RWTH Aachen

Department of Computer Science
Technical Report

Time-Bounded Reachability in Continuous-Time Markov Decision Processes

Martin R. Neuhäußer and Lijun Zhang

ISSN 0935–3232 · Aachener Informatik Berichte · AIB-2009-12

RWTH Aachen · Department of Computer Science · May 2009

Time-Bounded Reachability in Continuous-Time Markov Decision Processes ⋆

Martin R. Neuhäußer[1,3] and Lijun Zhang[2]

[1] Software Modeling and Verification Group, RWTH Aachen University, Germany
[2] Dependable Systems and Software Group, Universität des Saarlands, Germany
[3] Formal Methods and Tools Group, University of Twente, The Netherlands

Abstract. This paper solves the problem of computing the maximum and minimum probability to reach a set of goal states within a given time bound for locally uniform continuous-time Markov decision processes (CTMDPs). As this model allows for nondeterministic choices between exponentially delayed transitions, we define total time positional (TTP) schedulers which rely on the CTMDP's current state and the total elapsed time when taking a decision. In this paper, TTP schedulers are proved to be sufficient to maximize timed reachability even w.r.t. fully time- and history-dependent schedulers; further, they allow us to derive a fixed point characterization of the optimal timed-reachability probability. The main contribution of this paper is a discretization technique which, for an a priori given error bound ε, induces a discrete-time MDP that approximates the optimal timed reachability probability in the underlying CTMDP up to ε.

1 Introduction

Continuous-time Markov decision processes (CTMDPs) [3,14] are a stochastic model which allows for nondeterministic choices between transitions whose delay is governed by negative exponential distributions. Therefore, CTMDPs can be seen as an extension of continuous-time Markov chains (CTMCs) [10,11] by nondeterministic choices as well as an extension of Markov decision processes (MDPs) [14] by exponentially distributed delays.

To obtain a unique stochastic process, we follow the MDP approach [14] and use schedulers to resolve the CTMDP's nondeterministic choices. Intuitively, depending on the trajectory that led into the current state, a scheduler chooses the next action to be performed in a randomised way. Accordingly, the stochastic behaviour of a CTMDP is described by the upper and lower probability bounds which are induced by a given (usually uncountable) class of schedulers.

For general CTMDPs, the sojourn time distribution of the current state depends on the action that is chosen by the scheduler; this dependency requires the scheduler to decide early, that is, when entering the current state. However, in this paper we restrict to the subclass of locally uniform CTMDPs which share the property that the state residence time distribution does not depend on the scheduler's choice. As shown in [13], local uniformity allows to delay the scheduling decision until the current state is left; the resulting late schedulers, which are well-defined only for locally uniform CTMDPs, perform at least as good as any early scheduler and generally induce strictly better probability bounds [13].

We solve the timed reachability problem by computing the maximum and minimum probability to reach a set of goal states G within a given time bound z

⋆ Supported by the NWO project QUPES (612.000.420) and by the EU grant FP7-ICT-2007-1 (QUASIMODO).

under all late schedulers. More precisely, we prove that for timed reachability, it suffices to consider *total time positional deterministic schedulers* (*TTPD*) which base their decision only on the total amount of time that has passed and on the current state. Exploiting this result, we characterize the maximum timed reachability probability as the least fixed point of a higher-order operator which involves integration over the time domain. With this characterization at hand, we reduce the timed reachability problem for locally uniform CTMDPs to the problem of computing step-bounded reachability probabilities in discrete-time MDPs. More precisely, we approximate the behaviour of the CTMDP up to an a priori specified accuracy ε by defining its discretized MDP such that its step-bounded reachability probabilities which can be computed by standard methods like value iteration [4], coincide (up to ε) with the timed-reachability probabilities of the underlying CTMDP. The complexity of our approach lies in $\mathcal{O}(m\cdot(\lambda z)^2/\varepsilon)$ where m denotes the size of the input model and λ is its maximal exit rate.

Besides the theoretical contribution, we believe that the results of this paper have many practical applications: In fact, CTMDPs are the semantic basis for many stochastic models like stochastic activity networks [15], generalized stochastic Petri nets [6,12] and Markovian process algebras [9]. So far, their analysis is restricted to those instances (called "well-specified" or "non-confused") which do not exhibit non-deterministic choices, that is, where the underlying stochastic process is a CTMC. However, if nondeterminism is present in those models, their semantics becomes more intricate and, in fact, is a locally uniform CTMDP. So far, model checking of CTMDPs has received scant attention. Most of the existing analysis methods focus on optimizing criteria such as the expected total reward [8] or the expected long-run average reward [7,8]. In [3], the authors provide an algorithm to compute timed-reachability probabilities in globally uniform CTMDPs. However, its applicability is severely restricted, as global uniformity — which requires the sojourn time in any state of the model to be identically distributed — is hard to achieve. Further, the approach is based on time-abstract schedulers, which are strictly less powerful than time-dependent ones [3]. Therefore, our paper extends the results of [3] in two ways: We consider fully time- and history-dependent schedulers and lift the restriction to globally uniform CTMDPs by allowing different states to have different sojourn time distributions.

Although the results of this paper refer to the maximum timed reachability probability, all proofs can be adapted to the dual problem, namely determining the minimum timed reachability probability.

2 Continuous-time Markov decision processes

We consider CTMDPs with a finite set of states $\mathcal{S} = \{s_0, s_1, \dots\}$ and a finite set of actions $Act = \{\alpha, \beta, \dots\}$; accordingly, we use $Distr(\mathcal{S})$ and $Distr(Act)$ to denote the sets of probability distributions over $\mathcal{S}$ and Act, respectively.

Definition 1 (Continuous-time Markov decision process). *A* continuous-time Markov decision process *(CTMDP) is a tuple* $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ *where* $\mathcal{S}$ *and Act are finite, nonempty sets of* states *and* actions, $\mathbf{R} : \mathcal{S} \times Act \times \mathcal{S} \to \mathbb{R}_{\geq 0}$ *is a three-dimensional* rate matrix *and* $\nu \in Distr(\mathcal{S})$ *is an* initial distribution.

Intuitively, a CTMDP behaves as follows: If $\mathbf{R}(s, \alpha, s') = \lambda$ and $\lambda > 0$, an α-transition leads from state s to state s'. Here, λ is the *rate* of an exponential distribution which governs the transition's delay. If action α is chosen, the transition executes in time interval $[a, b]$ with probability $\eta_\lambda\big([a, b]\big) = \int_a^b \lambda e^{-\lambda t}\, dt$. If multiple α-transitions emanate a state s, a *race condition* occurs: The α-transition whose delay expires first executes and thereby determines the sojourn time of state s under action α. As the minimum of independent exponentials is again exponentially distributed with the sum of their rates, the sojourn-time of state s under action α is distributed with *exit rate* $E(s, \alpha) = \sum_{s' \in \mathcal{S}} \mathbf{R}(s, \alpha, s')$.

An action α is *enabled* in state s iff $E(s, \alpha) > 0$; accordingly, the set $Act(s) = \{\alpha \in Act \mid E(s, \alpha) > 0\}$ is the set of *enabled* actions in state s and determines the available nondeterministic choices in state s. A CTMDP is *well-formed* iff $Act(s) \neq \emptyset$ for all $s \in \mathcal{S}$. As this is easily achieved by adding self-loops, we restrict to well-formed CTMDPs. A well-formed CTMDP $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ induces the *embedded* discrete time MDP $\mathcal{M} = (\mathcal{S}, Act, \mathbf{P}, \nu)$ where the (time abstract) transition probability matrix $\mathbf{P}$ satisfies $\mathbf{P}(s, \alpha, s') = \frac{\mathbf{R}(s,\alpha,s')}{E(s,\alpha)}$ if $E(s, \alpha) > 0$ and 0, otherwise. In fact, $\mathbf{P}(s, \alpha, s')$ is the *time-abstract* probability to move from state s to state s' if action α is selected in state s. For the initial distribution ν, we use $\nu_s = \{s \mapsto 1\}$ to denote a single initial state.

Example 1. Consider the CTMDP in Fig. 1. If α is chosen in state s_0, the two emanating α-transitions to s_2 and s_3, resp., compete for execution. The sojourn time in s_0 is exponentially distributed with rate $E(s_0, \alpha) = 3$. Further, the probability to reach state s_2 is given by $\mathbf{P}(s_0, \alpha, s_2) = \frac{\mathbf{R}(s_0,\alpha,s_2)}{E(s_0,\alpha)} = \frac{1}{3}$.

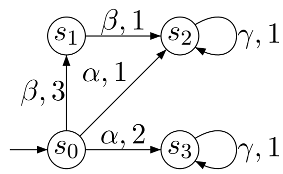


Fig. 1. Example CTMDP

A CTMDP only defines a stochastic process if the nondeterministic choices between actions are resolved by a scheduler. As shown in [13], locally uniform CTMDPs allow to define schedulers that cannot be defined for general CTMDPs and which perform strictly better than general schedulers. As local uniformity is commonly found in queuing systems, GSPNs and SANs, we restrict our analysis to the subclass of locally uniform CTMDPs:

Definition 2 (Local uniformity). *Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a well formed CTMDP. $\mathcal{C}$ is* locally uniform *iff $\forall s \in \mathcal{S}.\ \forall \alpha, \beta \in Act(s).\ E(s, \alpha) = E(s, \beta)$.*

Local uniformity ensures that the sojourn time in any state does not depend on the action chosen in that state. Hence, we may use $\lambda(s) = E(s, \alpha)$ for some $\alpha \in Act(s)$ to denote the exit rate of state s.

2.1 The probability space

Finite *paths* in a CTMDP $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ are sequences $\pi = s_0 \xrightarrow{t_0,\alpha_0} s_1 \xrightarrow{t_1,\alpha_1} \cdots \xrightarrow{t_{n-1},\alpha_{n-1}} s_n$ where $s_i \in \mathcal{S}$, $\alpha_i \in Act$ and $t_i \in \mathbb{R}_{\geq 0}$ for $i \leq n$; n is the length of π, denoted $|\pi|$. We use $\pi[k] = s_k$ and $\delta(\pi, k) = t_k$ to refer to the k-th state on π and its associated sojourn time. Further, let $\Delta(\pi) = \sum_{k=0}^{n-1} t_k$ be the total time spent on π and $\pi{\downarrow} = s_n$ be the last state of π.

Any path π is a concatenation of a state with a sequence of *combined transitions* from the set $\Omega = \mathbb{R}_{\geq 0} \times Act \times \mathcal{S}$; hence, $\pi = s_0 \circ m_0 \circ m_1 \circ \cdots \circ m_{n-1}$

with $m_i = (t_i, \alpha_i, s_{i+1}) \in \Omega$. Thus $Paths^n(\mathcal{C}) = \mathcal{S} \times \Omega^n$ is the set of paths of length n in $\mathcal{C}$; further, let $Paths^\star(\mathcal{C})$, $Paths^\omega(\mathcal{C})$ and $Paths(\mathcal{C})$ denote the sets of finite, infinite and all paths of $\mathcal{C}$. To simplify notation, we omit the reference to $\mathcal{C}$ wherever possible. The following measure-theoretic concepts are mentioned only briefly; we refer to [13,16] for an in-depth discussion. Events in $\mathcal{C}$ are measurable sets of paths; as paths are Cartesian products of combined transitions, we first define the σ-field $\mathfrak{F} = \sigma\big(\mathfrak{B} \times \mathfrak{F}_{Act} \times \mathfrak{F}_{\mathcal{S}}\big)$ on subsets of Ω where $\mathfrak{B}$ denotes the Borel σ-field over $\mathbb{R}_{\geq 0}$ and $\mathfrak{F}_{\mathcal{S}} = 2^{\mathcal{S}}$ and $\mathfrak{F}_{Act} = 2^{Act}$. Based on $(\Omega, \mathfrak{F})$, we derive the product σ-field $\mathfrak{F}_{Paths^n} = \sigma\big(\{S_0 \times M_0 \times \cdots \times M_{n-1} \mid S_0 \in \mathfrak{F}_{\mathcal{S}}, M_i \in \mathfrak{F}\}\big)$ of measurable subsets of $Paths^n$. Finally, the cylinder-set construction [1] extends this uniquely to a σ-field over infinite paths: A set $B \in \mathfrak{F}_{Paths^n}$ is called a *base* of an infinite *cylinder* C if $C = Cyl(B) = \{\pi \in Paths^\omega \mid \pi[0..n] \in B\}$. Now the desired σ-field $\mathfrak{F}_{Paths^\omega}$ is generated by the set of all cylinders, i.e. $\mathfrak{F}_{Paths^\omega} = \sigma\big(\bigcup_{n=0}^{\infty} \{Cyl(B) \mid B \in \mathfrak{F}_{Paths^n}\}\big)$.

2.2 Schedulers

To reason about probabilities in CTMDPs, we need to introduce schedulers that resolve the nondeterminism that occurs in states with multiple enabled actions: If $Act(s) = \{\alpha_1, \ldots, \alpha_n\}$, a scheduler yields a distribution over $\alpha_1, \ldots, \alpha_n$ and thereby quantifies the nondeterminism. In the classical setting, the scheduler immediately decides for an action when entering a state [3,14,16]. However, in locally uniform CTMDPs the state's sojourn time distribution is independent of the choice of a particular action. Thus, we can define a strictly better class of schedulers which delay their decision up to the point when the current state is left and which incorporate the current state's sojourn time into their decision:

Definition 3 (Generic measurable scheduler). *A* generic scheduler *for a locally uniform CTMDP $(\mathcal{S}, Act, \mathbf{R}, \nu)$ is a mapping $D : Paths^\star \times \mathbb{R}_{\geq 0} \times \mathfrak{F}_{Act} \to [0,1]$ where $D(\pi, t, \cdot) \in Distr(Act(\pi{\downarrow}))$ for all $t \in \mathbb{R}_{\geq 0}$ and $\pi \in Paths^\star$. D is* measurable *iff the functions $D(\cdot, \cdot, A) : Paths^\star \times \mathbb{R}_{\geq 0} \to [0,1]$ are measurable for all $A \in \mathfrak{F}_{Act}$.*

Formally, the measurability condition states that for all $A \in \mathfrak{F}_{Act}$ and $B \in \mathfrak{B}\left([0,1]\right)$, i.e. measurable subsets of $[0,1]$, it holds that $\{(\pi, t) \mid D(\pi, t, A) \in B\} \in \sigma\left(\mathfrak{F}_{Paths^\star} \times \mathfrak{B}\right)$; intuitively, it excludes schedulers which resolve the nondeterminism in a way that induces non-measurable sets (like Vitali sets, see [1, p.34]), i.e., sets of paths that cannot be assigned a probability. Let π be a finite path ending in state s with $|Act(s)| > 1$. If s is left after t units of time, then $D(\pi, t, \cdot)$ defines a probability distribution over $Act(s)$ which resolves the nondeterminism in state s for *history* π and sojourn time t. We call a gm-scheduler D *deterministic* iff for all $\pi \in Paths^\omega$ and $t \in \mathbb{R}_{\geq 0}$ the distribution $D(\pi, t, \cdot)$ is degenerate. We use GM (and GMD) to denote the class of generic measurable (deterministic) schedulers.

Definition 4 (Total time positional scheduler). *Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a locally uniform CTMDP. A mapping $D : \mathcal{S} \times \mathbb{R}_{\geq 0} \to Distr(Act)$ is a* total time positional scheduler *iff $\forall s \in \mathcal{S}.\ \forall t \in \mathbb{R}_{\geq 0}.\ D(s,t)(\alpha) > 0 \Longrightarrow \alpha \in Act(s)$.*

As $D(s,t)$ in general is a distribution over actions, we use $TTPR$ to denote the set of all *randomized* total time positional schedulers; accordingly, $TTPD \subset TTPR$

denotes the subset of all *deterministic* *TTPR* schedulers where $D(s,t)(\alpha) = 1$ for some $\alpha \in Act(s)$. Note that any $D \in TTPR$ naturally induces the generic scheduler D' where $D'(\pi, t, A) = \sum_{\alpha \in A} D(\pi{\downarrow}, \Delta(\pi){+}t)(\alpha)$ for all $\pi \in Paths^\star$, $t \in \mathbb{R}_{\geq 0}$. Accordingly, D is measurable iff D' is a gm-scheduler. Intuitively, *TTPR*-scheduler base their decision only on the current state and the total amount of time that has passed; later, Thm. 1 proves that *TTPD* schedulers suffice to optimize timed-reachability objectives.

2.3 Probability measures

A gm-scheduler D uniquely determines the stochastic process of a CTMDP. We define its induced probability measure on the measurable space $(Paths^\omega, \mathfrak{F}_{Paths^\omega})$. Following the lines of [16,13], we start by deriving the probability of measurable sets of combined transitions:

Definition 5 (Probability of combined transitions). *Let $\mathcal{C}{=}(\mathcal{S}, Act, \mathbf{R}, \nu)$ be a locally uniform CTMDP and D a gm-scheduler on $\mathcal{C}$. For all $\pi \in Paths^\star$, define the probability measure $\mu_D(\pi, \cdot) : \mathfrak{F} \to [0,1]$ where*

$$\mu_D(\pi, M) = \int_{\mathbb{R}_{\geq 0}} \eta_{\lambda(\pi\downarrow)}(dt) \int_{Act} D(\pi, t, d\alpha) \int_{\mathcal{S}} \mathbf{I}_M(t, \alpha, s') \;\; \mathbf{P}(s, \alpha, ds'). \qquad (1)$$

Here, $\eta_{\lambda(\pi\downarrow)}$ is the exponential distribution that refers to the sojourn time of the state $\pi{\downarrow}$ which is distributed with rate $\lambda(\pi{\downarrow})$; further, $\mathbf{I}_M$ is the characteristic function of $M \in \mathfrak{F}$. In fact, $\mu_D(\pi, M)$ is the probability to continue with some combined transition in M, given that we hit the current state $\pi{\downarrow}$ along the trajectory π. Having the probability measures $\mu_D(\pi, \cdot)$ at hand, we now can define the probabilities of measurable sets of paths from the sets $\mathfrak{F}_{Paths^n}$:

Definition 6 (Probability measure). *Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a locally uniform CTMDP and D a gm-scheduler on $\mathcal{C}$. For $n \geq 0$, we inductively define the probability measures $Pr^n_{\nu,D}$ on the measurable space $(Paths^n, \mathfrak{F}_{Paths^n})$:*

$$Pr^0_{\nu,D} : \mathfrak{F}_{Paths^0} \to [0,1] \;\; : \Pi \mapsto \sum_{s \in \Pi} \nu(\{s\}) \quad \textit{and for } n > 0$$

$$Pr^n_{\nu,D} : \mathfrak{F}_{Paths^n} \to [0,1] \;\; : \Pi \mapsto \int_{Paths^{n-1}} Pr^{n-1}_{\nu,D}(d\pi) \int_{\Omega} \mathbf{I}_\Pi(\pi \circ m) \; \mu_D(\pi, dm).$$

Intuitively, $Pr^n_{\nu,D}$ measures a set of paths Π of length n by multiplying the probabilities $Pr^{n-1}_{\nu,D}(d\pi)$ of path prefixes π (of length $n{-}1$) with the probability $\mu_D(\pi, dm)$ of a combined transition $m \in M$ which extends π to a path in Π. Together, the measures $Pr^n_{\nu,D}$ extend to a unique measure on $\mathfrak{F}_{Paths^\omega}$: if $B \in \mathfrak{F}_{Paths^n}$ is a measurable base and $C = Cyl(B)$, we define $Pr^\omega_{\nu,D}(C) = Pr^n_{\nu,D}(B)$. Due to the inductive definition of $Pr^n_{\nu,D}$, the Ionescu–Tulcea extension theorem [1] is applicable and yields a unique extension of $Pr^\omega_{\nu,D}$ from cylinders to sets in $\mathfrak{F}_{Paths^\omega}$.

3 A fixed point characterization for timed reachability

Reasoning about a CTMDP's behaviour is relative to an a priori fixed class of schedulers. In this paper, we aim at computing the upper (and lower) bounds on the probability to reach a set of goal states G within a given time bound z (denoted $\Diamond^{[0,z]}G$) w.r.t. the class of GM-schedulers. As proved in [13], the class of GM-schedulers is the most general one:

Definition 7 (Maximum timed reachability). *Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a locally uniform CTMDP, $G \subseteq \mathcal{S}$, $s \in \mathcal{S}$ and $z \in \mathbb{R}_{\geq 0}$. Then*

$$p^G_{max} : \mathcal{S} \times \mathbb{R}_{\geq 0} \to [0,1] : (s,z) \mapsto \sup_{D \in GM} Pr^{\omega}_{\nu_s,D}\Big(\Diamond^{[0,z]}G\Big)$$

is the maximum timed reachability *for the set of* goal states G *and time bound* z*.*

A scheduler $D \in GM$ is *optimal* for the set of goal states G and time bound z iff $Pr^{\omega}_{\nu_s,D}\left(\Diamond^{[0,z]}G\right) = p^G_{max}(s,z)$ for all $s \in \mathcal{S}$. Further, for $\varepsilon > 0$, $D \in GM$ is ε-*optimal* for G and z iff $\left|Pr^{\omega}_{\nu_s,D}\left(\Diamond^{[0,z]}G\right) - p^G_{max}(s,z)\right| \leq \varepsilon$ for all $s \in \mathcal{S}$.

We now construct an optimal $TTPD$ scheduler which computes p^G_{max}; this is possible, as for timed reachability, we only need to consider $TTPD$ schedulers:

Definition 8 (Optimal $TTPD$ scheduler). *Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a locally uniform CTMDP, $G \subseteq \mathcal{S}$ a set of goal states and $z \in \mathbb{R}_{\geq 0}$ a time bound. We call a TTPD scheduler D^z* optimal *iff it satisfies for all $s \in \mathcal{S}$, $t \leq z$ and $\alpha \in Act$ that*

$$D^z(s,t) = \alpha \quad \Longrightarrow \quad \forall \beta \in Act(s).\ f(s, z-t, \beta) \leq f(s, z-t, \alpha),$$

where

$$f(s, z', \gamma) := \sum_{s' \in \mathcal{S}} \mathbf{P}(s, \gamma, s') \sup_{D' \in GM} Pr^{\omega}_{\nu_{s'},D'}\left(\Diamond^{[0,z']}G\right).$$

We use $f(s, z-t, \beta)$ to express the maximum probability to reach G via action β when the sojourn time t in state s has already expired and $z-t$ time units remain.

Theorem 1 (Optimality). *Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a locally uniform CTMDP, $G \subseteq \mathcal{S}$ a set of goal states and $z \in \mathbb{R}_{\geq 0}$ a time bound. Then*

$$p^G_{max}(s,z) = Pr^{\omega}_{\nu_s,D^z}\left(\Diamond^{[0,z]}G\right).$$

Proof. We split the event $\Diamond^{[0,z]}G$ into disjoint subsets as follows: Let $\Pi(z,n) = \left\{\pi \in Paths^{\omega} \mid \pi[n] \in G \wedge \pi[i] \notin G \text{ for } i < n \wedge \sum_{i=0}^{n-1} \delta(\pi,i) \leq z\right\}$. Then $\Pi(z,n) \cap \Pi(z,m) = \emptyset$ for $n \neq m$ and $\Diamond^{[0,z]}G = \biguplus_{n=0}^{\infty} \Pi(z,n)$. Hence, for all $\nu \in Distr(\mathcal{S})$ and $D \in GM$ it holds that

$$Pr^{\omega}_{\nu,D}\big(\Diamond^{[0,z]}G\big) = Pr^{\omega}_{\nu,D}\Big(\biguplus_{n=0}^{\infty} \Pi(z,n)\Big) = \sum_{n=0}^{\infty} Pr^{\omega}_{\nu,D}\big(\Pi(z,n)\big).$$

By induction on the number n of transitions needed to reach G, it can be proved that $Pr^{\omega}_{\nu,D^z}\big(\Pi(z,n)\big) \geq Pr^{\omega}_{\nu,D}\big(\Pi(z,n)\big)$ for all $n \in \mathbb{N}$, $z \in \mathbb{R}_{\geq 0}$, $\nu \in Distr(\mathcal{S})$ and $D \in GM$: For the base case $n = 0$, it holds that $Pr^{\omega}_{\nu,D^z}\big(\Pi(z,n)\big) = \nu(G) =$

$Pr^{\omega}_{\nu,D}\big(\Pi(z,n)\big)$ and the claim follows. In the induction step, the induction hypothesis yields

$$\begin{aligned}
&Pr^{\omega}_{\nu,D^z}\big(\Pi(z,n+1)\big)\\
&= \sum_{s\in\mathcal{S}} \nu(s) \int_0^z \lambda(s)e^{-\lambda(s)t} \sum_{s'\in\mathcal{S}} \mathbf{P}\big(s, D^z(s,t), s'\big) \cdot Pr^{\omega}_{\nu_{s'},D^{z-t}}\big(\Pi(z-t,n)\big)\ dt\\
&\geq \sum_{s\in\mathcal{S}} \nu(s) \int_0^z \lambda(s)e^{-\lambda(s)t} \sum_{s'\in\mathcal{S}} \mathbf{P}\big(s, D^z(s,t), s'\big) \cdot Pr^{\omega}_{\nu_{s'},D'}\big(\Pi(z-t,n)\big)\ dt
\end{aligned}$$

for all schedulers $D' \in GM$. Hence

$$\begin{aligned}
&Pr^{\omega}_{\nu,D^z}\big(\Pi(z,n+1)\big)\\
&\geq \sum_{s\in\mathcal{S}} \nu(s) \int_0^z \lambda(s)e^{-\lambda(s)t} \sum_{s'\in\mathcal{S}} \mathbf{P}\big(s, D^z(s,t), s'\big) \sup_{D'\in GM} Pr^{\omega}_{\nu_{s'},D'}\big(\Pi(z-t,n)\big)\ dt\\
&= \sum_{s\in\mathcal{S}} \nu(s) \int_0^z \lambda(s)e^{-\lambda(s)t} \max_{\alpha\in Act} \sum_{s'\in\mathcal{S}} \mathbf{P}(s,\alpha,s') \sup_{D'\in GM} Pr^{\omega}_{\nu_{s'},D'}\big(\Pi(z-t,n)\big)\ dt\\
&\geq \sum_{s\in\mathcal{S}} \nu(s) \int_0^z \lambda(s)e^{-\lambda(s)t} \max_{\alpha\in Act} \sum_{s'\in\mathcal{S}} \mathbf{P}(s,\alpha,s') \cdot Pr^{\omega}_{\nu_{s'},D(s\xrightarrow{t,\alpha}\cdot)}\big(\Pi(z-t,n)\big)\ dt\\
&\geq \sum_{s\in\mathcal{S}} \nu(s) \int_0^z \lambda(s)e^{-\lambda(s)t} \sum_{\alpha\in Act} D(s,t,\{\alpha\}) \sum_{s'\in\mathcal{S}} \mathbf{P}\big(s,\alpha,s'\big)\\
&\qquad\qquad\qquad\qquad\qquad\qquad \cdot Pr^{\omega}_{\nu_{s'},D(s\xrightarrow{t,\alpha}\cdot)}\big(\Pi(z-t,n)\big)\ dt\\
&= Pr^{\omega}_{\nu,D}\big(\Pi(z,n+1)\big)
\end{aligned}$$

holds for all $D \in GM$. This implies that for all $D \in GM$

$$Pr^{\omega}_{\nu,D}\big(\Diamond^{[0,z]}G\big) = \sum_{n=0}^{\infty} Pr^{\omega}_{\nu,D}\big(\Pi(z,n)\big) \leq \sum_{n=0}^{\infty} Pr^{\omega}_{\nu,D^z}\big(\Pi(z,n)\big) = Pr^{\omega}_{\nu,D^z}\big(\Diamond^{[0,z]}G\big).$$

Thus $\sup_{D\in GM} Pr^{\omega}_{\nu_s,D}\big(\Diamond^{[0,z]}G\big) \leq Pr^{\omega}_{\nu_s,D^z}\big(\Diamond^{[0,z]}G\big)$. As $D^z \in GM$, the other direction is trivial and the claim follows. □

Corollary 1. *The class of TTPD schedulers suffices to optimize time-bounded reachability probabilities in locally uniform CTMDPs.*

Based on the class of *TTPD* schedulers, we are now ready to derive a fixed point characterization of the maximum timed reachability probability in locally uniform CTMDPs. A similar technique has been used in [2, Thm. 1] to derive the probability of time-bounded until formulas in CTMCs.

Theorem 2 (Fixed point characterization of timed reachability). *Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a locally uniform CTMDP and $G \subseteq \mathcal{S}$ a set of goal states. Then p^G_{max} is the least fixed point of the higher-order operator $\Omega : (\mathcal{S} \times \mathbb{R}_{\geq 0} \to [0,1]) \to (\mathcal{S} \times \mathbb{R}_{\geq 0} \to [0,1])$ which is defined for all $s \in \mathcal{S}$ and $z \in \mathbb{R}_{\geq 0}$ such that if $s \notin G$,*

$$\Omega(F)(s,z) = \int_0^z \max_{\alpha\in Act} \lambda(s)e^{-\lambda(s)t} \sum_{s'\in\mathcal{S}} \mathbf{P}(s,\alpha,s') \cdot F(s', z-t)\ dt \qquad (2)$$

and $\Omega(F)(s,z) = 1$, otherwise.

Proof. First, we prove that p^G_{max} is a fixed point of Ω: If $s \in G$, then $p^G_{max}(s,z) = 1 = \Omega\left(p^G_{max}\right)(s,z)$ and the claim follows. If $s \notin G$, we proceed as follows:

$$\begin{aligned}
\Omega\big(p^G_{max}\big)(s,z) &= \int_0^z \max_{\alpha\in Act} \lambda(s)e^{-\lambda(s)t} \sum_{s'\in\mathcal{S}} \mathbf{P}(s,\alpha,s')\cdot p^G_{max}(s',z-t)\;dt \\
&= \int_0^z \lambda(s)e^{-\lambda(s)t} \underbrace{\max_{\alpha\in Act} \sum_{s'\in\mathcal{S}} \mathbf{P}(s,\alpha,s')\cdot p^G_{max}(s',z-t)}_{f(s',z-t,\alpha)}\;dt \\
&= \int_0^z \lambda(s)e^{-\lambda(s)t} \sum_{s'\in\mathcal{S}} \mathbf{P}(s,D^z(s,t),s')\cdot \sup_{D\in GM} Pr^{\omega}_{\nu_{s'},D}\big(\Diamond^{[0,z-t]}G\big)\;dt \\
&= \int_0^z \lambda(s)e^{-\lambda(s)t} \sum_{\alpha\in Act} D^z(s,t,\{\alpha\}) \sum_{s'\in\mathcal{S}} \mathbf{P}(s,\alpha,s')\cdot \sup_{D\in GM} Pr^{\omega}_{\nu_{s'},D}\big(\Diamond^{[0,z-t]}G\big)\;dt \\
&= \sup_{D\in GM} Pr^{\omega}_{\nu_s,D}\big(\Diamond^{[0,z]}G\big) = p^G_{max}(s,z).
\end{aligned}$$

It remains to show that p^G_{max} is the least fixed point of Ω: Let $\Pi(s,z,n) \subseteq \Pi(z,n)$ be the set of all infinite paths $\pi = s_0 \xrightarrow{t_0,\alpha_0} s_1 \xrightarrow{t_1,\alpha_1} \cdots$ such that $s_0 = s$, $s_n \in G$, $s_i \notin G$ for all $i < n$ and $\sum_{i=0}^{n-1} t_i \leq z$. Further, let $p^{G,n}_{max}(s,z) = \sup_{D\in GM} Pr^{\omega}_{\nu_s,D}\left(\biguplus_{i=0}^{n} \Pi(s,z,i)\right)$ be the upper bound for timed reachability of G within time bound z and with at most n transitions. From the first part, we know that p^G_{max} is a fixed point of Ω; further, from the definition of Ω, it follows that $p^{G,n+1}_{max}(s,z) = \Omega(p^{G,n}_{max})(s,z)$. Therefore, $\lim_{n\to\infty} p^{G,n}_{max} = p^G_{max}$. Now, let $F : \mathcal{S}\times\mathbb{R}_{\geq 0} \to [0,1]$ be another fixed point of Ω. By induction on n, we show that $p^{G,n}_{max}(s,z) \leq F(s,z)$ for all $n\in\mathbb{N}$. For the base case $p^{G,0}_{max}(s,z) = 1 = F(s,z)$ if $s \in G$ and $p^{G,0}_{max}(s,z) = 0 \leq F(s,z)$, otherwise. Further,

$$\begin{aligned}
p^{G,n+1}_{max}(s,z) &= \int_0^z \lambda(s)e^{-\lambda(s)t} \max_{\alpha\in Act} \sum_{s'\in\mathcal{S}} \mathbf{P}(s,\alpha,s')\cdot p^{G,n}_{max}(s',z-t)\;dt \\
&\leq \int_0^z \lambda(s)e^{-\lambda(s)t} \max_{\alpha\in Act} \sum_{s'\in\mathcal{S}} \mathbf{P}(s,\alpha,s')\cdot F(s',z-t)\;dt = F(s,z).
\end{aligned}$$

Hence, $F(s,z) \geq \lim_{n\to\infty} p^{G,n}_{max}(s,z) = p^G_{max}(s,z)$ and the claim follows. $\square$

3.1 Piecewise constant schedulers

In Def. 4, the upper bound $p^G_{\max}$ on timed reachability of a set of goal states G is defined w.r.t. the class of *GM* schedulers. Corollary 1 allows us to only consider the subclass of *TTPD* schedulers to compute p^G_{max}, i.e. we restrict to schedulers of the form $D : \mathcal{S}\times\mathbb{R}_{\geq 0} \to Act$. Note however, that also *TTPD* schedulers are continuous in their second argument, i.e. they may yield different actions $D(s,t)$ for any point in time $t \in \mathbb{R}_{\geq 0}$. Thus, the set *TTPD* is still uncountable.

Definition 9 (Piecewise constant *TTPD* scheduler). *Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a CTMDP and $D : \mathcal{S}\times\mathbb{R}_{\geq 0} \to Act$ a TTPD scheduler. D is* piecewise constant *iff for all $s\in\mathcal{S}$ and $\alpha \in Act(s)$ there exist disjoint intervals $A^0_{s,\alpha}, A^1_{s,\alpha}, A^2_{s,\alpha}, \dots \subseteq \mathbb{R}_{\geq 0}$ such that $D(s,t) = \alpha \iff t \in \bigcup_{i=0}^{\infty} A^i_{s,\alpha}$. A piecewise constant scheduler D is* non-Zeno *iff for all $z\in\mathbb{R}_{\geq 0}$, $s\in\mathcal{S}$ and $\alpha\in Act$: $\left|\left\{A^i_{s,\alpha} \mid z > \inf A^i_{s,\alpha}\right\}\right| < \infty$.*

We use *PCD* to denote the set of all piecewise constant and non-Zeno *TTPD* schedulers. Intuitively, for state $s \in \mathcal{S}$ and a given time-bound z, a *PCD*-scheduler changes its decision for an action only finitely many times: The intervals $A^i_{s,\alpha}$ in Def. 9 exactly describe the time-periods in which the scheduler constantly chooses action α. The non-Zeno assumption implies that only finitely many decision epochs occur up to time z.

Theorem 3 (*PCD* schedulers suffice for timed reachability). *Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a CTMDP, $G \subseteq \mathcal{S}$ and $z \in \mathbb{R}_{\geq 0}$. Then*

$$\sup_{D \in PCD} Pr^{\omega}_{\nu,D}\left(\Diamond^{[0,z]}G\right) = \sup_{D \in TTPD} Pr^{\omega}_{\nu,D}\left(\Diamond^{[0,z]}G\right).$$

Proof. We prove that the *TTPD* scheduler D^z from the proof of Thm. 1 can be approximated arbitrarily closely by *PCD* schedulers: Let $s \in \mathcal{S}$, $\alpha \in Act$ and define $A_{s,\alpha} = D^z(s,\cdot)^{-1}(\alpha)$. By definition, D^z is a measurable scheduler. Hence $A_{s,\alpha} \in \mathfrak{B}$. Now let $\mathfrak{B}_0$ be a field of subsets of $\mathbb{R}_{\geq 0}$ that generates the σ-field $\mathfrak{B}$, i.e. let $\sigma(\mathfrak{B}_0) = \mathfrak{B}$. By the Approximation Theorem [1, Thm. 1.3.11], given $\varepsilon > 0$, we can approximate the set $A_{s,\alpha}$ by a set $B_{s,\alpha} \in \mathfrak{B}_0$ up to an error of ε. More precisely, let $\theta : \mathfrak{B} \to \mathbb{R}_{\geq 0}$ be the Lebesgue measure defined by the distribution function $\Theta(x) = x$ for $x \in \mathbb{R}_{\geq 0}$ and let $A \triangle B = (A \setminus B) \cup (B \setminus A)$ denote set difference. Then $\theta\left(A_{s,\alpha} \triangle B_{s,\alpha}\right) < \varepsilon$.

For $\mathfrak{B}_0$, we choose the set of finite disjoint unions of right semi-closed intervals; as $\mathfrak{B}_0$ is a field and $\sigma(\mathfrak{B}_0) = \mathfrak{B}$, this is a valid choice (see also [1, Ex. 1.2.4]). As $B_{s,\alpha} \in \mathfrak{B}_0$, there exist $n_{s,\alpha} \in \mathbb{N}$ and disjoint intervals $B^0_{s,\alpha}, \ldots, B^{n_{s,\alpha}}_{s,\alpha}$ such that $B_{s,\alpha} = \biguplus_{i=0}^{n_{s,\alpha}} B^i_{s,\alpha}$. Now we are ready to construct a scheduler D^z_ε which approximates D^z up to an error of ε as follows: $D^z_\varepsilon(s,t) = \alpha \iff t \in \biguplus_{i=0}^{n_{s,\alpha}} B^i_{s,\alpha}$. By definition, D^z_ε is a piecewise constant and a non-Zeno scheduler. Thus $D^z_\varepsilon \in PCD$ for all $\varepsilon > 0$; further, as $\theta\left(\{t \in \mathbb{R}_{\geq 0} \mid D^z(s,t) \neq D^z_\varepsilon(s,t)\}\right) < \varepsilon$, we obtain that $\lim_{\varepsilon \to 0} \mu_{D^z_\varepsilon} = \mu_{D^z}$ (cf. Def. 5). Hence

$$\lim_{\varepsilon \to 0} Pr^{\omega}_{\nu_s,D^z_\varepsilon}\left(\Diamond^{[0,z]}G\right) = Pr^{\omega}_{\nu_s,D^z}\left(\Diamond^{[0,z]}G\right) = \sup_{D \in TTPD} Pr^{\omega}_{\nu_s,D^z}\left(\Diamond^{[0,z]}G\right).$$

Therefore, we derive

$$\begin{aligned}\sup_{D \in TTPD} Pr^{\omega}_{\nu,D}\left(\Diamond^{[0,z]}G\right) &= Pr^{\omega}_{\nu,D^z}\left(\Diamond^{[0,z]}G\right) = \sum_{s \in \mathcal{S}} \nu(s) \cdot Pr^{\omega}_{\nu_s,D^z}\left(\Diamond^{[0,z]}G\right) \\ &= \sum_{s \in \mathcal{S}} \nu(s) \cdot \lim_{\varepsilon \to 0} Pr^{\omega}_{\nu_s,D^z_\varepsilon}\left(\Diamond^{[0,z]}G\right) = \lim_{\varepsilon \to 0} Pr^{\omega}_{\nu,D^z_\varepsilon}\left(\Diamond^{[0,z]}G\right).\end{aligned}$$

Now the claim follows, as $D^z_\varepsilon \in TTPD$ for all $\varepsilon > 0$. □

In order to derive our discretization for CTMDPs, we restrict to a subclass of piecewise constant non-Zeno schedulers, called τ-schedulers:

Definition 10 (τ-scheduler). *Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a CTMDP, $\tau \in \mathbb{R}_{>0}$ and $D \in PCD$ a non-Zeno scheduler. D is a τ-scheduler iff for all $s \in \mathcal{S}$, $k \in \mathbb{N}$:*

$$\exists \alpha \in Act(s).\ \forall t \in [k\tau, (k+1)\tau)\ .\ D(s,t) = \alpha.$$

Intuitively, a *PCD* scheduler is a τ-scheduler if its choices remain constant on intervals of length τ. We now prove that for $\tau \to 0$, the probabilities of *PCD* and τ-schedulers coincide:

Theorem 4 (Limiting τ-scheduler). *Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a CTMDP, $z \in \mathbb{R}_{\geq 0}$ and $G \subseteq \mathcal{S}$. For any $D \in PCD$, there exist τ-schedulers D_τ such that*

$$\lim_{\tau \to 0} Pr^{\omega}_{\nu, D_\tau}\left(\Diamond^{[0,z]} G\right) = Pr^{\omega}_{\nu, D}\left(\Diamond^{[0,z]} G\right).$$

Proof. As $D \in PCD$, there exists $n_{s,\alpha} \in \mathbb{N}$ and disjoint intervals $B^0_{s,\alpha}, \ldots, B^{n_{s,\alpha}}_{s,\alpha}$ for all $s \in \mathcal{S}$ and $\alpha \in Act$ such that $D(s,t) = \alpha$ iff $t \in B^i_{s,\alpha}$ for some $i \leq n_{s,\alpha}$. If $\tau \to 0$, we can approximate those intervals arbitrarily closely, that is, there exist schedulers D_τ such that $D_\tau(s,\cdot)^{-1}(\alpha) \to D(s,\cdot)^{-1}(\alpha)$. Similar to the proof of Thm. 3, this implies that $\lim_{\tau \to 0} \mu_{D_\tau} = \mu_D$ and therefore

$$\lim_{\tau \to 0} Pr^{\omega}_{\nu, D_\tau}\left(\Diamond^{[0,z]} G\right) = Pr^{\omega}_{\nu, D}\left(\Diamond^{[0,z]} G\right)$$

proving the claim. $\square$

4 Discretization for CTMDPs

In this section, we discuss how to compute the maximum time-bounded reachability probability for CTMDPs using discretization. Moreover, we also discuss how to bound the error induced during discretization. Given a CTMDP $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$, a set $G \subseteq \mathcal{S}$ of goal states, a starting state s and a time-bound $z \in \mathbb{R}_{\geq 0}$, let $p^G_{max}(s,z)$ be defined as in Def. 7. Then, p^G_{max} is the least fixed point of the higher-order operator $\Omega : (\mathcal{S} \times \mathbb{R}_{\geq 0} \to [0,1]) \to (\mathcal{S} \times \mathbb{R}_{\geq 0} \to [0,1])$, as defined in Thm. 2. The case that $s \in G$ is trivial, as $p^G_{max}(s,z) = 1$ for all $z \in \mathbb{R}_{\geq 0}$. To derive $p^G_{max}(s,z)$ for $s \notin G$, we consider the first sub-interval $[0,\tau]$ of the integral in Eq. (2) and split it accordingly:

$$\begin{aligned} p^G_{max}(s,z) = &\int_0^\tau \lambda(s) e^{-\lambda(s)t} \max_{\alpha \in Act} \sum_{s' \in \mathcal{S}} \mathbf{P}(s,\alpha,s') \cdot p^G_{max}(s', z-t)\, dt \\ &+ \int_\tau^z \lambda(s) e^{-\lambda(s)t} \max_{\alpha \in Act} \sum_{s' \in \mathcal{S}} \mathbf{P}(s,\alpha,s') \cdot p^G_{max}(s', z-t)\, dt. \end{aligned} \tag{3}$$

Now, let $A(s,z)$ and $B(s,z)$ denote the first, resp. second summand in (3). Shifting the range of integration in $B(s,z)$ by $(-\tau)$, we derive

$$\begin{aligned} B(s,z) &= \int_0^{z-\tau} \lambda(s) e^{-\lambda(s)(t+\tau)} \max_{\alpha \in Act} \sum_{s' \in \mathcal{S}} \mathbf{P}(s,\alpha,s') \cdot p^G_{max}(s', z-(t+\tau))\, dt \\ &= \int_0^{z-\tau} \lambda(s) e^{-\lambda(s)t} e^{-\lambda(s)\tau} \max_{\alpha \in Act} \sum_{s' \in \mathcal{S}} \mathbf{P}(s,\alpha,s') \cdot p^G_{max}(s', z-(t+\tau))\, dt \\ &= e^{-\lambda(s)\tau} \underbrace{\int_0^{z-\tau} \lambda(s) e^{-\lambda(s)t} \max_{\alpha \in Act} \sum_{s' \in \mathcal{S}} \mathbf{P}(s,\alpha,s') \cdot p^G_{max}(s', (z-\tau)-t)\, dt}_{p^G_{max}(s, z-\tau)} \\ &= e^{-\lambda(s)\tau} \cdot p^G_{max}(s, z-\tau). \end{aligned}$$

Thus, $B(s,z)$ is the maximum probability of the event that starting from state s, the set G is reached within z time units while no transition occurs in the time interval $[0,\tau]$. To be more precise, let $\#_{[0,\tau]} : Paths^{\omega} \to \mathbb{N}$ be the random variable for the number of transitions that occur in interval $[0,\tau]$. Then, it holds that $B(s,z) = \sup_{D\in TTPD} Pr^{\omega}_{\nu_s,D}\left(\Diamond^{[0,z]}G \cap \#_{[0,\tau]} = 0\right)$. With the same reasoning, the first summand $A(s,z)$ of (3) is the maximum probability to reach G within time z with at least one transition taking place in $[0,\tau]$. Hence, $A(s,z) = \sup_{D\in TTPD} Pr^{\omega}_{\nu_s,D}\left(\Diamond^{[0,z]}G \cap \#_{[0,\tau]} \geq 1\right)$.

In the next step, we decompose the underlying event of $A(s,z)$ as follows:

$$\left(\Diamond^{[0,z]}G \cap \#_{[0,\tau]} \geq 1\right) = \biguplus_{n=1}^{\infty} \left(\Diamond^{[0,z]}G \cap \#_{[0,\tau]} = n\right).$$

Now let $A_n(s,z)$ be the maximum probability to reach G in z time units with exactly n transitions occurring in the first time slice $[0,\tau]$. If we maximize each of the events $\left(\Diamond^{[0,z]}G \cap \#_{[0,\tau]} = n\right)$ separately, we obtain the probability

$$A_n(s,z) = \sup_{D\in TTPD} Pr^{\omega}_{\nu_s,D}\left(\Diamond^{[0,z]}G \cap \#_{[0,\tau]} = n\right). \tag{4}$$

To relate $A(s,z)$ with the probabilities $A_n(s,z)$, observe that

$$\begin{aligned}
A(s,z) &= \sup_{D\in TTPD} Pr^{\omega}_{\nu_s,D}\left(\Diamond^{[0,z]}G \cap \#_{[0,\tau]} \geq 1\right)\\
&= \sup_{D\in TTPD} Pr^{\omega}_{\nu_s,D}\left(\biguplus_{n=1}^{\infty}\left(\Diamond^{[0,z]}G \cap \#_{[0,\tau]} = n\right)\right)\\
&\leq \sum_{n=1}^{\infty}\left(\sup_{D\in TTPD} Pr^{\omega}_{\nu_s,D}\left(\Diamond^{[0,z]}G \cap \#_{[0,\tau]} = n\right)\right) = \sum_{n=1}^{\infty} A_n(s,z).
\end{aligned} \tag{5}$$

For our discretization, we approximate $A(s,z)$ from below via $A_1(s,z)$: Intuitively, for small τ, the probability that more than one transition occurs in interval $[0,\tau]$ is negligibly small. We prove that $A_1(s,z)$ converges to $A(s,z)$ for $\tau \to 0$:

- First, note that $A_1(s,z) \leq A(s,z)$. Thus $A_1(s,z)$ is a lower bound on $A(s,z)$.
- To establish an upper bound, first recall that for an exponential distribution with rate λ and a time interval $[0,\tau]$, the Poisson distribution $\rho(n,\lambda\tau) = e^{-\lambda\tau}\cdot\frac{(\lambda\tau)^n}{n!}$ expresses the probability that n transitions occur within $[0,\tau]$. If $\lambda = \max_{s\in\mathcal{S}}\lambda(s)$ is the maximum exit rate over all states, this allows to derive an upper bound for $A_n(s,z)$:

$$\begin{aligned}
A_n(s,z) &= \sup_{D\in TTPD}\left(\Diamond^{[0,z]}G \cap \#_{[0,\tau]} = n\right) \leq \sup_{D\in TTPD} Pr^{\omega}_{\nu_s,D}\left(\#_{[0,\tau]} = n\right)\\
&\leq \rho(n,\lambda\tau) = e^{-\lambda\tau}\cdot\frac{(\lambda\tau)^n}{n!}.
\end{aligned} \tag{6}$$

  Moreover, by maximality of λ, the probability that more than n transitions occur in any state $s \in \mathcal{S}$ within time interval $[0,\tau]$ is at most

$$\sum_{i=n+1}^{\infty}\rho(i,\lambda\tau) = e^{-\lambda\tau}\sum_{i=n+1}^{\infty}\frac{(\lambda\tau)^i}{i!} = e^{-\lambda\tau}\cdot R_n\big(\lambda\tau\big), \tag{7}$$

where $R_n(x) = \sum_{i=n+1}^{\infty} \frac{x^i}{i!}$ is the remainder term of the Taylor expansion of $f(x) = e^x$ for the point $a = 0$. By Taylor's Theorem, there exists $\xi \in [0, \lambda\tau]$ such that

$$R_n\big(\lambda\tau\big) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \cdot (\lambda\tau)^{n+1} = \frac{e^{\xi}}{(n+1)!} \cdot (\lambda\tau)^{n+1}. \tag{8}$$

Hence, we obtain an upper bound for the error that is induced by approximating $A(s, z)$ by only considering $A_1(s, z)$:

$$A(s,z) \overset{(5)}{\leq} \sum_{n=1}^{\infty} A_n(s,z) \overset{(6)}{\leq} A_1(s,z) + \sum_{n=2}^{\infty} \rho(n,\lambda\tau) \overset{(7)}{=} A_1(s,z) + e^{-\lambda\tau} R_1\big(\lambda\tau\big).$$

By Eq. (8), there exists $\xi \in [0, \lambda\tau]$ such that $R_1(\lambda\tau) = \frac{e^{\xi}}{2} \cdot (\lambda\tau)^2$. To derive an upper bound, choose ξ maximal in $[0, \lambda\tau]$. Then

$$A_1(s,z) + e^{-\lambda\tau} R_1\big(\lambda\tau\big) \leq A_1(s,z) + e^{-\lambda\tau} \cdot \frac{e^{\lambda\tau}}{2} (\lambda\tau)^2 = A_1(s,z) + \frac{(\lambda\tau)^2}{2}.$$

Therefore, we have that $A(s, z) \leq A_1(s, z) + \frac{(\lambda\tau)^2}{2}$.

Taking the lower and upper bound together, we finally obtain

$$A_1(s,z) \leq A(s,z) \leq A_1(s,z) + \frac{(\lambda\tau)^2}{2}. \tag{9}$$

Therefore, we may approximate $A(s, z)$ from below via $A_1(s, z)$, allowing for an error of at most $\frac{(\lambda\tau)^2}{2}$. Exploiting the fact that in this approximation, there is exactly one transition within $[0, \tau]$, the following lemma simplifies $A_1(s, z)$:

Lemma 1. *Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a locally uniform CTMDP, $G \subseteq \mathcal{S}$ a set of goal states and $z \in \mathbb{R}_{\geq 0}$ a time bound. For $A_1(s, z)$ as defined in Eq. (4), it holds*

$$A_1(s,z) = \Big(1 - e^{-\lambda(s)\tau}\Big) \max_{\alpha \in Act} \sum_{s' \in \mathcal{S}} \mathbf{P}(s,\alpha,s') \cdot p^G_{max}(s', z-\tau). \tag{10}$$

Proof. First, let $D^z \in TTPD$ be the optimal $TTPD$ scheduler as defined in Def. 8. By Thm. 1, it holds that $Pr^{\omega}_{\nu_s,D^z}\big(\Diamond^{[0,z]}G\big) = p^G_{max}(s, z)$. As D^z is measurable, there exists a countable sequence $0 = \tau_0 < \tau_1 < \cdots \in \mathbb{R}_{\geq 0}$ with $\tau_i \to \tau$ and actions $\alpha_0, \alpha_1, \ldots$ such that $D^z(s, t) = \alpha_i$ for all $t \in [\tau_i, \tau_{i+1})$ and $i \in \mathbb{N}$. Hence, we can split the integral in $A(s, z)$ as follows:

$$\begin{aligned} A(s,z) =& \sum_{i=0}^{\infty} \int_{\tau_i}^{\tau_{i+1}} \lambda(s) e^{-\lambda(s)t} \sum_{s' \in \mathcal{S}} \mathbf{P}(s,\alpha_i,s') \cdot p^G_{max}(s', z-t)\, dt \\ =& \sum_{i=0}^{\infty} \int_{\tau_i}^{\tau_{i+1}} \lambda(s) e^{-\lambda(s)t} \max_{\alpha \in Act} \sum_{s' \in \mathcal{S}} \mathbf{P}(s,\alpha,s') \cdot p^G_{max}(s', z-t)\, dt. \end{aligned}$$

Recall that $A_1(s, z)$ is the maximum probability to reach G in z time units with exactly one transition occurring in $[0, \tau]$. Therefore, we obtain $A_1(s, z)$ from the term above, by replacing $p^G_{max}(s', z - t)$ with $p^G_{max}(s', z - \tau)$ (cf. Eq. (3)):

$$A_1(s,z) = \sum_{i=0}^{\infty} \int_{\tau_i}^{\tau_{i+1}} \lambda(s) e^{-\lambda(s)t} \max_{\alpha \in Act} \sum_{s' \in \mathcal{S}} \mathbf{P}(s,\alpha,s') \cdot p^G_{max}(s', z-\tau)\, dt.$$

Now, the term $\max_{\alpha \in Act} \sum_{s' \in \mathcal{S}} \mathbf{P}(s, \alpha, s') \cdot p^G_{max}(s', z - \tau)$ is independent of i and t; thus, it can be taken out of the sum and the integral:

$$\begin{aligned} A_1(s,z) &= \left(\max_{\alpha \in Act} \sum_{s' \in \mathcal{S}} \mathbf{P}(s,\alpha,s') \cdot p^G_{max}(s', z-\tau) \right) \cdot \sum_{i=0}^{\infty} \left(e^{-\lambda(s)\tau_i} - e^{-\lambda(s)\tau_{i+1}} \right) \\ &= \left(1 - e^{-\lambda(s)\tau}\right) \max_{\alpha \in Act} \sum_{s' \in \mathcal{S}} \mathbf{P}(s,\alpha,s') \cdot p^G_{max}(s', z-\tau). \end{aligned}$$

□

Based on $A_1(s,z)$ and $B(s,z)$ we are now ready to derive a *discretization* for $p^G_{max}(s,z)$ on CTMDPs with respect to a *step duration* τ:

Definition 11 (Discretization). *Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a locally uniform CTMDP, and let τ be a step duration. The induced MDP $\mathcal{C}_\tau = (\mathcal{S}, Act, \mathbf{P}_\tau, \nu)$ is defined such that for all $s, s' \in \mathcal{S}$ and $\alpha \in Act(s)$:*

$$\mathbf{P}_\tau(s,\alpha,s') = \begin{cases} \left(1 - e^{-\lambda(s)\tau}\right) \cdot \mathbf{P}(s,\alpha,s') & \text{if } s \neq s' \\ \left(1 - e^{-\lambda(s)\tau}\right) \cdot \mathbf{P}(s,\alpha,s') + e^{-\lambda(s)\tau} & \text{if } s = s'. \end{cases} \tag{11}$$

Further, for all $\alpha \notin Act(s)$, we define $\mathbf{P}_\tau(s,\alpha,s') = 0$.

In the resulting MDP $\mathcal{C}_\tau$, each step corresponds to one time slice of length τ in the original CTMDP. For a single step and a fixed successor state $s' \neq s$, $\mathbf{P}_\tau(s,\alpha,s')$ equals the probability that a transition to s' occurs within τ time units, given that α is chosen. In case of $s' = s$, the first summand of $\mathbf{P}_\tau(s,\alpha,s)$ is the probability to take the loop back to s; the second summand denotes the probability that no transition occurs within time τ and thus $s = s'$.

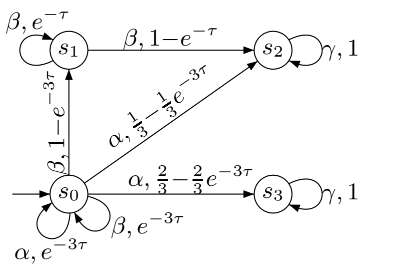


Fig. 2. Discretization of Fig. 1.

Example 2. Reconsider the locally uniform CTMDP $\mathcal{C}$ in Fig. 1. To compute the maximum probability to reach $G = \{s_2\}$ within z time units up to a precision of ε, choose $k \in \mathbb{N}$ such that $\varepsilon \geq \frac{(\lambda z)^2}{2k}$. The discretization step $\tau = \frac{z}{k}$ then induces the discretized MDP $\mathcal{C}_\tau$ which is depicted in Fig. 2 where $\lambda = \max_{s \in \mathcal{S}} \lambda(s) = 3$.

Let $p^{\mathcal{C}_\tau}_{max}(s,k)$ be the maximum probability to reach G starting from s in at most k steps in the MDP $\mathcal{C}_\tau$. Therefore, $p^{\mathcal{C}_\tau}_{max}(s,k) = 1$ if $s \in G$ and $p^{\mathcal{C}_\tau}_{max}(s,0) = 0$ if $s \notin G$. Further, for $s \notin G$ and $k > 0$, $p^{\mathcal{C}_\tau}_{max}$ is defined recursively:

$$p^{\mathcal{C}_\tau}_{max}(s,k) = \max_{\alpha \in Act} \sum_{s' \in \mathcal{S}} \mathbf{P}_\tau(s,\alpha,s') \cdot p^{\mathcal{C}_\tau}_{max}(s', k-1). \tag{12}$$

The next theorem proves that the probabilities in $\mathcal{C}_\tau$ to reach G from state s within at most $k = \frac{z}{\tau}$ steps converge from below (for $\tau \to 0$) to the corresponding time-bounded reachability probability in $\mathcal{C}$:

Theorem 5. *Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a CTMDP, $G \subseteq \mathcal{S}$ a set of goal states, $z \in \mathbb{R}_{\geq 0}$ a time bound and $k > 0$ the number of time steps, i.e. $\tau = \frac{z}{k}$. For all $s \in S$, it holds:*

$$p^{\mathcal{C}_\tau}_{max}(s,k) \leq p^G_{max}(s,z) \leq p^{\mathcal{C}_\tau}_{max}(s,k) + \frac{(\lambda z)^2}{2k}. \tag{13}$$

Proof. Firstly, recall it holds that $p^G_{max}(s,z) = A(s,z) + B(s,z)$ and $A_1(s,z) \leq A(s,z) \leq A_1(s,z) + \frac{(\lambda\tau)^2}{2}$ (see Eq. (9)). The Inequality (13) follows by induction on k. For $k = 1$, we have $z = \tau$. If $s \in G$, then $p^{\mathcal{C}_\tau}_{max}(s,1) = 1 = p^G_{max}(s,\tau)$ and Inequality 13 follows. If $s \notin G$, the lower bound is established as $p^{\mathcal{C}_\tau}_{max}(s,1) = \max_{\alpha\in Act}\left(1 - e^{-\lambda(s)\tau}\right) \cdot \mathbf{P}(s,\alpha,G) = A_1(s,\tau) \leq p^G_{max}(s,\tau)$. For the upper bound, note that $s \notin G$ implies $B(s,\tau) = 0$, thus $p^G_{max}(s,\tau) \leq \frac{(\lambda\tau)^2}{2} + A_1(s,\tau) = p^{\mathcal{C}_\tau}_{max}(s,\tau) + \frac{(\lambda\tau)^2}{2}$. For the induction step, together with Lemma 1 and the definition of $\mathcal{C}_\tau$, we have that:

$$A_1(s,z) + B(s,z) = \max_{\alpha\in Act} \sum_{s'\in\mathcal{S}} \mathbf{P}_\tau(s,\alpha,s') \cdot p^G_{max}(s', z-\tau). \tag{14}$$

First we consider the lower bound on the left part of Inequality 13: By induction hypothesis it holds that $p^{\mathcal{C}_\tau}_{max}(s',k-1) \leq p^G_{max}(s',z-\tau)$ for all $s' \in \mathcal{S}$. Thus

$$\begin{aligned} p^G_{max}(s,z) &\geq A_1(s,z) + B(s,z) \\ &\overset{(14)}{=} \max_{\alpha\in Act} \sum_{s'\in\mathcal{S}} \mathbf{P}_\tau(s,\alpha,s') \cdot p^G_{max}(s',z-\tau) \\ &\overset{i.h.}{\geq} \max_{\alpha\in Act} \sum_{s'\in\mathcal{S}} \mathbf{P}_\tau(s,\alpha,s') \cdot p^{\mathcal{C}_\tau}_{max}(s',k-1) = p^{\mathcal{C}_\tau}_{max}(s,k). \end{aligned}$$

For the upper bound: $A(s,z) \leq A_1(s,z) + \frac{(\lambda\tau)^2}{2}$ and Eq. (14) imply that

$$p^G_{max}(s,z) \leq \frac{(\lambda\tau)^2}{2} + \max_{\alpha\in Act} \sum_{s'\in\mathcal{S}} \mathbf{P}_\tau(s,\alpha,s') \cdot p^G_{max}(s',z-\tau).$$

Applying the induction hypothesis, we obtain

$$p^G_{max}(s,z) \overset{i.h.}{\leq} \frac{(\lambda\tau)^2}{2} + \max_{\alpha\in Act} \sum_{s'\in\mathcal{S}} \mathbf{P}_\tau(s,\alpha,s') \left(p^{\mathcal{C}_\tau}_{max}(s',k-1) + \frac{(\lambda(z-\tau))^2}{2(k-1)} \right).$$

By inspection, the right part can be simplified to $p^{\mathcal{C}_\tau}_{max}(s,k) + \frac{(\lambda z)^2}{2k}$, completing the proof for the upper bound. □

4.1 Algorithm and complexity

Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a locally uniform CTMDP, G a set of goal states and z a time bound. For some accuracy $\varepsilon > 0$, let k be the number of steps needed to satisfy $\varepsilon \geq \frac{(\lambda z)^2}{2k}$. Then $\tau = \frac{z}{k}$ induces the discretized MDP $\mathcal{C}_\tau$ of $\mathcal{C}$ with discretization step τ. By Thm. 13, the maximum probability to reach G within z time units in $\mathcal{C}$ can be approximated (up to ε) by maximizing the step-bounded reachability of G in $\mathcal{C}_\tau$ within k steps. In $\mathcal{C}_\tau$, this probability can easily be computed by the well-known *value iteration* approach [4]. Briefly, it starts with a probability vector $\mathbf{v}_0$ with $\mathbf{v}_0(s) = 1$ if $s \in G$ and 0, otherwise. In each iteration, $\mathbf{v}_i$ is obtained from $\mathbf{v}_{i-1}$ according to Eq. 12. In each round, i denotes the number of steps in the MDP $\mathcal{C}_\tau$; hence, $\mathbf{v}_i(s)$ equals $p^{\mathcal{C}_\tau}_{max}(s,i)$.

The value iteration approach yields the following complexity. For $s \in \mathcal{S}$ and $\alpha \in Act(s)$, let $post(s,\alpha) = \{s' \in \mathcal{S} \mid \mathbf{R}(s,\alpha,s') > 0\}$ be the set of α-successors of state s. The *size* of $\mathcal{C}$ is denoted by $m = \sum_{s\in\mathcal{S}} \sum_{\alpha\in Act} |post(s,\alpha)|$. In the worst case, the discretized MDP $\mathcal{C}_\tau$ is obtained by adding a self-loop for each state $s \in \mathcal{S}$ and action $\alpha \in Act(s)$. Thus, the size of $\mathcal{C}_\tau$ is bounded by $2m$. For a given accuracy ε, it is easy to derive the number k of value-iteration step: By Thm. 5, $\left|p^G_{max}(s,z) - p^{\mathcal{C}_\tau}_{max}(s,k)\right| \leq \frac{(\lambda z)^2}{2k}$. Letting $\frac{(\lambda z)^2}{2k} \leq \varepsilon$, we conclude that the smallest k to guarantee ε is $\frac{(\lambda z)^2}{2\varepsilon}$. In each step of the value iteration, the update of the vector $\mathbf{v}_i$ takes time $2m$. Thus, the time complexity of our approach is $\mathcal{O}(m \cdot (\lambda z)^2/\varepsilon)$.

4.2 Generation of ε-optimal schedulers

Let $\mathcal{C}$, G, z, τ and $\mathcal{C}_\tau$ be as before. A byproduct of the value iteration on the discretized MDP is an ε-optimal scheduler for the set of goal states G and time bound z. More precisely, in any of the i value iteration steps, for each state $s \in \mathcal{S}$, an action $\alpha_{s,i}$ is chosen according to Eq. 12. In this way, we obtain a history dependent (or *step dependent*) scheduler in $\mathcal{C}_\tau$. This scheduler induces a τ-scheduler of the original CTMDP $\mathcal{C}$, denoted D_τ, as follows: $D_\tau(s,t) = \alpha_{s,k}$ if $t \in [k\tau, (k+1)\tau)$. The following theorem shows that D_τ is an ε-optimal scheduler:

Theorem 6 (ε-optimal scheduler). *The scheduler D_τ is an ε-optimal scheduler for $\mathcal{C}$ w.r.t. the maximum timed reachability probability.*

Proof. Let $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ be a locally uniform CTMDP, G a set of goal states and z a time bound. For some accuracy $\varepsilon > 0$, let k be the number of steps needed to satisfy $\varepsilon \geq \frac{(\lambda z)^2}{2k}$. Let $\mathcal{C}_\tau$ be the induced MDP with $\tau = \frac{z}{k}$, and D_τ be the τ-scheduler as described. To show D_τ is an ε-optimal scheduler for $\mathcal{C}$ w.r.t. the maximum timed reachability probability, we show for all state $s \in S$, it holds that

$$\left|Pr^\omega_{\nu_s,D_\tau}\left(\Diamond^{[0,z]}G\right) - p^{\mathcal{C}_\tau}_{max}(s,k)\right| \leq \varepsilon.$$

It is sufficient to show the following Inequality:

$$p^{\mathcal{C}_\tau}_{max}(s,k) \leq Pr^\omega_{\nu_s,D_\tau}\left(\Diamond^{[0,z]}G\right) \leq p^{\mathcal{C}_\tau}_{max}(s,k) + \varepsilon. \tag{15}$$

By Theorem 5, the upper bound can be shown directly:

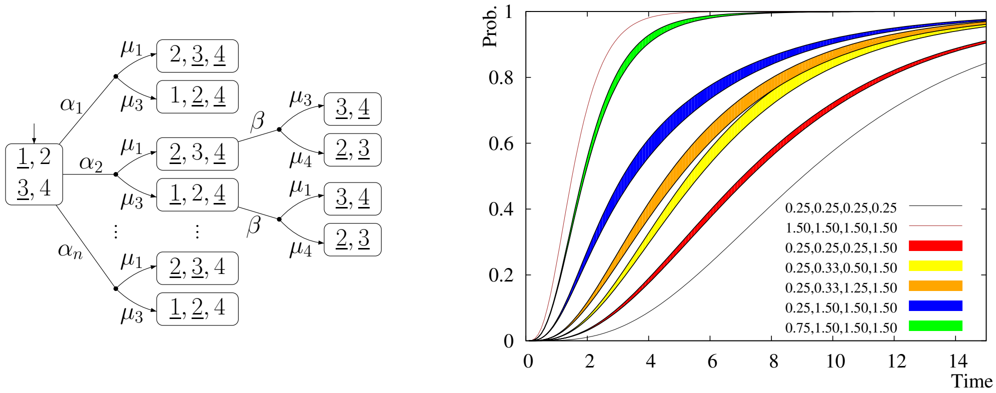
$$Pr^\omega_{\nu_s,D_\tau}\left(\Diamond^{[0,z]}G\right) \leq p^G_{max}(s,z) \leq p^{\mathcal{C}_\tau}_{max}(s,k) + \frac{(\lambda z)^2}{2k} \leq p^{\mathcal{C}_\tau}_{max}(s,k) + \varepsilon.$$

Now we discuss how to show the lower bound of Inequality (15). Firstly, we observe that under any *TTPD* scheduler D the CTMDP $\mathcal{C}$ is totally stochastic, and for $s \notin G$, the probability $Pr^\omega_{\nu_s,D}\left(\Diamond^{[0,z]}G\right)$ can be computed by:

$$Pr^\omega_{\nu_s,D}\left(\Diamond^{[0,z]}G\right) = \int_0^z \lambda(s)e^{-\lambda(s)t} \sum_{s'\in\mathcal{S}} \mathbf{P}(s, D(s,t), s') \cdot Pr^\omega_{\nu_{s'},D}\left(\Diamond^{[0,z-t]}G\right)\, dt$$

Note D_τ is a *TTPD* scheduler, thus it holds that

$$Pr^\omega_{\nu_s,D_\tau}\left(\Diamond^{[0,z]}G\right) = \int_0^z \lambda(s)e^{-\lambda(s)t} \sum_{s'\in\mathcal{S}} \mathbf{P}(s, D_\tau(s,t), s') \cdot Pr^\omega_{\nu_{s'},D_\tau}\left(\Diamond^{[0,z-t]}G\right)\, dt$$



(a) Fragment of the CTMDP model. (b) Best and worst case probabilities.

Fig. 3. Modelling and analysis of the stochastic job scheduling problem.

This integral can then be split into two parts $A(s,z)$ and $B(s,z)$ at time $t=\tau$: it follows in a similar way as Eq. (3) with the difference of taking the action $D_\tau(s,t)$ instead of the maximum over all $\alpha \in Act$. The lower bound can then be established by induction on k, by adapting the lower bound proof of Inequality (13) of Theorem 5 appropriately. $\square$

5 Case study

To show the applicability of our approach, we consider the stochastic job scheduling problem (sJSP) from [5]. In their paper, the authors analyze the *expected* time to complete a set of stochastic jobs on a number of identical processors under a preemptive scheduling policy. An instance of the sJSP is a tuple (m,n,μ) where $m \geq 2$ is the number of processors, n is the number of stochastic jobs and $\mu(i) \in \mathbb{R}_{\geq 0}$ specifies the rate of job i. Each time a job finishes, the preemptive scheduling allows to assign each processors one of the k remaining jobs, giving rise to $\binom{k}{m}$ nondeterministic choices.

Let X_k $(k=1,\ldots,n)$ be the time spent between decision epochs k and $k+1$. In [5], the authors prove that the *longest expected processing time first* (LEPT) policy optimizes the expectation of the sum $\sum_{k=1}^{n} X_k$. In the following we show how to model the sJSP as a locally uniform CTMDP. Applying the results from Sec. 4, we are now able to algorithmically compute the upper and lower probability bounds to finish all jobs within some time bound z.

A *configuration* of the sJSP is a tuple $(R,W) \in 2^J \times 2^J$, where R and W are the sets of running, resp. waiting jobs. When a job $j \in R$ completes, the decision which jobs to schedule next is nondeterministic. An action $\alpha \in Act\left((R,W)\right)$ is a preemptive schedule: If job $j \in R$ finishes first and $\alpha : R \to 2^{R\cup W}$ is chosen, the set $\alpha(j)$ defines the jobs to execute next. In each configuration (R,W), let $Act\left((R,W)\right) = \left\{\alpha : R \to 2^{R\cup W} \,|\, \forall j \in R. j \notin \alpha(j) \wedge |\alpha(j)| \leq m \wedge \alpha(j) \text{ maximal}\right\}$. For $\alpha \in Act\left((R,W)\right)$, we define the $\alpha(j)$-successor (R',W') of (R,W), denoted $(R,W) \xrightarrow{\alpha(j)} (R',W')$, such that $R' = \alpha(j)$ and $W' = (R \cup W) \setminus (\{j\} \cup \alpha(j))$:

Definition 12 (Modelling the sJSP as a CTMDP). *Let $J=(m,n,\mu)$ be a sJSP and (R,W) a configuration. The induced CTMDP $(\mathcal{S}, Act, \mathbf{R}, \nu)$ is defined* *such that* $\mathcal{S} = 2^J \times 2^J$, $\nu = \{(R,W) \mapsto 1\}$, $Act = \bigcup_{(R,W)\in\mathcal{S}} Act\,((R,W))$ *and*

$$\mathbf{R}\big((R,W),\alpha,(R',W')\big) = \begin{cases} \mu(j) & \text{if } (R,W) \xrightarrow{\alpha(j)} (R',W') \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Thus, given configuration (R,W), for every job $j \in R$ and action α, there exists an α-transition with the rate $\mu(j)$ of job j that leads to the $\alpha(j)$-successor (R',W').

Lemma 2 (The sJSP is locally uniform). *For any sJSP* $J = (m,n,\mu)$ *and initial configuration* (R,W)*, the induced CTMDP is locally uniform.*

Proof. From the definition of the induced CTMDP, it directly follows

$$\begin{aligned} E\,((R,W),\alpha) &= \sum_{(R',W')\in\mathcal{S}} \mathbf{R}\,\big((R,W),\alpha,(R',W')\big) \\ &= \sum_{(R,W)\xrightarrow{\alpha(j)}(R',W')} \mathbf{R}\,\big((R,W),\alpha,(R',W')\big) = \sum_{j\in R} \mu(j). \end{aligned}$$

Hence, $E\,((R,W),\alpha) = E\,((R,W),\beta)$ for all $\alpha,\beta \in Act\,((R,W))$. □

Fig. 3(a) depicts a fragment of the CTMDP induced by the $(2,4,\mu)$ sJSP with initial configuration (R,W) where R is given by the underlined process identifiers (i.e. $R = \{1,3\}$) and $W = \{2,4\}$. Action α_1 represents a replacement strategy where jobs $\{3,4\}$ are executed next if job $1 \in R$ finishes first and otherwise, the next jobs are $\{2,4\}$. In Fig. 3(b), we plot the upper and lower probability bounds to finish jobs $\{1,\ldots,4\}$ over a time bound $z \in [0,15]$ for different values of μ. Clearly, for equally distributed job durations, i.e. if $\mu(i) = \mu(k)$ for all i,k, the upper and lower probability bounds coincide. However, if $\mu(i) \neq \mu(k)$ for some i,k, the probabilities clearly depend on the scheduling policy.

6 Conclusion

We propose a discretization which reduces the problem of computing time bounded reachability probabilities in locally uniform CTMDPs to the problem of step bounded reachability in discrete-time MDPs. Future work includes the extension of our results towards reward reachability as well as the open problem of computing timed reachability probabilities for early schedulers in CTMDPs that are not locally uniform.

Acknowledgement The authors thank Holger Hermanns and Joost-Pieter Katoen for their comments and for many fruitful discussions.

References

1. Ash, R. B., Doléans-Dade, C. A.: *Probability & Measure Theory.* 2nd edn. Academic Press (2000)
2. Baier, C., Haverkort, B. R., Hermanns, H., Katoen, J.-P.: Model-checking algorithms for continuous-time Markov chains. *IEEE TSE* **29** (2003) 524–541

3. Baier, C., Hermanns, H., Katoen, J.-P., Haverkort, B. R.: Efficient computation of time-bounded reachability probabilities in uniform continuous-time Markov decision processes. *Theor. Comp. Sci.* **345** (2005) 2–26
4. Bertsekas, D. P.: *Dynamic Programming and Optimal Control*. Vol. II. Athena Scientific (1995)
5. Bruno, J. L., Downey, P. J., Frederickson, G. N.: Sequencing tasks with exponential service times to minimize the expected flow time or makespan. *Journal of the ACM* **28** (1981) 100–113
6. Chiola, G., Marsan, M. A., Balbo, G., Conte, G.: Generalized stochastic Petri nets: A definition at the net level and its implications. *IEEE TSE* **19** (1993) 89–107
7. de Alfaro, L.: *Formal Verification of Probabilistic Systems*. PhD thesis, Stanford University (1997)
8. Guo, X., Hernández-Lerma, O., Prieto-Rumeau, T.: A survey of recent results on continuous-time Markov decision processes. *TOP* **14** (2006) 177–261
9. Hermanns, H.: *Interactive Markov Chains: The Quest for Quantified Quality*. *LNCS*, Vol. 2428. Springer (2002)
10. Howard, R. A.: *Dynamic Probabilistic Systems*. John Wiley and Sons (1971)
11. Kulkarni, V. G.: *Modeling and Analysis of Stochastic Systems*. Chapman & Hall (1995)
12. Molloy, M. K.: Performance analysis using stochastic Petri nets. *IEEE Transactions on Computers* **31** (1982) 913–917
13. Neuhäußer, M. R., Stoelinga, M., Katoen, J.-P.: Delayed nondeterminism in continuous-time Markov decision processes. In: *FOSSACS*. *LNCS*, Vol. 5504. Springer (2009) 364–379
14. Puterman, M. L.: *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley and Sons (1994)
15. Sanders, W. H., Meyer, J. F.: Stochastic activity networks: Formal definitions and concepts. In: *FMPA*. *LNCS*, Vol. 2090. Springer (2000) 315–343
16. Wolovick, N., Johr, S.: A characterization of meaningful schedulers for continuous-time Markov decision processes. In: *FORMATS*. *LNCS*, Vol. 4202. Springer (2006) 352–367

Aachener Informatik-Berichte

This list contains all technical reports published during the past five years. A complete list of reports dating back to 1987 is available from `http://aib.informatik.rwth-aachen.de/`. To obtain copies consult the above URL or send your request to: Informatik-Bibliothek, RWTH Aachen, Ahornstr. 55, 52056 Aachen, Email: `biblio@informatik.rwth-aachen.de`

2004-01 * Fachgruppe Informatik: Jahresbericht 2003
2004-02 Benedikt Bollig, Martin Leucker: Message-Passing Automata are expressively equivalent to EMSO logic
2004-03 Delia Kesner, Femke van Raamsdonk, Joe Wells (eds.): HOR 2004 – 2nd International Workshop on Higher-Order Rewriting
2004-04 Slim Abdennadher, Christophe Ringeissen (eds.): RULE 04 – Fifth International Workshop on Rule-Based Programming
2004-05 Herbert Kuchen (ed.): WFLP 04 – 13th International Workshop on Functional and (Constraint) Logic Programming
2004-06 Sergio Antoy, Yoshihito Toyama (eds.): WRS 04 – 4th International Workshop on Reduction Strategies in Rewriting and Programming
2004-07 Michael Codish, Aart Middeldorp (eds.): WST 04 – 7th International Workshop on Termination
2004-08 Klaus Indermark, Thomas Noll: Algebraic Correctness Proofs for Compiling Recursive Function Definitions with Strictness Information
2004-09 Joachim Kneis, Daniel Mölle, Stefan Richter, Peter Rossmanith: Parameterized Power Domination Complexity
2004-10 Zinaida Benenson, Felix C. Gärtner, Dogan Kesdogan: Secure Multi-Party Computation with Security Modules
2005-01 * Fachgruppe Informatik: Jahresbericht 2004
2005-02 Maximillian Dornseif, Felix C. Gärtner, Thorsten Holz, Martin Mink: An Offensive Approach to Teaching Information Security: "Aachen Summer School Applied IT Security"
2005-03 Jürgen Giesl, René Thiemann, Peter Schneider-Kamp: Proving and Disproving Termination of Higher-Order Functions
2005-04 Daniel Mölle, Stefan Richter, Peter Rossmanith: A Faster Algorithm for the Steiner Tree Problem
2005-05 Fabien Pouget, Thorsten Holz: A Pointillist Approach for Comparing Honeypots
2005-06 Simon Fischer, Berthold Vöcking: Adaptive Routing with Stale Information
2005-07 Felix C. Freiling, Thorsten Holz, Georg Wicherski: Botnet Tracking: Exploring a Root-Cause Methodology to Prevent Distributed Denial-of-Service Attacks
2005-08 Joachim Kneis, Peter Rossmanith: A New Satisfiability Algorithm With Applications To Max-Cut
2005-09 Klaus Kursawe, Felix C. Freiling: Byzantine Fault Tolerance on General Hybrid Adversary Structures
2005-10 Benedikt Bollig: Automata and Logics for Message Sequence Charts
2005-11 Simon Fischer, Berthold Vöcking: A Counterexample to the Fully Mixed Nash Equilibrium Conjecture

2005-12 Neeraj Mittal, Felix Freiling, S. Venkatesan, Lucia Draque Penso: Efficient Reductions for Wait-Free Termination Detection in Faulty Distributed Systems

2005-13 Carole Delporte-Gallet, Hugues Fauconnier, Felix C. Freiling: Revisiting Failure Detection and Consensus in Omission Failure Environments

2005-14 Felix C. Freiling, Sukumar Ghosh: Code Stabilization

2005-15 Uwe Naumann: The Complexity of Derivative Computation

2005-16 Uwe Naumann: Syntax-Directed Derivative Code (Part I: Tangent-Linear Code)

2005-17 Uwe Naumann: Syntax-directed Derivative Code (Part II: Intraprocedural Adjoint Code)

2005-18 Thomas von der Maßen, Klaus Müller, John MacGregor, Eva Geisberger, Jörg Dörr, Frank Houdek, Harbhajan Singh, Holger Wußmann, Hans-Veit Bacher, Barbara Paech: Einsatz von Features im Software-Entwicklungsprozess - Abschlußbericht des GI-Arbeitskreises "Features"

2005-19 Uwe Naumann, Andre Vehreschild: Tangent-Linear Code by Augmented LL-Parsers

2005-20 Felix C. Freiling, Martin Mink: Bericht über den Workshop zur Ausbildung im Bereich IT-Sicherheit Hochschulausbildung, berufliche Weiterbildung, Zertifizierung von Ausbildungsangeboten am 11. und 12. August 2005 in Köln organisiert von RWTH Aachen in Kooperation mit BITKOM, BSI, DLR und Gesellschaft fuer Informatik (GI) e.V.

2005-21 Thomas Noll, Stefan Rieger: Optimization of Straight-Line Code Revisited

2005-22 Felix Freiling, Maurice Herlihy, Lucia Draque Penso: Optimal Randomized Fair Exchange with Secret Shared Coins

2005-23 Heiner Ackermann, Alantha Newman, Heiko Röglin, Berthold Vöcking: Decision Making Based on Approximate and Smoothed Pareto Curves

2005-24 Alexander Becher, Zinaida Benenson, Maximillian Dornseif: Tampering with Motes: Real-World Physical Attacks on Wireless Sensor Networks

2006-01 * Fachgruppe Informatik: Jahresbericht 2005

2006-02 Michael Weber: Parallel Algorithms for Verification of Large Systems

2006-03 Michael Maier, Uwe Naumann: Intraprocedural Adjoint Code Generated by the Differentiation-Enabled NAGWare Fortran Compiler

2006-04 Ebadollah Varnik, Uwe Naumann, Andrew Lyons: Toward Low Static Memory Jacobian Accumulation

2006-05 Uwe Naumann, Jean Utke, Patrick Heimbach, Chris Hill, Derya Ozyurt, Carl Wunsch, Mike Fagan, Nathan Tallent, Michelle Strout: Adjoint Code by Source Transformation with OpenAD/F

2006-06 Joachim Kneis, Daniel Mölle, Stefan Richter, Peter Rossmanith: Divide-and-Color

2006-07 Thomas Colcombet, Christof Löding: Transforming structures by set interpretations

2006-08 Uwe Naumann, Yuxiao Hu: Optimal Vertex Elimination in Single-Expression-Use Graphs

2006-09 Tingting Han, Joost-Pieter Katoen: Counterexamples in Probabilistic Model Checking

- 2006-10 Mesut Günes, Alexander Zimmermann, Martin Wenig, Jan Ritzerfeld, Ulrich Meis: From Simulations to Testbeds - Architecture of the Hybrid MCG-Mesh Testbed
- 2006-11 Bastian Schlich, Michael Rohrbach, Michael Weber, Stefan Kowalewski: Model Checking Software for Microcontrollers
- 2006-12 Benedikt Bollig, Joost-Pieter Katoen, Carsten Kern, Martin Leucker: Replaying Play in and Play out: Synthesis of Design Models from Scenarios by Learning
- 2006-13 Wong Karianto, Christof Löding: Unranked Tree Automata with Sibling Equalities and Disequalities
- 2006-14 Danilo Beuche, Andreas Birk, Heinrich Dreier, Andreas Fleischmann, Heidi Galle, Gerald Heller, Dirk Janzen, Isabel John, Ramin Tavakoli Kolagari, Thomas von der Maßen, Andreas Wolfram: Report of the GI Work Group "Requirements Management Tools for Product Line Engineering"
- 2006-15 Sebastian Ullrich, Jakob T. Valvoda, Torsten Kuhlen: Utilizing optical sensors from mice for new input devices
- 2006-16 Rafael Ballagas, Jan Borchers: Selexels: a Conceptual Framework for Pointing Devices with Low Expressiveness
- 2006-17 Eric Lee, Henning Kiel, Jan Borchers: Scrolling Through Time: Improving Interfaces for Searching and Navigating Continuous Audio Timelines
- 2007-01 * Fachgruppe Informatik: Jahresbericht 2006
- 2007-02 Carsten Fuhs, Jürgen Giesl, Aart Middeldorp, Peter Schneider-Kamp, René Thiemann, and Harald Zankl: SAT Solving for Termination Analysis with Polynomial Interpretations
- 2007-03 Jürgen Giesl, René Thiemann, Stephan Swiderski, and Peter Schneider-Kamp: Proving Termination by Bounded Increase
- 2007-04 Jan Buchholz, Eric Lee, Jonathan Klein, and Jan Borchers: coJIVE: A System to Support Collaborative Jazz Improvisation
- 2007-05 Uwe Naumann: On Optimal DAG Reversal
- 2007-06 Joost-Pieter Katoen, Thomas Noll, and Stefan Rieger: Verifying Concurrent List-Manipulating Programs by LTL Model Checking
- 2007-07 Alexander Nyßen, Horst Lichter: MeDUSA - MethoD for UML2-based Design of Embedded Software Applications
- 2007-08 Falk Salewski and Stefan Kowalewski: Achieving Highly Reliable Embedded Software: An empirical evaluation of different approaches
- 2007-09 Tina Kraußer, Heiko Mantel, and Henning Sudbrock: A Probabilistic Justification of the Combining Calculus under the Uniform Scheduler Assumption
- 2007-10 Martin Neuhäußer, Joost-Pieter Katoen: Bisimulation and Logical Preservation for Continuous-Time Markov Decision Processes
- 2007-11 Klaus Wehrle (editor): 6. Fachgespräch Sensornetzwerke
- 2007-12 Uwe Naumann: An L-Attributed Grammar for Adjoint Code
- 2007-13 Uwe Naumann, Michael Maier, Jan Riehme, and Bruce Christianson: Second-Order Adjoints by Source Code Manipulation of Numerical Programs

2007-14 Jean Utke, Uwe Naumann, Mike Fagan, Nathan Tallent, Michelle Strout, Patrick Heimbach, Chris Hill, and Carl Wunsch: OpenAD/F: A Modular, Open-Source Tool for Automatic Differentiation of Fortran Codes

2007-15 Volker Stolz: Temporal assertions for sequential and concurrent programs

2007-16 Sadeq Ali Makram, Mesut Güneç, Martin Wenig, Alexander Zimmermann: Adaptive Channel Assignment to Support QoS and Load Balancing for Wireless Mesh Networks

2007-17 René Thiemann: The DP Framework for Proving Termination of Term Rewriting

2007-18 Uwe Naumann: Call Tree Reversal is NP-Complete

2007-19 Jan Riehme, Andrea Walther, Jörg Stiller, Uwe Naumann: Adjoints for Time-Dependent Optimal Control

2007-20 Joost-Pieter Katoen, Daniel Klink, Martin Leucker, and Verena Wolf: Three-Valued Abstraction for Probabilistic Systems

2007-21 Tingting Han, Joost-Pieter Katoen, and Alexandru Mereacre: Compositional Modeling and Minimization of Time-Inhomogeneous Markov Chains

2007-22 Heiner Ackermann, Paul W. Goldberg, Vahab S. Mirrokni, Heiko Röglin, and Berthold Vöcking: Uncoordinated Two-Sided Markets

2008-01 * Fachgruppe Informatik: Jahresbericht 2007

2008-02 Henrik Bohnenkamp, Marielle Stoelinga: Quantitative Testing

2008-03 Carsten Fuhs, Jürgen Giesl, Aart Middeldorp, Peter Schneider-Kamp, René Thiemann, Harald Zankl: Maximal Termination

2008-04 Uwe Naumann, Jan Riehme: Sensitivity Analysis in Sisyphe with the AD-Enabled NAGWare Fortran Compiler

2008-05 Frank G. Radmacher: An Automata Theoretic Approach to the Theory of Rational Tree Relations

2008-06 Uwe Naumann, Laurent Hascoet, Chris Hill, Paul Hovland, Jan Riehme, Jean Utke: A Framework for Proving Correctness of Adjoint Message Passing Programs

2008-07 Alexander Nyßen, Horst Lichter: The MeDUSA Reference Manual, Second Edition

2008-08 George B. Mertzios, Stavros D. Nikolopoulos: The λ-cluster Problem on Parameterized Interval Graphs

2008-09 George B. Mertzios, Walter Unger: An optimal algorithm for the k-fixed-endpoint path cover on proper interval graphs

2008-10 George B. Mertzios, Walter Unger: Preemptive Scheduling of Equal-Length Jobs in Polynomial Time

2008-11 George B. Mertzios: Fast Convergence of Routing Games with Splittable Flows

2008-12 Joost-Pieter Katoen, Daniel Klink, Martin Leucker, Verena Wolf: Abstraction for stochastic systems by Erlang's method of stages

2008-13 Beatriz Alarcón, Fabian Emmes, Carsten Fuhs, Jürgen Giesl, Raúl Gutiérrez, Salvador Lucas, Peter Schneider-Kamp, René Thiemann: Improving Context-Sensitive Dependency Pairs

2008-14 Bastian Schlich: Model Checking of Software for Microcontrollers

2008-15 Joachim Kneis, Alexander Langer, Peter Rossmanith: A New Algorithm for Finding Trees with Many Leaves

2008-16 Hendrik vom Lehn, Elias Weingärtner and Klaus Wehrle: Comparing recent network simulators: A performance evaluation study

2008-17 Peter Schneider-Kamp: Static Termination Analysis for Prolog using Term Rewriting and SAT Solving

2008-18 Falk Salewski: Empirical Evaluations of Safety-Critical Embedded Systems

2009-03 Alexander Nyßen: Model-Based Construction of Embedded Real-Time Software - A Methodology for Small Devices

2009-05 George B. Mertzios, Ignasi Sau, Shmuel Zaks: A New Intersection Model and Improved Algorithms for Tolerance Graphs

2009-06 George B. Mertzios, Ignasi Sau, Shmuel Zaks: The Recognition of Tolerance and Bounded Tolerance Graphs is NP-complete

2009-07 Joachim Kneis, Alexander Langer, Peter Rossmanith: Derandomizing Non-uniform Color-Coding I

2009-08 Joachim Kneis, Alexander Langer: Satellites and Mirrors for Solving Independent Set on Sparse Graphs

2009-11 Kyriaki Ioannidou, George B. Mertzios, Stavros D. Nikolopoulos: The Longest Path Problem is Polynomial on Interval Graphs

* These reports are only available as a printed version.

Please contact `biblio@informatik.rwth-aachen.de` to obtain copies.