Modeling and Verification of Probabilistic Systems

Joost-Pieter Katoen

Lehrstuhl für Informatik 2
Software Modeling and Verification Group

http://moves.rwth-aachen.de/teaching/ws-1819/movep18/

December 10, 2018

Overview

1. Recall: continuous-time Markov chains
2. Probability measure on CTMC paths
3. Reachability probabilities
   - Untimed reachability
   - Timed reachability
   - Reduction to transient analysis
   - Bisimulation and timed reachability
4. Summary

Continuous-time Markov chain

Continuous-time Markov chain

A CTMC is a tuple $(S, \mathbf{P}, r, \iota_{\text{init}}, AP, L)$ where

- $(S, \mathbf{P}, \iota_{\text{init}}, AP, L)$ is a DTMC, and
- $r : S \to \mathbb{R}_{>0}$, the exit-rate function

Let $\mathbf{R}(s, s') = \mathbf{P}(s, s') \cdot r(s)$ be the transition rate of transition (s, s')

Interpretation

- residence time in state s is exponentially distributed with rate $r(s)$.
- phrased alternatively, the average residence time of state s is $\frac{1}{r(s)}$.

CTMC semantics

Enabledness

The probability that transition $s \to s'$ is *enabled* in $[0, t]$ is $1 - e^{-\mathbf{R}(s,s') \cdot t}$.

State-to-state timed transition probability

The probability to *move* from non-absorbing s to s' in $[0, t]$ is:

$$\frac{\mathbf{R}(s, s')}{r(s)} \cdot \left(1 - e^{-r(s) \cdot t}\right).$$

Residence time distribution

The probability to *take some* outgoing transition from s in $[0, t]$ is:

$$\int_0^t r(s) \cdot e^{-r(s) \cdot x} \, dx \;=\; 1 - e^{-r(s) \cdot t}$$

Paths in a CTMC

Timed paths

Paths in CTMC $\mathcal{C}$ are maximal (i.e., infinite) paths of alternating states and time instants:

$$\pi \;=\; s_0 \xrightarrow{t_0} s_1 \xrightarrow{t_1} s_2 \cdots$$

such that $s_i \in S$ and $t_i \in \mathbb{R}_{>0}$. Let $Paths(\mathcal{C})$ be the set of paths in $\mathcal{C}$ and $Paths^*(\mathcal{C})$ the set of finite prefixes thereof.

Time instant t_i is the amount of time spent in state s_i.

Notations

- Let $\pi[i] := s_i$ denote the $(i{+}1)$-st state along the timed path π.
- Let $\pi\langle i\rangle := t_i$ the time spent in state s_i.
- Let $\pi@t$ be the state occupied in π at time $t \in \mathbb{R}_{\geqslant 0}$, i.e. $\pi@t := \pi[i]$ where i is the smallest index such that $\sum_{j=0}^{i} \pi\langle j\rangle \;>\; t$.

Overview

1. Recall: continuous-time Markov chains
2. Probability measure on CTMC paths
3. Reachability probabilities
   - Untimed reachability
   - Timed reachability
   - Reduction to transient analysis
   - Bisimulation and timed reachability
4. Summary

Paths and probabilities

To reason quantitatively about the behavior of a CTMC, we need to define a probability space over its paths.

Intuition

For a given state s in CTMC $\mathcal{C}$:

- Sample space := set of all interval-timed paths $s_0\, I_0 \ldots I_{k-1}\, s_k$ with $s = s_0$
- Events := sets of interval-timed paths starting in s
- Basic events := cylinder sets
- Cylinder set of finite interval-timed paths := set of all infinite timed paths with a prefix in the finite interval-timed path

Timed cylinder sets

Timed cylinder set

Let $s_0, \ldots, s_k \in S$ with $\mathbf{P}(s_i, s_{i+1}) > 0$ for $0 \leqslant i < k$ and $I_0, \ldots, I_{k-1}$ non-empty intervals in $\mathbb{R}_{\geqslant 0}$ with rational bounds. The *cylinder set* of $s_0\, I_0\, s_1\, I_1 \ldots I_{k-1}\, s_k$ is defined by:

$$\begin{aligned} Cyl(s_0, I_0, \ldots, I_{k-1}, s_k) \;\;=\;\; \big\{\, \pi \in Paths(\mathcal{C}) \mid\; & \forall 0 \leqslant i \leqslant k.\, \pi[i] = s_i \\ & \text{and } i < k \;\Rightarrow\; \pi\langle i\rangle \in I_i \,\big\} \end{aligned}$$

The cylinder set spanned by $s_0, I_0, \ldots, I_{k-1}, s_k$ thus consists of all infinite timed paths that have a prefix $\hat{\pi}$ that lies in $s_0, I_0, \ldots, I_{k-1}, s_k$. Cylinder sets serve as basic events of the smallest σ-algebra on $Paths(\mathcal{C})$.

σ-algebra over timed cylinders

The σ-algebra associated with CTMC $\mathcal{C}$ is the smallest σ-algebra $\mathcal{F}(Paths(s_0))$ that contains all cylinder sets $Cyl(s_0, I_0, \ldots, I_{k-1}, s_k)$ where $s_0 \ldots s_k$ is a path in the state graph of $\mathcal{C}$ (starting in s_0) and $I_0, \ldots, I_{k-1}$ range over all sequences of non-empty intervals in $\mathbb{R}_{\geqslant 0}$.

Probability measure on CTMCs

Cylinder set

The *cylinder set* $Cyl(s_0, I_0, \ldots, I_{k-1}, s_k)$ of $s_0\, I_0 \ldots I_{k-1}\, s_k$ is defined by:

$$\{\, \pi \in Paths(\mathcal{C}) \mid \forall 0 \leqslant i \leqslant k.\, \pi[i] = s_i \text{ and } i < k \;\Rightarrow\; \pi\langle i \rangle \in I_i \,\}$$

Probability measure

Pr is the unique *probability measure* on the σ-algebra $\mathcal{F}(Paths(s_0))$ defined by induction on k as follows: $Pr(Cyl(s_0)) \;=\; \iota_{\text{init}}(s_0)$ and for $k > 0$:

$$\begin{aligned} Pr(Cyl(s_0, I_0, \ldots, I_{k-1}, s_k)) \;&=\; Pr(Cyl(s_0, I_0, \ldots, I_{k-2}, s_{k-1}))\cdot \\ &\quad \int_{I_{k-1}} \mathbf{R}(s_{k-1}, s_k)\cdot e^{-r(s_{k-1})\cdot \tau}\, d\tau. \end{aligned}$$

Solving the integral

$$Pr(Cyl(s_0, I_0, \ldots, I_{k-2}, s_{k-1})) \cdot \mathbf{P}(s_{k-1}, s_k)\cdot \left(e^{-r(s_{k-1})\cdot \inf I_{k-1}} - e^{-r(s_{k-1})\cdot \sup I_{k-1}}\right).$$

Zeno theorem

Zeno path

Path $s_0 \xrightarrow{t_0} s_1 \xrightarrow{t_1} s_2 \xrightarrow{t_2} s_3 \ldots\ldots$ is called Zeno [1] if $\sum_i t_i$ converges.

Intuition

In case $\sum_i t_i$ does not diverge, the timed path represents an"unrealistic" computation where infinitely many transitions are taken in a finite amount of time. Example:

$$s_0 \xrightarrow{1} s_1 \xrightarrow{\frac{1}{2}} s_2 \xrightarrow{\frac{1}{4}} s_3 \ldots s_i \xrightarrow{\frac{1}{2^i}} s_{i+1} \ldots$$

In real-time systems, such executions are typically excluded from the analysis. Thanks to the following theorem, Zeno paths do not harm for CTMCs.

Zeno theorem

For all states s in any CTMC, $Pr\{\, \pi \in Paths(s) \mid \pi \text{ is Zeno} \,\} = 0$.

[1] Zeno of Elea (490–430 BC), philosopher, famed for his paradoxes.

Proof of Zeno theorem

Zeno theorem

For all states s in any CTMC, $Pr\{\, \pi \in Paths(s) \mid \pi \text{ is Zeno} \,\} = 0$.

Proof:

On the blackboard.

Overview

1 Recall: continuous-time Markov chains

2 Probability measure on CTMC paths

3 Reachability probabilities
- Untimed reachability
- Timed reachability
- Reduction to transient analysis
- Bisimulation and timed reachability

4 Summary

Reachability events

Let CTMC $\mathcal{C}$ with (possibly infinite) state space S.

(Simple) reachability

Eventually reach a state in $G \subseteq S$. Formally:

$$\Diamond G \;=\; \{\, \pi \in \mathit{Paths}(\mathcal{C}) \mid \exists i \in \mathbb{N}.\, \pi[i] \in G \,\}$$

Invariance, i.e., always stay in state in G:

$$\Box G \;=\; \{\, \pi \in \mathit{Paths}(\mathcal{C}) \mid \forall i \in \mathbb{N}.\, \pi[i] \in G \,\} \;=\; \overline{\Diamond \overline{G}}.$$

Constrained reachability

Or "reach-avoid" properties where states in $F \subseteq S$ are forbidden:

$$\overline{F} \,\mathrm{U}\, G \;=\; \{\, \pi \in \mathit{Paths}(\mathcal{C}) \mid \exists i \in \mathbb{N}.\, \pi[i] \in G \wedge \forall j < i.\, \pi[j] \notin F \,\}$$

Measurability

Measurability theorem

Events $\Diamond G$, $\Box G$, $\overline{F} \,\mathrm{U}\, G$, $\Box\Diamond G$ and $\Diamond\Box G$ are measurable on any CTMC.

Proof:

Consider $\Diamond G$. $\Diamond G$ is the union of all cylinders $\mathit{Cyl}(s_0, [0,\infty), \ldots, [0,\infty), s_n)$ where $s_0, \ldots, s_{n-1} \notin G$ and $s_n \in G$. As the set of state sequences $s_0 \ldots s_n$ is countable, $\Diamond G$ is a countable union of cylinders. Thus $\Diamond G$ is measurable. The proof for $\Box\Diamond G$ goes along similar lines, using the proof principle for DTMCs.

Reachability probabilities in finite CTMCs

Problem statement

Let $\mathcal{C}$ be a CTMC with finite state space S, $s \in S$ and $G \subseteq S$.

Aim: determine $\mathit{Pr}(s \models \Diamond G) \;=\; \mathit{Pr}_s(\Diamond G) \;=\; \mathit{Pr}_s\{\, \pi \in \mathit{Paths}(s) \mid \pi \models \Diamond G \,\}$

where Pr_s is the probability measure in $\mathcal{C}$ with single initial state s.

Characterisation of reachability probabilities

- Let variable $x_s = \mathit{Pr}(s \models \Diamond G)$ for any state s
  - if G is not reachable from s, then $x_s = 0$
  - if $s \in G$ then $x_s = 1$
- For any state $s \in \mathit{Pre}^*(G) \setminus G$:

$$x_s \;=\; \underbrace{\sum_{t \in S \setminus G} \mathbf{P}(s,t) \cdot x_t}_{\text{reach } G \text{ via } t \in S \setminus G} \;+\; \underbrace{\sum_{u \in G} \mathbf{P}(s,u)}_{\text{reach } G \text{ in one step}}$$

Verifying CTMCs

Verifying untimed properties

So, computing reachability probabilities is exactly the same as for DTMCs. The same holds for constrained reachability, persistence and repeated reachability. In fact, all PCTL and LTL formulas can be checked on the embedded DTMC $(S, \mathbf{P}, \iota_{\mathrm{init}}, AP, L)$ using the techniques described before in these lecture slides.

Justification:

As the above temporal logic formulas or events do not refer to elapsed time, it is not surprising that they can be checked on the embedded DTMC.

Timed reachability events

Let CTMC $\mathcal{C}$ with (possibly infinite) state space S.

(Simple) timed reachability

Eventually reach a state in $G \subseteq S$ in the interval I. Formally:

$$\Diamond^{I}\, G \;=\; \{\, \pi \in \mathit{Paths}(\mathcal{C}) \mid \exists t \in I.\, \pi@t \in G \,\}$$

Invariance, i.e., always stay in state in G in the interval I:

$$\Box^{I}\, G \;=\; \{\, \pi \in \mathit{Paths}(\mathcal{C}) \mid \forall t \in I.\, \pi@t \in G \,\} \;=\; \overline{\Diamond^{I}\, \overline{G}}.$$

Constrained timed reachability

Or "reach-avoid" properties where states in $F \subseteq S$ are forbidden:

$$\overline{F}\, \mathrm{U}^{I}\, G \;=\; \{\, \pi \in \mathit{Paths}(\mathcal{C}) \mid \exists t \in I.\, \pi@t \in G \wedge \forall d < t.\, \pi@d \notin F \,\}$$

Measurability

Measurability theorem

Events $\Diamond^{I}\, G$, $\Box^{I}\, G$, and $\overline{F}\, \mathrm{U}^{I}\, G$ are measurable on any CTMC.

Proof (sketch):

Consider $\Diamond^{I}\, G$ where $I = [0, t]$. $\Diamond^{\leqslant t} G$ is the union of $\mathit{Cyl}(s_0, I_0, \ldots, I_{n-1}, s_n)$ with $s_0, \ldots, s_{n-1} \notin G$, $s_n \in G$, and $\sup(I_0) + \ldots \sup(I_{n-1}) \leqslant t$. The set of state sequences $s_0 \ldots s_n$ is countable and the set of rational bounded intervals $I_0, \ldots, I_{n-1}$ is countable. Thus $\Diamond^{\leqslant t} G$ is a countable union of cylinders, and thus is measurable. The proof for the remaining case $\overline{F}\, \mathrm{U}^{I}\, G$ is similar and left as an exercise.

Timed reachability probabilities in finite CTMCs

Problem statement

Let $\mathcal{C}$ be a CTMC with finite state space S, $s \in S$, $t \in \mathbb{R}_{\geqslant 0}$ and $G \subseteq S$.

Aim: $Pr(s \models \Diamond^{\leqslant t}\, G) \;=\; Pr_s(\Diamond^{\leqslant t}\, G) \;=\; Pr_s\{\, \pi \in \mathit{Paths}(s) \mid \pi \models \Diamond^{\leqslant t}\, G \,\}$

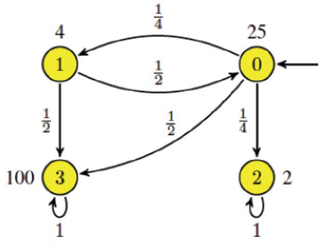
where Pr_s is the probability measure in $\mathcal{C}$ with single initial state s.

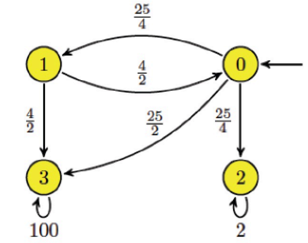
Characterisation of timed reachability probabilities

- Let function $x_s(t) = Pr(s \models \Diamond^{\leqslant t}\, G)$ for any state s
  - if G is not reachable from s, then $x_s(t) = 0$ for all t
  - if $s \in G$ then $x_s(t) = 1$ for all t
- For any state $s \in \mathit{Pre}^*(G) \setminus G$:

$$x_s(t) \;=\; \int_0^t \sum_{s' \in S} \underbrace{\mathbf{R}(s, s') \cdot e^{-r(s) \cdot x}}_{\text{probability to move to state } s' \text{ at time } x} \;\cdot\; \underbrace{x_{s'}(t{-}x)}_{\text{prob. to fulfill } \Diamond^{\leqslant t-x}\, G \text{ from } s'} \; dx$$

Timed reachability probabilities





Integral equations for $\Diamond^{\leqslant 10}\, 2$:

- $x_3(d) = 0$ and $x_2(d) = 1$ for all d
- $x_0(d) = \displaystyle\int_0^d {}^{25}\!/_{4} \cdot e^{-25 \cdot x} \cdot x_1(d{-}x) \;+\; {}^{25}\!/_{4} \cdot e^{-25 \cdot x} \cdot x_2(d{-}x)\; dx$
- $x_1(d) = \displaystyle\int_0^d {}^{4}\!/_{2} \cdot e^{-4 \cdot x} \cdot x_0(d{-}x) \;+\; {}^{4}\!/_{2} \cdot e^{-4 \cdot x} \cdot x_3(d{-}x)\; dx$

Reachability

Reachability probabilities in finite DTMCs and CTMCs

Can be obtained by solving a system of linear equations for which many efficient techniques exists.

Timed reachability probabilities in finite CTMCs

Can be obtained by solving a system of Volterra integral equations. This is in general a non-trivial issue, inefficient, and has several pitfalls such as numerical stability.

Solution

Reduce the problem of computing $Pr(s \models \Diamond^{\leqslant t} G)$ to an alternative problem for which well-known efficient techniques exist: computing transient probabilities (see previous lecture).

Timed reachability probabilities = transient probabilities

Aim

Compute $Pr(s \models \Diamond^{\leqslant t} G)$ in CTMC $\mathcal{C}$. Observe that once a path π reaches G within t time, then the remaining behaviour along π is not important. This suggests to make all states in G absorbing.

Let CTMC $\mathcal{C} = (S, \mathbf{P}, r, \iota_{\text{init}}, AP, L)$ and $G \subseteq S$. The CTMC $\mathcal{C}[G] = (S, \mathbf{P}_G, r, \iota_{\text{init}}, AP, L)$ with $\mathbf{P}_G(s,t) = \mathbf{P}(s,t)$ if $s \notin G$ and $\mathbf{P}_G(s,s) = 1$ if $s \in G$.

All outgoing transitions of $s \in G$ are replaced by a single self-loop at s.

Lemma

$$\underbrace{Pr(s \models \Diamond^{\leqslant t} G)}_{\text{timed reachability in } \mathcal{C}} = \underbrace{Pr(s \models \Diamond^{= t} G)}_{\text{timed reachability in } \mathcal{C}[G]} = \underbrace{\sum_{s' \in G} p_{s'}(t) \text{ with } \underline{p}(0) = \mathbf{1}_s}_{\text{transient prob. in } \mathcal{C}[G]}$$

Example

Constrained timed reachability probabilities

Problem statement

Let $\mathcal{C}$ be a CTMC with finite state space S, $s \in S$, $t \in \mathbb{R}_{\geqslant 0}$ and $G, F \subseteq S$.

Aim: $Pr(s \models \overline{F} \, \mathsf{U}^{\leqslant t} \, G) = Pr_s(\overline{F} \, \mathsf{U}^{\leqslant t} \, G) = Pr_s\{\pi \in Paths(s) \mid \pi \models \overline{F} \, \mathsf{U}^{\leqslant t} \, G\}$.

Characterisation of timed reachability probabilities

- Let function $x_s(t) = Pr(s \models \overline{F} \, \mathsf{U}^{\leqslant t} \, G)$ for any state s
  - if G is not reachable from s via $\overline{F}$, then $x_s(t) = 0$ for all t
  - if $s \in G$ then $x_s(t) = 1$ for all t
- For any state $s \in Pre^*(G) \setminus (F \cup G)$:

$$x_s(t) = \int_0^t \sum_{s' \in S} \underbrace{\mathbf{R}(s,s') \cdot e^{-r(s) \cdot x}}_{\substack{\text{probability to move to} \\ \text{state } s' \text{ at time } x}} \cdot \underbrace{x_{s'}(t-x)}_{\substack{\text{prob. to fulfill} \\ \overline{F} \, \mathsf{U}^{\leqslant t-x} \, G \text{ from } s'}} \, dx$$

Constrained timed reachability = transient probabilities

Aim

Compute $Pr(s \models \overline{F}\,\mathrm{U}^{\leqslant t}\, G)$ in CTMC $\mathcal{C}$. Observe (as before) that once a path π reaches G within time t via $\overline{F}$, then the remaining behaviour along π is not important. Now also observe that once $s \in F \setminus G$ is reached within time t, then the remaining behaviour along π is not important. This suggests to make all states in G and $F \setminus G$ absorbing.

Lemma

$$\underbrace{Pr(s \models \overline{F}\,\mathrm{U}^{\leqslant t}\, G)}_{\text{timed reachability in } \mathcal{C}} \;=\; \underbrace{Pr(s \models \Diamond^{=t} G)}_{\text{timed reachability in } \mathcal{C}[F \cup G]} \;=\; \underbrace{\sum_{s' \in G} p_{s'}(t) \text{ with } \underline{p}(0) = \mathbf{1}_s}_{\text{transient prob. in } \mathcal{C}[F \cup G]} .$$

Strong and weak bisimulation

Bisimulation preserves timed reachability events

Let $\mathcal{C}$ be a CTMC with state space S, $s, u \in S$, $t \in \mathbb{R}_{\geqslant 0}$ and $G, F \subseteq S$. Then:

1. $s \sim_m u$ implies $Pr(s \models \overline{F}\,\mathrm{U}^{\leqslant t}\, G) \;=\; Pr(u \models \overline{F}\,\mathrm{U}^{\leqslant t}\, G)$
2. $s \approx_m u$ implies $Pr(s \models \overline{F}\,\mathrm{U}^{\leqslant t}\, G) \;=\; Pr(u \models \overline{F}\,\mathrm{U}^{\leqslant t}\, G)$

provided F and G are closed under $\sim_m$ and $\approx_m$, respectively.

Proof:

Left as an exercise.

Example

Other Properties on CTMCs

- Expected time objectives

  Can be characterised as solution of set of linear equations

- Long-run average objectives
  1. Determine the limiting distribution in any terminal SCC
  2. Take weighted sum with reachability probabilities terminal SCCs

- Probabilistic timed CTL model checking

  recursive descent over parse tree

- Deterministic timed automata objectives
  1. Take product of the MC and the Zone automaton of the DTA[2]
  2. Determine the probability to reach an accepting zone

[2] This yields a piecewise deterministic Markov process.

Overview

1. Recall: continuous-time Markov chains
2. Probability measure on CTMC paths
3. Reachability probabilities
   - Untimed reachability
   - Timed reachability
   - Reduction to transient analysis
   - Bisimulation and timed reachability
4. Summary

Summary

Main points

- Cylinder sets in a CTMC are paths that share interval-timed path prefixes.
- Reachability, persistence and repeated reachability can be checked as on DTMCs.
- Timed reachability probabilities can be characterised as Volterra integral equation system.
- Computing timed reachability probabilities can be reduced to transient probabilities.
- Weak and strong bisimilarity preserve timed reachability probabilities.