Stochastic Transition Systems for Continuous State Spaces and Non-determinism*

Stefano Cattani[1], Roberto Segala[2], Marta Kwiatkowska[1], and Gethin Norman[1]

[1] School of Computer Science, The University of Birmingham
Birmingham B15 2TT, United Kingdom
{stc,mzk,gxn}@cs.bham.ac.uk
[2] Dipartimento di Informatica, Università di Verona
Strada Le Grazie 15, Ca' Vignal 2 37134 Verona, Italy
roberto.segala@univr.it

Abstract. We study the interaction between non-deterministic and probabilistic behaviour in systems with continuous state spaces, arbitrary probability distributions and uncountable branching. Models of such systems have been proposed previously. Here, we introduce a model that extends probabilistic automata to the continuous setting. We identify the class of schedulers that ensures measurability properties on executions, and show that such measurability properties are preserved by parallel composition. Finally, we demonstrate how these results allow us to define an alternative notion of weak bisimulation in our model.

1 Introduction

Current trends in ubiquitous computing, such as mobility, portability, sensor and wireless ad hoc networks, place an increasing emphasis on the need to model and analyse complex stochastic behaviours. For example, network traffic demands continuously distributed durations, sensors may generate real-valued data, and the geographical mobility of agents typically involves movement in space and time with stochastic trajectories. The presence of the distributed computation scenario creates a requirement to model non-determinism, in addition to such stochastic features.

Several models capable of representing probabilistic behaviour have been proposed in the literature, see e.g. [1, 10, 13, 16, 24]. Particular attention has been paid to the nature of interaction between probabilistic and non-deterministic behaviour; though these can be seen as orthogonal, the way they interact in the model has led to fundamental distinctions. In the discrete state, discrete time model different variants have been proposed. In some models randomisation replaces non-determinism [10], while elsewhere [11] states are either probabilistic or non-deterministic, such that probabilistic and deterministic choices alternate. Furthermore, one can replace conventional transitions with probabilistic transitions (transitions whose target is a distribution over states); in the resulting

* Supported in part by EPSRC grants GR/N22960, GR/S46727 and GR/S11107, MURST project CoVer and FIRB project SPY-Mod.

V. Sassone (Ed.), Proc. Foundations of Software Science and Computation Structures (FOSSCAS'05), LNCS, 2005, to appear.

model of *probabilistic automata* [21, 22], both non-deterministic and probabilistic choices are present at each step. Each of these variants can be endowed with appropriate relations, e.g. bisimulation, simulation or trace equivalence relations.

More recently, the analysis of probabilistic systems has been extended to continuous spaces. Such models can represent systems whose progress, for instance, depends on continuously distributed real-time or geographical position information. Stochastic process algebras [13] are an extension of process algebras in which delays are distributed according to some probability distribution over the reals. Initially, only exponential distributions were considered, as they are easier to handle because of their memoryless property. The usual convention is to replace non-determinism with race condition, leading to continuous time Markov chains, but non-determinism can be kept (e.g. interactive Markov chains [12]). Generally distributed delays have also been introduced, both in the case in which non-determinism is replaced with race conditions (e.g. generalised semi-Markov processes), and in the case in which it is retained [7, 4]. Labelled Markov processes [8] are extensions of transition systems to continuous state spaces and general distributions, but have no non-determinism, in the sense that the choice of action determines the next transition. Operational models with non-determinism have already been proposed, e.g. [7, 4, 5], and the notions of bisimulation and parallel composition have been studied for such systems; [4] also defines weak bisimulation.

When considering continuous distributions and state spaces, the notion of measurability of executions plays an important role, and a departure from pointwise consideration of behaviour is needed since the probability of reaching an individual state is often 0. This leads to the central topic of this paper: here, we investigate the measurability issues that arise from the interaction between non-determinism and continuous state spaces. By allowing the most general setting, the behaviour of a system can become mathematically intractable when studying the properties of a system over several steps of execution. We introduce a model for continuous states spaces, called *stochastic transition systems*, which can be seen as an extension of probabilistic automata to a fully continuous setting: both the set of states and the set of action labels can be continuous. This model also encapsulates labelled Markov processes by the addition of non-determinism, and it can serve as an operational model for stochastic process algebras, since states can record the passage through 3D space and/or time, and labels can include real-valued delays, as well as discrete actions.

As in the discrete case, we use the notion of scheduler as the entity that resolves non-determinism. The power of schedulers has to be restricted since arbitrary schedulers could generate executions that are not tractable from a mathematical point of view. For this reason, we define the class of *measurable* schedulers and show that it identifies the set of schedulers that generate all and only the "good" executions that are measurable. Under this restriction, we can define a probability measure on executions, thus enabling us to reason about global properties of a run of a system.

We also introduce the notion of parallel composition for stochastic transition systems and show that our measurability properties are compositional: if there exists a "good" scheduler for the composition, then there must exist two "good" schedulers on the components that give rise to the same behaviour. This property is important because it both allows for compositional reasoning and serves as a sanity check on the correctness of the definition of measurable schedulers that we have given. As a final remark, we show how we can define the notion of weak transitions for stochastic transition systems. The measurability conditions introduced in the paper are needed to define the target probability of several steps of silent transitions. Based on such transitions, we also give an alternative notion of weak bisimulation for our model.

The main contribution of this paper is the study of measurability properties of stochastic transition systems with non-determinism and continuous state spaces. We identify the class of measurable schedulers that generate tractable runs, confirming the choice originally made in [15]; this restriction enables the definition of a measure on executions. We also show that such measurability properties are preserved through parallel composition.

Structure of the paper. In Section 2 we review the basic notions of measure theory used in this paper. Section 3 introduces the model of stochastic transition systems, and in Section 4 we study the class of schedulers that guarantees measurability of executions. Section 5 introduces a CSP-style parallel operator and analyses the compositionality properties of stochastic transition systems. In Section 6 weak transitions and weak bisimulation are defined. Finally, Section 7 discusses possible future work.

2 Preliminaries

In this section we review the basic definitions and results of measure theory that are necessary for the remainder of the paper. A basic knowledge of topology and metric spaces is assumed. Most results can be found in standard textbooks, e.g. [2]; [18] serves as a good introduction to measure theory.

Basic definitions. Given a set X, an *algebra* over X is a family F_X of subsets of X that includes X and is closed under complementation and finite union; $\mathcal{F}_X$ is a *σ-algebra* over X if we additionally require closure under countable union. A *measurable space* is a pair $(X, \mathcal{F}_X)$, where $\mathcal{F}_X$ is a σ-algebra over X. The elements of $\mathcal{F}_X$ are called *measurable sets*. We abuse the notation and refer to X as a measurable space whenever the corresponding σ-algebra is clear from the context. The σ-algebra generated by a family G of subsets of X is the *smallest* σ-algebra including G. The *product space* of two measurable spaces $(X, \mathcal{F}_X)$ and $(Y, \mathcal{F}_Y)$ is the measurable space $(X \times Y, \mathcal{F}_X \otimes \mathcal{F}_Y)$, where $\mathcal{F}_X \otimes \mathcal{F}_Y$ is the σ-algebra generated by the *rectangles* $A \times B = \{(x, y) \mid x \in A, y \in B\}$, for all $A \in \mathcal{F}_X$ and $B \in \mathcal{F}_Y$; we alternatively denote $\mathcal{F}_X \otimes \mathcal{F}_Y$ by $\mathcal{F}_{X \times Y}$. The union of two measurable spaces is the measurable space $(X \cup Y, \mathcal{F}_{X \cup Y})$, where $\mathcal{F}_{X \cup Y}$ is the σ-algebra generated by the union of $\mathcal{F}_X$ and $\mathcal{F}_Y$. The *Borel* σ-algebra for a topological space (X, T) is the σ-algebra generated by the open sets and is denoted by $\mathcal{B}(X)$.

Given a measurable space $(X, \mathcal{F}_X)$, a *measure* over $(X, \mathcal{F}_X)$ is a function $\mu : \mathcal{F}_X \to \mathbb{R}^{\geq 0}$ such that $\mu(\emptyset) = 0$ and, for every countable family of pairwise disjoint measurable sets $\{A_i\}_{i \in I}$, $\mu(\cup_{i \in I} A_i) = \sum_{i \in I} \mu(A_i)$; the triple $(X, \mathcal{F}_X, \mu)$ is called a *measure space*. A *probability* (resp., *sub-probability*) measure μ over $(X, \mathcal{F}_X)$ is a measure such that $\mu(X) = 1$ (resp., $\mu(X) \leq 1$). A measurable set whose complement has probability 0 is called a *support* for a measure μ. If μ is a (sub-)probability measure, $(X, \mathcal{F}_X, \mu)$ is called a *(sub-)probability space*. We denote the set of probability (resp., sub-probability) measures over $(X, \mathcal{F}_X)$ by $\mathcal{D}(X, \mathcal{F}_X)$ (resp, $sub\mathcal{D}(X, \mathcal{F}_X)$). The product probability space for two probability spaces $(X, \mathcal{F}_X, \mu_X)$ and $(Y, \mathcal{F}_Y, \mu_Y)$, is $(X \times Y, \mathcal{F}_X \otimes \mathcal{F}_Y, \mu_X \otimes \mu_Y)$, where $\mu_X \otimes \mu_Y$ is the unique probability measure such that $(\mu_X \otimes \mu_Y)(A \times B) = \mu_X(A) \cdot \mu_Y(B)$, for all $A \in \mathcal{F}_X$ and $B \in \mathcal{F}_Y$.

A function $f : (X, \mathcal{F}_X) \to (Y, \mathcal{F}_Y)$ is *measurable* if the pre-image of every measurable set is measurable, that is, if $f^{-1}(B) = \{x \in X \mid f(x) \in B\} \in \mathcal{F}_X$ for all $B \in \mathcal{F}_Y$. Given a measurable space $(X, \mathcal{F}_X)$, the indicator function for a measurable set $A \in \mathcal{F}_X$ is the measurable function $I_A(x) = 1$ if $x \in A$, 0 otherwise. Let $(X, \mathcal{F}_X, \mu)$ be a probability space, $(Y, \mathcal{F}_Y)$ a measurable space and f a measurable function from X to Y. The induced probability measure for f over $(Y, \mathcal{F}_Y)$ is given by $f(\mu)$ defined as $f(\mu)(B) = \mu(f^{-1}(B))$ for all $B \in \mathcal{F}_Y$.

We call a family S of subsets of a set X a *semi-ring* if S includes $\emptyset$, is closed under finite intersection, and if, whenever $A, B \in S$, there exists a finite family $\{A_i\}_{i \in \{0 \dots n\}}$ of pairwise disjoint elements of S such that $A \setminus B = \cup_{i=0}^{n} A_i$.

Theorem 1. *Every sub-probability measure defined over a semi-ring S can be uniquely extended to a sub-probability measure over the σ-algebra generated by S.*

Theorem 2. *Let $(X, \mathcal{F}_X)$ and $(Y, \mathcal{F}_Y)$ be two measurable spaces and f a real-valued nonnegative measurable function on $X \times Y$. Assume we have a function $\nu : Y \times \mathcal{F}_X \to \mathbb{R}^{\geq 0}$ such that $\nu(y, \cdot)$ is a measure on $(X, \mathcal{F}_X)$ for all $y \in Y$ and $\nu(\cdot, A)$ is measurable for all $A \in \mathcal{F}_X$. Then $\int_X f(x, y)\nu(y, dx)$ exists and is a measurable function of Y.*

Regular conditional probabilities. As we will demonstrate later, our construction will require conditional probabilities. In the discrete case, we can define the probability of an event A given B as $P(A|B) = P(A \cap B)/P(B)$, which is defined only when $P(B) > 0$. Unfortunately, this cannot be done in general for the continuous case, as it is still meaningful to condition with respect to events of probability 0. Consider for example a measure defined on $\mathbb{R}^2$; even if the probability of a given x can be zero, it can be interesting to study the probability measure on $\mathbb{R}$ for such given x. It is therefore necessary to extend the concept of conditional probabilities.

Definition 1. *Let $(X, \mathcal{F}_X, \mu)$ be a probability space, $(Y, \mathcal{F}_Y)$ a measurable space and $f : X \to Y$ a measurable function. A* **regular conditional probability** *for μ with respect to f is a function $\nu : Y \times \mathcal{F}_X \to [0, 1]$ such that:*

1. *$\nu(y, \cdot)$ is a probability measure on $\mathcal{F}_X$, for each $y \in Y$;*
2. *$\nu(\cdot, A)$ is a measurable function on $(Y, \mathcal{F}_Y)$, for each $A \in \mathcal{F}_X$;*
3. *$\mu(A \cap f^{-1}(B)) = \int_B \nu(y, A)\, f(\mu)(dy)$.*

Regular conditional probabilities do not exist for all probability spaces. It is necessary to impose restrictions on the kind of measurable spaces we consider. A *Polish space* is the topological space underlying a complete separable metric space. Given a Polish space X, $(X, \mathcal{F}_X)$ is a standard Borel space if $\mathcal{F}_X$ is the Borel σ-algebra generated by the topology. Finally, given a standard Borel space $(X, \mathcal{F}_X)$, $Y \subseteq X$ is an *analytic set* if it is the continuous image of some Polish space. The space $(Y, \mathcal{F}_Y)$ is an *analytic space* if it is measurably isomorphic to an analytic set in a Polish space, that is, if there exists a measurable bijection whose inverse is also measurable. Note that singleton sets are measurable in Polish and analytic spaces. Examples of analytic sets are the discrete spaces and any open or closed subset of the reals equipped with the Borel σ-algebra. Analytic sets are closed under union and Cartesian product. Thus, analytic sets are quite general; for instance, the semantic model of timed systems is given by the product of a discrete set (the graph-theoretic representation of a system) and the possible values of time (the real numbers).

Theorem 3. *Let $(Y, \mathcal{F}_Y)$be an analytic spac and $f : (X, \mathcal{F}_X, \mu) \to (Y, \mathcal{F}_Y)$ a measurable function. Then there exists a regular conditional probability ν for f.*

A σ-algebra on probability measures. In the following, we define probability distributions on sets of probabilistic transitions whose targets are probability measures on states. We therefore need to define a σ-algebra on sets of probability measures; we use the standard construction, due to Giry [9]. Let $(X, \mathcal{F}_X)$ be a measurable set and $\mathcal{D}(X, \mathcal{F}_X)$ the set of probability measures on X. We build a σ-algebra on the set of probability measures $\mathcal{D}(X, \mathcal{F}_X)$ as follows: for each $A \in \mathcal{F}_X$, define a function $p_A : \mathcal{D}(X, \mathcal{F}_X) \to [0, 1]$ by $p_A(\nu) = \nu(A)$. The σ-algebra on $\mathcal{D}(X, \mathcal{F}_X)$, denoted by $\mathcal{F}_{\mathcal{D}(X,\mathcal{F}_X)}$ is the *least* σ-algebra such that all the p_A's are measurable. The generators of the σ-algebra are the sets of probability measures $D_{A,I} = p_A^{-1}(I) = \{\mu \in \mathcal{D}(X, \mathcal{F}_X) \mid \mu(A) \in I\}$, for all $A \in \mathcal{F}_X$ and $I \in \mathcal{B}([0, 1])$.

3 Stochastic Transition Systems

In this section we introduce our model, called *stochastic transition systems*, which features both non-deterministic and probabilistic behaviour. The model can be seen as an extension of probabilistic automata [21] to continuous state and label spaces and to continuous probability measures. Stochastic transition systems are fully non-deterministic, and thus also generalise labelled Markov processes [8].

In this section we introduce the fundamental concepts of our continuous model, most of which are an adaptation of [21] to the continuous setting.

Definition 2. *A* **stochastic transition system** *(STS) S is a tuple $((Q, \mathcal{F}_Q), \overline{q}, (L, \mathcal{F}_L), \rightarrow)$, where*

- *$(Q, \mathcal{F}_Q)$ is the analytic space of states;*
- *$\overline{q} \in Q$ is the initial state;*
- *$(L, \mathcal{F}_L)$ is the analytic space of labels;*
- *$\rightarrow \subseteq Q \times L \times \mathcal{D}(Q, \mathcal{F}_Q)$ is the set of probabilistic transitions.*

We say that a transition (q, a, μ) is labelled by a and enabled from q, and denote it by $q \xrightarrow{a} \mu$; transitions are ranged over by t. We denote the set of possible transitions by $\mathcal{T} = Q \times L \times \mathcal{D}(Q, \mathcal{F}_Q)$ and define a σ-algebra on it as the product of the σ-algebras of the components, that is, $\mathcal{F}_\mathcal{T} = \mathcal{F}_Q \otimes \mathcal{F}_L \otimes \mathcal{F}_{\mathcal{D}(Q, \mathcal{F}_Q)}$. The set of transitions enabled from a state q is denoted by $\mathcal{T}(q) = \{(q', a, \mu) \in \rightarrow \mid q = q'\}$. We denote the elements of an STS S by Q, $\mathcal{F}_Q$, $\overline{q}$, L, $\mathcal{F}_L$ and $\rightarrow$ and we propagate indices when necessary; thus, the elements of S_i are Q_i, $\mathcal{F}_{Q_i}$, $\overline{q}_i$, L_i, $\mathcal{F}_{L_i}$ and $\rightarrow_i$.

Combined transitions. Following [21], since we resolve non-determinism in a randomised way, we combine the transitions leaving a state q in order to obtain a new transition. Similarly to the discrete case, this induces a probability measure on the set of transitions leaving state q, that is, a measure π on $\mathcal{T}$ with a support contained in $\mathcal{T}(q)$. Since different transitions have in general different labels, the combination of the transitions leaving a state results in a new distribution on both labels and target states.

Definition 3. *Given a state q and a sub-probability measure π over $(\mathcal{T}, \mathcal{F}_\mathcal{T})$ with a support contained in $\mathcal{T}(q)$, the* **combined transition for π from** *q is the pair (q, μ_π) (denoted by $q \rightarrow \mu_\pi$), where μ_π is the sub-probability measure over $(L \times Q, \mathcal{F}_L \otimes \mathcal{F}_Q)$ defined as follows:*

$$\mu_\pi(A \times X) = \int_{(q,a,\mu) \in \mathcal{T}} I_A(a) \mu(X) d\pi$$

The integral above is well defined for the σ-algebra $\mathcal{F}_\mathcal{T}$ on transitions. It is easy to show that μ_π is a sub-probability measure. Observe that we require π to be a sub-probability measure, therefore it is possible that no transition is scheduled with positive probability. We let this denote the probability to stop, which is defined as $\mu_\pi(\perp) = 1 - \mu_\pi(L \times Q)$.

Executions. Given an STS S, a possibly infinite alternating sequence of states and actions $\alpha = q_0 a_1 q_1 \cdots$ is called an *execution*. We denote the set of executions by *Exec*, the set of finite executions ending with a state by $Exec^*$ and the set of infinite executions by $Exec^\omega$. Given a finite execution α, $\alpha[\downarrow]$ denotes its last state. The *length* of an execution α, denoted by $|\alpha|$, is the number of occurrences of actions in α; if α is infinite $|\alpha| = \infty$. We denote a finite execution α that has terminated by $\alpha\perp$, where an execution α terminates if $\perp$ is scheduled from the last state of α.

A σ-algebra on executions. We define the σ-algebra $\mathcal{F}_{Exec}$ over the set of executions. This is necessary to study the properties of system runs. In the discrete case, $\mathcal{F}_{Exec}$ is the σ-algebra generated by cones, that is, the set of executions that extend some finite prefix. This concept is generalised to the continuous case by using sets of executions called basic sets. Formally, given a non empty finite sequence of measurable sets $\Lambda = X_0 A_1 X_1 \cdots A_n X_n$, $A_i \in \mathcal{F}_L$, $i \in \{1..n\}$, and $X_i \in \mathcal{F}_Q$, $i \in \{0..n\}$, the *basic set with base* Λ is defined as:

$$\mathcal{C}_\Lambda = \{q_0 a_1 \cdots q_n \alpha \mid \forall i \in \{0..n\}\ q_i \in X_i \text{ and } \forall i \in \{1..n\}\ a_i \in A_i \text{ and } \alpha \in Exec\}$$

The length of a basic set $\mathcal{C}_\Lambda$ is given by the number of occurrences of elements of $\mathcal{F}_L$ in Λ. Observe that basic sets form a semi-ring. $\mathcal{F}_{Exec}$ is the σ-algebra generated by basic sets.

We define the σ-algebra $\mathcal{F}_{Exec^*}$ on finite executions in a similar way as the σ-algebra generated by the sets of the form $Q_0 A_1 \cdots Q_n = \{\alpha = q_0 a_1 \cdots q_n \mid q_i \in Q_i \text{ for all } i \in \{0 \cdots n\} \text{ and } a_j \in A_j \text{ for all } j \in \{1 \cdots n\}\}$, where $Q_0 \ldots Q_n \in \mathcal{F}_Q$ and $A_1 \ldots A_n \in \mathcal{F}_L$. $(Exec^*, \mathcal{F}_{Exec^*})$ is the measurable set of finite executions. Note that $\mathcal{F}_{Exec^*}$ is the restriction of $\mathcal{F}_{Exec}$ to finite executions.

Schedulers. We use schedulers as the entities that resolve non-determinism. Given a history in the form of a sequence of states and labels that the system has visited, a scheduler chooses the next transition from the current state by assigning a sub-probability measure to the enabled transitions.

Definition 4. *A* **scheduler** *is a function* $\eta : Exec^* \to sub\mathcal{D}(\mathcal{T})$*, such that, for all* $\alpha \in Exec^*$*,* $\mathcal{T}(\alpha[\downarrow])$ *is a support for* $\eta(\alpha)$*.*

We denote the set of schedulers by $\mathcal{A}$. Since a scheduler η returns a distribution on transitions for each finite execution α, it induces a combined transition $(\alpha[\downarrow], \mu_{\eta(\alpha)})$ leaving the last state of each execution. Note that we use *randomised* schedulers; originally introduced for discrete systems, they have been shown to have important properties, for example the probabilistic temporal logic PCTL is preserved by bisimulation under randomised schedulers [22]. Randomised schedulers are also necessary to obtain compositionality under parallel compositions (see Section 5). Non-randomised (deterministic) schedulers can be seen as the subclass of schedulers that return a Dirac distribution after each execution.

According to the above definition, a scheduler can make arbitrary choices at each point of the computation. We define a class of schedulers whose global behaviour respects measurability properties.

Definition 5. *A scheduler* η *is measurable if the function* $f_\eta(\alpha) = \mu_{\eta(\alpha)}$ *(called the* flattening *of* η*) is a measurable function from* $(Exec^*, \mathcal{F}_{Exec^*})$ *to* $(sub\mathcal{D}(L \times Q), \mathcal{F}_{sub\mathcal{D}(L\times Q)})$*. We denote the class of measurable schedulers by* $\mathcal{A}_{meas}$*.*

Probabilistic executions. The interaction of an STS S and a scheduler η results in a system with no non-determinism, i.e. a purely probabilistic process. We call this object a *probabilistic execution* following [21].

Definition 6. *Given an STS S and a scheduler η, the* **probabilistic execution** $P_{S,\eta}$ *for S and η is the tuple $(Exec^*, \mathcal{F}_{L\times Q}, \mu)$, where $\mu : Exec^* \times \mathcal{F}_{L\times Q} \to [0,1]$ such that for each $\alpha \in Exec^*$ $\mu(\alpha, \cdot)$ is a sub-probability measure over $L \times Q$ defined by $\mu_{\eta(\alpha)}$.*

A probabilistic execution defines the transitions induced by the scheduler η: given a finite execution it returns the combined transition scheduled by η. We write μ_X, $X \in \mathcal{F}_{L\times Q}$, whenever we fix X and μ is a function on *Exec**. Similarly, we write μ_α (or $\mu_{\eta(\alpha)}$) whenever we fix α and μ is a measure on $\mathcal{F}_{L\times Q}$.

Not all probabilistic executions are "good"; our objective is to define a measure on executions, essential to define weak transitions, and the measurability of the function μ is necessary for this purpose. In the purely probabilistic case (no non-determinism) this problem is solved by using Markov kernels (e.g. [8]). We adapt this idea to our setting by defining measurable probabilistic executions and by studying the conditions under which they are generated.

Definition 7. *A probabilistic execution $P_{S,\eta} = (Exec^*, \mathcal{F}_{L\times Q}, \mu)$ is measurable if $\mu : Exec^* \times \mathcal{F}_{L\times Q} \to [0,1]$ is such that $\mu(\cdot, X)$ is a measurable function for each $X \in \mathcal{F}_{L\times Q}$.*

When μ has such a measurability property, we can see it as a generalisation of Markov kernels to the history-dependent case.

Related models. As stated above, stochastic transition systems are an extension of probabilistic automata to the continuous case; the latter correspond to the subset of STSs with discrete σ-algebras on states. Labelled Markov processes (LMPs) [8] correspond to the case where there is no non-determinism on actions, that is, for every action there is exactly one distribution from each state. The measurability problem is solved in LMPs by using, for each action, a Markov kernel to denote the probability transition function. An extension of probabilistic automata with continuous distributions and real-valued time labels is proposed in [5], but measurability properties are not considered. Similar models for the continuous setting can be found in [7], proposing an alternating model, where states are either non-deterministic or probabilistic, and in [4], proposing a model where each state enables one probabilistic distribution or arbitrarily many transitions labelled with actions or time. Again, neither of these papers considers the problem of measurability of executions.

4 Measurability and Schedulers

We aim to extend the results of the discrete case to stochastic transition systems and define the measure on executions induced by a scheduler. This requires the corresponding probabilistic execution to be measurable. In this section we show that the class of measurable schedulers identifies all and only the measurable probabilistic executions. The following example shows that arbitrary schedulers could produce "bad" executions and explains why considering the point-wise

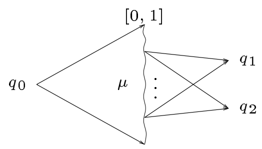

Fig. 1. A simple stochastic transition system illustrating the need for measurable schedulers.

behaviour of a scheduler is not enough in the continuous setting; instead, it is necessary to consider its global behaviour.

Example 1. Consider the system of Figure 1: the initial state q_0 enables a single transition with some measure μ on the interval $[0,1]$. From each state in $[0,1]$ two Dirac transitions are enabled: one to q_1 and the other to q_2. Labels are not relevant. The probability of moving to q_1 after two steps under a scheduler η is given by $\int_{[0,1]} \mu_{\eta(q)}(\{q_1\})\mu(dq)$, that is, the probability of reaching any state q multiplied by the probability of reaching q_1 from each q. Let η be the scheduler that chooses q_1 from a non-measurable subset A of $[0,1]$, and q_2 from its complement. The integral above is not defined as $\mu_{\eta(q)}(\{q_1\})$ is not a measurable function, that is, the probabilistic execution is not measurable. We want to rule out such a probabilistic execution as "pathological" and disallow the scheduler generating it.

4.1 Measurable Schedulers and Probabilistic Executions

We restrict our analysis to measurable executions only, as they represent "well behaved", feasible, schedulers and allow us to define probability measure on paths. We think this is not an unreasonable restriction since schedulers that produce non-measurable executions represent pathological cases and thus can be discarded. A similar approach has been adopted in [15], where only the schedulers that preserve the measurability of logical formulae are considered, though without studying the nature of such schedulers.

Proposition 1. *Given an STS S, and a scheduler η, η is measurable if and only if $P_{S,\eta}$ is measurable.*

Proof outline. We prove the two directions:

– *If:* Let f_η be the flattening of η as in Definition 5. We have to show that $f_\eta^{-1}(D) \in \mathcal{F}_{Exec^*}$ for all $D \in \mathcal{F}_{sub\mathcal{D}(L\times Q)}$. Firstly, we prove it for the generators $D_{X,I}$ of $\mathcal{F}_{sub\mathcal{D}(L\times Q)}$, for all $X \in \mathcal{F}_{L\times Q}$ and $I \in \mathcal{B}([0,1])$. Consider one such $D_{X,I}$. Since $P_{S,\eta}$ is measurable, we get $\mu_X^{-1}(I) = Y \in \mathcal{F}_{Exec^*}$ by hypothesis. We show that $Y = f_\eta^{-1}(D_{X,I})$:
  - $f_\eta^{-1}(D_{X,I}) \supseteq Y$: consider $\alpha \in Y$, then $\mu(\alpha, X) \in I$; this is equivalent to $\mu_{\eta(\alpha)}(X) \in I$, which implies $\mu_{\eta(\alpha)}(X) \in D_{X,I}$. It follows that $\alpha \in f_\eta^{-1}(D_{X,I})$.

- $f_\eta^{-1}(D_{X,I}) \subseteq Y$: consider $\alpha \in f_\eta^{-1}(D_{X,I})$. Then $\mu_{\eta(\alpha)}(X) \in I$, that is, $\mu(\alpha, X) \in I$. This, of course, means that $\alpha \in \mu_X^{-1}(I) = Y$.

  The result is extended to the σ-algebra $\mathcal{F}_{sub\mathcal{D}(L\times Q)}$ by standard arguments.
- *Only if:* consider $P_{S,\eta} = (\mathit{Exec}^*, \mathcal{F}_{L\times Q}, \mu)$; we have to show that for all $X \in \mathcal{F}_{L\times Q}$ and for all $I \in \mathcal{B}([0,1])$, $\mu_X^{-1}(I) \in \mathcal{F}_{\mathit{Exec}^*}$. It is easy to observe that $\mu_X^{-1}(I)$ corresponds to all the executions from which a distribution in the generator $D_{X,I}$ of σ-algebra on distributions (see Section 2) is scheduled. The measurability of the scheduler ensures that such set of executions is in $\mathcal{F}_{\mathit{Exec}^*}$, as required. □

The proposition above shows that measurable schedulers generate all and only the measurable probabilistic executions, that is, the probabilistic executions that we are interested in. We can therefore disallow non-measurable schedulers.

4.2 Measure on Executions

We can now define the measure on $(\mathit{Exec}, \mathcal{F}_{\mathit{Exec}})$ induced by a scheduler and show that it is defined only for measurable schedulers. Being able to define such a measure is important in order to study global properties on paths, such as the extension of trace distributions [20] to our setting, or if we want to use stochastic transition systems as a model for a stochastic extension of temporal or modal logic [8, 15].

We define the measure $\delta_{\eta,q}$ on basic sets induced by a scheduler η from a start state q inductively on the length of the basic sets as follows:

$$\delta_{\eta,q}(\mathcal{C}_X) = \begin{cases} 1 & \text{if } q \in X \\ 0 & \text{otherwise} \end{cases}$$

$$\delta_{\eta,q}(\mathcal{C}_{\Lambda A X}) = \int_{\alpha\in\Lambda} \mu_{\eta(\alpha)}(A, X)\delta_{\eta,q}(d\alpha)$$

The integral above is defined when the function $f(\alpha) = \mu_{\eta(\alpha)}(A, X)$ is measurable from the measure space of finite executions to $[0, 1]$. From Proposition 1, this is true whenever we deal with measurable schedulers. The measure $\delta_{\eta,q}$ extends uniquely to $\mathcal{F}_{\mathit{Exec}}$ since basic sets form a semi-ring (Theorem 1). We get the following Proposition.

Proposition 2. *Given an STS S and a scheduler η, the measure $\delta_{\eta,q}$ is defined for all basic sets if and only if η is measurable.*

Proof outline. The proof is a consequence of the definition of $\delta_{\eta,q}$ and of Proposition 1. □

Using the measure defined above, and since schedulers use sub-probability distributions, we can define the probability of a set of finite executions that have terminated as the probability to stop after each execution. Formally, given a sequence $\Lambda = X_0A_1\cdots A_nX_n$ of measurable sets of states and actions, we define the probability to stop after Λ as $\delta_{\eta,q}(\mathcal{C}_{\Lambda\perp}) = \int_{\alpha\in\Lambda} \mu_{\eta(\alpha)}(\perp)$. The cones $\mathcal{C}_{\Lambda\perp}$ are in fact the generators of $\mathcal{F}_{Exec^*}$. The probability of eventually terminating is the probability of finite executions, which can be defined as the countable union of disjoint basic sets as follows: $Exec^* = \cup_{i\geq 0}\mathcal{C}_{Q(LQ)^i\perp}$.

5 Parallel Composition and Measurability

In this section we introduce a CSP-style parallel operator [14], under which two STSs synchronise on a common interface alphabet, and study the compositionality properties of schedulers and measurable executions.

Given an STS S, we partition its label space into two measurable sets L^p and L^i of private and interface labels, respectively. We say that two STS S_1 and S_2 are *compatible* if $L_1^p \cap L_2 = \emptyset$ and $L_2^p \cap L_1 = \emptyset$. We denote the union of the measurable spaces of labels by $(L, \mathcal{F}_L)$. We can now define the parallel composition between two compatible STSs.

Definition 8. *Let S_1 and S_2 be two compatible labelled stochastic transition systems. The* **parallel composition** *$S_1 \mid\mid S_2$ of S_1 and S_2 is the stochastic transition system $S = ((Q, \mathcal{F}_Q), \overline{q}, (L, \mathcal{F}_L), \rightarrow)$, where:*

- $(Q, F_Q) = (Q_1 \times Q_2, \mathcal{F}_{Q_1} \otimes \mathcal{F}_{Q_2})$.
- $\overline{q} = (\overline{q}_1, \overline{q}_2)$.
- $(L, \mathcal{F}_L)$ *is the union of the labels of the components.*
- $\rightarrow \subseteq Q \times L \times \mathcal{D}(Q)$ *such that* $((q_1, q_2), a, \mu_1 \otimes \mu_2) \in\rightarrow$ *iff, for* $i \in \{1, 2\}$*:*
  - *if* $a \in L_i$*, then* $(q_i, a, \mu_i) \in\rightarrow_i$*, or*
  - *if* $a \notin L_i$*, then* $\mu_i = Dirac(q_i)$*.*

Observe that $S_1 \mid\mid S_2$ is a well-defined STS given the closure properties of analytic spaces. Next we define two families of functions, π_1 and π_2, to be the left and right projections respectively. Given a state q of S, the projection π_i returns the i-th component of q. For an execution α of S, define the projection $\pi_i(\alpha)$ as the execution of S_i obtained from α by projecting all the states and removing all the actions not in L_i together with the subsequent state. Given a distribution μ on $Q_1 \times Q_2$, the projection $\pi_i(\mu)$ is the distribution on Q_i induced by π_i; $\pi_i(\mu)$ exists since π_i is a measurable function. Finally, given a transition $t = ((q_1, q_2), a, \mu)$, its projection $\pi_i(t)$ is $(q_i, a, \pi_i(\mu))$. If $a \notin L_i$ the projection $\pi_i(t)$ is still defined but it does not correspond to a possible transition of S_i. Note that all the variants of π_1 and π_2 are measurable functions. The following two theorems are important for compositional reasoning.

Theorem 4. *Let S_1 and S_2 be two compatible STSs and α an execution of $S_1 \mid\mid S_2$. Then $\pi_i(\alpha)$ is an execution of S_i, for $i \in \{1, 2\}$.*

Theorem 5. *Let S_1 and S_2 be two compatible STSs and η a measurable scheduler for $S_1 \mid\mid S_2$. Then there exists a measurable scheduler η_1 such that $\delta_{\eta_1,\overline{q}_1} = \pi_1(\delta_{\eta,\overline{q}})$.*

Proof outline. We define the scheduler η_1 on the first component as follows

$$\eta_1(\alpha_1)(T) = \int_{\alpha \in \pi_1^{-1}(\alpha_1)} \eta(\alpha)(\pi_1^{-1}(T))\ \nu(\alpha_1, d\alpha) \tag{1}$$

for all $T \in \mathcal{F}_{\mathcal{T}}$, where $\nu(\alpha_1, d\alpha)$ is the regular conditional probability for $\delta_{\eta,\overline{q}}$ with respect to π_1, whose existence follows from Theorem 3. It is easy to show that η_1 defines a legal scheduler for S_1 and its measurability follows from Theorem 2. In order to prove that $\delta_{\eta_1,\overline{q}_1} = \pi_1(\delta_\eta, \overline{q})$, we need to show that

$$\delta_{\eta_1,\overline{q}_1}(\mathcal{C}_\Lambda) = \delta_{\eta,\overline{q}}(\pi^{-1}(\mathcal{C}_\Lambda)) \tag{2}$$

for all basic sets $\mathcal{C}_\Lambda$. Equation (2) is proved by algebraic arguments and by exploiting the properties of regular conditional probabilities. Since the two measures agree on the basic sets, which form a semi-ring, they extend to the same measure by Theorem 1. □

Theorem 5 shows that the action of a scheduler on S can be derived from the action of the corresponding schedulers on each component since the properties of an execution can be derived from the properties of its components. This allows us to analyse systems in a compositional way. At the same time, this result also confirms that the notion of measurable schedulers and measurable executions is well-defined, since it respects the important requirement of compositionality.

Remark 1. Theorem 5 extends the analogous result for the discrete case [21]. In particular, Equation (1) can be rewritten in a more familiar form as:

$$\eta_1(\alpha_1)(t) = \sum_{\alpha \in \pi_1^{-1}(\alpha_1)} \delta(\mathcal{C}_\alpha \mid \pi_1^{-1}(C_{\alpha_1})) \cdot \eta(\alpha)(\pi_1^{-1}(t)).$$

In the discrete case, we can define the probability for a single transition t. The equation above shows the intuition behind the definition of η_1: each transition is assigned the weighted probability of its inverse image under projection after each execution in the parallel composition, conditioned on being in an execution whose projection is α_1.

6 Weak Transitions and Weak Bisimulation

In this section, we show how the results of the previous sections enable us to define weak transitions and weak bisimulation in our model. A weak transition [17] abstracts from internal computation and considers sequences of actions of the form $\tau^* a \tau^*$, where τ denotes a generic internal action. In the case of probabilistic automata, this is achieved by considering sequences of transitions that form a probabilistic execution where only executions whose trace is of the form $\tau^* a \tau^*$ have positive probabilities [21]. We wish to extend such approach to stochastic transition systems. Of course, in order to do this, we must be able to define the target probability over several steps of executions and we need to restrict to measurable schedulers. Our definition of weak transitions would not be possible without the restrictions on schedulers and the construction of the measure on cones described in Section 4.

We assume the existence of another partitioning of the label space L into two measurable sets, L^e and L^τ, to denote visible and invisible actions, respectively. We denote generic internal actions by τ. A weak transition is defined as a probabilistic execution which terminates with probability 1 and with a support contained in the set of executions containing exactly one visible action. Let $\mathcal{W}_A$ denote the executions whose visible trace is exactly one action $a \in A \subseteq L^e$ and $\mathcal{W} = \mathcal{W}_L$. $\mathcal{W}_A$ is measurable as it can be constructed from basic sets and it also contains infinite executions.

Definition 9. *The pair (q, μ), $q \in Q$ and $\mu \in \mathcal{D}(L \times Q)$, is a* **weak transition** *(denoted by $q \Rightarrow \mu$) if there exists a measurable scheduler η such that $\delta_{\eta,q}(Exec^*) = 1$, $\delta_{\eta,q}(\mathcal{W}) = 1$ and μ is defined as follows: $\mu(A, X) = \delta_{\eta,q}((\cup_{i>0} \mathcal{C}_{Q((A\cup\{\tau\})Q)^i X\perp}) \cap \mathcal{W}_A)$ for all $A \in \mathcal{F}_L$, $A \subseteq L^e$, and $X \in \mathcal{F}_Q$.*

It is easy to show that, under the termination condition $\delta_{\eta,q}(Exec^*) = 1$, μ is a probability measure on $L \times Q$. Weak transitions only consider the local behaviour from one state, and therefore do not preserve measurability properties that are defined on sets of states. For this reason, we use the more general notion of weak hyper-transitions [23], defined as transitions from a distribution over states to a distribution over states and labels.

Definition 10. *Let μ be a probability measure on $(Q, \mathcal{F}_Q)$ and for each $q \in Q$ let $q \Rightarrow \mu_q$ be a weak transition. Define $\mu'(A, X) = \int_Q \mu_q(A, X)\mu(dq)$ if the integral is defined for all $A \in \mathcal{F}_L$ and $X \in \mathcal{F}_Q$. Then we say that $\mu \Rightarrow \mu'$ is a* **weak hyper-transition**.

Hyper transitions are used in the discrete case to prove linear-time properties of systems, such as the fact that bisimulation preserves trace semantics [21]. Note that, in the discrete case, a meaure defined on a set of states and a set of transitions enabled from each of such states always induce a hyper-transition, while in the continuous case this is not always true, because of the usual problems of measurability. This is the reason why we strengthen our notion of weak bisimulation and define it in terms of weak hyper-transitions.

Weak bisimulation. We extend the notion of weak bisimulation to stochastic transition systems. Bisimulation relations, first introduced in the context of CCS [17], are fundamental relations for concurrent systems, and have been extended to the probabilistic setting, both for discrete (strong and weak bisimulation) [16, 11, 22, 3, 19] and continuous state spaces (strong bisimulation) [7, 8, 5]. A notion of weak bisimulation for the continuous setting was introduced in [4], where the problem of defining a measure on paths for weak transitions was not considered, since a weak transition was defined as a succession of τ-labelled non probabilistic transitions followed by a probabilistic transition. The notion of weak transition defined in this paper is suitable for our more general case of several probabilistic steps. Strong bisimulation could be easily defined as it does not abstract from internal computation and therefore can be defined without restrictions to measurable schedulers.

Given an equivalence relation $\mathcal{R}$ on a measurable space $(Q, \mathcal{F}_Q)$, we say that $X \in \mathcal{F}_Q$ is $\mathcal{R}$-closed if it is the union of equivalence classes. Two probability measures μ_1 and μ_2 on Q are $\mathcal{R}$-equivalent ($\mu_1 \mathcal{R} \mu_2$) if $\mu_1(X) = \mu_2(X)$ for all $\mathcal{R}$-closed $X \in \mathcal{F}_Q$, while two probability measures μ_1 and μ_2 on $Q \times L$ are $\mathcal{R}$-equivalent if $\mu_1(A, X) = \mu_2(A, X)$ for all $\mathcal{R}$-closed $X \in \mathcal{F}_Q$ and for all $A \in \mathcal{F}_L$.

Definition 11. *Let S_1 and S_2 be two STSs with the same space of labels. An equivalence relation $\mathcal{R}$ on the union of their sets of states is a* **weak bisimulation** *between S_1 and S_2 if:*

1. *$\overline{q}_1 \mathcal{R} \overline{q}_2$ and*
2. *for all μ_1 and μ_2 $\mathcal{R}$-equivalent measures on states, whenever there is a hyper-transition $\mu_1 \to \mu_1'$, there exists a weak hyper-transition $\mu_2 \Rightarrow \mu_2'$ s.t. $\mu_1' \mathcal{R} \mu_2'$.*

7 Conclusions

We have introduced an operational model for non-deterministic systems with continuous state spaces and continuous probability distributions, thus generalising existing models. We have studied a framework where it is possible to assign probabilities to sets of executions, defined weak bisimulation relation and a parallel composition operator. The relationship between our notion of bisimulation and trace distributions is currently being investigated. Stochastic transition systems are also used as a semantic model for a stochastic process algebra [6]. Further work would include a logical characterisation of our equivalence relations, approximation and metrics.

Acknowledgements. We would like to thank Prakash Panangaden for the helpful discussions.

References

1. L. d. Alfaro. *Formal Verification of Probabilistic Systems.* PhD thesis, Stanford University, 1997. Available as Technical report STAN-CS-TR-98-1601.
2. R. B. Ash. *Real Analysis and Probability.* Academic Press, 1972.
3. C. Baier and H. Hermanns. Weak bisimulation for fully probabilistic processes. In *Proc. 9th International Conference on Computer Aided Verification (CAV'97)*, volume 1254 of *Lecture Notes in Computer Science*, pages 119–130, 1997.
4. M. Bravetti. *Specification and Analysis of Stochastic Real-Time Systems.* PhD thesis, Università di Bologna, Padova, Venezia, 2002.
5. M. Bravetti and P. D'Argenio. Tutte le algebre insieme: Concepts, discussions and relations of stochastic process algebras with general distributions. In C. Baier, B. Haverkort, H. Hermanns, J.-P. Katoen, M. Siegle, and F. Vaandrager, editors, *Validation of Stochastic Systems: A Guide to Current Research*, volume 2925 of *Lecture Notes in Computer Science (Tutorial Volume).* Springer, 2004.

6. S. Cattani. *Trace-based Process Algebras for Real-time Probabilistic Systems*. PhD thesis, School of Computer Science, The University of Birmingham, 2005. Forthcoming.
7. P. R. D'Argenio. *Algebras and Automata for Timed and Stochastic Systems*. PhD thesis, Department of Computer Science, University of Twente, Nov. 1999.
8. J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation for labelled markov processes. *Information and Computation*, 179(2):163–193, 2002.
9. M. 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.
10. R. v. Glabbeek, S. Smolka, and B. Steffen. Reactive, generative, and stratified models of probabilistic processes. *Information and Computation*, 121(1):59–80, 1995.
11. H. Hansson. *Time and Probability in Formal Design of Distributed Systems*, volume 1 of *Real-Time Safety Critical Systems*. Elsevier, 1994.
12. H. Hermanns. *Interactive Markov Chains: The Quest for Quantified Quality*, volume 2428 of *Lecture Notes in Computer Science*. Springer, 2002.
13. J. Hillston. *A Compositional Approach to Performance Modelling*. Cambridge University Press, 1996.
14. C. Hoare. *Communicating Sequential Processes*. Prentice-Hall International, Englewood Cliffs, 1985.
15. M. Kwiatkowska, G. Norman, R. Segala, and J. Sproston. Verifying quantitative properties of continuous probabilistic real-time graphs. In C. Palamidessi, editor, *Proceedings of CONCUR 2000*, volume 1877 of *Lecture Notes in Computer Science*, pages 132–137. Springer, 2000.
16. K. Larsen and A. Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94(1):1–28, Sept. 1991.
17. R. Milner. *Communication and Concurrency*. Prentice-Hall International, Englewood Cliffs, 1989.
18. P. Panangaden. Measure and probability for concurrency theorists. *Theoretical Comput. Sci.*, 253(2):287–309, 2001.
19. A. Philippou, I. Lee, and O. Sokolsky. Weak bisimulation for probabilistic systems. In C. Palamidessi, editor, *Proceedings of CONCUR 2000*, volume 1877 of *Lecture Notes in Computer Science*, pages 334–349. Springer, 2000.
20. R. Segala. A compositional trace-based semantics for probabilistic automata. In I. Lee and S. Smolka, editors, *Proceedings of CONCUR 95*, volume 962 of *Lecture Notes in Computer Science*, pages 234–248. Springer-Verlag, 1995.
21. R. Segala. *Modeling and Verification of Randomized Distributed Real-Time Systems*. PhD thesis, MIT, Dept. of Electrical Engineering and Computer Science, 1995. Also appears as technical report MIT/LCS/TR-676.
22. R. Segala and N. Lynch. Probabilistic simulations for probabilistic processes. *Nordic Journal of Computing*, 2(2):250–273, 1995.
23. M. Stoelinga. *Alea jacta est: verification of probabilistic, real-time and parametric systems*. PhD thesis, University of Nijmegen, the Netherlands, Apr. 2002.
24. S. Wu, S. Smolka, and E. Stark. Composition and behaviors of probabilistic I/O automata. *Theoretical Comput. Sci.*, 176(1-2):1–38, 1999.