Proving Positive Almost Sure Termination Under Strategies

Olivier Bournez and Florent Garnier

LORIA/INRIA, 615 Rue du Jardin Botanique
54602 Villers lès Nancy Cedex, France

Abstract In last RTA, we introduced the notion of probabilistic rewrite systems and we gave some conditions entailing termination of those systems within a finite mean number of reduction steps.
Termination was considered under arbitrary unrestricted policies. Policies correspond to strategies for non-probabilistic rewrite systems.
This is often natural or more useful to restrict policies to a subclass. We introduce the notion of positive almost sure termination under strategies, and we provide sufficient criteria to prove termination of a given probabilitic rewrite system under strategies. This is illustrated with several examples.

1 Introduction

As discussed in several papers such as [7,22,15], when specifying probabilistic systems, it is rather natural to consider that the firing of a rewrite rule can be subject to some probabilistic law.

Considering rewrite rules subject to probabilities leads to numerous questions about the underlying notions and results. In [7], we introduced probabilistic abstract reduction systems, and we introduced notions like almost-sure termination or probabilistic confluence, with relations between all these notions. In [6], we proved that, unlike what happens for classical rewriting logic, there is no hope to build a sound and complete proof system with probabilities in the general case [6]. In [5], we argue that positive almost sure termination is a better notion than simple almost sure termination for probabilistic systems and we provide necessary and sufficient criteria entailing positive almost sure termination.

In this paper, we pursue the investigation, by considering positive almost sure termination under strategies. As we show through several examples, it is often natural to restrict strategies to a subset of strategies. Many simple probabilistic rewrite systems do not terminate under arbitrary strategies, whereas they terminate if strategies are restricted to natural strategies.

The idea of adding probabilities to high level models of reactive systems is not new, and has also been explored for models like Petri Nets [3,26], automata based models [10,27], or process algebra [16]. There is now a rather important literature about model-checking techniques for probabilistic systems: see example [21] and the references there. Computer Tools like PRISM [20], APMC [19], do exist. Observe however, that most of the studies and techniques restrict to finite state systems.

Termination of probabilistic concurrent programs has already been investigated. In particular, in [18] it has been argued that this is important to restrict to fair schedulers, and techniques for proving termination under fair schedulers have been provided. These techniques have been extended to infinite systems in [17]. Compared to our work, they focus on almost sure termination, whereas we focus on positive almost sure termination. Furthermore, we deal with probabilistic abstract reduction systems or rewrite systems, whereas these two papers are focusing on concurrent programs, where strategies correspond to schedulers.

Several notions of fairness have been introduced for concurrent programs, and in particular for probabilistic concurrent programs. In particular, Pnueli [23], and Pnueli and Zuck have introduced extreme fairness and α-fairness [24]. Hart, Sharir and Pnueli [18] and Vardi [27] consider probabilistic systems in which the choice of actions at the states is subject to fairness requirements, and proposed model checking algorithms. A survey and discussion of several fairness notions for probabilistic systems can be found in chapter 8 of [10].

Probabilistic abstract reduction systems and probabilistic rewrite systems do correspond to classical abstract reduction systems and classical rewrite systems where probabilities can only be 0 or 1 [5]. Therefore, any technique for proving termination of a probabilistic system must have a counterpart for classical systems. In particular, any technique for proving termination of probabilistic rewrite systems under strategies is an extension of a technique for proving termination of classical rewrite systems under strategies. The termination of rewrite systems under strategies has been investigated in e.g. [12,13]. Since the extension to the probabilistic case of very basic techniques already yields several problems discussed in this paper, we do not consider so general strategies.

The paper is organized as follows: in Section 2, we recall probabilistic abstract reduction systems, and probabilistic rewrite systems, as well as several concepts and results from [5]. In Section 3, we introduce positive almost sure termination under strategies, and we discuss several examples of systems that are non positively almost surely terminating but which are positively almost surely terminating under some strategies. In Section 4, we derive some techniques to prove positive almost sure termination under strategies. In Section 5, we discuss several applications of our results.

2 Probabilistic Abstract Reduction Systems and Probabilistic Rewrite Systems

A *stochastic sequence on a set* A is a family $(X_i)_{i\in\mathbb{N}}$, of random variables defined on some fixed probability space (Ω, σ, P) with values on A. It is said to be *Markovian* if its conditional distribution function satisfies the so-called Markov property, that is for all n and $s \in A$,

$$P(X_n = s | X_0 = \pi_0, X_1 = \pi_1, \dots, X_{n-1} = \pi_{n-1}) = P(X_n = s | X_{n-1} = \pi_{n-1}),$$

and *homogeneous* if furthermore this probability is independent of n.

Probabilistic abstract reduction systems (PARS) were introduced in [5]. In the same way that abstract reduction systems are also called *transition systems* in other contexts, PARS correspond[1][2] to *Markov Decision Processes* [25].

Definition 1 (PARS). *Given some denumerable set S, we note $Dist(S)$ for the set of probability distributions on S: $\mu \in Dist(S)$ is a function $S \to [0,1]$ that satisfies $\sum_{i \in S} \mu(i) = 1$.*

A probabilistic abstract reduction system (PARS) is a pair $\mathcal{A} = (A, \to)$ consisting of a countable set A and a relation $\to \subset A \times Dist(A)$. A state $a \in A$ with no μ such that $a \to \mu$ is said terminal.

A PARS is said deterministic if, for all a, there is at most one μ with $a \to \mu$. We denote $Dist(\mathcal{A})$ for the set of distributions μ with $a \to \mu$ for some a.

We now need to explain how such systems evolve: a *history* is a finite sequence $a_0 a_1 \cdots a_n$ of elements of the state space A. It is non-terminal if a_n is.

Definition 2 (Deterministic Policy/Strategy). *A (deterministic) policy ϕ, that can also be called a (deterministic) strategy, is a function that maps non-terminal histories to distributions in such a way that $\phi(a_0 a_1 \cdots a_n) = \mu$ is always one (of the possibly many) distribution μ with $a_n \to \mu$. A history is said realizable, if for all $i < n$, if μ_i denotes $\phi(a_0 a_1 \cdots a_i)$, one has $\mu_i(a_{i+1}) > 0$.*

Actually, previous definition assumes that strategies must be deterministic (μ is a deterministic function of the history). If we want to be very general, we can also allow the strategy to be itself random (μ is selected among the possible μ with $a_n \to \mu$ in a random fashion).

Definition 3 (Randomized Policy/Strategy). *A randomized policy ϕ, that can also be called a randomized strategy, is a function that maps non-terminal histories to $Dist(M)$, where M is the set of μ with $a_n \to \mu$.*

Following the classification from [25], one can also distinguish history dependent strategies (the general case) from Markovian strategies (the value of the function on a history $a_0, \cdots, a_n$ depends only on a_n), to get the classes HD, HR, MD, MD, where H is for history dependent, M for Markovian, D for deterministic, R for randomized. In what follows, when we talk about strategies, it may mean a strategy of any of these classes.

A *derivation of $\mathcal{A}$* is then a stochastic sequence where the non-deterministic choices are given by some policy ϕ, and the probabilistic choices are governed by the corresponding distributions.

---

[1] The only true difference with [25] is that here action names are omitted.

[2] We prefer to keep to the terminology of [5], since we think that PARS indeed correspond to a probabilistic extension of Abstract Reduction Systems (ARS), Markov Decision Processes indeed correspond to a probabilistic extension of transition systems, and hence that the question of the best terminology is related to the question of the best terminology for ARS/transition systems, i.e. a cultural question.

Definition 4 (Derivations). *A derivation* π *of* $\mathcal{A}$ *over policy* ϕ *is a stochastic sequence* $\pi = (\pi_i)_{i \in \mathbb{N}}$ *on set* $A \cup \{\perp\}$ *(where* $\perp$ *is a new element:* $\perp \notin A$*) such that for all* n*,*

$$P(\pi_{n+1} = \perp | \pi_n = \perp) = 1,$$

$$P(\pi_{n+1} = \perp | \pi_n = s) = 1 \text{ if } s \in A \text{ is terminal},$$

$$P(\pi_{n+1} = \perp | \pi_n = s) = 0 \text{ if } s \in A \text{ is non-terminal},$$

and for all $t \in A$*.*

$$P(\pi_{n+1} = t | \pi_n = a_n, \pi_{n-1} = a_{n-1}, \ldots, \pi_0 = a_0) = \mu(t)$$

whenever $a_0 a_1 \cdots a_n$ *is a realizable non-terminal history and* $\mu = \phi(a_0 a_1 \cdots a_n)$*.*

If a derivation is such that $\pi_n = \perp$ for some n, then $\pi_{n'} = \perp$ almost surely for all $n' \geq n$. Such a derivation is said to be *terminating*. In other words, a non-terminating derivation is such that $\pi_n \in A$ ($\pi_n \neq \perp$) for all n.

The following two notions were introduced in [5]:

Definition 5 (Almost Sure Termination). *A PARS* $\mathcal{A} = (A, \rightarrow)$ *will be said almost surely (a.s) terminating iff for any policy* ϕ*, the probability that a derivation* $\pi = (\pi_i)_{i \in \mathbb{N}}$ *under policy* ϕ *terminates is* 1*: i.e. for all* ϕ*,* $P(\exists n | \pi_n = \perp) = 1$*.*

Definition 6 (Positive Almost Sure Termination). *A PARS* $\mathcal{A} = (A, \rightarrow)$ *will be said positively almost surely (+a.s.) terminating if for all policies* ϕ*, for all states* $a \in A$*, the mean number of reduction steps before termination under policy* ϕ *starting from* a*, denoted by* $T[a, \phi]$*, is finite.*

The following was proved in [5].

Theorem 1. *A PARS* $\mathcal{A} = (A, \rightarrow)$ *is +a.s. terminating if there exist some function* $V : A \rightarrow \mathbb{R}$*, with* $\inf_{i \in A} V(i) > -\infty$*, and some* $\epsilon > 0$*, such that, for all states* $a \in A$*, for all* μ *with* $a \rightarrow \mu$*, the drift in* a *according to* μ *defined by*

$$\Delta_\mu V(a) = \sum_i \mu(i) V(i) - V(a)$$

satisfies

$$\Delta_\mu V(a) \leq -\epsilon.$$

The technique was proved complete for finitely branching systems in [5]: such a function V always exists for +a.s. terminating finitely branching systems.

In [5], we also introduce the following notion, that covers classical (i.e. non-probabilistic) rewrite systems, and also Markov chains over finite spaces. It follows in particular that all examples that have been modeled in literature using finite Markov chains (for e.g. in model-checking contexts [21,20]) can be modeled as probabilistic rewrite systems.

Definition 7 (Probabilistic Rewrite system). *Given a signature Σ and a set of variables X, the set of terms over Σ and X is denoted by $T(\Sigma, X)$.*

A probabilistic rewrite rule is an element of $T(\Sigma, X) \times Dist(T(\Sigma, X))$. A probabilistic rewrite system is a finite set $\mathcal{R}$ of probabilistic rewrite rules.

To a probabilistic rewrite system is associated a probabilistic abstract reduction system $(T(\Sigma, X), \to_{\mathcal{R}})$ over the set of terms $T(\Sigma, X)$ where $\to_{\mathcal{R}}$ is defined as follows: When $t \in T(\Sigma, X)$ is a term, let $Pos(t)$ be the set of its positions. For $\rho \in Pos(t)$, let $t|_\rho$ be the subterm of t at position ρ, and let $t[s]_\rho$ denote the replacement of the subterm at position ρ in t by s. The set of all substitutions is denoted by Sub.

Definition 8 (Reduction relation). *To a probabilistic rewrite system $\mathcal{R}$ is associated the following PARS $(T(\Sigma, X), \to)$ over terms: $t \to_{\mathcal{R}} \mu$ iff there is a rule $(g, M) \in \mathcal{R}$, some position $p \in Pos(t)$, some substitution $\sigma \in Sub$, such that $t|_p = \sigma(g)$, and, for all t', $\mu(t') = \sum_{d | t' = t[\sigma(d)]_p} M(d)$.*

For example, a probabilistic rewrite rule can be $f(x, y) \mapsto \begin{cases} g(a) : 1/2 \\ y \quad\; : 1/2 \end{cases}$, where right hand side denotes the distribution with value $1/2$ on $g(a)$ and value $1/2$ on y. Then $f(b, c)$ rewrites to $g(a)$ with probability $1/2$, and to c with probability $1/2$. Now, $f(b, g(a))$ rewrites to $g(a)$ with probability 1.

Example 1. Consider[3] the following probabilistic rewrite system, with two rules R_1 and R_2 (of course, we assume $0 \leq p_1 \leq 1$, $0 \leq p_2 \leq 1$).

$$\begin{array}{ll} X \odot (Y \oplus Z) & \to \begin{cases} (X \odot Y) \oplus (X \odot Z) : p_1 \\ X \odot (Y \oplus Z) \qquad\quad : 1 - p_1 \end{cases} \\ ((X \odot Y) \oplus (X \odot Z)) \oplus X & \to \begin{cases} (X \odot (Y \oplus Z)) \oplus X : p_2 \\ X \odot ((Y \oplus Z) \oplus X) : 1 - p_2 \end{cases} \end{array}$$

Consider the polynomial interpretation of symbols $\{\oplus, \odot\}$ given by $[X \oplus Y] = 2[X] + [Y] + 1$ and $[X \odot Y] = [X] * [Y]$, where $[P] \in \mathbb{N}[X_1, \ldots, X_n]$ denotes the polynomial interpretation of a term P of arity n.

Fix some integer $n_0 \geq 2$, yet to be determined. The set of integers $\geq n_0$ is preserved by the polynomials $[P]$. Consider function V that maps any term P to $[P](n_0, \cdots, n_0)$. Denote also $V(P)$ by $\{P\}$.

We have $[X \odot (Y \oplus Z)] = 2[X][Y] + [X][Z] + [X]$, $[(X \odot Y) \oplus (X \odot Z)] = 2[X][Y] + [X][Z] + 1$, and hence $\{X \odot (Y \oplus Z)\} = 2\{X\}\{Y\} + \{X\}\{Z\} + \{X\}$, $\{(X \odot Y) \oplus (X \odot Z)\} = 2\{X\}\{Y\} + \{X\}\{Z\} + 1$, and the drift of the first rule (see [5]) is given by $\Delta_{R_1} V(X \odot (Y \oplus Z)) = p_1 \times \{(X \odot Y) \oplus (X \odot Z)\} + (1 - p_1)\{X \odot (Y \oplus Z)\} - \{X \odot (Y \oplus Z)\} = p_1 \times (1 - \{X\})$. This is negative, and any

[3] Example obtained by modifying an example discussed in [2] about polynomial interpretations. As far as we know, this is the first time that a polynomial interpretation is used to prove termination of a probabilistic system (the examples from [5] used only linear interpretation functions).

substitution on X can only decrease it: R_1 is substitution decreasing following the terminology of [5].

Considering the second rule, we have $[((X \odot Y) \oplus (X \odot Z)) \oplus X] = 4[X][Y] + 2[X][Z] + [X] + 3$, $[(X \odot (Y \oplus Z)) \oplus X] = 4[X][Y] + 2[X][Z] + 3[X] + 1$, $[X \odot ((Y \oplus Z) \oplus X)] = 2[X][Y] + 2[X][Z] + X + 2$, and hence $\Delta_{R_2} V(((X \odot Y) \oplus (X \odot Z)) \oplus X) = p_2 \times \{(X \odot (Y \oplus Z)) \oplus X\} + (1 - p_2)\{X \odot ((Y \oplus Z) \oplus X)\} - \{((X \odot Y) \oplus (X \odot Z)) \oplus X\} = 2(p_2 - 1)\{X\}\{Y\} + 2p_2\{X\} - p_2 - 1$. This drift is not necessarily negative: in particular for $p_2 = 1$, it is positive. However, assume $p_2 < 1$. If we take, $n_0 \geq p_2/(1 - p_2)$, we can be sure that it becomes negative, since $2(p_2 - 1)\{X\}\{Y\} \leq -2p_2\{X\}$. For such an n_0, it is substitution decreasing.

Now, observing the form of the interpretation of symbols $\{\oplus, \odot\}$, which are linear in each of their variables with integer positive coefficients, a context can only decrease a drift. We get that the probabilistic rewrite system is +a.s. terminating for $p_2 < 1$.

This is a fortiori true for the following system, since the drift of the third rule is $-1 - 2\{X\}$, and hence negative.

Example 2. Consider the following probabilistic rewrite system, with three rules R_1, R_2, R_3:

$$
\begin{array}{ll}
X \odot (Y \oplus Z) & \rightarrow \begin{cases} (X \odot Y) \oplus (X \odot Z) : p_1 \\ X \odot (Y \oplus Z) \qquad\quad : 1 - p_1 \end{cases} \\
((X \odot Y) \oplus (X \odot Z)) \oplus X & \rightarrow \begin{cases} (X \odot (Y \oplus Z)) \oplus X : p_2 \\ X \odot ((Y \oplus Z) \oplus X) : 1 - p_2 \end{cases} \\
(X \oplus Y) \oplus Z & \rightarrow \{ X \oplus (Y \oplus Z) : 1
\end{array}
$$

3 Positive Almost Sure Termination Under Strategies

Positive almost sure termination means that for all starting term the mean number of rewrite steps to reach a terminal state is finite under any policy/strategy. In particular, non termination can happen with a single very specific strategy.

In many examples, one is often tempted not to consider arbitrary strategies, but to restrict to a subset of strategies. Whatever the considered class of strategies is, the following notion is rather natural.

Definition 9 (Positive Almost Sure Termination Under Strategies). *Fix a class Φ of strategies (i.e. policies);*

A PARS $\mathcal{A} = (A, \rightarrow)$ will be said positively almost surely (+a.s.) terminating under Φ if for all strategy (i.e. policy) $\phi \in \Phi$, for all states $a \in A$, the mean number of reduction steps before termination under ϕ starting from a, denoted by $T[a, \phi]$, is finite.

Example 3. Consider the following probabilistic rewrite system, with two rules.

$$
\begin{array}{l}
a \rightarrow \{a : 1 \\
a \rightarrow \{b : 1
\end{array}
$$

This system is clearly not (almost surely) terminating , since there is the infinite derivation $a \rightarrow a \rightarrow \cdots a \rightarrow \cdots$.

However, it is +a.s. terminating under Markovian non-deterministic[4] randomized strategies: indeed, in state a, such a strategy selects either the first rule with probability p_1, or the second with probability $1 - p_1$, for some fixed $p_1 < 1$. The system is then equivalent to the probabilistic rewrite system

$$a \rightarrow \begin{cases} a : p_1 \\ b : 1 - p_1 \end{cases}$$

whose positive almost sure termination can be established easily, for example using previous theorem and $V(a) = 1$, $V(b) = 0$.

Example 4. Consider the following probabilistic rewrite system, with two rules named *red* and *green*: see Figure 1.

$$s(x) \rightarrow \begin{cases} x & : p_1 \\ s(s(x)) : 1 - p_1 \end{cases}$$
$$s(x) \rightarrow \begin{cases} x & : p_2 \\ s(s(x)) : 1 - p_2 \end{cases}$$

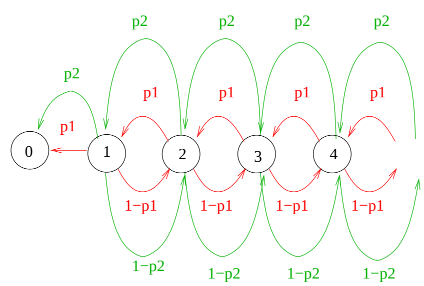

Figure 1. Example 4.

The red (respectively: green) rule[5] is easily shown to be +a.s. terminating iff $p_1 > 1/2$ (resp. $p_2 > 1/2$).

Suppose that $p_1 < 1/2$, $p_2 > 1/2$. The whole system is not +a.s. terminating: consider the strategy that always selects the red rule.

However, it is +a.s. terminating under the strategy that always selects the green rule.

Intuitively, in a more general case, its +a.s. termination depends on the ratio of selection of the red versus green rule. Indeed, if we focus on Markovian

[4] we want to avoid $p_1 = 1$.

[5] That is to say: the probabilistic rewrite system made of this rule alone.

randomized strategies that select the red (respectively green) rule with a fixed probability p (resp. $1-p$), the whole system is equivalent to

$$s(x) \to \begin{cases} x & : p_1 * p + p_2 * (1-p) \\ s(s(x)) & : (1-p_1) * p + (1-p_2) * (1-p) \end{cases}$$

which is easily shown to be +a.s. terminating iff $p_1 * p + p_2 * (1-p) > 1/2$, i.e. $p < (1-2p_2)/(2(p_2-p_1))$.

Example 5. Consider the following probabilistic rewrite system, over signature $\Sigma = \{A, B, C\}$, with four rules, where we assume $p_1 > 0$, $p_2 > 0$.

$$\begin{array}{l} A \to \begin{cases} B : p_1 \\ A : 1-p_1 \end{cases} \\ B \to \begin{cases} A : p_2 \\ B : 1-p_2 \end{cases} \\ A \to \{C \quad : 1 \\ B \to \{C \quad : 1 \end{array}$$

We have only states A and B, and in each of these states, a strategy can either select the rule among the two first that applies or the rule among the two last that applies. It is easy to see that with probability one, an infinite derivation is made of a sequence of A and B, each of them appearing infinitely often.

This probabilistic rewrite system is not +a.s. terminating: consider the strategy ϕ_∞ that always excludes the second possibility (i.e. never choose third or fourth rule).

However, it is clearly +a.s. terminating under Φ, for any class Φ that does not contain this specific strategy ϕ_∞.

This example illustrates that one may want to restrict to fair strategies, for some or one's preferred notion of fairness: in this example, since third and fourth rule can fire infinitely often, one may want that they fire at least once (or with positive probability).

In literature, several notions of fairness have been introduced: see [23,24,27,10] and references in the introduction of this paper. Termination of probabilistic systems under fairness constraints has been investigated, in particular in [18] for probabilistic finite state systems, and in [17] for probabilistic infinite state systems.

Next section will be devoted to provide techniques to prove positive almost sure termination of a probabilistic rewrite system under strategies. These results can be applied with classes of strategies constrained by several of these notions of fairness. The following results can also be seen as an extension of the two papers [18,17] to deal with +a.s. termination (and not only almost sure termination).

4 Proving +a.s. Termination Under Strategies

A slight generalization of Theorem 1 yields rather directly:

Theorem 2. *Fix a class of strategies Φ.*

A PARS $\mathcal{A} = (A, \to)$ is +a.s. terminating under Φ if there exist some function $V : A \to \mathbb{R}$, with $\inf_{i \in A} V(i) > -\infty$, and some $\epsilon > 0$, such that, for all realizable non-terminal history $h = a_0 a_1 \cdots a_n$, for all $\phi \in \Phi$, the drift in h according to ϕ defined by

$$\Delta_\phi V(h) = \sum_i \phi(h)(i) V(i) - V(a_n)$$

satisfies

$$\Delta_\phi V(h) \leq -\epsilon.$$

Fortunately, we can do better in many cases.

Consider a PARS $\mathcal{A} = (A, \to)$. Assume that $Dist(\mathcal{A})$ (see Definition 1) can be partitioned into finitely many subsets $Dist(\mathcal{A}) = D_1 \cup D_2 \ldots \cup D_k$. Intuitively, when $\mathcal{A}$ is corresponding to a PARS associated to some probabilistic rewrite system with k probabilistic rewrite rules $R_1, \ldots, R_k$, each D_i corresponds to rewrite rule R_i: D_i is made of distributions μ obtained by varying position p, and substitution σ in the distribution of rule R_i, according to Definition 8.

We assume that for any strategy $\phi \in \Phi$, $\phi^{-1}(D_i)$ is measurable. The expectation of a random variable X is denoted by $E[X]$.

Definition 10 (Next Selection of a Rule). *Fix some D_i.*

Fix some deterministic policy ϕ and some realizable non-terminal history $h = a_0 a_1 \cdots a_n$. Let $(\pi_i)_{i \in \mathbb{N}}$ be a derivation starting from h: $(\pi_i)_{i \in \mathbb{N}}$ is a stochastic sequence as in Definition 4 with $\pi_0 = a_0, \cdots, \pi_n = a_n$.

Let τ be the random variable denoting the first index greater than n at which D_i is selected, or a terminal state is reached (set $\tau = \infty$ if there is no such index). I.e. $\tau = m$ iff $\phi(\pi_0, \cdots, \pi_m) \in D_i$, and $\phi(\pi_0, \cdots, \pi_{m'}) \notin D_i$ for $n < m' < m$, or $\pi_m = \perp$ and $\pi_{m'} \neq \perp$ for $n < m' < m$.

Let $\tau_{D_i,\pi,\phi,h}$ denote the τ for the corresponding D_i, π, ϕ and h.

Each random variable $\tau_{D_i,\pi,\phi,h}$ is a stopping time with respect to derivation π (see e.g. [9]): it is a random variable taking its value in $\mathbb{N} \cup \{\infty\}$, such that for all integers $m \geq 0$, the event $\{\tau = m\}$ can be expressed in terms of $\pi_0, \pi_1, \ldots, \pi_m$.

Remark 1. One must understand that even if the policy is deterministic, and hence not depending on any random choice, each $\tau_{D_i,\pi,\phi,h}$ is random. Indeed, when $h = a_0 \cdots a_n$ is fixed, the choice of a_{n+1} is made according to distribution $\phi(a_0 \cdots a_n)$, and hence random; the choice of a_{n+2} is then made according to distribution $\phi(a_0 \cdots a_n a_{n+1})$, and hence random. And so on. The event D_i is selected or a terminal state is reached at time n is then random.

Definition 11 (Bounded Mean Selection). *A class of strategy Φ has bounded mean selection $\alpha \in \mathbb{R}$ for D_i, if for any strategy $\phi \in \Phi$, for any history, the expected time to wait before reaching a final state or selecting a rule from D_i is less than α. I.e. for any realizable non-terminal history $h = a_0 \cdots a_n$, for any policy $\phi \in \Phi$, for any derivation π starting from h, $\tau_{D_i,\pi,\phi,h}$ has a finite mean with*

$$E[\tau_{D_i,\pi,\phi,h}] \leq n + \alpha.$$

Observe that a variable taking values in $\mathbb{N} \cup \{\infty\}$ with a finite mean is necessarily almost surely finite: in other words, when the conditions of the previous definition hold, one knows that almost surely starting from any history h, one reaches either a final state, or one selects a rule from D_i.

Definition 12 (Expected Value of V At Time τ). *Let $V : A \to \mathbb{R}$ be some function. Let $\tau \in \mathbb{N} \cup \{\infty\}$ denotes some stopping time with respect to derivation π, which is almost surely finite: $P(\tau < \infty) = 1$. Fix some policy ϕ, and a corresponding derivation $(\pi_i)_{i \in \mathbb{N}}$.*

We denote by $E_\tau V$ the expected value of V at time τ: formally

$$E_\tau V = E[V(\pi_\tau)]$$

when it exists.

We claim:

Theorem 3 (Almost Sure Termination Under Strategy). *Fix a class of strategies Φ.*

A PARS $\mathcal{A} = (A, \to)$ is almost surely terminating under strategies Φ if there exist some function $V : A \to \mathbb{R}$, with $\inf_{i \in A} V(i) > -\infty$, some $\epsilon > 0$, and some D_i such that for all strategy $\phi \in \Phi$, for all realizable non-terminal history $h = a_0 \dots a_n$, for all derivation π starting from h,

1. *the stopping time $\tau_{D_i,\pi,\phi,h}$ is almost surely finite,*
2. *and*

$$E_{\tau_{D_i,\pi,\phi,h}} V \leq V(a_n) - \epsilon.$$

This follows from the following result from Martingale theory: See [11] for a proof (1_A denotes the characteristic function of a set A).

Proposition 1. *Let $(\Omega, \mathcal{F}, P)$ be a given probability space, and $\{\mathcal{F}_n, n \geq 0\}$ an increasing family of σ-algebra.*

Consider a sequence $(S_i)_{i \in \mathbb{N}}$ of real non-negative random variables, such that S_i is $\mathcal{F}_i$-measurable, for all i. Assume S_0 to be constant, w.l.o.g.

Denote by τ the $\mathcal{F}_n$-stopping time representing the epoch of the first entry into $[0, C]$, for some C: $\tau = \inf\{i \geq 1, S_i \leq C\}$.

Introduce the stopped sequence $S'_i = S_{min(i,\tau)}$.

Assume $S_0 > C$, and for some $\epsilon > 0$, and for all $n \geq 0$, almost surely

$$E[S'_{i+1} | \mathcal{F}_{n-1}] \leq S'_i - \epsilon 1_{\tau > i}. \tag{1}$$

Then:

- *Almost surely τ is finite.*
- *$E[\tau] < S_0/\epsilon$.*

Proof (of Theorem 3). Replacing function V by $V+K$ for some constant K if needed, we can assume without loss of generality that $V(a) \geq 2\epsilon$ for all $a \in A$. Extend function V on $A \cup \{\perp\}$ by $V(\perp) = 0$.

Fix a strategy $\phi \in \Phi$, a realizable non-terminal history h, and a derivation $(\pi_i)_{i\in\mathbb{N}}$ starting from h.

From Condition 1., one can build a sequence of random functions $(\psi_n)_{n\in\mathbb{N}}$ such that almost surely, for all $n > 1$, either $\pi_{\psi(n)} = \perp$, or D_i is selected at rank $\psi(n)$. Indeed: Take $\psi(0) = 0$; when $\psi(n)$ is built, build $\psi(n+1)$ as $\psi(n)$ if $\pi_{\psi(n)} = \perp$ and as $\psi(n) + \tau_{D_i,\pi,\phi,h}$ otherwise.

Consider the increasing family of σ-algebra $\mathcal{F}_n$ where $\mathcal{F}_n$ is the σ-algebra generated by $\pi_0, \cdots, \pi_n$. Condition 2. implies almost surely $E[S'_{n+1}|\mathcal{F}_{n-1}] \leq S'_n - \epsilon 1_{\pi_{\psi(n)} \neq \perp}$, where $S'_n = V(\pi_{\psi(n)})$ for all n. By Proposition 1 above with $C = \epsilon$, almost surely there must exist some n with $\pi_n = \perp$.

In other words, the PARS is almost surely terminating under Φ.

Remark 2. Previous hypotheses yield almost sure termination, but not positive almost sure termination. Indeed, the proof build a subsequence of indexes $\psi(n)$ yielding almost surely to termination. But there is no reason that $\psi(n+1) - \psi(n)$ stay bounded, and hence the original derivation can be non positively almost surely terminating (such an example is easy to build).

Actually, weaker conditions entailing almost sure termination have been derived in [17]: in particular ϵ can be taken as 0. However, for +a.s. almost sure derivation, we claim:

Theorem 4 (+ A.S. Termination Under Strategy). *Fix a class of strategies Φ.*

A PARS $\mathcal{A} = (A, \rightarrow)$ is +a.s. terminating under strategies Φ if there exist some function $V : A \rightarrow \mathbb{R}$, with $\inf_{i\in A} V(i) > -\infty$, some $\epsilon > 0$, and some D_i such that Φ has bounded mean selection for D_i, and such that for all strategy $\phi \in \Phi$, for all realizable non-terminal history $h = a_0 \ldots a_n$, for all derivation π starting from h,

$$E_{\tau_{D_i,\pi,\phi,h}} V \leq V(a_n) - \epsilon.$$

Proof. By previous discussion, the fact that Φ has bounded mean selection for D_i entails Condition 1. of previous theorem, and hence we have almost sure termination. Even if we did not mention it, the application of Proposition 1 in the proof of previous Theorem also yields that the random variable N giving the smallest n with $\pi_n = \perp$ has a finite mean with $E[N] \leq V(a_n)/\epsilon$.

Now, since Φ has bounded mean selection α for some $\alpha > 0$, we can bound $E[\psi(N)]$ by $\alpha V(a_0)/\epsilon$ using following Lemma, whose proof can easily be established (for example by adapting the proof of Wald's Lemma in [9]).

Lemma 1. *Consider a stochastic sequence $(X_i)_{i\in\mathbb{N}}$ taking non-negative values. Let N be an integer-valued random variable, with a finite expectation. Assume there exists some constant M such that $0 \leq X_{n+1} - X_n \leq M$ almost surely for all n. The random variable X_N has an expectation bounded by $E[X_0] + M * E[N]$.*

5 Applications

We first derive one simple case:

Proposition 2. *Consider a PARS $\mathcal{A} = (A, \to)$ so that there exists $V : A \to \mathbb{R}$, with $\inf_{i \in A} V(i) > -\infty$, some (possibly positive) α, some $\epsilon > 0$, such that $Dist(\mathcal{A})$ can be partitioned into $Dist(\mathcal{A}) = D_1 \cup D_2$ such that for all $a \to \mu$, we have*

1. *$\Delta_\mu V(a) \leq \alpha$ whenever $\mu \in D_1$.*
2. *$\Delta_\mu V(a) \leq -\epsilon$ whenever $\mu \in D_2$.*

Assume that a rule of the form $a \to \mu$, with $\mu \in D_1$ never lead to a terminal: for all $a \to \mu$, $\mu \in D_1$, for all a' with $\mu(a') > 0$, a' is not a terminal.

Assume that ϕ selects D_2 at least once every k steps for some constant k: for any $h = a_0 \cdots a_n$, for any $\phi \in \Phi$, for any π, we assume that $\tau_{D_2,\pi,\phi,h}$ exists and satisfies $\tau_{D_2,\pi,\phi,h} \leq n + k$.

Assume that $(k-1)\alpha - \epsilon < 0$.

Then $\mathcal{A}$ is +a.s. terminating under strategies Φ.

Proof. It is easy to see that we always have $E_{\tau_{D_2,\pi,\phi,h}} V \leq V(a_n) + (k-1)\alpha - \epsilon$ in this case: Indeed, a derivation starting from h must either reach a terminal or lead to a state where D_2 is selected. In any case, the last applied rule will be a rule from D_2, and hence V will decrease in mean of at least ϵ, after the at most $k-1$ first rules that can make it increase in mean of at most $(k-1)\alpha$. We can then apply previous theorem.

Example 6. Consider the following probabilistic rewrite system, with three rules R_1, R_2', R_3:

$$
\begin{array}{ll}
X \odot (Y \oplus Z) & \to \begin{cases} (X \odot Y) \oplus (X \odot Z) : p_1 \\ X \odot (Y \oplus Z) \qquad\quad : 1 - p_1 \end{cases} \\
((X \odot Y) \oplus (X \odot Z)) \oplus X & \to \begin{cases} (X \odot (Y \oplus Z)) \oplus X \qquad\quad : p_2 \\ ((X \odot Y) \oplus (X \odot Z)) \oplus X : 1 - p_2 \end{cases} \\
(X \oplus Y) \oplus Z & \to \begin{cases} X \oplus (Y \oplus Z) : 1 \end{cases}
\end{array}
$$

This probabilistic rewrite system is not positively almost surely terminating. Indeed, for the policy which always apply the first two rules and never the third, we have an infinite derivation with terms $((X \odot Y) \oplus (X \odot Z)) \oplus X$ and $(X \odot (Y \oplus Z)) \oplus X$, each of them appearing almost surely infinitely often.

The drift of the rules R_1 and R_3 have been computed in Example 2. Now, the drift of the rule R_2' is $\Delta_{R_2'} V(((X \odot Y) \oplus (X \odot Z)) \oplus X) = 2p_2 \times (\{X\} - 1)$, and hence positive.

If we choose a policy ϕ with a bounded mean selection for the rewrite rule R_3, and if ϕ always reduce the term of a cycle $((X \odot Y) \oplus (X \odot Z)) \oplus X \to (X \odot (Y \oplus Z)) \oplus X$ until it can be broken by firing rule R_3, then conditions of Theorem 4 are satisfied, because, for all histories $h = a_0, \ldots, a_n$ such that a_n contains a subterm which is an instance of $((X \odot Y) \oplus (X \odot Z)) \oplus X$, then $E_{\tau_{D_3},\pi,\phi,h} V \leq V(a_n) - 2 \times \{X\} - 1$.

Example 7. Let's now consider the following term rewrite system, coming from the model of [8] of a simulator for the CSMA-CA protocol [1]. The rules rewrite lists of couples. Each couple is made of two positive integers. The *sort* operator triggers a rule based sort algorithm, which sorts in increasing order the list in function of the value of the first field. The first rule will take the head of the list, replace the first field by a random value between 1 and p following an uniform law with probability μ and decrease the value of the second field with probability $1 - \mu$.

$$
\begin{array}{ll}
(\Delta t, n+1), \ldots, (\Delta_k, n_k) & \rightarrow \begin{cases} (U(1,\ldots,p), n+1), \ldots, (\Delta_k, n_k) & : \mu \\ (U(1,\ldots,p), n), \ldots, (\Delta_k - \Delta t, n_k) & : 1-\mu \end{cases} \\
(\Delta t, n+1), \ldots, (\Delta_k, n_k) & \rightarrow sort((\Delta t, n+1), \ldots, (\Delta_k, n_k)) \\
sort((\Delta_t, n_t), X) & \rightarrow sort1((\Delta_t, n_t), nil, X) \\
sort1((\Delta_t, n_t), l, (\Delta'_t, n'_t).X) & \rightarrow sort1(((\Delta_t, n_t), l.(\Delta'_t, n'_t), X)) \text{ If } \Delta_t < \Delta'_t \\
sort1((\Delta_t, n_t), l, (\Delta'_t, n'_t).X) & \rightarrow sort1(((\Delta'_t, n'_t), l.(\Delta_t, n_t), X)) \text{ If } \Delta_t > \Delta'_t \\
sort1((\Delta_t, n_t), l, (\Delta'_t, n'_t).nil) & \rightarrow (\Delta_t, n_t).sort(l.(\Delta'_t, n'_t)) \text{ If } \Delta'_t > \Delta_t \\
sort1((\Delta_t, n_t), l, (\Delta'_t, n'_t).nil) & \rightarrow (\Delta'_t, n'_t).sort(l.(\Delta_t, n_t)) \text{ If } \Delta'_t < \Delta_t
\end{array}
$$

where, X, l are some lists of couples of integers, the operator "."denotes the concatenation of lists and *nil* is the empty list. $U(1, \ldots, p)$ is a random integer variable following an uniform law on $\{1, \ldots, p\}$. n and $n_{i i \in 1, \ldots, k}$ are non negative integers.

This PRS is easily seen not +a.s. terminating: For example the first two rules always apply on every list or sublists.

Now let's build a policy under which the PRS positively almost surely terminates. Let's start with a_0 a list of length n, $\phi(a_0)$ is the rule that rewrites a_0 to $sort(a_0)$. The length of the sorting process is $n(n-1)$, and the policy ϕ chooses only the rules coding the sort algorithm during the sort process. If the first element of the list has a zero second field, there's no rule matching this list and this term is terminal. Otherwise, the policy ϕ will choose again the rule that triggers the sort of the list, and later apply the rule number one, and so one since no terminal state is reached.

To show this system is +a.s. terminating, let's consider the function $V : T(\Sigma, X) \rightarrow \mathbb{N}$ computing the sum of the second field of each element of a list, and apply Proposition 2.

An alternative proof is the following: We can apply Theorem 4, because ϕ has bounded mean selection for the first rule rewrite relation D_1, because such a rule is triggered between two sorts of length $n(n-1)$ and $E_{\tau_{D_1}, \pi, \phi, h} V = V(a_n) + \mu - 1$, because V does not change during the sorting process since the values of the second field are not touched, and the only variation of the mean is induced by the rule D_1 whose drift is $\mu - 1$. V, as the sum of positive value, is lower bounded.

6 Conclusion and Future Work

In this paper, we introduced positive almost sure termination under strategies, and we provide sufficient criteria to prove positive almost sure termination of a given probabilitic rewrite system under strategies.

We plan to apply our techniques on industrially motivated examples of bigger size. It may be possible to weaken the hypotheses of our theorems since they mainly use a special case of Proposition 1. As mentioned in the introduction, any technique to deal with probabilistic systems, must work for classical ones, since probabilities can be 0/1. The classical (non-probabilistic) counterpart of our framework for proving termination under strategies is very poor: the question of understanding which of the techniques from literature for non-probabilistic systems can be extended to deal with probabilistic systems seems fascinating.

References

1. Ieee csma/ca 802.11 working group home page. http://www.ieee802.org/11/.
2. Franz Baader and Tobias Nipkow. *Term Rewriting and All That*. Cambridge University Press, 1998.
3. Gianfranco Balbo. Introduction to stochastic Petri nets. *Lecture Notes in Computer Science*, 2090:84, 2001.
4. P. Borovanský, C. Kirchner, and H. Kirchner. Controlling Rewriting by Rewriting. In J Meseguer, editor, *Proceedings of 1st International Workshop on Rewriting Logic*, volume 4, Asilomar (CA, USA), September 1996. Electronic Notes in Theoretical Computer Science.
5. Olivier Bournez and Florent Garnier. Proving positive almost sure termination. In Jürgen Giesl, editor, *16th International Conference on Rewriting Techniques and Applications (RTA'2005)*, volume 3467 of *Lecture Notes in Computer Science*, page 323, Nara, Japan, 2005. Springer.
6. Olivier Bournez and Mathieu Hoyrup. Rewriting logic and probabilities. In Robert Nieuwenhuis, editor, *Rewriting Techniques and Applications, 14th International Conference, RTA 2003, Valencia, Spain, June 9-11, 2003, Proceedings*, volume 2706 of *Lecture Notes in Computer Science*, pages 61–75. Springer, June 2003.
7. Olivier Bournez and Claude Kirchner. Probabilistic rewrite strategies: Applications to ELAN. In Sophie Tison, editor, *Rewriting Techniques and Applications*, volume 2378 of *Lecture Notes in Computer Science*, pages 252–266. Springer-Verlag, July 22-24 2002.
8. Olivier Bournez Florent Garnier Claude Kirchner. Termination in finite mean time of a csma/ca rule-based model. Technical report, LORIA, Nancy, 2005.
9. Pierre Brémaud. *Markov Chains, Gibbs Fields, Monte Carlo Simulation, and Queues*. Springer-Verlag, New York, 2001.
10. Luca de Alfaro. *Formal Verification of Probabilistic Systems*. PhD thesis, Stanford University, 1998.
11. Guy Fayolle, Vadim A. Malyshev, and Mikhail V. Menshikov. *Topics in constructive theory of countable Markov chains.* Cambridge University Press, 1995.
12. Olivier Fissore, Isabelle Gnaedig, and Hélène Kirchner. Simplification and termination of strategies in rule-based languages. In *PPDP '03: Proceedings of the 5th ACM SIGPLAN international conference on Principles and practice of declaritive programming*, pages 124–135, New York, NY, USA, 2003. ACM Press.

13. Olivier Fissore, Isabelle Gnaedig, and Hélène Kirchner. A proof of weak termination providing the right way to terminate. In *Theoretical Aspects of Computing – ICTAC 2004: First International Colloquium*, Lecture Notes in Computer Science. Springer, 2004.
14. F. G. Foster. On the stochastic matrices associated with certain queuing processes. *The Annals of Mathematical Statistics*, 24:355–360, 1953.
15. Thom Frühwirth, Alexandra Di Pierro, and Herbert Wiklicky. Toward probabilistic constraint handling rules. In Slim Abdennadher and Thom Frühwirth, editors, *Proceedings of the third Workshop on Rule-Based Constraint Reasoning and Programming (RCoRP'01)*, Paphos, Cyprus, December 2001. Under the hospice of the International Conferences in Constraint Programming and Logic Programming.
16. H. Hansson. *Time and Probability in Formal Design of Distributed Systems.* Series in Real-Time Safety Critical Systems. Elsevier, 1994.
17. Sergiu Hart and Micha Sharir. Concurrent probabilistic programs, or: How to schedule if you must. *SIAM Journal on Computing*, 14(4):991–1012, November 1985.
18. Sergiu Hart, Micha Sharir, and Amir Pnueli. Termination of probabilistic concurrent program. *ACM Transactions on Programming Languages and Systems*, 5(3):356–380, July 1983.
19. Thomas Hérault, Richard Lassaigne, Frédéric Magniette, and Sylvain Peyronnet. Approximate probabilistic model checking. In *VMCAI*, pages 73–84, 2004.
20. Marta Kwiatkowska, Gethin Norman, and David Parker. PRISM: Probabilistic symbolic model checker. *Lecture Notes in Computer Science*, 2324:200, 2002.
21. Marta Z. Kwiatkowska. Model checking for probability and time: from theory to practice. In *LICS*, page 351, 2003.
22. Kumar Nirman, Koushik Sen, Jose Meseguer, and Gul Agha. A rewriting based model for probabilistic distributed object systems. In *Proceedings of 6th IFIP International Conference on Formal Methods for Open Object-based Distributed Systems (FMOODS'03)*, volume 2884 of *Lecture Notes in Computer Science*, pages 32–46. Springer, Paris, France, November 2003.
23. Amir Pnueli. On the extremely fair treatment of probabilistic algorithms. In *STOC: ACM Symposium on Theory of Computing (STOC)*, 1983.
24. Amir Pnueli and Lenore D. Zuck. Probabilistic verification. *Information and Computation*, 103(1):1–29, March 1993.
25. Martin L. Puternam. *Markov Decision Processes - Discrete Stochastic Dynamic Programming.* Wiley series in probability and mathematical statistics. John Wiley & Sons, 1994.
26. William H. Sanders and John F. Meyer. Stochastic activity networks: Formal definitions and concepts. *Lecture Notes in Computer Science*, 2090:315, 2001.
27. Moshe Y. Vardi. Automatic verification of probabilistic concurrent finite-state programs. In *26th Annual Symposium on Foundations of Computer Science*, pages 327–338, Portland, Oregon, 21–23 October 1985.