Axiomatizations for probabilistic finite-state behaviors

Yuxin Deng, Catuscia Palamidessi

Axiomatizations for probabilistic finite-state behaviors ⋆

Yuxin Deng [a], Catuscia Palamidessi [b]

[a] *Shanghai Jiao Tong University, China*

[b] *INRIA Futurs and LIX, École Polytechnique, France*

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 recursion 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.

Key words: Process calculus, Probability, Nondeterminism, Axiomatizations, Behavioral equivalences

1 Introduction

The last decade has witnessed increasing interest in the area of formal methods for the specification and analysis of probabilistic systems [22,5,3,20,26,7]. In [28] van Glabbeek *et al.* classified probabilistic models into *reactive, generative*

⋆ The work of Yuxin Deng was done when he was doing his PhD study at INRIA and Université Paris 7, France, under the support of the EU project PROFUNDIS. The work of Catuscia Palamidessi was partially supported by the Project Rossignol of the ACI Sécurité Informatique (Ministère de la recherche et nouvelles technologies). An extended abstract of this paper appeared at FOSSACS 2005.

Email addresses: `deng-yx@cs.sjtu.edu.cn` (Yuxin Deng), `catuscia@lix.polytechnique.fr` (Catuscia Palamidessi).

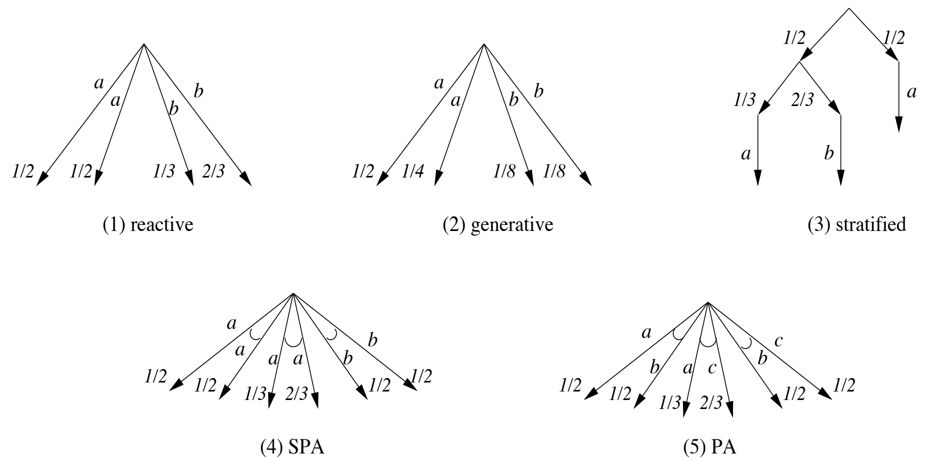

Fig. 1. Probabilistic models

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 [22] 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. 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. [24] provides a detailed comparison between the various models, and argues that PA subsume all other models above except for the stratified ones.

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 [1]. Axiom systems are important both at the theoretical level, as they help to gain insight into the calculus and establish its foundations, and at the practical level, as tools for systems specification and verification. Our calculus is basically a probabilistic version of the core CCS used by Milner to express finite-state behaviors [13,15].

We shall consider four types of behavioral equivalences: two strong bisimulation equivalences, a weak equivalence sensitive to divergence, and observational equivalence. 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.

The reason why we are interested in studying a model which expresses both nondeterministic and probabilistic behavior, and an equivalence sensitive to divergence, 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 enables one to deal 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 are interested, in particular, in developing a fully distributed implementation of the (synchronous) π-calculus (π) [16,21] using a probabilistic asynchronous π-calculus (π_{pa}) [12] as an intermediate language. The reason why we need a probabilistic calculus is that it has been shown impossible to implement certain mechanisms of the π-calculus without using randomization [18]. We need also the nondeterministic dimension for the usual reasons: the implementation should be portable and in particular make no assumption about the scheduler. Some preliminary initial results of this project appeared in [19], where preliminary results on implementation were reported. We are now investigating a more realistic and efficient implementation.

We consider it important that an implementation does not introduce livelocks (or other kinds of unintended outcomes), thus the translation from π to π_{pa} should preserve livelock-freedom (see [19] for a discussion on the subject), and hence the semantics should be sensitive to divergence. For this reason, the second author chose (a probabilistic version of) testing semantics in [19]. However, it turned out that probabilistic testing semantics, at least the version

[1] 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.

invented in [19], was rather difficult to use. The correctness proofs were ad-hoc, by hand, and rather complicated. For the realistic (and necessarily more sophisticated) implementation, we need proof methods feasible and (at least in part) automatic. For this reason, we are investigating here a divergence-sensitive *bisimulation-like* semantics. In the future, we plan to extend the results of this paper to π_{pa}.

Related work

In [13] and [15] Milner gave complete axiomatizations for strong bisimulation and observational equivalence, respectively, for a core CCS [14]. 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 5.1 and Section 6.2 we extend [13] and [15] (for guarded expressions), respectively, to the setting of probabilistic process algebra.

In [25] Stark and Smolka gave a probabilistic version of the results of [13]. So, our paper extends [25] in that we consider also nondeterminism. Note that, when nondeterministic choice is added, Stark and Smolka's technique of proving soundness of axioms can no longer be used. (See the discussion at the beginning of Appendix A.) The same remark applies also to [1] which follows the approach of [25] 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 [17] Mislove *et al* presented a domain model for a process algebra with both probabilistic and nondeterministic choice. Their model is fully abstract with respect to a strong bisimilarity, for which they provided a complete axiomatization. However, weak behavioural equivalences are not considered in that paper.

In [6] Bandini and Segala axiomatized both strong and weak behavioral equivalences for process calculi corresponding to SPA and to an alternating model version of SPA. As their process calculus with non-alternating semantics corresponds to SPA, our results in Section 7 can be regarded as an extension of that work to PA.

For probabilistic process algebra of based on ACP, 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].

Contribution of this work

The original contributions of this paper are:

- A complete axiomatization of a calculus which contains both nondeterministic and probabilistic choice, and recursion. We axiomatize both strong and weak behavioral equivalences. This is the first time, as far as we know, that a complete axiomatization of weak behavioral equivalences is presented for a language of this kind.
- The development and the axiomatization of a (probabilistic) weak behavioral equivalence sensitive to divergence.

Plan of the paper

In the next section we briefly recall some basic concepts and definitions about probability distributions. In Section 3 we introduce the calculus, with its syntax and operational semantics. In Section 4 we define the four behavioral equivalences we are interested in, and we extend the "bisimulation up to" technique of [14] to the probabilistic case. This technique is used extensively for the proofs of soundness of some axioms, especially in the case of the weak equivalences. In Sections 5 and 6 we give complete axiomatizations for the strong equivalences and for the weak equivalences respectively, restricted to guarded expressions in the second case. Section 7 gives complete axiomatizations for the four equivalences in the case of the finite fragment of the language. The interest of this section is that we use different and much simpler proof techniques than those in Sections 5 and 6. Finally, Section 8 concludes and illustrates our research plans.

2 Preliminaries

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 $r_1, ..., r_n \in [0, 1]$ such that $\sum_{i\in 1..n} r_i = 1$, 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 $s \in S$.

3 Probabilistic process calculus

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 second 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 $E\{F_1, ..., F_n/X_1, ..., X_n\}$ or $E\{\widetilde{F}/\widetilde{X}\}$ for the result of simultaneously substituting F_i for each occurrence of X_i in E $(1 \le i \le n)$, renaming bound variables if necessary.

Definition 1 *The variable X is* weakly guarded *(resp.* guarded*) in E if every*

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 $j \in 1..m$

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 $(Var \cup Act_\tau) \times \mathcal{E}$. 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.

As in [6], 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 [26] 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

$$E \stackrel{\text{def}}{=} \mu_X(\frac{1}{2}a \oplus \frac{1}{2}\tau.X) \tag{1}$$

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

$$E \Rightarrow \{(a, \mathbf{0} : \frac{3}{4}), (\tau, E : \frac{1}{4})\}. \tag{2}$$

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

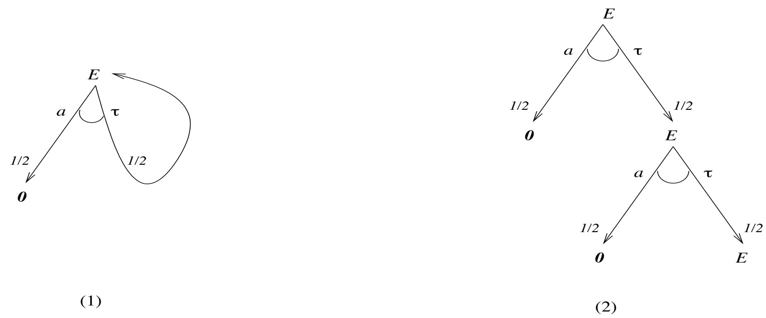

Fig. 2. A weak transition

Table 2
Weak transitions

wea1	$$\frac{E \to \eta}{E \Rightarrow \eta}$$
wea2	$$\frac{E \Rightarrow \{(u_i, E_i : p_i)\}_i \uplus \{(u, F : p)\} \quad F \Rightarrow \{(\tau, F_j : q_j)\}_j}{E \Rightarrow \{(u_i, E_i : p_i)\}_i \uplus \{(u, F_j : pq_j)\}_j}$$
wea3	$$\frac{E \Rightarrow \{(u_i, E_i : p_i)\}_i \uplus \{(\tau, F : p)\} \quad F \Rightarrow \{(v_j, F_j : q_j)\}_j}{E \Rightarrow \{(u_i, E_i : p_i)\}_i \uplus \{(v_j, F_j : pq_j)\}_j}$$
wea4	$$\frac{E \Rightarrow \{(\tau, E_i : p_i)\}_i \quad \forall i : E_i \Rightarrow \vartheta(X)}{E \Rightarrow \vartheta(X)}$$

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.

As an example of applying these transition rules, we consider the expression E in (1). Using rules rec and wea1, we can infer the following transitions.

$$\cfrac{\cfrac{\frac{1}{2}a \oplus \frac{1}{2}\tau.E \to \{(a, \mathbf{0} : \frac{1}{2}), (\tau, E : \frac{1}{2})\}}{E \to \{(a, \mathbf{0} : \frac{1}{2}), (\tau, E : \frac{1}{2})\}}}{E \Rightarrow \{(a, \mathbf{0} : \frac{1}{2}), (\tau, E : \frac{1}{2})\}}$$

Note that $\{(a, \mathbf{0} : \frac{1}{2}), (\tau, E : \frac{1}{2})\} = \{(a, \mathbf{0} : \frac{1}{2})\} \uplus \{(\tau, E : \frac{1}{2})\}$, so we can appeal to wea3 and do the following inference.

$$\frac{E \Rightarrow \{(a, \mathbf{0} : \frac{1}{2})\} \uplus \{(\tau, E : \frac{1}{2})\} \quad E \Rightarrow \{(a, \mathbf{0} : \frac{1}{2}), (\tau, E : \frac{1}{2})\}}{E \Rightarrow \{(a, \mathbf{0} : \frac{1}{2})\} \uplus \{(a, \mathbf{0} : \frac{1}{4}), (\tau, E : \frac{1}{4})\}}$$

Since $\{(a, \mathbf{0} : \frac{1}{2})\} \uplus \{(a, \mathbf{0} : \frac{1}{4}), (\tau, E : \frac{1}{4})\} = \{(a, \mathbf{0} : \frac{3}{4}), (\tau, E : \frac{1}{4})\}$, we have established (2).

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 $E \stackrel{\epsilon}{\Rightarrow} \delta(E)$. We also define a *weak combined transition*: $E \stackrel{\epsilon}{\Rightarrow}_c \eta$ 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 $E \stackrel{\epsilon}{\Rightarrow} \eta_i$. 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 $i \leq n$.

4 Behavioral equivalences

In this section we define the behavioral equivalences that we mentioned in the introduction, namely, strong bisimulation, strong probabilistic bisimulation, divergence-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.

To define behavioral equivalences in probabilistic process algebra, it is customary to consider equivalence of distributions with respect to equivalence relations on processes.

4.1 Equivalence of distributions

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*

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

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

Two expressions E, F *are* strong bisimilar*, written* $E \sim F$*, 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* $E \rightarrow \eta_1$*, there exists* η_2 *such that* $F \rightarrow_c \eta_2$ *and* $\eta_1 \equiv_{\mathcal{R}} \eta_2$.

We write $E \sim_c F$*, 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 be the following one.

Definition (Tentative). *An equivalence relation* $\mathcal{R} \subseteq \mathcal{E} \times \mathcal{E}$ *is a* weak bisimulation *if* $E\ \mathcal{R}\ F$ *implies:*
- *whenever* $E \rightarrow \eta_1$*, then there exists* η_2 *such that either* $F \Rightarrow \eta_2$ *or* $F \stackrel{\epsilon}{\Rightarrow} \eta_2$*, and* $\eta_1 \equiv_{\mathcal{R}} \eta_2$.

E *and* F *are* weak bisimilar*, written* $E \asymp F$*, whenever there exists a weak bisimulation* $\mathcal{R}$ *s.t.* $E\ \mathcal{R}\ F$.

Unfortunately the above definition is incorrect because it defines a relation which is not transitive. That is, there exist E, F and G with $E \asymp F$ and $F \asymp G$ but $E \not\asymp G$. For example, consider the following expressions and relations:

$$E \stackrel{\text{def}}{=} \tau.a + (\tfrac{1}{2}\tau.(a+a) \oplus \tfrac{1}{2}a)$$
$$F \stackrel{\text{def}}{=} \tfrac{1}{2}\tau.(a+a) \oplus \tfrac{1}{2}\tau.a$$
$$G \stackrel{\text{def}}{=} \tau.a$$

$$\mathcal{R}_1 \stackrel{\text{def}}{=} \{(E,F),(F,E),(E,E),(F,F),(a+a,a+a),(a+a,a),(a,a+a),(a,a),(\mathbf{0},\mathbf{0})\}$$
$$\mathcal{R}_2 \stackrel{\text{def}}{=} \{(F,G),(G,F),(F,F),(G,G),(a+a,a+a),(a+a,a),(a,a+a),(a,a),(\mathbf{0},\mathbf{0})\}$$

It can be checked that $\mathcal{R}_1$ and $\mathcal{R}_2$ are weak bisimulations according to the tentative definition. However we have $E \not\approx G$. To see this, consider the transition $E \rightarrow \eta$, where $\eta = \{(\tau, a+a : \frac{1}{2}),(a,\mathbf{0} : \frac{1}{2})\}$. There are only three possible weak transitions from $G : G \stackrel{\epsilon}{\Rightarrow} \delta(G)$, $G \Rightarrow \eta_1$ and $G \Rightarrow \eta_2$ where $\eta_1 = \{(\tau, a : 1)\}$ and $\eta_2 = \{(a, \mathbf{0} : 1)\}$. Now, among the three distributions η_1, η_2 and $\delta(G)$, none is equivalent to η. Therefore, E and G are not bisimilar. Nevertheless, if we consider the weak combined transition: $G \Rightarrow_c \eta'$ where $\eta' = \frac{1}{2}\eta_1 + \frac{1}{2}\eta_2$, we observe that $\eta \equiv \eta'$.

The above example suggests that for a "good" definition of weak bisimulation it is necessary to use combined transitions. So we cannot give a weak variant of Definition 3, but only of Definition 4, called weak probabilistic bisimulation.

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 $E \rightarrow \eta_1$, there exists η_2 such that $F \stackrel{\epsilon}{\Rightarrow}_c \eta_2$ and $\eta_1 \equiv_{\mathcal{R}} \eta_2$.*

We write $E \approx F$ whenever 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, F are* observationally equivalent*, written $E \simeq F$, if*

(1) whenever $E \rightarrow \eta_1$, there exists η_2 such that $F \Rightarrow_c \eta_2$ and $\eta_1 \equiv_{\approx} \eta_2$.
(2) whenever $F \rightarrow \eta_2$, 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 divergence. We therefore introduce a variant which does not have this shortcoming.

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

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

We write $E \doteq F$ *whenever there exists a divergence-sensitive equivalence* $\mathcal{R}$ *s.t.* $E \; \mathcal{R} \; F$.

It is easy to see that $\doteq$ 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 $\doteq$ (this shows also that $\doteq$ is sensitive to divergence), while $\tau.a$ and $\tau.a + a$ are related by $\doteq$ but not by $\sim_c$.

One can check that all the relations defined above (except for $\asymp$) are indeed equivalence relations and we have the inclusion ordering: $\sim \subsetneq \sim_c \subsetneq \doteq \subsetneq \simeq \subsetneq \approx$.

4.3 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, whose usefulness will be exhibited in the next subsection.

In the following definitions, for a binary relation $\mathcal{R}$ we denote the relation $(\mathcal{R} \cup \sim)^*$ by $\mathcal{R}_\sim$. Similarly for other relations such as $\mathcal{R}_\approx$ and $\mathcal{R}_\simeq$.

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

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

A strong bisimulation up to $\sim$ is not necessarily an equivalence relation. It is just an ordinary binary relation included in $\sim$, as shown by the next proposition.

Proposition 9 *If $\mathcal{R}$ is a strong bisimulation up to $\sim$, then $\mathcal{R} \subseteq \sim$.*

One can also define a strong probabilistic bisimulation up to $\sim_c$ relation and show that it is included in $\sim_c$. For weak probabilistic bisimulation, the "up to" relation can be defined as well, but we need to be careful.

Definition 10 *A binary relation $\mathcal{R}$ is a* weak probabilistic bisimulation up to $\approx$ *if $E \ \mathcal{R} \ F$ implies:*

(1) whenever $E \Rightarrow \eta_1$, there exists η_2 such that $F \stackrel{\epsilon}{\Rightarrow}_c \eta_2$ and $\eta_1 \equiv_{\mathcal{R}_\approx} \eta_2$.
(2) whenever $F \Rightarrow \eta_2$, there exists η_1 such that $E \stackrel{\epsilon}{\Rightarrow}_c \eta_1$ and $\eta_1 \equiv_{\mathcal{R}_\approx} \eta_2$.

In the above definition, we are not able to replace the first double arrow in each clause by a simple arrow. Otherwise, the resulting relation is not included in $\approx$.

Proposition 11 *If $\mathcal{R}$ is a weak probabilistic bisimulation up to $\approx$, then $\mathcal{R} \subseteq \approx$.*

Definition 12 *A binary relation $\mathcal{R}$ is an* observational equivalence up to $\simeq$ *if $E \ \mathcal{R} \ F$ implies:*

(1) whenever $E \Rightarrow \eta_1$, there exists η_2 such that $F \Rightarrow_c \eta_2$ and $\eta_1 \equiv_{\mathcal{R}_\approx} \eta_2$.
(2) whenever $F \Rightarrow \eta_2$, there exists η_1 such that $E \Rightarrow_c \eta_1$ and $\eta_1 \equiv_{\mathcal{R}_\approx} \eta_2$.

As expected, observational equivalence up to $\simeq$ is useful because of the following property.

Proposition 13 *If $\mathcal{R}$ is an observational equivalence up to $\simeq$, then $\mathcal{R} \subseteq \simeq$.*

4.4 Some properties of behavioral equivalences

The "bisimulation up to" technique works well with Milner's transition induction technique [14], and by combining them we obtain the following results for the calculus introduced in Section 3.

Proposition 14 (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 $\sim_c$.

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

S1	$E + \mathbf{0} = E$
S2	$E + E = E$
S3	$\sum_{i\in I} E_i = \sum_{i\in I} E_{\rho(i)}$ $\quad\rho$ 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	$(\bigoplus_i p_i u_i.E_i) \oplus pu.E \oplus qu.E = (\bigoplus_i p_i u_i.E_i) \oplus (p+q)u.E$
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	$\mu_X(E + X) = \mu_X E$

Proposition 15 (Properties of $\simeq$ and $\doteqdot$) *(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 $\doteqdot$ as well.

Each property above is shown by exhibiting an equivalence up to the corresponding bisimulation relation. For instance, in Clause 3 of Proposition 15 we prove that the relation $\mathcal{R} = \{(G\{E/X\}, G\{\mu_X F/X\}) \mid \text{for any } G \in \mathcal{E}\}$ is an observational equivalence up to $\simeq$ by transition induction (see Appendix A for more details). We find it necessary to use the "bisimulation up to" technique particularly in the cases of Properties 1 and 3 of Proposition 15, since we are not able to directly construct an equivalence relation and prove that it is an observational equivalence. In all other cases the "up to" technique is optional.

5 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 $\mathcal{E}$.

5.1 Axiomatizing strong bisimulation

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.

The notation $\mathcal{A}_r \vdash E = F$ (and $\mathcal{A}_r \vdash \widetilde{E} = \widetilde{F}$ for a finite sequence of equations) means that the equation $E = F$ is derivable by applying the axioms and rules from $\mathcal{A}_r$. The following theorem shows that $\mathcal{A}_r$ is sound with respect to $\sim$.

Theorem 16 (Soundness of $\mathcal{A}_r$) *If $\mathcal{A}_r \vdash E = E'$ then $E \sim E'$.*

Proof. The soundness of the recursion axioms **R1-3** is shown in Proposition 14; the soundness of **S1-4** is obvious, and **S5** is a consequence of Definition 2. $\square$

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 [13,25], 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.

Definition 17 *Let $\widetilde{X} = \{X_1, ..., X_m\}$ and $\widetilde{W} = \{W_1, W_2, ...\}$ be disjoint sets of variables. Let $\widetilde{H} = \{H_1, ..., H_m\}$ be expressions with free variables in $\widetilde{X} \cup \widetilde{W}$. In the equation set $S : \widetilde{X} = \widetilde{H}$, we call $\widetilde{X}$* formal variables *and $\widetilde{W}$* free variables*. We say S is* standard *if each H_i takes the form $\sum_j E_{f(i,j)} + \sum_l W_{h(i,l)}$ where $E_{f(i,j)} = \bigoplus_k p_{f(i,j,k)} u_{f(i,j,k)}.X_{g(i,j,k)}$. We call S* weakly guarded *if there is no H_i s.t. $H_i \rightarrow \vartheta(X_i)$. We say that E* provably satisfies *S if there are expressions $\widetilde{E} = \{E_1, ..., E_m\}$, with $E_1 \equiv E$ and $fv(\widetilde{E}) \subseteq \widetilde{W}$, such that $\mathcal{A}_r \vdash \widetilde{E} = \widetilde{H}\{\widetilde{E}/\widetilde{X}\}$.*

We first recall the theorem of unique solution of equations, which originally appeared in [13]. Adding probabilistic choice does not affect the validity of this theorem.

Theorem 18 (Unique solution of equations I) *If S is a weakly guarded equation set with free variables in $\widetilde{W}$, then there is an expression E which provably satisfies S. Moreover, if F provably satisfies S and has free variables in $\widetilde{W}$, then $\mathcal{A}_r \vdash E = F$.*

Proof. Exactly the same as in [13]. $\square$

Below we give an extension of Milner's equational characterization theorem by accommodating probabilistic choice.

Theorem 19 (Equational characterization I) *For any expression E, with free variables in $\widetilde{W}$, there exist some expressions $\widetilde{E} = \{E_1, ..., E_m\}$, with*

$E_1 \equiv E$ and $fv(\widetilde{E}) \subseteq \widetilde{W}$, satisfying m equations

$$\mathcal{A}_r \vdash E_i = \sum_{j\in 1..n(i)} E_{f(i,j)} + \sum_{j\in 1..l(i)} W_{h(i,j)} \qquad (i \leq m)$$

where $E_{f(i,j)} \equiv \bigoplus_{k\in 1..o(i,j)} p_{f(i,j,k)} u_{f(i,j,k)}.E_{g(i,j,k)}$.

Proof. By induction on the structure of E, similar to the proof in [13]. □

The following completeness proof is closely analogous to that of [25]. It is complicated somewhat by the presence of nondeterministic choice. For example, to construct the formal equations, we need to consider a more refined relation $L_{iji'j'}$ underneath the relation $K_{ii'}$ while in [13,25] it is sufficient to just use $K_{ii'}$.

Theorem 20 (Completeness of $\mathcal{A}_r$) *If $E \sim E'$ then $\mathcal{A}_r \vdash E = E'$.*

Proof. Let E and E' have free variables in $\widetilde{W}$. By Theorem 19 there are provable equations such that $E \equiv E_1$, $E' \equiv E'_1$ and

$$\mathcal{A}_r \vdash E_i = \sum_{j\in 1..n(i)} E_{f(i,j)} + \sum_{j\in 1..l(i)} W_{h(i,j)} \qquad (i \leq m)$$

$$\mathcal{A}_r \vdash E'_{i'} = \sum_{j'\in 1..n'(i')} E'_{f'(i',j')} + \sum_{j'\in 1..l'(i')} W_{h'(i',j')} \qquad (i' \leq m')$$

with

$$E_{f(i,j)} \equiv \bigoplus_{k\in 1..o(i,j)} p_{f(i,j,k)} u_{f(i,j,k)}.E_{g(i,j,k)}$$

$$E'_{f'(i',j')} \equiv \bigoplus_{k'\in 1..o'(i',j')} p'_{f'(i',j',k')} u'_{f'(i',j',k')}.E'_{g'(i',j',k')}.$$

Let $I = \{\langle i,i'\rangle \mid E_i \sim E'_{i'}\}$. By hypothesis we have $E_1 \sim E'_1$, so $\langle 1,1\rangle \in I$. Moreover, for each $\langle i,i'\rangle \in I$, the following holds, by the definition of strong bisimilarity:

(1) There exists a total surjective relation $K_{ii'}$ between $\{1,...,n(i)\}$ and $\{1,...,n'(i')\}$, given by

$$K_{ii'} = \{\langle j,j'\rangle \mid \langle f(i,j), f'(i',j')\rangle \in I\}.$$

Furthermore, for each $\langle j,j'\rangle \in K_{ii'}$ there exists a total surjective relation $L_{iji'j'}$ between $\{1,...,o(i,j)\}$ and $\{1,...,o'(i',j')\}$, given by

$$L_{iji'j'} = \{\langle k,k'\rangle \mid u_{f(i,j,k)} = u'_{f'(i',j',k')} \text{ and } \langle g(i,j,k), g'(i',j',k')\rangle \in I\}.$$

(2) $\vdash \sum_{j\in 1..l(i)} W_{h(i,j)} = \sum_{j'\in 1..l'(i')} W_{h'(i',j')}$.

Now, let $L_{iji'j'}(k)$ denote the image of $k \in \{1, ..., o(i,j)\}$ under $L_{iji'j'}$ and $L^{-1}_{iji'j'}(k')$ the preimage of $k' \in \{1, ..., o'(i',j')\}$ under $L_{iji'j'}$. We write $[k]_{iji'j'}$ for the set $L^{-1}_{iji'j'}(L_{iji'j'}(k))$ and $[k']_{iji'j'}$ for $L_{iji'j'}(L^{-1}_{iji'j'}(k'))$. It follows from the definitions that

(1) If $\langle i, i'_1\rangle \in I$, $\langle i, i'_2\rangle \in I$, $\langle j, j'_1\rangle \in K_{ii'_1}$ and $\langle j, j'_2\rangle \in K_{ii'_2}$, then $[k]_{iji'_1j'_1} = [k]_{iji'_2j'_2}$.
(2) If $q_1 \in [k]_{iji'j'}$ and $q_2 \in [k]_{iji'j'}$, then $u_{f(i,j,q_1)} = u_{f(i,j,q_2)}$ and $E_{g(i,j,q_1)} \sim E_{g(i,j,q_2)}$.

Define $\nu_{ijk} = \sum_{q\in[k]_{iji'j'}} p_{f(i,j,q)}$ for any i', j' such that $\langle i, i'\rangle \in I$ and $\langle j, j'\rangle \in K_{ii'}$; define $\nu'_{i'j'k'} = \sum_{q'\in[k']_{iji'j'}} p'_{f'(i',j',q')}$ for any i, j such that $\langle i, i'\rangle \in I$ and $\langle j, j'\rangle \in K_{ii'}$. It is easy to see that whenever $\langle i, i'\rangle \in I$, $\langle j, j'\rangle \in K_{ii'}$ and $\langle k, k'\rangle \in L_{iji'j'}$ then $\nu_{ijk} = \nu'_{i'j'k'}$.

We now consider the formal equations, one for each $\langle i, i'\rangle \in I$:

$$X_{ii'} = \sum_{\langle j,j'\rangle\in K_{ii'}} H_{f(i,j),f'(i',j')} + \sum_{j\in 1..l(i)} W_{h(i,j)}$$

where

$$H_{f(i,j),f'(i',j')} \equiv \bigoplus_{\langle k,k'\rangle\in L_{iji'j'}} \left(\frac{p_{f(i,j,k)}p'_{f'(i',j',k')}}{\nu_{ijk}}\right)u_{f(i,j,k)}.X_{g(i,j,k),g'(i',j',k')}.$$

These equations are provably satisfied when each $X_{ii'}$ is instantiated to E_i, since $K_{ii'}$ and $L_{iji'j'}$ are total and the right-hand side differs at most by repeated summands from that of the already proved equation for E_i. Note that each probabilistic branch $p_{f(i,j,k)}u_{f(i,j,k)}.E_{g(i,j,k)}$ in E_i becomes the probabilistic summation of several branches like

$$\bigoplus_{q'\in[k']_{iji'j'}} \left(\frac{p_{f(i,j,k)}p'_{f'(i',j',q')}}{\nu_{ijk}}\right)u_{f(i,j,k)}.E_{g(i,j,k)}$$

in $H_{f(i,j),f'(i',j')}\{E_i/X_{ii'}\}_i$, where $\langle i, i'\rangle \in I$, $\langle j, j'\rangle \in K_{ii'}$ and $\langle k, k'\rangle \in L_{iji'j'}$. But they are provably equal because

$$\begin{aligned}\textstyle\sum_{q'\in[k']_{iji'j'}} \left(\frac{p_{f(i,j,k)}p'_{f'(i',j',q')}}{\nu_{ijk}}\right) &= \tfrac{p_{f(i,j,k)}}{\nu_{ijk}} \cdot \textstyle\sum_{q'\in[k']_{iji'j'}} p'_{f'(i',j',q')} \\ &= \tfrac{p_{f(i,j,k)}}{\nu_{ijk}} \cdot \nu_{i'j'k'} = p_{f(i,j,k)}\end{aligned}$$

and then the axiom **S5** can be used. Symmetrically, the equations are provably satisfied when each $X_{ii'}$ is instantiated to $E'_{i'}$; this depends on the surjectivity of $K_{ii'}$ and $J_{iji'j'}$.

Finally, we note that each $X_{ii'}$ is weakly guarded in the right-hand sides of the formal equations. It follows from Theorem 18 that $\vdash E_i = E'_{i'}$ for each $\langle i, i' \rangle \in I$, and hence $\vdash E = E'$. $\Box$

5.2 Axiomatizing strong probabilistic bisimulation

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 $\sum_{i \in 1..n} r_i = 1$. It is easy to show that the expressions on the left and right sides are strong probabilistic bisimilar. We denote $\mathcal{A}_r \cup \{\mathbf{C}\}$ by $\mathcal{A}_{rc}$.

Theorem 21 (Soundness and completeness of $\mathcal{A}_{rc}$) *$E \sim_c E'$ iff $\mathcal{A}_{rc} \vdash E = E'$.*

Proof. The soundness part follows immediately by the definition of $\rightarrow_c$. Below we focus on the completeness part.

Let E and E' have free variables in $\widetilde{W}$. By Theorem 19 there are provable equations such that $E \equiv E_1$, $E' \equiv E'_1$ and

$$\mathcal{A}_{rc} \vdash E_i = A_i \qquad (i \leq m)$$

$$\mathcal{A}_{rc} \vdash E'_{i'} = A'_{i'} \qquad (i' \leq m')$$

where $A_i \equiv \sum_{j \in 1..n(i)} E_{f(i,j)} + \sum_{j \in 1..l(i)} W_{h(i,j)}$ and

$$E_{f(i,j)} \equiv \bigoplus_{k \in 1..o(i,j)} p_{f(i,j,k)} u_{f(i,j,k)}.E_{g(i,j,k)}$$

Similarly for the form of $A'_{i'}$.

Next we shall use axiom **C** to saturate the right hand side of each equation with some summands so as to transform each A_i (resp. $A'_{i'}$) into a provably equal expression B_i (resp. $B'_{i'}$) which satisfies the following property:

(*) *For any $C_1, C_2 \in \widetilde{B} \cup \widetilde{B'}$ with $C_1 \sim_c C_2$, if $C_1 \rightarrow \eta_1$ then there exists some η_2 s.t. $C_2 \rightarrow \eta_2$ and $\eta_1 \equiv_{\sim_c} \eta_2$.*

Initially we set $\widetilde{B} = \widetilde{A}$ and $\widetilde{B'} = \widetilde{A'}$. Let $S = \{(C_1, C_2) \mid C_1 \sim_c C_2$ and $C_1, C_2 \in \widetilde{A} \cup \widetilde{A'}\}$. Clearly the set S is finite because there are finitely many expressions in $\widetilde{A} \cup \widetilde{A'}$. Without loss of generality, we take a pair (C_1, C_2) from S such that $C_1 \equiv A'_{i'} \in \widetilde{A'}$ and $C_2 \equiv A_i \in \widetilde{A}$ (we do similar manipulations for the other three cases, namely (i) $C_1, C_2 \in \widetilde{A}$; (ii) $C_1, C_2 \in \widetilde{A'}$; (iii) $C_1 \in \widetilde{A}$ and $C_2 \in \widetilde{A'}$). If $A'_{i'} \rightarrow \eta'$ then for some η we have $A_i \rightarrow_c \eta$ and $\eta \equiv_{\sim_c} \eta'$, by the definition of $\sim_c$. If $A_i \rightarrow \eta$ (obviously we are in this case if $\eta = \vartheta(X)$) we do nothing but go on to pick another pair from S to do the analysis. Otherwise η is a convex combination $\eta = r_1\eta_1 + ... + r_n\eta_n$ and $A_i \rightarrow \eta_j$ for each $j \leq n$. Hence, each η_j must be in the form $\{(u_{f(i,j,k)}, E_{g(i,j,k)} : p_{f(i,j,k)})\}_k$ and $E_{f(i,j)}$ is a summand of A_i (so it is also a summand of B_i). By axiom **C** we have

$$\mathcal{A}_{rc} \vdash B_i = B_i + \bigoplus_{j \in 1..n} \bigoplus_k r_j p_{f(i,j,k)} u_{f(i,j,k)}.E_{g(i,j,k)}.$$

Now we update B_i to be to the expression on the right hand side of last equation. To this point we have finished the analysis to the pair (C_1, C_2). We need to pick a different pair from S to iterate the above procedure. When all the pairs in S are exhausted, we end up with $\widetilde{B}$ and $\widetilde{B'}$ for which it is easy to verify that they satisfy property (*). Observe that only axiom **C** is involved when updating B_i, so we have the following results:

$$\mathcal{A}_{rc} \vdash E_i = B_i \qquad (i \leq m)$$

$$\mathcal{A}_{rc} \vdash E'_{i'} = B'_{i'} \qquad (i' \leq m')$$

From now on, by using the above equations as our starting point, the subsequent arguments are like those for Theorem 20, so we omit them. □

6 Axiomatizations for guarded expressions

Now we proceed with the axiomatizations of the two weak behavioral equivalences: $\doteq$ 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, see Section 8), 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).

6.1 Axiomatizing divergence-sensitive equivalence

We first study the axiom system for $\doteq$. As a starting point, let us consider the system $\mathcal{A}_{rc}$. Clearly, **S1-5** are still valid for $\doteq$, 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 $\doteq$ we use the probabilistic τ-laws **T1-3** shown in Table 4. Note that **T3** is the probabilistic extension of Milner's

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	$(\bigoplus_i p_i u_i.E_i) \oplus pu.(F + \bigoplus_j q_j\tau.F_j) + (\bigoplus_i p_i u_i.E_i) \oplus (\bigoplus_j pq_j u.F_j)$
	$= (\bigoplus_i p_i u_i.E_i) \oplus pu.(F + \bigoplus_j q_j\tau.F_j)$

third τ-law ([15] 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 $\simeq$. Below we let $\mathcal{A}_{gd}$ =$\{$**R2′, T1-3**$\}$ $\cup\mathcal{A}_{rc}\backslash\{$**R2-3**$\}$.

Theorem 22 (Soundness of $\mathcal{A}_{gd}$) *If $\mathcal{A}_{gd} \vdash E = E'$ then $E \simeq E'$.*

Proof. The rule **R2′** is shown to be sound in Proposition 15. 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, and which can be proved by transition induction, using the probabilistic τ-laws and the axiom **C**.

Lemma 23 (Saturation) *(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 $\mathcal{A}_{gd} \vdash E = E + X$.

To show the completeness of $\mathcal{A}_{gd}$, we need some notations. Let V be a set, we write $\mathcal{P}(V)$ for the set of all probability distributions over V. Given a standard equation set $S : \widetilde{X} = \widetilde{H}$, which has free variables $\widetilde{W}$, we define the relations $\rightarrow_S \subseteq \widetilde{X} \times \mathcal{P}((Var \cup Act_\tau) \times \widetilde{X})$ by $X_i \rightarrow_S \eta$ iff $H_i \rightarrow \eta$. From $\rightarrow_S$ we can define the weak transition $\Rightarrow_S$ in the same way as in Section 3. We write $X_i \rightsquigarrow_S X_k$ iff $X_i \Rightarrow_S \eta$, with $\eta = \{(u_j, X_j : p_j)\}_{j \in J}$, $k \in J$ and $u_k = \tau$. We shall call S *guarded* if there is no X_i s.t. $X_i \rightsquigarrow_S X_i$. We call S *saturated* if for all $X \in \widetilde{X}$, $X \Rightarrow_S \eta$ implies $X \rightarrow_S \eta$. The variable W is *guarded* in S if it is not the case that $X_1 \rightarrow_S \vartheta(W)$ or $X_1 \rightsquigarrow_S \rightarrow_S \vartheta(W)$.

For guarded expressions, the equational characterization theorem and the unique solution theorem given in last section can now be refined, as done in [15].

Theorem 24 (Equational characterization II) *Each guarded expression*

E with free variables in $\widetilde{W}$ provably satisfies a standard guarded equation set S with free variables in $\widetilde{W}$. Moreover, if W is guarded in E then W is guarded in S.

Proof. By induction on the structure of E. Consider the case that $E \equiv \bigoplus_{i\in I} p_i u_i.E_i$. For each $i \in I$, let X_i be the distinguished variable of the equation set S_i for E_i. We can define S as $\{X = \bigoplus_{i\in I} p_i u_i.X_i\} \cup \bigcup_{i\in I} S_i$, with the new variable X distinguished. All other cases are the same as in [15]. $\Box$

Lemma 25 *Assume E provably satisfies the standard guarded equation set S. Then there is a saturated, standard, and guarded equation set S' provably satisfied by E.*

Proof. By using Lemma 23, we show that if $X_i \Rightarrow \eta$ then $\mathcal{A}_{gd} \vdash E_i = E_i + \bigoplus_j p_j u_j.E_j$ when $\eta \equiv \{(u_j, X_j : p_j)\}_j$, and $\mathcal{A}_{gd} \vdash E_i = E_i + X$ when $\eta \equiv \vartheta(X)$. Note that the equation set S is guarded, so there are only finite number of different distributions η such that $X_i \Rightarrow \eta$. By repeating this step for all weak transitions of E_i, at last we get $\mathcal{A}_{gd} \vdash E_i = H'_i\{\widetilde{E}/\widetilde{X}\}$. Hence, we can take S' to be the equation set $\widetilde{X} = \widetilde{H'}$. $\Box$

Theorem 26 (Unique solution of equations II) *If S is a guarded equation set with free variables in $\widetilde{W}$, then there is an expression E which provably satisfies S. Moreover, if F provably satisfies S and has free variables in $\widetilde{W}$, then $\mathcal{A}_{gd} \vdash E = F$.*

Proof. Nearly the same as the proof of Theorem 18, just replacing the recursion rule **R2** with **R2′**. $\Box$

The completeness result can be proved in a similar way as Theorem 20. 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 23.

Theorem 27 (Completeness of $\mathcal{A}_{gd}$) *If E and E' are guarded expressions and $E \simeq E'$ then $\mathcal{A}_{gd} \vdash E = E'$.*

Proof. By Theorem 24 there are provable equations such that $E \equiv E_1$, $E' \equiv E'_1$ and

$$\mathcal{A}_{rc} \vdash E_i = A_i \qquad (i \leq m)$$

$$\mathcal{A}_{rc} \vdash E'_{i'} = A'_{i'} \qquad (i' \leq m')$$

For any $C \in \widetilde{A} \cup \widetilde{A'}$, we assume by Lemma 25 that C is saturated. Therefore, it is easy to show that $C \Rightarrow_c \eta$ implies $C \rightarrow_c \eta$. Let $C' \in \widetilde{A} \cup \widetilde{A'}$. We note the interesting property that if $C \simeq C'$ and $C \rightarrow \eta$ then there exists η' s.t. $C' \rightarrow_c \eta'$ and $\eta \equiv_{\simeq} \eta'$. Thanks to this property the remaining arguments are quite similar to that in Theorem 21, thus are omitted. $\Box$

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 $\tau.E = \tau.F$.

6.2 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 20, with the additional machinery required for dealing with observational equivalence, like in [15]. 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 [15] because of the subtlety introduced by the probabilistic dimension (cf: Theorem 31). 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 7.

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 [15]. In the same table **T5** is the additional rule mentioned above. We let $\mathcal{A}_{go} = \mathcal{A}_{gd} \cup \{$**T4-5**$\}$.

Theorem 28 (Soundness of $\mathcal{A}_{go}$) *If $\mathcal{A}_{go} \vdash E = F$ then $E \simeq F$.*

Proof. Rule **T5** is proved to be sound in Proposition 15. The soundness of **T4**, and therefore of $\mathcal{A}_{go}$, is straightforward. □

The rest of the section is devoted to the completeness proof of $\mathcal{A}_{go}$. First we need two basic properties of weak combined transitions.

Lemma 29 *(1) If $E \stackrel{\epsilon}{\Rightarrow}_c \eta$ then $\tau.E \Rightarrow_c \eta$;*
(2) If $E \stackrel{\epsilon}{\Rightarrow}_c \vartheta(X)$ then $E \Rightarrow \vartheta(X)$.

Proof. The first clause is easy to show. Let us consider the second one. If $\vartheta(X)$ is a convex combination of $\eta_1, .., \eta_n$ and $E \Rightarrow \eta_i$ for all $i \in 1..n$, then each η_i must assign probability 1 to $(X, \mathbf{0})$, thus $\eta_i = \vartheta(X)$. □

Lemma 30 *If $E \stackrel{\epsilon}{\Rightarrow}_c \eta$ with $\eta = \{(u_i, E_i : p_i)\}_i$ then $\mathcal{A}_{gd} \vdash \tau.E = \tau.E + \bigoplus_i p_i u_i.E_i$.*

Proof. It follows from Lemma 29 and Lemma 23. □

The following theorem plays a crucial role in proving the completeness of $\mathcal{A}_{go}$.

Theorem 31 *Let E provably satisfy S and F provably satisfy T, where both S and T are standard, guarded equation sets, and let $E \simeq F$. Then there is a standard, guarded equation set U satisfied by both E and F.*

Proof. Suppose that $\widetilde{X} = \{X_1, ..., X_m\}$, $\widetilde{Y} = \{Y_1, ..., Y_n\}$ and $\widetilde{W} = \{W_1, W_2, ...\}$ are disjoint sets of variables. Let

$$S : \widetilde{X} = \widetilde{H}$$

$$T : \widetilde{Y} = \widetilde{J}$$

with $fv(\widetilde{H}) \subseteq \widetilde{X} \cup \widetilde{W}$, $fv(\widetilde{J}) \subseteq \widetilde{Y} \cup \widetilde{W}$, and that there are expressions $\widetilde{E} = \{E_1, ..., E_m\}$ and $\widetilde{F} = \{F_1, ..., F_n\}$ with $E_1 \equiv E$, $F_1 \equiv F$, and $fv(\widetilde{E}) \cup fv(\widetilde{F}) \subseteq \widetilde{W}$, so that

$$\mathcal{A}_{go} \vdash \widetilde{E} = \widetilde{H}\{\widetilde{E}/\widetilde{X}\}$$
$$\mathcal{A}_{go} \vdash \widetilde{F} = \widetilde{J}\{\widetilde{F}/\widetilde{Y}\}.$$

Consider the least equivalence relation $\mathcal{R} \subseteq (\widetilde{X} \cup \widetilde{Y}) \times (\widetilde{X} \cup \widetilde{Y})$ such that

(1) whenever $(Z, Z') \in \mathcal{R}$ and $Z \rightarrow \eta$, then there exists η' s.t. $Z' \stackrel{\hat{\epsilon}}{\Rightarrow}_c \eta'$ and $\eta \equiv_{\mathcal{R}} \eta'$;
(2) $(X_1, Y_1) \in \mathcal{R}$ and if $X_1 \rightarrow \eta$ then there exists η' s.t. $Y_1 \Rightarrow_c \eta'$ and $\eta \equiv_{\mathcal{R}} \eta'$.

Clearly, $\mathcal{R}$ is a weak probabilistic bisimulation on the transition system over $\widetilde{X} \cup \widetilde{Y}$, determined by $\rightarrow \stackrel{\text{def}}{=} \rightarrow_S \cup \rightarrow_T$. Now for two given distributions $\eta = \{(u_i, X_i : p_i)\}_{i \in I}$, $\eta' = \{(v_j, Y_j : q_j)\}_{j \in J}$, with $\eta \equiv_{\mathcal{R}} \eta'$, we introduce the following notations:

$$K_{\eta,\eta'} = \{(i,j) \mid i \in I,\ j \in J,\ u_i = v_j \text{ and } (X_i, Y_j) \in \mathcal{R}\}$$
$$\nu_i = \textstyle\sum\{p_{i'} \mid i' \in I,\ u_{i'} = u_i,\ \text{and } (X_i, X_{i'}) \in \mathcal{R}\} \qquad \text{for } i \in I$$
$$\nu_j = \textstyle\sum\{p_{j'} \mid j' \in J,\ v_{j'} = v_j,\ \text{and } (Y_j, Y_{j'}) \in \mathcal{R}\} \qquad \text{for } j \in J$$

Since $\eta \equiv_{\mathcal{R}} \eta'$ it follows by definition that if $(i,j) \in K_{\eta,\eta'}$, for some η, η', then $\nu_i = \nu_j$. Thus, we can define the expression

$$G_{\eta,\eta'} \stackrel{\text{def}}{=} \bigoplus_{(i,j) \in K_{\eta,\eta'}} \frac{p_i q_j}{\nu_i} u_i.Z_{ij}$$

which will play the same role as the expression $H_{f(i,j),f'(i',j')}$ in the proof of Theorem 20. On the other hand, if $\eta = \eta' = \vartheta(X)$ we simply define the expression $G_{\eta,\eta'} \stackrel{\text{def}}{=} X$.

Based on the above $\mathcal{R}$ we choose a new set of variables $\widetilde{Z}$ such that

$$\widetilde{Z} = \{Z_{ij} \mid X_i \in \widetilde{X},\ Y_j \in \widetilde{Y} \text{ and } (X_i, Y_j) \in \mathcal{R}\}.$$

Furthermore, for each $Z_{ij} \in \widetilde{Z}$ we construct three auxiliary finite sets of expressions, denoted by A_{ij}, B_{ij} and C_{ij}, by the following procedure.

(1) Initially the three sets are empty.
(2) For each η with $X_i \rightarrow \eta$, arbitrarily choose one (and only one — the same principle applies in other cases too) η' (if it exists) satisfying $\eta \equiv_{\mathcal{R}} \eta'$ and $Y_j \Rightarrow_c \eta'$, construct the expression $G_{\eta,\eta'}$ and update A_{ij} to be $A_{ij} \cup \{G_{\eta,\eta'}\}$; Similarly for each η' with $Y_j \rightarrow \eta'$, arbitrarily choose one η (if it exists) satisfying $\eta \equiv_{\mathcal{R}} \eta'$ and $X_i \Rightarrow_c \eta$, construct $G_{\eta,\eta'}$ and update A_{ij} to be $A_{ij} \cup \{G_{\eta,\eta'}\}$.
(3) For each η with $X_i \rightarrow \eta$, arbitrarily choose one η' (if it exists) satisfying $\eta \equiv_{\mathcal{R}} \eta'$, $Y_j \stackrel{\epsilon}{\Rightarrow}_c \eta'$ but not $Y_j \Rightarrow_c \eta'$, construct the expression $G_{\eta,\eta'}$ and update B_{ij} to be $B_{ij} \cup \{G_{\eta,\eta'}\}$.
(4) For each η' with $Y_j \rightarrow \eta'$, arbitrarily choose one η (if it exists) satisfying $\eta \equiv_{\mathcal{R}} \eta'$, $X_i \stackrel{\epsilon}{\Rightarrow}_c \eta$ but not $X_i \Rightarrow_c \eta$, construct $G_{\eta,\eta'}$ and update C_{ij} to be $C_{ij} \cup \{G_{\eta,\eta'}\}$.

Clearly, the three sets constructed in this way are finite. Now we build a new equation set

$$U : \widetilde{Z} = \widetilde{L}$$

where U_{11} is the distinguished variable and

$$L_{ij} = \begin{cases} \sum_{G \in A_{ij}} G & \text{if } B_{ij} \cup C_{ij} = \emptyset \\ \tau.(\sum_{G \in A_{ij} \cup B_{ij} \cup C_{ij}} G) & \text{otherwise.} \end{cases}$$

We assert that E provably satisfies the equation set U. To see this, we choose expressions

$$G_{ij} = \begin{cases} E_i & \text{if } B_{ij} \cup C_{ij} = \emptyset \\ \tau.E_i & \text{otherwise} \end{cases}$$

and verify that $\mathcal{A}_{go} \vdash G_{ij} = L_{ij}\{\widetilde{G}/\widetilde{Z}\}$.

In the case that $B_{ij} \cup C_{ij} = \emptyset$, all those summands of $L_{ij}\{\widetilde{G}/\widetilde{Z}\}$ which are not variables are of the forms:

$$\bigoplus_{(i,j) \in K_{\eta,\eta'}} \frac{p_i q_j}{\nu_i} u_i.E_i \quad \text{or} \quad \bigoplus_{(i,j) \in K_{\eta,\eta'}} \frac{p_i q_j}{\nu_i} u_i.\tau.E_i.$$

By **T4** we can transform the second form into the first one. Then by some arguments similar to those in Theorem 20, together with Lemma 23, we can show that

$$\mathcal{A}_{go} \vdash L_{ij}\{\widetilde{G}/\widetilde{Z}\} = H_i\{\widetilde{E}/\widetilde{X}\} = E_i.$$

On the other hand, if $B_{ij} \cup C_{ij} \neq \emptyset$, we let $C_{ij} = \{D_1, ..., D_o\}$ ($C_{ij} = \emptyset$ is a special case of the following argument) and $D = \sum_{l\in 1..o} D_l\{\widetilde{G}/\widetilde{Z}\}$. As in last case we can show that

$$\mathcal{A}_{go} \vdash L_{ij}\{\widetilde{G}/\widetilde{Z}\} = \tau.(H_i\{\widetilde{E}/\widetilde{X}\} + D).$$

For any l with $1 \leq l \leq o$, let $D_l\{\widetilde{G}/\widetilde{Z}\} = \bigoplus_k p_k u_k.E_k$. It is easy to see that $E_i \overset{\epsilon}{\Rightarrow}_c \eta$ with $\eta = \{(u_k, E_k : p_k)\}_k$. So by Lemma 30 it holds that

$$\mathcal{A}_{go} \vdash \tau.E_i = \tau.E_i + D_l\{\widetilde{G}/\widetilde{Z}\}.$$

As a result we can infer

$$\mathcal{A}_{go} \vdash \tau.E_i = \tau.E_i + D = \tau.E_i + (E_i + D).$$

by Lemma 23. Similarly,

$$\mathcal{A}_{go} \vdash \tau.(E_i + D) = \tau.(E_i + D) + E_i.$$

Consequently it follows from **T5** that

$$\mathcal{A}_{go} \vdash \tau.E_i = \tau.(E_i + D) = \tau.(H_i\{\widetilde{E}/\widetilde{X}\} + D) = L_{ij}\{\widetilde{G}/\widetilde{Z}\}.$$

In the same way we can show that F provably satisfies U. At last U is guarded because S and T are guarded. □

To help understanding the proof of the above theorem, we illustrate the construction of the equation set U by a simple example. Consider the equation sets S and T as follows.

$$S: \quad X_1 = a.X_2 \qquad\qquad T: \quad Y_1 = \tfrac{1}{2}a.Y_2 \oplus \tfrac{1}{2}a.Y_3$$

$$X_2 = a.X_2 + \tfrac{1}{2}a.X_2 \oplus \tfrac{1}{2}\tau.X_1 \qquad\qquad Y_2 = a.Y_3 + \tau.Y_3$$

$$Y_3 = a.Y_2$$

Note that if E_1, E_2 provably satisfy S, and F_1, F_2, F_3 provably satisfy T, then $E_1 \simeq F_1 \simeq \mu_Z(a.Z)$.

Let $\mathcal{R}$ be the equivalence relation that has the only equivalence class $\{X_1, X_2, Y_1, Y_2, Y_3\}$. It is easy to check that $\mathcal{R}$ is a weak bisimulation on the transition system over $\widetilde{X} \cup \widetilde{Y}$. Now we take new variables $\{Z_{ij} \mid 1 \leq i \leq 2, 1 \leq j \leq 3\}$ and form the sets A_{ij}, B_{ij} and C_{ij} for each variable Z_{ij}, as displayed in Table 6, by using the procedure presented in the above proof. For example, consider the line for $(i, j) = (2, 1)$.

(1) Initially the sets A_{21}, B_{21} and C_{21} are empty.

Table 6
The construction of sets A_{ij}, B_{ij}, C_{ij}

(i,j)	A_{ij}	B_{ij}	C_{ij}
$(1,1)$	$\{\frac{1}{2}a.Z_{22} \oplus \frac{1}{2}a.Z_{23}\}$	$\emptyset$	$\emptyset$
$(1,2)$	$\{a.Z_{23}\}$	$\emptyset$	$\{\tau.Z_{13}\}$
$(1,3)$	$\{a.Z_{22}\}$	$\emptyset$	$\emptyset$
$(2,1)$	$\{\frac{1}{2}a.Z_{22} \oplus \frac{1}{2}a.Z_{23}\}$	$\{\frac{1}{4}a.Z_{22} \oplus \frac{1}{4}a.Z_{23} \oplus \frac{1}{2}\tau.Z_{11}\}$	$\emptyset$
$(2,2)$	$\{a.Z_{23},\ \frac{1}{2}a.Z_{23} \oplus \frac{1}{2}\tau.Z_{13}\}$	$\emptyset$	$\{\tau.Z_{23}\}$
$(2,3)$	$\{a.Z_{22}\}$	$\{\frac{1}{2}a.Z_{22} \oplus \frac{1}{2}\tau.Z_{13}\}$	$\emptyset$

(2) Let's see how to form the set A_{21}. From X_2 there are two outgoing transitions: $X_2 \rightarrow \eta_1$, with $\eta_1 = \{a, X_2 : 1\}$, and $X_2 \rightarrow \eta_2$, with $\eta_2 = \{(a, X_2 : \frac{1}{2}), (\tau, X_1 : \frac{1}{2})\}$. The first one is matched by the transition $Y_1 \rightarrow \eta_1'$, with $\eta_1' = \{(a, Y_2 : \frac{1}{2}), (a, Y_3 : \frac{1}{2})\}$, because $\eta_1 \equiv_{\mathcal{R}} \eta_1'$. The second one will not contribute anything to the set A_{21} because there is no η such that $Y_1 \Rightarrow_c \eta$ and $\eta_2 \equiv_{\mathcal{R}} \eta$. For the other direction, Y_1 has one outgoing transition $Y_1 \rightarrow \eta_1'$ which is matched by $X_2 \rightarrow \eta_1$. So we construct the expression $G_{\eta_1,\eta_1'} = \frac{1}{2}a.Z_{22} \oplus \frac{1}{2}a.Z_{23}$ and add it to the set A_{21}, which is updated to be $\{G_{\eta_1,\eta_1'}\}$.

(3) Let's see how to form the set B_{21}. We have just used one of the two transitions of X_2 to form A_{21}. The unused one is the only candidate to contribute to B_{21}. Indeed, there is an η_2' such that the transition $X_2 \rightarrow \eta_2$ is matched by $Y_1 \stackrel{\epsilon}{\Rightarrow}_c \eta_2'$ with $\eta_2 \equiv_{\mathcal{R}} \eta_2'$ but $Y_1 \not\Rightarrow_c \eta_2'$. To see this, we simply take $\eta_2' = \frac{1}{2}\eta_1' + \frac{1}{2}\delta(Y_1) = \{(a, Y_2 : \frac{1}{4}), (a, Y_3 : \frac{1}{4}), (\tau, Y_1 : \frac{1}{2})\}$. So we construct the expression $G_{\eta_2,\eta_2'} = \frac{1}{4}a.Z_{22} \oplus \frac{1}{4}a.Z_{23} \oplus \frac{1}{2}\tau.Z_{11}$ and add it to the set B_{21}, which now becomes $\{G_{\eta_2,\eta_2'}\}$.

(4) Let's see how to form the set C_{21}. From Y_1 there is only one outgoing transition $Y_1 \rightarrow \eta_1'$, but it has been used in forming A_{21}. Indeed, there is no η such that $\eta \equiv_{\mathcal{R}} \eta_1'$, $X_2 \stackrel{\epsilon}{\Rightarrow}_c \eta$ but $X_2 \not\Rightarrow_c \eta$. Therefore, we have nothing to add to the set C_{21}, which remains to be $\emptyset$.

We construct the equation set U, based on all expressions shown in Table 6.

$$
\begin{aligned}
U: \quad Z_{11} &= \tfrac{1}{2}a.Z_{22} \oplus \tfrac{1}{2}a.Z_{23} \\
Z_{12} &= \tau.(a.Z_{23} + \tau.Z_{13}) \\
Z_{13} &= a.Z_{22} \\
Z_{21} &= \tau.(\tfrac{1}{2}a.Z_{22} \oplus \tfrac{1}{2}a.Z_{23} + \tfrac{1}{4}a.Z_{22} \oplus \tfrac{1}{4}a.Z_{23} \oplus \tfrac{1}{2}\tau.Z_{11}) \\
Z_{22} &= \tau.(a.Z_{23} + \tfrac{1}{2}a.Z_{23} \oplus \tfrac{1}{2}\tau.Z_{13} + \tau.Z_{23}) \\
Z_{23} &= \tau.(a.Z_{22} + \tfrac{1}{2}a.Z_{22} \oplus \tfrac{1}{2}\tau.Z_{13})
\end{aligned}
$$

We can see that E_1 provably satisfies U by substituting E_1, $\tau.E_1$, E_1, $\tau.E_2$, $\tau.E_2$, $\tau.E_2$ for $Z_{11}, Z_{12}, Z_{13}, Z_{21}, Z_{22}, Z_{23}$, respectively; similarly F_1 provably satisfies U by substituting F_1, $\tau.F_2$, F_3, $\tau.F_1$, $\tau.F_2$, $\tau.F_3$ for these variables.

Theorem 32 (Completeness of $\mathcal{A}_{go}$) *If E and F are guarded expressions and $E \simeq F$, then $\mathcal{A}_{go} \vdash E = F$.*

Proof. A direct consequence by combining Theorem 24, 31 and 26. □

7 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 recursion.

$$\mathcal{A}_s \stackrel{\text{def}}{=} \{\textbf{S1-5}\} \qquad \mathcal{A}_{sc} \stackrel{\text{def}}{=} \mathcal{A}_s \cup \{\textbf{C}\}$$

$$\mathcal{A}_{fd} \stackrel{\text{def}}{=} \mathcal{A}_{sc} \cup \{\textbf{T1-3}\} \qquad \mathcal{A}_{fo} \stackrel{\text{def}}{=} \mathcal{A}_{fd} \cup \{\textbf{T4-5}\}$$

Theorem 33 (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 $\mathcal{A}_{fo} \vdash E = F$.

The soundness part is obvious. The completeness can be shown by following the lines of previous sections. However, since there is no recursion here, we have a much simpler proof which does not use the equational characterization theorem and the unique solution theorem. Roughly speaking, all the clauses are proved by induction on the depth of the expressions. We define the depth of a process, $d(E)$, as follows.

$$d(\mathbf{0}) = 0$$

$$d(X) = 1$$

$$d(\textstyle\bigoplus_i p_i u_i.E_i) = 1 + max\{E_i\}_i$$

$$d(\textstyle\sum_i E_i) = max\{d(E_i)\}_i$$

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 [14], 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

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

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 34 (Promotion) *For any $E, F \in \mathcal{E}_f$, if $E \approx F$ then $\mathcal{A}_{fo} \vdash \tau.E = \tau.F$.*

Proof. By induction on $d = d(E) + d(F)$. We consider the nontrivial case that $d > 0$.

If X is a nondeterministic summand of E, then $E \rightarrow \vartheta(X)$. Since $E \doteq F$ it holds that $F \stackrel{\epsilon}{\Rightarrow}_c \vartheta(X)$. By Lemma 29 we have $F \Rightarrow \vartheta(X)$. It follows from (the recursion-free version of) Lemma 23 that $\mathcal{A}_{fd} \vdash F = F + X$.

Let $\bigoplus_{i\in I} p_i u_i.E_i$ be any summand of E. Then we have $E \rightarrow \eta$, with $\eta = \{(u_i, E_i : p_i)\}_{i\in I}$. Since $E \approx F$, there exists η', with $\eta' = \{(v_j, F_j : q_j)\}_{j\in J}$ s.t. $F \stackrel{\epsilon}{\Rightarrow}_c \eta'$ and $\eta \equiv_{\approx} \eta'$. For any $k, l \in I$ with $u_k = u_l$ and $E_k \approx E_l$, it follows from **T4** and induction hypothesis that $\mathcal{A}_{fo} \vdash u_k.E_k = u_k.\tau.E_k = u_l.\tau.E_l = u_l.E_l$. By **S5** we can derive that $\mathcal{A}_{fo} \vdash \bigoplus_{i\in I} p_i u_i.E_i = \bigoplus_{i'\in I'} p'_{i'} u'_{i'}.E'_{i'}$, where the process on the right hand side is "compact", i.e., for any $k', l' \in I'$, if $u'_{k'} = u'_{l'}$ and $E'_{k'} = E'_{l'}$ then $k' = l'$. Similarly we can derive $\mathcal{A}_{fo} \vdash \bigoplus_{j\in J} q_j v_j.F_j = \bigoplus_{j'\in J'} q'_{j'} v'_{j'}.F'_{j'}$ with the process on the right hand side "compact". From $\eta \equiv_{\approx} \eta'$ and the soundness of $\mathcal{A}_{fd}$, it is easy to prove that $\mathcal{A}_{fo} \vdash \bigoplus_{i'\in I'} p'_{i'} u'_{i'}.E'_{i'} = \bigoplus_{j'\in J'} q'_{j'} v'_{j'}.F'_{j'}$ since each probabilistic branch of one process is provably equal to a unique branch of the other process. It follows that $\mathcal{A}_{fo} \vdash \bigoplus_{i\in I} p_i u_i.E_i = \bigoplus_{j\in J} q_j v_j.F_j$. By (a recursion-free version of) Lemma 30 we infer $\mathcal{A}_{fo} \vdash \tau.F = \tau.F + \bigoplus_{j\in J} q_j v_j.F_j = \tau.F + \bigoplus_{i\in I} p_i u_i.E_i$.

In summary $\mathcal{A}_{fo} \vdash \tau.F = \tau.F + E$. Symmetrically $\mathcal{A}_{fo} \vdash \tau.E = \tau.E + F$. Therefore, $\mathcal{A}_{fo} \vdash \tau.E = \tau.F$ by **T5**. $\square$

The promotion lemma is inspired by [10], where a similar result is proved for a language of mobile processes.

At last, the completeness part of Theorem 33 (4) can be proved as Lemma 34. Note that for any $k, l \in I$ with $u_k = u_l$ and $E_k \approx E_l$, we derive $\mathcal{A}_{fo} \vdash u_k.E_k = u_l.E_l$ by using **T4** and the promotion lemma instead of using induction hypothesis.

It is worth noticing that rule **T5** is necessary to prove Lemma 34. 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 observationally 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 this term is not provably equal to $\tau.a$, which has no probabilistic choice.

8 Concluding remarks and future work

In this work we have proposed a probabilistic process calculus which corresponds to Segala and Lynch's probabilistic automata. We have presented strong bisimulation, strong probabilistic bisimulation, divergence-sensitive equivalence and observational equivalence. Sound and complete inference systems for the four behavioral equivalences are summarized in Table 8.

Note that we have axiomatized divergence-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 observationally 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

Table 7
All the axioms and rules

S1	$E + \mathbf{0} = E$
S2	$E + E = E$
S3	$\sum_{i \in I} E_i = \sum_{i \in I} E_{\rho(i)}$ 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	$(\bigoplus_i p_i u_i.E_i) \oplus pu.E \oplus qu.E = (\bigoplus_i p_i u_i.E_i) \oplus (p+q)u.E$
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	$(\bigoplus_i p_i u_i.E_i) \oplus p\tau.(F + \bigoplus_j q_j \beta_j.F_j) + (\bigoplus_i p_i u_i.E_i) \oplus (\bigoplus_j p q_j \beta_j.F_j)$ $= (\bigoplus_i p_i u_i.E_i) \oplus p\tau.(F + \bigoplus_j q_j \beta_j.F_j)$
T3	$(\bigoplus_i p_i u_i.E_i) \oplus pu.(F + \bigoplus_j q_j \tau.F_j) + (\bigoplus_i p_i u_i.E_i) \oplus (\bigoplus_j p q_j u.F_j)$ $= (\bigoplus_i p_i u_i.E_i) \oplus pu.(F + \bigoplus_j q_j \tau.F_j)$
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 8
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
$\doteq$	$\mathcal{A}_{fd}$: **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′**

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 undecidability. For instance, the notion of weak probabilistic bisimulation defined in [22,7] 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 [22,7] 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 [7].

In our work we chose, 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 ([23]), has the advantage of being more formal, and in the case of guarded recursion it is equivalent to the one of [22,7]. In the case of unguarded recursion, however, we feel that it would be more natural to consider also the "limit" weak transitions of [22,7]. The axiomatization of the corresponding notion of observational equivalence is an open problem.

In the future it might be interesting to see how to refine our process calculus to allow for parallel composition. To do that it seems necessary to add some syntactic constraints, because parallel composition is hard to define for PA, as discussed in [22]. Having both recursion and parallel composition in a process calculus complicates the matters to establish a complete axiomatization, mostly because this can give rise to infinite-state systems even with the guardedness condition. In [9] we focus on SPA and require that free variables do not appear in the scope of parallel composition in order to achieve complete axiomatizations in a calculus that includes parallel composition and guarded recursion. A nice idea of admitting parallelism in generative models is presented in [8]. We would like to adapt that idea in PA and consider its effect on axiomatizations. Another interesting research direction is to develop some automated verification tool by exploiting the axioms and inference rules in Table 7. One possible approach is to extend μCRL [11,27] to the probabilistic setting, and use some rewriting rules based on axioms similar to ours in Table 7. Our long term goal, as explained in the introduction, is to develop verification techniques for the asynchronous probabilistic π-calculus and to apply them to the verification of distributed algorithms.

Acknowledgements

We thank the anonymous referees for their helpful comments.

Appendix

A Proof of Proposition 15(3)

In [25] Stark and Smolka use a special function f that associates a probability to a nonprobabilistic transition so as to form a probabilistic transition. For example, let $E \equiv \frac{1}{3}a \oplus \frac{2}{3}b$, then $f(E \xrightarrow{a} \mathbf{0}) = \frac{1}{3}$ and $f(E \xrightarrow{b} \mathbf{0}) = \frac{2}{3}$. The function f can be characterized as $f = \sup_{i \geq 0} f_i$ for some functions $f_0, f_1, ...$ that take nonprobabilistic transitions to probabilities and respect some ordering. Therefore, in the soundness proofs of some axioms, to show that $f(E \xrightarrow{a} E') \leq p$, it suffices to prove by induction on i that $f_i(E \xrightarrow{a} E') \leq p$ for all $i \geq 0$. In the presence of nondeterministic choice, however, this technique becomes unusable because now the probability with which an expression performs an action and evolves into another expression is not deterministic any more. For example, let $E \equiv (\frac{1}{3}a \oplus \frac{2}{3}b) + (\frac{1}{2}a \oplus \frac{1}{2}c)$, then what is the value of $f(E \xrightarrow{a} \mathbf{0})$? Should it be $\frac{1}{3}$, $\frac{1}{2}$, or some value in between? Now the meaning of the function f is unclear because it depends on how the nondeterminism is resolved. Nevertheless, our "bisimulation up to" technique works well with Milner's transition induction technique, as can be seen in the proof of Proposition 15(3) below.

Lemma 35 *If $\eta_1 \equiv_{\mathcal{R}_1} \eta_2$ and $\mathcal{R}_1 \subseteq \mathcal{R}_2$ then $\eta_1 \equiv_{\mathcal{R}_2} \eta_2$.*

Proof. Let $V \in \mathcal{E}/\mathcal{R}_2$. Since $\mathcal{R}_1$ is contained in $\mathcal{R}_2$, we know that V is the disjoint union of all elements in some set $\{V_i\}_i$, with $V_i \in \mathcal{E}/\mathcal{R}_1$ for each i. It follows from $\eta_1 \equiv_{\mathcal{R}_1} \eta_2$ that

$$\forall \alpha \in \mathit{Var} \cup \mathit{Act}_\tau : \eta_1(\alpha, V_i) = \eta_2(\alpha, V_i).$$

Therefore, we have

$$\eta_1(\alpha, V) = \textstyle\sum_i \eta_1(\alpha, V_i) = \sum_i \eta_2(\alpha, V_i) = \eta_2(\alpha, V).$$ □

Lemma 36 *Let $\eta = r_1\eta_1 + ... + r_n\eta_n$ and $\eta' = r_1\eta'_1 + ... + r_n\eta'_n$ with $\sum_{i \in 1..n} r_i = 1$. If $\eta_i \equiv_{\mathcal{R}} \eta'_i$ for each $i \leq n$, then $\eta \equiv_{\mathcal{R}} \eta'$.*

Proof. For any $V \in \mathcal{E}/\mathcal{R}$ and $\alpha \in \mathit{Var} \cup \mathit{Act}_\tau$, we have

$$\eta(\alpha, V) = \sum_{i \in 1..n} r_i \eta_i(\alpha, V) = \sum_{i \in 1..n} r_i \eta'_i(\alpha, V) = \eta'(\alpha, V).$$

Therefore, $\eta \equiv_{\mathcal{R}} \eta'$ by definition. □

Lemma 37 *Suppose $E \simeq F$. If $E \Rightarrow_c \eta$ then there exists η' s.t. $F \Rightarrow_c \eta'$ and $\eta \equiv_{\approx} \eta'$.*

Proof. By transition induction. □

We use a measure $d_X(E)$ to count the depth of guardedness of the free variable X in expression E.

$$\begin{aligned} d_X(X) &= 0 \\ d_X(Y) &= 0 \\ d_X(a.E) &= d_X(E) + 1 \\ d_X(\tau.E) &= d_X(E) \\ d_X(\bigoplus_i p_i u_i.E_i) &= min\{d_X(u_i.E_i)\}_i \\ d_X(\sum_i E_i) &= min\{d_X(E_i)\}_i \\ d_X(\mu_Y E) &= d_X(E) \end{aligned}$$

If $d_X(E) > 0$ then X is guarded in E.

Lemma 38 *Let $d_X(G) = n$ and $\eta = \{(u_i, G_i : p_i)\}_{i\in I}$. Suppose $G\{E/X\} \Rightarrow \eta$. For all $i \in I$, it holds that*

(1) If $n > 0$ and $u_i = \tau$ then $G_i = G'_i\{E/X\}$ and $d_X(G'_i) \geq n$;
(2) If $n > 1$ and $u_i \neq \tau$ then $G_i = G'_i\{E/X\}$ and $d_X(G'_i) \geq n - 1$.

Proof. By induction on the depth of the inference of $G\{E/X\} \Rightarrow \eta$. There are three cases, depending on the last rule used in the inference. A typical case is for Rule wea3. In this case $\eta = \{(u_i, G_i : p_i)\}_{i\in I} \uplus \{(v_j, H_j : q_j)\}_{j\in J}$ and $G\{E/X\} \Rightarrow \eta$ is derived from the shorter inferences of $G\{E/X\} \Rightarrow \{(u_i, G_i : p_i)\}_{i\in I} \uplus \{(\tau, G_0 : p_0)\}$ and $G_0 \Rightarrow \{(v_j, H_j : q_j)\}_{j\in J}$. By induction hypothesis, for each $i \in I \cup \{0\}$, it holds that

(1) If $n > 0$ and $u_i = \tau$ then $G_i = G'_i\{E/X\}$ and $d_X(G'_i) \geq n$;
(2) If $n > 1$ and $u_i \neq \tau$ then $G_i = G'_i\{E/X\}$ and $d_X(G'_i) \geq n - 1$.

Particularly for G_0 we have $G_0 = G'_0\{E/X\}$ and $d_X(G'_0) \geq n > 0$. By induction hypothesis on the transition of $G'_0\{E/X\}$, it follows that for each $j \in J$

(1) if $v_j = \tau$ then $H_j = H'_j\{E/X\}$ and $d_X(H'_j) \geq d_X(G'_0) \geq n$ for each $j \in J$;
(2) $n > 1$ and $v_j \neq \tau$ then $H_j = H'_j\{E/X\}$ and $d_X(H'_j) \geq d_X(G'_0) - 1 \geq n - 1$.

□

Lemma 39 *Suppose* $d_X(G) > 1$, $\eta = \{(u_i, G_i : p_i)\}_{i \in I}$ *and* $G\{E/X\} \Rightarrow \eta$. *Then* $G_i = G'_i\{E/X\}$ *for each* $i \in I$. *Moreover,* $G\{F/X\} \Rightarrow \eta'$ *and* $\eta \equiv_{\mathcal{R}^*} \eta'$, *where* $\eta' = \{(u_i, G'_i\{F/X\} : p_i)\}_{i \in I}$ *and*

$$\mathcal{R} = \{(G\{E/X\}, G\{F/X\}) \mid \text{ for any } G \in \mathcal{E}\}.$$

Proof. A direct consequence of Lemma 38. □

Lemma 40 *Let* $d_X(G) > 1$. *If* $G\{E/X\} \Rightarrow_c \eta$ *then* $G\{F/X\} \Rightarrow_c \eta'$ *such that* $\eta \equiv_{\mathcal{R}^*} \eta'$ *where* $\mathcal{R} = \{(G\{E/X\}, G\{F/X\}) \mid$ *for any* $G \in \mathcal{E}\}$.

Proof. Let $\eta = r_1\eta_1 + ... + r_n\eta_n$ and $G\{E/X\} \Rightarrow \eta_i$ for each $i \leq n$. By Lemma 39, for each $i \leq n$, there exists η'_i s.t. $G\{F/X\} \Rightarrow \eta'_i$ and $\eta_i \equiv_{\mathcal{R}^*} \eta'_i$. Now let $\eta' = r_1\eta'_1 + ... + r_n\eta'_n$, thus $G\{F/X\} \Rightarrow_c \eta'$. By lemma 36 it follows that $\eta \equiv_{\mathcal{R}^*} \eta'$. □

Proof of Proposition 15(3). We show that the relation

$$\mathcal{R} = \{(G\{E/X\}, G\{\mu_X F/X\}) \mid \text{ for any } G \in \mathcal{E}\}$$

is an observational equivalence up to $\simeq$. That is, we need to show the following assertions:

(1) if $G\{E/X\} \Rightarrow \eta$ then there exists η' s.t. $G\{\mu_X F/X\} \Rightarrow_c \eta'$ and $\eta \equiv_{\mathcal{R}_\approx} \eta'$;
(2) if $G\{\mu_X F/X\} \Rightarrow \eta'$ then there exists η s.t. $G\{E/X\} \Rightarrow_c \eta$ and $\eta \equiv_{\mathcal{R}_\approx} \eta'$;

We concentrate on the first clause as the second one is similar. The proof is carried out by induction on the depth of the inference of $G\{E/X\} \Rightarrow \eta$. There are several cases depending on the structure of G. As an example, here we consider the case that $G \equiv X$.

We write $G(E)$ for $G\{E/X\}$ and $G^2(E)$ for $G(G(E))$. Since $E \simeq F(E)$, we have $E \simeq F^2(E)$ since $\simeq$ is an congruence relation by Proposition 15. If $E \Rightarrow \eta$ then by Lemma 37 there exists θ_1 s.t. $F^2(E) \Rightarrow_c \theta_1$ and $\eta \equiv_\approx \theta_1$. Since X is guarded in F, i.e., $d_X(F) > 0$, then it follows that $d_X(F^2(X)) > 1$. By Lemma 40, there exists θ_2 s.t. $F^2(\mu_X F) \Rightarrow_c \theta_2$ and $\theta_1 \equiv_{\mathcal{R}^*} \theta_2$. From Proposition 14 we have $\mu_X F \sim F^2(\mu_X F)$, thus $\mu_X F \simeq F^2(\mu_X F)$. By Lemma 37 there exists η' s.t. $\mu_X F \Rightarrow_c \eta'$ and $\theta_2 \equiv_\approx \eta'$. From Lemma 35 and the transitivity of $\equiv_{\mathcal{R}_\approx}$ it follows that $\eta \equiv_{\mathcal{R}_\approx} \eta'$. □

References

[1] L. Aceto, Z. Ésik, and A. Ingólfsdóttir. Equational axioms for probabilistic bisimilarity (preliminary report). Technical Report RS-02-6, BRICS, 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] C. Baier and H. Hermanns. Weak bisimulation for fully probabilistic processes. In *Proceedings of the 9th International Conference on Computer Aided Verification*, volume 1254 of *LNCS*, pages 119–130. Springer, 1997.

[6] 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.

[7] 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.

[8] P. R. D'Argenio, H. Hermanns, and J.-P. Katoen. On generative parallel composition. *ENTCS*, 22, 1999.

[9] Y. Deng, C. Palamidessi, and J. Pang. Compositional reasoning for probabilistic finite-state behaviors. In *Processes, Terms and Cycles: Steps on the Road to Infinity, Essays Dedicated to Jan Willem Klop, on the Occasion of His 60th Birthday*, volume 3838 of *LNCS*, pages 309–337. Springer, 2005.

[10] Y. Fu and Z. Yang. Tau laws for pi calculus. *Theoretical Computer Science*, 308:55–130, 2003.

[11] J. Groote and A. Ponse. The syntax and semantics of μCRL. Technical Report CS-R9076, CWI, Amsterdam, 1990.

[12] O. M. Herescu and C. Palamidessi. Probabilistic asynchronous pi-calculus. Technical report, INRIA Futurs and LIX, 2004.

[13] R. Milner. A complete inference system for a class of regular behaviours. *Journal of Computer and System Science*, 28:439–466, 1984.

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

[15] R. Milner. A complete axiomatisation for observational congruence of finite-state behaviours. *Information and Computation*, 81:227–247, 1989.

[16] R. Milner. *Communicating and Mobile Systems: the π-Calculus*. Cambridge University Press, 1999.

[17] M. W. Mislove, J. Ouaknine, and J. Worrell. Axioms for probability and nondeterminism. *ENTCS*, 96:7–28, 2004.

[18] C. Palamidessi. Comparing the expressive power of the synchronous and the asynchronous pi-calculus. *Mathematical Structures in Computer Science*, 13(5):685–719, 2003.

[19] C. Palamidessi and O. M. Herescu. A randomized encoding of the π-calculus with mixed choice. Technical report, INRIA Futurs and LIX, 2004.

[20] A. Philippou, I. Lee, and O. Sokolsky. Weak bisimulation for probabilistic systems. In *Proceedings of the 11th International Conference on Concurrency Theory*, LNCS, pages 334–349. Springer, 2000.

[21] D. Sangiorgi and D. Walker. *The π-calculus: a Theory of Mobile Processes*. Cambridge University Press, 2001.

[22] R. Segala. Modeling and verification of randomized distributed real-time systems. Technical Report MIT/LCS/TR-676, PhD thesis, MIT, Dept. of EECS, 1995.

[23] R. Segala and N. Lynch. Probabilistic simulations for probabilistic processes. In *Proceedings of the 5th International Conference on Concurrency Theory*, volume 836 of *LNCS*, pages 481–496. Springer, 1994.

[24] 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.

[25] 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.

[26] M. Stoelinga. *Alea jacta est: verification of probabilistic, real-time and parametric systems*. PhD thesis, University of Nijmegen, 2002.

[27] J. van de Pol. A prover for the muCRL toolset with applications – version 0.1. Technical Report SEN-R0106, CWI, Amsterdam, 2001.

[28] 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.