Axiomatizations for Probabilistic Finite-State Behaviors

Yuxin Deng[1]* and Catuscia Palamidessi[2]**

[1] INRIA Sophia-Antipolis and Université Paris 7
[2] INRIA Futurs and LIX, École Polytechnique

Abstract. We study a process calculus which combines both nondeterministic and probabilistic behavior in the style of Segala and Lynch's probabilistic automata. We consider various strong and weak behavioral equivalences, and we provide complete axiomatizations for finite-state processes, restricted to guarded definitions in case of the weak equivalences. We conjecture that in the general case of unguarded recursion the "natural" weak equivalences are undecidable.

This is the first work, to our knowledge, that provides a complete axiomatization for weak equivalences in the presence of recursion and both nondeterministic and probabilistic choice.

1 Introduction

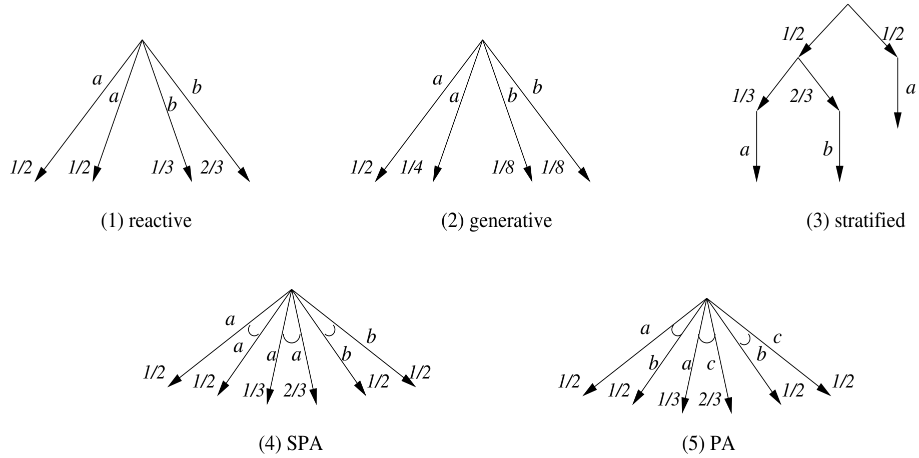
The last decade has witnessed increasing interest in the area of formal methods for the specification and analysis of probabilistic systems [11, 3, 15, 6]. In [16] van Glabbeek *et al.* classified probabilistic models into *reactive, generative* and *stratified*. In reactive models, each labeled transition is associated with a probability, and for each state the sum of the probabilities with the same label is 1. Generative models differ from reactive ones in that for each state the sum of the probabilities of all the outgoing transitions is 1. Stratified models have more structure and for each state either there is exactly one outgoing labeled transition or there are only unlabeled transitions and the sum of their probabilities is 1.

In [11] Segala pointed out that neither reactive nor generative nor stratified models capture real nondeterminism, an essential notion for modeling scheduling freedom, implementation freedom, the external environment and incomplete information. He then introduced a model, the *probabilistic automata* (PA), where both probability and nondeterminism are taken into account. Probabilistic choice is expressed by the notion of *transition*, which, in PA, leads to a probabilistic distribution over pairs (action, state) and deadlock. Nondeterministic choice, on the other hand, is expressed by the possibility of choosing different transitions.

* Supported by the EU project PROFUNDIS.
** Partially supported by the Project Rossignol of the ACI Sécurité Informatique (Ministère de la recherche et nouvelles technologies).

Segala proposed also a simplified version of PA called *simple probabilistic automata* (SPA), which are like ordinary automata except that a labeled transition leads to a probabilistic distribution over a set of states instead of a single state.

Figure 1 exemplifies the probabilistic models discussed above. In models where both probability and nondeterminism are present, like those of diagrams (4) and (5), a transition is usually represented as a bundle of arrows linked by a small arc. [13] provides a detailed comparison between the various models, and argues that PA subsume all other models above except for the stratified ones.



(1) reactive

(2) generative

(3) stratified

(4) SPA

(5) PA

Fig. 1. Probabilistic models

In this paper we are interested in investigating axiom systems for a process calculus based on PA, in the sense that the operational semantics of each expression of the language is a probabilistic automaton[3]. Axiom systems are important both at the theoretical level, as they help gaining insight of the calculus and establishing its foundations, and at the practical level, as tools for system specification and verification. Our calculus is basically a probabilistic version of the calculus used by Milner to express finite-state behaviors [8, 10].

We shall consider the two strong and the weak behavioral equivalences common in literature, plus one novel notion of weak equivalence having the advantage of being sensitive to divergency. For recursion-free expressions we provide complete axiomatizations of all the four equivalences. For the strong equivalences we also give complete axiomatizations for all expressions, while for the weak equivalences we achieve this result only for guarded expressions.

[3] Except for the case of deadlock, which is treated slightly differently: following the tradition of process calculi, in our case deadlock is a state, while in PA it is one of the possible components of a transition.

The reason why we are interested in studying a model which expresses both nondeterministic and probabilistic behavior, and an equivalence sensitive to divergency, is that one of the long-term goals of this line of research is to develop a theory which will allow us to reason about probabilistic algorithms used in distributed computing. In that domain it is important to ensure that an algorithm will work under any scheduler, and under other unknown or uncontrollable factors. The nondeterministic component of the calculus allows coping with these conditions in a uniform and elegant way. Furthermore, in many distributed computing applications it is important to ensure livelock-freedom (progress), and therefore we will need a semantics which does not simply ignore divergencies.

We end this section with a discussion about some related work. In [8] and [10] Milner gave complete axiomatizations for strong bisimulation and observational equivalence, respectively, for a core *CCS* [9]. These two papers serve as our starting point: in several completeness proofs that involve recursion we adopt Milner's *equational characterization theorem* and *unique solution theorem*. In Section 4 and Section 5.2 we extend [8] and [10] (for guarded expressions) respectively, to the setting of probabilistic process algebra.

In [14] Stark and Smolka gave a probabilistic version of the results of [8]. So, our paper extends [14] in that we consider also nondeterminism. Note that when nondeterministic choice is added, Stark and Smolka's technique of proving soundness of axioms is no longer usable. The same remark applies also to [1] which follows the approach of [14] but uses some axioms from iteration algebra to characterize recursion. In contrast, our probabilistic version of "bisimulation up to" technique works well when combined with the usual transition induction.

In [5] Bandini and Segala axiomatized both strong and weak behavioral equivalences for process calculi corresponding to SPA and to an alternated-model version of SPA. As their process calculus with non-alternating semantics corresponds to SPA, our results in Section 6 can be regarded as an extension of that work to PA.

For probabilistic process algebra of ACP-style, several complete axiom systems have appeared in the literature. However, in each of the systems either weak bisimulation is not investigated [4, 2] or nondeterministic choice is prohibited [4, 3].

2 Probabilistic Process Calculus

We begin with some preliminary notations. Let S be a set. A function $\eta : S \mapsto [0,1]$ is called a *discrete probability distribution*, or *distribution* for short, on S if the *support* of η, defined as $spt(\eta) = \{x \in S \mid \eta(x) > 0\}$, is finite or countably infinite and $\sum_{x\in S} \eta(x) = 1$. If η is a distribution with finite support and $V \subseteq spt(\eta)$ we use the set $\{(s_i : \eta(s_i))\}_{s_i \in V}$ to enumerate the probability associated with each element of V. To manipulate the set we introduce the operator $\uplus$ defined as follows.

$$\{(s_i : p_i)\}_{i\in I} \uplus \{(s : p)\} = \\ \begin{cases} \{(s_i : p_i)\}_{i\in I\setminus j} \cup \{s_j : (p_j + p)\} & \text{if } s = s_j \text{ for some } j \in I \\ \{(s_i : p_i)\}_{i\in I} \cup \{(s : p)\} & \text{otherwise.} \end{cases}$$

$$\{(s_i : p_i)\}_{i\in I} \uplus \{(t_j : p_j)\}_{j\in 1..n} = \\ (\{(s_i : p_i)\}_{i\in I} \uplus \{(t_1 : p_1)\}) \uplus \{(t_j : p_j)\}_{j\in 2..n}$$

Given some distributions $\eta_1, ..., \eta_n$ on S and some real numbers $r_1, ..., r_n \in [0, 1]$ with $\sum_{i\in 1..n} r_i = 1$, we define the *convex combination* $r_1\eta_1 + ... + r_n\eta_n$ of $\eta_1, ..., \eta_n$ to be the distribution η such that $\eta(s) = \sum_{i\in 1..n} r_i\eta_i(s)$, for each $s \in S$.

We use a countable set of variables, $\mathit{Var} = \{X, Y, ...\}$, and a countable set of atomic actions, $\mathit{Act} = \{a, b, ...\}$. Given a special action τ, we let $u, v, ...$ range over the set $\mathit{Act}_\tau = \mathit{Act} \cup \{\tau\}$, and let $\alpha, \beta, ...$ range over the set $\mathit{Var} \cup \mathit{Act}_\tau$. The class of expressions $\mathcal{E}$ is defined by the following syntax:

$$E, F ::= \bigoplus_{i\in 1..n} p_i u_i.E_i \;\Big|\; \sum_{i\in 1..m} E_i \;\Big|\; X \;\Big|\; \mu_X E$$

Here $\bigoplus_{i\in 1..n} p_i u_i.E_i$ stands for a *probabilistic choice* operator, where the p_i's represent positive probabilities, i.e., they satisfy $p_i \in (0, 1]$ and $\sum_{i\in 1..n} p_i = 1$. When $n = 0$ we abbreviate the probabilistic choice as $\mathbf{0}$; when $n = 1$ we abbreviate it as $u_1.E_1$. Sometimes we are interested in certain branches of the probabilistic choice; in this case we write $\bigoplus_{i\in 1..n} p_i u_i.E_i$ as $p_1u_1.E_1 \oplus \cdots \oplus p_nu_n.E_n$ or $(\bigoplus_{i\in 1..(n-1)} p_i u_i.E_i) \oplus p_nu_n.E_n$ where $\bigoplus_{i\in 1..(n-1)} p_i u_i.E_i$ abbreviates (with a slight abuse of notation) $p_1u_1.E_1 \oplus \cdots \oplus p_{n-1}u_{n-1}.E_{n-1}$. The construction $\sum_{i\in 1..m} E_i$ stands for *nondeterministic choice*, and occasionally we may write it as $E_1 + ... + E_m$. The notation μ_X stands for a recursion which binds the variable X. We shall use $\mathit{fv}(E)$ for the set of free variables (i.e., not bound by any μ_X) in E. As usual we identify expressions which differ only by a change of bound variables. We shall write $E\{F/X\}$ for the result of substituting F for each occurrence of X in E, renaming bound variables if necessary.

Definition 1. *The variable X is* weakly guarded *(resp.* guarded*) in E if every free occurrence of X in E occurs within some subexpression $u.F$ (resp. $a.F$), otherwise X is* weakly unguarded *(resp.* unguarded*) in E.*

The operational semantics of an expression E is defined as a probabilistic automaton whose states are the expressions reachable from E and the transition relation is defined by the axioms and inference rules in Table 1, where $E \to \eta$ describes a transition that leaves from E and leads to a distribution η over $(\mathit{Var} \cup \mathit{Act}_\tau) \times \mathcal{E}$. We shall use $\vartheta(X)$ for the special distribution $\{(X, \mathbf{0} : 1)\}$. It is evident that $E \to \vartheta(X)$ iff X is weakly unguarded in E.

The behavior of each expression can be visualized by a transition graph. For instance, the expression $(\frac{1}{2}a \oplus \frac{1}{2}b) + (\frac{1}{3}a \oplus \frac{2}{3}c) + (\frac{1}{2}b \oplus \frac{1}{2}c)$ exhibits the behavior drawn in diagram (5) of Figure 1.

Table 1. Strong transitions

var	$X \to \vartheta(X)$	psum	$\bigoplus_{i\in 1..n} p_i u_i.E_i \to \biguplus_{i\in 1..n}\{(u_i, E_i : p_i)\}$
rec	$\dfrac{E\{\mu_X E/X\} \to \eta}{\mu_X E \to \eta}$	nsum	$\dfrac{E_j \to \eta}{\sum_{i\in 1..m} E_i \to \eta}$ for some $j \in 1..m$

As in [5], we define the notion of *combined transition* as follows: $E \to_c \eta$ if there exists a collection $\{\eta_i, r_i\}_{i\in 1..n}$ of distributions and probabilities such that $\sum_{i\in 1..n} r_i = 1$, $\eta = r_1\eta_1 + ... + r_n\eta_n$ and $E \to \eta_i$, for each $i \in 1..n$.

We now introduce the notion of weak transitions, which generalizes the notion of *finitary weak transitions* in SPA [15] to the setting of PA. First we discuss the intuition behind it. Given an expression E, if we unfold its transition graph, we get a finitely branching tree. By cutting away all but one alternative in case of several nondeterministic candidates, we are left with a subtree with only probabilistic branches. A weak transition of E is a finite subtree of this kind, called *weak transition tree*, such that in any path from the root to a leaf there is at most one visible action. For example, let E be the expression $\mu_X(\frac{1}{2}a \oplus \frac{1}{2}\tau.X)$. It is represented by the transition graph displayed in Diagram (1) of Figure 2. After one unfolding, we get Diagram (2) which represents the weak transition $E \Rightarrow \eta$, where $\eta = \{(a, \mathbf{0} : \frac{3}{4}), (\tau, E : \frac{1}{4})\}$.

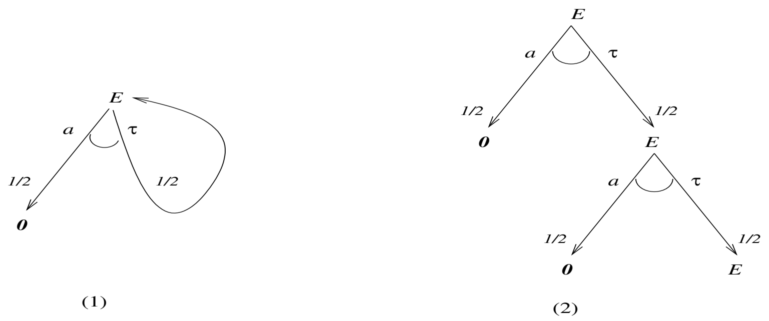


Fig. 2. A weak transition

Formally, weak transitions are defined by the rules in Table 2. Rule wea1 says that a weak transition tree starts from a bundle of labelled arrows derived from a strong transition. The meaning of Rule wea2 is as follows. Given two expressions E, F and their weak transition trees $tr(E), tr(F)$, if F is a leaf of $tr(E)$ and there is no visible action in $tr(F)$, then we can extend $tr(E)$ with $tr(F)$ at node F. If F_j is a leaf of $tr(F)$ then the probability of reaching F_j

from E is pq_j, where p and q_j are the probabilities of reaching F from E, and F_j from F, respectively. Rule wea3 is similar to Rule wea2, with the difference that we can have visible actions in $tr(F)$, but not in the path from E to F. Rule wea4 allows to construct weak transitions to unguarded variables. Note that if $E \Rightarrow \vartheta(X)$ then X is unguarded in E.

Table 2. Weak transitions

$$\mathsf{wea1}\ \ \frac{E \to \eta}{E \Rightarrow \eta}$$

$$\mathsf{wea2}\ \ \frac{E \Rightarrow \{(u_i, E_i : p_i)\}_i \uplus \{(u, F : p)\} \quad F \Rightarrow \{(\tau, F_j : q_j)\}_j}{E \Rightarrow \{(u_i, E_i : p_i)\}_i \uplus \{(u, F_j : pq_j)\}_j}$$

$$\mathsf{wea3}\ \ \frac{E \Rightarrow \{(u_i, E_i : p_i)\}_i \uplus \{(\tau, F : p)\} \quad F \Rightarrow \{(v_j, F_j : q_j)\}_j}{E \Rightarrow \{(u_i, E_i : p_i)\}_i \uplus \{(v_j, F_j : pq_j)\}_j}$$

$$\mathsf{wea4}\ \ \frac{E \Rightarrow \{(\tau, E_i : p_i)\}_i \quad \forall i : E_i \Rightarrow \vartheta(X)}{E \Rightarrow \vartheta(X)}$$

For any expression E, we use $\delta(E)$ for the unique distribution $\{(\tau, E : 1)\}$, called the *virtual distribution* of E. For any expression E, we introduce a special weak transition, called *virtual transition*, denoted by $E \overset{\epsilon}{\Rightarrow} \delta(E)$. We also define a *weak combined transition*: $E \overset{\epsilon}{\Rightarrow}_c \eta$ if there exists a collection $\{\eta_i, r_i\}_{i \in 1..n}$ of distributions and probabilities such that $\sum_{i \in 1..n} r_i = 1$, $\eta = r_1\eta_1 + ... + r_n\eta_n$ and for each $i \in 1..n$, either $E \Rightarrow \eta_i$ or $E \overset{\epsilon}{\Rightarrow} \eta_i$. We write $E \Rightarrow_c \eta$ if every component is a "normal" (i.e., non-virtual) weak transition, namely, $E \Rightarrow \eta_i$ for all $i \leq n$.

3 Behavioral Equivalences

In this section we define the behavioral equivalences that we mentioned in the introduction, namely, strong bisimulation, strong probabilistic bisimulation, divergency-sensitive equivalence and observational equivalence. We also introduce a probabilistic version of "bisimulation up to" technique to show some interesting properties of the behavioral equivalences.

3.1 Strong and Weak Equivalences

To define behavioral equivalences in probabilistic process algebra, it is customary to consider equivalence of distributions with respect to equivalence relations on processes. If η is a distribution on $S \times T$, $s \in S$ and $V \subseteq T$, we write $\eta(s, V)$ for $\sum_{t \in V} \eta(s, t)$. We lift an equivalence relation on $\mathcal{E}$ to a relation between distributions over $(Var \cup Act_\tau) \times \mathcal{E}$ in the following way.

Definition 2. *Given two distributions η_1 and η_2 over $(Var \cup Act_\tau) \times \mathcal{E}$, we say that they are equivalent w.r.t. an equivalence relation $\mathcal{R}$ on $\mathcal{E}$, written $\eta_1 \equiv_\mathcal{R} \eta_2$, if*

$$\forall \alpha \in Var \cup Act_\tau, \forall V \in \mathcal{E}/\mathcal{R} : \eta_1(\alpha, V) = \eta_2(\alpha, V).$$

Strong bisimulation is defined by requiring equivalence of distributions at every step. Because of the way equivalence of distributions is defined, we need to restrict to bisimulations which are equivalence relations.

Definition 3. *An equivalence relation $\mathcal{R} \subseteq \mathcal{E} \times \mathcal{E}$ is a* strong bisimulation *if $E\ \mathcal{R}\ F$ implies:*

- *whenever $E \longrightarrow \eta_1$, there exists η_2 such that $F \longrightarrow \eta_2$ and $\eta_1 \equiv_\mathcal{R} \eta_2$.*

We write $E \sim F$ if there exists a strong bisimulation $\mathcal{R}$ s.t. $E\ \mathcal{R}\ F$.

If we allow a strong transition to be matched by a strong combined transition, then we get a relation slightly weaker than strong bisimulation.

Definition 4. *An equivalence relation $\mathcal{R} \subseteq \mathcal{E} \times \mathcal{E}$ is a* strong probabilistic bisimulation *if $E\ \mathcal{R}\ F$ implies:*

- *whenever $E \longrightarrow \eta_1$, there exists η_2 such that $F \longrightarrow_c \eta_2$ and $\eta_1 \equiv_\mathcal{R} \eta_2$.*

$E \sim_c F$ if there exists a strong probabilistic bisimulation $\mathcal{R}$ s.t. $E\ \mathcal{R}\ F$.

We now consider the case of the weak bisimulation. The definition of weak bisimulation for PA is not at all straightforward. In fact, the "natural" weak version of Definition 3 would give rise to a relation which is not transitive. Therefore we only define the weak variant of Definition 4.

Definition 5. *An equivalence relation $\mathcal{R} \subseteq \mathcal{E} \times \mathcal{E}$ is a* weak probabilistic bisimulation *if $E\ \mathcal{R}\ F$ implies:*

- *whenever $E \longrightarrow \eta_1$, there exists η_2 such that $F \overset{\epsilon}{\Rightarrow}_c \eta_2$ and $\eta_1 \equiv_\mathcal{R} \eta_2$.*

$E \approx F$ if there exists a weak probabilistic bisimulation $\mathcal{R}$ s.t. $E\ \mathcal{R}\ F$.

As usual, observational equivalence is defined in terms of weak probabilistic bisimulation.

Definition 6. *Two expressions E and F are* observationally equivalent*, written $E \simeq F$, if*

1. *whenever $E \longrightarrow \eta_1$, there exists η_2 such that $F \Rightarrow_c \eta_2$ and $\eta_1 \equiv_\approx \eta_2$.*
2. *whenever $F \longrightarrow \eta_2$, there exists η_1 such that $E \Rightarrow_c \eta_1$ and $\eta_1 \equiv_\approx \eta_2$.*

Often observational equivalence is criticised for being insensitive to divergency. So we introduce a variant which has not this shortcoming.

Definition 7. *An equivalence relation $\mathcal{R} \subseteq \mathcal{E} \times \mathcal{E}$ is* divergency-sensitive *if $E\ \mathcal{R}\ F$ implies:*

- *whenever $E \longrightarrow \eta_1$, there exists η_2 such that $F \Rightarrow_c \eta_2$ and $\eta_1 \equiv_{\mathcal{R}} \eta_2$.*

$E \bumpeq F$ if there exists a divergency-sensitive equivalence $\mathcal{R}$ s.t. $E \mathrel{\mathcal{R}} F$.

It is easy to see that $\bumpeq$ lies between $\sim_c$ and $\simeq$. For example, we have that $\mu_X(\tau.X + a)$ and $\tau.a$ are related by $\simeq$ but not by $\bumpeq$ (this shows also that $\bumpeq$ is sensitive to divergency), while $\tau.a$ and $\tau.a + a$ are related by $\bumpeq$ but not by $\sim_c$.

One can check that all the relations defined above are indeed equivalence relations and we have the inclusion ordering: $\sim \varsubsetneq \sim_c \varsubsetneq \bumpeq \varsubsetneq \simeq \varsubsetneq \approx$.

3.2 Probabilistic "Bisimulation up to" Technique

In the classical process algebra, the conventional approach to show $E \sim F$, for some expressions E, F, is to construct a binary relation $\mathcal{R}$ which includes the pair (E, F), and then to check that $\mathcal{R}$ is a bisimulation. This approach can still be used in probabilistic process algebra, but things are more complicated because of the extra requirement that $\mathcal{R}$ must be an equivalence relation. For example we cannot use some standard set-theoretic operators to construct $\mathcal{R}$, because, even if $\mathcal{R}_1$ and $\mathcal{R}_2$ are equivalences, $\mathcal{R}_1\mathcal{R}_2$ and $\mathcal{R}_1 \cup \mathcal{R}_2$ may not be equivalences.

To avoid the restrictive condition and at the same time to reduce the size of the relation $\mathcal{R}$, we introduce the probabilistic version of "bisimulation up to" technique.

Definition 8. *A binary relation $\mathcal{R}$ is a* strong bisimulation up to $\sim$ *if $E \mathrel{\mathcal{R}} F$ implies:*

1. *whenever $E \longrightarrow \eta_1$, there exists η_2 such that $F \longrightarrow \eta_2$ and $\eta_1 \equiv_{\mathcal{R}_\sim} \eta_2$.*
2. *whenever $F \longrightarrow \eta_2$, there exists η_1 such that $E \longrightarrow \eta_1$ and $\eta_1 \equiv_{\mathcal{R}_\sim} \eta_2$.*

where $\mathcal{R}_\sim$ stands for the relation $(\mathcal{R} \cup \sim)^*$.

A strong bisimulation up to $\sim$ is not necessarily an equivalence relation. It is just an ordinary binary relation included in $\sim$.

Proposition 1. *If $\mathcal{R}$ is a strong bisimulation up to $\sim$, then $\mathcal{R} \subseteq \sim$.*

Similarly we can define strong probabilistic bisimulation up to $\sim_c$, weak probabilistic bisimulation up to $\approx$, etc. (some care is needed when dealing with weak equivalences). The "bisimulation up to" technique works well with Milner's transition induction technique [9], and by combining them we obtain the following results.

Proposition 2 (Properties of $\sim$ and $\sim_c$).

1. *$\sim$ is a congruence relation.*
2. *$\mu_X E \sim E\{\mu_X E/X\}$.*
3. *$\mu_X(E + X) \sim \mu_X E$.*
4. *If $E \sim F\{E/X\}$ and X weakly guarded in F, then $E \sim \mu_X F$.*

Properties 1-4 are also valid for $\sim_c$.

Proposition 3 (Properties of $\simeq$ and $\bumpeq$).

1. *$\simeq$ is a congruence relation.*
2. *If $\tau.E \simeq \tau.E + F$ and $\tau.F \simeq \tau.F + E$ then $\tau.E \simeq \tau.F$.*
3. *If $E \simeq F\{E/X\}$ and X is guarded in F then $E \simeq \mu_X F$.*

Properties 1-3 hold for $\bumpeq$ as well.

4 Axiomatizations for All Expressions

In this section we provide sound and complete axiomatizations for two strong behavioral equivalences: $\sim$ and $\sim_c$. The class of expressions to be considered is $\mathcal{E}$.

First we present the axiom system $\mathcal{A}_r$, which includes all axioms and rules displayed in Table 3. We assume the usual rules for equality (reflexivity, symmetry, transitivity and substitutivity), and the alpha-conversion of bound variables.

Table 3. The axiom system $\mathcal{A}_r$

S1 $E + \mathbf{0} = E$
S2 $E + E = E$
S3 $\sum_{i \in I} E_i = \sum_{i \in I} E_{\rho(i)}$ ρ is any permutation on I
S4 $\bigoplus_{i \in I} p_i u_i.E_i = \bigoplus_{i \in I} p_{\rho(i)} u_{\rho(i)}.E_{\rho(i)}$ ρ is any permutation on I
S5 $(\bigoplus_i p_i u_i.E_i) \oplus pu.E \oplus qu.E = (\bigoplus_i p_i u_i.E_i) \oplus (p+q)u.E$

R1 $\mu_X E = E\{\mu_X E/X\}$
R2 If $E = F\{E/X\}$, X weakly guarded in F, then $E = \mu_X F$
R3 $\mu_X(E + X) = \mu_X E$

The notation $\mathcal{A}_r \vdash E = F$ means that the equation $E = F$ is derivable by applying the axioms and rules from $\mathcal{A}_r$. The interest of $\mathcal{A}_r$ is that it characterizes exactly strong bisimulation, as shown by the following theorem.

Theorem 1 (Soundness and completeness of $\mathcal{A}_r$). *$E \sim E'$ iff $\mathcal{A}_r \vdash E = E'$.*

The soundness of $\mathcal{A}_r$ is easy to prove: **R1-3** correspond to clauses 2-4 of Proposition 2; **S1-4** are obvious, and **S5** is a consequence of Definition 2. For the completeness proof, the basic points are: (1) if two expressions are bisimilar then we can construct an equation set in a certain format (standard format) that they both satisfy; (2) if two expressions satisfy the same standard equation set, then they can be proved equal by $\mathcal{A}_r$. This schema is inspired by [8, 14], but in our case the definition of standard format and the proof itself are more complicated due to the presence of both probabilistic and nondeterministic dimensions.

The difference between $\sim$ and $\sim_c$ is characterized by the following axiom:

$$\mathbf{C} \quad \sum_{i\in 1..n} \bigoplus_j p_{ij} u_{ij}.E_{ij} = \sum_{i\in 1..n} \bigoplus_j p_{ij} u_{ij}.E_{ij} + \bigoplus_{i\in 1..n} \bigoplus_j r_i p_{ij} u_{ij}.E_{ij}$$

where $\sum_{i\in 1..n} r_i = 1$. We denote $\mathcal{A}_r \cup \{\mathbf{C}\}$ by $\mathcal{A}_{rc}$.

Theorem 2 (Soundness and completeness of $\mathcal{A}_{rc}$). *$E \sim_c E'$ iff $\mathcal{A}_{rc} \vdash E = E'$.*

5 Axiomatizations for Guarded Expressions

Now we proceed with the axiomatizations of the two weak behavioral equivalences: $\doteqdot$ and $\simeq$. We are not able to give a complete axiomatization for the whole set of expressions (and we conjecture that it is not possible), so we restrict to the subset of $\mathcal{E}$ consisting of *guarded expressions* only. An expression is guarded if for each of its subexpression of the form $\mu_X F$, the variable X is guarded in F (cf: Definition 1).

5.1 Axiomatizing Divergency-Sensitive Equivalence

We first study the axiom system for $\doteqdot$. As a starting point, let us consider the system $\mathcal{A}_{rc}$. Clearly, **S1-5** are still valid for $\doteqdot$, as well as **R1**. **R3** turns out to be not needed in the restricted language we are considering. As for **R2**, we replace it with its (strongly) guarded version, which we shall denote as **R2′** (see Table 4). As in the standard process algebra, we need some τ-laws to abstract from invisible steps. For $\doteqdot$ we use the probabilistic τ-laws **T1-3** shown in Table 4. Note that **T3** is the probabilistic extension of Milner's third τ-law ([10] page 231), and **T1** and **T2** together are equivalent, in the nonprobabilistic case, to Milner's second τ-law. However, Milner's first τ-law cannot be derived from **T1-3**, and it is actually unsound for $\doteqdot$. Below we let $\mathcal{A}_{gd}$ =$\{$**R2′, T1-3**$\}$ $\cup\mathcal{A}_{rc}\backslash\{$**R2-3**$\}$.

The rule **R2′** is shown to be sound in Proposition 3. The soundness of **T1-3**, and therefore of $\mathcal{A}_{gd}$, is evident. For the completeness proof, it is convenient to use the following saturation property, which relates operational semantics to term transformation.

Lemma 1. *1. If $E \Rightarrow_c \eta$ with $\eta = \{(u_i, E_i : p_i)\}_i$, then $\mathcal{A}_{gd} \vdash E = E + \bigoplus_i p_i u_i.E_i$.*
2. If $E \Rightarrow \vartheta(X)$ then $\mathcal{A}_{gd} \vdash E = E + X$.

The completeness result can be proved in a similar way as Theorem 1. The main difference is that here the key role is played by equation sets which are not only in standard format, but also saturated. The transformation of a standard equation set into a saturated one is obtained by using Lemma 1.

Theorem 3 (Soundness and completeness of $\mathcal{A}_{gd}$). *Let E and E' be two guarded expressions. Then $E \doteqdot E'$ iff $\mathcal{A}_{gd} \vdash E = E'$.*

Table 4. Some laws for the axiom system $\mathcal{A}_{gd}$

R2′ If $E = F\{E/X\}$, X guarded in F, then $E = \mu_X F$
T1 $\bigoplus_i p_i\tau.(E_i + X) = X + \bigoplus_i p_i\tau.(E_i + X)$
T2 $(\bigoplus_i p_i u_i.E_i) \oplus p\tau.(F + \bigoplus_j q_j\beta_j.F_j) + (\bigoplus_i p_i u_i.E_i) \oplus (\bigoplus_j pq_j\beta_j.F_j)$
$= (\bigoplus_i p_i u_i.E_i) \oplus p\tau.(F + \bigoplus_j q_j\beta_j.F_j)$
T3 $(\bigoplus_i p_i u_i.E_i) \oplus pu.(F + \bigoplus_j q_j\tau.F_j) + (\bigoplus_i p_i u_i.E_i) \oplus (\bigoplus_j pq_j u.F_j)$
$= (\bigoplus_i p_i u_i.E_i) \oplus pu.(F + \bigoplus_j q_j\tau.F_j)$

5.2 Axiomatizing Observational Equivalence

In this section we focus on the axiomatization of $\simeq$. In order to obtain completeness, we can follow the same schema as for Theorem 1, with the additional machinery required for dealing with observational equivalence, like in [10]. The crucial point of the proof is to show that, if $E \simeq F$, then we can construct an equation set in standard format which is satisfied by E and F. The construction of the equation is more complicated than in [10] because of the subtlety introduced by the probabilistic dimension. Indeed, it turns out that the simple probabilistic extension of Milner's three τ-laws would not be sufficient, and we need an additional rule for the completeness proof to go through. We shall further comment on this rule at the end of Section 6.

Table 5. Two τ-laws for the axiom system $\mathcal{A}_{go}$

T4 $u.\tau.E = u.E$
T5 If $\tau.E = \tau.E + F$ and $\tau.F = \tau.F + E$ then $\tau.E = \tau.F$.

The probabilistic extension of Milner's τ-laws are axioms **T1-4**, where **T1-3** are those introduced in previous section, and **T4**, defined in Table 5, takes the same form as Milner's first τ-law [10]. In the same table **T5** is the additional rule mentioned above. We let $\mathcal{A}_{go} = \mathcal{A}_{gd} \cup \{$**T4-5**$\}$.

Theorem 4 (Soundness and completeness of $\mathcal{A}_{go}$). *If E and F are guarded expressions then $E \simeq F$ iff $\mathcal{A}_{go} \vdash E = F$.*

6 Axiomatizations for Finite Expressions

In this section we consider the recursion-free fragment of $\mathcal{E}$, that is the class $\mathcal{E}_f$ of all expressions which do not contain constructs of the form $\mu_X F$. In other words all expressions in $\mathcal{E}_f$ have the form: $\sum_i \bigoplus_j p_{ij} u_{ij}.E_{ij} + \sum_k X_k$.

We define four axiom systems for the four behavioral equivalences studied in this paper. Basically $\mathcal{A}_s, \mathcal{A}_{sc}, \mathcal{A}_{fd}, \mathcal{A}_{fo}$ are obtained from $\mathcal{A}_r$, $\mathcal{A}_{rc}$, $\mathcal{A}_{gd}$, $\mathcal{A}_{go}$ respectively, by cutting away all those axioms and rules that involve recursions.

$$\begin{array}{ll} \mathcal{A}_s \overset{\text{def}}{=} \{\mathbf{S1\text{-}5}\} & \mathcal{A}_{sc} \overset{\text{def}}{=} \mathcal{A}_s \cup \{\mathbf{C}\} \\ \mathcal{A}_{fd} \overset{\text{def}}{=} \mathcal{A}_{sc} \cup \{\mathbf{T1\text{-}3}\} & \mathcal{A}_{fo} \overset{\text{def}}{=} \mathcal{A}_{fd} \cup \{\mathbf{T4\text{-}5}\} \end{array}$$

Theorem 5 (Soundness and completeness). *For any $E, F \in \mathcal{E}_f$,*

1. *$E \sim F$ iff $\mathcal{A}_s \vdash E = F$;*
2. *$E \sim_c F$ iff $\mathcal{A}_{sc} \vdash E = F$;*
3. *$E \doteq F$ iff $\mathcal{A}_{fd} \vdash E = F$;*
4. *$E \simeq F$ iff $\mathcal{A}_{fo} \vdash E = F$.*

Roughly speaking, all the clauses are proved by induction on the depth of the expressions. The completeness proof of $\mathcal{A}_{fo}$ is a bit tricky. In the classical process algebra the proof can be carried out directly by using Hennessy Lemma [9], which says that if $E \approx F$ then either $\tau.E \simeq F$ or $E \simeq F$ or $E \simeq \tau.F$. In the probabilistic case, however, Hennessy's Lemma does not hold. For example, let

$$E \overset{\text{def}}{=} a \quad \text{and} \quad F \overset{\text{def}}{=} a + (\frac{1}{2}\tau.a \oplus \frac{1}{2}a).$$

We can check that: (1) $\tau.E \not\simeq F$, (2) $E \not\simeq F$, (3) $E \not\simeq \tau.F$. In (1) the distribution $\{(\tau, E : 1)\}$ cannot be simulated by any distribution from F. In (2) the distribution $\{(\tau, a : \frac{1}{2}), (a, \mathbf{0} : \frac{1}{2})\}$ cannot be simulated by any distribution from E. In (3) the distribution $\{(\tau, F : 1)\}$ cannot be simulated by any distribution from E.

Fortunately, to prove the completeness of $\mathcal{A}_{fo}$, it is sufficient to use the following weaker property.

Lemma 2. *For any $E, F \in \mathcal{E}_f$, if $E \approx F$ then $\mathcal{A}_{fo} \vdash \tau.E = \tau.F$.*

It is worth noticing that rule **T5** is necessary to prove Lemma 2. Consider the following two expressions: $\tau.a$ and $\tau.(a + (\frac{1}{2}\tau.a \oplus \frac{1}{2}a))$. It is easy to see that they are observational equivalent. However, we cannot prove their equality if rule **T5** is excluded from the system $\mathcal{A}_{fo}$. In fact, by using only the other rules and axioms it is impossible to transform $\tau.(a + (\frac{1}{2}\tau.a \oplus \frac{1}{2}a))$ into an expression without a probabilistic branch $p\tau.a$ occurring in any subexpression, for some p with $0 < p < 1$. So it is not provably equal to $\tau.a$, which has no probabilistic choice.

7 Concluding Remarks

In this paper we have proposed a probabilistic process calculus which corresponds to Segala and Lynch's probabilistic automata. We have presented strong bisimulation, strong probabilistic bisimulation, divergency-sensitive equivalence

Table 6. All the axioms and rules

S1 $E + \mathbf{0} = E$
S2 $E + E = E$
S3 $\sum_{i\in I} E_i = \sum_{i\in I} E_{\rho(i)}$ $\quad \rho$ is any permutation on I
S4 $\bigoplus_{i\in I} p_i u_i.E_i = \bigoplus_{i\in I} p_{\rho(i)} u_{\rho(i)}.E_{\rho(i)}$ $\quad \rho$ is any permutation on I
S5 $(\bigoplus_i p_i u_i.E_i) \oplus pu.E \oplus qu.E = (\bigoplus_i p_i u_i.E_i) \oplus (p+q)u.E$

C $\sum_{i\in 1..n} \oplus_j p_{ij} u_{ij}.E_{ij} = \sum_{i\in 1..n} \oplus_j p_{ij} u_{ij}.E_{ij} + \oplus_{i\in 1..n} \oplus_j r_i p_{ij} u_{ij}.E_{ij}$

T1 $\bigoplus_i p_i\tau.(E_i + X) = X + \bigoplus_i p_i\tau.(E_i + X)$
T2 $(\bigoplus_i p_i u_i.E_i) \oplus p\tau.(F + \bigoplus_j q_j\beta_j.F_j) + (\bigoplus_i p_i u_i.E_i) \oplus (\bigoplus_j pq_j\beta_j.F_j)$
$= (\bigoplus_i p_i u_i.E_i) \oplus p\tau.(F + \bigoplus_j q_j\beta_j.F_j)$
T3 $(\bigoplus_i p_i u_i.E_i) \oplus pu.(F + \bigoplus_j q_j\tau.F_j) + (\bigoplus_i p_i u_i.E_i) \oplus (\bigoplus_j pq_j u.F_j)$
$= (\bigoplus_i p_i u_i.E_i) \oplus pu.(F + \bigoplus_j q_j\tau.F_j)$
T4 $u.\tau.E = u.E$
T5 If $\tau.E = \tau.E + F$ and $\tau.F = \tau.F + E$ then $\tau.E = \tau.F$.

R1 $\mu_X E = E\{\mu_X E/X\}$
R2 If $E = F\{E/X\}$, X weakly guarded in F, then $E = \mu_X F$
R2′ If $E = F\{E/X\}$, X guarded in F, then $E = \mu_X F$
R3 $\mu_X(E + X) = \mu_X E$

In **C**, there is a side condition $\sum_{i\in 1..n} r_i = 1$.

Table 7. All the inference systems

strong equivalences	finite expressions	all expressions
$\sim$	$\mathcal{A}_s$: **S1-5**	$\mathcal{A}_r$: **S1-5,R1-3**
$\sim_c$	$\mathcal{A}_{sc}$: **S1-5,C**	$\mathcal{A}_{rc}$: **S1-5,R1-3,C**

weak equivalences	finite expressions	guarded expressions
$\overset{\frown}{=}$	$\mathcal{A}_{fd}$: **S1-5,C,T1-3**	$\mathcal{A}_{gd}$: **S1-5,C,T1-3,R1,R2′**
$\simeq$	$\mathcal{A}_{fo}$: **S1-5,C,T1-5**	$\mathcal{A}_{go}$: **S1-5,C,T1-5,R1,R2′**

and observational equivalence. Sound and complete inference systems for the four behavioral equivalences are summarized in Table 7.

Note that we have axiomatized divergency-sensitive equivalence and observational equivalence only for guarded expressions. For unguarded expressions whose transition graphs include τ-loops, we conjecture that the two behavioral equivalences are undecidable and therefore not finitely axiomatizable. The reason is the following: in order to decide whether two expressions E and F are observational equivalent, one can compute the two sets

$$S_E = \{\eta \mid E \Rightarrow \eta\} \quad \text{and} \quad S_F = \{\eta \mid F \Rightarrow \eta\}$$

and then compare them to see whether each element of S_E is related to some element of S_F and vice versa. For guarded expressions E and F, the sets S_E and S_F are always finite and thus they can be compared in finite time. For unguarded expressions, these sets may be infinite, and so the above method does not apply. Furthermore, these sets can be infinite even when we factorize them with respect to an equivalence relation as required in the definition of probabilistic bisimulation. For example, consider the expression $E = \mu_X(\frac{1}{2}a \oplus \frac{1}{2}\tau.X)$. It can be proved that S_E is an infinite set $\{\eta_i \mid i \geq 1\}$, where

$$\eta_i = \{(a, \mathbf{0} : (1 - \frac{1}{2^i})), (\tau, E : \frac{1}{2^i})\}.$$

Furthermore, for each $i, j \geq 1$ with $i \neq j$ we have $\eta_i \not\equiv_{\mathcal{R}} \eta_j$ for any equivalence relation $\mathcal{R}$ which distinguishes E from $\mathbf{0}$. Hence the set S_E modulo $\mathcal{R}$ is infinite.

It should be remarked that the presence of τ-loops in itself does not necessarily cause non-decidability. For instance, the notion of weak probabilistic bisimulation defined in [11, 6] is decidable for finite-state PA. The reason is that in those works weak transitions are defined in terms of schedulers, and one may get some weak transitions that are not derivable by the (finitary) inference rules used in this paper. For instance, consider the transition graph of the above example. The definition of [11, 6] allows the underlying probabilistic execution to be infinite as long as that case occurs with probability 0. Hence with that definition one has a weak transition that leads to the distribution $\theta = \{(a, \mathbf{0} : 1)\}$. Thus each η_i becomes a convex combination of θ and $\delta(E)$, i.e. these two distributions are enough to characterize all possible weak transitions. By exploiting this property, Cattani and Segala gave a decision algorithm for weak probabilistic bisimulation in [6].

In this paper we have chosen, instead, to generate weak transitions via (finitary) inference rules, which means that only finite executions can be derived. This approach, which is also known in literature ([12]), has the advantage of being more formal, and in the case of guarded recursion it is equivalent to the one of [11, 6]. In the case of unguarded recursion, however, we feel that it would be more natural to consider also the "limit" weak transitions of [11, 6]. The axiomatization of the corresponding notion of observational equivalence is an open problem.

References

1. L. Aceto, Z. Ésik, and A. Ingólfsdóttir. Equational axioms for probabilistic bisimilarity (preliminary report). Technical Report RS-02-6, BRICS, Feb. 2002.
2. S. Andova. Process algebra with probabilistic choice. Technical Report CSR 99-12, Eindhoven University of Technology, 1999.
3. S. Andova and J. C. M. Baeten. Abstraction in probabilistic process algebra. In *Tools and Algorithms for the Construction and Analysis of Systems*, volume 2031 of *LNCS*, pages 204–219. Springer, 2001.
4. J. C. M. Baeten, J. A. Bergstra, and S. A. Smolka. Axiomatizing probabilistic processes: ACP with generative probabilities. *Information and Computation*, 121(2):234–255, 1995.
5. E. Bandini and R. Segala. Axiomatizations for probabilistic bisimulation. In *Proceedings of the 28th International Colloquium on Automata, Languages and Programming*, volume 2076 of *LNCS*, pages 370–381. Springer, 2001.
6. S. Cattani and R. Segala. Decision algorithms for probabilistic bisimulation. In *Proceedings of the 13th International Conference on Concurrency Theory*, volume 2421 of *LNCS*, pages 371–385. Springer, 2002.
7. P. R. D'Argenio, H. Hermanns, and J.-P. Katoen. On generative parallel composition. *ENTCS*, 22, 1999.
8. R. Milner. A complete inference system for a class of regular behaviours. *Journal of Computer and System Science*, 28:439–466, 1984.
9. R. Milner. *Communication and Concurrency*. Prentice-Hall, 1989.
10. R. Milner. A complete axiomatisation for observational congruence of finite-state behaviours. *Information and Computation*, 81:227–247, 1989.
11. R. Segala. Modeling and verification of randomized distributed real-time systems. Technical Report MIT/LCS/TR-676, PhD thesis, MIT, Dept. of EECS, 1995.
12. Roberto Segala and Nancy Lynch. Probabilistic simulations for probabilistic processes. In *Proceedings of the 5th International Conference on Concurrency Theory*, volume 836 of *LNCS*, pages 481–496. Springer-Verlag, 1994.
13. A. Sokolova and E. de Vink. Probabilistic automata: system types, parallel composition and comparison. In *Validation of Stochastic Systems: A Guide to Current Research*, volume 2925 of *LNCS*, pages 1–43. Springer, 2004.
14. E. W. Stark and S. A. Smolka. A complete axiom system for finite-state probabilistic processes. In *Proof, language, and interaction: essays in honour of Robin Milner*, pages 571–595. MIT Press, 2000.
15. M. Stoelinga. *Alea jacta est: verification of probabilistic, real-time and parametric systems*. PhD thesis, University of Nijmegen, 2002.
16. R. J. van Glabbeek, S. A. Smolka, and B. Steffen. Reactive, generative, and stratified models of probabilistic processes. *Information and Computation*, 121(1):59–80, 1995.