Confluence and convergence modulo equivalence in probabilistically terminating reduction systems

Kirkeby, Maja Hanne; Christiansen, Henning

Published in:
International Journal of Approximate Reasoning

DOI:
10.1016/j.ijar.2018.11.018

Publication date:
2019

Document Version
Early version, also known as pre-print

Citation for published version (APA):
Kirkeby, M. H., & Christiansen, H. (2019). Confluence and convergence modulo equivalence in probabilistically terminating reduction systems. *International Journal of Approximate Reasoning*, *105*, 217-228. https://doi.org/10.1016/j.ijar.2018.11.018

Confluence and Convergence Modulo Equivalence in Probabilistically Terminating Reduction Systems☆

Maja H. Kirkeby and Henning Christiansen

Roskilde University, Denmark
majaht@ruc.dk and henning@ruc.dk

Abstract

Convergence of an abstract reduction system is the property that the possible derivations from a given initial state all end in the same final state. Relaxing this by "modulo equivalence" means that these final states need not be identical, only equivalent wrt. a specified equivalence relation.

We generalize this notion for probabilistic abstract reduction systems, naming it almost-sure convergence modulo equivalence, such that the final states are reached with probability 1. We relate it to the well-studied properties of almost-sure termination and confluence/convergence of probabilistic and non-probabilistic systems. In addition, we provide a transformational approach for proving – or disproving – almost-sure convergence modulo equivalence of given systems.

Keywords: almost-sure convergence modulo equivalence, almost-sure termination, probabilistic abstract reduction systems, abstract reduction systems, confluence modulo equivalence
2010 MSC: 00-01, 99-00

1. Introduction

Abstract reduction systems (ARS) are ubiquitous in computer science and logic, as theoretical models of computational systems that develop over time in discrete steps. In their probabilistic version (PARS), the choice of successor state is governed by a probability distribution, which in turn induces a global, probabilistic behaviour of the system. We consider an important intersection of two classes of PARS, namely those having the properties of

- almost-sure termination: even if infinite reduction sequences exists, the probability of termination is 1;

☆This work is supported by The Danish Council for Independent Research, Natural Sciences, grant no. DFF 4181-00442.

- convergence: any reduction sequence from a given state must end in the same final state.

Convergence generalizes to almost-sure convergence: such a final state is reached with probability 1 (although there may be diverging reduction sequences).

In this paper, we generalize previous approaches to characterize and prove almost-sure convergence to take into account an equivalence on states. We introduce the notion of *almost-sure convergent modulo an equivalence relation* meaning that all equivalent initial states reduce to equivalent final states, and for each such initial state, the total probability of these reductions is 1.

This generalization opens up for studying a much larger class of systems with alternative final states which are "semantically" the same, but differs syntactically. One example is redundant data representations in which, e.g., the order of elements in a list are considered immaterial, or "the same" tree may be rebalanced. Other examples are systems that search for one of perhaps many acceptable solutions to a problem.

Related work

Abstract Reduction Systems and studies of their properties have emerged from a variety of areas, with the work on the lambda calculus by Church and Rosser [1] in the 1930s as a noticeable example.

The notion of confluence has been extensively studied for ARS, see, e.g., [2, 3] for overview. Confluence means that whenever alternative repeated reductions are possible from some state, these can be extended to join in a common state. While the focus in the present paper is on convergence (modulo equivalence), confluence (modulo equivalence) plays an important role since convergence can be defined as the combination of confluence and termination.

Newman's lemma [4] from 1942 is a central result for ARS: a terminating system is confluent if and only if it satisfies a simpler property of local confluence. The property of *confluence modulo equivalence* was introduced in 1972 by Aho et al [5]; a system is confluent modulo equivalence whenever every pair of equivalent states with alternative reduction sequences join in equivalent states. In 1980, Huét [3] generalized Newman's lemma for confluence modulo equivalence. It is well-known that these results do not generalize to non-terminating systems (and thus neither to almost-sure terminating ones); see, e.g., [3].

In 1991, Curien and Ghelli [6] described a powerful method for proving confluence of normalizing non-probabilistic systems, using suitable transformations from the original system into one, known to be confluent. This was extended by Kirkeby and Christiansen [7] that also allows disproving confluence using transformations into a non-confluent system.

In term rewriting [2], proving local confluence may be reduced to a finite number of cases, described by *critical pairs* (for a definition, see the reference), which may be checked automatically. Similar results have been shown for a subset of the programming language Constraint Handling Rules, CHR [8], in the 1990s by Abdennadher et al [9, 10], assuming a theoretical, logic-based semantics. Recent work by Christiansen and Kirkeby [11, 12] generalize these results to confluence modulo equivalence under a different semantics that reflects the implemented CHR systems (taking into account runtime errors and non-logical predicates).

Almost-sure convergence and almost-sure termination were introduced in an early 1983 paper [13] by Hart et al for a specific class of probabilistic programs with finite state spaces. Almost-sure convergence was introduced in the context of PARS by Kirkeby and Christiansen [7], and shown equivalent to almost-sure termination and confluence. The related notion of almost-sure confluence was introduced concurrently by Frühwirth et al. [14] – in the context of a probabilistic version of CHR – and by Bournez and Kirchner [15] in more generality for PARS. They also formulated the simple concept of PARS and almost-sure termination in that context. Bournez and Garnier [16] generalized PARS to nondeterministic PARS, introducing a nondeterministic choice of transition probabilities, leading to methods for proving almost-sure termination [16, 17, 18], and studies of uniqueness of limit distributions [19, 20].

Contributions

We broaden previous work on PARS by the introduction of state equivalence and investigate its consequences for convergence. More specifically, we propose the novel notion of *almost-sure convergence modulo equivalence* and provide results that may be used for proving or disproving this for given systems. In a previous paper [7], we studied almost-sure convergence (without equivalence) and gave precise and self-contained definitions for PARS with proofs of their basic results, which had been lacking in the literature so far. For completeness, we repeat some results of [7] together with their proofs, i.e., Propositions 5, 6, Lemmas 10, 12, Theorem 25, while the central results in the present paper are novel generalizations for modulo equivalence. This includes a suitable generalization of the result by Curien and Ghelli [6] explained above, and we demonstrate its application in relation to almost-sure convergence modulo equivalence.

Overview of this paper

In Section 2, we review definitions for abstract reduction systems and introduce and motivate our choices of definitions for their probabilistic counterparts; a proof that the defined probabilities indeed constitute a probability distribution is found in the Appendix. Section 3 formulates and proves important properties, relevant for showing almost-sure convergence modulo equivalence. Section 4 goes in detail with applications of the transformational approach [6] to (dis-) proving almost-sure convergence modulo equivalence, and in Section 5 we demonstrate the use of this approach for a naive probabilistic sorting algorithm for non-repeating lists and for repeating lists. Finally, we provide a summary and suggestions for future work in Section 6.

2. Basic definitions

The definitions for non-probabilistic systems are standard; see, e.g., [2, 3]. Basic definitions and properties for probabilistic abstract reduction systems are rephrased from [7].

Definition 1 (ARS). *An* Abstract Reduction System *is a pair* $R = (A, \rightarrow)$ *where the* reduction $\rightarrow$ *is a binary relation on a countable set* A*. An* equivalence *relation* $\approx$ *is a binary relation* $\approx \subseteq A \times A$ *that is reflexive, transitive and symmetric.*

Instead of $(s,t) \in \rightarrow$, we write $s \rightarrow t$ (or $t \leftarrow s$ when convenient). The relation $\leftrightarrow$ refers to $(\leftarrow \cup \rightarrow)$, $\stackrel{\approx}{\rightarrow}$ to $(\rightarrow \cup \approx)$, and $\stackrel{\approx}{\leftrightarrow}$ to $(\leftrightarrow \cup \approx)$. The notation $s \rightarrow^* t$ denotes the transitive reflexive closure of $\rightarrow$, and analogously for the other relations mentioned.

In the literature, an ARS is often required to have only finite branching. i.e., for any element s, the set $\{t \mid s \rightarrow t\}$ is finite. We do not require this, as the implicit restriction to countable branching is sufficient for our purposes.

The set of *normal forms* R_{NF} are those $s \in A$ for which there is no $t \in A$ such that $s \rightarrow t$. For given element s, the *normal forms of* s, are defined as the set $R_{NF}(s) = \{t \in R_{NF} \mid s \rightarrow^* t\}$. An element which is not a normal form is said to be *reducible*; i.e., an element s is reducible if and only if $\{s' \mid s \rightarrow s'\} \neq \emptyset$.

A *path* from an element s is a (finite or infinite) sequence of reductions $s \rightarrow s_1 \rightarrow s_2 \rightarrow \cdots$; a finite path $s \rightarrow s_1 \rightarrow s_2 \rightarrow \cdots \rightarrow s_n$ has *length* n ($n \geq 0$); in particular, we recognize an empty path (of length 0) from a given state to itself. For given elements s and $t \in R_{NF}(s)$, $\Delta(s,t)$ denotes the set of finite paths $s \rightarrow \cdots \rightarrow t$ (including the empty path); $\Delta^\infty(s)$ denotes the set of infinite paths from s. A system is

- *confluent* if for all $s_1 \leftarrow^* s \rightarrow^* s_2$ there is a t such that $s_1 \rightarrow^* t \leftarrow^* s_2$,
- *confluent modulo* $\approx$ if for all $s_1' \leftarrow^* s_1 \approx s_2 \rightarrow^* s_2'$ there exist t_1 and t_2 such that $s_1' \rightarrow^* t_1 \approx t_2 \leftarrow^* s_2'$,
- *terminating*[1] iff it has no infinite path,
- *convergent* iff it is terminating and confluent, and
- *normalizing*[2] iff every element s has a normal form, i.e., there is an element $t \in R_{NF}$ such that $s \rightarrow^* t$.

The following property indicates the complexity of the probability measures that are needed in order to cope with paths in probabilistic abstract reduction systems defined over countable sets.

Proposition 2. *Given an ARS as above and given elements* s *and* $t \in R_{NF}(s)$*, it holds that* $\Delta(s,t)$ *is countable, and* $\Delta^\infty(s)$ *may or may not be countable.*

---

[1]A terminating system is also called *strongly normalizing* elsewhere, e.g., [6].

[2]A normalizing system is also called *weakly normalizing* or *weakly terminating* elsewhere, e.g., [6].

Proof. For the first part, $\Delta(s,t)$ is isomorphic to a subset of $\bigcup_{n=1,2,\ldots} A^n$. A countable union of countable sets is countable, so $\Delta(s,t)$ is countable.

For the second part, consider the ARS $\langle \{0,1\}, \{i \to j \mid i,j \in \{0,1\}\}\rangle$. Each infinite path can be read as a real number in the unit interval, and any such real number can be described by an infinite path. The real numbers (in the unit interval) are uncountable.

This means that we can define discrete and summable probabilities over $\Delta(s,t)$, and – which we will avoid – considering probabilities over the space $\Delta^\infty(s)$ requires more advanced measure theory.

In the next definition, a path is considered a Markov process/chain, i.e., each reduction step is independent of the previous ones, and thus the probability of a path is defined as a product in the usual way.

Definition 3 (PARS). *A* Probabilistic Abstract Reduction System *is a pair $R^P = (R,P)$ where $R = (A, \to)$ is an ARS, and for each reducible element $s \in A \setminus R_{NF}$, $P(s \to \cdot)$ is a probability distribution over the reductions from s, i.e., $\sum_{s\to t} P(s \to t) = 1$; it is assumed, that for all s and t, $P(s \to t) > 0$ if and only if $s \to t$.*

The probability of a finite path $s_0 \to s_1 \to \ldots \to s_n$ *with $n \geq 0$ is given as*

$$P(s_0 \to s_1 \to \ldots \to s_n) = \prod_{i=1}^{n} P(s_{i-1} \to s_i).$$

For any element s and normal form $t \in R_{NF}(s)$, the probability of s reaching t*, written $P(s \to^* t)$, is defined as*

$$P(s \to^* t) = \sum_{\delta \in \Delta(s,t)} P(\delta);$$

the probability of s not reaching a normal form *(or* diverging*) is defined as*

$$P(s \to^\infty) = 1 - \sum_{t \in R_{NF}(s)} P(s \to^* t).$$

When referring to confluence, local confluence, termination*, and* normalization *of a PARS, we refer to these properties for the underlying ARS.*

Our definition of PARS is similar to that of [15] except that they allow transitions with zero probabilities. This gives no difference in the derived probabilities but the absence of zero probability transitions simplifies the connection between a PARS and its underlying ARS.

Notice that when s is a normal form then $P(s \to^* s) = 1$ since $\Delta(s,t)$ contains only the empty path with probability $\prod_{i=1}^{0} P(s_{i-1} \to s_i) = 1$. It is important that $P(s \to^* t)$ is defined only when t is a normal form of s since otherwise, the defining sum may be ≥ 1, as demonstrated by the following example.

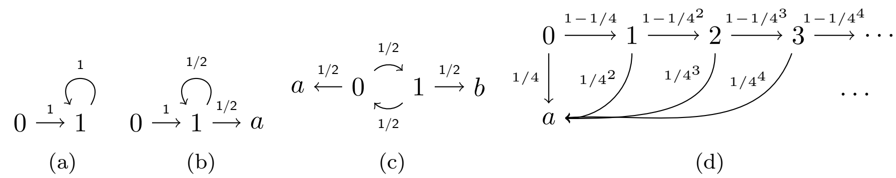


Figure 1: PARS with different properties, see Table 1.

Example 4. *Consider the PARS R^P given in Figure 1(a); formally, $R^P = ((\{0,1\}, \{0 \to 1, 1 \to 1\}),\ P)$ with $P(0 \to 1) = 1$ and $P(1 \to 1) = 1$. An attempt to define $P(0 \to^* 1)$ as in Def. 3, for the reducible element 1, does not lead to a probability, i.e., $P(0 \to^* 1) \not\leq 1$: $P(0 \to^* 1) = P(0 \to 1) + P(0 \to 1 \to 1) + P(0 \to 1 \to 1 \to 1) + \ldots = \infty$.*

The authors of [15] use a different (and more complicated) probability measure that allows to define the probability of visiting also non-normal forms, which seems necessary for defining a notion of almost-sure confluence. Our definitions are sufficient for characterizing and proving almost-surely convergence modulo equivalence.

The following proposition shows that P is indeed a probability distribution.

Proposition 5. *For an arbitrary finite path π, $1 \geq P(\pi) > 0$. For every element s, $P(s \to^* \cdot)$ and $P(s \to^\infty)$ comprise a probability distribution, i.e., $\forall t \in R_{NF}(s)$: $0 \leq P(s \to^* t) \leq 1$; $0 \leq P(s \to^\infty) \leq 1$; and $\sum_{t \in R_{NF}(s)} P(s \to^* t) + P(s \to^\infty) = 1$.*

Proof. The proofs are simple but lengthy and are given in the Appendix.

Proposition 6 justifies that we refer to $P(s \to^\infty)$ as a probability of divergence.

Proposition 6. *Consider a PARS which has an element s for which $\Delta^\infty(s)$ is countable (finite or infinite). Let $P(s_1 \to s_2 \to \cdots) = \prod_{i=1,2,\ldots} P(s_i \to s_{i+1})$ be the probability of an infinite path, then $P(s \to^\infty) = \sum_{\delta \in \Delta^\infty(s)} P(\delta)$.*

Proof. See Appendix.

We can now define *probabilistic* and *almost-sure* (abbreviated "a-s.") versions of important notions for derivation systems. A system is

- *almost-surely convergent* if for all $s_1 \leftarrow^* s \to^* s_2$ there is a normal form $t \in R_{NF}$ such that $s_1 \to^* t \leftarrow^* s_2$ and $P(s_1 \to^* t) = P(s_2 \to^* t) = 1$,
- *almost-surely convergent modulo* $\approx$ if for all $s_1' \leftarrow^* s_1 \approx s_2 \to^* s_2'$ there are two normal forms $t_1, t_2 \in R_{NF}$ such that $s_1 \to^* t_1 \approx t_2 \leftarrow^* s_2$ and $\sum_{t_1 \in R_{NF}(s_1')} P(s_1' \to^* t_1) = \sum_{t_2 \in R_{NF}(s_2')} P(s_2' \to^* t_2) = 1$,

	(a)	(b)	(c)	(d)	(d′)
Confl.	+	+	−	+	+
Term.	−	−	−	−	−
A-s. conv.	−	+	−	−	+
A-s. term.	−	+	+	−	+

Table 1: A property overview of the systems (a)–(d) in Figure 1 and (d′) with same ARS as (d), but with all probabilities replaced by 1/2.

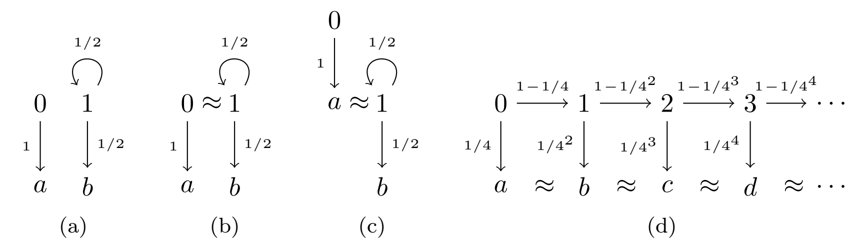


Figure 2: PARS with different "modulo ≈"-properties, see Table 2.

- *almost-surely terminating*[3] iff every element s has $P(s \to^{\infty}) = 0$, and
- *probabilistically normalizing* iff every element s has a normal form t such that $P(s \to^{*} t) > 0$.

Example 7. *The four probabilistic systems in Figure 1 demonstrate the properties of confluence, termination, a-s. termination and a-s. convergence. We notice that (b)–(d) are normalizing. Furthermore, they are all non-terminating: system (b) and (c) are a-s. terminating, which is neither the case for (a) nor (d); for element* 0 *in system (d) we have* $P(0 \to^{\infty}) = \prod_{i=1}^{\infty}(1-(1/4)^i) \approx 0.6885 > 0$.[4] *Table 1 summarizes their properties of (almost-sure) (local) confluence; (d′) refers to a PARS with the same underlying ARS as (d) and with all probabilities* $= 1/2$*. The difference between system (d) and (d′) emphasizes that the choice of probabilities do matter for whether or not different probabilistic properties hold. For any element* s *in (d′), the probability of reaching the normal form* a *is* $1/2 + 1/2^2 + 1/2^3 + \cdots = 1$.

Example 8. *The four probabilistic systems in Figure 2 demonstrate the properties confluence modulo* $\approx$*, (non)termination, a-s. termination and a-s. convergence modulo* $\approx$*. We notice that all systems are normalizing and non-terminating and that systems (a)–(c) are a-s. terminating, which is not the*

[3] Almost-sure termination is named *probabilistic termination* elsewhere, e.g., [21, 14].

[4] Verified by Mathematica. The exact result is $\left(\frac{1}{4};\frac{1}{4}\right)_{\infty}$; see [22] for the definition of this notation.

Figure 2:	(a)	(b)	(c)	(d)	(d′)
Confl. mod. $\approx$	+	−	−	+	+
Term.	−	−	−	−	−
A-s. conv. mod. $\approx$	+	−	−	−	+
A-s. term.	+	+	+	−	+

Table 2: A property overview of the systems (a)–(d) in Figure 2 and (d′) with same ARS as (d), but with all probabilities replaced by 1/2.

case for (d). The system (a) has "=" as $\approx$; (a) together with (b) demonstrates that different equivalence relations influence confluence modulo $\approx$ and a-s. convergence modulo $\approx$. Since $a \approx 1 \to b$ is in system (c) and $a \not\approx b$, it is not confluent modulo $\approx$. Table 2 summarizes their properties of (almost-sure) termination, confluence modulo $\approx$ and a-s. confluence modulo $\approx$; (d′) refers to a PARS with the same underlying ARS as (d) and with all probabilities $= 1/2$ and emphasizes that the choice of probabilities do matter. For any element s in (d′), the probability of reaching the normal form a is $1/2 + 1/2^2 + 1/2^3 + \cdots = 1$.

3. Properties of Probabilistic Abstract Reduction Systems

With a focus on almost-sure convergence modulo equivalence, we consider now relevant relationships between the properties of probabilistic and their underlying non-probabilistic systems. Lemmas 10 and 12, below, have been suggested by [15] without proofs. The most important properties are summarized as follows. For any PARS R^P:

- R^P is normalizing if and only if it is probabilistically normalizing (Lemma 10),
- if R^P is almost-surely terminating then it is normalizing (Lemma 11),
- if R^P is terminating then it is almost-surely terminating (Lemma 12),
- R^P is almost-surely terminating and confluent modulo $\approx$, if and only if it is almost-surely convergent modulo $\approx$ (Theorem 18),
- R^P is almost-surely terminating and confluent, if and only if it is almost-surely convergent (Corollary 19).

The following inductive characterization of the probabilities for reaching a given normal form is useful for the proofs that follow.

Proposition 9. *For any reducible element s, the following holds.*

$$\sum_{t \in R_{NF}} P(s \to^* t) = \sum_{s \to s'} \left(P(s \to s') \times \sum_{t \in R_{NF}} P(s' \to^* t) \right)$$

Proof. Any path from s to a normal form t will have the form $s \to s' \to \cdots \to t$, for some direct successor s' of s. The other way round, any normal form for a direct successor s' of s will also be a normal form of s. With this observation, the proposition follows directly from Definition 3 (prob. of path).

Lemma 10 ([15]). *A PARS is normalizing if and only if it is probabilistically normalizing.*

Proof. Every element s in a normalizing PARS has a normal form t such that $s \to^* t$ and by definition of PARS, $P(s \to^* t) > 0$, which makes it probabilistically normalizing. The other way round, the definition of probabilistic normalizing includes normalization.

Prob. normalization differs from the other properties by nature (requiring probability > 0 instead of $= 1$), and is the only one which is equivalent to its non-probabilistic counterpart. Thus, the existing results on proving and disproving normalization can be used directly to determine probabilistic normalization. The following lemma is a consequence of [15, Prop. 7.3,7.5].

Lemma 11. *If a PARS is almost-surely terminating then it is normalizing.*

Proof. For every element s in an a-s. terminating system, Proposition 5 gives that $\sum_{t \in R_{NF}} P(s \to^* t) = 1$, and hence s has at least one normal form t such that $P(s \to^* t) > 0$. By Lemma 10, the system is also normalizing.

The opposite is not the case, as demonstrated by system (d) in Figure 1; every element has a normal form, but the system is not almost-surely terminating.

Lemma 12 ([15]). *If a PARS is terminating then it is almost-surely terminating.*

Proof. In a terminating PARS, $\Delta^\infty(s) = \emptyset$ for any element s. By Proposition 6 we have $P(s \to^\infty) = 0$.

The opposite is not the case, as demonstrated by systems (b)–(d) in Figure 1.

The following lemma provides the first insight into the implications of almost-surely convergence modulo equivalence.

Lemma 13. *If a PARS is almost-surely convergent modulo $\approx$ then it is almost-surely terminating .*

Proof. Let $R^P = ((A, \to), P)$ be a-s. convergent modulo $\approx$, and s an arbitrary element of A. We may write $s \leftarrow^* s \approx s \to^* s$ and by definition of a-s. convergent modulo $\approx$, $\sum_{tR_{NF}(s)} P(s \to^* t) = 1$. A-s. termination follows from Proposition 5.

Lemma 14. *In a system $R = (A, \to)$ that is confluent modulo $\approx$, it holds for any $s \in A$ and $t_1, t_2 \in R_{NF}(s)$, that $t_1 \approx t_2$.*

Proof. Assume the notion of the lemma. We may write $t_1 \leftarrow^* s \approx s \rightarrow^* t_2$ and since t_1 and t_2 are normal forms, confluence modulo $\approx$ yields $t_1 \approx t_2$.

This immediately gives the following.

Proposition 15. *In a normalizing system $R = (A, \rightarrow)$ that is confluent modulo $\approx$, it holds for any $s \in A$, that $R_{NF}(s) \neq \emptyset$ and $t_1, t_2 \in R_{NF}(s) \Rightarrow t_1 \approx t_2$.*

The opposite is not the case, as demonstrated by the underlying ARS of (b) in Figure 2. However, if $\approx$ is reduced to "=", a normalizing system is confluent if only if every element has a unique normal form [7].

Lemma 16. *In a system $R^P = ((A, \rightarrow), P)$ that is almost-surely convergent modulo $\approx$, it holds for any $s \in A$, that $R_{NF}(s) \neq \emptyset$ and $t_1, t_2 \in R_{NF}(s) \Rightarrow t_1 \approx t_2$.*

Proof. This is a direct consequence of Lemmas 11 and 13, yielding that $R^P$is normalizing, and Proposition 15.

The opposite is not the case, as demonstrated by system (b) in Figure 2. Huét presented a lemma concerning confluence modulo equivalence and equivalence of normal forms for normalizing systems; this lemma is relevant later in Lemma 24.

Proposition 17 ([3, Lem. 2.6]). *A normalizing system is confluent modulo $\approx$ if and only if*

$$s_1' \leftarrow^* s_1(\leftrightarrow \cup \approx)^* s_2 \rightarrow^* s_2' \quad \Rightarrow \quad \forall t_1 \in R_{NF}(s_1'), \forall t_2 \in R_{NF}(s_2')\ t_1 \approx t_2.$$

The following theorem is essential when proving almost-sure convergence modulo equivalence.

Theorem 18. *A PARS is almost-surely terminating and confluent modulo $\approx$ if and only if it is almost-surely convergent modulo $\approx$.*

Corollary 19. *A PARS is almost-surely terminating and confluent if and only if it is almost-surely convergent.*

Thus, to prove almost-sure convergence of a given PARS, one may use the methods of [16, 17, 18] to prove almost-sure termination and prove classical confluence – referring to Newman's lemma (cf. our discussion in the Introduction), or using the method of mapping the system into another system, already known to be confluent, as described in Section 4, below.

Proof (Theorem 18). We split the proof into smaller parts, referring to properties that are shown below: "if": by Lemmas 13 and 20. "only if": by Lemma 21.

Lemma 20. *If a PARS is almost-surely convergent modulo $\approx$ then it is confluent modulo $\approx$*

Proof. Assume almost-sure convergence modulo $\approx$, then for each $s_1 \leftarrow^* s \rightarrow^* s_2$ there exist two t_1, t_2 (a normal form) such that $s_1 \rightarrow^* t_1 \approx t_2 \leftarrow^* s_2$.

Lemma 21. *If a PARS is almost-surely terminating and confluent modulo $\approx$ then it is almost-surely convergent modulo $\approx$.*

Proof. Let $R^P = ((A, \rightarrow), P)$ be a-s. terminating and confluent modulo $\approx$ and let $s_1' \leftarrow^* s_1 \approx s_2 \rightarrow^* s_2'$. According to Lemma 11, $R^P$is normalizing, and consequently there are normal forms t_1 and t_2 such that $t_1 \leftarrow^* s_1' \leftarrow^* s_1 \approx s_2 \rightarrow^* s_2' \rightarrow^* t_2$, obtaining $t_1 \leftarrow^* s_1 \approx s_2 \rightarrow^* t_2$. Confluence modulo $\approx$ implies $t_1 \approx t_2$ (as t_1 and t_2 are normal forms). A-s. terminating gives, via Proposition 5, that $\sum_{t \in R_{NF}(s_i')} P(s_i' \rightarrow^* t) = 1$ for $i \in \{1, 2\}$.

4. Showing Almost-sure Convergence Modulo $\approx$ by Transformation

Theorem 18 shows that a PARS cannot be almost-surely convergent modulo $\approx$ if it is not almost-surely terminating. To show almost-sure convergence modulo $\approx$ of an almost-surely terminating system, it needs to be shown confluent modulo equivalence. The following proposition is a weaker formulation and consequence of Theorem 18; it shows that (dis)proving confluence modulo $\approx$ for almost-surely terminating systems is crucial when (dis)proving almost-sure convergence modulo $\approx$.

Proposition 22. *An almost-surely terminating PARS is almost-surely convergent modulo $\approx$ if and only if it is confluent modulo $\approx$.*

Proof. This is a direct consequence of Theorem 18.

Curien and Ghelli [6] presented a method for proving confluence by transforming[5] the system of interest (under some restrictions) to a new system which is known to be confluent. We start by repeating their relevant result.

Lemma 23 ([6]). *Given two ARS $R = (A, \rightarrow_R)$ and $R' = (A', \rightarrow_{R'})$ and a mapping $G\colon A \rightarrow A'$, then R is confluent if the following holds.*

(C1) R' is confluent,

(C2) R is normalizing,

(C3) if $s \rightarrow_R t$ then $G(s) \leftrightarrow^*_{R'} G(t)$,

(C4) $\forall t \in R_{NF},\ G(t) \in R'_{NF}$, and

(C5) $\forall t, u \in R_{NF},\ G(t) = G(u) \Rightarrow t = u$

[5]This is also referred to as interpreting a system elsewhere, e.g., [6].

An extended method for proving and disproving confluence was presented by Kirkeby and Christiansen [7]; in that method a system R is confluent if and only if its transformed system R' is confluent. Both these results follow from the subsequent lemma.

Lemma 24. *Given two ARS $R = (A, \to_R)$ and $R' = (A', \to_{R'})$, an equivalence $\approx \subseteq A \times A$ and a mapping $G\colon A \to A'$, satisfying*

(C1′) (surjective) $\forall s' \in A',\ \exists s \in A, G(s) = s'$,

(C2′) R and R' are normalizing,

(C3′) if $s \overset{\approx}{\to}_R t$ then $G(s) \leftrightarrow^*_{R'} G(t)$, and
if $G(s) \leftrightarrow^*_{R'} G(t)$ then $s \overset{\approx}{\leftrightarrow}^*_R t$,

(C4′) $\forall t \in R_{NF},\ G(t) \in R'_{NF}$, and
$\forall t' \in R'_{NF},\ G^{-1}(t') \subseteq R_{NF}$, where $G^{-1}(t') = \{t \in A \mid G(t) = t'\}$

(C5′) $\forall t, u \in R_{NF},\ G(t) = G(u) \Leftrightarrow t \approx u$,

then R is confluent modulo $\approx$ iff R' is confluent.

Proof. "$\Rightarrow$": Assume that R is confluent modulo $\approx$ and R' is not confluent, i.e., there exist $s'_1 \leftarrow^*_{R'} s' \to^*_{R'} s'_2$ for which $\nexists t' \in R'\colon s'_1 \to^*_{R'} t' \leftarrow^*_{R'} s'_2$.

By (C2′): $\exists t'_1, t'_2 \in R'_{NF}\colon t'_1 \leftarrow^*_{R'} s'_1 \leftarrow^*_{R'} s' \to^*_{R'} s'_2 \to^*_{R'} t'_2$ obtaining $t'_1 \leftrightarrow^*_{R'} t'_2$ where $t'_1 \neq t'_2$.

By (C1′) and (C4′): $\exists t_1, t_2 \in R_{NF}\colon G(t_1) = t'_1 \wedge G(t_2) = t'_2$

By (C5′): $t'_1 \neq t'_2$ yields $t_1 \not\approx t_2$

By (C3′): $t'_1 \leftrightarrow^*_{R'} t'_2 \Rightarrow t_1 \overset{\approx}{\leftrightarrow}^*_R t_2$

Since R is confluent modulo $\approx$ we apply Lemma 17 and obtain $t_1 \approx t_2$ which contradicts $t_1 \not\approx t_2$.

"$\Leftarrow$": Assume that R' is confluent and R is not confluent modulo $\approx$, i.e., there exist $s_1 \leftarrow^*_R s_a \approx s_b \to^*_R s_2$ for which $\nexists t_1, t_2 \in R\colon s_1 \to^*_R t_1 \approx t_2 \leftarrow^*_R s_2$.

By (C2′): $\exists t_1, t_2 \in R_{NF}\colon t_1 \leftarrow^*_R s_1 \leftarrow^*_R s_a \approx s_b \to^*_R s_2 \to^*_R t_2$ where $t_1 \not\approx t_2$.

By (C3′): $G(t_1) \leftrightarrow^*_{R'} G(s_1) \leftrightarrow^*_{R'} G(s_a) \leftrightarrow^*_{R'} G(s_b) \leftrightarrow^*_{R'} G(s_2) \leftrightarrow^*_{R'} G(t_2)$ obtaining $G(t_1) \leftrightarrow^*_{R'} G(t_2)$.

By (C1′) and (C4′): $\exists t'_1, t'_2 \in R'_{NF}\colon G(t_1) = t'_1 \wedge G(t_2) = t'_2$, obtaining $t'_1 \leftrightarrow^*_{R'} t'_2$.

Since R' is confluent it also has the Church-Rosser property, i.e., $x \leftrightarrow^* y \Rightarrow \exists z, x \to^* z \leftarrow^* y$. Because t'_1 and t'_2 are normal forms they are irreducible, thus, $\exists z, x \to^0 z \leftarrow^0 y$ yielding $t'_1 = t'_2$.

By (C5′): $t_1 \not\approx t_2 \Rightarrow t'_1 \neq t'_2$ which contradicts $t'_1 = t'_2$.

To apply Lemma 24 for a given system R, one may search for an R' which is as simple or as small as possible to reduce the complexity of especially the second part of $C3'$, as demonstrated in the following examples.

We summarize the application of the above to probabilistic systems in Theorems 25 and 27.

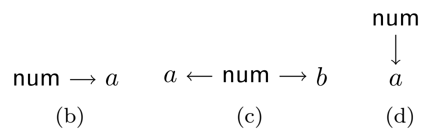
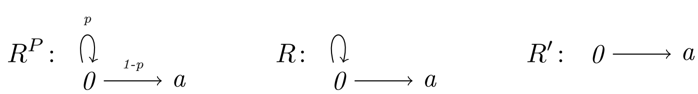


Figure 3: Systems (b)–(d) are transformed systems of the underlying ARS system in Figure 1 with the same names.

Theorem 25 ([7]). *An almost-surely terminating PARS* $R^P = ((A, \to_R), P)$ *is almost-surely convergent if there exists an ARS* $R' = (A', \to_{R'})$ *and a mapping* $G\colon A \to A'$ *which together with* $(A, \to_R)$ *satisfy (C1)–(C5).*

Proof. Since R^P is a-s. terminating, R is normalizing (Lemma 11). So, given an ARS R' and G be a mapping from R to R' satisfying (C1), (C3)–(C5), we can apply Lemma 23 and obtain that R and thereby R^P is confluent. A-s. convergence of R^P follows from Prop. 22 since R^P is confluent and a-s. terminating.

Example 26. *We consider the nonterminating, almost-surely terminating system* R^P *(below to the left) with the underlying normalizing system* R *(below, middle), the confluent system* R' *(below to the right) and the mapping* $G(0)= 0$, $G(a)= a$.



The systems R, R' *and the mapping* G *satisfy (C1)–(C5), and therefore we can conclude that* R^P *is almost-surely convergent.*

Theorem 27. *Given an almost-surely terminating PARS* $R^P = (R, P)$ *with* $R = (A, \to_R)$*, an equivalence relation* $\approx$ *over* A*, an ARS* $R' = (A', \to_{R'})$ *and a mapping* G *from* A *to* A' *which together with* R *satisfy (C1′)–(C5′), then system* R^P *is almost-surely convergent modulo* $\approx$ *if and only* R' *is confluent.*

Proof. Assume notation as above. Since R^P is a-s. terminating, R is normalizing (Lemma 11), satisfying the first part of (C2′). Thus, given an ARS R' and a mapping G which together with R and $\approx$ satisfy (C1′)–(C5′), we can apply Lemma 24 obtaining that R is confluent modulo $\approx$ iff R' is confluent. Prop. 22 gives that the a-s. terminating R^P is a-s. convergent iff R' is confluent.

Corollary 28. *Given an almost-surely terminating PARS* $R^P = (R, P)$ *with* $R = (A, \to_R)$*, an ARS* $R' = (A, \to_{R'})$ *and a mapping* G *from* A *to* A' *which together with* R *satisfy (C1′)–(C5′), then system* R^P *is almost-surely convergent if and only* R' *is confluent.*

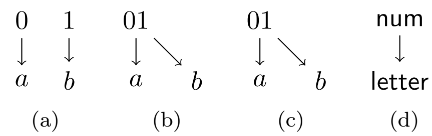


Figure 4: Examples of transformed systems where the original systems are those of Figure 2.

Example 29. *The systems (b)–(d) in Figure 3 are transformed systems of the underlying normalizing ARS in Figure 1 with the same names. The transformation is done by mapping all numbers to* num *and all letters to themselves. By Corollary 28 the confluence of the transformed systems (b), and (d) imply that the ditto systems in Figure 1 are confluent; and that non-confluence of (c) in Figure 3 implies that (c) in Figure 1 is not confluent.*

Example 30. *The systems (a)–(d) in Figure 4 are transformed systems of the underlying normalizing ARS in Figure 2 with the same names. There are three mappings;* $G_{(a)}$ *is used when transforming system (a): it maps the states to states of the same name, e.g.,* $G_{(a)}(1) = 1$*;* $G_{(b)(c}$ *is used when transforming (b) and (c) and is defined by* $G_{(b)(c)}(a) = a$*,* $G_{(b)(c)}(a) = b$*,* $G_{(b)(c)}(0) = G_{(b)(c)}(1)01$*; and the third mapping* $G_{(d)}$ *is used when transforming (d) and is defined such that all numbers are mapped to* num *and all normal forms to* letter*.*

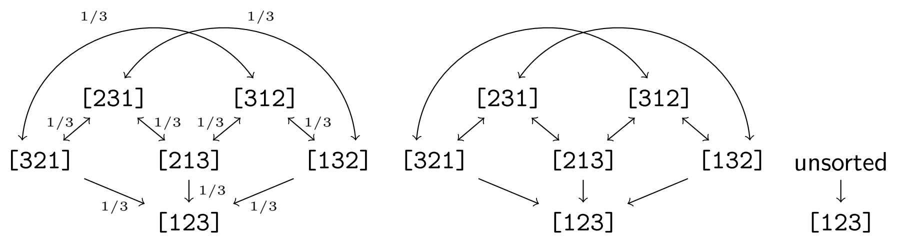
The confluence of transformed systems (a) and (d) imply that the ditto original systems in Figure 2 are confluent. The fact that (b) and (c) in Figure 4 are not confluent imply that their original systems are not confluent.

5. Examples

In the following we show almost-sure convergence (modulo equivalence) in two different cases that exemplify Theorem 27 and Corollary 28. As shown by [16, 17, 18], to prove that a PARS $R^P = ((A, \to), P)$ is almost-surely terminating, it suffices to show existence of a *Lyapunov ranking function*, i.e., a function $\mathcal{V} : A \to \mathbb{R}^+$ where $\forall s \in A$ there exists an $\epsilon > 0$ so the *inequality of s*, $\mathcal{V}(s) \geq \sum_{s \to s'} P(s \to s') \cdot \mathcal{V}(s') + \epsilon$ holds.

5.1. A naive sorting algorithm

We consider a naive sorting algorithm, which stops when the list is sorted and, otherwise, randomly chooses two elements and interchange them; all element pairs have an equal probability of being interchanged. We study two simple cases, one with non-repeating elements and one with repeating elements. For the repeating case we assume that each element has both a value and a unique identifier, for instance 2_a and 2_b both have value 2, but different identifiers a and b, respectively. We assume that the input list has three elements; in the non-repeating case the list contains `1`, `2`, and `3`, and in the repeating case the list contains $\mathtt{1_a}$, $\mathtt{2_b}$, and $\mathtt{2_c}$.



(a) Original R^P. (b) Underlying R. (c) Confl. R'

Figure 5: Naive sorting of a non-repeating list.

In the following we display the two PARS system described by the algorithm and the possible lists. Our objective is to show almost-sure convergence for the case with non-repeating elements and almost-sure convergence modulo an equivalence defining that any two value-ordered lists are equivalent, for instance $[\mathsf{1_a\,2_b\,2_c}] \approx [\mathsf{1_a\,2_c\,2_b}]$.

5.1.1. Non-repeating lists

We consider $R^P = (R, P)$, shown in Figure 5(a), and start by showing it a-s. terminating using a Lyapunov ranking function. Then, we show R^P a-s. convergent using Corollary 28; let $\approx$ be $=$ and provide a mapping G that satisfy (C1′)–(C5′).

The function $\mathcal{V}$ is defined by $\mathcal{V}(\texttt{[123]}) = 1$, $\mathcal{V}(\texttt{[321]}) = \mathcal{V}(\texttt{[213]}) = \mathcal{V}(\texttt{[132]}) = 3$, and $\mathcal{V}(\texttt{[231]}) = \mathcal{V}(\texttt{[312]}) = 4$. The following calculations show that $\mathcal{V}$ is in fact a Luyapunov ranking function for R^P.

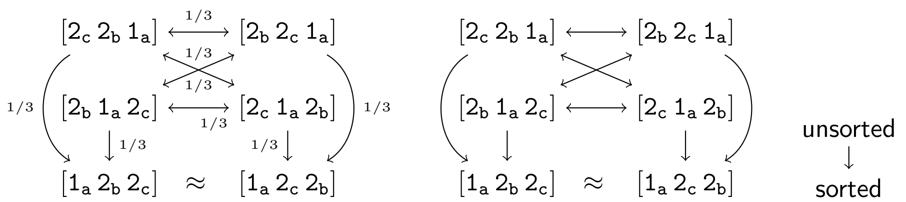
$$
\begin{aligned}
&4 = \mathcal{V}(\texttt{[231]}) > \tfrac{1}{3}\mathcal{V}(\texttt{[321]}) + \tfrac{1}{3}\mathcal{V}(\texttt{[213]}) + \tfrac{1}{3}\mathcal{V}(\texttt{[132]}) = \tfrac{3}{3} + \tfrac{3}{3} + \tfrac{3}{3} = 3\\
&4 = \mathcal{V}(\texttt{[312]}) > \tfrac{1}{3}\mathcal{V}(\texttt{[321]}) + \tfrac{1}{3}\mathcal{V}(\texttt{[213]}) + \tfrac{1}{3}\mathcal{V}(\texttt{[132]}) = \tfrac{3}{3} + \tfrac{3}{3} + \tfrac{3}{3} = 3\\
&3 = \mathcal{V}(\texttt{[321]}) > \tfrac{1}{3}\mathcal{V}(\texttt{[231]}) + \tfrac{1}{3}\mathcal{V}(\texttt{[312]}) + \tfrac{1}{3}\mathcal{V}(\texttt{[123]}) = \tfrac{4}{3} + \tfrac{4}{3} + \tfrac{1}{3} = 2 + \tfrac{2}{3}\\
&3 = \mathcal{V}(\texttt{[213]}) > \tfrac{1}{3}\mathcal{V}(\texttt{[231]}) + \tfrac{1}{3}\mathcal{V}(\texttt{[312]}) + \tfrac{1}{3}\mathcal{V}(\texttt{[123]}) = \tfrac{4}{3} + \tfrac{4}{3} + \tfrac{1}{3} = 2 + \tfrac{2}{3}\\
&3 = \mathcal{V}(\texttt{[132]}) > \tfrac{1}{3}\mathcal{V}(\texttt{[231]}) + \tfrac{1}{3}\mathcal{V}(\texttt{[312]}) + \tfrac{1}{3}\mathcal{V}(\texttt{[123]}) = \tfrac{4}{3} + \tfrac{4}{3} + \tfrac{1}{3} = 2 + \tfrac{2}{3}\\
&1 = \mathcal{V}(\texttt{[123]}) > 0
\end{aligned}
$$

Since R^P is a-s. terminating, it suffice to let $R' = (\{\mathsf{unsorted}, \texttt{[123]}\}, \mathsf{unsorted} \to \texttt{[123]})$, see Figure 5(c), and the mapping G defined by $G(\texttt{[231]}) = G(\texttt{[312]}) = G(\texttt{[321]}) = G(\texttt{[213]}) = G(\texttt{[132]}) = \mathsf{unsorted}$ and $G(\texttt{[123]}) = \texttt{[123]}$.

Because R^P is a-s. terminating, R' is (trivially) a confluent system, and the mapping G satisfies (C1′)–(C5′), we get (by Cor. 28) R^P is a-s. convergent.

5.1.2. Repeating lists

We consider $R^P = (R, P)$, shown in Figure 6(a), and proceed as above showing it a-s. terminating using a Lyapunov ranking function. Then, we show



(a) Original R^P. (b) Underlying R. (c) Confl. R'

Figure 6: Naive sorting of a repeating list.

R^P a-s. convergent using Theorem 27 by a mapping G that satisfy (C1′)–(C5′), and the mapped system is trivially confluent.

We define the function $\mathcal{V}$ by $\mathcal{V}([1_a\, 2_b\, 2_c]) = \mathcal{V}([1_a\, 2_c\, 2_b]) = 1$ and $\mathcal{V}([2_b\, 1_a\, 2_c]) = \mathcal{V}([2_c\, 1_a\, 2_b]) = \mathcal{V}([2_c\, 2_b\, 1_a]) = \mathcal{V}([2_b\, 2_c\, 1_a]) = 3$. The following calculations show that $\mathcal{V}$ is in fact a Luyapunov ranking function for R^P.

$$
\begin{aligned}
3 = \mathcal{V}([2_c\, 2_b\, 1_a]) &> \tfrac{1}{3}\mathcal{V}([2_b\, 2_c\, 1_a]) + \tfrac{1}{3}\mathcal{V}([2_c\, 1_a\, 2_b]) + \tfrac{1}{3}\mathcal{V}([1_a\, 2_b\, 2_c]) = 2 + \tfrac{1}{3}\\
3 = \mathcal{V}([2_b\, 2_c\, 1_a]) &> \tfrac{1}{3}\mathcal{V}([2_c\, 2_b\, 1_a]) + \tfrac{1}{3}\mathcal{V}([2_b\, 1_a\, 2_c]) + \tfrac{1}{3}\mathcal{V}([1_a\, 2_c\, 2_b]) = 2 + \tfrac{1}{3}\\
3 = \mathcal{V}([2_b\, 1_a\, 2_c]) &> \tfrac{1}{3}\mathcal{V}([2_b\, 2_c\, 1_a]) + \tfrac{1}{3}\mathcal{V}([2_c\, 1_a\, 2_b]) + \tfrac{1}{3}\mathcal{V}([1_a\, 2_b\, 2_c]) = 2 + \tfrac{1}{3}\\
3 = \mathcal{V}([2_c\, 1_a\, 2_b]) &> \tfrac{1}{3}\mathcal{V}([2_c\, 2_b\, 1_a]) + \tfrac{1}{3}\mathcal{V}([2_b\, 1_a\, 2_c]) + \tfrac{1}{3}\mathcal{V}([1_a\, 2_c\, 2_b]) = 2 + \tfrac{1}{3}\\
1 = \mathcal{V}([1_a\, 2_b\, 2_c]) &> 0\\
1 = \mathcal{V}([1_a\, 2_c\, 2_b]) &> 0
\end{aligned}
$$

Since R^P is a-s. terminating, it suffice to define $R' = (\{\mathsf{unsorted}, \mathsf{sorted}\}, \mathsf{unsorted} \to \mathsf{sorted})$, see Figure 6(c), and the mapping G defined by $G([2_c\, 2_b\, 1_a]) = G([2_b\, 2_c\, 1_a]) = G([2_b\, 1_a\, 2_c]) = G([2_c\, 1_a\, 2_b]) = \mathsf{unsorted}$ and $G([1_a\, 2_b\, 2_c]) = G([1_a\, 2_c\, 2_b]) = \mathsf{sorted}$.

Because R^P is a-s. terminating, R' is (trivially) a confluent system, and the mapping G satisfies (C1′)–(C5′), we get (by Thm. 27) that R^P is a-s. convergent modulo $\approx$.

6. Conclusion

We have introduced and characterized a novel notion of almost-sure convergence modulo equivalence for probabilistic abstract reduction systems. This generalization of earlier results without the equivalence [15, 7] is important, as it opens up for studying a much larger class of interesting systems having a convergent behaviour, and we gave some indicative examples of systems with this property.

Our main results aim to facilitate proving – or disproving – almost-sure convergence modulo equivalence of given systems. As a central theorem (Th. 18), we have shown that almost-sure convergence modulo equivalence is the same as almost-sure termination plus confluence modulo equivalence. The picture is completed by our generalization to modulo equivalence of an important result [6] on proving confluence by relating the system of interest to another, known to be confluent.

It would be useful to automate the construction of the transformations, especially for rule-based systems like term rewriting systems or CHR; as this would provide a fully automatic method for proving confluence modulo equivalence of such systems and a step towards proving almost-sure convergence modulo equivalence automatically.

A related property which have been disregarded is local confluence; local confluence is essential for proving confluence (with or without equivalence) for terminating systems [3, 4]. An almost-sure terminating system may not be terminating and thus Newman's Lemma cannot be applied directly, but it is highly relevant to investigate local properties for almost-surely terminating systems.

References

[1] A. Church, J. B. Rosser, Some properties of conversion, Transactions of the American Mathematical Society 39 (3) (1936) 472–482.

[2] F. Baader, T. Nipkow, Term rewriting and all that, Cambridge University Press, 1999.

[3] G. P. Huet, Confluent reductions: Abstract properties and applications to term rewriting systems, J. ACM 27 (4) (1980) 797–821. doi:10.1145/322217.322230.

[4] M. H. A. Newman, On theories with a combinatorial definition of "equivalence", Annals of Mathematics 43 (2) (1942) 223–243.

[5] A. V. Aho, R. Sethi, J. D. Ullman, Code Optimization and Finite Church-Rosser Systems, in: R. Rustin (Ed.), Design and Optimization of Compilers, Prentice-Hall, 1972, pp. 89–106.

[6] P. Curien, G. Ghelli, On confluence for weakly normalizing systems, in: RTA-91, 1991, pp. 215–225. doi:10.1007/3-540-53904-2_98.

[7] M. H. Kirkeby, H. Christiansen, Confluence and Convergence in Probabilistically Terminating Reduction Systems, in: F. Fioravanti, J. P. Gallagher (Eds.), LOPSTR 2017, Vol. 10855 of LNCS, Springer, 2018, pp. 164–179. doi:10.1007/978-3-319-94460-9_10.

[8] T. Fruhwirth, Constraint Handling Rules, Cambridge University Press, Cambridge, 2009. doi:10.1017/CBO9780511609886.

[9] S. Abdennadher, Operational semantics and confluence of constraint propagation rules, in: CP97, Vol. 1330 of LNCS, Springer, 1997, pp. 252–266. doi:10.1007/BFb0017444.

[10] S. Abdennadher, T. W. Frühwirth, H. Meuss, On confluence of Constraint Handling Rules, in: CP96, Vol. 1118 of LNCS, Springer, 1996, pp. 1–15.

[11] H. Christiansen, M. H. Kirkeby, Confluence Modulo Equivalence in Constraint Handling Rules, in: M. Proietti, H. Seki (Eds.), LOPSTR 2014. Revised Selected Papers, Vol. 8981 of LNCS, Springer, 2014, pp. 41–58.

[12] H. Christiansen, M. H. Kirkeby, On proving confluence modulo equivalence for Constraint Handling Rules, Formal Aspects of Computing 29 (1) (2017) 57–95. arXiv:1611.03628, doi:10.1007/s00165-016-0396-9.

[13] S. Hart, M. Sharir, A. Pnueli, Termination of probabilistic concurrent program, ACM Trans. Program. Lang. Syst. 5 (3) (1983) 356–380. doi:10.1145/2166.357214.

[14] T. W. Frühwirth, A. D. Pierro, H. Wiklicky, Probabilistic Constraint Handling Rules, Electr. Notes Theor. Comput. Sci. 76 (2002) 115–130. doi:10.1016/S1571-0661(04)80789-8.

[15] O. Bournez, C. Kirchner, Probabilistic rewrite strategies. Applications to ELAN, in: S. Tison (Ed.), RTA 2002, Vol. 2378 of LNCS, Springer, 2002, pp. 252–266. doi:10.1007/3-540-45610-4_18.

[16] O. Bournez, F. Garnier, Proving positive almost sure termination under strategies, in: F. Pfenning (Ed.), RTA 2006, Vol. 4098 of LNCS, Springer, 2006, pp. 357–371. doi:10.1007/11805618_27.

[17] L. M. F. Fioriti, H. Hermanns, Probabilistic termination: Soundness, completeness, and compositionality, in: POPL 2015, 2015, pp. 489–501. doi:10.1145/2676726.2677001.

[18] M. Avanzini, U. D. Lago, A. Yamada, On probabilistic term rewriting, in: J. P. Gallagher, M. Sulzmann (Eds.), FLOPS 2018, Vol. 10818 of LNCS, Springer, 2018, pp. 132–148. doi:10.1007/978-3-319-90686-7_9.

[19] C. Faggian, Probabilistic rewriting: Relations between normalization, termination, and unique normal forms, CoRR abs/1804.05578. arXiv:1804.05578.

[20] A. Díaz-Caro, G. Martinez, Confluence in probabilistic rewriting, in: 12th Workshop on Logical and Semantic Frameworks with Applications, LFSA 2017, 2017. arXiv:1708.03536.

[21] J. Sneyers, D. D. Schreye, Probabilistic termination of CHRiSM programs, in: LOPSTR 2011, 2011, pp. 221–236. doi:10.1007/978-3-642-32211-2_15.

[22] E. W. Weisstein, q-Pochhammer Symbol, MathWorld – A Wolfram Web Resource, 2017.
URL http://mathworld.wolfram.com/q-PochhammerSymbol.html

Appendix A. Selected proofs

Proposition 5. *For an arbitrary finite path* π, $1 \geq P(\pi) > 0$. *For every element* s, $P(s \to^* \cdot)$ *and* $P(s \to^\infty)$ *comprise a probability distribution, i.e.,* $\forall t \in R_{NF}(s)\colon 0 \leq P(s \to^* t) \leq 1$; $0 \leq P(s \to^\infty) \leq 1$; *and* $\sum_{t \in R_{NF}(s)} P(s \to^* t) + P(s \to^\infty) = 1$.

Proof. Part one follows by Definition 3. Part two is shown by defining a sequence of distributions $P^{(n)}$, $n \in \mathbb{N}$, only containing paths up to length n, and show that it converges to P. Let$\Delta^{(n)}(s,t)$ be the subset of $\Delta(s,t)$ with paths of length n or less, and $\Delta^{(n)}(s,\sharp)$ be the set of paths of length n, starting in s and ending in a reducible element.
We can now define $P^{(n)}$ over $\{\Delta^{(n)}(s,t) \mid t \in R_{NF}(s)\} \uplus \{\Delta^{(n)}(s,\sharp)\}$ as follows:

$$P^{(n)}(s \to^* t) = \textstyle\sum_{\delta \in \Delta^{(n)}(s,t)} P(\delta), \quad \text{and} \tag{A.1}$$

$$P^{(n)}(s \to^\infty) = \textstyle\sum_{\pi \in \Delta^{(n)}(s,\sharp)} P(\pi). \tag{A.2}$$

First, we prove by induction that $P^{(n)}$ is a distribution for all n. The $P^{(0)}$ is a distribution because: (i) If s is irreducible, $P^{(0)}(s \to^* s) = 1$ (the empty-path); and $P^{(0)}(s \to^\infty) = 0$ (a sum of zero elements). (ii) If s is reducible, $P^{(0)}(s \to^* s) = 0$; and $P^{(0)}(s \to^\infty) = \sum_{s \to t} P(s \to t) = 1$ by Definition 3.
The inductive step: The sets $\Delta^{(n+1)}(s,t)$, $t \in R_{NF}(s)$, and $\Delta^{(n+1)}(s,\sharp)$ can be constructed by, for each path in $\Delta^{(n)}(s,\sharp)$, create its possible extensions by one reduction. When an extension leads to a normal form t, it is added to $\Delta^{(n)}(s,t)$. Otherwise, i.e., if the new path leads to a reducible, it is included in $\Delta^{(n+1)}(s,\sharp)$. Formally, for any normal form t of s, we write:

$$\begin{aligned}
\Delta^{(n+1)}(s,t) &= \{(s \to \cdots \to u \to t) \mid (s \to \cdots \to u) \in \Delta^{(n)}(s,\sharp),\, u \to t\} \uplus \Delta^{(n)}(s,t) \\
\Delta^{(n+1)}(s,\sharp) &= \{(s \to \cdots \to u \to v) \mid (s \to \cdots \to u) \in \Delta^{(n)}(s,\sharp),\, u \to v,\, u \notin R_{NF}(s)\}
\end{aligned}$$

We show that for a given s, the probability mass added to the $\Delta^{(\cdot)}(s,t)$ sets is equal to the probability mass removed from $\Delta^{(\cdot)}(s,\sharp)$ as follows (where $\delta_{su} = (s \to \cdots \to u)$).

$$\begin{aligned}
&\sum_{t \in R_{NF}(s)} P^{(n+1)}(s \to^* t) + P^{(n+1)}(s \to^\infty) = \sum_{\substack{t \in R_{NF}(s) \\ \delta \in \Delta^{(n+1)}(s,t)}} P^{(n+1)}(\delta) + P^{(n+1)}(s \to^\infty) \\
&= \sum_{\substack{t \in R_{NF}(s) \\ \delta_{st} \in \Delta^{(n)}(s,t)}} P^{(n)}(\delta) \;+ \sum_{\substack{\delta_{su} \in \Delta^{(n)}(s,\sharp), \\ u \to v, v \in R_{NF}(s)}} P^{(n)}(\delta) P(u \to v) \;+ \sum_{\substack{\delta_{su} \in \Delta^{(n)}(s,\sharp), \\ u \to v, v \notin R_{NF}(s)}} P^{(n)}(\delta) P(u \to v) \\
&= \sum_{t \in R_{NF}(s)} P^{(n)}(s \to^* t) + \sum_{\substack{\delta_{su} \in \Delta^{(n)}(s,\sharp), \\ u \to v}} P^{(n)}(\delta) P(u \to v) = \sum_{t \in R_{NF}(s)} P^{(n)}(s \to^* t) + \sum_{\delta_{su} \in \Delta^{(n)}(s,\sharp)} P^{(n)}(\delta) \left(\sum_{u \to v} P(u \to v) \right) \\
&= \sum_{t \in R_{NF}(s)} P^{(n)}(s \to^* t) + P^{(n)}(s \to^\infty) = 1
\end{aligned}$$

Thus, for given s, $P^{(n+1)}$ defines a probability distribution. Notice also that the equations above indicate that $P^{(n+1)}(s \to^* t) \geq P^{(n)}(s \to^* t)$, for all $t \in R_{NF}(s)$.

Finally, for any s and $t \in R_{NF}(s)$, $\lim_{n\to\infty} \Delta^{(n)}(s,t) = \Delta(s,t)$, we get (as we consider increasing sequences of real numbers in a closed interval) $\lim_{n\to\infty} P^{(n)}(s \to^* t) = P(s \to^* t)$, and as a consequence of this, $\lim_{n\to\infty} P^{(n)}(s \to^\infty) = P(s \to^\infty)$. This finishes the proof.

Proposition 6. *Consider a PARS which has an element s for which $\Delta^\infty(s)$ is countable (finite or infinite). Let $P(s_1 \to s_2 \to \cdots) = \prod_{i=1,2,\ldots} P(s_i \to s_{i+1})$ be the probability of an infinite path, then $P(s \to^\infty) = \sum_{\delta \in \Delta^\infty(s)} P(\delta)$.*

Proof. We assume the characterization in the proof of Proposition 5 above, of P by the limits of the functions $P^{(n)}(s \to^* t)$ and $P^{(n)}(s \to^\infty)$ given by equations (A.1) and (A.2). When $\Delta^\infty(s)$ is countable, $\lim_{n\to\infty} P^{(n)}(s \to^\infty) = \sum_{\delta \in \Delta^\infty(s)} P(\delta)$.