity" natural transformations.

The case of the monad $(PbPo, \eta, Pb\epsilon Po)$ is a bit more involved. Let $R \subseteq X \times Y$ be a relation in $\mathbf{MSpan}(X, Y)$, then we have

$$PbPo(R) = Pb((X + Y)/R^*) = \{(x, y) \in X \times Y \mid (in_X(x), in_Y(y)) \in R^*\}$$

where R^* is the smallest equivalence relation on $X + Y$ containing $\{(in_X(x), in_Y(y)) \mid (x, y) \in R\}$. More explicitly, R^* can be characterized as the countable union $R^* = \bigcup_{n \in \mathbb{N}} R^*_n$, where

$$R^*_0 = \{(in_X(x), in_X(x)) \mid x \in X\} \cup \{(in_Y(y), in_Y(y)) \mid y \in Y\}$$
$$R^*_{n+1} = \{(in_X(x), in_Y(y')) \mid (x, y) \in R \wedge (in_X(x'), in_Y(y)) \in R^*_n \wedge (x', y') \in R\} \cup$$
$$\{(in_Y(y), in_X(x)) \mid (x, y) \in R \wedge (in_X(x'), in_Y(y)) \in R^*_n \wedge (x', y') \in R\}.$$

Therefore, $PbPo(R) = \bigcup_{n \in \mathbb{N}} \overline{R}_n$, where $\overline{R}_n = \{(x, y) \in X \times Y \mid (in_X(x), in_Y(y)) \in R^*_n\}$. Explicitly, we have that $\overline{R}_0 = \emptyset$ and, for all $n > 0$,

$$\overline{R}_1 = R,$$

$$\overline{R}_{n+1} = \{(x, y') \in X \times Y \mid (x, y) \in R \wedge (x', y) \in \overline{R}_n \wedge (x', y') \in R\}$$

Hence, $PbPo(R) = \overline{R}$, where $\overline{R}$ is the *z-closure of R*, that is, the smallest relation closed under the following rules

$$\frac{(x, y) \in R}{(x, y) \in \overline{R}} \qquad \frac{(x, y) \in \overline{R} \quad (x, y') \in \overline{R} \quad (x', y') \in \overline{R}}{(x, y') \in \overline{R}} \text{ (z-rule)}$$

Notice that $\overline{\overline{R}} = \overline{R}$, $R \subseteq \overline{R}$, and if $R \subseteq S$, then $\overline{R} \subseteq \overline{S}$, therefore $\overline{(\cdot)}$ is a closure operator. The functor $PbPo$ maps arrows $f \colon R \to S$ between relations $R, S \subseteq X \times Y$, to $\overline{f} \colon \overline{R} \to \overline{S}$, the z-closure extension of f, defined inductively on $\overline{R}_n$, in the obvious way. The unit $\eta \colon Id \Rightarrow PbPo$ is just the natural inclusion, and the multiplication $Pb\epsilon Po \colon PbPoPbPo \Rightarrow PbPo$ is the "identity" natural transformation.

As for the σ-algebra structures, $PoPb$ assigns to a quotient $(X+Y)/E$ in **ECospan**(X, Y) the final σ-algebra with respect to canonical injections. Note that, $PoPb((X+Y)/E)$ and $(X+Y)/E$, as measurable spaces, are not necessarily isomorphic, since the final σ-algebra on $PoPb((X+Y)/E)$ actually contains all the measurable sets of the quotient $(X+Y)/E$, but this inclusion may be strict. Dually, the functor $PbPo$ assigns to a relation $R \subseteq X \times Y$ in **MSpan**(X, Y) the measurable space $(\overline{R}, \Sigma_{\overline{R}})$, where $\overline{R}$ is the z-closure of R, $\Sigma_{\overline{R}}$ is the initial σ-algebra with respect to the canonical projections $\pi_X \colon \overline{R} \to X$ and $\pi_Y \colon \overline{R} \to Y$, that is, the σ-algebra generated by the sets $(E \times F) \cap \overline{R}$, for $E \in \Sigma_X$ and $F \in \Sigma_Y$.

5.3.2 Adjunction between Bisimulations and Cocongruences

We already seen that there is an adjunction $Po \dashv Pb$ between the categories of monic spans **MSpan**(X, Y) and epic cospans **ECospan**(X, Y) in **Meas**, and that this is given as an instance of a more general construction that applies in any category with pullbacks and pushouts. In this section, we show that this adjunction can be partially lifted to the categories **Bisim**$((X, \alpha), (Y, \beta))$ and **Cocong**$((X, \alpha), (Y, \beta))$ of $\Delta_{<\infty}^L$-bisimulations and $\Delta_{<\infty}^L$-cocongruences. We used the term "partially", since the adjunction actually works only restricting the category of cocongruences to the image given by the functor

$$Po \colon \textbf{Bisim}((X, \alpha), (Y, \beta)) \to \textbf{Cocong}((X, \alpha), (Y, \beta)),$$

namely, $Po(\textbf{Bisim}((X, \alpha), (Y, \beta)))$. This, because only for cocongruences that have been derived from bisimulations it is possible to define a coalgebra structure making the cospan an actual bisimulation. Moreover, we will also see that the there is an equivalence between the subcategories

$$PbPo(\textbf{Bisim}((X, \alpha), (Y, \beta))) \cong PoPb(\textbf{Cocong}((X, \alpha), (Y, \beta))),$$

given as the images of the functors $PbPo$ and $PoPb$, respectively. This exactly establish which the conditions under which the concept of bisimulation and cocongruence coincide.

The first step consists in the formal definition of the categories of bisimulations and cocongruences for the endofunctor $\Delta_{<\infty}^L \colon \textbf{Meas} \to \textbf{Meas}$.

Category of bisimulations. **Bisim**$((X, \alpha), (Y, \beta))$ has as objects bisimulations $((R, \gamma), f, g)$ between $\Delta_{<\infty}^L$-coalgebras (X, α) and (Y, β), and arrows $f \colon ((R, \gamma_R), r_1, r_2) \to ((S, \gamma_S), s_1, s_2)$ which are morphisms $f \colon K \to H$ in **Meas** making the following diagrams commute:

hence, $f \colon K \to H$ is both a morphism in $\Delta_{<\infty}^L$-**coalg** and in **MSpan**$_{\textbf{Meas}}(X, Y)$.

Category of cocongruences. $\mathbf{Cocong}((X,\alpha),(Y,\beta))$ has objects cocongruences $((K,\kappa),f,g)$ between $\Delta_{<\infty}^L$-coalgebras (X,α) and (Y,β), and arrows $f\colon ((K,\kappa_K),k_1,k_2) \to ((H,\kappa_H),h_1,h_2)$ are morphisms $f\colon K \to H$ in $\mathbf{Meas}$ such that the following diagrams commute:

therefore, $f\colon K \to H$ is both a morphism in $\Delta_{<\infty}^L$-$\mathbf{coalg}$ and in $\mathbf{ECospan_{Meas}}(X,Y)$.

Between these two categories we can define a functor

$$Po\colon \mathbf{Bisim}((X,\alpha),(Y,\beta)) \to \mathbf{Cocong}((X,\alpha),(Y,\beta)),$$

acting on $\Delta_{<\infty}^L$-bisimulations $((R,\gamma_R),r_1,r_2)$ has $Po(((R,\gamma),r_1,r_2)) = ((K,\kappa),k_1,k_2)$, where (K,k_1,k_2) is the pushout for the span (R,r_1,r_2), and $\kappa\colon K \to \Delta_{<\infty}^L K$ is the unique, given by the universal property of pushouts, making the following diagram on the left commute; and on arrows $f\colon ((R,\gamma_R),r_1,r_2) \to ((S,\gamma_S),s_1,s_2)$ as $Po(f)$, defined as the unique arrow, given by the universal property of pushouts, making the diagram on the right commute:

The arrow $Po(f)$ is obviously a morphism between the cospans $(Po(R),k_1,k_2)$ and $(Po(S),h_1,h_2)$, and can be proved to be also a morphisms of coalgebras exploiting the universal property of pushouts. Functoriality follows for similar reasons.

Remark 5.3.3 The above construction is standard an applies in any category with pushouts independently of the choice of the behaviour functor. ∎

It is well known that if the behavior functor preserves weak pullbacks, then cocongruences give rise to bisimulations (Proposition 3.5.4). Unfortunately, the functor $\Delta_{<\infty}$ does not preserves weak pullbacks, as proved by Viglizzo in [89, 68], hence we cannot define a functor

$$Pb\colon \mathbf{Cocong}((X,\alpha),(Y,\beta)) \to \mathbf{Bisim}((X,\alpha),(Y,\beta)).$$

However, it turns out that if we restrict our attention only to the subcategory of cocongruences that are images of the functor Po for some bisimulation, namely, $Po(\mathbf{Bisim}((X,\alpha),(Y,\beta)))$, it is possible to define such a functor.

Let $((R,\gamma),r_1,r_2)$ be an object and f be an arrow in $\mathbf{Bisim}((X,\alpha)(Y,\beta))$, then we define

$$Z\colon Po\big(\mathbf{Bisim}((X,\alpha),(Y,\beta))\big) \to \mathbf{Bisim}((X,\alpha),(Y,\beta))$$

$$Z(Po((R,\gamma),r_1,r_2)) = ((\overline{R},\overline{\gamma}),\overline{r}_1,\overline{r}_2)$$

$$Z(Po(f)) = \overline{f}\,,$$

where $PbPo((R,\gamma),r_1,r_2) = (\overline{R},\overline{r}_1,\overline{r}_2)$ and $PbPo(f) = \overline{f}$ are the span and span-morphism, respectively, given by the (well defined) endofunctor $PbPo\colon \mathbf{MSpan}(X,Y) \to \mathbf{MSpan}(X,Y)$ on

monic spans between X and Y in **Meas** (recall from Section 5.3.1 that, $\overline{R}$ and $\overline{f}$ are, respectively, the z-closure of R and the z-closure extension of f), and $\overline{\gamma}\colon \overline{R} \to \Delta^L_{<\infty}\overline{R}$ is defined, for each $r \in \overline{R}$, $a \in L$, as the (unique) measure on $(\overline{R}, \Sigma_{\overline{R}})$ such that

$$\overline{\gamma}(r)(a)(\langle \overline{r}_1, \overline{r}_2 \rangle^{-1}(E \times F)) = \alpha(\overline{r}_1(r))(a)(E) \qquad \text{for all } E \in \Sigma_X \text{ and } F \in \Sigma_Y,$$

given by Proposition 5.2.5 (notice that, $\Sigma_{\overline{R}}$ is generated by the sets $\langle \overline{r}_1, \overline{r}_2 \rangle^{-1}(E \times F)$, for $E \in \Sigma_X$ and $F \in \Sigma_Y$, hence the construction in Proposition 5.2.5 applies).

Next we show that this functor is well defined. To this end we have to prove that $((\overline{R}, \overline{\gamma}), \overline{r}_1, \overline{r}_2)$ is a $\Delta^L_{<\infty}$-bisimulation and that $\overline{f}$ is a morphism between bisimulations, i.e., a span-morphism (this already holds by definition) and a $\Delta^L_{<\infty}$-homomorphism.

Lemma 5.3.4 *Let (X, α) and (Y, β) be two $\Delta^L_{<\infty}$-coalgebras, (R, r_1, r_2) and (S, s_1, s_2) be spans between X and Y in **Meas**, with σ-algebras Σ_R and Σ_S, initial w.r.t. $\langle r_1, r_2 \rangle\colon R \to X \times Y$ and $\langle s_1, s_2 \rangle\colon S \to X \times Y$, respectively. Define $\gamma_R\colon R \to \Delta^L_{<\infty}R$ and $\gamma_S\colon S \to \Delta^L_{<\infty}S$, for $r \in R$, $s \in S$, $a \in L$, as the the unique measures on (R, Σ_R) and (S, Σ_S) such that*

$$\gamma_R(r)(a)(\langle r_1, r_2 \rangle^{-1}(E \times F)) = \alpha(r_1(r))(a)(E) \qquad \text{for all } E \in \Sigma_X \text{ and } F \in \Sigma_Y,$$
$$\gamma_S(s)(a)(\langle s_1, s_2 \rangle^{-1}(E \times F)) = \alpha(s_1(s))(a)(E) \qquad \text{for all } E \in \Sigma_X \text{ and } F \in \Sigma_Y,$$

following the procedure given in Proposition 5.2.5.

Then (R, γ_R) and (S, γ_S) are well-defined $\Delta^L_{<\infty}$-coalgebras, and any morphism $f\colon R \to S$ between the span (R, r_1, r_2) and (S, s_1, s_2), is also $\Delta^L_{<\infty}$-homomorphism between (R, γ_R) and (S, γ_S).

Proof. To prove that (R, γ_R) and (S, γ_S) are well defined, we need to show that γ_R and γ_S are measurable. We prove it for γ_R, the other follows similarly. Note that Σ_R can be generated by the boolean algebra of all finite unions of the form $S = \bigcup_{i=0}^{k} \langle r_1, r_2 \rangle^{-1}(E_i \times F_i)$, $E_i \in \Sigma_X$ and $F_i \in \Sigma_Y$, for $0 \leq i \leq k$. Without loss of generality, all $E_i \times F_i$ can be assumed to be disjoint (otherwise we can find a refinement that is so), thus also $\langle r_1, r_2 \rangle^{-1}(E_i \times F_i)$ must be disjoint. Hence, by Lemmas 2.2.8 and 2.3.3 it suffices to show that for any S as above $(\gamma_R(\cdot)(a))^{-1}(L_r(S)) \in \Sigma_R$.

$$
\begin{aligned}
(\gamma_R(\cdot)(a))^{-1}(L_r(S)) &= \{r \in R \mid \gamma_R(r)(a) \in L_r(S)\} &\text{(by inverse image)} \\
&= \{r \in R \mid \gamma_R(r)(a)(S) \geq r\} &\text{(by def. } L_r(E)) \\
&= \{r \in R \mid \textstyle\sum_{i=0}^{k} \gamma(r)(a)(\langle r_1, r_2 \rangle^{-1}(E_i \times F_i)) \geq r\} &\text{(by finite additivity)} \\
&= \{r \in R \mid \textstyle\sum_{i=0}^{k} \alpha(r_1(r))(a)(E_i) \geq r\} &\text{(by def. } \gamma_R) \\
&= \{r \in R \mid \alpha(r_1(r))(a)(\textstyle\bigcup_{i=0}^{k} E_i) \geq r\} &\text{(by finite additivity)} \\
&= \{r \in R \mid \alpha(r_1(r))(a) \in L_r(\textstyle\bigcup_{i=0}^{k} E_i)\} &\text{(by def. } L_r(E)) \\
&= \langle r_1, r_2 \rangle^{-1}(\alpha^{-1}(\textstyle\bigcup_{i=0}^{k} E_i) \times Y) &\text{(by inverse image)}
\end{aligned}
$$

Since α is measurable, $\alpha^{-1}(\bigcup_{i=0}^{k} E_i) \in \Sigma_X$, so that $(\gamma_R(\cdot)(a))^{-1}(L_r(S)) \in \Sigma_R$.

To prove that $f\colon R \to S$ is a morphism between the coalgebras (R, γ_R) and (S, γ_S), we have to show $\gamma_S \circ f = \Delta^L_{<\infty}f \circ \gamma_R$. Let $r \in R$, $a \in L$, $E \in \Sigma_X$, and $F \in \Sigma_Y$, then

$$
\begin{aligned}
(\gamma_S \circ f)(r)(a)(\langle s_1, s_2 \rangle^{-1}(E \times F)) &= \gamma_S(f(r))(a)(\langle s_1, s_2 \rangle^{-1}(E \times F)) &\text{(composition)} \\
&= \alpha(s_1 \circ f(r))(a)(E) &\text{(by def. } \gamma_S) \\
&= \alpha(r_1(r))(a)(E) &\text{(by } f \text{ span-morphism)} \\
&= \gamma_R(r)(a)(\langle r_1, r_2 \rangle^{-1}(E \times F)) &\text{(by def. } \gamma_R) \\
&= \gamma_R(r)(a)((\langle s_1, s_2 \rangle \circ f)^{-1}(E \times F)) &\text{(by } f \text{ span-morphism)} \\
&= \gamma_R(r)(a)(f^{-1} \circ \langle s_1, s_2 \rangle^{-1}(E \times F)) &\text{(by comp. inverses)} \\
&= (\Delta^L_{<\infty}f \circ \gamma_R)(a)(\langle s_1, s_2 \rangle^{-1}(E \times F)) &\text{(by def. } \Delta^L_{<\infty})
\end{aligned}
$$

Due to the uniqueness of the definition of γ_R and γ_S (see right hand sides of of lines 1 and 6 above) this equality is sufficient to prove the equality for all measurable sets in Σ_S. ∎

Lemma 5.3.5 ([34]) *Let* $\mathcal{M} = (X, \Sigma_X, \{\alpha_a\}_{a \in L})$ *and* $\mathcal{N} = (Y, \Sigma_Y, \{\beta_a\}_{a \in L})$ *be L-labelled Markov kernels, and* $\mathcal{R} \subseteq X \times Y$ *be a bisimulation between them. Then the z-closure of* $\mathcal{R}$ *is a bisimulation.*

Proof. Let $\overline{\mathcal{R}} \subseteq X \times Y$ be the z-closure of $\mathcal{R}$. We have to show that for all $(x, y) \in \overline{\mathcal{R}}$, $a \in A$, $E \in \Sigma_X$ and $F \in \Sigma_Y$, such that (E, F) are $\overline{\mathcal{R}}$-closed,

$$\alpha_a(x)(E) = \beta_a(y)(F).$$

Note that $\overline{\mathcal{R}} = \bigcup_{n \in \mathbb{N}} \overline{\mathcal{R}}_n$, for $\overline{\mathcal{R}}_0 = \mathcal{R}$, $\overline{\mathcal{R}}_{n+1} = \{(x, y') \mid (x, y) \in \mathcal{R} \wedge (x', y) \in \overline{\mathcal{R}}_n \wedge (x', y') \in \mathcal{R}\}$. We proceed by induction on $n \geq 0$. Base case $n = 0$: assume $(x, y) \in \mathcal{R}$. Since $\mathcal{R} \subseteq \overline{\mathcal{R}}$, by Lemma 5.1.3, (E, F) is $\mathcal{R}$-closed, so that $\alpha_a(x)(E) = \beta_a(y)(F)$ follows since $\mathcal{R}$ is a bisimulation. Inductive case $n > 0$: assume $(x, y) \in \overline{\mathcal{R}}_{n+1}$ and by inductive hypothesis that for all $(x', y') \in \overline{\mathcal{R}}_n$, $\alpha_a(x')(E) = \beta_a(y')(F)$. By definition of $\overline{\mathcal{R}}_{n+1}$, we have that there exists $x' \in X$ and $y' \in Y$, such that $(x, y'), (x', y) \in \mathcal{R}$ and $(x', y') \in \overline{\mathcal{R}}$. Then, by $\mathcal{R} \subseteq \overline{\mathcal{R}}$, and Lemma 5.1.3, (E, F) is $\mathcal{R}$-closed, thus $\alpha_a(x)(E) = \beta_a(y')(F)$ and $\alpha_a(x')(E) = \beta_a(y)(F)$, since $\mathcal{R}$ is a bisimulation. By inductive hypothesis, $\alpha_a(x')(E) = \beta_a(y')(F)$, therefore $\alpha_a(x)(E) = \beta_a(y)(F)$. ∎

Proposition 5.3.6 $Z \colon Po\big(\mathbf{Bisim}((X, \alpha), (Y, \beta))\big) \to \mathbf{Bisim}((X, \alpha), (Y, \beta))$ *is well-defined.*

Proof. It follows by Lemmas 5.3.4, 5.3.5, and Theorem 5.2.9, noticing that the coalgebra structure map is given following the definition in Proposition 5.2.6. Functoriality follows since the functor $PbPo$ is so. ∎

The functors Po and Z acts on spans and cospans and their morphisms exactly as the functors we have introduced at the beginning of Section 5.3. This means that many properties can be lifted also to this setting. In particular, we have that the two functors determines an adjunction:

Theorem 5.3.7 (Adjunction & equivalence) *Let* (X, α) *and* (Y, β) *be* $\Delta_{<\infty}^L$*-coalgebras. Then the following functors are are adjoint,* $Po \dashv Z$,

$$Po \colon \mathbf{Bisim}((X, \alpha), (Y, \beta)) \to Po\big(\mathbf{Bisim}((X, \alpha), (Y, \beta))\big)$$
$$Z \colon Po\big(\mathbf{Bisim}((X, \alpha), (Y, \beta))\big) \to \mathbf{Bisim}((X, \alpha), (Y, \beta))$$

Moreover, the subcategories $ZPo\big(\mathbf{Bisim}((X, \alpha), (Y, \beta))\big)$ *and* $PoZPo\big(\mathbf{Bisim}((X, \alpha), (Y, \beta))\big)$ *are equivalent with isomorphism defined by* Po *and* Z.

Proof. Due to the universal properties of pushouts, for any pair of objects (R, r_1, r_2) and (S, s_1, s_2) in $\mathbf{Bisim}((X, \alpha), (Y, \beta))$ it holds that,

$$Hom(Po(R, r_1, r_2), Po(S, s_1, s_2)) \cong Hom((R, r_1, r_2), ZPo(S, s_1, s_2)),$$

that is, Po is left adjoint to Z. The equivalence follows by Proposition 5.3.1(i). ∎

Historical note. In [31], Danos et al. proposed a notion alternative to bisimulations, the so called *event bisimulation*, being aware that it coincides to behavioral equivalence. Here we recall their definition and try to make a comparison between their results in connection to Theorem 5.3.7.

Definition 5.3.8 (Event bisimulation) *Let* $\mathcal{M} = (X, \Sigma, \{\theta_a\}_{a \in L})$ *be a L-labelled Markov kernel. A sub-σ-algebra* $\Lambda \subseteq \Sigma$ *is an* event bisimulation *if, for all* $a \in L$ *and* $E \in \Lambda$, $\theta_a^{-1}(E) \in \Lambda$.

Any σ-algebra Σ on X induces a notion of separability in the form of an (equivalence) relation $\Re(\Sigma) \subseteq X \times X$ defined by $\Re(\Sigma) = \{(x, y) \mid \forall E \in \Sigma. \, x \in \Sigma \text{ iff } y \in \Sigma\}$. Moreover, considering only equivalence relations $\mathcal{R} \subseteq X \times X$, they denoted by $\Sigma(\mathcal{R}) = \{E \in \Sigma \mid (E, E) \text{ is } \mathcal{R}\text{-closed}\}$ the set of measurable $\mathcal{R}$-closed sets, which is readily seen to be a σ-algebra on X. The "operator" $\Re(\cdot)$ maps σ-algebras to equivalence relations and, conversely, $\Sigma(\cdot)$ maps equivalence relations to σ-algebras.

Therefore, under some assumptions, they can also thought of as maps between event bisimulations and state bisimulations.

Identifying event bisimulations Λ with their associated equivalence $\Re(\Lambda)$, they proved ([31, Lemma 4.8]) that a state bisimulation $\mathcal{R}$ is an event bisimulation iff $\mathcal{R} = \Re(\Sigma(\mathcal{R}))$. Replacing the operator $\Sigma(\cdot)$ with our functor Po, and $\Re(\cdot)$ with Z, their result is in accordance with our restriction to the subcategory $Po\big(\mathbf{Bisim}((X, \alpha), (Y, \beta))\big)$ for the definition of the functor Z, and with the fact that the objects in $ZPo\big(\mathbf{Bisim}((X, \alpha), (Y, \beta))\big)$ are in bijection with cocongruences.

We do not know whether they recognized the adjunction and equivalence of categories of Theorem 5.3.7, but our results, although just informally, seems to agree with their. The main merits of our coalgebraic exposition are, not just to have extended the results to bisimulation relations that are not assumed to be equivalences, but also in having recognized the underlying universal property in terms of a formal adjunction. This, of course, gives a better understanding of the deep relations occurring between state and event bisimulations, remarkably, even without a weak-pullback preserving behavior functor.

Remark 5.3.9 (Bisimilarity is a behavioral equivalence) Any bisimulation $\mathcal{R} \subseteq X \times X$ can be extended to contain the identity relation on X and all symmetric pairs. The z-closure $\overline{\mathcal{R}}$ of a reflexive and symmetric relation $\mathcal{R}$ corresponds to the smallest equivalence relation containing it. Since z-closure corresponds exactly with the action on objects for the functor Z, we have that all bisimulations which are equivalences are objects in $Po\big(\mathbf{Bisim}((X, \alpha), (Y, \beta))\big)$, so that, they are in bijection with the cocongruence (actually, behavioral equivalence) obtained applying the functor Po to them.

Recall that (state) bisimilarity, $\sim$, is proved to be an equivalence (Theorem 5.1.9), hence it is in bijection with its associated behavioral equivalence "$Po(\sim)$", so that we can informally say that bisimilarity *is* a behavioral equivalence or, using the terminology in [31], an event bisimulation. Note that, this does not mean that it corresponds to event bisimilarity (i.e., the largest event bisimulation). Indeed, Terraf in [80] already proved this is not the case, showing a counter example in which two states of a Markov kernel are event bisimilar but not state bisimilar. Still, we can say, that state bisimilarity is included in event bisimilarity. ∎

6

Rule Formats for Continuous State Probabilistic and Stochastic Systems

Plotkin's *structural operational semantics* (SOS) [71] is generally recognized to be a successful tool for giving operational semantics to recursively defined process description languages in terms of labelled transition systems. An SOS specification system consists of a collection of derivation rules that allows for a simple description of the transitions of a labelled transition system following the syntactic structure of the terms of the programming language. The great success of the SOS paradigm is mainly due to the fact that many important semantic properties, such as congruence for bisimilarity, can be established simply by inspecting the syntactic format of the rules. Among all proposed rule formats, the most popular are the GSOS format [17] and the tyft/tyxt rules [51] (for an introduction on rule specification formats we recommend the survey by Aceto et al. [2]).

An abstract categorical formulation of well-behaved SOS specification formats has been proposed by Turi and Plotkin [83, 82], who recognized that rule specification systems can be formulated in terms of certain natural transformations, that is distributive laws of a monad over a comonad. Intuitively, the monad represents the syntax of the programming language and the comonad models the computations of the operational system. The models for these distributive laws are the so called *bialgebras*, that is, a pair consisting of an algebra (for the monad) and a coalgebra (for the comonad) over the same carrier, and such that they are well-behaved with respect to the distributive law. The main advantages of this abstract categorical formulation are that it allows for a deeper understanding of the theory leaving out all the technical details due to the specific model at hand, and, more importantly, it permits to instantiate the general framework to different kinds of systems behaviors.

In recent years, this approach has been applied also to *stochastic* and *probabilistic* systems, due to their important applications to performance evaluation, systems biology, etc [55, 21, 54, 41]. For example Bartels [15] has investigated rule formats both for simple discrete probabilistic systems and Segala systems, proving that probabilistic bisimilarity of Larsen and Skou is a congruence with respect his rule formats. Inspired by Bartel's results, Klin and Sassone [61, 60] have proposed rule formats for stochastic systems and, more generally, for weighted transition systems, providing evidence that many well-known stochastic extension of process calculi, such as PEPA [55], CSP [22], and CCS [65] fit their format.

However, these formats still do not cover the case of *continuous-state* probabilistic and stochastic systems, like calculi with spatial/geometric features introduced in last years [24, 12]. In these models, the behaviour of the system may be influenced by continuous data, which therefore is part of the state of the system. Typical examples are quantitive informations such as density, volumes, and spatial informations, such as the position of processes and where transitions take place; e.g., in wireless networks distance may affect data access, or in biological models diffusion alters the signaling pathways, etc. Surprisingly, none have been proposed rule formats for this kind of systems in the literature yet. So in this chapter we introduce the first well-behaved SOS-like specification formats for continuous state probabilistic and stochastic systems.

The operational models we are interested in are Markov processes (cf. Chapter 5), so that the notion of interest is no longer a measure on a discrete space, but a measure over a generic

measurable space. Categorically, this corresponds to move to the category **Meas** of measurable spaces and measurable functions, and to model the system behaviour by a $\Delta_{<\infty}^L$-coalgebra, where $\Delta_{<\infty} \colon \mathbf{Meas} \to \mathbf{Meas}$ is the measure functor introduced in Section 2.3, associating to a measurable space (X, Σ) the measurable space of measures over it. This leads to transitions of the form $x \xrightarrow{a} \mu$, where $x \in X$ is the current state of the system, $a \in L$ is an action label representing the interactions with an external environment, and μ is an actual *measure* over (X, Σ) measuring either the probability or the execution rate of the possible outcomes of x. Semantics with a similar transition format have been considered already by Cardelli and Mardare in [26, 10] for dealing with specific equational stochastic systems. However, differently from the case of discrete processes, the SOS specification given in [26, 10] are rather ad hoc, and they are not based on any general framework for operational descriptions.

In this chapter we will cover this gap introducing a new GSOS-like rule format for Markov processes over generic measurable spaces. To this end, we aim to apply the bialgebraic framework introduced by Turi and Plotkin [83] to distributive laws of type $S(Id \times \Delta_{<\infty}^L) \Rightarrow (\Delta_{<\infty} T^S)^L$, where $S \colon \mathbf{Meas} \to \mathbf{Meas}$ is a syntactic endofunctor over **Meas**, specifying the syntax of continuous data types, and T^S the corresponding free monad modeling the term language as an actual measurable space.

A "good" SOS rule format must be compositional, i.e., it has to define a system's behaviour in terms of those of its subsystems. In traditional GSOS format, this is reflected by the fact that the target of a transition is a term built from the components of the source process, and their corresponding semantics. In our settings, the target of a transition is not a term, but a measure over a generic measurable space, which do not have any syntactic structure to play with. In order to circumvent this problem, we propose to use transitions of the form $t \xrightarrow{a} \mu$ where μ is no more a measure but a syntactic expression intended to denote a measure, which we call *measure term*. The syntax of these measure terms, and their interpretation as actual measures, is part of the operational specification: a specification is given by a set of rules together with a description of how measures must be combined. We will show that this specification format, is general enough to cover interesting examples of process calculi dealing with continuous data. In particular, we show that any MGSOS specification leads to a distributive law of type $S(Id \times \Delta_{<\infty}^L) \Rightarrow (\Delta_{<\infty} T^S)^L$, and as a consequence, the induced behavioural equivalence is a congruence.

Remark 6.0.10 One may have noticed that the type of the coalgebraic behaviour functor we have considered above is $\Delta_{<\infty}$, hence we have restricted the attention only to finite measures. This is done because for the definition of a final operational semantic we need that the behaviour functor has final coalgebra. In Section 4.1.3 we proved that a final coalgebra exists for the finite measure functor, but we were not able to provide the same result for the general case Δ (see Remark 4.1.21).

This, however is not a dramatic limitation, since it just corresponds to impose that the stochastic transitions of the Markov processes have finite execution rate, a condition that is usually imposed in the literature (sometimes even without a specific theoretical motivation). ∎

6.1   Measurable Spaces of Terms

The development of a structural operational semantics over generalized Markov processes demands for an algebraic description of their measurable space states. This can be formalized by means of *measurable signatures*.

Definition 6.1.1 *A* measurable signature *is a triple* $(\mathcal{S}, ar, \{(X_s, \Sigma_s)\}_{s \in \mathcal{S}})$, *where* $\mathcal{S}$ *is a set of operator symbols,* $ar \colon \mathcal{S} \to \mathbb{N}$ *is an arity function, and, for* $s \in \mathcal{S}$, (X_s, Σ_s) *is a measurable space. An* interpretation *of* $(\mathcal{S}, ar, \{(X_s, \Sigma_s)\}_{s \in \mathcal{S}})$ *on a measurable set* (X, Σ) *is an* $\mathcal{S}$-*indexed collection of measurable functions* $(\llbracket s \rrbracket \colon X_s \times X^{ar(s)} \to X)_{s \in \mathcal{S}}$.

Differently, from standard set-signatures, each operator symbol $s \in \mathcal{S}$ is associated with a measurable space (X_s, Σ_s). This, for example, allows for the definition of actual measurable spaces of constant symbols (i.e., 0-ary operators) and for the specification of nontrivial measurable spaces of terms.

Definition 6.1.2 (Measurable terms) *Let* $(\mathcal{S}, ar, \{(X_s, \Sigma_s)\}_{s \in \mathcal{S}})$ *be a measurable signature and* (X, Σ) *be a measurable space of variables. The measurable space of terms (freely) generated over* (X, Σ) *and* $(\mathcal{S}, ar, \{(X_s, \Sigma_s)\}_{s \in \mathcal{S}})$ *is defined as* (TX, Σ_{TX}) *where, TX and Σ_{TX} are, respectively, the smallest set and the smallest σ-algebra satisfying the following rules, for all $s \in \mathcal{S}$*

$$\frac{x \in X}{x \in TX} \qquad \frac{t_1, \ldots, t_{ar(s)} \in TX \qquad k \in X_s}{s(k, t_1, \ldots, t_{ar(s)}) \in TX}$$

$$\frac{E \in \Sigma}{E \in \Sigma_{TX}} \qquad \frac{T_1, \ldots, T_{ar(s)} \in \Sigma_{TX} \qquad K \in \Sigma_s}{s(K, T_1, \ldots, T_{ar(s)}) \in \Sigma_{TX}}$$

where $s(K, T_1, \ldots, T_{ar(s)}) = \{s(k, t_1, \ldots, t_{ar(s)}) \mid k \in K, t_i \in T_i\}$.

In the definition above the number arguments for an operator symbol $s \in \mathcal{S}$ is augmented by one in the term $s(k, t_1, \ldots, t_{ar(s)})$. This notation is particularly convenient for discriminating between measurable terms which differ only on the σ-algebra structure (a situation that does not happen in standard set-signatures). For example, assume the measurable spaces $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and $(\mathbb{R}, \{\emptyset, \mathbb{R}\})$ are associated with constant operators c and c', respectively. If we had considered $r \in \mathbb{R}$ as a term, it would have been impossible to discriminate between the elements in c or c'. Adopting the notation above these problems are overcome, since these constants are denoted either by $c(r)$ or $c'(r)$.

6.1.1 Signature Interpretations as Algebras

Algebras are typically used to give an abstract categorical formalization to denotational semantics. Now we show how measurable signatures and their interpretations can be modeled as algebras for certain functors in **Meas**. Although this technique is standard in **Set**, the σ-algebra structures endowed with the objects in **Meas** makes the construction less easy to be treated.

For a measurable signature $(\mathcal{S}, ar, \{(X_s, \Sigma_s)\}_{s \in \mathcal{S}})$, the *syntactic endofunctor* in **Meas** associated with it is given by $S = \coprod_{s \in \mathcal{S}} (X_s, \Sigma_s) \times Id^{ar(s)}$. Explicitly S acts on objects (X, Σ) and arrows $f \colon (X, \Sigma_X) \to (Y, \Sigma_Y)$ as follows

$$S(X, \Sigma) = (\{\langle s, (k, x_1, \ldots, x_{ar(\sigma)}) \rangle \mid s \in \mathcal{S}, k \in X_s, x_1, \ldots, x_{ar(\sigma)} \in X\}, \Sigma_{S(X,\Sigma)}),$$

$$Sf = \left[\langle s, (k, x_1, \ldots, x_{ar(\sigma)}) \rangle \mapsto \langle s, (k, f(x_1), \ldots, f(x_{ar(\sigma)})) \rangle \right],$$

where $\langle s, (k, x_1, \ldots, x_{ar(\sigma)}) \rangle$ denotes $in_s^X(k, x_1, \ldots, x_{ar(\sigma)})$, and $\Sigma_{S(X,\Sigma)}$ is the final σ-algebra with respect to the injections $in_s^X \colon X_s \times X^{ar(s)} \to SX$, for each $s \in \mathcal{S}$, that is

$$\Sigma_{S(X,\Sigma)} = \bigcap_{s \in \mathcal{S}} \{A \subseteq SX \mid (in_s^X)^{-1}(A) \in \Sigma_{X_s \times X^{ar(s)}}\}. \tag{6.1.1}$$

The σ-algebra $\Sigma_{S(X,\Sigma)}$ can be characterized in a more convenient way by means of a generating family of sets having a structure that is simpler to handle than the subsets $A \subseteq SX$ occurring in Equation (6.1.1).

Proposition 6.1.3 *Let* $(\mathcal{S}, ar, \{(X_s, \Sigma_s)\}_{s \in \mathcal{S}})$ *be a measurable signature and S be the syntactic* **Meas**-*endofunctor associated with it. Then, for any measurable space (X, Σ), the σ-algebra of $S(X, \Sigma)$ is generated by*

$$\{\bigcup_{s \in \mathcal{S}} \langle s, (K_s, E_1, \ldots, E_{ar(s)}) \rangle \mid \forall s \in \mathcal{S}. \, K_s \in \Sigma_s \text{ and } E_i \in \Sigma\}.$$

where $\langle s, (K_s, E_1, \ldots, E_{ar(s)}) \rangle = \{\langle s, (k, x_1, \ldots, x_{ar(s)}) \rangle \mid k \in K_s \text{ and } x_i \in E_i\}$.

Proof. Let first notice that, for each $s \in \mathcal{S}$, the σ-algebra $\Sigma_{X_s \times X^{ar(s)}}$ is generated by $\mathcal{R}_s = \{K \times E_1 \times \cdots \times E_{ar(s)} \mid K \in \Sigma_s, E_i \in \Sigma\}$, that is, the family of measurable rectangles. So, we have that

$$\Sigma_{S(X,\Sigma)} = \bigcap_{s \in \mathcal{S}} \{A \subseteq SX \mid (in_s^X)^{-1}(A) \in \Sigma_{X_s \times X^{ar(s)}}\}$$

$$= \bigcap_{s \in \mathcal{S}} \{A \subseteq SX \mid (in_s^X)^{-1}(A) \in \sigma(\mathcal{R}_s)\}$$
$$= \bigcap_{s \in \mathcal{S}} \sigma\big(\{A \subseteq SX \mid (in_s^X)^{-1}(A) \in \mathcal{R}_s\}\big) \qquad \text{(by Prop. 2.2.10)}$$
$$= \sigma\big(\{\textstyle\bigcup_{s \in \mathcal{S}} \langle s, (K_s, E_1, \ldots, E_{ar(s)}) \rangle \mid \forall s \in \mathcal{S}.\, K_s \in \Sigma_s \text{ and } E_i \in \Sigma\}\big),$$

where the last equality holds since each $A \subseteq SX$ such that $(in_s^X)^{-1}(A) \in \mathcal{R}_s$ can always be given as a disjoint union of the form $\bigcup_{s \in \mathcal{S}} \langle s, (K_s, E_1, \ldots, E_{ar(s)}) \rangle$ for some $K_s \in \Sigma_s$ and $E_i \in \Sigma$, for $1 \leq i \leq ar(s)$. ∎

6.1.2 Term monad over measurable spaces

In this section we show how the measurable space of terms over a measurable signature can be elegantly modeled as the free monad over the syntactic endofunctor associated with the signature. As a consequence we obtain a principle of structural induction over measurable terms which extends the well-known principle of structural induction for terms over standard set-signatures.

To this end, for any measurable signature $(\mathcal{S}, ar, \{(X_s, \Sigma_s)\}_{s \in \mathcal{S}})$, we need to show that the forgetful functor from the category of algebras for the syntactic **Meas**-endofunctor S associated with the signature has a left adjoint. It is standard that, for endofunctors $F \colon \mathbf{C} \to \mathbf{C}$ in a category $\mathbf{C}$ with binary coproducts, the the forgetful functor $U^F \colon F\text{-}\mathbf{alg} \to \mathbf{C}$ has a left adjoint $L^F \colon \mathbf{C} \to F\text{-}\mathbf{alg}$, if for any object X in $\mathbf{C}$ the functor $X + F$ has initial algebra. Therefore, it suffices to show that for any measurable space (X, Σ) the functor $(X, \Sigma) + S$ has initial algebra.

Theorem 6.1.4 *Let $(\mathcal{S}, ar, \{(X_s, \Sigma_s)\}_{s \in \mathcal{S}})$ be a measurable signature, S be the syntactic endofunctor associated with it, and (X, Σ) be a measurable set. Then $((TX, \Sigma_{TX}), [\eta_X, \psi_X])$ is the initial $((X, \Sigma) + S)$-algebra, where (TX, Σ_{TX}) is the measurable space of terms generated over (X, Σ) and $(\mathcal{S}, ar, \{(X_s, \Sigma_s)\}_{s \in \mathcal{S}})$ and $\eta_X \colon X \to TX$ and $\psi_X \colon STX \to TX$ are defined as follows*

$$\eta_X(x) = x\,, \qquad\qquad \psi_X(\langle s, (k, t_1, \ldots, t_{ar(s)}) \rangle) = s(k, t_1 \ldots t_{ar(\sigma)})\,,$$

for all $x \in X$, $s \in \mathcal{S}$, $k \in X_s$, and $t_1, \ldots, t_{ar(s)} \in TX$.

Proof. By Proposition 6.1.3 it is immediate to see that both η_X and ψ_X are measurable, hence $[\eta_X, \psi_X]$ is a well-defined $((X, \Sigma) + S)$-algebra structure. We proceed first proving that $(X, \Sigma) + S$ has an initial algebra, then we prove that it is isomorphic to $((TX, \Sigma_{TX}), [\eta_X, \psi_X])$. Let $A \colon \mathbf{Ord} \to \mathbf{Meas}$ be the initial sequence of $(X, \Sigma) + S$, we prove that $A(\omega \to \omega{+}1)$ is an epimorphism. To this end, consider the **Set**-endofunctor $S' = \coprod_{s \in \mathcal{S}} X_s \times Id^{ar(s)}$. It is immediate to see that $(X + S')U = U((X, \Sigma) + S)$, where $U \colon \mathbf{Meas} \to \mathbf{Set}$ is the forgetful functor forgetting the σ-algebra structure of a measurable space.

We prove that $UA \colon \mathbf{Ord} \to \mathbf{Set}$ is the initial sequence for $X + S'$. Clearly, since both A and U preserves colimits, also UA preserves them. Moreover, for all ordinals $\gamma \leq \beta$ the following holds

$$UA(0) = U\mathbf{0} = \mathbf{0}$$
$$UA(\beta{+}1) = U((X, \Sigma) + S)A(\beta) = (X + S')UA(\beta)$$
$$UA(\gamma{+}1 \to \beta{+}1) = U((X, \Sigma) + S)A(\gamma \to \beta) = (X + S')UA(\gamma \to \beta)\,.$$

Therefore $UA \colon \mathbf{Ord} \to \mathbf{Set}$ is the initial sequence of $X + S'$. Recall that polynomial functors in **Set** are ω-cocontinuous, that is, preserves colimits of ω-sequences. Therefore the initial sequence UA of $X + S'$ stabilizes at ω, thus, $UA(\omega \to \omega{+}1)$ is an isomorphism and, in particular, also an epimorphism. Since U reflects epimorphisms, $A(\omega \to \omega{+}1)$ is an epic arrow in **Meas**.

Polynomial endofunctors in **Meas** preserves epimorphism, moreover **Meas** is cowell-powered and has $(\mathbf{Epic}, \mathbf{Emb})$ as a factorization system. Thus, by Lemma **??**, A stabilizes at some ordinal $\kappa \geq \omega$ hence, $(A(\kappa), A(\kappa \to \kappa{+}1)^{-1})$ is an initial $((X, \Sigma) + S)$-algebra (Theorem **??**).

To prove that $((TX, \Sigma_{TX}), [\eta_X, \psi_X])$ is isomorphic to $(A(\kappa), A(\kappa \to \kappa{+}1)^{-1})$, we exploit the connection between the the initial sequences A and UA. The key observation is that both $(TX, [\eta_X, \psi_X])$ and $(UA(\kappa), UA(\kappa \to \kappa{+}1)^{-1})$ are initial $(X + S')$-algebras, so that, denoting

$A(\kappa) = (I, \Sigma_I)$ and $\iota = A(\kappa \to \kappa{+}1)^{-1}$ we have that the $((X, \Sigma) + S)$-homomorphism $\varphi \colon I \to TX$ given by initiality of $((I, \Sigma_I), \iota)$ is a bijection:

$$
\begin{array}{ccc}
(X, \Sigma) + S(I, \Sigma_I) \xrightarrow{\ \iota\ } (I, \Sigma_I) & \quad & X + S'I \xrightarrow{\ \iota\ } I \\[2pt]
(X,\Sigma)+S'\varphi \downarrow \qquad\qquad \downarrow \varphi & \xmapsto{\ U\ } & {}_{X+S'\varphi} \downarrow \qquad\qquad \downarrow \varphi \\[2pt]
(X, \Sigma) + S(TX, \Sigma_{TX}) \xrightarrow[{[\eta_X, \psi_X]}]{} (TX, \Sigma_{TX}) & \quad & X + S'TX \xrightarrow[{[\eta_X, \psi_X]}]{} TX
\end{array}
$$

So $(I, \Sigma_I) \cong (\varphi(I), \varphi(\Sigma_I)) = (TX, \varphi(\Sigma_I))$, and $((TX, \varphi(\Sigma_I)), [\eta_X, \psi_X])$ is an isomorphic initial $(X, \Sigma) + S$-algebras. By initiality, there exists $\varphi' \colon TX \to TX$ such that the following diagrams commute

$$
\begin{array}{ccc}
(X, \Sigma) + S(TX, \varphi(\Sigma_I)) \xrightarrow{[\eta_X, \psi_X]} (TX, \varphi(\Sigma_I)) & \quad & X + S'TX \xrightarrow{[\eta_X, \psi_X]} TX \\[2pt]
(X,\Sigma)+S'\varphi' \downarrow \qquad\qquad \downarrow \varphi' & \xmapsto{\ U\ } & {}_{X+S'\varphi'} \downarrow \qquad\qquad \downarrow \varphi' \\[2pt]
(X, \Sigma) + S(TX, \Sigma_{TX}) \xrightarrow[{[\eta_X, \psi_X]}]{} (TX, \Sigma_{TX}) & \quad & X + S'TX \xrightarrow[{[\eta_X, \psi_X]}]{} TX
\end{array}
$$

but, by unicity of the initial $(X + S')$-homomorphism $\varphi' = id_{TX}$. Since φ' is measurable, we have that for all $E \in \Sigma_{TX}$, $id^{-1}(E) = \varphi^{-1}(E) \in \varphi(\Sigma_I)$, hence $\Sigma_{TX} \subseteq \varphi(\Sigma_I)$. By Lambek's lemma, $[\eta_X, \psi_X]$ is and isomorphism between $(X, \Sigma) + S(TX, \varphi(\Sigma_I))$ and $(TX, \varphi(\Sigma_I))$ so it is also an embedding, and $\varphi(\Sigma_I)$ is the smallest σ-algebra rendering $[\eta_X, \psi_X]$ measurable. Thus, by the fact that $[\eta_X, \psi_X] \colon (X, \Sigma) + S(TX, \Sigma_{TX}) \to (TX, \Sigma_{TX})$ is measurable and Σ_{TX} is contained in $\varphi(\Sigma_I)$, we have that $\Sigma_{TX} = \varphi(\Sigma_I)$. This proves that $((TX, \Sigma_{TX}), [\eta_X, \psi_X])$ is isomorphic to the initial $((X, \Sigma) + S)$-algebra, hence it is initial too. ∎

By Theorem 6.1.4, for any syntactic endofunctor S associated with a measurable signature $(\mathcal{S}, ar, \{(X_s, \Sigma_s)\}_{s \in \mathcal{S}})$, the forgetful functor $U^S \colon S\text{-}\mathbf{alg} \to \mathbf{Meas}$ has a left adjoint $L^S \colon \mathbf{Meas} \to S\text{-}\mathbf{alg}$ defined as follows, for any measurable space (X, Σ), and measurable function $f \colon (X, \Sigma_X) \to (Y, \Sigma_Y)$,

$$
\begin{aligned}
L^S(X, \Sigma) &= ((TX, \Sigma_{TX}), \psi_X) \\
L^S f &= (f^{\#} \colon TX \to TY)
\end{aligned}
$$

where $f^{\#}$ is the unique $((X, \Sigma) + S)$-homomorphism from the initial algebra $(TX, \Sigma_{TX}, [\eta_X, \psi_X])$ to $((TY, \Sigma_{TY}), [f \circ \eta_Y, \psi_Y])$.

Next we give a more explicit characterization of the monad arising from the adjunction $U^S \dashv L^S$. Remarkably, due to Theorem 6.1.4 this monad is essentially defined as the term monad freely generated by syntactic **Set**-endofunctors arising from set-signatures.

Definition 6.1.5 (Term monad) *Let $(\mathcal{S}, ar, \{(X_s, \Sigma_s)\}_{s \in \mathcal{S}})$ be a measurable signature and S be the syntactic **Meas**-endofunctor associated with it. The monad freely generated by S, called mea-surable term monad over S, is given by the triple (T^S, η^S, μ^S), where $T^S \colon \mathbf{Meas} \to \mathbf{Meas}$ is defined as follows, for (X, Σ) a measurable space and $f \colon (X, \Sigma_X) \to (Y, \Sigma_Y)$ a measurable map*

$$
T^S(X, \Sigma) = (TX, \Sigma_{TX})
$$

$$
T^S f(x) = f(x), \qquad T^S f(s(k, t_1, \ldots t_{ar(s)})) = s(k, T^S f(t_1), \ldots, T^S f(t_{ar(s)})),
$$

where $x \in X$, $s \in \mathcal{S}$, and $k \in X_s$, $t_1, \ldots, t_{ar(s)} \in TX$; with unit $\eta^S_X = \eta_X \colon (X, \Sigma) \to (TX, \Sigma_{TX})$ (the insertion-of-variables function); and multiplication $\mu^S_X \colon (TTX, \Sigma_{TTX}) \to (TX, \Sigma_{TX})$ (the operation which allows one to plug measurable terms into contexts) inductively defined as follows

$$
\mu^S_X(t) = t, \qquad \mu^S_X(s(k, C_1, \ldots C_{ar(s)})) = s(k, \mu^S(C_1), \ldots, \mu^S(C_{ar(s)})),
$$

for all $t \in TX$, $s \in \mathcal{S}$, $k \in X_s$, and $C_1, \ldots, C_{ar(s)} \in TTX$ (i.e., contexts).

6.2   Measure GSOS Specification Rule Format

We are now going to present a concrete well-behaved specification rule format for stochastic transition systems with continuos state space as a particular instance of abstract GSOS distributive laws of [83].

Briefly, our approach consists in instantiating the bialgebraic framework of Turi-Plotkin [83] to L-labelled Markov kernels, that is, to the coalgebras for the functor $\Delta_{<\infty}^L \colon \mathbf{Meas} \to \mathbf{Meas}$ (Proposition 5.2.2). The key intuition behind the bialgebraic framework is that rule specification systems can be formulated in terms of certain natural transformations, called *distributive laws*. The models for these distributive laws are *bialgebras*, that is, a pair consisting of a T-algebra $\alpha \colon TX \to X$ and a D-coalgebra $\beta \colon X \to DX$ on the same carrier and such that they are related by a distributive law $\lambda \colon TD \Rightarrow DT$ of a monad T over a comonad D as follows:

$$
\begin{array}{ccccc}
TX & \xrightarrow{\;\;\alpha\;\;} & X & \xrightarrow{\;\;\beta\;\;} & DX \\
{\scriptstyle T\beta}\big\downarrow & & & & \big\uparrow{\scriptstyle D\alpha} \\
TDX & & \xrightarrow{\;\;\lambda_X\;\;} & & DTX
\end{array}
$$

Intuitively, the monad T represents the syntax of the programming language and the comonad D models the shape of computations. The algebra $\alpha \colon TX \to X$ and coalgebra $\beta \colon X \to DX$, respectively, denote the denotational and operational models of the system, and the distributive law $\lambda \colon TD \Rightarrow DT$ explains how the syntax distributes over the computations, that is to say, how the computation of syntactic operator depends on the executions of its arguments. Bialgebras form a category, where any unique morphism from the initial object represent a denotational semantics, and any unique morphism to the final one an operational semantics. Hence, one can always find a canonical fully-abstract semantics: the universal morphism from the initial to the final bialgebras.

The distributive laws we are interested in are those of type $S(Id \times \Delta_{<\infty}^L) \Rightarrow (\Delta_{<\infty}T^S)^L$, that is, abstract GSOS distributive laws of a monad T^S over the copointed functor $(Id \times \Delta_{<\infty}^L)$, where T^S is freely generated by a syntactic functor $S \colon \mathbf{Meas} \to \mathbf{Meas}$ associated with some measurable signature. Our aim is to describe these distributive laws by means of a set of derivation rules similar to the well-know GSOS format of Bloom et al. [17].

In the GSOS format, the target of a transition is a term built from the components of the source process, and their corresponding semantics. In our settings, the target of a transition is not a term, but a (finite) measure over a generic measurable space, hence the derivations of the stochastic transitions becomes more complicated. We cope with this problem proposing labelled transitions of the form $t \xrightarrow{a} \mu$, where μ is no more a measure but a syntactic expression intended to denote a measure, which we call *measure term*.

Definition 6.2.1 *Let* $(\mathcal{S}, ar_S, \{(X_s, \Sigma_s)\}_{s \in \mathcal{S}})$ *and* $(\mathcal{M}, ar_M, \{(X'_m, \Sigma'_m)\}_{m \in \mathcal{M}})$ *be the measurable signatures for processes and measure terms, respectively. An* MGSOS *rule over them, for a finite set L of labels, is an expression of the form*

$$
\frac{\left\{ x_i \xrightarrow{a_{ij}} \mu_{ij} \right\}_{1 \le i \le ar_S(s),\, a_{ij} \in A_i}^{1 \le j \le m_i} \quad \left\{ x_i \xrightarrow{\;b\;}\!\!\!\!\!/ \;\; \right\}_{1 \le i \le ar_S(s)}^{b \in B_i}}{s(k, x_1, \ldots, x_{ar_S(s)}) \xrightarrow{c} \mu}
$$

where

- $s \in S$, $k \in X_s$*;*

- $\{x_i \mid 1 \le i \le ar_S(s)\}$ *and* $\{\mu_{ij} \mid 1 \le i \le ar_S(s),\, 1 \le j \le m_i\}$ *are pairwise distinct* process *and* measure *term variables, respectively;*

- $A_i \cap B_i = \emptyset$ *are disjoint subsets of labels in L, for all $1 \le i \le n$, and $c \in L$;*

- μ *is a measurable term for the signature* $(\mathcal{M}, ar_M, \{(X'_m, \Sigma'_m)\}_{m \in \mathcal{M}})$ *with variables in* $\{x_i \mid 1 \le i \le ar_S(s)\}$ *and* $\{\mu_{ij} \mid 1 \le i \le ar_S(s),\, 1 \le j \le m_i\}$.

Note that, differently from the standard GSOS rule format of [17], in the premises one is not allowed to use the same label twice for the same variable x_i.

Measure terms have to be interpreted as actual measure (over process terms) by means of some suitable interpretation function. Has one may aspect, not all interpretations guarantees that behavioral equivalence is a congruence. A condition which ensures that these are well-behaved is they are natural transformations of a particular type:

Definition 6.2.2 *Let $(\mathcal{S}, ar_S, \{(X_s, \Sigma_s)\}_{s \in \mathcal{S}})$ and $(\mathcal{M}, ar_M, \{(X'_m, \Sigma'_m)\}_{m \in \mathcal{M}})$ respectively be the measurable signatures for processes and measure terms, and $S, M \colon \mathbf{Meas} \to \mathbf{Meas}$ be the syntactic functors associate with them. A* measure term interpretation *over these signatures is a natural transformation of type $T^M \Delta_{<\infty} \Rightarrow \Delta_{<\infty} T^S$, where T^S and T^M are the free monads over S and M respectively.*

The operational specification is given by a set of rules together with a description of how measures must be combined.

Definition 6.2.3 *Let $(\mathcal{S}, ar_S, \{(X_s, \Sigma_s)\}_{s \in \mathcal{S}})$ and $(\mathcal{M}, ar_M, \{(X'_m, \Sigma'_m)\}_{m \in \mathcal{M}})$ be the measurable signatures for processes and measure terms, respectively.*

An MGSOS specification system *is a pair $(\mathcal{R}, \langle\!\langle \cdot \rangle\!\rangle)$, such that $\mathcal{R}$ is a set of image finite MGSOS rules and $\langle\!\langle \cdot \rangle\!\rangle \colon T^M \Delta_{<\infty} \Rightarrow \Delta_{<\infty} T^S$ is a measure term interpretation over the process and measure term signatures.*

Similarly to GSOS transition systems specifications, also MGSOS specification systems define a structural operational semantics, but in this particular case in the form of a labelled Markov kernel over the measurable space of process terms. Its definition can be summarized in two stages. First, an image finite labelled transition system $(T^S\mathbf{0}, \{\xrightarrow{\alpha} \subseteq T^S\mathbf{0} \times T^M(T^S\mathbf{0})\}_{\alpha \in L})$ is (inductively) defined from the set of MGSOS derivation rules, then the associated Markov kernel is obtained evaluating measure terms to measures. Formally, the associated $\Delta_{<\infty}^L$-coalgebra γ over T^S0 is defined, for $\alpha \in L$ and $t \in T^S\mathbf{0}$, by

$$\gamma(t)(\alpha) = \oplus_{T^S\mathbf{0}}\big(\{\langle\!\langle \mu \rangle\!\rangle_{T^S\mathbf{0}} \mid t \xrightarrow{\alpha} \mu\}\big).$$

where, for any given finite set of U measures over a measurable space (X, Σ), we define $\oplus_X(U) \colon \Sigma \to [0, \infty)$ by $\oplus_X(U)(E) = \sum_{\mu \in U} \mu(E)$, which is easily seen to be a finite measure over (X, Σ).

The main advantage in using this two staged definition is that Markov processes are defined syntactically by structural induction on process terms.

The next theorem shows how MGSOS specification systems induce natural transformation of type $S(Id \times \Delta_{<\infty}^L) \Rightarrow (\Delta_{<\infty}T^S)^L$, i.e., abstract GSOS laws of [83].

Theorem 6.2.4 *An MGSOS specification system $(\mathcal{R}, \langle\!\langle \cdot \rangle\!\rangle)$ over the measurable signatures for processe and measure terms $(\mathcal{S}, ar_S, \{(X_s, \Sigma_s)\}_{s \in \mathcal{S}})$ and $(\mathcal{M}, ar_M, \{(X'_m, \Sigma'_m)\}_{m \in \mathcal{M}})$, where S and M are the associated syntactic functors, and set of labels L, determines a natural transformation of type $S(Id \times \Delta_{<\infty}^L) \Rightarrow (\Delta_{<\infty}T^S)^L$.*

Proof. For any measurable space (X, Σ), define the (set) function $[\![\mathcal{R}]\!]_X$ as the composite:

$$S(X \times (\Delta_{<\infty}X)^L) \xrightarrow{\nu_X} (\mathcal{P}_{\mathrm{fin}}T^M\Delta_{<\infty}X)^L \xrightarrow{(\oplus T^S \circ \mathcal{P}_{\mathrm{fin}}\langle\!\langle \cdot \rangle\!\rangle)_X^L} (\Delta_{<\infty}T^SX)^L,$$

where, ν_X is defined as follows: for all $\mu' \in T^M \Delta_{<\infty}X$, $c \in L$, $s \in S$, $k \in X_s$, $x_i \in X$, and $\beta_i \in (\Delta_{<\infty}X)^L$, put

$$\mu' \in \nu_X(s(k, (x_1, \beta_1), \ldots, (x_{ar_S(s)}, \beta_{ar_S(s)})))(c)$$

if and only if there exists a (possibly renamed) rule in $\mathcal{R}$ of the form

$$\frac{\{x_i \xrightarrow{a_{ij}} \mu_{ij}\}_{1 \le i \le ar_S(s), \, a_{ij} \in A_i}^{1 \le j \le m_i} \quad \{x_i \xrightarrow{b} \!\!\!\!\!/\,\}_{1 \le i \le ar_S(s)}^{b \in B_i}}{s(k, x_1, \ldots, x_{ar_S(s)}) \xrightarrow{c} \mu}$$

such that $\beta_i(b) = \mathbf{0}$ (the zero measure), for $b \in B_i$, and $\mu' = (T^M\sigma)(\mu)$ for a substitution map σ such that $\sigma(x_i) = \delta_{x_i}$ (the Dirac measure at x_i) and $\sigma(\mu_{ij}) = \beta_i(a_{ij})$.

Naturality of $[\![\mathcal{R}]\!]$ is proved separately for the two components. To prove the naturality of ν one proceeds as in [83, Th. 1.1]. The composite $(\oplus T^S \circ \mathcal{P}_{\text{fin}} (\!\!\{\,\cdot\,\}\!\!))^L$, is natural since $(\!\!\{\,\cdot\,\}\!\!)$ and $\oplus$ are natural.

As for measurability of $[\![\mathcal{R}]\!]_X$, it suffices to check that $[\![\mathcal{R}]\!]_X^{-1}(U_\alpha[E])$ is measurable in $S(X \times (\Delta_{<\infty}X)^L)$, for $U_\alpha^E = \{\beta' \in (\Delta_{<\infty}T^SX)^L \mid \beta'(\alpha) \in E\}$, where $\alpha \in L$, and $E \in \Sigma_{\Delta_{<\infty}T^SX}$.

Let $s(k, \overline{(x,\beta)})$ abbreviates $s(k, (x_1, \beta_1), \ldots, (x_n, \beta_n)) \in S(X \times (\Delta_{<\infty}X)^L)$, then

$$
\begin{aligned}
[\![\mathcal{R}]\!]_X^{-1}(U_\alpha^E) &= \left\{ s(k, \overline{x,\beta}) \mid [\![\mathcal{R}]\!]_X(s(k, \overline{x,\beta})) \in U_\alpha^E \right\} \\
&= \left\{ s(k, \overline{x,\beta}) \mid ((\oplus T^S \circ \mathcal{P}_{\text{fin}}(\!\!\{\,\cdot\,\}\!\!))^L \circ \nu)_X(s(k, \overline{x,\beta})) \in U_\alpha^E \right\} \\
&= \left\{ s(k, \overline{x,\beta}) \mid ((\oplus T^S \circ \mathcal{P}_{\text{fin}}(\!\!\{\,\cdot\,\}\!\!))^L \circ \nu)_X(s(k, \overline{x,\beta}))(\alpha) \in E \right\} \\
&= \left\{ s(k, \overline{x,\beta}) \mid (\oplus T^S \circ \mathcal{P}_{\text{fin}}(\!\!\{\,\cdot\,\}\!\!) \circ \nu^\alpha)_X(s(k, \overline{x,\beta})) \in E \right\}
\end{aligned}
$$

where $\nu_X^\alpha \triangleq \nu_X(\cdot)(\alpha)$ is the specialization of ν_X on a fixed $\alpha \in L$,

$$
\begin{aligned}
&= (\oplus T^S \circ \mathcal{P}_{\text{fin}}(\!\!\{\,\cdot\,\}\!\!) \circ \nu^\alpha)_X^{-1}(E) \\
&= (\nu_X^\alpha)^{-1} \circ (\oplus T^S \circ \mathcal{P}_{\text{fin}}(\!\!\{\,\cdot\,\}\!\!))_X^{-1}(E)
\end{aligned}
$$

Now is easy to prove measurability. Since $(\!\!\{\,\cdot\,\}\!\!)$ is measurable and sums and products of measurable functions is measurable, $(\oplus T^S \circ \mathcal{P}_{\text{fin}}(\!\!\{\,\cdot\,\}\!\!))_X^{-1}(E) \in \Sigma_{\mathcal{P}_{\text{fin}}T^M\Delta_{<\infty}X}$.

To prove measurability of ν_X^α we need only to check that $(\nu_X^\alpha)^{-1}(\bigcup_{j=0}^k \{E_j\})$ is a measurable set in $S(X \times (\Delta_{<\infty}X)^L)$, for $E_j \in \Sigma_{T^M\Delta_{<\infty}X}$, $0 \le j \le k$.

$$
\begin{aligned}
(\nu_X^\alpha)^{-1}(\textstyle\bigcup_{j=0}^k \{E_j\}) &= \left\{ s(k, \overline{x,\beta}) \mid \nu_X^\alpha(s(k, \overline{x,\beta})) \in \textstyle\bigcup_{j=0}^k \{E_j\} \right\} \\
&= \textstyle\bigcup_{j=0}^k \left\{ s(k, \overline{x,\beta}) \mid \nu_X^\alpha(s(k, \overline{x,\beta})) = E_j \right\}
\end{aligned}
$$

But $\nu_X^\alpha(s(k, \overline{x,\beta})) = E_j$ iff there exists some rule in $\mathcal{R}$ with conclusion of the form $f(x_1, \ldots, x_n) \xrightarrow{\alpha} \mu$ and $(T^M\sigma)(\mu) \in E_j$ (obviously, all the other conditions given above have to be satisfied too). By construction, σ is defined by sums of Dirac measures δ_X and $ev_{a_{ij}}$, which are measurable. Therefore, $T^M\sigma$ is measurable, and as a consequence also ν_X^α is measurable. ∎

In the proof above, measure terms variables are interpreted as Dirac measures via the natural transformation $\delta \colon Id \Rightarrow \Delta_{<\infty}$ (hence, measurable in each component). This together with the assumption that $(\!\!\{\,\cdot\,\}\!\!) \colon T^M\Delta_{<\infty} \Rightarrow \Delta_{<\infty}T^S$ is a natural transformation are crucial to prove measurability of $[\![\mathcal{R}]\!]$.

Remark 6.2.5 To establish a correspondence between abstract GSOS distributive laws and concrete rule formats, in [15] Bartels proposed an elegant decomposition strategy to recover congruential specification systems from distributive laws. The method uses a collection of representation lemmas for distributive laws, which allows to explain natural transformations of complex type in terms of collections of natural transformations of simpler type. Using this technique Bartels has been able to give the first detailed proof of correspondence between natural transformations of type $S(Id \times \mathcal{P}_{fin}^L) \Rightarrow (\mathcal{P}_{fin}T^S)^L$ for labelled transition systems and GSOS specification systems. Moreover, he further extended the technique in order to derive from scratch a rule specification formats for discrete probabilistic transition systems, the so called PGSOS. The very same technique has been applied also by Klin and Sassone [61] in the definition of a rule format for stochastic transition systems (with discrete state space).

Unfortunately, this method applies only in the category **Set** and cannot be ported easily to **Meas**. This is due to the fact that many of the decomposition lemmas used in [15] require objects

to be represented as indexed coproducts of simpler canonical subobjects. This clearly works in **Set**, since, for any set X, the isomorphism $X \cong \coprod_{x \in X} \{x\}$ holds, but it does not work in **Meas**, due to the presence of the σ-algebra structure in the objects. ∎

In [83], it is shown that distributive laws $\rho \colon S(Id \times B) \Rightarrow BT^S$, where T^S is the free monad over S, for endofunctors S and B admitting initial S-algebra $(T^S\mathbf{0}, \alpha)$ and final B-coalgebra (F, ω), give rise to a unique B-coalgebra structure $\beta_\rho \colon T^S\mathbf{0} \to BT^S\mathbf{0}$ such that $(T^S\mathbf{0}, \alpha, \beta_\rho)$ is the initial ρ-bialgebra. Dually, there is a unique S-algebra structure $\alpha_\rho \colon SF \to F$ such that (F, α_ρ, β) is the final ρ-bialgebra. The unique (both by initiality and finality) homomorphism from the initial to the final ρ-model is both the initial and final semantics for ρ, and it is called *universal semantics* for ρ. Note, that two B-coalgebras have the same universal semantics if and only if they are behavioural equivalent, therefore B-behavioural equivalence is an S-congruence.

As for abstract GSOS laws of type $S(Id \times \Delta_{<\infty}^L) \Rightarrow (\Delta_{<\infty}T^S)^L$ we have the following result for behavioural equivalence on probabilistic processes on **Meas**.

Theorem 6.2.6 *Behavioural equivalence on the $\Delta_{<\infty}^L$-coalgebras over $T^S\mathbf{0}$ inductively induced by MGSOS specification systems over S and M, and set of labels L, is an S-congruence.*

Proof. By Theorem 6.1.4 every syntactic functor $S \colon \mathbf{Meas} \to \mathbf{Meas}$ have initial algebra. As for the endofunctor $\Delta_{<\infty}^L \colon \mathbf{Meas} \to \mathbf{Meas}$, the existence of the final coalgebra is given by Theorem 4.1.20 and Remark 4.1.22. The thesis follows by Theorem 6.2.4 and [83, Cor. 7.3]. ∎

6.3 Measure terms evaluations

MGSOS specifications require the evaluations $\langle\!\langle \cdot \rangle\!\rangle_X \colon T^M \Delta_{<\infty} X \to \Delta_{<\infty} T^S X$ to be both natural and measurable. This time consuming check is overcome in this section, where we provide a rather easy (and general) technique for defining natural measure terms evaluation functions. This technique in based on a proof principle dual to the "coiterative proof principle" described in [15].

We start giving a recursion lemma dual to [15], which generalizes the standard induction proof principle by means of a simple distributive law λ.

Lemma 6.3.1 *Let $S, B \colon \mathbf{C} \to \mathbf{C}$ be functors on a category with countable products, (A, α) be the initial S-algebra, and $\lambda \colon SB \Rightarrow BS$ be a (simple) distributive law. For any SB-algebra (X, φ) there exists a unique λ-iterative arrow $f \colon A \to X$ induced by φ, such that the following diagrams commute*

$$
\begin{array}{ccccc}
SBX & \xleftarrow{\ \ SBf\ \ } & SBA & \xleftarrow{\ S\beta_\lambda\ } & SA \\
{\scriptstyle \varphi}\downarrow & & & & \downarrow{\scriptstyle \alpha} \\
X & \xleftarrow{\quad\quad\quad f\quad\quad\quad} & & & A
\end{array}
\qquad
\begin{array}{ccccc}
SA & \xrightarrow{\ \alpha\ } & A & \xrightarrow{\ \beta_\lambda\ } & BA \\
{\scriptstyle S\beta_\lambda}\downarrow & & & & \uparrow{\scriptstyle B\alpha} \\
SBA & \xrightarrow{\quad\quad \lambda_A \quad\quad} & & & BSA
\end{array}
$$

Proof. Dualize [15, Theorem 4.2.2]. ∎

We denote this induction proof principle by λ-*iteration proof principle*, and we call $f \colon A \to X$ as the λ-*iterative arrow induced by* φ. Note that, the diagram on the right is the initial λ-model, and in particular β_λ is uniquely determined by (standard) induction on the S-algebra $(BA, B\alpha \circ \lambda_A)$.

The λ-iteration proof principle of Lemma 6.3.1 can be extended as a proof principle on the monad T^S freely generated by S as follows:

Proposition 6.3.2 (Structural λ-iteration) *Let $S, B \colon \mathbf{C} \to \mathbf{C}$ be functors on a category with binary coproducts and countable products, (T^S, η^S, μ^S) be the free monad over S, $\lambda \colon SB \Rightarrow BS$ be a distributive law, and $\psi_X \colon ST^S X \to T^S X$ be the free S-algebra structure over X. For any SB-algebra (Y, φ), B-coalgebra (X, k), and arrow $\phi \colon X \to Y$, there exists a unique arrow $f \colon T^S X \to Y$*

such that the following diagrams commute

$$
\begin{array}{ccccc}
SBY & \xleftarrow{\ SBf\ } & SBT^S X & \xleftarrow{\ S\beta_\lambda\ } & ST^S X \\
\varphi \downarrow & & & & \downarrow \psi_X \\
Y & \xleftarrow{\quad\quad f\quad\quad} & & T^S X & \\
& & & \uparrow \eta^S_X & \\
& \xleftarrow[\ \phi\ ]{} & & X &
\end{array}
\qquad
\begin{array}{ccccc}
X & \xrightarrow{\ \eta^S_X\ } & T^S X & \xleftarrow{\ \psi_X\ } & ST^S X \\
k \downarrow & & \downarrow \beta_\lambda & & \downarrow S\beta_\lambda \\
BX & \xrightarrow[\ B\eta^S_X\ ]{} & BT^S X & \xleftarrow[\ B\psi_X \circ \lambda_{T^S X}\ ]{} & SBT^S X
\end{array}
$$

Proof. First, notice that β_λ is the inductive extension of the $(X+S)$-algebra structure on $BT^S X$ given by the copair $[B\eta^S_X \circ k, B\psi_X \circ \lambda_{T^S X}]$, along the initial $(X+S)$-algebra structure $[\eta^S_X, \psi_X]$ on $T^S X$. Now, define the distributive law $\lambda'\colon (X+S)B \Rightarrow B(X+S)$ as $\lambda'_Y = [Bin_l, Bin_r] \circ (k + \lambda_Y)$ (the proof of naturality is straightforward). By definition of λ' we have

$$[B\eta^S_X \circ k, B\psi_X \circ \lambda_{T^S X}] = B[\eta^S_X, \psi_X] \circ \lambda'_{T^S X}.$$

Therefore, by unicity of the inductive extension, $(T^S X, \psi_X, \beta_\lambda)$ turns out to be a λ'-model on $T^S X$. This allows to apply Lemma 6.3.1 on the $(X+S)B$-algebra structure $(Y, [\phi, \varphi])$ obtaining a unique λ'-iterative arrow $f\colon T^S X \to Y$ making the required diagrams above commute. ∎

We denote this proof principle by *structural λ-induction proof principle*, and we say that f is the *λ-iterative extension of φ along the (pair of) valuations ϕ and k*. Note that, the diagram on the right define β_λ as the structural inductive extension of $B\psi_X \circ \lambda_{T^S X}$ along $B\eta^S_X \circ k$.

Proposition 6.3.2 can be turned into an induction proof principle on natural transformations in the following way:

Corollary 6.3.3 *Let $S, B, F, G\colon \mathbf{C} \to \mathbf{C}$ be functors on a category with binary coproducts and countable products, (T^S, η^S, μ^S) be the free monad over S, $\lambda\colon SB \Rightarrow BS$ be a (simple) distributive law. For any $\varphi\colon SBF \Rightarrow F$, $k\colon Id \Rightarrow B$, and $\phi\colon G \Rightarrow F$, there exist unique natural transformations $\beta_\lambda\colon T^S \Rightarrow BT^S$ and $f\colon T^S \Rightarrow F$ such that the following (natural) diagrams commute*

$$
\begin{array}{ccccc}
SBF & \xleftarrow{\ SBf\ } & SBT^S G & \xleftarrow{\ S\beta_\lambda G\ } & ST^S G \\
\varphi \downarrow & & & & \downarrow \psi G \\
F & \xleftarrow{\quad\quad f\quad\quad} & & T^S G & \\
& & & \uparrow \eta^S G & \\
& \xleftarrow[\ \phi\ ]{} & & G &
\end{array}
\qquad
\begin{array}{ccccc}
Id & \xrightarrow{\ \eta^S\ } & T^S & \xleftarrow{\ \psi\ } & ST^S \\
k \downarrow & & \downarrow \beta_\lambda & & \downarrow S\beta_\lambda \\
B & \xrightarrow[\ B\eta^S\ ]{} & BT^S & \xleftarrow[\ B\psi \circ \lambda T^S\ ]{} & SBT^S
\end{array}
$$

Proof. We first prove naturality of β_λ, i.e., that for any morphism $g\colon X \to Y$, $(\beta_\lambda)_Y \circ T^S g = BT^S g \circ (\beta_\lambda)_X$. The commuting diagrams

$$
\begin{array}{ccccc}
X & \xrightarrow{\ \eta^S_X\ } & T^S X & \xleftarrow{\ \psi_X\ } & ST^S X \\
k_X \downarrow & \text{\color{red}(by def.)} & (\beta_\lambda)_X \downarrow & \text{\color{red}(by def.)} & \downarrow S(\beta_\lambda)_X \\
BX & \xrightarrow{\ B\eta^S_X\ } & BT^S X & \xleftarrow{\ B\psi_X \circ \lambda_{T^S X}\ } & SBT^S X \\
Bg \downarrow & \text{\color{red}(nat. } \eta^S\text{)} & BT^S g \downarrow & \text{\color{red}(nat. } \psi \text{ and } \lambda\text{)} & \downarrow SBT^S g \\
BY & \xrightarrow[\ B\eta^S_Y\ ]{} & BT^S Y & \xleftarrow[\ B\psi_Y \circ \lambda_{T^S Y}\ ]{} & SBT^S Y
\end{array}
$$

and

$$
\begin{array}{ccccc}
X & \xrightarrow{\eta_X^S} & T^S X & \xleftarrow{\psi_X} & ST^S X \\
\downarrow g & \text{(nat. } \eta^S) & \downarrow T^S g & \text{(nat. } \psi) & \downarrow ST^S g \\
Y & \xrightarrow{\eta_Y^S} & T^S Y & \xleftarrow{\psi_Y} & ST^S Y \\
\downarrow k_Y & \text{(by def.)} & \downarrow (\beta_\lambda)_Y & \text{(by def.)} & \downarrow S(\beta_\lambda)_Y \\
BY & \xrightarrow{B\eta_Y^S} & BT^S Y & \xleftarrow{B\psi_Y \circ \lambda_{T^S Y}} & SBT^S Y
\end{array}
$$

assert that $(\beta_\lambda)_Y \circ T^S g$ and $BT^S g \circ (\beta_\lambda)_X$ are both inductive extensions of $B\psi_Y \circ \lambda_{T^S Y}$ along $B\eta_X^S \circ Bg \circ k_X$, hence they necessarily coincide. Naturality of f follows similarly by unicity of the λ-iterative extension.

Indeed, both $f_Y \circ Tg$ and $Fg \circ f_X$ are λ-iterative extensions of φ_Y along the pair of valuations $Fg \circ \phi_X$ and k_{GX}, as proved by the commutative diagrams above. ∎

6.4 Examples of MGSOS specifications

To illustrate the expressiveness of the MGSOS format, we present specifications for some simple process calculi which extend Milner's CCS [65] (without restriction) with continuous data information. Remarkably, the MGSOS format allows for a very simple presentation of continuos state stochastic semantics, ensuring important properties (e.g., congruence) that are notoriously difficult to obtain even in the discrete case.

6.4.1 Quantitative CCS

In this section we present a CCS-like process calculus able to model continuos occurrences of agents. The syntax is defined as follows:

$$
P, Q ::= \mathsf{nil} \mid \alpha.P \mid P + Q \mid P \parallel Q \mid c \text{ of } P
$$

$$\frac{}{\alpha.x \xrightarrow{\alpha} x} \qquad \frac{x \xrightarrow{\alpha} \mu}{x + x' \xrightarrow{\alpha} \mu} \qquad \frac{x' \xrightarrow{\alpha} \mu}{x + x' \xrightarrow{\alpha} \mu}$$

$$\frac{x \xrightarrow{\alpha} \mu}{c \text{ of } x \xrightarrow{\alpha} U_c(\mu)} \qquad \frac{x \xrightarrow{\alpha} \mu}{x \parallel x' \xrightarrow{\alpha} \mu \mid x'} \qquad \frac{x' \xrightarrow{\alpha} \mu}{x \parallel x' \xrightarrow{\alpha} x \mid \mu}$$

$$\frac{x \xrightarrow{a} \mu \quad x' \xrightarrow{\overline{a}} \mu'}{x \parallel x' \xrightarrow{\tau} \mu \mid \mu'} \qquad \frac{x \xrightarrow{\overline{a}} \mu \quad x' \xrightarrow{a} \mu'}{x \parallel x' \xrightarrow{\tau} \mu \mid \mu'}$$

Figure 6.1: MGSOS specification system for the Quantitative CCS.

$$\alpha ::= a \mid \overline{a} \mid \tau$$

where $c \in \mathbb{R}_{\geq 0}$ represents a positive concentration, and $a \in A$ is an action label taken from a finite set A. The nil operator denotes the null process, $\alpha.P$ the action prefix, $P + Q$ the stochastic choice operator, $P \parallel Q$ the parallel composition. The *concentration* operator c **of** P models a process with a continuous number c of occurrences of the agent P.

As for the semantics we aim to give to processes a stochastic baheviour which is faithful with the standard non-deterministic one of CCS, where the execution rate of each action depends on the availability of the agents that may perform it. The introduction of the concentration operator c **of** P opens several problems that cannot be solved by a discrete state semantics. Indeed, a discrete semantics forces to decide a priory the quantity of P to be consumed in c **of** P, with a rule of the following form

$$\frac{x \xrightarrow{\alpha[r]} x'}{c \text{ of } x \xrightarrow{\alpha[c' \cdot r]} c' \text{ of } x' \parallel (c - c') \text{ of } x} \tag{6.4.1}$$

where r denotes the execution rate of the stochastic a-transition in the premise, and c' denotes the concentration of the agent consumed by the transition. Having to deal with continuos concentrations, any fixed choice of $c' \leq c$ is unreasonable since the uniform probability of choosing the exact value of c' in the interval $[0, c]$ would be always zero. The only satisfactory choice is to give an actual continuous state operational semantics to the calculus. We do this by means of an MGSOS specification system.

We start by defining the measurable signatures for processes and measure terms in order to formally determine which are the σ-algebras associated with each operator symbols and, consequently, to the measurable spaces of terms freely generated over them. We do this directly by giving the syntactic functors $S, M \colon \mathbf{Meas} \to \mathbf{Meas}$ associated with the measurable signatures for processes and measure terms, respectively:

$$SX = \overbrace{\mathbf{1}}^{\text{nil}} + \overbrace{A \times X}^{a.x} + \overbrace{A \times X}^{\overline{a}.x} + \overbrace{X}^{\tau.x} + \overbrace{X \times X}^{x+x} + \overbrace{X \times X}^{x \parallel x} + \overbrace{\mathbb{R}_{\geq 0} \times X}^{c \text{ of } x},$$

$$MX = \overbrace{X \times X}^{x|x} + \overbrace{\mathbb{R}_{\geq 0} \times X}^{U_c(x)},$$

where A is the set of action labels endowed with the discrete σ-algebra, and $\mathbb{R}_{\geq 0}$ is the set of positive real numbers with Borel σ-algebra.

The stochastic semantics for the quantitative CCS is described by means of the MGSOS specification system $(\mathcal{R}^{\mathsf{q}}, \langle\!\langle \cdot \rangle\!\rangle^{\mathsf{q}})$, where the set $\mathcal{R}^{\mathsf{q}}$ of derivation rules is given in Figure 6.1, and the measure term evaluation $\langle\!\langle \cdot \rangle\!\rangle^{\mathsf{q}} \colon T^M \Delta_{<\infty} \Rightarrow \Delta_{<\infty} T^S$, is inductively defined, for $\mu, \mu' \in T^M \Delta_{<\infty} X$ and $\beta \in \Delta_{<\infty} X$, as follows:

$$\langle\!\langle \beta \rangle\!\rangle^{\mathsf{q}}_X = \beta$$

$$\langle\!\langle \mu \mid \mu' \rangle\!\rangle_X^{\mathsf{q}} = (\langle\!\langle \mu \rangle\!\rangle_X^{\mathsf{q}} \times \langle\!\langle \mu' \rangle\!\rangle_X^{\mathsf{q}}) \circ (\lambda(x, x').\, x \parallel x')^{-1}$$
$$\langle\!\langle U_c(\mu) \rangle\!\rangle_X^{\mathsf{q}} = c \cdot (U[0, c] \times \langle\!\langle \mu \rangle\!\rangle_X^{\mathsf{q}} \times \delta_X) \circ (\lambda(c', x', x).\, c' \text{ of } x' \parallel (c - c') \text{ of } x)^{-1} \, ,$$

where $U[0, c](E) = \int_{[0,c] \cap E} \frac{1}{c}\, dx$, for any Borel set $E \subseteq \mathbb{R}_{\geq 0}$, denotes the uniform probability measure over the interval $[0, c] \cap E$, δ_X is the Dirac measure, $\beta \times \beta'$ denotes the product measure, and $(r\beta)(E) := r \cdot \beta(E)$, for $0 \leq r \leq 1$.

All the rules in $\mathcal{R}^{\mathsf{q}}$ but the one for the quantitative operator are easy to understand. As for the case for c **of** x, the actual semantics is given by the measure interpretation $\langle\!\langle \cdot \rangle\!\rangle^{\mathsf{q}} \colon T^M \Delta_{<\infty} \Rightarrow \Delta_{<\infty} T^S$. Intuitively, for any measurable set E of processes, the pre-image of $(\lambda(c', x', x).\, c' \text{ of } x' \parallel (c - c') \text{ of } x)$ permits to select those triples (c', x', x) for which a term of the form $c' \text{ of } x' \parallel (c - c') \text{ of } x$ belongs to E. Since the function is measurable, the set of all such triples belongs to the σ-algebra of the product space $\mathbb{R}_{\geq 0} \times T^S X \times T^S X$, so that it can be measured by $(U[0, c] \times \langle\!\langle \mu \rangle\!\rangle_X^{\mathsf{q}} \times \delta_X)$. In this way, the concentration c acts as a formal parameter for the uniform distribution $U[0, c]$ (note the analogy with rule (6.4.1)). This solves the problem posed at the beginning of the section, when a discrete semantic was shown to be inadequate.

6.4.2 FlatCCS

In this section we introduce a simple yet paradigmatic calculus of agents living in the Euclidean plane $\mathbb{R}^2$, which we call *FlatCCS*[1]. The idea we aim to model is that the rate of communications between two agents depends on their distance (like, e.g., in wireless networks). To this end, FlatCCS extends CCS with a syntactic *frame operator* representing a process' displacement:

$$P, Q ::= \text{nil} \mid \alpha.P \mid P + Q \mid P \parallel Q \mid [P]_z$$
$$\alpha ::= a \mid \bar{a} \mid \tau$$

where $a \in A$ ranges over action labels, and z over the plane $\mathbb{R}^2$. Intuitively, if the P is in position $z' \in \mathbb{R}^2$, the process $[P]_z$ is in $z' + z$. If no frame operator occurs, processes are assumed to be in the origin $(0, 0)$. Thus, in $[P \parallel [Q]_{(0,1)}]_{(1,0)}$, P is (externally) seen to be in $(1, 0)$ and Q in $(1, 1)$.

As for the semantics, we assume that the communication probability decreases exponentially with the distance. Thus, we expect the FlatCCS process $a.\text{nil} \parallel [\bar{a}.\text{nil}]_{(r,0)}$ (with $r \in \mathbb{R}$) to perform an internal communication evolving into $\text{nil} \parallel [\text{nil}]_{(r,0)}$ at rate $|r|$ (hence, with probability $e^{-|r|}$).

We start by defining the measurable signatures for processes and measure terms in order to formally determine which are the σ-algebras associated with each operator symbols and, consequently, to the measurable spaces of terms freely generated over them. We do this directly by giving the syntactic functors $S, M \colon \mathbf{Meas} \to \mathbf{Meas}$ associated with the measurable signatures for processes and measure terms, respectively:

The syntactic functors $S, M \colon \mathbf{Meas} \to \mathbf{Meas}$ for FlatCCS processes and measure terms are given as follows

$$SX = \overbrace{1}^{\text{nil}} + \overbrace{A \times X}^{a.x} + \overbrace{A \times X}^{\bar{a}.x} + \overbrace{X}^{\tau.x} + \overbrace{X \times X}^{x+x} + \overbrace{X \times X}^{x \parallel x} + \overbrace{\mathbb{R}^2 \times X}^{[x]_z} \, ,$$

$$MX = \overbrace{X \times X}^{x \blacktriangleleft x} + \overbrace{X \times X}^{x \blacktriangleright x} + \overbrace{X \times X}^{x \blacktriangledown x} + \overbrace{\mathbb{R}^2 \times X}^{\langle x \rangle_z} \, .$$

where the set A is endowed with the discrete σ-algebra, and $\mathbb{R}^2$ is the real plane endowed with its Borel σ-algebra.

The stochastic semantics is defined by means of the specification system $(\mathcal{R}^{\mathsf{fl}}, \langle\!\langle \cdot \rangle\!\rangle^{\mathsf{fl}})$, where the set $\mathcal{R}^{\mathsf{fl}}$ of MGSOS rules is given in Figure 6.2. According to these rules, the term $\mu \blacktriangleleft \mu'$ indicates that an action has been performed on the left hand side (dually in $\mu \blacktriangleright \mu'$), $\mu \blacktriangledown \mu'$ denotes that the process succeeded in a synchronization, and $\langle \mu \rangle_z$ encodes the absolute position.

---

[1]Of course other variants can be considered, e.g. LineCCS, SpaceCCS, etc. [1].

$$\frac{}{\alpha.x \xrightarrow{\alpha} x} \qquad \frac{x \xrightarrow{\alpha} \mu}{x + x' \xrightarrow{\alpha} \mu} \qquad \frac{x' \xrightarrow{\alpha} \mu}{x + x' \xrightarrow{\alpha} \mu}$$

$$\frac{x \xrightarrow{\alpha} \mu}{[\,x\,]_z \xrightarrow{\alpha} \langle\,\mu\,\rangle_z} \qquad \frac{x \xrightarrow{\alpha} \mu}{x \parallel x' \xrightarrow{\alpha} \mu \blacktriangleleft x'} \qquad \frac{x' \xrightarrow{\alpha} \mu}{x \parallel x' \xrightarrow{\alpha} x \blacktriangleright \mu}$$

$$\frac{x \xrightarrow{a} \mu \quad x' \xrightarrow{\overline{a}} \mu'}{x \parallel x' \xrightarrow{\tau} \mu \blacktriangledown \mu'} \qquad \frac{x \xrightarrow{\overline{a}} \mu \quad x' \xrightarrow{a} \mu'}{x \parallel x' \xrightarrow{\tau} \mu \blacktriangledown \mu'}$$

Figure 6.2: MGSOS specification system for FlatCCS.

The measure term evaluation $\langle\!\langle\,\cdot\,\rangle\!\rangle^{\mathsf{fl}}\colon T^M\Delta_{<\infty} \Rightarrow \Delta_{<\infty}T^S$ completes the description of the semantics we have introduced at the beginning of the section. It is defined, for $\mu,\mu' \in T^M\Delta_{<\infty}X$ and $\beta \in \Delta_{<\infty}X$, as follows

$$\langle\!\langle\beta\rangle\!\rangle^{\mathsf{fl}}_X = \beta$$
$$\langle\!\langle\mu \blacktriangleleft \mu'\rangle\!\rangle^{\mathsf{fl}}_X = (\langle\!\langle\mu\rangle\!\rangle^{\mathsf{fl}}_X \times \langle\!\langle\mu'\rangle\!\rangle^{\mathsf{fl}}_X) \circ (\lambda(x,x').\, x \parallel x')^{-1}$$
$$\langle\!\langle\mu \blacktriangleright \mu'\rangle\!\rangle^{\mathsf{fl}}_X = (\langle\!\langle\mu\rangle\!\rangle^{\mathsf{fl}}_X \times \langle\!\langle\mu'\rangle\!\rangle^{\mathsf{fl}}_X) \circ (\lambda(x,x').\, x \parallel x')^{-1}$$
$$\langle\!\langle\mu \blacktriangledown \mu'\rangle\!\rangle^{\mathsf{fl}}_X = \Big(\|pos(\mu) - pos(\mu')\| \cdot (\langle\!\langle\mu\rangle\!\rangle^{\mathsf{fl}}_X \times \langle\!\langle\mu'\rangle\!\rangle^{\mathsf{fl}}_X)\Big) \circ (\lambda(x,x').\, x \parallel x')^{-1}$$
$$\langle\!\langle\langle\,\mu\,\rangle_z\rangle\!\rangle^{\mathsf{fl}}_X = \langle\!\langle\mu\rangle\!\rangle^{\mathsf{fl}}_X \circ (\lambda x.\, [\,x\,]_z)^{-1}\,,$$

where $pos\colon T^M\Delta_{<\infty}X \to \mathbb{R}^2$ determines the position of an action by inspecting the syntactic structure of μ, inductively defined by

$$pos(\beta) = (0,0) \qquad pos(\mu \blacktriangleleft \mu') = pos(\mu) \qquad pos(\mu \blacktriangleright \mu') = pos(\mu')$$
$$pos(\mu \blacktriangledown \mu') = \tfrac{1}{2}(pos(\mu) + pos(\mu')) \qquad pos(\langle\,\mu\,\rangle_z) = z + pos(\mu)\,.$$

Note that the use of an intermediate syntax for the measure terms occurring in the target of the transitions is essential in order to recover the absolute position of the action that has been performed.

The measure term evaluation $\langle\!\langle\,\cdot\,\rangle\!\rangle^{\mathsf{fl}}\colon T^M\Delta_{<\infty} \Rightarrow \Delta_{<\infty}T^S$, as defined above, arises as an instance of the structural λ-iteration of Corollary 6.3.3, for a suitable (simple) distributive law. Let $\lambda\colon M(\mathbb{R}^2 \times Id) \Rightarrow (\mathbb{R}^2 \times Id)M$ be, for $x,x' \in X$ and $z,z' \in \mathbb{R}^2$,

$$\lambda_X((z,x) \blacktriangleleft (z',x')) = (z, x \blacktriangleleft x')$$
$$\lambda_X((z,x) \blacktriangleright (z',x')) = (z', x \blacktriangleright x')$$
$$\lambda_X((z,x) \blacktriangledown (z',x')) = (\tfrac{1}{2}(z + z'), x \blacktriangledown x')$$
$$\lambda_X(\langle\,(z,x)\,\rangle_{z'}) = (z + z', \langle\,x\,\rangle_{z'})\,.$$

Then, define $k\colon Id \Rightarrow (\mathbb{R}^2 \times Id)$, and a $\varphi\colon M(\mathbb{R}^2 \times \Delta_{<\infty}T^S) \Rightarrow \Delta_{<\infty}T^S$ as follows, for $x \in X$, $z,z' \in \mathbb{R}^2$, and $\beta,\beta' \in \Delta_{<\infty}T^SX$,

$$k_X(x) = ((0,0),x)\,,$$

$$\varphi_X((z,\beta) \blacktriangleleft (z',\beta')) = (\beta \times \beta') \circ (\lambda(x,x').\, x \parallel x')^{-1}$$
$$\varphi_X((z,\beta) \blacktriangleright (z',\beta')) = (\beta \times \beta') \circ (\lambda(x,x').\, x \parallel x')^{-1}$$
$$\varphi_X((z,\beta) \blacktriangledown (z',\beta')) = \big(e^{-\|z-z'\|} \cdot (\beta \times \beta')\big) \circ (\lambda(x,x').\, x \parallel x')^{-1}$$
$$\varphi_X(\langle\,(z,\beta)\,\rangle_{z'}) = \beta \circ (\lambda x.\, [\,x\,]_{z'})^{-1}\,.$$

These are easily seen to be natural in X and measurable. Now, applying the structural λ-iteration proof principle of Proposition 6.3.2 we obtain

$$
\begin{array}{ccccc}
M(\mathbb{R}^2 \times \Delta_{<\infty} T^S X) & \xleftarrow{\;\;M(\mathbb{R}^2 \times \langle\!\!\!\mid \cdot \mid\!\!\!\rangle_X)\;\;} & M(\mathbb{R}^2 \times T^M \Delta_{<\infty} X) & \xleftarrow{\;\;M\langle pos, id\rangle_{\Delta_{<\infty}X}\;\;} & MT^M \Delta_{<\infty} X \\
\downarrow{\varphi_X} & & & & \downarrow{\psi_X} \\
\Delta_{<\infty} T^S X & \xleftarrow{\qquad\qquad\langle\!\!\!\mid \cdot \mid\!\!\!\rangle_X\qquad\qquad} & & & T^M \Delta_{<\infty} X \\
& & & & \uparrow{\eta^m_X} \\
& & \xleftarrow{\;\;\Delta_{<\infty}\eta^S_X\;\;} & & \Delta_{<\infty} X
\end{array}
$$

$$
\begin{array}{ccccc}
X & \xrightarrow{\;\;\eta^m_X\;\;} & T^M X & \xleftarrow{\qquad\psi_X\qquad} & MT^M X \\
\downarrow{k_X} & & \downarrow{\langle pos, id\rangle_X} & & \downarrow{M\langle pos, id\rangle_X} \\
\mathbb{R}^2 \times X & \xrightarrow{(\mathbb{R}^2 \times \eta^m_X)} & \mathbb{R}^2 \times T^M X & \xleftarrow{(\mathbb{R}^2 \times \psi_X) \circ \lambda_{(T^M X)}} & M(\mathbb{R}^2 \times T^M X)
\end{array}
$$

It is easy to check that the diagrams above commute also when $\langle\!\!\!\mid \cdot \mid\!\!\!\rangle^{\mathsf{fl}}_X$ is used in place of $\langle\!\!\!\mid \cdot \mid\!\!\!\rangle_X$, hence $\langle\!\!\!\mid \cdot \mid\!\!\!\rangle^{\mathsf{fl}}_X = \langle\!\!\!\mid \cdot \mid\!\!\!\rangle_X$. Naturality follows by Corollary 6.3.3.

We conclude the description of the MGSOS rule format by showing a practical example of how the construction in Theorem 6.2.4 applies to the specification system $(\mathcal{R}^{\mathsf{fl}}, \langle\!\!\!\mid \cdot \mid\!\!\!\rangle^{\mathsf{fl}})$ for FlatCCS. The abstract GSOS distributive law $\llbracket\cdot\rrbracket \colon S(Id \times \Delta^L_{<\infty}) \Rightarrow (\Delta_{<\infty}T^S)^L$ for $x, x' \in X$, $L = A \cup \overline{A} \cup \{\tau\}$, and $\beta, \beta' \in (\Delta_{<\infty}X)^L$, is given by

$$\llbracket \mathsf{nil} \rrbracket_X = \lambda\alpha.\, \mathbf{0}$$

$$\llbracket \alpha.\, \rrbracket_X(x, \beta) = \lambda\alpha'.\begin{cases} \langle\!\!\!\mid \delta_x \mid\!\!\!\rangle^{\mathsf{fl}}_X & \text{if } \alpha' = \alpha \\ \mathbf{0} & \text{otherwise} \end{cases}$$

$$(x, \beta)\llbracket + \rrbracket_X(x', \beta') = \lambda\alpha.\, \langle\!\!\!\mid \beta(\alpha) \mid\!\!\!\rangle^{\mathsf{fl}}_X \oplus \langle\!\!\!\mid \beta'(\alpha) \mid\!\!\!\rangle^{\mathsf{fl}}_X$$

$$(x, \beta)\llbracket \parallel \rrbracket_X(x', \beta') = \lambda\alpha.\begin{cases} \langle\!\!\!\mid \beta(\alpha) \blacktriangleleft \delta_{x'} \mid\!\!\!\rangle^{\mathsf{fl}}_X \oplus \langle\!\!\!\mid \delta_x \blacktriangleright \beta'(\alpha) \mid\!\!\!\rangle^{\mathsf{fl}}_X & \text{if } \alpha \neq \tau \\[2mm] \displaystyle\bigoplus_{a \in A} \left\{ \begin{array}{l} \langle\!\!\!\mid \beta(\alpha) \blacktriangleleft \delta_{x'} \mid\!\!\!\rangle^{\mathsf{fl}}_X, \langle\!\!\!\mid \delta_x \blacktriangleright \beta'(\alpha) \mid\!\!\!\rangle^{\mathsf{fl}}_X, \\ \langle\!\!\!\mid \beta(a) \blacktriangledown \beta'(\overline{a}) \mid\!\!\!\rangle^{\mathsf{fl}}_X, \langle\!\!\!\mid \beta(\overline{a}) \blacktriangledown \beta'(a) \mid\!\!\!\rangle^{\mathsf{fl}}_X \end{array} \right\} & \text{if } \alpha = \tau \end{cases}$$

$$\llbracket [\;]_z \rrbracket_X(x, \beta) = \lambda\alpha'.\langle\!\!\!\mid \beta(\alpha) \mid\!\!\!\rangle^{\mathsf{fl}}_X$$

where $\mathbf{0}$ denotes the null sub-probability measure assigning to each measurable set probability zero, and δ_x is the Dirac measure at $x \in X$.

Remark 6.4.1 When the measurable space of variables X is taken to be the space $T^S\mathbf{0}$ of ground measurable terms, the above definition gives rise to the "canonical" universal (initial and final) semantics for FlatCCS. $\blacksquare$

7

On-the-Fly Exact Computation of Bisimilarity Distances

Probabilistic bisimulation for Markov chains (MCs), introduced by Larsen and Skou [63], is the key concept for reasoning about the equivalence of probabilistic systems. However, when one focuses on quantitative behaviours it becames obvious that such an equivalence is too "exact" for many purposes as it only relates processes with identical behaviours. In various applications, such as systems biology [81], games [27], planning [30] or security [23], we are interested in knowing whether two processes that may differ by a small amount in the real-valued parameters (probabilities) have "sufficiently" similar behaviours. This motivated the development of the metric theory for MCs, initiated by Desharnais et al. [39] and greatly developed and explored by van Breugel, Worrell and others [88, 87]. It consists in proposing a *bisimilarity distance (pseudometric)*, which measures the behavioural similarity of two MCs. The pseudometric proposed by Desharnais et al. is parametric in a *discount factor* $\lambda \in (0, 1]$ that controls the significance of the future in the measurement.

Since van Breugel et al. have presented a fixed point characterization of the bisimilarity pseudometric, several iterative algorithms have been developed in order to compute approximations of δ_λ up to any degree of accuracy [46, 88, 87]. Recently, Chen et al. [28] proved that, for finite MCs with rational transition function, the bisimilarity pseudometrics can be computed exactly in polynomial time. The proof consists in describing the pseudometric as the solution of a linear program that can be solved using the *ellipsoid method*. Although the ellipsoid method is theoretically efficient, *"computational experiments with the method are very discouraging and it is in practice by no means a competitor of the, theoretically inefficient, simplex method"*, as stated in [75]. Unfortunately, in this case the simplex method cannot be used to speed up performances in practice, since the linear program to be solved may have an exponential number of constraints in the number of states of the MC.

In this chaper, we propose an alternative approach to this problem, which allows us to compute the pseudometric exactly and efficiently in practice. This is inspired by the characterization of the undiscounted pseudometric given in [28] based on the notion of *coupling*, which we extend to generic discount factors. A coupling defines a possible redistribution of the transition probabilities of each pair of states of a given MC; it is evaluated by the *discrepancy function* that measures the behavioural dissimilarities between the states revealed by the redistributions. In [28] it is shown that the bisimilarity pseudometric for a given MC is the minimum among the discrepancy functions corresponding to all possible couplings for that MC; moreover, the bisimilarity pseudometric is itself a discrepancy function corresponding to an *optimal coupling*. This suggests that the problem of computing the pseudometric can be reduced to the problem of finding a coupling with the least discrepancy function.

Our approach aims at finding an optimal coupling using a greedy strategy that starts from an arbitrary coupling and repeatedly looks for new couplings that improve the discrepancy function. This strategy will eventually find an optimal coupling. We use it to support the design of an on-the-fly algorithm for computing the exact behavioural pseudometric that can be either applied to compute all the distances in the model or to compute only some designated distances. The advantage of using an on-the-fly approach consists in the fact that we do not need to exhaustively

explore the state space nor to construct entire couplings but only those fragments that are needed in the local computation.

The efficiency of our algorithm has been evaluated on a significant set of randomly generated MCs. The results show that our algorithm performs orders of magnitude better than the corresponding iterative algorithms proposed, for instance in [46, 28]. Moreover, we provide empirical evidence for the fact that our algorithm enjoys good execution running times.

One of the main practical advantages of our approach consists in the fact that it can focus on computing only the distances between states that are of particular interest. This is useful in practice, for instance when large systems are considered and visiting the entire state space is expensive. A similar issue has been considered by Comanici et al., in [29], who have noticed that for computing the approximated pseudometric one does not need to update the current value for all the pairs at each iteration, but it is sufficient to only focus on the pairs where changes are happening rapidly. Our approach goes much beyond this idea. Firstly, we are not only looking to approximations of the bisimilarity distance, but we develop an exact algorithm; secondly, we provide a termination condition that can be checked locally, still ensuring that the local optimum corresponds to the global one. In addition, our method can be applied to decide whether two states of an MC are probabilistic bisimilar, to identify the bisimilarity classes for a given MC or to solve lumpability problems. Our approach can also be used with approximation techniques as, for instance, to provide a least over-approximation of the behavioural distance given over-estimates of some particular distances. This can be further integrated with other approximate algorithms having the advantage of the on-the-fly state space exploration.

7.1   Markov Chains and Bisimilarity Pseudometrics

In this section we give the definitions of *(discrete-time) Markov chains* (MCs) and *probabilistic bisimilarity* for MCs [63]. Then we recall the *bisimilarity pseudometric* of Desharnais et al. [39] and its fixed point characterization given by van Breugel et al. [87].

Definition 7.1.1 (Markov chain) *A* (discrete-time) Markov chain *is a tuple* $\mathcal{M} = (S, A, \pi, \ell)$ *consisting of a countable nonempty set* S *of* states, *a countable nonempty set* A *of* labels, *a* transition probability function $\pi \colon S \times S \to [0, 1]$ *such that, for arbitrary* $s \in S$, $\sum_{t \in S} \pi(s, t) = 1$, *and a* labelling function $\ell \colon S \to A$. $\mathcal{M}$ *is* finite *if its support set* S *is finite.*

Remark 7.1.2 The above definition differs form that of Markov kernel (see Definition 5.1.1) in that (i) the state space is discrete and (ii) labels are embedded into states rather than on the transitions. We chose to deal with discrete-time MCs for algorithmic reasons and also because we want to compare our algorithm with that in [28] where this kind models are used.   ■

Given an MC $\mathcal{M} = (S, A, \pi, \ell)$, we identify the transition probability function π with its *transition matrix* $(\pi(s, t))_{s, t \in S}$. For $s, t \in S$, we denote by $\pi(s, \cdot)$ and $\pi(\cdot, t)$, respectively, the probability distributions of exiting from s to any state and entering to t from any state.

The MC $\mathcal{M}$ induces an *underlying (directed) graph*, denoted by $\mathcal{G}(\mathcal{M})$, where the states act as vertices and (s, t) is an edge in $\mathcal{G}(\mathcal{M})$, if and only if, $\pi(s, t) > 0$. For a subset $Q \subseteq S$, we denote by $R_{\mathcal{M}}(Q)$ the set of states reachable from some $s \in Q$, and by $R_{\mathcal{M}}(s)$ we denote $R_{\mathcal{M}}(\{s\})$. The size of $\mathcal{M}$, denoted by $size(\mathcal{M})$, is the number of vertices plus the number of edges of $\mathcal{G}(\mathcal{M})$.

From a theoretical point of view, it is irrelevant whether the transition probability function of a given MC has rational values or not. However, for algorithmic purposes, in this chapter we assume that for arbitrary $s, t \in S$, $\pi(s, t) \in \mathbb{Q} \cap [0, 1]$. For similar reasons, we also restrict our investigation to finite MCs.

Definition 7.1.3 (Probabilistic Bisimulation) *Let* $\mathcal{M} = (S, A, \pi, \ell)$ *be a MC. An equivalence relation* $R \subseteq S \times S$ *is a* probabilistic bisimulation *if whenever* $s \, R \, t$, *then*
   (i) $\ell(s) = \ell(t)$ *and,*
   (ii) for each R-*equivalence class* E, $\sum_{u \in E} \pi(s, u) = \sum_{u \in E} \pi(t, u)$.

Two states $s, t \in S$ are bisimilar, *written $s \sim t$, if they are related by a probabilistic bisimulation.*

This definition is due to Larsen and Skou [63]. The intuition behind this definition is that, two states are bisimilar if they have the same label and their probability of moving by a single transition to any given equivalence class is always the same.

The notion of equivalence can be relaxed by means of a pseudometric, which tells how two elements are far apart from each other and whenever they are at zero distance they are equivalent. The bisimilarity pseudometric of Desharnais et al. [39] on MCs enjoys the property that two states are at zero distance if and only if they are bisimilar. This pseudometric was first defined using a real-valued semantics for a logic [39], then it has been characterized as the least fixed point of an operator based on the Kantorovich metric for comparing probability distributions [87]. Actually, for the purpose of this chapter we only require the fixed point characterization, but for the sake of completeness (and also clarity), in the rest of this section we recall both of them.

Logical Characterization. The bisimilarity pseudometric of [39] is given by means of a real-valued semantics. The logic $\mathcal{L}$ is defined by the following grammar, for $\sigma \in A$ and $q \in \mathbb{Q} \cap [0, 1]$

$$\phi ::= \sigma \mid \neg\phi \mid \phi \vee \phi \mid \phi \ominus q \mid \diamond\phi \,.$$

Given a labelled Markov process $\mathcal{M} = (S, A, \pi, \ell)$ and a parameter $\lambda \in (0, 1]$, the semantics of a formula ϕ is given by the function $[\![\phi]\!]_\lambda \colon S \to [0, 1]$ defined by

$$[\![\sigma]\!]_\lambda(s) = \begin{cases} 1 & \text{if } \ell(s) = \sigma \\ 0 & \text{otherwise} \end{cases} \qquad\qquad [\![\phi \vee \psi]\!]_\lambda(s) = \max\{[\![\phi]\!]_\lambda(s), [\![\psi]\!]_\lambda(s)\}$$

$$[\![\neg\phi]\!]_\lambda(s) = 1 - [\![\phi]\!]_\lambda(s) \qquad\qquad\qquad [\![\phi \ominus q]\!]_\lambda(s) = \max\{[\![\phi]\!]_\lambda(s) - q, 0\}$$

$$[\![\diamond\phi]\!]_\lambda(s) = \lambda \cdot \sum_{t \in S} \pi(s, t) \cdot [\![\phi]\!]_\lambda(t) \,.$$

The parameter λ in the real-valued semantics of $\mathcal{L}$ plays the role of a *discount factor*. Indeed, the smaller the value of λ, the more the future is discounted. If $\lambda = 1$ the future is not discounted and semantics is said *undiscounted*.

Definition 7.1.4 (Bisimilarity Pseudometric) *Let $\lambda \in (0, 1]$ be the discount factor, the (1-bounded) pseudometric $\delta_\lambda \colon S \times S \to [0, 1]$ assigns a distance to any given pair of states of a MC according to the following definition:*

$$\delta_\lambda(s, t) = \max_{\phi \in \mathcal{L}} |[\![\phi]\!]_\lambda(s) - [\![\phi]\!]_\lambda(t)| \,.$$

This characterisation illustrates the sense in which states that are close in the pseudometric satisfy similar behavioural properties. The following theorem justifies this intuition.

Theorem 7.1.5 (Soundness [39]) *For all $\lambda \in (0, 1]$, $\delta_\lambda(s, t) = 0$ iff $s \sim t$.*

Fixed-point characterization. In [87], van Breugel et al. characterized the bisimilarity pseudometric δ_λ as the least fixed point of an operator based on the Kantorovich metric for comparing probability distributions which makes use of the notion of *matching*.

Definition 7.1.6 (Matching) *Let $\mu, \nu \colon S \to [0, 1]$ be probability distributions on S. A* matching *for the pair (μ, ν) is a probability distribution $\omega \colon S \times S \to [0, 1]$ on $S \times S$ satisfying*

$$\forall\, u \in S. \ \textstyle\sum_{s \in S} \omega(u, s) = \mu(u)\,, \qquad\qquad \forall\, v \in S. \ \textstyle\sum_{s \in S} \omega(s, v) = \nu(v)\,. \qquad (7.1.1)$$

We call μ and ν, respectively, the left and the right mariginals *of ω.*

In the following, we denote by $\mu \otimes \nu$ the set of all matchings for (μ, ν).

Remark 7.1.7 Note that, for S finite, (7.1.1) describes the constraints of a homogeneous transportation problem (TP) [32, 47], where the vector $(\mu(u))_{u \in S}$ specifies the amounts to be shipped and $(\nu(v))_{v \in S}$ the amounts to be received. Thus, a matching ω for (μ, ν) induces a matrix $(\omega(u, v))_{u,v \in S}$ to be thought as a shipping schedule belonging to the polytope $\mu \otimes \nu$. Hereafter, we denote by $TP(c, \nu, \mu)$ the TP with cost matrix $(c(u, v))_{u,v \in S}$ and marginals ν and μ. ∎

Transportation Problem. In 1941 Hitchcock and, independently, in 1947 Koopmans considered the problem which is usually referred to as the (homogeneous) *transportation problem*. This problem can be intuitively described as: a homogeneous product is to be shipped in the amounts $a_1, \ldots, a_m$ respectively, from each of m shipping *origins* and received in amounts $b_1, \ldots, b_n$ respectively, by each of n shipping *destinations*. The cost of shipping a unit amount from the i-th origin to the j-th destination is $c_{i,j}$ and is known for all combinations (i, j). The problem is to determine an optimal *shipping schedule*, i.e. the amount $x_{i,j}$ to be shipped over all routes (i, j), which minimizes the total cost of transportation. This problem is easily formalized as a linear programming problem

$$
\begin{aligned}
\text{minimize } & \sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j} \cdot x_{i,j} \\
\text{such that } & \sum_{j=1}^{n} x_{i,j} = a_i & (i = 1, \ldots, m) \\
& \sum_{i=1}^{m} x_{i,j} = b_j & (j = 1, \ldots, n) \\
& x_{i,j} \geq 0 & (i = 1, \ldots, m \text{ and } j = 1, \ldots, n)
\end{aligned}
$$

The set of schedules feasible for a transportation problem, which is formalized as a conjunction of linear constraints, describes a (bounded) convex polytope in $\mathbb{R}^2$, often called transportation polytope. These polytopes have good geometrical properties, which makes the transportation problem one of the most studied optimization problems in literature. As reported in [59], for transportation problems of size $m \times n$ the number of vertices can be exponential in n and m, namely $\max\{m, n\}^{\min\{m,n\}-1}$. This result is due to Demuth [35]. Anyway the diagonal of the transportation polytope is linearly bounded by $8(m + n - 2)$ (see [20]).

There are several algorithms in literature which efficiently solve (not necessarily homogeneous) transportation problems. Among these we recall [32, 47].

For a Markov chain $\mathcal{M} = (S, A, \pi, \ell)$ and a discount factor $\lambda \in (0, 1]$, wed define the operator $\Delta_\lambda^{\mathcal{M}} \colon [0, 1]^{S \times S} \to [0, 1]^{S \times S}$ as follows, for $d \colon S \times S \to [0, 1]$ and $s, t \in S$:

$$
\Delta_\lambda^{\mathcal{M}}(d)(s, t) = \begin{cases} 1 & \text{if } \ell(s) \neq \ell(t) \\ \lambda \cdot \min_{\omega \in \pi(s, \cdot) \otimes \pi(t, \cdot)} \sum_{u,v \in S} d(u, v) \cdot \omega(u, v) & \text{if } \ell(s) = \ell(t) \end{cases}
$$

The set $[0, 1]^{S \times S}$ is endowed with the partial order $\sqsubseteq$ defined as $d \sqsubseteq d'$ iff $d(s, t) \leq d'(s, t)$ for all $s, t \in S$. This forms a complete lattice with bottom element $\mathbf{0}$ and top element $\mathbf{1}$, defined as $\mathbf{0}(s, t) = 0$ and $\mathbf{1}(s, t) = 1$, for all $s, t \in S$. For $D \subseteq [0, 1]^{S \times S}$, the least upper bound $\bigsqcup D$, and greatest lower bound $\bigsqcap D$ are given by $(\bigsqcup D)(s, t) = \sup_{d \in D} d(s, t)$ and $(\bigsqcap D)(s, t) = \inf_{d \in D} d(s, t)$ for all $s, t \in S$.

In [87], for any $\mathcal{M}$ and $\lambda \in [0, 1]$, $\Delta_\lambda^{\mathcal{M}}$ is proved to be monotonic, thus, by Tarski's fixed point theorem, it admits least and greatest fixed points. In particular the former characterizes the bisimilarity pseudometric.

Theorem 7.1.8 ([87]) *Let $\mathcal{M}$ be a MC and $\lambda \in (0, 1]$ be a discount factor. Then δ_λ corresponds to the least fixed point of $\Delta_\lambda^{\mathcal{M}}$.*

Hereafter, $\Delta_\lambda^{\mathcal{M}}$ and $\delta_\lambda^{\mathcal{M}}$ will be denoted simply by Δ_λ and δ_λ, respectively, when the Markov chain $\mathcal{M}$ is clear from the context.

7.2 Alternative Characterization of the Pseudometric

In [28], Chen et al. proposed an alternative characterization of δ_1, relating the pseudometric to the notion of *coupling*. In this section, we recall the definition of coupling, and generalize the characterization for generic discount factors.

Definition 7.2.1 (Coupling) *Let $\mathcal{M} = (S, A, \pi, \ell)$ be a finite Markov chain. A Markov chain of the form $\mathcal{C} = (S \times S, A \times A, \omega, l)$ is a coupling for $\mathcal{M}$ if, for all $s, t \in S$*

  (i) $\omega((s, t), \cdot) \in \pi(s, \cdot) \otimes \pi(t, \cdot)$;
  (ii) $l(s, t) = (\ell(s), \ell(t))$.

Intuitively, a coupling for $\mathcal{M}$ can be seen a probabilistic pairing of two copies of $\mathcal{M}$ running synchronously, although not necessarily independently. Couplings have been used to characterize weak ergodicity of arbitrary Markov chains [50], or to give upper bounds on convergence to stationary distributions [11, 67].

Given a coupling $\mathcal{C} = (S \times S, A \times A, \omega, l)$ for $\mathcal{M} = (S, A, \pi, \ell)$ we define $\Gamma_\lambda^{\mathcal{C}} \colon [0,1]^{S \times S} \to [0,1]^{S \times S}$ for $d \colon S \times S \to [0,1]$ and $s, t \in S$, as follows:

$$
\Gamma_\lambda^{\mathcal{C}}(d)(s,t) = \begin{cases} 1 & \text{if } \ell(s) \neq \ell(t) \\ \lambda \cdot \sum_{u,v \in S} d(u,v) \cdot \omega((s,t),(u,v)) & \text{if } \ell(s) = \ell(t) \end{cases}
$$

One should easily convice himself that, for any $\lambda \in (0,1]$, $\Gamma_\lambda^{\mathcal{C}}$ is well-defined and monotonic:

Lemma 7.2.2 *Let $\mathcal{C}$ be a coupling for $\mathcal{M} = (S, A, \pi, \ell)$ and $\lambda \in (0,1]$. Then, $\Gamma_\lambda^{\mathcal{C}}$ is well-defined and, whenever $d \sqsubseteq d'$, $\Gamma_\lambda^{\mathcal{C}}(d) \sqsubseteq \Gamma_\lambda^{\mathcal{C}}(d)$.*

Proof. Assume $\mathcal{C} = (S \times S, A \times A, \omega, l)$. We first prove that, given $d \in [0,1]^{S \times S}$ then $\Gamma_\lambda^{\mathcal{C}}(d) \in [0,1]^{S \times S}$, that is, for all $s, t \in S$, $0 \leq \Gamma_\lambda^{\mathcal{C}}(d)(s,t) \leq 1$.

If $\ell(s) \neq \ell(t)$, by definition, $\Gamma_\lambda^{\mathcal{C}}(d)(s,t) = 1$. If $\ell(s) = \ell(t)$, we have that $\Gamma_\lambda^{\mathcal{C}}(d)(s,t) = \lambda \cdot \sum_{u,v \in S} d(u,v) \cdot \omega((s,t),(u,v))$. By Definition 7.2.1, $\omega((s,t), \cdot)$ is a probability distribution, thus, for all $u, v \in S$, $\omega((s,t),(u,v)) \geq 0$, and $\sum_{u,v \in S} \omega((s,t),(u,v)) = 1$. By hypothesis, $\lambda \in (0,1]$ and, for all $u, v \in S$, $0 \leq d(u,v) \leq 1$, therefore

$$
0 \leq \lambda \cdot \sum_{u,v \in S} d(u,v) \cdot \omega((s,t),(u,v)) \leq \sum_{u,v \in S} \omega((s,t),(u,v)) = 1 \,.
$$

Therefore $0 \leq \Gamma_\lambda^{\mathcal{C}}(d)(s,t) \leq 1$.

Let $d, d' \colon S \times S \to [0,1]$, such that $d \sqsubseteq d'$, and $s, t \in S$. If $\ell(s) \neq \ell(t)$ then $\Gamma_\lambda^{\mathcal{C}}(d)(s,t) = 1 = \Gamma_\lambda^{\mathcal{C}}(d')(s,t)$. If $\ell(s) = \ell(t)$ we have that

$$
\Gamma_\lambda^{\mathcal{C}}(d)(s,t) = \lambda \sum_{u,v \in S} d(u,v) \cdot \omega(s,t)(u,v)
$$
$$
\leq \lambda \sum_{u,v \in S} d'(u,v) \cdot \omega(s,t)(u,v) = \Gamma_\lambda^{\mathcal{C}}(d)(s,t) \,,
$$

since, for all $u, v \in S$, $d(u,v) \leq d'(u,v)$ and $\omega(s,t)(u,v) \geq 0$. Since $\Gamma_\lambda^{\mathcal{C}}$ is monotonic and $\sqsubseteq$ is antisymmetric, it follows that $\Gamma_\lambda^{\mathcal{C}}$ is well-defined. ∎

By Tarski's fixed point theorem, $\Gamma_\lambda^{\mathcal{C}}$ admits least fixed point, which we denote by $\gamma_\lambda^{\mathcal{C}}$. We will see that, for any $s, t \in S$, $\gamma_1^{\mathcal{C}}(s,t)$ corresponds to the probability of reaching a state (u,v) with $\ell(u) \neq \ell(v)$ starting from the state (s,t) in the underling graph of $\mathcal{C}$. For this reason we will call $\gamma_\lambda^{\mathcal{C}}$ the λ-*discounted discrepancy* of $\mathcal{C}$ or simply the λ-discrepancy of $\mathcal{C}$.

Lemma 7.2.3 *Let $\mathcal{M}$ be a MC, $\mathcal{C}$ be a coupling for $\mathcal{M}$, and $\lambda \in (0,1]$ be a discount factor. If $d = \Gamma_\lambda^{\mathcal{C}}(d)$ then $\delta_\lambda \sqsubseteq d$.*

Proof. Assume $\mathcal{M} = (S, A, \pi, \ell)$ and $\mathcal{C} = (S \times S, A \times A, \omega, l)$. In order to prove $\delta_\lambda \sqsubseteq d$, it suffices to show that $\Delta_\lambda(d) \sqsubseteq d$. Indeed, by Tarski's fixed point theorem, δ_λ is a lower bound of $\{d \mid \Delta_\lambda(d) \sqsubseteq d\}$.

Let $s, t \in S$. If $\ell(s) \neq \ell(t)$, then $\Delta_\lambda(d)(s,t) = 1 = \Gamma_\lambda^{\mathcal{C}}(d)(s,t) = d(s,t)$. If $\ell(s) = \ell(t)$, $\Delta_\lambda(d)(s,t) = \lambda \cdot \min_{\omega' \in \pi(s,\cdot) \otimes \pi(t,\cdot)} \sum_{u,v \in S} d(u,v) \cdot \omega'(u,v)$, and $\Gamma_\lambda^{\mathcal{C}}(d)(s,t) = \lambda \cdot \sum_{u,v \in S} d(u,v) \cdot \omega((s,t),(u,v))$. Since $\omega((s,t),\cdot) \in \pi(s,\cdot) \otimes \pi(t,\cdot)$ (Definition 7.2.1), we have that $\Delta_\lambda(d)(s,t) \leq \Gamma_\lambda^{\mathcal{C}}(d)(s,t) = d(s,t)$. ∎

As a consequence of Lemma 7.2.3 we obtain the following characterization for δ_λ, which generalizes [28, Theorem 8] for generic discount factors.

Theorem 7.2.4 (Minimum coupling criterion) *Let $\mathcal{M}$ be a MC and $\lambda \in (0,1]$ be a discount factor. Then, $\delta_\lambda = \min \left\{ \gamma_\lambda^{\mathcal{C}} \mid \mathcal{C} \text{ coupling for } \mathcal{M} \right\}$.*

Proof. For any fixed $d \in [0,1]^{S \times S}$ there exists a coupling $\mathcal{C}$ for $\mathcal{M}$ such that $\Gamma_\lambda^{\mathcal{C}}(d) = \Delta_\lambda(d)$. Indeed we can take as transition function for $\mathcal{C}$, the joint probability distribution ω such that, for all $s, t \in S$, $\sum_{u,v \in S} d(u,v) \cdot \omega((s,t),(u,v))$ achieves the minimim value.

Let $\mathcal{D}$ be a coupling for $\mathcal{M}$ such that $\Gamma_\lambda^{\mathcal{D}}(\delta_\lambda) = \Delta_\lambda(\delta_\lambda)$. By Theorem 7.1.8, $\Delta_\lambda(\delta_\lambda) = \delta_\lambda$, therefore δ_λ is a fixed point for $\Gamma_\lambda^{\mathcal{D}}$. By Lemma 7.2.3, δ_λ is a lower bound of the set of fixed points of $\Gamma_\lambda^{\mathcal{D}}$, therefore $\delta_\lambda = \gamma_\lambda^{\mathcal{D}}$. By Lemma 7.2.3, we have also that, for any coupling $\mathcal{C}$ of $\mathcal{M}$, $\delta_\lambda \sqsubseteq \gamma_\lambda^{\mathcal{C}}$. Therefore, given the set $D = \left\{ \gamma_\lambda^{\mathcal{C}} \mid \mathcal{C} \text{ coupling for } \mathcal{M} \right\}$, it follows that $\delta_\lambda \in D$ and δ_λ is a lower bound for D. Hence, by antisymmetry of $\sqsubseteq$, $\delta_\lambda = \min D$. ∎

7.3   Exact Computation of Bisimilarity Distance

Inspired by the characterization given in Theorem 7.2.4, in this section we propose a procedure to exactly compute the bisimilarity pseudometric.

For $\lambda \in (0,1]$, the set of couplings for $\mathcal{M}$ can be endowed with the preorder $\trianglelefteq_\lambda$ defined as $\mathcal{C} \trianglelefteq_\lambda \mathcal{D}$, if and only if, $\gamma_\lambda^{\mathcal{C}} \sqsubseteq \gamma_\lambda^{\mathcal{D}}$. Theorem 7.2.4 suggests to look at all the couplings $\mathcal{C}$ for $\mathcal{M}$ in order to find an optimal one, that is, minimal with respect to $\trianglelefteq_\lambda$. Needless to say, an enumeration of all the couplings is unfeasible, therefore it is crucial to provide an efficient search strategy which prevents us to do that. Moreover we also need an efficient method for computing the λ-discrepancy.

In Subsection 7.3.1 the problem of computing the λ-discrepancy of a coupling $\mathcal{C}$ is reduced to the problem of computing reachability probabilities in $\mathcal{C}$. Then, Subsection 7.3.2 illustrates a greedy strategy that explores the set of couplings until and an optimal one eventually reached.

7.3.1   Computing the λ-discrepancy

In this section, we first recall the problem of computing the reachability probability for general MCs [11], then we instantiate it to compute the λ-discrepancy.

Let $\mathcal{M} = (S, A, \pi, \ell)$ be an MC, and x_s denote the probability of reaching $G \subseteq S$ from $s \in S$. The goal is to compute x_s for all $s \in S$. The following holds

$$x_s = 1 \quad \text{if } s \in G\,, \qquad\qquad x_s = \sum_{t \in S} x_t \cdot \pi(s,t) \quad \text{if } s \in S \setminus G\,, \qquad\qquad (7.3.1)$$

that is, either G is already reached, or by way of an other state. Equation (7.3.1) defines a linear equation system of the form $\vec{x} = \mathbf{A}\vec{x} + \vec{b}$, where $S_? = S \setminus G$, $\vec{x} = (x_s)_{s \in S_?}$, $\mathbf{A} = (\pi(s,t))_{s,t \in S_?}$, and $\vec{b} = (\sum_{t \in G} \pi(s,t))_{s \in S_?}$.

This linear equation system always admits a solution in $[0,1]^S$, howerver, it may not be unique. Since we are interested in the least solution, we address to this problem fixing each free variable to zero, so that we obtain a reduced system with a unique solution. This can be easily done inspecting the graph $\mathcal{G}(\mathcal{M})$: all variables with zero probability of reaching G are detected by checking that they cannot be reached from any state in G in the reverse graph of $\mathcal{G}(\mathcal{M})$.

Regarding the λ-discrepancy for a coupling $\mathcal{C}$, if $\lambda = 1$, one can directly instantiate the afore-mentioned method with $G = \{(s,t) \in S \times S \mid \ell(s) \neq \ell(t)\}$ and $S_? = (S \times S) \setminus G$. As for generic

$\lambda \in (0, 1]$, the discrepancy $\gamma_\lambda^\mathcal{C}$ can be formulated as the least solution in $[0, 1]^{S \times S}$ of the linear equation system

$$\vec{x} = \lambda \mathbf{A} \vec{x} + \lambda \vec{b}. \tag{7.3.2}$$

Remark 7.3.1 If one is interested in computing the λ-discrepancy for a particular pair of states (s, t), the method above can be applied on the least independent subsystem of Equation (7.3.2) containing the variable $x_{(s,t)}$. Moreover, assuming that for some pairs the λ-discrepancy is already known, the goal set can be extended with all those pairs with λ-discrepancy greater than zero. ∎

7.3.2 Greedy search strategy for computing an optimal coupling

In this section, we give a greedy strategy for moving toward an optimal coupling starting from a given one, Then we provide sufficient and necessary conditions for a coupling, ensuring that its associated λ-discrepancy coincides with δ_λ.

Hereafter, we fix a coupling $\mathcal{C} = (S \times S, A \times A, \omega, l)$ for $\mathcal{M} = (S, A, \pi, \ell)$. Let $s, t \in S$ and μ be a matching for $(\pi(s, \cdot), \pi(t, \cdot))$. We denote by $\mathcal{C}[(s, t)/\mu]$ the coupling for $\mathcal{M}$ with the same labeling function of $\mathcal{C}$ and transition function ω' defined by $\omega'((u, v), \cdot) = \omega((u, v), \cdot)$, for all $(u, v) \neq (s, t)$, and $\omega'((s, t), \cdot) = \mu$.

Lemma 7.3.2 *Let $\mathcal{C}$ be a coupling for $\mathcal{M}$, $s, t \in S$, $\omega' \in \pi(s, \cdot) \otimes \pi(t, \cdot)$, and $\mathcal{D} = \mathcal{C}[(s, t)/\omega']$. If $\Gamma_\lambda^\mathcal{D}(\gamma_\lambda^\mathcal{C})(s, t) < \gamma_\lambda^\mathcal{C}(s, t)$, then $\gamma_\lambda^\mathcal{D} \sqsubset \gamma_\lambda^\mathcal{C}$.*

Proof. It suffices to show that $\Gamma_\lambda^\mathcal{D}(\gamma_\lambda^\mathcal{C}) \sqsubset \gamma_\lambda^\mathcal{C}$, i.e., $\gamma_\lambda^\mathcal{C}$ is a strict post-fixed point of $\Gamma_\lambda^\mathcal{D}$. Then, the thesis follows by Tarski's fixed point theorem.

Assume $\bar{\omega}$ be the transition function of $\mathcal{D}$ and let $u, v \in S$. If $\ell(u) \neq \ell(v)$, then, by definition, $\Gamma_\lambda^\mathcal{D}(\gamma_\lambda^\mathcal{C})(u, v) = 1 = \Gamma_\lambda^\mathcal{C}(\gamma_\lambda^\mathcal{C})(u, v) = \gamma_\lambda^\mathcal{C}(u, v)$. Notice that, this also means that $\ell(s) = \ell(t)$, since $\Gamma_\lambda^\mathcal{D}(\gamma_\lambda^\mathcal{C})(s, t) < \gamma_\lambda^\mathcal{C}(s, t)$, by hypothesis. If $\ell(u) = \ell(v)$ and $(u, v) \neq (s, t)$, by definition of $\mathcal{D}$, we have that $\bar{\omega}((u, v), \cdot) = \omega((u, v), \cdot)$, hence $\Gamma_\lambda^\mathcal{C}(\gamma_\lambda^\mathcal{C})(u, v) = \Gamma_\lambda^\mathcal{D}(\gamma_\lambda^\mathcal{C})(u, v)$. ∎

Lemma 7.3.2 states that $\mathcal{C}$ can be improved w.r.t. $\trianglelefteq_\lambda$ by updating its transition function at (s, t), whenever we find a distribution $\omega' \in \pi(s, \cdot) \otimes \pi(t, \cdot)$ such that

$$\textstyle\sum_{u,v \in S} \gamma_\lambda^\mathcal{C}(u, v) \cdot \omega'(u, v) < \sum_{u,v \in S} \gamma_\lambda^\mathcal{C}(u, v) \cdot \omega((s, t), (u, v)).$$

Notice that, an optimal schedule ω' for $TP(\gamma_\lambda^\mathcal{C}, \pi(s, \cdot), \pi(t, \cdot))$ enjoys the above condition, so that, the update $\mathcal{C}[(s, t)/\omega']$ improves $\mathcal{C}$. This gives us a strategy for moving toward δ_λ by successive improvements on the couplings.

Now we proceed giving sufficient and necessary condition for termination. This is done by first giving two technical results, Lemma 7.3.3 and 7.3.4, then we will be able to give a sufficient condition for termination (Lemma 7.3.5).

Lemma 7.3.3 *Let $s, t \in S$, and $\gamma_1^\mathcal{C} = \Delta_1(\gamma_\lambda^\mathcal{C})$. $\gamma_1^\mathcal{C}(s, t) = 1$ iff $\delta_1(s, t) = 1$.*

Proof. ($\Leftarrow$) Follows by Theorem 7.2.4. ($\Rightarrow$) Assume ω be the transition function of $\mathcal{C}$. If $\ell(s) \neq \ell(t)$ the thesis follows trivially. Assume $\ell(s) = \ell(t)$.

$$\begin{aligned} 1 = \gamma_1^\mathcal{C}(s, t) &= \Gamma_1^\mathcal{C}(\gamma_1^\mathcal{C})(s, t) \\ &= \textstyle\sum_{u,v \in S} \gamma_1^\mathcal{C}(u, v) \cdot \omega((s, t), (u, v)) \\ &\leq \textstyle\sum_{u,v \in S} \omega((s, t), (u, v)) = 1 \end{aligned}$$

Thus whenever $\omega((s, t), (u, v)) > 0$, $\gamma_1^\mathcal{C}(u, v) = 1$. By hypothesis we have $\gamma_1^\mathcal{C} = \Delta_1(\gamma_\lambda^\mathcal{C})$, therefore $1 = \gamma_1^\mathcal{C}(s, t) = \min_{\omega' \in \pi(s, \cdot) \otimes \pi(t, \cdot)} \sum_{u,v \in S} \gamma_1^\mathcal{C}(u, v) \cdot \omega'(u, v)$. Hence there is no coupling that can improve the summation. Therefore, by Theorem 7.2.4, $\delta_1(s, t) = 1$. ∎

Lemma 7.3.4 *For any $\lambda \in (0, 1]$, if $\gamma_\lambda^\mathcal{C} = \Delta_\lambda(\gamma_\lambda^\mathcal{C})$ then $\delta_\lambda = \gamma_\lambda^\mathcal{C}$.*

Proof. By Theorem 7.1.8, it suffices to prove that if $\gamma_\lambda^\mathcal{C}$ is a fixed point for Δ_λ, it is also the least one. We distinguish two cases: when $\lambda < 1$ and $\lambda = 1$.

For $\lambda < 1$, [28, Theorem 6] states that Δ_λ has a unique fixed point. By hypothesis $\gamma_\lambda^\mathcal{C}$ is a fixed point for Δ_λ, therefore it is also the least one.

For $\lambda = 1$, we proceed by contradiction. Assume $\delta_1 \neq \gamma_1^\mathcal{C}$ and ω be the transition function of $\mathcal{C}$. By $\delta_1 \neq \gamma_1^\mathcal{C}$ and Theorem 7.2.4, we have that $\delta_1 \sqsubset \gamma_1^\mathcal{C}$. Let $\Delta''\colon [0,1]^{S \times S} \to [0,1]^{S \times S}$ defined by

$$\Delta''(d)(s,t) = \begin{cases} 0 & \text{if } \gamma_1^\mathcal{C}(s,t) = 0 \\ \Delta_1(d)(s,t) & \text{otherwise} \end{cases}$$

Since Δ_1 is monotonic so is Δ''. Thus it admits greatest fixed-point, say g. By $\delta_1 \sqsubset \gamma_1^\mathcal{C}$ there exists $s,t \in S$ such that $\delta_1(s,t) < \gamma_1^\mathcal{C}(s,t)$, so that $\gamma_1^\mathcal{C}(s,t) \neq 0$.

Suppose that $\left\{ (s,t) \mid \gamma_1^\mathcal{C}(s,t) = 0 \right\} = \sim$, by [28, Corollary 18], Δ'' has a unique fixed point which corresponds to δ_1. By $\gamma_1^\mathcal{C} = \Delta_1(\gamma_1^\mathcal{C})$, we have that $\gamma_1^\mathcal{C} = \Delta''(\gamma_1^\mathcal{C})$, which contradicts the hypothesis that $\delta_1 \neq \gamma_1^\mathcal{C}$. Therefore, there exist $s,t \in S$ such that $\gamma_1^\mathcal{C}(s,t) \neq 0$ and $s \sim t$. It can be shown that there exists $s,t \in S$ satisfying the previous conditions and $g(s,t) = 1$. By Lemma 7.3.3, and $\delta_1(s,t) = 0$ we have that $\gamma_1^\mathcal{C}(s,t) < 1$. Thus $\gamma_1^\mathcal{C} \sqsubset g$. Now, let m and M be defined as

$$m = \max \left\{ g(s,t) - \gamma_1^\mathcal{C}(s,t) \mid s,t \in S \right\}, \quad M = \left\{ (s,t) \mid g(s,t) - \gamma_1^\mathcal{C}(s,t) = m \right\}.$$

By $\gamma_1^\mathcal{C} \sqsubset g$, $m > 0$. We prove first two properties on M:

$$M \cap \{ (s,t) \mid \ell(s) \neq \ell(t) \} = \emptyset \tag{7.3.3}$$

$$M \cap \left\{ (s,t) \mid \gamma_1^\mathcal{C}(s,t) = 0 \right\} = \emptyset \tag{7.3.4}$$

(7.3.3) follows since, for all $\ell(u) \neq \ell(v)$, $\gamma_1^\mathcal{C}(u,v) = 1 = g(u,v)$, and $m > 0$. (7.3.4) follows by definition of Δ'' and $m > 0$.

Let $(s,t) \in M$, then

$$
\begin{aligned}
m &= g(s,t) - \gamma_1^\mathcal{C}(s,t) \\
&= \Delta''(g)(s,t) - \Gamma_1^\mathcal{C}(\gamma_1^\mathcal{C})(s,t) \\
&= \Delta_1(g)(s,t) - \Gamma_1^\mathcal{C}(\gamma_1^\mathcal{C})(s,t) \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \text{(by (7.3.4))} \\
&= \left(\min_{\omega' \in \pi(s,\cdot) \otimes \pi(t,\cdot)} \sum_{u,v \in S} g(u,v) \cdot \omega'(u,v) \right) - \sum_{u,v \in S} \gamma_1^\mathcal{C}(u,v) \cdot \omega((s,t),(u,v)) \\
&\leq \sum_{u,v \in S} g(u,v) \cdot \omega((s,t),(u,v)) - \sum_{u,v \in S} \gamma_1^\mathcal{C}(u,v) \cdot \omega((s,t),(u,v)) \\
&= \sum_{u,v \in S} \left(g(u,v) - \gamma_1^\mathcal{C}(u,v) \right) \cdot \omega((s,t),(u,v)).
\end{aligned}
$$

Since, for all $u,v \in S$, $g(u,v) - \gamma_1^\mathcal{C}(u,v) \leq m$ and $\sum_{u,v \in S} \omega((s,t),(u,v)) = 1$, we have that, whenever $\omega((s,t),(u,v)) > 0$, $g(u,v) - \gamma_1^\mathcal{C}(u,v) = m$. Thus ω has support contained in M. This means that, for all $(s,t) \in M$, $R_\mathcal{C}((s,t)) \subseteq M$. Thus, by (7.3.3), we have that $\gamma_1^\mathcal{C}(s,t) = 0$, but this contradicts (7.3.4). ∎

Lemma 7.3.5 *For any discount factor $\lambda \in (0,1]$, if $\gamma_\lambda^\mathcal{C} \neq \delta_\lambda$, then there exist $s,t \in S$ and a coupling $\mathcal{D} = \mathcal{C}[(s,t)/\omega']$ for $\mathcal{M}$ such that $\Gamma_\lambda^\mathcal{D}(\gamma_\lambda^\mathcal{C})(s,t) < \gamma_\lambda^\mathcal{C}(s,t)$.*

Proof. We proceed by contraposition. Assume that for all $s,t \in S$ and for all couplings $\mathcal{D}$ such that $\mathcal{D} = \mathcal{C}[(s,t)/\omega']$, $\Gamma_\lambda^\mathcal{D}(\gamma_\lambda^\mathcal{C})(s,t) \geq \gamma_\lambda^\mathcal{C}(s,t)$. This corresponds to say that $\gamma_\lambda^\mathcal{C} = \Delta_\lambda(\gamma_\lambda^\mathcal{C})$. Then the thesis follows by Lemma 7.3.4. ∎

The above result ensures that, unless $\mathcal{C}$ is optimal w.r.t $\trianglelefteq_\lambda$, the hypothesis of Lemma 7.3.2 are satisfied, so that, we can further improve $\mathcal{C}$.

The next statement proves that this search strategy is correct.

Theorem 7.3.6 $\delta_\lambda = \gamma_\lambda^{\mathcal{C}}$ *iff there is no coupling $\mathcal{D}$ for $\mathcal{M}$ such that $\Gamma_\lambda^{\mathcal{D}}(\gamma_\lambda^{\mathcal{C}}) \sqsubset \gamma_\lambda^{\mathcal{C}}$.*

Proof. We prove: $\delta_\lambda \neq \gamma_\lambda^{\mathcal{C}}$ iff there exists $\mathcal{D}$ such that $\Gamma_\lambda^{\mathcal{D}}(\gamma_\lambda^{\mathcal{C}}) \sqsubset \gamma_\lambda^{\mathcal{C}}$. ($\Rightarrow$) Assume $\delta_\lambda \neq \gamma_\lambda^{\mathcal{C}}$. By Lemma 7.3.5, there are $s, t \in S$ and $\omega' \in \pi(s, \cdot) \otimes \pi(t, \cdot)$ such that $\lambda \cdot \sum_{u,v \in S} \gamma_\lambda^{\mathcal{C}}(u, v) \cdot \omega'(u, v) < \gamma_\lambda^{\mathcal{C}}(s, t)$. As in the proof of Lemma 7.3.2, we have that $\mathcal{D} = \mathcal{C}[(s,t)/\omega']$ satisflies $\Gamma^{\mathcal{D}}(\gamma_\lambda^{\mathcal{C}}) \sqsubset \gamma_\lambda^{\mathcal{C}}$. ($\Leftarrow$) Let $\mathcal{D}$ be such that $\Gamma_\lambda^{\mathcal{D}}(\gamma_\lambda^{\mathcal{C}}) \sqsubset \gamma_\lambda^{\mathcal{C}}$. By Tarski's fixed point theorem $\gamma_\lambda^{\mathcal{D}} \sqsubset \gamma_\lambda^{\mathcal{C}}$. By Theorem 7.2.4, $\delta_\lambda \sqsubseteq \gamma_\lambda^{\mathcal{D}} \sqsubset \gamma_\lambda^{\mathcal{C}}$. ∎

Remark 7.3.7 (Termination) Note that, in general there could be an infinite number of couplings for a given MC, so it is not obvious that our strategy is terminating.

Let us call *vertex coupling*, a coupling for $\mathcal{M}$ having a transition function ω such that, for all $s, t \in S$, $\omega((s,t), \cdot)$ is a vertex of $\pi(s, \cdot) \otimes \pi(t, \cdot)$. Since for all $s, t \in S$ the transportation polytope $\pi(s, \cdot) \otimes \pi(t, \cdot)$ has a finite number of vertices, the set of vertex couplings is finite. For each fixed $d \in [0,1]^{S \times S}$, the linear function mapping $\mu \in \pi(s, \cdot) \otimes \pi(t, \cdot)$ to $\lambda \sum_{u,v \in S} d(u, v) \cdot \mu(u, v)$ achieves its minimum at some vertex in $\pi(s, \cdot) \otimes \pi(t, \cdot)$. Thus, using any optimal TP schedule for the update (which has not to be necessarily a vertex of the transportation polytope) we ensure the strategy is always terminating. Indeed, the couplings that are encountered during any computation can be immersed in the set of vertex couplings where $\vartriangleleft_\lambda$ is obviously well-founded. ∎

7.4 The On-the-fly Algorithm

In this section, we describe an on-the-fly algorithm for the exact computation of the bisimilarity pseudometric δ_λ making full use of the greedy strategy presented in Section 7.3.2.

Let $Q \subseteq S \times S$. Assume we want to compute $\delta_\lambda(s, t)$, for all $(s, t) \in Q$. The method proposed in Section 7.3.2 has the following key features:

1. the improvement of each coupling $\mathcal{C}$ is obtained by a *local* update of its transition function at some state (u, v) in $\mathcal{C}$;
2. the strategy does not depend on the choice of the state (u, v);
3. whenever a coupling $\mathcal{C}$ is considered, the over-approximation $\gamma_\lambda^{\mathcal{C}}$ of the distance can be computed by solving a system of linear equations.

Among them, only the last one requires a visit of the coupling. However, as noticed in Remark 7.3.1, the value $\gamma_\lambda^{\mathcal{C}}(s, t)$ can be computed without considering the entire linear system of Equation (7.3.2), but only its smallest independent subsystem containing the variable $x_{(s,t)}$, which is obtained by restricting on the variables $x_{(u,v)}$ such that $(u, v) \in R_C((s,t))$. This subsystem can be further reduced, by Gaussian elimination, when some values for δ_λ are known. The last observation suggests that, in order to compute $\gamma_\lambda^{\mathcal{C}}(s, t)$, we do not need to store the entire coupling, but it can be constructed on-the-fly.

The exact computation of the bisimilarity pseudometric is implemented by Algorithm 1. It takes as input an MC $\mathcal{M} = (S, A, \pi, \ell)$, a discout factor λ, and a query set Q. We assume the following variables to store:

- $\mathcal{C}$: the current partial coupling;
- d: the λ-discrepancy associated with $\mathcal{C}$;
- $ToCompute$: the pairs of states for which the distance has to be computed;
- $Exact$: the pairs of states (s, t) such that $d(s, t) = \delta_\lambda(s, t)$;
- $Visited$: the states of $\mathcal{C}$ considered so far.

At the beginning (line 1) both the coupling $\mathcal{C}$ and the discrepancy d are empty, there are no visited states, and no exact computed distances. While there are still pairs left to be computed (line 2), we pick one (line 3), say (s, t). According to the definition of δ_λ, if $\ell(s) \neq \ell(t)$ then $\delta_\lambda(s, t) = 1$; if $s = t$ then $\delta_\lambda(s, t) = 0$, so that, $d(s, t)$ is set accordingly, and (s, t) is added to $Exact$ (lines 4–7). Otherwise, if (s, t) was not previously visited, a matching $\omega \in \pi(s, \cdot) \otimes \pi(t, \cdot)$ is guessed, and the routine $SetPair$ updates the coupling $\mathcal{C}$ at (s, t) with ω (line 9), then the routine $Discrepancy$ updates d with the λ-discrepancy associated with $\mathcal{C}$ (line 10). According to the greedy strategy, $\mathcal{C}$ is successively improved and d is consequently updated, until no further

---

Algorithm 1 On-the-fly Bisimilarity Pseudometric

Input: MC $\mathcal{M} = (S, A, \pi, \ell)$; discout factor $\lambda \in (0, 1]$; query $Q \subseteq S \times S$.

 1. $\mathcal{C} \leftarrow$ empty; $d \leftarrow$ empty; $Visited \leftarrow \emptyset$; $Exact \leftarrow \emptyset$; $ToCompute \leftarrow Q$;   // Init.
 2. **while** $ToCompute \neq \emptyset$ **do**
 3.   pick $(s, t) \in ToCompute$
 4.   **if** $\ell(s) \neq \ell(t)$ **then**
 5. $d(s, t) \leftarrow 1$; $Exact \leftarrow Exact \cup \{(s, t)\}$; $Visited \leftarrow Visited \cup \{(s, t)\}$
 6.   **else if** $s = t$ **then**
 7. $d(s, t) \leftarrow 0$; $Exact \leftarrow Exact \cup \{(s, t)\}$; $Visited \leftarrow Visited \cup \{(s, t)\}$
 8.   **else**   // if (s, t) is nontrivial
 9.   **if** $(s, t) \notin Visited$ **then** pick $\omega \in \pi(s, \cdot) \otimes \pi(t, \cdot)$; $SetPair(\mathcal{M}, (s, t), \omega)$
10. $Discrepancy(\lambda, (s, t))$ // update d as the λ-discrepancy for $\mathcal{C}$
11.   **while** $\exists (u, v) \in R_{\mathcal{C}}((s, t))$. $\mathcal{C}[(u, v)]$ not opt. for $TP(d, \pi(u, \cdot), \pi(v, \cdot))$ **do**
12. $\omega \leftarrow$ optimal schedule for $TP(d, \pi(u, \cdot), \pi(v, \cdot))$
13. $SetPair(\mathcal{M}, (u, v), \omega)$ // improve the current coupling
14. $Discrepancy(\lambda, (s, t))$ // update d as the λ-discrepancy for $\mathcal{C}$
15.   **end while**
16. $Exact \leftarrow Exact \cup R_{\mathcal{C}}((s, t))$ // add new exact distances
17.   remove from $\mathcal{C}$ all edges exiting from nodes in $Exact$
18.   **end if**
19. $ToCompute \leftarrow ToCompute \setminus Exact$ // remove exactly computed pairs
20. **end while**
21. **return** $d{\restriction}_Q$ // return the distance for all pairs in Q

---

improvements are possible (lines 11–15). Each improvement is demanded by the existence of a better schedule for $TP(d, \pi(u, \cdot), \pi(u, \cdot))$ (line 11). Note that, each improvement actually affects the current value of $d(s, t)$. This is done by restricting our attention only on the pairs that are reachable from (s, t) in $\mathcal{G}(\mathcal{C})$. It is worth to note that, $\mathcal{C}$ is constantly updated, hence $R_{\mathcal{C}}((s, t))$ may differ from one iteration to another. When line 16 is reached, for each $(u, v) \in R_{\mathcal{C}}((s, t))$, we are guaranteed that $d(u, v) = \delta_\lambda(s, t)$, therefore $R_{\mathcal{C}}((s, t))$ is added to $Exact$, and these values can be used in successive computations, so the edges exiting from these states are removed from $\mathcal{G}(\mathcal{C})$. In line 19, the exact pairs computed so far are removed from $ToCompute$. Finally, if no more pairs need be considered, the exact distance on Q is returned (line 21).

Algorithm 1 calls the subroutines $SetPair$ and $Discrepancy$, respectively, to construct/update the coupling $\mathcal{C}$, and to update the current over-approximation d during the computation. Now we explain how they works.

$SetPair$ (Algorithm 2) takes as input an MC $\mathcal{M} = (S, A, \pi, \ell)$, a pair of states (s, t), and a matching $\omega \in \pi(s, \cdot) \otimes \pi(t, \cdot)$. In lines 1–2 the transition function of the coupling $\mathcal{C}$ is set to ω at (s, t), then (s, t) is added to $Visited$. The on-the-fly construction of the coupling is recursively propagated to the successors of (s, t) in $\mathcal{G}(\mathcal{C})$. During this construction, if some states with trivial distance are encountered, d and $Exact$ are updated accordingly (lines 5–6).

$Discrepancy$ (Algorithm 3) takes as input a discount factor λ and a pair of states (s, t). It constructs the smallest (reduced) independent subsystem of Equation 7.3.2 having the variable $x_{(s,t)}$ (lines 9–10). As noticed in Remark 7.3.1, the least solution is computed by fixing d to zero for all the pairs which cannot be reached from any pair in $Exact$ and such that its distance is greater than zero (lines 5–7). Then, the discrepancy is computed and d is consequently updated.

Next, we present a simple example of Algorithm 1, showing the main features of our method: (1) the on-the-fly construction of the (partial) coupling, and (2) the restriction only to those variables which are demanded for the solution of the system of linear equations.

Example 7.4.1 (On-the-fly computation) Let us compute the undiscounted distance between states 1 and 4 for the $\{red, blu\}$-labeled MC depicted in Figure 7.1.

Algorithm 2 $SetPair(\mathcal{M}, (s,t), \omega)$

Input: MC $\mathcal{M} = (S, A, \pi, \ell)$; $s, t \in S$; $\omega \in \pi(s, \cdot) \otimes \pi(t, \cdot)$

1. $\mathcal{C}[(s,t)] \leftarrow \omega$ // update the coupling at (s,t) with ω
2. $Visited \leftarrow Visited \cup \{(s,t)\}$ // set (s,t) as visited
3. **for all** $(u,v) \in \{(u',v') \mid \omega(u',v') > 0\} \setminus Visited$ **do**   // for all demanded pairs
4. $Visited \leftarrow Visited \cup \{(u,v)\}$
5.   **if** $u = v$ **then** $d(u,v) \leftarrow 0$; $Exact \leftarrow Exact \cup \{(u,v)\}$;
6.   **if** $\ell(u) \neq \ell(v)$ **then** $d(u,v) \leftarrow 1$; $Exact \leftarrow Exact \cup \{(u,v)\}$;
7.   // propagate the construction
8.   **if** $(u,v) \notin Exact$ **then**
9.   pick $\omega' \in \pi(u, \cdot) \otimes \pi(v, \cdot)$ // guess a matching
10. $SetPair(\mathcal{M}, (u,v), \omega')$
11.   **end if**
12. **end for**

Algorithm 3 $Discrepancy(\lambda, (s,t))$

Input: discount factor $\lambda \in (0,1]$; partial coupling $\mathcal{C}$; approx. distance d; $s, t \in S$

1. $Nonzero \leftarrow \emptyset$ // detect non-zero variables
2. **for all** $(u,v) \in R_{\mathcal{C}}((s,t)) \cap Exact$ **such that** $d(u,v) > 0$ **do**
3. $Nonzero \leftarrow Nonzero \cup \{(u',v') \mid (u,v) \rightsquigarrow (u',v') \text{ in } \mathcal{G}^{-1}(\mathcal{C})\}$
4. **end for**
5. **for all** $(u,v) \in R_{\mathcal{C}}((s,t)) \setminus Nonzero$ **do**   // set distance to zero
6. $d(u,v) \leftarrow 0$; $Exact \leftarrow Exact \cup \{(u,v)\}$
7. **end for**
8. // construct the reduced linear system over nonzero variables
9. $\mathbf{A} \leftarrow (C[(u,v)](u',v'))_{(u,v),(u',v') \in Nonzero}$
10. $\vec{b} \leftarrow \left(\sum_{(u',v') \in Exact} d(u',v') \cdot C[(u,v)](u',v') \right)_{(u,v) \in Nonzero}$
11. $\tilde{\vec{x}} \leftarrow$ solve $\vec{x} = \lambda \mathbf{A} \vec{x} + \lambda \vec{b}'$ // solve the reduced linear system
12. **for all** $(u,v) \in Nonzero$ **do**   // update distances
13. $d(u,v) \leftarrow \tilde{x}_{(u,v)}$
14. **end for**

Algorithm 1 guesses an initial coupling $\mathcal{C}_0$ with transition distribution ω_0. This is done considering only the pairs of states which are needed: starting from $(1,4)$, the distribution $\omega_0((1,4), \cdot)$ is guessed as in Figure 7.1, which demands for the exploration of $(3,4)$ and a guess $\omega_0((3,4), \cdot)$. Since no other pairs are demanded, the construction of $\mathcal{C}_0$ terminates. This gives the equation system:

$$\begin{cases} x_{1,4} = \frac{1}{3} \cdot \overbrace{x_{1,2}}^{=1} + \frac{1}{3} \cdot \overbrace{x_{2,3}}^{=1} + \frac{1}{6} \cdot x_{3,4} + \frac{1}{6} \cdot \overbrace{x_{3,6}}^{=1} = \frac{1}{6} \cdot x_{3,4} + \frac{5}{6} \\ x_{3,4} = \frac{1}{3} \cdot \overbrace{x_{1,2}}^{=1} + \frac{1}{6} \cdot \overbrace{x_{2,2}}^{=0} + \frac{1}{6} \cdot \overbrace{x_{2,3}}^{=1} + \frac{1}{3} \cdot \overbrace{x_{3,3}}^{=0} = \frac{1}{2} . \end{cases}$$

Note that the only variables appearing in the above equation system correspond to the pairs which have been considered so far. The least solution for it is given by $d^{\mathcal{C}_0}(1,4) = \frac{11}{12}$ and $d^{\mathcal{C}_0}(3,4) = \frac{1}{2}$.

Now, these solutions are taken as the costs of a TP, from which we get an optimal transportation schedule $\omega_1((1,4), \cdot)$ improving $\omega_0((1,4), \cdot)$. The distribution ω_1 is used to update $\mathcal{C}_0$ to $\mathcal{C}_1 = \mathcal{C}_0[(1,4)/\omega_1]$ (depicted in Figure 7.1), obtaining the following new equation system:

$$x_{1,4} = \frac{1}{3} \cdot \overbrace{x_{2,2}}^{=0} + \frac{1}{3} \cdot \overbrace{x_{3,3}}^{=0} + \frac{1}{6} \cdot x_{1,4} + \frac{1}{6} \cdot \overbrace{x_{1,6}}^{=1} = \frac{1}{6} \cdot x_{1,4} + \frac{1}{6} ,$$

which has $d^{\mathcal{C}_1}(1,4) = \frac{1}{5}$ as least solution. Note that, $(3,4)$ is no more demanded, thus we do not need to update it. Running again the TP on the improved over-approximation $d^{\mathcal{C}_1}$, we discover that

$\mathcal{C}_0$

(1,4)	1	2	3	
2	$\frac{1}{3}$			$\frac{1}{3}$
3		$\frac{1}{3}$		$\frac{1}{3}$
4			$\frac{1}{6}$	$\frac{1}{6}$
6			$\frac{1}{6}$	$\frac{1}{6}$
	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	

(3,4)	1	2	3	
2	$\frac{1}{3}$	$\frac{1}{6}$		$\frac{1}{2}$
3		$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$
	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	

$\mathcal{C}_1$

(1,4)	1	2	3	
2		$\frac{1}{3}$		$\frac{1}{3}$
3			$\frac{1}{3}$	$\frac{1}{3}$
4	$\frac{1}{6}$			$\frac{1}{6}$
6	$\frac{1}{6}$			$\frac{1}{6}$
	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	
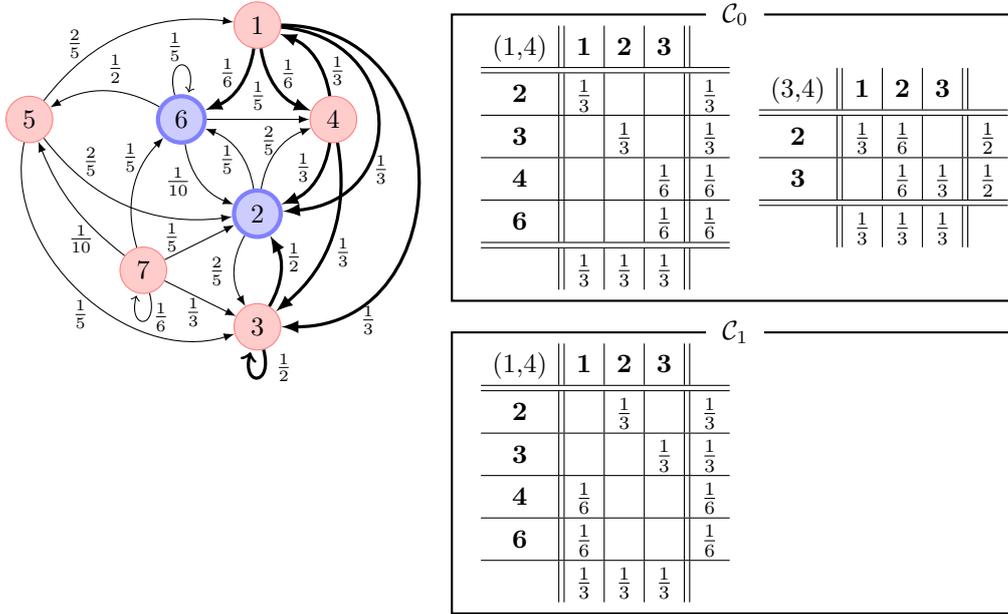
Figure 7.1: Execution trace for the computation of $\delta_1(1,4)$ (details in Example 7.4.1).

the coupling $\mathcal{C}_1$ cannot be further improved, hence we stop the computation, returning $\delta_1(1,4) = d^{\mathcal{C}_1}(1,4) = \frac{1}{5}$.

It is worth noticing that, Algorithm 1 does not explore the entire MC, not even all the reachable states from 1 and 4. The only edges in the MC which have been considered during the computation are highlighted in Figure 7.1.   ∎

Remark 7.4.2 Notably, Algorithm 1 can also be used for computing over-approximated distances. Indeed, assuming over-estimates for some particular distances are already known, they can be taken as inputs and used in our algorithm simply storing them in the variable d and treated as "exact" values. In this way our method will return the least over-approximation of the distance agreeing with the given over-estimates. This modification of the algorithm can be used to further decrease the exploration of the MC. Moreover, it can be employed in combination with other existing approximated algorithms, having the advantage of an on-the-fly state space exploration.   ∎

7.5   Experimental Results

In this section, we evaluate the performances of the on-the-fly algorithm on a collection of randomly generated MCs[1].

First, we compare the execution times of the on-the-fly algorithm with those of the iterative method proposed in [28] in the discounted case. Since the iterative method only allows for the computation of the distance for all state pairs at once, the comparison is (in fairness) made with respect to runs of our on-the-fly algorithm with input query the set of all state pairs. For each input instance, the comparison involves the following steps:

1. we run the on-the-fly algorithm, storing both execution time and the number of solved transportation problems,

---

[1] The tests have been made using a prototype implementation coded in Mathematica® (available at `http://people.cs.aau.dk/~mardare/projects/tools/mc_dist.zip`) running on an Intel Core-i7 3.4 GHz processor with 12GB of RAM.

# States	On-the-Fly (exact)		Iterating (approximated)			Approximation
	Time (s)	# TPs	Time (s)	# Iterations	# TPs	Error
5	0.019675	1.19167	0.0389417	1.73333	26.7333	0.139107
6	0.05954	3.04667	0.09272	1.82667	38.1333	0.145729
7	0.13805	6.01111	0.204789	2.19444	61.7278	0.122683
8	0.255067	8.5619	0.364019	2.30476	83.0286	0.11708
9	0.499983	12.0417	0.673275	2.57917	114.729	0.111104
10	1.00313	18.7333	1.27294	3.11111	174.363	0.0946047
11	2.15989	25.9733	2.66169	3.55667	239.557	0.0959714
12	4.64225	34.797	5.52232	4.04242	318.606	0.0865612
13	6.73513	39.9582	8.06186	4.63344	421.675	0.0977743
14	6.33637	38.0048	7.18807	4.91429	593.981	0.118971
17	11.2615	47.0143	12.8048	5.88571	908.61	0.13213
19	26.6355	61.1714	29.6542	6.9619	1328.6	0.14013
20	34.379	66.4571	38.2058	7.5381	1597.92	0.142834

Table 7.1: Comparison between the on-the-fly algorithm and the iterative method.

# States	outer-degree = 3		$2 \leq$ outer-degree $\leq$ # States*	
	Time (s)	# TPs	Time (s)	# TPs
5	0.00594318	0.272727	0.011654	0.657012
6	0.0115532	0.548936	0.0304482	1.66696
7	0.0168408	0.980892	0.0884878	3.67706
8	0.0247971	1.34606	0.164227	5.30112
9	0.0259426	1.29074	0.394543	8.16919
10	0.0583405	2.03887	1.1124	13.0961
11	0.0766988	1.82706	2.22016	18.7228
12	0.0428891	1.62038	4.94045	26.0965
13	0.06082	1.88134	10.3606	35.1738
14	0.0894778	2.79441	20.1233	46.0775
20	0.35631	6.36833	1.5266	13.1367
30	4.66113	17.3167	74.8146	76.2642
50	27.2147	30.8217	2234.54	225.394

Table 7.2: Average performances of the on-the-fly algorithm on single-pair queries. On the first to columns the outer-degree is 3; on the last two columns, the outer-degree varies from 2 to # States. (*) For 20, 30 and 50 states, outer-degree is 4;

2. then, on the same instance, we execute the iterative method until the running time exceeds that of step 1. We report the execution time, the number of iterations, and the number of solved transportation problems.

3. Finally, we calculate the approximation error between the exact solution δ_λ computed by our method at step 1 and the approximate result d obtained in step 2 by the iterative method, as $\max_{s,t \in S} \delta_\lambda(s,t) - d(s,t)$.

This has been made on a collection of MCs varying from 5 to 20 states. For each $n = 5, \ldots, 20$, we have considered 80 randomly generated MCs per outer-degree, varying from 2 to n. Table 7.1 reports the average results of the comparison.

As can be seen, the our use of a greedy strategy in the construction of the couplings leads to a significant improvement in the performances. We are able to compute the exact solution before the iterative method can under-approximate it with an error of ≈ 0.1, which is a considerable error for a value in $[0, 1]$.

So far, we only examined the case when the on-the-fly algorithm is run on all state pairs at once. Now, we show how the performance of our method is improved even further when the distance is computed only for single pairs of states. Table 7.2 shows the average execution times and number of solved transportation problems for (nontrivial) single-pair queries for randomly generated of MCs with number of states varying from 5 to 50. In the first two columns we consider MCs with outer-degree equal to 3, while the last two columns show the average values for outer-degrees
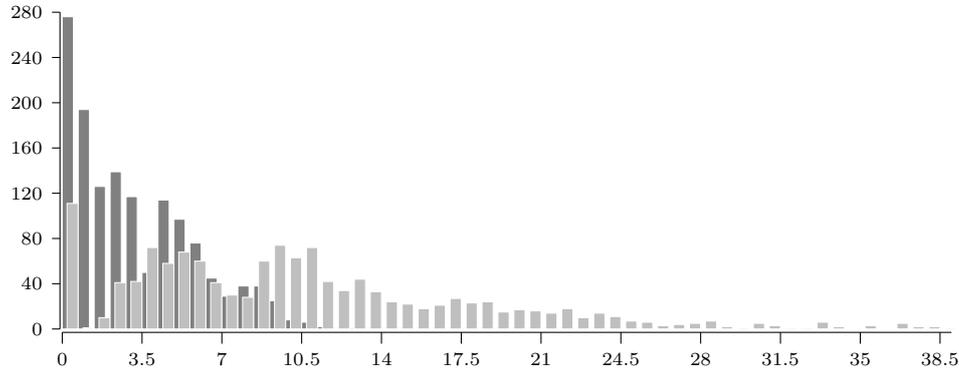
Figure 7.2: Distribution of the execution times (in seconds) for 1332 tests on randomly generated MCs with 14 states, out-degree 6 (darkest) and 8 (lightest).

varying from 2 to the number of states of the MCs. The results show that, when the outer-degree of the MCs is low, our algorithm performs orders of magnitude better than in the general case. This is illustrated in Figure 7.2, where the distributions of the execution times for outer-degree 6 and 8 are juxtaposed, in the case of MCs with 14 states. Each bar in the histogram represents the number of tests that terminate within the time interval indicated in the x-axsis.

Notably, our method may perform better on large queries then on single-pairs queries. This is due to the fact that, although the returned value does not depend on the order the queried pairs are considered, a different order may speed up the performances. So that, when the algorithm is run on more than a single pair, the way they are picked may increase the performances (e.g., compare the execution times in Tables 7.1 and 7.2 for MCs with 14 states).

8

Conclusions and Future Work

We conclude by briefly recalling the main contributions and techniques used in this thesis, and by listing some related work and possible further directions of research related to each subject that has been considered.

Initial and Final Sequences in Categories with Factorization Systems. In Chapter 4 we have considered alternative constructions for initial algebras and final coalgebras for endofunctors $F\colon \mathbf{C} \to \mathbf{C}$ in categories with factorization systems. To this end, we have exploited initial and final sequences in combination with the axiomatic properties of factorization systems. The key intuition behind the use of factorization systems relies on the fact that they are good generalizations of the notions of subobject and quotient, which always played a crucial rôle in the construction of final coalgebras and initial algebras.

Aczel and Mendler [4] obtained a final coalgebra as a quotient (by bisimilarity) of a coproduct of a set of coalgebras. Barr [14] showed that if a set functor T is accessible then the category of T-coalgebras as a set (not a proper class!) of generators. He then used the Special Adjoint Functor theorem, whose proof also involves a quotient of a sum-construction, to derive the existence of a final coalgebra.

In [92, 94, 93] Worrell adopted the approach of Adámek and Koubek [8] and Barr [13] using final sequences which generalizes the iterative construction of the greatest fixed point of a monotone function on a complete lattice. For **Set**-endofunctors *accessibility* seems to be a common denominator among some of the hypothesis involved in the various final coalgebra theorems in the literature, e.g. being *bounded* in [58] and *set based* in [4]. In [5] it has been shown that the assumption of boundedness of a **Set**-functor is equivalent to accessibility. Since then, accessible categories received much attention in order to give sufficient conditions for the existence of a final coalgebra.

Informally, accessibility describes a generalized notion of "smallness" for a particular set of objects in the category, the so called *representable objects*. These objects are colimits of a bounded set of objects (formally, a κ-filtered diagram), and they enjoy the property that any morphism to them can be factorized through some object in the colimit diagram. Intuitively, this amounts to say that these objects can be fully described by means of the objects belonging to the colimit diagram. Asking that such diagrams are small corresponds to say that the information carried by representable objects can be encoded by a set rather then a proper class. Accessible endofunctors are functors between accessible categories, such that they preserve κ-filtered colimits, for some ordinal κ. Although this is a very general and reasonable categorical notion of "smallness" for objects, many categories of interest fail to be accessible (an important example is **Top**). Moreover, it is usually hard to prove that a category is accessible and, following the final coalgebra construction in [94], functors have to be accessible as well.

Our approach has many similarities, but the notion of "smallness" is characterized in a much more simpler way by means of factorization systems. It would be interesting to relate all these techniques to our. A tentative in this direction has been already done in Theorem 4.2.12, which slightly generalizes a well-known theorem of stabilization for final sequences of **Set**-functors due to Worrell [92, Theorem 4.6]. As a side remark, note that accessibility can be weakened to a notion

of $\mathcal{M}$-accessibility, where $\mathcal{M}$ is a class of morphisms and representable objects must be colimits of (κ-filtered) diagrams with morphisms in **M**. With this notion **Top** is **Emb**-accessible, where **Emb** is the class of topological embeddings.

Bisimulation for Labelled Markov processes   In Chapter 5 we proved that bisimilarity on generalized Markov processes is an equivalence without assuming that the state space is analytic. The proof is given via a characterization of the coalgebraic bisimulation, introduced in order to overcome some technical problems that do not allow to apply the standard techniques usually employed in the theory of universal coalgebras. For example the functor Δ does not preserve weak pullbacks, hence transitivity for bisimilarity and the existence of a "maximal" coalgebraic bisimulation cannot be proved via standard (weak) universal constructions. Typically, in case the behavior functor does not preserves weak pullbacks, the existence of a maximal bisimulation can be proved without having recourse to universal properties, but requiring that the underlying category has well-behaved factorization systems for which the left class of morphisms has right inverse. This technique usually works in **Set**, but not in **Meas**, since epimorphisms do not have right inverse. An open question is if there are techniques which can be applied in order to prove the same result within the categorical language. A possible strategy could be to use natural factorization systems [49] instead of classical factorization systems, where left and right classes of morphisms are represented as comonads and monads in the category of arrows.

The other contribution of the chapter is a coalgebraic analysis of the relationships between bisimulation and cocongruences over Markov processes. This is done establishing an adjunction between the category of bisimulations and that of cocongruences (actually, only a subcategory of the latter). The adjunction gives rise to a closure operator (z-closure) that works as a kind of transitive closure for bisimulation relations. In the study of probabilistic systems coalgebraically, but also in coalgebraic modal logics in general, behavioral equivalence has advantages over bisimilarity. However, the good side of bisimilarity is that it is computable by efficient algorithms (in the case of labelled transition systems, (non-)deterministic automata, etc.), and traditionally many concrete types of systems come equipped with a concrete notion of bisimilarity. Hence formal techniques aimed at establishing a bridge between the notion of coalgebraic bisimulation and cocongruence are of particular interest.

Congruential Rule Formats for Markov Processes   In Chapter 6, we have introduced *Measure GSOS specification systems*, an SOS specification format for continuous state probabilistic and stochastic calculi. To show the expressivity of the proposed rule format, we have introduced two simple yet paradigmatic calculi where continuous data affect the operational description of processes: the *Quantitative CSS* and *FlatCCS*.

In this format, transitions have the form $t \xrightarrow{\alpha} \mu$, where t is a process term, and μ is a *measure term*, i.e., an expression over a specifically designed language aimed at denoting a (finite) measure over the measurable space of processes. An MGSOS specification is composed by a set of GSOS-like inference rules, and a measure term interpretation, i.e., a natural transformation taking measure terms to their denotation as an actual measure. The rule set yields a labelled transition system corresponding to the collective semantics of all the derivable measure terms for a given process. Then, each measure term is given a denotation via the measure terms interpretation, and the overall operational semantics is given by summing up the set of partial behaviors. It is interesting to compare this format with the usual GSOS. In particular, in a transition $t \xrightarrow{\alpha} \mu$, the source term t and the target term μ are from different languages, defined by two different syntactic monads. The connection between these languages is provided by the measure term interpretation, which is a kind of "distributive law" across two languages. Notably, the usual GSOS format can be seen as a special case, when the two languages for the source and target of transitions are the same, so that in this case, the interpretation is just the identity natural transformation. Thus, a possible future work is to investigate the use of different syntactic languages for the sources and targets of transitions, in combination with interpretations. This may help to give purely syntactical representations to operational semantics for which the behavior functors have a shape that it is

difficult to be represented syntactically.

MGSOS specifications yield well-behaved operational semantics for which the induced behavioral equivalence is a congruence. This is proved categorically, showing that MGSOS specifications give rise to abstract GSOS distributive laws of type $S(Id \times \Delta_{<\infty}^L) \Rightarrow (\Delta_{<\infty} T_S)^L$ and providing a canonical universal fully-abstract semantics as both initial and final morphism in the category of bialgebras for the distributive law. The bialgebraic framework of Turi and Plotkin [83], however, provides other formats as well. In particular, there is a categorical dual of abstract GSOS which, for example, captures formats that allow for look-ahead in the case of labelled transition systems, i.e. the premises of the derivation rules may refer to several successive transitions of the arguments instead of just the immediate outgoing transitions. It would be interesting to try to derive a rule format for this dual abstract distributive law also for labelled Markov processes. In the literature, this kind of dual formats has not received much attention yet, even in the case of LTSs. A reason for this, may be that the cofree comonad, on which the dual format is based, is much more difficult to work with than the free monad generated by a signature. This problem still holds in the case of labelled Markov processes, hence a careful examination on the structure on the cofree comonad over $\Delta_{<\infty}$ is another interesting direction for future work.

Metric Bisimulations In Chapter 7 we have proposed an on-the-fly algorithm for computing exactly the bisimilarity distance between Markov chains, introduced by Desharnais et al. in [39]. Our algorithm represents an important improvement of the state of the art in this field where, before our contribution, the known tools were only concerned with computing approximations of the bisimilarity distances and they were, in general, based on iterative techniques. We demonstrate that, using on-the-fly techniques, we cannot only calculate exactly the bisimilarity distance, but the computation time is improved with orders of magnitude with respect to the corresponding iterative approaches. Moreover, our technique allows for the computation on a set of target distances that might be done by only investigating a significantly reduced set of states, and for further improvement of speed.

Our algorithm can be practically used to address a large spectrum of problems. For instance, it can be seen as a method to decide whether two states of a given Markov chain are probabilistic bisimilar, to identify bisimilarity classes, or to solve lumpability problems. It is sufficiently robust to be used with approximation techniques as, for instance, to provide a least over-approximation of the behavioral distance given over-estimates of some particular distances. It can be integrated with other approximate algorithms, having the advantage of the efficient on-the-fly state space exploration.

Having a practically efficient tool to compute bisimilarity distances opens the perspective of new applications already announced in previous research papers. One of these is the state space reduction problem for Markov chains. Our technique can be used in this context as an indicator for the sets of neighbor states that can be collapsed due to their similarity; it also provides a tool to estimate the difference between the initial Markov chain and the reduced one, hence a tool for the approximation theory of Markov chains.

Bibliography

[1] Edwin Abbott Abbott. *Flatland: A Romance of Many Dimensions*. Blackwell, 1884.

[2] Luca Aceto, Wan Fokkink, and Chris Verhoef. Structural operational semantics. In Jan Bergstra, Alban Ponse, and Scott Smolka, editors, *Handbook of Process Algebra*, pages 197–292. Elsevier Science, 2001.

[3] Peter Aczel. Non-well-founded sets. CSLI Lecture Notes 14, Stanford University, 1988.

[4] Peter Aczel and Nax Mendler. A final coalgebra theorem. In D. H. Pitt, D. E. Rydeheard, P. Dybjer, A. M. Pitts, and A. Poigné, editors, *Proceedings 3rd Conf. on Category Theory and Computer Science, CTCS'89*, volume 389 of *Lecture Notes in Computer Science*, pages 357–365, Berlin, 1989. Springer-Verlag.

[5] Jirí Adámek and Hans-E. Porst. From varieties of algebras to covarieties of coalgebras. *Electronic Notes Theoretical Computer Science*, 44(1):27–46, 2001.

[6] Jiří Adámek. On a description of terminal coalgebras and iterative theories. *Electr. Notes Theor. Comput. Sci.*, 82(1):1–16, 2003.

[7] Jiří Adámek. On final coalgebras of continuous functors. *Theor. Comput. Sci.*, 294(1/2):3–29, 2003.

[8] Jiří Adámek and Václav Koubek. On the greatest fixed point of a set functor. *Theoretical Computer Science*, 150(1):57–75, 1995.

[9] Jiří Adámek and Vera Trnková. *Automata and Algebras in Categories*. Kluwer Academic Publishers, Norwell, MA, USA, 1st edition, 1990.

[10] Giorgio Bacci and Marino Miculan. Measurable stochastics for Brane Calculus. *Theoretical Computer Science*, 431:117–136, 2012.

[11] Christel Baier and Joost Pieter Katoen. *Principles of Model Checking*. MIT Press, 2008.

[12] Roberto Barbuti, Andrea Maggiolo-Schettini, Paolo Milazzo, and Giovanni Pardini. Spatial calculus of looping sequences. *Electr. Notes Theor. Comput. Sci.*, 229(1):21–39, 2009.

[13] Michael Barr. Algebraically compact functors. *Journal of Pure and Applied Algebra*, 82(3):211 – 231, 1992.

[14] Michael Barr. Terminal coalgebras in well-founded set theory. *Theoretical Computer Science*, 114(2):299–315, 1993.

[15] Falk Bartels. *On Generalised Coinduction and Probabilistic Specification Formats: Distributive Laws in Coalgebraic Modelling*. PhD thesis, CWI, Amsterdam, 2004. `http://homepages.cwi.nl/~bartels/Papers/Thesis.pdf`.

[16] Falk Bartels, Ana Sokolova, and Erik P. de Vink. A hierarchy of probabilistic system types. *Electr. Notes Theor. Comput. Sci.*, 82(1), 2003.

[17] Bard Bloom, Sorin Istrail, and Albert R. Meyer. Bisimulation can't be traced. *J. ACM*, 42(1):232–268, 1995.

[18] Richard Blute, Josee Desharnais, Abbas Edalat, and Prakash Panangaden. Bisimulation for labelled Markov processes. In *LICS*, pages 149–158. IEEE Computer Society, 1997.

[19] Francis Borceux. *Handbook of Categorical Algebra 1*. Number 50 in Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1994.

[20] Graham Brightwell, Jan van den Heuvel, and Leen Stougie. A Linear Bound On The Diameter Of The Transportation Polytope. *Combinatorica*, 26(2):133–139, April 2006.

[21] Linda Brodo, Pierpaolo Degano, Stephen Gilmore, Jane Hillston, and Corrado Priami. Performance evaluation for global computation. In Corrado Priami, editor, *Global Computing*, volume 2874 of *Lecture Notes in Computer Science*, pages 229–253. Springer, 2003.

[22] Stephen D. Brookes, C. A. R. Hoare, and A. W. Roscoe. A theory of communicating sequential processes. *Journal of ACM*, 31(3):560–599, 1984.

[23] Xiaojuan Cai and Yonggen Gu. Measuring Anonymity. In *ISPEC '09*, pages 183–194, Berlin, Heidelberg, 2009. Springer-Verlag.

[24] Luca Cardelli and Philippa Gardner. Processes in space. In Fernando Ferreira, Benedikt Löwe, Elvira Mayordomo, and Luís Mendes Gomes, editors, *CiE*, volume 6158 of *Lecture Notes in Computer Science*, pages 78–87. Springer, 2010.

[25] Luca Cardelli, Kim G. Larsen, and Radu Mardare. Modular markovian logic. In Luca Aceto, Monika Henzinger, and Jiri Sgall, editors, *ICALP (2)*, volume 6756 of *Lecture Notes in Computer Science*, pages 380–391. Springer, 2011.

[26] Luca Cardelli and Radu Mardare. The measurable space of stochastic processes. In *Proc. QEST*, pages 171–180. IEEE Computer Society, 2010.

[27] Krishnendu Chatterjee, Luca de Alfaro, Rupak Majumdar, and Vishwanath Raman. Algorithms for Game Metrics. *Logical Methods in Computer Science*, 6(3), 2010.

[28] Di Chen, Franck van Breugel, and James Worrell. On the Complexity of Computing Probabilistic Bisimilarity. In Lars Birkedal, editor, *FoSSaCS*, volume 7213 of *Lecture Notes in Computer Science*, pages 437–451. Springer, 2012.

[29] Gheorghe Comanici, Prakash Panangaden, and Doina Precup. On-the-fly algorithms for bisimulation metrics. In *Proceedings of the 9th International Conference on Quantitative Analysis of Systems*, QEST. to appear, 2012.

[30] Gheorghe Comanici and Doina Precup. Basis function discovery using spectral clustering and bisimulation metrics. In *AAMAS '11*, volume 3, pages 1079–1080, Richland, SC, 2011. International Foundation for Autonomous Agents and Multi-agent Systems.

[31] Vincent Danos, Josee Desharnais, François Laviolette, and Prakash Panangaden. Bisimulation and cocongruence for probabilistic systems. *Inf. Comput.*, 204(4):503–523, 2006.

[32] George B. Dantzig. Application of the Simplex method to a transportation problem. In T.C. Koopmans, editor, *Activity analysis of production and allocation*, pages 359–373. J. Wiley, New York, 1951.

[33] Erik P. de Vink and Jan J. M. M. Rutten. Bisimulation for probabilistic transition systems: A coalgebraic approach. In Pierpaolo Degano, Roberto Gorrieri, and Alberto Marchetti-Spaccamela, editors, *Proc. ICALP*, volume 1256 of *Lecture Notes in Computer Science*, pages 460–470. Springer, 1997.

[34] Erik P. de Vink and Jan J. M. M. Rutten. Bisimulation for probabilistic transition systems: A coalgebraic approach. *Theor. Comput. Sci.*, 221(1-2):271–293, 1999.

[35] O. Demuth. Poznámka k dopravnímu [Czech, with German summary; A remark on the transportation problem]. *Časopis pro Pěstování Matematiky*, 86:103–110, 1961.

[36] Josee Desharnais, Abbas Edalat, and Prakash Panangaden. Bisimulation for labelled Markov processes. *Inf. Comput.*, 179(2):163–193, 2002.

[37] Josee Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Approximating labeled markov processes. In *Proc. LICS*, pages 95–106, 2000.

[38] Josee Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Approximating labelled markov processes. *Inf. Comput.*, 184(1):160–200, 2003.

[39] Josee Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labelled Markov processes. *Theoretical Compututer Science*, 318(3):323–354, 2004.

[40] Josee Desharnais and Prakash Panangaden. Continuous stochastic logic characterizes bisimulation of continuous-time Markov processes. *J. Log. Algebr. Program.*, 56(1-2):99–115, 2003.

[41] Jie Ding and Jane Hillston. Structural analysis for stochastic process algebra models. In Michael Johnson and Dusko Pavlovic, editors, *AMAST*, volume 6486 of *Lecture Notes in Computer Science*, pages 1–27. Springer, 2010.

[42] Ernst-Erich Doberkat. *Stochastic relations: foundations for Markov transition systems*. Chapman & Hall/CRC studies in informatics series. Chapman & Hall/CRC, 2007.

[43] Ernst-Erich Doberkat. *Stochastic Coalgebraic Logic*. Springer, 2010.

[44] R. M. Dudley. *Real analysis and probability*. Cambridge University Press, 2002.

[45] Abbas Edalat. Semi-pullbacks and bisimulation in categories of Markov processes. *Mathematical Structures in Computer Science*, 9(5):523–543, 1999.

[46] Norm Ferns, Prakash Panangaden, and Doina Precup. Metrics for finite Markov decision processes. In *Proceedings of the 20th conference on Uncertainty in Artificial Intelligence*, UAI, pages 162–169. AUAI Press, 2004.

[47] Lester Randolph Ford and Delbert Ray Fulkerson. Solving the transportation problem. *Management Science*, 3(1):24–32, 1956.

[48] Michèle Giry. A categorical approach to probability theory. In Banaschewski, editor, *Categorical Aspects of Topology and Analysis*, volume 915 of *Lecture Notes in Mathematics*, pages 68–85. Springer Berlin/Heidelberg, 1982. 10.1007/BFb0092872.

[49] Marco Grandis and Walter Tholen. Natural weak factorization systems. *Archivum Mathematicum*, 42:397–408, 2006.

[50] David Griffeath. A maximal coupling for markov chains. *Probability Theory and Related Fields*, 31:95–106, 1975.

[51] Jan Friso Groote and Frits W. Vaandrager. Structured operational semantics and bisimulation as a congruence. *Information and Computation*, 100(2):202–260, 1992.

[52] H. Peter Gumm and Tobias Schr oder. Products of coalgebras. *Algebra Universalis*, 46:163–185, 2001.

[53] Paul R. Halmos. *Measure Theory*. Number 18 in Graduate Texts in Mathematics. Springer, 1974.

[54] Holger Hermanns, Ulrich Herzog, and Joost-Pieter Katoen. Process algebra for performance evaluation. *Theor. Comput. Sci.*, 274(1-2):43–87, 2002.

[55] Jane Hillston. Process algebras for quantitative analysis. In *LICS*, pages 239–248. IEEE Computer Society, 2005.

[56] Bart Jacobs and Jan J. M. M. Rutten. A Tutorial on (Co)Algebras and (Co)Induction. *EATCS Bulletin*, 62:62–222, 1997.

[57] Bart Jacobs and Ana Sokolova. Esemplaric expressivity of modal logics. *J. Log. Comput.*, 20(5):1041–1068, 2010.

[58] Yasuo Kawahara and Masao Mori. A small final coalgebra theorem. *Theoretical Computer Science*, 233:129–145, 1998.

[59] Victor Klee and Christoph Witzgall. Facets and vertices of transportation polytopes. In G. B. Dantzig and A. F. Veniott, editors, *Mathematics of the Decision Sciences, Part I*, chapter III, pages 257–282. American Mathematical Society, 1968.

[60] Bartek Klin. Structural operational semantics for weighted transition systems. In Jens Palsberg, editor, *Semantics and Algebraic Specification*, volume 5700 of *Lecture Notes in Computer Science*, pages 121–139. Springer, 2009.

[61] Bartek Klin and Vladimiro Sassone. Structural operational semantics for stochastic process calculi. In Roberto M. Amadio, editor, *FoSSaCS*, volume 4962 of *Lecture Notes in Computer Science*, pages 428–442. Springer, 2008.

[62] Alexander Kurz. *Logics for Coalgebras and Applications to Computer Science*. PhD thesis, Ludwig-Maximilians-Universität München, 2000.

[63] Kim Guldstrand Larsen and Arne Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94(1):1–28, 1991.

[64] Saunders Mac Lane. *Categories for the Working Mathematician*. Springer-Verlag, Berlin, 1971.

[65] Robin Milner. *A Calculus of Communicating Systems*, volume 92 of *Lecture Notes in Computer Science*. Springer, 1980.

[66] Robin Milner. *Communication and Concurrency*. Prentice-Hall, 1989.

[67] Michael Mitzenmacher and Eli Upfal. *Probability and Computing - randomized algorithms and probabilistic analysis*. Cambridge University Press, 2005.

[68] Lawrence S. Moss and Ignacio D. Viglizzo. Final coalgebras for functors on measurable spaces. *Inf. Comput.*, 204(4):610–636, 2006.

[69] Prakash Panangaden. *Labelled Markov Processes*. Imperial College Press, London, U.K., 2009.

[70] David Michael Ritchie Park. Concurrency and automata on infinite sequences. In Peter Deussen, editor, *Theoretical Computer Science*, volume 104 of *Lecture Notes in Computer Science*, pages 167–183. Springer, 1981.

[71] Gordon D. Plotkin. A structural approach to operational semantics. DAIMI FN-19, Computer Science Department, Århus University, Århus, Denmark, 1981.

[72] Jan J. M. M. Rutten. Universal coalgebra: a theory of systems. *TCS*, 249(1):3–80, 2000.

[73] Jan J. M. M. Rutten. Behavioural differential equations: a coinductive calculus of streams, automata, and power series. *Theor. Comput. Sci.*, 308(1-3):1–53, 2003.

[74] Jan J. M. M. Rutten and Daniele Turi. Initial algebra and final coalgebra semantics for concurrency. In J. W. Bakker, W. P. Roever, and G. Rozenberg, editors, *A Decade of Concurrency Reflections and Perspectives*, volume 803 of *Lecture Notes in Computer Science*, pages 530–582. Springer Berlin Heidelberg, 1994.

[75] Alexander Schrijver. *Theory of linear and integer programming.* John Wiley & Sons, Inc., New York, NY, USA, 1986.

[76] Krister Segerberg. *An essay in classical modal logic.* Number volume 1 in Filosofiska studier ut-givna av Filosofiska föreningen och Filosofiska institutionen vid Uppsala universitet. Filosofiska föreningen och Filosofiska institutionen vid Uppsala universitet, 1971.

[77] Michael B. Smyth and Gordon D. Plotkin. The category-theoretic solution of recursive domain equations. *SIAM J. Comput.*, 11(4):761–783, 1982.

[78] Sam Staton. Relating coalgebraic notions of bisimulation. *Logical Methods in Computer Science*, 7(1), 2011.

[79] Terence Tao. *An Introduction to Measure Theory.* Number 126 in Graduate Studies in Mathematics. American Mathematical Society, 2011.

[80] Pedro Sánchez Terraf. Unprovability of the logical characterization of bisimulation. *CoRR*, abs/1005.5142, 2010.

[81] David Thorsley and Eric Klavins. Approximating stochastic biochemical processes with Wasserstein pseudometrics. *IET Systems Biology*, 4(3):193–211, 2010.

[82] Daniele Turi. *Functorial Operational Semantics.* PhD thesis, CWI - University of Amsterdam, 1996.

[83] Daniele Turi and Gordon Plotkin. Towards a mathematical operational semantics. In *Proc. 12$^{\text{th}}$ LICS Conf.*, pages 280–291. IEEE Computer Society Press, 1997.

[84] Johan van Benthem. *Modal Correspondence Theory.* PhD thesis, Universiteit van Amsterdam, Instituut voor Logica en Grondslagenonderzoek van de Exacte Wetenschappen, Amsterdam, 1977.

[85] Johan van Benthem. *Modal logic and classical logic*, volume 3 of *Monographs in philosophical logic and formal linguistics.* Bibliopolis, Napoli, 1983.

[86] Franck van Breugel, Michael W. Mislove, Joël Ouaknine, and James Worrell. Domain theory, testing and simulation for labelled markov processes. *Theor. Comput. Sci.*, 333(1-2):171–197, 2005.

[87] Franck van Breugel, Babita Sharma, and James Worrell. Approximating a Behavioural Pseudometric without Discount for Probabilistic Systems. *Logical Methods in Computer Science*, 4(2):1–23, 2008.

[88] Franck van Breugel and James Worrell. Approximating and computing behavioural distances in probabilistic transition systems. *Theoretical Computer Science*, 360(1-3):373–385, 2006.

[89] Ignacio D. Viglizzo. *Coalgebras on Measurable Spaces.* PhD thesis, Department of Mathematics, IndianaUniversity, 2005.

[90] Ignacio D. Viglizzo. Final sequences and final coalgebras for measurable spaces. In José Luiz Fiadeiro, Neil Harman, Markus Roggenbach, and Jan J. M. M. Rutten, editors, *Proc. CALCO*, volume 3629 of *Lecture Notes in Computer Science*, pages 395–407. Springer, 2005.

[91] James Worrell. Toposes of coalgebras and hidden algebras. *Electronic Notes Theoretical Computer Science*, 11:212–230, 1998.

[92] James Worrell. Terminal sequences for accessible endofunctors. *Electronic Notes in Theoretical Computer Science*, 19:24–38, 1999.

[93] James Worrell. *On Coalgebras and Final Semantics*. PhD thesis, Computer Laborarory, Oxford University, 2000.

[94] James Worrell. On the final sequence of a finitary set functor. *Theoretical Computer Science*, 338(1-3):184–199, 2005.