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 4, 2018

Overview

Negative exponential distribution

Density of exponential distribution

The density of an *exponentially distributed* r.v. Y with *rate* $\lambda \in \mathbb{R}_{>0}$ is:

$$f_Y(x) = \lambda \cdot e^{-\lambda \cdot x} \quad \text{for } x > 0 \quad \text{and } f_Y(x) = 0 \text{ otherwise}$$

The cumulative distribution of r.v. Y with rate $\lambda \in \mathbb{R}_{>0}$ is:

$$F_Y(d) = \int_0^d \lambda \cdot e^{-\lambda \cdot x} \, dx = [-e^{-\lambda \cdot x}]_0^d = 1 - e^{-\lambda \cdot d}.$$

The rate $\lambda \in \mathbb{R}_{>0}$ uniquely determines an exponential distribution.

Variance and expectation

Let r.v. Y be exponentially distributed with rate $\lambda \in \mathbb{R}_{>0}$. Then:

- Expectation $E[Y] = \int_0^\infty x \cdot \lambda \cdot e^{-\lambda \cdot x} \, dx = \frac{1}{\lambda}$
- Variance $Var[Y] = \int_0^\infty (x - E[X])^2 \lambda \cdot e^{-\lambda \cdot x} \, dx = \frac{1}{\lambda^2}$

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:

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

Overview

Transient distribution of a CTMC

Transient state probability

Let $X(t)$ denote the state of a CTMC at time $t \in \mathbb{R}_{\geqslant 0}$. The probability to be in state s at time t is defined by:

$$\begin{aligned} p_s(t) &= Pr\{\, X(t) = s \,\} \\ &= \sum_{s' \in S} Pr\{\, X(0) = s' \,\} \cdot Pr\{\, X(t) = s \mid X(0) = s' \,\} \end{aligned}$$

Theorem: transient distribution as linear differential equation

The transient probability vector $\underline{p}(t) = (p_{s_1}(t), \ldots, p_{s_k}(t))$ satisfies:

$$\underline{p}'(t) \;=\; \underline{p}(t) \cdot (\mathbf{R} - \mathbf{r}) \quad \text{given} \quad \underline{p}(0)$$

where $\mathbf{r}$ is the diagonal matrix of vector $\underline{r}$.

Transient distribution theorem

Theorem: transient distribution as linear differential equation

The transient probability vector $\underline{p}(t) = (p_{s_1}(t), \ldots, p_{s_k}(t))$ satisfies:

$$\underline{p}'(t) \;=\; \underline{p}(t) \cdot (\mathbf{R} - \mathbf{r}) \quad \text{given} \quad \underline{p}(0)$$

where $\mathbf{r}$ is the diagonal matrix of vector $\underline{r}$.

Proof:

On the blackboard.

Computing transient probabilities

The transient probability vector $\underline{p}(t) = (p_{s_1}(t), \ldots, p_{s_k}(t))$ satisfies:

$$\underline{p}'(t) = \underline{p}(t) \cdot (\mathbf{R} - \mathbf{r}) \quad \text{given} \quad \underline{p}(0).$$

Solution using standard knowledge yields: $\underline{p}(t) \;=\; \underline{p}(0) \cdot e^{(\mathbf{R}-\mathbf{r}) \cdot t}$.

Computing a matrix exponential

First attempt: use Taylor-Maclaurin expansion. This yields

$$\underline{p}(t) \;=\; \underline{p}(0) \cdot e^{(\mathbf{R}-\mathbf{r}) \cdot t} \;=\; \underline{p}(0) \cdot \sum_{i=0}^{\infty} \frac{((\mathbf{R}-\mathbf{r}) \cdot t)^i}{i!}$$

But: numerical instability due to fill-in of $(\mathbf{R}-\mathbf{r})^i$ in presence of positive and negative entries in the matrix $\mathbf{R}-\mathbf{r}$.

Overview

Uniformization

Let CTMC $\mathcal{C} = (S, \mathbf{P}, r, \iota_{\text{init}}, AP, L)$ with S finite.

Uniform CTMC

CTMC $\mathcal{C}$ is uniform if $r(s) = r$ for all $s \in S$ for some $r \in \mathbb{R}_{>0}$.

Uniformization [Gross and Miller, 1984]

Let $r \in \mathbb{R}_{>0}$ such that $r \geqslant max_{s \in S}\ r(s)$. Then $unif(r, \mathcal{C})$ is the tuple $(S, \overline{\mathbf{P}}, \overline{r}, \iota_{\text{init}}, AP, L)$ with $\overline{r}(s) = r$ for all $s \in S$, and:

$$\overline{\mathbf{P}}(s, s') = \frac{r(s)}{r} \cdot \mathbf{P}(s, s') \text{ if } s' \neq s \quad \text{and} \quad \overline{\mathbf{P}}(s, s) = \frac{r(s)}{r} \cdot \mathbf{P}(s, s) + 1 - \frac{r(s)}{r}.$$

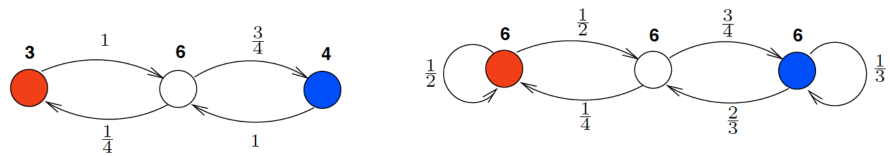
It follows that $\overline{\mathbf{P}}$ is a stochastic matrix and $unif(r, \mathcal{C})$ is a CTMC.

Uniformization: example

Uniformization

Let $r \in \mathbb{R}_{>0}$ such that $r \geqslant max_{s \in S}\ r(s)$. Then $unif(r, \mathcal{C}) = (S, \overline{\mathbf{P}}, \overline{r}, \iota_{\text{init}}, AP, L)$ with $\overline{r}(s) = r$ for all $s \in S$, and:

$$\overline{\mathbf{P}}(s, s') = \frac{r(s)}{r} \cdot \mathbf{P}(s, s') \text{ if } s' \neq s \quad \text{and} \quad \overline{\mathbf{P}}(s, s) = \frac{r(s)}{r} \cdot \mathbf{P}(s, s) + 1 - \frac{r(s)}{r}.$$



CTMC $\mathcal{C}$ and its uniformized counterpart $unif(6, \mathcal{C})$

Uniformization: intuition

Uniformization

Let $r \in \mathbb{R}_{>0}$ such that $r \geqslant max_{s \in S}\ r(s)$. Then $unif(r, \mathcal{C}) = (S, \overline{\mathbf{P}}, \overline{r}, \iota_{\text{init}}, AP, L)$ with $\overline{r}(s) = r$ for all $s \in S$, and:

$$\overline{\mathbf{P}}(s, s') = \frac{r(s)}{r} \cdot \mathbf{P}(s, s') \text{ if } s' \neq s \quad \text{and} \quad \overline{\mathbf{P}}(s, s) = \frac{r(s)}{r} \cdot \mathbf{P}(s, s) + 1 - \frac{r(s)}{r}.$$

Intuition

- Fix all exit rates to (at least) the maximal exit rate r occurring in CTMC $\mathcal{C}$.
- Thus, $\frac{1}{r}$ is the shortest mean residence time in the CTMC $\mathcal{C}$.
- Then normalize the residence time of all states with respect to r as follows:
  1. replace an average residence time $\frac{1}{r(s)}$ by a shorter (or equal) one, $\frac{1}{r}$
  2. decrease the transition probabilities by a factor $\frac{r(s)}{r}$, and
  3. increase the self-loop probability by a factor $\frac{r-r(s)}{r}$

  That is, slow down state s whenever $r(s) < r$.

Overview

Strong bisimulation on DTMCs

Probabilistic bisimulation [Larsen & Skou, 1989]

Let $\mathcal{D} = (S, \mathbf{P}, \iota_{\text{init}}, AP, L)$ be a DTMC and $R \subseteq S \times S$ an equivalence. Then: R is a *probabilistic bisimulation* on S if for any $(s, t) \in R$:

1. $L(s) = L(t)$, and
2. $\mathbf{P}(s, C) = \mathbf{P}(t, C)$ for all equivalence classes $C \in S/R$

where $\mathbf{P}(s, C) = \sum_{s' \in C} \mathbf{P}(s, s')$.

For states in R, the probability of moving by a single transition to some equivalence class is equal.

Probabilistic bisimilarity

Let $\mathcal{D}$ be a DTMC and s, t states in $\mathcal{D}$. Then: s is *probabilistically bisimilar* to t, denoted $s \sim_p t$, if there exists a probabilistic bisimulation R with $(s, t) \in R$.

Strong bisimulation on CTMCs

Probabilistic bisimulation [Buchholz, 1994]

Let $\mathcal{C} = (S, \mathbf{P}, r, \iota_{\text{init}}, AP, L)$ be a CTMC and $R \subseteq S \times S$ an equivalence. Then: R is a *probabilistic bisimulation* on S if for any $(s, t) \in R$:

1. $L(s) = L(t)$, and
2. $r(s) = r(t)$, and
3. $\mathbf{P}(s, C) = \mathbf{P}(t, C)$ for all equivalence classes $C \in S/R$

The last two conditions amount to $\mathbf{R}(s, C) = \mathbf{R}(t, C)$ for all equivalence classes $C \in S/R$.

Probabilistic bisimilarity

Let $\mathcal{C}$ be a CTMC and s, t states in $\mathcal{C}$. Then: s is *probabilistically bisimilar* to t, denoted $s \sim_m t$, if there exists a probabilistic bisimulation R with $(s, t) \in R$.

Weak bisimulation on DTMCs

Weak probabilistic bisimulation [Baier & Hermanns, 1996]

Let $\mathcal{D} = (S, \mathbf{P}, \iota_{\text{init}}, AP, L)$ be a DTMC and $R \subseteq S \times S$ an equivalence. Then: R is a *weak probabilistic bisimulation* on S if for any $(s, t) \in R$:

1. $L(s) = L(t)$, and
2. if $\mathbf{P}(s, [s]_R) < 1$ and $\mathbf{P}(t, [t]_R) < 1$, then:

$$\frac{\mathbf{P}(s, C)}{1 - \mathbf{P}(s, [s]_R)} = \frac{\mathbf{P}(t, C)}{1 - \mathbf{P}(t, [t]_R)} \quad \text{for all } C \in S/R, C \neq [s]_R = [t]_R.$$

3. s can reach a state outside $[s]_R$ iff t can reach a state outside $[t]_R$.

For states in R, the conditional probability of moving by a single transition to another equivalence class is equal. In addition, either all states in an equivalence class C almost surely stay there, or have an option to escape from C.

Weak bisimulation on DTMCs

Weak probabilistic bisimulation [Baier & Hermanns, 1996]

Let $\mathcal{D} = (S, \mathbf{P}, \iota_{\text{init}}, AP, L)$ be a DTMC and $R \subseteq S \times S$ an equivalence. Then: R is a *weak probabilistic bisimulation* on S if for any $(s, t) \in R$:

1. $L(s) = L(t)$, and
2. if $\mathbf{P}(s, [s]_R) < 1$ and $\mathbf{P}(t, [t]_R) < 1$, then:

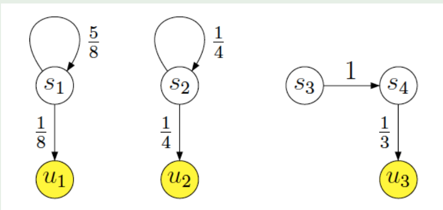
$$\frac{\mathbf{P}(s, C)}{1 - \mathbf{P}(s, [s]_R)} = \frac{\mathbf{P}(t, C)}{1 - \mathbf{P}(t, [t]_R)} \quad \text{for all } C \in S/R, C \neq [s]_R = [t]_R.$$

3. s can reach a state outside $[s]_R$ iff t can reach a state outside $[t]_R$.

Probabilistic weak bisimilarity

Let $\mathcal{D}$ be a DTMC and s, t states in $\mathcal{D}$. Then: s is *probabilistically weak bisimilar* to t, denoted $s \approx_p t$, if there exists a probabilistic weak bisimulation R with $(s, t) \in R$.

Weak bisimulation on DTMC: example



The equivalence relation R with $S/R = \{ \{s_1, s_2, s_3, s_4\}, \{u_1, u_2, u_3\} \}$ is a weak bisimulation. This can be seen as follows. For $C = \{ u_1, u_2, u_3 \}$ and s_1, s_2, s_4 with $\mathbf{P}(s_i, [s_i]_R) < 1$ we have:

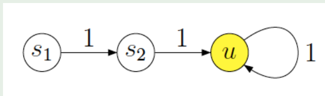
$$\frac{\mathbf{P}(s_1, C)}{1 - \mathbf{P}(s_1, [s_1])} = \frac{1/8}{1{-}5/8} = \frac{1/4}{1{-}1/4} = \frac{\mathbf{P}(s_2, C)}{1 - \mathbf{P}(s_2, [s_2])} = \frac{1/3}{1} = \frac{\mathbf{P}(s_4, C)}{1 - \mathbf{P}(s_4, [s_4])}.$$

Note that $\mathbf{P}(s_3, [s_3]_R) = 1$. Since s_3 can reach a state outside $[s_3]$ as s_1, s_2 and s_4, it follows that $s_1 \approx_p s_2 \approx_p s_3 \approx_p s_4$.

Reachability condition

Remark

Consider the following DTMC:



It is not difficult to establish $s_1 \approx_p s_2$. Note: $\mathbf{P}(s_1, [s_1]_R) = 1$, but $\mathbf{P}(s_2, [s_2]_R) < 1$. Both s_1 and s_2 can reach a state outside $[s_1]_R = [s_2]_R$. The reachability condition is essential to establish $s_1 \approx_p s_2$ and cannot be dropped: otherwise s_1 and s_2 would be weakly bisimilar to an equally labelled absorbing state.

Weak bisimulation on CTMCs

Weak probabilistic bisimulation [Bravetti, 2002]

Let $\mathcal{C} = (S, \mathbf{P}, r, \iota_{\text{init}}, AP, L)$ be a CTMC and $R \subseteq S \times S$ an equivalence.
Then: R is a *weak probabilistic bisimulation* on S if for any $(s, t) \in R$:

1. $L(s) = L(t)$, and
2. $\mathbf{R}(s, C) = \mathbf{R}(t, C)$ for all $C \in S/R$ with $C \neq [s]_R = [t]_R$

Weak probabilistic bisimilarity

Let $\mathcal{C}$ be a CTMC and s, t states in $\mathcal{C}$. Then: s is *weak probabilistically bisimilar* to t, denoted $s \approx_m t$, if there exists a weak probabilistic bisimulation R with $(s, t) \in R$.

A useful lemma

Let $\mathcal{C}$ be a CTMC and R an equivalence relation on S with $(s, t) \in R$, $\mathbf{P}(s, [s]_R) < 1$ and $\mathbf{P}(t, [t]_R) < 1$. Then: the following two statements are equivalent:

1. for all $C \in S/R$, $C \neq [s]_R = [t]_R$:

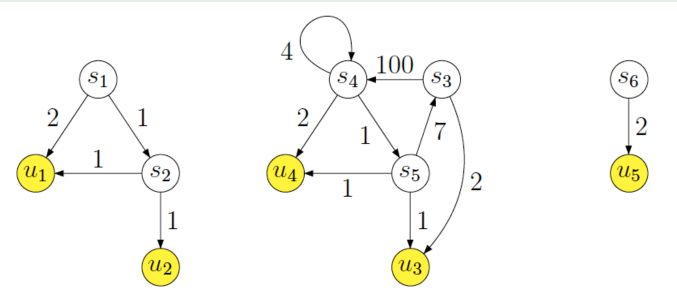
$$\frac{\mathbf{P}(s, C)}{1 - \mathbf{P}(s, [s]_R)} = \frac{\mathbf{P}(t, C)}{1 - \mathbf{P}(t, [t]_R)} \quad \text{and} \quad \mathbf{R}(s, S \setminus [s]_R) = \mathbf{R}(t, S \setminus [t]_R)$$

2. $\mathbf{R}(s, C) = \mathbf{R}(t, C)$ for all $C \in S/R$ with $C \neq [s]_R = [t]_R$.

Proof:

Left as an exercise.

Weak bisimulation on CTMCs: example



Equivalence relation R with $S/R = \{ \{s_1, s_2, s_3, s_4, s_5, s_6\}, \{u_1, u_2, u_3, u_4, u_5\} \}$ is a weak bisimulation on the CTMC depicted above. This can be seen as follows. For $C = \{ u_1, u_2, u_3, u_4, u_5 \}$, we have that all s-states enter C with rate 2. The rates between the s-states are not relevant.

Properties (without proof)

Strong and weak bisimulation in uniform CTMCs

For all uniform CTMCs $\mathcal{C}$ and states s, u in $\mathcal{C}$, we have:

$$s \sim_m u \quad \text{iff} \quad s \approx_m u \quad \text{iff} \quad s \sim_p u.$$

For any CTMC $\mathcal{C}$, we have: $\mathcal{C} \approx_m \mathit{unif}(r, \mathcal{C})$ with $r \geqslant \max_{s \in S} r(s)$.

Preservation of transient probabilities

For all CTMCs $\mathcal{C}$ with states s, u in $\mathcal{C}$ and $t \in \mathbb{R}_{\geqslant 0}$, we have:

$$s \approx_m u \quad \text{implies} \quad \underline{p}^s(t) = \underline{p}^u(t)$$

where $\underline{p}^s(0) = \mathbf{1}_s$ and $\underline{p}^u(0) = \mathbf{1}_u$ where $\mathbf{1}_s$ is the characteristic function for state s, i.e., $\mathbf{1}_s(s') = 1$ iff $s = s'$.

Overview

Computing transient probabilities

The transient probability vector $\underline{p}(t) = (p_{s_1}(t), \ldots, p_{s_k}(t))$ satisfies:

$$\underline{p}'(t) = \underline{p}(t) \cdot (\mathbf{R}-\mathbf{r}) \quad \text{given} \quad \underline{p}(0).$$

Standard knowledge yields: $\underline{p}(t) \;=\; \underline{p}(0) \cdot e^{(\mathbf{R}-\mathbf{r})\cdot t}$.

As uniformization preserves transient probabilities, we replace $\mathbf{R}-\mathbf{r}$ by its variant for the uniformized CTMC, i.e., $\overline{\mathbf{R}}-\overline{\mathbf{r}}$. We have:

$$\overline{\mathbf{R}}(s,s') = \overline{\mathbf{P}}(s,s') \cdot \overline{r}(s) = \overline{\mathbf{P}}(s,s') \cdot r \quad \text{and} \quad \overline{\mathbf{r}} = \mathbf{I} \cdot r.$$

Thus:

$$\underline{p}(0) \cdot e^{(\overline{\mathbf{R}}-\overline{\mathbf{r}})\cdot t} \;=\; \underline{p}(0) \cdot e^{(\overline{\mathbf{P}}\cdot r - \mathbf{I}\cdot r)\cdot t} \;=\; \underline{p}(0) \cdot e^{(\overline{\mathbf{P}}-\mathbf{I})\cdot r \cdot t} \;=\; \underline{p}(0) \cdot e^{-rt} \cdot e^{r\cdot t\cdot \overline{\mathbf{P}}}.$$

Computing transient probabilities

$$\underline{p}(t) \;=\; \underline{p}(0) \cdot e^{(\overline{\mathbf{R}}-\overline{\mathbf{r}})\cdot t} \;=\; \underline{p}(0) \cdot e^{(\overline{\mathbf{P}}\cdot r - \mathbf{I}\cdot r)\cdot t} \;=\; \underline{p}(0) \cdot e^{(\overline{\mathbf{P}}-\mathbf{I})\cdot r \cdot t} \;=\; \underline{p}(0) \cdot e^{-rt} \cdot e^{r\cdot t\cdot \overline{\mathbf{P}}}.$$

Computing a matrix exponential

Exploit Taylor-Maclaurin expansion. This yields:

$$\underline{p}(0) \cdot e^{-rt} \cdot e^{r\cdot t\cdot \overline{\mathbf{P}}} \;=\; \underline{p}(0) \cdot e^{-rt} \cdot \sum_{i=0}^{\infty} \frac{(r\cdot t)^i}{i!} \cdot \overline{\mathbf{P}}^i \;=\; \underline{p}(0) \cdot \sum_{i=0}^{\infty} \underbrace{e^{-r\cdot t}\frac{(r\cdot t)^i}{i!}}_{\text{Poisson prob.}} \cdot \overline{\mathbf{P}}^i$$

As $\overline{\mathbf{P}}$ is a stochastic matrix, computing the matrix exponential $\overline{\mathbf{P}}^i$ is numerically stable.

Intermezzo: Poisson distribution

Poisson distribution

The Poisson distribution is a discrete probability distribution that expresses the probability of a given number i of events occurring in a fixed interval of time $[0, t]$ if these events occur with a known average rate r and independently of the time since the last event. Formally, the pdf is:

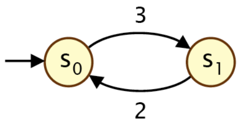
$$f(i; r\cdot t) \;=\; e^{-r\cdot t} \frac{(r\cdot t)^i}{i!}$$

where r is the mean of the Poisson distribution.

Remark

The Poisson distribution can be derived as a limiting case to the binomial distribution as the number of trials goes to infinity and the expected number of successes remains fixed.

Transient probabilities: example



$$\mathbf{P} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \underline{r} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \text{ and } \overline{\mathbf{P}}_3 = \begin{bmatrix} 0 & 1 \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

Let initial distribution $\underline{p}(0) = (1,0)$, and time bound $t{=}1$. Then:

$$\begin{aligned} \underline{p}(1) &= \underline{p}(0) \cdot \sum_{i=0}^{\infty} e^{-3} \frac{3^i}{i!} \cdot \overline{\mathbf{P}}^i \\ &= (1,0) \cdot e^{-3} \tfrac{1}{0!} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + (1,0) \cdot e^{-3} \tfrac{3}{1!} \cdot \begin{bmatrix} 0 & 1 \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} \\ &\quad + (1,0) \cdot e^{-3} \tfrac{9}{2!} \cdot \begin{bmatrix} 0 & 1 \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}^2 + \ldots\ldots \\ &\approx (0.404043, 0.595957) \end{aligned}$$

Truncating the infinite sum

Computing transient probabilities

$$\underline{p}(t) = \underline{p}(0) \cdot \sum_{i=0}^{\infty} e^{-r \cdot t} \frac{(r \cdot t)^i}{i!} \cdot \overline{\mathbf{P}}^i$$

- Summation can be truncated *a priori* for a given error bound $\varepsilon > 0$.
- The *error* that is introduced by truncating at summand k_ε is:

$$\left\| \sum_{i=0}^{\infty} e^{-rt} \frac{(rt)^i}{i!} \cdot \underline{p}(i) - \sum_{i=0}^{k_\varepsilon} e^{-rt} \frac{(rt)^i}{i!} \cdot \underline{p}(i) \right\| = \left\| \sum_{i=k_\varepsilon+1}^{\infty} e^{-rt} \frac{(rt)^i}{i!} \cdot \underline{p}(i) \right\|$$

- Strategy: choose k_ε minimal such that:

$$\sum_{i=k_{\varepsilon+1}}^{\infty} e^{-rt} \frac{(rt)^i}{i!} = \sum_{i=0}^{\infty} e^{-rt} \frac{(rt)^i}{i!} - \sum_{i=0}^{k_\varepsilon} e^{-rt} \frac{(rt)^i}{i!} = 1 - \sum_{i=0}^{k_\varepsilon} e^{-rt} \frac{(rt)^i}{i!} \leqslant \varepsilon$$

$\sum_{i=0}^{\infty} e^{-rt} \frac{(rt)^i}{i!} = 1$ due to the fact that $e^{-rt} \frac{(rt)^i}{i!}$ is a (Poisson) distribution

Overview

Summary

Main points

- Bisimilar states are equally labelled and their cumulative rate to any equivalence class coincides.
- Weak bisimilar states have equal conditional probabilities to move to some equivalence class, and can either both leave their class or both can't.
- Uniformization normalizes the exit rates of all states in a CTMC.
- Uniformization transforms a CTMC into a weak bisimilar one.
- Transient distribution are obtained by solving a system of linear differential equations.
- These equations can be solved conveniently on the uniformized CTMC.