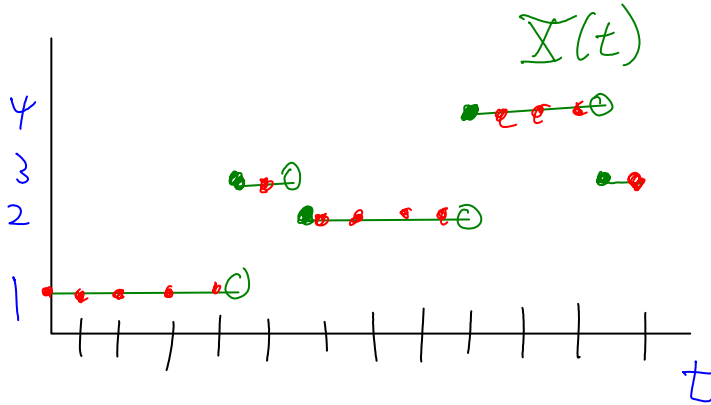


## Continuous-time Markov Chains

Monday, October 27, 2008  
12:00 PM

References: Lawler Ch. 3, Karlin and Taylor Ch. 4

We allow the state space to be finite or countable but now the time parameter is continuous over a parameter domain which is a subset of the real line.



The continuous-time Markov chain is piecewise constant with right-hand limits (right-continuous), jumping between states at random times which can be any value along the continuous time axis.  
(cadlag process)

Markov property in continuous time:

$$P(X(t) = j \mid X(t_1) = i_1, X(t_2) = i_2, \dots, X(t_n) = i_n) \\ = P(X(t) = j \mid X(t_n) = i_n) \\ \text{for any } 0 \leq t_1 < t_2 < \dots < t_n < t$$

This suffices for our present purposes but can also be generalized using filtrations to refer to all information available up to a given time  $t_n$ .

Most of the powerful tools for continuous-time Markov chains apply when the dynamics are time-homogenous.

$$P(X(t) = j \mid X(s) = i) = P_{ij}(t-s) \\ \text{for } s < t$$

What are the relationships and relative advantages of discrete and continuous-time Markov chains?

From the Markov property, one can show that observing a continuous-time

Markov chain at regular time intervals  $\{n \Delta t\}$  gives a discrete-time Markov chain.

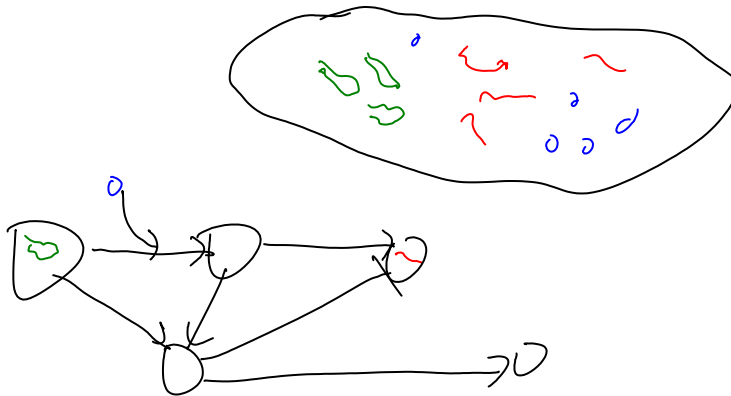
$X_n = X(n \Delta t)$  is discrete-time MC

This means in particular that continuous-time Markov chains can be approximated by discrete-time Markov chains, so what is the motivation for defining a continuous-time Markov chain model?

- When it's important to observe visits to important short-lived states.
- Of wider importance is simply the relative ease in modeling **instantaneous** changes rather than changes over finite times in physics, engineering, chemical applications.
  - One reason this is useful is that different influences on a system tend to behave additively instantaneously, but not additively over a finite time period.
- Continuous time is natural when one moves to continuous space
- Continuous-time Markov chains can actually be easier to simulate than discrete-time Markov chains!

Applications:

- statistical modeling in quantum mechanics
- chemical reactions, particularly biochemical networks in cells

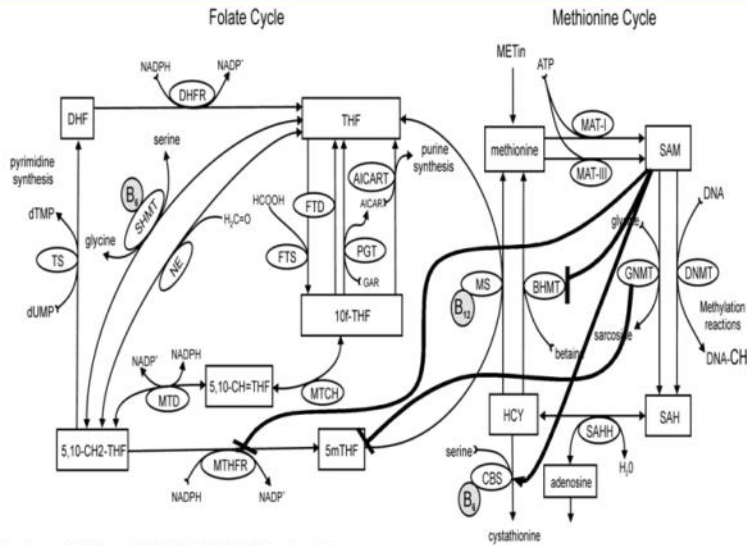


The novelty of biochemical reaction networks in the cell is that they tend to involve populations of biomolecules that are not large enough for the usual mass-action laws to be good models.

The scientists studying these networks are now increasingly turning toward continuous-time Markov chain models (master equation models).

State of the system is defined by the number of each relevant biomolecule (high-dimensional!).

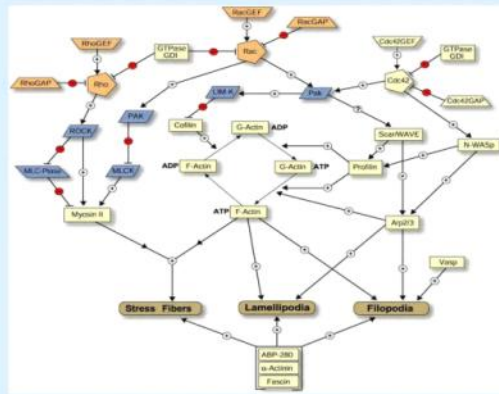
**The reaction scheme for Folate-Mediated One-Carbon Metabolism:  
Note nonlocal interactions!**



Reed, M. C. et al. J. Nutr. 2006;136:2653-2661, adapted as per outstanding survey talk of Michael Reed (Duke): <http://www.ima.umn.edu/videos/?id=392>  
Copyright ©2006 American Society for Nutrition

## Network Analysis

### Details of the actin control network



Hans Othmer, Minnesota

Another recent application of continuous-time Markov chain is to population and epidemic dynamics on networks.

Mathematical formulation for time-homogenous, continuous-time Markov chains

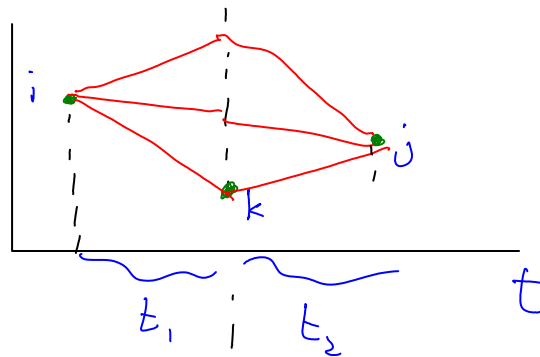
$$P_{ij}(t) = P(X(t+t') = j \mid X(t') = i)$$

The Markov property implies that we should be able to express these probability transition functions in terms of their behavior at infinitesimally small times.

Reason: **Chapman-Kolmogorov equation**

$$P_{ij}(t_1+t_2) = \sum_{k \in S} P_{ik}(t_1) P_{kj}(t_2)$$

for any  $t_1, t_2 > 0$



This would seem to allow us to break the calculation of the probability transition (matrix) function  $P_{ij}(t)$  in terms of its behavior on shorter and shorter time intervals.

So let's consider what happens over an infinitesimal time interval.

$$P(t) = \left\{ P_{ij}(t) \right\}_{\substack{i \in S \\ j \in S}}$$

Taylor-expand this probability transition function for small  $t$ :

$$P(t) = P(0) + P'(0)t + o(t)$$

$$P(0) = \mathbb{I} \quad \left( P_{ij}(0) = \delta_{ij} = \begin{cases} 1 & \text{for } i=j \\ 0 & \text{for } i \neq j \end{cases} \right)$$

↑  
identity matrix

$$P'(0) \equiv A \quad \text{will be determined by the model and is known as}$$

the **transition rate matrix** or **infinitesimal generator**.

Actually the dynamics of the continuous-time Markov chain are entirely encoded in this matrix  $A$ .

What are the entries of the transition rate matrix?

$$A = \lim_{\Delta t \downarrow 0} \frac{P(\Delta t) - I}{\Delta t}$$

Off-diagonal terms:

$$A_{ij} = \lim_{\Delta t \downarrow 0} \frac{P_{ij}(\Delta t)}{\Delta t} = \lim_{\Delta t \downarrow 0} \frac{P(X(t+\Delta t)=j \mid X(t)=i)}{\Delta t}$$

just describes the rate at which a transition is made from state  $i$  to state  $j$ .

Diagonal terms:

$$\begin{aligned} A_{ii} &= \lim_{\Delta t \downarrow 0} \frac{P_{ii}(\Delta t) - 1}{\Delta t} = \lim_{\Delta t \downarrow 0} \frac{P(X(t+\Delta t)=i \mid X(t)=i) - 1}{\Delta t} \\ &= \lim_{\Delta t \downarrow 0} - \frac{P(X(t+\Delta t) \neq i \mid X(t)=i)}{\Delta t} \\ &= \lim_{\Delta t \downarrow 0} - \sum_{\substack{j \in S \\ j \neq i}} \frac{P(X(t+\Delta t)=j \mid X(t)=i)}{\Delta t} \end{aligned}$$

(sum over partition of event  $\{X(t+\Delta t) \neq i\} = \bigsqcup_{\substack{j \in S \\ j \neq i}} \{X(t+\Delta t)=j\}$ )  
 commute limit + sum; should be careful

$$A_{ii} = - \sum_{\substack{j \in S \\ j \neq i}} A_{ij}$$

This can be interpreted as the **total transition rate** out of state  $i$ .

Note that this relationship shows that the row sums of the transition rate matrix is equal to zero:

$$\sum_{j \in S} A_{ij} = 0$$

$$P(0,t) = I + A_0 t + O((0,t)^2)$$

One can also see that the row sum of  $A$  must be zero from this relationship and the fact that the row sums of  $P(t)$  must be 1.

To define a continuous-time Markov chain model in an application:

- Model for all the transition rates  $i \rightarrow j$  neglecting the presence of all other processes, enter these values into the off-diagonal entries  $A_{ij}$

- Set the diagonal entries to minus the total transition rate out:

$$A_{ii} = - \sum_{j \neq i} A_{ij} \equiv - \bar{A}_i$$

- Check that the transition rate matrix satisfies the technical conditions described in [Karlin and Taylor Ch. 4](#) if you want to be sure that the model will be well-behaved. (Only a concern with an infinite number of states.)

Note that the entries of  $A$  have units of (1/time). They are rates, not probabilities. Another way to interpret these entries is that

$1/(A_{ij})$  is the average amount of time it would take for the transition  $i \rightarrow j$  to occur if all other transitions out of state  $i$  were disabled.

Also,  $1/\bar{A}_i$

turns out to be the average amount of time that the Markov chain spends in state  $i$ , with all transitions active.

Just supplement this dynamical description with initial data for the probability distribution of the initial state:

$$\phi_j = P(X(0) = j)$$

Example of transition rates in a biochemical reaction network:

## Simulating reaction networks with many random discrete events

### Heat shock model

The following reaction network is given as a model for the heat shock response in *E. Coli* by [Srivastava, Peterson and Bentley \[9\]](#)

Reaction	Intensity	Reaction	Intensity
$\emptyset \rightarrow S_0$	$4.00 \times 10^0$	$S_6 + S_8 \rightarrow S_9$	$3.62 \times 10^{-4} X_{S_6} X_{S_8}$
$S_2 \rightarrow S_3$	$7.00 \times 10^{-1} X_{S_2}$	$S_8 \rightarrow \emptyset$	$9.99 \times 10^{-8} X_{S_8}$
$S_2 \rightarrow S_2$	$1.30 \times 10^{-1} X_{S_2}$	$S_0 \rightarrow S_6 + S_8$	$4.40 \times 10^{-8} X_{S_0}$
$\emptyset \rightarrow S_2$	$7.00 \times 10^{-3} X_{S_2}$	$\emptyset \rightarrow S_1$	$1.40 \times 10^{-8}$
$\text{stuff} + S_2 \rightarrow S_2 + S_2$	$6.30 \times 10^{-3} X_{S_2}$	$S_1 \rightarrow \emptyset$	$1.40 \times 10^{-6} X_{S_1}$
$\text{stuff} + S_2 \rightarrow S_2 + S_2$	$4.88 \times 10^{-3} X_{S_2}$	$S_7 \rightarrow S_6$	$1.42 \times 10^{-6} X_{S_7} X_{S_7}$
$\text{stuff} + S_2 \rightarrow S_6 + S_2$	$4.88 \times 10^{-3} X_{S_2}$	$S_8 \rightarrow \emptyset$	$1.80 \times 10^{-8} X_{S_8}$
$S_7 \rightarrow S_2 + S_8$	$4.40 \times 10^{-4} X_{S_7}$	$S_6 \rightarrow \emptyset$	$6.40 \times 10^{-10} X_{S_6}$
$S_2 + S_6 \rightarrow S_7$	$3.62 \times 10^{-4} X_{S_2} X_{S_6}$	$S_4 \rightarrow \emptyset$	$7.40 \times 10^{-11} X_{S_4}$

# Simulating reaction networks with many random discrete events

## Heat shock model

The following reaction network is given as a model for the heat shock response in E. Coli by Srivastava, Peterson and Bently [9]

Reaction	Intensity	Reaction	Intensity
$\emptyset \rightarrow S_1$	$4.00 \times 10^0$	$S_6 + S_8 \rightarrow S_7$	$3.62 \times 10^{-4} X_{S_6} X_{S_8}$
$S_2 \rightarrow S_3$	$7.00 \times 10^{-1} X_{S_2}$	$S_8 \rightarrow \emptyset$	$9.99 \times 10^{-9} X_{S_8}$
$S_3 \rightarrow S_2$	$1.30 \times 10^{-1} X_{S_3}$	$S_9 \rightarrow S_6 + S_8$	$4.40 \times 10^{-8} X_{S_9}$
$\emptyset \rightarrow S_2$	$7.00 \times 10^{-3} X_{S_1}$	$\emptyset \rightarrow S_1$	$1.40 \times 10^{-6}$
$\text{stuff} + S_3 \rightarrow S_7 + S_2$	$6.30 \times 10^{-3} X_{S_3}$	$S_1 \rightarrow \emptyset$	$1.40 \times 10^{-6} X_{S_1}$
$\text{stuff} + S_3 \rightarrow S_4 + S_2$	$4.88 \times 10^{-3} X_{S_3}$	$S_7 \rightarrow S_6$	$1.42 \times 10^{-6} X_{S_4} X_{S_7}$
$\text{stuff} + S_3 \rightarrow S_6 + S_2$	$4.88 \times 10^{-3} X_{S_3}$	$S_5 \rightarrow \emptyset$	$1.80 \times 10^{-8} X_{S_5}$
$S_7 \rightarrow S_2 + S_8$	$4.40 \times 10^{-4} X_{S_7}$	$S_6 \rightarrow \emptyset$	$6.40 \times 10^{-10} X_{S_6}$
$S_2 + S_6 \rightarrow S_7$	$3.62 \times 10^{-4} X_{S_2} X_{S_6}$	$S_4 \rightarrow \emptyset$	$7.40 \times 10^{-11} X_{S_4}$

Thomas Kurtz, Wisconsin



[First](#)
[Prev](#)
[Next](#)
[Go To](#)
[Go Back](#)
[Full Screen](#)
[Close](#)
[Quit](#)