Probabilistic Model Checking

Lecture 7
Continuous-time Markov chains

Alessandro Abate

Department of Computer Science
University of Oxford

Time in DTMCs

- Time in a DTMC proceeds in discrete steps

- Two possible interpretations:
  - accurate model of (discrete) time units
    - e.g. clock ticks in model of an embedded device
  - time-abstract
    - 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

Overview

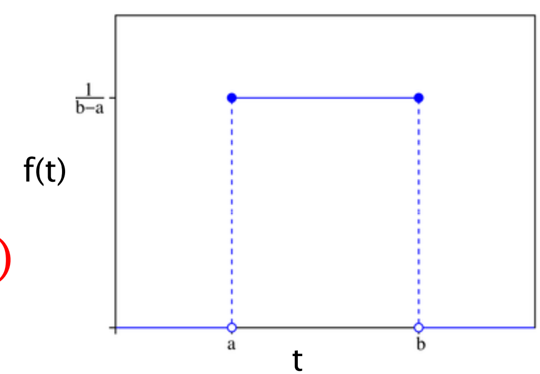
- Exponential distribution and its properties
- Continuous-time Markov chains (CTMCs)
  - definition, examples
  - race condition
  - embedded DTMC
  - generator matrix
- Paths and probabilities
  - probabilistic reachability

Continuous probability distributions

- Consider r.v. X defined by:
  - cumulative distribution function (cdf)

$$F(t) = \Pr(X \le t) = \int_{-\infty}^{t} f(x)\,dx$$

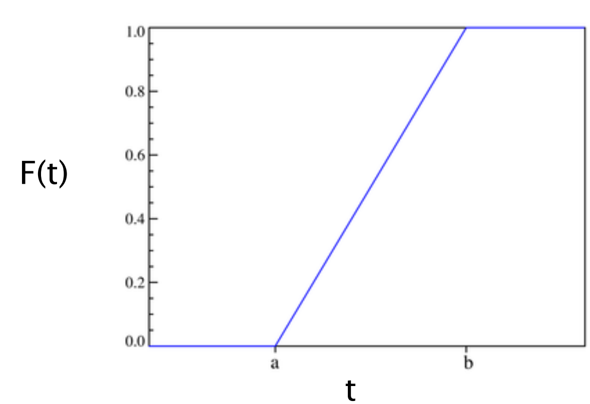
  - f being the probability density function (pdf)
  - $\Pr(X=t) = 0$ for all t



- Example: uniform distribution: U(a,b)

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

$$F(t) = \begin{cases} 0 & \text{if } t < a \\ \frac{t-a}{b-a} & \text{if } a \le t < b \\ 1 & \text{if } t \ge b \end{cases}$$



Exponential distribution

- A continuous random variable X is exponential with parameter $\lambda > 0$ if the density function is given by:

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

λ = "rate"

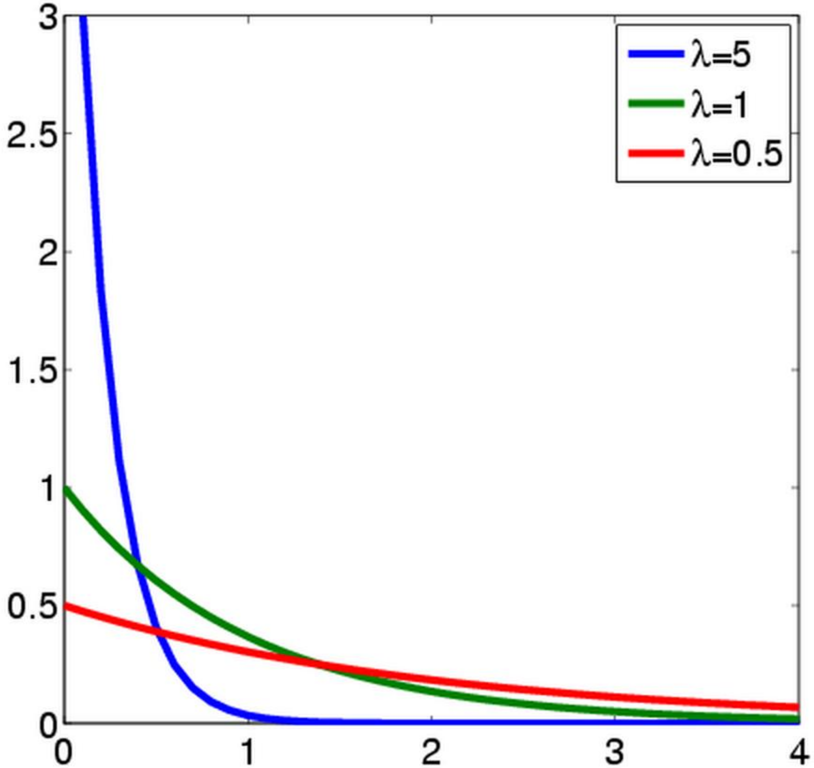
  - we write: $X \sim \text{Exponential}(\lambda)$
- Cumulative distribution function (for $t \geq 0$):

$$F(t) = \Pr(X \leq t) = \int_0^t \lambda \cdot e^{-\lambda \cdot x} dx = [-e^{-\lambda \cdot x}]_0^t = 1 - e^{-\lambda \cdot t}$$

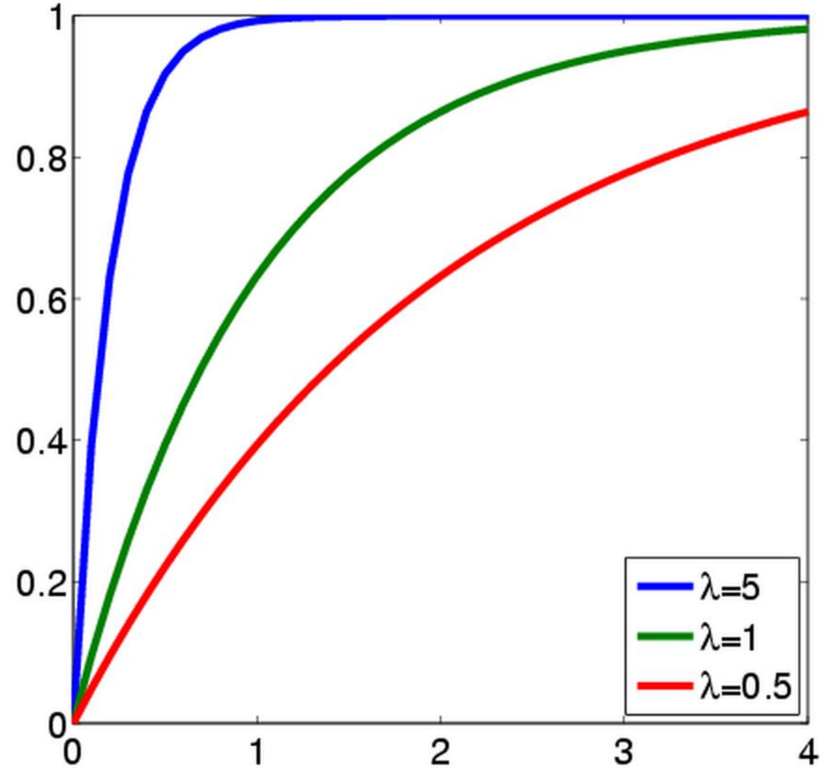
- Other properties:
  - negation: $\Pr(X > t) = e^{-\lambda \cdot t}$
  - mean (expectation): $E[X] = \int_0^\infty x \cdot \lambda \cdot e^{-\lambda \cdot x} dx = \frac{1}{\lambda}$
  - variance: $Var(X) = 1/\lambda^2$

Exponential distribution – Examples

Probability density function



Cumulative distribution function



- The larger the value of λ, the faster the c.d.f. approaches 1 (saturates)

Exponential distribution

- Adequate for modelling many real-life phenomena (constant rate, independent events)
  - Failures in process engineering
    - e.g. time before machine component fails
  - Inter-arrival times in communication engineering
    - e.g. time before next call/customer arrives to a call centre/shop
  - Biological/chemical systems
    - e.g. times within successive reactions between species

- Maximal entropy ("uncertainty") if just the mean is known
  - i.e. best approximation when only mean is known

- Can approximate general distributions arbitrarily closely
  - phase-type distributions

Exponential distribution – Property 1

- The exponential distribution has the memoryless property:
  - $\Pr(X>t_1+t_2 \mid X>t_1) = \Pr(X>t_2)$

- The exponential distribution is the only continuous distribution that is memoryless
  - discrete-time equivalent is the geometric distribution

Exponential distribution – Property 1

- The exponential distribution has the memoryless property:
  - $Pr(X>t_1+t_2 \mid X>t_1) = Pr(X>t_2)$

- $$\begin{aligned} Pr (X>t_1+t_2 \mid X>t_1) &= Pr(X>t_1+t_2 \wedge X>t_1) / Pr(X>t_1) \\ &= Pr(X>t_1+t_2) / Pr(X>t_1) \\ &= e^{-\lambda\cdot(t1+t2)} / e^{-\lambda\cdot t1} \\ &= (e^{-\lambda\cdot t1} \cdot e^{-\lambda\cdot t2}) / e^{-\lambda\cdot t1} \\ &= e^{-\lambda\cdot t2} \\ &= Pr(X>t_2) \end{aligned}$$

- The exponential distribution is the only continuous distribution that is memoryless
  - discrete-time equivalent is the geometric distribution

Exponential distribution – Property 2

- The minimum of two independent exponential distributions is an exponential distribution (parameter is sum)
  - $X_1 \sim \text{Exponential}(\lambda_1)$, $X_2 \sim \text{Exponential}(\lambda_2)$
  - $Y = \min(X_1, X_2)$
  - $Y \sim \text{Exponential}(\lambda_1 + \lambda_2)$
- Generalises to minimum of n distributions
- Maximum is not exponential

Exponential distribution – Property 2

- The minimum of two independent exponential distributions is an exponential distribution (parameter is sum)
  - $X_1 \sim \text{Exponential}(\lambda_1)$, $X_2 \sim \text{Exponential}(\lambda_2)$
  - $Y = \min(X_1, X_2)$

$$
\begin{aligned}
\Pr(Y \le t) &= \Pr(\min(X_1, X_2) \le t) \\
&= 1 - \Pr(\min(X_1, X_2) > t) \\
&= 1 - \Pr(X_1 > t \wedge X_2 > t) \\
&= 1 - \Pr(X_1 > t) \cdot \Pr(X_2 > t) \\
&= 1 - e^{-\lambda_1 \cdot t} \cdot e^{-\lambda_2 \cdot t} \\
&= 1 - e^{-(\lambda_1 + \lambda_2) \cdot t}
\end{aligned}
$$

  - $Y \sim \text{Exponential}(\lambda_1 + \lambda_2)$
- Generalises to minimum of n distributions
- Maximum is not exponential

Exponential distribution – Property 3

- Consider two independent exponential distributions
  - $X_1 \sim \text{Exponential}(\lambda_1)$, $X_2 \sim \text{Exponential}(\lambda_2)$
  - what is the probability that $X_1 < X_2$?

$$
\begin{aligned}
P(X_1 < X_2) &= P(\min\{X_1, X_2\} = X_1) \\
&= \int_0^\infty P(X_1 = x) P(X_2 > x) dx \\
&= \int_0^\infty \lambda_1 e^{-\lambda_1 x} e^{-\lambda_2 x} dx \\
&= \lambda_1 \int_0^\infty e^{-(\lambda_1 + \lambda_2)x} dx \\
&= \frac{\lambda_1}{\lambda_1 + \lambda_2}
\end{aligned}
$$

  - probability that $X_1 < X_2$ is $\lambda_1/(\lambda_1+\lambda_2)$
- Generalises to n distributions

Continuous-time Markov chains

- Continuous-time Markov chains (CTMCs)
  - labelled transition systems augmented with rates
  - discrete states
  - continuous time steps
  - delays exponentially distributed

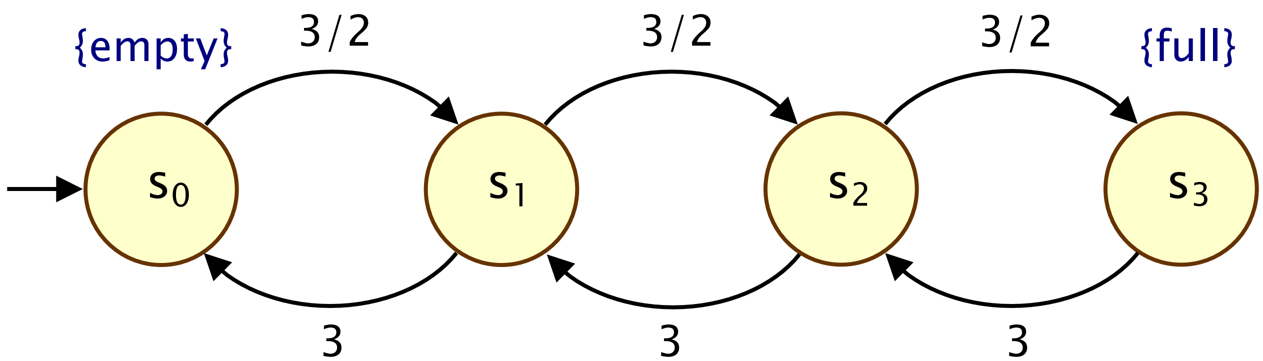
- Suited to modelling:
  - reliability/dependency models
  - control systems
  - queueing and communication networks
  - biological pathways
  - chemical reaction nets
  - DNA computing ...

Continuous-time Markov chains

- Formally, a CTMC C is a tuple $(S,s_{init},\mathbf{R},L)$ where:
  - S is a finite set of states ("state space")
  - $s_{init} \in S$ is the initial state
  - $\mathbf{R} : S \times S \to \mathbb{R}_{\geq 0}$ is the transition rate matrix
  - $L : S \to 2^{AP}$ is a labelling with atomic propositions
- Transition rate matrix assigns rates to each pair of states
  - used as a parameter to the exponential distribution
  - transition between s and s' when $\mathbf{R}(s,s')>0$
  - probability of transition before t time units: $1 - e^{-\mathbf{R}(s,s')\cdot t}$
- Assumption for this lecture
  - by convention, $\mathbf{R}(s,s)=0$ (can be generalised easily)

Simple CTMC example

- Modelling a queue of jobs
  - maximum size of the queue is 3
  - state space: $S = \{s_i\}_{i=0..3}$ where s_i indicates i jobs in queue
  - 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)



Race conditions

- What happens when there exists multiple s' with **R**(s,s')>0?
  - race condition: first transition triggered determines next state
  - two questions:
    - 1. How long is spent in s before a transition occurs?
    - 2. Which transition is eventually taken?

- 1. Time spent in a state before a transition
  - minimum of exponential distributions
  - exponential with parameter given by summation:

$$\mathrm{E}(s) = \sum_{s' \in S} \mathbf{R}(s, s')$$

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

Race conditions (cont'd)

- 2. Which transition is taken from state s?
  - the choice is independent of the time at which it occurs
  - e.g. if $X_1 \sim \text{Exponential}(\lambda_1)$, $X_2 \sim \text{Exponential}(\lambda_2)$
  - then the probability that $X_1 < X_2$ is $\lambda_1 / (\lambda_1 + \lambda_2)$
  - more generally, the probability is given by…
- The embedded DTMC: $emb(C) = (S, s_{init}, \mathbf{P}^{emb(C)}, L)$
  - state space, initial state and labelling as the CTMC
  - for any $s, s' \in S$

$$\mathbf{P}^{emb(C)}(s,s') = \begin{cases} R(s,s')/E(s) & \text{if } E(s) > 0 \\ 1 & \text{if } E(s) = 0 \text{ and } s = s' \\ 0 & \text{otherwise} \end{cases}$$

- Probability that next state from s is s' given by $\mathbf{P}^{emb(C)}(s,s')$

Two interpretations of a CTMC

- Consider a (non-absorbing) state $s \in S$ with multiple outgoing transitions, i.e. multiple $s' \in S$ with $\mathbf{R}(s,s')>0$

- 1. Race condition
  - each transition triggered after exponentially distributed delay
    - i.e. probability triggered before t time units: $1 - e^{-\mathbf{R}(s,s')\cdot t}$
  - first transition triggered determines the next state

- 2. Separate delay/transition
  - remain in s for delay exponentially distributed with rate E(s)
    - i.e. probability of taking an outgoing transition from s within [0,t] is given by $1-e^{-E(s)\cdot t}$
  - probability that next state is s' is given by $\mathbf{P}^{emb(C)}(s,s')$
    - i.e. $\mathbf{R}(s,s')/E(s) = \mathbf{R}(s,s') \ / \ \Sigma_{s'\in S} \mathbf{R}(s,s')$

More on CTMCs…

- Infinitesimal generator matrix Q

$$\mathbf{Q}(s,s') = \begin{cases} \mathbf{R}(s,s') & s \neq s' \\ -\sum_{s \neq s'} \mathbf{R}(s,s') & \text{otherwise} \end{cases}$$

- Alternative definition: a CTMC is:
  - a family of random variables $\{ X(t) \mid t \in \mathbb{R}_{\geq 0} \}$
  - $X(t)$ are observations made at time instant t
  - i.e. $X(t)$ is the state of the system at time instant t
  - which satisfies…

- Memoryless (Markov property)

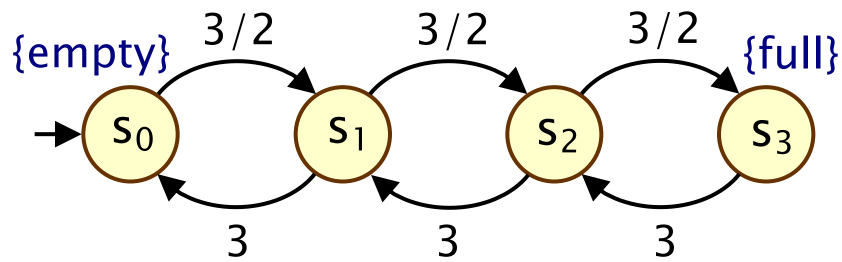
$$Pr(X(t_{k+1})=s_{k+1} \mid X(t_k)=s_k, \ldots, X(t_0)=s_0) = Pr(X(t_{k+1})=s_{k+1} \mid X(t_k)=s_k)$$

Simple CTMC 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 $L(s_3)=\{full\}$

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

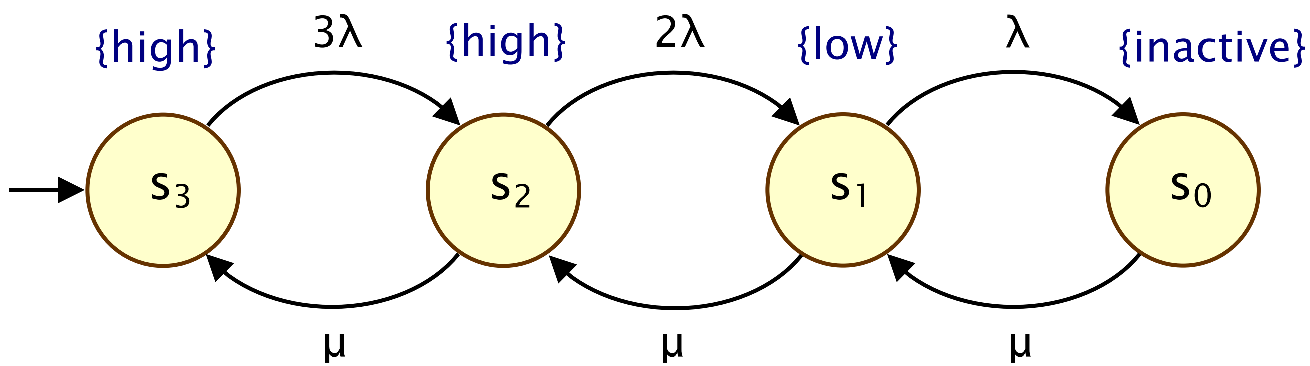
transition rate matrix

embedded DTMC

infinitesimal generator matrix

Example 2

- 3 machines, each can fail independently
  - delay modelled as exponential distributions
  - failure rate λ, i.e. mean-time to failure (MTTF) = 1/ λ
- One repair unit
  - repairs a single machine at rate μ (also exponential)
- State space:
  - $S = \{s_i\}_{i=0..3}$ where s_i indicates i machines operational

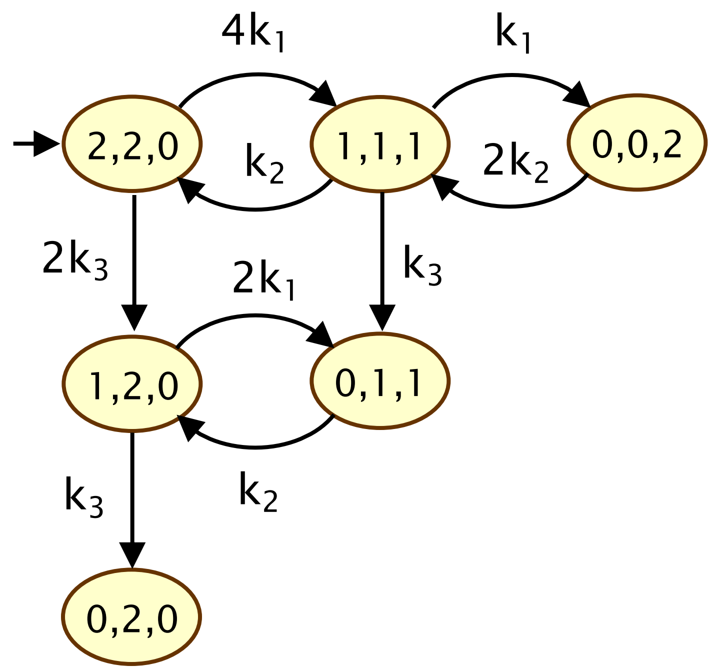


Example 3

- Chemical reaction system: two species A and B
- Two reactions:

$$A + B \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AB \qquad A \xrightarrow{k_3}$$

  - reversible reaction under which species A and B bind to form AB (forwards rate = $|A| \cdot |B| \cdot k_1$, backwards rate = $|AB| \cdot k_2$)
  - degradation of A (rate $|A| \cdot k_3$)
  - $|X|$ denotes number of molecules of species X



- CTMC with state space
  - $(|A|,|B|,|AB|)$
  - initially (2,2,0)

Paths of a CTMC

- An infinite path ω is a sequence $s_0t_0s_1t_1s_2t_2\ldots$ such that
  - $\mathbf{R}(s_i,s_{i+1}) > 0$ and $t_i \in \mathbb{R}_{>0}$ for all $i \in \mathbb{N}$
  - t_i denotes the amount of time spent in s_i
- or a sequence $s_0t_0s_1t_1s_2t_2\ldots t_{k-1}s_k$ such that
  - $\mathbf{R}(s_i,s_{i+1}) > 0$ and $t_i \in \mathbb{R}_{>0}$ for all $i<k$
  - where s_k is absorbing (i.e. $\mathbf{R}(s_k,s') = 0$ for all $s' \in S$)
  - i.e. it remains in state s_k indefinitely

- Path(s) denotes all infinite paths starting in state s
- Further notation:
  - time(ω,j) = amount of time spent in the jth state, i.e. t_j
  - ω@t = state occupied at time t
  - see e.g. [BHHK03, KNP07a] for precise definitions

Recall: Probability spaces

- A σ−algebra (or σ−field) on Ω is a set Σ of subsets of Ω closed under complementation and countable union, i.e.:
  - if $A \in \Sigma$, the complement $\Omega \setminus A$ is in Σ
  - if $A_i \in \Sigma$ for $i \in \mathbb{N}$, the union $\cup_i A_i$ is in Σ
  - the empty set ∅ is in Σ
- Elements of Σ are called measurable sets or events
- Theorem: For any set F of subsets of Ω, there exists a unique smallest σ−algebra on Ω containing F
- Probability space (Ω, Σ, Pr)
  - Ω is the sample space
  - Σ is the set of events: σ−algebra on Ω
  - Pr : Σ → [0,1] is the probability measure:

    $\Pr(\Omega) = 1$ and $\Pr(\cup_i A_i) = \Sigma_i \Pr(A_i)$ for countable disjoint A_i

Probability space

- Sample space: Path(s) (set of all inf. paths from a state s)
- Events: sets of infinite paths
- Basic events: cylinders
  - cylinders = sets of paths with common finite prefix
  - include time intervals in cylinders

- Finite prefix is a sequence $s_0,I_0,s_1,I_1,\ldots,I_{n-1},s_n$
  - $s_0,s_1,s_2,\ldots,s_n$ sequence of states where $\mathbf{R}(s_i,s_{i+1})>0$ for $i<n$
  - $I_0,I_1,I_2,\ldots,I_{n-1}$ sequence of non-empty intervals of $\mathbb{R}_{\geq 0}$

- Cylinder $Cyl(s_0,I_0,s_1,I_1,\ldots,I_{n-1},s_n)$ is the set of infinite paths:
  - $\omega(i)=s_i$ for all $i \leq n$ and $time(\omega,i) \in I_i$ for all $i < n$

Probability space

- Define probability measure over cylinders inductively
- $Pr_s(Cyl(s))=1$
- $Pr_s(Cyl(s,I,s_1,I_1,\ldots,I_{n-1},s_n,I',s'))$ equals:

$$Pr_s(Cyl(s,I,s_1,I_1,\ldots,I_{n-1},s_n)) \cdot P^{emb(C)}(s_n,s') \cdot \left(e^{-E(s_n)\cdot \inf I'} - e^{-E(s_n)\cdot \sup I'}\right)$$

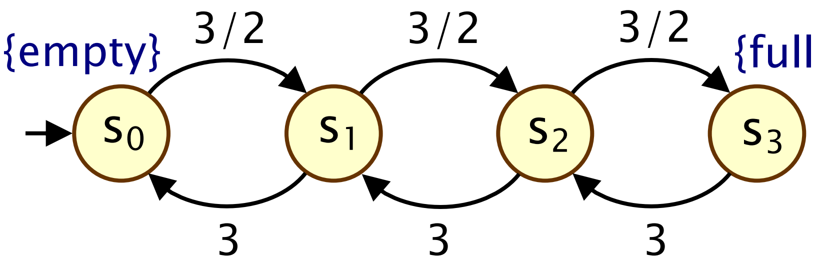
probability of transition from s_n to s' (defined using embedded DTMC)

probability of time spent in state s_n is within the interval I'

Probability space – Example

- Probability of leaving the initial state s_0 and moving to state s_1 within the first 2 time units of operation

- Cylinder $Cyl(s_0,(0,2],s_1)$



- $Pr_{s0}(Cyl(s_0,(0,2],s_1))$

$$= Pr_{s0}(Cyl(s_0)) \cdot \mathbf{P}^{emb(C)}(s_0,s_1) \cdot (e^{-E(s0)\cdot 0} - e^{-E(s0)\cdot 2})$$

$$= 1 \cdot \mathbf{P}^{emb(C)}(s_0,s_1) \cdot (e^{-E(s0)\cdot 0} - e^{-E(s0)\cdot 2})$$

$$= 1 \cdot 1 \cdot (e^{-3/2\cdot 0} - e^{-3/2\cdot 2})$$

$$= 1 - e^{-3}$$

$$\approx 0.95021$$

Probability space

- Probability space (Path(s), $\Sigma_{Path(s)}$, Pr_s)   (see [BHHK03])

- Sample space Ω = Path(s)
  - i.e. all infinite paths
- Event set $\Sigma_{Path(s)}$
  - least σ-algebra on Path(s) containing all cylinders sets $Cyl(s_0,I_0,\ldots,I_{n-1},s_n)$ where:
    - $s_0,\ldots,s_n$ ranges over all state sequences with $\mathbf{R}(s_i,s_{i+1})>0$ for all i
    - $I_0,\ldots,I_{n-1}$ ranges over all sequences of non-empty intervals in $\mathbb{R}_{\geq 0}$ (where intervals are bounded by rationals)
- Probability measure Pr_s
  - Pr_s extends uniquely from probability defined over cylinders

Probabilistic reachability

- Probabilistic reachability
  - the probability of reaching a target set $T \subseteq S$
  - measurability:
    - union of all basic cylinders $Cyl(s_0,(0,\infty),s_1,(0,\infty),\ldots,(0,\infty),s_n)$ where $s_n \in T$
    - set of state sequences $s_0 s_1 \ldots s_n$ is countable

- Time-bounded probabilistic reachability
  - the probability of reaching a target set $T \subseteq S$ within t time units
  - measurability:
    - union of all basic cylinders $Cyl(s_0,I_0,s_1,I_1,\ldots,I_{n-1},s_n)$ where $s_n \in T$ and $\sup(I_0)+\ldots+\sup(I_{n-1}) \leq t$
    - set of state sequences $s_0 s_1 \ldots s_n$ is countable
    - set of rational-bounded intervals is countable

Summing up…

- Exponential distribution
  - suitable for modelling failures, waiting times, reactions, …
  - nice mathematical properties

- Continuous-time Markov chains
  - transition delays modelled as exponential distributions
  - race condition
  - embedded DTMC
  - generator matrix

- Probability space over paths
  - (untimed and timed) probabilistic reachability