BRICS

Basic Research in Computer Science

BRICS RS-02-6 Aceto et al.: Equational Axioms for Probabilistic Bisimilarity

Equational Axioms for Probabilistic Bisimilarity

(Preliminary Report)

Luca Aceto
Zoltán Ésik
Anna Ingólfsdóttir

BRICS Report Series **RS-02-6**

ISSN 0909-0878 **February 2002**

Copyright © 2002, Luca Aceto & Zoltán Ésik & Anna Ingólfsdóttir.
BRICS, Department of Computer Science
University of Aarhus. All rights reserved.

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

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

BRICS
Department of Computer Science
University of Aarhus
Ny Munkegade, building 540
DK–8000 Aarhus C
Denmark

Telephone: +45 8942 3360
Telefax: +45 8942 3255
Internet: BRICS@brics.dk

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

`http://www.brics.dk`
`ftp://ftp.brics.dk`
This document in subdirectory `RS/02/6/`

Equational Axioms for Probabilistic Bisimilarity (Preliminary Report)

Luca Aceto, Zoltán Ésik*, and Anna Ingólfsdóttir

BRICS**
Department of Computer Science
University of Aalborg
Fredrik Bajers Vej 7E
DK-9220 Aalborg Ø

Abstract. This paper gives an equational axiomatization of probabilistic bisimulation equivalence for a class of finite-state agents previously studied by Stark and Smolka ((2000) *Proof, Language, and Interaction: Essays in Honour of Robin Milner*, pp. 571–595). The axiomatization is obtained by extending the general axioms of iteration theories (or iteration algebras), which characterize the equational properties of the fixed point operator on (ω-)continuous or monotonic functions, with three axiom schemas that express laws that are specific to probabilistic bisimilarity. Hence probabilistic bisimilarity (over finite-state agents) has an equational axiomatization relative to iteration algebras.

1 Introduction

Probabilistic variations on process algebras have been extensively studied in the literature, and concepts from concurrency theory have been extended to these languages and their underlying probabilistic models—see, e.g., [19] for a survey and many references to the original literature. As part of this research effort to lift process algebraic results to the probabilistic setting, several notions of probabilistic behavioural equivalences and preorders have been proposed in the literature over various models of probabilistic processes, and have been axiomatized over fragments of probabilistic process algebras. Works presenting complete axiomatizations of probabilistic semantic theories for processes are, e.g., [2, 4, 18, 20, 21, 25, 29]. Amongst the aforementioned references, the studies [18, 20, 21, 29] consider languages with finite-state recursive definitions, and offer implicational proof systems for probabilistic bisimulation equivalence. (Indeed, the language TPCCS studied by Hansson in [18] involves a combination of time and probabilities.)

* Permanent address: Dept. of Computer Science, University of Szeged, P.O.B. 652, 6701 Szeged, Hungary. Supported in part by the National Foundation of Hungary for Scientific Research.

** **B**asic **R**esearch **in** **C**omputer **S**cience, Centre of the Danish National Research Foundation.

In this paper, we contribute to the quest for complete axiomatizations of behavioural equivalences for probabilistic processes by offering a purely equational axiomatization of probabilistic bisimulation equivalence for a class of finite-state agents previously studied by, e.g., Stark and Smolka in [29]. The axiomatization is obtained by extending the general axioms of *iteration theories* (or *iteration algebras*) [6, 14], which characterize, among others, the equational properties of the fixed point operator on (ω-)continuous or monotonic functions, with three axiom schemas that express laws that are specific to probabilistic bisimilarity. Hence probabilistic bisimilarity (over finite-state agents) has an equational axiomatization relative to iteration algebras; this axiomatization is finite relative to iteration algebras if we allow for the use of equation schemas.

Historically, an implicational axiom system for probabilistic bisimilarity over finite-state processes was first proposed in [21], where its soundness and completeness were announced for a class of finite-state probabilistic agents with rational probabilities. The unpublished dissertation [20] offered a proof of the soundness and completeness result announced in [21], showing that the assumption of rational probabilities could be dropped. However, according to Stark and Smolka [29], some of the soundness proofs in [20] were flawed, and [29] is apparently the first study which gave a full proof of the soundness and completeness of the axiom system from [21] for a CCS-like language with (possibly unguarded) finite-state recursive definitions. In the meantime, Hansson [18] offered an implicational proof system for bisimilarity over the fragment of his language TPCCS with guarded finite-state recursion. Amongst all these original references, our technical developments in this paper have mostly been influenced by those in [29].

We believe that the results presented in this paper improve upon those previous axiomatizations of probabilistic bisimilarity for the language we consider. First of all, in light of the simplicity and foundational role played by equational logic, it is natural to look for purely equational axiomatizations of algebras of processes—as done, for instance, in the ACP family of process algebras (see, e.g., [16] for a textbook presentation). Moreover, whenever finite-state processes are concerned, implicational axiom systems based on variants on the unique fixed point induction rule for guarded terms, like those presented in the classic paper [26] and the aforementioned references, are somewhat unsatisfactory as they afford very few models. A classic example of this phenomenon is present in the long history of the quest for equational axiomatizations of the algebra of regular languages. Salomaa gave two complete axiomatizations of the algebra of regular languages in [28]. However, one of them contains an infinitary rule, and, as argued by Kozen in [22], the other is not sound in most common interpretations of regular expressions (such as binary relations) because it uses a version of the unique fixed point rule. Implicational axiomatizations for the equational theory of regular languages that are sound over a wealth of important nonstandard interpretations that arise in computer science have been given by Krob in [23] and Kozen in [22]. A purely equational axiomatization of regular languages has been offered by Krob in [23].

Giving a relative axiomatization of probabilistic bisimilarity with respect to iteration algebras has also the benefit of separating the general (embodied by the equations of iteration theories) from the specific (expressed by the equations that describe properties of probabilistic bisimilarity proper). This separation of concerns has at least two benefits. First of all, as a relative axiomatization of probabilistic bisimilarity can be given by adding three axiom schemas to those of iteration algebras, it follows that the nonfinite axiomatizability of the equational theory of probabilistic bisimilarity is due to that of iteration algebras (see, e.g., [12]). Secondly, any advance in the equational axiomatization of iteration algebras would yield an improved equational axiomatization for probabilistic bisimilarity.

That standard bisimulation equivalence is finitely axiomatizable relative to iteration algebras was shown in [6, Chapter 13].

2 Preliminaries

In order to make the paper self-contained, we now briefly review the basic notions from [29] and of iteration algebras that will be needed in this study. Moreover, we extend the operational semantics from *op. cit.* and the definition of probabilistic bisimilarity so that they apply to the whole language of probabilistic terms directly.

2.1 Probabilistic Finite-State Terms and Probabilistic Bisimilarity

We begin by presenting the syntax and the operational semantics of the language of finite-state probabilistic terms that will be studied in the remainder of the paper. Our presentation is based on that in [29], to which the reader is referred for more details and background information.

We use Var to stand for a countably infinite set of *agent variables*, ranged over by x, y, w, z possibly subscripted and/or superscripted, and Act to denote a nonempty collection of *atomic actions*, ranged over by a. The meta-variable α stands for an element of the set $\mathsf{Act} \cup \mathsf{Var}$.

The syntax of *probabilistic terms* (over Var and Act) is defined as follows:

$$t ::= x \mid at \mid t \;{}_p{+}\; t \mid \mu x.t \enspace ,$$

where $x \in \mathsf{Var}$, $a \in \mathsf{Act}$ and p is a real number in the open interval $(0, 1)$. The notions of *free* and *bound variables* are defined in the standard way—with $\mu x.\_$ as a binding construct—, and a variable x is *guarded* in term t if every free occurrence of x in t occurs within a subterm of the form at'. If $\overline{x} = (x_1, \ldots, x_n)$ is a vector of distinct variables, we shall sometimes write $t(\overline{x})$ to denote the fact that every free variable of t is in $\overline{x}$. (As usual, the free variables of t need not contain all of the variables in $\overline{x}$.) A term is *closed* if it does not contain any free variable. Throughout the paper, we consider two terms as syntactically identical if they are equal up to renaming of their bound variables. If $\overline{x} = (x_1, \ldots, x_n)$ is a vector of distinct variables, $\overline{t} = (t_1, \ldots, t_n)$ is a a vector of terms, and t is a term, then $t[\overline{t}/\overline{x}]$ denotes the term that results by substituting each occurrence of x_i in t with t_i. The definition of substitution in the presence of binders like μ is standard, and is therefore omitted. When writing terms, we assume that the scope of a $\mu x.\_$ extends to the right as far as possible. In the remainder of the paper, we use $\perp$ as an abbreviation for the closed term $\mu x.x$.

$$\frac{t \xrightarrow{\alpha} t'}{t \;{}_p\!+ u \xrightarrow{\alpha} t'} \qquad \frac{u \xrightarrow{\alpha} u'}{t \;{}_p\!+ u \xrightarrow{\alpha} u'} \qquad at \xrightarrow{a} t \qquad x \xrightarrow{x} \perp \qquad \frac{t[\mu x.t/x] \xrightarrow{\alpha} t'}{\mu x.t \xrightarrow{\alpha} t'}$$

Table 1. Transition rules for terms ($\alpha \in \mathsf{Act} \cup \mathsf{Var}$)

Following the approach adopted by Stark and Smolka in [29], we define the operational semantics for probabilistic terms in two steps. First, we give the transitions of terms using standard structural operational semantics [27]. (Cf. Table 1 for the rules. Note that, in our formulation of the operational semantics for terms, the statement $t \xrightarrow{x} \perp$ means that the variable x occurs unguarded in the term t—see, e.g., [1, 17] for similar semantics for fragments of regular CCS.) Next, we incorporate information about the probability of occurrence of transitions into the operational semantics by associating, with each triple (t, α, u) consisting of terms t, u and $\alpha \in \mathsf{Act} \cup \mathsf{Var}$, a transition probability $\mathrm{prob}(t, \alpha, u) \in [0, 1]$. Following Stark and Smolka, we shall use the more suggestive notation $\mathrm{prob}(t \xrightarrow{\alpha} u)$ in lieu of $\mathrm{prob}(t, \alpha, u)$. For the sake of completeness, we recall that the function prob is defined as the least solution (over the complete partial order of the set of all functions mapping triples (t, α, u) to the real numbers in the interval $[0, 1]$, ordered pointwise) of the recursive equation

$$\mathrm{prob} = \mathcal{P}(\mathrm{prob}) \;\; ,$$

where $\mathcal{P}$ is given by:

$$\begin{aligned}
\mathcal{P}(\mathrm{prob})(at \xrightarrow{\alpha} u) &= \begin{cases} 1 \text{ if } a = \alpha \text{ and } t = u \\ 0 \text{ otherwise} \end{cases} \\
\mathcal{P}(\mathrm{prob})(x \xrightarrow{\alpha} u) &= \begin{cases} 1 \text{ if } x = \alpha \text{ and } \perp = u \\ 0 \text{ otherwise} \end{cases} \\
\mathcal{P}(\mathrm{prob})(t_1 \;{}_p\!+ t_2 \xrightarrow{\alpha} u) &= p \cdot \mathrm{prob}(t_1 \xrightarrow{\alpha} u) + (1-p) \cdot \mathrm{prob}(t_2 \xrightarrow{\alpha} u) \\
\mathcal{P}(\mathrm{prob})(\mu x.t \xrightarrow{\alpha} u) &= \mathrm{prob}(t[\mu x.t/x] \xrightarrow{\alpha} u) \;\; .
\end{aligned}$$

For example, we have that, for every $p \in (0, 1)$,

$$\mathrm{prob}(\mu x.a\perp \;{}_p\!+ x \xrightarrow{a} \perp) = 1 = \mathrm{prob}(\mu x.x \;{}_p\!+ y \xrightarrow{y} \perp) \;\; .$$

We refer the interested reader to [29] for more information on the definition of the probability assigning function prob, and on its connections with the structural operational semantics in Table 1. Here we limit ourselves to recalling that $\mathrm{prob}(t \xrightarrow{\alpha} u)$ is positive if, and only if, $t \xrightarrow{\alpha} u$ can be inferred from the rules in Table 1 (cf. [29, Lem. 2.1]), and that, for every term t, set S of terms and $\alpha \in \mathsf{Act} \cup \mathsf{Var}$, the summation

$$\sum_{u \in S} \mathrm{prob}(t \xrightarrow{\alpha} u)$$

converges to a value between 0 and 1 (cf. [29, Propn. 2.2]). Following Stark and Smolka, we use $\mathrm{prob}(t \xrightarrow{\alpha} S)$ to denote this value.

Remark 1. Unlike Stark and Smolka in the developments in [29], we have presented the operational semantics for terms, possibly containing free variables, in the language we study. In our operational semantics, $\mathrm{prob}(t \xrightarrow{x} \bot)$ measures the "probability of unguardedness" of variable x in term t. Note that $\mathrm{prob}(t \xrightarrow{x} \bot)$ does *not* coincide with $\mathrm{unguard}_t(x)$, a measure of the total probability assigned to unguarded occurrences of variable x in term t defined by Stark and Smolka in [29, Page 587]. For example, if t is the term $\mu x.x \ {}_p{+}\ y$, with $p \in (0, 1)$, then

$$\mathrm{unguard}_t(y) = (1 - p) \neq 1 = \mathrm{prob}(t \xrightarrow{y} \bot) \ .$$

The reason for using the definition given here instead of the one by Stark and Smolka is that, unlike the one given in *op. cit.*, the probability of unguardedness of variables it gives is stable under behavioural equivalence.

In what follows, when we refer to the probabilistic labelled transition system determined by a term, we mean the fragment of the transition system generated by the aforementioned operational semantics that can be reached from it.

Notation 1 *In what follows, we shall sometimes use the suggestive notation* $t \xrightarrow{p;\alpha} u$ *to denote the fact that* $\mathrm{prob}(t \xrightarrow{\alpha} u) = p$ *and* $p > 0$*, i.e., that* t *can perform* α *with positive probability* p*, and become* u *in doing so.*

The notion of behavioural equivalence we shall consider in this paper is probabilistic bisimilarity. This we now proceed to present, extended to the whole set of terms.

Definition 1. *A* probabilistic bisimulation *is an equivalence relation* $\mathcal{R}$ *over terms that satisfies the following condition:*

> *Whenever* $t\mathcal{R}u$*, then for all* $\alpha \in \mathsf{Act} \cup \mathsf{Var}$ *and all equivalence classes* S *of* $\mathcal{R}$ *we have that*
>
> $$\mathrm{prob}(t \xrightarrow{\alpha} S) = \mathrm{prob}(u \xrightarrow{\alpha} S) \ .$$

Two terms t *and* u *are* probabilistically bisimilar*, written* $t \overset{\mathrm{pr}}{\sim} u$*, iff there is a probabilistic bisimulation that relates them.*

The relation $\overset{\text{pr}}{\sim}$ will henceforth be referred to as *probabilistic bisimilarity*.

Example 1. It is not hard to see that the least equivalence relation containing all the pairs of the form $(\mu x.x \;_p{+}\; y, y)$, with $p \in (0,1)$, is a probabilistic bisimulation. Thus, for every $p \in (0,1)$, it holds that

$$\mu x.x \;_p{+}\; y \overset{\text{pr}}{\sim} y \ .$$

Remark 2. The definition of probabilistic bisimilarity given above is an extension to the whole set of terms of the original one by Larsen and Skou in [24]. The definitions of this relation given in, e.g., [4, 19] are based on an extension of equivalence relations to probability distributions. The two definitions coincide.

The import of the following theorem is that the explicit definition of probabilistic bisimilarity for the whole language of probabilistic terms we have presented coincides with the one given by Stark and Smolka in [29].

Theorem 1. *Let $t(\overline{x})$ and $u(\overline{x})$ be terms. Then $t(\overline{x}) \overset{\text{pr}}{\sim} u(\overline{x})$ iff for all vectors of closed terms $\overline{t}$, it holds that $t[\overline{t}/\overline{x}] \overset{\text{pr}}{\sim} u[\overline{t}/\overline{x}]$.*

A proof of the following result, due to Stark and Smolka, may be found in [29, Sect. 4].

Proposition 1. *The relation of probabilistic bisimilarity is a congruence over the language of finite-state probabilistic terms. Thus, whenever t and u are probabilistically bisimilar, so are the terms*

- *at and au, for every action a;*
- *$t \;_p{+}\; t'$ and $u \;_p{+}\; t'$, for every term t';*
- *$t' \;_p{+}\; t$ and $t' \;_p{+}\; u$, for every term t'; and*
- *$\mu x.t$ and $\mu x.u$, for every variable x.*

2.2 Axioms of Iteration Algebras

The axioms of *iteration algebras* (or iteration theories) [6] capture the equational properties of the fixed point operation (be it least, unique, initial, etc.). Several (conditional) equational bases of identities for iteration algebras have been studied in the literature (cf., e.g., *op. cit.* and the references [7, 13, 15]). In this study, we shall specifically consider an equational axiomatization of iteration algebras obtained by the second author in [14]. This equational basis for iteration algebras consists of the so-called *Conway equations* [6, 7]

$$\mu x.t[t'/x] = t[\mu x.t'[t/x]/x] \tag{1}$$

$$\mu x.t[x/y] = \mu x.\mu y.t \ , \tag{2}$$

and of a set of equations containing one equation for each finite (simple) group. (Group equations for the language of μ-terms were introduced in [14] as a generalization of Conway's group equations for regular languages, cf. [8]. The completeness of the Conway equations and the group equations for iteration algebras extends Krob's result in [23], where he confirmed a long standing conjecture of Conway [8] about the axiomatization of the equational theory of regular sets.)

In the setting of monotonic and continuous functions, equations (1)–(2) above were established by de Bakker, Bekič, Scott and others (see, e.g., [3, 5]), and are sometimes referred to as the *composition identity* (also known as the *rolling identity*) and the *diagonal identity* (also known as the *double-dagger identity*), respectively. Note that the classic *fixed point equation*, viz.

$$\mu x.t = t[\mu x.t/x] \ , \tag{3}$$

is the instance of the composition identity obtained by taking t' to be the variable x.

To define the group equations, we need to extend the μ-notation to term vectors $\overline{t} = (t_1, \ldots, t_n)$. (Henceforth, we shall consider term vectors as ordinary terms.) Let $\overline{x} = (x_1, \ldots, x_n)$ be a vector of distinct variables. When $n = 1$, we use $\mu\overline{x}.\overline{t}$ to denote the term vector of dimension one whose unique component is $\mu x_1.t_1$. (We identify any term vector of dimension one with its component.) If $n > 1$, let $\overline{x}' = (x_1, \ldots, x_{n-1})$, $\overline{t}' = (t_1, \ldots, t_{n-1})$ and $\overline{s} = \overline{t}'[\mu x_n.t_n/x_n]$. (Substitution into a term vector is defined componentwise.) We define

$$\mu\overline{x}.\overline{t} \stackrel{\text{def}}{=} (\mu\overline{x}'.\overline{s}, (\mu x_n.t_n)[\mu\overline{x}'.\overline{s}/\overline{x}']) \ .$$

The definition is motivated by the Bekič-de Bakker-Scott rule [3, 5].

Suppose now that $(G, \cdot)$ is a finite group of order n, whose elements are the integers in the set $[n] = \{1, \ldots, n\}$. Given a vector $\overline{x} = (x_1, \ldots, x_n)$ of distinct variables and an integer $i \in [n]$, define $i \cdot \overline{x} = (x_{i \cdot 1}, \ldots, x_{i \cdot n})$. Thus, $i \cdot \overline{x}$ is obtained by permuting the components of $\overline{x}$ according to the ith row of the multiplication table of G. The *group equation associated with* G is

$$(\mu\overline{x}.(t[1 \cdot \overline{x}/\overline{x}], \ldots, t[n \cdot \overline{x}/\overline{x}]))_1 = \mu y.t[y/x_1, \ldots, y/x_n] \ , \tag{4}$$

where t is any μ-term, y is a variable, and where

$$(\mu\overline{x}.(t[1 \cdot \overline{x}/\overline{x}], \ldots, t[n \cdot \overline{x}/\overline{x}]))_1$$

is the first component of the term vector $\mu\overline{x}.(t[1 \cdot \overline{x}/\overline{x}], \ldots, t[n \cdot \overline{x}/\overline{x}])$.

The group equations are a special case of the *commutative identity*, which is entailed by the *weak functorial implication* (see, e.g., [6, Chapter 6, Sect. 4]). The weak functorial implication can be stated as follows:

> For terms $t_i(x_1, \ldots, x_n, \overline{y})$ ($i \in [n]$) and $t(x, \overline{y})$, if $t_i(x, \ldots, x, \overline{y}) = t(x, \overline{y})$ for every $i \in [n]$, then
>
> $$(\mu(x_1, \ldots, x_n).(t_1, \ldots, t_n))_1 = \mu x.t(x, \overline{y}) \ .$$

The above implication will play an important role in the technical developments to follow (see the appendix).

S1	$x\ {}_p\!+\ y = y\ {}_{1-p}\!+\ x$		
S2	$x\ {}_p\!+\ (y\ {}_q\!+\ z) = (x\ {}_r\!+\ y)\ {}_s\!+\ z$	if $C(p,q,r,s)$ holds	
S3	$x\ {}_p\!+\ x = x$		
R2	$\mu x.t\ {}_p\!+\ x = \mu x.t$		

Table 2. Stark and Smolka's Equational Axioms for Probabilistic Bisimilarity

3 An Equational Axiomatization of Probabilistic Bisimilarity

The main result of [29] is a complete implicational axiomatization for probabilistic bisimilarity over probabilistic finite-state terms. The axiomatization offered by Stark and Smolka in *op. cit.* consists of the fixed point equation (3) (axiom R1 in *op. cit.*), the unique fixed point rule for guarded terms and of the equations in Table 2. In equation S2, the condition $C(p,q,r,s)$ holds true whenever $p = rs$, $(1-p)q = (1-r)s$ and $(1-s) = (1-p)(1-q)$. We recall that the *unique fixed point rule* for guarded terms, axiom R3 in [29], states that:

> From $t = u[t/x]$, where all occurrences of x in u are guarded, infer that $t = \mu x.u$.

The main aim of this study is to show that the equational laws of probabilistic bisimilarity over finite-state probabilistic terms have a natural axiomatization over the equations of iteration algebras. To this end, we shall consider the purely equational axiom system Ax obtained by extending the equational basis of iteration algebras consisting of (1), (2) and the group equations (4) with axioms S1 and S2 from Table 2, and the following equation

$$\mu x.(x\ {}_p\!+\ y) = y\ . \tag{5}$$

The above equation expresses a strengthened form of idempotence of the ${}_p\!+$ operation. In fact, equation S3 in Table 2 follows from it and the fixed point equation (3) thus:

$$y = \mu x.x\ {}_p\!+\ y = (\mu x.x\ {}_p\!+\ y)\ {}_p\!+\ y = y\ {}_p\!+\ y\ .$$

Moreover, in the presence of axiom S1 and of the diagonal equation, (5) proves equation R2 in Table 2. Indeed:

$$\begin{aligned}
\mu x.t\ {}_p\!+\ x &= \mu x.\mu z.t\ {}_p\!+\ z \qquad (z \text{ fresh})\\
&= \mu x.((\mu z.y\ {}_p\!+\ z)[t/y]) \qquad (y \text{ fresh})\\
&= \mu x.(y[t/y])\\
&= \mu x.t\ .
\end{aligned}$$

The remainder of this paper will be devoted to a proof of the following soundness and completeness theorem:

Theorem 2. *The axiom system* Ax *completely axiomatizes probabilistic bisimilarity over the language of finite-state probabilistic terms, i.e., for all terms t, u, the equivalence $t \overset{\text{pr}}{\sim} u$ holds iff the equality $t = u$ is provable using the equations in* Ax*.*

4 Soundness

Our first step towards a proof of Theorem 2 will be to show the following theorem, to the effect that the axiom system Ax is sound with respect to probabilistic bisimilarity.

Theorem 3 (Soundness). *For all terms t and u, if* Ax *proves that $t = u$ then $t \overset{\text{pr}}{\sim} u$.*

Since Stark and Smolka have proven in [29, Sect. 4] that their axiom system is sound with respect to $\overset{\text{pr}}{\sim}$, the above theorem follows from the following result to the effect that the implicational axiom system due to Stark and Smolka entails Ax:

Theorem 4. *Every equation in* Ax *can be proven from the axiom system for probabilistic bisimilarity due to Stark and Smolka.*

Remark 3. Another, possibly more standard, approach to showing the soundness of Ax with respect to probabilistic bisimilarity is to identify the probabilistic transition systems associated with the left- and right-hand sides of all of the equations in Ax, and to exhibit appropriate probabilistic bisimulations between them—as we did for axiom (5) in Example 1. We have, however, plumped for an equational approach to such a proof because it can be better formalized for complex equations like the Conway and group equations. Moreover, Theorem 4 gives more information than the mere soundness of Ax with respect to probabilistic bisimilarity. For example, it entails that the models of the axiom system by Stark and Smolka are also models of Ax.

It is clear that axiom (5) is derivable from the axiom system by Stark and Smolka by using axioms S1 and R2 in Table 2, and the fixed point equation. It is much less clear that so are the Conway and group equations. The interested reader may find the details of the non-trivial equational arguments used in the proofs of these equations from the axiom system by Stark and Smolka in the appendix.

5 Normal Forms

The next step in the proof of our main result is the isolation of a suitable notion of normal form for finite-state probabilistic terms. Normal forms have a direct interpretation as finite-state probabilistic transition systems, and this will be crucial in establishing the completeness of our axiom system. We prove that, modulo Ax, every term is provably equal to one in normal form (Theorem 5).

We begin by introducing a useful notation that will help clarify the connection between terms in normal form, and the probabilistic transition systems they denote. In this notation, we use the notion of *stochastic vector*, which is a vector of real numbers in the interval $[0,1]$ that sum up to 1.

Notation 2 *Let $\{(p_i, t_i) \mid i \in [k]\}$ be a nonempty set of pairs, where each t_i is a term, and $(p_1, \ldots, p_k)$ is a stochastic vector. The notation*

$$\sum_{i \in [k]} p_i \cdot t_i$$

is defined recursively as follows:

$$\sum_{i \in [k]} p_i \cdot t_i \stackrel{def}{=} \begin{cases} t_k & \text{if } p_k = 1 \\ \sum_{i \in [k-1]} p_i \cdot t_i & \text{if } p_k = 0 \\ (\sum_{i \in [k-1]} (p_i/(1-p_k)) \cdot t_i) \ {}_{1-p_k}{+} \ t_k & \text{if } p_k \in (0,1) \end{cases} .$$

In what follows, we shall often write

$$p_1 \cdot t_1 + \cdots + p_k \cdot t_k$$

in lieu of $\sum_{i \in [k]} p_i \cdot t_i$, and t instead of $1 \cdot t$. The $\cdot$'s will often be omitted in equational derivations.

For example, we write

$$\frac{1}{4} \cdot a\bot + \frac{1}{2} \cdot x + \frac{1}{4} \cdot \bot$$

for the term

$$(a\bot \ {}_{1/3}{+} \ x) \ {}_{3/4}{+} \ \bot \ .$$

Definition 2. *A* simple term *in the variables $\overline{x} = (x_1, \ldots, x_n)$ and parameters $\overline{y} = (y_1, \ldots, y_m)$ is a term of the form*

$$p_1 \cdot t_1 + \cdots + p_k \cdot t_k + q_1 \cdot y_1 + \cdots + q_m \cdot y_m + q \cdot \bot \ , \tag{6}$$

where:

- *$k, m \geq 0$,*
- *$(p_1, \ldots, p_k, q_1, \ldots, q_m, q)$ is a stochastic vector,*
- *for every $i \in [k]$, the term t_i is of the form ax_ℓ for some action a and variable $x_\ell \in \{x_1, \ldots, x_n\}$, and*
- *for every $i_1, i_2 \in [k]$, if $i_1 \neq i_2$ then $t_{i_1} \neq t_{i_2}$.*

(If k or m are 0, then the corresponding part of a simple term is missing.)

A summand *of a simple term of the form (6) is a subterm of the form $p \cdot t$, where p is positive.*

Note that, in light of the last condition above on simple terms, modulo S1 and S2, axiom S3 in Table 2 cannot be applied to a simple term when used as a rewrite rule from left to right. For example, the term

$$\frac{2}{3} \cdot ax + \frac{1}{3} \cdot ax$$

is not simple, but using S3 from left to right it can be proven equal to the simple term $1 \cdot ax$, that is to ax.

Definition 3 (Normal Form Terms). *A* normal form term *in the parameters* $\overline{y} = (y_1, \ldots, y_m)$ *is a term*

$$\mu\overline{x}.\overline{t} = \mu(x_1, \ldots, x_n).(t_1, \ldots, t_n) \ ,$$

where each t_i *(*$i \in [n]$*) is a simple term in the variables* $\overline{x} = (x_1, \ldots, x_n)$ *and parameters* $\overline{y} = (y_1, \ldots, y_m)$*.*

Intuitively, the ith component of a term vector $\mu(x_1, \ldots, x_n).(t_1, \ldots, t_n)$ is the ith component of a distinguished solution of the list of equations

$$\begin{aligned} x_1 &= t_1 \\ &\vdots \\ x_n &= t_n \ . \end{aligned}$$

We shall sometimes identify a normal form term with its corresponding list of equations. The main reason for doing so is that the list of equations associated with a normal form term can be naturally viewed as a kind of probabilistic transition system. Indeed, the set of states of such a probabilistic transition system may be taken to be the integers in the set $[n]$—here, integer i stands for the ith component of the term vector determined by the list of equations—plus a distinguished $\perp$ state. If the simple term t_i has the form (6), the set of transitions out of state $i \in [n]$ is defined as follows:

- for every $j \in [n]$, there is a transition $i \stackrel{p,a}{\to} j$ if, and only if, $p \cdot ax_j$ is a summand of the simple term t_i—that is, when $t_i \stackrel{p,a}{\to} x_j$ holds; and
- for every variable x, there is a transition $i \stackrel{p,x}{\to} \perp$ if, and only if, $p \cdot x$ is a summand of the simple term t_i—that is, when $t_i \stackrel{p,x}{\to} \perp$ holds.

We use $\mathrm{ts}(\overline{t}(\overline{x}, \overline{y}))$ to denote this probabilistic transition system, and say that a normal form $\mu\overline{x}.\overline{t}(\overline{x}, \overline{y})$ is *accessible* if every state of $\mathrm{ts}(\overline{t}(\overline{x}, \overline{y}))$ is reachable from state 1.

Proposition 2. *The transition system* $\mathrm{ts}(\overline{t}(\overline{x}, \overline{y}))$ *is probabilistically bisimilar to* $\mu\overline{x}.\overline{t}(\overline{x}, \overline{y})$*, for every normal form* $\mu\overline{x}.\overline{t}(\overline{x}, \overline{y})$*.*

Remark 4. Actually, the transition systems $\mathrm{ts}(\overline{t}(\overline{x}, \overline{y}))$ and $\mu\overline{x}.\overline{t}(\overline{x}, \overline{y})$ are not just probabilistically bisimilar, but also strongly equivalent—in the sense that the two transition systems "unfold to the same tree" (cf. [11]).

The following normal form theorem is a version of Milner's equational characterization of regular CCS process, cf. [26]. A version of Milner's equational characterization theorem for the finite-state probabilistic terms we consider has been given by Stark and Smolka in [29, Thm. 2].

Theorem 5. *For every term $t(\overline{y})$, there is an accessible normal form term $\mu\overline{x}.\overline{t}(\overline{x},\overline{y})$ such that the equality $t(\overline{y}) = (\mu\overline{x}.\overline{t}(\overline{x},\overline{y}))_1$ is provable from the axioms in* Ax.

Proof. The proof follows the lines of similar arguments in, e.g., [6], and uses the Conway equations (1)–(2) and their vector forms, axioms S1 and S2 from Table 2 and equation (5). Related arguments may be found in, e.g., [10, 26, 28, 29]. □

¿From now on, we shall assume that terms in normal form are accessible. Moreover, we equip the transition system $\mathrm{ts}(\overline{t}(\overline{x},\overline{y}))$ with the initial state 1. Probabilistic bisimulations between two such transition systems will relate their initial states.

6 Completeness

In Sect. 4, we established the soundness of our axiom system Ax with respect to probabilistic bisimilarity by showing that the axiom system proposed by Stark and Smolka in [29] entails it. We now aim at proving that, like the one by Stark and Smolka, our axiom system is complete with respect to probabilistic bisimilarity. This is the import of the following theorem:

Theorem 6 (Completeness). *For all terms t and u, if $t \overset{\mathrm{pr}}{\sim} u$ then* Ax *proves that $t = u$.*

The remainder of this section will be devoted to a proof of the above result. Apart from the normal form theorem (Theorem 5), our proof of the completeness theorem consists of two main ingredients.

- First we shall show that, in a suitable technical sense, two probabilistically bisimilar normal forms are structurally related.
- Next, we use the structural relationship between probabilistically bisimilar normal forms to show that two equivalent normal forms are provably equal using Ax. It is this final step of the proof that relies upon the group equations and the full power of the axioms of iteration theories.

We now proceed to study the structural relation that exists between probabilistically bisimilar normal forms. In the remainder of this section, equality of terms is modulo the axioms S1–S3 in Table 2.

Definition 4. *Let $\overline{x} = (x_1,\ldots,x_n)$ and $\overline{z} = (z_1,\ldots,z_k)$ be two vectors, each consisting of distinct variables. Let, furthermore, ρ be a function mapping $[n]$ to $[k]$. For every term $t(\overline{x},\overline{y})$, we define the term $t \circ \rho$ thus:*

$$t \circ \rho \overset{def}{=} t[z_{\rho(1)}/x_1,\ldots,z_{\rho(n)}/x_n] \enspace .$$

If $\overline{t} = (t_1, \ldots, t_n)$ is a vector of terms over variables $\overline{x} = (x_1, \ldots, x_n)$, then we write $\overline{t} \circ \rho$ for the vector of terms $(t_1 \circ \rho, \ldots, t_n \circ \rho)$.

If $\overline{u} = (u_1, \ldots, u_k)$ is a vector of terms, then we write $\rho \circ \overline{u}$ for the vector of terms $(u_{\rho(1)}, \ldots, u_{\rho(n)})$.

Note that when the components of $\overline{t}$ are simple, then, modulo S1–S3, so are the components of $\overline{t} \circ \rho$.

Example 2. Consider the vectors of simple terms $\overline{t} = (t_1, t_2, t_3)$ and $\overline{u} = (u_1, u_2)$ over variables $\overline{x} = (x_1, x_2, x_3)$, where

$$\begin{aligned} t_1 &= \frac{1}{3} \cdot ax_2 + \frac{2}{3} \cdot ax_3 \\ t_2 &= ax_2 \\ t_3 &= \frac{1}{2} \cdot ax_2 + \frac{1}{2} \cdot ax_3 \\ u_1 &= ax_1 \qquad \text{and} \\ u_2 &= \text{an arbitrary simple term} \ . \end{aligned}$$

Let ρ map each $i \in [3]$ to 1. Then, modulo the axioms S1–S3 in Table 2,

$$\overline{t} \circ \rho = (ax_1, ax_1, ax_1) = \rho \circ \overline{u} \ .$$

Whenever $\mu\overline{x}.\overline{t}(\overline{x}, \overline{y})$ and $\mu\overline{z}.\overline{u}(\overline{z}, \overline{y})$ are two terms in normal form over variables $\overline{x} = (x_1, \ldots, x_n)$ and $\overline{z} = (z_1, \ldots, z_k)$, respectively, we say that a function $\rho : [n] \to [k]$ determines a probabilistic bisimulation from the probabilistic transition system $\mathrm{ts}(\overline{t}(\overline{x}, \overline{y}))$ to $\mathrm{ts}(\overline{u}(\overline{z}, \overline{y}))$ if, and only if, the least equivalence relation over the disjoint union of $[n]$ and $[k]$ containing the graph of ρ is a probabilistic bisimulation.

Proposition 3. *Let $\mu\overline{x}.\overline{t}(\overline{x}, \overline{y})$ and $\mu\overline{z}.\overline{u}(\overline{z}, \overline{y})$ be two terms in normal form over variables $\overline{x} = (x_1, \ldots, x_n)$ and $\overline{z} = (z_1, \ldots, z_k)$, respectively. A function $\rho : [n] \to [k]$ determines a probabilistic bisimulation from $\mathrm{ts}(\overline{t}(\overline{x}, \overline{y}))$ to $\mathrm{ts}(\overline{u}(\overline{z}, \overline{y}))$ if, and only if, the following two conditions hold:*

1. *$\rho(1) = 1$, and*
2. *$\overline{t} \circ \rho = \rho \circ \overline{u}$ modulo axioms S1–S3 in Table 2.*

Notation 3 *In what follows, we shall write $\overline{t}(\overline{x}, \overline{y}) \underset{\rho}{\to} \overline{u}(\overline{z}, \overline{y})$ when ρ determines a probabilistic bisimulation from $\mathrm{ts}(\overline{t}(\overline{x}, \overline{y}))$ to $\mathrm{ts}(\overline{u}(\overline{z}, \overline{y}))$.*

Proposition 4. *Let $\mu\overline{x}.\overline{t}(\overline{x}, \overline{y})$ and $\mu\overline{z}.\overline{u}(\overline{z}, \overline{y})$ be two terms in normal form over variables $\overline{x} = (x_1, \ldots, x_n)$ and $\overline{z} = (z_1, \ldots, z_k)$ respectively. Then the probabilistic transition system $\mathrm{ts}(\overline{t}(\overline{x}, \overline{y}))$ is probabilistically bisimilar to $\mathrm{ts}(\overline{u}(\overline{x}, \overline{y}))$ if, and only if, there are a normal form $\mu\overline{w}.\overline{r}(\overline{w}, \overline{y})$ (over variables $\overline{w} = (w_1, \ldots, w_\ell)$), and functions $\rho : [\ell] \to [n]$ and $\tau : [\ell] \to [k]$ such that*

$$\overline{t}(\overline{x}, \overline{y}) \underset{\rho}{\leftarrow} \overline{r}(\overline{w}, \overline{y}) \underset{\tau}{\to} \overline{u}(\overline{z}, \overline{y}) \ .$$

In light of Propositions 2–4, in order to prove the completeness of our axiom system with respect to probabilistic bisimilarity it would be sufficient to show that:

Proposition 5. *Whenever* $\mu\overline{x}.\overline{t}(\overline{x},\overline{y})$ *and* $\mu\overline{z}.\overline{u}(\overline{z},\overline{y})$ *are two terms in normal form over variables* $\overline{x} = (x_1,\ldots,x_n)$ *and* $\overline{z} = (z_1,\ldots,z_k)$*, respectively, and* $\rho : [n] \to [k]$ *is a function meeting the constraints in the statement of Proposition 3, then the axiom system* Ax *proves that*

$$(\mu\overline{x}.\overline{t})_1 = (\mu\overline{z}.\overline{u})_1 \ .$$

Indeed, using the above statement, we can prove Theorem 6 thus:

Proof of Theorem 6: Let t and u be two probabilistically bisimilar terms. By Theorem 5, there are normal forms $\mu\overline{x}.\overline{t}(\overline{x},\overline{y})$ and $\mu\overline{z}.\overline{u}(\overline{z},\overline{y})$ such that Ax proves that

$$t = (\mu\overline{x}.\overline{t}(\overline{x},\overline{y}))_1 \qquad \text{and} \qquad u = (\mu\overline{z}.\overline{u}(\overline{z},\overline{y}))_1 \ .$$

By Proposition 2 and the soundness of the axiom system Ax with respect to probabilistic bisimilarity, we have that the probabilistic transition system $\mathrm{ts}(\overline{t}(\overline{x},\overline{y}))$ is probabilistically bisimilar to $\mathrm{ts}(\overline{u}(\overline{z},\overline{y}))$. By Proposition 4, it follows that there are a normal form $\mu\overline{w}.\overline{r}(\overline{w},\overline{y})$, and functions ρ and τ such that

$$\overline{t}(\overline{x},\overline{y}) \underset{\rho}{\leftarrow} \overline{r}(\overline{w},\overline{y}) \underset{\tau}{\rightarrow} \overline{u}(\overline{z},\overline{y}) \ .$$

By Proposition 5, Ax proves that

$$(\mu\overline{x}.\overline{t}(\overline{x},\overline{y}))_1 = (\mu\overline{w}.\overline{r}(\overline{w},\overline{y}))_1 = (\mu\overline{z}.\overline{u})_1 \ .$$

By transitivity, we thus obtain $t = u$, which was to be shown. $\square$

Our order of business is now to prove Proposition 5. In fact, we establish the following strengthening of that statement:

Proposition 6. *Let* $\mu\overline{x}.\overline{t}(\overline{x},\overline{y})$ *and* $\mu\overline{z}.\overline{u}(\overline{z},\overline{y})$ *be two terms in normal form over variables* $\overline{x} = (x_1,\ldots,x_n)$ *and* $\overline{z} = (z_1,\ldots,z_k)$*, respectively. Assume that* $\rho : [n] \to [k]$ *is a function meeting the constraints in the statement of Proposition 3. Then* Ax *proves that*

$$\mu\overline{x}.\overline{t} = \rho \circ (\mu\overline{z}.\overline{u}) \ . \tag{7}$$

To prove the above result, we rely on the fact that (7) is implied by the identities of iteration algebras if the term vectors $\overline{t}(\overline{x},\overline{y})$ and $\overline{u}(\overline{z},\overline{y})$ satisfy the following condition: There exist a vector $\overline{w} = (w_1,\ldots,w_\ell)$ of variables, term vectors $r_j(\overline{w},\overline{y})$, $j \in \rho([n])$, and functions $\rho_i : [\ell] \to [n]$, $i \in [n]$, such that for all $i \in [n]$ and $j \in [k]$ with $\rho(i) = j$ it holds that

$$r_j \circ \rho_i = t_i$$

modulo axioms S1–S3 in Table 2, where r_j and t_i denote the jth component of $\overline{r}$ and the ith component of $\overline{t}$, respectively. (See the *generalized commutative identity* on p. 138 in [6].) We establish this condition by using:

Lemma 1. *Let c be a positive real number. Suppose furthermore that $\overline{c_i} = (c_{i1}, \ldots, c_{ik_i})$ ($i \in [n]$) are nonempty sequences of positive real numbers all summing up to c. Then there is a sequence $(d_1, \ldots, d_\ell)$ of positive real numbers such that, for every $i \in [n]$, there are positive integers $\ell_1 < \ell_2 < \cdots < \ell_{k_i - 1} < \ell$ with*

$$\begin{aligned} c_{i1} &= d_1 + \cdots + d_{\ell_1} \\ c_{i2} &= d_{\ell_1+1} + \cdots + d_{\ell_2} \\ &\vdots \\ c_{ik_i} &= d_{(\ell_{k_i-1}+1)} + \cdots + d_\ell \ . \end{aligned}$$

References

1. L. Aceto, W. J. Fokkink, R. J. van Glabbeek, and A. Ingólfsdóttir, *Axiomatizing prefix iteration with silent steps*, Inform. and Comput., 127 (1996), pp. 26–40.
2. J. C. M. Baeten, J. A. Bergstra, and S. A. Smolka, *Axiomatizing probabilistic processes: ACP with generative probabilities*, Inform. and Comput., 121 (1995), pp. 234–255.
3. J. W. de Bakker and D. Scott, *A theory of programs*, Technical Report, IBM Laboratory, Vienna, 1969.
4. E. Bandini and R. Segala, *Axiomatizations for probabilistic bisimulation*, in Proceedings of the 28th International Colloquium on Automata, Languages and Programming (ICALP) 2001, vol. 2076 of Lecture Notes in Computer Science, Springer-Verlag, July 2001, pp. 370–381.
5. H. Bekič, *Definable operations in general algebras, and the theory of automata and flowcharts*, Technical Report, IBM Laboratory, Vienna, 1969.
6. S. L. Bloom and Z. Ésik, *Iteration Theories*, Springer-Verlag, Berlin, 1993.
7. ———, *The equational logic of fixed points*, Theoret. Comput. Sci., 179 (1997), pp. 1–60.
8. J. H. Conway, *Regular Algebra and Finite Machines*, Mathematics Series (R. Brown and J. De Wet eds.), Chapman and Hall, London, United Kingdom, 1971.
9. S. Eilenberg, *Automata, Languages, and Machines. Vol. A*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York, 1974. Pure and Applied Mathematics, Vol. 58.
10. C. C. Elgot, *Monadic computation and iterative algebraic theories*, in Logic Colloquium '73 (Bristol, 1973), North-Holland, Amsterdam, 1975, pp. 175–230. Studies in Logic and the Foundations of Mathematics, Vol. 80.
11. C. C. Elgot, S. L. Bloom, and R. Tindell, *On the algebraic structure of rooted trees*, J. Comput. System Sci., 16 (1978), pp. 362–399.
12. Z. Ésik, *Independence of the equational axioms for iteration theories*, J. Comput. System Sci., 36 (1988), pp. 66–76.
13. ———, *Completeness of Park induction*, Theoret. Comput. Sci., 177 (1997), pp. 217–283. Mathematical foundations of programming semantics (Manhattan, KS, 1994).
14. ———, *Group axioms for iteration*, Inform. and Comput., 148 (1999), pp. 131–180.
15. ———, *The power of the group-identities for iteration*, Internat. J. Algebra Comput., 10 (2000), pp. 349–373.

16. W. J. Fokkink, *Introduction to Process Algebra*, Texts in Theoretical Computer Science, An EATCS Series, Springer-Verlag, 2000.
17. R. J. van Glabbeek, *A complete axiomatization for branching bisimulation congruence of finite-state behaviours*, in Mathematical Foundations of Computer Science 1993, Gdansk, Poland, A. Borzyszkowski and S. Sokolowski, eds., vol. 711 of Lecture Notes in Computer Science, Springer-Verlag, 1993, pp. 473–484.
18. H. Hansson, *Time and Probability in Formal Design of Distributed Systems*, vol. 1 of Real-Time Safety Critical Systems, Elsevier, 1994.
19. B. Jonsson, W. Yi, and K. G. Larsen, *Probabilistic extensions of process algebras*, in Handbook of Process Algebra, North-Holland, Amsterdam, 2001, pp. 685–710.
20. C.-C. Jou, *Aspects of Probabilistic Process Algebra*, PhD thesis, SUNY at Stony Brook, Stony Brook, New York, 1991.
21. C.-C. Jou and S. A. Smolka, *Equivalences, congruences, and complete axiomatizations for probabilistic processes*, in Proceedings CONCUR 90, Amsterdam, J. Baeten and J. Klop, eds., vol. 458 of Lecture Notes in Computer Science, Springer-Verlag, 1990, pp. 367–383.
22. D. Kozen, *A completeness theorem for Kleene algebras and the algebra of regular events*, Inform. and Comput., 110 (1994), pp. 366–390.
23. D. Krob, *Complete systems of B-rational identities*, Theoretical Comput. Sci., 89 (1991), pp. 207–343.
24. K. G. Larsen and A. Skou, *Bisimulation through probabilistic testing*, Information and Computation, 94 (1991), pp. 1–28.
25. ———, *Compositional verification of probabilistic processes*, in Proceedings CONCUR 92, Stony Brook, NY, USA, W. R. Cleaveland, ed., vol. 630 of Lecture Notes in Computer Science, Springer-Verlag, 1992, pp. 456–471.
26. R. Milner, *A complete inference system for a class of regular behaviours*, J. Comput. System Sci., 28 (1984), pp. 439–466.
27. G. Plotkin, *A structural approach to operational semantics*, Report DAIMI FN-19, Computer Science Department, Aarhus University, 1981.
28. A. Salomaa, *Two complete axiom systems for the algebra of regular events*, J. Assoc. Comput. Mach., 13 (1966), pp. 158–169.
29. 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, MIT Press, Cambridge, MA, 2000, pp. 571–595.

A Proof of the Soundness Theorem

For use in the technical developments to follow, we begin by noting that the set obtained by adding ∞ to the collection of nonnegative real numbers is a *complete semiring* in the sense of [9]. In such a semiring, there is a canonical *-operation defined by

$$p^* \stackrel{\text{def}}{=} \sum_{i=0}^{\infty} p^i \ .$$

In our setting, this operation becomes

$$p^* = \begin{cases} \frac{1}{1-p} & \text{if } 0 \leq p < 1, \\ \infty & \text{otherwise.} \end{cases}$$

¿From this definition we derive easily that:

1. $p^*(1-p) = 1$, for all $p \in [0,1)$,
2. $p^* = 1 + pp^*$,
3. $(pq)^*p = p(qp)^*$, and
4. $(p^*q)^*p^* = (p+q)^*$.

Note that, for every stochastic vector (p,q) whose components are both positive, term t and variable x, we have that:

$$\mu x.px + qt = \mu x.(p^*q)t \ . \tag{8}$$

(In the above equation, and in the remainder of this appendix, we use the notation introduced in Notation 2 in the main body of the paper. Note, moreover, that this equation remains valid also for $q = 0$, if we agree that $0t = \perp$.) Indeed, using axiom R2 in Table 2, the left-hand side of the above equation reduces to $\mu x.t$, which is equal to the right-hand side because $p^*q = 1$.

Furthermore for all terms t, s and variable x, we define:

$$t \cdot_x s \stackrel{\text{def}}{=} t[s/x] \ .$$

We use $\cdot$ instead of $\cdot_x$ if the meaning is clear from the context. Then we have that the following syntactic equalities hold:

1. $(t \cdot s) \cdot u = t \cdot (s \cdot u)$,
2. $x \cdot s = s \cdot x = s$, and
3. $(t + u) \cdot s = t \cdot s + u \cdot s$.

If we express the fixed point identities (1)–(3) in this formalism we get:

1. **Fixed point identity:** $\mu x.f(x) = f(x) \cdot (\mu x.f(x))$.
2. **Diagonal identity:** $\mu x.f(x,x) = \mu y.\mu x.f(x,y)$.
3. **Rolling identity:** $\mu x.(f(x) \cdot g(x)) = f(x) \cdot \mu x.(g(x) \cdot f(x))$.

Note that all the terms mentioned in the above equations, and in the proofs to follow, may contain any number of parameters apart from the explicitly mentioned recursion variables. We shall henceforth omit these parameters from terms for the sake of readability.

The following result will be used in the proofs of Propositions 7 and 9.

Lemma 2. *Let t be a term containing unguarded occurrences of the variables $x_1, \ldots, x_n$. Then, using axioms S1–S3, R2 and the fixed point equation, t can be written in the form*

$$p_1 x_1 + \cdots + p_n x_n + pu \ ,$$

where $(p_1, \ldots, p_n, p)$ is a stochastic vector, all of the p_i's ($i \in [n]$) are positive, and u is a term in which all of the variables $x_1, \ldots, x_n$ are guarded.

Proposition 7. *The diagonal and the rolling identities are derivable from the implicational axiom system by Stark and Smolka presented in Sect. 3.*

Proof. If the variables only occur guarded in the terms, the result follows from the developments in [6, Chapter 6, Sect. 4]. For the general case we proceed as follows.

Diagonal identity: We want to prove the equation

$$\mu x.g(x,x) = \mu y.\mu x.g(x,y) \ ,$$

where

$$g(x,y) = px + qy + rf(x,y) \ ,$$

(p,q,r) is a stochastic vector, and the variables x and y are guarded in $f(x,y)$. In what follows, we focus on the case in which all of the components of (p,q,r) are positive. All of the remaining cases can be dealt with in similar or simpler fashion.

We begin by noting that

$$\mu x.g(x,x) = \mu x.f(x,x) \ . \tag{9}$$

Indeed, by (8) we have:

$$\mu x.g(x,x) = \mu x.px + qx + rf(x,x) = \mu x.(p+q)^* rf(x,x) = \mu x.f(x,x) \ .$$

As x is guarded in $f(x,x)$, it is therefore sufficient to prove that $\mu y.\mu x.g(x,y)$ is a fixed point of $f(x,x)$ as, by (9) and the unique fixed point rule for guarded terms (axiom R3 in [29]) that would imply that

$$\mu y.\mu x.g(x,y) = \mu x.f(x,x) = \mu x.g(x,x) \ .$$

Thus we aim at proving that

$$f(x,x) \cdot_x (\mu y.\mu x.g(x,y)) = \mu y.\mu x.g(x,y) \ .$$

We have that

$$\begin{aligned}\mu x.g(x,y) &= \mu x.px + qy + rf(x,y)\\ &= \mu x.p^*qy + p^*rf(x,y)\\ &= p^*qy + (p^*rf(x,y) \cdot_x (\mu x.(p^*qy + p^*rf(x,y))))\\ &= p^*qy + (p^*rf(x,y) \cdot_x (\mu x.g(x,y)))\ \ .\end{aligned}$$

It now follows that:

$$\mu y.\mu x.g(x,y) = \mu y.f(x,y) \cdot_x \mu x.g(x,y)\ \ . \tag{10}$$

Indeed, this equality can be proven thus:

$$\begin{aligned}\mu y.\mu x.g(x,y) &= \mu y.p^*qy + (p^*rf(x,y) \cdot_x \mu x.g(x,y))\\ &= \mu y.(p^*q)^*p^*(rf(x,y) \cdot_x \mu x.g(x,y))\\ &= \mu y.(p+q)^*r(f(x,y) \cdot_x \mu x.g(x,y))\\ &= \mu y.f(x,y) \cdot_x \mu x.g(x,y)\ \ .\end{aligned}$$

We note that if ξ and η are fresh variables then

1. $f(x,x) \cdot_x t = (f(\xi,\eta) \cdot_\xi t) \cdot_\eta t$ and
2. $(f(x,y) \cdot_x t) \cdot_y u = (f(\xi,\eta) \cdot_\xi (t \cdot_y u)) \cdot_\eta u$.

These equalities, together with the fixed point equation and (10), imply

$$\begin{aligned}&f(x,x) \cdot_x \mu y.\mu x.g(x,y) =\\ &\quad (f(\xi,\eta) \cdot_\xi [\mu y.\mu x.g(x,y)]) \cdot_\eta \{\mu y.\mu x.g(x,y)\}\\ &= (f(\xi,\eta) \cdot_\xi [\mu x.g(x,y) \cdot_y \mu y.\mu x.g(x,y)]) \cdot_\eta \{\mu y.\mu x.g(x,y)\}\\ &= (f(x,y) \cdot_x [\mu x.g(x,y)]) \cdot_y \{\mu y.\mu x.g(x,y)\}\\ &= (f(x,y) \cdot_x \mu x.g(x,y)) \cdot_y \{\mu y.(f(x,y) \cdot_x \mu x.g(x,y))\}\\ &= \mu y.(f(x,y) \cdot_x \mu x.g(x,y))\\ &= \mu y.\mu x.g(x,y)\end{aligned}$$

as we wanted to prove.

Rolling identity: We want to show that

$$\begin{aligned}&\mu x.[(p_1x + p_2f(x)) \cdot (q_1x + q_2g(x))] =\\ &\quad (p_1x + p_2f(x)) \cdot \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))]\ \ ,\end{aligned} \tag{11}$$

where (p_1, p_2) and (q_1, q_2) are stochastic vectors and the variable x is guarded in both $f(x)$ and $g(x)$. Again, in what follows, we focus on the case in which all components of (p_1, p_2) and (q_1, q_2) are positive. All of the remaining cases can be dealt with in similar or simpler fashion.

The left-hand side of (11) can be rewritten as a fixed point of a guarded term as follows:

$$\begin{aligned}&\mu x.(p_1x + p_2f(x)) \cdot (q_1x + q_2g(x)) =\\ &\mu x.p_1q_1x + p_1q_2g(x) + p_2[f(x) \cdot (q_1x + q_2g(x))] =\\ &\mu x.(p_1q_1)^*(p_1q_2g(x) + p_2[f(x) \cdot (q_1x + q_2g(x))])\ \ .\end{aligned}$$

Therefore it is sufficient to prove that the right-hand side of (11) is a fixed point of the term

$$(p_1q_1)^*(p_1q_2g(x) + p_2f(x) \cdot (q_1x + q_2g(x))) \ .$$

To obtain this we proceed as follows:

$$\begin{array}{l}
(p_1q_1)^*(p_1q_2g(x) + p_2f(x) \cdot (q_1x + q_2g(x))) \\
\qquad\qquad \cdot(p_1x + p_2f(x)) \cdot \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))] = \\
\\
(p_1q_1)^*p_1q_2g(x) \cdot (p_1x + p_2f(x)) \cdot \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))]+ \\
(p_1q_1)^*p_2f(x) \cdot [(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))] \cdot \\
\qquad\qquad \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))] = \\
\\
(p_1q_1)^*p_1q_2g(x) \cdot (p_1x + p_2f(x)) \cdot \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))]+ \\
(p_1q_1)^*p_2f(x) \cdot \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))] = \\
\\
p_1(p_1q_1)^*q_2g(x) \cdot (p_1x + p_2f(x)) \cdot \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))]+ \\
(p_1(p_1q_1)^*q_1 + 1)p_2)f(x) \cdot \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))] = \\
\\
p_1(p_1q_1)^*q_2g(x) \cdot (p_1x + p_2f(x)) \cdot \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))]+ \\
p_1(p_1q_1)^*q_1p_2f(x) \cdot \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))]+ \\
p_2f(x) \cdot \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))] = \\
\\
p_1(p_1q_1)^*\{q_2g(x) \cdot (p_1x + p_2f(x)) + q_1p_2f(x)\} \\
\qquad\qquad \cdot\mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))]+ \\
p_2f(x) \cdot \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))] = \\
\\
p_1[(p_1q_1)^*\{q_1p_2f(x) + q_2g(x) \cdot (p_1x + p_2f(x))\} \\
\qquad\qquad \cdot\mu x.[(p_1q_1)^*\{q_1p_2f(x) + q_2g(x) \cdot (p_1x + p_2f(x))\}+ \\
p_2f(x) \cdot \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))] = \\
\\
p_1\mu x.[(p_1q_1)^*\{q_1p_2f(x) + q_2g(x) \cdot (p_1x + p_2f(x))\}]+ \\
p_2f(x) \cdot \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))] = \\
\\
p_1\mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))]+ \\
p_2f(x) \cdot \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))] = \\
\\
(p_1x + p_2f(x)) \cdot \mu x.[(q_1x + q_2g(x)) \cdot (p_1x + p_2f(x))]
\end{array}$$

which was to be shown. □

We are now left to prove that the group equations mentioned in Sect. 2.2 are provable from the axiom system proposed by Stark and Smolka in [29]. In fact, since the group equations are a special case of the commutative identity, which is entailed by the weak functorial implication, we shall show that the weak functorial implication is derivable from the implicational axiom system by Stark and

Smolka presented in Sect. 3. In the proof of this result, it will be convenient to have the generalization of (8) presented in the following proposition. In stating this generalization of (8), we use the fact that the semiring of all matrices whose entries are either nonnegative reals or ∞ is also complete, and thus comes equipped with its *-operation.

Proposition 8. *Let $\overline{x}$ be n-dimensional vector of distinct variables, and $\overline{t}$ be an n-dimensional vector of terms. Assume that $[A, E]$ is a stochastic matrix (i.e., a matrix all of whose rows are stochastic vectors), where A is an $n \times n$ matrix, and E is an n-dimensional vector. Then the implicational axiom system by Stark and Smolka presented in Sect. 3 proves that*

$$\mu\overline{x}.A\overline{x} + E\overline{t} = \mu\overline{x}.(A^*E)\overline{t} \ . \tag{12}$$

Recall from Sect. 2.2 that the weak functorial implication can be stated as follows:

> For terms $t_i(x_1, \ldots, x_n, \overline{y})$ $(i \in [n])$ and $t(x, \overline{y})$, if $t_i(x, \ldots, x, \overline{y}) = t(x, \overline{y})$ for every $i \in [n]$, then
>
> $$(\mu(x_1, \ldots, x_n).(t_1, \ldots, t_n))_1 = \mu x.t(x, \overline{y}) \ .$$

Proposition 9. *The weak functorial implication is derivable from the implicational axiom system by Stark and Smolka presented in Sect. 3.*

Proof. Throughout the proof, we shall write ρ for the unique mapping from $[n]$ to $[1]$, and $t^\dagger$ for $\mu x.t$. Following Defn. 4, we also use the following abbreviations:

$$\begin{aligned} \overline{t} \circ \rho &\stackrel{\text{def}}{=} (t_1(x, \ldots, x, \overline{y}), \ldots, t_n(x, \ldots, x, \overline{y})) \qquad \text{and} \\ \rho \circ t &\stackrel{\text{def}}{=} \underbrace{(t(x, \overline{y}), \ldots, t(x, \overline{y}))}_{n\text{-times}} \ . \end{aligned}$$

Using these notations, assume that

$$\overline{t} \circ \rho = \rho \circ t \ . \tag{13}$$

We aim at proving that

$$\mu\overline{x}.\overline{t} = \rho \circ t^\dagger$$

is derivable from the implicational axiom system by Stark and Smolka presented in Sect. 3.

First of all, note that, if t is guarded, then so is $\overline{t}$, and the implication follows from the unique fixed point induction rule for guarded terms. Indeed, we have that

$$\begin{aligned} \overline{t}[\underbrace{(t^\dagger, \ldots, t^\dagger)}_{n\text{-times}}/(x_1, \ldots, x_n)] &= (\overline{t} \circ \rho)[t^\dagger/x] \\ &= (\rho \circ t)[t^\dagger/x] \qquad \text{(By (13))} \\ &= \rho \circ (t[t^\dagger/x]) \\ &= \rho \circ t^\dagger \qquad \text{(By the fixed point equation)} \ . \end{aligned}$$

Thus, $\rho \circ t^\dagger$ is a fixed point of the vector of guarded terms $\overline{t}$, and the unique fixed point induction rule for guarded term vectors yields the claim (see, e.g., [29, Theorem 1]).

Suppose now that t is not guarded. Write

$$\overline{t} = A\overline{x} + E\overline{u} \ ,$$

where $[A, E]$ is an $n \times (n+1)$ stochastic matrix, and each of the terms in the vector $\overline{u}$ is guarded. Then, by (13), we have that

$$t = px + qs \ ,$$

where p is the sum of the entries of each row of A. Moreover, each component of E is equal to $q = (1-p)$. Note that, since x is not guarded in t, the real number p is positive. In what follows we focus on the case in which q is also positive. (The case in which q is zero is entailed by (12) taking E to be a vector whose entries are all 0.) In this case, we have that

$$\begin{aligned} A \circ \rho &= \rho \circ p \qquad \text{and} \\ E \circ \rho &= \rho \circ q \ . \end{aligned}$$

Moreover, by (13),

$$\overline{u} \circ \rho = \rho \circ s \ . \tag{14}$$

Using (8) and (12), it now follows that:

$$\begin{aligned} \mu\overline{x}.\overline{t} &= \mu\overline{x}.(A^*E)\overline{u} \qquad \text{and} \\ \mu x.t &= \mu x.(p^*q)s \ . \end{aligned}$$

Note now that $s = (p^*q)s$ is guarded, and so is each component of $(A^*E)\overline{u}$. Thus, by the previous analysis for the guarded case, we are done if we can show that

$$((A^*E)\overline{u}) \circ \rho = \rho \circ ((p^*q)s) \ .$$

The above equality, however, can be proven as follows:

$$\begin{aligned} ((A^*E)\overline{u}) \circ \rho &= (A^*E)(\overline{u} \circ \rho) \\ &= (A^*E)(\rho \circ s) \qquad \text{(By (14))} \\ &= A^*(\rho \circ (qs)) \\ &= \rho \circ ((p^*q)s) \ . \end{aligned}$$

The last equality in the above derivation follows from the fact that, since each row of A has sum p,

$$A^*\rho = (\sum_{i=0}^{\infty} A^i)\rho = \sum_{i=0}^{\infty}(A^i\rho) = \sum_{i=0}^{\infty}(\rho \circ p^i) = \rho \circ (\sum_{i=0}^{\infty} p^i) = \rho \circ p^* \ . \ \square$$

Recent BRICS Report Series Publications

RS-02-6 **Luca Aceto, Zoltán Ésik, and Anna Ingólfsdóttir.** ***Equational Axioms for Probabilistic Bisimilarity (Preliminary Report)*****. February 2002. 22 pp.**

RS-02-5 **Federico Crazzolara and Glynn Winskel.** ***Composing Strand Spaces*****. February 2002. 30 pp.**

RS-02-4 **Olivier Danvy and Lasse R. Nielsen.** ***Syntactic Theories in Practice*****. January 2002. 34 pp. This revised report supersedes the earlier BRICS report RS-01-31.**

RS-02-3 **Olivier Danvy and Lasse R. Nielsen.** ***On One-Pass CPS Transformations*****. January 2002. 18 pp.**

RS-02-2 **Lasse R. Nielsen.** ***A Simple Correctness Proof of the Direct-Style Transformation*****. January 2002.**

RS-02-1 **Claus Brabrand, Anders Møller, and Michael I. Schwartzbach.** ***The <bigwig> Project*****. January 2002. 36 pp. This revised report supersedes the earlier BRICS report RS-00-42.**

RS-01-55 **Daniel Damian and Olivier Danvy.** ***A Simple CPS Transformation of Control-Flow Information*****. December 2001.**

RS-01-54 **Daniel Damian and Olivier Danvy.** ***Syntactic Accidents in Program Analysis: On the Impact of the CPS Transformation*****. December 2001. 41 pp. To appear in the** ***Journal of Functional Programming*****. This report supersedes the earlier BRICS report RS-00-15.**

RS-01-53 **Zoltán Ésik and Masami Ito.** ***Temporal Logic with Cyclic Counting and the Degree of Aperiodicity of Finite Automata*****. December 2001. 31 pp.**

RS-01-52 **Jens Groth.** ***Extracting Witnesses from Proofs of Knowledge in the Random Oracle Model*****. December 2001. 23 pp.**

RS-01-51 **Ulrich Kohlenbach.** ***On Weak Markov's Principle*****. December 2001. 10 pp.**

RS-01-50 **Jiří Srba.** ***Note on the Tableau Technique for Commutative Transition Systems*****. December 2001. 19 pp. To appear in Nielsen and Engberg, editors,** ***Foundations of Software Science and Computation Structures*****, FoSSaCS '02 Proceedings, LNCS 2303, 2002.**