Disproving Reachability in Probabilistic Term Rewriting

Jan-Christoph Kassing[1][2], Moritz Leven Rosarius[1][3],
Henri Nagel[1][4], and Jürgen Giesl[1][5]

[1] RWTH Aachen University, Aachen, Germany
[2] kassing@cs.rwth-aachen.de
[3] leven@cs.rwth-aachen.de
[4] henri.nagel@rwth-aachen.de
[5] giesl@informatik.rwth-aachen.de

Abstract

Reachability is a central question in term rewriting: can a given target term (e.g., an error state) be reached from a start term? It is also an important property in confluence analysis, and corresponding tools compete in the annual confluence competition. An interesting generalization of this problem is handling programs that can make random choices during execution. For such probabilistic programs, reachability becomes a *quantitative* property instead of a *qualitative* one: instead of asking *whether* the target is reachable, one asks *with which probability* it is reached. We lift reachability analysis from ordinary term rewriting to probabilistic term rewrite systems. To do so, we formalize the maximal probability of reaching a target term and adapt two techniques for analyzing reachability (based on symbol transition graphs and on term orderings) to compute upper bounds on this probability.

1 Introduction

Term rewrite systems (TRSs) are a fundamental model for program execution, transformation, and verification. A TRS consists of a finite number of rewrite rules $\ell \to r$, indicating that parts of a term that are matched by ℓ can be replaced by the reduct r. A central question is the *reachability problem* "$s \to_{\mathcal{R}}^* t$?": can a target t (e.g., an error state) be reached from a start term s? Often, s and t are templates, and one asks whether t can be reached for *some* ground instantiation σ of their variables, i.e., "$\exists \sigma.\, s\sigma \to_{\mathcal{R}}^* t\sigma$?". Reachability for TRSs has been studied extensively, e.g., [6, 8, 9, 10].

We lift reachability analysis to probabilistic term rewrite systems (PTRSs), where the rule selection remains non-deterministic, but after selecting a rule, the reduct is chosen probabilistically. We formalize the maximal probability of reaching a target (Sect. 2) and adapt two existing techniques from the non-probabilistic setting to compute upper bounds on it. This yields upper bounds on the maximal reachability probability: we can prove that a target is reached with probability at most p, and thereby disprove reachability claims (e.g., almost-sure reachability whenever $p < 1$, or reachability at all whenever $p = 0$). First, the symbol transition graph turns into an over-approximating Markov decision process (MDP), whose reachability probability can be computed with tools like Storm [7] (Sect. 3). Second, we lift term orderings for disproving reachability to the probabilistic setting (Sect. 4). We plan to implement both techniques in AProVE [5] and to evaluate them experimentally.

2 Preliminaries

We assume the reader to be familiar with ordinary term rewriting as presented, e.g., in [1]. In the probabilistic setting, a single reduction step leads to a *finite multi-distribution* over possible

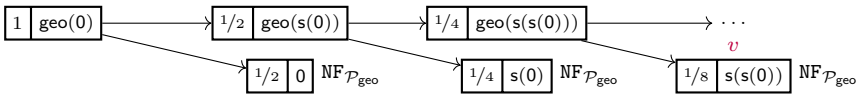

Figure 1: A $\mathcal{P}_{\mathsf{geo}}$-RT $\mathfrak{T}$ with $\mathrm{root}(\mathfrak{T}) = \mathsf{geo}(\mathsf{0})$. The node v is labeled by the probability $p_v = 1/8$ and the term $a_v = \mathsf{s}(\mathsf{s}(\mathsf{0}))$.

results rather than to a single result. A finite *multi-distribution* μ on a set $A \neq \varnothing$ is a finite multiset of pairs ${}^{p}a$, where $0 < p \leq 1$ is a probability and $a \in A$, such that $\sum_{{}^{p}a \in \mu} p = 1$. Let $\mathrm{FDist}(A)$ denote the set of all finite multi-distributions on A. For $\mu \in \mathrm{FDist}(A)$, its *support* is the multiset $\mathrm{Supp}(\mu) = \{a \mid {}^{p}a \in \mu\}$. A *probabilistic abstract reduction system* (PARS) is a pair $(A, \to)$ such that $\to \subseteq A \times \mathrm{FDist}(A)$.

Let $\mathcal{T} = \mathcal{T}(\Sigma, \mathcal{V})$ be the set of terms over a finite signature Σ and variable set $\mathcal{V}$. A *probabilistic term rewrite rule* $\ell \to \mu$ is a pair $(\ell, \mu) \in \mathcal{T} \times \mathrm{FDist}(\mathcal{T})$ such that $\ell \notin \mathcal{V}$ and $\mathcal{V}(r) \subseteq \mathcal{V}(\ell)$ for all $r \in \mathrm{Supp}(\mu)$, and a *probabilistic TRS* (PTRS) is a finite set of probabilistic term rewrite rules. Similar to TRSs, a PTRS $\mathcal{P}$ induces a PARS $(\mathcal{T}, \to_{\mathcal{P}})$ with $s \to_{\mathcal{P}} \{{}^{p_1}t_1, \ldots, {}^{p_k}t_k\}$ if there is a position $\pi \in \mathrm{Pos}(s)$, a rule $\ell \to \{{}^{p_1}r_1, \ldots, {}^{p_k}r_k\} \in \mathcal{P}$, and a substitution σ such that $s|_\pi = \ell\sigma$ and $t_j = s[r_j\sigma]_\pi$ for all $1 \leq j \leq k$. Often, we simply refer to $\mathcal{P}$ instead of $\to_{\mathcal{P}}$. Consider the PTRS $\mathcal{P}_{\mathsf{geo}}$ with the only rule $\mathsf{geo}(x) \to \{{}^{1/2}\mathsf{geo}(\mathsf{s}(x)),\ {}^{1/2}x\}$. When starting with $\mathsf{geo}(\mathsf{0})$, repeated rewriting yields $\mathsf{s}^k(\mathsf{0})$ with a probability of $(1/2)^{k+1}$, i.e., a geometric distribution.

To track rewrite sequences of a PARS $(A, \to)$ with their probabilities, we consider *reduction trees (RTs)*. The nodes v of a $\to$-RT are labeled by pairs ${}^{p_v}a_v$ of a probability $p_v \in (0, 1]$ and an object $a_v \in A$, where the probability at the root is 1. For each node v with successors $w_1, \ldots, w_k$, the edge relation represents a rewrite step, i.e., $a_v \to \{{}^{p_{w_1}/p_v}a_{w_1}, \ldots, {}^{p_{w_k}/p_v}a_{w_k}\}$. For a $\to$-RT $\mathfrak{T}$, $V(\mathfrak{T})$ denotes its set of nodes, $\mathrm{root}(\mathfrak{T})$ is the object at its root, and $\mathrm{Leaf}(\mathfrak{T})$ denotes its set of leaves. An example for a $\mathcal{P}_{\mathsf{geo}}$-RT is shown in Fig. 1.

To analyze the probability of a term s reaching a term t, we use *truncated* RTs. For an RT $\mathfrak{T}$ and a term u, the truncated RT $\mathfrak{T}_{|u}$ results from $\mathfrak{T}$ by removing all proper successors of every node v labeled with a pair ${}^{p_v}u$. Hence, every node labeled with u is a leaf of $\mathfrak{T}_{|u}$, and along each path of $\mathfrak{T}_{|u}$ there is at most one node labeled with u, which is then the last node of that path. Intuitively, this lets us collect the probability mass that reaches u *for the first time*, without counting paths that pass through u several times.

Definition 1 (Reachability Probability)**.** Let $\mathcal{P}$ be a PTRS and $s, t \in \mathcal{T}$. For a $\mathcal{P}$-RT $\mathfrak{T}$, the *probability of reaching t from s in $\mathfrak{T}$* is $\mathbb{P}_{\mathfrak{T}}(s \to_{\mathcal{P}}^* t) = \sum_{\substack{v \in \mathrm{Leaf}(\mathfrak{T}_{|t}) \\ a_v = t}} p_v$ if $\mathrm{root}(\mathfrak{T}) = s$ and $\mathbb{P}_{\mathfrak{T}}(s \to_{\mathcal{P}}^* t) = 0$ otherwise. The *maximal reachability probability* of reaching t from s is $\mathbb{P}^{\max}(s \to_{\mathcal{P}}^* t) = \sup_{\mathcal{P}\text{-RT } \mathfrak{T}} \mathbb{P}_{\mathfrak{T}}(s \to_{\mathcal{P}}^* t)$. For a PTRS $\mathcal{P}$ and terms $s, t \in \mathcal{T}$, we write $s \dashrightarrow_{\leq p}^{\mathcal{P}} t$ iff $\mathbb{P}^{\max}(s\sigma \to_{\mathcal{P}}^* t\sigma) \leq p$ holds for *every* ground substitution σ w.r.t. s and t.

The supremum in $\mathbb{P}^{\max}(s \to_{\mathcal{P}}^* t)$ ranges over all $\mathcal{P}$-RTs with root ${}^{1}s$, i.e., over all ways of resolving the nondeterministic choices. Thus, $\mathbb{P}^{\max}(s \to_{\mathcal{P}}^* t)$ is the tightest upper bound on the probability of reaching t from s under any evaluation strategy, and $s \dashrightarrow_{\leq p}^{\mathcal{P}} t$ states that this bound stays below p for *all* ground instantiations of the variables of s and t. As an example, reconsider $\mathcal{P}_{\mathsf{geo}}$ and the RT $\mathfrak{T}$ in Fig. 1. Its leaves are the normal forms $\mathsf{s}^k(\mathsf{0})$, each reached with probability $(1/2)^{k+1}$. Hence $\mathbb{P}^{\max}(\mathsf{geo}(\mathsf{0}) \to_{\mathcal{P}_{\mathsf{geo}}}^* \mathsf{s}(\mathsf{0})) = 1/4$. The term $\mathsf{s}(\mathsf{0})$ only occurs once as a leaf (with probability $1/4$), and it cannot be reached again once we have rewritten to $\mathsf{geo}(\mathsf{s}(\mathsf{s}(\mathsf{0})))$ or stopped at $\mathsf{0}$. Moreover, there is only one $\mathcal{P}_{\mathsf{geo}}$-RT with root ${}^{1}\mathsf{geo}(\mathsf{0})$. We therefore have $\mathsf{geo}(\mathsf{0}) \dashrightarrow_{\leq 1/4}^{\mathcal{P}_{\mathsf{geo}}} \mathsf{s}(\mathsf{0})$, and this bound is tight. Similarly, $\mathsf{geo}(x) \dashrightarrow_{\leq 1/2}^{\mathcal{P}_{\mathsf{geo}}} x$ holds.

3 Upper Bounds on Reachability via MDPs

We now lift the *symbol transition graph* technique of [9] to the probabilistic setting. The idea is to over-approximate the (in general infinite) rewrite relation $\to_{\mathcal{P}}$ by a *finite* MDP and to read off an upper bound on $\sup_\sigma \mathbb{P}^{\max}(s\sigma \to_{\mathcal{P}}^* t\sigma)$ from the maximal reachability probability of that MDP, computable by probabilistic model checkers such as Storm [7]. The abstraction has two ingredients: (i) we only keep the top of each term, cutting everything below a fixed height n, so that finitely many terms remain; and (ii) since a cut term stands for *all* of its instances, a rule may apply to some instance even if its left-hand side does not match the cut term, so we replace matching by *unification*, i.e., the edges of the MDP are *probabilistic narrowing steps*. The *height* of a term is $h(x) = 0$ for $x \in \mathcal{V}$ and $h(f(t_1, \dots, t_k)) = 1 + \max\{h(t_1), \dots, h(t_k)\}$.

Definition 2 (Cut)**.** For $t \in \mathcal{T}$ and $n \in \mathbb{N}$, the *cut* $\lceil t \rceil_n$ replaces every subterm $t|_\pi$ at a position $\pi \in \mathrm{Pos}(t)$ with $|\pi| = n$ by a fresh variable (distinct positions get distinct fresh variables).

Thus $h(\lceil t \rceil_n) \le n$, and $\lceil t \rceil_n = t$ (up to renaming) if and only if $h(t) \le n$. Up to renaming there are only finitely many terms of height at most n, and a cut term u represents the set $\gamma(u) = \{u\sigma \mid \sigma \text{ a substitution}\}$ of all its instances (so $t \in \gamma(\lceil t \rceil_n)$ for all terms t), e.g., $\lceil \mathsf{geo}(\mathsf{s}(\mathsf{s}(0))) \rceil_2 = \mathsf{geo}(\mathsf{s}(x))$.

Recall that an MDP $\mathcal{M} = (S, Act, \mathbb{P})$ consists of a countable set of *states* S, finitely many *actions* Act, and a *transition probability function* $\mathbb{P}\colon S \times Act \times S \to [0,1]$ with $\sum_{s' \in S} \mathbb{P}(s, \alpha, s') \in \{0, 1\}$ for all states $s \in S$ and actions $\alpha \in Act$. The action α is *enabled* in s if this sum is 1. A *scheduler* chooses an enabled action in each state (possibly depending on the history), inducing a probability measure on the runs of $\mathcal{M}$. For a set $T \subseteq S$ of *target* states, $\mathbb{P}^{\max}_{\mathcal{M}}(s, T)$ denotes the maximal probability of eventually reaching T from s over all schedulers.

Definition 3 (n-Cut MDP)**.** Let $\mathcal{P}$ be a PTRS with maximal arity m and $n \in \mathbb{N}$. The *n-cut MDP* $\mathcal{M}_n(\mathcal{P}) = (S_n, Act, \mathbb{P})$ has states $S_n = \{t \in \mathcal{T} \mid h(t) \le n\}$ (up to renaming) and actions $Act = \{1, \dots, m\}^{\le n} \times \mathcal{P}$ (a position together with a rule, representing a possible rewrite step). An action $(\pi, \ell \to \mu)$ with $\mu = \{{}^{p_1}r_1, \dots, {}^{p_k}r_k\}$ is *enabled* in a term t (where w.l.o.g. $\mathcal{V}(t) \cap \mathcal{V}(\ell) = \varnothing$) iff $\pi \in \mathrm{Pos}(t)$ and $t|_\pi$ unifies with ℓ with mgu σ. It then performs a probabilistic narrowing step followed by a cut, yielding $\mathbb{P}(t, (\pi, \ell \to \mu), s') = \sum_{1 \le j \le k,\, \lceil (t\sigma)[r_j\sigma]_\pi \rceil_n = s'} p_j$ and $\mathbb{P}(t, (\pi, \ell \to \mu), \cdot) = 0$ if it is not enabled.

As an example, the 2-Cut MDP $\mathcal{M}_2(\mathcal{P}_{\mathsf{geo}})$ is shown in Fig. 2. Here, we only present the nodes that are reachable from $\mathsf{geo}(0)$.

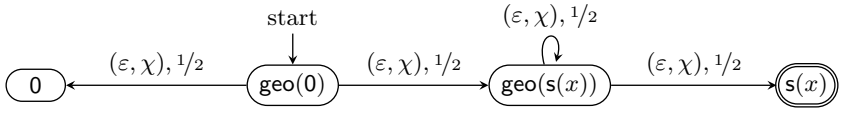

Figure 2: The reachable part of $\mathcal{M}_2(\mathcal{P}_{\mathsf{geo}})$ for $\mathcal{P}_{\mathsf{geo}}$ (both actions apply the rule at the root). A double border marks the target $\mathsf{s}(x) \in T^2_{\mathsf{s}(0)}$, and $\mathsf{0}$ is a normal form. By χ we denote the only rule in $\mathcal{P}_{\mathsf{geo}}$, namely $\mathsf{geo}(x) \to \{{}^{1/2}\mathsf{geo}(\mathsf{s}(x)),\ {}^{1/2}x\}$.

Applying the mgu σ to the *whole* term t (not only to the redex $t|_\pi$) is precisely a narrowing step $t \rightsquigarrow (t\sigma)[r_j\sigma]_\pi$, since σ may also instantiate variables of t outside of π. To obtain a bound on $\sup_\sigma \mathbb{P}^{\max}(s\sigma \to_{\mathcal{P}}^* t\sigma)$, we start in $\lceil s \rceil_n$ and use the target set $T^n_t = \{t' \in S_n \mid t' \text{ and } \lceil t \rceil_n \text{ are unifiable}\}$, i.e., all cut terms with a common instance with t.

Theorem 4 (Upper Bounds on $\sup_\sigma \mathbb{P}^{\max}(s\sigma \to_{\mathcal{P}}^* t\sigma)$)**.** *Let $\mathcal{P}$ be a PTRS, $s, t \in \mathcal{T}$, and $n \in \mathbb{N}$. Then*

$$\mathbb{P}^{\max}_{\mathcal{M}_n(\mathcal{P})}(\lceil s \rceil_n, T_t^n) \leq p \quad \Longrightarrow \quad \sup_\sigma \mathbb{P}^{\max}(s\sigma \to_{\mathcal{P}}^* t\sigma) \leq p$$

Proof Sketch. Fix a substitution σ and a $\to_{\mathcal{P}}$-RT $\mathfrak{T}$ with $\mathrm{root}(\mathfrak{T}) = s\sigma$, and map every node labeled with a term u to the state $\lceil u \rceil_n$. A concrete step at a position π with $|\pi| < n$ corresponds to the action $(\pi, \ell \to \mu)$ with the *same* probabilities p_j, while a step with $|\pi| \geq n$ occurs inside a cut-off subterm and leaves $\lceil u \rceil_n$ unchanged. So $\mathfrak{T}$ induces a scheduler of $\mathcal{M}_n(\mathcal{P})$ starting in $\lceil s\sigma \rceil_n$ that reaches T_t^n with probability at least $\mathbb{P}_{\mathfrak{T}}(s\sigma \to^* t\sigma)$, as every leaf labeled $t\sigma$ maps to $\lceil t\sigma \rceil_n \in T_t^n$; thus $\mathbb{P}_{\mathfrak{T}}(s\sigma \to^* t\sigma) \leq \mathbb{P}^{\max}_{\mathcal{M}_n(\mathcal{P})}(\lceil s\sigma \rceil_n, T_t^n)$. Since $\lceil s\sigma \rceil_n \in \gamma(\lceil s \rceil_n)$ and a more general state over-approximates its instances (every narrowing step enabled in an instance is also enabled, up to instantiation, in the more general term), we have $\mathbb{P}^{\max}_{\mathcal{M}_n(\mathcal{P})}(\lceil s\sigma \rceil_n, T_t^n) \leq \mathbb{P}^{\max}_{\mathcal{M}_n(\mathcal{P})}(\lceil s \rceil_n, T_t^n) \leq p$. Taking the supremum over all $\mathfrak{T}$ yields the claim. □

Increasing the cutoff n keeps more structure of the PTRS, so the bounds get tighter.

Theorem 5 (Higher n Yields Tighter Bounds)**.** *For all $n \leq n'$, we have*

$$\sup_\sigma \mathbb{P}^{\max}(s\sigma \to_{\mathcal{P}}^* t\sigma) \quad \leq \quad \mathbb{P}^{\max}_{\mathcal{M}_{n'}(\mathcal{P})}(\lceil s \rceil_{n'}, T_t^{n'}) \quad \leq \quad \mathbb{P}^{\max}_{\mathcal{M}_n(\mathcal{P})}(\lceil s \rceil_n, T_t^n) \quad \leq \quad 1.$$

Proof Sketch. The first inequality is Thm. 4 (taking the supremum over all σ). For the second one, consider the map $t' \mapsto \lceil t' \rceil_n$ of $\mathcal{M}_{n'}(\mathcal{P})$ into $\mathcal{M}_n(\mathcal{P})$. Every narrowing step in $\mathcal{M}_{n'}(\mathcal{P})$ is matched by one in $\mathcal{M}_n(\mathcal{P})$, and $T_t^{n'}$ maps into T_t^n. So every scheduler of $\mathcal{M}_{n'}(\mathcal{P})$ corresponds to one scheduler of $\mathcal{M}_n(\mathcal{P})$ reaching the target with at least the same probability. □

Example 6. Reconsider $\mathcal{P}_{\mathsf{geo}}$ with $\mathsf{geo}(x) \to \{{}^1\!/_2\mathsf{geo}(\mathsf{s}(x)),\ {}^1\!/_2 x\}$ and the reachability problem $\mathbb{P}^{\max}(\mathsf{geo}(0) \to_{\mathcal{P}}^* \mathsf{s}(0))$, whose best upper bound is $1/4$. The part of $\mathcal{M}_2(\mathcal{P}_{\mathsf{geo}})$ that is reachable from the start term $\mathsf{geo}(0)$ is shown in Fig. 2. In $\mathsf{geo}(\mathsf{s}(x))$, the right branch $\mathsf{s}(x)$ unifies with $\lceil \mathsf{s}(0) \rceil_2 = \mathsf{s}(0)$ and is thus a target, while the left branch returns to $\mathsf{geo}(\mathsf{s}(x))$ since the second s is cut off. This yields $\mathbb{P}^{\max}_{\mathcal{M}_2(\mathcal{P}_{\mathsf{geo}})}(\mathsf{geo}(0), T^2_{\mathsf{s}(0)}) = 1/2$, proving $\mathsf{geo}(0) \dashrightarrow^{\mathcal{P}_{\mathsf{geo}}}_{\leq 1/2} \mathsf{s}(0)$. The bound over-approximates $1/4$ because $\mathsf{geo}(\mathsf{s}(x))$ merges $\mathsf{geo}(\mathsf{s}(0))$ with all deeper terms $\mathsf{geo}(\mathsf{s}^k(0))$. Refining to $n = 3$ separates further terms, see Fig. 3. Now $\mathsf{geo}(\mathsf{s}(0))$ and $\mathsf{geo}(\mathsf{s}(\mathsf{s}(x)))$ are distinct states, and only the former reaches the target $\mathsf{s}(0)$, while the latter loops and escapes to $\mathsf{s}(\mathsf{s}(x))$. Hence $\mathbb{P}^{\max}_{\mathcal{M}_3(\mathcal{P}_{\mathsf{geo}})}(\mathsf{geo}(0), T^3_{\mathsf{s}(0)}) = 1/2 \cdot 1/2 = 1/4$, the exact value.

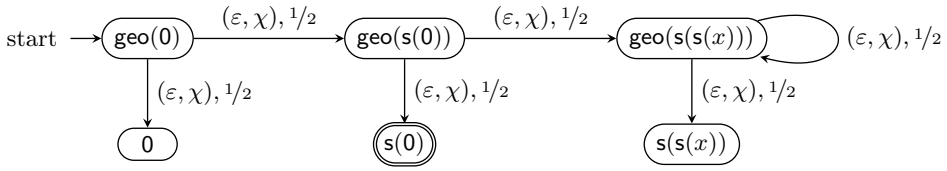

Figure 3: The reachable part of $\mathcal{M}_3(\mathcal{P}_{\mathsf{geo}})$ for $\mathcal{P}_{\mathsf{geo}}$. Here, $\mathsf{geo}(\mathsf{s}(0))$ and $\mathsf{geo}(\mathsf{s}(\mathsf{s}(x)))$ are distinct states and only the former reaches the target $\mathsf{s}(0) \in T^3_{\mathsf{s}(0)}$. Thus, $\mathbb{P}^{\max}_{\mathcal{M}_3(\mathcal{P}_{\mathsf{geo}})}(\mathsf{geo}(0), T^3_{\mathsf{s}(0)}) = 1/4$. As before, by χ we denote the only rule in $\mathcal{P}_{\mathsf{geo}}$.

4 Upper Bounds on Reachability via Interpretations

Next, we lift the notion of *term orderings* [10] to the probabilistic setting. Similar ideas have been used for probabilistic imperative programs in, e.g., [2, 3]. Compared to the classical setting,

we have to use interpretations that allow the definition of an expected value and that induce a supermartingale. Then, we can obtain an upper bound on $\sup_\sigma \mathbb{P}^{\max}(s\sigma \to_{\mathcal{P}}^* t\sigma)$ instead of simple qualitative answers whether $s \to^* t$ holds or not. For simplicity we only consider weight functions instead of more complex term orderings as in [10]. A *weight function* is a mapping $W : \Sigma \to \mathbb{N}$ and extends to a function on terms $W : \mathcal{T}(\Sigma, \mathcal{V}) \to \mathbb{N}[\mathcal{V}]$ by defining $W(x) = x$ for all $x \in \mathcal{V}$, and $W(f(t_1, \dots, t_n)) = W(f) + W(t_1) + \cdots + W(t_n)$ otherwise. The relation $\leq$ on $\mathbb{N}[\mathcal{V}]$ is defined as usual: $p(x_1, \dots, x_n) \leq p'(x_1, \dots, x_n)$ if this holds for all instantiations of $x_1, \dots x_n$ with natural numbers. The expected weight of a finite multi-distribution $\mu = \{{}^{p_1}t_1, \dots, {}^{p_k}t_k\}$ on terms is defined as $\mathbb{E}_W(\mu) = \sum_{1 \leq i \leq k} p_i \cdot W(t_i)$. For example, consider the PTRS $\mathcal{P}_{\mathsf{rw}}$ with a unary function symbol s and a constant symbol $\mathsf{0}$ with the only rules $\mathsf{s}(x) \to \{{}^{1/2}x, {}^{1/2}\mathsf{s}^2(x)\}$, which describes a one dimensional symmetric *random walk* over natural numbers in Peano-notation. A weight function $W : \{\mathsf{s}, \mathsf{0}\} \to \mathbb{N}$ is given by $W(\mathsf{0}) = 0$ and $W(\mathsf{s}) = 1$. Then $W(\mathsf{s}(x)) = x + 1$ and $\mathbb{E}_W(\{{}^{1/2}x, {}^{1/2}\mathsf{s}^2(x)\}) = 1/2 \cdot x + 1/2 \cdot (x + 2) = x + 1$ as well.

We now sketch how a weight function W that induces a supermartingale can be used to infer an upper bound on $\sup_\sigma \mathbb{P}^{\max}(s\sigma \to_{\mathcal{P}}^* t\sigma)$. The main idea is to approximate all rewrite sequences starting in $s\sigma$ by a supermartingale that tracks the weight of the current term. To this end, we require that W does not increase in expectation along rewrite steps, i.e.,

$$W(\ell) \geq \mathbb{E}_W(\mu) \qquad \text{for all rules } \ell \to \mu \in \mathcal{P}. \tag{1}$$

The inequation (1) lifts from rules to rewrite steps: whenever $t \to_{\mathcal{P}} \mu$, we have $W(t) \geq \mathbb{E}_W(\mu)$. Intuitively, this means that with every rewrite step we either move away from the target or stay at the same distance in expectation, i.e., we have

$$\begin{array}{cccccccccc} & \{{}^1 s\} & \to_{\mathcal{P}} & \mu_1 & \to_{\mathcal{P}} & \mu_2 & \to_{\mathcal{P}} & \mu_3 & \to_{\mathcal{P}} & \cdots \\ \implies & \mathbb{E}_W(\{{}^1 s\}) & \geq & \mathbb{E}_W(\mu_1) & \geq & \mathbb{E}_W(\mu_2) & \geq & \mathbb{E}_W(\mu_3) & \geq & \cdots \end{array}$$

If we have $W(t) > W(s)$, then we can bound the reachability probability $\sup_\sigma \mathbb{P}^{\max}(s\sigma \to_{\mathcal{P}}^* t\sigma)$: we cannot move with probability 1 from s to t, since this would imply that the expected value has to increase eventually. The Stopping Theorem for supermartingales yields a concrete bound.

We now give a formal definition of the induced supermartingale. Let σ be a ground substitution for s and t, and let $\mathfrak{T}$ be a $\to$-RT with $\mathrm{root}(\mathfrak{T}) = s\sigma$. We extend W to nodes by $W(v) = W(a_v)$ and to a fresh symbol $\bot \notin V(\mathfrak{T})$ by $W(\bot) = 0$. We turn $\mathfrak{T}$ into a stochastic process $\{X_n\}_{n \in \mathbb{N}}$ over $V(\mathfrak{T}) \uplus \{\bot\}$: (1) X_0 is the root of $\mathfrak{T}$; (2) if $X_n = v$ has children $w_1, \dots, w_k$, then $X_{n+1} = w_i$ with probability ${}^{p_{w_i}}/_{p_v}$. By definition of RTs, $\{{}^{p_{w_1}/p_v}a_{w_1}, \dots, {}^{p_{w_k}/p_v}a_{w_k}\}$ is exactly the distribution μ of the rewrite step $a_v \to \mu$; and (3) if $X_n = v$ is a leaf or $X_n = \bot$, then $X_{n+1} = \bot$ with probability 1. Thus, the realizations of $\{X_n\}_{n \in \mathbb{N}}$ are exactly the maximal paths of $\mathfrak{T}$, weighted by their probabilities. Let $Y_n = W(X_n)$ track the weight along the path. Then $\{Y_n\}_{n \in \mathbb{N}}$ is non-negative, and a *supermartingale*: the expected weight of the successor X_{n+1} never exceeds the current weight Y_n.[1] Indeed, for an inner node $X_n = v$ with $a_v \to \mu$, this expected weight is $\mathbb{E}_W(\mu) \leq W(a_v) = Y_n$ by inequation (1), and for a leaf or $\bot$ it is $0 \leq Y_n$.

To capture the event of reaching a term of weight at least $W(t\sigma)$, we use the *stopping time* $N = \inf\{\, n \in \mathbb{N} \mid Y_n \geq W(t\sigma) \,\} \in \mathbb{N} \cup \{\infty\}$ (with $\inf \varnothing = \infty$), i.e., we stop once we see a weight greater than or equal to $W(t\sigma)$. Since the event $\{Y_n \geq W(t\sigma)\}$ depends only on $X_0, \dots, X_n$, N is indeed a stopping time. Intuitively, we run the process until the current term has weight at least $W(t\sigma)$, and run forever otherwise. Now the following stopping theorem yields $\mathbb{E}(Y_0) \geq \mathbb{E}(Y_N)$ and thereby our upper bound on $\mathbb{P}^{\max}(s\sigma \to_{\mathcal{P}}^* t\sigma)$.

[1] Formally, $\mathbb{E}(Y_{n+1} \mid \mathfrak{F}_n) \leq Y_n$ for the natural filtration $\{\mathfrak{F}_n\}_{n \in \mathbb{N}}$ of $\{X_n\}_{n \in \mathbb{N}}$.

Theorem 7 (Stopping Theorem (Theorem 4.8.4 in [4])). *Let $\{Y_n\}_{n\in\mathbb{N}}$ be a non-negative supermartingale and $N \in \mathbb{N}\cup\{\infty\}$ a stopping time. Then $\mathbb{E}(Y_0) \geq \mathbb{E}(Y_N)$.*

Corollary 8. *Let s,t be terms in a PTRS $\mathcal{P}$, W be a weight function with $W(\ell) \geq \mathbb{E}_W(\mu)$ for all rules $\ell \to \mu \in \mathcal{P}$, and σ be a ground substitution for s and t. If $W(t\sigma) \geq W(s\sigma)$ and $W(t\sigma) > 0$ then $\mathbb{P}^{\max}(s\sigma \to^*_{\mathcal{P}} t\sigma) \leq W(s\sigma)\cdot W(t\sigma)^{-1}$.*

Proof Sketch. We instantiate the construction above for some $\to$-RT $\mathfrak{T}$ with $\mathrm{root}(\mathfrak{T}) = s\sigma$, yielding the non-negative supermartingale $\{Y_n\}_{n\in\mathbb{N}}$ and the stopping time N. By construction, N is the first time the process hits a node ${}^{p_N}a_N$ in $\mathfrak{T}$ where $W(a_N) \geq W(t\sigma)$. Using Thm. 7 we can infer that $\mathbb{E}(Y_0) \geq \mathbb{E}(Y_N)$. Note that our process always starts in $Y_0 = W(s\sigma)$. Moreover, $p^* := \mathbb{P}[N < \infty]$ is the probability that a path in $\mathfrak{T}$ ever hits a node ${}^{p_N}a_N$ with $W(a_N) \geq W(t\sigma)$. Since every path that reaches a node labeled with $t\sigma$ hits a node of weight $W(t\sigma)$ there, we have $\mathbb{P}_{\mathfrak{T}}(s\sigma \to^* t\sigma) \leq p^*$.

Finally, recall that $Y_N \geq W(t\sigma)$ with probability p^* and $Y_N \geq 0$ otherwise. Hence, $\mathbb{E}(Y_N) \geq p^*\cdot W(t\sigma) + (1-p^*)\cdot 0 = p^*\cdot W(t\sigma)$, and together with $\mathbb{E}(Y_0) \geq \mathbb{E}(Y_N)$ from above we obtain $W(s\sigma) \geq p^*\cdot W(t\sigma)$. So we can conclude that $\mathbb{P}_{\mathfrak{T}}(s\sigma \to^* t\sigma) \leq p^* \leq W(s\sigma)\cdot W(t\sigma)^{-1}$. Note that we can obtain the same upper bound for all $\to$-RTs $\mathfrak{T}$ with root ${}^1s\sigma$. Hence, we can ultimately conclude that $\mathbb{P}^{\max}(s\sigma \to^*_{\mathcal{P}} t\sigma) \leq W(s\sigma)\cdot W(t\sigma)^{-1}$ which proves the claim. □

Now we can combine all results so far to obtain the main result of this section.

Theorem 9 (Upper Bounds on Max Reachability via Weight Functions). *Let s,t be terms in a PTRS $\mathcal{P}$, and W be a weight function with $W(\ell) \geq \mathbb{E}_W(\mu)$ for all rules $\ell \to \mu \in \mathcal{P}$ such that $W(t) > 0$. Then $\sup_\sigma \mathbb{P}^{\max}(s\sigma \to^*_{\mathcal{P}} t\sigma) \leq p$ for*

$$p = \begin{cases} \sup_{\alpha\colon \mathcal{V}\to\mathbb{N}}\big(W(s)[v/\alpha(v) \mid v\in\mathcal{V}]\cdot(W(t)[v/\alpha(v)\mid v\in\mathcal{V}])^{-1}\big), & W(t\sigma)\geq W(s\sigma) \text{ for all } \sigma,\\ 1, & \text{otherwise.}\end{cases}$$

Proof Sketch. If $W(t\sigma) < W(s\sigma)$ for some ground substitution σ, then $p = 1$ is a trivial upper bound and the claim is immediate. Otherwise, $W(t\sigma) \geq W(s\sigma)$ holds for all ground substitutions σ, and Cor. 8 yields $\mathbb{P}^{\max}(s\sigma \to^*_{\mathcal{P}} t\sigma) \leq W(s\sigma)\cdot W(t\sigma)^{-1}$ for each such σ. Finally note that $\sup_{\alpha\colon \mathcal{V}\to\mathbb{N}}\big(W(s)[v/\alpha(v)\mid v\in\mathcal{V}]\cdot(W(t)[v/\alpha(v)\mid v\in\mathcal{V}])^{-1}\big) \geq \sup_\sigma(W(s\sigma)\cdot W(t\sigma)^{-1})$, since $W(u\sigma) = W(u)[v/W(v\sigma)\mid v\in\mathcal{V}]$ and α ranges freely over $\mathcal{V}\to\mathbb{N}$, while $W(v\sigma)$ ranges only over a subset of $\mathbb{N}$. Hence $\mathbb{P}^{\max}(s\sigma \to^*_{\mathcal{P}} t\sigma) \leq p$ for all σ, which proves the claim. □

Example 10. Consider the PTRS $\mathcal{P}_{\mathsf{rw}}$ with the only rules $\mathsf{s}(x) \to \{{}^{1/2}x, {}^{1/2}\mathsf{s}^2(x)\}$ again. Using the weight function $W(\mathsf{0}) = 0$ and $W(\mathsf{s}) = 1$, we get $W(\mathsf{s}(x)) = x+1 = 1/2\cdot x + 1/2\cdot(x+2) = \mathbb{E}_W(\{{}^{1/2}x, {}^{1/2}\mathsf{s}^2(x)\})$. Therefore, we can infer $\mathsf{s}(\mathsf{0}) \dashrightarrow^{\mathcal{P}_{\mathsf{rw}}}_{\leq 1/3} \mathsf{s}^3(\mathsf{0})$, since $W(\mathsf{s}(\mathsf{0})) = 1$, $W(\mathsf{s}^3(\mathsf{0})) = 3$, and $W(s)\cdot W(t)^{-1} = 1/3$.

In the future, we will investigate the relation between both presented techniques.

Acknowledgements We thank Éléanore Meyer for feedback on an earlier version, and the anonymous reviewers for their comments.

References

[1] Franz Baader and Tobias Nipkow. *Term Rewriting and All That.* Cambridge University Press, 1998.

[2] Krishnendu Chatterjee, Amir Kafshdar Goharshady, Tobias Meggendorfer, and Đorđe Žikelić. Sound and Complete Certificates for Quantitative Termination Analysis of Probabilistic Programs. In *Proc. CAV '22*, LNCS 13371, pages 55–78, 2022.

[3] Krishnendu Chatterjee, Petr Novotný, and Đorđe Žikelić. Stochastic invariants for probabilistic termination. In *Proc. POPL '17*, ?, pages 145–160, 2017.

[4] Rick Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 5 edition, 2019.

[5] Jürgen Giesl, Cornelius Aschermann, Marc Brockschmidt, Fabian Emmes, Florian Frohn, Carsten Fuhs, Jera Hensel, Carsten Otto, Martin Plücker, Peter Schneider-Kamp, Thomas Ströder, Stephanie Swiderski, and René Thiemann. Analyzing Program Termination and Complexity Automatically with AProVE. *Journal of Automated Reasoning*, 58(1):3–31, 2017.

[6] Raúl Gutiérrez and Salvador Lucas. Automatically Proving and Disproving Feasibility Conditions. In *Proc. IJCAR '20*, LNCS 12167, pages 416–435, 2020.

[7] Christian Hensel, Sebastian Junges, Joost-Pieter Katoen, Tim Quatmann, and Matthias Volk. The Probabilistic Model Checker Storm. *International Journal on Software Tools for Technology Transfer*, 24:589–610, 2022.

[8] Salvador Lucas and Raúl Gutiérrez. Use of Logical Models for Proving Infeasibility in Term Rewriting. *Information Processing Letters*, 136:90–95, 2018.

[9] Christian Sternagel and Akihisa Yamada. Reachability Analysis for Termination and Confluence of Rewriting. In *Proc. TACAS '19*, LNCS 11429, pages 262–278, 2019.

[10] Akihisa Yamada. Term Orderings for Non-reachability of (Conditional) Rewriting. In *Proc. IJCAR '22*, LNCS 13385, pages 248–267, 2022.