

30th October, 2017

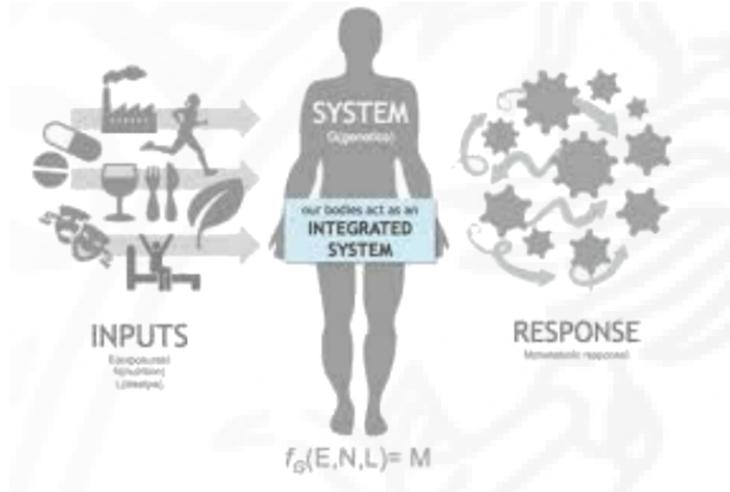
Systems Biology: Petri net formalism

Marco Beccuti

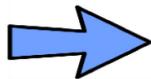
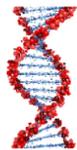
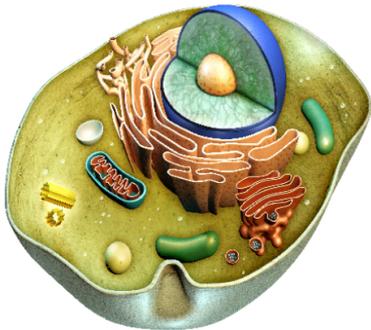
Department of Computer Science

University of Turin

A biological system is a complex system



In Systems Biology mathematical and computational modeling is exploited to help scientists in the study of biological systems



$$\frac{dU_x}{dt} = \theta H - \rho_f \sigma_f C_f \frac{U_x}{N} - \rho_f \sigma_p D \frac{U_x}{N} - \rho_f \sigma_c C_c \frac{U_x}{N}$$

$$\frac{dH}{dt} = \rho_f \sigma_f C_f \frac{U_x}{N} + \rho_f \sigma_p D \frac{U_x}{N} + \rho_f \sigma_c C_c \frac{U_x}{N} - \theta H$$

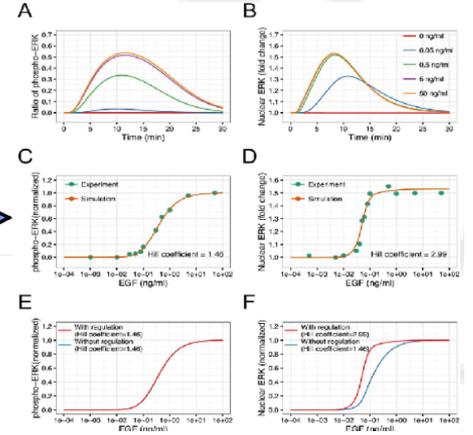
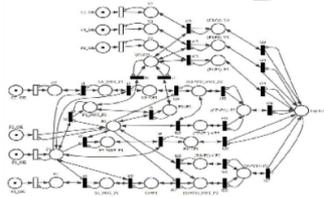
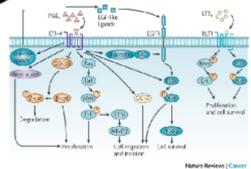
$$\frac{dU_c}{dt} = \frac{1}{\omega} U_x - \alpha U_c - \rho_f \mu_f \frac{H}{N} - \mu_f \sigma_f \frac{C_f}{N} - \mu_f \sigma_p \frac{D}{N} - \mu_f \sigma_c \frac{C_c}{N} - \theta U_c + v_{c1} (\theta P + \zeta D)$$

$$\frac{dU_x}{dt} = \alpha U_c - \frac{1}{\omega} U_x - \rho_f \mu_f \frac{H}{N} - \mu_f \sigma_f \frac{C_f}{N} - \mu_f \sigma_p \frac{D}{N} - \mu_f \sigma_c \frac{C_c}{N} - \theta U_x + v_{c2} (\theta P + \zeta D)$$

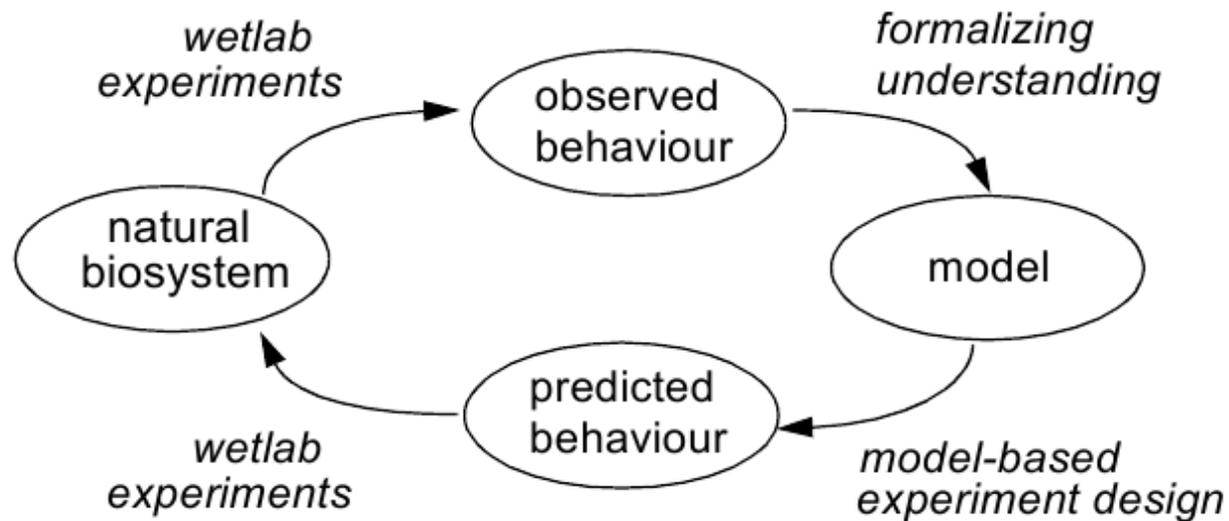
$$\frac{dC_c}{dt} = \frac{1}{\omega} C_c + \gamma \theta D + \rho_f \mu_f \frac{H}{N} - \mu_f \sigma_f \frac{C_f}{N} + \mu_f \sigma_p \frac{D}{N} - \alpha C_c - \theta C_c + v_{c3} (\theta P + \zeta D)$$

$$\frac{dC_f}{dt} = \alpha C_c + \gamma (1 - \theta) D + \rho_f \mu_f \frac{H}{N} + \mu_f \sigma_f \frac{C_f}{N} + \mu_f \sigma_p \frac{D}{N} - \frac{1}{\omega} C_f - \kappa C_f - \theta C_f + v_{c4} (\theta P + \zeta D)$$

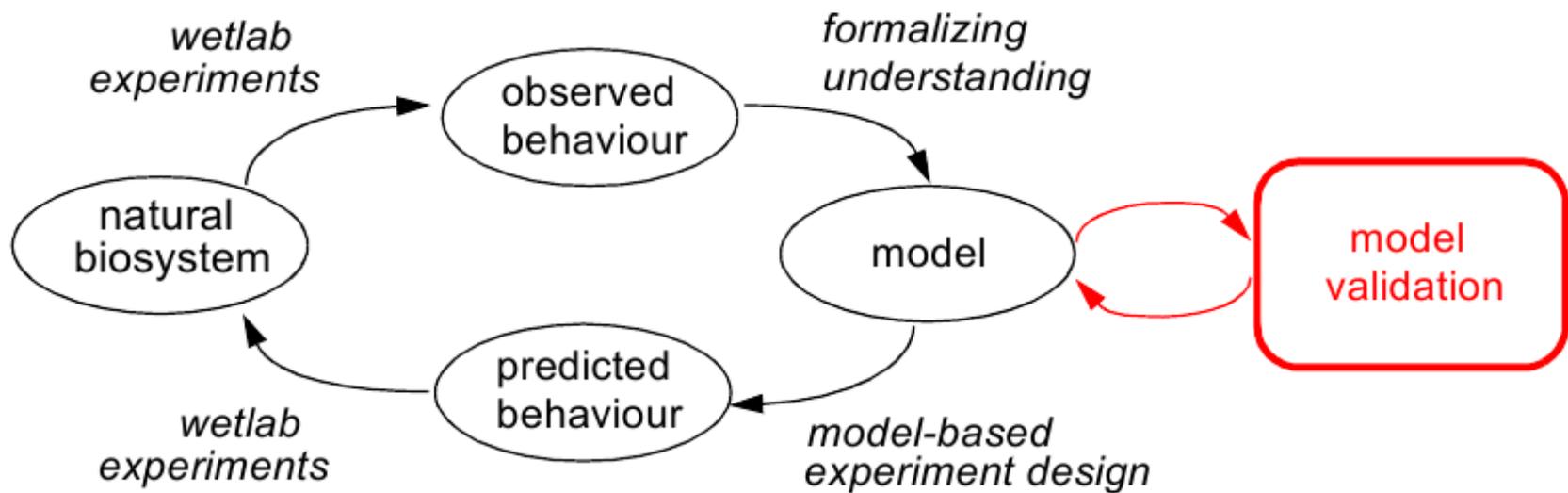
$$\frac{dD}{dt} = \kappa C_f + \kappa T C_x - \gamma \theta D - \gamma (1 - \theta) D - \zeta D$$



Modeling = formal knowledge representation

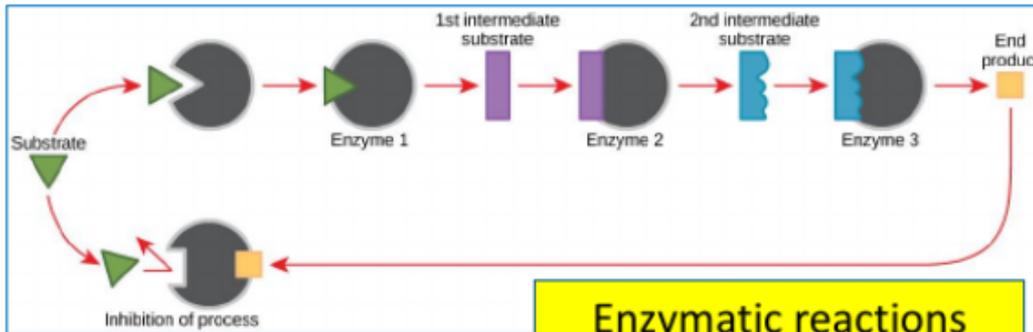


Modelling = formal knowledge representation

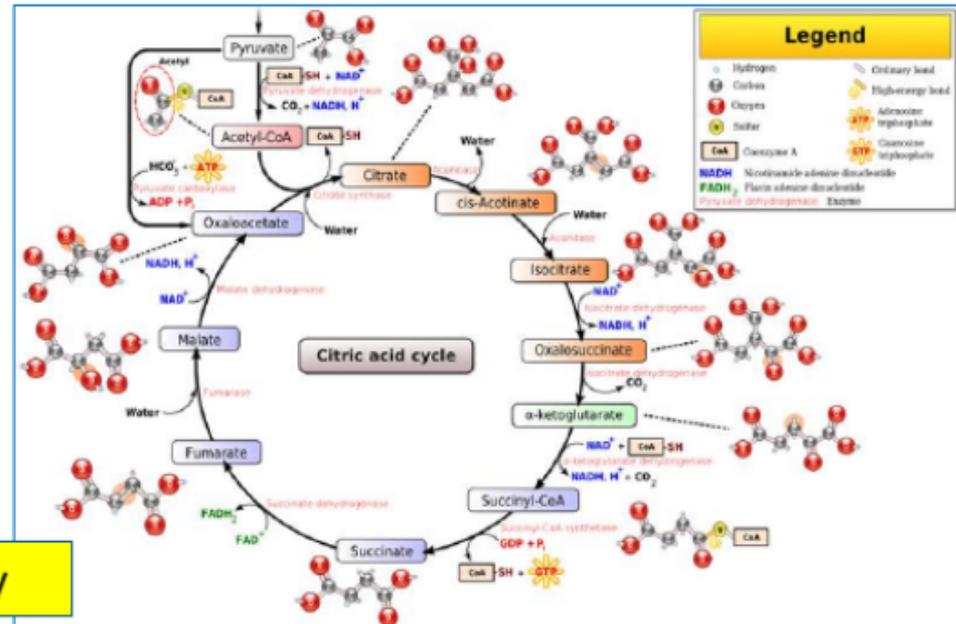


MODEL VALIDATION = CONFIDENCE INCREASE

Different biological systems...

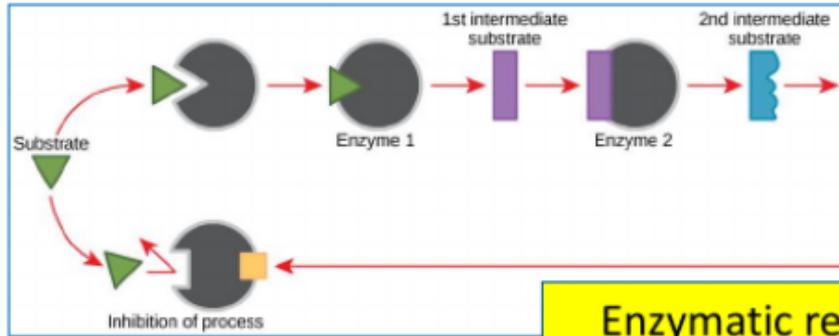


Enzymatic reactions with negative feedback



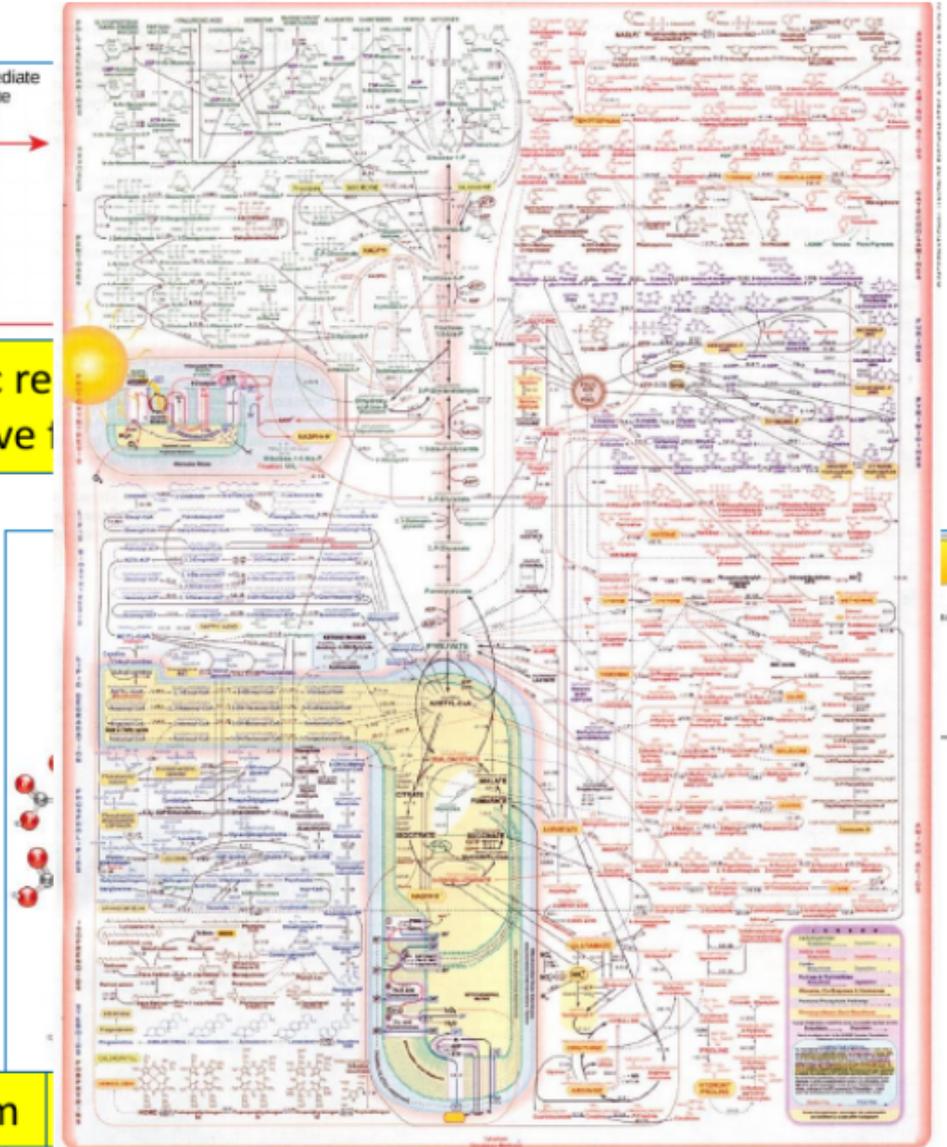
Metabolic pathway

Different biological systems...

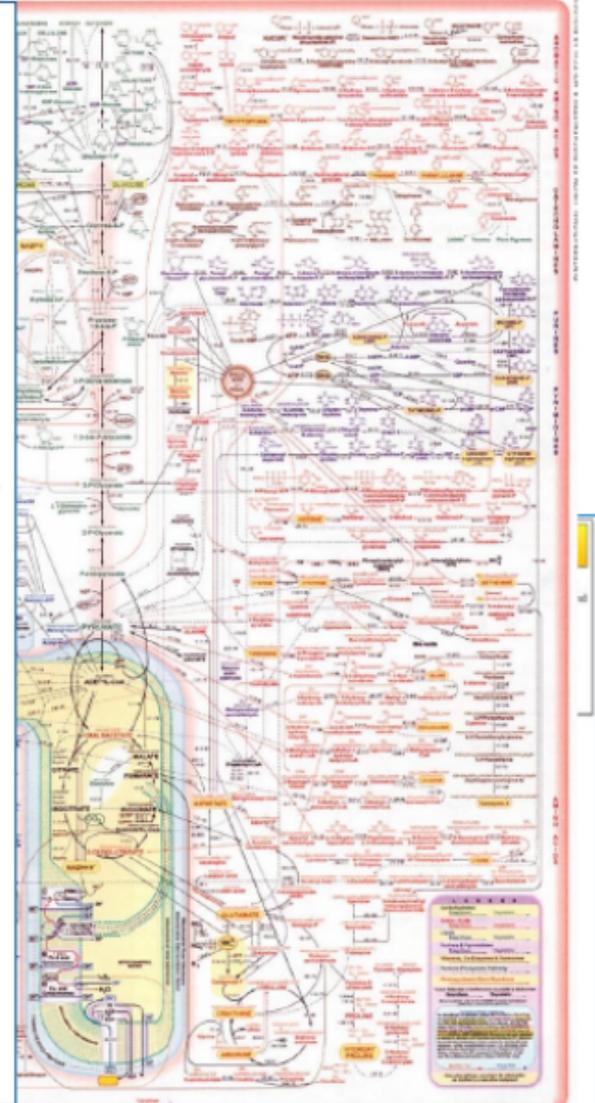
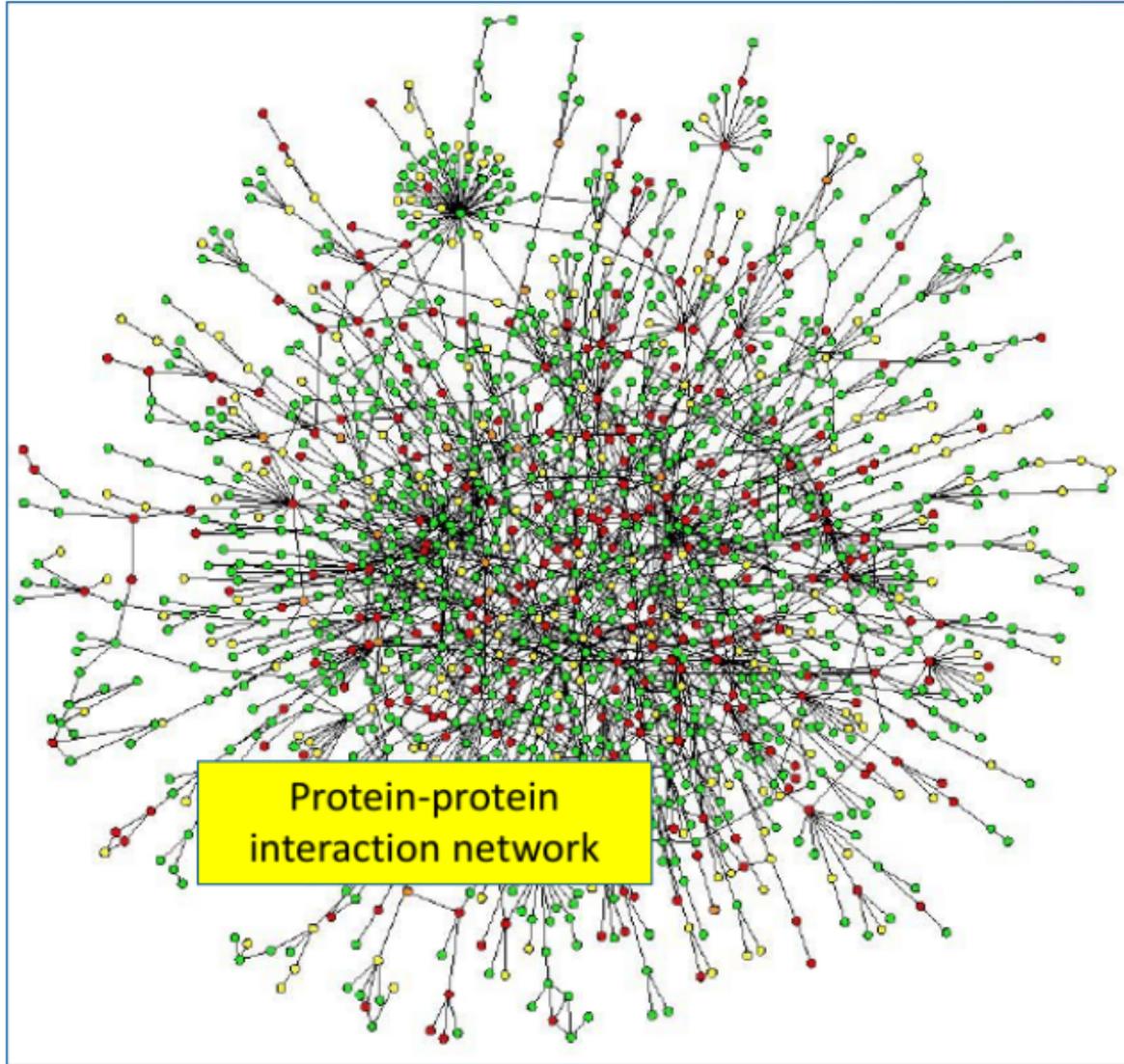


Enzymatic regulation with negative feedback

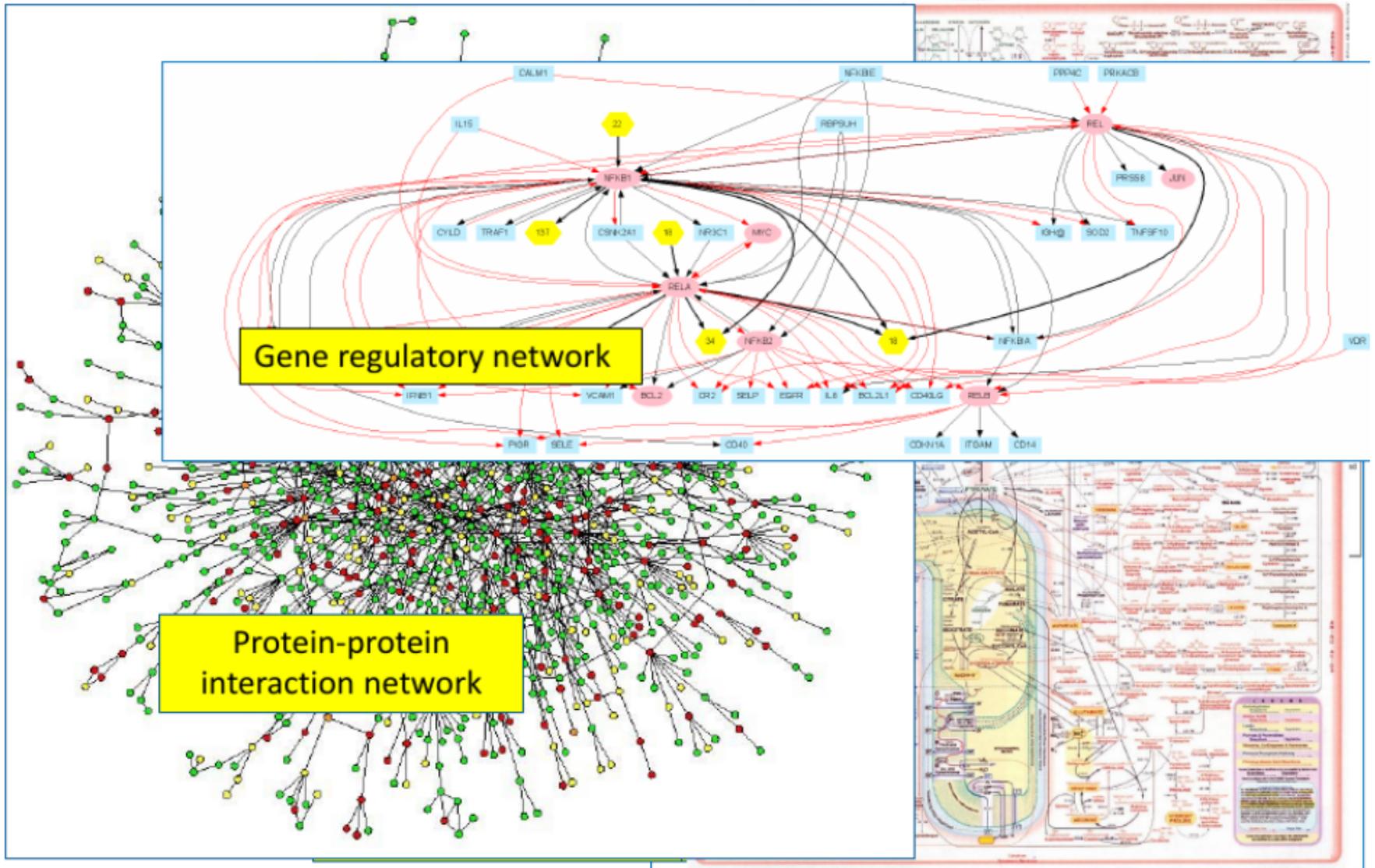
Whole-cell metabolism



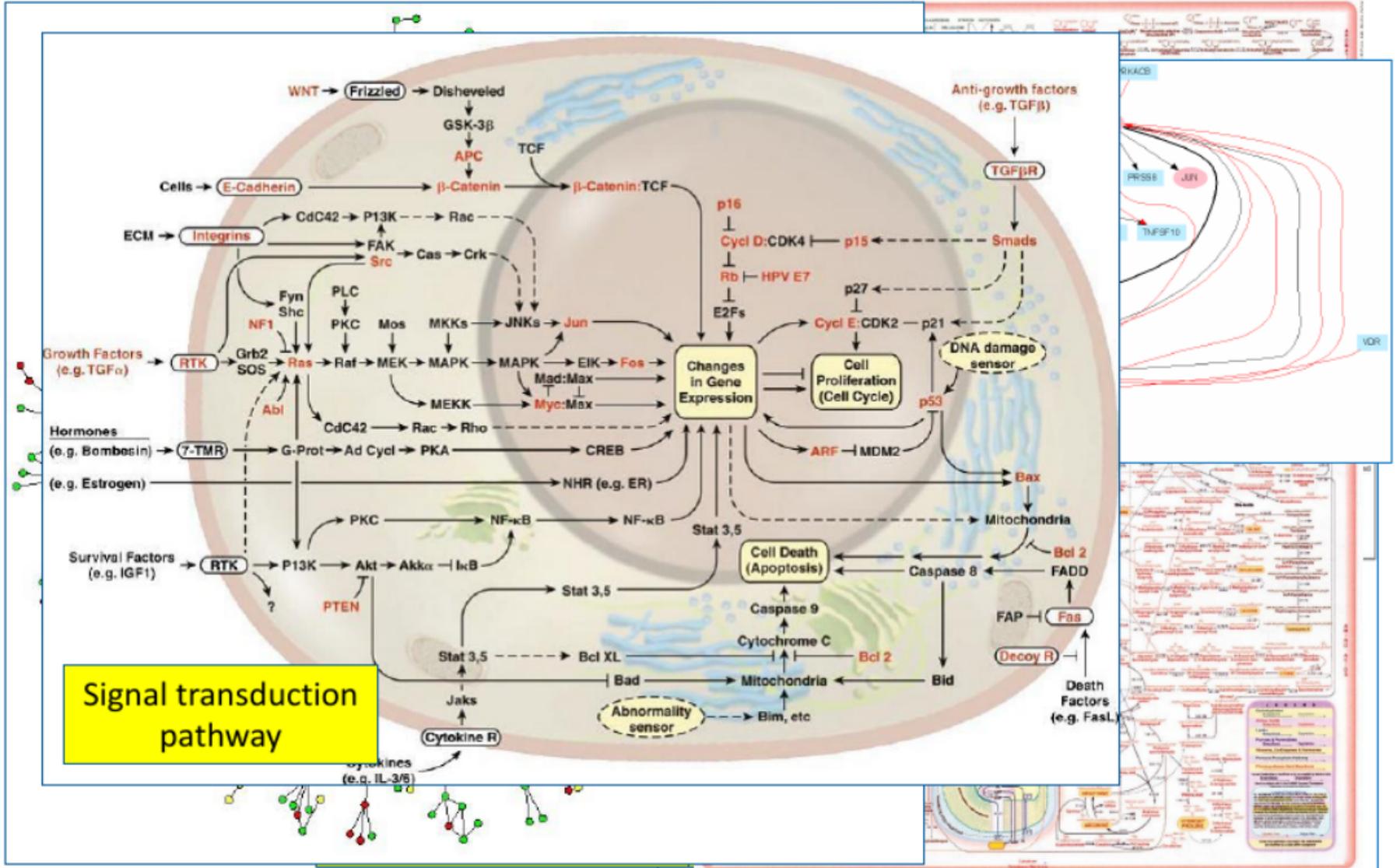
Different biological systems...

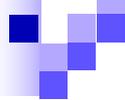


Different biological systems...



Different biological systems...





... require different modeling approaches

- **interaction-based approaches**

- ▶ protein-protein interaction network;
- ▶ gene regulatory network.

based on graph theory → *topological analysis*.

- **constraint-based approaches**

- ▶ metabolic network.

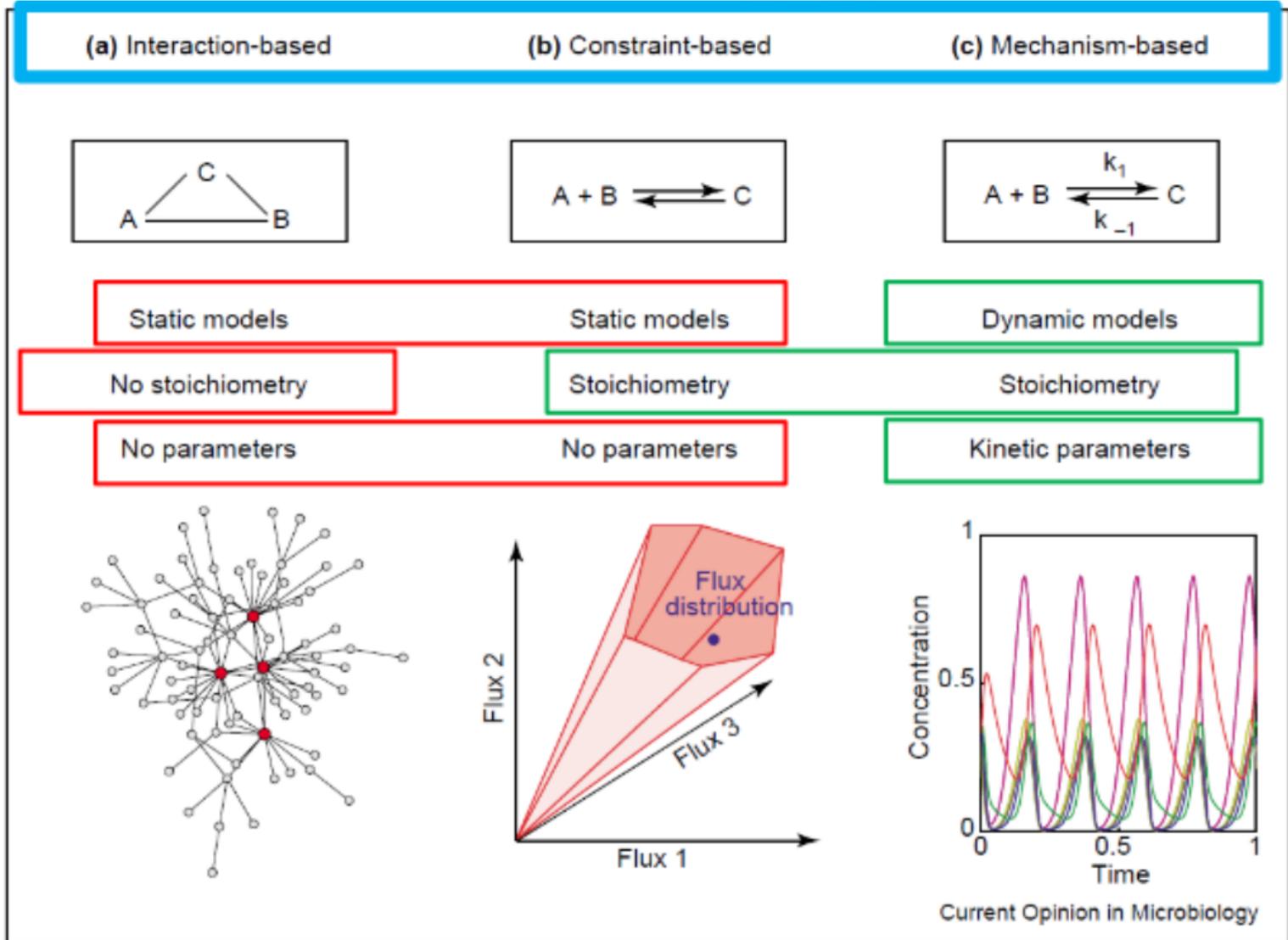
based on linear algebra and optimization with linear programming → *Flux balance analysis*.

- **mechanism-based approaches**

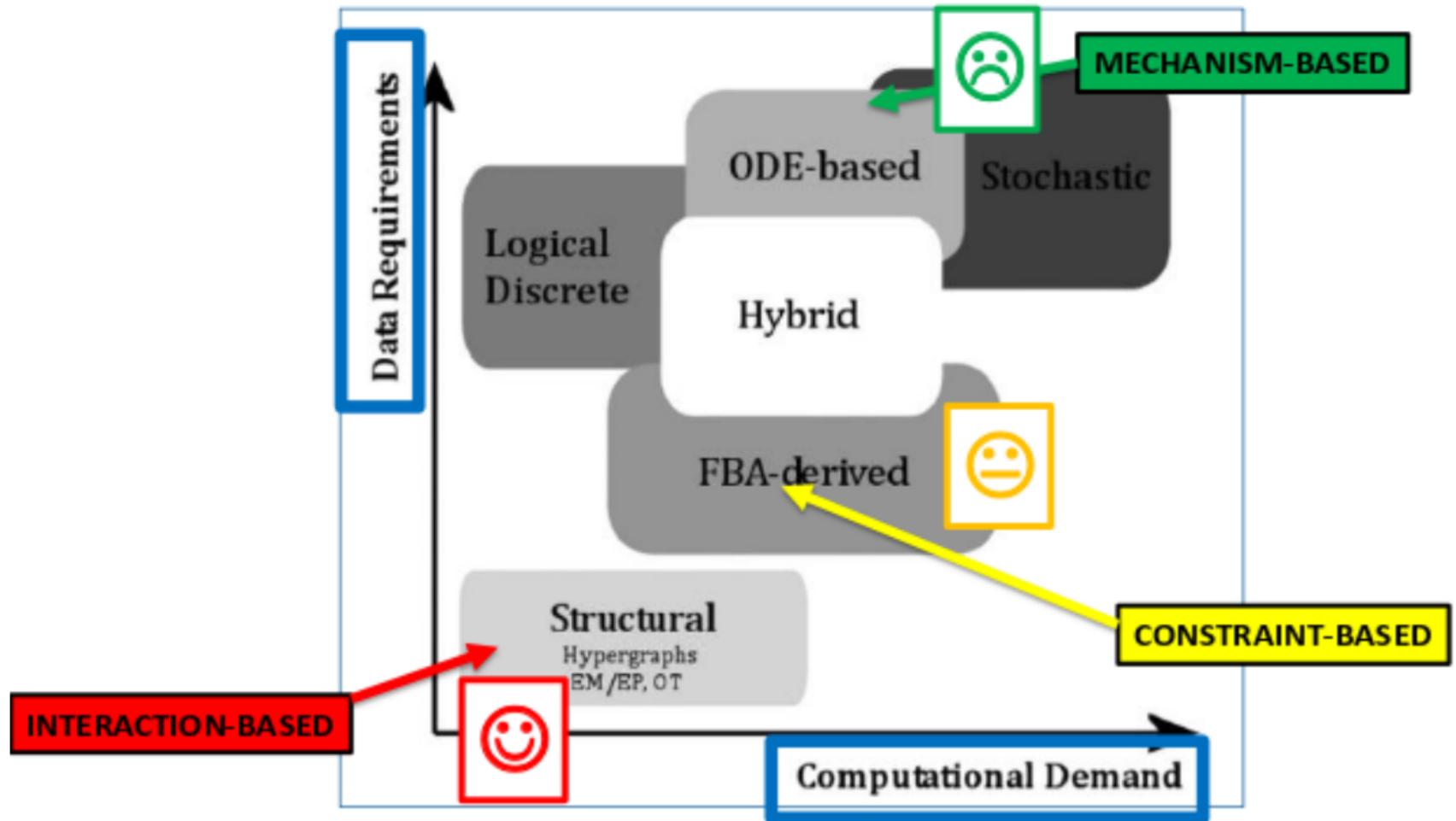
- ▶ metabolic pathways;
- ▶ signal transduction pathways;
- ▶ cell population.

based ordinary/stochastic differential equations, Monte-Carlo simulation
... → *dynamic behaviour analysis*.

Different modeling approaches:



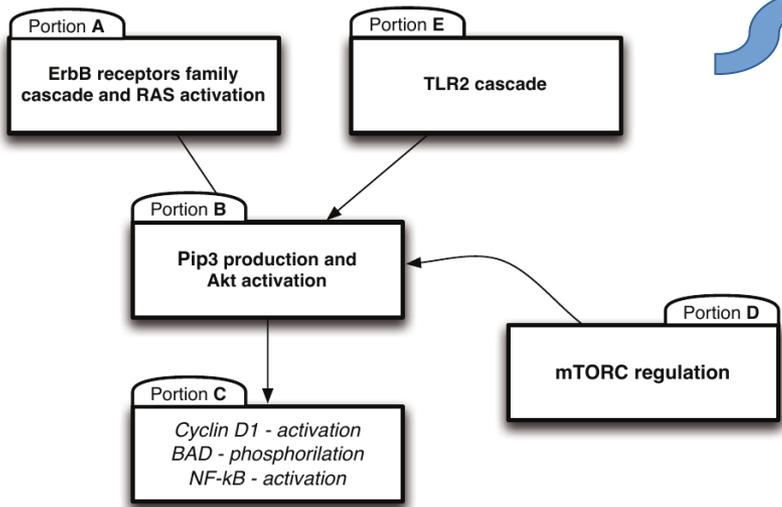
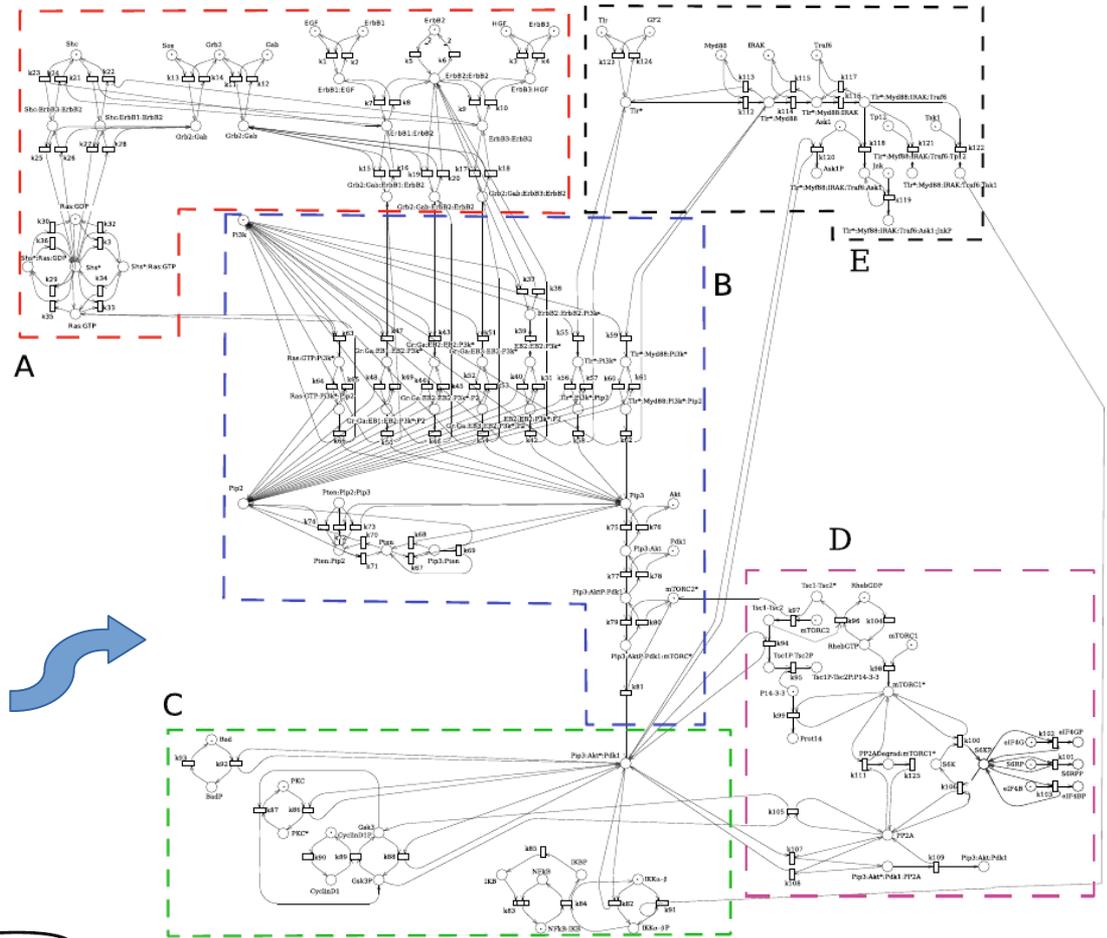
Data requirements VS Computational demand:



Petri Nets

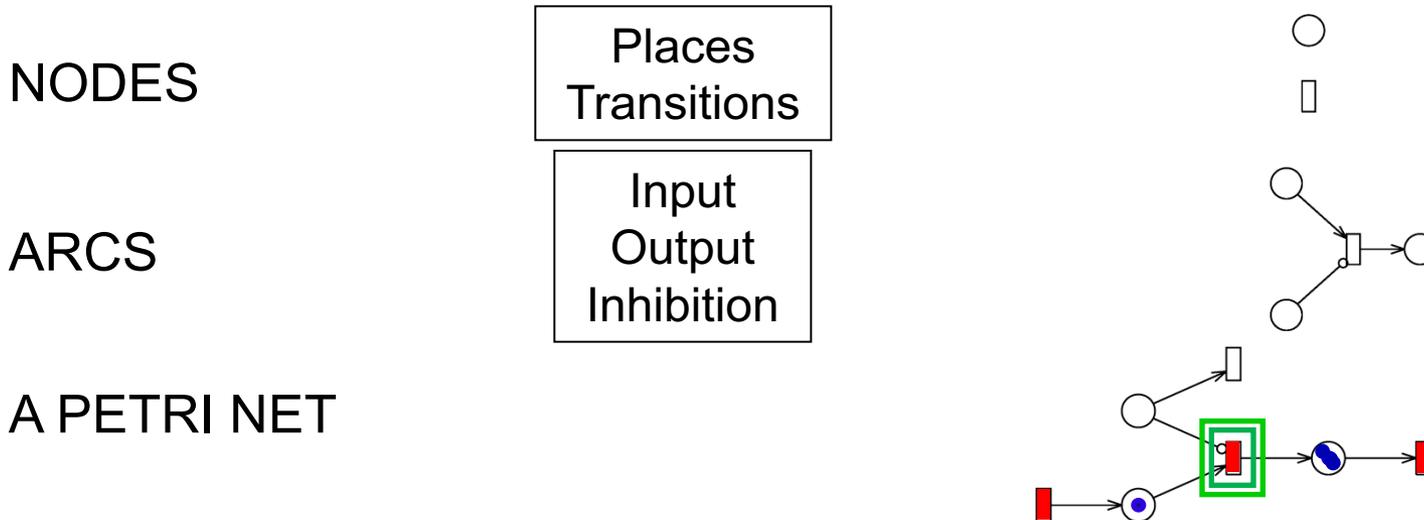
Petri nets are graphical modeling formalisms which are becoming quite popular to build models of biological systems.

They can be used to represent in a simple and intuitive manner many important features of biological system easily to understand also by non-mathematicians and non-computer scientists.

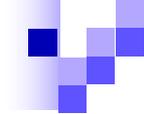


Petri Nets: Definition, Notation and Rules

Petri nets are bipartited directed graphs



- A marking M is an assignment of tokens to places
- A transition is *enabled* if at least one token exists in each of its *input* places, and no tokens exist in its *inhibition* places
- A transition *may fire* if it is *enabled*
- A Petri nets *executes* by *firing* transitions
 - A transition *fires* by *removing* tokens from each of its input places and *depositing* tokens in each of its output places
- Dynamic properties of Petri nets result from their *execution* controlled by the position and movement of tokens



- ❑ **metabolic networks**

 - signal transduction networks**

 - gene regulatory networks**

- ❑ **transitions**

 - > *(reversible, stoichiometric, enzyme-catalyzed) chemical reactions,*

 - > *conversions/transport of metabolites, proteins, . . .*

 - > *complexations/decomplexations, de-/phosphorylations, . . .*

- ❑ **places**

 - > *(primary, secondary) chemical compounds,*

 - > *(various states of) proteins, protein complex, genes, . . .*

- ❑ **tokens**

 - > *molecules, moles, . . .*

 - > *concentration levels, gene expression levels, . . .*

 - e.g., high/low = present/not present, or any finite integer number*

Typical basic structures

$$A \rightarrow B + C$$

$$A + B \rightarrow C$$

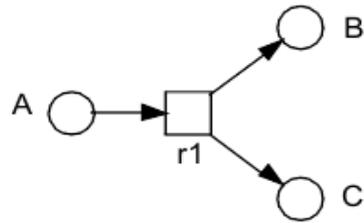
$$\begin{aligned} A &\rightarrow B, \\ A &\rightarrow C \end{aligned}$$

$$\begin{aligned} A &\rightarrow C, \\ B &\rightarrow C \end{aligned}$$

$$A \xrightleftharpoons{E} B$$

Typical basic structures

$$A \rightarrow B + C$$



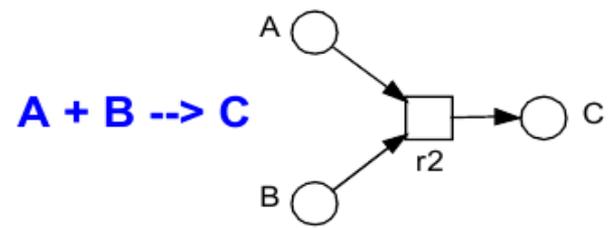
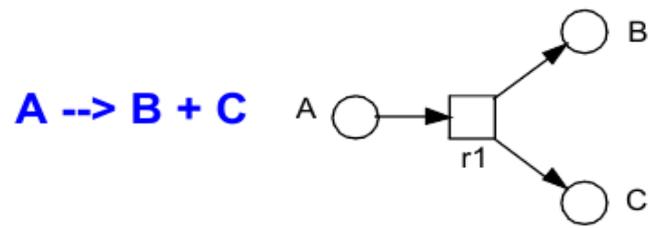
$$A + B \rightarrow C$$

$$A \rightarrow B, \\ A \rightarrow C$$

$$A \rightarrow C, \\ B \rightarrow C$$

$$A \overset{E}{\leftrightarrow} B$$

Typical basic structures



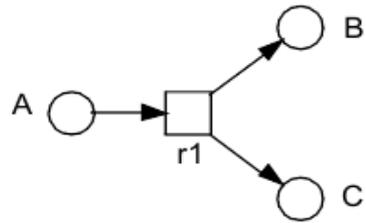
$A \rightarrow B,$
 $A \rightarrow C$

$A \rightarrow C,$
 $B \rightarrow C$

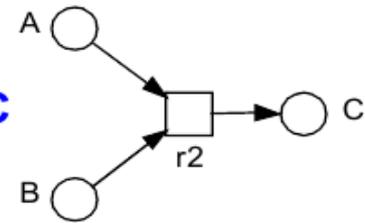
$A \xrightleftharpoons{E} B$

Typical basic structures

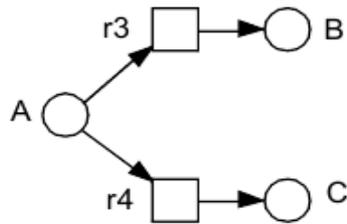
$A \rightarrow B + C$



$A + B \rightarrow C$



$A \rightarrow B,$
 $A \rightarrow C$

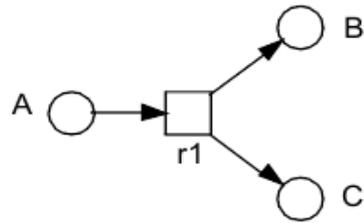


$A \rightarrow C,$
 $B \rightarrow C$

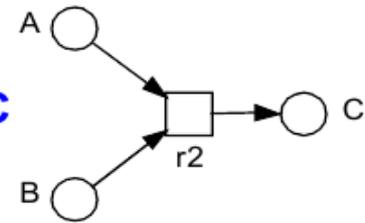
$A \overset{E}{\leftrightarrow} B$

Typical basic structures

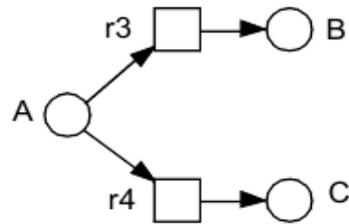
$A \rightarrow B + C$



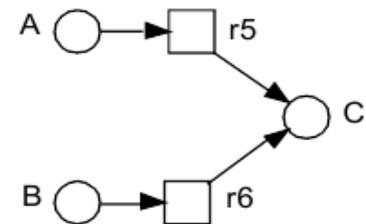
$A + B \rightarrow C$



$A \rightarrow B,$
 $A \rightarrow C$



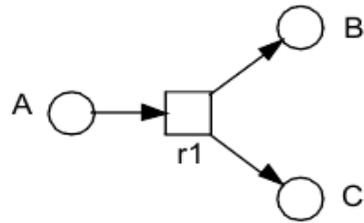
$A \rightarrow C,$
 $B \rightarrow C$



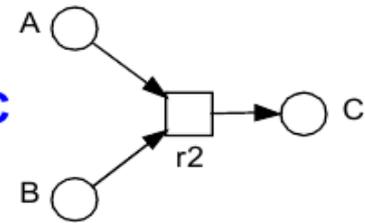
$A \xrightleftharpoons{E} B$

Typical basic structures

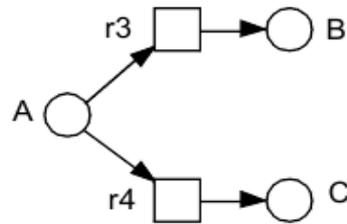
$A \rightarrow B + C$



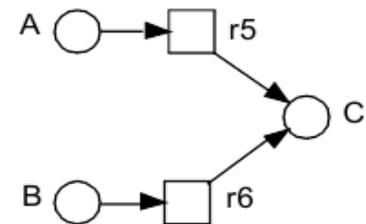
$A + B \rightarrow C$



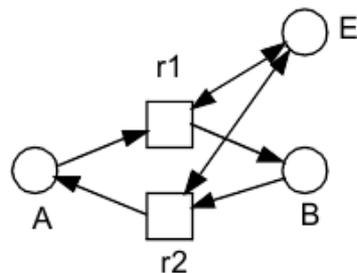
$A \rightarrow B,$
 $A \rightarrow C$



$A \rightarrow C,$
 $B \rightarrow C$

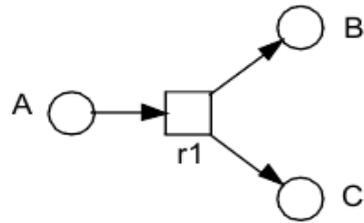


$A \overset{E}{\leftrightarrow} B$

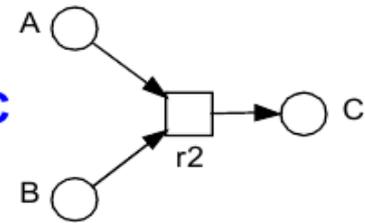


Typical basic structures

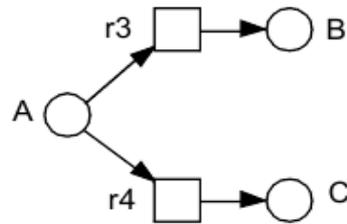
$A \rightarrow B + C$



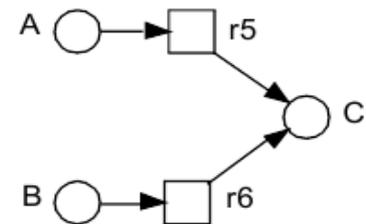
$A + B \rightarrow C$



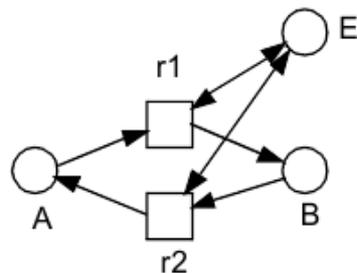
$A \rightarrow B,$
 $A \rightarrow C$



$A \rightarrow C,$
 $B \rightarrow C$

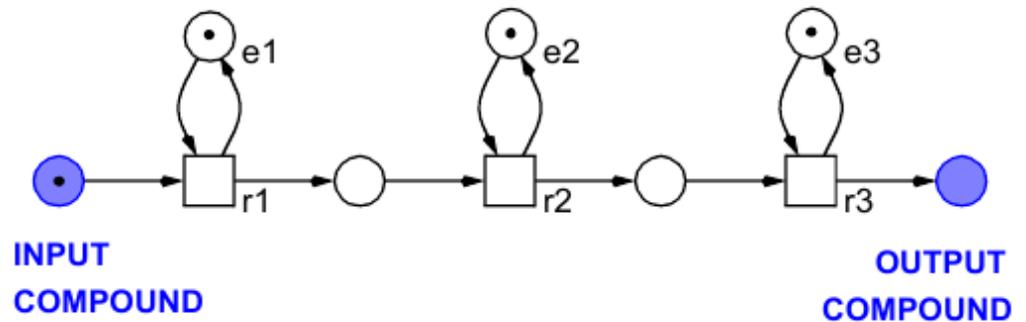


$A \overset{E}{\leftrightarrow} B$



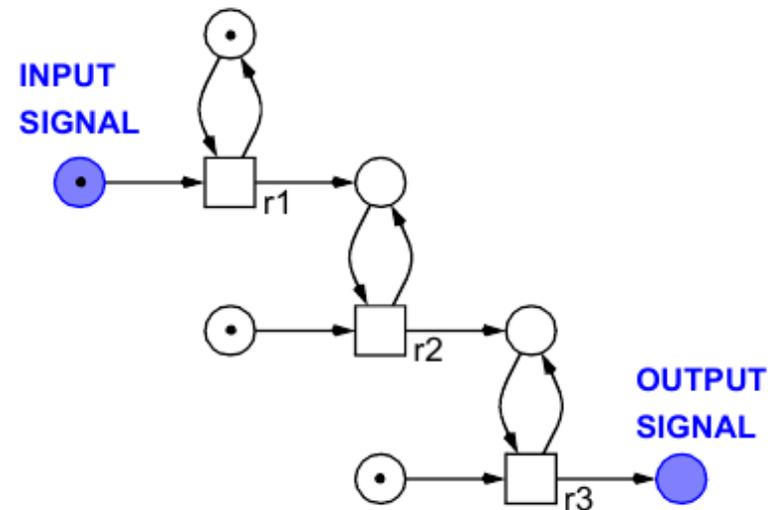
□ **metabolic networks**

-> *substance flows*



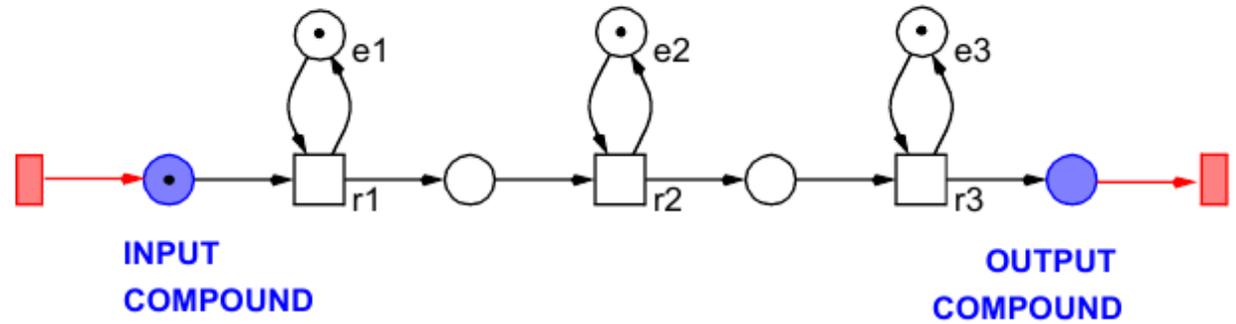
□ **signal transduction networks**

-> *signal flows*



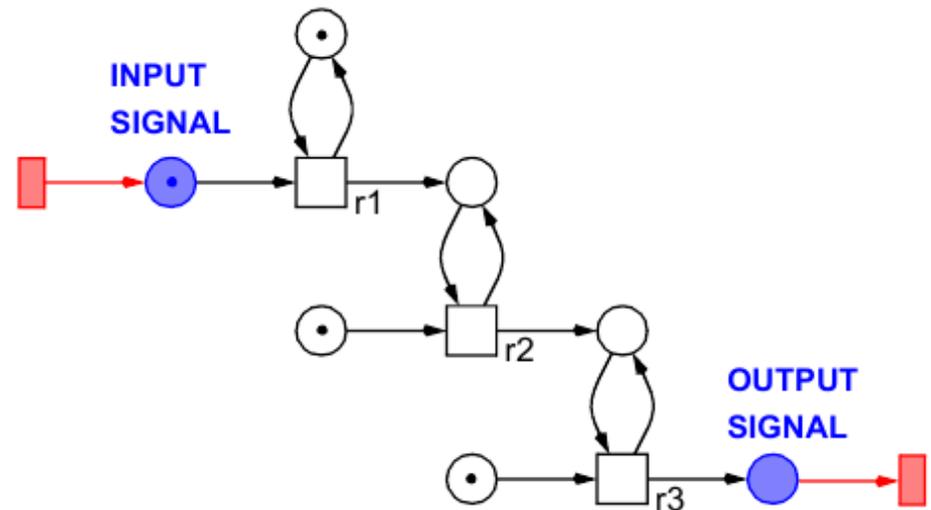
□ **metabolic networks**

-> *substance flows*



□ **signal transduction networks**

-> *signal flows*



-> **OPEN / CLOSED SYSTEMS**

Petri Nets: Formal Definition

A marked Petri net is formally defined by the following tuple

$$PN = (P, T, F, W, M_0)$$

where

$P = (p_1, p_2, \dots, p_P)$ is the set of places

$T = (t_1, t_2, \dots, t_T)$ is the set of transitions

$F = (P \rightarrow T) \& (T \rightarrow P)$ is the set of arcs

$W : f(t,p)$ is a weight function

$M_0 = (m_{01}, m_{02}, \dots, m_{0P})$ is the initial marking

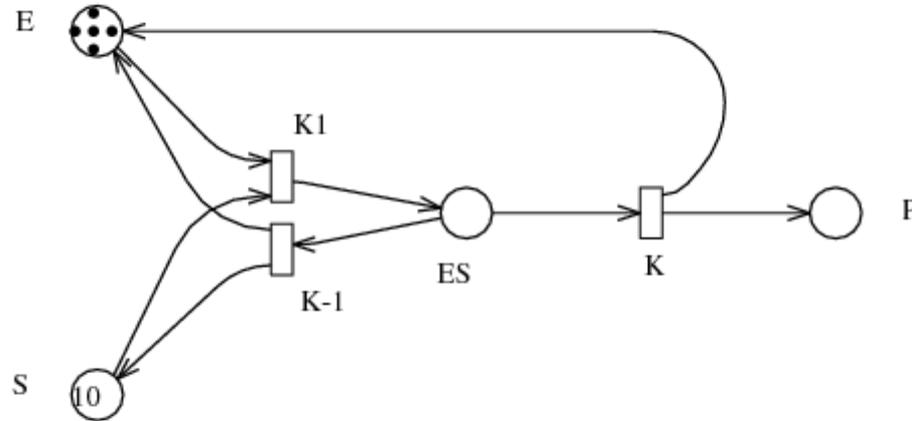
Combining the information provided by the flow relations and by the weight function, we obtain the *Incidence Matrix*

$$C = \begin{array}{c} \text{p} \\ \text{l} \\ \text{a} \\ \text{c} \\ \text{e} \\ \text{s} \end{array} \begin{array}{c} \text{transitions} \\ \hline C_{pt} \end{array}$$

with $c_{pt} = c_{pt}^+ + c_{pt}^- = w(t,p) - w(p,t)$

A simple example: Michaelis-Menten kinetics

Petri net model:



Set of places: $P = ?$

Set of transitions: $T = ?$

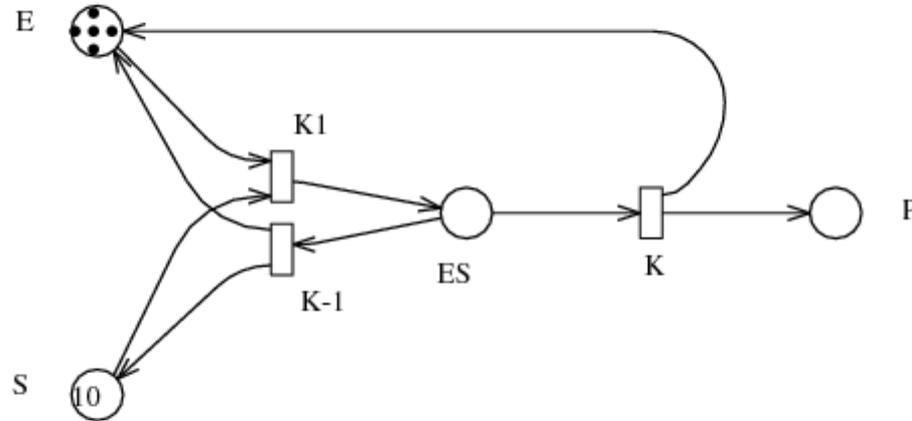
Incidence matrix:

$$C = \begin{matrix} & K1 & K-1 & K \\ \begin{matrix} E \\ S \\ ES \\ P \end{matrix} & \begin{bmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix} \end{matrix}$$

Initial marking: $M_0 = ?$

A simple example: Michaelis-Menten kinetics

Petri net model:



Set of places:

$$P = (E, S, ES, P)$$

Set of transitions:

$$T = (K1, K-1, K)$$

Incidence matrix:

$$C = \begin{array}{c} \\ \\ \\ \\ \end{array} \begin{array}{|c|c|c|c|} \hline & K1 & K-1 & K \\ \hline E & -1 & +1 & +1 \\ \hline S & -1 & +1 & \\ \hline ES & +1 & -1 & -1 \\ \hline P & & & +1 \\ \hline \end{array}$$

Initial marking:

$$M_0 = (5, 10, 0, 0)$$

Petri Nets: Basic Definitions

$RS(M_0)$

Set of markings reachable from M_0

$E(M)$

Set of transitions enabled in marking M

$M \xrightarrow{\sigma} M'$

M' is reachable from M by firing a sequence σ of transitions

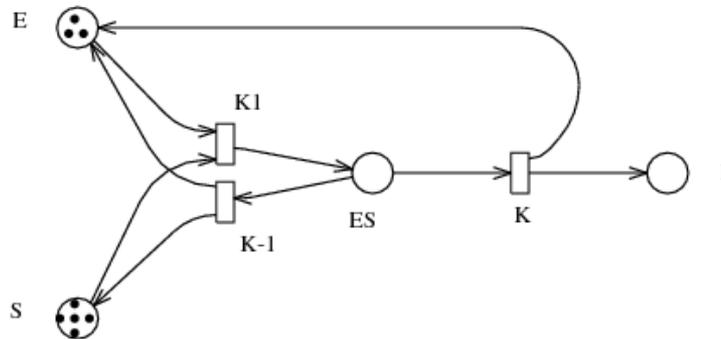
a transition t_r is enabled in marking M iff

$$M \geq [C_{t_r}^-]^T$$

$$M \xrightarrow{t_r} M' \equiv M - [C_{t_r}^-]^T + [C_{t_r}^+]^T = M'$$

A simple example: Michaelis-Menten kinetics

Petri net model:



Reachability graph :

	K1	K-1	K
E	-1	+1	+1
S	-1	+1	
ES	+1	-1	-1
P			+1

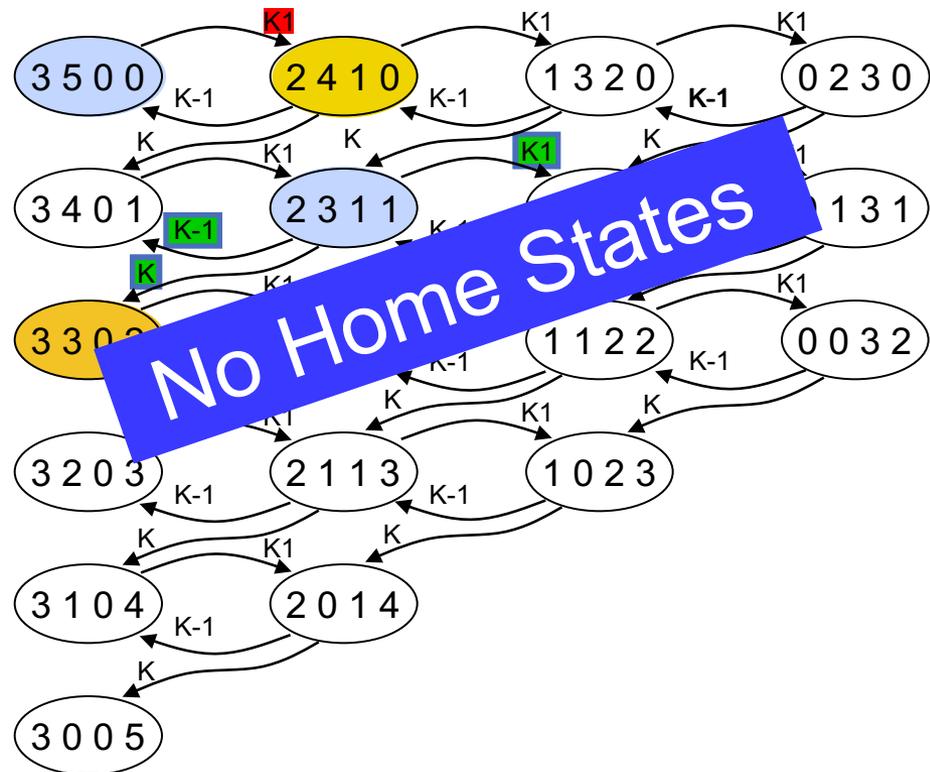
$$M + [C_r]^T = M'$$

$$(3,5,0,0) + (-1,-1,+1,0) = (2,4,1,0)$$

$$(2,3,1,1) + (+1,0,-1,+1) = (3,3,0,2)$$

$$E(2,3,1,1) = \{K_1, K_{-1}, K\}$$

$$E(3,3,0,2) = \{K_1\}$$





Petri Nets: Structural and Behavioural Properties

Structural properties of Petri nets are obtained from the incidence matrix, independently of the initial marking

Behavioural properties of Petri nets depend on the initial marking and are obtained from the reachability graph (finite case) of the net or from the covering tree (infinite case)

Petri Nets: P Semiflows

A Petri net is *strictly conservative* (or strictly invariant) iff

$$\sum_{p=1}^P m_p = \sum_{p=1}^P m_{0p}, \quad \forall M \in RS(M_0)$$

A Petri net is *conservative* (or P invariant) iff

$$\begin{aligned} &\exists Y = (y_1, y_2, \dots, y_P) > 0 \text{ such that} \\ &\sum_{p=1}^P y_p m_p = \sum_{p=1}^P y_p m_{0p} \quad \forall M \in RS(M_0) \end{aligned}$$

The integer solution Y of the equation

is called a P Semiflow

$$YC = 0$$

Petri Nets: Boundedness

A place p_i is *bounded* (k -bounded) iff

$$\forall M \in RS(M_0), \exists k : m_i \leq k$$

A Petri net is *bounded* (k -bounded) iff

$$\exists k : (\forall p_i \in P : is\ k - bounded)$$

A net covered by P -semiflows
Is bounded

Petri Nets: T Semiflows

The integer solution X of the equation

$$CX = 0$$

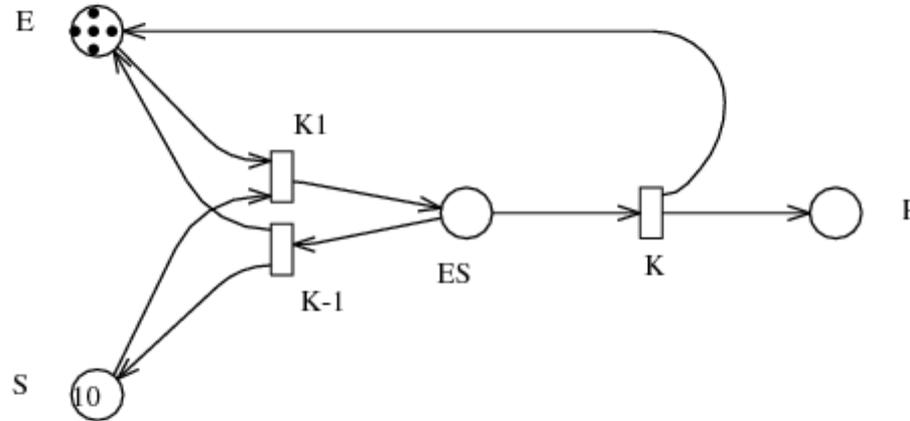
is called a T -Semiflow

A net covered by T -semiflows
may have *home states*

A net with home states is covered by T -semiflows

A simple example: Michaelis-Menten kinetics

Petri net model:



Set of places:

$$P = (E, S, ES, P)$$

Set of transitions:

$$T = (K1, K-1, K)$$

Incidence matrix:

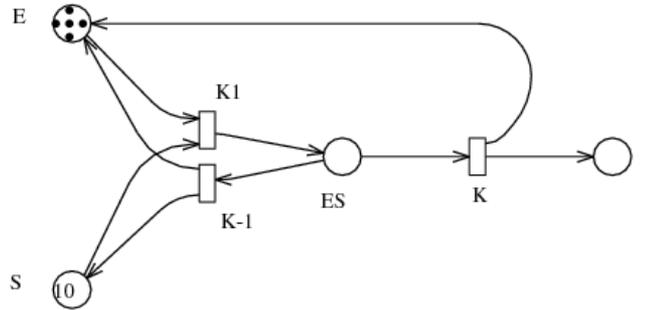
$$C = \begin{array}{c} \\ \\ \\ \end{array} \begin{array}{c} \\ \\ \\ \end{array} \begin{array}{ccc} K1 & K-1 & K \\ \hline E & -1 & +1 & +1 \\ S & -1 & +1 & \\ ES & +1 & -1 & -1 \\ P & & & +1 \end{array}$$

Initial marking:

$$M_0 = (3, 5, 0, 0)$$

A simple example: Michaelis-Menten kinetics

Petri net model:



Incidence matrix:

$$C = \begin{array}{c} \begin{array}{ccc} & K1 & K-1 & K \\ E & -1 & +1 & +1 \\ S & -1 & +1 & \\ ES & +1 & -1 & -1 \\ P & & & +1 \end{array} \end{array}$$

P semiflows ($YC = 0$):

$$y = (0, 1, 1, 1)$$

$$y = (1, 0, 1, 0)$$

The net is covered by P -semiflows, thus is bounded

T semiflows ($CX = 0$):

$$x = (1, 1, 0)$$

The net is not covered by T -semiflows, thus is not live

❑ **How many tokens can reside at most in a given place ?**

-> $(0, 1, k, \infty)$

-> **BOUNDEDNESS**

❑ **How often can a transition fire ?**

-> $(0\text{-times}, n\text{-times}, \infty\text{-times})$

-> **LIVENESS**

❑ **How often can a system state be reached ?**

-> *never*

-> **UNREACHABLE** -> **SAFETY PROPERTIES**

-> *n-times*

-> **REPRODUCIBLE**

-> *∞ -times*

-> **REVERSIBILITY**

❑ **Are there behaviourally invariant net structures ?**

-> *token conservation*

-> **P - INVARIANTS**

-> *token distribution reproduction*

-> **T - INVARIANTS**

❑ **static analyses** → **no state space construction**

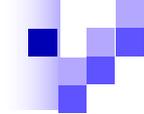
-> *structural properties (graph theory)*

-> *P / T - invariants (linear algebra)*

❑ **dynamic analyses** → **total/ partial state space construction**

-> *analysis of **general** behavioural system properties,
e.g. boundedness, liveness, reversibility, . . .*

-> *model checking of **special** behavioural system properties,
e.g. reachability of a given (sub-) system state (with constraints),
reproducability of a given (sub-) system state (with constraints)*



Stochastic (Exponential) Petri Nets



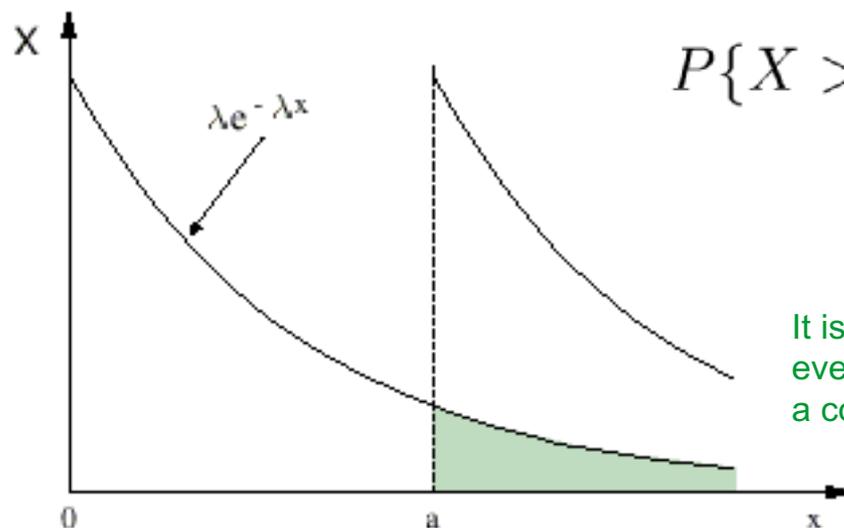
Stochastic (Exponential) Petri Nets

- The delay of a transition is a random variable
- Timed Transition PN with atomic firing and race policy in which transition delays are random variables *exponentially* distributed are called Stochastic Petri Nets (SPN)
- SPN is the name chosen by Molloy in 1982, but a more adequate one is Exponential Petri Nets

Exponential distributions

- The exponential pdf is $f_X(x) = \lambda e^{-\lambda x} \quad (x \geq 0)$

it is the only continuous distribution for which the *memoryless property* holds



$$P\{X > x + \alpha | X > \alpha\} = P\{X > x\}$$

It is used to describe a process in which events occur continuously and independently at a constant average rate λ

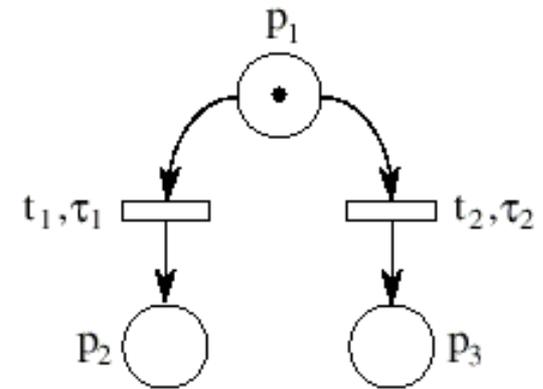
Why Exponential distributions ?

- If X is the random variable for t_1 and Y is the random variable for t_2

if the race policy is assumed, then
the random variable that describes:

how long the system stays in marking $1 \cdot p_1$

is defined as $Z = \min(X, Y)$





Markov chains

- Continuous Time Markov chain – CTMC

is a simple type of stochastic process with discrete state space

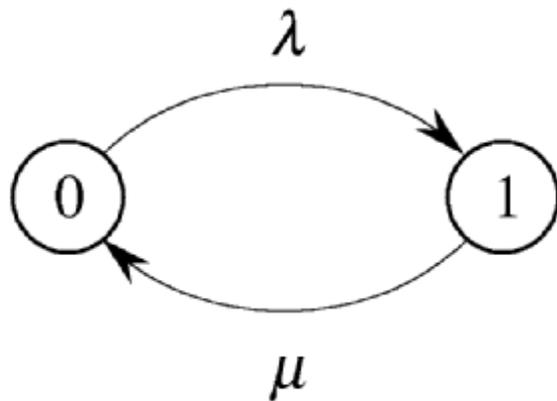
- Sojourn times in states are exponentially distributed random variables

and

- future evolution only depends on the present state
(there is no need to keep history information)

Markov chains

- CTMC can be described as direct graph with labeled transitions; the value of the label describes the rate associated with that change of state



State transition rate diagram

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}$$

Infinitesimal generator

Markov chains

- The solution of the CTMC at time t amounts to the computation of the solution of a set of differential equations called forward Chapman-Kolmogorov equations (as many equations as there are states in the CTMC)

Variation of probability in the interval \longrightarrow

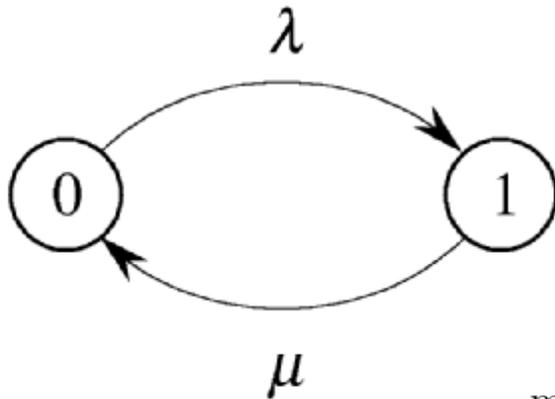
$$\frac{d\pi_j(t)}{dt} = \sum_{i \neq j} \pi_i(t)q_{ij} - \pi_j(t)q_{jj}$$

$$\frac{d\boldsymbol{\pi}(t)}{dt} = \boldsymbol{\pi}(t)Q$$

Matrix form where Q is the infinitesimal generator

- The solution is represented by the state probability vector at time t

Markov chains



$$Q = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}$$

$\pi(0) = (1, 0)$
meaning that the system is in state 0 at time 0 with probability 1.

Chapman-Kolmogorov equations

$$\frac{d(\pi_0)(t)}{dt} = -\lambda * \pi_0(t) + \mu * \pi_1(t)$$

$$\frac{d(\pi_1)(t)}{dt} = -\mu * \pi_1(t) + \lambda * \pi_0(t)$$

$$\pi_0(0) = 1$$

$$\pi_1(0) = 0$$

ODE solution in a nutshell

$$\begin{array}{l} \frac{dy}{dx} = F(x) \\ dy = F(x) dx \end{array} \quad \longrightarrow \quad y = \int^x F(\lambda) d\lambda + C$$

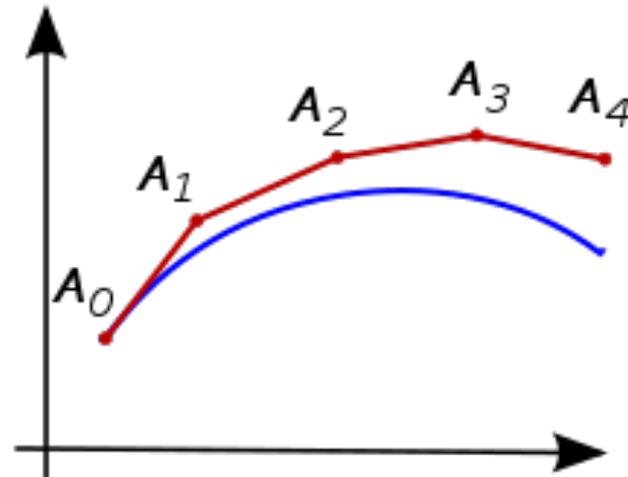
It is not always possible to evaluate indefinite integrals

To cope with this numerical methods can be used.

Euler method:

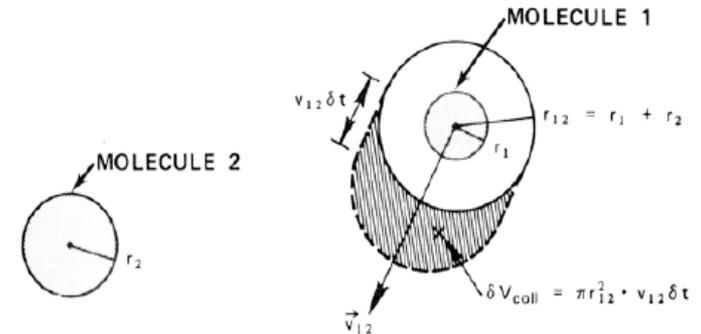
Its basic idea is to use a known point as a "starter," and then use the tangent line through this known point to jump to a new point

$$y_{n+1} = y_n + hf(t_n, y_n).$$



Relevance of CTMCs in biological models

In 1977 Daniel T. Gillespie developed a theory based on the hypothesis that *collisions* among molecules in constant volumes and temperatures are random



He used this theory to show that the kinetics of the chemical reactions deriving from these collisions corresponds to an underlying stochastic process that is a **Continuous Time Markov Chain (CTMC)**.

SPN definition

- An SPN is defined as a 7-tuple

$$\text{SPN} = (\text{PN}, R(.))$$

where

$\text{PN} = (P, T, F, W, M_0)$ is the P/T system underlying the SPN

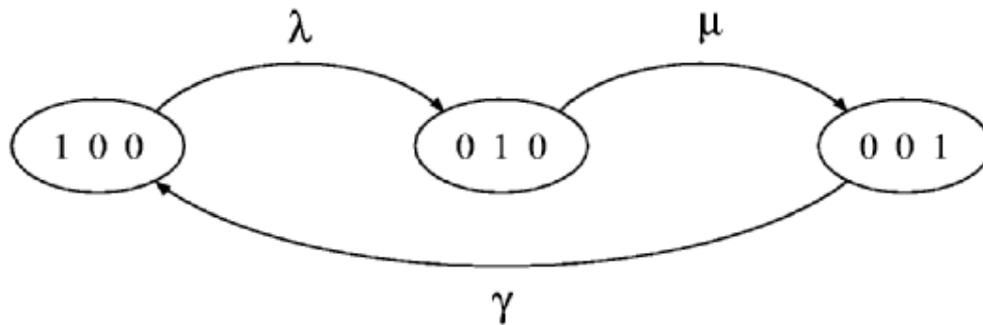
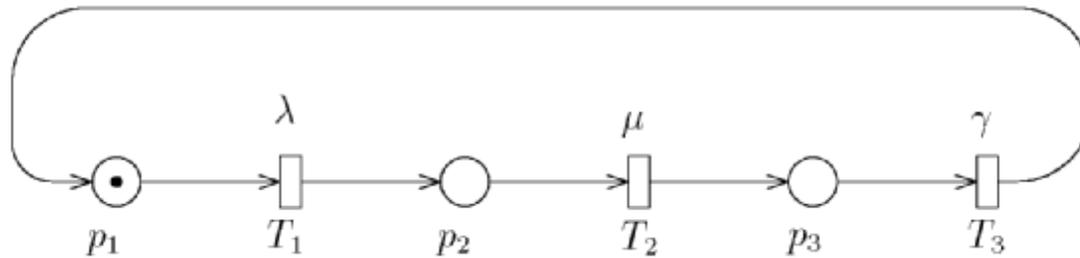
- Transitions have an exponentially distributed delay
- $R: T \rightarrow \mathcal{Real}$ assigns a rate to each transition
(inverse of the mean firing time)



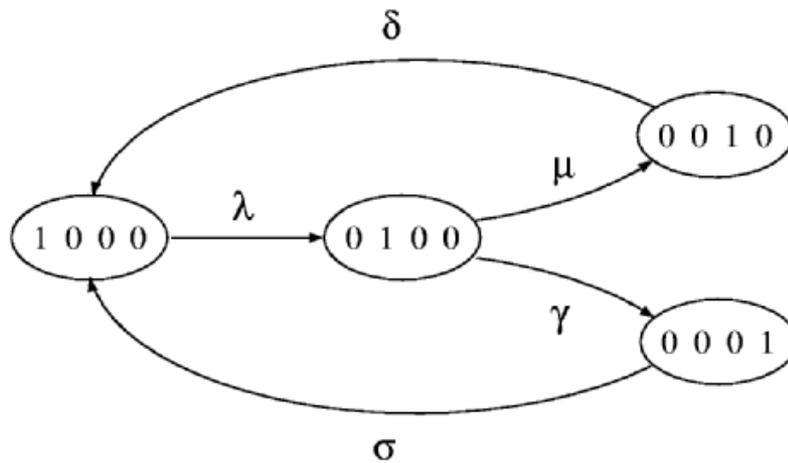
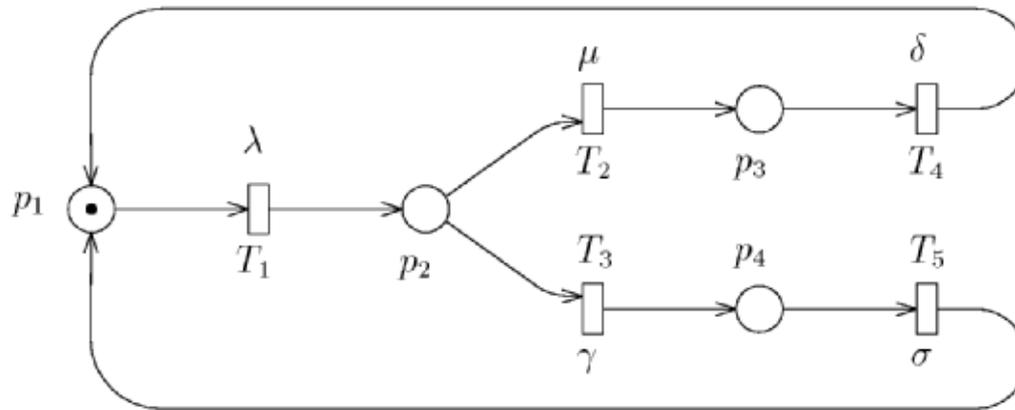
SPN definition

- The stochastic process underlying an SPN is a CTMC in which
 - the state transition rate diagram is isomorphic to the reachability graph
 - the transition labels are computed from the $R()$ functions of the transitions enabled in a state

SPN w/o synchronization and choices



SPN with choices



SPN to CTMC

- When the model is very complex (a huge number of places and tokens) → we have the state space explosion

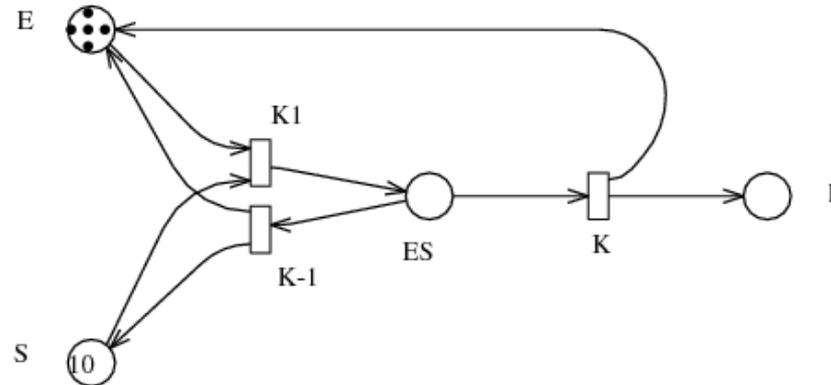
Then, the Chapman-Kolmogorov equations cannot be written since we have too much equations (one for each state)

- To deal with this we approximate the stochastic process as a deterministic one in which each system quantity is defined by an ODE.

$$\begin{aligned}\frac{dm_{p_i}(\nu)}{d\nu} &= \sum_{j:C(p_i,t_j)\neq 0} C(p_i,t_j)r(t_j) \prod_{h:C^-(p_h,t_j)\neq 0} m_{p_h}(\nu)^{|C^-(p_h,t_j)|} \\ m_{p_i}(0) &= m_{p_i,0}\end{aligned}$$

where ν is time, t_i the transition i th and m_{p_i} number of tokens/molecules in p_i

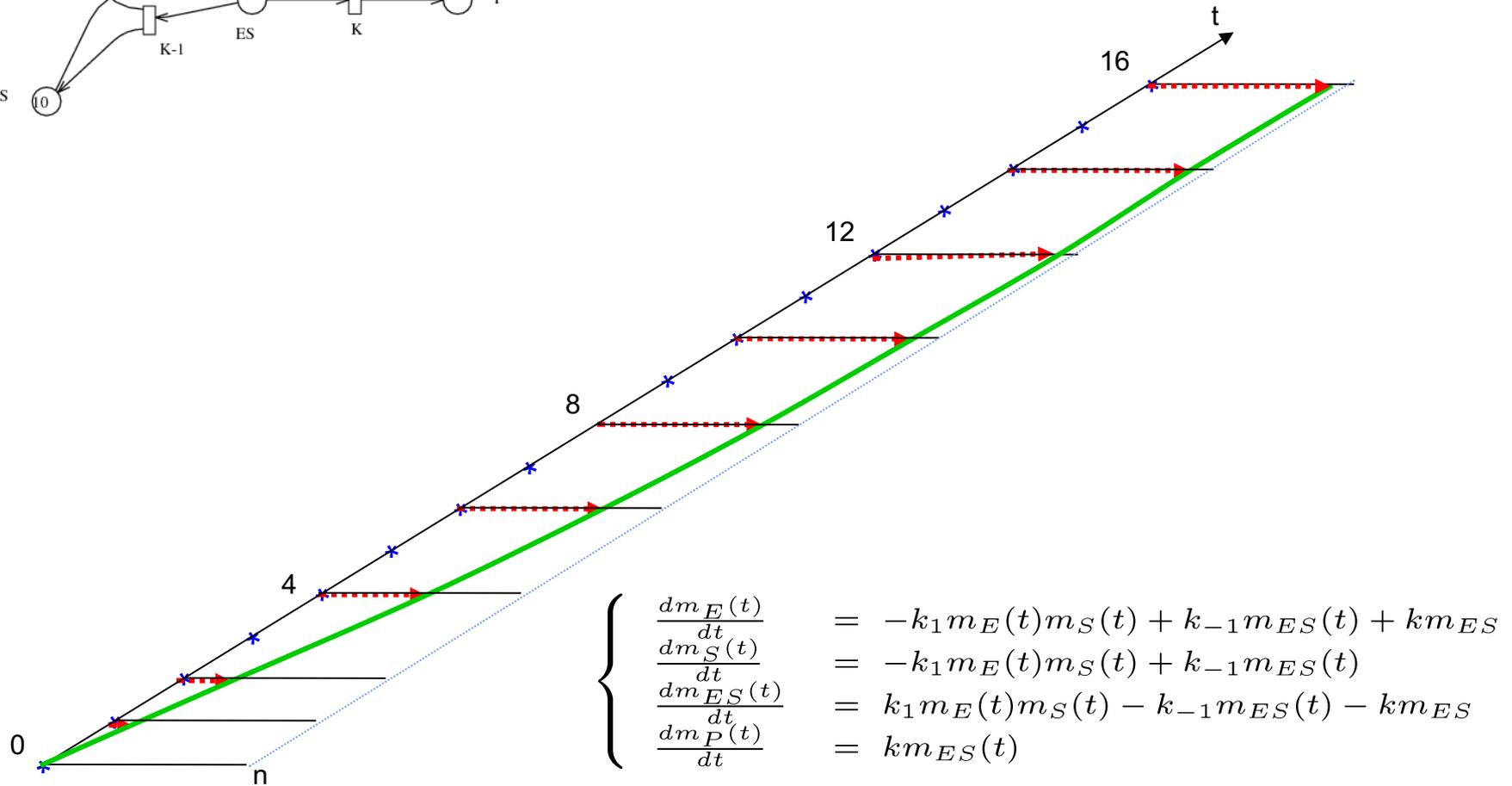
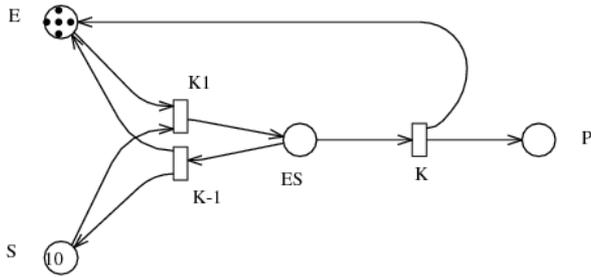
Simple example – Deterministic approximation



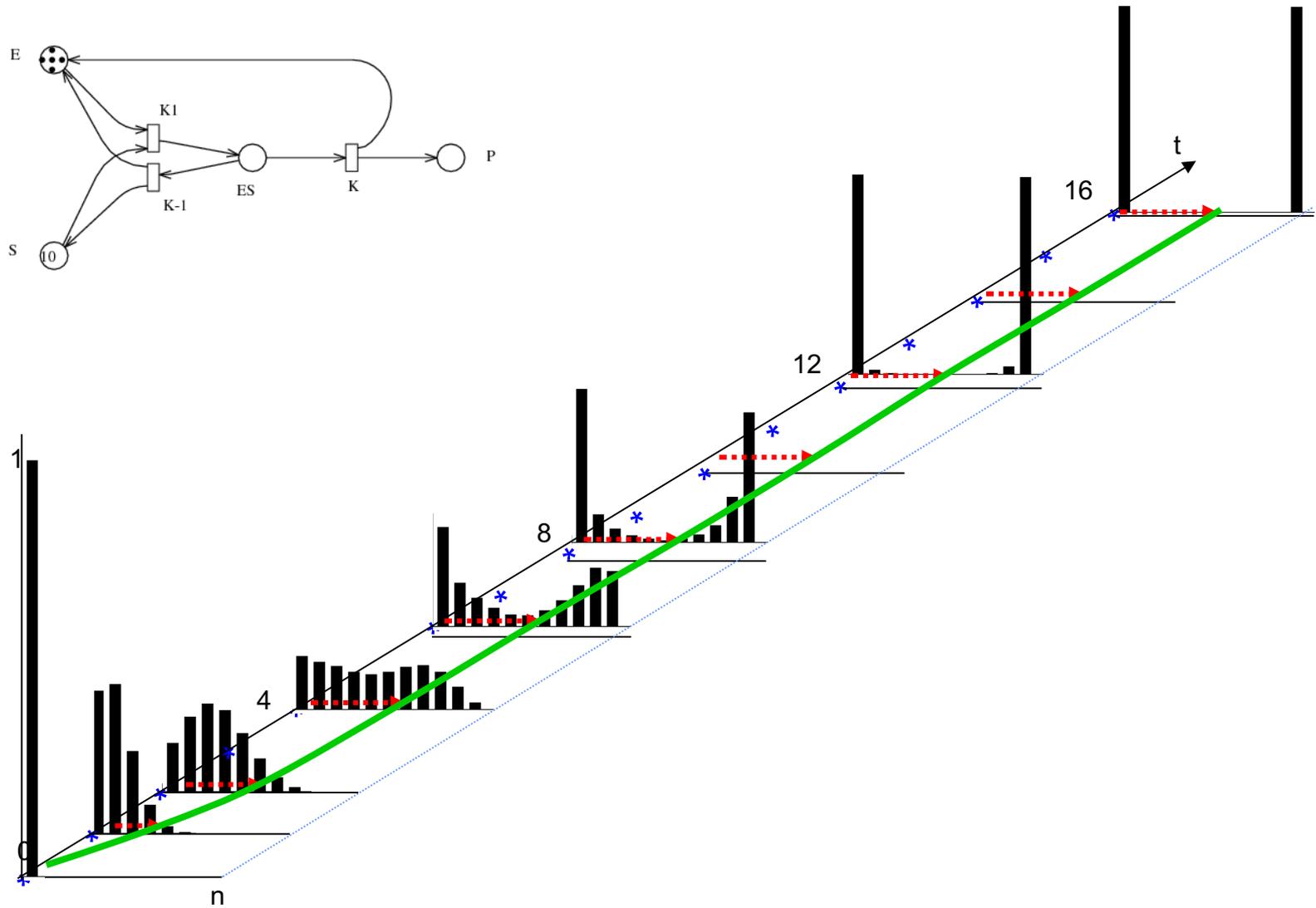
$$\left\{ \begin{array}{l} \frac{dm_E(t)}{dt} = -k_1 m_E(t) m_S(t) + k_{-1} m_{ES}(t) + k m_{ES} \\ \frac{dm_S(t)}{dt} = -k_1 m_E(t) m_S(t) + k_{-1} m_{ES}(t) \\ \frac{dm_{ES}(t)}{dt} = k_1 m_E(t) m_S(t) - k_{-1} m_{ES}(t) - k m_{ES} \\ \frac{dm_P(t)}{dt} = k m_{ES}(t) \end{array} \right.$$

4 ODEs using Deterministic approximation vs 51 ODEs using Chapman-Kolmogorov equations

Simple example – ODE Analysis



More complex example – CTMC Analysis



Summary

❑ **representation of bionetworks by Petri nets**

- > *partial order representation*
- > *formal semantics*
- > *unifying view*

- > *better comprehension*
- > *sound analysis techniques*

❑ **purposes**

- > *animation*
- > *model validation against consistency criteria*
- > *qualitative / quantitative behaviour prediction*

- > *to experience the model*
- > *to increase confidence*
- > *experiment design, new insights*

❑ **step-wise model development**

- > *qualitative model*
- > *discrete quantitative model*
- > *continuous quantitative model*

- > *discrete Petri nets*
- > *stochastic Petri nets*
- > *continuous Petri nets = ODEs*



SPN with (self) concurrency

- If there can be more than one token in the initial marking then it is necessary to consider:
 - transitions that are enabled in the same state but that are not in conflict
 - transitions with k- and infinite server semantics

CTMC analysis (1)

- Let $\{X(t), t, 0\}$ to be the mathematical representation of the CTMC.
- Denote with $p_{ij}(t)$ the probability of finding the CTMC in state s_j at time $z+t$, given that it was in state s_i at time z , independently of z
- In matrix notation we define

$$\mathbf{H}(t) \triangleq [p_{ij}(t)]$$

- Using the Chapman-Kolmogorov equations it is possible to show that

$$\frac{d\mathbf{H}(t)}{dt} = \mathbf{H}(t) \mathbf{Q}$$

- From which we obtain

$$\mathbf{H}(t) = e^{\mathbf{Q}t}$$

CTMC analysis (2)

- Let

$$\pi_j(t) \triangleq Pr\{\mathbf{X}(t) = j\} \quad \text{for all } s_j \in \mathcal{S}$$

be the probability of finding the CTMC in state s_j at time t , and

$$\pi_j(0) \triangleq Pr\{\mathbf{X}(0) = j\} \quad \text{for all } s_j \in \mathcal{S}$$

be the initial distribution, we have

$$\pi(t) = \pi(0)\mathbf{H}(t)$$

which becomes

$$\pi(t) = \pi(0)e^{\mathbf{Q}t}$$

- This is also the solution of the following differential equation

$$\frac{d\pi(t)}{dt} = \pi(t)\mathbf{Q}$$

- that in detailed form is

$$\frac{d\pi_j(t)}{dt} = q_{jj}\pi_j(t) + \sum_{k \neq j} q_{kj}\pi_k(t)$$

CTMC solution – finite state space

■ Transient

$$\pi(t) = \pi(0)\mathbf{H}(t) \quad \mathbf{H}(t) = e^{\mathbf{Q}t} \quad e^{\mathbf{Q}t} = \sum_{k=0}^{\infty} \frac{(\mathbf{Q}t)^k}{k!}$$

- Direct computation of \mathbf{Q}^k is numerically unstable
- There are methods (e.g., uniformization) for efficient and effective computation of the result

■ Steady state

- If the CTMC is ergodic then a limit exist for the state probability

$$\pi = \lim_{t \rightarrow \infty} \pi(t)$$

- In this case the result is obtained from the solution of a system of linear equations

$$0 = \pi\mathbf{Q}$$