Proceedings Paper:

Beohar, H. orcid.org/0000-0001-5256-1334 and Küpper, S. (2017) On path-based coalgebras and weak notions of bisimulation. In: Bonchi, F. and König, B., (eds.) 7th Conference on Algebra and Coalgebra in Computer Science (CALCO 2017). 7th Conference on Algebra and Coalgebra in Computer Science (CALCO 2017), 14-16 Jun 2017, Ljubljana, Slovenia. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik , 6:1-6:17. ISBN 9783959770330

https://doi.org/10.4230/LIPIcs.CALCO.2017.6

On path-based coalgebras and weak notions of bisimulation

Harsh Beohar[1] and Sebastian Küpper[1]

1 **Theoretical Computer Science Group**
Universität Duisburg-Essen
{harsh.beohar,sebastian.kuepper}@uni-due.de

Abstract

It is well known that the theory of coalgebras provides an abstract definition of behavioural equivalence that coincides with strong bisimulation across a wide variety of state-based systems. Unfortunately, the theory in the presence of so-called silent actions is not yet fully developed. In this paper, we give a coalgebraic characterisation of branching (delay) bisimulation in the context of labelled transition systems (fully probabilistic systems). It is shown that recording executions (up to a notion of stuttering), rather than the set of successor states, from a state is sufficient to characterise the respected bisimulation relations in both cases.

1998 ACM Subject Classification F.3.2 Semantics of Programming Languages

Keywords and phrases Paths, Executions, Branching bisimulation, Coalgebras

Digital Object Identifier 10.4230/LIPIcs.CALCO.2017.6

1 Introduction

Since its inception, coalgebra-based modelling of systems provides a simple and abstract definition of behavioural equivalence that coincides with the so-called strong bisimulation relations across a wide variety of dynamical systems (see [19] for an introduction). Two states are said to be behaviourally equivalent if they are mapped to a common point by a coalgebra homomorphism. Unfortunately, the theory in the presence of so-called *silent* actions is not yet well developed, albeit some general constructions (with varying level of generality) characterising Milner's weak bisimulation [16] are proposed in the literature (see, for instance, [8, 9, 10, 12, 22] and the references therein).

Another refinement of strong bisimulation is *branching bisimulation* proposed by van Glabbeek and Weijland [25], which is the coarsest equivalence (in the van Glabbeek spectrum [23]) preserving the branching structure of a state [24]. In this context, we are unaware of any prior work that captured branching bisimulation in the framework of coalgebras. Moreover, a natural notion of behavioural equivalence should preserve the branching structure of a state just like strong bisimulation does in the absence of silent action.

Bonchi et al. [7] have considered silent transitions coalgebraically by removing them all together by considering the labels as words rather than single letters. This approach is not useful when characterising branching bisimulation (or even weak bisimulation) because not all silent transitions can always be removed from the system without violating the transfer properties of branching (weak) bisimulation. In [8, 9, 10, 12], weak bisimulation is captured in two phases: first, a given coalgebra is transformed into a coalgebra (possibly over a different base category) which captures the "saturation" effect of a silent action; second, it is shown that the notion of behavioural equivalence on this transformed coalgebra coincides

7th Conference on Algebra and Coalgebra in Computer Science (CALCO 2017).
Editors: Filippo Bonchi and Barbara König; Article No. 6; pp. 6:1–6:16
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

with weak bisimulation on the corresponding dynamical system. In [22] the authors used Aczel-Mendler style formulation of strong bisimulation in the latter step.

Nevertheless, it is well known that the saturation step (i.e. adding strong a-transitions for each weak a-transition to the transition system) is not sound with respect to branching bisimulation even in the case of labelled transition systems [25]; thus, a different approach to characterise it is required. The reason is that τ-steps can enable or disable choice in an observable behaviour, something that is hidden by the saturation step. For weak bisimulation this point is irrelevant; however, branching bisimulation (which is finer than weak bisimulation) requires that τ-steps which are required to answer an observable action may only lead to states that are still in bisimulation relation with the original state. For instance, consider the states x_1, y_1 described in Figure 1. They are *not* branching bisimilar because the transition $y_1 \xrightarrow{a} y_4$ can be simulated by the transitions $x_1 \xrightarrow{\tau} x_2 \xrightarrow{a} x_4$, but the intermediate state x_2 cannot be related with the state y_1 because y_1 can fire a b-transition which x_2 cannot simulate.

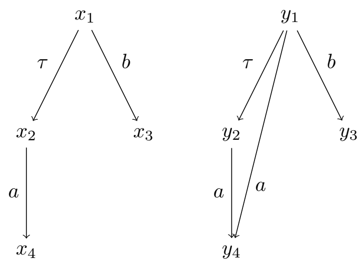


Figure 1 The states x_1, y_1 are weakly bisimilar, but not branching bisimilar.

Our research hypothesis is that recording executions (up to a notion of stuttering) generated by a state (instead of the set of successor states) is sufficient to capture branching bisimulation across different classes of systems. Stated differently, it is the set of executions (not the set of successor states) which specifies the branching structure of a state in the presence of silent actions. In particular, we will substantiate this claim for the class of labelled transition systems and fully probabilistic systems in this paper.

In lieu of the above hypothesis, we restructure the classical coalgebraic machinery in the following way. We begin by studying a notion of *paths* on an arbitrary set X (denoted $\mathrm{Path}(X)$) in Section 2, which is general enough to specify the executions of a labelled transition system and a fully probabilistic system. Intuitively, a path on X can be viewed as a finite sequence that alternates between the elements of X and an action in the alphabet $A_\tau = A \uplus \{\tau\}$, where $\tau \notin A$ is the silent action. Now for every path $p \in \mathrm{Path}(X)$ there is a unique stutter invariant path $p^\dagger \in \mathrm{Path}(X)$ associated with it, which intuitively can be constructed by removing the τ self-loops. This is reminiscent of a coloured trace from [25], which is obtained from a concrete coloured trace in a process graph whose nodes are labelled by a fixed set of colours. In the sequel, stutter invariance induces an equivalence relation $\sim$ on the set $\mathrm{Path}(X)$ whose quotient is denoted as $\mathrm{Path}_\sim(X)$. Furthermore, it turns out that both the mappings $\mathrm{Path}(X), \mathrm{Path}_\sim(X)$ are endofunctors on the category of sets $\mathbf{Set}$.

Now what is missing in our approach is the type of dynamics (also known as the branching type in the theory of coalgebras). For instance, a labelled transition system can be viewed as a coalgebra of type $\mathcal{P} \circ (A_\tau \times \mathrm{id})$ over the base category $\mathbf{Set}$. Here $\mathcal{P}$ is the covariant powerset functor and $A_\tau \times \mathrm{id}$ is the product functor whose left component is fixed. In other words, the branching type of labelled transition system is nondeterministic. Therefore, to characterise branching bisimulation, we consider coalgebras of type $\mathcal{P} \circ \mathrm{Path}_\sim$ over the base category $\mathbf{Set}$. In Section 3, we show that behavioural equivalence in this coalgebra coincides with the traditional branching bisimulation relation [25]. Moreover, this framework can also be used to characterise the weak bisimulation, delay bisimulation, and eta-bisimulation relations; however, for reasons of space, this is worked out in [6].

Nevertheless, the situation is not so straightforward in the case of a fully probabilistic system. Often such systems are modelled as coalgebras of type $\mathcal{D} \circ (A_\tau \times \mathrm{id})$ over the base category **Set**, where $\mathcal{D}$ is the sub-distribution functor. It turns out that one needs a notion of measurable space and a measure on the set of maximal executions[1] in order to define branching (weak) bisimulation relations over the states of a fully probabilistic system (cf. [4, 21]). Thus, it is natural to consider fully probabilistic systems as 'weighted' coalgebras of type $\mathcal{G} \circ (A_\tau \times \mathrm{id})$ over the base category of measurable spaces **Meas**. Here, $\mathcal{G}$ is the well-known Giry monad of probability measures [17].

In Section 5, just like in the discrete case, we consider coalgebras of type $\mathcal{G} \circ \mathrm{Path}_\sim$ over the base category **Meas** to characterise probabilistic delay bismulation, which was mistakenly [20] called probabilistic branching bisimulation in [21, 22]. The crux of the matter is in defining $\mathrm{Path}(X)$ and $\mathrm{Path}_\sim(X)$ as endofunctors on the category **Meas**. In other words, we need to resolve the following issues: first, which subsets of $\mathrm{Path}(X)$ and $\mathrm{Path}_\sim(X)$ are measurable; second, whether $\mathrm{Path}(X) \xrightarrow{\mathrm{Path}(f)} \mathrm{Path}(Y)$ (for a given $X \xrightarrow{f} Y$ in **Set**) is a measurable function or not; third, constructions of measures on the sets $\mathrm{Path}(X)$ and $\mathrm{Path}_\sim(X)$. These issues are explored in Section 4, for which some preliminary knowledge on topology, domain theory, and measure theory is required. In Section 6, we discuss future directions for research and present some concluding remarks. An extended version of this paper containing all the complete proofs pertaining to each section can be found in [6].

2 Preliminaries

This section is devoted to formally introduce a notion of path and stutter equivalent path on a set X, which will be used throughout the paper. As mentioned earlier, a path on X can be intuitively viewed as a finite sequence that alternates between the elements of the set X and an action in the alphabet A_τ. However, we abstain from this operational view in favour of Definition 2.1 to reason about paths from a functional perspective.

Let A_τ^* be the set of finite words with $\varepsilon \in A_\tau^*$ denoting the empty sequence. We write $\preceq$ to denote the prefix ordering on words and let $\downarrow\sigma = \{\sigma' \mid \sigma' \preceq \sigma\}$.

▶ **Definition 2.1.** A *path* p on a set X is a function whose codomain is X and domain is the set of all prefixes of some word in A_τ^*.

Let $\mathrm{Path}(X)$ be the set of all paths on a given set X. Then, this lifts to an endofunctor on the category of sets **Set** by letting: $\mathrm{Path}(f)(p) = f \circ p$, for every function $X \xrightarrow{f} Y$.

Every path $p \in \mathrm{Path}(X)$ has a *trace* associated with it. Moreover, every path $p \in \mathrm{Path}(X)$ reaches a *last* element from the set X. Symbolically, we write

$$\mathrm{trace}(p) = \max \mathrm{dom}(p) \quad \text{and} \quad \mathrm{last}(p) = p(\mathrm{trace}(p)).$$

▶ **Proposition 2.2.** *Let $X \xrightarrow{f} Y$ be a function. Then, for every path $p \in \mathrm{Path}(X)$ we have $\mathrm{trace}(p) = \mathrm{trace}(fp)$ and $f\mathrm{last}(p) = \mathrm{last}(fp)$.*

The above proposition states that the functor Path preserves the trace of a path, which is quite strong for our purpose. To exemplify this, consider two labelled transition systems and a function between the states as shown in Figure 2. Since the τ-step from the state x_1 does not disable

$$\begin{tikzcd}
x_1 \arrow[rr, mapsto, "f"] \arrow[d, "\tau"'] & & y_1 \arrow[d, "a"] \\
x_2 \arrow[urr, mapsto, "f"] \arrow[d, "a"'] & & y_2 \\
x_3 \arrow[urr, mapsto, "f"] & &
\end{tikzcd}$$

Figure 2 Two systems having same branching structure.

[1] An execution of a fully probabilistic system is *maximal* if it is an infinite execution or it stops in a state with the sum of probabilities of all outgoing transitions as 0.

any choice of observable action offered by x_1, we would like to declare the states x_1 and x_2 as equivalent. In other words, we would like to assert that f is a homomorphism between the coalgebras $(X = \{x_1, x_2, x_3\}, \alpha)$ and $(Y = \{y_1, y_2\}, \beta)$, where the functions α, β return all the generated executions. However, we note that this is not the case because the f-image of the execution $p = \langle x_1\ \tau\ x_2\ a\ x_3\rangle \in \alpha(x_1)$ is $\langle y_1\ \tau\ y_1\ a\ y_2\rangle$ which is not an execution from y_1. Thus, the key observation is to relate the executions of two systems up to stuttering, which leads to the following definition.

▶ **Definition 2.3.** A function ϕ with $\mathrm{dom}(\phi) = \mathrm{dom}(p)$ and $\mathrm{cod}(\phi) = A_\tau^*$ is a *stutter basis* for a path $p \in \mathrm{Path}(X)$ if it can be constructed inductively by the following rules:

1. $\phi(\varepsilon) = \varepsilon$.
2. if $\sigma'\tau \in \mathrm{dom}(p)$ and $p(\sigma'\tau) = p(\sigma')$ then $\phi(\sigma'\tau) = \phi(\sigma')$.
3. if $\sigma'\tau \in \mathrm{dom}(p)$ and $p(\sigma'\tau) \neq p(\sigma')$ then $\phi(\sigma'\tau) = \phi(\sigma')\tau$.
4. if $\sigma' a \in \mathrm{dom}(p)$ and $a \in A$ then $\phi(\sigma' a) = \phi(\sigma')a$.

As an example, consider a path $p = \langle x_1\ \tau\ x_1\ a\ x_2\rangle$. Then, the function ϕ defined as $\phi(\tau) = \phi(\varepsilon) = \varepsilon$ and $\phi(\tau a) = a$ is a stutter basis ϕ for the path p. However, if $p = \langle x_1\ \tau\ x_2\ a\ x_3\rangle$ with $x_1 \neq x_2$, then $\phi = \mathrm{id}$ is a stutter basis for p.

▶ **Theorem 2.4.** *For any path there is a unique stutter basis.*

▶ **Lemma 2.5.** *Let ϕ be the stutter basis for a path $p \in \mathrm{Path}(X)$ with $\mathrm{dom}(p) = \downarrow\sigma$, for some $\sigma \in A_\tau^*$. Then, $\phi(\downarrow\sigma) = \downarrow\phi(\sigma)$.*

▶ **Definition 2.6.** Given a path $p \in \mathrm{Path}(X)$ and its corresponding stutter basis ϕ, then a function $\phi(\mathrm{dom}(p)) \xrightarrow{p^\dagger} X$ is the *stutter invariant* path relative to p if $p^\dagger \circ \phi = p$.

Notice that for the function $p^\dagger$ to be a path its domain should be a prefix closed subset of a word, which follows directly from Lemma 2.5.

The notion of stutter invariant path induces an equivalence relation on the set of all paths as follows. Two paths $p, q \in \mathrm{Path}(X)$ are said to be *stutter equivalent*, denoted $p \sim q$, if and only if they have the identical stutter invariant path, i.e., $p^\dagger = q^\dagger$. Let $\mathrm{Path}_\sim(X)$ be the set of all the paths up to stutter equivalence. This lifts to a functor as well:

$$\mathrm{Path}_\sim(f)[p]_\sim = [f \circ p]_\sim \qquad \text{for any } p \in \mathrm{Path}(X) \text{ and } X \xrightarrow{f} Y. \tag{1}$$

To prove that the above map is well-defined, we need the following lemma.

▶ **Lemma 2.7.** *For any $p \in \mathrm{Path}(X)$ and any $X \xrightarrow{f} Y$, we have $f \circ p^\dagger \sim f \circ p$.*

▶ **Theorem 2.8.** *The mapping in* (1) *is well defined and* $\mathrm{Path}_\sim$ *is an endofunctor on* **Set***.*

Notation We write π_X for the quotient map that maps a path $p \in \mathrm{Path}(X)$ to $[p]_\sim$.

We end this subsection with few properties on (stutter equivalent) paths.

▶ **Lemma 2.9.** *Let $p \in \mathrm{Path}(X), q \in \mathrm{Path}(Y)$, and $X \xrightarrow{f} Y$.*

1. *If $fp \sim q$ and $\mathrm{trace}(q) \in \tau^* a \tau^*$ then $q(\varepsilon) = fp(\varepsilon) \wedge \mathrm{trace}(p) \in \tau^* a \tau^*$.*
2. $\mathrm{last}(p) = \mathrm{last}(p^\dagger)$.
3. *If $fp \sim q$, then $f(\mathrm{last}(p)) = \mathrm{last}(q)$.*
4. $\pi_Y \circ \mathrm{Path}(f) = \mathrm{Path}_\sim(f) \circ \pi_X$.
5. *Let $V_i \subseteq \mathrm{Path}_\sim(Y)$ (for $i \in I$) be a family of pairwise disjoint sets. Then,*

$$\mathrm{Path}_\sim(f)^{-1}\Big(\bigcup_{i\in I} V_i\Big) = \bigcup_{i\in I} \mathrm{Path}_\sim(f)^{-1}(V_i) \ .$$

Coalgebraic preliminaries

▶ **Definition 2.10.** Let $\mathbf{C}$ be a category and let $\mathbf{C} \xrightarrow{F} \mathbf{C}$ be an endofunctor. An *F-coalgebra* over the base category $\mathbf{C}$ is a tuple (X, α), where X is an object in $\mathbf{C}$ and $X \xrightarrow{\alpha} FX$ is an arrow in $\mathbf{C}$. Given two objects (X, α) and (Y, β), an F-coalgebra homomorphism is an arrow $X \xrightarrow{f} Y$ in $\mathbf{C}$ such that $Ff \circ \alpha = \beta \circ f$.

▶ **Definition 2.11.** Let $\mathbf{C}$ be a concrete category over the category of sets $\mathbf{Set}$, i.e., there is a faithful functor $\mathbf{C} \xrightarrow{|\_|} \mathbf{Set}$. Let (X, α) be an F-coalgebra over the concrete category $\mathbf{C}$. Then, two points $x, x' \in |X|$ are said to be *F-behaviourally equivalent* if and only if there is an F-coalgebra (Y, β) and an F-coalgebra homomorphism $X \xrightarrow{f} Y$ such that $f(x) = f(x')$.

In Section 3, we will let $\mathbf{C} = \mathbf{Set}$ and $|\_| = \mathrm{id}$; however, in Section 5, we will let $\mathbf{C} = \mathbf{Meas}$ and the faithful functor $|\_|$ to be the forgetful functor which forgets the sigma algebras associated with the underlying sets.

3 Branching bisimulation on labelled transition systems

The goal is to characterise branching bisimulation of van Glabbeek and Weijland [25] using a coalgebraic approach based on paths as outlined in the introduction.

▶ **Definition 3.1.** A *labelled transition system* is a triple $(X, A_\tau, \to)$, where X is a set of states, A_τ a set of actions, and $\to \subseteq X \times A_\tau \times X$ is the so-called *transition relation*.

As usual, we write $x \xrightarrow{a} x'$ and $\twoheadrightarrow \; \subseteq X \times A^* \times X$ to denote an element $(x, a, x') \in \to$ and the weak reachability relation, respectively. The latter is defined as the smallest relation satisfying the following inference rules: $\dfrac{}{x \overset{\varepsilon}{\twoheadrightarrow} x}$ $\quad \dfrac{x \overset{\sigma}{\twoheadrightarrow} x' \quad x' \xrightarrow{a} x''}{x \overset{\sigma a}{\twoheadrightarrow} x''}$.

▶ **Definition 3.2.** Let $(X, A_\tau, \to)$ be a labelled transition system. A symmetric relation $R \subseteq X \times X$ is called a *branching bisimulation* relation [25] if and only if for any $x, y, x' \in X$ and $a \in A_\tau$, if $x \xrightarrow{a} x' \wedge xRy$ then $(x'Ry \wedge a = \tau) \vee \exists_{y', y''}\ y \overset{\varepsilon}{\twoheadrightarrow} y' \xrightarrow{a} y'' \wedge xRy' \wedge x'Ry''$. Two states $x \in X$ and $x' \in X$ are *branching bisimlar* if and only if there exists a branching bisimulation relation R such that xRx'.

Next, we construct a $\mathcal{P} \circ \mathrm{Path}_\sim$-coalgebra based on paths, where $\mathcal{P}$ is the covariant power set endofunctor on the category of sets $\mathbf{Set}$.

An *execution* starting from a state $x \in X$ of a labelled transition system $(X, A_\tau, \to)$ is a path $p \in \mathrm{Path}(X)$ such that $p(\sigma) \xrightarrow{a} p(\sigma a)$, for all $\sigma a \in \mathrm{dom}(p)$. Let $\mathrm{Exec}(x)$ be the set of all executions starting from x. Such a transition system can be modelled as a coalgebra $(X, \pi_X \circ \alpha)$, where transition function α is given as:

$$\alpha(x) = \{p \mid p \in \mathrm{Exec}(x) \wedge \mathrm{trace}(p) \in \tau^* a\} \cup \{p \mid p \in \mathrm{Exec}(x) \wedge \mathrm{trace}(p) \in \tau^*\}.$$

▶ Remark. At this stage, we would like to highlight the distinction between a path and an execution made in this paper. It should be noted that all executions of a system (under investigation) are paths; however, the converse may not be true. This is not unusual because after all the executions of a system are generated on the basis of how behaviour of the system is specified (for instance, by the transition relation in the case of labelled transition systems and by the transition function in the case of fully probabilistic system).

Next, we state the main result of this section.

▶ **Theorem 3.3.** *Let* $(X, A_\tau, \rightarrow)$ *be a labelled transition system and* $(X, \pi_X \circ \alpha)$ *be the corresponding* $\mathcal{P} \circ \mathrm{Path}_\sim$*-coalgebra. Then, two states* $x, x' \in X$ *are branching bisimilar if and only if the states* x, x' *are* $\mathcal{P} \circ \mathrm{Path}_\sim$*-behaviourally equivalent.*

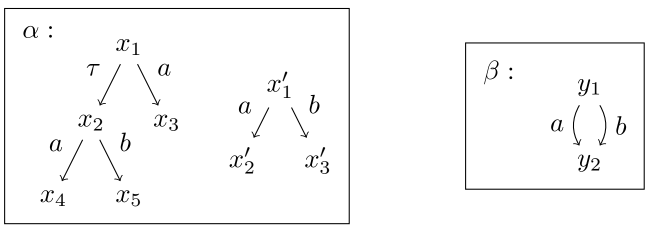
Proof. $\boxed{\Rightarrow}$ Let $R \subseteq X \times X$ be the largest branching bisimulation on the given labelled transition system. Then, from [25] we know that R is an equivalence relation. So let $X \xrightarrow{f} X/R$ be the quotient map. Now to show that f is indeed the required $\mathcal{P} \circ \mathrm{Path}_\sim$-coalgebra homomorphism, we first construct a coalgebra $X/R \xrightarrow{\beta} \mathcal{P}\mathrm{Path}_\sim(X/R)$:

$$\beta(f(x)) = \{\mathrm{Path}_\sim(f)(p) \mid p \in \alpha(x)\}, \qquad \text{for all } x \in X \ .$$

Clearly, β is a total function because f is surjective. Next, we claim that β is well-defined, i.e., independent of the chosen representative. Let $x, x' \in X$ such that $f(x) = f(x')$. Then, we need to show that $\beta(f(x)) = \beta(f(x'))$. Suppose $[fp]_\sim \in \beta(f(x))$ with $p \in \alpha(x)$. Then, by structural induction on the word $\sigma \in \mathrm{dom}(p)$ we show that there is a path $p' \in \alpha(x')$ such that $f \circ (p|_\sigma) \sim f \circ p'$. Here, we write $p|_\sigma$ to denote the restriction of the function p to the sub-domain $\downarrow \sigma$. To see this, without loss of generality, let $\sigma a \in \mathrm{dom}(p)$. Then by the induction hypothesis we find an execution $p' \in \alpha(x')$ such that $f \circ (p|_\sigma) \sim f \circ p'$. Note that $p(\sigma) \xrightarrow{a} p(\sigma a)$ and using Lemma 2.9(3) we get $f \circ (p|_\sigma) \sim f \circ p' \implies p(\sigma)\ R\ \mathrm{last}(p')$. Let $a \in A$. Then, using the transfer property of branching bisimulation we get $\mathrm{last}(p') \overset{\varepsilon}{\twoheadrightarrow} y \xrightarrow{a} y'$ such that $p(\sigma)Ry$ and $p(\sigma a)Ry'$ since $p(\sigma)$ and $\mathrm{last}(p')$ are branching bisimilar. Moreover, from the stuttering lemma [25] we know that any intermediate state visited in the path $\mathrm{last}(p') \overset{\varepsilon}{\twoheadrightarrow} y$ is also R-related to $p(\sigma)$. Therefore, there is a path $p'' \in \mathrm{Path}(X)$ which extends p' such that $f \circ (p|_{\sigma a}) \sim f \circ p''$. In addition, if $a = \tau$ then we either have $p(\sigma\tau)\ R\ \mathrm{last}(p')$ or $\mathrm{last}(p') \overset{\varepsilon}{\twoheadrightarrow} y \xrightarrow{\tau} y'$, for some y, y', with $p(\sigma)Ry$ and $p(\sigma\tau)Ry'$. Suppose the former is true, then clearly we have $f \circ (p|_{\sigma\tau}) \sim f \circ p'$. The latter case is similar to the case when $a \in A$. Thus, for every $p \in \alpha(x)$ there is a path $p' \in \alpha(x')$ such that $f \circ p \sim f \circ p'$. Likewise, we can show the symmetric property when the role of x and x' is interchanged. This completes the proof of the above claim. Clearly, we have $\beta \circ f = \mathcal{P}\mathrm{Path}_\sim(f) \circ \alpha$.

$\boxed{\Leftarrow}$ Let (Y, β) be a $\mathcal{P} \circ \mathrm{Path}_\sim$-coalgebra and $X \xrightarrow{f} Y$ be a $\mathcal{P} \circ \mathrm{Path}_\sim$-coalgebra homomorphism. Below we rather illustrate why the relation $xRx' \iff f(x) = f(x')$ is a witnessing branching bisimulation. The complete proof can be found in [6]. ◀

Consider the two labelled transition systems drawn below enclosed inside the two rectangles. Here, $X \xrightarrow{\alpha} \mathcal{P}\mathrm{Path}_\sim(X)$ and $Y \xrightarrow{\beta} \mathcal{P}\mathrm{Path}_\sim(Y)$ denote the corresponding path-based



coalgebras with $X = \{x_i, x'_j \mid i \in \{1,2,3,4,5\}, j \in \{1,2,3\}\}$ and $Y = \{y_1, y_2\}$. Furthermore, let $X \xrightarrow{f} Y$ be a function defined as $f(x) = y_1$ if $x \in \{x_1, x'_1, x_2\}$; otherwise $f(x) = y_2$. To illustrate why R (as defined above) is a witnessing branching bisimulation, consider the transition $x'_1 \xrightarrow{b} x'_3$ and $x_1 R x'_1$. Clearly, $\langle x'_1\ b\ x'_3 \rangle \in \alpha(x_1)$, which further implies that $\langle y_1\ b\ y_2 \rangle \in \beta(y_1)$. Since $\mathcal{P}\mathrm{Path}_\sim(f) \circ \alpha = \beta \circ f$ we know that there is an execution p such that $f \circ p$ is stutter equivalent to $\langle y_1\ b\ y_2 \rangle$. And by inspection we note that $p = \langle x_1\ \tau\ x_2\ b\ x_5 \rangle$

is such an execution. Moreover, $x_2 R y_1$ and $x_5 R y_2$ which is required by the transfer property of a branching bisimulation relation.

In hindsight, using the terminology of [24], a $\mathcal{P}\mathrm{Path}_{\sim}$-coalgebra homomorphism preserves the branching structure of states. As a consequence, two behaviourally equivalent states have the same set of executions under the image of a $\mathcal{P}\mathrm{Path}_{\sim}$-coalgebra homomorphism up to stutter invariance. For instance, in the above example, the sets of all executions having trace $\tau^* a$ from the states x_1 and x_1' are $\{\langle x_1\ a\ x_3\rangle, \langle x_1\ \tau\ x_2\ a\ x_4\rangle\}$ and $\{\langle x_1'\ a\ x_2'\rangle\}$, respectively. Notice that the f-image of these two sets are equivalent up to stutter invariance. A similar argument can be observed for the set of executions from x_1, x_1' having trace $\tau^* b$.

Though we have focussed on branching bisimulation, this approach can also be used to capture weak, η and delay bisimulation, by defining α differently, saturating τ leading transitions, trailing τ transitions or both, respectively. This is made explicit in [6].

4 A measurable space on paths

As mentioned in the introduction, we will consider coalgebras of type $\mathcal{G} \circ \mathrm{Path}_{\sim}$ over the base category **Meas** to characterise probabilistic delay bisimulation. However, before we do so, we have to fix which subsets of the sets $\mathrm{Path}(X)$ and $\mathrm{Path}_{\sim}(X)$ are measurable together with the construction of a measure on the space of paths, which can be a challenging issue in its own right. In this section, we resolve these fundamental issues by first recalling some basic definitions of measure theory taken from [17].

▶ **Definition 4.1.** A set $\Sigma_X \subseteq \mathcal{P}(X)$ of subsets of X is a *sigma-algebra* on X if and only if $X \in \Sigma_X$ and Σ_X is closed under the set complements and countable unions. Then, the tuple (X, Σ_X) is called a *measurable space*. A *measure* space is a measurable space (X, Σ_X) with a measure $\Sigma_X \xrightarrow{\mu_X} [0, \infty]$, i.e., μ_X is a function satisfying $\mu(\emptyset) = 0$ and the *sigma-additivity* property: for any countable family of pairwise disjoint sets $U_i \in \Sigma_X$ (for $i \in I$) we have

$$\mu_X\big(\bigcup_{i\in I} U_i\big) = \sum_{i\in I} \mu_X(U_i) \ .$$

A *probability* space (X, Σ_X, μ_X) is a measure space with $\mu_X(X) = 1$. A *discrete* space is a measure space such that X is countable and $\Sigma_X = \mathcal{P}(X)$.

Here, the arbitrary sum of a family $\{r_i \mid i \in I\}$ of nonnegative real numbers is defined as $\sum_{i\in I} r_i = \sup\{\sum_{i\in J} r_i \mid J \subseteq_{\mathrm{f}} I\}$ (cf. [21]), where $J \subseteq_{\mathrm{f}} I \iff J \subseteq I \wedge J$ is a finite set.

We want to endow a notion of measurability on the set $\mathrm{Path}_{\sim}(X)$; however, for simplicity we first restrict ourselves to the set of all paths on X, i.e., $\mathrm{Path}(X)$. It turns out that the set of all paths carries a topological structure (precisely, they form what is known as Alexandroff topology [2]) and also satisfies the so-called Kolmogorov separability axiom. Once we have a topological space, the convention is to consider the smallest sigma-algebra generated by the set of all open sets (also known as the Borel sigma-algebra) as the set of measurable sets.

▶ **Definition 4.2.** A *topology* on a set X consists of a set of open sets $\mathcal{O}_X \subseteq \mathcal{P}(X)$ such that: first, the empty set and the whole space are in $\mathcal{O}_X$; second, the set $\mathcal{O}_X$ is closed under finite intersection and arbitrary unions. A topological space $(X, \mathcal{O}_X)$ is an *Alexandroff* space if the set $\mathcal{O}_X$ is closed under arbitrary intersection. A topological space $(X, \mathcal{O}_X)$ satisfies the *Kolmogorov separability axiom* (X is a T_0 space) if any two distinct points are topologically distinguishable, i.e., $\forall_{x,x'\in X}\ x \neq x' \implies \exists_{U\in\mathcal{O}_X}\ (x \in U \wedge x' \notin U) \vee (x \notin U \wedge x' \in U)$.

It is well-known that the set of all upward closed subsets generated by a poset forms a T_0 Alexandroff space. In particular, our set of paths $\mathrm{Path}(X)$ carries the following order:

$$p \preceq q \iff \mathrm{dom}(p) \subseteq \mathrm{dom}(q) \wedge \forall_{\sigma \in \mathrm{dom}(p)}\ q(\sigma) = p(\sigma) \ .$$

Actually, the above ordering is a prefix order in the sense of Cuijpers [11].

▶ **Definition 4.3.** A *prefix order* is a partial order whose every principal ideal is a totally ordered set.

▶ **Proposition 4.4.** *The history of a path* $p \in \mathrm{Path}(X)$ *is downward total, i.e., the set* $\downarrow p = \{p' \mid p' \preceq p\}$ *is a totally ordered set.*

▶ **Proposition 4.5.** *The set of all paths* $\mathrm{Path}(X)$ *on a set* X *forms a* T_0 *Alexandroff space, whose open sets are upward closed subsets of* $\mathrm{Path}(X)$*, i.e.,* $\mathcal{O}_{\mathrm{Path}(X)} = \{U \subseteq \mathrm{Path}(X) \mid U = \uparrow U\}$*. Here, the set* $\uparrow U = \{p' \mid \exists\ p \in U \wedge p \preceq p'\}$ *denotes future of paths in the set* U*.*

At this stage, we note the following relationship between stutter paths and the order $\preceq$.

▶ **Lemma 4.6.** *Let* X *be a set. Then we have the following property: for any two paths* $p_1, p_2 \in \mathrm{Path}(X)$*, if* $p_1^\dagger \preceq p_2^\dagger$ *then* $\exists_{p \in \mathrm{Path}(X)}\ p \sim p_1 \wedge p \preceq p_2$*.*

Every point $x \in X$ in an Alexandroff space has a special neighbourhood associated with it, often called the *smallest neighbourhood* of x, denoted $\mathcal{N}(x) = \bigcap\{U \mid U \in \mathcal{O}_X \wedge x \in U\}$. In particular, this structure, in the case of paths, is the principal filter generated by a path.

▶ **Proposition 4.7.** *For a path* $p \in \mathrm{Path}(X)$*, the smallest neighbourhood of* p *represents the future of the path* p*, i.e.,* $\uparrow p = \mathcal{N}(p)$*. In contrast, the closure* $\mathrm{cl}(p)$ *of a path* $p \in \mathrm{Path}(X)$ *– the smallest closed set that contains* p *– represents the history of* p*, i.e.,* $\mathrm{cl}(p) = \downarrow p$*.*

The next proposition states that the subsets of paths which belong to the Borel sigma-algebra $\mathcal{B}(\mathcal{O}_{\mathrm{Path}(X)})$ are measurable.

▶ **Proposition 4.8.** *The tuple* $(\mathrm{Path}(X), \Sigma_{\mathrm{Path}(X)})$ *is a measurable space, where* $\Sigma_{\mathrm{Path}(X)} = \mathcal{B}(\mathcal{O}_{\mathrm{Path}(X)})$*. Here,* $\mathcal{B}(\mathcal{X})$ *denotes the smallest sigma-algebra generated by* $\mathcal{X} \subseteq \mathcal{P}(X)$*.*

Next, we establish that the $\mathrm{Path}(X) \xrightarrow{\mathrm{Path}(f)} \mathrm{Path}(Y)$ (for a given $X \xrightarrow{f} Y$) is measurable, i.e., if $V \in \Sigma_{\mathrm{Path}(Y)}$ then $f^{-1}V \in \Sigma_{\mathrm{Path}(X)}$. For this, we need the following result.

▶ **Theorem 4.9.** *For any* $X \xrightarrow{f} Y$*, the function* $\mathrm{Path}(f)$ *is an order embedding, i.e., for any* $p, p' \in \mathrm{Path}(X)$ *we have* $p \preceq p' \iff f \circ p \preceq f \circ p'$*.*

Since every order preserving function is continuous and every continuous function is Borel measurable, it follows that, in particular, $\mathrm{Path}(f)$ is Borel measurable.

▶ **Corollary 4.10.** *For any* $X \xrightarrow{f} Y$*, the function* $\mathrm{Path}(f)$ *is measurable.*

In hindsight, the function $\mathrm{Path}(f)$ is an arrow in the category **Meas**. Next, we turn our attention on constructing a measurable space on the set $\mathrm{Path}_\sim(X)$. The idea is to first define an order on the quotient space $\mathrm{Path}_\sim(X)$, which can be inherited from the underlying space of paths $\mathrm{Path}(X)$ by simply letting: $[p]_\sim \preceq [q]_\sim \iff p^\dagger \preceq q^\dagger$, for all $p, q \in \mathrm{Path}(X)$.

▶ **Lemma 4.11.** *The relation* $\preceq$ *on the set* $\mathrm{Path}_\sim(X)$ *is a well-defined partial order. Furthermore, the relation* $\preceq$ *on the set* $\mathrm{Path}_\sim(X)$ *is also a prefix order.*

Once we have an order on the quotient space, we can establish that the quotient maps are order preserving (or continuous in the topological sense).

▶ **Theorem 4.12.** *The quotient function* $\mathrm{Path}(X) \xrightarrow{\pi} \mathrm{Path}_\sim(X)$ *is order preserving. Consequently, the quotient function* $\mathrm{Path}(X) \xrightarrow{\pi} \mathrm{Path}_\sim(X)$ *is Borel measurable, where the sigma-algebra on paths is given by* $\Sigma_{\mathrm{Path}_\sim(X)} = \mathcal{B}(\mathcal{O}_{\mathrm{Path}_\sim(X)})$.

Next, we state the main theorem of this section.

▶ **Theorem 4.13.** *For any* $X \xrightarrow{f} Y$*, the function* $\mathrm{Path}_\sim(f)$ *is order preserving. Thus, the function* $\mathrm{Path}_\sim(f)$ *is Borel measurable.*

Constructing measures on the space of paths

Often, measures on a space are constructed in a top-down manner by identifying a measurable set of building blocks and defining a set-function on this collection (for example, in the case of Lebesgue measures on $\mathbb{R}$, a semi-closed interval $[r, r')$ with $r \leq r'$ is one such building block and the set-function maps every interval of the form $[r, r')$ to the value $r' - r$). In turn, measure extension theorems (for instance, the well-known Carathéodory-Hahn extension theorem; see [18, pp 356]) are invoked to lift the set-function on building blocks to a measure on the whole measurable space. In this paper, we will follow a similar recipe; our building blocks will be open subsets of paths. As for measure extension theorems, we will use a result (cf. Theorem 4.17) established by Alvarez-Manilla [3]. Below, we recall some definitions on a topological space necessary to state this result.

▶ **Definition 4.14.** Let $(X, \mathcal{O}_X)$ be a topological space. A function $\mathcal{O}_X \xrightarrow{\mu} [0, \infty]$ is a *valuation* if and only if the following conditions are satisfied.

1. The function μ is strict, i.e., $\mu(\emptyset) = 0$
2. The function μ is order preserving, i.e., for any two open sets $U, U' \in \mathcal{O}_X$, we have $U \subseteq U'$ implies $\mu(U) \leq \mu(U')$.
3. The function μ is modular, i.e., for any two open sets $U, U' \in \mathcal{O}_X$, we have $\mu(U)+\mu(U') = \mu(U \cup U') + \mu(U \cap U')$.

A valuation μ is *Scott-continuous* if and only if for any directed family of open sets $(U_i)_{i \in I}$ we have $\mu(\bigcup_{i\in I} U_i) = \sup_{i \in I} \mu(U_i)$. Lastly, a valuation μ is *locally finite* if and only if every point has a finitely valued open neighbourhood.

▶ **Definition 4.15.** A space $(X, \mathcal{O}_X)$ is *locally compact* if and only if for every point x and open set U with $x \in U$, there is a compact subset $V \subseteq X$ such that $x \in \mathrm{int}(V)$ and $V \subseteq U$. Here, $\mathrm{int}(V)$ denotes the interior of $V \subseteq X$.

▶ **Definition 4.16.** A topological space $(X, \mathcal{O}_X)$ is *sober* if and only if every irreducible closed set is a closure of a unique point. A closed set C is *irreducible* if and only if C is nonempty and it cannot be expressed as union of two smaller closed subsets, i.e., if $C = C_1 \cup C_2$ and C_1, C_2 are closed sets, then $C = C_1$ or $C = C_2$.

We call a subset $C \subseteq X$ *non-sober* if C is irreducible, C is closed, and it cannot be stated as a closure of point (i.e., $\nexists_{x \in X}\ C = \mathrm{cl}(x)$).

▶ **Theorem 4.17** ([3])**.** *Every locally finite and Scott-continuous valuation on a locally compact sober space extends uniquely to a Borel measure.*

The restrictions on $\mathcal{O}_X \xrightarrow{\mu} [0, \infty]$ imposed by the above theorem are not unreasonable; atleast for our purpose. In Section 5, we will construct a locally finite and a Scott-continuous valuation on open subsets of paths, which is induced by a given fully-probabilistic transition system. Nevertheless, we cannot immediately apply Theorem 4.17 because our space $\mathrm{Path}(X)$ is not a sober space, even though it is locally compact, i.e., every path $p \in \mathrm{Path}(X)$ has a compact neighbourhood (since $p \in \uparrow p$). As a result, in the following, we first 'soberify' our space $\mathrm{Path}(X)$ and use Theorem 4.17 to construct a Borel measure on $\mathrm{Path}(X)$ by lifting a given locally finite and Scott-continuous valuation $\mathcal{O}_{\mathrm{Path}(X)} \xrightarrow{\mu} [0, \infty]$.

▶ Remark. By inspection, we note that our space $\mathrm{Path}(X)$ is non-sober. For instance, if X is non-empty then unfolding a τ-loop results in an infinite chain of paths without any maximum since the domain of a path is a set of prefixes generated by some *finite* word.

Recall that, for a set X, both sets of paths $\mathrm{Path}(X)$ and stutter-equivalent paths $\mathrm{Path}_\sim(X)$ are prefix orders. We want to construct measures on both kinds of spaces, therefore below we work with a class of *simple* prefix orders which generalises both the structures.

▶ **Definition 4.18.** A prefix order is *simple* if the history of every point is a finite set.

For example, the sets $\mathrm{Path}(X)$ and $\mathrm{Path}_\sim(X)$ are simple prefix orders.

▶ **Proposition 4.19.** *A directed subset of a prefix order is always totally ordered. In addition, an irreducible downward closed subset of a prefix order is always totally ordered.*

Next, we construct a space X^∞ consisting of all points from X in which the non-sober sets (w.r.t. Alexandroff topology) are added as limit points.

$$X^\infty = X \cup \{\infty_C \mid C \subseteq X \text{ is a non-sober set w.r.t. Alexandroff topology}\}.$$
$$\preceq' = \preceq \cup \{(\infty_C, \infty_C) \mid \infty_C \in X^\infty\} \cup \{(x, \infty_C) \mid x \in C\}.$$

As an example, consider the prefix order $(\mathbb{N}, \leq)$ with their natural ordering. The sober space $\mathbb{N}^\infty = \mathbb{N} \cup \{\infty_\mathbb{N}\}$ is isomorphic to the well-known set of extended natural numbers $\mathbb{N}_\omega$.

▶ **Lemma 4.20.** *The set X^∞ is prefix ordered by the relation $\preceq'$, if $(X, \preceq)$ is a prefix order.*

Henceforth, we do not distinguish between the relation $\preceq$ and $\preceq'$. Notice that being sober is a topological property and therefore, we need a 'right' notion of topology on X^∞ to qualify it as sober. For instance, if we take upward closed sets as open sets (just like in the case of X) we find that the space X^∞ is still non-sober; as a result, X^∞ is non-sober w.r.t. Alexandroff topology. However, if we endow X^∞ with a Scott topology then the space becomes sober w.r.t. this finer topology. For example, in the case of extended natural numbers, the problematic case of the directed set $\mathbb{N}$ (which was non-sober w.r.t. Alexandroff topology) is actually not a Scott-closed set[2] since $\sup \mathbb{N} = \infty_\mathbb{N}$ and $\infty_\mathbb{N} \notin \mathbb{N}$.

▶ **Proposition 4.21.** *Let $(X, \preceq)$ be a simple prefix order. A subset $U \subseteq X^\infty$ is* Scott *open if and only if U is upward closed and it is inaccessible by directed joins, i.e., for any directed set $D \subseteq X^\infty$ if* $\sup D$ *exists and* $\sup D \in U$ *then $D \cap U \neq \emptyset$. Let $\mathcal{S}_{X^\infty}$ denote the collection of Scott open subsets of X^∞. Then, the space $(X^\infty, \mathcal{S}_{X^\infty})$ is a sober space.*

To apply Theorem 4.17, we need to first construct a locally finite and Scott-continuous valuation on our new sober space $\mathrm{Path}^\infty(X)$. In the following theorem, we will construct one such valuation on $\mathrm{Path}^\infty(X)$ from an old valuation $\mathcal{O}_{\mathrm{Path}(X)} \xrightarrow{\mu} [0, \infty]$.

[2] A subset $C \subseteq X$ of a prefix order $(X, \preceq)$ is Scott closed if and only if C is downward closed and for any directed set $D \subseteq C$, if $\sup D$ exists then $\sup D \in C$.

▶ **Theorem 4.22.** *Let $(X, \preceq)$ be a prefix order and let $\mathcal{O}_X \xrightarrow{\mu} [0,\infty]$ be a locally finite and Scott-continuous valuation. Then, the function $\mathcal{S}_{X^\infty} \xrightarrow{\tilde{\mu}} [0,\infty]$ defined as follows:*

$$\tilde{\mu}(V) = \mu(V \cap X) \quad \textit{(for every Scott-open set } V \in \mathcal{S}_{X^\infty}\textit{)}$$

is a Scott-continuous valuation. If X is simple then $\tilde{\mu}$ is locally finite.

As a result, the function $\tilde{\mu}$ lifts to a unique Borel measure on the sigma-algebra $\Sigma_{\mathrm{Path}^\infty(X)} = \mathcal{B}(\mathcal{S}_{\mathrm{Path}^\infty(X)})$ due to Theorem 4.17. However, in order to reflect back this measure on the original sigma-algebra $\Sigma_{\mathrm{Path}(X)}$, it is sufficient to establish that $\Sigma_{\mathrm{Path}(X)}$ is contained in the Borel sigma-algebra induced by Scott-open sets, i.e., $\Sigma_{\mathrm{Path}(X)} \subseteq \Sigma_{\mathrm{Path}^\infty(X)} = \mathcal{B}(\mathcal{S}_{\mathrm{Path}^\infty(X)})$. The next theorem states under what conditions the set-containment between the two sigma-algebras $\Sigma_{\mathrm{Path}(X)}, \Sigma_{\mathrm{Path}^\infty(X)}$ is possible.

▶ **Theorem 4.23.** *For countable sets A and X, the sigma-algebra $\Sigma_{\mathrm{Path}(X)}$ is contained in the Borel sigma-algebra $\Sigma_{\mathrm{Path}^\infty(X)}$. Moreover, we also have $\Sigma_{\mathrm{Path}_\sim(X)} \subseteq \Sigma_{\mathrm{Path}^\infty_\sim(X)}$.*

▶ **Corollary 4.24.** *Suppose the sets A and X are countable. Then every locally finite and Scott-continuous valuations $\mathcal{O}_{\mathrm{Path}(X)} \xrightarrow{\mu} [0,\infty]$ and $\mathcal{O}_{\mathrm{Path}_\sim(X)} \xrightarrow{\mu} [0,\infty]$ lifts to a unique Borel measure $\Sigma_{\mathrm{Path}(X)} \xrightarrow{\tilde{\mu}} [0,\infty]$ and $\Sigma_{\mathrm{Path}_\sim(X)} \xrightarrow{\tilde{\mu}} [0,\infty]$, respectively.*

5 Probabilistic delay bisimulation

In this section, we use the concepts developed in the previous section to characterise probabilistic delay bisimulation relation between the states of a fully probabilistic system.

▶ **Definition 5.1.** A *(fully) probabilistic transition system* is a triple (X, A_τ, P) consisting of a countable set of states X, a countable set of actions A_τ, and a *probability transition function* $X \times A_\tau \times X \xrightarrow{P} [0,1]$ such that for every $x \in X$, the set $\{(a, x') \mid 0 < P(x, a, x')\}$ is finite and $\sum_{(a,x') \in A_\tau \times X} P(x, a, x') \in \{0, 1\}$.

Given a probabilistic transition system (X, A_τ, P), an *execution* p is a path on X such that $\forall_{\sigma a \in \mathrm{dom}(p)}\ 0 < P(p(\sigma), a, p(\sigma a))$. Let $\mathrm{Exec}(x)$ be the set of all executions starting from the state x. We write $\hat{a} = \varepsilon$ if $a = \tau$ and $\hat{a} = a$ if $a \in A$.

▶ **Definition 5.2.** An equivalence relation $R \subseteq X \times X$ on a probabilistic transition system (X, A_τ, P) is a *probabilistic delay bisimulation* [4, 21] if and only if

$$\forall_{x,x' \in X}\ xRx' \implies \forall_{x'' \in X, a \in A_\tau}\ \ P(x, \tau^*\hat{a}, [x'']_R) = P(x', \tau^*\hat{a}, [x'']_R)\ .$$

Two states $x, x' \in X$ are *probabilistic delay bisimilar* if and only if there is a probabilistic delay bisimulation R such that xRx'.

Here, the probabilities associated with weak transitions are defined (originally given in [22]) in the following way. For $x \in X, Y \subseteq X, L \subseteq A_\tau^*$, we let

- $x \xrightarrow{L} Y = \{p \in \mathrm{Exec}(x) \mid \mathrm{trace}(p) \in L \wedge \mathrm{last}(p) \in Y \wedge$
  $\forall_q\ (q \prec p \wedge \mathrm{trace}(q) \in L) \implies \mathrm{last}(q) \notin Y\}$.
- $P(x, L, Y) = \sum_{p \in x \xrightarrow{L} Y} \mu_P(p)$, where $\mathrm{Path}(X) \xrightarrow{\mu_P} [0,1]$ is defined as:

$$\mu_P(p) = \begin{cases} \prod_{\sigma a \in \mathrm{dom}(p)} P(p(\sigma), a, p(\sigma a)), & \text{if } p \in \mathrm{Exec}(x),\ \text{for some } x \in X \\ 0, & \text{otherwise} \end{cases}.$$

▶ **Proposition 5.3.** *For a given fully probabilistic system (X, A_τ, P), the induced function μ_P on paths is order reversing. Moreover, $\mu_P(\varepsilon_x) = 1$ (for any $x \in X$).*

In contrast to Section 3, our base category will be rather the category of measurable spaces and measurable functions **Meas**.

▶ **Definition 5.4.** Below we recall the well known Giry functor $\mathcal{G}$ (see e.g. [17]):

- Let (X, Σ_X) be a measurable space. Then, $\mathcal{G}(X, \Sigma_X) = (\mathcal{G}X, \Sigma_{\mathcal{G}X})$, where $\mathcal{G}X$ is the set of all probability measures on the measurable space (X, Σ_X). The sigma-algebra $\Sigma_{\mathcal{G}X}$ is the smallest sigma-algebra such that the evaluation maps $\mathcal{G}X \xrightarrow{\epsilon_U} [0,1]$ are Borel measurable, for every $U \in \Sigma_X$.
- For any arrow $X \xrightarrow{f} Y$ in **Meas**, we let $\mathcal{G}(f)(\mu) = \mu \circ f^{-1}$.

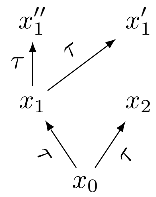


Figure 3 An example motivating separation closure.

To motivate the next definition, consider the transition system depicted in Figure 3 and assume a nonzero probability with each of the drawn transitions. Furthermore, let p_1, p_1', and p_2 be the three executions that reach the states x_1, x_1', and x_2, resp., from the state x_0. Notice that $P(x_0, \tau^*, \{x_1, x_1', x_2\})$ is the sum of the probabilities associated only with the executions p_1, p_2. The execution p_1' is not considered in the above computation because one can reach the set of target states $\{x_1, x_1', x_2\}$ with the execution p_1 which is a prefix of p_1'. This means, the execution p_1' is redundant and neglected while computing the probability to reach the above target states. Such redundancies at the level of paths are identified by the following notion of separation closure.

▶ **Definition 5.5.** Let $(X, \mathcal{O}_X)$ be an Alexandroff space. The *separation closure* of a subset $U \subseteq X$ is the set $U^\star = \{x \in U \mid \mathrm{cl}(x) \cap U = \{x\}\}$. A subset $U \subseteq X$ is *separated* if $U = U^\star$.

In the context of the previous example (Figure 3), let $U = \{p_1, p_1', p_2\}$. Then, we find that $\downarrow x \cap U = \{x\}$, for $x \in \{p_1, p_2\}$, while for the execution p_1' we find that $\downarrow p_1' \cap U = \{p_1, p_2\}$. Thus, $U^\star = \{p_1, p_2\}$ which were the only executions needed to compute the probability to reach one of the target states. Incidentally, a separated subset of paths $U \subseteq \mathrm{Path}(X)$ (i.e., $U = U^\star$) is minimal in the sense that any two distinct paths $p, p' \in U$ are not in the prefix relation $\preceq$, i.e., $p \not\preceq p'$ and $p' \not\preceq p$. The following proposition asserts this.

▶ **Proposition 5.6.** *In an Alexandroff space $(X, \mathcal{O}_X)$, a separated subset $U \subseteq X$ (i.e., $U = U^\star$) has topologically distinguishable points. Moreover,*

1. *the separation closure of a set is always separated, i.e., $U^\star = U^{\star\star}$, for any $U \subseteq X$.*
2. *the collection of separated sets is hereditary, i.e., if $U_1 \subseteq U_2$ and $U_2 = U_2^\star$, then $U_1 = U_1^\star$.*
3. *for any subset $U \subseteq X$, we have $(\uparrow U)^\star \subseteq U^\star$. Moreover the converse also holds, if the underlying space X is a T_0 space. Here, upward closure is w.r.t. the specialisation order $\preceq$, i.e., $x \preceq x' \iff \mathrm{cl}(x) \subseteq \mathrm{cl}(x')$, for any $x, x' \in X$.*

Thus, separation closure provides an alternative way to compute $P(x, L, Y)$.

▶ **Lemma 5.7.** *For a given system (X, A_τ, P), define the set $x \overset{L}{\twoheadrightarrow} Y = \{p \in \mathrm{Path}(X) \mid p(\varepsilon) = x \wedge \mathrm{trace}(p) \in L \wedge \mathrm{last}(p) \in Y\}$. Then, $\sum_{p \in x \overset{L}{\rightarrow} Y} \mu_P(p) = \sum_{p \in (x \overset{L}{\twoheadrightarrow} Y)^\star} \mu_P(p)$.*

It should be noted that for a given probabilistic transition system (X, A_τ, P), we have $x \overset{L}{\rightarrow} Y \subseteq (x \overset{L}{\twoheadrightarrow} Y)^\star$; however, the converse is not true in general.

The next theorem highlights a property characteristic to fully probabilistic systems. It states that if a separated subset of paths U has a lower bound p, i.e., $\forall_{q \in U}\ p \preceq q$, then the sum of probabilities associated with each path in U is bounded by the weight of p. This property is due to the order reversing nature of the function μ_P (cf. Proposition 5.3).

▶ **Theorem 5.8.** *Given a system (X, A_τ, P), a path $p \in \mathrm{Path}(X)$, and a separated set of paths $U \subseteq \mathrm{Path}(X)$ such that $p \preceq U$, i.e., $\forall_{q\in U}\ p \preceq q$. Then, $\sum_{q\in U} \mu_P(q) \leq \mu_P(p)$.*

Next we focus on the construction of probability measures on the space $\mathrm{Path}(X)$. From Corollary 4.24, it suffices to construct a locally-finite and Scott-continuous valuation on the open subsets of paths. The following theorem extends the function μ_P (induced by a given fully probabilistic system (X, A_τ, P)) to such a valuation on paths.

▶ **Theorem 5.9.** *Given a system (X, A_τ, P), then the function $\mathcal{O}_{\mathrm{Path}(X)} \xrightarrow{\tilde{\mu}_P} [0, \infty]$ defined as $\tilde{\mu}_P(U) = \mu_P(U^\star)$ (for every open set U) is a locally finite and Scott-continuous valuation.*

Now we have all the technical machinery to encode a given probabilistic transition system (X, A_τ, P) as a coalgebra $X \xrightarrow{\alpha} \mathcal{G}\mathrm{Path}_\sim(X)$, where $\Sigma_X = \mathcal{P}(X)$ (since X is countable). The transition system α is defined, coalgebraically, as follows:

$$\alpha(x)(U) = \tilde{\mu}_P({\pi_X}^{-1}(U) \cap \uparrow\varepsilon_x), \qquad \text{for every } U \in \Sigma_{\mathrm{Path}_\sim(X)} \ . \tag{2}$$

Here, we abuse notation by using $\tilde{\mu}_P$ to denote a measure on $\Sigma_{\mathrm{Path}(X)}$. Note that this measure is rather constructed by extending the valuation given in Theorem 5.9.

▶ **Proposition 5.10.** *The mapping in (2) is a probability measure.*

Now we are ready to state the main result of this section.

▶ **Theorem 5.11.** *Two states are probabilistic delay bisimilar if and only if they are $\mathcal{G} \circ \mathrm{Path}_\sim$-behaviourally equivalent.*

Proof. $\boxed{\Rightarrow}$ Let R be a probabilistic delay bisimulation, let (X, α) be the $\mathcal{G} \circ \mathrm{Path}_\sim$-coalgebra induced by (X, A_τ, P), and let $X \xrightarrow{f} X/R$ be the quotient map. We will construct a coalgebra on the quotient set X/R in two stages. First, we construct a measure $\nu_{f(x)}$ (for each $x \in X$) on the space $\mathrm{Path}_\sim(X/R)$ using the extension result (cf. Corollary 4.24) such that it coincides with the pushforward measure $(\alpha(x))_*$ on the open subsets $V \in \mathcal{O}_{\mathrm{Path}_\sim(X/R)}$. Second, we invoke the well-known application (taken from [17, Proposition 2.10]) of Dynkin's $\lambda - \pi$ theorem to conclude that $\nu_{f(x)} = (\alpha(x))_*$.

Let $x \in X$. Define a function $\mathcal{O}_{\mathrm{Path}_\sim(X/R)} \xrightarrow{\nu_{f(x)}} [0, \infty]$ as follows:

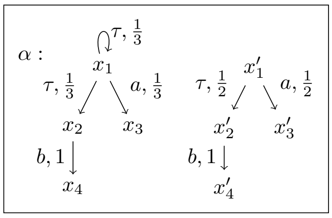
$$\nu_{f(x)}(V) = \tilde{\mu}_P(\pi^{-1}\mathrm{Path}_\sim(f)^{-1}V \cap \uparrow\varepsilon_x), \qquad \text{for every } V \in \mathcal{O}_{\mathrm{Path}_\sim(X/R)}. \tag{3}$$

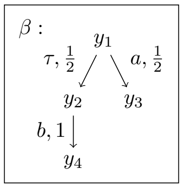
We need a technical result proven in [21, Lemma 24] in order to show that $\nu_{f(x)}$ is well defined, i.e., for any xRx' and open set $V \in \mathcal{O}_{\mathrm{Path}_\sim(X/R)}$ we have $\nu_{f(x)}(V) = \nu_{f(x')}(V)$. (See [6] for the proof of this claim). Moreover, the function $\nu_{f(x)}$ is a valuation, which immediately follows from (3) and the fact that μ is a valuation. Therefore, from Corollary 4.24, the valuation $\nu_{f(x)}$ extends to a Borel measure $\tilde{\nu}_{f(x)}$ on the space $\mathrm{Path}_\sim(X/R)$.

Recall that the pushforward measure $(\alpha(x))_*(V) = \alpha(x)(\mathrm{Path}_\sim(f)^{-1}(V))$ (for each Borel set $V \in \Sigma_{\mathrm{Path}_\sim(X/R)}$), is also a measure on the space $\mathrm{Path}_\sim(X/R)$. Clearly, for any $V \in \mathcal{O}_{\mathrm{Path}_\sim(X/R)}$, we have $(\alpha(x))_*(V) = \nu_{f(x)}(V) = \tilde{\nu}_{f(x)}(V)$ due to Equation (2). Since both $\tilde{\nu}_{f(x)}$ and $(\alpha(x))_*$ are probability measures, so from [17, Proposition 2.10] we get $\tilde{\nu}_{f(x)} = (\alpha(x))_*$. Now letting $\beta(f(x)) = \tilde{\nu}_{f(x)}$, we find that f is a coalgebra homomorphism because $\beta(f(x))(V) = \tilde{\nu}_{f(x)}(V) = (\alpha(x))_*(V) = \alpha(x)(\mathrm{Path}_\sim(f)^{-1}V)$, for every $V \in \Sigma_{\mathrm{Path}_\sim(X/R)}$.

$\boxed{\Leftarrow}$ Let (X, A_τ, P) be a fully probabilistic system and (X, α) be the corresponding $\mathcal{G} \circ \mathrm{Path}_\sim$-coalgebra. Moreover, let (Y, β) be a $\mathcal{G} \circ \mathrm{Path}_\sim$-coalgebra and $X \xrightarrow{f} Y$ be a $\mathcal{G} \circ \mathrm{Path}_\sim$ coalgebra homomorphism. In [6], we show that the equivalence relation $xRx' \iff f(x) = f(x')$ is a probabilistic delay bisimulation. Below we rather illustrate why R is a witnessing probabilistic delay bisimulation relation. ◀

Consider the two probabilistic transition systems drawn below,





together with the path-based coalgebras $X \xrightarrow{\alpha} \mathcal{G}\mathrm{Path}_\sim(X)$ and $Y \xrightarrow{\beta} \mathcal{G}\mathrm{Path}_\sim(Y)$ where $X = \{x_i, x'_j \mid i, j \in \{1,2,3,4\}\}$ and $Y = \{y_i \mid i \in \{1,2,3,4\}\}$. Furthermore, let $X \xrightarrow{f} Y$ be a function defined as $f(z_i) = y_i$, where $z \in \{x, x'\}$ and $i \in \{1,2,3,4\}$. To see why the relation R (as defined above) is a witnessing bisimulation, consider the equation

$$\alpha(x_1)\Big(\bigcup_{p \in x_1 \xrightarrow{\tau^* b} [x_4]_R} \uparrow [p]_\sim\Big) = \alpha(x'_1)\Big(\bigcup_{p \in x'_1 \xrightarrow{\tau^* b} [x_4]_R} \uparrow [p]_\sim\Big), \tag{4}$$

which can be derived from the facts $x_1 R x'_1$ and $\beta \circ f = \mathcal{G}\mathrm{Path}_\sim(f) \circ \alpha$ (see [6] for the proof in the general case). The two terms in Equation 4 denote the probabilities of reaching the equivalence class $[x_4]_R$ from the states x_1 and x'_1. This can be seen, for instance, by deriving $\alpha(x_1)(\bigcup_{p \in x_1 \xrightarrow{\tau^* b} [x_4]_R} \uparrow [p]_\sim) = P(x_1, \tau^* b, [x_4]_R)$ using Equation 2, definition of $\tilde{\mu}_P$, Proposition 3, and Lemma 5.7. Moreover, the probability to reach the equivalence class $[x_4]_R$ from x_1, x'_1 is $\frac{1}{2}$ because $P(x_1, \tau^* b, [x_4]_R) = \sum_{i=1}^{\infty} (\frac{1}{3})^i = \frac{1}{2} = P(x'_1, \tau^* b, [x_4]_R)$.

6 Discussion and conclusion

The main message of this paper is that behavioural equivalence in a path-based coalgebra is sufficient to capture branching bisimulation. In particular, we considered coalgebras of type $F \circ \mathrm{Path}_\sim$ over a concrete category $\mathbf{C}$, where F is an endofunctor modelling the branching type of the system under investigation. We showed that behavioural equivalence when $F = \mathcal{P}$ and $\mathbf{C} = \mathbf{Set}$ coincides with the traditional branching bisimulation [25]. In a similar spirit, we also showed that behavioural equivalence when $F = \mathcal{G}$ and $\mathbf{C} = \mathbf{Meas}$ coincides with the probabilistic delay bisimulation [4, 21, 22].

Interestingly, in the case of labelled transition systems, we can use the final chain based algorithm presented in [1] to minimise the system with respect to branching bisimulation. The following prerequisites for this algorithm are satisfied in this context: first, a terminal object exists in **Set**; second, **Set** is equipped with a (epi,mono)-factorisation structure; third, the functor $\mathcal{P} \circ \mathrm{Path}_\sim$ preserves monomorphisms. However, for the probabilistic case, more research is required to find out whether the above conditions are valid or not.

In retrospect, our paper comes short in one regard when comparing with the recent works [8, 10, 12] on capturing weak bisimulation; namely, there is no abstract construction given to construct our path-based coalgebras from the system under study. In particular, we would like to construct a path-based coalgebra, for instance, $X \xrightarrow{\alpha'} F\mathrm{Path}_{\sim}(X)$ from a given coalgebra of type $X \xrightarrow{\alpha} F(A_\tau \times X)$.

In this regard, it might be interesting to extend the initial work of Jacobs and Sokolova [13]: Given a system $X \longrightarrow TFX$ over **Set** (T is a monad modelling the branching type and F is an endofunctor modelling the transition type), then the traces (executions) can be described as an arrow $X \longrightarrow I$ ($X \times I \longrightarrow I$) in the Kleisli category of T with I being the initial algebra of F. Note that this insight of [13] works under some technical requirements and it is unclear whether these requirements hold in a more general setting of **Meas**. This was already voiced by Kerstan and König [15] in conjunction with generic trace semantics for probabilistic systems that were modelled over the base category **Meas**.

Another way to generalise the result of this paper is to consider the executions of a system as first-class citizens from the onset. Such a venture is carried out by Cuijpers [11] under the banner of prefix orders. Prefix orders are partially ordered sets whose principal ideals are totally ordered sets. The homomorphisms on such structures are called history preserving functions, those order preserving functions that preserve the principal ideals of the underlying ordered sets. Beohar and Cuijpers [5] extended the theory of open maps [14] to the concrete category setting to get a characterisation of traditional branching bisimulation. Therefore, it will be worthwhile to study whether the measure theoretic concepts proposed here can be lifted to the more general setting of prefix orders to capture probabilistic branching bisimulation. Lastly, it will be interesting to construe a notion of behavioural equivalence in the open map approach akin to the theory of coalgebras, where the notion of bisimulation is parametric to a functor modelling the branching type of system under study.

Acknowledgements

The authors thank the anonymous reviewers of the conference CALCO'17 for their feedbacks on an earlier draft of this paper. The authors also thank Barbara König for various discussions regarding this work and for her earlier comments which led to significant improvements of this paper. The authors would also like to thank Pieter Cuijpers for his various feedbacks and continuing support over the course of this work.

This work has been carried out as part of the "Behavioural Equivalences: Environmental Aspects, Metrics and Generic Algorithms" (BEMEGA) project supported by Deutsche Forschungsgemeinschaft (DFG).

References

[1] J. Adámek, F. Bonchi, M. Hülsbusch, B. König, S. Milius and A. Silva. 'A Coalgebraic Perspective on Minimization and Determinization'. In: *Proc. of FOSSACS '12*. LNCS/ARCoSS 7213. Springer, 2012, pp. 58–73.

[2] P. Alexandroff. 'Diskrete Räume'. In: *Rec. Math. N.S.* 2(44).3 (1937), pp. 501–519.

[3] M. Alvarez-Manilla. 'Extension of valuations on locally compact sober spaces'. In: *Topology and its Applications* 124.3 (2002), pp. 397 –433.

[4] C. Baier and H. Hermanns. 'Weak Bisimulation for Fully Probabilistic Processes'. In: *Proc. of CAV*. London, UK: Springer, 1997, pp. 119–130.

[5] H. Beohar and P.J.L. Cuijpers. 'Open Maps in Concrete Categories and Branching Bisimulation for Prefix Orders'. In: *ENTCS* 319 (2015), pp. 51 –66.

[6] H. Beohar and S. Küpper. 'On path-based coalgebras and weak notions of bisimulation'. In: *ArXiv e-prints* (May 2017). eprint: `1705.08715`. URL: `https://arxiv.org/abs/1705.08715`.

[7] F. Bonchi, S. Milius, A. Silva and F. Zanasi. 'Killing epsilons with a dagger: A coalgebraic study of systems with algebraic label structure'. In: *TCS* 604 (2015). Coalgebraic Methods in Computer Science, pp. 102 –126.

[8] T. Brengos. 'On Coalgebras with Internal Moves'. In: *Proc. of CMCS*. Ed. by M.M. Bonsangue. Springer, 2014, pp. 75–97.

[9] T. Brengos. 'Weak bisimulation for coalgebras over order enriched monads'. In: *Logical Methods in Computer Science* 11.2 (2015).

[10] T. Brengos, M. Miculan and M. Peressotti. 'Behavioural equivalences for coalgebras with unobservable moves'. In: *JLAP* 84.6 (2015), pp. 826 –852.

[11] P.J.L. Cuijpers. 'Prefix orders as a general model of dynamics'. In: *Proc. of Developments in Computation Models*. Ed. by I. Mackie, M. Ayala-Rincón and E. Bonelli. DCM'13. Buenos Aires, Argentina, 2013, pp. 25–29.

[12] S. Goncharov and D. Pattinson. 'Coalgebraic Weak Bisimulation from Recursive Equations over Monads'. In: *Proc. of ICALP, Part II*. Ed. by J. Esparza, P. Fraigniaud, T. Husfeldt and E. Koutsoupias. Springer, 2014, pp. 196–207.

[13] B. Jacobs and A. Sokolova. 'Traces, Executions and Schedulers, Coalgebraically'. In: *CALCO'09*. Ed. by A. Kurz, M. Lenisa and A. Tarlecki. Springer, pp. 206–220.

[14] A. Joyal, M. Nielson and G. Winskel. 'Bisimulation and open maps'. In: *Proc. of 8th Annual IEEE Symposium on Logic in Computer Science*. 1993, pp. 418–427.

[15] H. Kerstan and B. König. 'Coalgebraic Trace Semantics for Continuous Probabilistic Transition Systems'. In: *Logical Methods in Computer Science* 9.4 (2013).

[16] R. Milner. *A Calculus of Communicating Systems*. Vol. 92. Lecture Notes in Computer Science. Berlin Heidelberg: Springer, 1980.

[17] P. Panangaden. *Labelled Markov Processes*. London: Imperial College Press, 2009.

[18] H. L. Royden and P. M. Fitzpatrick. *Real analysis*. 4^{th} edition. Pearson, 2010.

[19] J.J.M.M. Rutten. 'Universal coalgebra: a theory of systems'. In: *TCS* 249.1 (2000), pp. 3 –80.

[20] A. Sokolova. 'Personal communication'. 2016.

[21] A. Sokolova, E. de Vink and H. Woracek. 'Coalgebraic Weak Bisimulation for Action-Type Systems'. In: *SACS* 19 (2009), pp. 93–144.

[22] A. Sokolova, E. de Vink and H. Woracek. 'Weak Bisimulation for Action-Type Coalgebras'. In: *ENTCS* 122 (2005), pp. 211 –228.

[23] R.J. van Glabbeek. 'The linear time — Branching time spectrum II'. In: *Proc. of CONCUR*. Ed. by E. Best. Springer, 1993, pp. 66–81.

[24] R.J. van Glabbeek. 'What is Branching Time Semantics and Why to Use It?' In: *Current Trends in Theoretical Computer Science*. Ed. by B. Păun, G. Rozenberg and A. Salomaa. River Edge, USA: World Scientific Publishing, 2001, pp. 469–479.

[25] R.J. van Glabbeek and W.P. Weijland. 'Branching Time and Abstraction in Bisimulation Semantics'. In: *J. ACM* 43.3 (May 1996), pp. 555–600.