Model checking conditional CSL for continuous-time Markov chains

Gao, Yang; Xu, Ming; Zhan, Naijun; Zhang, Lijun

Published in:
Information Processing Letters

Link to article, DOI:
10.1016/j.ipl.2012.09.009

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Model checking conditional CSL for continuous-time Markov chains

Yang Gao\textsuperscript{a,1}, Ming Xu\textsuperscript{b,2}, Naijun Zhan\textsuperscript{a,1}, Lijun Zhang\textsuperscript{c,*,3}

\textsuperscript{a} State Key Lab. of Comp. Sci., Institute of Software, Chinese Academy of Sciences, China
\textsuperscript{b} Department of Computer Science and Technology, East China Normal University, China
\textsuperscript{c} Technical University of Denmark, DTU Informatics, Denmark

\textbf{ABSTRACT}

In this paper, we consider the model-checking problem of continuous-time Markov chains (CTMCs) with respect to conditional logic. To the end, we extend Continuous Stochastic Logic introduced in Aziz et al. (2000)\cite{1} to Conditional Continuous Stochastic Logic (CCSL) by introducing a conditional probabilistic operator. CCSL allows us to express a richer class of properties for CTMCs. Based on a parameterized product obtained from the CTMC and an automaton extracted from a given CCSL formula, we propose an approximate model checking algorithm and analyse its complexity.

\begin{quote}
Crown Copyright © 2012 Published by Elsevier B.V. All rights reserved.
\end{quote}

1. Introduction

Continuous-time Markov chains (CTMC) have received considerable attentions in network performance analysis, model checking, and system biology. In [1], Continuous Stochastic Logic (CSL) has been introduced, that has been widely used to specify properties over CTMCs.

In the paper [1], Aziz et al. focused on the decidability of the model-checking of CSL. Later, Baier et al. [2] presented an approximate model checking algorithm for the case restricted to binary until formulas. Recently, the approximate algorithm has been extended to handle nested until formulas in [3]. The main idea is to exploit the notion of stratified CTMCs, which are a subclass of CTMCs that have the nice feature allowing one to obtain the desired probability using a sequence of transient analysis. Then, the product of the CTMC and a deterministic finite automaton (DFA) obtained from the nested until formula is constructed, which is guaranteed to be stratified by construction. The product CTMC can then be analyzed efficiently, in a similar manner as the approach in [2].

In this paper, we propose the conditional continuous stochastic logic (CCSL), an extension of CSL with a conditional probabilistic operator. CCSL allows one to express a richer class of properties, such as:

\[ P \geq 0.1 \land \left[ 10, 20 \right] f \land g \lor f_1 U \left[ 10, 20 \right] f_2 U \left[ 10, 20 \right] \cdots f_k \]

Such property can be expressed as a state formula of the form \( P_{\geq 0.1} \left( \phi_{\left[ 10, 20 \right]} \right) \land g \lor f_1 U \left[ 10, 20 \right] f_2 U \left[ 10, 20 \right] \cdots f_k \) where \( f, g, f_1, \ldots, f_k \) are appropriate atomic propositions. We believe that such conditional properties are an important extension because of the important role of conditional probabilities in stochastic models [4].

Essentially, the model checking for the conditional probabilistic operator deals with binary conjunction of CCSL path formulas, which is not allowed in the classical CSL, see [1,2]. Thus, in this paper, we extend the logic CSL.
with binary conjunction and disjunction operators for the path formulas.

We discuss how to compute the probability of a conjunctive path formula, and then present an approximate model checking algorithm, following the approach in [3]. First, a DFA $\mathcal{A}_\psi$ is constructed for a CCSL path formula $\psi$. The next step is to construct the DFA $\mathcal{A}_{\psi,\psi'}$ for the conjunction from the automata $\mathcal{A}_\psi$. A plain product construction turns out to be insufficient: We have to pay special attention to whether some conjuncts of the formula have been satisfied. We propose a notion of parameterized product construction. The probability is then computed on this product. The size of the automaton could be exponential in the number of binary operators in the path formulas, arising from the product construct, and the approximation calculation for the transient probability distributions is linear in the size of the product.

Related work. There is a rich literature on model checking techniques for CTMCs, see [1,2,5-8]. In [5,6], deterministic timed automata (DTA) are used for specifying path properties. As discussed in [9], nested until CSL path formulas can be expressed in DTA as well, however with a much larger number of states. Real-time is considered in [8], with exponential complexity both in the size of the formula and the size of the product. The control prob is then computed on this product construct, and the approximation calculation for the transient probability distributions is linear in the size of the product.

2. Preliminaries

In this section, we define some basic notions that will be used later. For details please refer to [2]. For convenience, we fix a set of propositions AP in the sequel, ranged by $f_1, f_2, g_1, g_2, \ldots$.

Definition 1. A labeled continuous-time Markov chain (CTMC) is a tuple $\mathcal{C} = (S, R, L, \alpha)$ where $S$ is a finite set of states, $\alpha : S \rightarrow [0, 1]$ is the initial distribution satisfying $\sum_{s \in S} \alpha(s) = 1$, $R : S \times S \rightarrow \mathbb{R}_{\geq 0}$ is a rate matrix, and $L : S \rightarrow 2^{AP}$ is a labeling function.

For $A \subseteq S$, define $R(s, A) := \sum_{s' \in A} R(s, s')$. We denote the exit rate of $s$ by $E(s) := R(s, S)$. A state $s$ is called absorbing if $E(s) = 0$. If $R(s, s') > 0$, we say that there is a transition from $s$ to $s'$.

Consider the CTMC in Fig. 1. If $s_1$ is the current state of the CTMC, the probability that some transition will be triggered within time $t$ is $1 - e^{-\lambda t}$. Furthermore, there is a competition between the transitions to $s_2$ and $s_3$: the probability to take the transition to $s_2$ is $R(s_1, s_2) \cdot (1 - e^{-\lambda t})$. The labeling function $L$ assigns to each state $s$ a set of atomic propositions $L(s) \subseteq AP$ which are valid in $s$.

Transient probability. Starting with distribution $\alpha$ of $\mathcal{C}$, the transient probability vector at time $t$, denoted by $\pi^S(\alpha, t)$, is the probability distribution over states at time $t$. For $t > 0$, the transient probability [11] is given by: $\pi^S(\alpha, t) = \pi^S(\alpha, 0) e^{E t}$ where $Q := R - \text{diag}(E)$ is the generator matrix and $\text{diag}(E)$ denotes the diagonal matrix with $\text{diag}(E)(s, s) = E(s)$.

Paths and probabilistic measures. A right continuous step function $\rho : \mathbb{R}_{\geq 0} \rightarrow S$ is called a step function (or an infinite (sample) path), where $\rho(t)$ stands for the state at $t$. For a given step function $\rho$ and $i \in \mathbb{N}$, we denote by $\rho[s_i] = s_j$ the state at the $t$-th step, and by $\rho[t_i]$ the time spent at $\rho[s_i]$, i.e., the length of the step segment starting with $\rho[s_i]$. Let Path$^C$ be the set of all infinite paths, and Path$^C_i(s)$ be the subset of those paths starting from $s$.

Let $I_0, \ldots, I_{k-1}$ be nonempty intervals in $\mathbb{R}_{\geq 0}$. The cylinder set $\text{Cyl}(s_0, I_0, s_1, I_1, \ldots, s_{k-1}, I_{k-1}, s_k)$ is defined by:

$$\{\rho \in \text{Path}^C | \forall 0 \leq i \leq k, \rho[s_i] = s_i \land \forall 0 \leq i < k, \rho[t_i] \in I_i\}.$$  

Let $\mathcal{F}(\text{Path}^C)$ denote the smallest $\sigma$-algebra on $\text{Path}^C$ containing all cylinder sets. For initial distribution $\alpha : S \rightarrow [0, 1]$, a probability measure (denoted $\text{Pr}_i^C$) on this $\sigma$-algebra is introduced as follows: $\text{Pr}_i^C$ is the unique measure that satisfies: $\text{Pr}_i^C(\text{Cyl}(s)) = \alpha(s)$, and for $k > 0$, $\text{Pr}_i^C(\text{Cyl}(s_0, I_0, \ldots, s_{k-1}, s_k))$ equals

$$\text{Pr}_i^C(\text{Cyl}(s_0, I_0, \ldots, s_{k-1})) \cdot \frac{R(s_{k-1}, s_k)}{E(s_{k-1})} \cdot \eta(I_{k-1}),$$

where $\eta(I_{k-1}) := e^{-E(s_{k-1})\inf I_{k-1}} - e^{-E(s_{k-1})\sup I_{k-1}}$ is the probability to take a transition during $I_{k-1}$. If $\alpha(s) = 1$ for some state $s \in S$, we sometimes simply write $\text{Pr}_i^C$ instead of $\text{Pr}_i^C$.

3. Conditional continuous stochastic logic (CCSL)

This section is devoted to introducing a conditional continuous stochastic logic (CCSL) by extending the Continuous Stochastic Logic (CSL) introduced by Aziz et al. [1] with a conditional probabilistic operator. Let $I_i$ be a nonempty left-closed and right-open interval on $\mathbb{R}_{\geq 0}$. Let $\Delta := (\ll, \leq, \geq, >, 0 < p < 1, \text{and } K > 1)$. The syntax of CCSL is defined as:
\( \Phi := f \mid \neg \Phi \mid \Phi \land \Phi \mid \mathcal{P}_{\text{ap}}(\varphi) \mid \mathcal{P}_{\text{ap}}(\varphi \mid \varphi) \)

\( \varphi := \varphi \land \varphi \mid \varphi \lor \varphi \mid \Phi_1 \ U_1 \Phi_2 \ U_2 \cdots \Phi_K \)

where \( f \in \text{AP} \) is an atomic proposition. The syntax of CCSL consists of state formulas and path formulas. We use \( \Phi, \varphi \) and their indexed versions for state formulas. The path formula \( \Phi_1 \ U_1 \Phi_2 \ U_2 \cdots \Phi_K \) with \( K > 1 \) is referred to as the \textit{atomic path formula}. Obviously, each path formula can be expressed into a disjunctive normal form (DNF) \( \varphi = V \cup / \cap \psi \) where \( \psi \) are atomic path formulas. We use \( \psi \) for atomic path formulas and \( \varphi \) for general path formulas in DNF.

Let \( \mathcal{C} = (S, R, L, a) \) be a CTMC with \( s \in S \). The semantics of CCSL state formulas is standard: \( s \models \text{true} \) for all \( s \in S \), \( s \models a \) if \( a \in L(s) \), \( s \models \neg \Phi \) iff \( s \not\models \Phi \), \( s \models \Phi \land \Psi \) iff \( s \models \Phi \) and \( s \models \Psi \). For probabilistic formulas, we have:

\[
s \models \mathcal{P}_{\text{ap}}(\varphi) \iff Pr_s((\rho \in \text{Path} \mid \rho \models \varphi)) = p, \\
s \models \mathcal{P}_{\text{ap}}(\varphi_1 \mid \varphi_2) \iff Pr_s((\rho \in \text{Path} \mid \rho \models \varphi_1 \land \varphi_2)) = p, \\
\]

where \( Pr_s((\rho \in \text{Path} \mid \rho \models \varphi)) \), or \( Pr_s(\varphi) \) for short, denotes the probability measure of the set of all paths which start from \( s \) and satisfy \( \varphi \). Similarly, \( Pr_s(\varphi_1 \mid \varphi_2) \) denotes the conditional probability \( \frac{Pr_s(\varphi_1 \land \varphi_2)}{Pr_s(\varphi_2)} \) under the premise \( Pr_s(\varphi_2) \neq 0 \).

The semantics for the Boolean operators is standard, and the semantics of the atomic path formula is given by [1,3]:

\[
\rho \models \varphi = \Phi_1 \ U_1 \Phi_2 \ U_2 \cdots \Phi_K \text{ iff there exist real numbers } 0 < t_1 \leq t_2 \leq \cdots \leq t_{K-1} \text{ such that } \rho(t_{K-1}) = \Phi_K, \text{ and for each integer } 0 < i < K \text{ we have } (t_i \in I_i) \land (\forall t' \in [t_{i-1}, t_i). \rho(t') = \Phi_i) \text{, where } t_0 \text{ is } \text{defined to be } 0 \text{ for notational convenience.}
\]

4. Model checking algorithm for CCSL

In this section, we present an algorithm for checking CCSL properties. We first recall the deterministic finite automaton construction for the atomic path formula \( \psi \).

Then, we extend the construction to the conjunctive path formula by introducing the notion of a parameterized product construction for the given CTMC and the conjunctive path formula. This is the key for computing the probability of the set of paths satisfying the conjunctive path formula. We further show how to compute the probabilities of general path formulas. Finally, we describe an algorithm for model checking CCSL and analyze its complexity.

In the rest of the paper, let \( \psi^i = f_1^i \ U_1^i f_2^i \ U_2^i \cdots f_k^i \) with \( i = 1, \ldots, n \) be \( n \) special atomic path formulas. For simplicity, as in [3,9] we assume that \( a_k^i \leq a_i^j \) and \( b_k^i \leq b_i^j \) for any two intervals \( I_k^i = [a_k^i, b_k^i] \) and \( I_i^j = [a_i^j, b_i^j] \) with \( k < i \); ii) all \( f_k^i \) are pairwise distinct for \( i = 1, \ldots, n \) and \( k = 1, \ldots, K_i \). We will drop the subscript in case \( n = 1 \).

4.1. Formula automata

In this subsection, we recall how to construct a deterministic finite automaton (DFA) for \( \bigwedge_{i=1}^n \psi^i \). Firstly, we consider the simple case when \( n = 1 \). So, the atomic path formula \( \psi^i \) describes the required order of \( f_1^i, \ldots, f_k^i \) states.

**Definition 2** (Atomic path formula automaton). (See [3].) The atomic path formula automaton \( A_\psi = (\Sigma, Q, q_{in}, \delta, F) \) is defined as follows:

- \( \Sigma = 2\{f_1, \ldots, f_k\} \).
- \( Q = \{q_1, \ldots, q_K, \bot\} \) with \( q_{in} = q_1 \) and \( F = \{q_1, \ldots, q_K\} \).
- For every \( a \in \Sigma \), the transition relation \( \delta \) is given by \( \delta(q_K, a) = q_K \), \( \delta(\bot, a) = \bot \), and for the rest \( q_i \in Q \setminus \{q_K, \bot\} \).

\[
\delta(q_i, a) = \begin{cases} 
q_j & \text{if } j \geq i \land f_i \neq a \lor f_j \in a; \\
\bot & \text{otherwise.}
\end{cases}
\]

Both \( q_K \) and \( \bot \) are absorbing states, i.e., with only transitions leading to themselves. The former state is referred to a good absorbing state, the latter a bad absorbing state.

The words accepted by \( A_\psi \) are finite traces \( w \in \Sigma^* \), such that they can be extended to a trace \( w \psi' \in \Sigma^{\omega} \) that satisfies the time–abstract (LTL) formula of the form \( f_1 U (f_2 U (\cdots (f_{k-1} U f_k) \cdots)) \).

Transitions in \( A_\psi \) go always from lower goal states to higher goal states. The good state \( q_K \) implies that any path traversing it satisfies the atomic path formula \( \psi \) under suitable timing constraint; while the bad state \( \bot \) implies that any path traversing it refutes the atomic path formula \( \psi \).

Below we define the automaton for the conjunction of several atomic path formulas, which is essentially the product construction.

**Definition 3** (Conjunctive path formula automaton). Let \( \psi = \bigwedge_{i=1}^n \psi^i \) and \( A_{\psi_i} = (\Sigma, Q_i, q_{in}^i, \delta_i, F_i) \) be the formula automaton for \( \psi^i \) respectively for \( i = 1, \ldots, n \). Then the conjunctive path formula automaton \( A_{\psi} = (\Sigma, Q, q_{in}, \delta, F) \) is defined as follows:

- \( \Sigma = 2\{f_1^1, \ldots, f_k^1\} \).
- \( Q = \bigcup_{i=1}^n (Q_i \times \cdots \times Q_n) \) with \( q_{in} = (q_{in}^1, \ldots, q_{in}^n) \), and \( F = F_i^1 \times \cdots \times F_n^i \).
- \( \delta((q_1^1, \ldots, q_n^1), a) = (\delta(q_1^1, a^1), \ldots, \delta(q_n^1, a^n)) \),

where \( a^i \) is the projection of \( a \) onto \( \Sigma^i \).

The state \( (q_1^1, \ldots, q_n^1) \) is good if all elements \( q_k^i \) are good; it is bad if at least one component \( q_k^i \) is bad. Both good and bad states are absorbing.
4.2. Product construction

We have defined the conjunctive path formula automaton. Following the approach in [3], the next step would be to construct the product of the CTMC and the automaton. This step turns out to be more involved. Thus we first need to construct the product from the CTMC and the automaton $A_{\psi^1 \land \psi^2}$ with $\psi^1 = f_1 U f_2$ and $\psi^2 = g_1 U g_2$ (see Fig. 2). The initial state is $(q_1^1, q_1^2)$, final states are marked with a double circle. The labeling function indicates which subsets of AP are acceptable. For example, we have $\delta((q_1^1, q_1^2), (f_2, g_1)) = (q_2^1, q_2^2)$, and $(q_3^1, q_3^2)$ is good. The node labeled with $\perp$ represents five bad absorbing states $(\perp, q_5^7), (\perp, q_5^2), (q_1^1, \perp), (q_3^2, \perp)$ and $(\perp, \perp)$. Transitions out of these two kinds of absorbing states are omitted.

Example 2. Consider the CTMC in Fig. 1, and the conjunctive path formula automaton $A_{\psi^1 \land \psi^2}$ with $\psi^1 = f_1 U f_2$ and $\psi^2 = g_1 U g_2$. We construct the product from the CTMC and the automaton $A_{\psi^1 \land \psi^2}$ in a straightforward way: Its reachable part is shown in Fig. 3. Notice that there is a transition from the state $(s_1, q_2^1, q_2^2)$ to $(s_2, \perp, q_2^2)$, since $\delta((q_1^2, q_1^2), (s_2)) = (\delta_1(q_2^1, 0), \delta_2(q_1^2, (s_1))) = (\perp, q_2^2)$. Let us explain why this plain product is not sufficient for our purpose. The valid path $\rho = s_0, t_0, s_1, t_1, s_2, t_2, s_4, \ldots, s_2, \perp, q_1^1$ is marked as a bad absorbing state.

The information missing in the product $C_{\psi^1 \land \psi^2}$ is whether one of the atomic path formulas $\psi^1$ (or $\psi^2$) is already satisfied, and the other still needs to be checked. This motivates the definition of the parameterized product CTMC, in which the parameter identifies such relevant information.

**Definition 4 (Parameterized product).** Let $C = (S, R, L, \alpha)$ be a CTMC and $A_{\psi}$ be the formula automaton for $\psi = \bigwedge_{i=1}^n \psi^i$. Let $\Lambda$ be a subset of all atomic path formulas $\psi^i$. Then the parameterized product CTMC $C^\Lambda_{\psi} = (\hat{S}, \hat{R}, \hat{L}, \hat{\alpha})$ is defined as follows:

- $\hat{S} = S \times \{\psi\}$.
- The rate $\hat{R}(s, q, (s', q'))$ equals $R(s, s')$ if $q = (q_1^1, \ldots, q_n^1)$ is not bad, $q' = (q_1^2, \ldots, q_n^2)$, and for each $i = 1, \ldots, n$,$$q_{ki}^i = \begin{cases} q_{ki}^1, & \text{if } s' = f_{ki}^i \text{ and } \psi^i \in \Lambda; \\ \delta^i(q_{ki}^1, L(s')) \cap \{f_{ki}^1, \ldots, f_{ki}^n\}, & \text{otherwise}. \end{cases}$$
  All other elements of $\hat{R}$ are zero.
- The labeling function $\hat{L}(s, q)$ is defined in two steps:
  1. If $q = (q_1^1, \ldots, q_n^1)$ is not bad, $L(s) = L(s) \cap \bigcup_{i=1}^n \{f_{ki}^1, \ldots, f_{ki}^n\}$; otherwise.
  2. For $i = 1, \ldots, n$, add the label $f_{ki}^i$ to those states $(s, q)$ with $q_{ki}^i = q_{ki}^1$.
- The initial distribution $\hat{\alpha} : S \times \{\psi\} \rightarrow [0, 1]$ is given by:
  - $\hat{\alpha}(s, q) = \alpha(s)$ if $i = 1, \ldots, n$.
  - $q_{ki}^i = \begin{cases} q_{ki}^1, & \text{if } s = f_{ki}^i \text{ and } \psi^i \in \Lambda; \\ \delta^i(q_{ki}^1, L(s) \cap \{f_{ki}^1, \ldots, f_{ki}^n\}), & \text{otherwise}, \end{cases}$
  where $q = (q_1^1, \ldots, q_n^1)$.
  - All other elements of $\hat{\alpha}(s, q)$ are zero.

In the product, each state is of the form $(s, (s_1^1, \ldots, s_1^n), \ldots, (s_n^1, \ldots, s_n^n))$, in which $q_{ki}^i$ from the path formula automaton of $\psi_i$ is relevant to the atomic propositions in $\psi^i$ that $s$ satisfies. The parameter $\Lambda$ is a set of path formulas $\psi^i$. If a transition in the product leads to a state $s'$ which satisfies $f_{ki}$ and $\Lambda$ contains $\psi^i$, the corresponding component of the state should be marked with the good absorbing state $q_{ki}^i$ of $A_{\psi^i}$, i.e., $(s', (s_1^i, \ldots, s_n^i), \ldots)$). In what follows, we will show that the set $\Lambda$ plays an important role in keeping track of the path formulas which have been satisfied during the probability computation.

Example 3. The product $C^\Lambda_{\psi^1 \land \psi^2}$ is shown in Fig. 1. Since both $\psi^1$ and $\psi^2$ can be satisfied during $[2, 3]$, we need to construct $C^\Lambda_{\psi^1 \land \psi^2}$. Its reachable part is shown in Fig. 4. Note that the previous state $(s_1, q_1^1, q_1^2)$ is now renamed to $(s_1, q_1^1, q_1^2)$ according to the initial distribution. It is easy to
see that the path $\rho = s_0, t_0, s_1, t_1, s_2, t_2, s_4, \ldots$, which was recognized as a bad path in Example 2, is captured by the new parameterized product CTMC during [2, 3] (assuming timing constraints are satisfied).

### 4.3. Probability computation

As in [3,9], the product CTMC stratifies the original CTMC in the sense that (time–abstract) bad paths will be uniformly directed towards the bad states. This allows us to reduce the computation by standard transient probability computation for CTMCs, which will be discussed in this section.

We fix a conjunctive path formula $\psi = \bigwedge_{i=1}^{n} \psi^i$ together with a CTMC $C = (S, R, L, \alpha)$. Now we focus on how to compute the probability of such a path formula starting from an arbitrary initial distribution $\alpha$ in a forward way. We first introduce some notation for convenience:

- For an interval $I$ and a positive number $h$, let $I \odot h$ denote the set $\{t - h \mid t \in I \wedge t \geq h\}$, and let $\psi \odot h$ denote the formula $f_1 U_{t_1 \odot h} f_2 U_{t_2 \odot h} \cdots f_k$.
- For $1 \leq i \leq j \leq K$, define $f_{i\ldots j} : = \bigvee_{k=i}^{j} f_k$.
- For $1 \leq i < K$, define $f^{(i)} : = f_i U_i f_{i+1} U_{i+1} \cdots f_k$.
- For a state formula $\Phi$, let $C[\Phi]$ be derived from $C$ by making its states that satisfy $\Phi$ absorbing.

**Definition 5 (Indicator matrix).** Given a state formula $\Phi$ and a subset $A$ of $\{\psi^i \mid i = 1, \ldots, n\}$, the indicator matrix $I^C_\Phi$ is defined by:

- If $(s, (q^1_1, \ldots, q^n_1)) \models \Phi$, then $I^C_\Phi((s, (q^1_1, \ldots, q^n_1)), (s, (q^1_2, \ldots, q^n_2))) = 1$, where $q^i_j$ is $q^i_j$, if $\psi^i \in A$; $q^i_j$, otherwise.
- All other entries of $I^C_\Phi$ are zero.

Now we show how to compute the probability $Pr^C_\alpha(\bigwedge_{i=1}^{n} \psi^i)$ in a forward way.

**Theorem 1 (Probability computation).** Given a conjunctive path formula $\psi = \bigwedge_{i=1}^{n} \psi^i$ and a CTMC $C = (S, R, L, \alpha)$ equipped with the following notations:

- Let $h$ be the least nonzero endpoint of all intervals $I^k$ occurring in $\psi$, and let $A$ be the set $\{\psi^i \mid h > a^i_{K-1}\}$.
- If $h < \max_{i=1}^{n} (b^i_{K-1})$, let $h'$ be the least endpoint of all intervals $I^k$ greater than $h$, and let $A'$ be the set $\{\psi^i \mid h' > a^i_{K-1}\}$.
- For $i = 1, \ldots, n$, let $a^i_0 = 0, a^i_0 = a^i_1, a^i_{K-1} = b^i_{K-1}, b^i_1 = \infty$ and $f^i_{h+1} = f^i_1$, and let $k^i_1, k^i_{i-1}$ be the unique indices such that $b^i_{k^i-1} \leq h < b^i_{k^i}$ and $a^i_{k^i-1} \leq h \leq a^i_{k^i}$.

The probability $Pr^C_\alpha(\psi) = Pr^C_\alpha(\psi)$ is computed as follows:

1. If $h < \max_{i=1}^{n} (b^i_{K-1})$, then
   \[
   Pr^C_\alpha(\psi) = \pi_{\bigwedge_{i=1}^{n} f^i_{h+1}}(\widehat{\alpha}, h) \cdot \Pi_{A'} \left( \bigwedge_{i=1}^{n} f^i_{h+1} \right),
   \]
   where $Pr^C_\alpha(\psi)$ stands for the vector $(Pr^C_\alpha(\psi))_{i \in S}$.
2. Otherwise $h = \max_{i=1}^{n} (b^i_{K-1})$, then
   \[
   Pr^C_\alpha(\psi) = \pi_{\psi^1} \widehat{C}_1(h) \cdot \Pi_{A'} \left( \bigwedge_{i=1}^{n} f^i_{h+1} \right) \begin{pmatrix} 1, \ldots, 1 \end{pmatrix}^T.
   \]

The proof is given in Appendix A for completeness, which follows the same idea as the proof in [9]. Intuitively, the computation is performed by traversing through the time intervals in a forward way in the product CTMC. The time is partitioned into finitely many intervals using endpoints appearing in the formula. With the initial distribution $\widehat{\alpha}$, we compute the probability distribution of all states at the time point $h$ for the first interval $[0, h]$, which determines the parameter $A$. The indicator matrix filters out all paths dissatisfying $\psi$ at the time point $h$. The parameter $A'$ is determined by the next interval $[h, h')$. We recursively compute the probability distribution at the time point $h'$, and repeat this until the last time point.

The parameter $A$ will be repeatedly adjusted when we push the time forward. We illustrate the theorem by computing $Pr^C_\alpha(\psi_1 \wedge \psi_2)$ in the following example.

**Example 4.** Let $\psi_1 = f_1 U_{(0,2)} f_2 U_{(2,3)} f_3$ and $\psi_2 = g_1 U_{(1,3)} g_2$. We compute $Pr^C_\alpha(\psi_1 \wedge \psi_2)$ in three phases.

Initially, $a^1_2 = 2$ and $b^1_2 = a^2_2 = 2$ are the first two nonzero endpoints, so $h = 1$ and $h' = 2$. Accordingly, $A = \emptyset$ and $A' = \{\psi_2\}$. By Theorem 1, $k_1 = 1$, $l_1 = 2$, $k_2 = 1$ and $l_2 = 1$, which means we should pick out the states satisfying $f_1 \lor f_2$ and $g_1$ at time point $h$. Then, we compute $\pi_{\psi_1} \widehat{C}_1(\psi_1 \wedge \psi_2 \rightarrow (f_1 \lor f_2) \wedge g_1) (\alpha, h)$ — the probability distribution at time point $h = 1$ w.r.t. the product CTMC equipped with parameter $\emptyset$ showed in Fig. 3. The indicator matrix has only two nonzero entries $I^2_{\psi_1 \wedge \psi_2}(1) = 1$ and $I^2_{\psi_1 \wedge \psi_2}(2) = 1$. After this step, we push forward the time, so $\psi_1$ and $\psi_2$ become $f_1 U_{(0,1)} f_2 U_{(1,2)} f_3$ and $g_1 U_{(0,1)} g_2$, respectively.

In the next phase, $h = 1$, $h' = 2$, accordingly, $A = \{\psi_2\}$ and $A' = \{\psi_1, \psi_2\}$. Then, we compute the probability distribution at $h = 1$ w.r.t. the same CTMC in Fig. 3. At the next phase, $\psi_1$ and $\psi_2$ may be fully satisfied, so we should relocate the state $(s_1, q^1_2, q^2_1)$ to $(s_1, q^1_2, q^2_1)$, so the indicator matrix has only one nonzero entry
Assume that the sets $\text{Sat}$ proved that it is decidable whether such an expression is the bound we compute the probability distribution at $f$ by fresh (pairwise disjoint) atomic propositions is equivalent to checking whether $\Pr_\Psi\psi$ calculated recursively. We replace5 if any state subformula is undefined, our algorithm will then report undefined.

In the last phase, $h = \max\{b^1_{K_1-1}, b^2_{K_2-1}\} = 1$. Let $A = (\psi^1, \psi^2)$, and we get to the end of the computation. Then, we compute the probability distribution at $h = 1$ w.r.t. the CTMC in Fig. 4. Now the indicator matrix has only two nonzero entries $I_{f_3\wedge g_2}^{(3)}((s_1, q_1^1, q_2^1), (s_3, q_3^1, q_3^1)) = 1$ and $I_{f_3\wedge g_2}^{(3)}((s_4, q_4^1, q_2^1), (s_4, q_3^1, q_2^2)) = 1$.

The total computational results are

<table>
<thead>
<tr>
<th>$t$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(s_0, q_1^1, q_1^1)$</td>
<td>$e^{-2}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$(s_1, q_1^1, q_2^1)$</td>
<td>0</td>
<td>$2e^{-2}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$(s_1, q_1^1, q_1^1)$</td>
<td>0</td>
<td>0</td>
<td>$4e^{-4}$</td>
<td>0</td>
</tr>
<tr>
<td>$(s_3, q_2^3, q_2^2)$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\frac{5}{2}e^{-4} - 2e^{-5} - \frac{7}{2}e^{-7}$</td>
</tr>
<tr>
<td>$(s_4, q_1^1, q_2^2)$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\frac{5}{2}e^{-4} - 2e^{-5} + \frac{7}{2}e^{-7}$</td>
</tr>
</tbody>
</table>

Therefore, we have $\Pr_{s_0}(\psi_1 \land \psi_2) = 4e^{-4} - 4e^{-5}$ by collecting all desired probabilities.

**Corollary 1.** Given a path formula $\varphi = \bigwedge_{i=1}^m \Lambda_j \psi^j$ in DNF and a CTMC $C$, the probability $\Pr^C_{s_0}(\varphi)$ can be computed by inclusion–exclusion principle.

4.4. Model checking algorithm and its complexity

Let $C = (S, R, L, \alpha)$ be a CTMC, $s \in S$, and $\Phi$ be a CCSL state formula. The model checking problem is to check whether $s \models \Phi$. The standard algorithm to solve CTL-like model checking problems recursively computes the sets of states satisfying $\Phi$, denoted by $\text{Sat}(\Phi)$, for all state subformulas $\Psi$ of $\Phi$. For CCSL, the cases where $\Phi$ is an atomic proposition, a negation or a conjunction are standard as for CTL. The case when $\Phi$ is a (conditional) probabilistic formula is the challenging part. The model checking algorithm for $\Pr_{s_0}(\varphi)$ has been discussed in [19], thus below we discuss the case of $\Pr_{s_0}(\psi)$.

Let $\Psi = \Pr_{s_0}(\varphi | \psi)$ with $\varphi = \Phi_1 U_1 j_1 \Phi_2 U_2 j_2 \ldots \Phi_k$ and $\psi = \Psi_1 U_1 j_1 \Psi_2 U_2 j_2 \ldots \Psi_l$. By definition, checking $\Psi$ is equivalent to checking whether $Pr_{s_0}(\varphi | \psi) = 0$, i.e., whether the quotient of $Pr_{s_0}(\psi \land \varphi)$ and $Pr_{s_0}(\psi)$ meets the bound $\neq 0$. Now we focus on the conjunction part. Assume that the sets $\text{Sat}(\Phi_1)$ and $\text{Sat}(\Psi_1)$ have been calculated recursively. We replace $\Phi_1, \ldots, \Phi_k$ and $\Psi_1, \ldots, \Psi_l$ by fresh (pairwise disjoint) atomic propositions $f_1, \ldots, f_k$ and $g_1, \ldots, g_l$, and add the label $f_1$ (resp. $g_j$) to the state $s$ if $s \in \text{Sat}(\Phi_1)$ (resp. $s \in \text{Sat}(\Psi_1)$). Thus, after applying Theorem 1 and Corollary 1 a finite number of times, $Pr_{s_0}(\varphi | \psi)$ is reduced to a product of transient probabilities. We can now apply the results in [1] as follows: By definition, $Pr_{s_0}(\varphi | \psi)$ can be expressed as a quotient of finite sum of the form $\sum_{i} \eta_i e^{\alpha_i}$ (with algebraic $\eta_i$ and $\alpha_i$). Aziz et al. proved that it is decidable whether such an expression is $\gg p$, for $p \in Q$, which implies directly the decidability of the model checking problem of CCSL.

Finally we discuss the complexity of the approach for approximating $\Pr_{s_0}(\varphi)$. The size of the product CTMC is exactly the product of the sizes of the original CTMC and the automaton obtained from the conjunctive path formula, i.e. $|C| \cdot \prod_i |\Psi^i|$. Thus, the usual numerical algorithm can be used to approximate the transient distributions, for instance via uniformization [11], or Runge–Kutta method, which is linear in the size of the product, the largest exit rate and the largest finite time bounds. Hence, the complexity for computing the probability of the conjunctive path formula is linear in the product of the sizes of the product CTMC and the conjunctive path formula, i.e. $O(I(\sum_j |\Psi^j|) \cdot |C| \cdot \prod_j |\Psi^j|)$. Furthermore, the complexity for computing the probability of the path formula in DNF as in Corollary 1 is bounded by $O(I(\sum_j |\Psi^j|) \cdot |C| \cdot \prod_j |\Psi^j|)$. It is also the worst case complexity of our model checking algorithm.

### Appendix A. Proof of Theorem 1

The equation $\Pr^C_{s_0}(\varphi) = \Pr^C_{s_0}(\psi)$ can be proven by establishing mapping of the cylinder sets, similar as the proof in [9] for atomic path formula. We provide the proof of Eq. (1), by extending the proof in [9] for the parametrized product CTMC. For $s' \in S$, define the event $Z(s') := [\sigma | \sigma@h = s' \land \forall \in [0, h]$). $\sigma@t] = [1_{t_i = 1_{f_{i}, t_i - h}}$, where $\sigma@h$ stands for the state of the path $\sigma$ at time $h$. The following inclusion holds:

$$\{\sigma | \sigma \models \varphi\} \subseteq \bigcup_{s' \models \Lambda_{n=1} f_{i} \land t_{i} - h} Z(s').$$

Note this property holds for the product CTMC, but not for general CTMCs. Intuitively, the deterministic automaton stratifies the original CTMC in a way, such that those paths $\sigma$ with $\sigma \nvdash \varphi$ will be directed to the bad absorbing state $\bot$. This is the crucial property allowing us to perform a forward transient analysis. The formal argument is done using the notion of stratification, and we refer to [3, 9] for details.

Now we fix first $\delta t$, as an initial distribution with $\delta t = 1$ and $s = \Lambda_{n=1} f_{i} \land t_{i} - 1$. By the law of total probability, we have

$$\Pr^C_{s_0}(\varphi) = \sum_{s' \models \Lambda_{n=1} f_{i} \land t_{i} - 1} \Pr^C_{s_0}(Z(s')) \cdot \Pr^C_{s_0}(\varphi | Z(s')).$$

By definition of $Z(s')$, we have

$$\Pr^C_{s_0}(Z(s')) = \pi_{C_{\varphi}^{\Lambda_{n=1} f_{i} \land t_{i} - 1}}(s'(h)).$$

Now let $\sigma \in Z(s')$ be a path. $\sigma \models \varphi$ implies that at time $h$, $\sigma$ has reached a state in a phase from $q_1, q_2, \ldots, q_n$. So the suffix path of $\sigma$ starting at time $h$ satisfies $\Lambda_{n=1} f_{i} \land t_{i} \sqcap \bot$. From the point time $h$, the labels $f_{i} \land t_{i} - 1$, $f_{i} \land t_{i} - 1$, $\ldots$, $f_{i} \land t_{i} - 1$ have been irrelevant for checking the formula $\psi$. Thus we could reconstruct the product CTMC with parameter $\Lambda'$ for the next phase and put forward the formula. By the Markov property of CTMCs, we have

5 If any state subformula is undefined, our algorithm will then report undefined.
\[ \Pr^{\sigma}_{Z}(s, h)(s') = \sum_{s' \in S} \pi^{C_{\psi}}_{Z} \left( (\bigwedge_{i=1}^{n} f_{i_{1}:i_{n}}) \right)(s, h)(s') \cdot \Pr^{\sigma}_{Z} \left( \bigwedge_{i=1}^{n} f_{i_{1}:i_{n}} \right) \cdot \Pr^{\sigma}_{Z} \left( \bigcup_{i=1}^{n} f_{i_{1}:i_{n}} \right) \cdot 1. \]

Eq. (2) for general initial distribution \( \sigma \) follows as the step 2 of this proof.

- Otherwise \( h = \max_{i} \{ b_{i_{1}:i_{n}} \} = \infty \), then \( \Pr^{\sigma}_{Z}(s, h)(s') \) is the probability to reach the states labeled with \( f_{K_{1}}^{1}, f_{K_{2}}^{2}, \ldots, f_{K_{n}}^{n} \) eventually. So we just need to pass the \( h \) to \( \infty \) to obtain the probability. □

References