Verifying Qualitative Properties of Probabilistic Programs

Benedikt Bollig\textsuperscript{1}\textsuperscript{*} and Martin Leucker\textsuperscript{2}\textsuperscript{**}

\textsuperscript{1} Lehrstuhl f"{u}r Informatik II, RWTH Aachen, Germany
bollig@informatik.rwth-aachen.de
\textsuperscript{2} IT department, Uppsala University, Sweden
Martin.Leucker@it.uu.se

Abstract. In this chapter, we present procedures for checking linear temporal logic and automata specifications of sequential and concurrent probabilistic programs. We follow two different approaches: For LTL and sequential probabilistic programs, our method proceeds in a tableau style fashion, while the remaining procedures are based on automata theory.

1 Introduction

Randomization techniques have been employed to solve numerous problems of computing both sequentially and in parallel. Examples of probabilistic algorithms that are asymptotically better than their deterministic counterparts in solving various fundamental problems abound. It has also been shown that they allow solutions of problems which cannot be solved deterministically \cite{7}. They have the advantages of simplicity and better performance both in theory and often in practice. An overview of the domain of randomized algorithms is given in \cite{12}.

As for any kind of hardware or software system, it is important to develop formal methods and tools for verifying their correctness, thus also for probabilistic programs. Model checking, introduced independently in \cite{2} and \cite{14}, turned out to be one fruitful approach to automatically verify systems (see \cite{3} for an overview of model checking in practice). In the model-checking approach, usually a finite system $M$, often an abstraction of a real system, and a property, usually expressed as a temporal-logic formula $\varphi$ or as an automaton describing the computations that adhere to the property, are given. The model-checking procedure decides whether the set or tree formed of all computations of $M$ satisfy $\varphi$, or, in other words, whether the given system satisfies the required property. Temporal logics used for expressing requirements are usually linear temporal logic (LTL) (as initially proposed by Pnueli in \cite{10}) or branching time logics like the computation-tree logics CTL and CTL*. 

\textsuperscript{*} Part of this work was done during the author's stay at the IT department, Uppsala University, Sweden. He is grateful for the hospitality and the overall support.

\textsuperscript{**} This author is supported by the European Research Training Network "Games".
In the probabilistic setting, one is no longer interested in all but only almost all computations. Thus, (sets or paths of trees of) computations with zero probability are ignored when deciding whether a given property is satisfied. More general, one is also interested whether a given property is satisfied with some given probability.

In this chapter, we describe methods for checking whether almost all runs of a finite sequential or concurrent probabilistic program in the sense of [5] and [8] satisfy a given LTL formula or all accepting runs of an automaton describing the underlying property. Thus, we are concerned about quality aspects of a given probabilistic program rather than in estimating quantity measures for given properties or in temporal logics allowing to express quantitative aspects of properties. These issues are addressed in one of the subsequent chapters.

We present a procedure for checking LTL specifications of sequential probabilistic programs, which requires running time that is exponential in the size of the formula and linear in the size of the program. It can be shown that this is optimal.

Furthermore, we describe a procedure for checking automata specifications of sequential and concurrent probabilistic programs. It can be shown that this problem can be solved in time polynomial in the size of the program and exponential in the size of the formula. However, to simplify our presentation, we present a less technical construction based on the same idea, however, with a slightly inferior complexity than the one with optimal bounds. As LTL specifications can easily be transformed into automata specifications involving an exponential blow-up [17], the procedure can also be used to check LTL specifications (in time double exponential in the size of the formula).

This chapter is mainly based on [4], in which additionally methods for an extension of LTL—known as ETL [18]—are shown and proofs showing that all constructions are optimal are given. Since we only want to give an introduction to the field of verification of qualitative properties of probabilistic programs, we refer to this paper for a more detailed study of these extensions and proof ideas.

This chapter is organized as follows. In the next section, we introduce the necessary concepts and notation of automata, linear temporal logic, and probabilistic programs. Furthermore, we recall some mathematical background in the domain of graph and probability theory. In Section 3, we describe a model-checking procedure for LTL formulas and sequential probabilistic programs. Section 4 studies this question for automata specifications and sequential as well as concurrent probabilistic programs. We draw our conclusions in Section 5. This chapter ends giving suggestions for further reading in Section 6.

2 Preliminaries

2.1 Words and Graphs

Given an alphabet \( \Sigma \), \( \Sigma^* \) denotes the set of finite and \( \Sigma^w \) the set of infinite words over \( \Sigma \). For a word \( w = a_0a_1 \ldots \in \Sigma^w \) and a natural \( i \), let \( w(i) \) denote
automata of the form $\omega$-automata. A run of an automaton $A$ on a word $w = a_0 a_1 \ldots$ is a sequence of states $s_0, s_1, \ldots, s_n \in S$ such that.

\[ s_0 = s_0, s_1 = \delta(s_0, a_1), \ldots, s_n = \delta(s_{n-1}, a_n) \] 

A run is accepting if there is a state $s_f \in F$ such that $s_n = s_f$. The set $\mathbb{L}(A)$ of all accepting runs of $A$ on words over $\Sigma$ is called the language of $A$.

### 2.3 Propositional Linear Temporal Logic

In the following, let $P$ be a nonempty finite set of propositions. The set $\mathbb{L}(P)$ of (propositional) Linear Temporal Logic (LTL) formulas over $P$ is inductively defined as follows: $P$ is a subset of $\mathbb{L}(P)$, and, for $\mathbb{L}(P)$ formulas $\varphi$ and $\psi$, $\neg \varphi$, $\varphi \lor \psi$, $\varphi \land \psi$, $\varphi \rightarrow \psi$, $\varphi \leftrightarrow \psi$, $\exists \varphi$, and $\forall \varphi$. For $\mathbb{L}(P)$ formulas as well, an $\mathbb{L}(P)$ formula is inductively interpreted over $\pi = \pi_0 \pi_1 \ldots \in (2^P)^\omega$ as follows:

- $\pi \models p \iff p \in \pi_0$
- $\pi \models \neg \varphi \iff \pi \not\models \varphi$
- $\pi \models \varphi \lor \psi \iff \pi \models \varphi$ or $\pi \models \psi$
- $\pi \models \varphi \rightarrow \psi \iff \pi \models \varphi$ if $\pi \models \psi$
- $\pi \models \varphi \land \psi \iff \pi(1) \models \varphi$
- $\pi \models \varphi \land \psi \iff$ there is an $i \geq 0$ such that $\pi(i) \models \psi$ and, for each $j \in \{0, \ldots, i-1\}$, $\pi(j) \models \varphi$

The size $|A|$ of an automaton $A$ is defined to be the size of its (transition) graph.
The language of a formula $\varphi \in \text{LTL}(P)$, denoted by $L(\varphi)$, is defined to be the set \( \{ \pi \in (2^P)^\omega \mid \pi \models \varphi \} \).

### 2.4 Probability Spaces and Probabilistic Programs

A nonempty set of possible outcomes of an experiment of chance is called sample space. Let $\Omega$ be a sample space. A set $\mathcal{B} \subseteq 2^\Omega$ is called Borel field over $\Omega$ if it contains $\Omega$, $\Omega \setminus E$ for each $E \in \mathcal{B}$, and the union of any countable sequence of sets from $\mathcal{B}$. A Borel field $\mathcal{B}$ is generated by an at most countable set $\mathcal{E}$, denoted by $\mathcal{B} = \langle \mathcal{E} \rangle$, if $\mathcal{B}$ is the closure of $\mathcal{E}$’s elements under complement and countable union.

A probability space is a triple $\mathcal{PS} = (\Omega, \mathcal{B}, \mu)$ where $\Omega$ is a sample space, $\mathcal{B}$ is a Borel field over $\Omega$, and $\mu$ is a mapping $\mathcal{B} \rightarrow [0, 1]$ such that $\mu(\Omega) = 1$ and $\mu(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i)$ for any sequence $E_1, E_2, \ldots$ of pairwise disjoint sets from $\mathcal{B}$. We call $\mu$ a probability measure. An event $E \in \mathcal{B}$ is said to occur almost surely if $\mu(E) = 1$.

A concurrent probabilistic program over $\mathcal{P}$, the set of propositions, is a tuple $\mathcal{M} = (Q, N, R, \Delta, P_0, P, \nu)$ where

- $Q$ is its nonempty finite set of states, which is partitioned into sets $N$ and $R$ of nondeterministic and randomizing states, respectively,
- $\Delta \subseteq Q \times Q$ is the set of transitions with $\Delta(q) := \{ q' \mid (q, q') \in \Delta \} \neq \emptyset$ for each $q \in Q$,
- $P_0 : Q \rightarrow [0, 1]$ and $P : \{ (q, q') \in \Delta \mid q \in R \} \rightarrow (0, 1]$ are the initial and transition probability distributions\(^1\), respectively, where $\sum_{q \in Q} P_0(q) = 1$ and, for each $q \in R$, $\sum_{q' \in \Delta(q)} P_{q, q'} = 1$, and
- $\nu : Q \rightarrow 2^P$ is the valuation function, which is extended to words over $Q$ as expected.

A concurrent probabilistic program over $\{p\}$ is shown in Figure 1. Hereby, the nondeterministic states are represented by circles, whereas the randomizing ones are given by rectangles.

\[\text{Fig. 1. A concurrent probabilistic program}\]

\(^1\) We usually write $P_{q, q'}$ instead of $P((q, q'))$. 

---

The nature of the language and the probability measures allows for a rich set of properties to be investigated. Understanding these properties can provide insights into the behavior of the programs and the likelihood of certain outcomes.
In case that a concurrent probabilistic program $\mathcal{M}$ has no nondeterministic states, i.e., $N$ is the empty set, we call it a \textit{sequential probabilistic program} over $\mathcal{P}$ and identify $\mathcal{M}$ with $(Q, \Delta, P_0, P, \mathcal{V})$. $\mathcal{M}$ induces a probability space $\mathcal{PS}_\mathcal{M} = (\Omega_\mathcal{M}, \mathcal{B}_\mathcal{M}, \mu_\mathcal{M})$ as follows: $\Omega_\mathcal{M} = \{q_1 q_2 \ldots \in Q^* \mid P_0(q_1) > 0 \text{ and, for all } i \geq 1, (q_i, q_{i+1}) \in \Delta\}$ is the set of trajectories of $\mathcal{M}$, $\mathcal{B}_\mathcal{M}$ is generated by $\{C_\mathcal{M}(x) \mid x \in Q^*\}$ where $C_\mathcal{M}(x) = \{x' \in \Omega_\mathcal{M} \mid x \text{ is a prefix of } x'\}$ is the \textit{basic cylinder set} of $x$ wrt. $\mathcal{M}$. The measure $\mu_\mathcal{M}$ is uniquely given by $\mu_\mathcal{M}(\Omega_\mathcal{M}) = 1$ and, for $n \geq 1$, $\mu_\mathcal{M}(C_\mathcal{M}(q_1 \ldots q_n)) = P_0(q_1) \cdot P_{q_1,q_2} \cdot \ldots \cdot P_{q_{n-1},q_n}$ (where we assume $P_{q_n,q_{n+1}} = 0$ if $(q_n, q_{n+1}) \notin \Delta$).

A \textit{scheduler} of a concurrent probabilistic program $\mathcal{M} = (Q, N, R, \Delta, P_0, P, \mathcal{V})$ over $\mathcal{P}$ is a mapping $u : Q^* N \to Q$ such that $u(xq) = q'$ implies $(q, q') \in \Delta$.

![Fig. 2. A scheduler](image)

Similarly to the sequential case, a scheduler $u$ of a concurrent program $\mathcal{M} = (Q, N, R, \Delta, P_0, P, \mathcal{V})$ induces a probability space $\mathcal{PS}_\mathcal{M,u} = (\Omega_{\mathcal{M,u}}, \mathcal{B}_{\mathcal{M,u}}, \mu_{\mathcal{M,u}})$ as follows: $\Omega_{\mathcal{M,u}} = \{q_1 q_2 \ldots \in Q^* \mid P_0(q_1) > 0 \text{ and, for all } i \geq 1, (q_i \in N \text{ and } u(q_1 \ldots q_i) = q_{i+1})\}$ or $(q_i \in R \text{ and } u(q_1 \ldots q_i) = q_{i+1})\}$ is the set of trajectories of $\mathcal{M}$ wrt. $u$, $\mathcal{B}_{\mathcal{M,u}} = \{C_{\mathcal{M,u}}(x) \mid x \in Q^*\}$ where $C_{\mathcal{M,u}}(x) = \{x' \in \Omega_{\mathcal{M,u}} \mid x \text{ is a prefix of } x'\}$ is the \textit{basic cylinder set} of $x$ wrt. $\mathcal{M}$ and $u$. The measure $\mu_{\mathcal{M,u}}$ is uniquely given by $\mu_{\mathcal{M,u}}(\Omega_{\mathcal{M,u}}) = 1$ and, for $n \geq 1$, $\mu_{\mathcal{M,u}}(C_{\mathcal{M,u}}(q_1 \ldots q_n)) = P_0(q_1) \cdot P_{q_1,q_2} \cdot \ldots \cdot P_{q_{n-1},q_n}$ where

$$P'_{q_i,q_{i+1}} = \begin{cases} 0 & \text{if } (q_i, q_{i+1}) \notin \Delta \\ P_{q_i,q_{i+1}} & \text{if } (q_i, q_{i+1}) \in \Delta \text{ and } q_i \in R \\ 1 & \text{if } (q_i, q_{i+1}) \in \Delta \text{ and } q_i \in N \end{cases}$$
When we concentrate on topological aspects of probabilistic programs, we consider a program to be a graph where the states become (nondeterministic or randomizing) nodes and the transitions become edges. In the following, unless the context requires, a program will also denote its graph. From this point of view, we call a state absorbing if it forms a (nontrivial) bottom SCC. Otherwise, it is called transient.

3 Checking LTL Specifications of Sequential Programs

For a concurrent probabilistic program \( \mathcal{M} \) over \( \mathcal{P} \) with valuation function \( \mathcal{V} \), a scheduler \( u \) of \( \mathcal{M} \), and a formula \( \varphi \in \text{LTL}(\mathcal{P}) \), let \( L_{\mathcal{M}, u}(\varphi) := \{ x \in \Omega_{\mathcal{M}, u} \mid \mathcal{V}(x) \in L(\varphi) \} \). Similarly, for a sequential program \( \mathcal{M} \) with valuation function \( \mathcal{V} \), we define \( L_{\mathcal{M}}(\varphi) := \{ x \in \Omega_{\mathcal{M}} \mid \mathcal{V}(x) \in L(\varphi) \} \). It can be shown, using structural induction, that these sets are measurable:

**Proposition 1** ([16]). Given a concurrent (sequential) probabilistic program \( \mathcal{M} \) over \( \mathcal{P} \) and a formula \( \varphi \in \text{LTL}(\mathcal{P}) \), we have \( L_{\mathcal{M}, u}(\varphi) \in \mathbb{B}_{\mathcal{M}, u} \) for each scheduler \( u \) of \( \mathcal{M} \) (\( L_{\mathcal{M}}(\varphi) \in \mathbb{B}_{\mathcal{M}} \)).

**Definition 1.** Given a formula \( \varphi \in \text{LTL}(\mathcal{P}) \), a concurrent probabilistic program \( \mathcal{M} \) over \( \mathcal{P} \) is said to satisfy \( \varphi \) if, for all schedulers \( u \) of \( \mathcal{M} \), \( \mu_{\mathcal{M}, u}(L_{\mathcal{M}, u}(\varphi)) = 1 \) (we say that \( u \) satisfies \( \varphi \)). A sequential probabilistic program is said to satisfy \( \varphi \) if \( \mu_{\mathcal{M}}(L_{\mathcal{M}}(\varphi)) = 1 \).

For example, the concurrent probabilistic program \( \mathcal{M} \) shown in Figure 1 and, in particular, the scheduler of \( \mathcal{M} \) from Figure 2, satisfy the LTL(\( \{p\} \)) formula \( \diamond \Box p. \) As usual, \( \Diamond \varphi \) is an abbreviation for \( \text{True}\varphi \) and \( \Box \varphi \) for \( \neg \Diamond \neg \varphi. \)

**Theorem 1.** We can test whether a sequential probabilistic program \( \mathcal{M} \) satisfies a formula \( \varphi \) in time \( O([\mathcal{M}]^{2^{\varphi}}) \), or in space polynomial in \( \varphi \) and polylogarithmic in \( \mathcal{M} \). We can compute the probability of satisfaction \( \mu_{\mathcal{M}}(L_{\mathcal{M}}(\varphi)) \) in time exponential in \( \varphi \) and polynomial in \( \mathcal{M} \).

In the remainder of this section, we prove the previous theorem by formulating a procedure checking whether a sequential probabilistic program \( \mathcal{M} \) satisfies an LTL formula \( \varphi \). The procedure processes \( \varphi \) according to its inductive structure in a bottom-up manner, eliminating its temporal connectives one-by-one and transforms \( \mathcal{M} \) in each step to preserve the probability of satisfiability.

To simplify our presentation, we sometimes say that a state satisfies an LTL formula instead of saying that all trajectories starting from this state satisfy this formula.

There are two transformations \( \mathcal{T}_U \) and \( \mathcal{T}_X \) corresponding to the two temporal connectives \( \mathcal{U} \) and \( \mathcal{X} \). We describe \( \mathcal{T}_U \) in detail and list the modifications for \( \mathcal{T}_X \). We illustrate the procedure with the program given in Figure 3.
Transformation $T_{\phi}$. Let $p \land \phi$ be the innermost subformula of $\phi$. Let $W^+$ contain all states satisfying $p'$ and $W^-_1$ contain all states satisfying $\neg p$ and $\neg p'$. Let $W^-_2$ comprise all states of bottom strongly connected components (of $M$ considered as a graph) which have no state in $W^+$. Thus, states of $W^-_2$ possibly satisfy $p$ but no state satisfying $p'$ is reachable. Let $W^- = W^-_1 \cup W^-_2$. Clearly, all trajectories starting in $W^+$ satisfy $p \land \phi$ and all trajectories starting in $W^-$ satisfy $\neg(p \land \phi)$.

For our example (Figure 3), we get $W^+ = \{q_5\}$, $W^- = W^-_2 = \{q_4\}$.

Let $Q^+$ ($Q^-$) be the states $q$ such that for all trajectories starting at $q$ the first state in $W^+ \cup W^-$ is in $W^+$ ($W^-$). Note that $W^+ \subseteq Q^+$ and $W^- \subseteq Q^-$. For $M$ of Figure 3, we have $Q^+ = \{q_3, q_5\}$ and $Q^- = \{q_2, q_4\}$.

Almost surely, every trajectory reaches eventually a bottom strongly connected component. Trajectories starting in a state of $Q^+$ satisfy $p \land \phi'$ and those starting in a state of $Q^-$ never reach a state satisfying $p'$ before reaching a state not satisfying $p$ or never reach a state satisfying $p'$ at all. Using $P_q$ as a shorthand for the measure of trajectories starting in $q$ satisfying $p \land \phi'$, we conclude that $P_q = 1$ for all $q \in Q^+$ and $P_q = 0$ for $q \in Q^-$. Let $Q^? = \{q\}$ in our example. As $q \in Q^?$ satisfies $p$ but not $p'$, a trajectory $\pi$ starting in $q$ satisfies $p \land \phi'$ iff $\pi(1)$ satisfies $p \land \phi'$. Thus, $P_q = \sum_{q'} P_{q,q'} P_{q'}$. We summarize:

**Lemma 1.**

\[
P_q = \sum_{q'} P_{q,q'} P_{q'} \quad \text{if } q \in Q^?
\]
\[
P_q = 1 \quad \text{if } q \in Q^+
\]
\[
P_q = 0 \quad \text{if } q \in Q^-
\]

This set of equations has a unique solution.

**Proof.** It remains to show that the set of equations has a unique solution. Clearly, any two solutions can only differ on $Q^?$. Let us form a new sequential probabilistic program $M^?$ whose state space consists of $Q^?$ and one additional absorbing state $\hat{q}$ with a transition to itself with probability 1. For pairs of states of $Q^?$,

\[\text{as Boolean combinations of propositions can be evaluated in each state, we simplify our life and think of } p \text{ and } p' \text{ as propositions.}\]
the transitions and their probabilities agree with those of $\mathcal{M}$. Furthermore, for each $q \in Q^+$ with transitions to states $q' \notin Q^+$ we add a single transition $q \to q'$ with probability $\sum_{q \in Q^+} P_{q,q'}$. The state $q$ forms the only bottom s.c.c.: Since all states of $Q^+$ satisfy $p$ and $\neg p'$ any further bottom s.c.c. would be in $W_2^-$, thus not in $Q^+$, hence, every trajectory will reach state $q$ with probability one.

Let $T$ be the matrix of transition probabilities $P_{q,q'}$ restricted to states of $Q^+$. Furthermore, for $T^k$, the $k$th power of the matrix $T$, the entry $T^k_{q,q'}$ is equal to the probability that a path starting in $q$ reaches $q'$ in $k$ steps. Since every trajectory goes eventually to $q$ (and stays there) with probability one, we get $\lim_{k \to \infty} T^k = 0$. For any two solutions $a = (a_q)_{q \in Q^+}$ and $b = (b_q)_{q \in Q^+}$ of the equation system, we have $a - b = T(a - b)$, thus $a - b = T^k(a - b)$, and taking the limit as $k$ tends to infinity we conclude that $a = b$. 

For the example shown in Figure 3, we get $P_{21} = P_{22} = 1$ and $P_{23} = P_{32} = 0$. Furthermore, $P_{31} = P_{4,31} + P_{4,3} + P_{4,4} = 0.2 P_{31} + 0.3 + 0.5 = \frac{2}{3}$.

We are ready for the construction of the new program $\mathcal{M}'$ and the new formula $\psi'$. $\mathcal{M}'$ has a larger state space $Q'$ and is defined over the set of atomic propositions $\mathcal{P}$ enriched with $\xi$ which serves in $\psi'$ as a substitute for $\mu \cup P'$. Let

- $Q' = \{ (q,\xi) \mid q \in Q^+ \cup Q^- \} \cup \{ (q,-\xi) \mid q \in Q^- \cup Q^- \}$.
- $\mathcal{V}'((q,\xi)) = \mathcal{V}(q) \cup \{ \xi \}$ and $\mathcal{V}'((q,-\xi)) = \mathcal{V}(q)$.

- we define the transitions of $\mathcal{M}'$ implicitly by giving non-zero transition probabilities $\xi_1$ and $\xi_2$ stand for $\xi$ or $-\xi$ and $P'$ for $1 - P$. We distinguish the following cases based on the states of $\mathcal{M}$:

  - $q, q' \in Q^+ \cup Q^-$. There is exactly one state $(q,\xi_1)$ and $(q',\xi_2)$ in $Q'$. We set $P'_{(q,\xi_1),(q',\xi_2)} = P_{q,q'}$.

  - $q \in Q^+ \cup Q^-$, $q' \in Q'$. There is exactly one state with first component $q$ and two states with first component $q'$. We set $P'_{(q,\xi_1),(q',\xi)} = P_{q,q'} P_{q'}$ and $P'_{(q,\xi_1),(q',-\xi)} = P_{q,q'} P_{\overline{q}}$.

  - $q \in Q^+$. If $q' \in Q^-$, we define two transitions by $P'_{(q,\xi_1),(q',-\xi)} = P_{q,q'} P_{\overline{q}}/P_{q'}$ and $P'_{(q,\xi_1),(q',\xi)} = P_{q,q'} P_{q'}/P_{\overline{q}}$. If $q' \in Q^+$ $(q \in Q^-)$, we define a single transition by $P'_{(q,\xi_1),(q',\xi)} = P_{q,q'} P_{q'}/P_{\overline{q}}$ (resp. $P'_{(q,\xi_1),(q',-\xi)} = P_{q,q'} P_{\overline{q}}/P_{q}$).

- $P_0'$ is defined by $P_0'(q,\xi) = R_0(q) P_q$ and $P_0'(q,-\xi) = R_0(q) P_{\overline{q}}$ for all $(q,\xi), (q,-\xi) \in Q'$.

The new sequential probabilistic program is $\mathcal{M}' = (Q', \Delta', P', P_0', \mathcal{V}')$ and the formula $\psi'$ is obtained by substituting $\mu \cup P'$ by $\xi$ in $\varphi$.

For the sequential probabilistic program $\mathcal{M}$ of Figure 3 and $\varphi = \mu \cup P'$, we obtain $\mathcal{M}'$ as given in Figure 4 and $\varphi' = \xi$.

Let us show that the probability of $\mathcal{M}$ satisfying $\varphi$ coincides with the one of $\mathcal{M}'$ satisfying $\varphi'$. Therefore, we analyze our construction in more detail.

Let $G^+$ be the subgraph of $\mathcal{M}'$ induced by states $(q,\xi)$ with $q \in Q^+$, $G^-\xi$ be the one induced by all states $(q,-\xi)$ with $q \in Q^-$. Furthermore, let $G'^{+\xi}$ be the subgraph induced by states $(q,\xi)$ and $G'^{-\xi}$ the one induced by states $(q,-\xi)$, for $q \in Q^+$. By construction $G'^{+\xi}$ ($G'^{-\xi}$) is isomorphic to $G^+$ ($G^-$), the subgraph
of $\mathcal{M}$ induced by states in $Q^+$ ($Q^-$). Furthermore, $G^x_+\xi$ is isomorphic to $G^y_\xi$, the subgraph of $\mathcal{M}$ induced by the states of $Q^x_\xi$.

Additionally, every transition from a state of $G^x_+\xi$ ($G^y_\xi$) leading to a state not in $G^x_+\xi$ ($G^y_\xi$), leads to a state in $G^x_\xi$ ($G^y_\xi$). Thus, if $\xi$ holds in a state of $Q^x$, almost surely every trajectory will eventually reach a state satisfying $p'$ and every state in between satisfies $p$, and, if $\neg\xi$, almost all trajectories satisfy $\neg(p\cup p')$. In other words, $\xi \equiv pl(p')$ holds almost surely in every state of $\mathcal{M}'$.

Let us see how the transition probabilities of $\mathcal{M}$ and $\mathcal{M}'$ relate: For every finite path $q_1' \cdots q_m'$ in $\mathcal{M}'$, we have by construction that $\mu_{\mathcal{M}'}(q_1' \cdots q_m') = P_{q_1,q_2} \cdots P_{q_{m-1},q_m} \cdot P_{q_m,\neg\xi}$, or $\mu_{\mathcal{M}'}(q_1' \cdots q_m') = P_{q_1,q_2} \cdots P_{q_{m-1},q_m}$, depending on whether $q_m' = (q_m,\xi)$ or $q_m' = (q_m,\neg\xi)$.

Let $g$ be the mapping that projects a trajectory of $\mathcal{M}'$ to $\mathcal{M}$ by projecting to the first component. If $g(q_m') \in Q^+ \cup Q^-$ then $q_1' \cdots q_m'$ is the only trajectory with projection to $q_1 \cdots q_m$. If $g(q_m') \in Q^\xi$, there are exactly two trajectories with projection $q_1 \cdots q_m$: $(q_1,\xi) \cdots (q_m,\xi)$ and $(q_1,\neg\xi) \cdots (q_m,\neg\xi)$.

We conclude that for every measurable set $E$ of trajectories of $\mathcal{M}$ the set $\{w' \in Q^{\xi'} \mid g(w) \in E\}$ is measurable and both sets have the same measure. This implies that $\mu_{\mathcal{M}'}(L_{\mathcal{M}'}(\varphi)) = \mu_{\mathcal{M}}(L_{\mathcal{M}}(\varphi))$. Since $pl(p') \equiv \xi$ for almost all trajectories in $\mathcal{M}'$, we get

**Proposition 2.**

$$\mu_{\mathcal{M}'}(L_{\mathcal{M}'}(\varphi')) = \mu_{\mathcal{M}}(L_{\mathcal{M}}(\varphi))$$

We conclude the presentation of $T_d$ analyzing its complexity. Let us first consider its time complexity. The construction of the new formula can be carried out in linear time with respect to $|\varphi|$. The information whether a state of $\mathcal{M}$ belongs to $Q^+$, $Q^-$, or $Q^\xi$ can be evaluated in linear time, as we show. We assume that $\mathcal{M}$ is given as a graph. Note that a Boolean combination of atomic propositions can be evaluated in time linear in its size in every state, so that our treatment of $p$ and $p'$ as propositions does not change the complexity.

1. Mark every state of $\mathcal{M}$ by $Q^+$ in which $p'$ holds and by $Q^-$ in which neither $p$ nor $p'$ holds.

![Fig. 4. A sequential probabilistic program](image-url)
2. Partition the remaining graph into strongly connected components. Process those in a bottom-up manner, i.e., when a s.c.c. $D$ is processed all s.c.c. that are reachable from $D$ have already been processed. Assign the nodes of $D$ in the following manner:

(a) if there is no edge leading to a state in $Q^+$ then add the nodes of $D$ to $Q^-$.
(b) if all edges leaving $D$ lead to states in $Q^+$ then add the nodes of $D$ to $Q^+$.
(c) otherwise add the nodes of $D$ to $Q^-$.

It is easy to see that the algorithm is correct. (1) corresponds to identifying $W^+$ and $W^-$, it can be done in linear time using a standard depth-first search. Partitioning a graph into strongly connected components can also be done in linear time. In (2), the bottom strongly connected components are processed first, which also finds the states of $W^+$. The further processing is obviously correct. The classification of every s.c.c. can be carried out in time linear in the size of the component.

For constructing $M'$ we have to solve the linear equation system of Lemma 1, which can be done in polynomial time in the size of $M$. For checking whether the formula is satisfied, however, the exact transition probabilities are not important, but it has to be decided whether they are 0 or $\geq 0$. Thus, we need either linear or polynomial time for construction $M'$ depending on whether we are interested in the question of satisfaction or in the exact measure.

Let us consider the space complexity of $T_{\Phi}$. First note that checking whether there is a path from a node $u$ to a node $v$ can be tested in space $\log^2 |M|$. For every state $q$ of $M$, we have to check whether $(q, q')$ or $(q, q')$ are states of $M'$. $(q, q') \in Q'$ iff $q \in Q^+ \cup Q^-$. This is the case iff $q$ satisfies $p$ or there is a node $q'$ reachable satisfying $p'$ by a path of which every node satisfies $p$. This can be checked in space $\log^2 |M|$, $(q, q') \in Q'$ iff $q \in Q^- \cup Q^+$. This is the case if $q$ satisfies $p$ and $p'$ or satisfies $p$ but not $p'$ and there is a path by states satisfying $p$ to a state $q' \in Q^- \cup Q^+$ (tested as described above). Again, this can be tested in space $\log^2 |M|$. The edges of $M'$ can be constructed using similar ideas.

Transformation $T_X$ Transformation $T_X$ for an innermost formula of the form $\lambda p$ is patterned after the transformation $T_{\Phi}$. We first partition the states of $M$ into three disjoint subsets $Q^+$, $Q^-$, and $Q^0$, defined as follows: $Q^+$ comprises all states $q$ for which all transitions are into states satisfying $p$, $Q^-$ contains the ones for which all transitions yield states not satisfying $p$, and the remaining states having transitions to states satisfying $p$ as well as transitions to states not satisfying $p$ are assigned to $Q^0$. Again, we will see that for a trajectory starting in a state of $Q^+$ ($Q^-$), it will satisfy $\lambda p$ with probability 1 (0, respectively). For trajectories starting in states of $Q^0$ both events have a non-zero probability.

As before, let $P_q$ denote the probability that $\lambda p$ is satisfied starting from state $q$, which is given as
Lemma 2.

\[ P_q = \sum_{p \in \mathcal{V}(q')} P_{q,p} \quad \text{if } q \in Q' \]
\[ P_q = 1 \quad \text{if } q \in Q^+ \]
\[ P_q = 0 \quad \text{if } q \in Q^- \]

As there is no recursion involved, the previous set of equations trivially has a unique solution.

The new sequential probabilistic program \( \mathcal{M}' \) has a larger state space \( Q' \) and is defined over the set of atomic propositions \( \mathcal{P} \) enriched with \( \xi \) which serves in \( \varphi' \) as a substitute for \( Xp \). The state space and the propositions valid in each state are similarly defined as in the case of \( \mathcal{T}_q \). We define the transitions of \( \mathcal{M}' \) implicitly by giving non-zero transition probabilities, as follows:

Let \( \xi_1 \) and \( \xi_2 \) stand for \( \xi \) or \( \neg \xi \) and \( \overline{P} \) for \( 1 - P \). We distinguish the following cases based on the states of \( \mathcal{M} \):

- \( q, q' \in Q^+ \cup Q^- \). There is exactly one state \((q, \xi_1)\) and \((q', \xi_2)\) in \( Q' \). We set \( P'_{(q, \xi_1), (q', \xi_2)} = P_{q,q'} \).
- \( q \in Q^+ \cup Q^- \), \( q' \in Q^2 \). There is exactly one state with first component \( q \) and two states with first component \( q' \). We set \( P'_{(q, \xi_1), (q', \xi_2)} = P_{q,q'} P_q \) and \( P'_{(q, \xi_1), (q', \neg \xi)} = P_{q,q'} \overline{P}_q \).
- \( q \in Q^2 \).
  - If \( q' \in Q^2 \) and \( q' \) satisfies \( p \), we define two transitions by \( P'_{(q, \xi_1), (q', \xi_2)} = P_{q,q'} P_{q'} / P_q \) and \( P'_{(q, \xi_1), (q', \neg \xi)} = P_{q,q'} \overline{P}_{q'} / P_q \).
  - If \( q' \in Q^2 \) and \( q' \) satisfies \( \neg p \), we define two transitions by \( P'_{(q, \xi_1), (q', \xi_2)} = P_{q,q'} P_{q'} / P_q \) and \( P'_{(q, \xi_1), (q', \neg \xi)} = P_{q,q'} \overline{P}_{q'} / P_q \).
  - If \( q' \in Q^\ast \cup Q^- \) and \( q' \) satisfies \( p \), we define a single transition by \( P'_{(q, \xi_1), (q', \xi_2)} = P_{q,q'} / P_q \) where \((q', \xi_2)\) is the unique state of \( \mathcal{M}' \) with first component \( q' \).
  - If \( q' \in Q^\ast \cup Q^- \) and \( q' \) satisfies \( \neg p \), we define a single transition by \( P'_{(q, \xi_1), (q', \xi_2)} = P_{q,q'} \overline{P}_q \) where \((q', \xi_2)\) is the unique state of \( \mathcal{M}' \) with first component \( q' \).

The initial distribution \( P_0' \) is defined by \( P_0'((q, \xi)) = R_0(q)P_q \) and \( P_0'((q, \neg \xi)) = R_0(q)\overline{P}_q \), for all \((q, \xi), (q, \neg \xi) \in Q'\).

The new sequential probabilistic program is \( \mathcal{M}' = (Q', A', P', P_0', \mathcal{V}') \) and the formula \( \varphi' \) is obtained by substituting \( Xp \) by \( \xi \) in \( \varphi \). Similar as in case of \( \mathcal{T}_q \), we get

**Proposition 3.**

\[ \mu_{\mathcal{M}'}(L_{\mathcal{M}'}(\varphi')) = \mu_{\mathcal{M}}(L_{\mathcal{M}}(\varphi)) \]

Using similar arguments as in the case of \( \mathcal{T}_q \), we learn that \( \mathcal{T}_q \) can be carried out within the same time complexity of \( \mathcal{T}_q \).
The overall procedure If $\varphi$ has $k$ temporal operators, we can compute the measure $\mu_M(L_M(\varphi))$ as follows: We apply $k$ times the appropriate transformations $T_M$ or $T_K$ and obtain the sequence $\varphi^1, \mathcal{M}_1, \ldots, \varphi^k, \mathcal{M}_k$, where $\varphi^k$ is a simple propositional formula. Then $\mu_M(L_M(\varphi)) = \mu_M(L_M(\varphi^k))$, which is simply the sum of the initial probabilities in $\mathcal{M}_k$ over all states satisfying $\varphi^k$.

To conclude the proof of Theorem 1, it remains to show that our procedure can be carried out within the given time and space boundaries. As both transformations can be carried out in time linear in the size of the program (or polynomial if we are interested in the exact transition probabilities) and in every step the state space is at most doubled, we get $O(2^{\varphi} |\mathcal{M}|)$ as an upper bound for checking whether $\mathcal{M}$ satisfies $\varphi$. When we are interested in the exact measure, we need time polynomially in the size of $\mathcal{M}$ and exponentially in the size of $\varphi$. Note that it was shown in [16] that this problem is PSPACE-hard (wrt. the size of the formula). Thus, it is PSPACE-complete (as in the non-probabilistic case) and the given procedure is optimal.

It can be shown that the space required to carry out this computation is proportional to $|\varphi|(|\varphi| + \log^2(2^{\varphi}|\mathcal{M}|))$. Thus, the space complexity is $O(|\varphi| \log^2 |\mathcal{M}|)$.

Note that the presented procedure can easily be extended to deal with past tense connectives defined in [9]. Using the same ideas, one can achieve also the same time and space boundaries.

4 Model Checking Probabilistic Programs against Automata Specifications

Let $A$ be an $\omega$-automaton over $2^P$. Given a concurrent probabilistic program $\mathcal{M}$ over $\mathcal{P}$ with valuation function $\mathcal{V}$ and a scheduler $u$ of $\mathcal{M}$, we denote by $L_{\mathcal{M}, u}(A)$, similarly to the case of LTL, the set $\{x \in \Omega_{\mathcal{M}, u} \mid \mathcal{V}(x) \in L(A)\}$. In the same manner, for a sequential program $\mathcal{M}$ over $\mathcal{P}$ with valuation function $\mathcal{V}$, we define $L_{\mathcal{M}}(A)$ to be $\{x \in \Omega_{\mathcal{M}} \mid \mathcal{V}(x) \in L(A)\}$.

In the following, probabilistic programs will be defined over $\mathcal{P}$ while $\omega$-automata will be defined over $2^P$. It is a precondition for model checking programs against automata specifications that languages defined by automata are measurable. This is the case:

**Proposition 4** ([16]). For a concurrent probabilistic program $\mathcal{M}$, a scheduler $u$ of $\mathcal{M}$, and a Büchi automaton $A$, $L_{\mathcal{M}, u}(A) \in 2^{\Omega_{\mathcal{M}, u}}$.

Of course, the corresponding holds for sequential probabilistic programs.

Given a concurrent probabilistic program $\mathcal{M}$ and a Büchi automaton $A$, the concurrent emptiness problem (for automata) is to decide whether it holds $\mu_{\mathcal{M}, u}(L_{\mathcal{M}, u}(A)) = 0$ for all schedulers $u$ of $\mathcal{M}$. In contrast, the concurrent universality problem is to decide whether $\mu_{\mathcal{M}, u}(L_{\mathcal{M}, u}(A)) = 1$ for all schedulers $u$. Sequential emptiness and sequential universality for automata are defined analogously.
From the computational point of view, the above problems become easier to handle if the automaton specification at hand is a deterministic one. So, the transformation of a Büchi automaton into an equivalent $\omega$-automaton that behaves deterministically forms the basis of forthcoming algorithms.

**Theorem 2 (Safra [15]).** For a Büchi automaton $B$ over $\Sigma$ with $n$ states, there is a deterministic Streett (Rabin) automaton $A$ over $\Sigma$ with $2^{O(n \log n)}$ states and $O(n)$ pairs in the acceptance component such that $L(A) = L(B)$.

### 4.1 Checking Concurrent Probabilistic Programs

We assume an automata specification to represent undesired behaviour. Thus, we are interested in solving the emptiness problem for automata.

Let us make some assumptions which make life easier in the following. Without loss of generality, we can assume that, in a concurrent probabilistic program $M = (Q, \Sigma, R, A, P_0, P, \mathcal{V})$, it holds $Q \subseteq 2^P$ and, for each $q \in Q$, $\mathcal{V}(q) = q$. This can be achieved by sufficiently adding propositions to the program (to distinguish states) and accordingly adding transitions to the automaton. Then $M$ can easily be transformed into an equivalent (in a sense described below) concurrent probabilistic program $M'$

- that has no longer an initial probability distribution but a (randomizing) initial state,
- whose transitions are additionally labelled with elements of $2^P$, and
- that can be viewed as a quasi-deterministic $\omega$-automaton over $2^P$.

The transformation proceeds as follows:

1. Add a new randomizing (initial) state $q_M$ to $M$ and, for each state $q$ of $M$ with $P_0(q) > 0$, add a transition $q_M \rightarrow q$, for which we set $P'_{q_M, q} = P_0(q)$.
2. Including the new ones, label each transition $q \rightarrow q'$ with $\mathcal{V}(q')$.

For an example how to transform a probabilistic program, look at Figure 5 where, for the sake of clarity, we assume $a$ and $b$ to be (different) sets of propositions.

A trajectory of $M'$ is no longer a sequence of states but a sequence of transition labellings $\mathcal{V}(q)$ that $M'$ finds along a path starting at $q_M$. We will freely use the notation introduced for probabilistic programs also in the setting of such transformed programs giving them the obvious meaning. Regardless of the probabilities some transitions are equipped with, $M'$ can be viewed as a quasi-deterministic Büchi automaton with single initial state $q_M$ and acceptance component $Q$. The product of $M'$ and a deterministic $\omega$-automaton $A$, denoted by $M' \times A$, is defined as usual and at least quasi-deterministic. On the other hand, preserving the probabilities, we can view $M' \times A$ also as a concurrent probabilistic program where a state is defined to be nondeterministic or randomizing according to its first component and trajectories are sequences of transition labelings. More precisely, $(q, s) \rightarrow (q', s')$ becomes a transition of $M' \times A$ if $q \rightarrow q'$ and $s \rightarrow s'$ are $\mathcal{V}(q')$-labelled transitions of $M'$ and $A$, respectively. In that case,
Verifying Qualitative Properties of Probabilistic Programs

\[ (q, s) \rightarrow (q', s') \] is also labelled \( V(q') \) and, if \( q \) is randomizing, equipped with probability \( P_{q, q'} \). The product of the concurrent probabilistic program \( M \) and the deterministic Streett automaton \( A \) from Figure 5 with acceptance condition \( \{(s_0, s_2), \{s_1\}\} \) is illustrated in Figure 6 (note that, for the sake of clarity, the probabilities of transitions that go out from randomizing states are omitted).

\[ M' \times A \]

Fig. 6. The product of a program and an automaton

Given a probabilistic program \( M \) and a deterministic Streett automaton \( A \) with acceptance component \( F \), we want to mark some SCCs of \( M' \times A \) to be good in some sense. We call a set \( D \) of its states accepting if, for all pairs \( (U, V) \in F \), the following holds:

\[ \{ s \mid (q, s) \in D \text{ for some } q \} \cap U \neq \emptyset \] implies \( \{ s \mid (q, s) \in D \text{ for some } q \} \cap V \neq \emptyset \)

Otherwise, \( D \) is called rejecting. We say that a state \((r, f)\) of a rejecting set \( D \) is rejecting if there is a pair \((U, V) \in F\) such that \( f \in U \) and \( D \) contains no state \((q, s)\) with \( s \in V \).
The following proposition is crucial in this subsection and immediately leads to an algorithm that solves the concurrent emptiness problem for automata.

**Proposition 5.** For a concurrent program $\mathcal{M} = (Q, N, R, \Delta, P_0, P, \mathcal{V})$ and a deterministic Streett automaton $A = (S, \{s_0\}, \delta, \mathcal{F})$, there exists a scheduler $u$ of $\mathcal{M}$ with $\mu_{\mathcal{M}, u}(L_{\mathcal{M}, u}(A)) > 0$ iff there is a set of states of $\mathcal{M} \times A$ satisfying the following:

1. $(\mathcal{M} \times A)[D]$ is nontrivial and strongly connected,
2. $D$ is accepting and reachable from $(q_{\mathcal{M}}, s_0)$, and
3. for all transitions $(q, s) \rightarrow (q', s')$ of $\mathcal{M} \times A$ with $(q, s) \in D$ and $(q', s') \not\in D$, $(q, s)$ is nondeterministic, i.e., it holds $q \in N$.

**Proof.** It is easy to see that $\mu_{\mathcal{M}, u}(L_{\mathcal{M}, u}(A)) > 0$ for a scheduler $u$ of $\mathcal{M}$ iff there is a scheduler $u'$ of $\mathcal{M}'$ such that $\mu_{\mathcal{M}', u'}(L_{\mathcal{M}', u'}(A)) > 0$.

$(\Rightarrow)$ We fix a path $\beta \in ((Q \cup \{q_{\mathcal{M}}\}) \times S)^*$ through $\mathcal{M} \times A$ from $(q_{\mathcal{M}}, s_0)$ to a state of $D$. Let $\beta'$ be the projection of $\beta$ onto the first component. The scheduler $u'$ of $\mathcal{M}'$ satisfying $\mu_{\mathcal{M}', u'}(L_{\mathcal{M}', u'}(A)) > 0$ follows $\beta'$ taking $\mathcal{M} \times A$ from the initial state to $D$ and, henceforth, forces the trajectory both to stay within $D$ and to almost surely visit each state of $D$ infinitely often. This can be accomplished by, for a given nondeterministic state $(q, s)$, alternatingly choosing the transitions $(q, s) \rightarrow (q', s')$ of $\mathcal{M} \times A$ with $(q', s') \in D$ (recall that the history of a trajectory is at the scheduler's disposal.) Clearly, $\mu_{\mathcal{M}', u'}(L_{\mathcal{M}', u'}(A))$ is nonzero. Given $L_{\mathcal{M}', u'}(\beta')$, the conditional probability that $\mathcal{M}'$, wrt. $u'$, follows a trajectory that visits exactly the states of $D$ infinitely often is one. As such a trajectory is contained in $L_{\mathcal{M}', u'}(A)$, we conclude $\mu_{\mathcal{M}', u'}(L_{\mathcal{M}', u'}(A)) > 0$.

$(\Leftarrow)$ Note that a trajectory $\bar{x}$ of $\mathcal{M}'$ wrt. $u'$ unambiguously defines a path $\bar{x}$ through $\mathcal{M}' \times A$ starting from $(q_{\mathcal{M}}, s_0)$. This is due to the fact that $A$ is deterministic and that a transition of $\mathcal{M}'$ has a unique label. Let $D$ contain the subsets $D$ of states of $\mathcal{M} \times A$ such that $(\mathcal{M} \times A)[D]$ is strongly connected. Furthermore, for $D \in D$, let $E(D) := \{x \in \Omega_{\mathcal{M}', u'} \mid \inf(\bar{x}) = D\}$. Now suppose that $\mu_{\mathcal{M}', u'}(L_{\mathcal{M}', u'}(A)) > 0$ for a scheduler $u'$ of $\mathcal{M}'$. As

$$L_{\mathcal{M}', u'}(A) = \bigcup_{D \in D \text{ is accepting}} E(D),$$

we can find an accepting set $D \in D$ that satisfies $\mu_{\mathcal{M}', u'}(E(D)) > 0$. (Otherwise, the probability of the countable union $L_{\mathcal{M}', u'}(A)$ of events would be zero.) As $D$ is the infinity set of at least one infinite path through $\mathcal{M} \times A$ starting from $(q_{\mathcal{M}}, s_0)$, it is reachable from the initial state, thus satisfies condition (2), and forms a nontrivial (strongly connected) subgraph of $\mathcal{M}' \times A$, satisfying condition (1). Now suppose there is a transition $(q, s) \rightarrow (q', s')$ of $\mathcal{M} \times A$ with $(q, s) \in D$, $(q', s') \not\in D$, and $q \in R$. As, for every trajectory $\bar{x} \in E(D)$, $\bar{x}$ visits $(q, s)$ infinitely often (and each time the probability to exit $D$ is nonzero), it will almost surely leave $D$ infinitely often so that we have $\mu_{\mathcal{M}', u'}(E(D)) = 0$ contradicting our assumption. It follows that $D$ also satisfies condition (3) from Proposition 5, which concludes our proof. □
Note that, in the above proof, we explicitly make use of the fact that a trajectory of $\mathcal{M}'$ determines exactly one corresponding run of $\mathcal{M}' \times A$ starting from $(q_M, s_0)$ (recall that $A$ is deterministic).

Let us fall back on the example contributed by Figures 5 and 6, respectively. There is an accepting set $D = \{(q_1, s_1), (q_2, s_0)\}$ of states of $\mathcal{M}' \times A$ that is reachable from $(q_M, s_0)$ via $\beta = (q_M, s_0)(q_1, s_1)$ and forms a nontrivial and strongly connected subgraph of $\mathcal{M}' \times A$. The only transition leading from a state of $D$ to a state outside of $D$ is $(q_1, s_1) \rightarrow (q_1, s_2)$, thus, arises from a nondeterministic state. In fact, for the scheduler $u'$ of $\mathcal{M}'$ with $u'(x) = q_2$ for every $x \in (Q \cup \{q_M\})^*$ that ends in a nondeterministic state (i.e. in $q_1$), it holds $\mu_{\mathcal{M}', u'}(L_{\mathcal{M}', u'}(A)) > 0$.

Based on Proposition 5, we now provide an algorithm that solves the concurrent emptiness problem, i.e., decides, given a concurrent probabilistic program $\mathcal{M}$ and a Büchi automaton $B$, whether there is a scheduler $u$ of $\mathcal{M}$ such that $\mu_{\mathcal{M}, u}(L_{\mathcal{M}, u}(B)) > 0$:

1. From $B$, build a deterministic Streett automaton $A$ with $L(A) = L(B)$.
2. Compute (the graph of) $\mathcal{M}' \times A$, remove those states that are not reachable from $(q_M, s_0)$, and let $G$ denote the resulting graph.
3. Repeat
   (a) Determine the sets $\mathcal{A}C$ of nontrivial and accepting and $\mathcal{R}C$ of nontrivial and rejecting SCCs of $G$, respectively.
   (b) For each $C \in \mathcal{R}C$, remove the transitions going out from rejecting states.
   (c) For each $C \in \mathcal{A}C$, do the following:
      i. Find the set $H$ of states $(q, s) \in C$ with randomizing $q$ from where there is a transition leaving $C$.
      ii. If $H$ is the empty set, then return “Yes” (thus, halt with success).
         Otherwise, remove the transitions going out from the states of $H$.
   until $\mathcal{A}C \cup \mathcal{R}C = \emptyset$.
4. Return “No”.

Obviously, the algorithm terminates. Furthermore, it returns the answer “Yes” iff there is a scheduler $u$ of $\mathcal{M}$ such that $\mu_{\mathcal{M}, u}(L_{\mathcal{M}, u}(B)) > 0$. Let us discuss its complexity in the following. Starting from a Büchi automaton with $n$ states, construct an equivalent deterministic Streett automaton with $2^{O(n \log n)}$ states and $O(n)$ pairs in the acceptance component. Assume that $\mathcal{M}'$ has $m$ states. The number of states of $\mathcal{M}' \times A$ is not greater than $m \cdot 2^{O(n \log n)}$. Thus, steps (a), (b), and (c) are repeated at most $m \cdot 2^{O(n \log n)}$-times, respectively. Determining the SCCs of $G$ can be done in time linear in the size of $\mathcal{M}' \times A$. Overall, the algorithm runs in time $O(m^3 \cdot 2^{O(n \log n)})$, i.e., it is quadratic in $|M|$ and exponential in $|A|$.

**Theorem 3.** Given a concurrent probabilistic program $\mathcal{M}$ and a Büchi automaton $A$, the concurrent emptiness problem can be solved in time $O(|\mathcal{M}|^2 \cdot 2^{O(|A|)})$.

The complexity of the above algorithm can be improved by constructing, instead of a (fully) deterministic $\omega$-automaton, a Büchi automaton that is deterministic in the limit [4], i.e., that behaves deterministically as soon as it reaches...
a final state. Such an automaton only requires $2^{O(n)}$ states and finally leads to an $O(m^3 \cdot 2^{O(n)})$-complexity of the algorithm. But as it does not provide further insights in the principles of probabilistic model checking, we omit the corresponding construction. Courcoubetis and Yannakakis furthermore show that to decide the concurrent emptiness problem requires exponential time in the total input size $|\mathcal{M}| + |\mathcal{A}|$ [4].

The concurrent universality problem for automata can be solved analogously, but, from the computational point of view, it seems to be more appropriate to construct a deterministic Rabin automaton from a Büchi automaton. Accordingly, we call a set $D$ of states of $\mathcal{M} \times \mathcal{A}$ accepting if there is a pair $(U, V)$ from the acceptance component of $\mathcal{A}$ and a state $(r, f) \in D$ such that $f \in U$ and $D$ contains no state $(q, s)$ with $s \in V$. Otherwise, we call $D$ rejecting.

**Proposition 6.** Given a concurrent program $\mathcal{M} = (Q, N, R, \Delta, P_0, P, V)$ and a deterministic Rabin automaton $\mathcal{A} = (S, \{s_0\}, \delta, F)$, there is a scheduler $u$ of $\mathcal{M}$ with $\mu_{\mathcal{M}, u}(L_{\mathcal{M}, u}(A)) < 1$ if there is a set $D$ of states of $\mathcal{M} \times \mathcal{A}$ satisfying the following:

1. $(\mathcal{M} \times \mathcal{A})[D]$ is nontrivial and strongly connected,
2. $D$ is rejecting and reachable from $(q_0, s_0)$, and
3. for all transitions $(q, s) \rightarrow (q', s')$ of $\mathcal{M} \times \mathcal{A}$ with $(q, s) \in D$ and $(q', s') \notin D$, it holds $q \in N$.

Proving Proposition 6 mainly applies techniques used to prove Proposition 5. The algorithm for the concurrent universality problem as suggested by Proposition 6 runs in time $O(|\mathcal{M}|^2 \cdot 2^{O(|\mathcal{A}|)})$.

However, to solve the concurrent universality problem, our algorithm cannot fall back on using ω-automata that are deterministic in the limit for the following reason: We then cannot be sure that a path leading to $D$ does not have an “accepting” counterpart elsewhere. In [4], it is mentioned that the concurrent universality problem for Büchi automata that are deterministic in the limit requires exponential time.

### 4.2 Checking Sequential Probabilistic Programs

Sequential probabilistic programs were defined as a special case of concurrent ones. Thus, the above algorithms carry over to the case the probabilistic program has no nondeterministic state. Let us formulate an algorithm explicitly for this case. Two basic properties wrt. a sequential probabilistic program $\mathcal{M}$ and a deterministic ω-automaton $\mathcal{A}$ will play a crucial role in both qualitative and quantitative analysis:

- Almost surely, a trajectory of $\mathcal{M}'$ will take $\mathcal{M}' \times \mathcal{A}$ to a bottom SCC.
- Almost surely, a trajectory of $\mathcal{M}'$ that leads $\mathcal{M}' \times \mathcal{A}$ into a bottom SCC $D$ will visit each state of $D$ infinitely often.
Wrt. the emptiness problem, a sequential probabilistic program does not satisfy its automata specification (and thus contains undesired behaviour), iff, within the product of the transformed program and the specification, one can find an accepting bottom SCC that is reachable from the initial state. The following characterization immediately follows from Proposition 5.

**Proposition 7.** Given a sequential probabilistic program \( M \) and a deterministic Streett automaton \( A \), it holds \( \mu_M(L_M(A)) > 0 \) iff there is a reachable (from the initial state) and accepting bottom SCC of \( M' \times A \).

Thus, after having computed a deterministic \( \omega \)-automaton and a product automaton, the algorithm for solving the sequential emptiness problem only needs to determine whether there is an accepting bottom SCC that is reachable from the initial state. This can be done in time linear in the size of the product automaton. The algorithm works in time \( O(m^2 \cdot 2^{O(n \log n)}) \) where \( m \) is assumed to be the number of states of the program and \( n \) the number of states of the Buchi automaton. Again, the usage of \( \omega \)-automata that are deterministic in the limit reduces the time complexity to \( O(m^2 \cdot 2^{O(n)}) \).

Checking universality wrt. sequential programs is closely related to checking emptiness.

**Proposition 8.** For a sequential probabilistic program \( M \) and a deterministic Streett automaton \( A \), we have \( \mu_M(L_M(A)) = 1 \) iff all the reachable bottom SCCs of \( M' \times A \) are accepting.

**Proof.** Clearly, \( \mu_M(L_M(A)) = 1 \) iff \( \mu_{M'}(L_{M'}(A)) = 1 \).

(\( \iff \)) Almost surely, a trajectory of \( M' \) will lead \( M' \times A \) into a bottom SCC. Let \( x \) be a finite sequence of states of \( M' \) that takes \( M' \times A \) from the initial state to a bottom SCC \( D \). As \( D \) is nontrivial and accepting, \( \mu_M'(L_{M'}(A) \mid \mathcal{C}_{M'}(x)) = 1 \). Overall, we get \( \mu_{M'}(L_{M'}(A)) = 1 \).

(\( \Rightarrow \)) Assume that \( \mu_{M'}(L_{M'}(A)) = 1 \) and suppose there is a reachable bottom SCC \( D \) of \( M' \times A \) that is rejecting. We can find a finite sequence \( x \) of states of \( M' \) with \( \mu_{M'}(\mathcal{C}_{M'}(x)) > 0 \) that takes \( M' \times A \) from the initial state to \( D \). As \( D \) is rejecting, \( \mu_{M'}(L_{M'}(A) \mid \mathcal{C}_{M'}(x)) = 0 \). Thus, with nonzero probability, \( M' \) will produce a trajectory that is not contained in \( L_M(A) \), contradicting the assumption. \( \square \)

For example, the product of a (transformed) sequential probabilistic program \( M' \) and a deterministic Streett automaton \( A \) may have the structure as shown in Figure 7 where the bottom SCCs \( \{7_4\} \) and \( \{7_5\} \) are both assumed to be accepting. As \( \{7_4\} \) and \( \{7_5\} \) are the only bottom SCCs, it follows from Proposition 8 that \( \mu_M(L_M(A)) = 1 \).

### 4.3 Quantitative Analysis

So far, we provided algorithms for checking probabilistic programs against qualitative properties, i.e., that do not compute the exact probability of satisfaction.
We have seen that the algorithms amount to simple graph analyses as satisfaction of qualitative properties does not depend on exact probabilities but only on the topology of the respective program and its specification at hand. In the following, we generalize the above results.

**Proposition 9.** Given a sequential probabilistic program $\mathcal{M}$ and a deterministic Streett automaton $A$, $\mu_M(L_M(A))$ is equal to the probability that a trajectory of $\mathcal{M}'$ takes $\mathcal{M}' \times A$ from the initial state to an accepting bottom SCC.

To determine $\mu_M(L_M(A))$, we compute, starting from $\mathcal{M}' \times A$, a new sequential probabilistic program $abs(\mathcal{M}' \times A)$ by

- removing all states that lie in a bottom SCC $D$ and do not have an immediate predecessor outside of $D$,
- removing all transitions inside a bottom SCC, and
- adding, for each remaining state $(q, s)$ of a bottom SCC, a transition from $(q, s)$ to itself, which is equipped with probability 1, respectively.

The resulting program $abs(\mathcal{M}' \times A)$ is absorbing, i.e., each of its bottom SCCs consists of a single absorbing state. Employing the fundamental matrix $N$ of $abs(\mathcal{M}' \times A)$ [6], we can easily compute the probability that a trajectory of $abs(\mathcal{M}' \times A)$ ends in an absorbing state. Suppose $\{q_0, \ldots, q_{n-1}\}$ is the set of states of $abs(\mathcal{M}' \times A)$ with only initial state $q_0$. Without loss of generality, $\{q_0, \ldots, q_{m-1}\}$ is the set of transient states for a natural $m < n$. Let $T = (t_{ij})_{i,j=0,\ldots,n-1}$ be the transition matrix of $abs(\mathcal{M}' \times A)$, i.e., $t_{ij}$ is the transition probability of going from $q_i$ to $q_j$. It is of the form

$$T = \begin{pmatrix} Q & R \\ 0 & I_{n-m} \end{pmatrix}$$

where $I_{n-m}$ is the $(n-m) \times (n-m)$ matrix that has 1’s on the main diagonal and 0’s elsewhere, $Q$ is the matrix of dimension $(n-m) \times m$ that has all components 0, the $m \times (n-m)$ matrix $R$ gives information about the transition probabilities of going from a transient to an absorbing state, and $Q$ is of dimension $m \times m$ and contains the transition probabilities exclusively respecting the set of transient states. For example, assume $abs(\mathcal{M}' \times A)$ to be partly shown in Figure 7 (note
that, in fact, \( \text{abs}(\mathcal{M}' \times \mathcal{A}) \) is absorbing) and suppose the only state that emanates from an accepting bottom SCC is \( q_4 \). Its transition matrix is given by

\[
T = \begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.6 & 0.4 & 0 & 0 \\
0 & 0 & 0 & 0.8 & 0.2 & 0 \\
0 & 0.5 & 0 & 0 & 0.5 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}.
\]

The fundamental matrix of \( \text{abs}(\mathcal{M}' \times \mathcal{A}) \) now arises as follows:

\[
N = (n_{ij})_{i,j \in \{0, \ldots, m-1\}} = (I_m - Q)^{-1}
\]

Hereby, \( n_{ij} \) tells us how often a trajectory of \( \text{abs}(\mathcal{M}' \times \mathcal{A}) \) starting at \( q_i \) is expected to be in \( q_j \) before absorption. In our example,

\[
Q = \begin{pmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 0.6 & 0.4 \\
0 & 0 & 0 & 0 \\
0 & 0.5 & 0 & 0
\end{pmatrix}, \quad I_4 - Q = \begin{pmatrix}
1 & -1 & 0 & 0 \\
0 & 1 & -0.6 & -0.4 \\
0 & 0 & 1 & 0 \\
0 & -0.5 & 0 & 1
\end{pmatrix}, \quad \text{and}
\]

\[
N = (I_4 - Q)^{-1} = \begin{pmatrix}
1 & 1.25 & 0.75 & 0.5 \\
0 & 1.25 & 0.75 & 0.5 \\
0 & 0 & 1 & 0 \\
0 & 0.625 & 0.375 & 1.25
\end{pmatrix}
\]

Finally,

\[
B = (b_{ij})_{i \in \{0, \ldots, m-1\}, j \in \{0, \ldots, n-m-1\}} = NR
\]

contains the probabilities \( b_{ij} \) that \( \text{abs}(\mathcal{M}' \times \mathcal{A}) \), starting at \( q_i \), reaches the absorbing state \( q_{j+m} \). Let \( A \) be the subset of \( \{m, \ldots, n-1\} \) such that \( q_j \) is an absorbing state of \( \text{abs}(\mathcal{M}' \times \mathcal{A}) \) that emanates from an accepting bottom SCC of \( \mathcal{M}' \times \mathcal{A} \) iff \( j \in A \). If \( A \) is empty, \( \mu_{\mathcal{M}}(L_{\mathcal{M}}(\mathcal{A})) = 0 \). Otherwise,

\[
\mu_{\mathcal{M}}(L_{\mathcal{M}}(\mathcal{A})) = \sum_{j \in A} b_{0(j-m)}.
\]

To complete our example,

\[
B = NR = \begin{pmatrix}
1 & 1.25 & 0.75 & 0.5 \\
0 & 1.25 & 0.75 & 0.5 \\
0 & 0 & 1 & 0 \\
0 & 0.625 & 0.375 & 1.25
\end{pmatrix} \begin{pmatrix}
0 & 0 \\
0 & 0 \\
0.8 & 0.2 \\
0 & 0.5
\end{pmatrix} = \begin{pmatrix}
0.6 & 0.4 \\
0.6 & 0.4 \\
0.8 & 0.2 \\
0.3 & 0.7
\end{pmatrix}
\]

lets us come to the conclusion that the probability the underlying sequential probabilistic program will produce a trajectory accepted by the automata specification is 0.6.

Under consideration of the matrix operations, we get an algorithm that determines the exact probability of satisfaction in time single exponential in the size of the Büchi automaton \( \mathcal{B} \) and polynomial in the size of the sequential probabilistic program \( \mathcal{M} \):
1. Construct a deterministic Streett automaton \( A \) with \( L(A) = L(B) \).
2. Build \( M' \times A \) and determine its accepting bottom SCCs.
3. Compute \( \text{abs}(M' \times A) \) as well as its fundamental matrix \( N \).
4. Compute \( B = NR \) and, within the first of its rows, sum up the probabilities that belong to states respectively emanating from an accepting bottom SCC of \( M' \times A \).

We can summarize the above observations as follows:

**Theorem 4.** For a sequential probabilistic program \( M \) and a Büchi automaton \( A \), sequential emptiness and sequential universality can be solved in time \( O(|M| \cdot 2^{O(|A|)}) \), respectively. Furthermore, we are able to determine the exact probability \( \mu_M(L_M(A)) \) in time single exponential in the size of \( A \) and polynomial in \( |M| \).

Note that both sequential emptiness and sequential universality for automata are PSPACE-complete in the size of the automaton ([16], [4]).

## 5 Conclusion

In this chapter, we presented procedures for checking linear temporal logic specifications and automata specifications of sequential and concurrent probabilistic programs. The complexities of these problems are summarized in Table 1. We followed two different approaches: For LTL and sequential probabilistic programs, our method proceeded in a tableau style manner, while for the remaining problems, we followed the automata theoretic approach.

<table>
<thead>
<tr>
<th></th>
<th>LTL formulas</th>
<th>( \omega )-automata</th>
</tr>
</thead>
<tbody>
<tr>
<td>sequential</td>
<td></td>
<td>( O(</td>
</tr>
<tr>
<td>concurrent</td>
<td></td>
<td>( O(</td>
</tr>
</tbody>
</table>

Table 1. Complexity of model checking problems

To gain further insight of probabilistic programs, one might be interested in specifying boundaries for the probabilities of properties. These extension to temporal logics are studied in the next chapters.

## 6 Bibliographic remarks and further reading

Initial papers studying temporal logics for probabilistic programs are [8] and [5]. Concurrent probabilistic programs were first studied in [16].

Our chapter is mainly based on [4]. The model checking procedure for LTL and sequential probabilistic programs (Theorem 1) is presented in [4]. PSPACE-hardness (wrt. length of formula) of this problem was first shown in [16]. Our algorithms for checking automata specifications are presented or at least suggested
Verifying Qualitative Properties of Probabilistic Programs

in [4]. Note that for solving the sequential emptiness problem, the authors make use of a slightly different kind of product construction. First complexity results were achieved by Vardi [16] and extended by [4].

A different approach to study probabilistic systems is to understand randomness as a kind of fairness. Let us reconsider the example shown in Figure 1, and recall, that nondeterministic states are denoted by circles while probabilistic states are drawn as squares. Intuitively, nondeterminism abstracts that it is not known which possible transition the system under consideration takes. For example, it is possible that the given system chooses to move to state $q_3$ whenever it is in state $q_1$. For random states, however, we know that the system chooses a transition with a fixed probability. For example, with probability 0.3, the transition from $q_3$ to $q_1$ is taken. This implies that it is unlikely (with probability 0) that the system will choose transition $q_3$ to $q_1$ infinitely often. In almost all runs, the system will be fair and eventually take the other possible outcome in state $q_3$ and move to state $q_5$ (if it moves to state $q_3$ at all).

The idea of using concepts of fairness to study probabilistic systems first appeared in [11] where the notion of extreme fairness was introduced. In [13], $\alpha$-fairness was defined to deal with checking a restricted version of linear temporal logic (but extended with past-tense modalities) specifications of concurrent probabilistic programs. A notion of $\gamma$-fairness is introduced in [1] to analyze parameterized probabilistic systems.

References


Note that Pnueli follows a different model of sequential and concurrent probabilistic programs. Thus, [13] speaks of sequential programs which correspond to concurrent programs in our terminology.