Delayed Nondeterminism in Continuous-Time Markov Decision Processes*

Martin R. Neuhäußer[1,2], Mariëlle Stoelinga[2], and Joost-Pieter Katoen[1,2]

[1] MOVES Group, RWTH Aachen University, Germany
[2] FMT Group, University of Twente, The Netherlands

Abstract. Schedulers in randomly timed games can be classified as to whether they use timing information or not. We consider continuous-time Markov decision processes (CTMDPs) and define a hierarchy of positional (P) and history-dependent (H) schedulers which induce strictly tighter bounds on quantitative properties on CTMDPs. This classification into time abstract (TA), total time (TT) and fully time-dependent (T) schedulers is mainly based on the kind of timing details that the schedulers may exploit. We investigate when the resolution of nondeterminism may be deferred. In particular, we show that TTP and TAP schedulers allow for delaying nondeterminism for all measures, whereas this does neither hold for TP nor for any TAH scheduler. The core of our study is a transformation on CTMDPs which unifies the speed of outgoing transitions per state.

1 Introduction

Continuous-time Markov decision processes (CTMDPs) which are also known as controlled Markov chains, have originated as continuous-time variants of finite-state probabilistic automata [1], and have been used for, among others, the control of queueing systems, epidemic, and manufacturing processes. The analysis of CTMDPs is mainly focused on determining optimal schedulers for criteria such as expected total reward and expected (long-run) average reward, cf. the survey [2].

As in discrete-time MDPs, nondeterminism in CTMDPs is resolved by schedulers. An important criterion for CTMDP schedulers is whether they use timing information or not. For time-bounded reachability objectives, e.g., timed schedulers are optimal [3]. For simpler criteria such as unbounded reachability or average reward, time-abstract (TA) schedulers will do. For such objectives, it suffices to either abstract the timing information in the CTMDP (yielding an "embedded" MDP) or to transform the CTMDP into an equivalent discrete-time MDP, see e.g., [4, p. 562] [2]. The latter process is commonly referred to as uniformization. Its equivalent on continuous-time Markov chains, a proper subclass of CTMDPs, is pivotal to probabilistic model checking [5].

The main focus of this paper is on defining a hierarchy of positional (P) and history-dependent (H) schedulers which induce strictly tighter bounds on quantitative properties on CTMDPs. This hierarchy refines the notion of generic measurable schedulers [6]. An important distinguishing criterion is the level of detail of timing information the

* Supported by the NWO projects QUPES (612.000.420) and FOCUS/BRICKS (642.000.505) (MOQS) and by the EU grants IST-004527 (ARTIST2) and FP7-ICT-2007-1 (QUASIMODO).

L. de Alfaro (Ed.): FOSSACS 2009, LNCS 5504, pp. 364–379, 2009.

schedulers may exploit, e.g., the delay in the last state, total time (TT), or all individual state residence times (T).

In general, the delay to jump to a next state in a CTMDP is determined by the action selected by the scheduler on entering the current state. We investigate under which conditions this resolution of nondeterminism may be deferred. Rather than focusing on a specific objective, we consider this delayed nondeterminism for generic (measurable) properties. The core of our study is a transformation —called local uniformization— on CTMDPs which unifies the speed of outgoing transitions per state. Whereas classical uniformization [7,8,9] adds self-loops to achieve this, local uniformization uses auxiliary copy-states. In this way, we enforce that schedulers in the original and uniformized CTMDP have (for important scheduler classes) the same power, whereas classical loop-based uniformization allows a scheduler to change its decision when re-entering a state through the added self-loop. Therefore, locally uniform CTMDPs allow to defer the resolution of nondeterminism, i.e., they dissolve the intrinsic dependency between state residence times and schedulers, and can be viewed as MDPs with exponentially distributed state residence times.

In particular, we show that TTP and TAP schedulers allow to delay nondeterminism for all measures. As TTP schedulers are optimal for time-bounded reachability objectives, this shows that local uniformization preserves the probability of such objectives. Finally, we prove that TP and TAH schedulers do not allow for delaying nondeterminism. This results in a hierarchy of time-dependent schedulers and their inclusions. Moreover, we solve an open problem in [3] concerning TAP schedulers.

The paper is organized as follows: Sec. 2 introduces CTMDPs and a general notion of schedulers which is refined in Sec. 3. In Sec. 4, we define local uniformization and prove its correctness. Sec. 5 summarizes the main results and Sec. 6 proves that deferring nondeterministic choices induces strictly tighter bounds on quantitative properties.

2 Continuous-Time Markov Decision Processes

We consider CTMDPs with finite sets $\mathcal{S}=\{s_0,s_1,\dots\}$ and $Act=\{\alpha,\beta,\dots\}$ of states and actions; $Distr(\mathcal{S})$ and $Distr(Act)$ are the respective sets of probability distributions.

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.

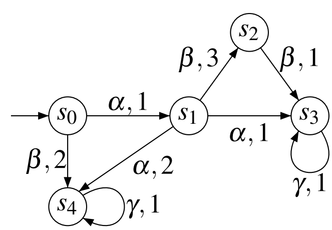


Fig. 1. A well-formed CTMDP

If $\mathbf{R}(s,\alpha,s') = \lambda$ and $\lambda > 0$, an α-transition leads from state s to state s'. λ is the *rate* of an exponential distribution which defines the transition's delay. Hence, it executes in time interval $[a,b]$ with probability $\eta_\lambda([a,b]) = \int_a^b \lambda e^{-\lambda t}\,dt$; note that η_λ directly extends to the Borel σ-field $\mathfrak{B}(\mathbb{R}_{\geq 0})$. Further, $Act(s) = \{\alpha \in Act \mid \exists s' \in \mathcal{S}.\ \mathbf{R}(s,\alpha,s') > 0\}$ is the set of *enabled* actions in state s and

$E(s,\alpha) = \sum_{s' \in \mathscr{S}} \mathbf{R}(s,\alpha,s')$ is its *exit rate* under action α. A CTMDP is *well-formed* if $Act(s) \neq \emptyset$ for all $s \in \mathscr{S}$. As this is easily achieved by adding self-loops, we restrict to well-formed CTMDPs. The time-abstract *branching probabilities* are captured by a matrix $\mathbf{P}$ where $\mathbf{P}(s,\alpha,s') = \frac{\mathbf{R}(s,\alpha,s')}{E(s,\alpha)}$ if $E(s,\alpha) > 0$ and $\mathbf{P}(s,\alpha,s') = 0$ otherwise.

Example 1. If α is chosen in state s_0 of the CTMDP in Fig. 1, we enter state s_1 after a delay which is exponentially distributed with rate $\mathbf{R}(s_0,\alpha,s_1) = E(s_0,\alpha) = 1$. For state s_1 and action α, a *race* decides which of the two α-transitions executes; in this case, the *sojourn time* of state s_1 is exponentially distributed with rate $E(s_1,\alpha) = 3$. The time-abstract probability to move to state s_3 is $\frac{\mathbf{R}(s_1,\alpha,s_3)}{E(s_1,\alpha)} = \frac{1}{3}$.

2.1 The Probability Space

In a CTMDP $\mathscr{C} = (\mathscr{S}, Act, \mathbf{R}, \nu)$, a *finite path* π of length n (denoted $|\pi| = n$) is a sequence $\pi = s_0 \xrightarrow{\alpha_0,t_0} s_1 \xrightarrow{\alpha_1,t_1} \cdots \xrightarrow{\alpha_{n-1},t_{n-1}} s_n$ where $s_i \in \mathscr{S}$, $\alpha_i \in Act$ and $t_i \in \mathbb{R}_{\geq 0}$. With $\pi[k] = s_k$ and $\delta(\pi,k) = t_k$ we refer to its k-th state and the associated sojourn time. Accordingly, $\Delta(\pi) = \sum_{k=0}^{n-1} t_k$ is the total time spent on π. Finally, $\pi\downarrow = s_n$ denotes the last state of π and $\pi[i..k]$ is the path infix $s_i \xrightarrow{\alpha_i,t_i} \cdots \xrightarrow{\alpha_{k-1},t_{k-1}} s_k$. The path π is built by a state and a sequence of *combined transitions* from the set $\Omega = Act \times \mathbb{R}_{\geq 0} \times \mathscr{S}$: It is the concatenation $s_0 \circ m_0 \circ m_1 \cdots \circ m_{n-1}$ where $m_i = (\alpha_i, t_i, s_{i+1}) \in \Omega$. Thus $Paths^n(\mathscr{C}) = \mathscr{S} \times \Omega^n$ yields the set of paths of length n in $\mathscr{C}$ and analogously, $Paths^\star(\mathscr{C})$, $Paths^\omega(\mathscr{C})$ and $Paths(\mathscr{C})$ denote the sets of finite, infinite and all paths of $\mathscr{C}$. We use $abs(\pi) = s_0 \xrightarrow{\alpha_0} s_1 \xrightarrow{\alpha_2} \cdots \xrightarrow{\alpha_{n-1}} s_n$ to refer to the *time-abstract* path induced by π and define $Paths^n_{abs}(\mathscr{C})$ accordingly. For simplicity, we omit the reference to $\mathscr{C}$ wherever possible.

Events in $\mathscr{C}$ are measurable sets of paths; as paths are sequences of combined transitions, we first define the σ-field $\mathfrak{F} = \sigma\big(\mathfrak{F}_{Act} \times \mathfrak{B}(\mathbb{R}_{\geq 0}) \times \mathfrak{F}_{\mathscr{S}}\big)$ on subsets of Ω where $\mathfrak{F}_{\mathscr{S}} = 2^{\mathscr{S}}$ and $\mathfrak{F}_{Act} = 2^{Act}$. Based on $(\Omega, \mathfrak{F})$, we derive the product σ-field $\mathfrak{F}_{Paths^n} = \sigma\big(\{S_0 \times M_0 \times \cdots \times M_{n-1} \mid S_0 \in \mathfrak{F}_{\mathscr{S}}, M_i \in \mathfrak{F}\}\big)$ for paths of length n. Finally, the cylinder-set construction [10] allows to extend this to a σ-field over infinite paths: A set $B \in \mathfrak{F}_{Paths^n}$ is a *base* of the 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. $\mathfrak{F}_{Paths^\omega} = \sigma\big(\bigcup_{n=0}^\infty \{Cyl(B) \mid B \in \mathfrak{F}_{Paths^n}\}\big)$. For an in-depth discussion, we refer to [10,11,6].

2.2 Probability Measure

The probability measures on $\mathfrak{F}_{Paths^n}$ and $\mathfrak{F}_{Paths^\omega}$ are defined using schedulers that resolve the nondeterminism in the underlying CTMDP.

Definition 2 (Generic measurable scheduler). *Let $\mathscr{C}$ be a CTMDP with actions in Act. A* generic scheduler *on $\mathscr{C}$ is a mapping $D : Paths^\star \times \mathfrak{F}_{Act} \to [0,1]$ where $D(\pi,\cdot) \in Distr(Act(\pi\downarrow))$. It is* measurable *(gm-scheduler) iff the functions $D(\cdot,A) : Paths^\star \to [0,1]$ are measurable for all $A \in \mathfrak{F}_{Act}$.*

On reaching state s_n via path π, $D(\pi,\cdot)$ defines a distribution over $Act(s_n)$ and thereby resolves the nondeterminism in state s_n. The measurability condition in Def. 2 states that $\{\pi \in Paths^\star | D(\pi,A) \in B\} \in \mathfrak{F}_{Paths^\star}$ for all $A \in \mathfrak{F}_{Act}$ and $B \in \mathfrak{B}([0,1])$; it is required for the Lebesgue-integral in Def. 4 to be well-defined.

To define a probability measure on sets of paths, we proceed stepwise and first derive a probability measure on sets of combined transitions:

Definition 3 (Probability on combined transitions). *Let $\mathscr{C}=(\mathscr{S},Act,\mathbf{R},\nu)$ be a CTMDP and D a gm-scheduler on $\mathscr{C}$. For all $\pi \in Paths^\star(\mathscr{C})$, define the probability measure $\mu_D(\pi,\cdot) : \mathfrak{F} \to [0,1]$ where*

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

Here, $\mathbf{I}_M$ denotes the characteristic function of $M \in \mathfrak{F}$. A proof that $\mu_D(\pi,\cdot)$ is indeed a probability measure can be found in [6, Lemma 1]. Intuitively, $\mu_D(\pi,M)$ is the probability to continue on path π under scheduler D with a combined transition in M. With $\mu_D(\pi,\cdot)$ and ν, we can define the probability of sets of paths:

Definition 4 (Probability measure). *Let $\mathscr{C} = (\mathscr{S},Act,\mathbf{R},\nu)$ be a CTMDP and D a gm-scheduler on $\mathscr{C}$. For $n \geq 0$, we define the probability measures $Pr^n_{\nu,D}$ on the measurable space $(Paths^n,\mathfrak{F}_{Paths^n})$ inductively as follows:*

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

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

Intuitively, we measure sets of paths Π of length n by multiplying the probability $Pr^{n-1}_{\nu,D}(d\pi)$ of path prefixes π with the probability $\mu_D(\pi,dm)$ of a combined transition m that 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 [10] is applicable and yields a unique extension of $Pr^\omega_{\nu,D}$ from cylinders to arbitrary sets in $\mathfrak{F}_{Paths^\omega}$.

As we later need to split a set of paths into a set of prefixes I and a set of suffixes Π, we define the set of path prefixes of length $k > 0$ by $PPref^k = (\mathfrak{F}_{\mathscr{S}} \times \mathfrak{F}_{Act} \times \mathfrak{B}(\mathbb{R}_{\geq 0}))^k$ and provide a probability measure on its σ-field $\mathfrak{F}_{PPref^k}$:

Definition 5 (Prefix measure). *Let $\mathscr{C} = (\mathscr{S},Act,\mathbf{R},\nu)$ be a CTMDP and D a gm-scheduler on $\mathscr{C}$. For $I \in \mathfrak{F}_{PPref^k}$ and $k > 0$, define*

$$\mu^k_{\nu,D}(I) = \int_{Paths^{k-1}} Pr^{k-1}_{\nu,D}(d\pi) \int_{Act} D(\pi,d\alpha) \int_{\mathbb{R}_{\geq 0}} \mathbf{I}_I(\pi \xrightarrow{\alpha,t}) \; \eta_{E(\pi\downarrow,\alpha)}(dt).$$

As $Pr^{k-1}_{\nu,D}$ is a probability measure, so is $\mu^k_{\nu,D}$. If $I \in \mathfrak{F}_{PPref^k}$ and $\Pi \in \mathfrak{F}_{Paths^n}$, their concatenation is the set $I \times \Pi \in \mathfrak{F}_{Paths^{k+n}}$; its probability $Pr^{k+n}_{\nu,D}(I \times \Pi)$ is obtained by multiplying the measure of prefixes $i \in I$ with the suffixes in Π:

Lemma 1. *Let $\Pi \in \mathfrak{F}_{Paths^n}$ and $I \in \mathfrak{F}_{PPref^k}$. If $i = s_0 \xrightarrow{\alpha_0,t_0} \cdots s_{k-1} \xrightarrow{\alpha_{k-1},t_{k-1}}$, define $\nu_i = \mathbf{P}(s_{k-1},\alpha_{k-1},\cdot)$ and $D_i(\pi,\cdot) = D(i \circ \pi,\cdot)$. Then*

$$Pr^{k+n}_{\nu,D}(I \times \Pi) = \int_{PPref^k} \mu^k_{\nu,D}(di) \int_{Paths^n} \mathbf{I}_{I\times\Pi}(i \circ \pi) \;\; Pr^n_{\nu_i,D_i}(d\pi).$$

Lemma 1 justifies to split sets of paths and to measure the components of the resulting Cartesian product; therefore, it abstracts from the inductive definition of $Pr^n_{\nu,D}$.

3 Scheduler Classes

Section 2.2 defines the probability of sets of paths w.r.t. a gm-scheduler. However, this does not fully describe a CTMDP, as a single scheduler is only one way to resolve nondeterminism. Therefore we define scheduler classes according to the information that is available when making a decision. Given an event $\Pi \in \mathfrak{F}_{Paths^\omega}$, a scheduler class induces a set of probabilities which reflects the CTMDP's possible behaviours. In this paper, we investigate which classes in Fig. 2 preserve minimum and maximum probabilities if nondeterministic choices are delayed.

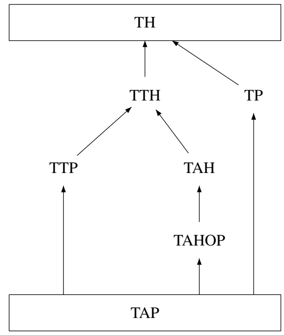


Fig. 2. Scheduler classes

As proved in [6], the most general class is the set of all gm-schedulers: If paths $\pi_1, \pi_2 \in Paths^\star$ of a CTMDP end in state s, a gm-scheduler $D : Paths^\star \times \mathfrak{F}_{Act} \to [0,1]$ may yield different distributions $D(\pi_1, \cdot)$ and $D(\pi_2, \cdot)$ over the next action, depending on the entire histories π_1 and π_2. We call this the class of *timed, history dependent (TH)* schedulers.

On the contrary, D is a *time-abstract positional* (TAP) scheduler, if $D(\pi_1, \cdot) = D(\pi_2, \cdot)$ for all $\pi_1, \pi_2 \in Paths^\star$ that end in the same state. As $D(\pi, \cdot)$ only depends on the current state, it is specified by a mapping $D : \mathscr{S} \to Distr(Act)$.

Example 2. For TAP scheduler D with $D(s_0) = \{\alpha \mapsto 1\}$ and $D(s_1) = \{\beta \mapsto 1\}$, the induced stochastic process of the CTMDP in Fig. 1 is the CTMC depicted in Fig. 3. Note that in general, randomized schedulers do not yield CTMCs as the induced sojourn times are hyper-exponentially distributed.

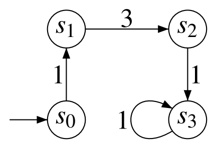


Fig. 3. Induced CTMC

For TAHOP schedulers, the decision may depend on the current state s and the length of π_1 and π_2 (*hop-counting schedulers*); accordingly, they are isomorphic to mappings $D : \mathscr{S} \times \mathbb{N} \to Distr(Act)$. Moreover, D is a *time-abstract history-dependent* scheduler (TAH), if $D(\pi_1, \cdot) = D(\pi_2, \cdot)$ for all histories $\pi_1, \pi_2 \in Paths^\star$ with $abs(\pi_1) = abs(\pi_2)$; given history π, TAH schedulers may decide based on the sequence of states and actions in $abs(\pi)$. In [3], the authors show that TAHOP and TAH induce the same probability bounds for timed reachability which are tighter than the bounds induced by TAP.

Time-dependent scheduler classes generally induce probability bounds that exceed those of the corresponding time-abstract classes [3]: If we move from state s to s', a *timed positional* scheduler (TP) yields a distribution over $Act(s')$ which depends on s' and the time to go from s to s'; thus TP extends TAP with information on the delay of the last transition.

Similarly, *total time history-dependent* schedulers (TTH) extend TAH with information on the time that passed up to the current state: If $D \in TTH$ and $\pi_1, \pi_2 \in Paths^\star$ are histories with $abs(\pi_1) = abs(\pi_2)$ and $\Delta(\pi_1) = \Delta(\pi_2)$, then $D(\pi_1, \cdot) = D(\pi_2, \cdot)$. Note that $TTH \subseteq TH$, as TTH schedulers may depend on the accumulated time but not on sojourn times in individual states of the history. Generally the probability bounds of TTH are less strict than those of TH.

Table 1. Proposed scheduler classes for CTMDPs

	scheduler class	scheduler signature
time abstract	positional (TAP)	$D : \mathscr{S} \to Distr(Act)$
	hop-counting (TAHOP)	$D : \mathscr{S} \times \mathbb{N} \to Distr(Act)$
	time abstract history dependent (TAH)	$D : Paths^{\star}_{abs} \to Distr(Act)$
time dependent	timed history dependent (TH)	full timed history $D : Paths^{\star} \to Distr(Act)$
	total time history dependent (TTH)	sequence of states & total time $D : Paths^{\star}_{abs} \times \mathbb{R}_{\geq 0} \to Distr(Act)$
	total time positional (TTP)	last state & total time $D : \mathscr{S} \times \mathbb{R}_{\geq 0} \to Distr(Act)$
	timed positional (TP)	last state & delay of last transition $D : \mathscr{S} \times \mathbb{R}_{\geq 0} \to Distr(Act)$

In this paper, we focus on *total time positional* schedulers (TTP) which are given by mappings $D : \mathscr{S} \times \mathbb{R}_{\geq 0} \to Distr(Act)$. They are similar to TTH schedulers but abstract from the state-history. For $\pi_1, \pi_2 \in Paths^{\star}$, $D(\pi_1, \cdot) = D(\pi_2, \cdot)$ if π_1 and π_2 end in the same state and have the same simulated time $\Delta(\pi_1) = \Delta(\pi_2)$. TTP schedulers are of particular interest, as they induce optimal bounds w.r.t. timed reachability: To see this, consider the probability to reach a set of goal states $G \subseteq \mathscr{S}$ within t time units. If state s is reached via $\pi \in Paths^{\star}$ (without visiting G), the maximal probability to enter G is given by a scheduler which maximizes the probability to reach G from state s within the remaining $t - \Delta(\pi)$ time units. Obviously, a TTP scheduler is sufficient in this case.

Example 3. For $t \in \mathbb{R}_{\geq 0}$, let the TTP-scheduler D for the CTMDP of Fig. 1 be given by $D(s_0, 0) = \{\alpha \mapsto 1\}$ and $D(s_1, t) = \{\alpha \mapsto 1\}$ if $t \leq 0.64$ and $D(s_1, t) = \{\beta \mapsto 1\}$, otherwise. It turns out that D maximizes the probability to reach s_3 within time t. For now, we only note that the probability induced by D is obtained by the gm-scheduler $D'(\pi) = D(\pi{\downarrow}, \Delta(\pi))$.

Note that we can equivalently specify any gm-scheduler $D : Paths^{\star} \times \mathfrak{F}_{Act} \to [0,1]$ as a mapping $D' : Paths^{\star} \to Distr(Act)$ by setting $D'(\pi)(A) = D(\pi, A)$ for all $\pi \in Paths^{\star}$ and $A \in \mathfrak{F}_{Act}$; to further simplify notation, we also use $D(\pi, \cdot)$ to refer to this distribution.

Definition 6 (Scheduler classes). *Let $\mathscr{C}$ be a CTMDP and D a gm-scheduler on $\mathscr{C}$. For $\pi, \pi' \in Paths^{\star}(\mathscr{C})$, the scheduler classes are defined as follows:*

$$D{\in}TAP \iff \forall \pi, \pi'. \pi{\downarrow} = \pi'{\downarrow} \Rightarrow D(\pi, \cdot) = D(\pi', \cdot)$$
$$D{\in}TAHOP \iff \forall \pi, \pi'. \big(\pi{\downarrow} = \pi'{\downarrow} \wedge |\pi| = |\pi'|\big) \Rightarrow D(\pi, \cdot) = D(\pi', \cdot)$$
$$D{\in}TAH \iff \forall \pi, \pi'. abs(\pi) = abs(\pi') \Rightarrow D(\pi, \cdot) = D(\pi', \cdot)$$
$$D{\in}TTH \iff \forall \pi, \pi'. \big(abs(\pi){=}abs(\pi') \wedge \Delta(\pi){=}\Delta(\pi')\big) \Rightarrow D(\pi, \cdot){=}D(\pi', \cdot)$$
$$D{\in}TTP \iff \forall \pi, \pi'. \big(\pi{\downarrow}{=}\pi'{\downarrow} \wedge \Delta(\pi){=}\Delta(\pi')\big) \Rightarrow D(\pi, \cdot){=}D(\pi', \cdot)$$
$$D{\in}TP \iff \forall \pi, \pi'. \big(\pi{\downarrow}{=}\pi'{\downarrow} \wedge \delta(\pi, |\pi{-}1|){=}\delta(\pi', |\pi'{-}1|)\big) \Rightarrow D(\pi, \cdot){=}D(\pi', \cdot).$$

Def. 6 justifies to restrict the domain of the schedulers to the information the respective class exploits. In this way, we obtain the characterization in Table 1. We now come to the transformation on CTMDPs that unifies the speed of outgoing transitions and thereby allows to defer the resolution of nondeterministic choices.

4 Local Uniformization

Generally, the exit rate of a state depends on the action that is chosen by the scheduler when entering the state. This dependency requires that the scheduler decides directly when entering a state, as otherwise the state's sojourn time distribution is not well-defined. An exception to this are *locally uniform* CTMDPs which allow to delay the scheduler's choice up to the point when the state is left:

Definition 7 (Local uniformity). *A CTMDP* $(\mathcal{S}, Act, \mathbf{R}, \nu)$ *is* locally uniform *iff there exists* $u : \mathcal{S} \to \mathbb{R}_{>0}$ *such that* $E(s,\alpha) = u(s)$ *for all* $s \in \mathcal{S}, \alpha \in Act(s)$.

In locally uniform CTMDPs the exit rates are state-wise constant with rate $u(s)$; hence, they do not depend on the action that is chosen. Therefore locally uniform CTMDPs allow to delay the scheduler's decision until the current state is left. To generalize this idea, we propose a transformation on CTMDPs which attains local uniformity; further, in Sec. 4.2 we investigate as to which scheduler classes local uniformization preserves quantitative properties.

Definition 8 (Local uniformization). *Let* $\mathcal{C} = (\mathcal{S}, Act, \mathbf{R}, \nu)$ *be a CTMDP and define* $u(s) = \max\{E(s,\alpha) \mid \alpha \in Act\}$ *for all* $s \in \mathcal{S}$. *Then* $\overline{\mathcal{C}} = (\overline{\mathcal{S}}, Act, \overline{\mathbf{R}}, \overline{\nu})$ *is the locally uniform CTMDP induced by* $\mathcal{C}$ *where* $\overline{\mathcal{S}} = \mathcal{S} \dot\cup \mathcal{S}_{cp}$, $\mathcal{S}_{cp} = \{s^\alpha \mid E(s,\alpha) < u(s)\}$ *and*

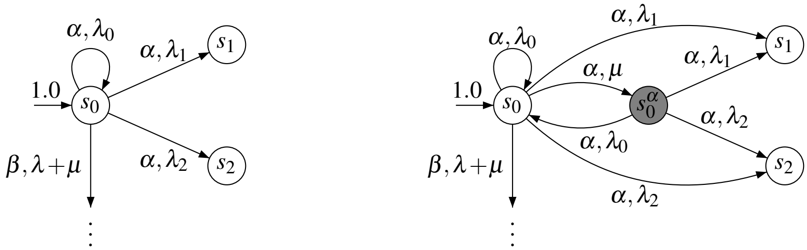
$$\overline{\mathbf{R}}(s,\alpha,s') = \begin{cases} \mathbf{R}(s,\alpha,s') & \text{if } s,s' \in \mathcal{S} \\ \mathbf{R}(t,\alpha,s') & \text{if } s = t^\alpha \wedge s' \in \mathcal{S} \\ u(s) - E(s,\alpha) & \text{if } s \in \mathcal{S} \wedge s' = s^\alpha \\ 0 & \text{otherwise.} \end{cases}$$

Further, $\overline{\nu}(s) = \nu(s)$ *if* $s \in \mathcal{S}$ *and* 0, *otherwise.*

Local uniformization is done for each state s separately with uniformization rate $u(s)$. If the exit rate of s under action α is less than $u(s)$, we introduce a copy-state s^α and an α-transition which carries the missing rate $\mathbf{R}(s,\alpha,s^\alpha) = u(s) - E(s,\alpha)$. Regarding s^α, only the outgoing α-transitions of s carry over to s^α. Hence s^α is deterministic in the sense that $Act(s^\alpha) = \{\alpha\}$.

Example 4. Consider the fragment CTMDP in Fig. 4(a) where $\lambda = \sum \lambda_i$ and $\mu > 0$. It is not locally uniform as $E(s_0,\alpha) = \lambda$ and $E(s_0,\beta) = \lambda + \mu$. Applying our transformation, we obtain the locally uniform CTMDP in Fig. 4(b).

Local uniformization of $\mathcal{C}$ introduces new states and transitions in $\overline{\mathcal{C}}$. The paths in $\overline{\mathcal{C}}$ reflect this and differ from those of $\mathcal{C}$; more precisely, they may contain sequences of transitions $s \xrightarrow{\alpha,t} s^\alpha \xrightarrow{\alpha,t'} s'$ where s^α is a copy-state. Intuitively, if we identify s and s^α, this corresponds to a single transition $s \xrightarrow{\alpha,t+t'} s'$ in $\mathcal{C}$. To formalize this correspondence,



(a) Fragment of a non-uniform CTMDP. (b) Local uniformization of state s_0.

Fig. 4. How to obtain locally uniform CTMDPs by introducing copy states

we derive a mapping *merge* on all paths $\overline{\pi} \in Paths^\star(\overline{\mathscr{C}})$ with $\overline{\pi}[0], \overline{\pi}\downarrow \in \mathscr{S}$: If $|\overline{\pi}| = 0$, $merge(\overline{\pi}) = \overline{\pi}[0]$. Otherwise, let

$$merge\big(s \xrightarrow{\alpha,t} \overline{\pi}\big) = \begin{cases} s \xrightarrow{\alpha,t} merge(\overline{\pi}) & \text{if } \overline{\pi}[0] \in \mathscr{S} \\ merge(s \xrightarrow{\alpha,t+t'} \overline{\pi}') & \text{if } \overline{\pi} = s^\alpha \xrightarrow{\alpha,t'} \overline{\pi}'. \end{cases}$$

Naturally, *merge* extends to infinite paths if we do not require $\overline{\pi}\downarrow \in \mathscr{S}$; further, merging a set of paths $\overline{\Pi}$ is defined element-wise and denoted $merge(\overline{\Pi})$.

Example 5. Let $\overline{\pi} = s_0 \xrightarrow{\alpha_0,t_0} s_0^{\alpha_0} \xrightarrow{\alpha_0,t_0'} s_1 \xrightarrow{\alpha_1,t_1} s_2 \xrightarrow{\alpha_2,t_2} s_2^{\alpha_2} \xrightarrow{\alpha_2,t_2'} s_3$ be a path in $\overline{\mathscr{C}}$. Then $merge(\overline{\pi}) = s_0 \xrightarrow{\alpha_0,t_0+t_0'} s_1 \xrightarrow{\alpha_1,t_1} s_2 \xrightarrow{\alpha_2,t_2+t_2'} s_3$.

For the reverse direction, we map sets of paths in $\mathscr{C}$ to sets of paths in $\overline{\mathscr{C}}$; formally, if $\Pi \subseteq Paths(\mathscr{C})$ we define $extend(\Pi) = \{\overline{\pi} \in Paths(\overline{\mathscr{C}}) \mid merge(\overline{\pi}) \in \Pi\}$.

Lemma 2. *Let $\mathscr{C}$ be a CTMDP and $\Pi_1, \Pi_2, \dots \subseteq Paths(\mathscr{C})$. Then*

1. $\Pi_1 \subseteq \Pi_2 \Longrightarrow extend(\Pi_1) \subseteq extend(\Pi_2)$,
2. $\Pi_1 \cap \Pi_2 = \emptyset \Longrightarrow extend(\Pi_1) \cap extend(\Pi_2) = \emptyset$ *and*
3. $\bigcup extend(\Pi_k) = extend\big(\bigcup \Pi_k\big)$.

Our goal is to construct gm-schedulers such that the path probabilities in $\mathscr{C}$ and $\overline{\mathscr{C}}$ are equal. Therefore, we first adopt a local view and prove that the probability of a single step in $\mathscr{C}$ equals the probability of the corresponding steps in $\overline{\mathscr{C}}$.

4.1 One-Step Correctness of Local Uniformization

Consider the CTMDP in Fig. 4(a) where $\lambda = \sum \lambda_i$. Assume that action α is chosen in state s_0; then $\frac{\mathbf{R}(s_0,\alpha,s_i)}{E(s_0,\alpha)} = \frac{\lambda_i}{\lambda}$ is the probability to move to state s_i (where $i \in \{0,1,2\}$). Hence the probability to reach s_i in time interval $[0,t]$ is

$$\frac{\lambda_i}{\lambda} \int_0^t \eta_\lambda(dt_1). \tag{2}$$

Let us compute the same probability for $\overline{\mathscr{C}}$ depicted in Fig. 4(b): The probability to go from s_0 to s_i directly (with action α) is $\frac{\overline{\mathbf{R}}(s_0,\alpha,s_i)}{\overline{E}(s_0,\alpha)} = \frac{\lambda_i}{\lambda+\mu}$; however, with probability

$\frac{\overline{\mathbf{R}}(s_0,\alpha,s_0^\alpha)}{\overline{E}(s_0,\alpha)} \cdot \frac{\overline{\mathbf{R}}(s_0^\alpha,\alpha,s_i)}{\overline{E}(s_0^\alpha,\alpha)} = \frac{\mu}{\lambda+\mu} \cdot \frac{\lambda_i}{\lambda}$ we instead move to state s_0^α and only then to s_i. In this case, the probability that in time interval $[0,t]$ an α-transition of s_0 executes, followed by one of s_0^α is $\int_0^t (\lambda+\mu)e^{-(\lambda+\mu)t_1} \int_0^{t-t_1} \lambda e^{-\lambda t_2}\, dt_2\, dt_1$. Hence, we reach state s_i with action α in at most t time units with probability

$$\frac{\lambda_i}{\lambda+\mu}\int_0^t \eta_{\lambda+\mu}(dt_1) + \frac{\mu}{\lambda+\mu}\cdot\frac{\lambda_i}{\lambda}\int_0^t \eta_{\lambda+\mu}(dt_1)\int_0^{t-t_1}\eta_\lambda(dt_2). \tag{3}$$

It is easy to verify that (2) and (3) are equal. Thus the probability to reach a (non-copy) successor state in $\{s_0,s_1,s_2\}$ is the same for $\mathscr{C}$ and $\overline{\mathscr{C}}$. It can be computed by replacing λ_i with $\sum\lambda_i$ in (2) and (3). This straightforwardly extends to the Borel σ-field $\mathfrak{B}(\mathbb{R}_{\geq 0})$; further, the equality of (2) and (3) is preserved even if we integrate over a Borel-measurable function $f:\mathbb{R}_{\geq 0}\to[0,1]$. In the following, we consider the probability to reach an arbitrary non-copy state within time $T\in\mathfrak{B}(\mathbb{R}_{\geq 0})$; thus in the following lemma, we replace λ_i with $\sum\lambda_i=\lambda$:

Lemma 3 (One-step timing). *Let $f:\mathbb{R}_{\geq 0}\to[0,1]$ be a Borel measurable function and $T\in\mathfrak{B}(\mathbb{R}_{\geq 0})$. Then*

$$\int_T f(t)\,\eta_\lambda(dt) = \frac{\lambda}{\lambda+\mu}\int_T f(t)\,\eta_{\lambda+\mu}(dt) + \frac{\mu}{\lambda+\mu}\int_{\mathbb{R}_{\geq 0}}\eta_{\lambda+\mu}(dt_1)\int_{T\ominus t_1} f(t_1+t_2)\,\eta_\lambda(dt_2)$$

where $T\ominus t = \{t'\in\mathbb{R}_{\geq 0}\mid t+t'\in T\}$.

The equality of (2) and (3) proves that the probability of a single step in $\mathscr{C}$ equals the probability of one or two transitions (depending on the copy-state) in $\overline{\mathscr{C}}$. In the next section, we lift this argument to sets of paths in $\mathscr{C}$ and $\overline{\mathscr{C}}$.

4.2 Local Uniformization Is Measure Preserving

We prove that for any gm-scheduler D (on $\mathscr{C}$) there exists a gm-scheduler $\overline{D}$ (on $\overline{\mathscr{C}}$) such that the induced probabilities for the sets of paths Π and $extend(\Pi)$ are equal. However, as $\overline{\mathscr{C}}$ differs from $\mathscr{C}$, we cannot use D to directly infer probabilities on $\overline{\mathscr{C}}$. Instead, given a history $\overline{\pi}$ in $\overline{\mathscr{C}}$, we define $\overline{D}(\overline{\pi},\cdot)$ such that it mimics the decision that D takes in $\mathscr{C}$ for history $merge(\overline{\pi})$: For all $\overline{\pi}\in Paths^\star(\overline{\mathscr{C}})$,

$$\overline{D}(\overline{\pi},\cdot) = \begin{cases} D(\pi,\cdot) & \text{if } \overline{\pi}[0],\overline{\pi}\!\downarrow\,\in\mathscr{S}\wedge merge(\overline{\pi})=\pi \\ \{\alpha\mapsto 1\} & \text{if } \overline{\pi}\!\downarrow\,= s^\alpha\in\mathscr{S}_{\text{cp}} \\ \gamma_{\overline{\pi}} & \text{otherwise,}\end{cases}$$

where $\gamma_{\overline{\pi}}$ is an arbitrary distribution over $Act(\overline{\pi}\!\downarrow)$: If $merge$ is applicable to $\overline{\pi}$ (i.e. if $\overline{\pi}[0],\pi\!\downarrow\,\in\mathscr{S}$), then $\overline{D}(\overline{\pi},\cdot)$ is the distribution that D yields for path $merge(\overline{\pi})$ in $\mathscr{C}$; further, if $\overline{\pi}\!\downarrow\,=s^\alpha$ then $Act(s^\alpha)=\{\alpha\}$ and thus $\overline{D}$ chooses action α. Finally, $\overline{\mathscr{C}}$ contains paths that start in a copy-state s^α. But as $\overline{\nu}(s^\alpha)=0$ for all $s^\alpha\in\mathscr{S}_{\text{cp}}$, they do not contribute any probability, independent of $\overline{D}(\overline{\pi},\cdot)$.

Based on this, we consider a measurable base B of the form $B=S_0\times A_0\times T_0\times\ldots\times S_n$ in $\mathscr{C}$. This corresponds to the set $extend(B)$ of paths in $\overline{\mathscr{C}}$. As $extend(B)$ contains paths of different lengths, we resort to its induced (infinite) cylinder $Cyl(extend(B))$ and prove that its probability equals that of B :

Lemma 4 (Measure preservation under local uniformization). *Let $\mathscr{C} = (\mathscr{S}, Act, \mathbf{R}, \nu)$ be a CTMDP, D a gm-scheduler on $\mathscr{C}$ and $B = S_0 \times A_0 \times T_0 \times \cdots \times S_n \in \mathfrak{F}_{Paths^n(\mathscr{C})}$. Then there exists a gm-scheduler $\overline{D}$ such that*

$$Pr^n_{\nu,D}(B) = \overline{Pr}^{\omega}_{\overline{\nu},\overline{D}}\big(Cyl(extend(B))\big)$$

where $\overline{Pr}^{\omega}_{\overline{\nu},\overline{D}}$ is the probability measure induced by $\overline{D}$ and $\overline{\nu}$ on $\mathfrak{F}_{Paths^{\omega}(\overline{\mathscr{C}})}$.

Proof. To shorten notation, let $\overline{B} = extend(B)$ and $\overline{C} = Cyl(\overline{B})$. In the induction base $B = S_0$ and $Pr^0_{\nu,D}(B) = \sum_{s\in B} \nu(s) = \sum_{s\in\overline{B}} \overline{\nu}(s) = \overline{Pr}^0_{\overline{\nu},\overline{D}}(\overline{B}) = \overline{Pr}^{\omega}_{\overline{\nu},\overline{D}}(\overline{C})$. In the induction step, we extend B with a set of initial path prefixes $I = S_0 \times A_0 \times T_0$ and consider the base $I \times B$ which contains paths of length $n+1$:

$$\begin{aligned}
Pr^{n+1}_{\nu,D}(I \times B) &= \int_I Pr^n_{\nu_i,D_i}(B)\, \mu^1_{\nu,D}(di) && \text{by Lemma 1}\\
&= \int_I \overline{Pr}^{\omega}_{\overline{\nu_i},\overline{D_i}}(\overline{C})\, \mu^1_{\nu,D}(di) && \text{by ind. hyp.}\\
&= \sum_{s\in S_0} \nu(s) \sum_{\alpha\in A_0} D(s,\alpha) \int_{T_0} \overline{Pr}^{\omega}_{\overline{\nu_i},\overline{D_i}}(\overline{C})\, \eta_{E(s,\alpha)}(dt) && \text{where } i = (s,\alpha,t)\\
&= \sum_{s\in S_0} \overline{\nu}(s) \sum_{\alpha\in A_0} \overline{D}(s,\alpha) \int_{T_0} \underbrace{\overline{Pr}^{\omega}_{\overline{\nu_i},\overline{D_i}}(\overline{C})}_{f(s,\alpha,t)}\, \eta_{E(s,\alpha)}(dt) && \text{by Def. of } \overline{\nu}, \overline{D}.
\end{aligned}$$

The probabilities $\overline{Pr}^{\omega}_{\overline{\nu_i},\overline{D_i}}(\overline{C})$ define a measurable function $f(s,\alpha,\cdot) : \mathbb{R}_{\geq 0} \to [0,1]$ where $f(s,\alpha,t) = \overline{Pr}^{\omega}_{\overline{\nu_i},\overline{D_i}}(\overline{C})$ if $i = (s,\alpha,t)$. Therefore we can apply Lemma 3 and obtain

$$\begin{aligned}
Pr^{n+1}_{\nu,D}(I \times B) = \sum_{s\in S_0} \overline{\nu}(s) \sum_{\alpha\in A_0} \overline{D}(s,\alpha) \cdot \Big[&\overline{\mathbf{P}}(s,\alpha,\mathscr{S}) \int_{T_0} f(s,\alpha,t)\, \eta_{\overline{E}(s,\alpha)}(dt)\\
&+ \overline{\mathbf{P}}(s,\alpha,s^{\alpha}) \int_{\mathbb{R}_{\geq 0}} \eta_{\overline{E}(s,\alpha)}(dt_1) \int_{T_0 \ominus t_1} f(s,\alpha,t_1+t_2)\, \eta_{\overline{E}(s^{\alpha},\alpha)}(dt_2)\Big].
\end{aligned} \tag{4}$$

To rewrite this further, note that any path prefix $i = (s,\alpha,t)$ in $\mathscr{C}$ induces the sets of path prefixes $\overline{I}_1(i) = \big\{s \xrightarrow{\alpha,t}\big\}$ and $\overline{I}_2(i) = \big\{s \xrightarrow{\alpha,t_1} s^{\alpha} \xrightarrow{\alpha,t_2} \mid t_1 + t_2 = t\big\}$ in $\overline{\mathscr{C}}$, where $\overline{I}_1(i)$ corresponds to directly reaching a state in $\mathscr{S}$, whereas in $\overline{I}_2(i)$ the detour via copy-state s^{α} is taken. As defined in Lemma 1, $\nu_i(s') = \mathbf{P}(s,\alpha,s')$ is the probability to go to state s' when moving along prefix i in $\mathscr{C}$. Similarly, for $\overline{\mathscr{C}}$ we define $\overline{\nu}_{\overline{i}}(s')$ as the probability to be in state $s' \in \mathscr{S}$ after a path prefix $\overline{i} \in \overline{I}_1(i) \cup \overline{I}_2(i)$: If $\overline{i} \in \overline{I}_1(i)$ then we move to a state $s' \in \mathscr{S}$ directly and do not visit copy-state s^{α}. Thus $\overline{\nu}_{\overline{i}}(s') = \overline{\mathbf{P}}(s,\alpha,s')$ for $\overline{i} \in \overline{I}_1(i)$. Further, $\mathbf{P}(s,\alpha,s')$ in $\mathscr{C}$ equals the conditional probability $\frac{\overline{\mathbf{P}}(s,\alpha,s')}{\overline{\mathbf{P}}(s,\alpha,\mathscr{S})}$ to enter s' in $\overline{\mathscr{C}}$ given that we move there directly. Therefore $\overline{\nu}_{\overline{i}}(s') = \overline{\mathbf{P}}(s,\alpha,\mathscr{S}) \cdot \nu_i(s')$ if $\overline{i} \in \overline{I}_1(i)$.

If instead $\overline{i} \in \overline{I}_2(i)$ then $\overline{i}$ has the form $s \xrightarrow{\alpha,t_1} s^{\alpha} \xrightarrow{\alpha,t_2}$ and $\overline{\nu}_{\overline{i}}(s') = \overline{\mathbf{P}}(s^{\alpha},\alpha,s')$ is the probability to end up in state s' after $\overline{i}$. By the definition of s^{α}, this is equal to the probability to move from state s to s' in $\mathscr{C}$. Hence $\overline{\nu}_{\overline{i}}(s') = \nu_i(s')$ if $\overline{i} \in \overline{I}_2(i)$.

As defined in Lemma 1, $D_i(\pi,\cdot) = D(i \circ \pi,\cdot)$ and $\overline{D}_{\overline{i}}(\overline{\pi},\cdot) = \overline{D}(\overline{i} \circ \overline{\pi},\cdot)$. From the definition of $\overline{D}$ we obtain $D_i(\pi,\cdot) = \overline{D}_{\overline{i}}(\overline{\pi},\cdot)$ for all $\overline{i} \in \overline{I}_1(i) \cup \overline{I}_2(i)$ and $\overline{\pi} \in extend(\pi)$. Hence it follows that if $i = (s,\alpha,t)$ and $\overline{i} \in \overline{I}_1(i) \cup \overline{I}_2(i)$ it holds

$$\overline{Pr}^{\omega}_{\overline{\nu}_{\bar{i}},\overline{D}_{\bar{i}}}(\overline{C}) = \begin{cases} \overline{\mathbf{P}}(s,\alpha,\mathscr{S}) \cdot \overline{Pr}^{\omega}_{\overline{\nu}_i,\overline{D}_i}(\overline{C}) & \text{if } \bar{i} \in \overline{I}_1(i) \\ \overline{Pr}^{\omega}_{\overline{\nu}_i,\overline{D}_i}(\overline{C}) & \text{if } \bar{i} \in \overline{I}_2(i). \end{cases} \tag{5}$$

Note that the first summand in (4) corresponds to the set $\overline{I_1}(s,\alpha,t)$ and the second to $\overline{I}_2(s,\alpha,t_1+t_2)$. Applying equality (5) to the right-hand side of (4) we obtain

$$Pr^{n+1}_{\nu,D}(I \times B) = \sum_{s\in S_0} \overline{\nu}(s) \sum_{\alpha\in A_0} \overline{D}(s,\alpha) \int_{T_0} \overline{Pr}^{\omega}_{\overline{\nu}_{\bar{i}},\overline{D}_{\bar{i}}}(\overline{C})\ \eta_{\overline{E}(s,\alpha)}(dt)$$
$$+ \sum_{s\in S_0} \overline{\nu}(s) \sum_{\alpha\in A_0} \overline{D}(s,\alpha) \cdot \overline{\mathbf{P}}(s,\alpha,s^{\alpha}) \int_{\mathbb{R}_{\geq 0}} \eta_{\overline{E}(s,\alpha)}(dt_1) \int_{T_0 \ominus t_1} \overline{Pr}^{\omega}_{\overline{\nu}_{\bar{i}},\overline{D}_{\bar{i}}}(\overline{C})\ \eta_{\overline{E}(s^{\alpha},\alpha)}(dt_2).$$

Applying Def. 5 allows to integrate over the sets of path prefixes $\overline{I}_1 = \bigcup_{i\in I} \overline{I}_1(i)$ and $\overline{I}_2 = \bigcup_{i\in I} \overline{I}_2(i)$ which are induced by $I = S_0 \times A_0 \times T_0$ and to obtain

$$Pr^{n+1}_{\nu,D}(I \times B) = \int_{\overline{I}_1} \overline{Pr}^{\omega}_{\overline{\nu}_{\bar{i}},\overline{D}_{\bar{i}}}(\overline{C})\ \overline{\mu}^1_{\overline{\nu},\overline{D}}(d\bar{i}) + \int_{\overline{I}_2} \overline{Pr}^{\omega}_{\overline{\nu}_{\bar{i}},\overline{D}_{\bar{i}}}(\overline{C})\ \overline{\mu}^2_{\overline{\nu},\overline{D}}(d\bar{i}).$$

Rewriting the right-hand side yields $Pr^{n+1}_{\nu,D}(I \times B) = \overline{Pr}^{\omega}_{\overline{\nu},\overline{D}}\big(Cyl(extend(I \times B))\big)$. □

Lemma 4 holds for all measurable rectangles $B = S_0 \times A_0 \times T_0 \times \ldots \times S_n$; however, we aim at an extension to arbitrary bases $B \in \mathfrak{F}_{Paths^n(\mathscr{C})}$. Thus let $\mathfrak{G}_{Paths^n(\mathscr{C})}$ be the class of all finite disjoint unions of measurable rectangles. Then $\mathfrak{G}_{Paths^n(\mathscr{C})}$ is a *field* [10, p. 102]:

Lemma 5. *Let $\mathscr{C} = (\mathscr{S}, Act, \mathbf{R}, \nu)$ be a CTMDP, D a gm-scheduler on $\mathscr{C}$ and $n \in \mathbb{N}$. Then $Pr^n_{\nu,D}(B) = \overline{Pr}^{\omega}_{\overline{\nu},\overline{D}}\big(Cyl(extend(B))\big)$ for all $B \in \mathfrak{G}_{Paths^n(\mathscr{C})}$.*

With the monotone class theorem [10], the preservation property extends from $\mathfrak{G}_{Paths^n}$ to the σ-field $\mathfrak{F}_{Paths^n}$: A class $\mathfrak{C}$ of subsets of $Paths^n$ is a monotone class if it is closed under in- and decreasing sequences: if $\Pi_k \in \mathfrak{C}$ and $\Pi \subseteq Paths^n$ such that $\Pi_0 \subseteq \Pi_1 \subseteq \cdots$ and $\bigcup_{k=0}^{\infty} \Pi_k = \Pi$, we write $\Pi_k \uparrow \Pi$ (similary for $\Pi_k \downarrow \Pi$). Then $\mathfrak{C}$ is a monotone class iff for all $\Pi_k \in \mathfrak{C}$ and $\Pi \subseteq Paths^n$ with $\Pi_k \uparrow \Pi$ or $\Pi_k \downarrow \Pi$ it holds that $\Pi \in \mathfrak{C}$.

Lemma 6 (Monotone class). *Let $\mathscr{C} = (\mathscr{S}, Act, \mathbf{R}, \nu)$ be a CTMDP with gm-scheduler D. The set $\mathfrak{C} = \Big\{ B \in \mathfrak{F}_{Paths^n(\mathscr{C})} \mid Pr^n_{\nu,D}(B) = \overline{Pr}^{\omega}_{\overline{\nu},\overline{D}}\big(Cyl(extend(B))\big) \Big\}$ is a monotone class.*

Lemma 7 (Extension). *Let $\mathscr{C} = (\mathscr{S}, Act, \mathbf{R}, \nu)$ be a CTMDP, D a gm-scheduler on $\mathscr{C}$ and $n \in \mathbb{N}$. Then $Pr^n_{\nu,D}(B) = \overline{Pr}^{\omega}_{\overline{\nu},\overline{D}}\big(Cyl(extend(B))\big)$ for all $B \in \mathfrak{F}_{Paths^n(\mathscr{C})}$.*

Proof. By Lemma 6, $\mathfrak{C}$ is a monotone class and by Lemma 5 it follows that $\mathfrak{G}_{Paths^n(\mathscr{C})} \subseteq \mathfrak{C}$. Thus, the Monotone Class Theorem [10, Th. 1.3.9] applies and $\mathfrak{F}_{Paths^n} \subseteq \mathfrak{C}$. Hence $Pr^n_{\nu,D}(B) = \overline{Pr}^{\omega}_{\overline{\nu},\overline{D}}\big(Cyl(extend(B))\big)$ for all $B \in \mathfrak{F}_{Paths^n}$. □

Lemma 4 and its measure-theoretic extension to the σ-field are the basis for the major results of this work as presented in the next section.

5 Main Results

The first result states the correctness of the construction of scheduler $\overline{D}$, i.e. it asserts that D and $\overline{D}$ assign the same probability to corresponding sets of paths.

Theorem 1. *Let $\mathscr{C} = (\mathscr{S}, Act, \mathbf{R}, \nu)$ be a CTMDP and D a gm-scheduler on $\mathscr{C}$. Then $Pr^{\omega}_{\nu,D}(\Pi) = \overline{Pr}^{\omega}_{\overline{\nu},\overline{D}}\big(extend(\Pi)\big)$ for all $\Pi \in \mathfrak{F}_{Paths^{\omega}}$.*

Proof. Each cylinder $\Pi \in \mathfrak{F}_{Paths^{\omega}(\mathscr{C})}$ is induced by a measurable base [10, Thm. 2.7.2]; hence $\Pi = Cyl(B)$ for some $B \in \mathfrak{F}_{Paths^{n}(\mathscr{C})}$ and $n \in \mathbb{N}$. But then, $Pr^{\omega}_{\nu,D}(\Pi) = Pr^{n}_{\nu,D}(B)$ and $Pr^{n}_{\nu,D}(B) = \overline{Pr}^{\omega}_{\overline{\nu},\overline{D}}\big(extend(\Pi)\big)$ by Lemma 7. □

With Lemma 4 and its extension, we are now ready to prove that local uniformization does not alter the CTMDP in a way that we leak probability mass with respect to the most important scheduler classes:

Theorem 2. *Let $\mathscr{C} = (\mathscr{S}, Act, \mathbf{R}, \nu)$ be a CTMDP and $\Pi \in \mathfrak{F}_{Paths^{\omega}(\mathscr{C})}$. For scheduler classes $\mathfrak{D} \in \{TH, TTH, TTP, TAH, TAP\}$ it holds that*

$$\sup_{D \in \mathfrak{D}(\mathscr{C})} Pr^{\omega}_{\nu,D}(\Pi) \leq \sup_{D' \in \mathfrak{D}(\overline{\mathscr{C}})} \overline{Pr}^{\omega}_{\overline{\nu},D'}(extend(\Pi)). \tag{6}$$

Proof. By Thm. 1, the claim follows for the class of all gm-schedulers, that is, for $\mathfrak{D} = TH$. For the other classes, it remains to check that the gm-scheduler $\overline{D}$ used in Lemma 4 also falls into the respective class. Here, we state the proof for TTP: If $D : \mathscr{S} \times \mathbb{R}_{\geq 0} \to Distr(Act) \in TTP$, define $\overline{D}(s,\Delta) = D(s,\Delta)$ if $s \in \mathscr{S}$ and $\overline{D}(s^{\alpha},\Delta) = \{\alpha \mapsto 1\}$ for $s^{\alpha} \in \mathscr{S}_{\text{cp}}$. Then Lemma 4 applies verbatim. □

Thm. 4 proves that (6) does not hold for TP and $TAHOP$. Although we obtain a gm-scheduler $\overline{D}$ on $\overline{\mathscr{C}}$ for any $D \in TP(\mathscr{C}) \cup TAHOP(\mathscr{C})$ by Thm. 1, $\overline{D}$ is generally not in $TP(\overline{\mathscr{C}})$ (or $TAHOP(\overline{\mathscr{C}})$, resp.). For the main result, we identify the scheduler classes, that do not gain probability mass by local uniformization:

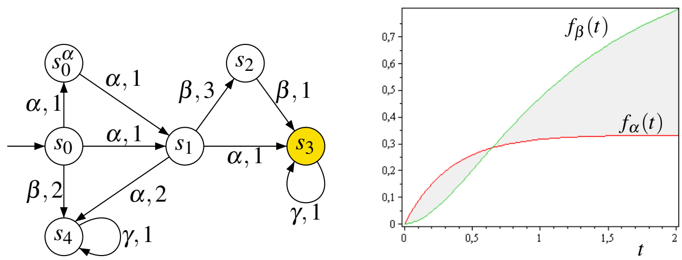
Theorem 3. *Let $\mathscr{C} = (\mathscr{S}, Act, \mathbf{R}, \nu)$ be a CTMDP and $\Pi \in \mathfrak{F}_{Paths^{\omega}(\mathscr{C})}$. Then*

$$\sup_{D \in \mathfrak{D}(\mathscr{C})} Pr^{\omega}_{\nu,D}(\Pi) = \sup_{D' \in \mathfrak{D}(\overline{\mathscr{C}})} \overline{Pr}^{\omega}_{\overline{\nu},D'}(extend(\Pi)) \quad \textit{for } \mathfrak{D} \in \{TTP, TAP\}.$$

Proof. Thm. 2 proves the direction from left to right. For the reverse, let $D' \in TTP(\overline{\mathscr{C}})$ and define $D \in TTP(\mathscr{C})$ such that $D(s,\Delta) = D'(s,\Delta)$ for all $s \in \mathscr{S}, \Delta \in \mathbb{R}_{\geq 0}$. Then $\overline{D} = D'$ and $\overline{Pr}^{\omega}_{\overline{\nu},D'}(extend(\Pi)) = Pr^{\omega}_{\nu,D}(\Pi)$ by Thm. 1. Hence the claim for TTP follows; analogue for $D' \in TAP(\overline{\mathscr{C}})$. □

Conjecture 1. We conjecture that Thm. 3 also holds for TH and TTH. For $D' \in TH(\overline{\mathscr{C}})$, we aim at defining a scheduler $D \in TH(\mathscr{C})$ that induces the same probabilities on $\mathscr{C}$. However, a history $\pi \in Paths^{\star}(\mathscr{C})$ corresponds to the uncountable set $extend(\pi)$ in $\overline{\mathscr{C}}$ s.t. $D'(\overline{\pi},\cdot)$ may be different for each $\overline{\pi} \in extend(\pi)$. As D can only decide once on history π, in order to mimic D' on $\overline{\mathscr{C}}$, we propose to weigh each distribution $D'(\overline{\pi},\cdot)$ with the conditional probability of $d\overline{\pi}$ given $extend(\pi)$.

In the following, we disprove (6) for TP and $TAHOP$ schedulers. Intuitively, TP schedulers rely on the sojourn time in the last state; however, local uniformization changes the exit rates of states by adding transitions to copy-states.



(a) Local uniformization of Fig. 1 (b) From state s_1 to state s_3.

Fig. 5. Timed reachability of state s_3 (starting in s_1) in $\mathscr{C}$ and $\overline{\mathscr{C}}$

Theorem 4. *For $\mathfrak{G} \in \{TP, TAHOP\}$, there exists $\mathscr{C}$ and $\Pi \in \mathfrak{F}_{Paths^\omega(\mathscr{C})}$ such that*

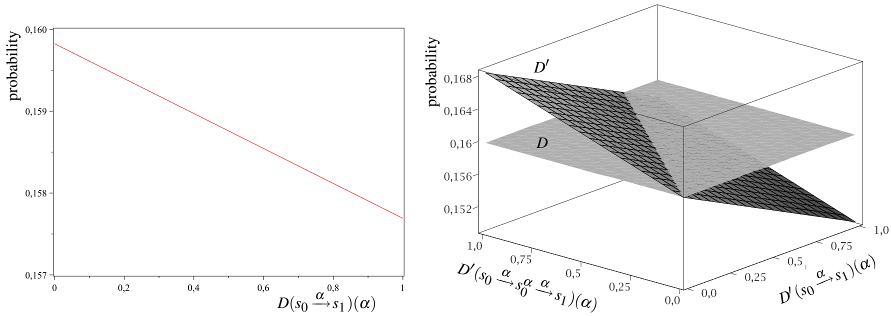
$$\sup_{D \in \mathfrak{G}(\mathscr{C})} Pr^\omega_{\nu,D}(\Pi) > \sup_{D' \in \mathfrak{G}(\overline{\mathscr{C}})} \overline{Pr}^\omega_{\overline{\nu},D'}\big(extend(\Pi)\big).$$

Proof. We give the proof for TP: Consider the CTMDPs $\mathscr{C}$ and $\overline{\mathscr{C}}$ in Fig. 1 and Fig. 5(a), resp. Let $\Pi \in \mathfrak{F}_{Paths^\omega(\mathscr{C})}$ be the set of paths in $\mathscr{C}$ that reach state s_3 in 1 time unit and let $\overline{\Pi} = extend(\Pi)$. To optimize $Pr^\omega_{\nu,D}(\Pi)$ and $\overline{Pr}^\omega_{\overline{\nu},D'}(\overline{\Pi})$, any scheduler D (resp. D') must choose $\{\alpha \mapsto 1\}$ in state s_0. Nondeterminism only remains in state s_1; here, the optimal distribution over $\{\alpha, \beta\}$ depends on the time t_0 that was spent to reach state s_1: In $\mathscr{C}$ and $\overline{\mathscr{C}}$, the probability to go from s_1 to s_3 in the remaining $t = 1 - t_0$ time units is $f_\alpha(t) = \frac{1}{3} - \frac{1}{3}e^{-3t}$ for α and $f_\beta(t) = 1 + \frac{1}{2}e^{-3t} - \frac{3}{2}e^{-t}$ for β. Fig. 5(b) shows the cdfs of f_α and f_β; as any convex combination of α and β results in a cdf in the shaded area of Fig. 5(b), we only need to consider the extreme distributions $\{\alpha \mapsto 1\}$ and $\{\beta \mapsto 1\}$ for maximal reachability. Let d be the unique solution (in $\mathbb{R}_{>0}$) of $f_\alpha(t) = f_\beta(t)$, i.e. the point where the two cdfs cross. Then $D_{opt}(s_0 \xrightarrow{\alpha,t_0} s_1, \cdot) = \{\alpha \mapsto 1\}$ if $1 - t_0 \leq d$ and $\{\beta \mapsto 1\}$ otherwise, is an optimal gm-scheduler for Π on $\mathscr{C}$ and $D_{opt} \in TP(\mathscr{C}) \cap TTP(\mathscr{C})$ as it depends only on the delay of the last transition.

For $\overline{\Pi}$, D' is an optimal gm-scheduler on $\overline{\mathscr{C}}$ if $D'(s_0 \xrightarrow{\alpha,t_0} s_1, \cdot) = D_{opt}(s_0 \xrightarrow{\alpha,t_0} s_1, \cdot)$ as before and $D'(s_0 \xrightarrow{\alpha,t_0} s_0^\alpha \xrightarrow{\alpha,t_1} s_1, \cdot) = \{\alpha \mapsto 1\}$ if $1 - t_0 - t_1 \leq d$ and $\{\beta \mapsto 1\}$ otherwise. Note that by definition, $D' = \overline{D_{opt}}$ and $\overline{D_{opt}} \in TTP(\overline{\mathscr{C}})$, whereas $D' \notin TP(\overline{\mathscr{C}})$ as any $TP(\overline{\mathscr{C}})$ scheduler is independent of t_0. For history $\pi = s_0 \xrightarrow{\alpha,t_0} s_0^\alpha \xrightarrow{\alpha,t_1} s_1$, the best approximation of t_0 is the expected sojourn time in state s_0, i.e. $\frac{1}{E(s_0,\alpha)}$. For the induced scheduler $D'' \in TP(\overline{\mathscr{C}})$, it holds $D''(s_1, t_1) \neq D'(s_0 \xrightarrow{\alpha,t_0} s_0^\alpha \xrightarrow{\alpha,t_1} s_1)$ almost surely. But as $\overline{D_{opt}}$ is optimal, there exists $\varepsilon > 0$ such that $\overline{Pr}^\omega_{\overline{\nu},D''}(\overline{\Pi}) = \overline{Pr}^\omega_{\overline{\nu},\overline{D_{opt}}}(\overline{\Pi}) - \varepsilon$. Therefore

$$\sup_{D'' \in TP(\overline{\mathscr{C}})} \overline{Pr}^\omega_{\overline{\nu},D''}(\overline{\Pi}) < \overline{Pr}^\omega_{\overline{\nu},\overline{D_{opt}}}(\overline{\Pi}) = Pr^\omega_{\nu,D_{opt}}(\Pi) = \sup_{D \in TP(\mathscr{C})} Pr^\omega_{\nu,D}(\Pi).$$

For $TAHOP$, a similar proof applies that relies on the fact that local uniformization changes the number of transitions needed to reach a goal state. $\square$



(a) *TAH*-schedulers on $\mathscr{C}$ (b) *TAH*-schedulers on $\overline{\mathscr{C}}$

Fig. 6. Optimal *TAH*-schedulers for time-bounded reachability

This proves that by local uniformization, essential information for *TP* and *TAHOP* schedulers is lost. In other cases, schedulers from *TAH* and *TAHOP* gain information by local uniformization:

Theorem 5. *There exists CTMDP $\mathscr{C} = (\mathscr{S}, Act, \mathbf{R}, \nu)$ and $\Pi \in \mathfrak{F}_{Paths^\omega(\mathscr{C})}$ such that*

$$\sup_{D \in \mathfrak{S}(\mathscr{C})} Pr^\omega_{\nu,D}(\Pi) < \sup_{D' \in \mathfrak{S}(\overline{\mathscr{C}})} \overline{Pr}^\omega_{\overline{\nu},D'}\big(extend(\Pi)\big) \quad for\ \mathfrak{S} = \{TAH, TAHOP\}.$$

Proof. Consider the CTMDPs $\mathscr{C}$ and $\overline{\mathscr{C}}$ in Fig. 1 and Fig. 5(a), resp. Let Π be the time-bounded reachability property of state s_3 within 1 time unit and let $\overline{\Pi} = extend(\Pi)$. We prove the claim for *TAH*: Therefore we derive $D \in TAH(\mathscr{C})$ such that $Pr^\omega_{\nu,D}(\Pi) = \sup_{D' \in TAH(\mathscr{C})} Pr^\omega_{\nu,D'}(\Pi)$. For this, $D(s_0) = \{\alpha \mapsto 1\}$ must obviously hold. Thus the only choice is in state s_1 for time-abstract history $s_0 \xrightarrow{\alpha} s_1$ where $D(s_0 \xrightarrow{\alpha} s_1) = \mu$, $\mu \in Distr(\{\alpha, \beta\})$. For initial state s_0, Fig. 6(a) depicts $Pr^\omega_{\nu,D}(\Pi)$ for all $\mu \in Distr(\{\alpha, \beta\})$; obviously, $D(s_0 \xrightarrow{\alpha} s_1) = \{\beta \mapsto 1\}$ maximizes $Pr^\omega_{\nu,D}(\Pi)$. On $\overline{\mathscr{C}}$, we prove that there exists $D' \in TAH(\overline{\mathscr{C}})$ such that $Pr^\omega_{\nu,D}(\Pi) < \overline{Pr}_{\overline{\nu},D'}(\overline{\Pi})$: To maximize $\overline{Pr}^\omega_{\overline{\nu},D'}(\overline{\Pi})$, define $D'(s_0) = \{\alpha \mapsto 1\}$. Note that D' may yield different distributions for the time-abstract paths $s_0 \xrightarrow{\alpha} s_1$ and $s_0 \xrightarrow{\alpha} s_0^\alpha \xrightarrow{\alpha} s_1$; for $\mu, \mu_c \in Distr(\{\alpha, \beta\})$ such that $\mu = D'(s_0 \xrightarrow{\alpha} s_1)$ and $\mu_c = D'(s_0 \xrightarrow{\alpha} s_0^\alpha \xrightarrow{\alpha} s_1)$ the probability of $\overline{\Pi}$ under D' is depicted in Fig. 6(b) for all $\mu, \mu_c \in Distr(\{\alpha, \beta\})$. Clearly, $\overline{Pr}^\omega_{\overline{\nu},D'}(\overline{\Pi})$ is maximal if $D'(s_0 \xrightarrow{\alpha} s_1) = \{\beta \mapsto 1\}$ and $D'(s_0 \xrightarrow{\alpha} s_0^\alpha \xrightarrow{\alpha} s_1) = \{\alpha \mapsto 1\}$. Further, Fig. 6(b) shows that with this choice of D', $\overline{Pr}^\omega_{\overline{\nu},D'}(\overline{\Pi}) > Pr^\omega_{\nu,D}(\Pi)$ and the claim follows. For *TAHOP*, the proof applies analogously. □

6 Delaying Nondeterministic Choices

To conclude the paper, we show how local uniformization allows to derive the class of *late schedulers* which resolve nondeterminism only when leaving a state. Hence, they may exploit information about the current state's sojourn time and, as a consequence, induce more accurate probability bounds than gm-schedulers.

More precisely, let $\mathscr{C} = (\mathscr{S}, Act, \mathbf{R}, \nu)$ be a locally uniform CTMDP and D a gm-scheduler on $\mathscr{C}$. Then $E(s,\alpha) = u(s)$ for all $s \in \mathscr{S}$ and $\alpha \in Act$ (cf. Def. 7). Thus the measures $\eta_{E(s,\alpha)}$ in Def. 3 do not depend on α and we may exchange their order of integration in (1) by applying [10, Thm. 2.6.6]. Hence for locally uniform CTMDPs let

$$\mu_D(\pi,M) = \int_{\mathbb{R}_{\geq 0}} \eta_{u(\pi\downarrow)}(dt) \int_{Act} D(\pi,d\alpha) \int_{\mathscr{S}} \mathbf{I}_M(\alpha,t,s')\ \mathbf{P}(s,\alpha,ds'). \tag{7}$$

Formally, (7) allows to define *late schedulers* as mappings $D : Paths^{\star}(\mathscr{C}) \times \mathbb{R}_{\geq 0} \times \mathfrak{F}_{Act} \to [0,1]$ that extend gm-schedulers with the sojourn-time in $\pi\downarrow$. Note that local uniformity is essential here: In the general case, the measures $\eta_{E(s,\alpha)}(dt)$ and a late scheduler $D(\pi,t,d\alpha)$ are inter-dependent in t and α; hence, in Def. 3, $\mu_D(\pi,\cdot)$ is not well-defined for late-schedulers. Intuitively, the sojourn time t of the current state s depends on D while D depends on t.

Let *LATE* and *GM* denote the classes of late and gm-schedulers, respectively. For all $\Pi \in Paths^{\omega}(\mathscr{C})$:

$$\sup_{D \in GM} Pr^{\omega}_{\nu,D}(\Pi) \leq \sup_{D \in LATE} Pr^{\omega}_{\nu,D}(\Pi) \tag{8}$$

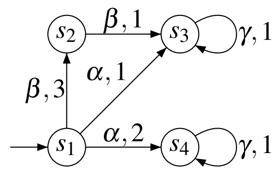


Fig. 7. Example

holds as $GM \subseteq LATE$. By Thm. 3, *TTP* and *TAP* preserve probability bounds; hence, late-schedulers are well-defined for those classes and yield better probability bounds than gm-schedulers, i.e., in general inequality (8) is strict: Let $\mathscr{C}$ be as in Fig. 7 and Π be timed-reachability for s_3 in 1 time unit. Then $\sup_{D \in GM} Pr^{\omega}_{\nu,D}(\Pi) = 1 + \frac{1}{2}e^{-3} - \frac{3}{2}e^{-1}$. On the other hand, the optimal late scheduler is given by $D(s_1,t,\cdot) = \{\beta \mapsto 1\}$ if $t < 1 + \ln 2 - \ln 3$ and $\{\alpha \mapsto 1\}$ otherwise. Then $Pr^{\omega}_{\nu,D}(\Pi) = 1 + \frac{19}{24}e^{-3} - \frac{3}{2}e^{-1}$ and the claim follows.

7 Conclusion

We studied a hierarchy of scheduler classes for CTMDPs, and investigated their sensitivity for general measures w.r.t. local uniformization. This transformation is shown to be measure-preserving for TAP and TTP schedulers. In addition, in contrast to TP and TAHOP schedulers, TH, TTH, and TAH schedulers cannot lose information to optimize their decisions. TAH and TAHOP schedulers can also gain information. We conjecture that our transformation is also measure-preserving for TTH and TH schedulers. Finally, late schedulers are shown to be able to improve upon generic schedulers [6].

References

1. Knast, R.: Continuous-time probabilistic automata. Inform. and Control 15, 335–352 (1969)
2. Guo, X., Hernández-Lerma, O., Prieto-Rumeau, T.: A survey of recent results on continuous-time Markov decision processes. TOP 14, 177–261 (2006)
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(1), 2–26 (2005)
4. Puterman, M.L.: Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley and Sons, Chichester (1994)

5. Baier, C., Haverkort, B.R., Hermanns, H., Katoen, J.P.: Model-checking algorithms for continuous-time Markov chains. IEEE TSE 29(6), 524–541 (2003)
6. Wolovick, N., Johr, S.: A characterization of meaningful schedulers for continuous-time Markov decision processes. In: Asarin, E., Bouyer, P. (eds.) FORMATS 2006. LNCS, vol. 4202, pp. 352–367. Springer, Heidelberg (2006)
7. Grassmann, W.K.: Finding transient solutions in Markovian event systems through randomization. In: Stewart, W.J. (ed.) Numerical Solutions of Markov Chains, pp. 357–371 (1991)
8. Gross, D., Miller, D.R.: The randomization technique as a modeling tool and solution procedure for transient Markov processes. Oper. Res. 32(2), 343–361 (1984)
9. Jensen, A.: Markov chains as an aid in the study of Markov processes. Skand. Aktuarietidskrift 3, 87–91 (1953)
10. Ash, R., Doléans-Dade, C.: Probability & Measure Theory, 2nd edn. Academic Press, London (2000)
11. Neuhäußer, M.R., Katoen, J.P.: Bisimulation and logical preservation for continuous-time Markov decision processes. In: Caires, L., Vasconcelos, V.T. (eds.) CONCUR 2007. LNCS, vol. 4703, pp. 412–427. Springer, Heidelberg (2007)