Continuous-Time Markov Chains

Lecture #20 of Probabilistic Models for Concurrency

Joost-Pieter Katoen

Software Modeling and Verification Group

E-mail: katoen@cs.rwth-aachen.de

July 16, 2009

© JPK

Overview Lecture #20

- Exponential distribution and its properties
- Continuous-time Markov chains (CTMCs)
  - definition and examples
  - race condition
  - transient distribution: uniformization
  - steady-state distribution

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 negative exponential distributions

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 of X is a continuous interval

- X is *continuously distributed* if there exists a function $f(x)$ such that:

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

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

  - $F_X(d) = \Pr\{X \leqslant d\}$ is the *(cumulative) probability distribution function*
  - $f(x)$ is the *probability density function*

Example

Exponential distribution

Continuous r.v. X is *exponential* with parameter $\lambda > 0$ if its density is

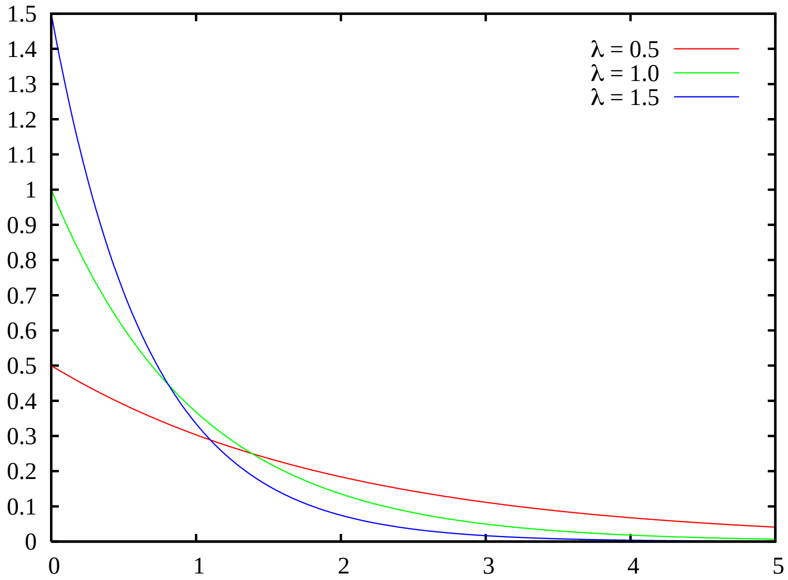
$$f(x) = \lambda{\cdot}e^{-\lambda \cdot x} \quad \text{for } x > 0 \quad \text{and } 0 \text{ otherwise}$$

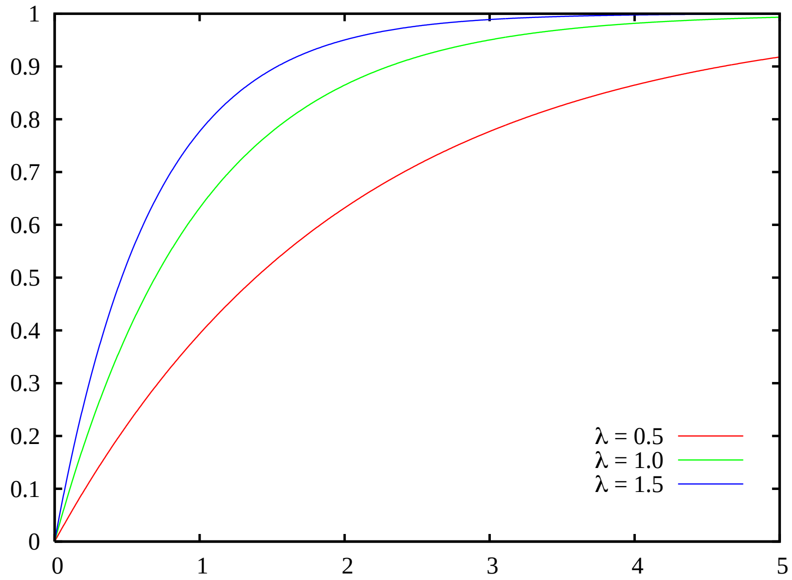
Cumulative distribution of X:

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

- $\Pr\{X > d\} = e^{-\lambda \cdot d}$
- expectation $E[X] = \int_0^\infty x{\cdot}\lambda{\cdot}e^{-\lambda \cdot x}\, dx = \frac{1}{\lambda}$
- variance *Var*$[X] = \frac{1}{\lambda^2}$

Exponential pdf and cdf





the higher λ, the faster the cdf approaches 1

Exponential distributions

- Have *nice mathematical* properties (cf. next slide)
- Are *adequate* for many real-life phenomena
  - the time until a radioactive particle decays
  - the constant hazard rate portion of the bathtub curve in reliability theory
  - the time it takes before your next telephone call
  - times for reactions between proteins to occur
- Can *approximate* general distributions arbitrarily closely
  - phase-type distributions
- Maximal *entropy* probability distribution if just the mean is known

Properties

- An exponential distribution possesses the *memory-less property*

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

  - exponential distributions are the only memoryless continuous distributions

- Let X and Y be exponential random variables with rate λ and μ
- $\min(X, Y)$ is exponentially distributed with rate $\lambda+\mu$
- $\Pr\{X = \min(X, Y)\} = \frac{\lambda}{\lambda+\mu}$
- $\max(X, Y)$ is not an exponential, but a phase-type distribution

$\Rightarrow$ *exponential distributions are closed under* $\min$*, but not under* $\max$

Proofs

Continuous-time Markov chain

- A *time-homogeneous continuous-time Markov chain* (CTMC) is
  - a Markov process
  - with continuous parameter T and discrete state space $X(t)$
  - which is time-homogeneous
- $p_s(t) = \Pr\{ X(t) = s \}$ probability to be in state s at time instant t
  - this depends on the starting distribution
- Probability of being in state s' at time t when in s at step $t' < t$:

$$
\begin{aligned}
p_{s,s'}(t', t) &= \Pr\{ X(t) = s' \mid X(t') = s \} \\
&= \Pr\{ X(t-t') = s' \mid X(0) = s \}
\end{aligned}
$$

Another perspective

A *continuous-time Markov chain* (CTMC) is a tuple $(S, \mathbf{R})$ where:

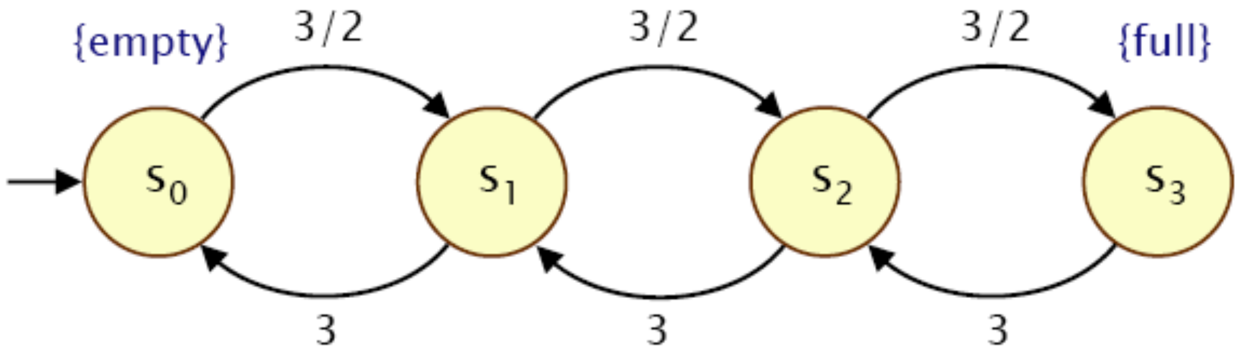
- S is a countable set of states
- $\mathbf{R} : S \times S \rightarrow \mathbb{R}_{\geqslant 0}$, a *rate matrix*
  - $\mathbf{R}(s, s') = \lambda$ means that the average speed of going from s to s' is $\frac{1}{\lambda}$
- $E(s) = \sum_{s' \in S} \mathbf{R}(s, s') = \mathbf{R}(s, S)$ is the *exit rate* of state s
  - s is called *absorbing* when $E(s) = 0$

a CTMC is a transition system (unlabeled transitions)
where transitions are equipped with continuous probabilities

A CTMC modeling a simple queue

Modelling a queue of jobs

- 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 queue



Modelling techniques for CTMCs

- Stochastic Petri nets [Molloy 1977]
- Markovian queueing networks [Kleinrock 1975]
- Stochastic automata networks [Plateau 1985]
- Stochastic activity networks [Meyer & Sanders 1985]
- Stochastic process algebra [Herzog *et al.*, Hillston 1993]
- Probabilistic input/output automata [Smolka *et al.* 1994]

and many variants thereof

Interpretation

- The probability that transition $s \longrightarrow s'$ is *enabled* in $[0, t]$:

$$1 - e^{-\mathbf{R}(s,s') \cdot t}$$

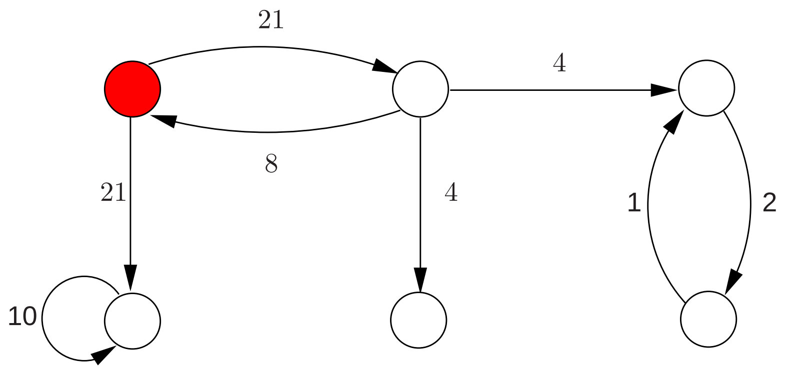
- The probability to *move* from non-absorbing s to s' in $[0, t]$ is:

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

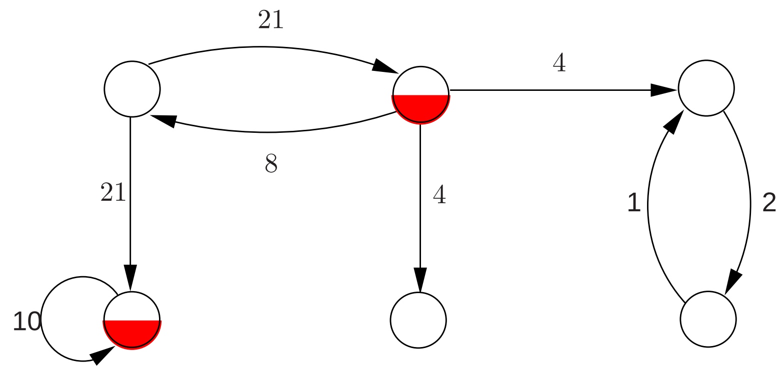
- The probability to take an outgoing transition from s within $[0, t]$ is:

$$1 - e^{-E(s) \cdot t}$$

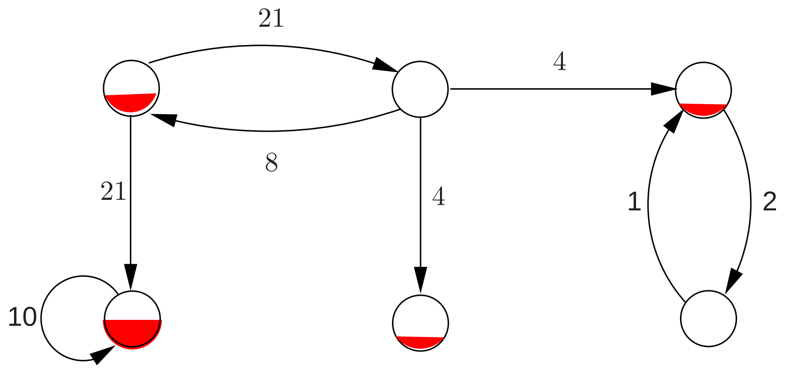
Time-abstract evolution of an example CTMC



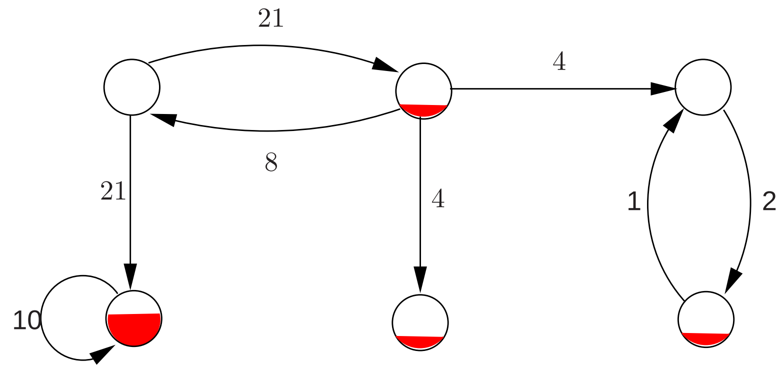
zero-th epoch



first epoch

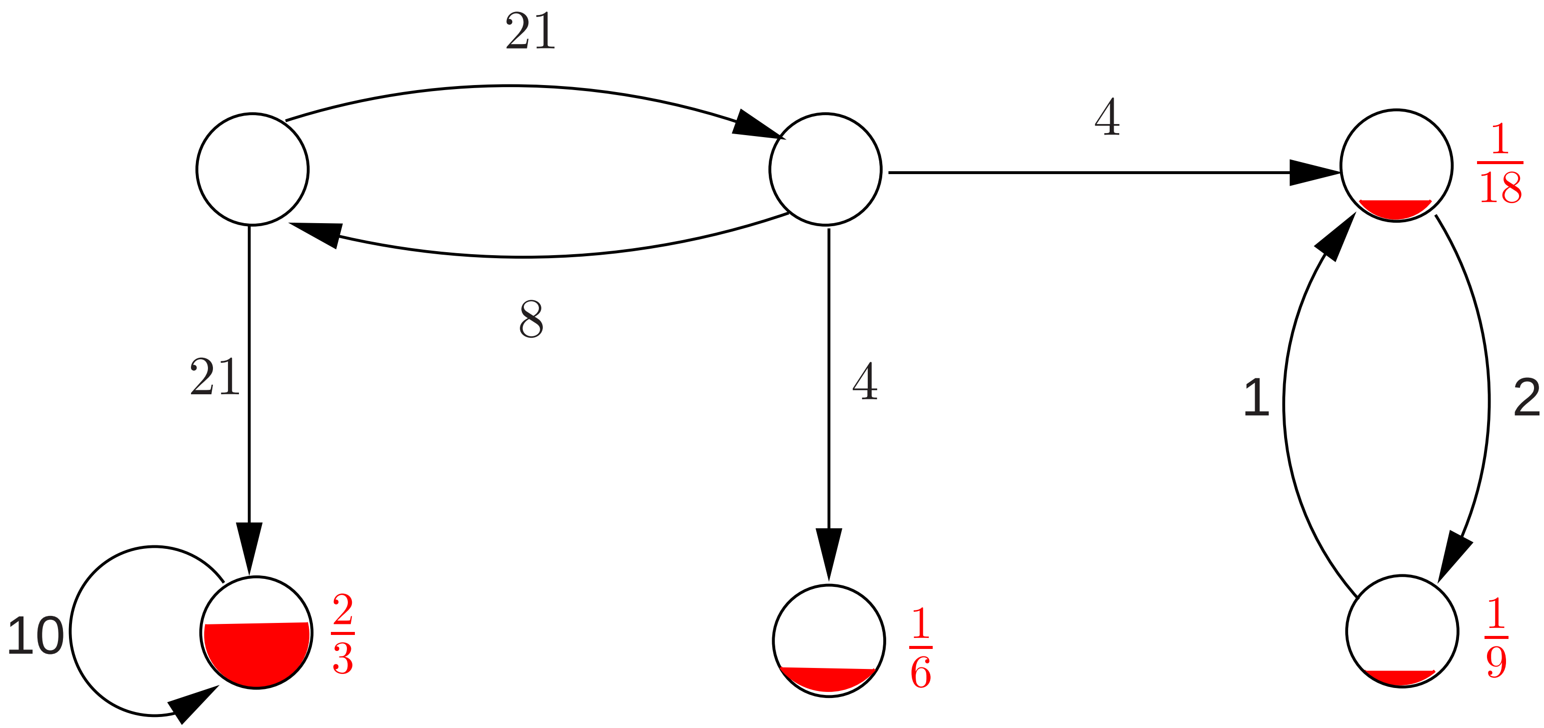


second epoch



third epoch

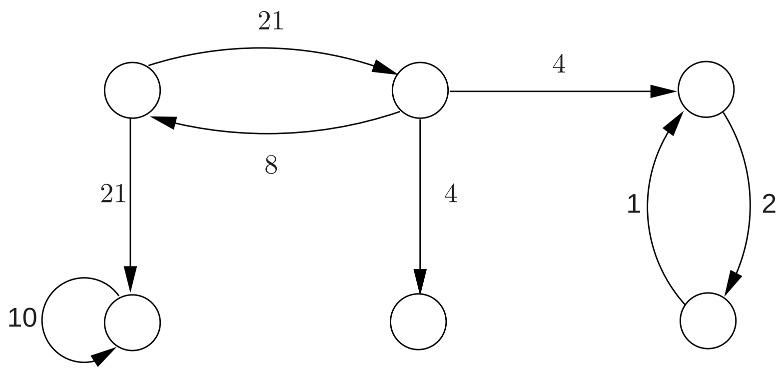
On the long run



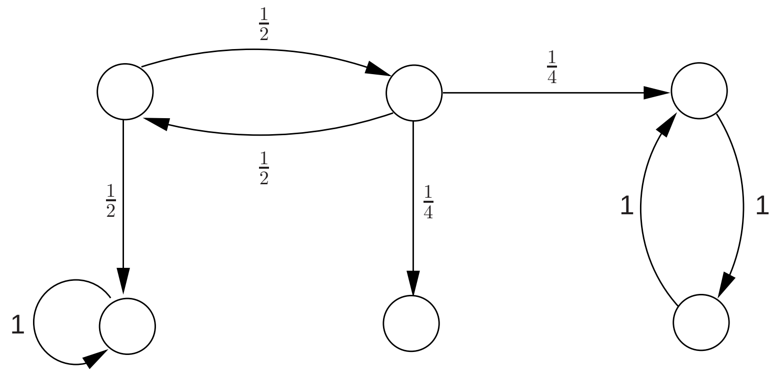
Embedded DTMC

The *embedded* DTMC of the CTMC $(S, \mathbf{R})$ is $(S, \mathbf{P})$ where

$$\mathbf{P}(s, s') = \begin{cases} \frac{\mathbf{R}(s,s')}{E(s)} & \text{if } E(s) > 0 \\ 0 & \text{otherwise} \end{cases}$$

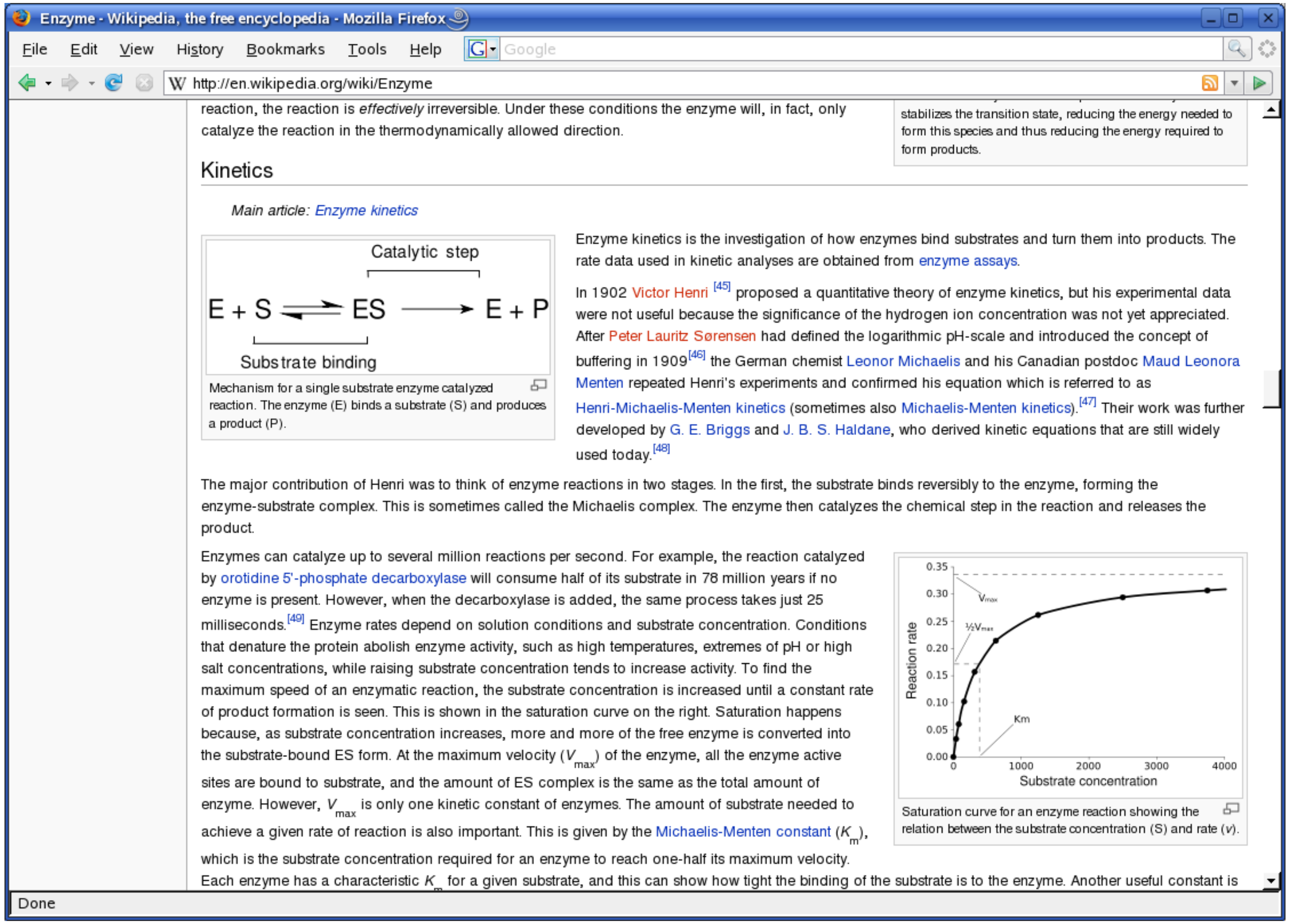


a CTMC



iits embedded DTMC

Enzyme-catalysed substrate conversion



reaction, the reaction is *effectively* irreversible. Under these conditions the enzyme will, in fact, only catalyze the reaction in the thermodynamically allowed direction.

stabilizes the transition state, reducing the energy needed to form this species and thus reducing the energy required to form products.

Kinetics

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 obtained from enzyme assays.

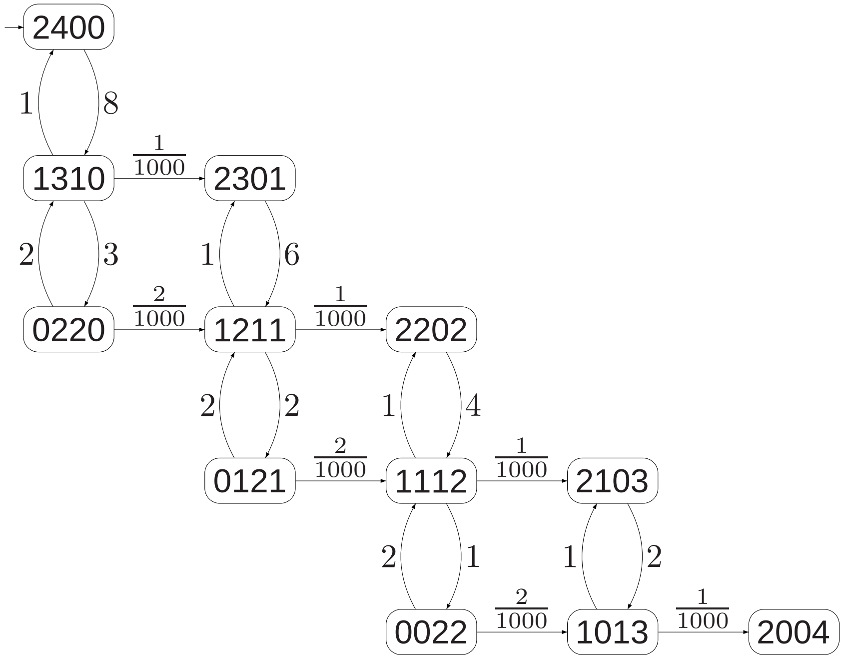
In 1902 Victor Henri [45] proposed a quantitative theory of enzyme kinetics, but his experimental data were not useful because the significance of the hydrogen ion concentration was not yet appreciated. After Peter Lauritz Sørensen had defined the logarithmic pH-scale and introduced the concept of buffering in 1909[46] 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 (sometimes also Michaelis-Menten kinetics).[47] Their work was further developed by G. E. Briggs and J. B. S. Haldane, who derived kinetic equations that are still widely used today.[48]

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.

Enzymes can catalyze up to several million reactions per second. For example, the reaction catalyzed by orotidine 5'-phosphate decarboxylase will consume half of its substrate in 78 million years if no enzyme is present. However, when the decarboxylase is added, the same process takes just 25 milliseconds.[49] Enzyme rates depend on solution conditions and substrate concentration. Conditions that denature the protein abolish enzyme activity, such as high temperatures, extremes of pH or high salt concentrations, while raising substrate concentration tends to increase activity. To find the maximum speed of an enzymatic reaction, the substrate concentration is increased until a constant rate of product formation is seen. This is shown in the saturation curve on the right. Saturation happens because, as substrate concentration increases, more and more of the free enzyme is converted into the substrate-bound ES form. At the maximum velocity (V_{max}) of the enzyme, all the enzyme active sites are bound to substrate, and the amount of ES complex is the same as the total amount of enzyme. However, V_{max} is only one kinetic constant of enzymes. The amount of substrate needed to achieve a given rate of reaction is also important. This is given by the Michaelis-Menten constant (K_m), which is the substrate concentration required for an enzyme to reach one-half its maximum velocity. Each enzyme has a characteristic K_m for a given substrate, and this can show how tight the binding of the substrate is to the enzyme. Another useful constant is

Saturation curve for an enzyme reaction showing the relation between the substrate concentration (S) and rate (*v*).

Enzyme-catalysed substrate conversion as a CTMC



States:

	init	*goal*
enzymes	2	2
substrate molecules	4	0
complex molecules	0	0
product molecules	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 $x_C > 0$

Transient distribution of a CTMC

Probability to be in state s at time t:

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

Using $\underline{p}(t) = (p_{s_0}(t), p_{s_1}(t), \ldots, p_{s_k}(t))$ we obtain in matrix form:

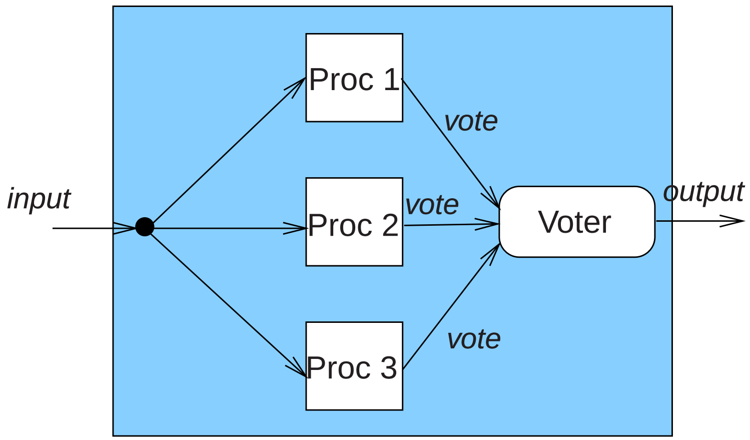
$$\underline{p}'(t) = \underline{p}(t) \cdot \mathbf{Q} \quad \text{given} \quad \underline{p}(0)$$

where $\mathbf{Q} = \mathbf{R} - diag(E)$ is the infinitesimal generator matrix

$\underline{p}(t)$ is the transient-state probability vector at time t

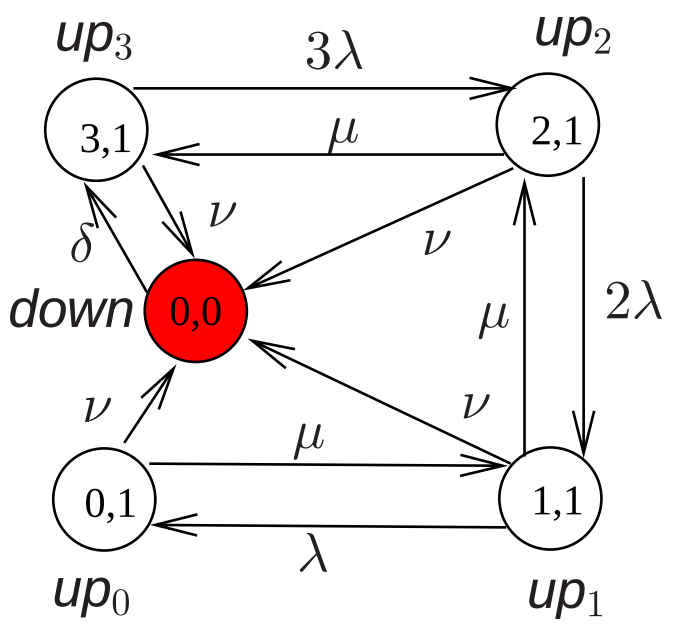
A triple modular redundant system

- 3 processors and a single voter:
  - processors run same program; voter takes a majority vote
  - each component (processor and voter) is failure-prone
  - there is a single repairman for repairing processors and voter



- Modelling assumptions:
  - if voter fails, entire system goes down
  - after voter-repair, system starts "as new"
  - state $= (\#\text{processors}, \#\text{voters})$

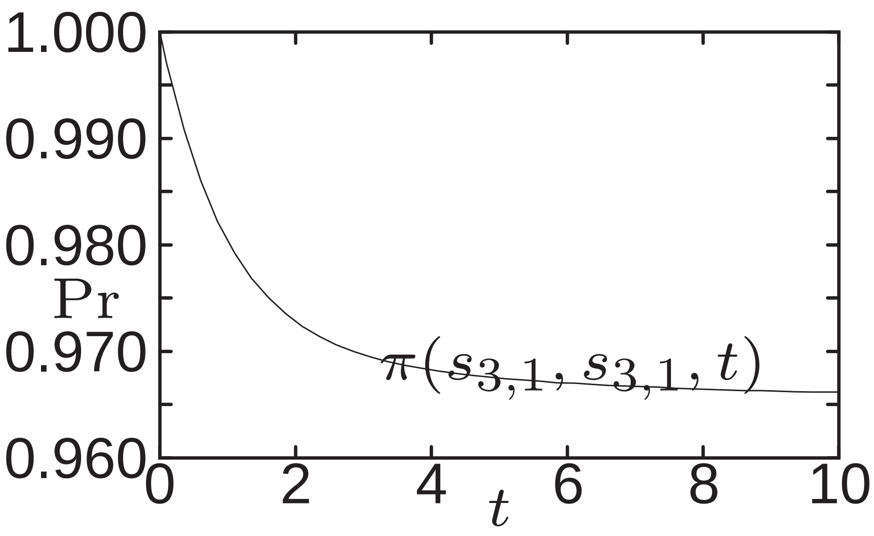
Modelling a TMR system as a CTMC



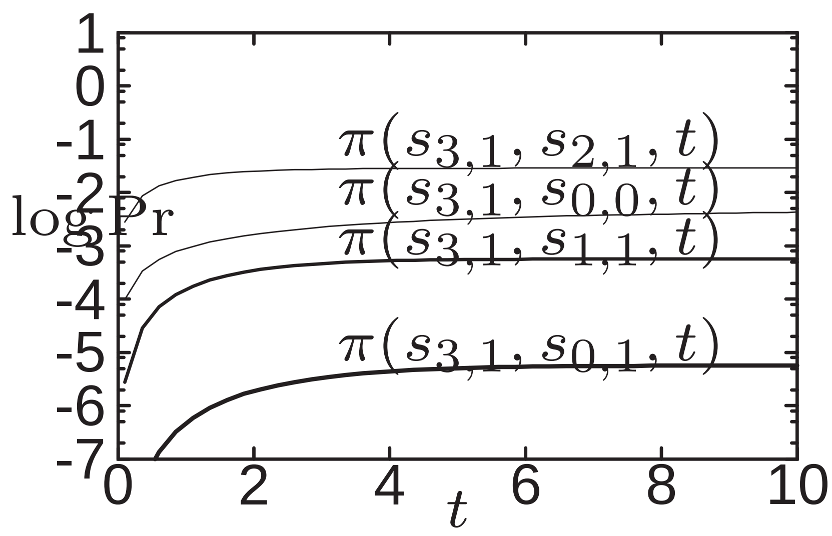
- processor failure rate is λ fph; its repair rate is μ rph
- voter failure rate is ν fph; its repair rate is δ rph
- rate matrix: e.g., $\mathbf{R}((3,1),(2,1)) = 3\lambda$
- exit rates: e.g., $E(3,1) = 3\lambda{+}\nu$
- probability matrix: e.g.,

$$\mathbf{P}((3,1),(2,1)) = \frac{3\lambda}{3\lambda{+}\nu}$$

Transient probabilities



$p_{s_{3,1}}(t)$ for first 10 hours



$p(t)$ for first 10 hours (logscale)

$\lambda = 0.01$ failures per hour (fph), $\nu = 0.001$ fph
$\mu = 1$ repairs per hour (rph) and $\delta = 0.2$ rph

(© B.R. Haverkort)

Steady-state distribution of a CTMC

Assuming a stationary distribution exists (e.g., finite and irreducible):

$$p_s \; = \; \lim_{t \to \infty} p_s(t) \quad \Leftrightarrow \quad \lim_{t \to \infty} p'_s(t) \; = \; 0 \quad \Leftrightarrow \quad \lim_{t \to \infty} p_s(t) \cdot \mathbf{Q} \; = \; 0$$

Using $\underline{p} = (p_{s_0}, p_{s_1}, \ldots, p_{s_k})$ we obtain in matrix form:

$$\underline{p} \cdot \mathbf{Q} \; = \; 0 \quad \text{where} \quad \textstyle\sum_{s \in S} p_s = 1$$

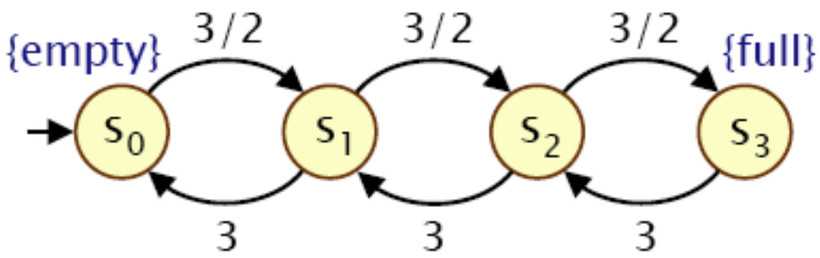
$\underline{p}$ is the steady-state probability vector

© JPK

Steady-state distribution: example

- Solve: $\underline{\pi} \cdot \mathbf{Q} = 0$ and $\Sigma\, \underline{\pi}(s) = 1$

$$\mathbf{Q} = \begin{bmatrix} -3/2 & 3/2 & 0 & 0 \\ 3 & -9/2 & 3/2 & 0 \\ 0 & 3 & -9/2 & 3/2 \\ 0 & 0 & 3 & -3 \end{bmatrix}$$



$$\begin{aligned}
-3/2 \cdot \underline{\pi}(s_0) + 3 \cdot \underline{\pi}(s_1) &= 0 \\
3/2 \cdot \underline{\pi}(s_0) - 9/2 \cdot \underline{\pi}(s_1) + 3 \cdot \underline{\pi}(s_2) &= 0 \\
3/2 \cdot \underline{\pi}(s_1) - 9/2 \cdot \underline{\pi}(s_2) + 3 \cdot \underline{\pi}(s_3) &= 0 \\
3/2 \cdot \underline{\pi}(s_2) - 3 \cdot \underline{\pi}(s_3) &= 0 \\
\underline{\pi}(s_0) + \underline{\pi}(s_1) + \underline{\pi}(s_2) + \underline{\pi}(s_3) &= 1
\end{aligned}$$

$\underline{\pi} = [\, 8/15,\ 4/15,\ 2/15,\ 1/15 \,]$

Steady-state distribution

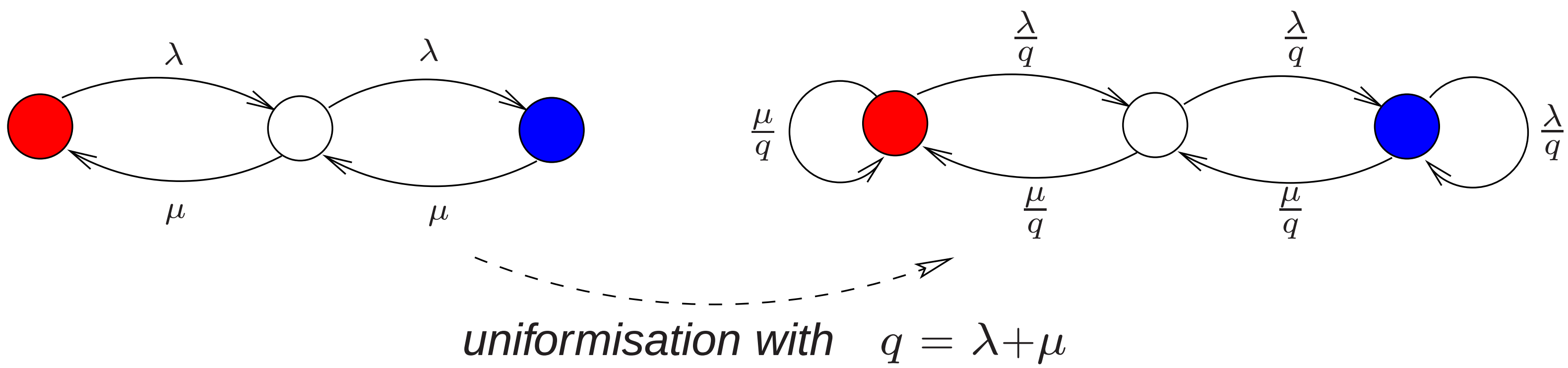
s	$s_{3,1}$	$s_{2,1}$	$s_{1,1}$	$s_{0,1}$	$s_{0,0}$
$p(s)$	$9.655 \cdot 10^{-1}$	$2.893 \cdot 10^{-2}$	$5.781 \cdot 10^{-4}$	$5.775 \cdot 10^{-6}$	$4.975 \cdot 10^{-3}$

The probability of $\geqslant$ two processors and the voter are up is 0.994

$\lambda = 0.01$ failures per hour (fph), $\nu = 0.001$ fph
$\mu = 1$ repairs per hour (rph) and $\delta = 0.2$ rph

Computing transient probabilities

- Solution to $\underline{p}'(t) = \underline{p}(t)\cdot\mathbf{Q}$ is: $\underline{p}(t) = \underline{p}(0)\cdot e^{\mathbf{Q}\cdot t} = \underline{p}(0) \cdot \sum_{i=0}^{\infty} \frac{(\mathbf{Q}\cdot t)^i}{i!} \quad (*)$
- Main problems: infinite summation + numerical instability due to
  - $\mathbf{Q}^i$ becomes non-sparse with positive and negative entries
- Solution: transform CTMC $(S, \mathbf{R})$ into DTMC $(S, \mathbf{U})$ with
  - $\mathbf{U} := \mathbf{I} + \frac{\mathbf{Q}}{q}$ with $q \geqslant \max_i \{ E(s_i) \}$



Uniformization

- Uniformised DTMC unif(C) of CTMC C $=(S,s_{init},\mathbf{R},L)$:
  - unif(C) $= (S,s_{init},\mathbf{P}^{unif(C)},L)$
  - set of states, initial state and labelling the same as C
  - $\mathbf{P}^{unif(C)} = \mathbf{I} + \mathbf{Q}/q$
  - $\mathbf{I}$ is the $|S|\times|S|$ identity matrix
  - $q \geq \max \{ E(s) \mid s \in S \}$ is the uniformisation rate

- Each time step (epoch) of uniformised DTMC corresponds to one exponentially distributed delay with rate q
  - if E(s)=q transitions the same as embedded DTMC (residence time has the same distribution as one epoch)
  - if E(s)<q add self loop with probability 1−E(s)/q (residence time longer than 1/q so one epoch may not be 'long enough')

Computing transient probabilities

- Now (*): $\underline{p}(t) = \underline{p}(0) \cdot e^{q(\mathbf{U}-\mathbf{I})t} = \underline{p}(0) \cdot e^{-qt} \cdot e^{qt\mathbf{U}} = \sum_{i=0}^{\infty} \underbrace{e^{-qt} \frac{(qt)^i}{i!}}_{\text{Poisson prob.}} \underline{p}(i)$

- Summation can be truncated *a priori* for a given error bound ε:

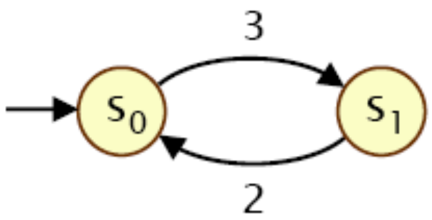
$$\left\| \sum_{i=0}^{\infty} e^{-qt} \frac{(qt)^i}{i!} \underline{p}(i) - \sum_{i=0}^{k_\varepsilon} e^{-qt} \frac{(qt)^i}{i!} \underline{p}(i) \right\| = \left\| \sum_{i=k_\varepsilon+1}^{\infty} e^{-qt} \frac{(qt)^i}{i!} \underline{p}(i) \right\|$$

- Choose k_ε minimal s.t.: $\sum_{i=k_{\varepsilon+1}}^{\infty} e^{-qt} \frac{(qt)^i}{i!} = 1 - \sum_{i=0}^{k_\varepsilon} e^{-qt} \frac{(qt)^i}{i!} \leqslant \varepsilon$

$\Rightarrow$ Transient analysis of a CTMC $\approx$ transient analysis of a DTMC

Transient probabilities: example

- CTMC C, uniformised DTMC for q=3



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

- Initial distribution: $\underline{\pi}_{s0,0} = [\,1, 0\,]$
- Transient probabilities for time $t = 1$:

$$\underline{\pi}_{s0,1} = \sum_{i=0}^{\infty} \gamma_{q\cdot t,i} \cdot \underline{\pi}_{s0,0} \cdot \left(\mathbf{P}^{\text{unif}(C)}\right)^{i}$$

$$= \gamma_{3,0} \cdot [1,0] \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \gamma_{3,1} \cdot [1,0] \cdot \begin{bmatrix} 0 & 1 \\ 2/3 & 1/3 \end{bmatrix} + \gamma_{3,2} \cdot [1,0] \cdot \begin{bmatrix} 0 & 1 \\ 2/3 & 1/3 \end{bmatrix}^{2} + \ldots$$

$$\approx [\,0.404043, 0.595957\,]$$