EECS 144/244: System Modeling, Analysis, and Optimization

Stochastic Systems

Lecture: Continuous Time Stochastic Systems

Alexandre Donzé

University of California, Berkeley

April 26, 2013

Probabilistic Models

	Fully Probabilistic	Nondeterministic
Discrete Time	Discrete-Time Markov Chains (DTMCs)	Markov Decision Processes (MDPs)
Continuous Time	Continuous-Time Markov Chains (DTMCs)	CTMDPs, Probabilistic Timed Automaton (PTAs)

Probabilistic Models

	Fully Probabilistic	Nondeterministic
Discrete Time	Discrete-Time Markov Chains (DTMCs)	Markov Decision Processes (MDPs)
Continuous Time	Continuous-Time Markov Chains (DTMCs)	CTMDPs, Probabilistic Timed Automaton (PTAs)

From DTMC to CTMCs

Time in a DTMC proceeds in discrete steps

- accurate model of (discrete) time units (e.g. clock ticks)
- or, no information assumed about the time transitions take

Continuous-time Markov chains (CTMCs): dense model of time

- transitions can occur at any (real-valued) time instant
- modelled using exponential distributions
- suits modelling of: performance/reliability (e.g. of computer networks, manufacturing systems, queueing networks), biological pathways, chemical reactions, ...

Continuous Probability Distribution

Defined by

- a cumulative distribution function : $F(t) = Pr(X \leq t) = \int_{-\infty}^{t} f(x)dx$

Continuous Probability Distribution

Defined by

- a cumulative distribution function : $F(t) = Pr(X \leq t) = \int_{-\infty}^{t} f(x)dx$
- where f is the probability density function

Continuous Probability Distribution

Defined by

- a cumulative distribution function : $F(t) = Pr(X \leq t) = \int_{-\infty}^{t} f(x)dx$
- where f is the probability density function
- Note: for all t, $Pr(X = t) = ?$

Continuous Probability Distribution

Defined by

- a cumulative distribution function : $F(t) = Pr(X \leq t) = \int_{-\infty}^{t} f(x)dx$
- where f is the probability density function
- Note: for all t, $Pr(X = t) = 0$!

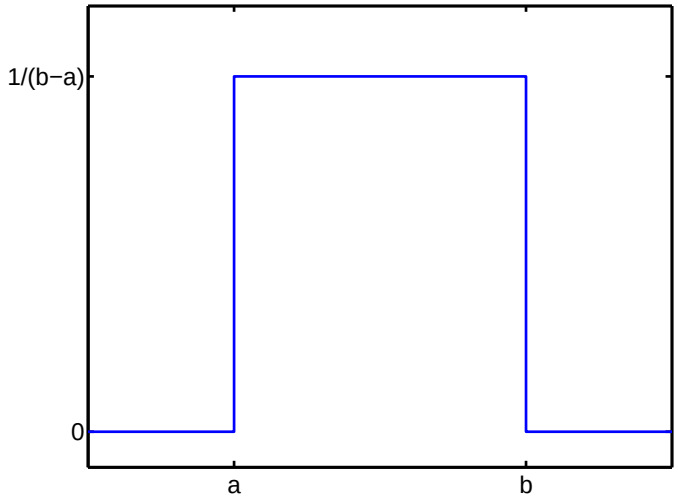
Continuous Probability Distribution

Defined by

- a cumulative distribution function : $F(t) = Pr(X \leq t) = \int_{-\infty}^{t} f(x)dx$
- where f is the probability density function
- Note: for all t, $Pr(X = t) =$ 0 !

Example: uniform distribution

$$f(t) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq t \leq b \ , \\ 0 & \text{otherwise.} \end{cases}$$



Continuous Probability Distribution

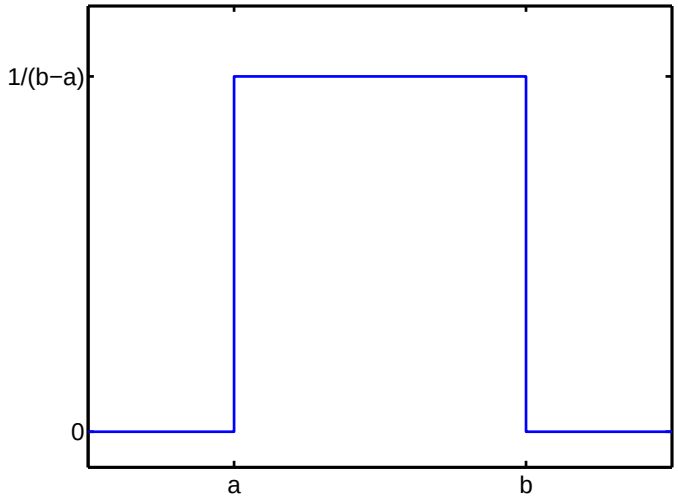
Defined by

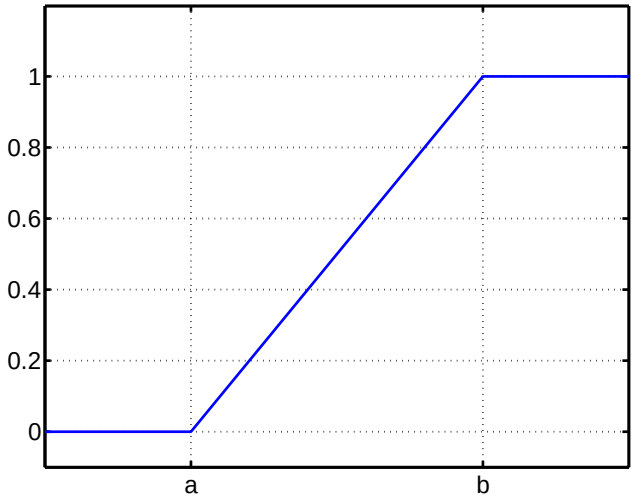
- a cumulative distribution function : $F(t) = Pr(X \leq t) = \int_{-\infty}^{t} f(x)dx$
- where f is the probability density function
- Note: for all t, $Pr(X = t) = 0$!

Example: uniform distribution

$$f(t) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq t \leq b \ , \\ 0 & \text{otherwise.} \end{cases}$$

$$F(t) = \begin{cases} 0 & \text{if } t \leq a, \\ \frac{t-a}{b-a} & \text{if } a \leq t \leq b, \\ 1 & \text{if } t > b. \end{cases}$$





Exponential Distribution

A continuous random variable X is exponential with parameter λ if its density functions is

$$f(t) = \begin{cases} \lambda \cdot e^{-\lambda t} & \text{if } t > 0 \ , \\ 0 & \text{otherwise} \ . \end{cases}$$

Exponential Distribution

A continuous random variable X is exponential with parameter λ if its density functions is

$$f(t) = \begin{cases} \lambda \cdot e^{-\lambda t} & \text{if } t > 0 \ , \\ 0 & \text{otherwise} \ . \end{cases}$$

$\lambda =$ "rate"

Exponential Distribution

A continuous random variable X is exponential with parameter λ if its density functions is

$$f(t) = \begin{cases} \lambda \cdot e^{-\lambda t} & \text{if } t > 0 \ , \\ 0 & \text{otherwise} \ . \end{cases}$$

$\lambda =$ "rate"

Cumulative distribution function

$$F(t) = Pr(X \leq t) = \int_0^t \lambda \cdot e^{-\lambda x} dx = 1 - e^{-\lambda t}$$

Exponential Distribution

A continuous random variable X is exponential with parameter λ if its density functions is

$$f(t) = \begin{cases} \lambda \cdot e^{-\lambda t} & \text{if } t > 0 \ , \\ 0 & \text{otherwise} \ . \end{cases}$$

$\lambda =$ "rate"

Cumulative distribution function

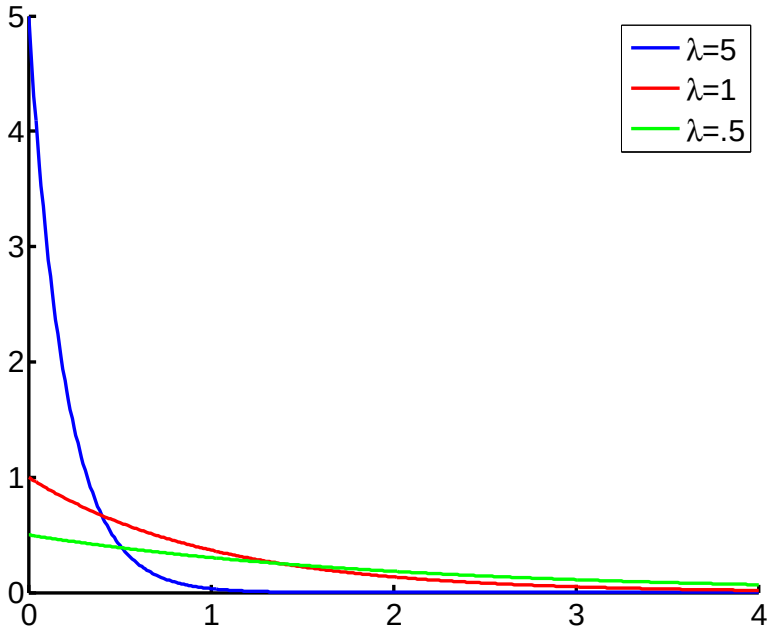
$$F(t) = Pr(X \leq t) = \int_0^t \lambda \cdot e^{-\lambda x} dx = 1 - e^{-\lambda t}$$

Other properties

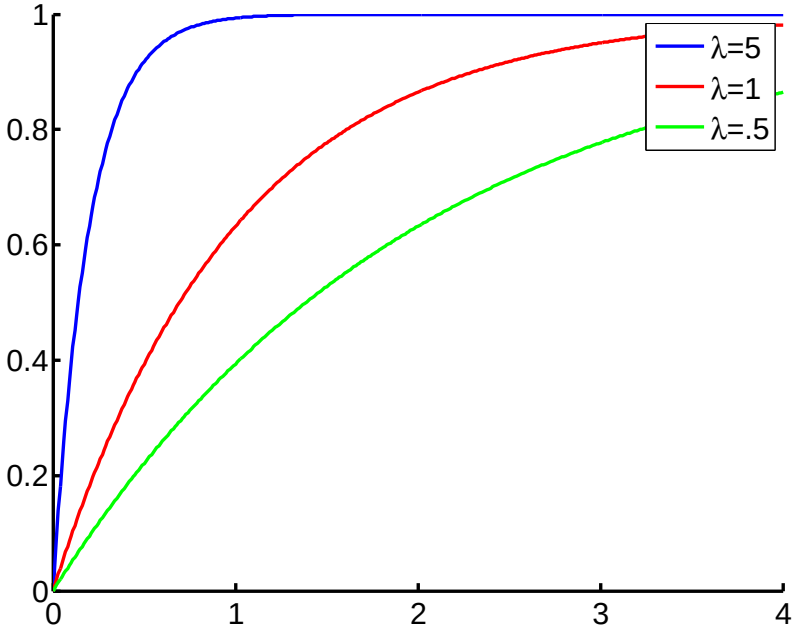
- negation : $Pr(X > t) = e^{-\lambda t}$
- mean : $E[X] = \int_0^\infty x \cdot \lambda \cdot e^{-\lambda x} dx = \frac{1}{\lambda}$
- variance: $Var(X) = \frac{1}{\lambda^2}$

Exponential Distribution - Examples

Probability distribution function



Cumulative distribution function



The more λ increases, the faster the c.d.f. approaches 1

Exponential distribution

Adequate for modelling many real-life phenomena

- failures
  - e.g. time before machine component fails
- inter-arrival times
  - e.g. time before next call arrives to a call center
- biological systems
  - e.g. times for reactions between proteins to occur

Useful Properties

The exponential distribution is memoryless

- $Pr(X > t_1 + t_2 | X > t_1) = Pr(X > t_2)$
- it is the only memoryless continuous distribution

Useful Properties

The exponential distribution is memoryless

- $Pr(X > t_1 + t_2 | X > t_1) = Pr(X > t_2)$
- it is the only memoryless continuous distribution
- the discrete time equivalent is the geometric distribution:

  Probability of a failure after k trials knowing the probability p of failure of one trial:

$$P(X = k) = (1 - p)^k p$$

Useful Properties

The minimum of two exponential distributions is an exponential distribution

- $X_1 \sim Exp(\lambda_1)$, $X_2 \sim Exp(\lambda_2)$
- $Y = min(X_1, X_2) \sim Exp(\lambda_1 + \lambda_2)$
- generalises to minimum of n distributions

Comparison of two exponential distributions

The probability of $X_1 < X_2$ is given by

$$Pr(X_1 < X_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

Continuous-time Markov Chains

Informally

- labelled transition systems, augmented with rates
- continuous time delays, exponentially distributed

Definition

A CTMC is a tuple (S, s_0, R, L) where

- S is a set of states
- $s_0 \in S$ is the initial state
- $R : S \times S \to \mathbb{R}^+$ is the transition rate matrix
- $L : S \to 2^{AP}$ is a labelling with atomic propositions in AP

CTMCs Semantics

The transition rate matrix assigns rates to each pair of states

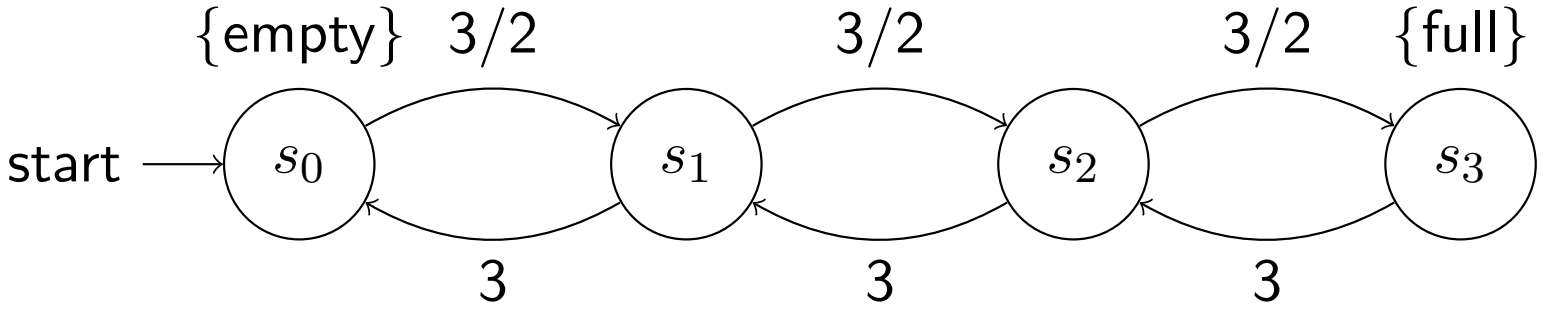
- used as a parameter to an exponential distribution
- transition between s and s' when $R(s, s') > 0$
- probability triggered before t time units: $1 - e^{-R(s,s')t}$

Race condition

If there exists multiple s' with $R(s, s') > 0$, the first transition to trigger determines the next state.

Example - Modeling a queue of jpbs

- Initially the queue is empty
- jobs arrive with rate 3/2 (i.e. mean inter-arrival time is 2/3
- jobs are served with rate 3 (i.e. mean service time is 1/3)
- maximum size of the queue is 3
- state-space: $S : \{s_i\}_{i=0...3}$ where s_i indicates i jobs in the queue



Interesting Questions for CTMCs

How long is spent in s before a transition occurs ?

Interesting Questions for CTMCs

How long is spent in s before a transition occurs ?

- minimum of exponential distributions of outgoing transitions
- i.e. exponential distribution with sum of outgoing rates

$$\text{Exit rate } E(s) = \sum_{s' \in S} R(s, s')$$

Interesting Questions for CTMCs

How long is spent in s before a transition occurs ?

- minimum of exponential distributions of outgoing transitions
- i.e. exponential distribution with sum of outgoing rates

$$\text{Exit rate } E(s) = \sum_{s' \in S} R(s, s')$$

Note:

- the probability of leaving a state s within $[0, s]$ is $1 - e^{-E(s)t}$
- s is called absorbing if $E(s) = 0$ (no outgoing transitions)

Interesting Questions for CTMCs

Which transition is eventually taken from state s ?

- Recall that $Pr(X_1 < X_2) = \frac{\lambda_1}{\lambda_1+\lambda_2}$
- This can generalized to $Pr(X_1 = \min_{i=1\ldots n} X_i) = \frac{\lambda_1}{\sum \lambda_i}$
- Thus the probability that next state from s is s' is given by

$$P_R(s,s') = \begin{cases} \frac{R(s,s')}{E(s)} & \text{if } E(s) > 0, \\ 1 & \text{if } E(s) = 0 \text{ and } s = s', \\ 0 & \text{if } E(s) = 0 \text{ and } s \neq s'. \end{cases}$$

Embedded DTMC

The transition target state is independent from the time the transition occurs.

I.e. we can define a DTMC that abstracts a CTMC by describing only states transitions without time information

Embedded DTMC

The transition target state is independent from the time the transition occurs.

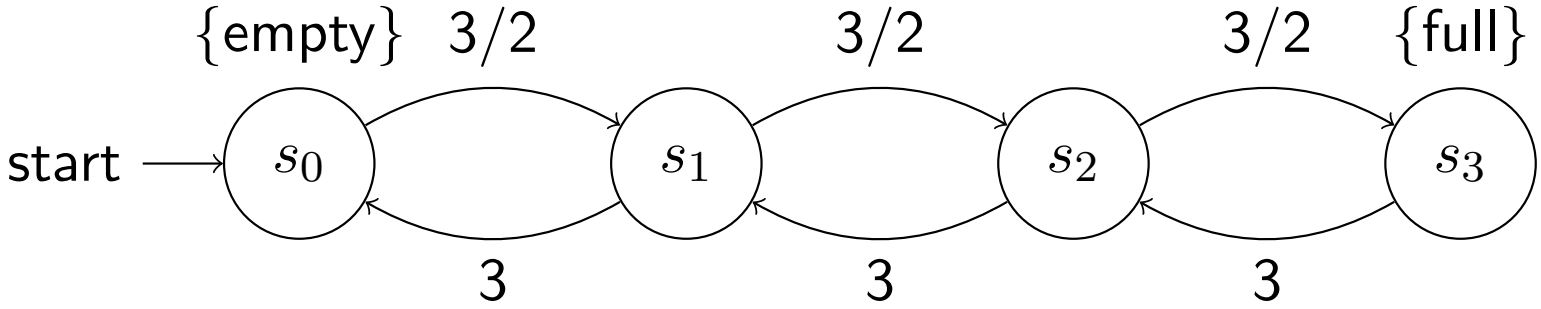
I.e. we can define a DTMC that abstracts a CTMC by describing only states transitions without time information

Definition

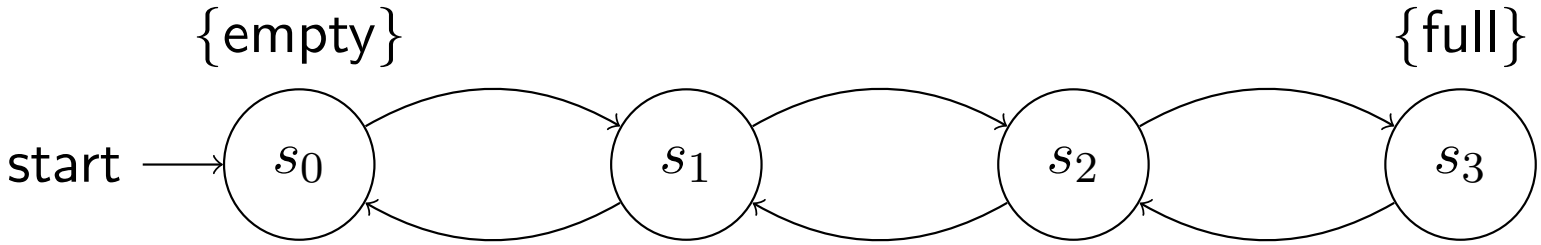
The embedded DTMC of a CTMC (S, s_0, R, L) is the DMTC (S, s_0, P_R, L)

Embedded DTMC - Example

What is the embedded DTMC of

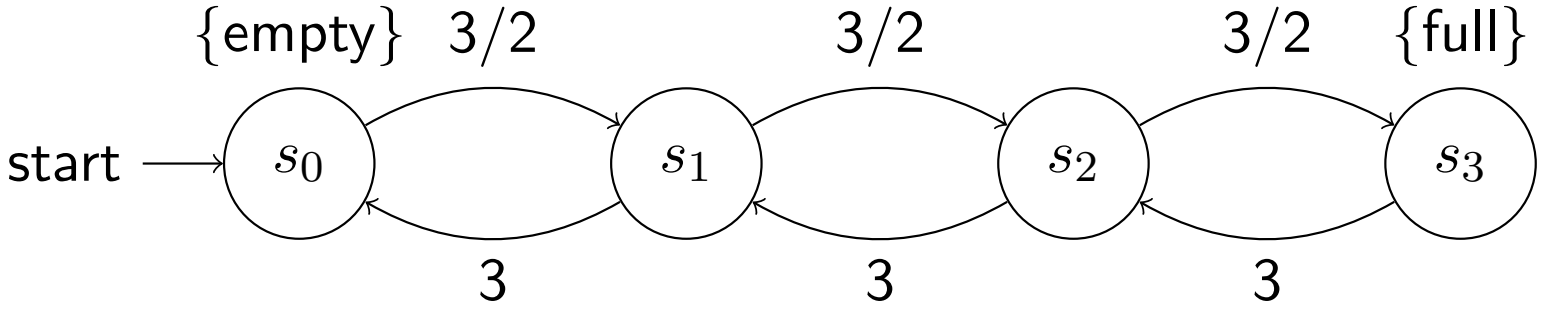


?

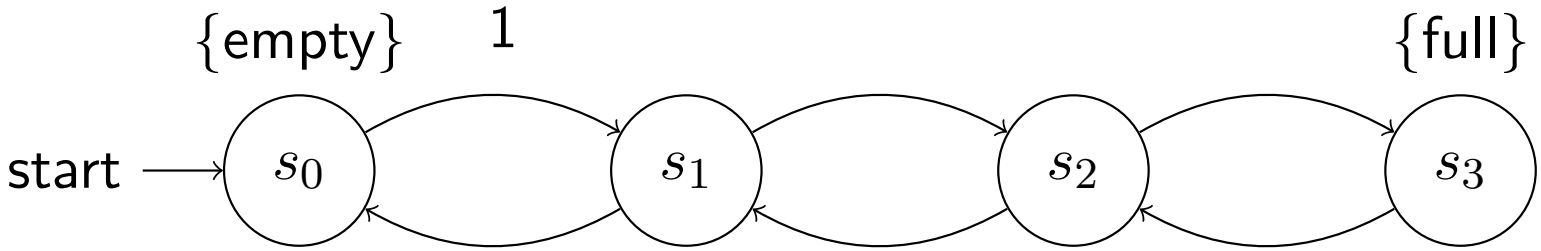


Embedded DTMC - Example

What is the embedded DTMC of

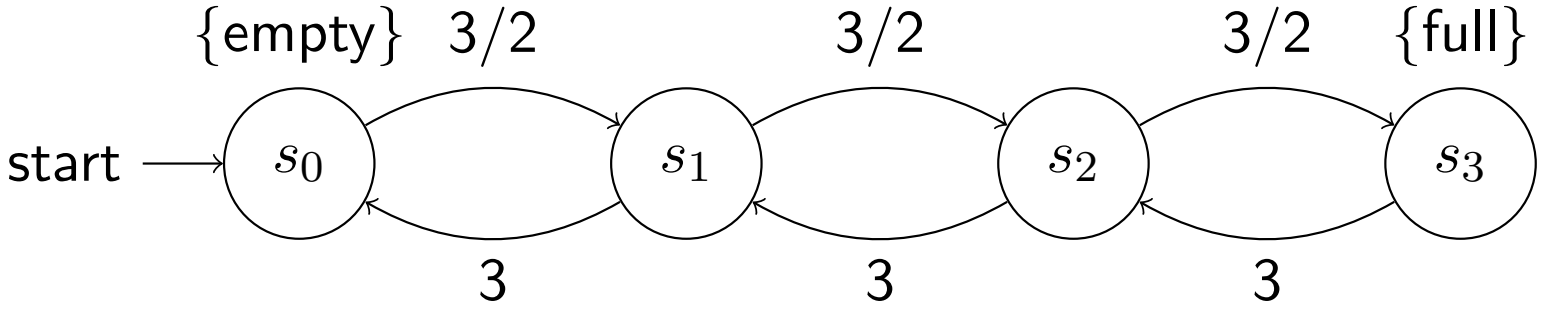


?

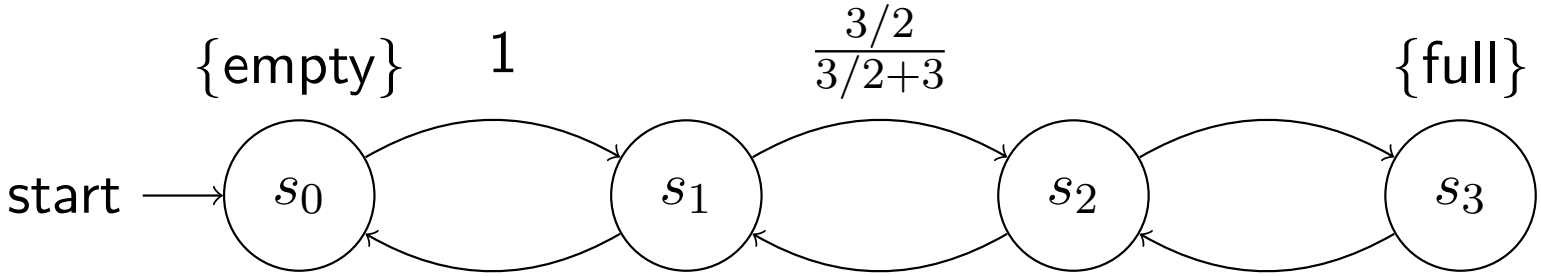


Embedded DTMC - Example

What is the embedded DTMC of

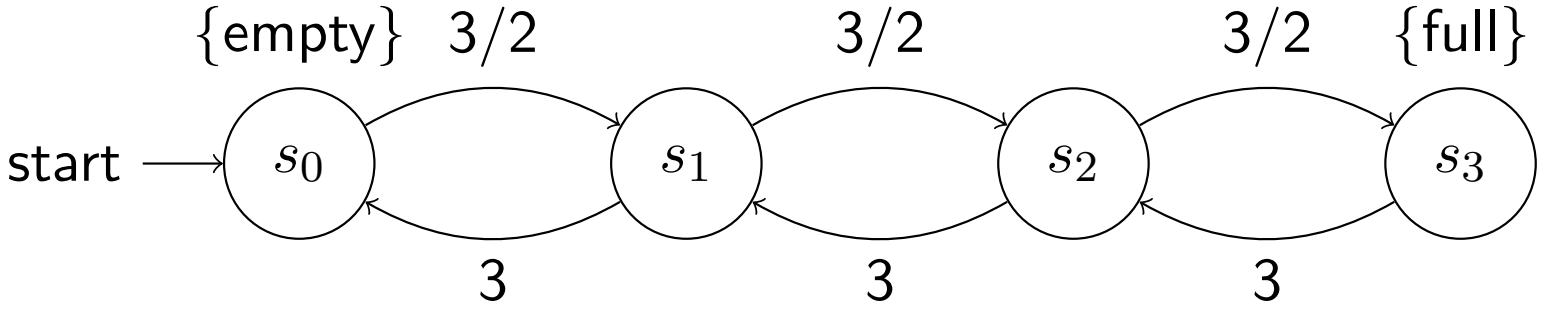


?

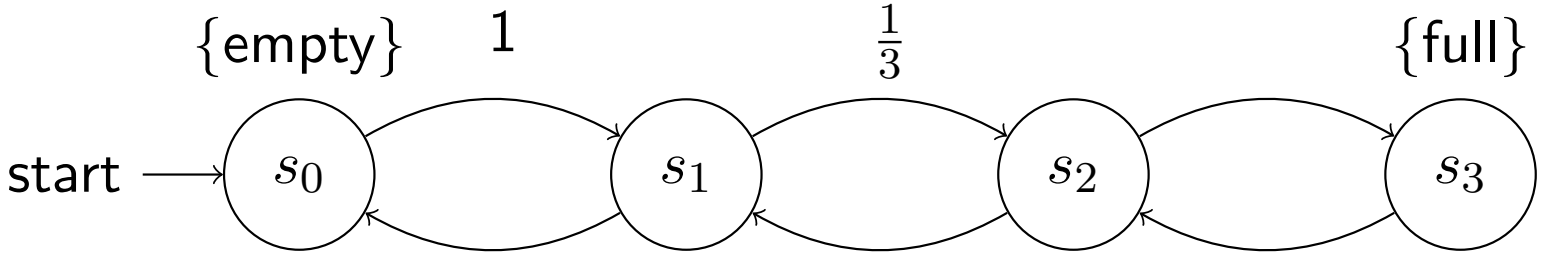


Embedded DTMC - Example

What is the embedded DTMC of

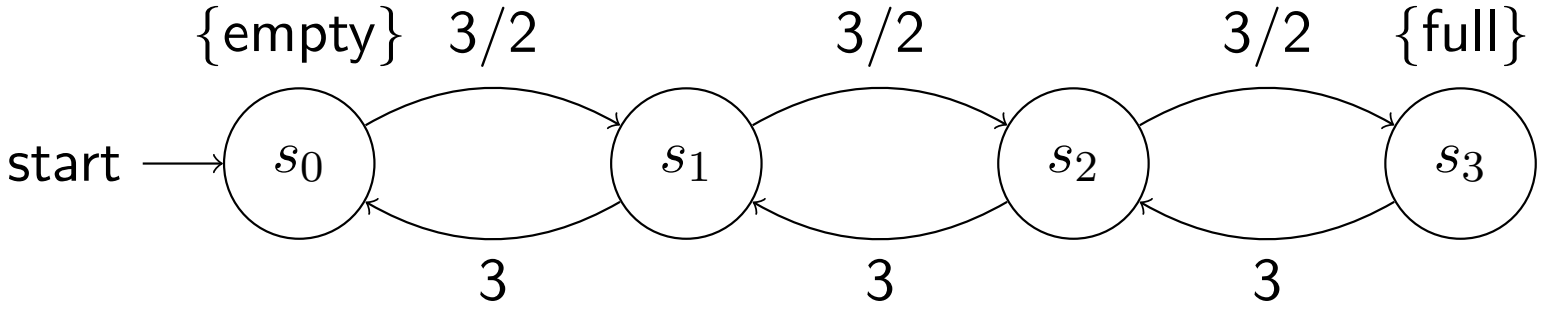


?

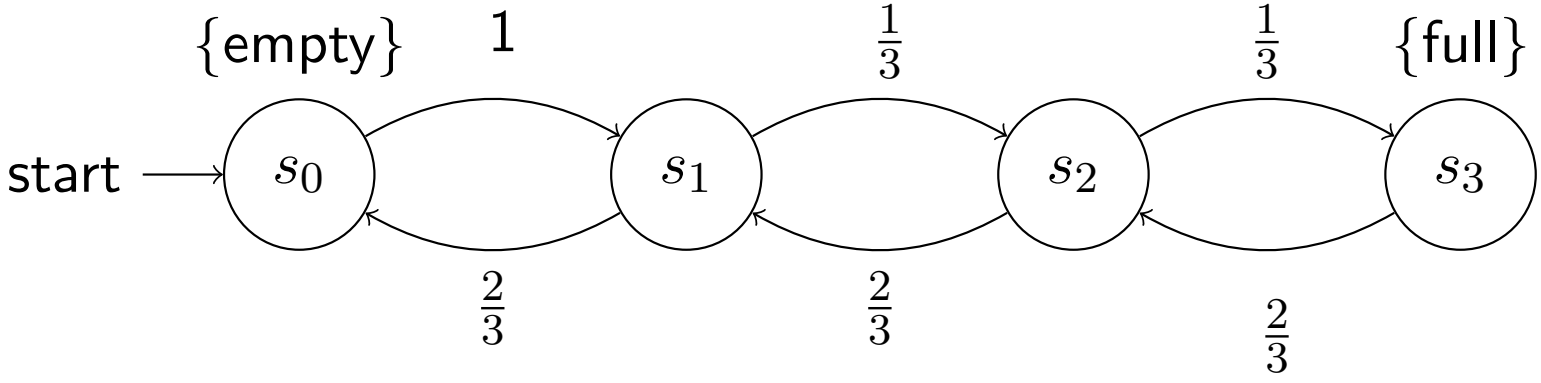


Embedded DTMC - Example

What is the embedded DTMC of



?



Interesting Question

What is the probability of being in state s_j at time t starting in s_i ?

Define

$$P(t) = \begin{pmatrix} P_{11}(t) & \dots & P_{1n}(t) \\ \vdots & \dots & \vdots \\ P_{n1}(t) & \dots & P_{nn}(t) \end{pmatrix} \text{ with } P_{ij}(t) = Pr(s(t) = s_j | s(t = 0) = s_i)$$

Interesting Question

What is the probability of being in state s_j at time t starting in s_i ?
Define

$$P(t) = \begin{pmatrix} P_{11}(t) & \dots & P_{1n}(t) \\ \vdots & \dots & \vdots \\ P_{n1}(t) & \dots & P_{nn}(t) \end{pmatrix} \text{ with } P_{ij}(t) = Pr(s(t) = s_j | s(t=0) = s_i)$$

and the infinitesimal generator matrix Q

$$Q(t) = \begin{pmatrix} Q_{11}(t) & \dots & Q_{1n}(t) \\ \vdots & \dots & \vdots \\ Q_{n1}(t) & \dots & Q_{nn}(t) \end{pmatrix} \text{ with } Q_{ij} = \begin{cases} R(s_i, s_j) & \text{if } i \neq j, \\ -\sum_{k \neq i} R(s_i, s_k) & \text{if } i = j. \end{cases}$$

Interesting Question

What is the probability of being in state s_j at time t starting in s_i ?
Define

$$P(t) = \begin{pmatrix} P_{11}(t) & \dots & P_{1n}(t) \\ \vdots & \dots & \vdots \\ P_{n1}(t) & \dots & P_{nn}(t) \end{pmatrix} \text{ with } P_{ij}(t) = Pr(s(t) = s_j | s(t=0) = s_i)$$

and the infinitesimal generator matrix Q

$$Q(t) = \begin{pmatrix} Q_{11}(t) & \dots & Q_{1n}(t) \\ \vdots & \dots & \vdots \\ Q_{n1}(t) & \dots & Q_{nn}(t) \end{pmatrix} \text{ with } Q_{ij} = \begin{cases} R(s_i, s_j) & \text{if } i \neq j, \\ -\sum_{k \neq i} R(s_i, s_k) & \text{if } i = j. \end{cases}$$

Then one can show that $P(t)$ satisfies the linear ODE:

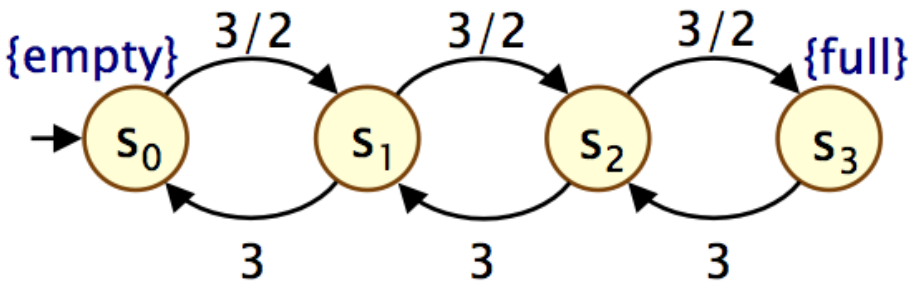
$$\dot{P}(t) = P(t)'Q, P(0) = I$$

Simple Example

$C = (S, s_{init}, R, L)$

$S = \{s_0, s_1, s_2, s_3\}$

$s_{init} = s_0$



$AP = \{empty, full\}$

$L(s_0)=\{empty\}$, $L(s_1)=L(s_2)=\varnothing$ and $L(s_3)=\{full\}$

$$R = \begin{bmatrix} 0 & 3/2 & 0 & 0 \\ 3 & 0 & 3/2 & 0 \\ 0 & 3 & 0 & 3/2 \\ 0 & 0 & 3 & 0 \end{bmatrix} \quad P^{emb(C)} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 2/3 & 0 & 1/3 & 0 \\ 0 & 2/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad Q = \begin{bmatrix} -3/2 & 3/2 & 0 & 0 \\ 3 & -9/2 & 3/2 & 0 \\ 0 & 3 & -9/2 & 3/2 \\ 0 & 0 & 3 & -3 \end{bmatrix}$$

transition rate matrix

embedded DTMC

infinitesimal generator matrix

CTMC path

A path ω is a sequence $s_0 t_0 s_1 t_1 s_2 t_2 ...$ such that

- $\forall i,\ R(s_i, s_{i+1})\, 0$ and $t_i \in \mathbb{R}_{>0}$
- t_i is the time spent in s_i

The path ω is finite if for some k, the state s_k is absorbing (i.e. $R(s, s') =$

Path(s) denotes all paths starting in s.

Simulation Algorithm

Main Idea

As the next state probability is independent from the time when the transition takes place, use two independent stochastic processes for s_i and t_i.

1. **Init** $i = 0$, $s_i = s_0$
2. **loop**
3. Pick $t_i \in \mathbb{R}_{>0}$ using exponential distribution with rate $E(s)$
4. Pick s_{i+1} using discrete distribution $P_R(s_i, s')$ of embedded DTMCs
5. $i = i + 1$
6. **end loop**

Sometimes referred as Gillespie's algorithm, and used by its author for stochastic simulation of chemical reactions

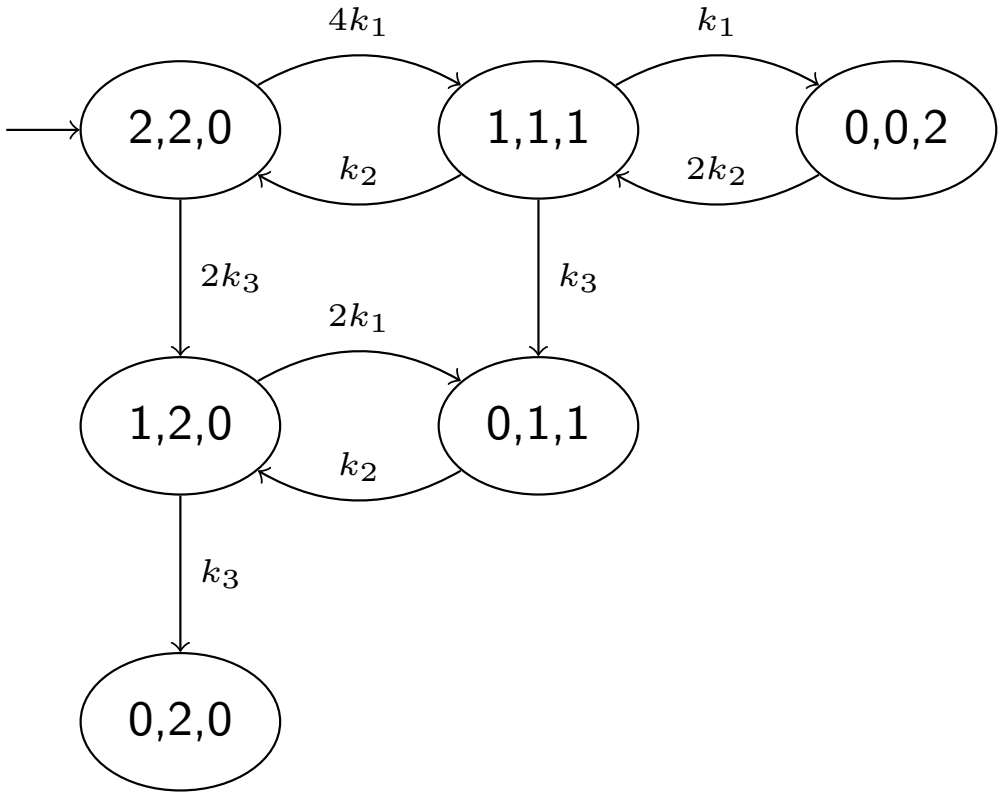
Example: A Chemical Reaction

- Three species: A, B and AB, three reactions:
  - A and B collide and produce AB $\quad A + B \xrightarrow{k_1} AB$
  - AB breaks into A and B $\quad AB \xrightarrow{k_2} A + B$
  - degradation of A $\quad A \xrightarrow{k_3} \emptyset$

Example: A Chemical Reaction

- Three species: A, B and AB, three reactions:
  - A and B collide and produce AB $\quad A + B \xrightarrow{k_1} AB$
  - AB breaks into A and B $\quad AB \xrightarrow{k_2} A + B$
  - degradation of A $\quad A \xrightarrow{k_3} \emptyset$

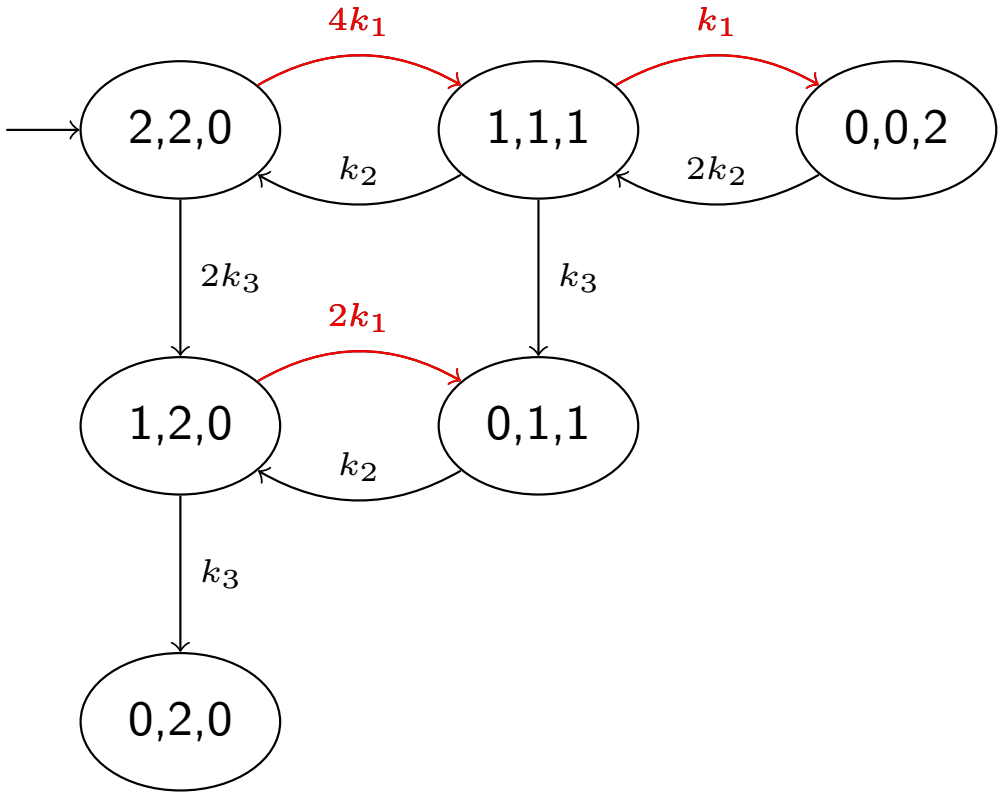
- CTMC with state-space $(\sharp A, \sharp B, \sharp AB)$



Example: A Chemical Reaction

- Three species: A, B and AB, three reactions:
  - A and B collide and produce AB $\quad A + B \xrightarrow{k_1} AB$
  - AB breaks into A and B $\quad AB \xrightarrow{k_2} A + B$
  - degradation of A $\quad A \xrightarrow{k_3} \emptyset$

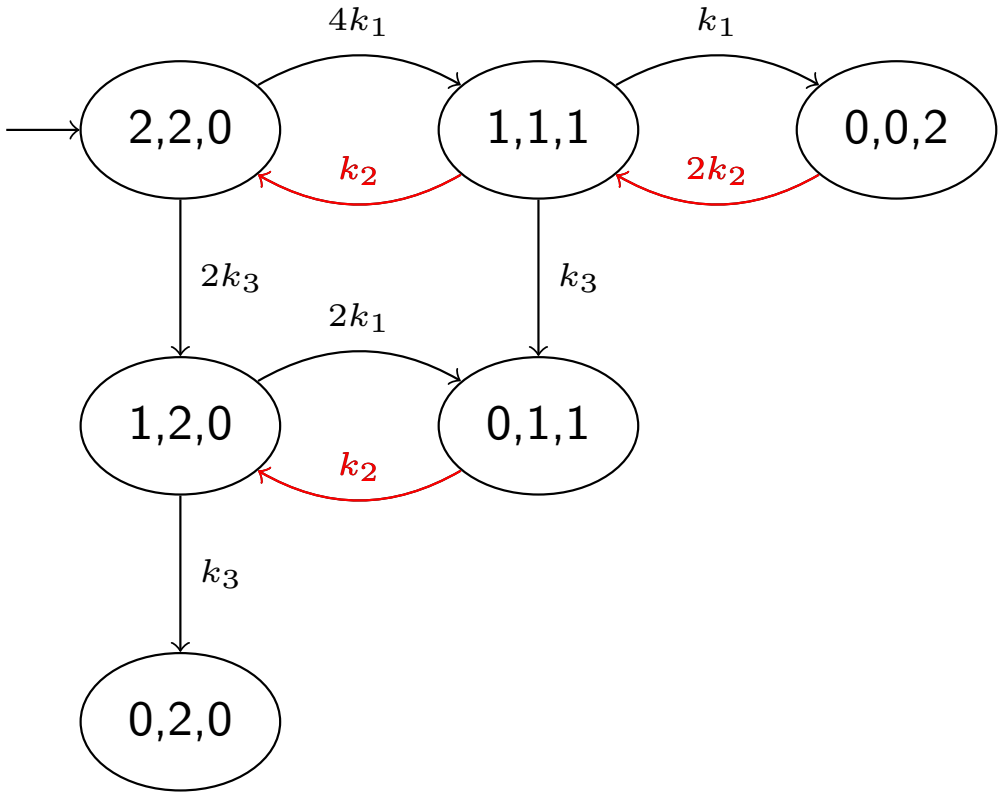
- CTMC with state-space $(\sharp A, \sharp B, \sharp AB)$
- A and B collide at rate $k_1 \sharp A \sharp B$



Example: A Chemical Reaction

- Three species: A, B and AB, three reactions:
  - A and B collide and produce AB $\quad A + B \xrightarrow{k_1} AB$
  - AB breaks into A and B $\quad AB \xrightarrow{k_2} A + B$
  - degradation of A $\quad A \xrightarrow{k_3} \emptyset$

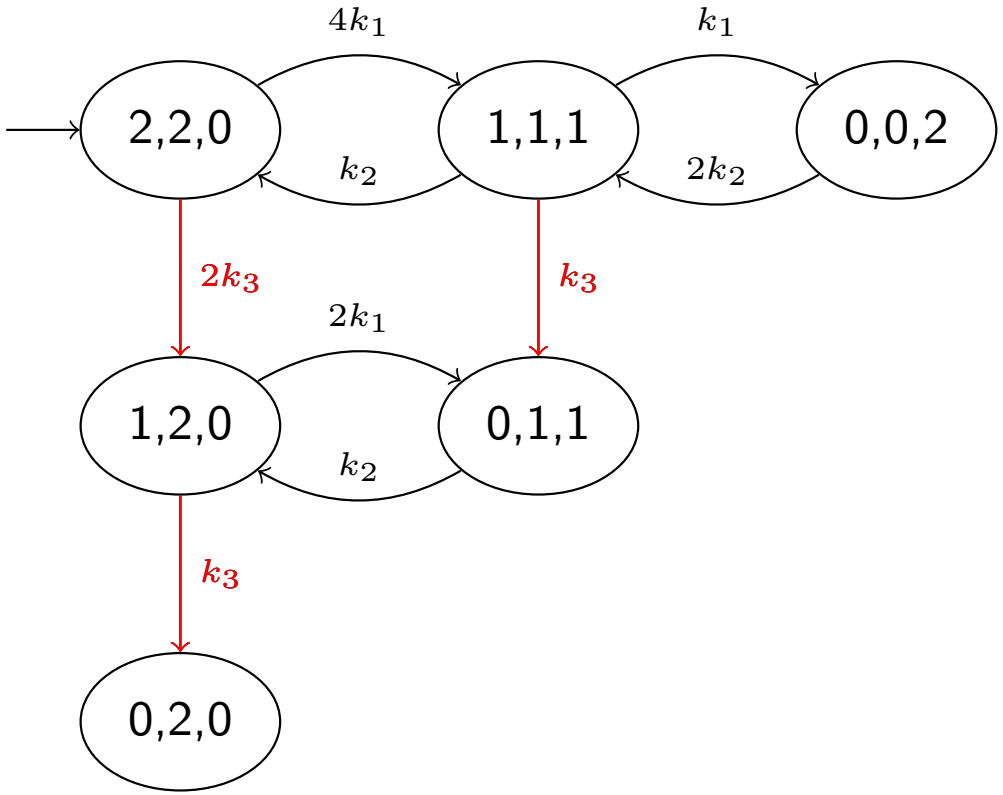
- CTMC with state-space $(\sharp A, \sharp B, \sharp AB)$
- A and B collide at rate $k_1 \sharp A \sharp B$
- AB breaks at rate $k_2 \sharp AB$



Example: A Chemical Reaction

- Three species: A, B and AB, three reactions:
  - A and B collide and produce AB $\quad A + B \xrightarrow{k_1} AB$
  - AB breaks into A and B $\quad AB \xrightarrow{k_2} A + B$
  - degradation of A $\quad A \xrightarrow{k_3} \emptyset$

- CTMC with state-space $(\sharp A, \sharp B, \sharp AB)$
- A and B collide at rate $k_1 \sharp A \sharp B$
- AB breaks at rate $k_2 \sharp AB$
- A degrades at rate $k_3 \sharp A$



CTMCs for Chemical Reactions

How many states does a CTMC of a chemical reaction has ?

CTMCs for Chemical Reactions

How many states does a CTMC of a chemical reaction has ?

It depends on

- The number of reactions
- The number species types
- The initial number of each specie

CTMCs for Chemical Reactions

How many states does a CTMC of a chemical reaction has ?

It depends on

- The number of reactions
- The number species types
- The initial number of each specie

If there is a production reaction, e.g., $\emptyset \to A$, the number can be infinite

Stochastic versus deterministic models of chemical reactions

Recall that we can formulate ODE to describe a deterministic evolution of the number of molecules (using mass-action laws)

Stochastic versus deterministic models of chemical reactions

Recall that we can formulate ODE to describe a deterministic evolution of the number of molecules (using mass-action laws)

The stochastic (CTMC) model is believed is to be more realistic, but can be quickly intractable.

Stochastic versus deterministic models of chemical reactions

Recall that we can formulate ODE to describe a deterministic evolution of the number of molecules (using mass-action laws)

The stochastic (CTMC) model is believed is to be more realistic, but can be quickly intractable.In general,

- For large populations of molecules the deterministic model is used
- For small populations use the stochastic model

CSL (slides: David Parker)

- Temporal logic for describing properties of CTMCs
  - CSL = Continuous Stochastic Logic [ASSB00,BHHK03]
  - extension of (non-probabilistic) temporal logic CTL
  - transient, steady-state and path-based properties
- Key additions:
  - probabilistic operator P (like PCTL)
  - steady state operator S
- Example: down $\rightarrow P_{>0.75}$ [$\neg$fail $U^{\leq[1,2.5]}$ up]
  - when a shutdown occurs, the probability of a system recovery being completed between 1 and 2.5 hours without further failure is greater than 0.75
- Example: $S_{<0.1}$[insufficient_routers]
  - in the long run, the chance that an inadequate number of routers are operational is less than 0.1

CSL Syntax (slides: David Parker)

- CSL syntax:
  - $\phi ::= \text{true} \mid a \mid \phi \wedge \phi \mid \neg\phi \mid P_{\sim p}[\psi] \mid S_{\sim p}[\phi]$ (state formulae)
  - $\psi ::= X\,\phi \mid \phi\, U^{I}\, \phi$ (path formulae)

ψ is true with probability ~p

"next"

"time bounded until"

in the "long run" ϕ is true with probability ~p

  - where a is an atomic proposition, I interval of $\mathbb{R}_{\geq 0}$, $p \in [0,1]$, and $\sim \in \{<,>,\leq,\geq\}$
  - unbounded until U is a special case: $\phi_1\, U\, \phi_2 \equiv \phi_1\, U^{[0,\infty)}\, \phi_2$

- Quantitative properties: $P_{=?}[\,\psi\,]$ and $S_{=?}[\,\phi\,]$
  - where P/S is the outermost operator

30

CSL Semantics (slides: David Parker)

- CSL formulae interpreted over states of a CTMC
  - $s \models \phi$ denotes ϕ is "true in state s" or "satisfied in state s"
- Semantics of state formulae:
  - for a state s of the CTMC (S, s_{init}, R, L):

  - $s \models a \quad \Leftrightarrow \quad a \in L(s)$
  - $s \models \phi_1 \wedge \phi_2 \quad \Leftrightarrow \quad s \models \phi_1$ and $s \models \phi_2$
  - $s \models \neg\phi \quad \Leftrightarrow \quad s \models \phi$ is false
  - $s \models P_{\sim p}[\psi] \quad \Leftrightarrow \quad Prob(s, \psi) \sim p$
  - $s \models S_{\sim p}[\phi] \quad \Leftrightarrow \quad \Sigma_{s' \models \phi} \underline{\pi}_s(s') \sim p$

Probability of, starting in state s, satisfying the path formula ψ

Probability of, starting in state s, being in state s' in the long run

CSL Semantics (slides: David Parker)

- Prob(s, ψ) is the probability, starting in state s, of satisfying the path formula ψ
  - $\text{Prob}(s, \psi) = \Pr_s \{\omega \in \text{Path}_s \mid \omega \models \psi\}$

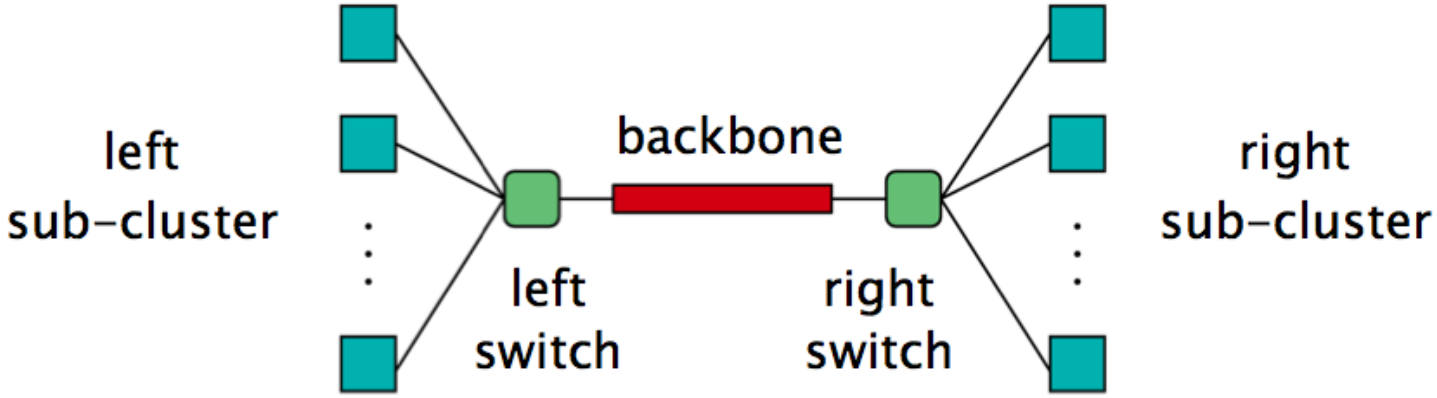
if ω(0) is absorbing, ω(1) not defined

- Semantics of path formulae:
  - for a path ω of the CTMC:
  - $\omega \models X\,\phi \quad \Leftrightarrow \quad \omega(1)$ is defined and $\omega(1) \models \phi$
  - $\omega \models \phi_1\, U^I\, \phi_2 \quad \Leftrightarrow \quad \exists t \in I.\ (\omega@t \models \phi_2 \wedge \forall t' < t.\ \omega@t' \models \phi_1)$

there exists a time instant in the interval I where ϕ_2 is true and ϕ_1 is true at all preceding time instants

CSL Example (slides: David Parker)

- Case study: Cluster of workstations [HHK00]
  - two sub-clusters (N workstations in each cluster)
  - star topology with a central switch
  - components can break down, single repair unit



  - minimum QoS: at least ¾ of the workstations operational and connected via switches
  - premium QoS: all workstations operational and connected via switches

33

CSL Example (slides: David Parker)

- $S_{=?}$ [minimum]
  - the probability in the long run of having minimum QoS
- $P_{=?}$ [$F^{[t,t]}$ minimum]
  - the (transient) probability at time instant t of minimum QoS
- $P_{<0.05}$ [$F^{[0,10]}$ $\neg$minimum]
  - the probability that the QoS drops below minimum within 10 hours is less than 0.05
- $\neg$minimum $\rightarrow$ $P_{<0.1}$ [$F^{[0,2]}$ $\neg$minimum]
  - when facing insufficient QoS, the chance of facing the same problem after 2 hours is less than 0.1

CSL Example (slides: David Parker)

- $\text{minimum} \rightarrow P_{>0.8}[\ \text{minimum}\ U^{[0,t]}\ \text{premium}\ ]$
  - the probability of going from minimum to premium QoS within t hours without violating minimum QoS is at least 0.8

- $P_{=?}[\ \neg\text{minimum}\ U^{[t,\infty)}\ \text{minimum}\ ]$
  - the chance it takes more than t time units to recover from insufficient QoS

- $\neg\text{r\_switch\_up} \rightarrow P_{<0.1}[\neg\text{r\_switch\_up}\ U\ \neg\text{l\_switch\_up}\ ]$
  - if the right switch has failed, the probability of the left switch failing before it is repaired is less than 0.1

- $P_{=?}[\ F^{[2,\infty)}\ S_{>0.9}[\ \text{minimum}\ ]\ ]$
  - the probability of it taking more than 2 hours to get to a state from which the long-run probability of minimum QoS is >0.9

35

CSL Model Checking

(See PRISM litterature for more details...)

- For untimed operators, equivalent to PCTL on embedded DTMCs
- For timed operators, can be reduced to computation of transient probabilities (such as matrix $P(t)$) -¿ complex
- An alternative is using Statistical Model Checking, approximate but more scalable

Statistical Model Checking

Assume we can decide $\omega \models \phi$. Based on simulations, decide hypothesis $H_1 : Pr_{\geq\theta}(\omega \models \phi)$ against $H_0 : Pr_{<\theta}(\omega \models \phi)$

Bounds on the error of choosing H_1 instead of H_0, depending on the number of positive runs