Modeling and Automation of Industrial Processes

Modelação e Automação de Processos Industriais / MAPI

Stochastic Analysis

Markov Chain Modelling

http://www.isr.tecnico.ulisboa.pt/~jag/courses/mapi2223

Prof. José Gaspar, rev. 2022/2023

Syllabus:

Chap. 2 – Discrete Event Systems

Chap. 3 – Stochastic Models

Stochastic Timed Automata (STA)
Stochastic Queueing Networks (SQN)
Stochastic Petri Nets (SPN)
Generalized Stochastic Petri Nets (GSPN)

Chap. 4 – Stochastic Analysis

Markov chain (MC) modelling

Chap. 5 – Supervision of DESs

Markov Chains

[Luenberger'79] "Introduction to dynamic systems: theory, models, and applications", David G. Luenberger, Vol. 1. New York: Wiley, 1979.

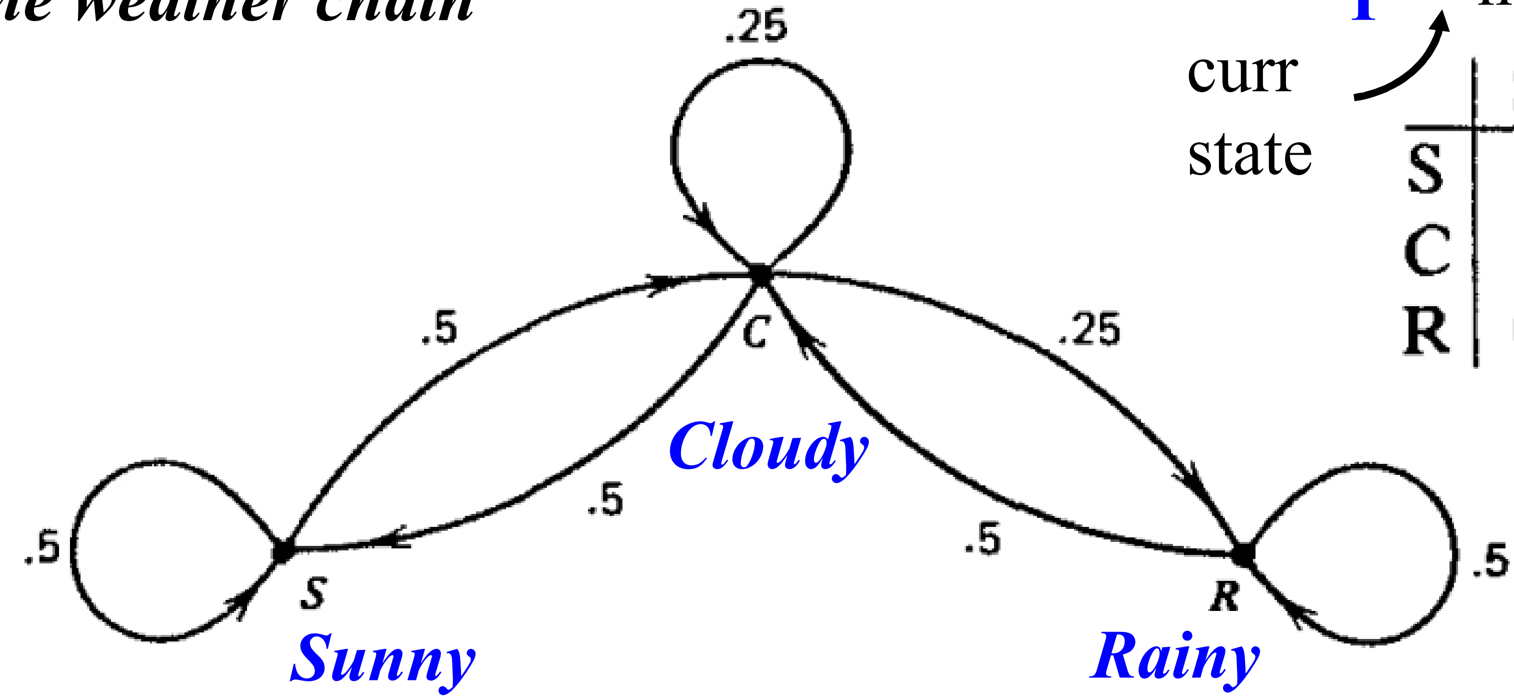
[Zurawski'94] "Petri nets and industrial applications: A tutorial." R. Zurawski, M. Zhou. IEEE Trans. on Industrial Electronics 41.6 (1994): 567-583. https://www.researchgate.net/publication/3217035

[Cassandras'08] Introduction to discrete event systems, Christos Cassandras and Stéphane Lafortune, Springer 2008 – chapter 6, section 6.4.

Markov Chain (MC), example 1 [Luenberger'79]

Transition probabilities

The weather chain



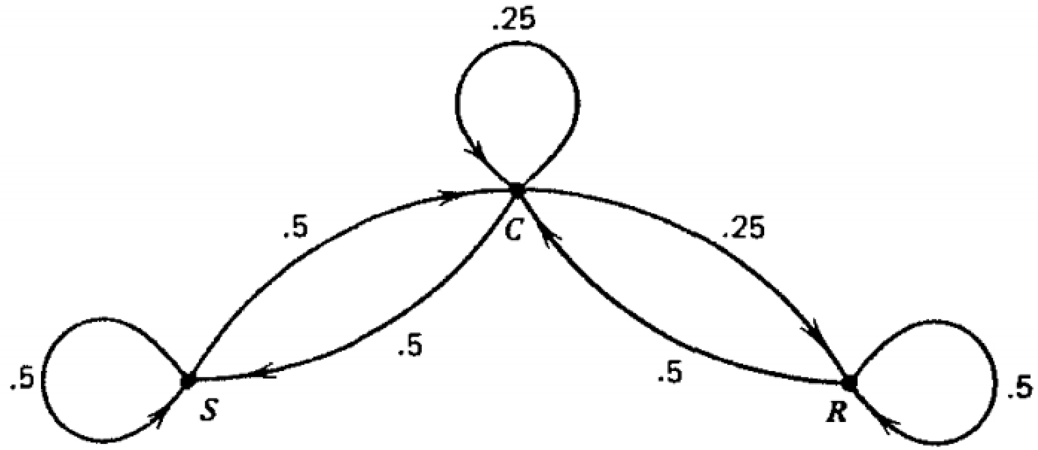
P — curr state → next state

P	S	C	R
S	$\frac{1}{2}$	$\frac{1}{2}$	0
C	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$
R	0	$\frac{1}{2}$	$\frac{1}{2}$

In this example, *on a sunny day, chances are 50% it continues sunny or 50% changes to cloudy (not rainy)*.

P always ***sums 1 at every row,*** is always a ***right-stochastic matrix***.

Markov Chain, example 1 [Luenberger'79]
The weather chain



Transition probabilities

next state

curr state

	S	C	R
S	$\frac{1}{2}$	$\frac{1}{2}$	0
C	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$
R	0	$\frac{1}{2}$	$\frac{1}{2}$

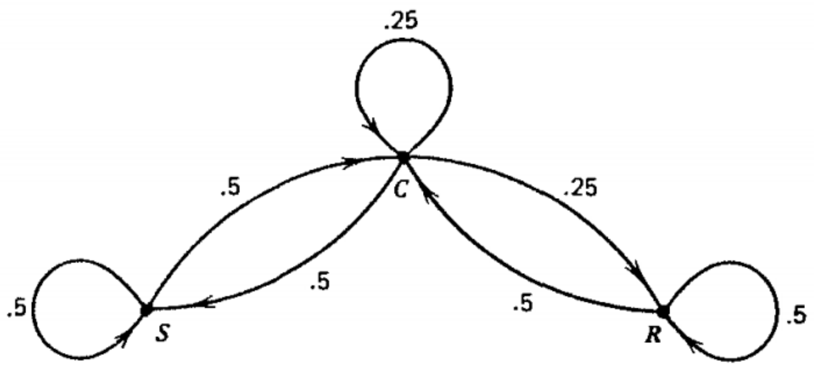
If one starts at a sunny day, $x(0)^T = [1\ 0\ 0]$, what are the chances of sunny / cloudy / rainy after 1, 2, … days?

$$P = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.25 & 0.25 \\ 0 & 0.5 & 0.5 \end{bmatrix}$$

$$\mathrm{x}(1)^{\mathrm{T}} = x(0)^T P = [0.5 \quad 0.5 \quad 0]$$

$$\mathrm{x}(2)^{\mathrm{T}} = x(1)^T P = x(0)^T P^2 = [0.5 \quad 0.375 \quad 0.125]$$

State Holding Times



Initial state $x(0)^T = [1\ 0\ 0]$
transition matrix

$$P = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.25 & 0.25 \\ 0 & 0.5 & 0.5 \end{bmatrix}$$

What are the chances of sunny / cloudy / rainy after 1, 2, … days?

```
x0T= [1 0 0];
P = [.5 .5 0; .5 .25 .25; 0 .5 .5];
for n= 0:20
    fprintf('x%dT = [%.3f %.3f %.3f]\n', ...
                    n, x0T*P^n);
end
```

```
x0T = [1.000 0.000 0.000]
x1T = [0.500 0.500 0.000]
x2T = [0.500 0.375 0.125]
x3T = [0.438 0.406 0.156]
x4T = [0.422 0.398 0.180]
x5T = [0.410 0.400 0.189]
x6T = [0.405 0.400 0.195]
x7T = [0.403 0.400 0.197]
x8T = [0.401 0.400 0.199]
x9T = [0.401 0.400 0.199]
x10T = [0.400 0.400 0.200]
x11T = [0.400 0.400 0.200]
x12T = [0.400 0.400 0.200]
x13T = [0.400 0.400 0.200]
x14T = [0.400 0.400 0.200]
x15T = [0.400 0.400 0.200]
x16T = [0.400 0.400 0.200]
x17T = [0.400 0.400 0.200]
x18T = [0.400 0.400 0.200]
x19T = [0.400 0.400 0.200]
x20T = [0.400 0.400 0.200]
```

Positive Linear System [Luenberger'79]

A discrete-time linear system $\boldsymbol{x}(k+1) = \mathbf{A}\boldsymbol{x}(k)$ is defined to be positive (or nonnegative) if the elements of **A** are all nonnegative.

If $\mathbf{A} = [a_{ij}]$ is a matrix, we write:
(i) $\mathbf{A} > \mathbf{0}$ if $a_{ij} > 0$ for all i,j
(ii) $\mathbf{A} \geq \mathbf{0}$ if $a_{ij} \geq 0$ for all i,j and $a_{ij} > 0$ for at least one element
(iii) $\mathbf{A} \geqq \mathbf{0}$ if $a_{ij} \geq 0$ for all i,j.

Terminology:
(i) **A** is ***strictly positive*** if all its elements are strictly greater than zero
(ii) **A** is ***positive*** or ***strictly nonnegative*** if all elements of A are nonnegative but at least one element IS nonzero, and
(iii) **A** is ***nonnegative*** if all elements are nonnegative.

[Luenberger'79] "Introduction to dynamic systems: theory, models, and applications", David G. Luenberger, Vol. 1. New York: Wiley, 1979.

***Theorem 1* (Frobenius-Perron).** *If* $\mathbf{A} > \mathbf{0}$, *then there exists* $\lambda_0 > 0$ *and* $\mathbf{x}_0 > \mathbf{0}$ *such that* (a) $\mathbf{A}\mathbf{x}_0 = \lambda_0 \mathbf{x}_0$; (b) if $\lambda \neq \lambda_0$ is *any other eigenvalue of* $\mathbf{A}$, *then* $|\lambda| < \lambda_0$; (c) λ_0 *is an eigenvalue of geometric and algebraic multiplicity* 1.

Theorem 2. *Let* $\mathbf{A} \geq \mathbf{0}$ *and suppose* $\mathbf{A}^m > \mathbf{0}$ *for some positive integer* m. *Then conclusions* (a), (b), *and* (c) *of Theorem* 1 *apply to* $\mathbf{A}$.

Theorem 3. *Let* $\mathbf{A} \geqq \mathbf{0}$. *Then there exists* $\lambda_0 \geq 0$ *and* $\mathbf{x}_0 \geq \mathbf{0}$ *such that* (a) $\mathbf{A}\mathbf{x}_0 = \lambda_0 \mathbf{x}_0$; (b) *if* $\lambda \neq \lambda_0$ *is any other eigenvalue of* $\mathbf{A}$, *then* $|\lambda| \leq \lambda_0$.

Note: in the following instead of writing

$$\boldsymbol{x}(k+1) = \mathbf{A}\,\boldsymbol{x}(k)$$

it is going to be written the equation transposed, as it is common in Markov chains

$$\boldsymbol{x}^T(k+1) = \boldsymbol{x}^T(k)\,\boldsymbol{P}$$

hence $\boldsymbol{A} = \boldsymbol{P}^T$.

Definition. A Markov chain is said to be *regular* if $\mathbf{P}^m > \mathbf{0}$ for some positive integer m.

Theorem 4 (Basic Limit Theorem for Markov Chains) [Luenberger'79]

Let **P** *be the transition matrix of a regular Markov chain. Then:*

(a) *There is a unique probability vector* $\mathbf{p}^{\mathrm{T}} > 0$ *such that*

$$\mathbf{p}^{\mathrm{T}}\,\mathbf{P} = \mathbf{p}^{\mathrm{T}}$$

(b) *For any initial state i (corresponding to an initial probability vector equal to the ith coordinate vector* $\mathbf{e}_i^T$*) the limit vector*

$$\mathbf{v}^{\mathrm{T}} = \lim_{m\to\infty} \mathbf{e}_i^T\,\mathbf{P}^m$$

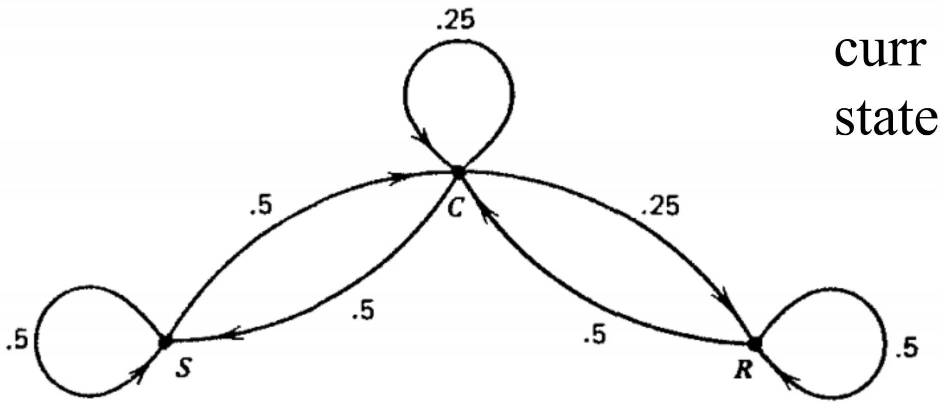
exists and is independent of i. Furthermore, $\mathbf{v}^{\mathrm{T}}$ *is equal to the eigenvector* $\mathbf{p}^{\mathrm{T}}$.

(c) $\lim_{\mathbf{m}\to\infty} \mathbf{P}^{\mathrm{m}} = \overline{\mathbf{P}}$, *where* $\overline{\mathbf{P}}$ *is the* $n \times n$ *matrix, each of whose rows is equal to* $\mathbf{p}^{\mathrm{T}}$.

Proof starts at the Frobenius-Perron theorem and the fact that the dominant eigenvalue is $\lambda_0 = 1$.

Markov Chain, example 1 [Luenberger'79]

The weather chain



next state / curr state

	S	C	R
S	$\frac{1}{2}$	$\frac{1}{2}$	0
C	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$
R	0	$\frac{1}{2}$	$\frac{1}{2}$

$$P = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.25 & 0.25 \\ 0 & 0.5 & 0.5 \end{bmatrix}$$

What are the probabilities of the states after an infinite running time?

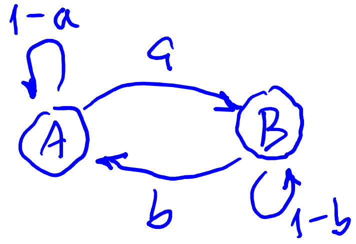
```
P= [0.5 0.5  0
    0.5 0.25 0.25
    0    0.5  0.5 ];

P2= P*P

P2 =
    0.5000    0.3750    0.1250
    0.3750    0.4375    0.1875
    0.2500    0.3750    0.3750
```

Good news $\boldsymbol{P}^m > 0$ for $m = 2 > 0$ we can use Theorem 2.

```
P= [.5 .5 0; .5 .25 .25; 0 .5 .5];
[A,B]= eig( P' ); eigenvalues= diag(B)', eigenvectors= A

eigenvalues =
   -0.2500    1.0000    0.5000

eigenvectors =
   -0.5345   -0.6667   -0.7071
    0.8018   -0.6667   -0.0000
   -0.2673   -0.3333    0.7071

p_inf= eigenvectors(:,2)'/sum(eigenvectors(:,2))
p_inf =
    0.4000    0.4000    0.2000
```

Markov Chain, example to *do by hand*

Two states MC



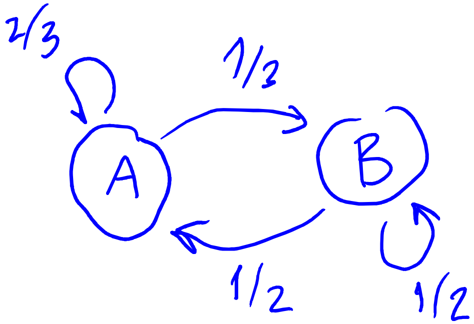
	A	B
A	1-a	a
B	b	1-b

Matlab:

```
syms a b
[V,E]= eig( [1-a b; a 1-b] )

V =
[ b/a, -1]
[   1,  1]

E =
[ 1,          0]
[ 0, 1 - b - a]
```

$\begin{bmatrix} b \\ a \end{bmatrix} / (b+a)$ — sum 1 vector

Using values



$$P = \begin{bmatrix} 2/3 & 1/3 \\ 1/2 & 1/2 \end{bmatrix}$$

Good news: $P>0$, hence $\exists^1_{\lambda=1}$

$$[p_A, p_B]\, P = [p_A, p_B]$$

$$\left[\tfrac{2}{3}p_A + \tfrac{1}{2}p_B ,\ \tfrac{1}{3}p_A + \tfrac{1}{2}p_B\right] = [p_A, p_B]$$

$$\begin{cases} -\tfrac{1}{3}p_A + \tfrac{1}{2}p_B = 0 \\ \tfrac{1}{3}p_A - \tfrac{1}{2}p_B = 0 \end{cases} \xrightarrow{\text{if } p_B = 1} \begin{matrix} p_A = 3/2 \\ p_B = 1 \end{matrix}$$

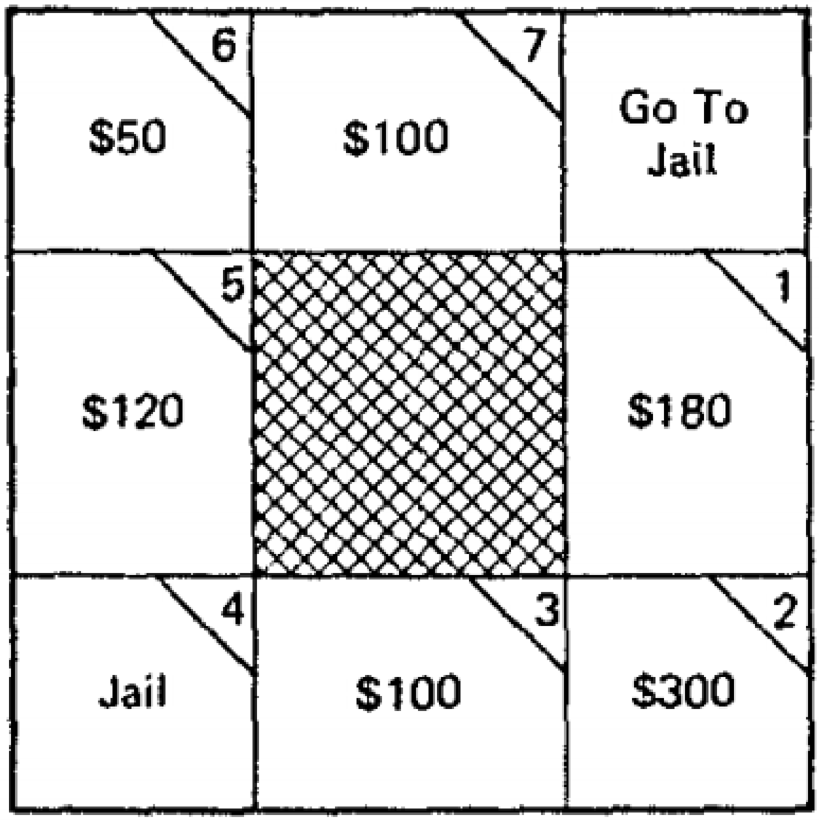
sum $= 3/2 + 1 = 5/2$

Normalize to sum = 1:

$$\begin{bmatrix} p_A^* \\ p_B^* \end{bmatrix} = \begin{bmatrix} 3/2 \\ 1 \end{bmatrix} / \tfrac{5}{2} = \begin{bmatrix} 3/5 \\ 2/5 \end{bmatrix}$$

Homework, 7-state MC

Simplified Monopoly [Luenberger'79]



Toss a coin, face1 move 1 step, face2 mode 2 steps. Stepping into *Go to jail* implies going to *Jail*.

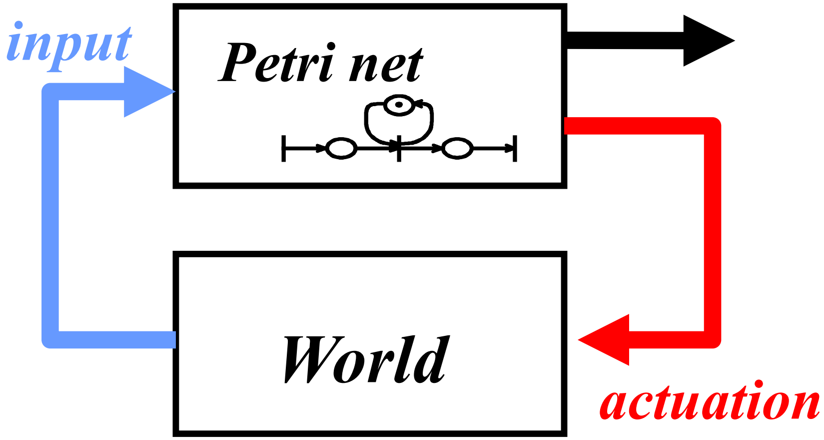
$$\mathbf{P} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \end{bmatrix}$$

Homework: Simulate the simplified Monopoly as an SPN. Find the most visited place. Find the place that provides the maximum profit.

Stochastic Timed Petri nets (STPN), Markov Chains (MC)

STPN to MC, and MC to STPN

STPN to MC

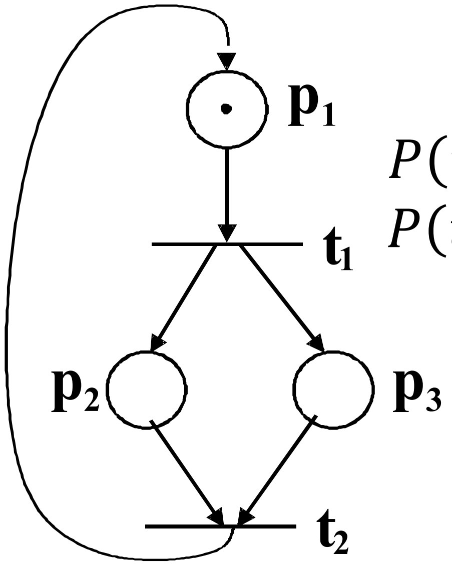


Given the dynamics

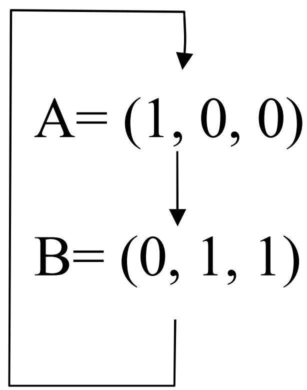
$$\mu(k+1) = \mu(k) + D\,\boldsymbol{q}(\boldsymbol{k}) \quad \text{and}$$
$$D^{-}\,q(k) \leq \mu(k)$$

$\boldsymbol{q}(\boldsymbol{k})$ *has a statistical characterization*

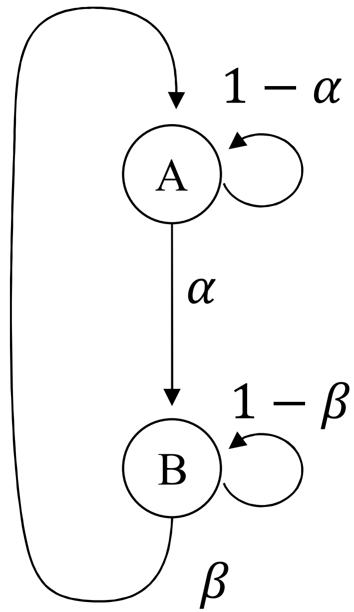
1. STPN (example)



2. Reachability graph

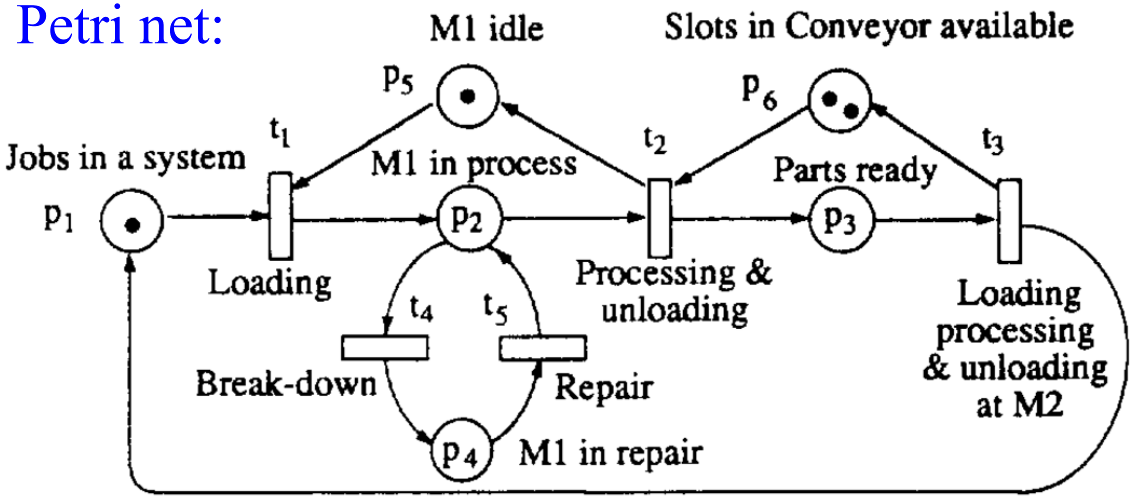


3. Markov chain



Example of an STPN (continued), steady state computation using an MC

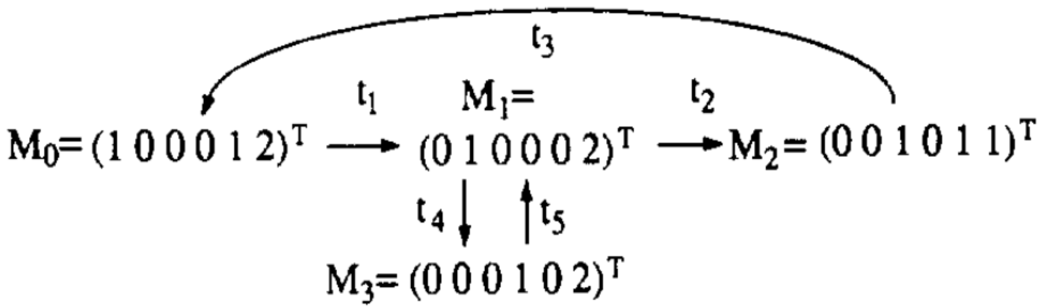
Petri net:



Statistical information about transition firing:

$P(t_1) = \alpha_1 = 0.982591$
$P(t_2) = \alpha_2 = 0.122824$
$P(t_3) = \alpha_3 = 0.098259$
$P(t_4) = \alpha_4 = 0.089308$
$P(t_5) = \alpha_5 = 0.714460$

Reachability graph / Markov chain:



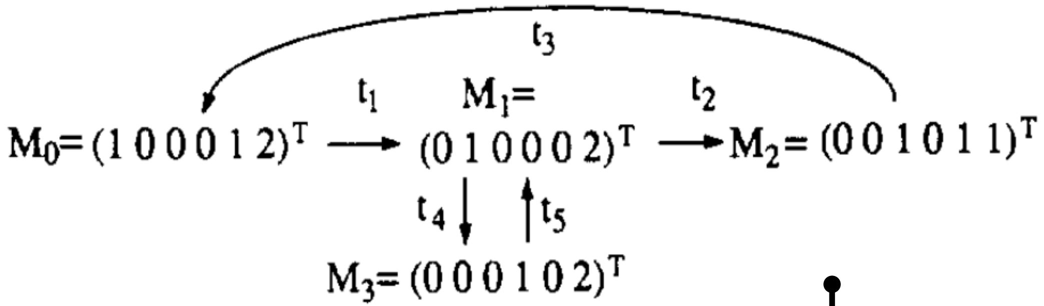
Markov chain matrix:

curr state → next state

$$P = \begin{bmatrix} 1-\alpha_1 & \alpha_1 & 0 & 0 \\ 0 & 1-\alpha_2-\alpha_4 & \alpha_2 & \alpha_4 \\ \alpha_3 & 0 & 1-\alpha_3 & 0 \\ 0 & \alpha_5 & 0 & 1-\alpha_5 \end{bmatrix}$$

Example of an STPN (continued), steady state computation using an MC

Reachability graph / Markov chain:



$$P = \begin{bmatrix} 1-\alpha_1 & \alpha_1 & 0 & 0 \\ 0 & 1-\alpha_2-\alpha_4 & \alpha_2 & \alpha_4 \\ \alpha_3 & 0 & 1-\alpha_3 & 0 \\ 0 & \alpha_5 & 0 & 1-\alpha_5 \end{bmatrix}$$

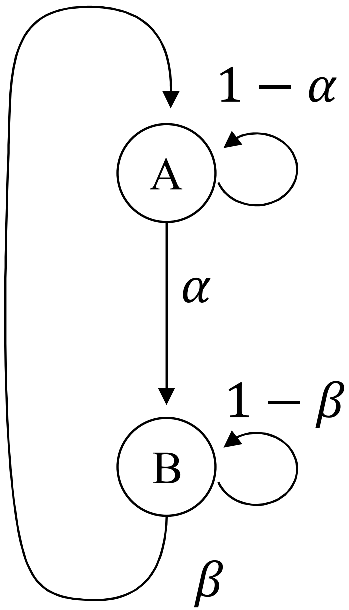
```
P =
    0.0174    0.9826         0         0
         0    0.7879    0.1228    0.0893
    0.0983         0    0.9017         0
         0    0.7145         0    0.2855

P*P*P =
    0.0119    0.6864    0.2060    0.0957
    0.0206    0.6197    0.2712    0.0885
    0.0815    0.1648    0.7451    0.0086
    0.0086    0.7081    0.1733    0.1100
```

Good news $\boldsymbol{P}^m > 0$ *for* $m = 3 > 0$
we can use Theorem 2.

```
[eigenvectors, B]= eig(P'); evalues=diag(B)',
eigenvectors

evalues =
    1.0000    0.7842    0.0467    0.1617

eigenvectors =
    0.0776   -0.0915    0.4088    0.0905
    0.6209    0.6832   -0.8472   -0.8006
    0.7762   -0.7141    0.1217    0.1329
    0.0776    0.1224    0.3167    0.5772

p_inf= eigenvectors(:,1)'/sum(eigenvectors(:,1))
p_inf =
    0.05      0.40      0.50      0.05
```

MC to STPN

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix} \qquad p^T P = p^T \Rightarrow p_\infty = \begin{bmatrix} \beta \\ \alpha \end{bmatrix} / \alpha+\beta$$



To obtain 70% holding time in A one has

$$\begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} = \frac{\begin{bmatrix} \beta \\ \alpha \end{bmatrix}}{\alpha+\beta}$$

$p(t_1) = 30\%$

$p(t_2) = 70\%$

Transient Analysis and Classification of States

Transient Analysis and Classification of States

What happens if there does NOT exist $m > 0$ such that $\boldsymbol{P}^m > 0$?

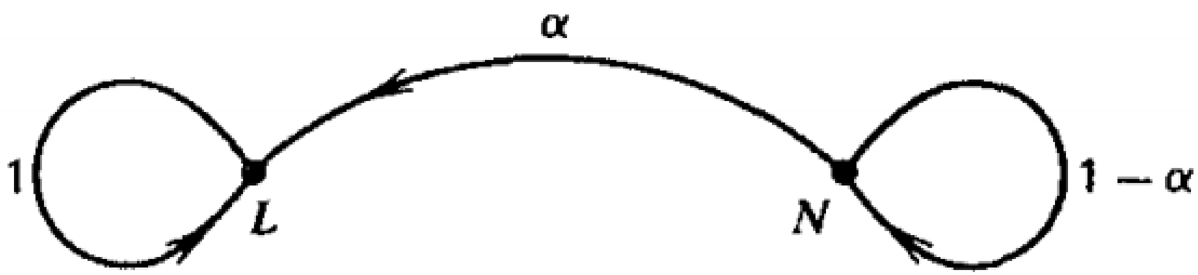
- *Some states of the MC will not have a stationary probability in]0, 1], those states will have **0 stationary probability***

- *In other words, as $t \rightarrow \infty$ some states will not be visited after some time.*

- *Those states will have transitory visits, will be transitory states, within the operation time.*

Markov Chain, example 2 [Luenberger'79]

Learning model



Learnt
(will not forget)

Not learnt
(yet to learn)

	L	N
L	1	0
N	α	$1-\alpha$

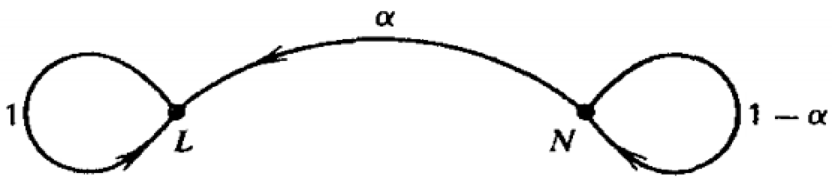
Specific example $\alpha = 0.5$:

```
P =
    1.0000         0
    0.5000    0.5000

P^2 =
    1.0000         0
    0.7500    0.2500

P^3 =
    1.0000         0
    0.8750    0.1250

P^10 =
    1.0000         0
    0.9990    0.0010

P^20 =
    1.0000         0
    1.0000    0.0000
```

Always finding zeros in P^m

Markov Chain, example 2 [Luenberger'79]

Learning model



Learnt
(will not forget)

Not learnt
(yet to learn)

	L	N
L	1	0
N	α	$1-\alpha$

Continuing the specific example $\alpha = 0.5$:

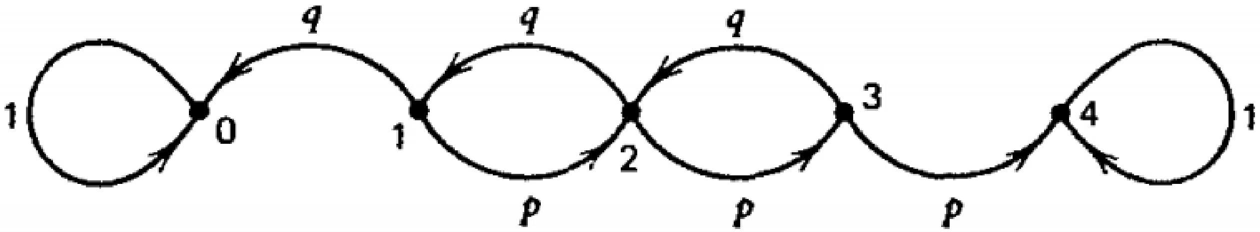
```
P=[1 0; .5 .5]

[eigenvectors,B]= eig( P' );
evalues=diag(B)', eigenvectors

evalues =
    1.0000    0.5000

eigenvectors =
    1.0000   -0.7071
         0    0.7071
```

One concludes:

- state L, the *learnt* state, is an **absorbing state**
- state N, the *yet to learn* state, is a **transient state**

Markov Chain, example 3 [Luenberger'79]

Gambler's ruin - *Two players have 2 coins each one. Player A has probability p of winning one coin from player B. Thus, player B has probability* $q = 1 - p$ *of being him to win a coin from A. (In the following is used* $p = q = 0.5$ *.)*



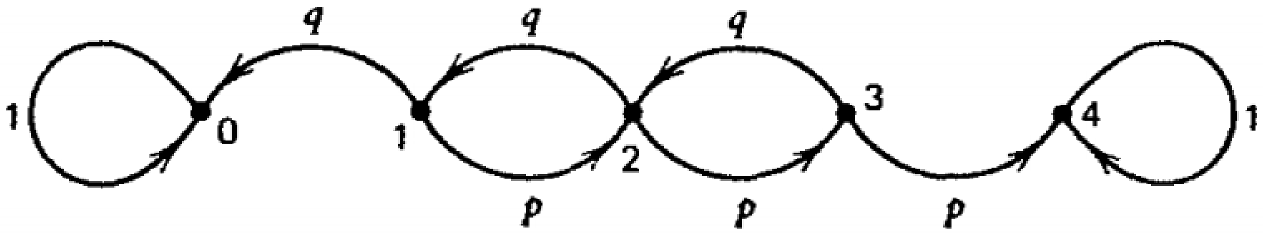
	0	1	2	3	4
0	1	0	0	0	0
1	q	0	p	0	0
2	0	q	0	p	0
3	0	0	q	0	p
4	0	0	0	0	1

```
P=[1 0 0 0 0; q 0 p 0 0; 0 q 0 p 0; 0 0 q 0 p; 0 0 0 0 1]

P^2 =
    1.0000         0         0         0         0
    0.5000    0.2500         0    0.2500         0
    0.2500         0    0.5000         0    0.2500
         0    0.2500         0    0.2500    0.5000
         0         0         0         0    1.0000

P^100 =
    1.0000         0         0         0         0
    0.7500    0.0000         0    0.0000    0.2500
    0.5000         0    0.0000         0    0.5000
    0.2500    0.0000         0    0.0000    0.7500
         0         0         0         0    1.0000
```

Always finding zeros in P^m

Markov Chain, example 3 [Luenberger'79]

Gambler's ruin



	0	1	2	3	4
0	1	0	0	0	0
1	q	0	p	0	0
2	0	q	0	p	0
3	0	0	q	0	p
4	0	0	0	0	1

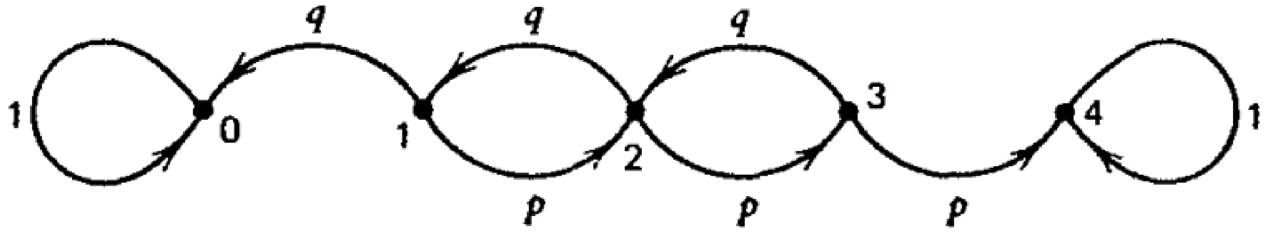
```
[eigenvectors,B]= eig( P' ); evalues=diag(B)', eigenvectors

evalues =

    1.0000    1.0000    0.7071   -0.7071         0

eigenvectors =

    1.0000         0   -0.5445    0.1434   -0.3162
         0         0    0.3190   -0.4896    0.6325
         0         0    0.4511    0.6924    0.0000
         0         0    0.3190   -0.4896   -0.6325
         0    1.0000   -0.5445    0.1434    0.3162
```

States 0 and 4 are ***absorbing states***.

States, 1, 2 and 3, are ***transient states***.

Classification of States

Gambler's ruin




```
[eigenvectors,B]= eig( P' ); evalues=diag(B)', eigenvectors

evalues =

    1.0000    1.0000    0.7071   -0.7071         0

eigenvectors =

    1.0000         0   -0.5445    0.1434   -0.3162
         0         0    0.3190   -0.4896    0.6325
         0         0    0.4511    0.6924    0.0000
         0         0    0.3190   -0.4896   -0.6325
         0    1.0000   -0.5445    0.1434    0.3162
```

States 0 and 4 are ***absorbing states****.*

States, 1, 2 and 3, are ***transient states****.*