2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Reachability Probabilities in Markovian Timed Automata

Taolue Chen, Tingting Han, Joost-Pieter Katoen, and Alexandru Mereacre

***Abstract*— We propose a novel stochastic extension of timed automata, i.e. *Markovian Timed Automata*. We study the problem of optimizing time-bounded reachability probabilities in this model, i.e., the maximum likelihood to hit a set of goal locations within a given deadline. We propose Bellman equations to characterize the probability, and provide two approaches to solve the Bellman equations, namely, a discretization and a reduction to Hamilton-Jacobi-Bellman equations.**

I. INTRODUCTION

This paper introduces **Markovian Timed Automata** (MTA), a novel extension of timed automata [1] with exponentially distributed location residence times. To give some motivation, let's look at the following example: A robot moves on a 3×3-grid (Fig. 1), starting from A and trying to reach B in T_2 units of time. At each cell, it can move up (u), down (d), left (l) and right (r) (when applicable). Cells are associated with rates, which intuitively represent the speed of the robot. As soon as it moves to a cell, the robot decides a direction to move to the next cell, and then waits in the current cell for a certain amount of time, which is governed by an exponential distribution with a given rate λ, i.e., the probability of leaving the cell within time t is $1 - e^{-\lambda t}$.[1] The robot is allowed to stay in consecutive dark cells for at most T_1 units of time, while there is no time constraint for the bright cells. At each cell, the robot has a probability 0.1 to break down, apart from moving to the neighboring one. Since there are different ways to reach B, each of which is associated with some probability, one natural question is: What's the *maximum probability* to reach B from A within T_2 units of time? This problem can be readily formulated as a controller synthesis problem for MTA depicted in Fig. 2, where each location in $\{\ell_0, \ldots, \ell_8\}$ corresponds to a cell and x, y are *clocks* to specify the time constraints. (N.B. the failure location and all the transitions leading to it are omitted for the clearer illustration. For the same reason the representation of the MTA here is a bit different from the one in Fig. 3.)

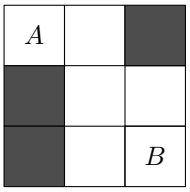

Fig. 1. Robot example

This work is partially supported by the DFG research training group 1295 AlgoSyn, the ERC Advanced Grant VERIWARE, and the FP7 Project MoVeS.

T. Chen, T. Han and A. Mereacre are with the Department of Computer Science, Oxford University. Email: {tchen, than, mereacre}@cs.ox.ac.uk.

J.-P. Katoen is with the Software Modelling and Verification Group, RWTH Aachen University. Email: katoen@cs.rwth-aachen.de.

[1]This seems to be restrictive. However, we note that in practice, one can approximate any distribution by phase-type distributions, resulting in series-parallel combinations of exponential distributions [15].

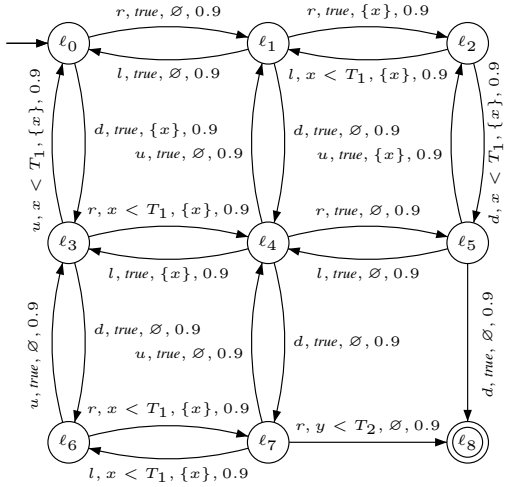

Fig. 2. MTA for the robot example

In general, controller synthesis problems for MTA are to determine the sequence of actions that maximize the probability[2] to reach certain goal locations in MTA within a given deadline (*time-bounded reachability*). To solve this issue, we first apply the standard region construction [1] to MTA. Then we characterize maximum (time-bounded) reachability probabilities by a variant of the Bellman equation [3]. This provides the basis for two approaches to compute such probabilities. The first approach uses discretization, and shows that max-reachability probabilities can be reduced to maximum reachability probabilities in a finite state Markov decision process (MDP), for which various efficient algorithms, such as value iteration [3], exist. We show that the accuracy of our result is $(1-e^{-\lambda h})\cdot(1-e^{-\lambda T})$, where h is the discretization step, T is the deadline, and λ is the maximal rate of all exponential distributions in the MTA. The second approach is based on partial differential equations (PDEs), in particular Hamilton-Jacobi-Bellman equations [13].

We point out that MTA are rather expressive: Zero-clock MTA correspond to a subclass of CTMDPs [2][6], whereas probabilistic timed automata (PTAs) [14] are obtained by basically ignoring the exit rates in any location in the MTA. In earlier work [9], [10], [4], we have used *deterministic* MTA as specification formalism for linear real-time properties over stochastic processes. Some more related works are in order: In [5] the authors consider stochastic timed games, which contain two types of locations: probabilistic ones and locations belonging to one of the players. The authors have addressed the reachability problem for this model: It was shown that the quantitative reachability problem is

[2]The minimal one can be treated in a completely dual fashion. We omitted it for the sake of simplicity.

undecidable in general (for $2\frac{1}{2}$-player games), while the qualitative question "$= 0$" or "$= 1$" can be solved in PTIME for $1\frac{1}{2}$-player games with a single clock. MTA are essentially $1\frac{1}{2}$-player stochastic timed games. However, our focus is on *quantitative* analysis rather than on qualitative analysis or decidability issues. In [7], a game extension of SMDP was considered and the winning objective was specified by a deterministic timed automaton. Again, only the *qualitative* question were addressed.

As a by-product of our work we obtain two procedures to compute maximum time-bounded reachability probabilities in *locally uniform* CTMDPs. This problem has also been treated in [2][6][17], where [2][6] mainly addressed time-abstract schedulers which are not necessarily optimal, while here time-dependent schedulers are considered, as well as [17]. Comparing to [17], we discretize both the time and the state space. Moreover a system of PDEs is derived in order to characterize the maximum reachability probability. The error bound that we obtain improves the error bound given in [17] yielding a substantial reduction in the required number of iterations.

Due to space restriction, all proofs and some further explanation are omitted here. We refer the reader to the technical report [11] for the full details.

II. Markovian Timed Automata

Given a set H, let $\Pr : \mathcal{F}(H) \to [0,1]$ be a probability measure on the measurable space $(H, \mathcal{F}(H))$, where $\mathcal{F}(H)$ is a σ-algebra over H. Let $\mathit{Distr}(H)$ denote the set of probability measures on this measurable space.

A. Markov Decision Processes

Definition 1: [MDP] A (continuous-state) *Markov decision process* is a tuple $\mathcal{D} = (\mathit{Act}, S, s_0, \mathcal{P})$ where

- *Act* is a *denumerable* set of *actions*;
- S is a set of *states*;
- $s_0 \in S$ is the *initial state*;
- $\mathcal{P} : S \times \mathit{Act} \times \mathcal{F}(S) \to [0,1]$ is the *transition probability function*, where $\mathcal{P}(s, \alpha, \cdot)$ is a probability measure over $\mathcal{F}(S)$ for any $s \in S$ and $\alpha \in \mathit{Act}$, such that $\mathcal{P}(\cdot, \cdot, A)$ is measurable for any $A \in \mathcal{F}(S)$.

The measure $\mathcal{P}(s, \alpha, A)$ is the one-step transition probability from state $s \in S$ to the set of states $A \in \mathcal{F}(S)$ by taking action $\alpha \in \mathit{Act}$. Notice that in general one can extend the MDP model to uncountably many actions (see [18]). In this paper we will consider only MDPs which have finitely many actions, i.e., finitely-branching MDPs.

B. Markovian Timed Automata

Let $\mathcal{X}=\{x_1, ..., x_n\}$ be a set of *nonnegative* variables in $\mathbb{R}$, called *clocks*. A clock-valuation is a function $\eta:\mathcal{X}\to\mathbb{R}_{\geqslant 0}$ assigning to each variable x a value $\eta(x)$. Let $\mathcal{V}(\mathcal{X})$ denote the set of all clock-valuations over $\mathcal{X}$. A *clock constraint* on $\mathcal{X}$, denoted by g, is a conjunction of expressions of the form $x \bowtie c$ for clock $x \in \mathcal{X}$, comparison operator $\bowtie \in \{<, \leqslant, >, \geqslant\}$ and $c \in \mathbb{N}$. Let $\mathcal{B}(\mathcal{X})$ denote the set of clock constraints over $\mathcal{X}$. A clock valuation η *satisfies* constraint $x \bowtie c$, denoted $\eta \models x \bowtie c$, if and only if $\eta(x) \bowtie c$; it satisfies a conjunction of such expressions if and only if η satisfies all of them. Let $\vec{0}$ denote the valuation that assigns 0 to all clocks. For a subset $X \subseteq \mathcal{X}$, the reset of X, denoted $\eta[X := 0]$, is the valuation η' such that $\forall x \in X.\ \eta'(x) := 0$ and $\forall x \notin X.\ \eta'(x) := \eta(x)$. For $\delta \in \mathbb{R}_{\geqslant 0}$ and $\mathcal{X}$-valuation η, $\eta+\delta$ is the $\mathcal{X}$-valuation η'' such that $\forall x \in \mathcal{X}.\ \eta''(x) := \eta(x)+\delta$, which implies that all clocks proceed at the same speed.

Definition 2: [MTA] A *Markovian timed automaton* is a tuple $\mathcal{M} = (\mathit{Act}, \mathcal{X}, \mathit{Loc}, \ell_0, E, \rightsquigarrow)$, where

- *Act* is a finite set of *actions*;
- $\mathcal{X}$ is a finite set of *clocks*;
- *Loc* is a finite set of *locations*;
- $\ell_0 \in \mathit{Loc}$ is the *initial location*;
- $E : \mathit{Loc} \to \mathbb{R}_{>0}$ is the *exit rate function* and
- $\rightsquigarrow \subseteq \mathit{Loc} \times \mathit{Act} \times \mathcal{B}(\mathcal{X}) \times \mathit{Distr}(2^{\mathcal{X}} \times \mathit{Loc})$ is the *edge relation*.

For simplicity we abbreviate $(\ell, \alpha, g, \zeta) \in \rightsquigarrow$ by $\ell \overset{\alpha,g}{\rightsquigarrow} \zeta$, where ζ is a probability distribution over $2^{\mathcal{X}} \times \mathit{Loc}$. Here we don't include location invariants, as in [1], and we don't require edge relation $\rightsquigarrow$ to be total, e.g., there might be some clock constraints g for which $\rightsquigarrow$ is not defined.

Example 1: An example MTA is shown in Fig. 3, where there are 6 locations with ℓ_0 the initial location. In ℓ_0 (resp. ℓ_2), there is a decision to be made between actions α_1 and α_2 (resp. γ_1 and γ_2) when the clock valuation is $\eta(x) \in [0,1)$ (resp. $\eta(x) \in (1,2)$). ℓ_3 is the *goal* location, which will be used later. For edge $\ell_0 \overset{\alpha_2, x\geqslant 0}{\rightsquigarrow} \zeta$, $\zeta(\varnothing, \ell_2) = 0.2$ and $\zeta(\varnothing, \ell_5) = 0.8$.

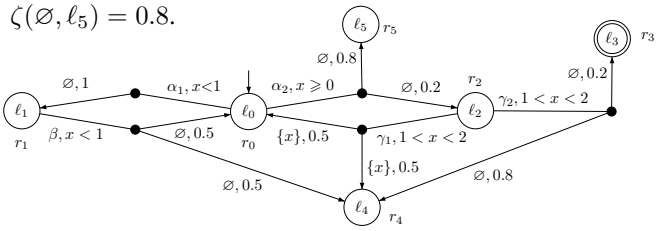

Fig. 3. An example MTA

a) Semantics: Intuitively, an MTA behaves as follows. Consider location ℓ_1 in Fig. 3. As soon as location ℓ_1 is entered with clock valuation η, action β is chosen and a waiting time τ in ℓ_1 is sampled from the probability distribution $E(\ell_1)e^{-E(\ell_1)\tau}$. If $\eta + \tau \models x < 1$, there will be a jump to location ℓ_0 with probability 0.5 or to location ℓ_4 with probability 0.5, otherwise no jump occurs and MTA remains in location ℓ_1. When for instance the next location ℓ_0 is entered with clock valuation η' ($= \eta + \tau$ in this example), action α_1 or α_2 is chosen and the waiting time τ' in ℓ_0 is sampled from the probability distribution $E(\ell_0)e^{-E(\ell_0)\tau'}$. Suppose α_2 is picked and $\eta' + \tau' \models x \geqslant 0$ (the guard of action α_2), the MTA jumps to the next location according to the probability distribution associated with α_2. The MTA follows the same behavior continuously.

The semantics of an MTA with clock set $\mathcal{X}$ is given as a continuous-state MDP, where *states* are of the form (ℓ, η) where $\ell \in \mathit{Loc}$ and $\eta \in \mathcal{V}(\mathcal{X})$ is a clock valuation.

Definition 3: [Semantics] Let $\mathcal{M} = (Act, \mathcal{X}, Loc, \ell_0, E, \rightsquigarrow)$ be an MTA. The MDP associated with $\mathcal{M}$ is $\mathcal{D}(\mathcal{M}) = (Act, S, s_0, \mathcal{P})$ where $S = Loc \times \mathcal{V}(\mathcal{X})$; $s_0 = (\ell_0, \vec{0})$; and

- for each edge $\ell \overset{\alpha,g}{\rightsquigarrow} \zeta$ in $\mathcal{M}$ with $\zeta(X, \ell') = p > 0$,

$$\mathcal{P}((\ell, \eta), \alpha, A) := \int_0^\infty E(\ell) e^{-E(\ell)\tau} \cdot \mathbf{1}_g(\eta + \tau) \cdot p \; d\tau, \quad (1)$$

where $A = \{(\ell', \eta') \mid \exists \tau \in \mathbb{R}_{\geqslant 0}.\ \eta' = (\eta + \tau)[X := 0]$ and $\eta + \tau \models g\}$ and $\mathbf{1}_g(\cdot)$ is the characteristic function, i.e., $\mathbf{1}_g(\eta + \tau) = 1$ if $\eta + \tau \models g$; 0, otherwise.

We emphasize that MTA is a Markovian model with *decisions* (so it is *nondeterministic*) instead of a pure stochastic model like *deterministic* MTA (DMTA) studied in [9]. To state it alternatively, any DMTA coincides with an MTA with $Act = \{\alpha\}$. Moreover, for DMTA, the edge relation is defined as $\rightsquigarrow \subseteq Loc \times \mathcal{B}(\mathcal{X}) \times 2^{\mathcal{X}} \times Distr(Loc)$, while the MTA model allows for each set of transitions to reset their clocks differently. This has also been used in *probabilistic timed automata* (PTA, [14]). In this sense, our model can be considered as a continuous-time extension of PTA, due to the presence of exponential distributions. Any MTA where the exit rate of any location is zero is a PTA. Locally uniform *continuous-time* MDP*s* (CTMDPs) [16] (the exit rate of each location does *not* depend on actions) with *finite* state space are zero-clock MTAs (i.e., $\mathcal{X} = \varnothing$). We note that as in MDPs, it is assumed that action labels from any location in MTA are pairwise different.

Finite paths in MTA $\mathcal{M}$ are of the form $\ell_0 \xrightarrow{\alpha_0, t_0} \ell_1 \xrightarrow{\alpha_1, t_1} \cdots \xrightarrow{\alpha_{n-1}, t_{n-1}} \ell_n$, where for each edge $\ell_i \overset{\alpha_i, g_i}{\rightsquigarrow} \zeta_i$ of $\mathcal{M}$ with $\zeta_i(X_i, \ell_{i+1}) > 0$ ($\ell_i \in Loc$, $\alpha_i \in Act$, $t_i \in \mathbb{R}_{\geqslant 0}$, $X_i \subseteq \mathcal{X}$ and $0 \leqslant i < n$), we have that η_i is a valid clock valuation on *entering* location ℓ_i satisfying $\eta_0 = \vec{0}$, $(\eta_i + t_i) \models g_i$, and $\eta_{i+1} = (\eta_i + t_i)[X_i := 0]$. Let $Paths(\mathcal{M})$ (resp. $Paths_{\ell,\eta}(\mathcal{M})$) denote the set of finite paths (resp. starting in location ℓ with initial clock valuation η) in $\mathcal{M}$. Given a set of locations $G \subseteq Loc$, we write $RPaths_{\ell,\eta}(G) \subseteq Paths_{\ell,\eta}(\mathcal{M})$ as the set of paths reaching G from location ℓ and clock-valuation η.

b) Schedulers: The decision of which action to choose in an MTA is resolved by schedulers.[3] A scheduler must have enough "knowledge" to make such a decision which might be the current location and clock valuation (memoryless/positional schedulers), or the path from the initial to the current location (history dependent schedulers). We use $\mathcal{I}(\ell) \in Act$ to denote the set of actions enabled in location ℓ.

Definition 4: [Schedulers] Let $\mathcal{M} = (Act, \mathcal{X}, Loc, \ell_0, E, \rightsquigarrow)$ be an MTA. A scheduler for $\mathcal{M}$ is a measurable function $\theta : Paths(\mathcal{M}) \to Act$ such that for $n \in \mathbb{N}$,

$$\theta(\ell_0 \xrightarrow{\alpha_0, t_0} \ell_1 \xrightarrow{\alpha_1, t_1} \cdots \xrightarrow{\alpha_{n-1}, t_{n-1}} \ell_n) \in \mathcal{I}(\ell_n). \quad (2)$$

In the above definition we assume that the scheduler makes the decision as soon as a location is entered. These are called *early* schedulers [16]. In contrast, a *late* scheduler will decide which action to take upon leaving a location, i.e., besides the history it will consider also the waiting time. In this paper we will mainly consider early schedulers, although the theory can be adapted to late schedulers easily (see *Remark* 1).

Given any initial location ℓ_0, the initial clock valuation η, and scheduler θ, one obtains a *probability measure* $\Pr_{\ell_0,\eta,\theta}$ over a set of paths in a standard way (see [11]).

c) Maximum reachability: We are mainly interested in computing the maximum probability to reach a set of goal locations $G \subseteq Loc$ from the initial location ℓ_0.

Definition 5: Let $\mathcal{M} = (Act, \mathcal{X}, Loc, \ell_0, E, \rightsquigarrow)$ be an MTA and $Sched$ the set of all schedulers. The maximum probability to reach a set of goal locations $G \subseteq Loc$ from a location ℓ and clock valuation η is the function $p^{\mathcal{M},G}_{\max} : Loc \times \mathcal{V}(\mathcal{X}) \to [0, 1]$ defined as

$$p^{\mathcal{M},G}_{\max}(\ell, \eta) = \sup_{\theta \in Sched} \Pr_{\ell,\eta,\theta}(RPaths_{\ell,\eta}(G)).$$

In words, the maximum probability is the maximal one among all the schedulers.

The next theorem says that for MTA and maximum reachability probabilities, it suffices to consider *positional* schedulers instead of history dependent schedulers defined in Def. 4, i.e., the decision depends only on the current location and clock valuation.

Theorem 1: [Reachability in MTA] Let $\mathcal{M} = (Act, \mathcal{X}, Loc, \ell_0, E, \rightsquigarrow)$ be an MTA and $G \subseteq Loc$ a set of goal locations. The maximum reachability probability $p^{\mathcal{M},G}_{\max}$ is the least fixpoint of the integral operator $\mathcal{F} : \big(Loc \times \mathcal{V}(\mathcal{X}) \to [0, 1]\big) \to \big(Loc \times \mathcal{V}(\mathcal{X}) \to [0, 1]\big)$, where for the given function $\Pr : Loc \times \mathcal{V}(\mathcal{X}) \to [0, 1]$, location $\ell \in Loc$ and clock valuation η, $\mathcal{F}(\Pr)(\ell, \eta) = 1$ if $\ell \in G$, and $\mathcal{F}(\Pr)(\ell, \eta) = 0$ if $RPaths_{\ell,\eta}(G) = \varnothing$. For all other $\ell \notin G$ we have $\mathcal{F}(\Pr)(\ell, \eta) =$

$$\max_{\alpha \in \mathcal{I}(\ell)} \left\{ \int_0^\infty E(\ell) e^{-E(\ell)\tau} \cdot \sum_{\ell \overset{\alpha,g}{\underset{p,X}{\longmapsto}} \ell'} \mathbf{1}_g(\eta + \tau) \cdot p \cdot \Pr(\ell', \eta') \; d\tau \right\}, \quad (3)$$

where transition $\ell \overset{\alpha,g}{\underset{p,X}{\longmapsto}} \ell'$ is defined by transition $\ell \overset{\alpha,g}{\rightsquigarrow} \zeta$, $\zeta(X, \ell') = p$ and $\eta' = (\eta + \tau)[X := 0]$.

Remark 1: One can transform Thm. 1 to deal with the class of *late* schedulers. This is obtained by moving the "max" term inside the integral of (3), i.e. $\mathcal{F}(\Pr)(\ell, \eta) =$

$$\int_0^\infty E(\ell) e^{-E(\ell)\tau} \cdot \max_{\alpha \in \mathcal{I}(\ell)} \left\{ \sum_{\ell \overset{\alpha,g}{\underset{p,X}{\longmapsto}} \ell'} \mathbf{1}_g(\eta + \tau) \cdot p \cdot \Pr(\ell', \eta') \right\} d\tau.$$

C. Region construction for MTA

A main step in computing maximum reachability probability or the least fixpoint of the operator defined in Thm. 1 is to apply the *region construction* [1] in a similar way as for standard TA. Formally, a region is an equivalence class under $\cong$, an equivalence relation on clock valuations, which can be characterized by a specific form of a clock constraint. Let c_{x_i} be the largest constant with which $x_i \in \mathcal{X}$ is compared in some guard in the MTA. Clock evaluations $\eta, \eta' \in \mathcal{V}(\mathcal{X})$ are *clock-equivalent*, denoted $\eta \cong \eta'$, if and only if either

[3] In control engineering, one tends to use another terminology, i.e., *controllers*, while in CS community, schedulers, adversaries, policies, strategies, *etc* are more common. We do *not* distinguish them in the current paper.

1) for any $x \in \mathcal{X}$ it holds: $\eta(x) > c_x$ and $\eta'(x) > c_x$, or
2) for any $x_i, x_j \in \mathcal{X}$ with $\eta(x_i), \eta'(x_i) \leqslant c_{x_i}$ and $\eta(x_j)$, $\eta'(x_j) \leqslant c_{x_j}$ it holds: $\eta(x_j) \leqslant \eta'(x_j)$ iff $\lfloor \eta(x_i) \rfloor = \lfloor \eta'(x_i) \rfloor$ and $\{\eta(x_i)\} \leqslant \{\eta'(x_i)\}$, where $\lfloor d \rfloor$ ($\{d\}$) is the integral (fractional) part of $d \in \mathbb{R}$.

This clock equivalence is coarser than the traditional definition [1] by merging the "non-delayable" regions (those with point constraints like "$x = 0$") into the "delayable" regions (those only with interval constraints like "$0 < y < 1$"). For instance, for $\mathcal{X} = \{x_1, x_2\}$, the non-delayable regions $(x_1 = 0, x_2 = 0)$, $(0 < x_1 < 1, x_2 = 0)$ and $(x_1 = 0, 0 < x_2 < 1)$ are merged with the delayable region $(0 < x_1 < 1, 0 < x_2 < 1)$ yielding $(0 \leqslant x_1 < 1, 0 \leqslant x_2 < 1)$. The reason for this slight change will become clear later. We define the boundary of a region Θ as $\partial\Theta = \overline{\Theta} \setminus \mathring{\Theta}$, where $\overline{\Theta}$ is the closure and $\mathring{\Theta}$ is the interior of Θ, respectively. For instance, a region $\Theta = (x_1 \leqslant 1, x_2 > 0)$ has its closure $\overline{\Theta} = (x_1 \leqslant 1, x_2 \geqslant 0)$, its interior $\mathring{\Theta} = (x_1 < 1, x_2 > 0)$ and its boundary $\partial\Theta = (x_1 = 1, x_2 = 2)$. Here Θ is viewed as a set of elements from $\mathcal{V}(\mathcal{X})$.

Let $\mathcal{R}e(\mathcal{X})$ be the set of regions over the set $\mathcal{X}$ of clocks. For $\Theta, \Theta' \in \mathcal{R}e(\mathcal{X})$, Θ' is the *successor region* of Θ if for all $\eta \models \Theta$ there exists $\delta \in \mathbb{R}_{>0}$ such that $\eta + \delta \models \Theta'$ and $\forall \delta' < \delta$. $\eta + \delta' \models \Theta \vee \Theta'$. The region Θ *satisfies* the guard g, denoted $\Theta \models g$, iff $\forall \eta \models \Theta$. $\eta \models g$. The *reset operation* on region Θ is defined as $\Theta[X := 0] := \big\{\eta[X := 0] \mid \eta \models \Theta\big\}$.

Notation: Given a tuple $v = (v_1, \cdots, v_n)$ with n components, by $v\lfloor_k$ we denote the k-th component of v. In particular, for a node $v = (\ell, \Theta)$, $v\lfloor_1$ returns location ℓ, while $v\lfloor_2$ returns the associated region Θ.

Definition 6: [Region graph of MTA] The *region graph* of MTA $\mathcal{M} = (\mathit{Act}, \mathcal{X}, \mathit{Loc}, \ell_0, E, \rightsquigarrow)$ with the set of goal locations $G \subseteq \mathit{Loc}$ is $\mathcal{G}(\mathcal{M}) = (\mathit{Act}, V, v_0, \Lambda, \hookrightarrow)$, where

- $V = \mathit{Loc} \times \mathcal{R}e(\mathcal{X})$ is a finite set of *vertices* with *initial vertex* $v_0 = (\ell_0, \Theta_0)$, where Θ_0 is the initial region such that $\vec{0} \in \Theta_0$;
- $\Lambda : V \to \mathbb{R}_{\geqslant 0}$ is the *exit rate function* where:

$$\Lambda(v) \;=\; \begin{cases} E(v\lfloor_1) & \text{if } v \overset{\alpha,p,X}{\hookrightarrow} v' \text{ for some } v' \in V \\ 0 & \text{otherwise.} \end{cases}$$

- $\hookrightarrow \;\subseteq\; V \times \big(\big(\mathit{Act} \times [0,1] \times 2^{\mathcal{X}}\big) \cup \{\delta\}\big) \times V$ is the *transition* (*edge*) *relation*, such that:
  - $v \overset{\delta}{\hookrightarrow} v'$ if $v\lfloor_1 = v'\lfloor_1$, and $v'\lfloor_2$ is the successor region of $v\lfloor_2$;
  - $v \overset{\alpha,p,X}{\hookrightarrow} v'$ if $v\lfloor_1 \overset{\alpha,g}{\underset{p,X}{\longmapsto}} v'\lfloor_1$ with $v\lfloor_2 \models g$, and $v\lfloor_2[X := 0] = v'\lfloor_2$.

We also define $\mathbb{G} = \big\{v \in V \mid v\lfloor_1 \in G\big\}$ to be the set of *goal vertices* in $\mathcal{G}(\mathcal{M})$.

Any vertex in the region graph is a pair consisting of a location and a region. For a vertex $v \in V$ and clock valuation $\eta \in \mathcal{V}(\mathcal{X})$ we define the boundary function $\flat(v, \eta) = \inf\{\delta \mid \eta + \delta \in \partial v\lfloor_2\}$, which is the minimum time (if it exists) to "hit" the boundary of the region corresponding to vertex v starting from a clock valuation η. Edges of the form $v \overset{\delta}{\hookrightarrow} v'$ are called *delay edges (jump)*, whereas those of the form $v \overset{\alpha,p,X}{\hookrightarrow} v'$ are called *Markovian edges (jump)*. Note that Markovian edges emanating from a vertex corresponding to a non-delayable region do *not* contribute to the reachability probability. The waiting time in such vertex is always zero. Therefore, we can safely remove all the Markovian edges emanating from vertices with non-delayable regions and combine each such non-delayable region with its unique delayable (direct) successor. In the sequel, by slight abuse of notation, we refer to this *simplified region graph* as $\mathcal{G}(\mathcal{M})$. Note that then $v\lfloor_2[X := 0] \subseteq v'\lfloor_2$ in the last item of Def. 6. An example region graph is shown in Fig. 4.

Example 2: For the MTA $\mathcal{M}$ in Fig. 3, the reachable part (forward reachable from the initial vertex and backward reachable from the accepting vertices) of the region graph $\mathcal{G}(\mathcal{M})$ is shown in Fig. 4.

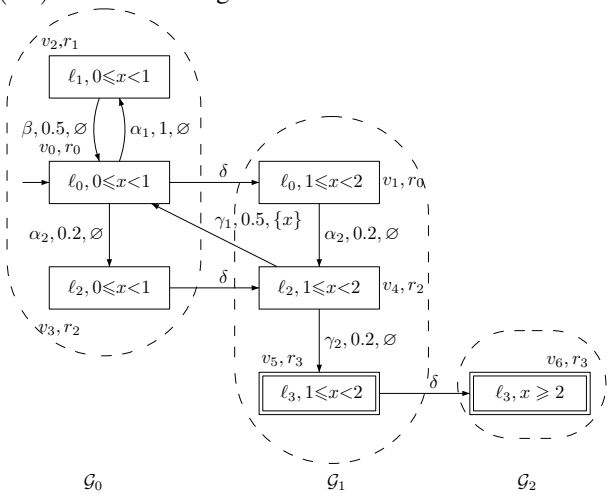

Fig. 4. Reachable region graph $\mathcal{G}(\mathcal{M})$

Now we define a system of integral equations on the region graph $\mathcal{G}(\mathcal{M})$ which will help compute the maximum reachability probability from Thm. 1.

Definition 7: Given the region graph $\mathcal{G}(\mathcal{M}) = (\mathit{Act}, V, v_0, \Lambda, \hookrightarrow)$ of the MTA $\mathcal{M} = (\mathit{Act}, \mathit{Loc}, \mathcal{X}, \ell_0, E, \rightsquigarrow)$ and the set of goal vertices $\mathbb{G}$, for the function $\mathit{Prob}_v(\eta) : V \times \mathcal{V}(\mathcal{X}) \to [0,1]$ let the operator $\widetilde{\mathcal{F}} : (V \times \mathcal{V}(\mathcal{X}) \to [0,1]) \to (V \times \mathcal{V}(\mathcal{X}) \to [0,1])$ be defined as $\widetilde{\mathcal{F}}(\mathit{Prob}_v(\eta)) = 1$ if $v \in \mathbb{G}$; $\widetilde{\mathcal{F}}(\mathit{Prob}_v(\eta)) = 0$ if $v \notin \mathbb{G}$ and $\mathbb{G}$ cannot be reached from v; and for $v \notin \mathbb{G}$ we have

$$\widetilde{\mathcal{F}}(\mathit{Prob}_v(\eta)) = \mathit{Prob}_{v,\delta}(\eta) + \max_{\alpha \in \mathcal{I}(v\lfloor_1)} \mathit{Prob}_{v,\alpha}(\eta),$$

$$\mathit{Prob}_{v,\alpha}(\eta) = \int_0^{\flat(v,\eta)} \Lambda(v) \cdot e^{-\Lambda(v)\tau} \cdot \sum_{v \overset{\alpha,p,X}{\hookrightarrow} v'} p \cdot \mathit{Prob}_{v'}\big((\eta + \tau)[X := 0]\big)\, d\tau,$$

$$\mathit{Prob}_{v,\delta}(\eta) = e^{-\Lambda(v)\flat(v,\eta)} \cdot \mathit{Prob}_{v'}\big(\eta + \flat(v,\eta)\big),$$

where $\mathit{Prob}_{v,\alpha}(\eta)$ denotes the probability to reach $\mathbb{G}$ by taking a Markovian jump and $\mathit{Prob}_{v,\delta}(\eta)$ the probability to reach $\mathbb{G}$ through vertex v' by taking the delay jump $v \overset{\delta}{\hookrightarrow} v'$.

Theorem 2: Let $\mathcal{M} = (\mathit{Act}, \mathit{Loc}, \mathcal{X}, \ell_0, E, \rightsquigarrow)$ be an MTA with the set of goal locations G and $\mathit{Prob}_v(\eta)$ be the least fixpoint of the operator $\widetilde{\mathcal{F}}$, then for $v\lfloor_1 = \ell$ we have

$$\mathit{Prob}_v(\eta) = p^{\mathcal{M},G}_{\max}(\ell, \eta).$$

Based on this theorem, we will focus on the efficient computation of maximum time-bounded reachability probabilities in region graphs (instead of in MTAs) in the following section.

III. Time-Bounded Reachability

In this section, we concentrate on maximizing *time-bounded reachability probabilities* in a region graph $\mathcal{G}(\mathcal{M}) = (\mathit{Act}, V, v_0, \Lambda, \hookrightarrow)$, i.e., given a set $\mathbb{G}$ of goal vertices and time bound $T \in \mathbb{N}$, we are interested in maximizing the probability to reach $\mathbb{G}$ *within T time units*. To this end, we introduce a *fresh* clock t, denoting the *global* time which is initialized to zero and never reset. To distinguish the role of t, we write the state of $\mathcal{G}(\mathcal{M})$ as (v, η, t). As in this case time bounds to reach $\mathbb{G}$ have to be considered, $t \leqslant T$ should be added to each vertex of $\mathcal{G}(\mathcal{M})$. Formally, by notation abuse we overload $\flat(v, \eta, t)$ to be $\min\{\flat(v, \eta), T - t\}$, i.e. the minimum time to hit the boundary $\partial v\lfloor_2$ *at time* $t \leqslant T$. For instance, let $\partial v\lfloor_2$ be $x = 1 \wedge y = 2$, $T = 100$, and suppose $\eta(x) = 0.5$, $\eta(y) = 1.7$. If $t = 2.5$, then $\flat(v, \eta, t) = \min\{1 - 0.5, 2 - 1.7, 100 - 2.5\} = 0.3$; if $t = 99.9$, then $\flat(v, \eta, t) = \min\{1 - 0.5, 2 - 1.7, 100 - 99.9\} = 0.1$;

The following Bellman (dynamic programming) equations derived from Def. 7 play an essential role in solving the time-bounded reachability problem. Let $P(v, \eta, t)$ be the maximum probability at time t for state (v, η) to reach $\mathbb{G}$ within time bound T. $P(v, \eta, t) = 1$ if $v \in \mathbb{G}$ and $t \leqslant T$; 0 if $t > T$ or $\mathbb{G}$ is *not* reachable from v; and otherwise $P(v, \eta, t) =$

$$\underbrace{\max_{\alpha \in \mathcal{I}(v)} \left\{ \sum_{v \overset{\alpha,p,X}{\hookrightarrow} v'} \int_0^{\flat(v,\eta,t)} \underbrace{\Lambda(v) e^{-\Lambda(v)\tau} p}_{(\star)} \cdot P(v', \eta', t+\tau)\ d\tau \right\}}_{\text{(I)}}$$
$$+ \underbrace{\underbrace{e^{-\Lambda(v)\flat(v,\eta,t)}}_{(\star\star)} P\Big(v'',\ \eta + \flat(v, \eta, t),\ t + \flat(v, \eta, t)\Big)}_{\text{(II)}}, \qquad (4)$$

where $\mathcal{I}(v)$ is the set of actions enabled in v, $\eta' = (\eta + \tau)[X := 0]$ and $v \overset{\delta}{\hookrightarrow} v''$, where v'' is the time successor of v. Term (I) represents the maximum reachability probability (among all enabled actions) by taking a Markovian jump $v \overset{\alpha,p,X}{\hookrightarrow} v'$ and (II) represents the probability of taking the boundary jump $v \overset{\delta}{\hookrightarrow} v''$. Note that $(\star)$ is the density function of taking $v \overset{\alpha,p,X}{\hookrightarrow} v'$ at time τ and $(\star\star)$ is the probability not to leave location v within $\flat(v, \eta, t)$ time units.

We will now provide two ways to solve (4): one by discretization (4) and the other based on the Hamilton-Jacobi-Bellman equation [13].

A. Discretization

Our first approach is to *discretize the continuous variables* in the Bellman equation. Using a discretization step $h = \frac{1}{N}$ ($N \in \mathbb{N}_{>0}$), the aim is to obtain a *finite-state* MDP $\mathcal{D}(\mathcal{M})$ from $\mathcal{G}(\mathcal{M})$. For this MDP, a similar Bellman equation can be derived and solved efficiently e.g. by value iteration [3]. Intuitively, h is the length of time in which a *single* Markovian jump takes place from a given location. By h, each Markovian jump in $\mathcal{G}(\mathcal{M})$ can be approximated by a Markovian jump which only takes place at time points $\{0, h, \ldots, NT\}$. This gives rise to an MDP:

Definition 8: Given $\mathcal{G}(\mathcal{M}) = (\mathit{Act}, V, v_0, \Lambda, \hookrightarrow)$ and discretization step $h = \frac{1}{N}$ ($N \in \mathbb{N}_{>0}$), the MDP $\mathcal{D}_h(\mathcal{M}) = (\mathit{Act} \cup \{\bot\}, S, s_0, \mathcal{P})$ is defined as follows:

- $S = \{(v, \eta, t) \mid v \in V \wedge \eta \in v\lfloor_2 \wedge t \leqslant T\}$;
- $s_0 = (v_0, \vec{0}, 0)$;
- $\bot$ is a *fresh* action encoding the "delay" in $\mathcal{G}(\mathcal{M})$;

For each $(v, \eta, t) \in S$ we distinguish three cases:

(i) If $h < \flat(v, \eta, t)$ and $v \overset{\alpha,p,X}{\hookrightarrow} v'$ then

$$\mathcal{P}\big((v, \eta, t), \alpha, (v', \eta[X := 0], t)\big) = p \cdot (1 - e^{-\Lambda(v)h});$$
$$\mathcal{P}\big((v, \eta, t), \alpha, (v, \eta + h, t + h)\big) = e^{-\Lambda(v)h};$$

(ii) If $h < \flat(v, \eta, t)$ and $\mathit{Act}(v) = \varnothing$ then

$$\mathcal{P}\big((v, \eta, t), \bot, (v, \eta + h, t + h)\big) = e^{-\Lambda(v)h};$$

(iii) If $h \geqslant \flat(v, \eta, t)$ and $v \overset{\delta}{\hookrightarrow} v'$ then

$$\mathcal{P}\big((v, \eta, t), \bot, (v', \eta + \flat(v, \eta, t), t + \flat(v, \eta, t))\big) = e^{-\Lambda(v)\flat(v,\eta,t)}.$$

Each state in the MDP $\mathcal{D}_h(\mathcal{M})$ has an outgoing transition of type (i), (ii) or (iii).

Example 3: Fig. 5 depicts the first half (till $2h$) of the reachable part of the MDP $\mathcal{D}_h(\mathcal{M})$ for the region graph $\mathcal{G}(\mathcal{M})$ in Fig. 4 with the step size $h = \frac{1}{2}$ and time bound $T = 2$. This means that in the MDP, the goal state(s) should be reached within 4 steps. Therefore, the second half ($3h$ and $4h$) is not necessary anyway. We add (i), (ii) and (iii) onto the edge labels to indicate which type of transition it is.

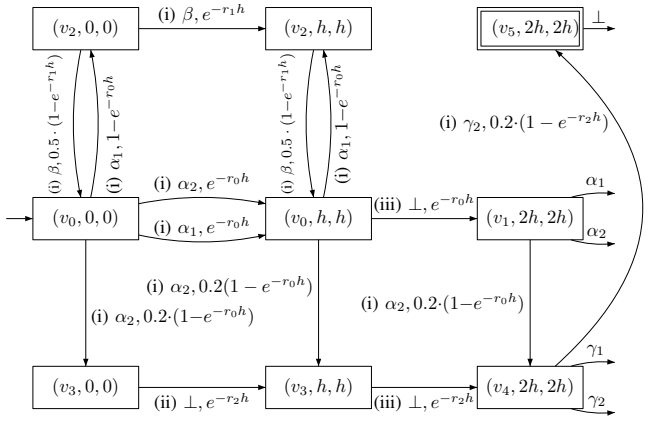

Fig. 5. MDP obtained from the region graph with discretization step h.

Let $Y(v, \eta, t)$ be the maximum reachability probability in the MDP $\mathcal{D}_h(\mathcal{M})$. Then $Y(v, \eta, t) =$

$$\begin{cases} e^{-\Lambda(v)\flat(v,\eta,t)} Y\big(v'',\ \eta + \flat(v, \eta, t),\ t + \flat(v, \eta, t)\big), \text{if } h \geqslant \flat(v, \eta, t) \\ \max\limits_{\alpha \in \mathcal{I}(v)} \Big\{ \sum\limits_{v \overset{\alpha,p,X}{\hookrightarrow} v'} (1 - e^{-\Lambda(v)h}) p \cdot Y(v', \widetilde{\eta}', t) \Big\} + e^{-\Lambda(v)h} Y(v, \eta + h, t + h), o/w, \end{cases} \qquad (5)$$

where $\widetilde{\eta}' = \eta[X := 0]$ and $v \overset{\delta}{\hookrightarrow} v''$.

By the discretization step h, the regions of $\mathcal{G}(\mathcal{M})$ contain finitely many points (one point is a state in the MDP). To be more exact, each region has maximally h points for each clock. Together with the fact that there are only finitely many regions of interest in $\mathcal{G}(\mathcal{M})$, the number of states in the MDP is finite.

Theorem 3: [Error bound] For any state (v, η), time bound T, discretization step $h = \frac{1}{N}$ and $\lambda = \max_{v \in V}\{\Lambda(v)\}$ which is the maximum rate of all exponential distributions appearing in the region graph:

$$\sup_{t \in [0,T]} |P(v,\eta,t) - Y(v,\eta,t)| \leq (1 - e^{-\lambda h})(1 - e^{-\lambda T}).$$

B. Hamilton-Jacobi-Bellman Equations

As in traditional control theory [3], the dynamic programming principles lead to a first-order integro-differential equation, which is the *Hamilton-Jacobi-Bellman* (HJB) *partial differential equation* (PDE).

Given the region graph $\mathcal{G}(\mathcal{M}) = (Act, V, v_0, \Lambda, \hookrightarrow)$ let $f(v,\eta,t) := P(v,\eta,t)$ be the maximum time-bounded reachability probability at time t in $\mathcal{G}(\mathcal{M})$. For every $v \in V$ and $\eta \in \mathcal{V}(\mathcal{X})$ and $t \leqslant T$ let $f(v,\eta,t)$ be given as follows:

$$\frac{\partial f(v,\eta,t)}{\partial t} + \sum_{i=1}^{|\mathcal{X}|} \frac{\partial f(v,\eta,t)}{\partial \eta^{(i)}} =$$

$$\max_{\alpha \in \mathcal{I}(v)} \left\{ \Lambda(v) \cdot \sum_{v \overset{\alpha,X,p}{\hookrightarrow} v'} p\left(f(v,\eta,t) - f\left(v', \eta[X := 0], t\right)\right) \right\},$$

where $\eta^{(i)}$ is the i'th clock variable. The initial conditions of the above PDE are $f(v,\eta,T) = \mathbf{1}_{\mathbb{G}}(v,\eta)$ for any $v \in V$ and $\eta \in v\lfloor_2$. Moreover, for every $\eta \in \partial v\lfloor_2$ and transition $v \overset{\delta}{\hookrightarrow} v'$, the boundary conditions take the form $f(v,\eta,t) = f(v',\eta,t)$.

Several methods can be used to solve the above HJB equation, e.g., the finite volume method [19] or the time and state space discretization technique [8]. Note that for zero-clock MTA, i.e., CTMDPs, we obtain a system of ODEs instead of PDEs; for details see [11].

IV. Conclusion

We have defined an extension of timed automata with exponentially distributed durations on locations. We constructed region graphs of such automata based on which the time-bounded reachability problems were investigated. We proposed Bellman equations to characterize the probability, and presented two approaches to solve these equations, namely, by discritization and a reduction to HJB PDEs.

Another interesting question is the *time-unbounded* reachability problem, i.e., the deadline of reaching the goal locations is absent. We refer the readers to [12], [11] for the solution of this question. In a nutshell, we show that, for single-clock MTA, they can be characterized as the solution of a system of linear equations whose coefficients are maximum reachability probabilities in CTMDPs, i.e., zero-clock MTA. Moreover, for general MTA, Bellman equations, which are a variant of Eq. (4), are proposed to characterize the probability. We remark that in this paper only the locally uniform model (i.e., the exit rate of each location solely depends on the location instead of actions) was addressed, however the general case can be treated without any difficulty.

Many future works remain to be done. For example, we plan to extend the MTA model to rewards; also one can consider more general variables than clocks, resulting in Markovian *hybrid* automata.

References

[1] R. Alur and D. L. Dill. A theory of timed automata. *Theor. Comput. Sci.*, 126(2):183–235, 1994.

[2] C. Baier, H. Hermanns, J.-P. Katoen, and B. R. H. M. Haverkort. Efficient computation of time-bounded reachability probabilities in uniform continuous-time Markov decision processes. *Theor. Comput. Sci.*, 345(1):2–26, 2005.

[3] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 1995.

[4] B. Barbot, T. Chen, T. Han, J.-P. Katoen, and A. Mereacre. Efficient CTMC Model Checking of Linear Real-Time Objectives. In *TACAS'11*, LNCS 6605, pp. 128-142. Springer, 2011.

[5] P. Bouyer and V. Forejt. Reachability in stochastic timed games. In *ICALP (2)*, LNCS 5556, pp. 103–114. Springer, 2009.

[6] T. Brázdil, V. Forejt, J. Krcal, J. Kretínský, and A. Kucera. Continuous-time stochastic games with time-bounded reachability. In *FSTTCS*, pp. 61–72, 2009.

[7] T. Brázdil, J. Krcál, J. Kretínský, A. Kucera, Vojtech Rehák: Stochastic real-time games with qualitative timed automata objectives. In *CONCUR*, LNCS 6269, pp. 207-221. Springer, 2010.

[8] F. Camilli. Approximation of integro-differential equations associated with piecewise deterministic process. *Optimal Control Applications and Methods*, 18(6):423–444, 1997.

[9] T. Chen, T. Han, J.-P. Katoen, and A. Mereacre. Quantitative model checking of continuous-time Markov chains against timed automata specifications. In *LICS*, pp. 309–318. IEEE Computer Society, 2009.

[10] T. Chen, T. Han, J.-P. Katoen, and A. Mereacre. Model checking of continuous-time Markov chains against timed automata specifications. *Logical Methods in Computer Science* 7(1): 1–34, 2011.

[11] T. Chen, T. Han, J.-P. Katoen, and A. Mereacre. Reachability probabilities in Markovian timed automata. Technical Report RR-11-02, OUCL, 2011. Available from `http://www.cs.ox.ac.uk/people/taolue.chen/pub-papers/cdc_full.pdf`.

[12] T. Chen, T. Han, J.-P. Katoen, and A. Mereacre. Observing Continuous-Time MDPs by 1-Clock Timed Automata. In *RP*, LNCS 6945, pp. 2-25. Springer, 2011.

[13] M. H. A. Davis. *Markov Models and Optimization*. Chapman and Hall, 1993.

[14] M. Z. Kwiatkowska, G. Norman, R. Segala, and J. Sproston. Automatic verification of real-time systems with discrete probability distributions. *Theor. Comput. Sci.*, 282(1):101–150, 2002.

[15] M. Neuts. *Matrix-Geometric solutiions in stochastic models; An algorithimic approach.* John Hopkins University Press, Baltimore, 1981.

[16] M. R. Neuhäußer, M. Stoelinga, and J.-P. Katoen. Delayed nondeterminism in continuous-time Markov decision processes. In *FOSSACS*, LNCS 5504, pp. 364–379. Springer, 2009.

[17] M. R. Neuhäußer and L. Zhang. Time-bounded reachability probabilities in continuous-time Markov decision processes. In *QEST*, pp. 209–218, 2010.

[18] M. L. Puterman. *Markov Decision Processes*. Wiley, 1994.

[19] S. Wang, L. S. Jennings, and K. L. Teo. Numerical solution of Hamilton-Jacobi-Bellman equations by an upwind finite volume method. *J. of Global Optimization*, 27(2-3):177–192, 2003.