Modeling and Verification of Probabilistic Systems

Lecture 14: Continuous-Time Markov Chains

Joost-Pieter Katoen

Lehrstuhl für Informatik 2
Software Modeling and Verification Group

http://www-i2.informatik.rwth-aachen.de/i2/mvps11/

June 20, 2011

Overview

Overview

Time in discrete-time Markov chains

The advance of time in DTMCs

- Time in a DTMC proceeds in discrete steps
- Two possible interpretations:
  1. accurate model of (discrete) time units
     - e.g., clock ticks in model of an embedded device
  2. time-abstract
     - no information assumed about the time transitions take
- State residence time is geometrically distributed

Continuous-time Markov chains

- dense model of time
- transitions can occur at any (real-valued) time instant
- state residence time is (negative) exponentially distributed

Continuous random variables

- X is a random variable (r.v., for short)
  - on a sample space with probability measure Pr
  - assume the set of possible values that X may take is dense
- X is *continuously distributed* if there exists a function $f(x)$ such that:

$$F_x(d) \;=\; Pr\{X \leqslant d\} \;=\; \int_{-\infty}^{d} f(x)\,dx \quad \text{for each real number } d$$

where f satisfies: $f(x) \geqslant 0$ for all x and $\int_{-\infty}^{\infty} f(x)\,dx = 1$

- $F_X(d)$ is the *(cumulative) probability distribution function*
- $f(x)$ is the *probability density function*

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:

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

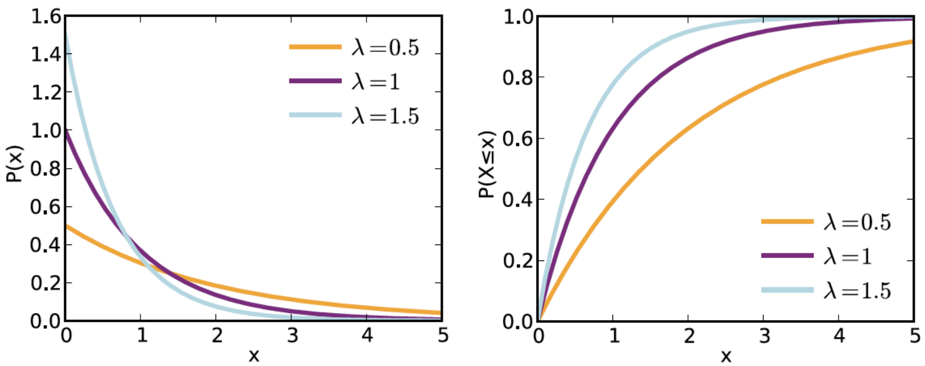
The rate $\lambda \in \mathbb{R}_{>0}$ uniquely determines an exponential distribution.

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

Exponential pdf and cdf



The higher λ, the faster the cdf approaches 1.

Why exponential distributions?

- Are *adequate* for many real-life phenomena
  - the time until a radioactive particle decays
  - the time between successive car accidents
  - inter-arrival times of jobs, telephone calls in a fixed interval
- Are the continuous counterpart of the *geometric* distribution
- Heavily used in physics, performance, and reliability analysis
- Can *approximate* general distributions arbitrarily closely
- Yield a *maximal entropy* if only the mean is known

Memoryless property

Theorem

1. For any exponentially distributed random variable X:

$$Pr\{X > t + d \mid X > t\} \;=\; Pr\{X > d\} \text{ for any } t, d \in \mathbb{R}_{\geqslant 0}.$$

2. Any cdf which is memoryless is a negative exponential one.

Proof:

Proof of 1. : Let λ be the rate of X's distribution. Then we derive:

$$Pr\{X > t + d \mid X > t\} \;=\; \frac{Pr\{X > t{+}d \cap X > t\}}{Pr\{X > t\}} \;=\; \frac{Pr\{X > t{+}d\}}{Pr\{X > t\}}$$

$$= \frac{e^{-\lambda\cdot(t+d)}}{e^{-\lambda\cdot t}} \;=\; e^{-\lambda\cdot d} \;=\; Pr\{X > d\}.$$

Proof of 2. : By contraposition, using the total law of probability.

Closure under minimum

Minimum closure theorem

For independent, exponentially distributed random variables X and Y with rates $\lambda, \mu \in \mathbb{R}_{>0}$, the r.v. $\min(X, Y)$ is exponentially distributed with rate $\lambda{+}\mu$, i.e.,:

$$Pr\{\min(X, Y) \leqslant t\} \;=\; 1 - e^{-(\lambda+\mu)\cdot t} \quad \text{for all } t \in \mathbb{R}_{\geqslant 0}.$$

Proof

Let λ (μ) be the rate of X's (Y's) distribution. Then we derive:

$$Pr\{\min(X, Y) \leqslant t\} = Pr_{X,Y}\{(x, y) \in \mathbb{R}^2_{\geqslant 0} \mid \min(x, y) \leqslant t\}$$

$$= \int_0^\infty \left(\int_0^\infty \mathbf{I}_{\min(x,y)\leqslant t}(x, y) \cdot \lambda e^{-\lambda x} \cdot \mu e^{-\mu y}\, dy \right) dx$$

$$= \int_0^t \int_x^\infty \lambda e^{-\lambda x} \cdot \mu e^{-\mu y}\, dy\, dx + \int_0^t \int_y^\infty \lambda e^{-\lambda x} \cdot \mu e^{-\mu y}\, dx\, dy$$

$$= \int_0^t \lambda e^{-\lambda x} \cdot e^{-\mu x}\, dx + \int_0^t e^{-\lambda y} \cdot \mu e^{-\mu y}\, dy$$

$$= \int_0^t \lambda e^{-(\lambda+\mu)x}\, dx + \int_0^t \mu e^{-(\lambda+\mu)y}\, dy$$

$$= \int_0^t (\lambda{+}\mu) \cdot e^{-(\lambda+\mu)z}\, dz \;=\; 1 - e^{-(\lambda+\mu)t}$$

Closure under minimum

Minimum closure theorem for several exponentially distributed r.v.'s

For independent, exponentially distributed random variables $X_1, X_2, \ldots, X_n$ with rates $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}_{>0}$ the r.v. $\min(X_1, X_2, \ldots, X_n)$ is exponentially distributed with rate $\sum_{0<i\leqslant n} \lambda_i$, i.e.,:

$$Pr\{\min(X_1, X_2, \ldots, X_n) \leqslant t\} \;=\; 1 - e^{-\sum_{0<i\leqslant n} \lambda_i \cdot t} \quad \text{for all } t \in \mathbb{R}_{\geqslant 0}.$$

Proof:

Generalization of the proof for the case of two exponential distributions.

Winning the race with two competitors

The minimum of two exponential distributions

For independent, exponentially distributed random variables X and Y with rates $\lambda, \mu \in \mathbb{R}_{>0}$, it holds:

$$Pr\{X \leqslant Y\} = \frac{\lambda}{\lambda+\mu}.$$

Proof

Let λ (μ) be the rate of X's (Y's) distribution. Then we derive:

$$\begin{aligned} Pr\{X \leqslant Y\} &= Pr_{X,Y}\{(x,y) \in \mathbb{R}^2_{\geqslant 0} \mid x \leqslant y\} \\ &= \int_0^\infty \mu e^{-\mu y} \left(\int_0^y \lambda e^{-\lambda x}\, dx \right) dy \\ &= \int_0^\infty \mu e^{-\mu y} \left(1 - e^{-\lambda y}\right) dy \\ &= 1 - \int_0^\infty \mu e^{-\mu y} \cdot e^{-\lambda y}\, dy \;=\; 1 - \int_0^\infty \mu e^{-(\mu+\lambda) y}\, dy \\ &= 1 - \frac{\mu}{\mu+\lambda} \cdot \underbrace{\int_0^\infty (\mu+\lambda) e^{-(\mu+\lambda)y}\, dy}_{=1} \\ &= 1 - \frac{\mu}{\mu+\lambda} \;=\; \frac{\lambda}{\mu+\lambda} \end{aligned}$$

Winning the race with many competitors

The minimum of several exponentially distributed r.v.'s

For independent, exponentially distributed random variables $X_1, X_2, \ldots, X_n$ with rates $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}_{>0}$ it holds:

$$Pr\{X_i = \min(X_1, \ldots, X_n)\} = \frac{\lambda_i}{\sum_{j=1}^n \lambda_j}.$$

Proof:

Generalization of the proof for the case of two exponential distributions.

Overview

Continuous-time Markov chain

Continuous-time Markov chain

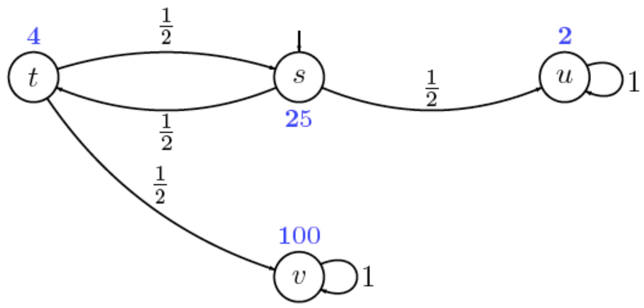
A CTMC is a tuple $(S, \mathbf{P}, r, \iota_{\text{init}}, AP, L)$ where

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

Interpretation

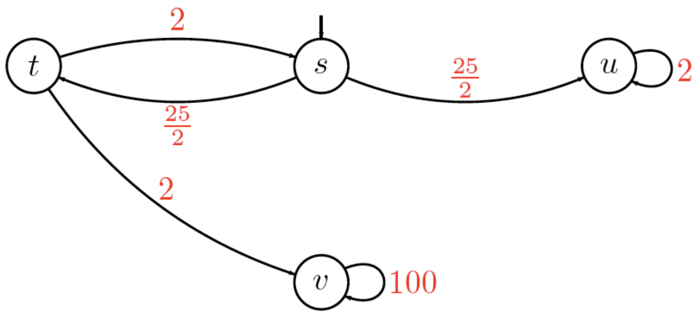
- residence time in state s is exponentially distributed with rate $r(s)$.
- phrased alternatively, the average residence time of state s is $\frac{1}{r(s)}$.
- thus, the higher the rate $r(s)$, the shorter the average residence time in s.

Example



$r(s) = 25,\ r(t) = 4,\ r(u) = 2$ and $r(v) = 100$

Example: a classical perspective

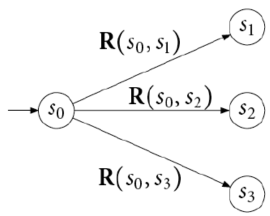


$r(s) = 25,\ r(t) = 4,\ r(u) = 2$ and $r(v) = 100$

The transition rate $\mathbf{R}(s, s') = \mathbf{P}(s, s') \cdot r(s)$

We use $(S, \mathbf{P}, r, \iota_{\text{init}}, AP, L)$ and $(S, \mathbf{R}, \iota_{\text{init}}, AP, L)$ interchangeably.

CTMC semantics by example



CTMC semantics

- Transition $s \to s'$:= r.v. $X_{s,s'}$ with rate $\mathbf{R}(s, s')$
- Probability to go from state s_0 to, say, state s_2 is:

$$Pr\{X_{s_0,s_2} \leqslant X_{s_0,s_1} \cap X_{s_0,s_2} \leqslant X_{s_0,s_3}\}$$
$$=$$
$$\frac{\mathbf{R}(s_0, s_2)}{\mathbf{R}(s_0, s_1) + \mathbf{R}(s_0, s_2) + \mathbf{R}(s_0, s_3)} = \frac{\mathbf{R}(s_0, s_2)}{r(s_0)}$$

- Probability of staying at most t time in s_0 is:

$$Pr\{\min(X_{s_0,s_1}, X_{s_0,s_2}, X_{s_0,s_3}) \leqslant t\}$$
$$=$$
$$1 - e^{-(\mathbf{R}(s_0,s_1)+\mathbf{R}(s_0,s_2)+\mathbf{R}(s_0,s_3)) \cdot t} = 1 - e^{-r(s_0) \cdot t}$$

CTMC semantics

Enabledness

The probability that transition $s \rightarrow s'$ 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:

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

CTMC semantics

State-to-state timed transition probability

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

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

Proof:

On the blackboard.

CTMC semantics

Residence time distribution

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

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

Proof:

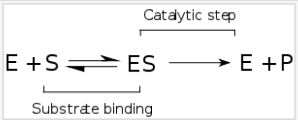
On the blackboard.

Enzyme-catalysed substrate conversion

Kinetics

[edit]

Main article: Enzyme kinetics



Mechanism for a single substrate enzyme catalyzed reaction. The enzyme (E) binds a substrate (S) and produces a product (P).

Enzyme kinetics is the investigation of how enzymes bind substrates and turn them into products. The rate data used in kinetic analyses are commonly obtained from enzyme assays, where since the 90s, the dynamics of many enzymes are studied on the level of individual molecules.

In 1902 Victor Henri[57] proposed a quantitative theory of enzyme kinetics, but his experimental data were not useful because the significance of the hydrogen ion concentration was not yet appreciated. After Peter Lauritz Sørensen had defined the logarithmic pH-scale and introduced the concept of buffering in 1909[58] the German chemist Leonor Michaelis and his Canadian postdoc Maud Leonora Menten repeated Henri's experiments and confirmed his equation which is referred to as Henri-Michaelis-Menten kinetics (termed also Michaelis-Menten kinetics).[59] Their work was further developed by G. E. Briggs and J. B. S. Haldane, who derived kinetic equations that are still widely considered today a starting point in solving enzymatic activity.[60]

The major contribution of Henri was to think of enzyme reactions in two stages. In the first, the substrate binds reversibly to the enzyme, forming the enzyme-substrate complex. This is sometimes called the Michaelis complex. The enzyme then catalyzes the chemical step in the reaction and releases the product. Note that the simple Michaelis Menten mechanism for the enzymatic activity is considered today a basic idea, where many examples show that the enzymatic activity involves structural dynamics. This is incorporated in the enzymatic mechanism while introducing several Michaelis Menten pathways that are connected with fluctuating rates [44][45][46]. Nevertheless, there is a mathematical relation connecting the behavior obtained from the basic Michaelis Menten mechanism (that was indeed proved correct in many experiments) with the generalized Michaelis Menten mechanisms involving dynamics and activity; [61] this means that the measured activity of enzymes on the level of many enzymes may be explained with the simple Michaelis-Menten equation, yet, the actual activity of enzymes is richer and involves structural dynamics.

Source: wikipedia (June 2011)

Stochastic chemical kinetics

- Types of reaction described by stochiometric equations:

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} E + P$$

- N different types of molecules that randomly collide

  where state $X(t) = (x_1, \ldots, x_N)$ with $x_i = \#$ molecules of sort i

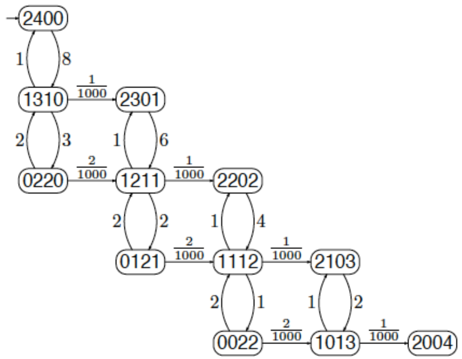
- Reaction probability within infinitesimal interval $[t, t{+}\Delta)$:

$$\alpha_m(\vec{x}) \cdot \Delta \;=\; Pr\{\text{reaction } m \text{ in } [t, t{+}\Delta) \mid X(t) = \vec{x}\} \text{ where}$$

$$\alpha_m(\vec{x}) = k_m \cdot \# \text{ possible combinations of reactant molecules in } \vec{x}$$

- This process is a continuous-time Markov chain.

Enzyme-catalyzed substrate conversion as a CTMC



States:	*init*	*goal*
enzymes	2	2
substrates	4	0
complex	0	0
products	0	4

Transitions: $E + S \underset{1}{\overset{1}{\rightleftharpoons}} C \xrightarrow{0.001} E + P$

e.g., $(x_E, x_S, x_C, x_P) \xrightarrow{0.001 \cdot x_C} (x_E + 1, x_S, x_C - 1, x_P + 1)$ for $x_C > 0$

CTMCs are omnipresent!

- Markovian queueing networks (Kleinrock 1975)
- Stochastic Petri nets (Molloy 1977)
- Stochastic activity networks (Meyer & Sanders 1985)
- Stochastic process algebra (Herzog *et al.*, Hillston 1993)
- Probabilistic input/output automata (Smolka *et al.* 1994)
- Calculi for biological systems (Priami *et al.*, Cardelli 2002)

CTMCs are one of the most prominent models in performance analysis

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.

Overview

Summary

Main points

- Exponential distributions are closed under minimum.
- The probability to win a race amongst several exponential distributions only depends on their rates.
- A CTMC is a DTMC where state residence times are exponentially distributed.
- CTMC semantics distinguishes between enabledness and taking a transition.
- Transient distribution are obtained by solving a system of linear differential equations.
- CTMCs are frequently used as semantical model for high-level formalisms.