Fundamenta Informaticae XX (2013) 1–17
DOI 10.3233/FI-2012-0000
IOS Press

The Measurable Space of Stochastic Processes

Luca Cardelli

Microsoft Research Cambridge, UK – luca@microsoft.com

Radu Mardare

Dept. of Computer Science, Aalborg University, Denmark – mardare@cs.aau.dk

Abstract. We introduce a stochastic extension of CCS endowed with structural operational semantics expressed in terms of measure theory. The set of processes is organised as a measurable space by the sigma-algebra generated by structural congruence. The structural operational semantics associates to each process a set of measures over the space of processes. The measures encode the rates of the transitions from a process (state of a system) to a measurable set of processes. We prove that stochastic bisimulation is a congruence that extends structural congruence. In addition to an elegant operational semantics, our calculus provides a canonic way to define metrics on processes that measure that measure how similar two processes are in terms of behaviour.

Keywords: Markov processes, stochastic process algebras, structural operational semantics

1. Introduction

The investigation of the fundamental relations between computational (theoretical) models and real, physical systems goes back to the first works of Turing on morphogenesis [37] and artificial intelligence [36]. With the development of computer science, the questions related to the existence of the "real computation" in nature specialised and revealed concrete computational phenomena in the real world. The Romanian mathematical school had significant contributions in this direction, being mainly centred on the pioneering work of Prof. Solomon Marcus and his collaborators [6, 28].

This paper treats a similar subject looking into the algebraic foundations of stochastic process algebras that are nowadays used to model complex bio-chemical systems. This paper is a revised version of [7] that was presented at QEST2010. At the time of its first publication, this paper opened a new research direction revealing the necessity to involve measure theory in the research on structural operational semantics.

Process algebras (PAs) [2] are formalisms designed to describe the evolution of concurrent communicating systems. For capturing observable behaviors, PAs are conceptualised along two orthogonal axes. From an algebraic point of view, they are endowed with construction principles in the form of algebraic operations that allow composing larger processes from more basic ones; a process is identified by its algebraic term. On the other hand, there exists a notion of nondeterministic evolution, described by a coalgebraic structure, in the form of a transition system. The algebraic and coalgebraic structures are not independent: Structural Operational Semantics (SOS) defines the behavior of a process inductively on its syntactic structure. In this way, classic PAs are supported by an easy and appealing underlying theory that guarantees their success.

In the past decades *probabilistic* and *stochastic* behaviors have also become of central interest due to the applications in performance evaluation and computational systems biology. *Stochastic process algebras* such as TIPP [18], PEPA [21, 22], EMPA [3] and stochastic π-calculus [32] have been defined as extensions of classic PAs, by considering more complex coalgebraic structures. The label of a stochastic transition contains, in addition to the name of the action, the rate of an exponentially distributed random variable that characterizes the duration of the transition. Consequently, SOS associates a non-negative rate value to each tuple $\langle$*state, action, state*$\rangle$. This additional information imposes important modifications in the SOS format, such as the *multi-transition system* approach of PEPA or the *proved SOS* approach of stochastic π-calculus, mainly because the nondeterminism is replaced by the race policy.

With the intention of developing a stochastic process calculus for applications in systems biology, in this paper we propose a stochastic version of CCS [29] based on the *mass action law* [5] and equipped with an SOS particularly suited to a domain where an equational theory and a measure of similarity of behaviours is important. At the same time we aim to avoid the complicated labeling and counting of previous approaches and to provide an operational semantics that resembles the ones for non-deterministic process algebras, by lifting process-results to measure-results. For doing this, our SOS rules are not given in the pointwise style, but using constructions based on measure theory. We organise the set of processes as a measurable space and associate to each process an indexed set of measures. Thus, for an action a and a measurable set S of processes, the measure μ_a associated to a process P specifies the rate $\mu_a(S) \in \mathbb{R}^+$ of a-transitions from P to (elements of) S. In this way, difficult instance-counting problems that otherwise require complicated versions of SOS can be solved by exploiting the properties of measures (e.g. additivity). Similar ideas have been proposed for probabilistic automata [25, 34] and Markov processes [23, 8, 31]. Following the *transition-systems-as-coalgebras* paradigm [12, 33], this approach follows naturally in the sequence started by nondeterministic and probabilistic transition systems.

The novelty of our approach derives firstly from the structure of the measurable space of stochastic processes. This space is organised by *structural congruence*, an equivalence that equates processes that are indistinguishable from a modeling perspective. For instance, if we model the parallel evolution of two processes, say Q and R, we expect no difference between $Q|R$, $R|Q$ and $R|Q|0$ (0 denotes an inactive process). This relation is required in systems biology where it models chemical mixing: structural congruence was invented in the first place from a chemical analogy [1]. In effect, our σ-algebra is generated by the structural congruence classes and our stochastic transitions are defined from processes to measurable (structural congruence-closed) sets of processes. In this way, if P can perform an action a with a rate r to $Q|R$, written $P \xrightarrow{a,r} Q|R$, we can also derive $P \xrightarrow{a,r} R|Q$ and $P \xrightarrow{a,r} R|(Q|0)$. Otherwise, the alternative approach of considering any set of processes measurable permits to calculate the rate of the a-transitions from P to the set $\{Q|R, R|Q, R|(Q|0)\}$ and obtain the undesired result

$P \xrightarrow{a,3r} \{Q|R, R|Q, R|(Q|0)\}$; to avoid such problems, in the literature have been proposed complicated variants of SOS that make the underlying theory heavy and problematic.

Our choice of developing a stochastic process algebra under the restrictions of the equational theory of structural congruence is sustained by an elegant SOS that supports a smooth development of the basic theory and the definition of metrics for stochastic behaviour. The structures we work with, simply called *Markov processes* (MPs), are particular cases of continuous Markov processes defined in [16]; they extend the notions of *labelled Markov process* [4, 15, 14] and *Harsanyi type space* [20, 30] on to the stochastic level. However, MPs are not continuous-time Markov chains because only the transitions to measurable sets are permitted (which are never singletons) and cannot be described in a pointwise style.

We also introduce a notion of *stochastic bisimulation* for MPs, along the lines of [26, 15, 14, 16]. It generalizes *rate aware bisimulation* of [11], being defined for arbitrary measurable spaces and closed to an equational theory. We prove that stochastic bisimulation is a congruence that extends structural congruence.

Another advantage of our approach consists in the fact that it can be naturally extended to define a class of metrics on stochastic processes which measure the similarity of process behaviours. This result has considerable practical application. The standard notion of bisimulation for probabilistic or stochastic systems cannot distinguish between two processes that are substantially different and two processes that differ by only a small amount in a real valued parameter. It is often more useful to say how similar two processes are than to say whether they are exactly the same. This is precisely what our metrics do: stochastic bisimilar processes are at distance zero, processes that differ by small values of rates are closer than the processes with bigger differences.

The paper is organised as follows. A preliminary section establishes the basic concepts and notations. Section 3 defines the general concept of Markov process (MP) and the stochastic bisimulation of MPs. Section 4 introduces the syntax of our process algebra and the axiomatization of structural congruence; we prove that the space of processes can be organised as a Markov kernel and that each process is an MP. These results guide us, in Section 5, to the definition of a structural operational semantics which induces a notion of behavioural equivalence that coincides with the bisimulation of MPs. In Section 6 we show that the bisimulation behaves well with respect to the algebraic structure of processes: stochastic bisimulation is a congruence. This relation is extended in Section 7 with a class of metrics on the space of processes that measure how similar two processes are. We also have a section dedicated to related work and a concluding section.

2. Preliminaries

In this section we recall the notions of measure theory and we introduce the notations used in the paper.

For arbitrary sets A and B, 2^A denotes the powerset of A, $A \uplus B$ their disjoint union and both $[A \to B]$ and B^A denote the class of functions from A to B. If $f \in B^A$, $f^{-1} \colon 2^B \to 2^A$ denotes the pre-image of f. For an arbitrary function $f : A \to B$, the *kernel of* f is the relation $ker(f) = \{(x, y) \in A \times A \mid f(y) = f(x)\}$.

Given a set M, a family $\Sigma \subseteq 2^M$ that contains M and is closed under complement and countable union is a *σ-algebra* over M; (M, Σ) is called a *measurable space*, the elements of Σ *measurable sets* and M the *support-set*. A set $\Omega \subseteq 2^M$ is a *generator for the σ-algebra* Σ on M if Σ is the smallest σ-algebra over M that contains Ω, written $\overline{\Omega} = \Sigma$. A generator with disjoint elements is called *base*.

A *measure* on a measurable space (M, Σ) is a function $\mu\colon \Sigma \to \mathbb{R}^+$ such that $\mu(\emptyset) = 0$ and for any $\{N_i \mid i \in I \subseteq \mathbb{N}\} \subseteq \Sigma$ with pairwise disjoint elements, $\mu(\bigcup_{i\in I} N_i) = \sum_{i\in I} \mu(N_i)$. The *null measure* on (M, Σ) is the measure ω such that $\omega(M) = 0$; for $m \in M$ and $r \in \mathbb{R}^+$, the *r-Dirac measure at m* on (M, Σ) is the measure $D(r, m)$ such that, for $S \in \Sigma$, $D(r, m)(S) = r$ if $m \in S$, and $D(r, m)(S) = 0$ otherwise. If Ω is a base for (M, Σ), any function $\mu'\colon \Omega \to \mathbb{R}^+$ can be (uniquely) extended to a measure μ on (M, Σ) as $\mu(\cup_{i\in I} N_i) = \sum_{i\in I} \mu'(N_i)$, for any $N_i \in \Omega$ and $i \in I \subseteq \mathbb{N}$. Let $\Delta(M, \Sigma)$ be the measurable space of the measures on (M, Σ) with the σ-algebra generated, for arbitrary $S \in \Sigma$ and $r > 0$, by the sets $\{\mu \in \Delta(M, \Sigma) : \mu(S) \geq r\}$.

Given two measurable spaces (M, Σ) and (N, Θ), a mapping $f\colon M \to N$ is *measurable* if for any $T \in \Theta$, $f^{-1}(T) \in \Sigma$. We use $[\![M \to N]\!]$ to denote the class of measurable mappings from (M, Σ) to (N, Θ).

Given a set X, a *pseudometric on* X is a function $d\colon X \times X \to \mathbb{R}^+$ such that for all $x, y, z \in X$

$$d(x, x) = 0\,, \qquad d(x, y) = d(y, x)\,, \qquad d(x, y) \leq d(x, z) + d(y, z)\,.$$

The pair (X, d) is a *pseudometric space*; it is a *metric space* if in addition whenever $d(x, y) = 0$, then $x = y$. Given a pseudometric on X, one can define an equivalence on X by pairing the elements at distance zero.

3. Continuous Markov processes

In this section we define a general notion of *Markov process* (MP) that encapsulates various notions of Markovian stochastic systems such as *Markov chain* with discrete or continuous time [23], *labelled Markov process* [31] as well as the most general case of *continuous-space and continuous-time Markov process* introduced in [16]. A Markov process is a coalgebraic structure that encodes stochastic behaviors and can be seen, following the *transition-systems-as-coalgebras* paradigm [33, 12], as a generalisation of the notion of transition system: a transition system associates with each state of a system an action-indexed set of functions over the state space; functions with *boolean values* define *labelled transition systems* while *probabilistic distributions* define labelled Markov processes [4, 15, 14, 31] and *Harsanyi type spaces* [20, 30]. This paradigm is particularly appropriate when one is interested in systems with a complex state space where transitions cannot be represented from one state to another, but from a state to a measurable set of states or to a (topological) neighbourhood.

An MP involves a set A of labels. The labels $\alpha \in A$ represent types of interactions with the environment. If m is the current state of the system and N is a measurable set of states, $\theta(\alpha)(m)$ is a measure on the state space and $\theta(\alpha)(m)(N) \in \mathbb{R}^+$ represents the *rate* of an exponentially distributed random variable that characterizes the duration of an α-transition from m to arbitrary $n \in N$. Indeterminacy is solved by races between events executing at different rates.

Definition 3.1. (Markov kernels and Markov processes)
Let (M, Σ) be a measurable space[1] and A a denumerable set. An *A-Markov kernel* is a tuple $\mathcal{M} = (M, \Sigma, \theta)$, with

$$\theta : A \to [\![M \to \Delta(M, \Sigma)]\!].$$

[1] In [31] the support space is required to be an analytic space. Because the spaces we will work with are in most of the cases discrete (or can be discretized by taking simple quotients) and we do not face the problems of continuous state space, we have chosen to avoid this requirement. For a detailed discussion on this issue the reader is referred to [31] (Section 7.5) or to [17] (Section 4.4).

If $m \in M$ then the tuple (M, Σ, θ, m) is an *A-Markov process* of $\mathcal{M}$ and m is its *initial state*.

Notice that $\theta(\alpha)$ is defined as a measurable mapping between (M, Σ) and the measurable space $\Delta(M, \Sigma)$ of the measures on (M, Σ). This condition is equivalent to the conditions on the two-variable *rate function* used in [16, 31] (see, e.g. Proposition 2.9, of [17]). If $\mathcal{M} = (M, \Sigma, \theta)$, we sometimes denote the process (M, Σ, θ, m) by $(\mathcal{M}, m)$.

We define the stochastic bisimulation relation on MPs following the similar definitions of [14, 16, 31].

Given a binary relation $\mathfrak{R} \subseteq M \times M$ on a set M, we call a subset $N \subseteq M$ $\mathfrak{R}$-closed iff

$$\{m \in M \mid \exists n \in N, (n, m) \in \mathfrak{R}\} \subseteq N.$$

If (M, Σ) is a measurable space, $\Sigma(\mathfrak{R})$ denotes the set of measurable $\mathfrak{R}$-closed subsets of M.

Definition 3.2. (Stochastic bisimulation)
Let $\mathcal{M} = (M, \Sigma, \theta)$ be an A-Markov kernel. A *rate-bisimulation relation* is an equivalence relation $\mathfrak{R} \subseteq M \times M$ such that $(m, n) \in \mathfrak{R}$ iff for any $C \in \Sigma(\mathfrak{R})$ and any $\alpha \in A$,

$$\theta(\alpha)(m)(C) = \theta(\alpha)(n)(C).$$

Two MPs $(\mathcal{M}, m)$ and $(\mathcal{M}, n)$ are *stochastic bisimilar*, written $m \sim_{\mathcal{M}} n$, if m and n are related by a rate-bisimulation relation.

4. A minimal Stochastic Process Algebra

In this section we introduce a stochastic extension of CCS without replication [29]. As usual in stochastic process algebras, each transition a has associated a *rate* in $\mathbb{R}^+$ representing the parameter of an exponentially distributed random variable that characterizes the duration of an a-action. In addition, we also consider synchronizations of actions. As in CCS, the set of actions is equipped with an *involution* that associates with each action a its paired action $\overline{a}$; the paired actions have the same rates. The synchronization of $(a, \overline{a})$ counts as an *internal τ-action* with the rate satisfying the *mass action law* [5].

Formally, the set of *labels (actions)* is a countable set $\mathbb{A}$ endowed with (i) an *involution* associating to each $a \in \mathbb{A}$ an element $\overline{a} \in \mathbb{A}$ such that $a \neq \overline{a}$ and $\overline{\overline{a}} = a$; (ii) a *weight function* $\iota : \mathbb{A} \to \mathbb{Q}^+$, such that for any $a \in \mathbb{A}$, $\iota(a) = \iota(\overline{a})$. In what follows we use two extensions of $\mathbb{A}$ defined for the *internal action* $\tau \notin \mathbb{A}$. On syntactic level we involve the set $\mathbb{A}^* = \mathbb{A} \cup \{\tau_r \mid r \in \mathbb{Q}^+\}$, where indexed internal actions will be used for modelling delays in a system[2]; we extend ι to $\mathbb{A}^*$ by $\iota(\tau_r) = r$. For operational semantics we use the set $\mathbb{A}^+ = \mathbb{A} \cup \{\tau\}$ of labels. In what follows a, a', a_i denote arbitrary elements of $\mathbb{A}$, $\varepsilon, \varepsilon', \varepsilon_i$ denote arbitrary elements of $\mathbb{A}^*$ and $\alpha, \alpha', \alpha_i$ denote arbitrary elements of $\mathbb{A}^+$.

Definition 4.1. $\mathbb{A}$-*stochastic processes* are defined, for arbitrary $\varepsilon \in \mathbb{A}^*$, as follows

$$P ::= 0 \mathrel{\vdots} \varepsilon.P \mathrel{\vdots} P|P \mathrel{\vdots} P + P\,.$$

[2] In practice we cannot measure nor specify models with irrational rates and for this reason we have chosen $\iota(\varepsilon) \in \mathbb{Q}^+$ for all $\varepsilon \in \mathbb{A}^*$. However, the technical development does not change if $\iota : \mathbb{A} \to \mathbb{R}^+$.

We denote by $\mathbb{P}$ the set of stochastic processes.

An essential notion for processes is the *structural congruence relation* which equates processes that, in spite of their different syntactic form, represent the same systems.

Definition 4.2. (Structural congruence)
Structural congruence $\equiv\subseteq\mathbb{P}\times\mathbb{P}$ is the smallest equivalence relation satisfying, for arbitrary $P,Q,R\in\mathbb{P}$ and $\varepsilon\in\mathbb{A}^*$ the following conditions:

$$\begin{array}{ccc} P|Q\equiv Q|P & (P|Q)|R\equiv P|(Q|R) & P|0\equiv P \\ P+Q\equiv Q+P & (P+Q)+R\equiv P+(Q+R) & P+0\equiv P \\ P\equiv Q\implies P|R\equiv Q|R & P\equiv Q\implies P+R\equiv Q+R & P\equiv Q\implies \varepsilon.P\equiv\varepsilon.Q \end{array}$$

Let $\mathbb{P}^{\equiv}$ be the set of $\equiv$-equivalence classes on $\mathbb{P}$. For arbitrary $P\in\mathbb{P}$, we denote by $P^{\equiv}$ the $\equiv$-equivalence class of P. The set of stochastic processes is organized as a measurable space $(\mathbb{P},\Pi)$, where Π is the σ-algebra generated by $\mathbb{P}^{\equiv}$. Note that $\mathbb{P}^{\equiv}$ is a base for $(\mathbb{P},\Pi)$ and the measurable sets are (possibly denumerable) unions of $\equiv$-equivalence classes on $\mathbb{P}$. In what follows we use $\mathcal{P},\mathcal{P}_i,\mathcal{R},\mathcal{Q}$ to denote arbitrary measurable sets of Π.

Consider the following operations on the sets of Π, for arbitrary $\mathcal{P},\mathcal{Q}\in\Pi$ and $P\in\mathbb{P}$.

$$\mathcal{P}|\mathcal{Q}=\bigcup_{P\in\mathcal{P},Q\in\mathcal{Q}}(P|Q)^{\equiv}\quad\text{and}\quad\mathcal{P}_P=\bigcup_{P|R\in\mathcal{P}}R^{\equiv}.$$

Notice that $\mathcal{P}|\mathcal{Q}$ and $\mathcal{P}_P$ are measurable sets.

In what follows we show that the measurable space $(\mathbb{P},\Pi)$ of stochastic processes can be organised as an $\mathbb{A}^+$-Markov kernel. Next we define a function $\theta\colon\mathbb{A}^+\to[\![\mathbb{P}\to\Delta(\mathbb{P},\Pi)]\!]$ which organizes $(\mathbb{P},\Pi,\theta)$ as an $\mathbb{A}^+$-Markov kernel. For arbitrary $P\in\mathbb{P},\mathcal{P}\in\Pi$ and $\alpha\in\mathbb{A}^+$, $\theta(\alpha)(P)(\mathcal{P})$ represents the total rate of the α actions from P to (elements of) $\mathcal{P}$.

Definition 4.3. Let $\theta\colon\mathbb{A}^+\to[\mathbb{P}\to\Delta(\mathbb{P},\Pi)]$ be defined, by induction on the structure of $P\in\mathbb{P}$, as

The case $P=0$: For any $\alpha\in\mathbb{A}^+$, let $\theta(\alpha)(0)=\omega$.

The case $P=\varepsilon.Q$, $\varepsilon\in\mathbb{A}^*$: For arbitrary $a\in\mathbb{A}$, let

$$\theta(\tau)(\varepsilon.Q)=\begin{cases}D(\iota(\varepsilon),Q), & \varepsilon\notin\mathbb{A}\\ \omega, & \varepsilon\in\mathbb{A}\end{cases}\qquad\theta(a)(\varepsilon.Q)=\begin{cases}D(\iota(\varepsilon),Q), & \varepsilon=a\\ \omega, & \varepsilon\neq a\end{cases}$$

The case $P=Q+R$: For any $\alpha\in\mathbb{A}^+$ and $\mathcal{P}\in\Pi$,

$$\theta(\alpha)(Q+R)(\mathcal{P})=\theta(\alpha)(Q)(\mathcal{P})+\theta(\alpha)(R)(\mathcal{P}).$$

The case $P=Q|R$: For any $a\in\mathbb{A}$ and $\mathcal{P}\in\Pi$,

$$\theta(a)(Q|R)(\mathcal{P})=\theta(a)(R)(\mathcal{P}_Q)+\theta(a)(Q)(\mathcal{P}_R),$$

$$\theta(\tau)(Q|R)(\mathcal{P})=\theta(\tau)(R)(\mathcal{P}_Q)+\theta(\tau)(Q)(\mathcal{P}_R)+\sum_{\mathcal{P}_1|\mathcal{P}_2\subseteq\mathcal{P}}^{a\in\mathbb{A},\iota(a)\neq0}\frac{\theta(a)(Q)(\mathcal{P}_1)\cdot\theta(\overline{a})(R)(\mathcal{P}_2)}{2\cdot\iota(a)}.$$

If we define the *set of active actions of a process* $P \in \mathbb{P}$ by $act(0) = \emptyset$, $act(a.P) = \{a\}$, $act(P + Q) = act(P|Q) = act(P) \cup act(Q)$, then any process has only a finite set of active actions. Notice that $\theta(a)(P) \neq \omega$ iff $a \in act(P)$. This means that for any $a \notin act(P)$ and any $\mathcal{R} \in \Pi$, $\theta(a)(P)(\mathcal{R}) = 0$. Consequently, the infinitary sum involved in Definition 4.3 has a finite number of non-zero summands.

The next theorem states that the space of processes with the function defined above is an $\mathbb{A}^+$-Markov kernel. It implicitly states the correctness of the previous definition: for each $\alpha \in \mathbb{A}^+$ and each $P \in \mathbb{P}$, $\theta(\alpha)(P) \in \Delta(\mathbb{P}, \Pi)$. It follows that for each $\alpha \in \mathbb{A}^+$, $\theta(\alpha) \in [\![\mathbb{P} \to \Delta(\mathbb{P}, \Pi)]\!]$.

Theorem 4.4. $(\mathbb{P}, \Pi, \theta)$ is an $\mathbb{A}^+$-Markov kernel.

A consequence of the previous theorem is that for each $P \in \mathbb{P}$, $(\mathbb{P}, \Pi, \theta, P)$ is a Markov process. In effect, we can define a stochastic bisimulation for the elements of our process algebra simply as stochastic bisimulation of Markov processes in $(\mathbb{P}, \Pi, \theta)$.

5. Structural Operational Semantics

In this section we introduce the structural operational semantics for the minimal process algebra, with the intention to induce a behavioural equivalence on processes that coincides with their bisimulation as MPs. In this case we do not associate to each tuple *(process, action, process)* a rate, as usual in stochastic process algebras, because a transition in our case is not between two processes, but from a process to an infinite measurable set of processes. However, our intention is to maintain "the spirit" of process algebras and for this reason we use "generalised" transitions of type $P \to \mu$ where $\mu \colon \mathbb{A}^+ \to \Delta(\mathbb{P}, \Pi)$ is a function defining a class of $\mathbb{A}^+$-indexed measures on $(\mathbb{P}, \Pi)$.

For simplifying the rules of the operational semantics, we first define some operations on the functions in $\Delta(\mathbb{P}, \Pi)^{\mathbb{A}^+}$ and analyze their mathematical structures and properties.

We say that a function $\mu \in \Delta(\mathbb{P}, \Pi)^{\mathbb{A}^+}$ has *finite support* if $\mathbb{A} \setminus \mu^{-1}(\omega)$ is finite or empty.

Definition 5.1.

1. Let $\overline{\omega} \colon \mathbb{A}^+ \to \Delta(\mathbb{P}, \Pi)$ be defined by $\overline{\omega}(\alpha) = \omega$, for any $\alpha \in \mathbb{A}^+$.
2. For arbitrary $\varepsilon \in \mathbb{A}^*$ and $P \in \mathbb{P}$ let $[^{\varepsilon}_{P}] \colon \mathbb{A}^+ \to \Delta(\mathbb{P}, \Pi)$ be defined, for arbitrary $a \in \mathbb{A}$, by

$$[^{\varepsilon}_{P}](a) = \begin{cases} D(\iota(\varepsilon), P), & a = \varepsilon \\ \omega, & a \neq \varepsilon \end{cases} \qquad [^{\varepsilon}_{P}](\tau) = \begin{cases} D(\iota(\varepsilon), P), & \varepsilon \notin \mathbb{A} \\ \omega, & \varepsilon \in \mathbb{A} \end{cases}$$

3. For arbitrary $\mu', \mu'' \in \Delta(\mathbb{P}, \Pi)^{\mathbb{A}^+}$, let $\mu' \oplus \mu'' \colon \mathbb{A}^+ \to \Delta(\mathbb{P}, \Pi)$ be defined, for $\alpha \in \mathbb{A}^+$, by

$$(\mu' \oplus \mu'')(\alpha) = \mu'(\alpha) + \mu''(\alpha)\,.$$

4. For arbitrary $\mu', \mu'' \in \Delta(\mathbb{P}, \Pi)^{\mathbb{A}^+}$ with finite support and $P, Q \in \mathbb{P}$, let $\mu' \;{}_{P}{\otimes}_{Q}\, \mu'' \colon \mathbb{A}^+ \to \Delta(\mathbb{P}, \Pi)$ be defined by,

$$(\mu' \;{}_{P}{\otimes}_{Q}\, \mu'')(a)(\mathcal{R}) = \mu'(a)(\mathcal{R}_Q) + \mu''(a)(\mathcal{R}_P) \text{ for } a \in \mathbb{A}$$

$$(\mu' \;{}_{P}{\otimes}_{Q}\, \mu'')(\tau)(\mathcal{R}) = \mu'(\tau)(\mathcal{R}_Q) + \mu''(\tau)(\mathcal{R}_P) + \sum_{\mathcal{P}_1|\mathcal{P}_2 \subseteq \mathcal{R}}^{a \in \mathbb{A}, \iota(a) \neq 0} \frac{\mu'(a)(\mathcal{P}_1) \cdot \mu''(\overline{a})(\mathcal{P}_2)}{2 \cdot \iota(a)}\,.$$

$$(Null)\ \frac{}{0 \to \overline{\omega}} \qquad\qquad \frac{}{\varepsilon.P \to [^{\varepsilon}_{P}]}\ (Guard)$$

$$(Sum)\ \frac{P \to \mu' \qquad Q \to \mu''}{P+Q \to \mu' \oplus \mu''} \qquad\qquad \frac{P \to \mu' \qquad Q \to \mu''}{P|Q \to \mu'\ {}_{P}\!\otimes_{Q}\, \mu''}\ (Par)$$

Table 1. Structural Operational Semantics

Observe that because μ' and μ'' have finite support, the sum involved in the definition of ${}_P\otimes_Q$ has a finite number of non-zero summands.

The next lemma proves that the definitions of $\oplus$ and ${}_P\otimes_Q$ for arbitrary $P, Q \in \mathbb{P}$ are correct; it also states some basic properties of these operators.

Lemma 5.2.

1. For arbitrary $\mu, \mu', \mu'' \in \Delta(\mathbb{P}, \Pi)^{\mathbb{A}^+}$, $\mu \oplus \mu' \in \Delta(\mathbb{P}, \Pi)^{\mathbb{A}^+}$ and
   (*a*) $\mu \oplus \mu' = \mu' \oplus \mu$,
   (*b*) $(\mu \oplus \mu') \oplus \mu'' = \mu \oplus (\mu' \oplus \mu'')$,
   (*c*) $\mu = \mu \oplus \overline{\omega}$.
2. For arbitrary $P, Q, R \in \mathbb{P}$ and $\mu', \mu'', \mu''' \in \Delta(\mathbb{P}, \Pi)^{\mathbb{A}^+}$ with finite support, $\mu\, {}_P\otimes_Q \mu' \in \Delta(\mathbb{P}, \Pi)^{\mathbb{A}^+}$ and
   (*a*) $\mu'\ {}_P\otimes_Q\ \mu'' = \mu''\ {}_Q\otimes_P\ \mu'$,
   (*b*) $(\mu'\ {}_P\otimes_Q\ \mu'')\ {}_{P|Q}\otimes_R\ \mu''' = \mu'\ {}_P\otimes_{Q|R}\ (\mu''\ {}_Q\otimes_R\ \mu''')$,
   (*c*) $\mu'\ {}_P\otimes_0\ \overline{\omega} = \mu'$.
3. For arbitrary $P, P', Q, Q' \in \mathbb{P}$, $\varepsilon \in \mathbb{A}^+$ and $\mu', \mu'' \in \Delta(\mathbb{P}, \Pi)^{\mathbb{A}^+}$ with finite support,
   (*a*) if $P \equiv P'$ and $Q \equiv Q'$, then $\mu'\ {}_P\otimes_Q\ \mu'' = \mu'\ {}_{P'}\otimes_{Q'}\ \mu''$,
   (*b*) if $P \equiv Q$, then $[^{\varepsilon}_{P}] = [^{\varepsilon}_{Q}]$.

The rules of the structural operational semantics, given for arbitrary $P, Q \in \mathbb{P}$ and $\varepsilon \in \mathbb{A}^+$, are listed in Table 1. The *stochastic transition relation* is the smallest relation $\to\ \subseteq \mathbb{P} \times \Delta(\mathbb{P}, \Pi)^{\mathbb{A}^+}$ satisfying these rules. The operational semantics associates with each process $P \in \mathbb{P}$ a mapping $\mu \in \Delta(\mathbb{P}, \Pi)^{\mathbb{A}^+}$. For each $\equiv$-closed set of processes $\mathcal{P} \in \Pi$ and each $\alpha \in \mathbb{A}^+$, $\mu(\alpha)(\mathcal{P}) \in \mathbb{R}^+$ represents the total rate of the α-reductions of P to some arbitrary element of $\mathcal{P}$.

The next lemma guarantees the consistency of the relation $\to$ and of our operational semantics.

Lemma 5.3. For any $P \in \mathbb{P}$ there exists a unique $\mu \in \Delta(\mathbb{P}, \Pi)^{\mathbb{A}^+}$ such that $P \to \mu$; moreover, μ has finite support.

The next lemma guarantees that the operational semantics does not differentiate the structural congruent processes.

Lemma 5.4. If $P \equiv Q$ and $P \to \mu$, then $Q \to \mu$.

Notice that the signature of $\mathbb{P}$ does not corresponds to the signature of $\Delta(\mathbb{P}, \Pi)^{\mathbb{A}^+}$ – a fact that differentiates our approach from the other GSOS [35] or SGSOS [24] formats. For instance, to the parallel operator "|" corresponds, in the domain of functions, a denumerable class of binary operators indexed by processes $- \,_P\otimes_Q$. This situation is a consequence of the fact that $\equiv\, \subsetneq\, \approx$. Consider the processes $P = a.0|b.0$ and $Q = a.b.0 + b.a.0$ for $a, b \in \mathbb{A}$ and $\{a, \overline{a}\} \cap \{b, \overline{b}\} = \emptyset$. Then $P \to (\mu_1 = \left[\begin{smallmatrix}a\\0\end{smallmatrix}\right] \,_{a.0}\otimes_{b.0} \left[\begin{smallmatrix}b\\0\end{smallmatrix}\right])$ and $Q \to (\mu_2 = \left[\begin{smallmatrix}a\\b.0\end{smallmatrix}\right] \oplus \left[\begin{smallmatrix}b\\a.0\end{smallmatrix}\right])$. One can verify that $\mu_1 = \mu_2$, however $P \not\equiv Q$. This shows that for some $R \to \nu$, we can have $\mu_1 \,_P\otimes_R\, \nu \neq \mu_1 \,_Q\otimes_R\, \nu$.

Notice also that the "generalised" transition system induced by our SOS is image-finite. The importance of this property was motivated from the perspective of GSOS in [35] where it is observed that image-finite GSOS are in one-to-one correspondence with the *distributive laws* that ensure the cooperation between the algebraic and the coalgebraic structures of the class of processes. The next lemma shows that our system has a similar property. We write $P \Longrightarrow Q$ if there exists $\alpha \in \mathbb{A}^+$ and $r \neq 0$ such that $P \xrightarrow{\alpha,r} Q^{\equiv}$ and let $\Longrightarrow^*$ be the transitive closure of $\Longrightarrow$.

Lemma 5.5. For an arbitrary process $P \in \mathbb{P}$, $\{\alpha \in \mathbb{A}^+ \mid P \xrightarrow{\alpha,r} \mathbb{P}, r \neq 0\}$, $\{Q^{\equiv} \in \Pi \mid P \Longrightarrow Q\}$ and $\{Q^{\equiv} \in \Pi \mid P \Longrightarrow^* Q\}$ are finite.

6. Stochastic bisimulation is a congruence

This section is dedicated to the study of stochastic bisimulation for the minimal stochastic process algebra. In the pointwise approach, since the operational semantics requires various mathematical artifacts such as the multi-transition systems [21, 22] or the proved SOS [32], the problem of stochastic bisimulation is difficult to trace. Recently, an elegant solution was proposed in [24] for the case when there are no equational restrictions on the algebraic level. As argued before, our algebra is endowed with an equational theory of structural congruence that organizes the measurable space of processes and consequently, stochastic bisimulation requires a different treatment.

We introduce the stochastic bisimulation for the minimal process algebra as the stochastic bisimulation on the Markov kernel $(\mathbb{P}, \Pi, \theta)$. We show that it behaves well both on coalgebraic and on algebraic levels: processes that have associated the same functions by our SOS are bisimilar and the bisimulation is a congruence that extends the structural congruence.

Lemma 5.3 shows that the operational semantics induces a function $\vartheta\colon \mathbb{P} \to \Delta(\mathbb{P}, \Pi)^{\mathbb{A}^+}$ defined by

$$\vartheta(P) = \mu \qquad \text{iff} \qquad P \to \mu\,.$$

There exists an obvious relation between ϑ and the function θ that organises $\mathbb{P}$ as a Markov kernel. It reflects the similarity between Definitions 4.3 and 5.1.

Lemma 6.1. If $(\mathbb{P}, \Pi, \theta)$ is the Markov kernel of processes and $\vartheta\colon \mathbb{P} \to \Delta(\mathbb{P}, \Pi)^{\mathbb{A}^+}$ is the function induced by SOS, then for any $P \in \mathbb{P}$, $\alpha \in \mathbb{A}^+$ and $\mathcal{P} \in \Pi$, $\theta(\alpha)(P)(\mathcal{P}) = \vartheta(P)(\alpha)(\mathcal{P})$.

Recall that for a Markov kernel (M, Σ, θ), $\sim_{(M,\Sigma,\theta)}$ denotes the stochastic bisimulation on it. The next result is a direct consequence of the previous lemma stating that $\sim_{(\mathbb{P},\Pi,\theta)}$ is an extension of the kernel of ϑ.

Corollary 6.2. For arbitrary $P, Q \in \mathbb{P}$, if $P \to \mu$ and $Q \to \mu$, then $P \sim_{(\mathbb{P},\Pi,\theta)} Q$.

This result guarantees that we can safely define the stochastic bisimulation for our process algebra as the stochastic bisimulation on $(\mathbb{P}, \Pi, \theta)$.

Definition 6.3. (Stochastic bisimulation on processes)
A *rate bisimulation relation* on processes is an equivalence relation $\mathfrak{R} \subseteq \mathbb{P} \times \mathbb{P}$ such that for arbitrary $P, Q \in \mathbb{P}$ with $P \to \mu$ and $Q \to \mu'$, $(P, Q) \in \mathfrak{R}$ iff for any $C \in \Pi(\mathfrak{R})$ and any $\alpha \in \mathbb{A}^+$,

$$\mu(\alpha)(C) = \mu'(\alpha)(C).$$

$P, Q \in \mathbb{P}$ are stochastic bisimilar, written $P \sim Q$, if there exists a rate bisimulation relating them.

The next theorem provides a characterization of stochastic bisimulation.

Theorem 6.4. The stochastic bisimulation $\sim$ is the smallest equivalence relation on $\mathbb{P}$ such that for arbitrary $P, Q \in \mathbb{P}$ with $P \to \mu$ and $Q \to \mu'$, $P \sim Q$ iff for any $C \in \Pi(\sim)$ and any $\alpha \in \mathbb{A}^+$, $\mu(\alpha)(C) = \mu'(\alpha)(C)$.

Denote by $\mathbb{P}^{\sim}$ the set of $\sim$-equivalence classes on $\mathbb{P}$, and by $P^{\sim}$ the $\sim$-equivalence class of $P \in \mathbb{P}$. If P and Q are not stochastic bisimilar, we write $P \not\sim Q$.

In what follows we show some bisimilar processes. The first example is a general rule for concurrent Markovian processes (see Section 4.1.2 of [23]).

Example 6.5. (i) If $a, b \in A$ such that $\overline{a} \neq b$, then for any $P, Q \in \mathbb{P}$, $a.P|b.Q \sim a.(P|b.Q) + b.(a.P|Q)$. Indeed, $a.P|b.Q \to [^a_P]\, {}_{a.P}\!\otimes_{b.Q} [^b_Q]$, $a.(P|b.Q) + b.(a.P|Q) \to [^a_{P|b.Q}] \oplus [^b_{a.P|Q}]$ and for arbitrary $C \in \mathbb{P}^{\sim}$,

$$[^a_P]\ {}_{a.P}\!\otimes_{b.Q}\ [^b_Q](x)(C) = [^a_{P|b.Q}] \oplus [^b_{a.P|Q}](x)(C) = \begin{cases} \iota(a) & \text{if } x = a \text{ and } P|b.Q \in C\,, \\ \iota(b) & \text{if } x = b \text{ and } a.P|Q \in C\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Example 6.6. Let $b, c \in \mathbb{A}$ be such that $\overline{b} \neq c$. In Example 6.5 we have seen that $b.0|c.0 \sim b.c.0 + c.b.0$. Consider the processes $P = \tau_r.(b.0|c.0) + \tau_r.(b.c.0 + c.b.0)$, $Q = \tau_r.(b.0|c.0) + \tau_r.(b.0|c.0)$, and $R = \tau_r.(b.c.0 + c.b.0) + \tau_r.(b.c.0 + c.b.0)$. If C is the $\sim$-equivalence class that contains $b.0|c.0$ and $b.c.0 + c.b.0$, then $P \xrightarrow{\tau,2r} C$, $Q \xrightarrow{\tau,2r} C$, $R \xrightarrow{\tau,2r} C$ and for any other $\sim$-equivalence class C', $P \xrightarrow{\tau,0} C'$, $Q \xrightarrow{\tau,0} C'$ and $R \xrightarrow{\tau,0} C'$. Consequently, $P \sim Q \sim R$. On the other hand, if we consider the pointwise semantics, we obtain

$$\begin{array}{lll} P \xrightarrow{\tau,r} b.0|c.0 & \quad Q \xrightarrow{\tau,2r} b.0|c.0 & \quad R \xrightarrow{\tau,0} b.0|c.0 \\ P \xrightarrow{\tau,r} b.c.0 + c.b.0 & \quad Q \xrightarrow{\tau,0} b.c.0 + c.b.0 & \quad R \xrightarrow{\tau,2r} b.c.0 + c.b.0\,. \end{array}$$

Notice that they are not agreeing on any "pointwise" transition. This emphasizes the difficulties with pointwise semantics.

The relation $\sim$ on $\mathbb{P}$ can be lifted to $\Delta(P)^{\mathbb{A}^+}$ by defining, for arbitrary $\mu, \mu' \in \Delta(\mathbb{P})^{\mathbb{A}^+}$, $\mu \sim \mu'$ iff for any $C \in \mathbb{P}^{\sim}$ and any $\alpha \in \mathbb{A}^+$, $\mu(\alpha)(C) = \mu'(\alpha)(C)$. Notice that $\sim \subseteq \Delta(\mathbb{P})^{\mathbb{A}^+} \times \Delta(\mathbb{P})^{\mathbb{A}^+}$ is an equivalence relation. We denote by $(\Delta(\mathbb{P})^{\mathbb{A}^+})^{\sim}$ the set of $\sim$-equivalence classes on $\Delta(\mathbb{P})^{\mathbb{A}^+}$ and for an arbitrary $\mu \in \Delta(\mathbb{P})^{\mathbb{A}^+}$ we denote by $\mu^{\sim}$ the $\sim$-equivalence class of μ.

With this notation, from Theorem 6.4 we derive the next corollary.

Corollary 6.7. Given $P, Q \in \mathbb{P}$, if $P \to \mu$ and $Q \to \mu'$, then $P \sim Q$ iff $\mu \sim \mu'$.

A consequence of $\sim$ being an equivalence on $\Delta(\mathbb{P})^{\mathbb{A}^+}$ is the next theorem that shows that our processes behave "correctly" with respect to structural congruence.

Theorem 6.8. Given $P, Q \in \mathbb{P}$, if $P \equiv Q$, then $P \sim Q$.

In addition, notice that $\sim$ is strictly larger than $\equiv$, because for arbitrary $a, b \in \mathbb{A}$ with $\overline{a} \neq b$, we have $a.0|b.0 \sim a.b.0 + b.a.0$ and $a.0|b.0 \not\equiv a.b.0 + b.a.0$.

We now state the main theorem of this section.

Theorem 6.9. (Congruence)
Stochastic bisimulation on $\mathbb{P}$ is a congruence, i.e. for arbitrary $P, P', Q, Q' \in \mathbb{P}$ and $\varepsilon \in \mathbb{A}^*$, if $P \sim P'$ and $Q \sim Q'$, then $\varepsilon.P \sim \varepsilon.P'$, $P + Q \sim P' + Q'$ and $P|Q \sim P'|Q'$.

Because $ker(\vartheta) \subseteq \sim$ and $\sim$ is a congruence for processes, we deduce that if $P \sim P'$, $Q \sim Q'$, $P \to \mu$, $P' \to \mu'$, $Q \to \nu$ and $Q' \to \nu'$, then $\mu \;{}_{P}\!\otimes_{Q}\; \nu \sim \mu' \;{}_{P'}\!\otimes_{Q'}\; \nu'$ and for any $\varepsilon \in \mathbb{A}^*$, $[{}^{\varepsilon}_{P}] \sim [{}^{\varepsilon}_{P'}]$. This shows that the quotients of $\sim$ on processes and functions produce identical signatures for both domains and an SOS format in the style of [35, 24].

7. Metrics for stochastic processes

In the case of stochastic and probabilistic systems, bisimulation is a strict concept: it verifies whether two processes have identical behaviours. In applications we need more. For instance, we want to know whether two processes that may differ by only a small amount in real-valued parameters (rates or probabilities) are behaving in a similar way. To solve this problem we define some pseudometrics on the set of processes of the minimal process algebra that will measure how much two processes are alike in terms of behaviour. In this sense, two processes are at distance zero iff they are bisimilar. Thus, the pseudometrics will be quantitative extensions of the notion of bisimulation. Similar metrics were proposed in [13, 31], exploiting the logical characterization of discrete-time Markov processes.

The behaviours of stochastic processes can be compared from two main points of view: the immediate transition rates and their future behaviour. The metrics that we propose in this section take both aspects into account. For this reason, our metrics $d^c \colon \mathbb{P} \times \mathbb{P} \to \mathbb{R}^+$ are indexed with the parameter $c \in [0, 1]$. d^1 captures only the differences between the transition rates of processes, giving equal importance to the differences between the immediate transitions and the differences that arise deeper in the evolution of the processes. On the other hand, a metric d^c with $c \in (0, 1)$ gives more weight to the rate differences that arise earlier in the evolution of the processes; as c approaches 0, the future gets discounted more, being completely ignored for $c = 0$.

The intuition behind these definitions is as follows. Assume we want to measure the distance between the processes P and Q that have the immediate transitions as represented below, where $\mathcal{P}_i, \mathcal{Q}_i \in \Pi(\sim)$ are bisimulation classes and the transitions are all by $\alpha \in \mathbb{A}^+$.

$$\begin{tikzcd} & P \arrow[dl, "\alpha{,}r_1"'] \arrow[d, "\alpha{,}r_2"] \arrow[dr, "\alpha{,}r_3"] & \\ \mathcal{P}_1 & \mathcal{P}_2 & \mathcal{P}_3 \end{tikzcd} \qquad \begin{tikzcd} & Q \arrow[dl, "\alpha{,}s_1"'] \arrow[dr, "\alpha{,}s_2"] & \\ \mathcal{Q}_1 & & \mathcal{Q}_2 \end{tikzcd}$$

For calculating the distance $d^c(P,Q)$, we first pair classes $\mathcal{P}_i$ and $\mathcal{Q}_j$ and then sum the differences between the rates of going from P and Q to $\mathcal{P}_i$ and $\mathcal{Q}_j$, respectively, and the weighted distance between arbitrary processes $P_i \in \mathcal{P}_i$ and $Q_j \in \mathcal{Q}_j$. We thus obtain, for the pair $(\mathcal{P}_i, \mathcal{Q}_j)$, the value $|r_i - s_j| + c \cdot d^c(P_i, Q_j)$. There are various ways in which one can take these pairs: d^c is the infimum of the values one can get taking all possible pairings of bisimulation classes. However, these pairings have to be one-to-one and onto on $\mathbb{P}^\sim$ and for this reason we will use the possible bijections on $\mathbb{P}^\sim$. Another observation is that we only need to consider the pairs $(\mathcal{P}_i, \mathcal{Q}_j)$, such that either P can do an α-transition to $\mathcal{P}_i$ with non-zero rate, or Q can do an α-transition to $\mathcal{Q}_j$ with non-zero rate.

In what follows we formalize these intuitions. We introduce two families of metrics on $\mathbb{P}$, $\mathbb{D}_\alpha$ for $\alpha \in \mathbb{A}^+$ and $\mathbb{D}$. The first family contains measures that concern only α-transitions, while the second considers all the transitions.

As before, for arbitrary $P, Q \in \mathbb{P}$, we write $P \Longrightarrow Q$ if there exists $\alpha \in \mathbb{A}^+$ and $r \neq 0$ such that $P \xrightarrow{\alpha,r} Q^{\equiv}$. Let $\mathcal{D}(P) = \bigcup\{Q^\sim \mid P \Longrightarrow Q\}$ be the set of derivatives of P. Let $\mathcal{B}$ be the set of bijections $\sigma : \mathbb{P}^\sim \to \mathbb{P}^\sim$. For arbitrary $P, Q, R, S \in \mathbb{P}$ and $\sigma \in \mathcal{B}$ we write $R[\sigma_Q^P]S$ if $R \in \mathcal{D}(P) \cup \sigma^{-1}(\mathcal{D}(Q))$ and $S^\sim = \sigma(R^\sim)$.

Definition 7.1. For arbitrary $\alpha \in \mathbb{A}^+$ consider the family $\mathbb{D}_\alpha$ of functions $d_\alpha^c : \mathbb{P} \times \mathbb{P} \to \mathbb{R}^+$, $c \in [0,1]$, defined, for $P, Q \in \mathbb{P}$ with $P \to \mu$ and $Q \to \eta$, by $d_\alpha^c(P,Q) = \inf_{\sigma \in \mathcal{B}}\{\sigma_\alpha(P,Q)\}$, where[3]

$$\sigma_\alpha(P,Q) = \sum_{(R^\sim, S^\sim)}^{R[\sigma_Q^P]S} \left(|\mu(\alpha)(R^\sim) - \eta(\alpha)(S^\sim)| + c \cdot d_\alpha^c(R,S)\right).$$

The correctness of this definition derives from Lemma 5.5 and from the fact that the transition tree of a derivative of a process P is strictly less complex than the transition tree of P. The same arguments guarantee that the infimum considered before is well defined.

The parameter $c \in [0,1]$ is used to associate a weight with each transition step. For instance if $a \in \mathbb{A}$, then $d_\tau^c(\tau_1.0, \tau_1.0+\tau_1.0) = |1-2| = 1$ because the first is doing a τ-transition with rate 1 and the second with rate 2, $d_\tau^c(\tau_1.0, \tau_1.0 + a.0) = |1-1| = 0$ because both are doing τ-transitions with rate 1 and for similar reasons $d_\tau^c(\tau_1.0 + \tau_1.0, \tau_1.0 + a.0) = |2-1| = 1$.

$$\begin{tikzcd}
\tau_1.0 \arrow[d, "\tau{,}1"'] & \tau_1.0+\tau_1.0 \arrow[d, "\tau{,}2"'] & \tau_1.0+a.0 \arrow[d, "\tau{,}1"'] \arrow[dr, "a{,}\iota(a)"] & \\
0^\sim & 0^\sim & 0^\sim & 0^\sim
\end{tikzcd}$$

If we prefix these three processes, we will see their difference only at the second level transitions and this will influence the measure. Thus, $d_\tau^c(\tau_2.\tau_1.0, \tau_2.(\tau_1.0 + \tau_1.0)) = |2-2| + c \cdot |2-1| = c$, $d_\tau^c(\tau_2.\tau_1.0, \tau_2.(\tau_1.0 + a.0)) = |2-2| + c \cdot |1-1| = 0$, and $d_\tau^c(\tau_2.(\tau_1.0+\tau_1.0), \tau_2.(\tau_1.0 + a.0)) = |2-2| + c \cdot |2-1| = c$.

We can use various values of $c \in [0,1]$ to give a certain weight to each transition step. Thus d_a^1 gives equal importance to the differences at each transition step, while d_a^0 is the measure that only looks to the immediate transitions. Notice also that for $a \in \mathbb{A}$ the values of d_a^c are of type $k_0 + k_1 \cdot c + k_2 \cdot c^2 + \ldots$

[3]The sum is for all pairs $(R^\sim, S^\sim)$ such that $R[\sigma_Q^P]S$.

where k_i are multiples of $\iota(a)$. This is not particularly significant, as our main issue is not the absolute value of the metric, but the significance of zero distance or the relative distance of processes.

The next lemma states that, indeed, our functions are pseudometrics.

Lemma 7.2. For any $c \in [0,1]$ and any $\alpha \in \mathbb{A}^+$, d^c_α is a pseudometric on $\mathbb{P}$.

The next theorem states that the distance between bisimilar processes is always zero. It also says that if for a fixed $c \neq 0$ the distances d^c_α between two given processes are zero for all $\alpha \in \mathbb{A}^+$, then the processes are bisimilar.

Theorem 7.3. Let $P, Q \in \mathbb{P}$.
(i) If $P \sim Q$, then for any $c \in [0,1]$ and any $\alpha \in \mathbb{A}^+$, $d^c_\alpha(P,Q) = 0$.
(ii) If there exists $c \in (0,1]$, such that for any $\alpha \in \mathbb{A}^+$, $d^c_\alpha(P,Q) = 0$, then for any $c' \in [0,1]$ and any $\alpha \in \mathbb{A}^+$ $d^{c'}_\alpha(P,Q) = 0$ and, moreover, $P \sim Q$.

Notice that the elements of $\mathbb{D}_\alpha$ measure only α-transitions and for this reason their utility is limited. Our main intention is to introduce a metric on processes that can characterize the bisimulation. For achieving this goal, in what follows we will introduce a family of metrics which consider all the transitions. The intuition is that the "general" distance d^c between two processes is the supremum of the distances d^c_α for all $\alpha \in \mathbb{A}^+$.

Definition 7.4. Consider the family $\mathbb{D}$ of functions $d^c \colon \mathbb{P} \times \mathbb{P} \to \mathbb{R}^+$ with $c \in [0,1]$, defined for arbitrary $P', P'' \in \mathbb{P}$ by $d^c(P', P'') = \sup_{\alpha \in \mathbb{A}^+}\{d^c_\alpha(P', P'')\}$.

Consider the processes $P = a.a.0 + \tau_r.\tau_r.0$ and $Q = a.(a.0 + a.0) + (\tau_r.0|\tau_r.0)$ represented below. For calculating $d^c(P,Q)$, we first observe that $d^c_a(P,Q) = |\iota(a) - \iota(a)| + c \cdot |2\iota(a) - \iota(a)| = c \cdot \iota(a)$, $d^c_\tau(P,Q) = |2r - r| + c \cdot |r - r| = r$ and for any $\alpha \notin \{a, \tau\}$, $d^c_\alpha(P,Q) = 0$.

$$
\begin{tikzcd}
 & P \arrow[dl, "{a,\iota(a)}"'] \arrow[dr, "{\tau,r}"] & \\
a.0^\sim \arrow[d, "{a,\iota(a)}"'] & & \tau_r.0^\sim \arrow[d, "{\tau,r}"'] \\
0^\sim & & 0^\sim
\end{tikzcd}
\qquad
\begin{tikzcd}
 & Q \arrow[dl, "{a,\iota(a)}"'] \arrow[dr, "{\tau,2r}"] & \\
a.0 + a.0^\sim \arrow[d, "{a,2\iota(a)}"'] & & \tau_r.0^\sim \arrow[d, "{\tau,r}"'] \\
0^\sim & & 0^\sim
\end{tikzcd}
$$

Consequently, $d^c(P,Q) = max\{c \cdot \iota(a), r\}$.

Lemma 7.5. For any $c \in [0,1]$, d^c is a pseudometric on $\mathbb{P}$.

The next theorem states that, indeed, the pseudometrics d^c generalise the bisimulation of processes. Lifted on the level of bisimulation classes, the pseudometrics became metrics and consequently, they organize the space $\mathbb{P}^\sim$ as a metric space.

Theorem 7.6. Let $P, Q \in \mathbb{P}$.
(i) If $P \sim Q$, then for any $c \in [0,1]$, $d^c(P,Q) = 0$.
(ii) If for some $c \in (0,1]$, $d^c(P,Q) = 0$, then for any $c' \in [0,1]$ $d^{c'}(P,Q) = 0$ and $P \sim Q$.

For concluding this section, we notice that the metrics are influenced by the algebraic structure of the processes. The next lemma reveals such a relation for the case of prefixing. However, we believe that more complex relations can be identified and we intend to return to this problem in future works. The possibility of computing the distance between two processes from the relative distances of their sub-processes is an idea that can find interesting applications especially in the case of large systems where it is more convenient to focus on subsystems.

Lemma 7.7. For arbitrary $P, Q \in \mathbb{P}$ and $\epsilon \in \mathbb{A}^*$, if $d^c(P,Q) = r$, then $d^c(\epsilon.P, \epsilon.Q) = max\{2{\cdot}\iota(\epsilon), c{\cdot}r\}$.

8. Related work

There is considerable existing research on probabilistic and stochastic process algebras. *Probabilistic process algebras* solve non-determinism by labeling the transitions with probabilities [39, 26, 2]. In extension, *stochastic process algebras* [18, 21, 22, 3, 32, 11] and interactive Markov chain algebra [23, 8] consider probabilistic distributions in the definition of transitions. In all these cases the transitions are defined pointwise, relying on semantics in terms of continuous-time Markov chains and probabilistic distributions over discrete spaces of processes. For correctly describing the stochastic behaviours with SOS rules of type $P \stackrel{label}{\longrightarrow} Q$, these calculi involve complex mathematical machineries for labeling and counting. By contrast, our process algebra is based on the measurable space of processes generated by structural congruence classes. Our rules are of type $P \longrightarrow \mu$ where μ is a class of distributions on the measurable space of processes. This allows us to propose an elegant SOS, similar to that of non-deterministic PAs, that maps process-structures into distribution-structures.

The idea of defining probabilistic transitions by a function that associates a probability distribution to each state of a system has been considered previously and advocated in the context of probabilistic automata [25, 34]. More recently, the *transition-systems-as-coalgebras* paradigm [12, 33] exploits the same idea, providing a general and uniform mathematical characterisation of transition systems.

The underlying Markovian structures used by our process algebra are more general than continuous-time Markov chains [23] due to the structure imposed by the equational theory. To handle this structure, we preferred to use a notion of a Markov process defined for a general measurable space and continuous time, similar to the one defined in [16]. Markov processes for arbitrary analytic spaces have been studied by Panangaden et al. in a series of papers [4, 9, 15, 14, 31] where also a notion of stochastic bisimulation, that extends the probabilistic bisimulation of [26], is defined and studied. A similar probabilistic model – Harsanyi type space – has been studied in the context of belief systems [20, 30]. Our definition of MP combines these two concepts, relying on results from [17, 31]. On the lines of [31, 14] we define the bisimulation of MPs.

The theory of GSOS [35] has been extended for the case of stochastic systems in [24], where general congruence formats for stochastic GSOS (SGSOS) are studied. The SGSOS framework, as well as GSOS, focuses on the *monads freely generated* by the algebraic signature of a process calculus. Our case is different: we have an *equational monad* because the structural congruence provides extra structure for the class of processes and thus we get a different type of SOS. In our format, for instance, the algebraic signature of processes is different from the algebraic signature of behaviors. Using a non-discrete σ-algebra makes our approach different, while considering the measurable sets closed to some congruence relation makes it more appropriate for modeling and for extensions to other *equational theories*.

Metrics for measuring the similarity of probabilistic systems in terms of behaviours have been proposed in [27, 38, 10]. In [13, 31] such metrics are introduced using a set of functional expressions, that generalise formulas of Hennessy-Milner logic, in a similar way to which Kantorovich metrics are defined by Lipschitz functions. These metrics are designed for general measurable spaces but the transitions are in discrete time. However, they have been extended in [19] for continuous-time systems (generalized semi-Markov processes). Our metrics are similar to Desharnais-Panangaden metrics, for instance in the way they explore the transition systems, but they are simpler, being particularly designed for our process algebra. We do not consider any functional expressions for calculating the distance, but we propose a direct approach.

9. Concluding remarks

In this paper we develop a stochastic extension of CCS. We propose a structural operational semantics based on measure theory and particularly suited to a domain where a measure of similarity of behaviours is important. For organizing the set of processes as a measurable space, we have chosen the σ-algebra generated by the structural congruence classes of processes and we base the theory on top of it. This choice is motivated by practical modelling reasons: the calculus is meant to be used for applications in computational systems biology. In this context, the structural congruence and the distributions over the space of congruence classes play a key role. The congruence classes represent chemical "soups" and the various syntactic representations of the same soup need to be identified. In fact, structural congruence was inspired by a chemical analogy [1].

The stochastic behaviour is defined using a general concept of Markov process that encapsulates most of the Markovian models, including continuous ones, as well as other models of probabilistic systems, e.g., Harsanyi type spaces. This concept is based on unspecified analytic (hence, measurable) spaces and generalizes rate transition systems [24, 11]. Consequently, we obtain a general definition of stochastic bisimulation similar to the one used in [16].

We also define quantitative extensions of stochastic bisimulation in the form of two classes of metrics that measure the distance between processes in terms of similar behaviours: two processes are at distance zero iff they are bisimilar; two processes are close if their behaviours are similar.

The novelty of this work consists in the fact that the measurable space of processes is axiomatized by structural congruence and the operational semantics reflects the interrelation between this space and the space of distributions on it. Our technology is appropriate for practical modelling purposes where various congruences can be relevant. It will help design (more complex) stochastic process algebras in a uniform way, possibly involving different equational axiomatizations, while avoiding the heavy techniques for counting of reductions. The organisation of the space of processes as a metric space is also a novelty. It can be extended to other calculi and used in applications, for example, to appreciate the quality of approximations of models or to characterise quantitatively the concept of robustness. For future work we intend to extend this calculus to include other algebraical operations, such as recursion or the new name quantification, and to define a general SOS format for these calculi. Another research direction that we intend to follow is the logical characterisation of bisimulation and of the metrics.

References

[1] G. Berry, G. Boudol. *The Chemical Abstract Machine*, In Proc. POPL 1990:81-94.

[2] J.A. Bergstra et al (Eds.) *Handbook of Process Algebra.* Elsevier, 2001.

[3] M. Bernardo, R. Gorrieri. *A tutorial on EMPA: A theory of concurrent processes with nondeterminism, priorities, probabilities and time.* TCS 202(1-2):1-54, 1998.

[4] R. Blute et al. *Bisimulation for labeled Markov processes.* In Proc. LICS'97, IEEE Press, 1997.

[5] M. Calder et al. *Automatically deriving ODEs from process algebra models of signaling pathways.* In Proc. CMSB'05, 2005.

[6] C. Calude, S. Marcus, G. Paun, *The universal grammar as a hypothetical brain*, Revue Roumaine de Linguistique 24 (5), 479-489, 1979

[7] L. Cardelli, R. Mardare. *The Measurable Space of Stochastic Processes.* In Proc. QEST'10 2010:171-180, IEEE Press.

[8] M. Bravetti et al. *YMCA: Why Markov Chain Algebra?*, ENTCS 162:107-112, 2006.

[9] V. Danos et al. *Bisimulation and Coconguence for Probabilistic Systems*, Inf. and Comp. 204(4):503-523, 2006.

[10] L. de Alfaro et al. *Discounting the future in systems theory.* LNCS 2719:1022-1037, 2003.

[11] R. De Nicola et al. *Rate-Based Transition Systems for Stochastic Process Calculi*. LNCS 5556:435-446, 2009

[12] E.P. de Vink, J. Rutten, *Bisimulation for probabilistic transition systems: A coalgebaic approach*. TCS 221(1-2):271-293, 1999.

[13] J. Desharnais, Labelled Markov Processes. PhD thesis, McGill University, 1999.

[14] J. Desharnais et al. *A logical characterization of bisimulation for labeled Markov processes*. In Proc LICS'98, IEEE Press, 1998.

[15] J. Desharnais et al. *Bisimulation for labeled Markov Processes*. Inf. and Comp., 179(2):163-193, 2002.

[16] J. Desharnais, P. Panangaden. *Continous Stochastic Logic Characterizes Bisimulation of Continuous-time Markov Processes*. JLAP. 56:99-115, 2003.

[17] E. Doberkat. *Stochastic Relations. Foundations for Markov Transition Systems*. Chapman and Hall/CRC, 2007.

[18] N. Gotz et al. *TIPP - A language for timed processes and performance evaluation.* Tech.Rep. 4/92 IMMD VII, University of Erlangen-Nurnberg.

[19] V. Gupta et al. *Approximate reasoning for real-time probabilistic processes.* LMCS 2(1:4) 2006.

[20] J. Harsanyi. *Games with incomplete information played by bayesian players, part one.* Management Sci., 14:159-182, 1967.

[21] J. Hillston. *A compositional approach to performance modelling.* Cambridge University Press, 1996.

[22] J. Hillston. *Process algebras for quantitative analysis.* In Proc. LICS'05, IEEE Press, 2005.

[23] H. Hermanns. *Interactive Markov Chains.* LNCS 2428, 2008.

[24] B. Klin, V. Sassone. *Structural Operational Semantics for Stochastic Process Calculi*. LNCS 4968:428-442, 2008.

[25] M. Kwiatkowska et al. *Automatic Verification of Real-Time Systems With Discrete Probability Distributions.* LNCS 1601:79-95, 1999.

[26] K. G. Larsen and A. Skou. *Bisimulation through probabilistic testing.* Inf. and Comp., 94:1-28, 1991.

[27] P. D. Lincoln et al. *A probabilistic poly-time framework for protocol analysis.* ACM(CCS-5), 1998.

[28] S. Marcus, *Membranes Versus DNA*, Fundam. Inform., 49(1-3), pp 223-227, 2002

[29] R. Milner. *A calculus of communicating systems.* J. of ACM, 1980.

[30] L.S. Moss, I.D. Viglizzo. *Harsanyi type spaces and final coalgebras constructed from satisfied theories.* ENTCS 106:279-295, 2004.

[31] P. Panangaden. *Labelled Markov Processes.* Imperial College Press, 2009.

[32] C. Priami. *Stochastic π-Calculus.* Computer Journal, 38(7):578-589, 1995.

[33] J. Rutten. *Universal coalgebra: a theory of systems*, TCS, 249:3-80, 2000.

[34] R. Segala, N. Lynch. *Probabilistic Simulations for Probabilistic Processes*, Nordic J. of Comp., 2(2):250-273, 1995.

[35] D. Turi, G.D. Plotkin. *Towards a mathematical operational semantics*, In Proc. Lics'97, IEEE Press, 1997.

[36] A.M. Turing. *Computing Machinery and Intelligence*, Mind 59, pp 433-460, 1950

[37] A.M. Turing. *The Chemical Basis of Morphogenesis*, Phil. Trans. R. Soc. London B 237 pp 37-72, 1952

[38] F. van Breugel, J. Worrell. *An algorithm for quantitative verification of probabilistic systems.* LNCS 2154:336-350, 2001.

[39] R. J. van Glabbeek et al. *Reactive, Generative and Stratified Models of Probabilistic Processes*. Inf. and Comp. 121(1):59-80, 1995.