Basic Research in Computer Science

Bisimulation for Labelled Markov Processes

Richard Blute
Josée Desharnais
Abbas Edalat
Prakash Panangaden

BRICS Report Series **RS-97-4**

ISSN 0909-0878 **March 1997**

Copyright © 1997, BRICS, Department of Computer Science University of Aarhus. All rights reserved.

Reproduction of all or part of this work is permitted for educational or research use on condition that this copyright notice is included in any copy.

See back inner page for a list of recent BRICS Report Series publications. Copies may be obtained by contacting:

BRICS
Department of Computer Science
University of Aarhus
Ny Munkegade, building 540
DK–8000 Aarhus C
Denmark
Telephone: +45 8942 3360
Telefax: +45 8942 3255
Internet: BRICS@brics.dk

BRICS publications are in general accessible through the World Wide Web and anonymous FTP through these URLs:

`http://www.brics.dk`
`ftp://ftp.brics.dk`
This document in subdirectory `RS/97/4/`

Bisimulation for Labelled Markov Processes

Richard Blute*
Department of Mathematics
University of Ottawa
Ottawa, Ontario, Canada
rblute@mathstat.uottawa.ca

Josée Desharnais*
School of Computer Science
McGill University
Montreal, Quebec, Canada
desharna@cs.mcgill.ca

Abbas Edalat†
Department of Computing
Imperial College
London, UK
ae@doc.ic.ac.uk

Prakash Panangaden*‡
BRICS
Aarhus University
Aarhus, Denmark
prakash@cs.mcgill.ca

Full version of the paper presented at LICS 97

Abstract

In this paper we introduce a new class of labelled transition systems - Labelled Markov Processes - and define bisimulation for them. Labelled Markov processes are probabilistic labelled transition systems where the state space is not necessarily discrete, it could be the reals, for example. We assume that it is a Polish space (the underlying topological space for a complete separable metric space). The mathematical theory of such systems is completely new from the point of view of the extant literature on probabilistic process algebra; of course, it uses classical ideas from measure theory and Markov process theory. The notion of bisimulation builds on the ideas of Larsen and Skou and of Joyal, Nielsen and Winskel. The main result that we prove is that a notion of bisimulation for Markov processes on Polish spaces, which extends the Larsen-Skou definition for discrete systems, is indeed an equivalence relation. This turns out to be a rather hard mathematical result which, as far as we know, embodies a new result in pure probability theory. This work heavily uses continuous mathematics which is becoming an important part of work on hybrid systems.

*Research supported in part by NSERC.
†Research supported by EPSRC, UK.
‡On Leave from McGill University.

1 Introduction

Computer Science has been increasingly expanding its borders to include subjects normally considered part of physics, dynamical systems or control theory, most notably in areas like "hybrid systems". One can look at volume 1066 of Springer-Verlag Lecture Notes In Computer Science [AHS96] for examples of recent work on Hybrid systems. Ideas from continuous mathematics: differential equations, probability theory, stochastic processes, systems theory and others are becoming part of the mathematical toolkit of concurrency theorists. In the area of performance evaluation of computer systems the use of stochastic Petri nets is quite widespread, one of the early papers is [Sif77]. Hillston has pioneered the use of process algebra in performance evaluation [Hil94]. Systems like Hytech [HHWT95] and the Uppaal system [BLL+96] have appeared based on hybrid systems and real-time systems respectively and use process equivalences very fruitfully.

A new ingredient offered by computer science (apart from the manifest idea of computability or effectiveness) to these subjects is *compositionality*. Thus while the theory of stochastic processes [CM65] is concerned with a detailed analysis of the time evolution of systems behaving according to probabilistic laws very little is ever done to analyze the behaviour of coupled systems in a systematic way. Computer scientists have stressed compositionality as a way to attack the formidable intricacy of the systems they have dealt with. For example, in Hillston's work, compositionality is the key contribution of her approach to performance modelling.

The notion of bisimulation is central to the study of concurrent systems. While there are a bewildering variety of different equivalence relations between processes (two-way simulation, trace equivalence, failures equivalence and many more) bisimulation enjoys some fundamental mathematical properties, most notably its characterization as a fixed-point, which make it the most discussed process equivalence. Of course there are a lot of different variants of bisimulation itself! In the present paper we are not so much concerned with adjudicating between the rival claims of all these relations, but rather, we are concerned with showing how to extend these ideas to the world of continuous state spaces. As we shall see below, new mathematical techniques (from the point of view of extant work in process algebra) have to be incorporated to do this. Once the model and the new mathematical ideas have been assimilated, the whole gamut of process equivalences can be studied and argued about.

From an immediate practical point of view, bisimulation can be used to reason about probabilistic, continuous state-space systems (henceforth Markov processes) in the following simple way. One often "discretizes" a continuous system by partitioning the state space into a few equivalence classes. Usually one has some intuition that the resulting discrete system "behaves like" the original continuous system. This can be made precise by our notion of bisimulation. It is also the case that some systems cannot be discretized, once again, one can formalize what this means via bisimulation.

The present paper develops a notion of *labelled Markov process* which is meant ultimately to be part of a theory of interacting dynamical systems. The key technical contribution is the development of a notion of bisimulation for processes which have continuous state spaces but make discrete temporal steps. These are called discrete-time Markov processes. If the state space is also discrete the phrase "Markov chain" is used. The adjective "Markovian" signifies that the transitions are entirely governed by the present state rather than by the past history of the system. The interaction is governed by "labels" in the manner now familiar from process algebra [Hoa85, Mil80, Mil89].

In brief, a labelled Markov process is as follows. There is a set of states and a set of labels. The system is in a state at a point in time and moves between states. Which state it moves to is governed by which interaction with the environment is taking place and this is indicated by the labels. The system evolves according to a probabilistic law. If the system interacts with the environment by synchronizing on a label it makes a transition to a new state governed by a transition probability distribution. So far, this is essentially the model developed by Larsen and Skou [LS91] in their very important and influential work on probabilistic bisimulation. They specify the transitions by giving, for each label, a probability for going from one state to another. Bisimulation then amounts to matching the moves with matching probabilities as well.

In the case of a continuous state space, however, one cannot just specify transition probabilities from one state to another. In most interesting systems all such transition probabilities would be zero! Instead one must work with probability densities. In so doing, one has to confront the major issues that arose when probability theory was first formalized, such as the existence of subsets for which the notion of probability does not make sense. In the appendix we provide a rapid recapitulation of ideas from probability and measure theory. In the present case we have to introduce a notion of sets

for which "probabilities make sense" (i.e. a σ-algebra) and instead of talking about probabilities of going from a state s to another state s', we have to talk about going from a state s to a *set of* states A.

The notion of bisimulation for these systems is a generalization of the definition of Larsen and Skou, which, as we shall argue in the next section, is a compelling, natural notion. Unfortunately this definition cannot be adapted in any simple way to the continuous case. One cannot even use their basic terminology for bisimulation because their definition is inextricably wound up in the notion of state-to-state transition probabilities. Furthermore, once the correct generalization is given it turns out to be a formidable technical problem to even show that bisimulation is an equivalence relation. This is solved by a construction due to Edalat and is given in detail in an appendix. In fact the construction heavily relies on properties that are not true for measure spaces in general. We have assumed a Polish space structure. A Polish space is a topological space which can be given a metric structure which generates the topology and such that the metric is complete and separable, i.e. there is a countable basis for the topology. In the classical study of Markov processes metric ideas play a significant role [Par67]. In any example of physical interest, the spaces will have this Polish structure, indeed they will usually come as metric spaces. Any discrete space is Polish and any of the closed subspaces of $\mathbb{R}^n$ will be as well.

The definition of bisimulation is inspired by the paper of Joyal, Nielsen and Winskel [JNW96] which provides a general categorical view of what bisimulation is in terms of certain special morphisms called *open maps*. It is not straightforward to adapt this to the probabilistic case. For discrete systems this has been done by Cheng and Nielsen [CN95] using infinitesimals[1]. Unfortunately one still needs to know how to construct pullbacks in the underlying category and for this one has to rely on the basic construction given in [Eda96]. It also turns out that our notion of bisimulation is precisely the notion of coalgebra homomorphism [dVR].

The rest of this paper is organized as follows. We describe some examples of systems with continuous state spaces in the next section. We then recapitulate the work of Larsen and Skou on discrete systems. The next two sections are concerned with the definition of labelled Markov processes and bisimulation. The last sections discuss related work and conclusions.

[1] One can actually do this without talking about infinitesimals [Win96].

2 Two Examples of Processes

We begin with a simple example, more for introducing terminology than for any intrinsic interest. Imagine a system with two labels $\{a, b\}$. The state space is the real plane, $\mathbb{R}^2$. When the system makes an a-move from state (x_0, y_0) it jumps to (x, y_0) where the probability distribution for x is given by the density $K_\alpha \exp(-\alpha(x - x_0)^2)$, where $K_\alpha = \sqrt{\alpha/\pi}$ is the normalizing factor. When it makes a b-move it jumps from state (x_0, y_0) to (x_0, y) where the distribution of y is given by the density function $K_\beta \exp(-\beta(y - y_0)^2)$. The meaning of these densities is as follows. The probability of jumping from (x_0, y_0) to a state with x-coordinate in the interval $[s, t]$ under an a-move is $\int_s^t K_\alpha \exp(-\alpha(x - x_0)^2)dx$. Note that the probability of jumping to any given point is, of course, 0. In this system the interaction with the environment controls whether the jump is along the x-axis or along the y-axis but the actual extent of the jump is governed by a probability distribution. Interestingly, this system is bisimilar to a one-state system which can make a or b moves. Thus, from the point of view of an external observer, this system has an extremely simple behaviour. The more complex internal behaviour is not externally visible. The point of a theory of bisimulation that encompasses such systems is to say this sort of thing. Of course this example is already familiar from the nonprobabilistic setting; if there is a system in which all trasitions are always enabled it will be bisimilar (in the traditional sense) to a system with one state.

Now we consider a system which cannot be reduced to a discrete system. There are three labels $\{a, b, c\}$. The state space is $\mathbb{R}$. The following physical description is for "local colour" and is mathematically unimportant. The state gives the pressure of a gaseous mixture in a tank in a chemical plant. The environment can interact by (a) simply measuring the pressure, or (b)it can inject some gas into the tank, or (c) it can pump some gas from the tank. The pressure fluctuates according to some thermodynamic laws depending on the reactions taking place in the tank. With each interaction, the pressure changes according to three different probability density functions, say $f(p_0, p), g(p_0, p)$ and $h(p_0, p)$ respectively, with nontrivial dependence on p_0. There are in addition two threshold values p_h and p_l. When the pressure rises above p_h the interaction labelled b is disabled, and when the pressure drops below p_l the interaction labelled c is disabled. It is tempting to model this as a three state system, with the continuous state space partitioned by the threshold values. Unfortunately one cannot assign unique transition proba-

bilities to these sets of states for any choices of f, g and h; only if very special uniformity conditions are obeyed can one do this.

3 Discrete Probabilistic Systems

In this section we recapitulate the Larsen-Skou definition of probabilistic bisimulation [LS91]. The systems that they consider will be referred to as labelled Markov chains in the present paper.

Definition 3.1 *A* **labelled Markov chain** *is a quadruple* $(S, \mathcal{L}, C_l, P_l)$, *where* S *is a countable set of states,* $\mathcal{L}$ *is a set of labels, and for each* $l \in \mathcal{L}$ *we have a subset* C_l *of* S, *a function,* P_l, *called a* **transition probability matrix**,

$$P_l : C_l \times S \longrightarrow [0, 1]$$

satisfying the normalization condition

$$\forall l \in \mathcal{L}, s \in C_l . \Sigma_{s' \in S} P_l(s, s') = 1.$$

If we have the weaker property

$$\forall l \in \mathcal{L}, s \in C_l . \Sigma_{s' \in S} P_l(s, s') \leq 1$$

we call the system a **partial** *labelled Markov chain.*

The sets C_l are the sets of states that *can* do an l-action. If we have partial labelled Markov chains then we can just dispense with the C_l sets. In what follows we suppress the label set, i.e. we assume a fixed label set given once and for all.

Definition 3.2 *Let* $T = (S, P_l)$ *be a labelled Markov chain. Then a* **probabilistic bisimulation** $\equiv_p$, *is an equivalence on* S *such that, whenever* $s \equiv_p t$, *the following holds:*

$$\forall l \in \mathcal{L} . \forall A \in S/\equiv_p,\ \Sigma_{s' \in A} P_l(s, s') = \Sigma_{s' \in A} P_l(t, s').$$

Two states s *and* t *are said to be* **probabilistically bisimilar** *(*$s \sim_{LS} t$*) in case* (s, t) *is contained in some probabilistic bisimulation.*

Intuitively we can read this as saying that two states are bisimilar if we get the same probability when we add up the transition probabilities to all the states in an equivalence class of bisimilar states. The adding up is crucial – the probabilities are not just another label. The subtlety in the definition is that one has to somehow know what states are probabilistically bisimilar in order to know what the equivalence classes are, which in turn one needs in order to compute the probabilities to match them appropriately. In fact a very natural notion of probabilistic synchronization trees yields a model of a probabilistic version of CCS with both probabilistic branching and nondeterministic branching. If one looks just at probabilistic branching, equality is precisely the Larsen-Skou notion of bisimulation [BK96].

The paper by Larsen and Skou does much more than just define bisimulation. They introduce the notion of testing a probabilistic process and associating probabilities with the possible outcomes. They then introduce a notion of testable properties. The link with probabilistic bisimulation is that two processes are probabilistically bisimilar precisely when they agree with the results of all tests. They also introduce a probabilistic modal logic and show that bisimulation holds exactly when two processes satisfy the same formulas.

One slight awkwardness in the Larsen-Skou definition is that when one compares two processes one has to combine the state sets and define probabilistic bisimulation on the combined states. This is a minor point but the reader should keep this in mind when reading the proofs below.

4 A Category of Markov Processes

A Markov process is a transition system with the property that the transition probabilities depend only on the current state and not on the past history of the process. We will consider systems where there is an interaction with the environment described by a set of labels as in process algebra. For each fixed label the system may undergo a transition governed by a transition probability. One could have a new set of possible states at every instant but, for simplicity, we restrict to a single state space.

We will organize the theory in categorical terms with objects being transition systems and morphisms being simulations. Bisimulation is most easily thought of in these terms. This presentation will also allow us to compare the theory with the more traditional theory of non-probabilistic processes,

see, for example, the handbook article by Winskel and Nielsen [WN95].

In formulating the notion of Markov processes we need to refine two concepts that were used in the discrete case. First we cannot just define transition probabilities between states; except in rare cases such transition probabilities are zero. We have to define transition probabilities between a state and a set of states. Second, we cannot define transition probabilities to any arbitrary set of states; we need to identify a family of sets for which transition probabilities can be sensibly defined. These are the *measurable sets*. Thus in addition to specifying a set of states we need to specify a σ-algebra on the set of states [Ash72, Bil79, Hal74, KT66, Rud66]

A key ingredient in the theory is the transition probability function.

Definition 4.1 *A* transition probability function *on a measurable space* (X, Σ) *is a function* $T : X \times \Sigma \longrightarrow [0, 1]$ *such that for each fixed* $x \in X$*, the set function* $T(x, \cdot)$ *is a (sub)probability measure and for each fixed* $A \in \Sigma$ *the function* $T(\cdot, A)$ *is a measurable function.*

One interprets $T(x, A)$ as the probability of the system starting in state x making a transition into one of the states in A. The transition probability is really a *conditional probability*; it gives the probability of the system being in one of the states of the set A after the transition, *given* that it was in the state x before the transition.

It will turn out to be convenient to work with *sub-probability* functions; i.e. with functions where $T(x, X) \leq 1$ rather than $T(x, X) = 1$. The mathematical results go through in this extended case and the resulting categories are often nicer but the stochastic systems studied in the literature are usually only the very special version where $T(x, X)$ is either 1 or 0. In fact what is often done is that a state x with no possibility of making a transition is modelled by having a transition back to itself. For questions concerning which states will eventually be reached (the bulk of the analysis in the traditional literature) this is convenient. If, however, we are modelling the interactions that the system has with its environment it is essential that we make a distinction between a state which can make a transition and one which cannot.

The key mathematical construction requires a Polish space structure on the set of states. Thus instead of imposing an arbitrary σ-algebra structure on the set of states we will require that the set of states be a Polish space and the σ-algebra be the Borel algebra generated by the topology.

Definition 4.2 *A* **partial, labelled, Markov process** *with label set* $\mathcal{L}$ *is a structure* $(S, \Sigma, \{k_l \mid l \in \mathcal{L}\})$*, where* S *is the set of states, which is assumed*

to be a Polish space, and Σ *is the Borel* σ*-algebra on* S*, and*

$$\forall l \in \mathcal{L}, k_l : S \times \Sigma \longrightarrow [0,1]$$

is a transition sub-probability function. We are usually interested in the following special case called a **labelled Markov process**. *We have a partial, labelled, Markov process as above and a predicate* **Can** *on* $S \times \mathcal{L}$ *such that for every* $(x,l) \in$ **Can** *we have* $k_l(x,X) = 1$ *and for every* $(x,l) \notin$ **Can** *we have* $k_l(x,X) = 0$.

We will fix the label set to be some $\mathcal{L}$ once and for all. The resulting theory is not seriously restricted by this. We will write just (S, Σ, k_l) for partial, labelled, Markov processes instead of the more precise $(S, \Sigma, \forall l \in \mathcal{L}.k_l)$. In case we are talking about discrete systems we will use the phrase "labelled Markov chain" rather than "discrete, labelled, Markov process".

In a (partial), labelled, Markov *chain* the set of states is countable, one can easily define a metric so that each point is an open set and the space is complete and obviously seperable, the Borel σ-algebra is then the entire powerset and the transition probabilities are given by a $\mathcal{L}$-indexed family of functions $\forall l \in \mathcal{L}.P_l : S \times S \rightarrow [0,1]$ satisfying the conditions required of a (sub)probability distribution. From this presentation we can construct the k_l in the following way $k_l(s,\sigma) \stackrel{\text{def}}{=} \sum_{s' \in \sigma} P_l(s,s')$. We use the phrase "transition function" for an object of type $S \times \Sigma \rightarrow [0,1]$ and "transition matrix" for an object of type $S \times S \rightarrow [0,1]$. A probabilistic transition system as defined by Larsen and Skou is precisely a labelled Markov chain. The partial notion has been used by Cheng and Nielsen [CN95] to give an open maps [JNW96] presentation of bisimulation.

In order to define a category of Markov processes we define simulation morphisms between processes. Intuitively a simulation says that a simulating process can make all the transitions of the simulated process with greater probability than in the process being simulated.

Definition 4.3 *A* **simulation morphism** f *between two partial, labelled, Markov processes,*
(S, Σ, k_l) *and* (S', Σ', k'_l) *is a measurable function* $f : (S, \Sigma) \rightarrow (S', \Sigma')$ *such that*

$$\forall l \in \mathcal{L}. \forall x \in S. \forall \sigma' \in \Sigma'. k_l(x, f^{-1}(\sigma')) \leq k'_l(f(x), \sigma').$$

One could have defined simulation to be continuous, however, the simulations play a secondary role in the present paper and we will just require them to be measurable [2] We do at least require this for the definition to make sense since without f being measurable we would not be guaranteed that $f^{-1}(\sigma')$ is measurable.

This notion extends the standard notion of simulation of labelled transition systems [WN95] in the following way. Given a partial, labelled, Markov chain, (S, P_l) we can define a labelled transition system (lts) with the same label set as follows. We take the same set of states S and we define a labelled transition relation $\rightarrow \subseteq (S \times \mathcal{L} \times S)$ by $(s, l, s') \in \rightarrow \iff P_l(s, s') > 0$. Given two labelled transition systems, $(S_1, \rightarrow_1)$ and $(S_2, \rightarrow_2)$, a function $f : S_1 \rightarrow S_2$ is a simulation morphism if $\forall s \in S_1 . s \xrightarrow{l} s' \Rightarrow f(s) \xrightarrow{l} f(s')$. We cannot do this for Markov processes because we can easily have systems where all the point-to-point transition probabilities are zero but the Markov process is nontrivial because the transition probabilities are nonzero to "larger" sets.

Proposition 4.4 *Given two partial, labelled, Markov chains, a simulation morphism between them is also a simulation morphism between the associated labelled transition systems.*

Proof (sketch). Suppose that we have two partial, labelled, Markov chains (S, Σ, k_l) and (S', Σ', k'_l) with f a simulation morphism from S to S'. Now suppose that in the associated lts the transition $s_1 \xrightarrow{l} s_2$ is possible. This means that $k_l(s_1, \{s_2\}) > 0$ so since f is a morphism we must have that $k'_l(f(s_1), \{f(s_2)\}) \geq k_l(s, f^{-1}(f(s_2))) \geq k_l(s_1, \{s_2\}) > 0$; hence in the lts $s_1 \xrightarrow{l} s_2$ is possible. ∎

5 Bisimulation for Markov Processes

The definition of bisimulation is very heavily influenced by the ideas of Joyal, Nielsen and Winskel [JNW96]. The idea is to identify a class of special systems called "observations" or "observable paths" or better still "observable

[2] In older texts, such as Halmos [Hal74] or Rudin [Rud66] measurable is defined to mean that the inverse image of an *open* set is measurable. This means that the composite of two measurable functions need not be measurable. Our definitions are standard and, of course with this definition the composite of two measurable functions is measurable.

path shapes", and to define bisimulation as a relation satisfying a kind of path lifting property, the so-called "open maps".

What one can prove for ordinary labelled transition systems is that if we take paths to be labelled paths in the usual sense then the open maps are morphisms that satisfy a condition called the "zigzag" condition[3]. We essentially want to say that there is a "zigzag relation". One can talk about relations by talking about **spans**. A span in any category between an object S_1 and another object S_2 is a third object T together with morphisms from T to both S_1 and S_2. One can think of this in the category fo **Sets** as viewing a relation as a set of ordered pairs with the morphisms being the projections.

Bisimulation is then defined to hold between two systems if they are connected by a span of zigzags. In our case the zigzag condition is easy to state and it is easy to see that it corresponds to Larsen-Skou bisimulation in the case of partial, labelled Markov chains. For partial, labelled Markov processes it will be our definition of bisimulation. From now on, we assume that all systems are partial and we will stop writing the adjective "partial" explicitly.

Definition 5.1 *The objects of the category* **LMP** *are labelled Markov processes, having $\mathcal{L}$ as set of labels, with simulations as the morphisms. The category of labelled, Markov chains is written* **LMC** *and is the full subcategory of* **LMP** *that includes only the labelled, Markov chains as objects.*

The key concept is the following.

Definition 5.2 *A morphism f from (S, Σ, k_l) to (S', Σ', k'_l) is* a zigzag morphism *if it satisfies the properties:*

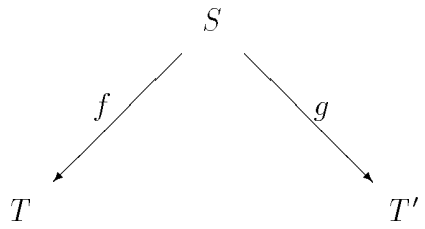
1. *f is surjective;*
2. *f is continuous;*
3. $\forall l \in \mathcal{L}, s \in S, \sigma' \in \Sigma', \;\; k_l(s, f^{-1}(\sigma')) = k'_l(f(s), \sigma')$.

Asking f to be surjective allows us to avoid introducing initial states and worrying about reachable states. Note that we are now taking the topological structure seriously and requiring zigzag morphisms to be continuous. One can immediately check that the identity morphism is a zigzag.

[3]The name arises from modal logic, see, for example, [Pop94].

We often refer to a labelled Markov process by its set of states. Following Joyal, Nielsen and Winskel ([JNW96]) we define bisimulation as the existence of a span of zigzag morphisms.

Definition 5.3 *Let T and T' be two labelled Markov processes. T is* probabilistically bisimilar to T' *(written $T \sim T'$) if there is a span of zigzag morphisms between them, i.e. there exists a labelled Markov process S and zigzag morphisms f and g such that*



Notice that if there is a zigzag morphism between two systems, they are bisimilar since the identity is a zigzag morphism.

It is interesting to note that we can take a coalgebraic view of bisimulation [AM89, Rut95, dVR] as well. We can view a labelled Markov process as a coalgebra of a suitable functor; in fact it is a functor introduced by Giry [Gir81] in order to define a monad on **Mes** analogous to the powerset monad. From this point of view, bisimulation is a span of coalgebra homomorphisms. But if one checks what this means, these are precisely zigzag morphisms.

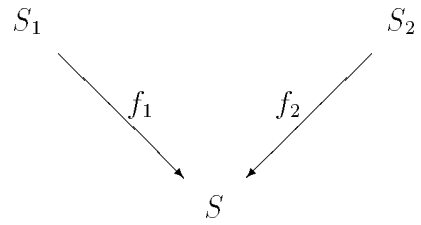
We want bisimulation to be an equivalence, so we need to prove transitivity of the existence of span, since it is obviously reflexive and symmetric. Proving transitivity presents formidable difficulties. In particular, it probably isn't true for probabilistic transition systems without the assumption of Polish structure. The proof works only for Polish spaces[4] The main result proved in detail in the appendix is the following.

Theorem 5.4 *The category with labelled Markov processes as objects and zigzag morphisms as the morphisms has pullbacks.*

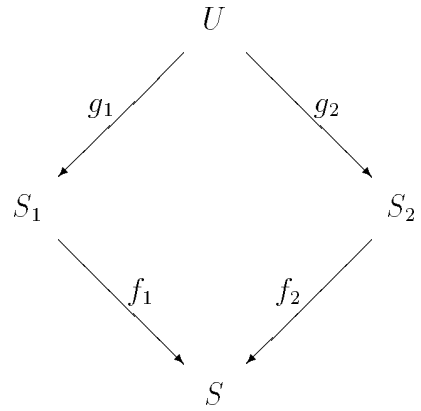
Corollary 5.5 *Bisimulation defined as above (i.e. with Polish space structure on the sets of states and with zigzag morphisms assumed to be continuous) on labelled Markov processes is an equivalence relation.*

[4]We have recently extended it to analytic spaces.

Before outlining the proof it is worth discussing why this isn't trivial. A naive idea might be to mimic the construction for the discrete case but this is too closely tied to the idea of there being probabilities defined for point-to-point transitions. Consider the situation in the diagram below.



where f_1 and f_2 are zigzags. Then we want to show that we can find a system, say U, such that the following diagram commutes.



It is natural to mimic the pullback construction in **Set**, thus, we take U to be the collection of all pairs (s_1, s_2) from $S_1 \times S_2$ such that $f_1(s_1) = f_2(s_2)$, and g_1 and g_2 to be the obvious projections, π_1 and π_2. The set U inherits a σ-algebra from the product σ-algebra on $S_1 \times S_2$. Now we have to define a transition probability $h : U \times \Sigma_U \to [0, 1]$. The requirement that π_1 and π_2 are zigzags forces us to define $h((s_1, s_2), (\sigma_1 \times S_2) \cap U) = k_1(s_1, \sigma_2)$ and similarly for π_2. Unfortunately sets of the form $\sigma_1 \times S_2$ and $S_1 \times \sigma_2$ are not enough to determine a measure on all of Σ_U. Roughly speaking we need the values on all the rectangles not just on the "strips". Worse yet, given a set of the form $(\sigma_1 \times \sigma_2) \cap U$ we do not have σ_1 and σ_2 unambiguously determined; there could be two unrelated sets σ_1' and σ_2' with

$$(\sigma_1 \times \sigma_2) \cap U = (\sigma_1' \times \sigma_2') \cap U$$

so we cannot really use σ_1 and σ_2 to define h.

The actual construction uses two key ideas. The first is to forget about U for the moment and define a transition probability on $S_1 \times S_2$ but instead of just taking the product system we condition the probabilities on the transitions agreeing when their images are looked at in S. Now we can define the transition probability on "rectangles" and extend to all of the space. The conditioning automatically makes the measure live just on U rather than on all of $S_1 \times S_2$. This finesses the ambiguities of the naive approach.

Proof. The fact that the identity morphism is a zigzag immediately implies that bisimulation is reflexive and the fact that it is defined by spans gives symmetry trivially. In order to prove transitivity we need to show that given two spans we can construct a composite span. If we had pullbacks we can of course compose spans. In fact pullbacks do exist if we look at the (obviously nonfull) subcategory of zigzag morphisms. The proof of this is given in detail in the appendix. It should not be assumed that the complete proof is uninteresting or unimportant; indeed it can be formulated as a theorem of pure probability theory.

Three objects

$$S_1 = (S_1, \Sigma_1, k_1 : S_1 \times \Sigma_1 \to [0,1]),$$

$$S_2 = (S_2, \Sigma_2, k_2 : S_2 \times \Sigma_2 \to [0,1]),$$

$$S = (S, \Sigma, k : S \times \Sigma \to [0,1])$$

and morphisms $f_1 : S_1 \to S$ and $f_2 : S_2 \to S$ are given.

Let $U = \{(s_1, s_2) \in S_1 \times S_2 \mid f_1(s_1) = f_2(s_2)\}$ equipped with the subspace topology of the product topology on $S_1 \times S_2$. The Borel σ-algebra Σ_U on U is generated by the set $\{(\sigma_1 \times \sigma_2) \cap U \mid \sigma_1 \in \Sigma_1, \sigma_2 \in \Sigma_2\}$. Let $\pi_1 : U \to S_1$ and $\pi_2 : U \to S_2$ be the projection maps. Since f_1 and f_2 are surjective, U is not empty. We want to construct $h : U \times \Sigma_U \to [0,1]$ so that (U, Σ_U, h) is an object and $\pi_1 : U \to S_1$ and $\pi_2 : U \to S_2$ are morphisms. For an element x of a set X, an element y of a set Y, subsets $A \subseteq X$, $B \subseteq Y$ and a function $f : X \to Y$, we sometimes write x instead of $\{x\}$, fx instead of $f(x)$, fA instead of $f(A)$, and $f^{-1}B$ instead of $f^{-1}(B)$.

We fix $(s_1, s_2) \in U$, i.e. $s_1 \in S_1$ and $s_2 \in S_2$ with $f_1 s_1 = f_2 s_2$, throughout this proof. The index i always takes the values 1 and 2. The variable ω_i always runs through S_i whereas the variable s runs through S.

For $i = 1, 2$, we have the probability measures $k_i(s_i, -) : \Sigma_i \to [0, 1]$ on the space (S_i, Σ_i). Also $\Sigma_i' = \{f_i^{-1}\sigma \mid \sigma \in \Sigma\} \subseteq \Sigma_i$ is a sub-σ-algebra of Σ_i $(i = 1, 2)$. We therefore have, for a given $\sigma_i \in \Sigma_i$, the conditional probability distribution $P((s_i, \sigma_i)|\Sigma_i') : S_i \to [0, 1]$ of the probability measure $k_i(s_i, -)$ given the sub-σ-algebra $\Sigma_i' \subseteq \Sigma_i$. Since S_i is a Polish space we can assume [Dud89, Theorem 10.2.2] that $P((s_i, \sigma_i)|\Sigma_i')(-)$ is a *regular* conditional probability distribution, i.e.

(i) $P((s_i, \sigma_i)|\Sigma_i') : S_i \to [0, 1]$ is Σ_i' measurable and integrable.

(ii) For all $\gamma_i' \in \Sigma_i'$, we have

$$\int_{\gamma_i'} P((s_i, \sigma_i)|\Sigma_i')(\omega_i) k_i(s_i, d\omega_i) = k_i(s_i, \sigma_i \cap \gamma_i').$$

(iii) For $k_i(s_i, -)$-almost all $\omega_i \in S_i$, $P((s_i, -)|\Sigma_i')(\omega_i) : \Sigma_i \to [0, 1]$ is a probability measure on S_i.

A regular conditional probability distribution, like a conditional probability distribution, is unique up to a set of measure zero, i.e. any two functions satisfying the above three properties are equal for $k_i(s_i, -)$-almost all $\omega_i \in S_i$.

One should think of $P((s_i, \sigma_i)|\Sigma_i')(\omega_i)$ as the probability that s_i makes a transition to σ_i given that s_i makes a transition to $f_i^{-1} f_i \omega_i$, or equivalently as the probability that s_i makes a transition to σ_i given that $f_i s_i$ makes a transition to $f_i \omega_i$.

By a standard method we can obtain the conditional probability distribution

$$P((s_i, \sigma_i)|\Sigma) : S \to [0, 1].$$

Here, $P((s_i, \sigma_i)|\Sigma)(s)$ gives the probability that s_i makes a transition to σ_i given that s_i makes a transition to some ω_i with $f_i(\omega_i) = s$ or equivalently the probability that s_i makes a transition to σ_i given that $f_i s_i$ makes a transition to s.

In order to define $h((s_1, s_2), -) : \Sigma_U \to [0, 1]$ we first define a probability measure $g((s_1, s_2), -) : \Sigma_{S_1 \times S_2} \to [0, 1]$ on the product space $(S_1 \times S_2, \Sigma_{S_1 \times S_2})$ where $\Sigma_{S_1 \times S_2}$ is the Borel σ-algebra of $S_1 \times S_2$. We will then show in the next section that $g((s_1, s_2), -)$ is supported on U, i.e. $g((s_1, s_2), U) = 1$. We will finally define $h((s_1, s_2), \alpha \cap U) = g((s_1, s_2), \alpha)$.

Let $\mathcal{R} = \Sigma_1 \times \Sigma_2$ be the collection of all "rectangles" $\sigma_1 \times \sigma_2 \in \Sigma_{S_1 \times S_2}$; $\mathcal{R}$ is a semi-ring. For $\sigma_1 \times \sigma_2 \in \mathcal{R}$ and $s \in S$, put

$$\rho((s_1, s_2), \sigma_1 \times \sigma_2)(s) = P((s_1, \sigma_1)|\Sigma)(s) P((s_2, \sigma_2)|\Sigma)(s).$$

Since for $k(f_1 s_1, -)$-almost all $s \in S$, $P((s_1, -)|\Sigma)(s)$ and $P((s_2, -)|\Sigma)(s)$ are probability measures on S_1 and S_2 respectively, we have:

Proposition 5.6 *[Dud89, Theorem 4.4.1]*
For $k(f_1 s_1, -)$-almost all $s \in S$, $\rho((s_1, s_2), -)(s)$ is countably additive on $\mathcal{R}$.

For $\sigma_1 \times \sigma_2 \in \mathcal{R}$, put

$$g((s_1, s_2), \sigma_1 \times \sigma_2) = \int_S P((s_1, \sigma_1)|\Sigma)(s) P((s_2, \sigma_2)|\Sigma)(s) k(f_1 s_1, ds). \quad (1)$$

Now we can show the following result.

Proposition 5.7 *$g((s_1, s_2), -)$ has a unique extension to a probability measure on $(S_1 \times S_2, \Sigma_{S_1 \times S_2})$.*

The measure g is defined on all of $S_1 \times S_2$ not just on U. But the following remarkable proposition says that in fact g is only supported on U.

Proposition 5.8 *For any $\sigma_1 \in \Sigma_1$ and $\sigma_2 \in \Sigma_2$ with $U \cap (\sigma_1 \times \sigma_2) = \emptyset$ we have $g((s_1, s_2), \sigma_1 \times \sigma_2) = 0$.*

This proof used some subtle facts about Borel sets and continuous functions from Polish spaces to Hausdorff spaces and in given in detail in the appendix. From this result we can easily define

$$h((s_1, s_2), \alpha \cap U) = g((s_1, s_2), \alpha).$$

We can now complete the proof by checking the following statement. $(U, \Sigma_U, h : U \times \Sigma_U \to [0, 1])$ is an object of the category and $\pi_i : U \to S_i$ are morphisms.

The verification of the last statement is as follows. Clearly U is a Polish space as it is a closed subset of the Polish space $S_1 \times S_2$. We have already

shown that $h((s_1, s_2), -)$ is a probability measure on U for any $(s_1, s_2) \in U$. Finally, for any $\sigma_1 \in \Sigma_1$ we have:

$$
\begin{aligned}
& h((s_1, s_2), \pi_1^{-1}\sigma_1) \\
& = h((s_1, s_2), (\sigma_1 \times S_2) \cap U) \\
& = g((s_1, s_2), \sigma_1 \times S_2) \\
& = \int_S P((s_1, \sigma_1)|\Sigma)(s) P((s_2, S_2)|\Sigma)(s) k(f_1 s_1, ds) \\
& = \int_S P((s_1, \sigma_1)|\Sigma)(s) k(f_1 s_1, ds) \qquad\qquad P((s_2, S_2)|\Sigma)(s) = 1 \text{ a.e.} \\
& = k_1(s_1, f_1^{-1}(s) \cap \sigma_1) = k_1(s_1, \sigma_1).
\end{aligned}
$$

Therefore π_1 is a morphism. Similarly π_2 is a morphism. ∎

We close this section with some examples. The first example is technically trivial but conceptually crucial.

Example 5.9 *We let the label set be the one element set. Consider a system (S, Σ, k) with S an arbitrarily complicated states space and Σ a σ-algebra generated by some Polish space structure on S. For example, S could be $\mathbb{R}$, the reals with the Borel algebra. We define the transition function, $k(s, \sigma)$ in any manner we please subject only to the conditions of the definition of a transition function and to the condition that $\forall s \in S.k(s, S) = 1$; i.e. for every s, the distribution $k(s, \cdot)$ is a probability measure. Consider the trivial labelled Markov chain with just one label, one state and one transition from the state to itself with probability 1. These two systems are bisimilar!*

This example allows us to clarify the discussion in the introduction. All of conventional stochastic process theory is described by systems like the first system above. From our point of view they are trivial. This is to be expected, as we are modelling *interaction* and all such systems are indeed trivial from the point of view of interaction. In order to get nontrivial examples one has to consider systems with richer label sets, and which are not always capable of making transitions with every label. Note also that many continuous systems are bisimilar to discrete ones. If we know this we can use the (presumably simpler) discrete system when reasoning about composite systems. Without our framework this would not be possible.

Example 5.10 *Consider the labelled Markov chain over a one-element label set* $(\{a,b,c\},P)$ *where the transition matrix,* P *is given by* $P(a,b) = P(a,c) = 1/2$ *with all other entries being zero. Consider another system over the same label set with* 4 *states* $\{w,x,y,z\}$ *and with transition matrix* Q *given by* $Q(w,x) = Q(w,y) = Q(w,z) = 1/3$*. These two systems are bisimilar. One can easily verify this by constructing the obvious "product" system and checking that the projections are zigzags. On the other hand there is no zigzag morphism between the two systems. These systems are small enough to check this by hand.*

The last example shows that bisimulation really does have to be defined in terms of spans, i.e. it really is a relation.

Example 5.11 *The next example illustrates a continuous system. Consider the labelled Markov process, over the trivial label set, defined as follows* $S = (\mathbb{R},\mathcal{B},k)$*, i.e. the states are real numbers, the measurable sets are Borel sets and the transition function is defined on intervals (and then extended to arbitrary Borel sets) as follows:*

$$k(x,[r,s]) = \begin{cases} \lambda/2 \int_r^s e^{-\lambda|x-y|}dy & \text{if } x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

where the constant factor of $\lambda/2$ *chosen to make* k *be* 1 *on the whole space. Intuitively this is a system where a particle makes random jumps with probability exponentially distributed with the length. However, there is an "absorbing wall" at the point* $x = 0$ *so that if the system jumps to the left of this point it gets stuck there. Note that every positive state has a different probability density for jumping to a negative state. Now consider the system* $U = (\mathbb{R}^2,\mathcal{B}^2,h)$ *defined as*

$$h((x,y),[r,s]\times[p,q]) = k(x,[r,s])P([p,q]),$$

where P *is some arbitrary probability measure over* $\mathbb{R}$*. This system should behave "observably" just like the first system because, roughly speaking, the first coordinate behaves just like the first system and the second system has trivial dynamics, i.e. it is bisimilar to the one-state, one-transition system. Indeed these two systems are bisimilar with the projection from the second to the first being a zigzag.*

The next example illustrates a possible objection to our definition.

Example 5.12 *Suppose that we have two systems,* $(\mathbb{R}, \mathcal{B}, \{a, b\}, k_l)$ *and* $(\mathbb{R}, \mathcal{B}, \{a, b\}, h_l)$. *In the first system we have the following transitions*

$$k_a(x, S) = \mu(S \cap [x - 0.5, x + 0.5])$$

and

$$k_b(x, S) = \mu(S \cap [x + 0.5, x + 1.5])$$

where S *is a Borel set and* μ *is Lebesgue measure. For the other system we have*

$$h_a(x, S) = \begin{cases} \mu(S \cap [x - 0.5, x + 0.5]) & \text{if } x \text{ is irrational} \\ 0 & \text{if } x \text{ is rational.} \end{cases}$$

the b transitions are the same as those for the first system. These two systems are not bisimilar by our definition. The first one is bisimilar to the trivial one-state system with both a and b enabled all the time while the second one is has states in which a gets disabled. However, the probability of landing in one of these states is 0. *Thus, in some sense, the difference is visible only on a set of probability* 0. *Should they be distinguished?*

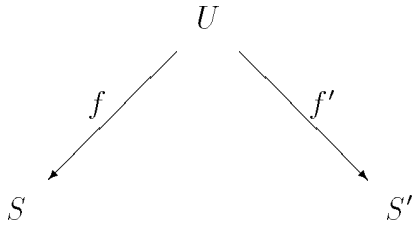
6 Discrete Systems Revisited

In this section we reconsider discrete systems (Markov chains) from the point of view of the bisimulation notion that we have defined for general partial, labelled, Markov processes. We show that the Larsen-Skou definition coincides with the "span of zigzags" definition. First, we have to say what it means for two Markov chains to be Larsen-Skou bisimilar, since the Larsen-Skou definition involves states of a single process rather than states of two different processes, and so doesn't apply without an appropriate interpretation. Roughly speaking, if there were an initial state in every Markov chain, then we would say that two Markov chains are Larsen-Skou bisimilar if and only if in a system "combining" them, their initial states are Larsen-Skou bisimilar. We have chosen not to equip Markov chains with initial states, instead we will use the following as a definition of Larsen-Skou bisimulation between two Markov processes.

Definition 6.1 Let (S_1, P_l^1), (S_2, P_l^2), be two Markov chains and (T, H_l), their disjoint union, that is, T is the disjoint union of the two sets S_1 and S_2, and for $s, t \in T$, $H_l(s,t) = P_l^i(s,t)$ if s, t are both in S_i $(i = 1, 2)$ and zero otherwise. We will say that $S_1 \sim_{LS} S_2$ if in T there is a probabilistic bisimulation $\equiv_p$ such that for every state s_i of S_i there is a state s_j of S_j such that $s_i \equiv_p s_j$ $(i, j \in \{1, 2\}, i \neq j)$.

We assume that all systems are partial in this section. Recall that we can always define a topology in which all the sets are open and hence all functions are continuous. Since the space is countable the topology is clearly Polish. Thus, we can forget about the topology in this section but we keep in mind that discrete systems are a special case of the formalism of the previous section.

Proposition 6.2 *Let (S, P_l) and (S', P_l') be two labelled Markov chains. $(S, P_l) \sim_{LS} (S', P_l')'$ if and only if there exists a span of zigzag morphisms f and f':*



Proof. We write k_l and k_l' for the two transition functions induced by the transition matrices P_l and P_l' respectively.

$\Leftarrow$: We first show that if $S \xrightarrow{f} S'$, where f is a zigzag morphism, then $S \sim_{LS} S'$. Let (t, H_l) be the disjoint union of (S, P_l) and (S', P_l'), and h_l be the transition function induced by the transition matrix H_l. Now f defines the following equivalence relation, R, on T:

$$s_1 R s_2 \iff (s_1 = s_2) \vee (f(s_1) = s_2) \vee (f(s_1) = f(s_2)).$$

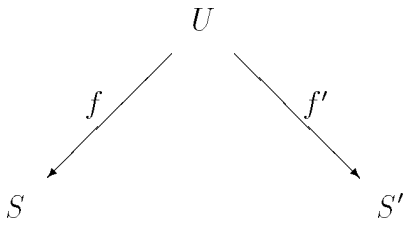
The equivalence classes are of the form $\{s'\} \cup f^{-1}(s')$ for each $s' \in S'$; thus each equivalence class can be represented uniquely by an element of S'. Let $l \in \mathcal{L}, s_1, s_2 \in T$ such that $s_1 R s_2$, and choose any $t' \in S'$ that represents the equivalence class $\{t'\} \cup f^{-1}(t')$. We want to show that $h_l(s_1, -)$ and $h_l(s_2, -)$ agree on $\{t'\} \cup f^{-1}(t')$.

First assume $s_1 \in S$ and $s_2 \in S'$, meaning that $f(s_1) = s_2$. Then

$$\begin{aligned} h_l(s_1, \{t'\} \cup f^{-1}(t')) &= k_l(s_1, f^{-1}(t')) \text{ since } s_1 \in S \\ &= k'_l(f(s_1), \{t'\}) \text{ since } f \text{ is a zigzag morphism} \\ &= h_l(s_2, \{t'\} \cup f^{-1}(t')), \end{aligned}$$

which is precisely the condition for Larsen-Skou bisimulation. Now if s_1 and s_2 are both in S (and still R-related), they have the same image, so $s_1 \, R \, f(s_1) = f(s_2) \, R \, s_2$ and we can simply apply the above calculation. Finally we have the trivial case where s_1 and s_2 are both in S', they are then equal and we are done since, f being a surjective function, every state of S is R-equivalent to a state of S' are vice versa. Since we know that Larsen-Skou bisimulation is an equivalence relation it follows that whenever we have a span of zigzags connecting two labelled Markov chains they are Larsen-Skou bisimilar.

$\Rightarrow$: Assume $(S, P_l) \sim_{LS} (S', P'_l)$, with $\equiv_p$ the probabilistic bisimulation over (T, H_l) the disjoint union of (S, P_l) and (S', P'_l). We need to construct a span of zigzag morphisms



To do this, let $U = (U, (Q_l)_{l \in \mathcal{L}})$ where $U = \{(s, s') \in S \times S' : s \sim_{LS} s' \text{ in } T\}$ and where the transition matrix Q is given by, for $l \in \mathcal{L}$,

$$Q_l((s, s'), (t, t')) = \frac{P_l(s, t) \, P'_l(s', t')}{h_l(s, [t]_{\sim_{LS}})}$$

where $[t]_{\sim_{LS}}$ denotes the equivalence class containing t in T, and h_l is the transition function induced by the transition matrix H_l. Since $s \sim_{LS} s'$ and $t \sim_{LS} t'$, we have by definition of $\sim_{LS}$ that $h_l(s, [t]_{\sim_{LS}}) = h_l(s', [t]_{\sim_{LS}}) = h_l(s', [t']_{\sim_{LS}})$.

To prove that U is a partial, labelled, Markov chain, we need that for any $(s, s') \in U$,

$$\sum_{(t,t') \in U} Q_l((s, s'), (t, t')) \leq 1.$$

This will follow from the proof that we have zigzag morphisms from U to S and S'. As morphisms $f : U \to S$ and $f' : U \to S'$, we simply take the left and right projections which are surjective by definition of Larsen-Skou probabilistic bisimulation. We prove that they are zigzag morphisms. We write q_l for the transition function derived from the transition matrix Q_l. First note that

$$\forall t \in S. f^{-1}(t) = \{t\} \times (S' \cap [t]_{\sim_{LS}}).$$

For any $l \in \mathcal{L}, (s, s') \in U, t \in S$, we have

$$\begin{aligned}
q_l((s,s'), f^{-1}(t)) &= \sum_{t' \in S' \cap [t]_{\sim_{LS}}} Q_l((s,s'),(t,t')) \\
&= \sum_{t' \in S' \cap [t]_{\sim_{LS}}} \frac{P_l(s,t)\, P'_l(s',t')}{h_l(s, [t]_{\sim_{LS}})} \\
&= \frac{P_l(s,t)}{h_l(s', [t]_{\sim_{LS}})} \sum_{t' \in S' \cap [t]_{\sim_{LS}}} P'_l(s',t') \\
&= \frac{P_l(s,t)}{k'_l(s, [t]_{\sim_{LS}} \cap S)} k'_l(s', [t]_{\sim_{LS}} \cap S') \\
&= P_l(s,t) = k_l(f(s,s'), \{t\}).
\end{aligned}$$

and f is thus a zigzag morphism. The same argument applies to f'. ∎

7 Testing equivalences

In this section we examine the relationship between a notion of testing equivalence, 2-way simulation and probabilistic bisimulation.

The tests we consider are simple sequences of experiments, each just described by a label. As usual, let $\mathcal{L}$ be the fixed set of labels describing the interaction of processes with their environments.

Definition 7.1 *A* **test over** $\mathcal{L}$ *is a finite sequence of labels of* $\mathcal{L}$*. Given a labelled Markov process* $S = (S, \Sigma_S, k_l)$*, a state* s_0 *of* S *and a test* T*, the*

probability $p(s_0;T)$ of success is defined in the following way;

$$p(s_0;\epsilon) = 1$$
$$p(s_0;a.T) = \int_S p(s;T)k_a(s_0,ds)$$

where ϵ is the empty string. Let (S,Σ,k_l) and (S',Σ',k'_l) be two labelled Markov processes. We say that S test-simulates *S' if, for every $s' \in S'$, there exists $s \in S$ such that for all test T,*

$$p(s';T) \leq p(s;T).$$

We say that S and S' are **test-equivalent**, *written $S \sim_T S'$ if they test-simulate each other.*

It is easily seen that $p(\cdot;T) : S \to \Re$ is a measurable function and that test-equivalence is indeed an equivalence relation.

Proposition 7.2 *Let (S,Σ,k_l) and (S',Σ',k'_l) be labelled Markov processes. If there is a simulation morphism from S to S', then S' test-simulates S. Moreover, if the morphism is a zigzag morphism, then S and S' are test-equivalent.*

Proof. We prove the second part since the first one can be easily proved in almost the same way, essentially by replacing the equalities by inequalities. Let $f : S \to S'$ be a zigzag morphism. We show that for all $s \in S$, for all tests T, $p(s;T) = p(f(s);T)$ by induction on the length of tests. Since f is a surjective function, this will prove that $S \sim_T S'$. We trivially have $p(s;\epsilon) = p(f(s);\epsilon) = 1$ for all $s \in S$. Now assume the assertion is true for tests of length n. Let $s \in S$ and let aT be a test of length $n+1$. We have by induction hypothesis and from the definition of zigzag expressed in the following from,

$$k_a(s,\cdot) \circ f^{-1} = k'_a(f(s),\cdot),$$

that

$$\begin{aligned} p(s;a.T) &= \int_S p(t;T)k_a(s,dt) \\ &= \int_S p(f(t);T)k_a(s,dt) \\ &= \int_{S'} p(s';T)k'_a(f(s),ds') \\ &= p(f(s);a.T). \end{aligned}$$

■

Corollary 7.3 *If there is a 2-way simulation between S and S', then S and S' are test-equivalent.*

Corollary 7.4 *If S and S' are bisimilar, then S and S' are test-equivalent.*

Note that the converse is not true, that is there exist two stochastic transition systems that are test-equivalent but not stochastic bisimilar.

These results are not meant to be a serious analysis of testing notions. They are only meant to give some feeling for what zigzag means. In future work we intend to analyze testing thoroughly.

8 Towards A Modal Logic for Bisimulation

In this section we begin the study of logics that characterize bisimulation. For discrete systems, indeed even if only one system is discrete, we can define a simple logic which plays the role of Hennessy-Milner logic. In practice one often wants to know that a continuous system can be discretized, i.e. is bisimilar to some discrete system, and this fact will be useful to determine that. We follow the treatment of Larsen and Skou [LS91] closely in terms of the definition of the logic but not in terms of proofs. The key difference is that we use negation.

We take as the syntax the following formulas:

$$T|\neg\phi|\phi_1 \wedge \phi_2|\langle a\rangle_q\phi$$

where a is an action from the fixed set of actions $\mathcal{L}$ and q is a rational number. Given a labelled Markov process (S, Σ, k_a) we write $s \models \phi$ to mean that the state s satisfies the formula ϕ. The definition of the relation $\models$ is given by induction on formulas. The definition is obvious for the propositional constants and connectives. We say $s \models \langle a\rangle_q\phi$ iff $\exists A \in \Sigma.\forall s' \in A.s' \models \phi \wedge k_a(s, A) \geq q$. In other words, s can make an a-move to a state that satisfies ϕ with probability greater than q. We write $[\![\phi]\!]$ for the set $\{s \in S | s \models \phi\}$. If we have two systems , say S, S', in mind and we want to distinguish them we write $[\![\phi]\!]_S$.

The first proposition below says that sets of states definable by formulas are always measurable.

Proposition 8.1 *For all formulas ϕ, we have $[\![\phi]\!] \in \Sigma$.*

Proof. We proceed by structural induction on ϕ. The base cases corresponding to T is trivial since $S \in \Sigma$. The propositional connectives are trivial because a σ-algebra is closed under countable union and complementation, by definition. Finally we have $[\![\langle a \rangle_q \phi]\!] = k_a(\cdot, [\![\phi]\!])^{-1}([q,1]) \in \Sigma$. To justify this first note that, by hypothesis, $[\![\phi]\!] \in \Sigma$ so $k_a(s, [\![\phi]\!])$ is meaningful. Secondly, k_a is a measurable function in its first argument and finally intervals are Borel. ∎

The next proposition links zigzag morphisms with formulas in the logic.

Proposition 8.2 *If f is a zigzag morphism from S to S' then for all formulas ϕ,*

$$s \models \phi \iff f(s) \models \phi.$$

Proof. We proceed by structural induction on ϕ. The only nontrivial cases correspond to the modal formulas. We proceed as follows; $s \models \langle a \rangle_q \phi$ means $k_a(s, [\![\phi]\!]_S) \geq q$. By the inductive hypothesis we have that $f^{-1}([\![\phi]\!]_{S'}) = [\![\phi]\!]_S$. Thus we have

$$q \leq k_a(s, [\![\phi]\!]_S) = k_a(s, f^{-1}([\![\phi]\!]_{S'})) = k'_a(f(s), [\![\phi]\!]_{S'}).$$

But the last equality means $f(s) \models \langle a \rangle_q \phi$. ∎

From this we get the immediate corollary, but first we need to define bisimilarity of states. We say $s \in S$ and $s' \in S'$ are bisimilar if S and S' are bisimilar and for *some* span of morphisms $f : T \to S$ and $f' : T \to S'$ there exists $t \in T$ with $f(t) = s$ and $f'(t) = s'$. Now this defines an equivalence relation. In anticipation of this we write, $s \sim s'$ when two states are bisimilar. To check the claim suppose that for $s \in S$, $s' \in S'$ and $s'' \in S''$ we have $s \sim s'$ and $s' \sim s''$. Then we have $f : T \to S$ and $f' : T \to S'$ and $g : T' \to S'$ and $g' : T' \to S''$ with the required properties and we can use the pullback $(h : U \to T, k : U \to T')$ of $f' : T \to S'$ and $g : T' \to S'$ to obtain $p : U \to S$ and $q : U \to S''$. In order to show that $s \sim s''$ we proceed as follows. By assumption there exists $t \in T$ and $t' \in T'$ such that $f(t) = s$, $f'(t) = s'$, $g(t') = s'$ and $g'(t') = s''$. Now $(t, t') \in U$ since $f'(t) = s'$ and $g(t') = s'$. Therefore, putting $p = f \circ h$ and $q = g' \circ k$ we have $p((t,t')) = f(t) = s$ and $q((t,t')) = g'(t') = s''$. Thus $s \sim s''$.

Corollary 8.3 *If s and s' are bisimilar states of S and S' respectively then they satisfy the same formulas.*

Now one can show that the transition probabilities to definable sets are determined completely by the formulas, independently of the system. Suppose that (S, Σ, k_a) and (S', Σ', k'_a) are two systems. We say that they satisfy *all the same formulas* if $\forall s \in S \exists s' \in S'$ such that s and s' satisfy all the same formulas and the same with s and s' interchanged. We write $S \bowtie S'$ if this is the case.

Proposition 8.4 *Suppose that $(S, \Sigma, k_a) \bowtie (S', \Sigma', k'_a)$ then for all formulas ϕ and all pairs (s, s') such that s and s' satisfy all the same formulas, we have $k_a(s, [\![\phi]\!]_S) = k'_a(s', [\![\phi]\!]_{S'})$.*

Proof. Suppose that the equation does not hold. Then, say, for some ϕ, $k_a(s, [\![\phi]\!]_S) < k'_a(s', [\![\phi]\!]_{S'})$. Now we choose a rational number q between these values. Now it follows that $s' \models \langle a \rangle_q \phi$ but $s \not\models \langle a \rangle_q \phi$, which contradicts the assumption that s and s' satisfy all the same formulas. ∎

Now we would like to show that if two systems satisfy all the same formulas they must be bisimilar. Instead of defining a span of zigzags directly, we can define a cospan and use the pullback property to infer that a span must exist. Given two systems $(S, \Sigma, k_a) \bowtie (S', \Sigma', k'_a)$ we first construct a system, (T, Σ_T, j_a), called the *direct sum* of S and S', as follows. We set $T = S \uplus S'$ with the evident σ-algebra. We define the transition probabilities as follows: $j_a(s, A \uplus A') = k_a(s, A)$ if $s \in S$ and $j_a(s', A \uplus A') = k'_a(s', A')$ if $s' \in S'$ where $A \in \Sigma$ and $A' \in \Sigma'$. There are the evident canonical injections ι, ι' which are *not* zigzags because they are not surjective. Now we define an equivalence relation on the states of T by saying that two states are equivalent iff they satisfy all the same formulas. We write $s \approx s'$. We define a quotient system (V, Σ_V, h_a) as follows. The states of V are the equivalence classes and the topology is the greatest one making the canonical surjection $r : T \to V$ continuous (and hence measurable). The set T is a Polish space since it is discrete. The composites $r \circ \iota$ and $r \circ \iota'$ are continuous and surjective, henceforth we call them q and q' respectively. To see that q, q' are surjective we recall that an equivalence class must, by hypothesis, include members of both S and S'. If we can define h_a so as to make q, q' both zigzag we will be done.

Proposition 8.5 *If we have* $(S, \Sigma, k_a) \bowtie (S', \Sigma', k'_a)$ *and* $s \approx s'$ *then for all* $t \in T$*, defined as above, we have* $k_a(s, [t] \cap S) = k'_a(s', [t] \cap S')$.

Proof. Let $\{\phi_i | i \in \mathbb{N}\}$ be an enumeration of the formulas satisfied by the members of $[t]$. If s is any state such that $s \bowtie t$, s must satisfy any formula satisfied by t, so s is in the intersection. Conversely, if $s \not\bowtie t$, there is a formula ϕ that distinguishes s and t, hence either $s \notin [\![\phi]\!]$ and $t \in [\![\phi]\!]$ or $s \in [\![\phi]\!]$ and $t \notin [\![\phi]\!]$; since the latter is equivalent to saying $s \notin [\![\neg\phi]\!]$ and $t \in [\![\neg\phi]\!]$, we have in both cases that s is not in the intersection of the ϕ_i's.

Let $\forall i \in \mathbb{N}.\psi_i \stackrel{\text{def}}{=} \bigwedge_{l=1}^{i} \phi_l$. We show that $[t] = \cap_{i=1}^{\infty} [\![\psi_i]\!]$. The sets $[\![\psi_i]\!]_S$ and $[\![\psi_i]\!]_{S'}$ are both measurable for all i and decrease as i increases. Furthermore, these decreasing chains converge to $[t] \cap S$ and $[t] \cap S'$ respectively. By the monotone convergence property of measures we have

$$k_a(s, [t] \cap S) = \lim_{i \longrightarrow \infty} k_a(s, [\![\psi_i]\!]_S),$$

by the previous proposition we have

$$= \lim_{i \longrightarrow \infty} k'_a(s', [\![\psi_i]\!]_{S'}) = k'_a(s', [t] \cap S').$$

■

We can define h_a as follows. We only need to specify the point to point transition probabilities since, in the discrete case, these determine all the transition probabilities. Now we set $h_a([s], [t]) \stackrel{\text{def}}{=} k_a(s, q^{-1}([t]))$. Clearly the representative of the $[s]$ equivalence class does not matter. Furthermore by proposition 8.5, we can see that $h_a([s], [t]) = k'_a(s', q'^{-1}([t]))$. The morphisms are now zigzags. Thus we have proved the following proposition

Proposition 8.6 *If we have two systems and one of them is discrete then they are bisimilar iff they obey the same formulas.*

Note that this implies the result for discrete systems without using the minimum deviation assumption used by Larsen and Skou.

9 Related Work

There has been a substantial amount of work on probabilistic transition systems and their associated equivalences. As far as we are aware, none of them have looked at bisimulation for continuous state spaces.

The starting point of work in the area of probabilistic semantics are the fundamental papers of Saheb-Djahromi [SD78, SD80] and of Kozen [Koz81, Koz85]. These are concerned with domain theory and programming languages rather than with process equivalences but they both introduced non-trivial measure-theoretic ideas. Kozen also noticed a very interesting duality between state-transformer semantics as described by stochastic kernels and a probabilistic predicate-transformer semantics in which programs are seen as inducing linear continuous maps on the Banach algebra of bounded measurable functions.

The one paper with an abstract categorical approach to stochastic processes is by Giry [Gir81] but she studies categorical constructions rather than process equivalences. Of course her work, inspired originally by Lawvere, does provide some of the mathematical underpinnings of our ideas. In particular she shows that the stochastic kernels (conditional probability distributions) that we use to define transition probabilities arise as the Kleisli category of a monad, which is a natural generalization of the powerset monad to the probabilistic case. If we recall that the category **Rel** of sets and relations is the Kleisli category of the powerset monad, we see that the stochastic kernels can reasonably be viewed as the probabilistic analogues of relations. This makes the analogy between labelled Markov processes and ordinary transitions quite striking.

The starting point of our investigations has been the pioneering paper of Larsen and Skou [LS91] which gave the first compelling analysis of probabilistic bisimulation. As we have mentioned before their work is couched in terms of state-to-state transition probabilities so there were both conceptual as well as technical obstacles to overcome. The work of Joyal, Nielsen and Winskel [JNW96] and of Cheng and Nielsen [CN95] provided vital clues. Following Joyal et. al. we define probabilistic bisimulation as spans of zigzags.

From the point of view of applications there have been a number of very interesting results. The most interesting work, in our opinion, is the work of Jane Hillston [Hil94] on developing a process algebra for performance evaluation. Her work is not comparable to ours at present because she works with continuous-time Markov chains, i.e. with continuous time and a discrete state space. Nevertheless it is our hope that as our own work evolves we will make contact with her approach.

The group at Oxford has been developing an extensive theory of probabilistic systems; see the collection of reports available from the web [Pro]. The focus has been on equational laws satisfied by processes. From the semantic

point of view they have extensively developed and enriched Kozen's [Koz85] predicate-transformer view. They have looked at continuous state-space systems and have incorporated nondeterminism in their framework. As with the earlier work of Kozen [Koz81, Koz85] stochastic kernels play an important role.

Finally we mention a sample of the large amount of work on various notions of testing [CSZ92, Chr90, JY95], simulations [SL94], specification [JL91], model checking [HK96], generative processes [vGSST90] and other variations which, for example, allow one to have both probability and nondeterminism. It should be possible to develop analogous notions for a variety of continuous state-space systems once the mathematical foundations are in place. With the present work the mathematical apparatus needed is in place though of course a significant amount of work remains to determine which models and equivalences are useful in practice.

10 Conclusions

In this paper we have two main contributions. First, we develop a notion of labelled Markov process as a model of interacting Markov processes. The state space may have complex structure but the interactions are simple synchronizations. Unlike the classical theory of Markov processes, which asks about the evolution of the system in state space, we study behavioural equivalences from the point of view of observing interactions. This kind of system is exactly what is studied in much of hybrid systems – simple finite-state controllers interacting with a complex dynamical system. Of course the stochastic versions of such systems have not been studied very much as yet.

Secondly we develop a notion of bisimulation for probabilistic processes which applies to systems where the state space can be continuous. The essential limitation, met in all examples known in physics, control theory or computing, is that the state space be equipped with the structure of a Polish space. The fundamental technical contribution is the proof that this is indeed an equivalence relation. This requires a fairly hard proof which can be seen as a theorem of Markov processes in its own right. In particular no previous study of probabilistic bisimulation has had to deal with these technical questions. We expect that future studies of such systems will use either this technique or ones very like it.

The secondary results of this paper are

1. that we have shown how this notion of bisimulation extends the definition of Larsen and Skou [LS91] for the discrete case

2. and that we have a logical characterization of bisimulation which works even when only one of the systems is discrete and *does not use the minimum deviation assumption.*

The important questions to answer for the future are the following.

1. The notion of bisimulation in the probabilistic case can be, and should be, debated. For example suppose there are two systems which have different behaviours but these differences are only visible with probability 0, do we want to say that they are or are not bisimilar. In the present paper we would say that they are not bisimilar.

2. What is the general logical characterization of bisimulation? We might have to adopt the point of view that formulas are measurable functions and that the satisfaction "relation" has to be generalized to mean integration. Kozen has already used such ideas [Koz85] in his study of probabilistic PDL.

3. What is the right notion of testing and can we exhibit bisimulation as a testing equivalence? The bisimulation relation that we introduce in this paper is, like Larsen and Skou's, sensitive to small changes in the probabilities. This makes the development of the appropriate testing notions important in order to discuss what one might plausibly test for in a continuous state space setting. In particular it should allow us to initiate a serious study of stability of systems.

4. What is the right notion of simulation? What will happen to simulation and bisimulation in the presence of nondeterminism? These questions have been studied extensively in the discrete case.

5. What are the right calculi for such systems? This is completely wide open at this point.

6. Can we extend these ideas to continuous time? This certainly should be possible and really bring this work into contact with extant work on hybrid systems.

Acknowledgements

Abbas Edalat would like to thank Terry Lyons for helpful discussions and EPSRC (UK) for support. Prakash Panangaden would like to thank Samson Abramsky, Christel Baier, Luca Cattani, Jørgen Hoffman-Jørgensen, Alan Jeffrey, Jan Rutten, Ian Stark, Colin Stirling and Glynn Winskel for helpful discussions and BRICS for its hospitality during the course of this work. Richard Blute, Josée Desharnais and Prakash Panangaden thank NSERC (Canada) for its support. Richard Blute and Prakash Panangaden acknowledge the support of an Ontario-Québec Bilateral Exchange grant.

A A Review of Measure Theory

For completeness we give the basic definitions in this section. We only discuss positive measures. We assume, however, that the reader knows the basic ideas of measure theory and probability as expounded in, for example "Probability and Measure" by Billingsley [Bil79] or "Real Analysis and Probability" by Ash [Ash72] or the book with the same title by Dudley [Dud89] or "Introduction to Measure and Probability" by Kingman and Taylor [KT66].

The basic concept is that of a σ-algebra on a set. These are the sets for which a notion of "measure" or "size" can be defined.

Definition A.1 *A* **σ-algebra** *on a set X is a family of subsets of X which includes X itself and which is closed under complementation and countable unions.*

A set equipped with a σ-algebra is called a *measurable space*. For simple situations we can usually take the σ-algebra to be all the subsets but when we get to sets like the real line (or closed intervals of the real line) we cannot any longer work with this σ-algebra because one cannot consistently define probabilities (measures) sensibly on all subsets.

Fact A.2 *The intersection of an arbitrary collection of σ-algebras is a σ-algebra.*

Thus, given an arbitrary collection of subsets of a set we can define the σ-algebra generated by the collection. For example, given a topological space $(X, \mathcal{T})$, we can define the σ-algebra, often written $\mathcal{B}$, generated by the open sets. This is usually called the *Borel algebra*.

The notion of *measure* generalizes the concept of "length of an interval" to arbitrary σ-algebras.

Definition A.3 *Given a σ-algebra (X, Σ), a* **(probability) measure** *on X is a ($[0,1]$-valued) $[0,\infty]$-valued set function, μ, defined on Σ such that*

- $\mu(\emptyset) = 0$,
- *for a pairwise disjoint, countable collection of sets, $\{A_i | i \in I\}$, in Σ, we require*

$$\mu(\bigcup_{i \in I} A_i) = \sum_{i \in I} \mu(A_i).$$

In addition for probability measures we require $\mu(X) = 1$.

The second property above is called *countable additivity*. A *probability space* is a set X equipped with a σ-algebra Σ and a measure P such that $P(X) = 1$.

Example A.4 *The key example in the subject is provided by Lebesgue measure on the real line. Let* $\mathbf{R}$ *stand for the real numbers. The elements of the σ-algebra generated by the open intervals are called the Borel sets. There is a measure defined on these sets which on intervals agrees with the intuitive notion of length. This is the Lebesgue measure. It may be calculated as follows. Suppose that A is a Borel set. We take a family of intervals, $\mathcal{J}$, (not necessarily pairwise disjoint) which covers A. We define the length of $\mathcal{J}$ as the sum of the lengths of all the intervals in $\mathcal{J}$. Now we take the inf of the lengths over all $\mathcal{J}$ which cover A. One can verify that single points have measure zero and that any countable set has measure zero. The Cantor set gives an example of an uncountable set of measure zero. The set of all irrationals between 0 and 1 has measure 1. This measure is translation invariant and is in fact the unique translation-invariant measure which assigns 1 to the interval $[0,1]$. The fact that makes this entire theory so complicated is the existence of nonmeasurable sets.*

Functions between measurable spaces are called measurable functions.

Definition A.5 *A function $f : (X, \Sigma_X) \to (Y, \Sigma_Y)$ between measurable spaces is said to be* **measurable** *if $\forall B \in \Sigma_Y . f^{-1}(B) \in \Sigma_X$.*

This resembles the topological notion of continuity. The fundamental difference is that limits of sequences of measurable functions is measurable; the analogous property certainly does not hold for continuity. Measurable spaces and measurable functions form a category called **Mes**. This category has countable coproducts and all finite limits (in particular products and equalizers).

We now define Lebesgue integration. We assume that we are given a measure space (X, Σ, μ). Let $f : X \to \mathbb{R}$ be a real-valued measurable function defined on X. Instead of carving up the domain space as is done in Riemann integration we carve up the range. More precisely we define a *simple*, real-valued function on X to be a function whose range is finite. Any simple function, say s, can be expressed as a linear combination of characteristic functions

$$s = \sum_{i=i}^{n} r_i \chi_{A_i}$$

where r_i are real numbers (the values assumed by s) and A_i are the sets on which s assumes the value r_i. The A_i have to be measurable for s to be measurable. It is clear what the integral of such a simple function should be

$$\int s d\mu \stackrel{\text{def}}{=} \sum_{i=1}^{n} r_i \mu(A_i).$$

Let us assume that f, the function we wish to integrate is positive. Now a fundamental (but easy) result of measure theory is that a positive measurable functions f can be approximated pointwise by an *increasing* sequence of simple functions

$$s_1 \leq s_2 \leq \ldots \leq s_n \leq \ldots \leq f.$$

We use this as the basis of the definition of the integral of f

$$\int f d\mu = \lim_{j \longrightarrow \infty} \int s_j d\mu$$

where the limit on the right is assumed to exist. If the limit does exist we say that the function is *integrable* and the value of the limit is the value of the integral. It is not difficult to show that the value is independent of

the approximating sequence of simple functions. The integral of a general function is defined by breaking it up into positive and negative pieces and defining the two integrals separately. The Lebesgue theory has much better convergence behaviour than the Riemann theory. In particular we have the

Theorem A.6 (Monotone Convergence Theorem) *If* $\{f_i | i \in I\}$ *is a sequence of measurable functions with* $\forall i. f_i \leq f_{i+1}$ *and* $f = sup_{i \in I} f_i$ *then*

$$\int f d\mu = \lim_{i \longrightarrow \infty} \int f_i d\mu.$$

There are much stronger convergence theorems available.

A key result in classical measure theory is the Radon-Nikodym theorem. We write $\mu << \nu$ if $\nu(A) = 0 \Rightarrow \mu(A) = 0$ for any measurable set A. We say that μ is *absolutely continuous* with respect to ν. The Radon-Nikodym says that if the measures are finite[5] then there is a measurable function h such that for any measurable set A we have

$$\mu(A) = \int_A h d\nu.$$

Furthermore this function is unique *almost everywhere*; in other words, if there is another function h' with the same property it can only differ from h on a set of ν measure 0. This theorem gives a notion of differentiating one measure by another. One often calls h the *Radon-Nikodym derivative* of μ with respect to ν.

An important application of the Radon-Nikodym theorem is to construct conditional probability distributions. Conditional probabilities relate probabilistic information with definite information and are the key to probabilistic reasoning. In the discrete case the conditional probability can be defined as follows

$$P(A|B) \stackrel{\text{def}}{=} \frac{P(A \cap B)}{P(B)}.$$

This should be read as "the probability of A being true *given that* B is true". Of course, this makes sense only if $P(B) \neq 0$. In the continuous case most of the probabilities are 0, so conditional probabilities have to be defined more subtly.

[5] Actually a much weaker condition suffices but the present result is all one needs for probability.

Suppose that we have a probability space (X, Σ, P). Suppose that Σ' is a sub σ-algebra of Σ. Suppose that A is a measurable set in Σ. We wish to know the following kind of information. We do an experiment and find out in which subset of Σ' a point lies; now we want to know how to estimate probabilities that the point lies in A. We can define a new probability measure Q by $Q(B) = P(A \cap B)$ for $B \in \Sigma'$. Now using the Radon-Nikodym theorem we can define a conditional probability density function $P(A|\Sigma') : X \to [0, 1]$ with the properties:

1. $P(A|\Sigma')(\cdot)$ is measurable with respect to Σ' and integrable as well;
2. for a set B in Σ' we have

$$\int_B P(A|\Sigma')(x)dP(x) = P(A \cap B).$$

This density replaces the naive conditional probabilities of the previous paragraph.

B Proof that Spans of Zigzags Can Be Composed

In this section we give a self-contained proof that we can compose spans of zigzag morphisms. In fact we will prove slightly more. In the category where the morphisms are zigzags we can construct pullbacks.

B.1 Introduction

Consider the following category. Objects are of the form

$$S = (S, \Sigma, k : S \times \Sigma \to [0, 1]),$$

where S is a Polish space, i.e. a topologically complete separable metrizable space, Σ is the Borel σ-algebra on S and k is a transition probability function, i.e. for each $s \in S$, $k(s, -) : \Sigma \to [0, 1]$ is a bounded Borel measure which we assume is extended to all measurable subsets of S. A morphism $f : S \to S'$, where $S = (S, \Sigma, k : S \times \Sigma \to [0, 1])$ and $S' = (S', \Sigma', k' : S' \times \Sigma' \to [0, 1])$, is a continuous surjection which preserves transition probabilities, i.e. it

satisfies, $k(x, f^{-1}(\sigma')) = k'(f(x), \sigma')$ for all $x \in S$ and $\sigma' \in \Sigma'$. It is easy to see that this gives a category. The composition of two morphisms, when it exists, is given by the composition of the two continuous maps. A diagram of morphisms will commute if and only if the corresponding diagram commutes in **Sets**.

Three objects

$$S_1 = (S_1, \Sigma_1, k_1 : S_1 \times \Sigma_1 \to [0,1]),$$

$$S_2 = (S_2, \Sigma_2, k_2 : S_2 \times \Sigma_2 \to [0,1]),$$

$$S = (S, \Sigma, k : S \times \Sigma \to [0,1])$$

and morphisms $f_1 : S_1 \to S$ and $f_2 : S_2 \to S$ are given. We will show that the pullback of f_1 and f_2 exists.

Let $U = \{(s_1, s_2) \in S_1 \times S_2 \mid f_1(s_1) = f_2(s_2)\}$ equipped with the subspace topology of the product topology on $S_1 \times S_2$. The Borel σ-algebra Σ_U on U is generated by the set $\{(\sigma_1 \times \sigma_2) \cap U \mid \sigma_1 \in \Sigma_1, \sigma_2 \in \Sigma_2\}$. Let $\pi_1 : U \to S_1$ and $\pi_2 : U \to S_2$ be the projection maps. Since f_1 and f_2 are surjective, U is not empty. We want to construct $h : U \times \Sigma_U \to [0,1]$ so that (U, Σ_U, h) is an object and $\pi_1 : U \to S_1$ and $\pi_2 : U \to S_2$ are morphisms. For an element x of a set X, an element y of a set Y, subsets $A \subseteq X$, $B \subseteq Y$ and a function $f : X \to Y$, we sometimes write x instead of $\{x\}$, fx instead of $f(x)$, fA instead of $f(A)$, and $f^{-1}B$ instead of $f^{-1}(B)$.

B.2 The conditional probabilities

We fix $(s_1, s_2) \in U$, i.e. $s_1 \in S_1$ and $s_2 \in S_2$ with $f_1 s_1 = f_2 s_2$, throughout this paper. The index i always takes the values 1 and 2. The variable ω_i always runs through S_i whereas the variable s runs through S.

For $i = 1, 2$, we have the probability measures $k_i(s_i, -) : \Sigma_i \to [0,1]$ on the space (S_i, Σ_i). Also $\Sigma_i' = \{f_i^{-1}\sigma \mid \sigma \in \Sigma\} \subseteq \Sigma_i$ is a sub-σ-algebra of Σ_i $(i = 1, 2)$. We therefore have, for a given $\sigma_i \in \Sigma_i$, the conditional probability distribution $P((s_i, \sigma_i)|\Sigma_i') : S_i \to [0,1]$ of the probability measure $k_i(s_i, -)$ given the sub-σ-algebra $\Sigma_i' \subseteq \Sigma_i$. Since S_i is a Polish space we can assume [Dud89, Theorem 10.2.2] that $P((s_i, \sigma_i)|\Sigma_i')(-)$ is a *regular* conditional probability distribution, i.e.

(i) $P((s_i, \sigma_i)|\Sigma_i') : S_i \to [0,1]$ is Σ_i' measurable and integrable.

(ii) For all $\gamma_i' \in \Sigma_i'$, we have

$$\int_{\gamma_i'} P((s_i, \sigma_i)|\Sigma_i')(\omega_i) k_i(s_i, d\omega_i) = k_i(s_i, \sigma_i \cap \gamma_i').$$

(iii) For $k_i(s_i, -)$-almost all $\omega_i \in S_i$, $P((s_i, -)|\Sigma_i')(\omega_i) : \Sigma_i \to [0,1]$ is a probability measure on S_i.

A regular conditional probability distribution, like a conditional probability distribution, is unique up to a set of measure zero, i.e. any two functions satisfying the above three properties are equal for $k_i(s_i, -)$-almost all $\omega_i \in S_i$.

One should think of $P((s_i, \sigma_i), \Sigma_i')(\omega_i)$ as the probability that s_i makes a transition to σ_i given that s_i makes a transition to $f_i^{-1} f_i \omega_i$, or equivalently as the probability that s_i makes a transition to σ_i given that $f_i s_i$ makes a transition to $f_i \omega_i$.

By the standard method in the theory of conditional probability (see for example [Bil79, Exercise 33.5]), we can obtain the conditional probability distribution

$$P((s_i, \sigma_i)|\Sigma) : S \to [0,1].$$

Here, $P((s_i, \sigma_i)|\Sigma)(s)$ gives the probability that s_i makes a transition to σ_i given that s_i makes a transition to some ω_i with $f_i(\omega_i) = s$ or equivalently the probability that s_i makes a transition to σ_i given that $f_i s_i$ makes a transition to s. For the sake of completeness, we explain how this is obtained.

Since $P((s_i, \sigma_i)|\Sigma_i') : S_i \to [0,1]$ is constant on $f_i^{-1} s$ for any $\sigma \in S$ we can define

$$P((s_i, \sigma_i)|\Sigma) : S \to [0,1]$$

by $P((s_i, \sigma_i)|\Sigma)(s) = P((s_i, \sigma_i)|\Sigma_i')(\omega_i)$ for any $\omega_i \in f_i^{-1} s$. Hence, $P((s_i, \sigma_i)|\Sigma) \circ f_i = P((s_i, \sigma_i)|\Sigma_i')$, and $P((s_i, \sigma_i)|\Sigma)$ has the following properties which are equivalent to the above three properties for $P((s_i, \sigma_i)|\Sigma_i')$:

(i) $P((s_i, \sigma_i)|\Sigma)$ is Σ-measurable and integrable.

(ii) Since $k_i(s_i, -) \circ f_i^{-1} = k(f_i s_i, -)$, we have, for all $\sigma \in \Sigma$:

$$\int_{\sigma} P((s_i, \sigma_i)|\Sigma) k(f_i s_i, ds) = \int_{f_i^{-1}\sigma} P((s_i, \sigma_i)|\Sigma_i') k_i(s_i, d\omega_i) = k_i(s_i, \sigma_i \cap f_i^{-1}\sigma).$$

(iii) For $k(f_i s_i, -)$-almost all s, we have a probability measure

$$P((s_i, -)|\Sigma)(s) : \Sigma_i \to [0, 1]$$

on S_i.

As before, $P((s_i, \sigma_i)|\Sigma)$ is unique up to a set of $k(f_i s_i, -)$-measure zero.

Note that one can define

$$P((s_i, \sigma_i)|\Sigma) : S \to [0, 1]$$

directly with the Radon-Nykodym derivative rather than using

$$P((s_i, \sigma_i)|\Sigma_i') : S_i \to [0, 1]$$

.

B.3 Construction of a probability measure on $S_1 \times S_2$

In order to define $h((s_1, s_2), -) : \Sigma_U \to [0, 1]$ we first define a probability measure $g((s_1, s_2), -) : \Sigma_{S_1 \times S_2} \to [0, 1]$ on the product space $(S_1 \times S_2, \Sigma_{S_1 \times S_2})$ where $\Sigma_{S_1 \times S_2}$ is the Borel σ-algebra of $S_1 \times S_2$. We will then show in the next section that $g((s_1, s_2), -)$ is supported on U, i.e. $g((s_1, s_2), U) = 1$. We will finally define $h((s_1, s_2), \alpha \cap U) = g((s_1, s_2), \alpha)$.

Let $\mathcal{R} = \Sigma_1 \times \Sigma_2$ be the collection of all "rectangles" $\sigma_1 \times \sigma_2 \in \Sigma_{S_1 \times S_2}$; $\mathcal{R}$ is a semi-ring. For $\sigma_1 \times \sigma_2 \in \mathcal{R}$ and $s \in S$, put

$$\rho((s_1, s_2), \sigma_1 \times \sigma_2)(s) = P((s_1, \sigma_1)|\Sigma)(s) P((s_2, \sigma_2)|\Sigma)(s).$$

Since for $k(f_1 s_1, -)$-almost all $s \in S$, $P((s_1, -)|\Sigma)(s)$ and $P((s_2, -)|\Sigma)(s)$ are probability measures on S_1 and S_2 respectively, we have:

Proposition B.1 *[Dud89, Theorem 4.4.1]*

For $k(f_1 s_1, -)$-almost all $s \in S$, $\rho((s_1, s_2), -)(s)$ is countably additive on $\mathcal{R}$.

For $\sigma_1 \times \sigma_2 \in \mathcal{R}$, put

$$g((s_1, s_2), \sigma_1 \times \sigma_2) = \int_S P((s_1, \sigma_1)|\Sigma)(s) P((s_2, \sigma_2)|\Sigma)(s) k(f_1 s_1, ds). \qquad (2)$$

Proposition B.2 *$g((s_1, s_2), -)$ has a unique extension to a probability measure on $(S_1 \times S_2, \Sigma_{S_1 \times S_2})$.*

Proof. We check that $g((s_1, s_2), -)$ is countably additive on $\mathcal{R}$. Let $\sigma_1 \times \sigma_2 = \bigcup_{j \in J} \sigma_{j1} \times \sigma_{j2}$ be the disjoint union of a countable set of rectangles. Then,

$$\begin{aligned}
& g((s_1, s_2), \sigma_1 \times \sigma_2) \\
&= \int_S P((s_1, \sigma_1)|\Sigma)(s) P((s_2, \sigma_2)|\Sigma)(s) k(f_1 s_1, ds) \\
&= \int_S \sum_{j \in J} P((s_1, \sigma_{j1})|\Sigma)(s) P((s_2, \sigma_{j2})|\Sigma)(s) k(f_1 s_1, ds) \qquad \text{by Proposition B.1} \\
&= \textstyle\sum_{j \in J} \displaystyle\int_S P((s_1, \sigma_{j1})|\Sigma)(s) P((s_2, \sigma_{j_2})|\Sigma)(s) k(f_1 s_1, ds) \quad \text{monotone convergence} \\
&= \textstyle\sum_{j \in J} g((s_1, s_2), \sigma_{j1} \times \sigma_{j2}).
\end{aligned}$$

This establishes countable additivity. We next extend $g((s_1, s_2), -)$ in the usual way [Dud89, page 102] to the ring $\mathcal{A}$ generated by the rectangles in $\mathcal{R}$, i.e to the collection of all unions of finitely many disjoint elements of $\mathcal{R}$: For a finite set of disjoint rectangles $\sigma_{j1} \times \sigma_{j2}$, $j = 1, \cdots, n$, we put

$$g((s_1, s_2), \bigcup_{1 \le j \le n} \sigma_{j1} \times \sigma_{j2}) = \sum_{1 \le j \le n} g((s_1, s_2), \sigma_{j1} \times \sigma_{j2}).$$

In fact $\mathcal{A}$ is an algebra of subsets of $S_1 \times S_2$ since it contains $S_1 \times S_2$. Furthermore,

$$\begin{aligned}
& g((s_1, s_2), S_1 \times S_2) \\
&= \int_S P((s_1, S_1)|\Sigma)(s) P((s_2, S_2)|\Sigma)(s) k(f_1 s_1, ds) \\
&= \int_S k(f_1 s_1, ds) \qquad \text{since } P((s_2, S_2)|\Sigma)(s) = 1 \text{ a.e.} \\
&= 1.
\end{aligned}$$

Therefore by standard results in measure theory [Dud89, Theorems 3.1.4 and 3.1.10], $g((s_1, s_2), -)$ has a unique extension to $\Sigma_{S_1 \times S_2}$, i.e., the σ-algebra generated by $\mathcal{R}$ or $\mathcal{A}$. ∎

B.4 Construction of $h((s_1, s_2), -)$

We will now show that $g((s_1, s_2), U^c) = 0$, where U^c is the complement of the set U. First note that $U \subseteq S_1 \times S_2$ is a closed subset of the product topology on $S_1 \times S_2$. In fact, the diagonal $D = \{(s, s) \mid s \in S\}$ is a closed subset of $S \times S$ equipped with the product topology and the function

$$(f_1, f_2) : S_1 \times S_2 \to S \times S$$

with $(f_1, f_2)(\omega_1, \omega_2) = (f_1\omega_1, f_2\omega_2)$ is continuous. Hence, U is closed since it is the inverse image of the closed set D under the continuous map (f_1, f_2).

It follows that $(S_1 \times S_2) \setminus U$ is an open subset of $S_1 \times S_2$. Since S_i is a Polish space, it has a countable basis $B_i \subseteq \Omega_i \subseteq \Sigma_i$ of open subsets, where Ω_i is the lattice of open subsets of S_i $(i = 1, 2)$. Then $B_1 \times B_2$ forms a countable basis of open subsets of $S_1 \times S_2$. Therefore, $(S_1 \times S_2) \setminus U$ is the countable union of open rectangles $\sigma_1 \times \sigma_2$ with $\sigma_i \in B_i$ $(i = 1, 2)$. Hence, in order to show that U has full $g((s_1, s_2), -)$-measure, it is sufficient to prove

Proposition B.3 *For any $\sigma_1 \in \Sigma_1$ and $\sigma_2 \in \Sigma_2$ with $U \cap (\sigma_1 \times \sigma_2) = \emptyset$ we have $g((s_1, s_2), \sigma_1 \times \sigma_2) = 0$.*

We prove the above Proposition by using the following technical lemmas. For the time being, we fix $\sigma_1 \in \Sigma$ and $\sigma_2 \in \Sigma_2$. For convenience we use the notation $\mu(-) = k(f_1 s_1, -)$.

Lemma B.4 *The set $f_i\sigma_i$ (and therefore its complement $S_i \setminus f_i\sigma_i$) is μ-measurable $(i = 1, 2)$.*

Proof. This follows from the fact that the continuous image of any Borel subset of a Polish space into any Hausdorff space Y is measurable (in particular) with respect to any Borel measure on Y [Fed69, statement 2.2.13]. ∎

Lemma B.5 *There exists a measurable subset $B_i \subseteq S \setminus f_i\sigma_i$ with $\mu(B_i) = \mu(S \setminus f_i\sigma_i)$ such that for all $s \in B_i$ we have $P((s_i, \sigma_i)|\Sigma)(s) = 0$.*

Proof. Since, by Lemma B.4, $S \setminus f_i\sigma_i$ is μ-measurable, there exists a Borel subset $C_i \in \Sigma$ with $C_i \subseteq S \setminus f_i\sigma_i$ such that $\mu(C_i) = \mu(S \setminus f_i\sigma_i)$. Hence,

$$\begin{aligned}\int_{C_i} P((s_i, \sigma_i)|\Sigma)(s)\mu(ds) &= \int_{C_i} P((s_i, \sigma_i)|\Sigma)(s)k(f_i s_i, ds) \\ &= k_i(s_i, \sigma_i \cap f_i^{-1}C_i) \\ &= 0,\end{aligned}$$

since $(\sigma_i \cap f_i^{-1}C_i) \subseteq (\sigma_i \cap f_i^{-1}(S \setminus f_i\sigma_i)) = \emptyset$. Therefore, there exists a μ-measurable subset B_i with $B_i \subseteq C_i$ and $\mu(B_i) = \mu(C_i) = \mu(S \setminus f_i\sigma_i)$ such that $P((s_i, \sigma_i)|\Sigma)(s) = 0$ for all $s \in B_i$. ∎

Proof of Proposition B.3 Since $(\sigma_1 \times \sigma_2) \cap U = \emptyset$, we have $f_1\sigma_1 \cap f_2\sigma_2 = \emptyset$ or equivalently

$$(S \setminus f_1\sigma_1) \cup (S \setminus f_2\sigma_2) = S.$$

It follows, by Lemma B.5, that $\mu(S \setminus (B_1 \cup B_2)) = 0$. Therefore, for μ-almost all s we have $P((s_1, \sigma_1)|\Sigma)(s) = 0$ or $P((s_2, \sigma_2)|\Sigma)(s) = 0$. In other words, for μ-almost all s, we have $P((s_1, \sigma_1)|\Sigma)(s)P((s_2, \sigma_2)|\Sigma)(s) = 0$. Hence,

$$g((s_1, s_2), \sigma_1 \times \sigma_2) = \int_S P((s_1, \sigma_1)|\Sigma)(s)P((s_2, \sigma_2)|\Sigma)(s)\mu(ds) = 0,$$

as required. ∎

Corollary B.6 $g((s_1, s_2), (S_1 \times S_2) \setminus U) = 0$.

Proof. By Proposition B.3, $g((s_1, s_2), (S_1 \times S_2) \setminus U) = 0$, since $(S_1 \times S_2) \setminus U)$ is the countable union of rectangles disjoint from U. ∎

Any element of Σ_U is of the form $\alpha \cap U$ with $\alpha \in \Sigma_{S_1 \times S_2}$. We now define $h((s_1, s_2), \alpha \cap U) = g((s_1, s_2), \alpha)$.

Theorem B.7 *$(U, \Sigma_U, h : U \times \Sigma_U \to [0,1])$ is an object of the category and $\pi_i : U \to S_i$ are morphisms.*

Proof. Clearly U is a Polish space as it is a closed subset of the Polish space $S_1 \times S_2$. We have already shown that $h((s_1, s_2), -)$ is a probability measure on U for any $(s_1, s_2) \in U$. Finally, for any $\sigma_1 \in \Sigma_1$ we have:

$$
\begin{aligned}
& h((s_1, s_2), \pi_1^{-1}\sigma_1) \\
& = h((s_1, s_2), (\sigma_1 \times S_2) \cap U) \\
& = g((s_1, s_2), \sigma_1 \times S_2) \\
& = \int_S P((s_1, \sigma_1)|\Sigma)(s) P((s_2, S_2)|\Sigma)(s) k(f_1 s_1, ds) \\
& = \int_S P((s_1, \sigma_1)|\Sigma)(s) k(f_1 s_1, ds) \qquad\qquad P((s_2, S_2)|\Sigma)(s) = 1 \text{ a.e.} \\
& = k_1(s_1, f_1^{-1}(s) \cap \sigma_1) = k_1(s_1, \sigma_1).
\end{aligned}
$$

Therefore π_1 is a morphism. Similarly π_2 is a morphism. ∎

Remark B.8 If we are dealing with discrete processes, i.e. if S_1, S_2 and S are countable discrete spaces, then Equation 2 reduces to a sum and we get

$$
\begin{aligned}
h((s_1, s_2), (\sigma_1 \times \sigma_2) \cap U) &= g((s_1, s_2), \sigma_1 \times \sigma_2) \\
&= \sum_{s \in S} P((s_1, \sigma_1)|\Sigma)(s) P((s_2, \sigma_2)|\Sigma)(s) k(f_1 s_1, s).
\end{aligned}
$$

Since, for $k(f_i s_i, s) \neq 0$ we have

$$
P((s_i, \sigma_i)|\Sigma)(s) = \frac{k_i(s_i, \sigma_i)}{k(f_i s_i, s)},
$$

we obtain

$$
P((s_1, \sigma_1)|\Sigma)(s) P((s_2, \sigma_2)|\Sigma)(s) k(f_1 s_1, s) = \begin{cases} \dfrac{k_1(s_1, \sigma_1) k_2(s_2, \sigma_2)}{k(f_1 s_1, s)} & \text{if } k(f_1 s_1, s) \neq 0 \\ 0 & \text{otherwise.} \end{cases}
$$

B.5 The pullback property

We show that $(U, \Sigma_U, h : U \times \Sigma_U \to [0,1])$ and $\pi_i : U \to S_i$ $(i = 1,2)$ give the pullback of $f_i : S_i \to S$ $(i = 1,2)$.

Let $(T, \Sigma_T, r : T \times \Sigma_T \to [0,1])$ be an object and $g_i : T \to S_i$ $(i = 1,2)$ be morphisms such that $f_1 \circ g_1 = f_2 \circ g_2$. Let $t \in T$ and $\sigma_i \in \Sigma_i$ $(i = 1,2)$.

Lemma B.9 *For $k(f_i g_i t, -)$-almost all $s \in S$,*

$$P((g_i t, \sigma_i)|\Sigma)(s) = P((t, g_i^{-1}\sigma_i)|\Sigma)(s).$$

Proof. We have, for any $\sigma \in \Sigma$,

$$\int_\sigma P((g_i t, \sigma_i)|\Sigma)(s) k(f_i g_i t, ds) = k_i(g_i t, \sigma_i \cap f_i^{-1}\sigma),$$

and also

$$\begin{aligned} \int_\sigma P((t, g_i^{-1}\sigma_i)|\Sigma)(s) k(f_i g_i t, ds) &= r(t, (g_i^{-1}\sigma_i) \cap (g_i^{-1} f_i^{-1}\sigma)) \\ &= r(t, g_i^{-1}(\sigma_i \cap f_i^{-1}\sigma)) \\ &= k_i(g_i t, \sigma_i \cap f_i^{-1}\sigma). \end{aligned}$$

It follows that the conditional probability $P((g_i t, \sigma_i)|\Sigma)(-)$ is a version of $P((t, g_i^{-1}\sigma_i)|\Sigma)(-)$ and hence they are equal for $k(f_i g_i t, -)$-almost all $s \in S$. ∎

Theorem B.10 *$(U, \Sigma_U, h(-,-))$ with $\pi_i : U \to S_i$ $(i = 1,2)$ is the pullback of $f_i : S_i \to S$ $(i = 1,2)$.*

Proof. Let $q : T \to U$ be given by $q(t) = (g_1 t, g_2 t)$. We show that $q : (T, \Sigma_T, r) \to (U, \Sigma_U, h)$ is a morphism and is the unique mediating morphism making the pullback diagram commute. To show that q is a morphism, we need to check that

$$r(t, q^{-1}\alpha) = h((g_1 t, g_2 t), \alpha), \tag{3}$$

for all $\alpha \in \Sigma_U$. The sub-σ-algebra $\Sigma'_T = \{q^{-1}\alpha | \alpha \in \Sigma_U\} \subseteq \Sigma_T$ is isomorphic with Σ_U. By Proposition B.2, $h((s_1, s_2), -)$ is uniquely determined by its

values on $\Sigma'_U = \{(\sigma_1 \times \sigma_2) \cap U | \sigma_i \in \Sigma_i\}$. Hence, $r(t,-) : \Sigma'_T \to [0,1]$ is uniquely determined by its values on $q^{-1}(\Sigma'_U)$. It therefore suffices to show that Equation (3) holds for all $\alpha \in \Sigma'_U$. Let $\alpha = (\sigma_1 \times \sigma_2) \cap U$ with $\sigma_1 \in \Sigma_1$ and $\sigma_2 \in \Sigma_2$. Note that

$$q^{-1}(\alpha) = \{t \in T | g_1 t \in \sigma_1, g_2 t \in \sigma_2\} = (g_1^{-1}\sigma_1) \cap (g_2^{-1}\sigma_2).$$

Therefore,

$$\begin{aligned}
& r(t, q^{-1}\alpha) \\
&= r(t, g_1^{-1}\sigma_1 \cap g_2^{-1}\sigma_2) \\
&= \int_S P((t, g_1^{-1}\sigma_1)|\Sigma)(s) P((t, g_2^{-1}\sigma_2)|\Sigma)(s) k(f_1 g_1 t, ds) \\
&= \int_S P((g_1 t, \sigma_1)|\Sigma)(s) P((g_2 t, \sigma_2)|\Sigma)(s) k(f_1 g_1 t, ds) \qquad \text{by Lemma B.9} \\
&= h((g_1 t, g_2 t), \alpha).
\end{aligned}$$

This shows that q is a morphism. Its uniqueness follows from the corresponding uniqueness in **Sets**. Finally, $g_i = \pi_i \circ q$ $(i = 1, 2)$ follows from the corresponding equality in **Sets**. ∎

References

[AHS96] R. Alur, T. Henzinger, and E. Sontag, editors. *Hybrid Systems III*, number 1066 in Lecture Notes in Computer Science. Springer-Verlag, 1996.

[AM89] P. Aczel and N. Mendler. A final-coalgebra theorem. In *Category Theory and Computer Science*, Lecture Notes In Computer Science, pages 357–365, 1989.

[Ash72] R. B. Ash. *Real Analysis and Probability*. Academic Press, 1972.

[Bil79] P. Billingsley. *Probability and Measure*. Wiley-Interscience, 1979.

[BK96] C. Baier and M. Kwiatkowska. Domain equations for probabilistic processes. Available from URL http://www.cs.bham.ac.uk/ mzk/, March 1996.

[BLL+96] J. Bengtsson, K. G. Larsen, F. Larsson, P. Pettersson, and W. Yi. Uppaal: A tool suite for automatic verification of real-time systems. In R. Alur, T. Henzinger, and E. Sontag, editors, *Hybrid Systems III*, number 1066 in Lecture Notes In Computer Science, pages 232–243, 1996.

[Chr90] I. Christoff. Testing equivalences and fully abstract models for probabilistic processes. In *CONCUR 90*, number 458 in Lecture Notes In Computer Science. Springer-Verlag, 1990.

[CM65] D. R. Cox and H. Miller. *An Introduction to the Theory of Stochastic Processes*. Dunno, 1965.

[CN95] A. Cheng and M. Nielsen. Open maps at work. In P. S. Thiagarajan, editor, *Proceedings of the 15th Annual Conference on Foundations of Software Technology and Theoretical Computer Science*, volume 1026 of *Lecture Notes In Computer Science*. Springer-Verlag, 1995.

[CSZ92] R. Cleaveland, S. Smolka, and A. Zwarico. Testing preorders for probabilistic processes. In *Proceedings of the International Colloquium On Automata Languages And Programming 1992*, number 623 in Lecture Notes In Computer Science. Springer-Verlag, 1992.

[Dud89] R. M. Dudley. *Real Analysis and Probability*. Wadsworth and Brookes/Cole, 1989.

[dVR] E. de Vink and J. J. M. M. Rutten. Probabilistic bisimulation for metric transition systems. In preparation.

[Eda96] A. Edalat. The category of markov processes on polish spaces with transition probability preserving continuous surjections has pullbacks. Available from `theory.doc.ic.ac.uk` in the directory `/people/Edalat/markov.ps`, 1996.

[Fed69] H. Federer. *Geometric Measure Theory*. Springer-Verlag, 1969.

[Gir81] Michèle Giry. A categorical approach to probability theory. In B. Banaschewski, editor, *Categorical Aspects of Topology and*

Analysis, number 915 in Lecture Notes In Mathematics, pages 68–85. Springer-Verlag, 1981.

[Hal74] P. Halmos. *Measure Theory.* Number 18 in Graduate Texts in Mathematics. Springer-Verlag, 1974. Originally published in 1950.

[HHWT95] T. Henzinger, P.-H. Ho, and H. Wong-Toi. Hytech: the next generation. In *Proceedings of the 16th Annual Real-time Systems Symposium*, pages 55–65. IEEE Computer Society Press, 1995.

[Hil94] Jane Hillston. *A Compositional Approach to Performance Modelling.* PhD thesis, University of Edinburgh, 1994. To be published as a Distinguished Dissertation by Cambridge University Press.

[HK96] M. Huth and M. Kwiatkowska. On probabilistic model checking. Technical Report CSR-96-15, University of Birmingham, 1996. available from http://www.cs.bham.ac.uk/ mzk/.

[Hoa85] C. A. R. Hoare. *Communicating Sequential Processes.* Series in Computer Science. Prentice-Hall International, London, 1985.

[JL91] B. Jonsson and K. Larsen. Specification and refinement of probabilistic processes. In *Proceedings of the 6th Annual IEEE Symposium On Logic In Computer Science*, 1991.

[JNW96] A. Joyal, M. Nielsen, and G. Winskel. Bisimulation from open maps. *Information and Computation*, 127(2):164–185, 1996.

[JY95] B. Jonsson and W. Yi. Compositional testing preorders for probabilistic processes. In *Proceedings of the 10th Annual IEEE Symposium On Logic In Computer Science*, 1995.

[Koz81] D. Kozen. Semantics of probabilistic programs. *Journal of Computer and Systems Sciences*, 22:328–350, 1981.

[Koz85] D. Kozen. A probabilistic PDL. *Journal of Computer and Systems Sciences*, 30(2):162–178, 1985.

[KT66] J. F. C. Kingman and S. J. Taylor. *Introduction to Measure and Probability.* Cambridge University Press, 1966.

[LS91] K. G. Larsen and A. Skou. Bisimulation through probablistic testing. *Information and Computation*, 94:1–28, 1991.

[Mil80] R. Milner. *A Calculus for Communicating Systems*, volume 92 of *Lecture Notes in Computer Science*. Springer-Verlag, 1980.

[Mil89] R. Milner. *Communication and Concurrency*. Prentice-Hall, 1989.

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

[Pop94] Sally Popkorn. *First Steps in Modal Logic*. Cambridge University Press, 1994.

[Pro] Probabilistic systems group, collected reports. Available from `www.comlab.ox.ac.uk` in the directory `/oucl/groups/probs/bibliography.html`.

[Rud66] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, 1966.

[Rut95] J. J. M. M. Rutten. A calculus of transition systems (towards universal coalgebra). In A. Ponse, M. de Rijke, and Y. Venema, editors, *Modal Logic and Process Algebra, a bisimulation perspective*, number 53 in CSLI Lecture Notes, 1995. Available electronically from `www.cwi.nl/~janr`.

[SD78] N. Saheb-Djahromi. Probabilistic LCF. In *Mathematical Foundations Of Computer Science*, number 64 in Lecture Notes In Computer Science. Springer-Verlag, 1978.

[SD80] N. Saheb-Djahromi. Cpos of measures for nondeterminism. *Theoretical Computer Science*, 12(1):19–37, 1980.

[Sif77] J. Sifakis. Use of petri nets for performance evaluation. In H. Beilner and E. Gelenbe, editors, *Measuring, Modeling and Evaluating Computer Systems*, pages 75–93. North-Holland, 1977.

[SL94] R. Segala and N. Lynch. Probabilistic simulations for probabilistic processes. In B. Jonsson and J. Parrow, editors, *Proceedings*

of CONCUR94, number 836 in Lecture Notes In Computer Science, pages 481–496. Springer-Verlag, 1994.

[vGSST90] R. van Glabbeek, S. Smolka, B. Steffen, and C. Tofts. Reactive generative and stratified models for probabilistic processes. In *Proceedings of the 5th Annual IEEE Symposium On Logic In Computer Science*, 1990.

[Win96] G. Winskel. Unpublished notes. Private Communication, 1996.

[WN95] G. Winskel and M. Nielsen. *Handbook of Logic in Computer Science*, volume 4, chapter Models for Concurrency. Oxford University Press, 1995.

Recent BRICS Report Series Publications

RS-97-4 **Richard Blute, Josée Desharnais, Abbas Edalat, and Prakash Panangaden.** ***Bisimulation for Labelled Markov Processes*****. March 1997. 48 pp. To appear in** ***Twelfth Annual IEEE Symposium on Logic in Computer Science*****, LICS '97 Proceedings.**

RS-97-3 **Carsten Butz and Ieke Moerdijk.** ***A Definability Theorem for First Order Logic*****. March 1997. 10 pp.**

RS-97-2 **David A. Schmidt.** ***Abstract Interpretation in the Operational Semantics Hierarchy*****. March 1997. 33 pp.**

RS-97-1 **Olivier Danvy and Mayer Goldberg.** ***Partial Evaluation of the Euclidian Algorithm (Extended Version)*****. January 1997. 16 pp. To appear in the journal** ***Lisp and Symbolic Computation*****.**

RS-96-62 **P. S. Thiagarajan and Igor Walukiewicz.** ***An Expressively Complete Linear Time Temporal Logic for Mazurkiewicz Traces*****. December 1996. i+13 pp. To appear in** ***Twelfth Annual IEEE Symposium on Logic in Computer Science*****, LICS '97 Proceedings.**

RS-96-61 **Sergei Soloviev.** ***Proof of a Conjecture of S. Mac Lane*****. December 1996. 53 pp. Extended abstract appears in Pitt, Rydeheard and Johnstone, editors,** ***Category Theory and Computer Science: 6th International Conference*****, CTCS '95 Proceedings, LNCS 953, 1995, pages 59–80.**

RS-96-60 **Johan Bengtsson, Kim G. Larsen, Fredrik Larsson, Paul Pettersson, and Wang Yi.** UPPAAL ***in 1995*****. December 1996. 5 pp. Appears in Margaria and Steffen, editors,** ***Tools and Algorithms for The Construction and Analysis of Systems: 2nd International Workshop*****, TACAS '96 Proceedings, LNCS 1055, 1996, pages 431–434.**

RS-96-59 **Kim G. Larsen, Paul Pettersson, and Wang Yi.** ***Compositional and Symbolic Model-Checking of Real-Time Systems*****. December 1996. 12 pp. Appears in** ***16th IEEE Real-Time Systems Symposium*****, RTSS '95 Proceedings, 1995.**

RS-96-58 **Johan Bengtsson, Kim G. Larsen, Fredrik Larsson, Paul Pettersson, and Wang Yi.** UPPAAL **—** ***a Tool Suite for Automatic Verification of Real–Time Systems*****. December 1996. 12 pp. Appears in Alur, Henzinger and Sontag, editors,** ***DIMACS***