

# Petri Nets for Systems & Synthetic Biology

in memory of Nadia Busi

Monika Heiner  
Brandenburg University of Technology Cottbus, Dept. of CS

- MODELLING WITH PETRI NETS

- MODEL-BASED SYSTEM ANALYSIS

- .-.-> *qualitative Petri nets*
- .-.-> *stochastic Petri nets*
- .-.-> *continuous Petri nets*

- CASE STUDY

- > *model checking in the three paradigms*

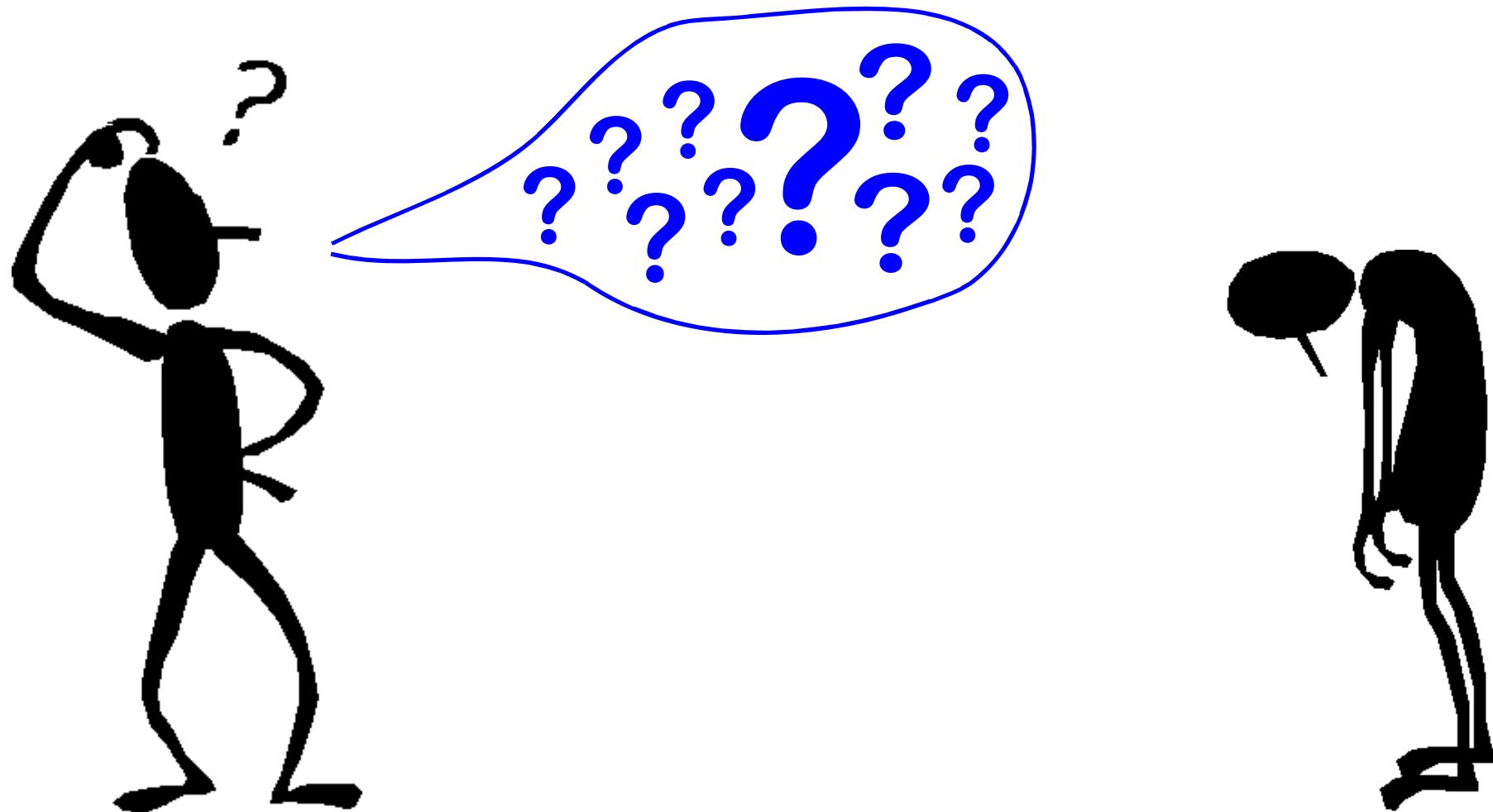
- SUMMARY

- > *challenges / open questions*

MEDICAL TREATMENT



MEDICAL TREATMENT, APPROACH 1- TRIAL-AND-ERROR DRUG PRESCRIPTION



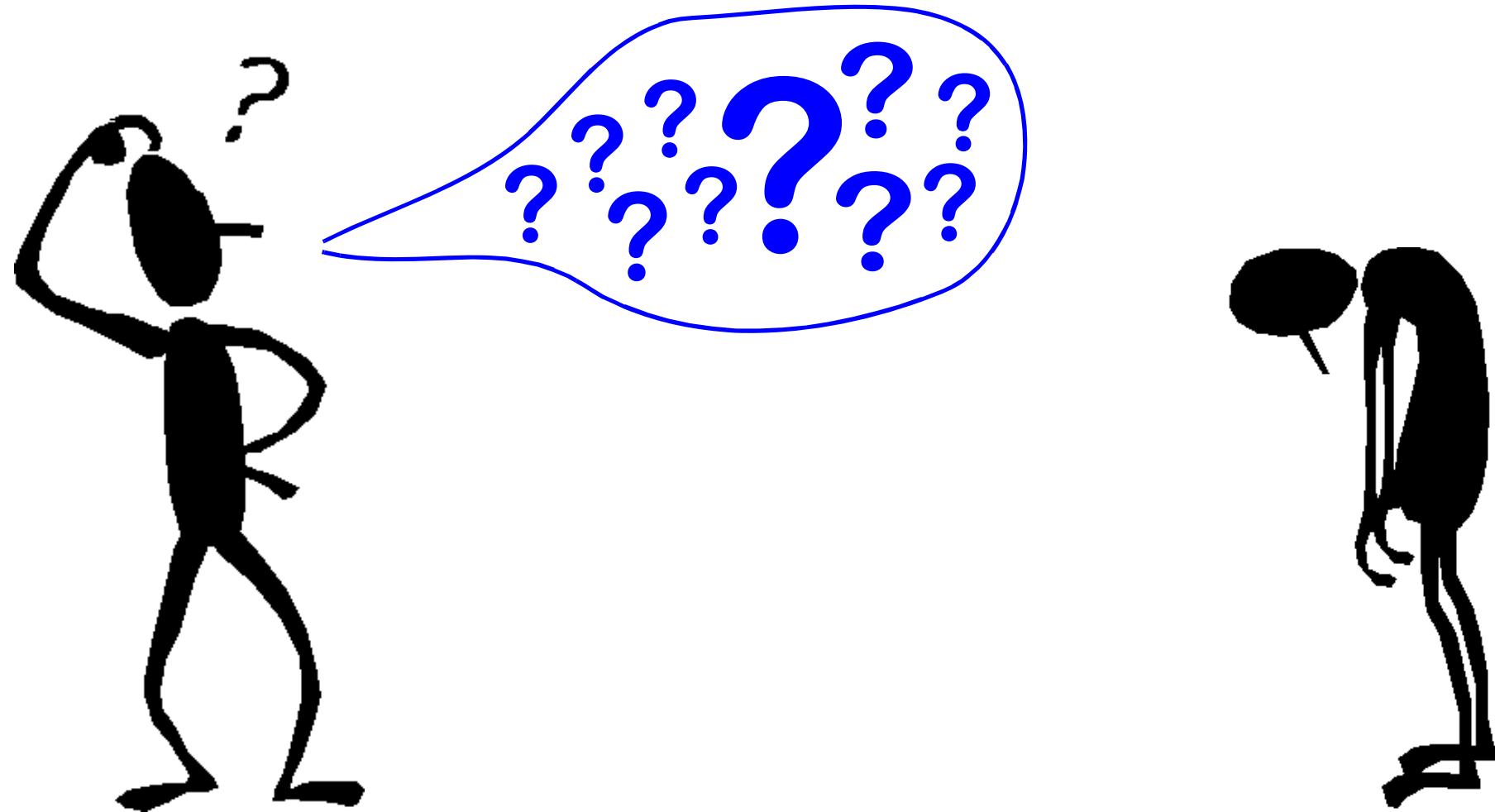
MEDICAL TREATMENT, APPROACH 1- TRIAL-AND-ERROR DRUG PRESCRIPTION



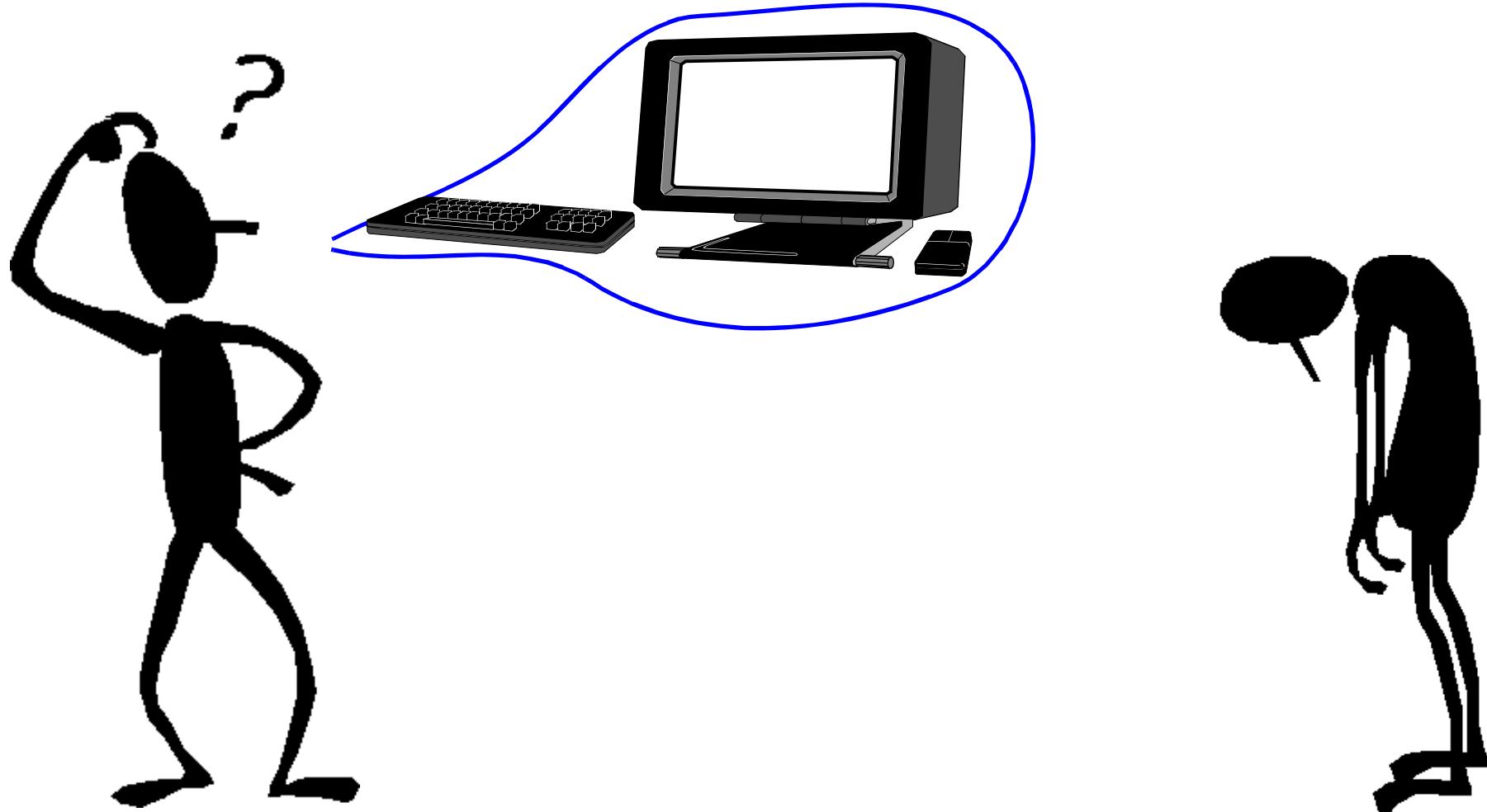
MEDICAL TREATMENT, APPROACH 1- TRIAL-AND-ERROR DRUG PRESCRIPTION



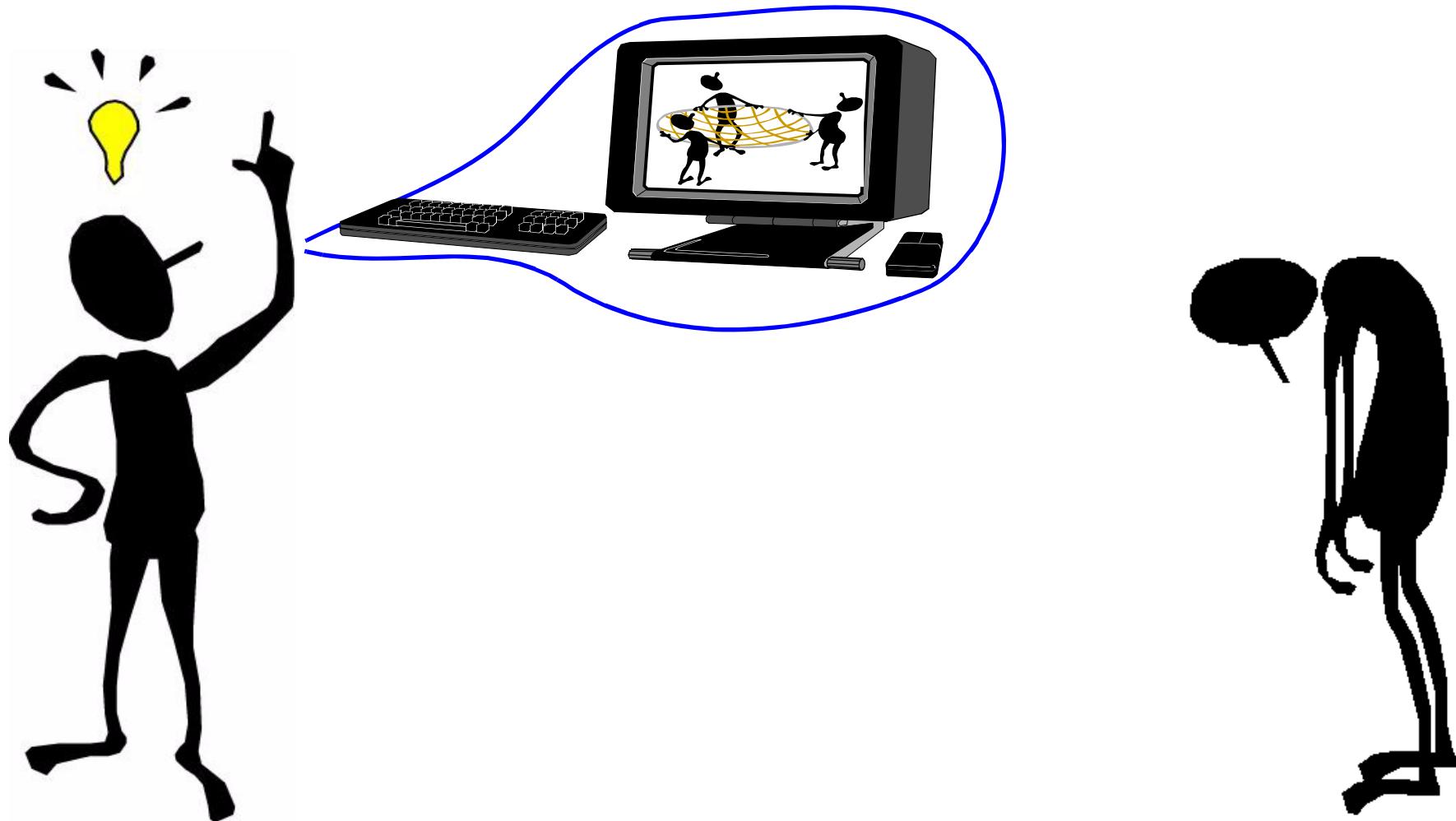
MEDICAL TREATMENT, APPROACH 2



MEDICAL TREATMENT, APPROACH 2 - MODEL-BASED DRUG PRESCRIPTION



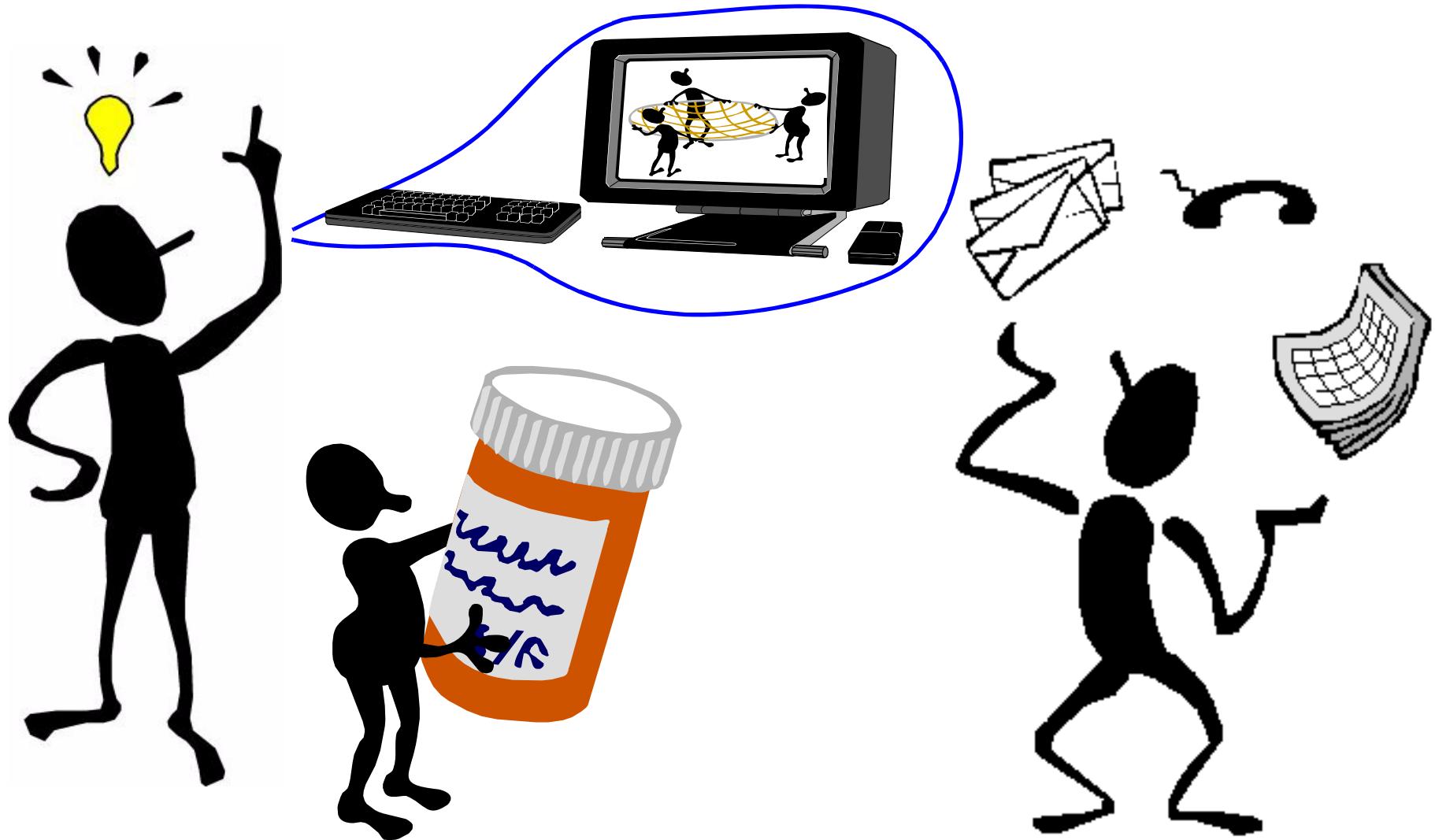
MEDICAL TREATMENT, APPROACH 2 - MODEL-BASED DRUG PRESCRIPTION



MEDICAL TREATMENT, APPROACH 2 - MODEL-BASED DRUG PRESCRIPTION



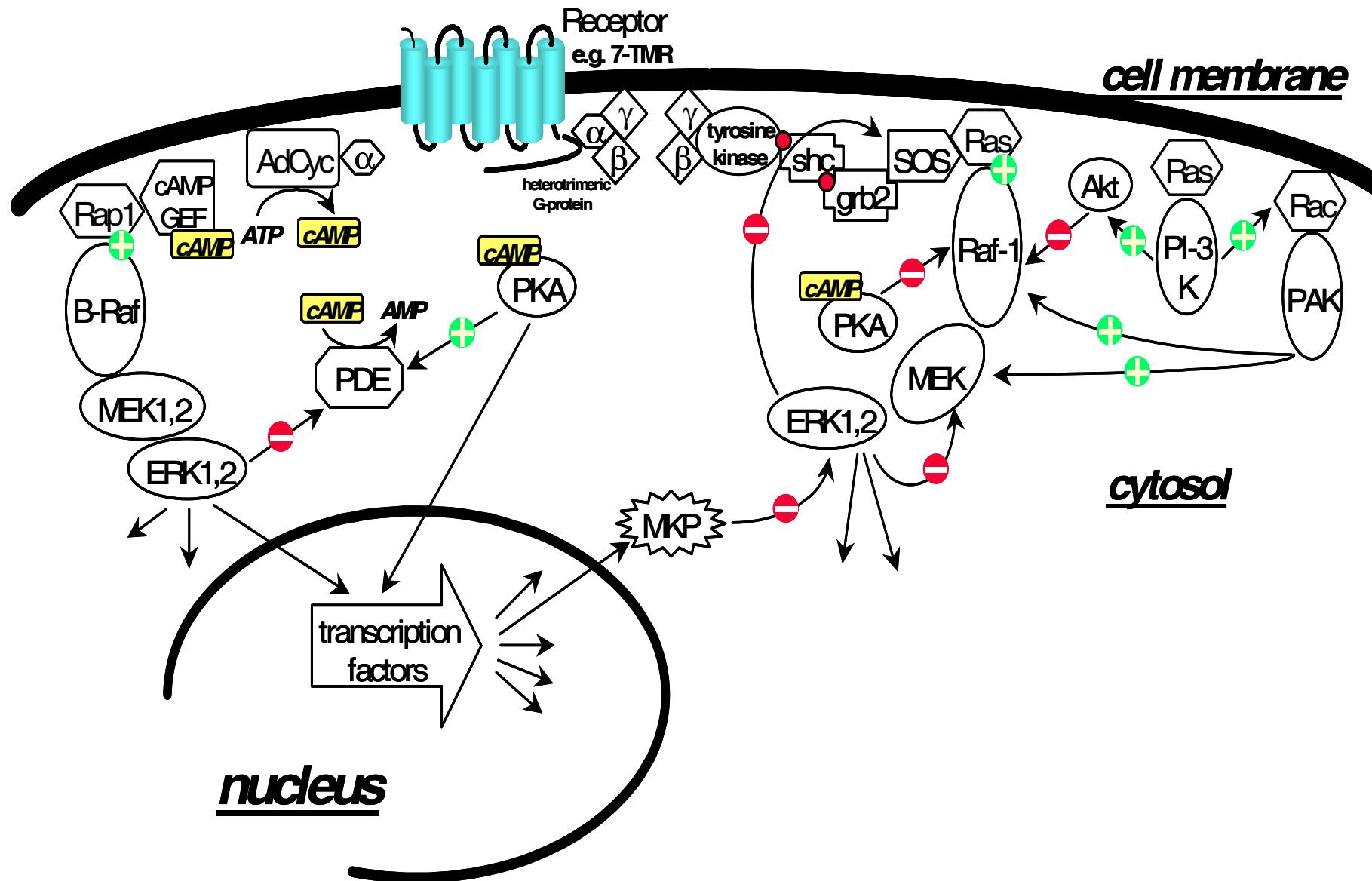
MEDICAL TREATMENT, APPROACH 2 - MODEL-BASED DRUG PRESCRIPTION

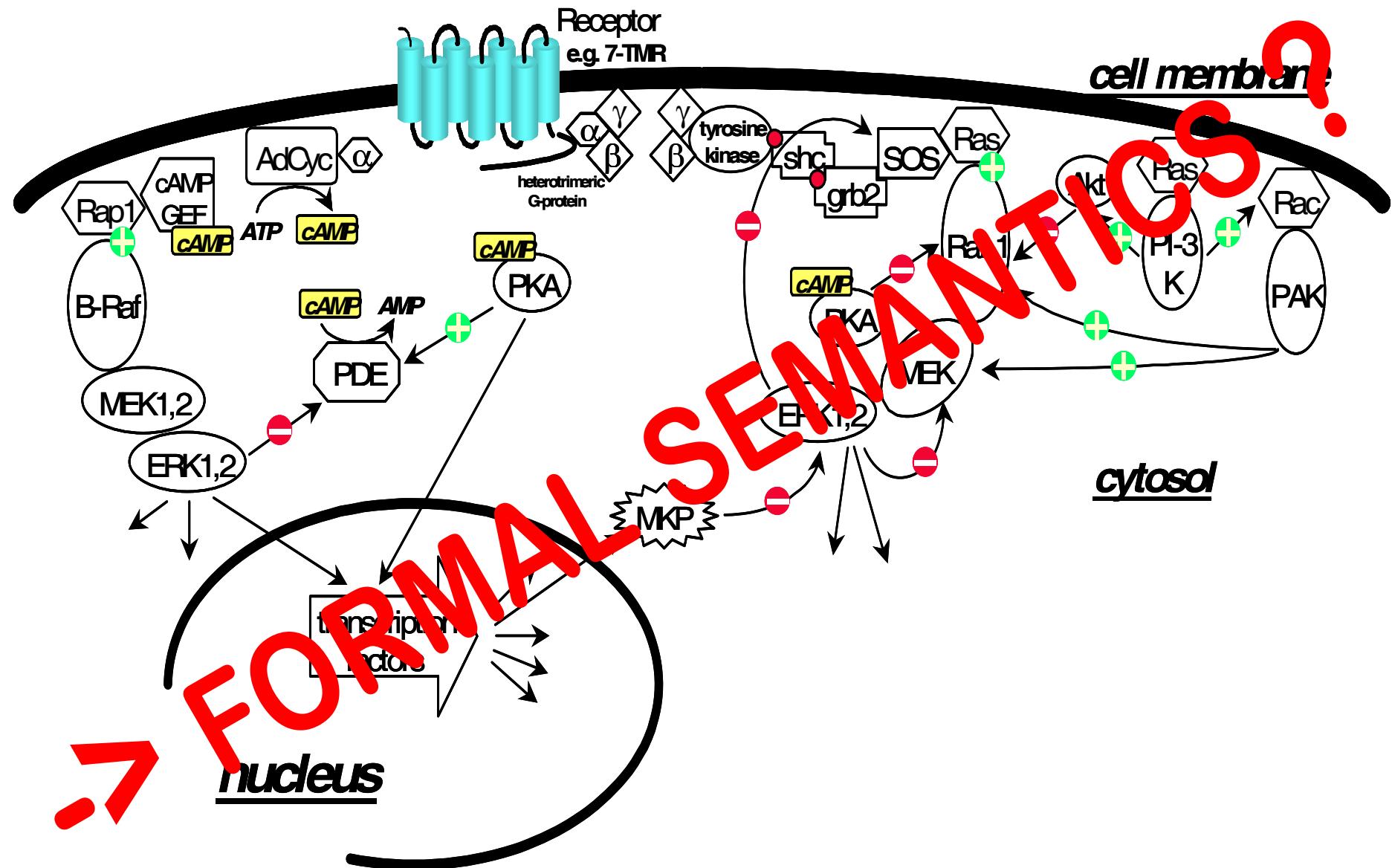


# **WHAT KIND OF MODEL SHOULD BE USED?**

# BIO NETWORK REPRESENTATIONS, Ex1

PN & Systems Biology





## BIO NETWORK REPRESENTATIONS, Ex2

PN & Systems Biology

$$\begin{aligned}
 \frac{d\alpha}{dt} &= -v_1 \\
 \frac{d\text{Ste2}}{dt} &= -v_2 + v_3 - v_5 \\
 \frac{d\text{Ste2}_{\text{active}}}{dt} &= v_2 - v_3 - v_4 \\
 \frac{d\text{Sst2}_{\text{active}}}{dt} &= v_{46} - v_{47} \\
 \frac{dG\alpha\beta\gamma}{dt} &= -v_6 + v_9 \\
 \frac{dG\alpha\text{GTP}}{dt} &= v_6 - v_7 - v_8 \\
 \frac{dG\alpha\text{GDP}}{dt} &= v_7 + v_8 - v_9 \\
 \frac{dG\beta\gamma}{dt} &= v_6 - v_9 - v_{10} + v_{11} + v_{21} + v_{23} + v_{25} + v_{27} + v_{32} \\
 &\quad - v_{42} + v_{43} \\
 \frac{d\text{Ste5}}{dt} &= -v_{12} + v_{13} + v_{17} + v_{21} + v_{23} + v_{25} + v_{27} + v_{32} \\
 \frac{d\text{Ste11}}{dt} &= -v_{12} + v_{13} + v_{17} + v_{21} + v_{23} + v_{25} + v_{27} + v_{32} \\
 \frac{d\text{Ste7}}{dt} &= -v_{14} + v_{15} + v_{17} + v_{21} + v_{23} + v_{25} + v_{27} + v_{32} \\
 \frac{d\text{Fus3}}{dt} &= -v_{14} + v_{15} + v_{17} + v_{21} + v_{23} + v_{25} + v_{27} - v_{29} \\
 &\quad + v_{30} + v_{33} \\
 \frac{d\text{Ste20}}{dt} &= -v_{18} + v_{19} + v_{21} + v_{23} + v_{25} + v_{27} + v_{32}
 \end{aligned}$$

$$\begin{aligned}
 v_1 &= \alpha[t] \cdot \text{Bar1}_{\text{active}}[t] \cdot k_1 \\
 v_2 &= \text{Ste2}[t] \cdot \alpha[t] \cdot k_2 \\
 v_3 &= \text{Ste2}_{\text{active}}[t] \cdot k_3 \\
 v_4 &= \text{Ste2}_{\text{active}}[t] \cdot k_4 \\
 v_5 &= \text{Ste2}[t] \cdot k_5 \\
 v_6 &= \text{Ste2}_{\text{active}}[t] \cdot G\alpha\beta\gamma[t] \cdot k_6 \\
 v_7 &= G\alpha\text{GTP}[t] \cdot k_7 \\
 v_8 &= G\alpha\text{GTP}[t] \cdot \text{Sst2}_{\text{active}}[t] \cdot k_8 \\
 v_9 &= G\alpha\text{GDP}[t] \cdot G\beta\gamma[t] \cdot k_9 \\
 v_{10} &= G\beta\gamma[t] \cdot C[t] \cdot k_{10} \\
 v_{11} &= D[t] \cdot k_{11} \\
 v_{12} &= \text{Ste5}[t] \cdot \text{Ste11}[t] \cdot k_{12} \\
 v_{13} &= A[t] \cdot k_{13} \\
 v_{14} &= \text{Ste7}[t] \cdot \text{Fus3}[t] \cdot k_{14} \\
 v_{15} &= B[t] \cdot k_{15} \\
 v_{16} &= A[t] \cdot B[t] \cdot k_{16} \\
 v_{17} &= C[t] \cdot k_{17} \\
 v_{18} &= D[t] \cdot \text{Ste20}[t] \cdot k_{18}
 \end{aligned}$$

## BIO NETWORK REPRESENTATIONS, Ex2

PN & Systems Biology

$$\begin{aligned}
 \frac{d\alpha}{dt} &= -v_1 \\
 \frac{d\text{Ste2}}{dt} &= -v_2 + v_3 - v_5 \\
 \frac{d\text{Ste2}_{\text{active}}}{dt} &= v_2 - v_3 - v_4 \\
 \frac{d\text{Sst2}_{\text{active}}}{dt} &= v_{46} - v_{47} \\
 \frac{dG\alpha\beta\gamma}{dt} &= -v_6 + v_9 \\
 \frac{dG\alpha\text{GTP}}{dt} &= v_6 - v_7 - v_8 \\
 \frac{dG\alpha\text{GDP}}{dt} &= v_7 + v_8 - v_9 \\
 \frac{dG\beta\gamma}{dt} &= v_6 - v_9 - v_{10} + v_{11} + v_{21} + v_{23} + v_{25} + v_{26} + v_{32} \\
 &\quad - v_{42} + v_{43} \\
 \frac{d\text{Ste5}}{dt} &= -v_{12} + v_{13} + v_{17} + v_{21} + v_{23} + v_{25} + v_{27} + v_{32} \\
 \frac{d\text{Ste11}}{dt} &= -v_{12} + v_{13} + v_{17} + v_{21} + v_{23} + v_{25} + v_{27} + v_{32} \\
 \frac{d\text{Ste7}}{dt} &= -v_{14} + v_{15} + v_{17} + v_{21} + v_{23} + v_{25} + v_{27} + v_{32} \\
 \frac{d\text{Fus3}}{dt} &= -v_{14} + v_{15} + v_{17} + v_{21} + v_{23} + v_{25} + v_{27} - v_{29} \\
 &\quad + v_{30} + v_{33} \\
 \frac{d\text{Ste20}}{dt} &= -v_{18} + v_{19} + v_{21} + v_{23} + v_{25} + v_{27} + v_{32}
 \end{aligned}$$

$$\begin{aligned}
 v_1 &= \alpha[t] \cdot \text{Bar1}_{\text{active}}[t] \cdot k_1 \\
 v_2 &= \text{Ste2}[t] \cdot \alpha[t] \cdot k_2 \\
 v_3 &= \text{Ste2}_{\text{active}}[t] \cdot k_3 \\
 v_4 &= \text{Ste2}_{\text{active}}[t] \cdot k_4 \\
 v_5 &= \text{Ste2}[t] \cdot k_5 \\
 v_6 &= \text{Sst2}_{\text{active}}[t] \cdot G\alpha\beta\gamma[t] \cdot k_6 \\
 v_7 &= G\alpha\text{GTP}[t] \cdot k_7 \\
 v_8 &= G\alpha\text{GTP}[t] \cdot \text{Sst2}_{\text{active}}[t] \cdot k_8 \\
 v_9 &= G\alpha\text{GDP}[t] \cdot G\beta\gamma[t] \cdot k_9 \\
 v_{10} &= G\beta\gamma[t] \cdot C[t] \cdot k_{10} \\
 v_{11} &= D[t] \cdot k_{11} \\
 v_{12} &= \text{Ste5}[t] \cdot \text{Ste11}[t] \cdot k_{12} \\
 v_{13} &= A[t] \cdot k_{13} \\
 v_{14} &= \text{Ste7}[t] \cdot \text{Fus3}[t] \cdot k_{14} \\
 v_{15} &= B[t] \cdot k_{15} \\
 v_{16} &= A[t] \cdot B[t] \cdot k_{16} \\
 v_{17} &= C[t] \cdot k_{17} \\
 v_{18} &= D[t] \cdot \text{Ste20}[t] \cdot k_{18}
 \end{aligned}$$

**READABILITY?**

- knowledge → **PROBLEM 1**
  - > *uncertain*
  - > *growing, changing*
  - > *distributed over independent data bases, papers, journals, . . .*
- various, mostly ambiguous representations → **PROBLEM 2**
  - > *verbose descriptions*
  - > *diverse graphical representations*
  - > *contradictory and / or fuzzy statements*
- network structures → **PROBLEM 3**
  - > *tend to grow fast*
  - > *dense, apparently unstructured*
  - > *hard to read*

knowledge

- > *uncertain*
- > *growing, changing*
- > *distributed over independent data bases, papers, journals ...*

-> **PROBLEM 1**

various, mostly ambiguous representations

- > *verbose descriptions*
- > *diverse graphical representations*
- > *contradictory and / or fuzzy statements*

-> **PROBLEM 2**

network structures

- > *tend to grow fast*
- > *dense, apparently unstructured*
- > *hard to read*

-> **PROBLEM 3**

- **readable**

- > *fault avoidance*
- > *informal = cartoon-like representations ?*

- **analysable**

- > *formal = mathematical representations*

- **executable**

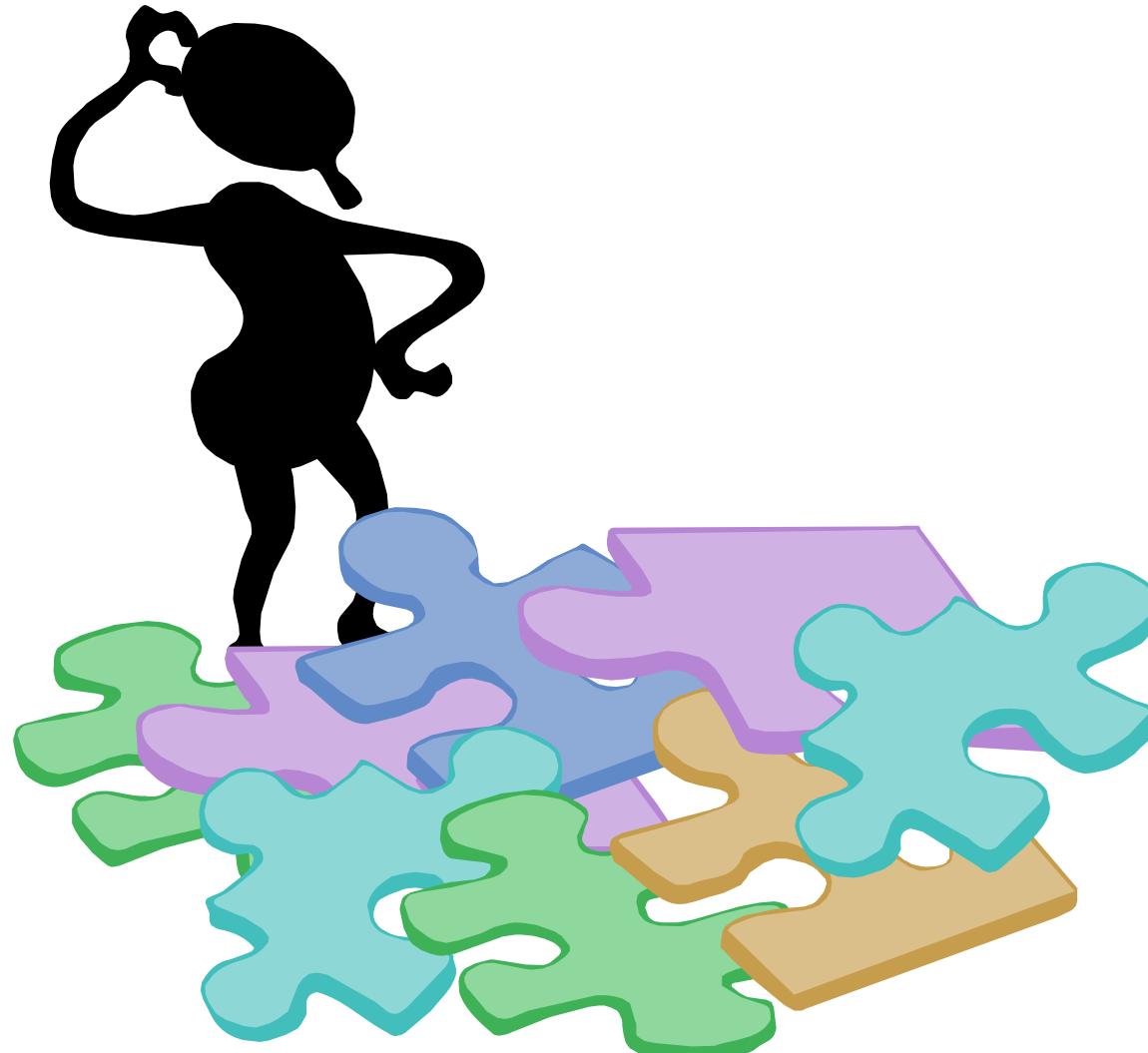
- > *to experience the model*

- **unifying power**

- > *high-level description for various analysis approaches*

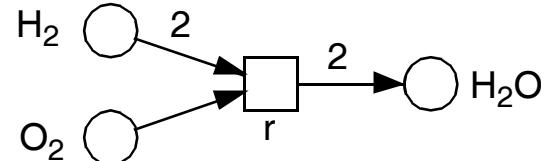
# WHAT KIND OF MODEL TO CHOSE?

PN & Systems Biology

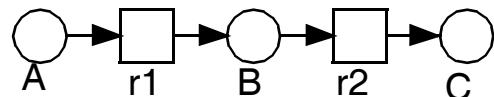


**... ARE  
NETWORKS OF  
(BIO-) CHEMICAL REACTIONS**

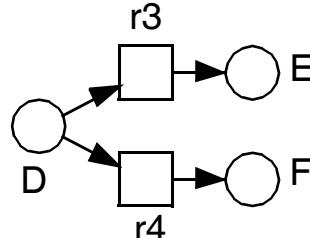
- ❑ bipartite - species & reactions



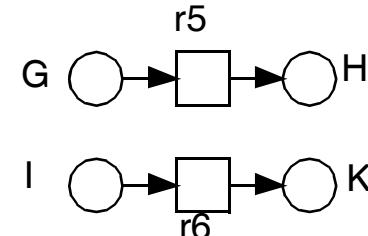
- ❑ reactions - sequential,



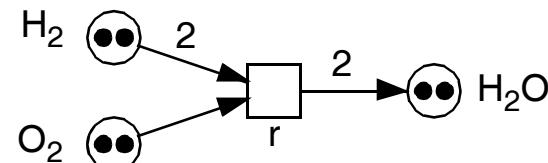
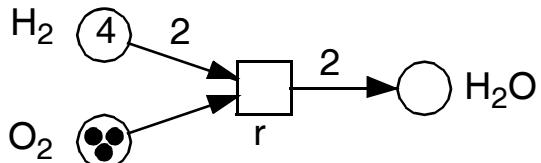
- ❑ reactions - alternative,

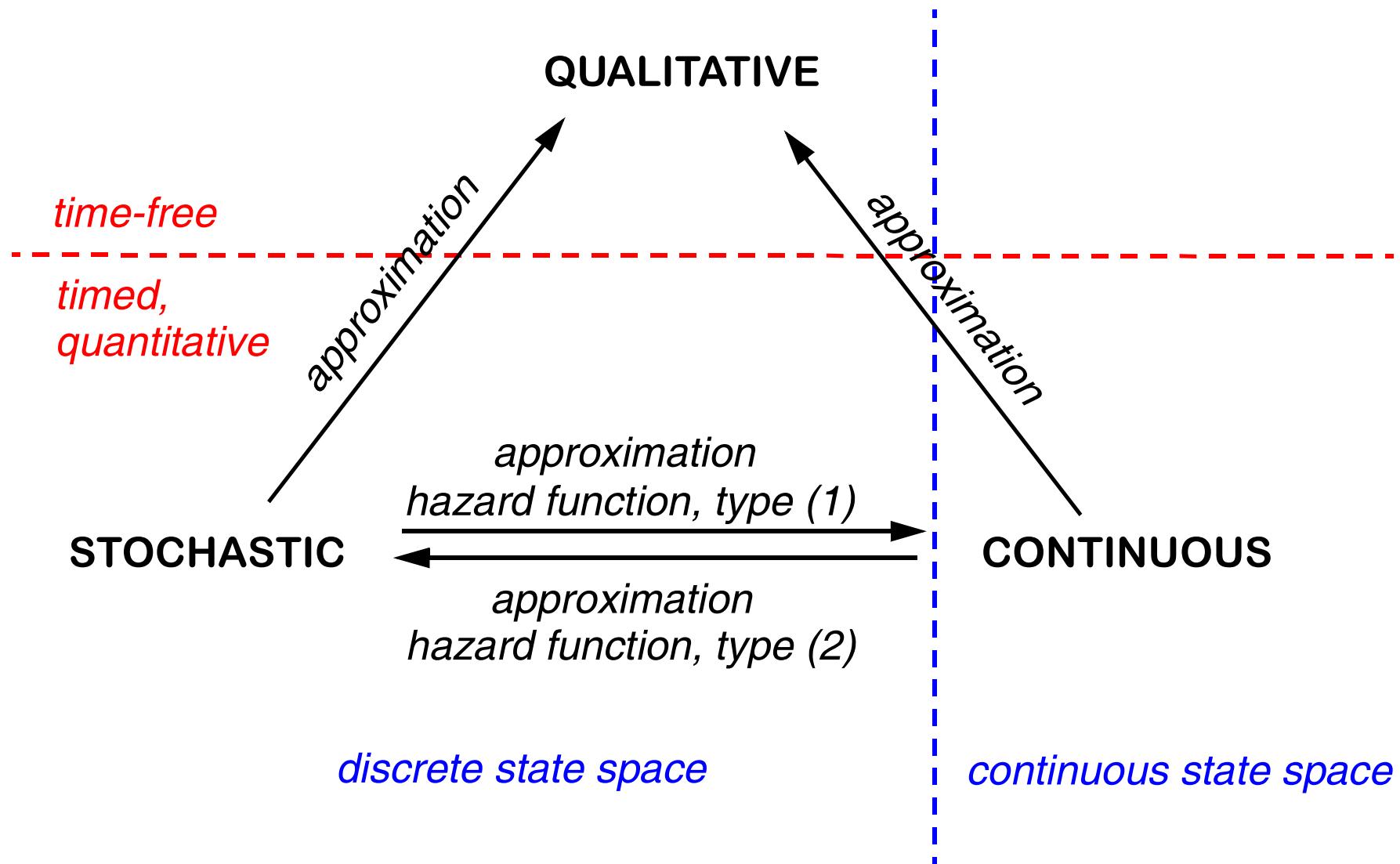


- ❑ reactions - concurrent

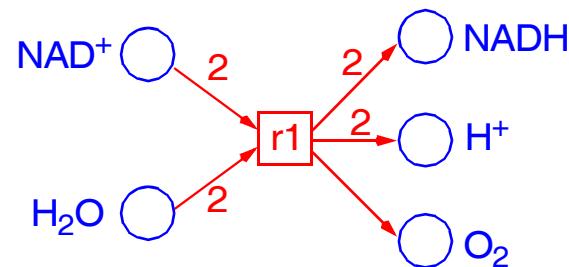
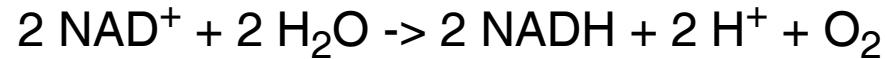


- ❑ behaviour - stochastic

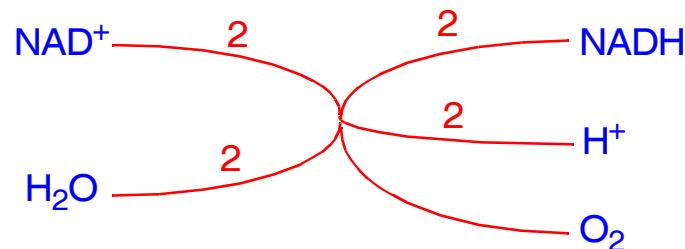




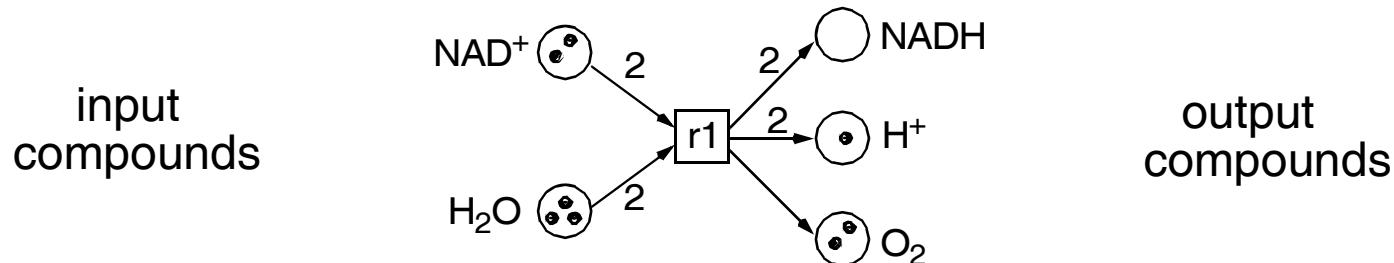
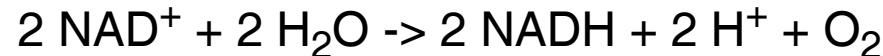
# PETRI NETS - AN INFORMAL CRASH COURSE



hyper-arcs



atomic actions      -> transitions      -> chemical reactions



local conditions      -> places      -> chemical compounds

multiplicities      -> arc weights      -> stoichiometric relations

condition's state      -> token(s)      -> available amount (e.g. mol)

system state      -> marking      -> compounds distribution

**PN = (P, T, F, m<sub>0</sub>)**,    F: ((P × T) ∪ (T × P)) → N<sub>0</sub>, m<sub>0</sub>: P → N<sub>0</sub>

□ **an action may happen, if** -> prerequisite

-> *all preconditions are fulfilled  
(corresponding to the arc weights)*

□ **if an action happens, then** -> firing behaviour

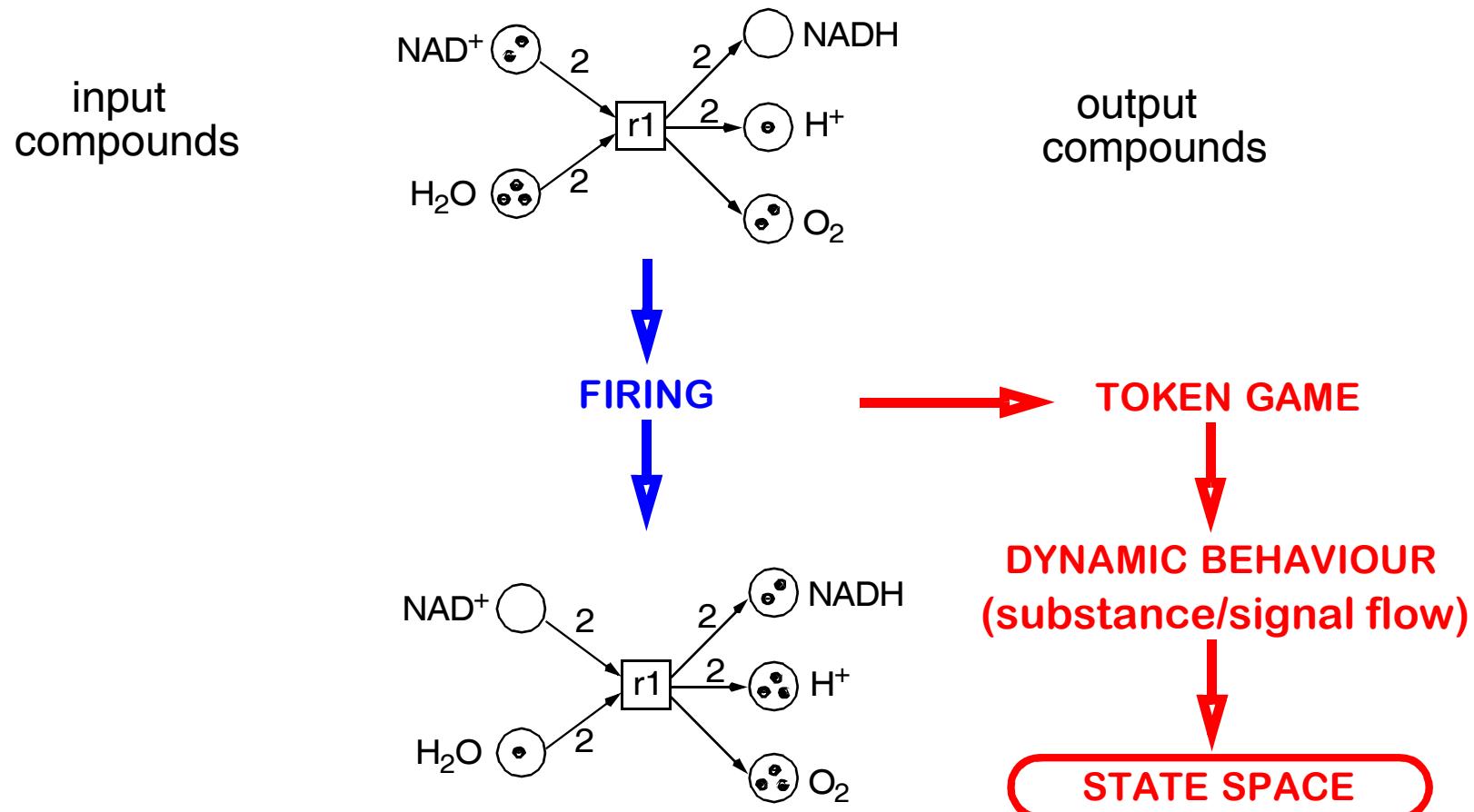
-> *tokens are removed from all preconditions  
(corresponding to the arc weights), and*  
-> *tokens are added to all postconditions  
(corresponding to the arc weights)*

□ **action happens (firing of a transition)** -> model assumptions

-> *atomic*  
-> *no time consumption*

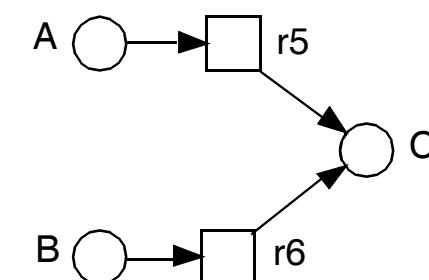
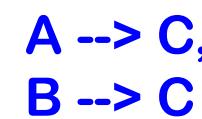
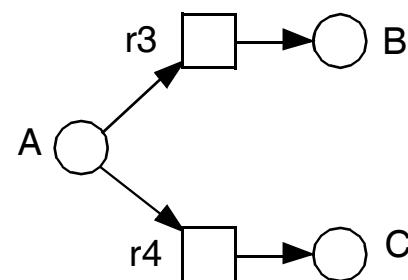
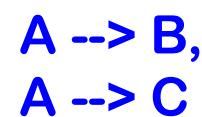
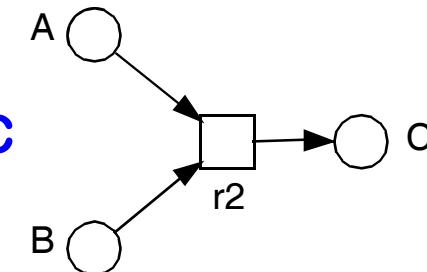
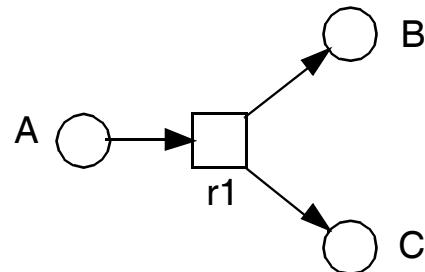
**-> TIME-FREE MODEL, WHICH CONSIDERS ALL POSSIBLE TIMING BEHAVIOUR**

□ atomic actions      -> transitions      -> chemical reactions

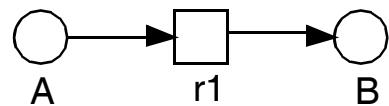


## TYPICAL BASIC STRUCTURES I

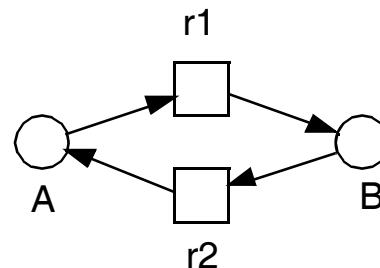
PN & Systems Biology



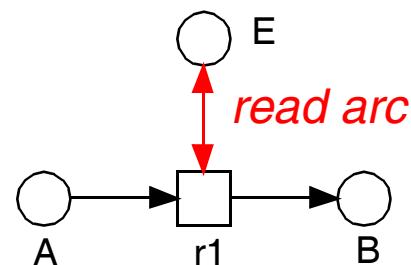
$A \rightarrow B$



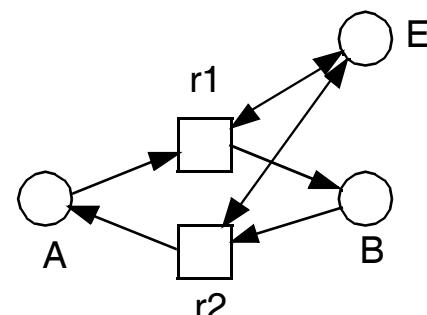
$A \leftrightarrow B$



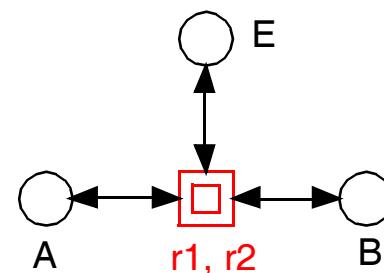
$A \xrightarrow{E} B$



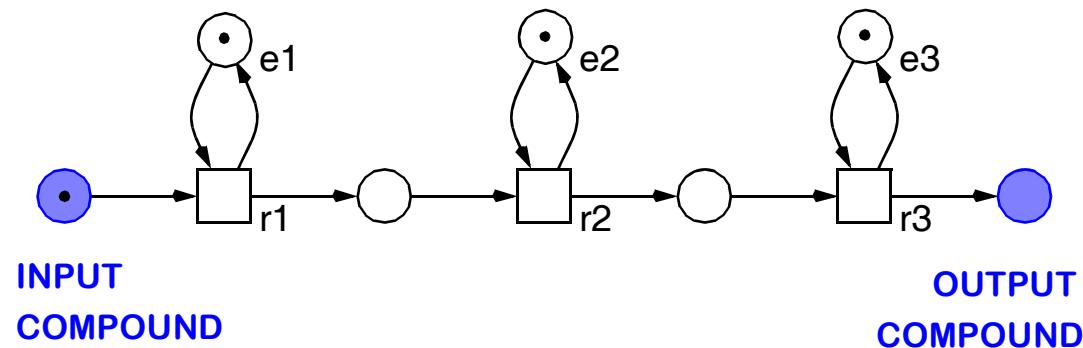
$A \xleftrightarrow{E} B$



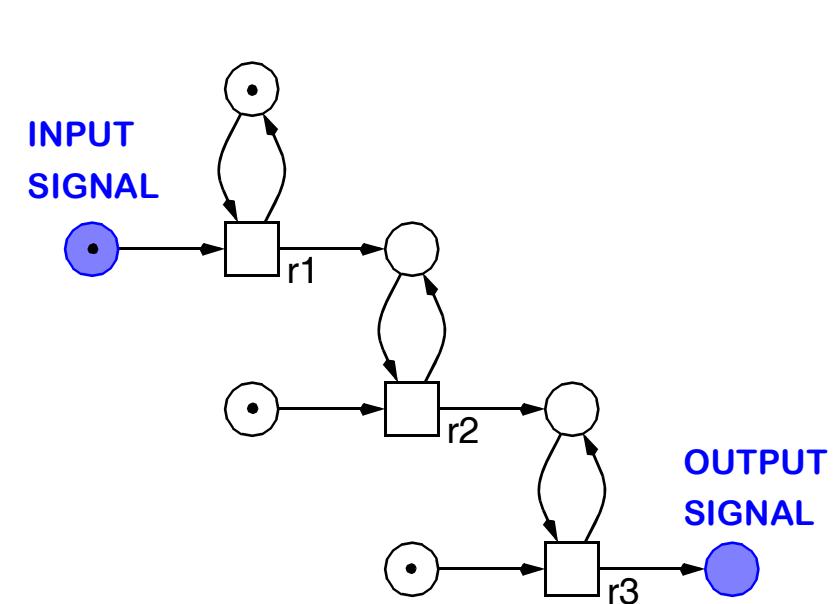
*macro transition*



- metabolic networks  
-> *substance flows*

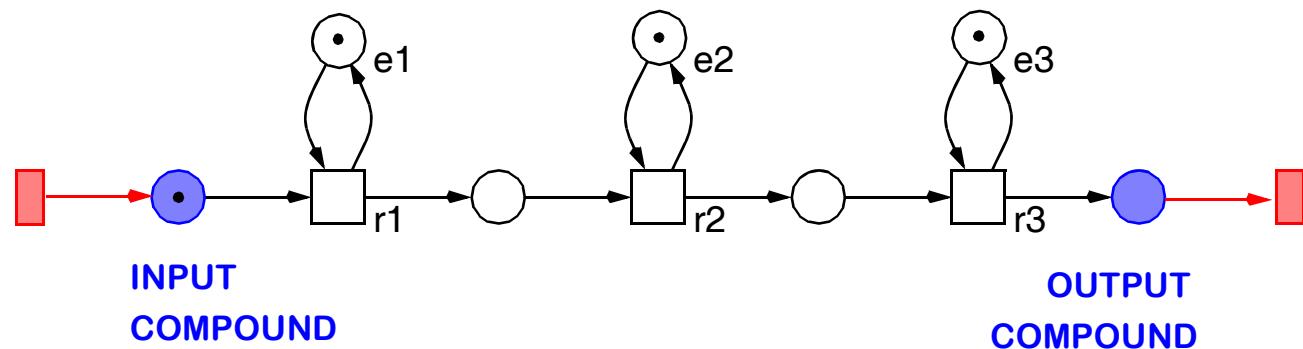


- signal transduction networks  
-> *signal flows*



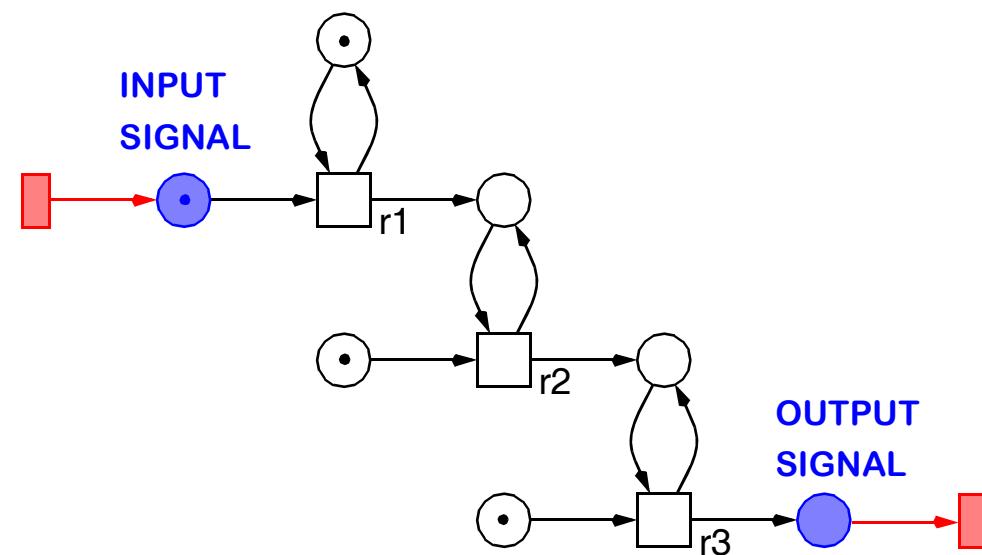
□ metabolic networks

-> *substance flows*



□ signal transduction networks

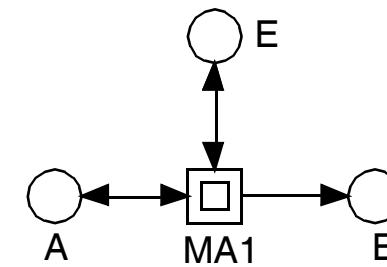
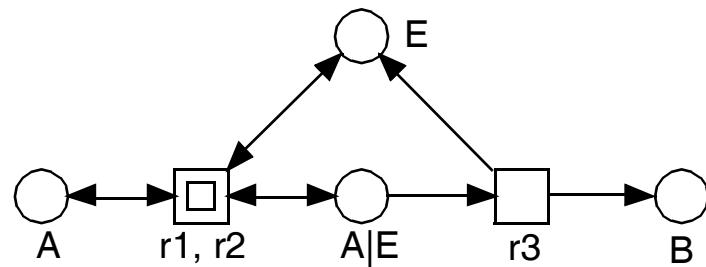
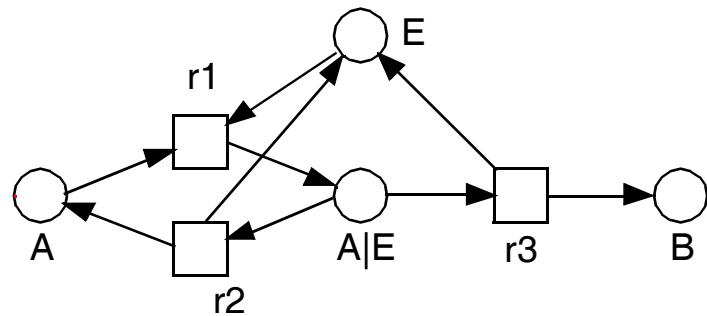
-> *signal flows*



-> OPEN / CLOSED SYSTEMS

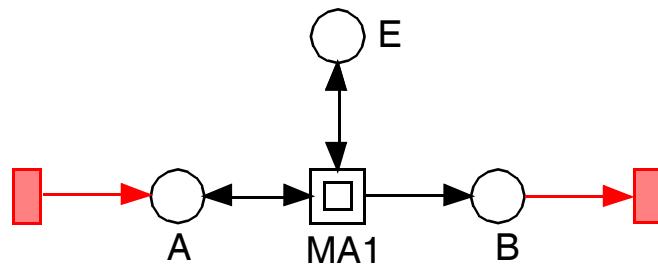
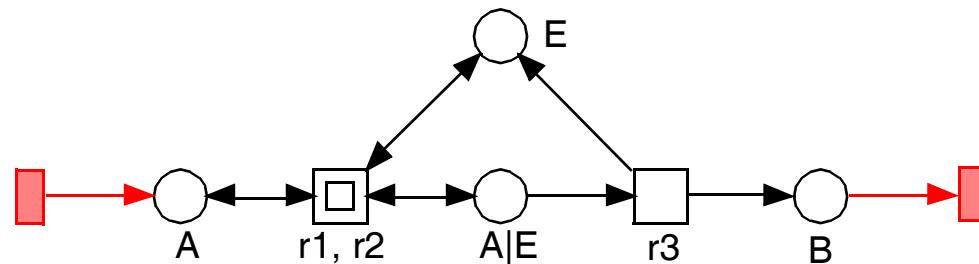
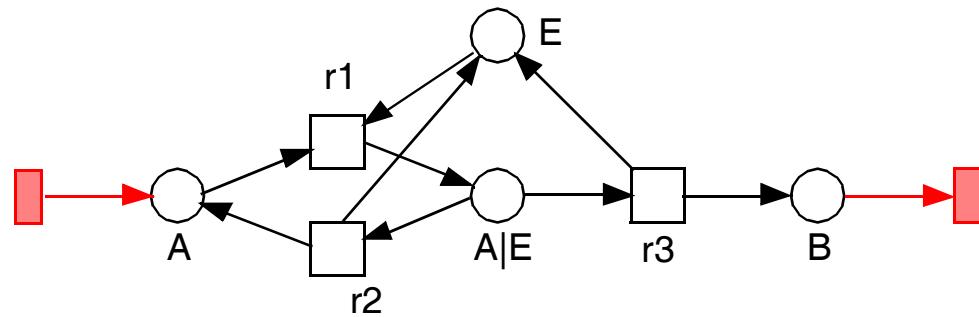


*enzymatic reaction,  
mass-action approach 1*

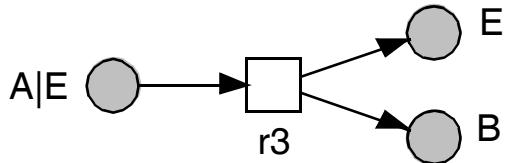
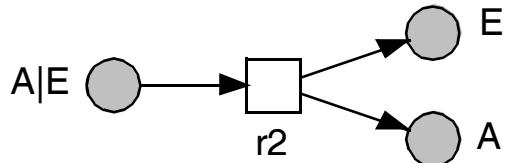
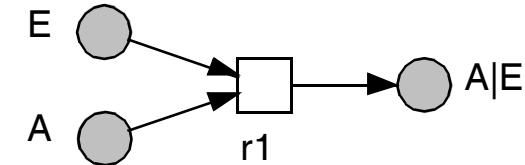




*enzymatic reaction,  
mass-action approach 1*

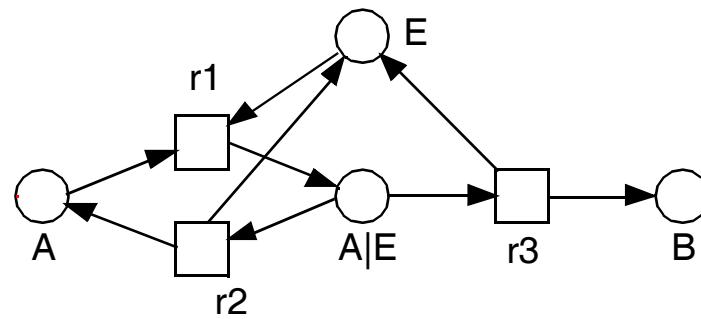


*reaction-centred view*

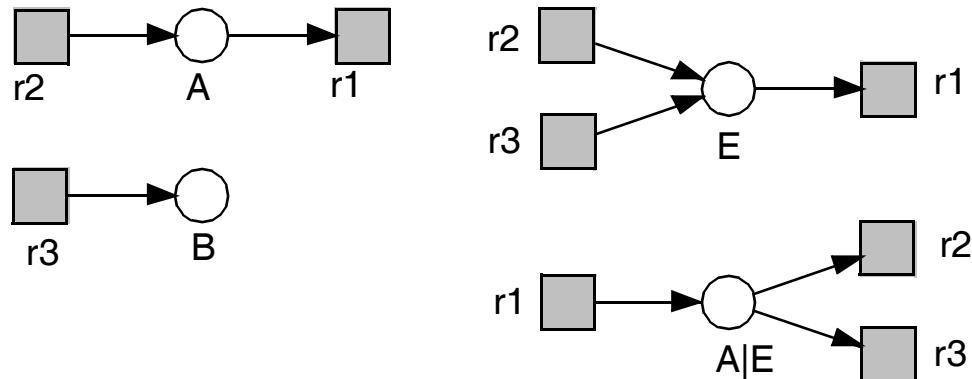


*logical nodes  
(fusion nodes)*

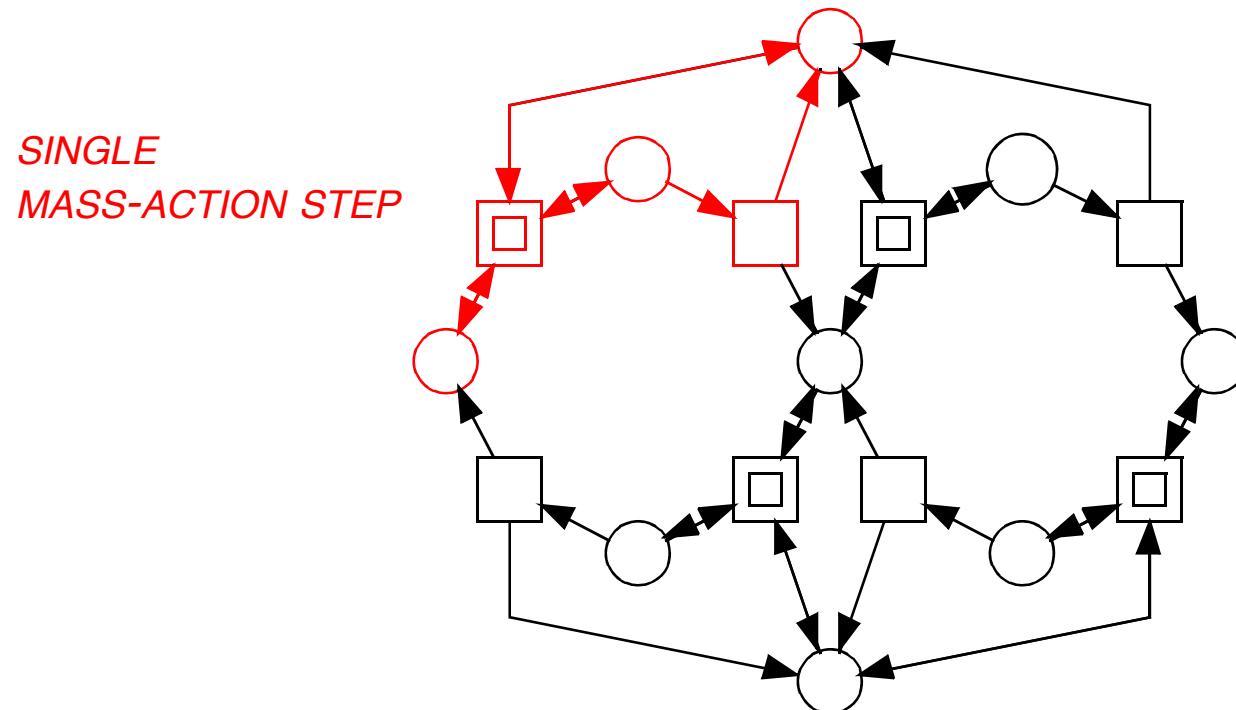
*process-oriented view*



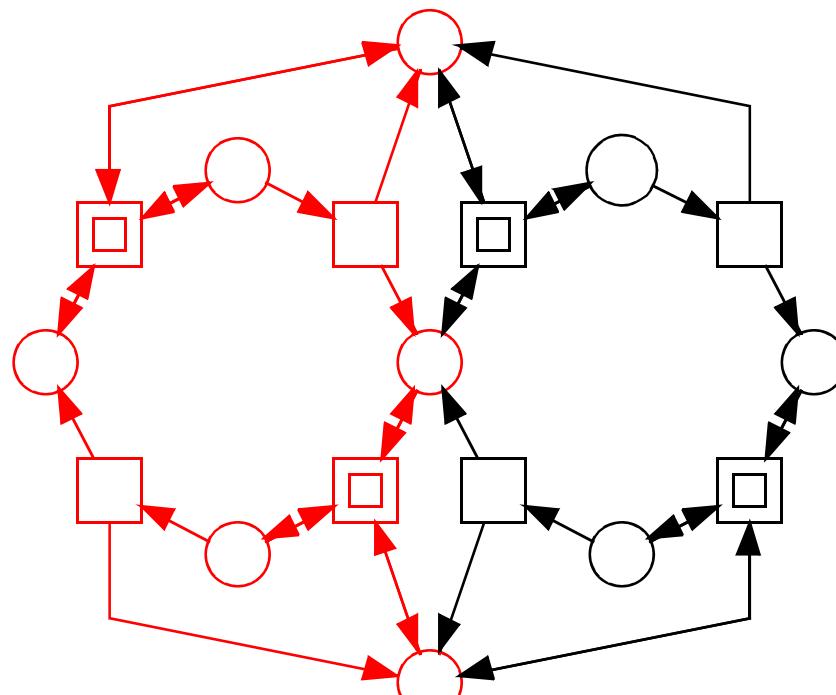
*species-centred view*



## DOUBLE PHOSPHOYLATION/DEPHOSPHORYLATION

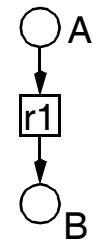


## DOUBLE PHOSPHOYLATION / DEPHOSPHORYLATION



*SINGLE  
PHOSPHOYLATION / DEPHOSPHORYLATION*

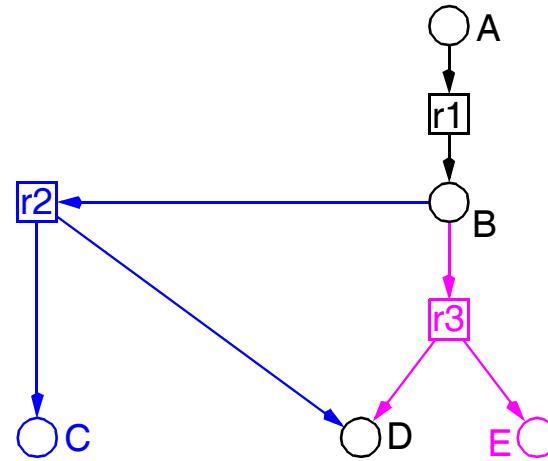
r1: A → B



r1: A → B

r2: B → C + D

r3: B → D + E



-> *alternative reactions*

r1: A → B

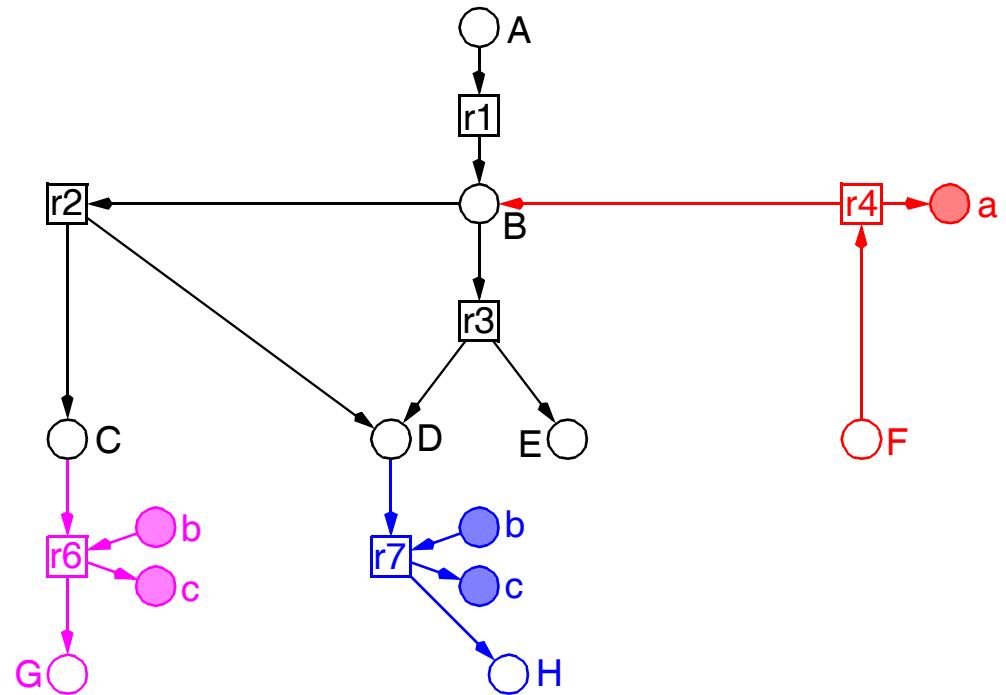
r2: B → C + D

r3: B → D + E

r4: F → B + a

r6: C + b → G + c

r7: D + b → H + c



-> concurrent reactions

r1: A → B

r2: B → C + D

r3: B → D + E

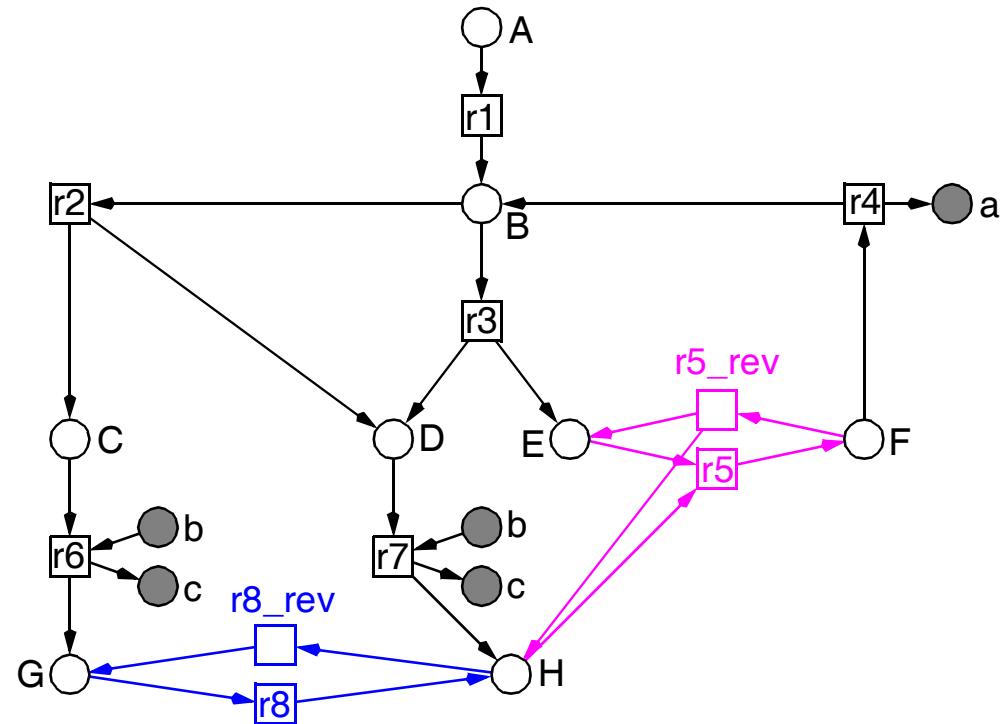
r4: F → B + a

r5: E + H  $\leftrightarrow$  F

r6: C + b → G + c

r7: D + b → H + c

r8: H  $\leftrightarrow$  G



-> reversible reactions

r1: A → B

r2: B → C + D

r3: B → D + E

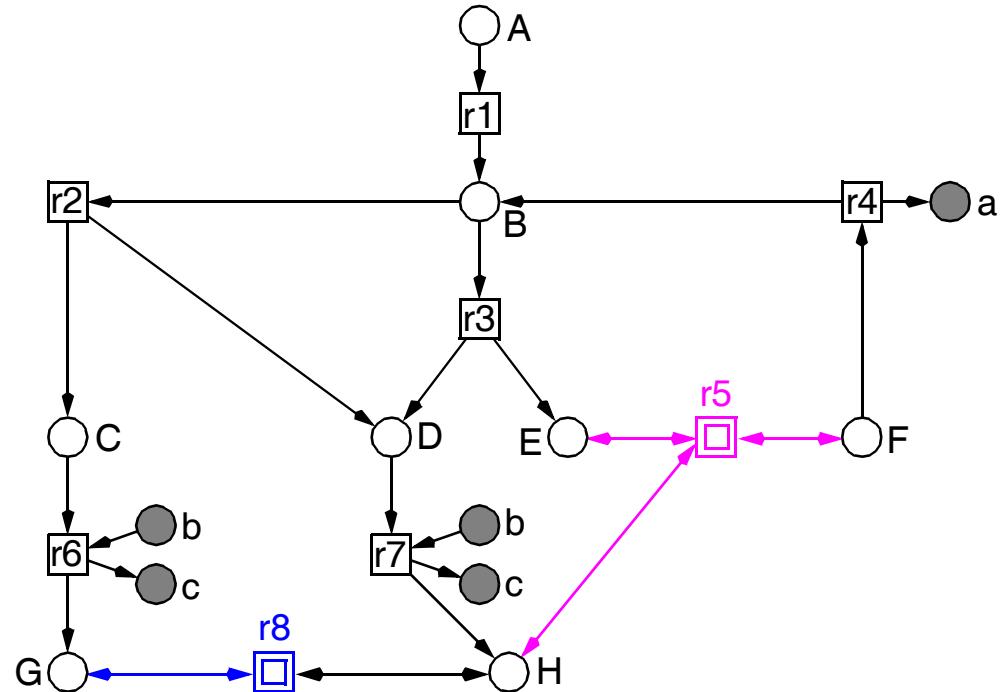
r4: F → B + a

r5: E + H <-> F

r6: C + b → G + c

r7: D + b → H + c

r8: H <-> G



-> reversible reactions  
- hierarchical nodes

r1: A → B

r2: B → C + D

r3: B → D + E

r4: F → B + a

r5: E + H <-> F

r6: C + b → G + c

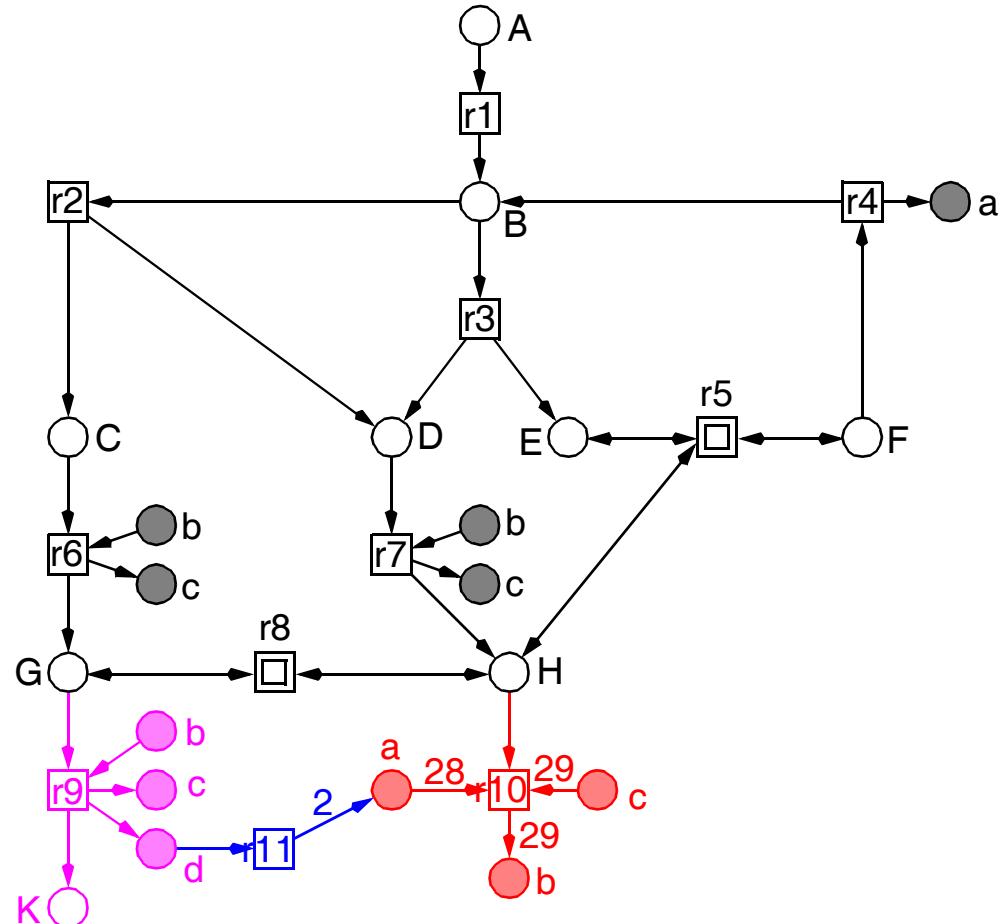
r7: D + b → H + c

r8: H <-> G

r9: G + b → K + c + d

r10: H + 28a + 29c → 29b

r11: d → 2a



r1: A → B

r2: B → C + D

r3: B → D + E

r4: F → B + a

r5: E + H <-> F

r6: C + b → G + c

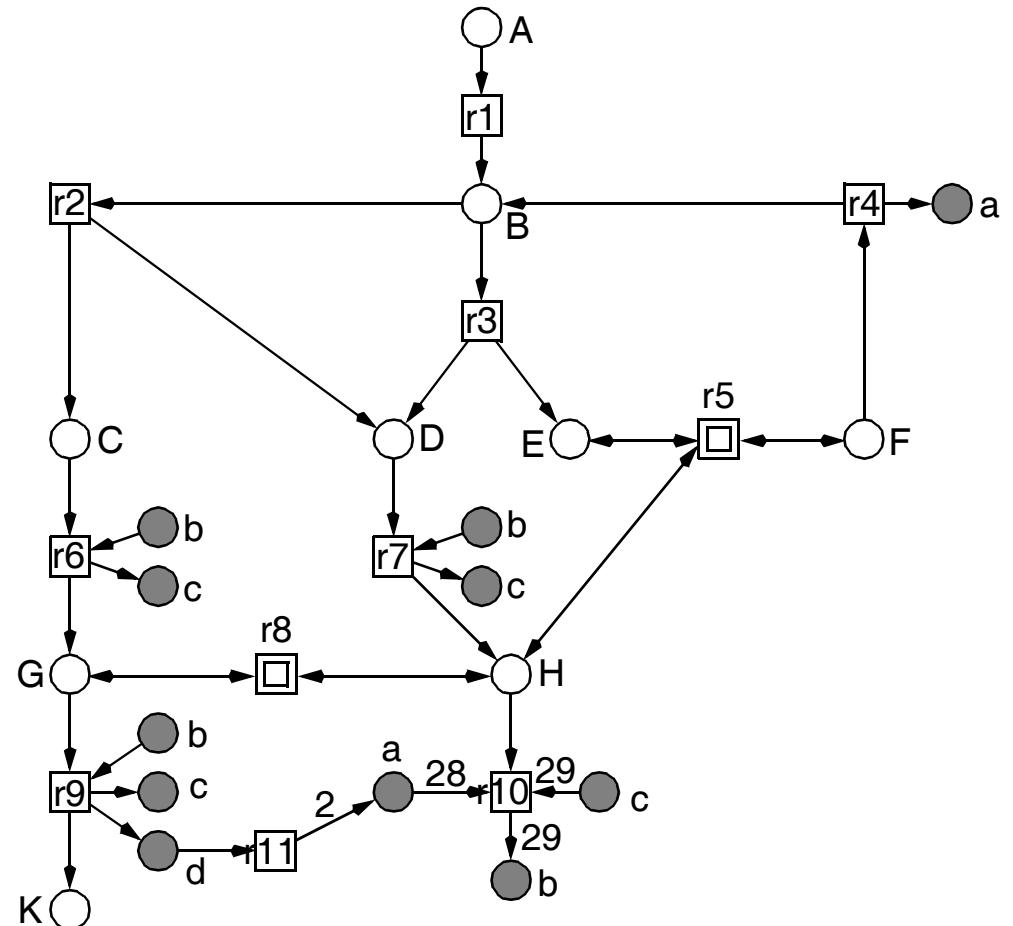
r7: D + b → H + c

r8: H <-> G

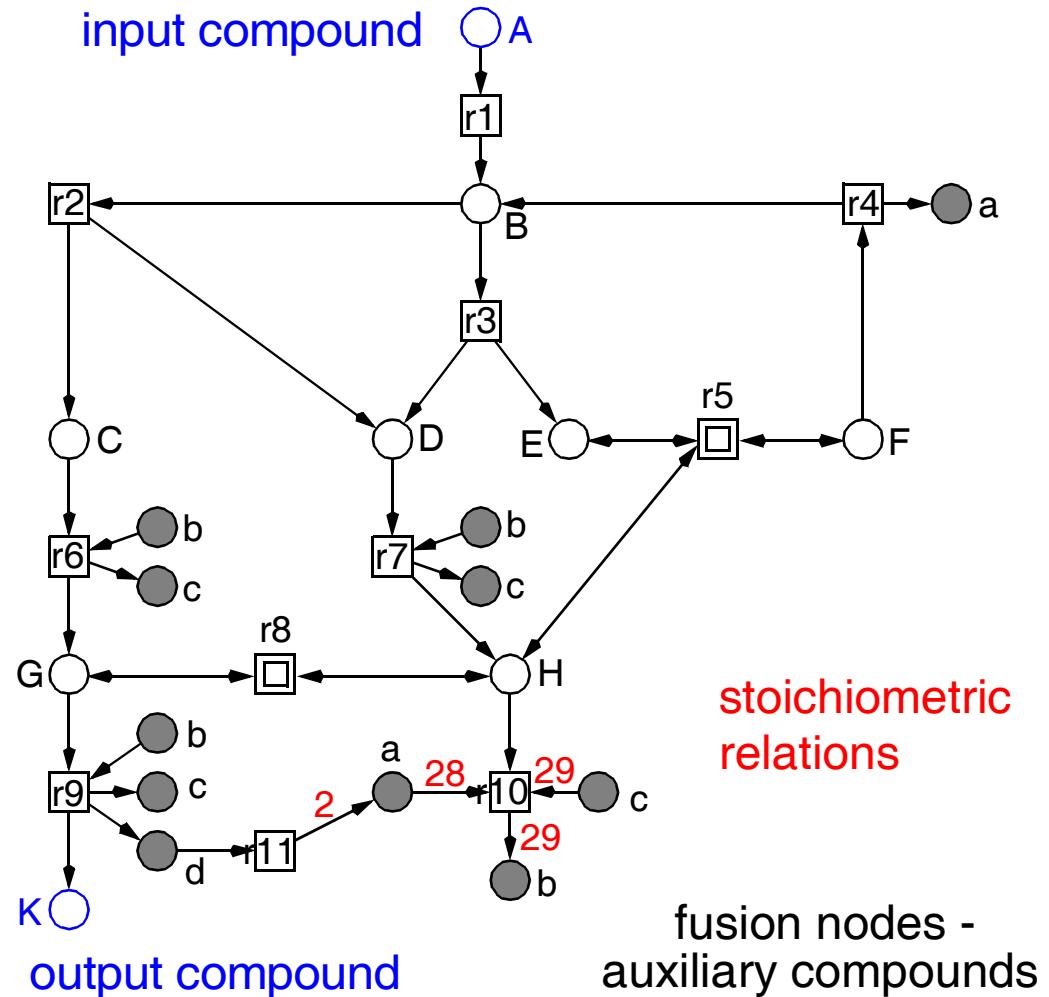
r9: G + b → K + c + d

r10: H + 28a + 29c → 29b

r11: d → 2a



- r1: A → B
- r2: B → C + D
- r3: B → D + E
- r4: F → B + a
- r5: E + H ⇌ F
- r6: C + b → G + c
- r7: D + b → H + c
- r8: H ⇌ G
- r9: G + b → K + c + d
- r10: H + 28a + 29c → 29b
- r11: d → 2a



r1: A → B

r2: B → C + D

r3: B → D + E

r4: F → B + a

r5: E + H <-> F

r6: C + b → G + c

r7: D + b → H + c

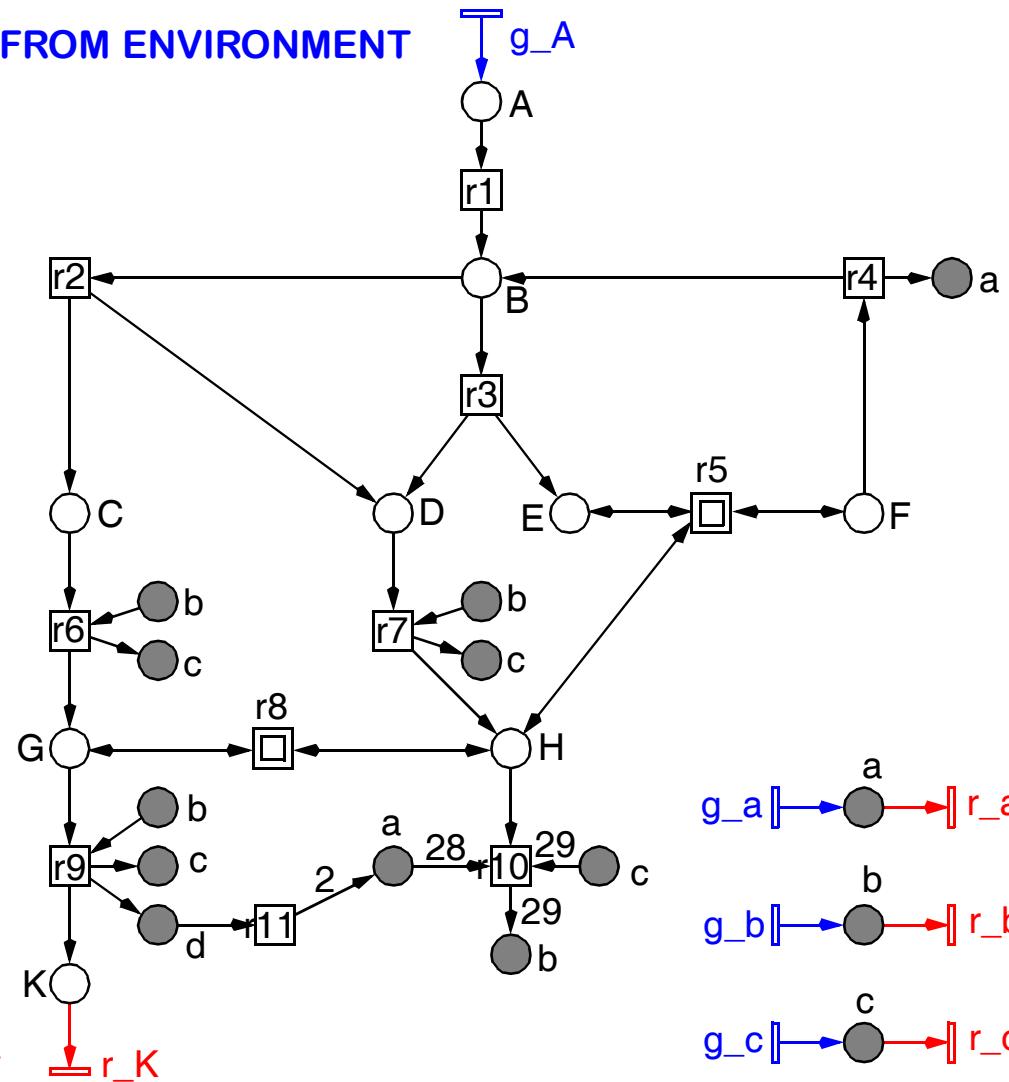
r8: H <-> G

r9: G + b → K + c + d

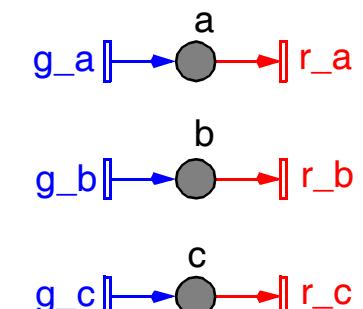
r10: H + 28a + 29c → 29b

r11: d → 2a

## INPUT FROM ENVIRONMENT



## OUTPUT TO ENVIRONMENT



## □ METABOLIC NETWORKS

... SIGNAL TRANSDUCTION NETWORKS

... GENE REGULATORY NETWORKS

## □ transitions

- > *(reversible, stoichiometric) chemical reactions,*
- > *enzyme-catalysed conversions 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 number)*

- biochemical networks

- > *networks of (abstract) chemical reactions*

- biochemically interpreted Petri net

- > *partial order sequences of chemical reactions (= elementary actions)  
transforming input into output compounds / signals  
[ respecting the given stoichiometric relations, if any ]*

- > *set of all pathways  
from the input to the output compounds / signals  
[ respecting the stoichiometric relations, if any ]*

- pathway

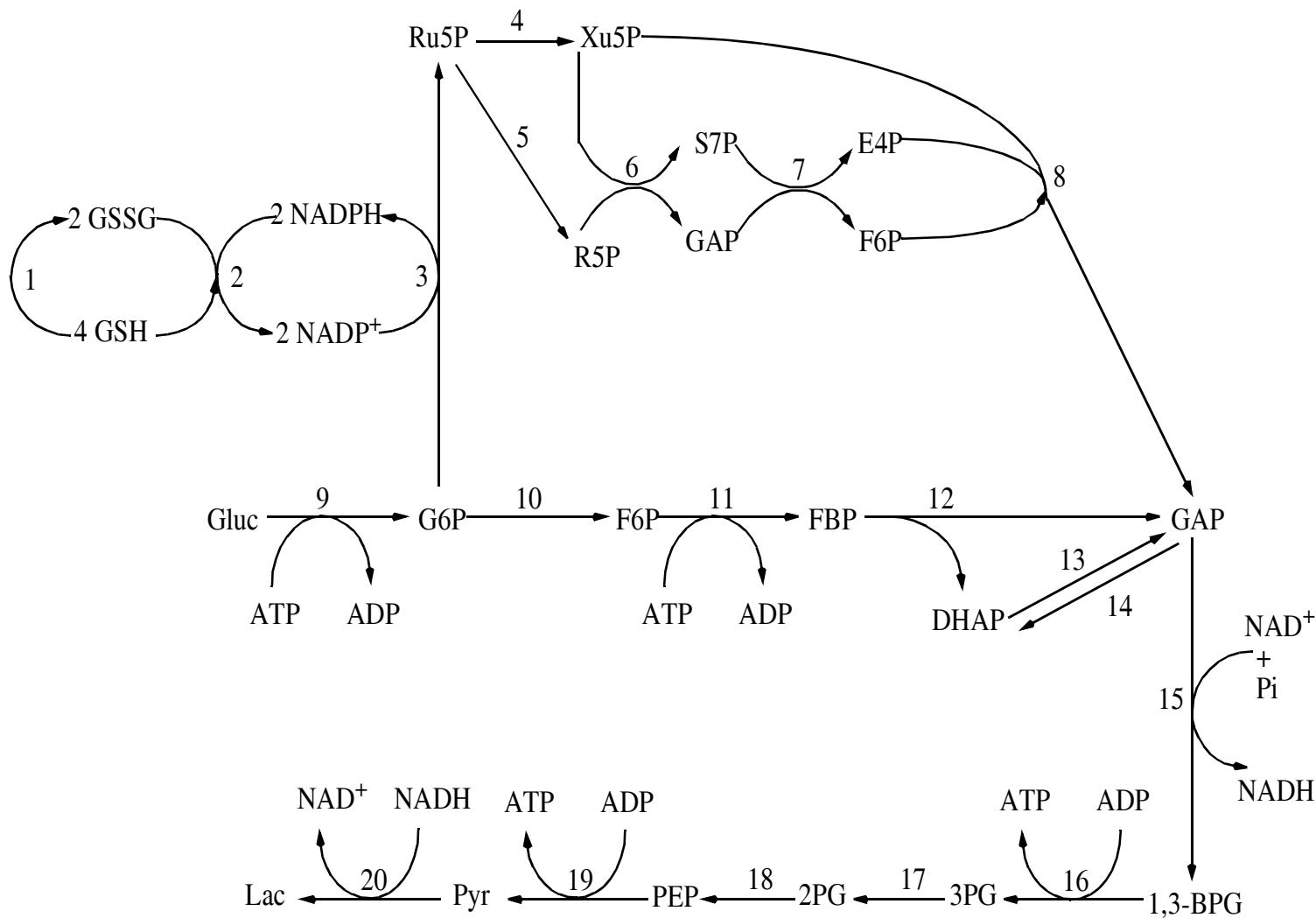
- > *self-contained partial order sequence of elementary (re-) actions*

# BIO PETRI NETS - SOME EXAMPLES

# Ex1 - Glycolysis and Pentose Phosphate Pathway

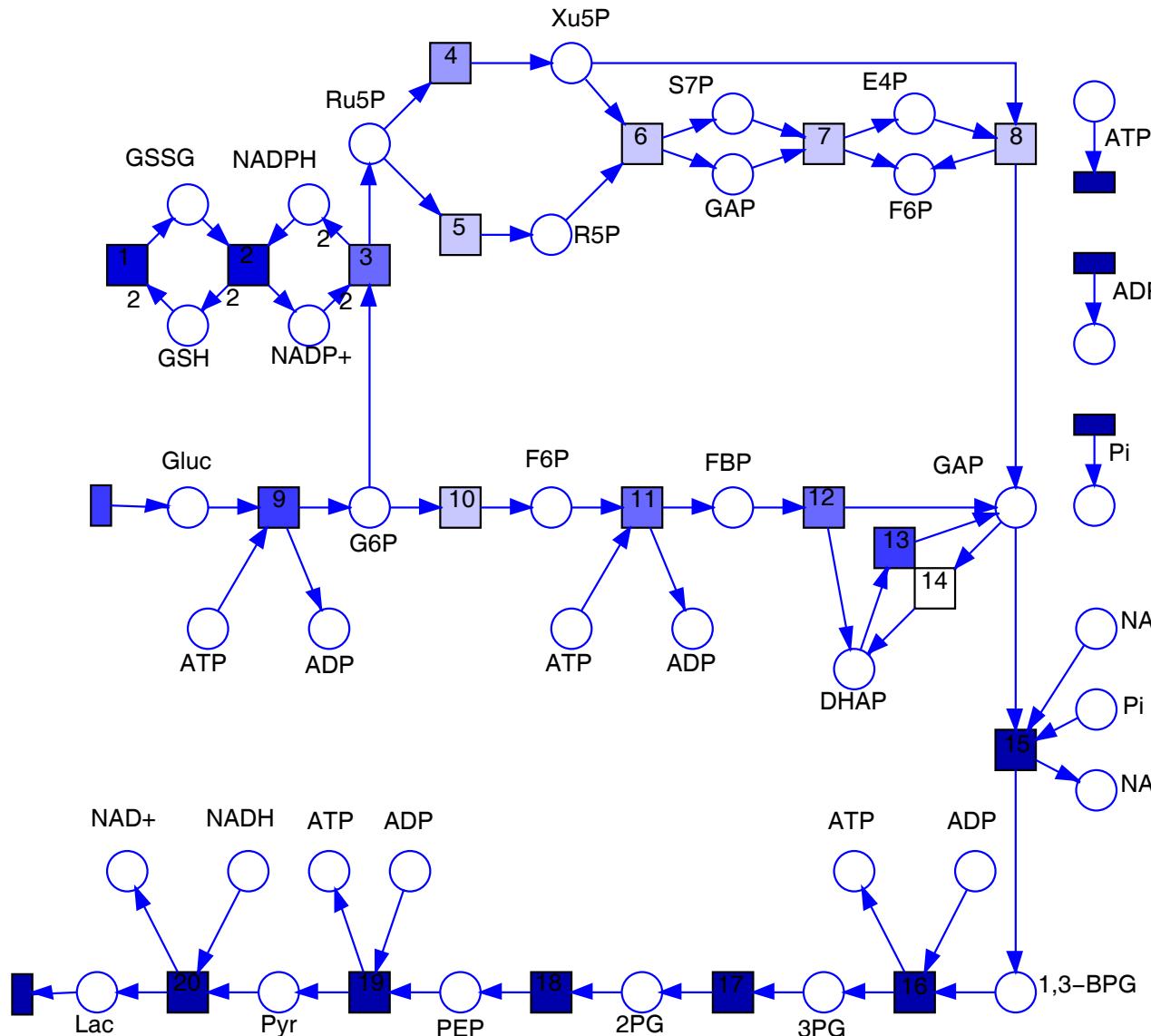
PN & Systems Biology

[Reddy 1993]



# Ex1 - Glycolysis and Pentose Phosphate Pathway

PN & Systems Biology

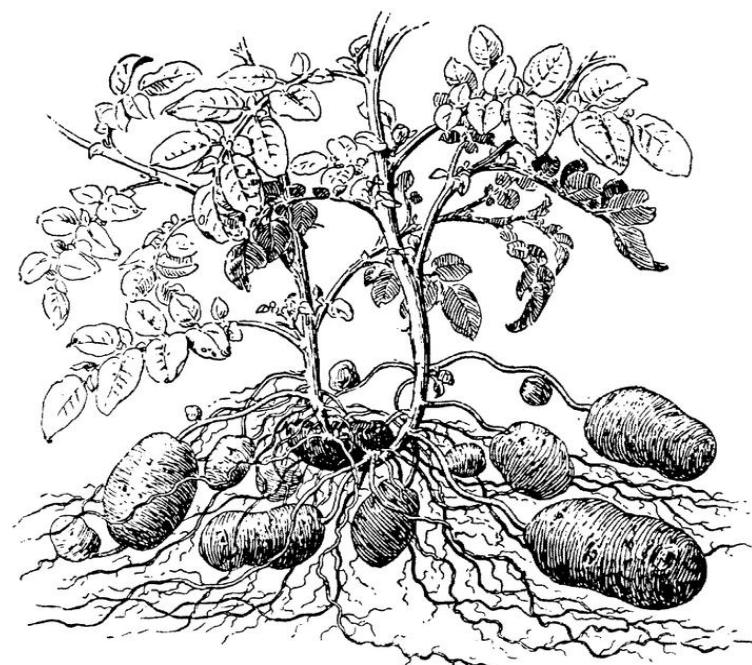


[Reddy 1993]

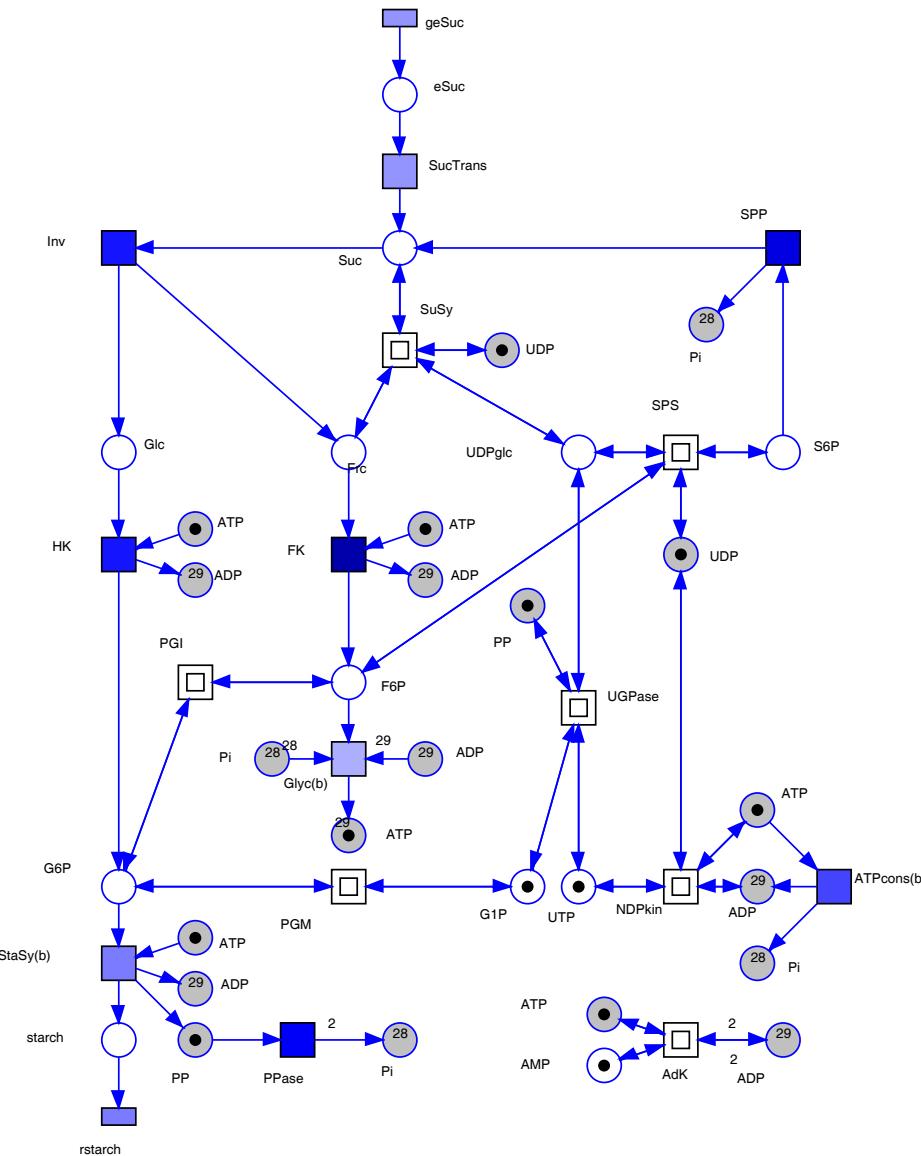
[Koch,  
HEINER 2008]

## Ex2 - Carbon Metabolism in Potato Tuber

PN & Systems Biology

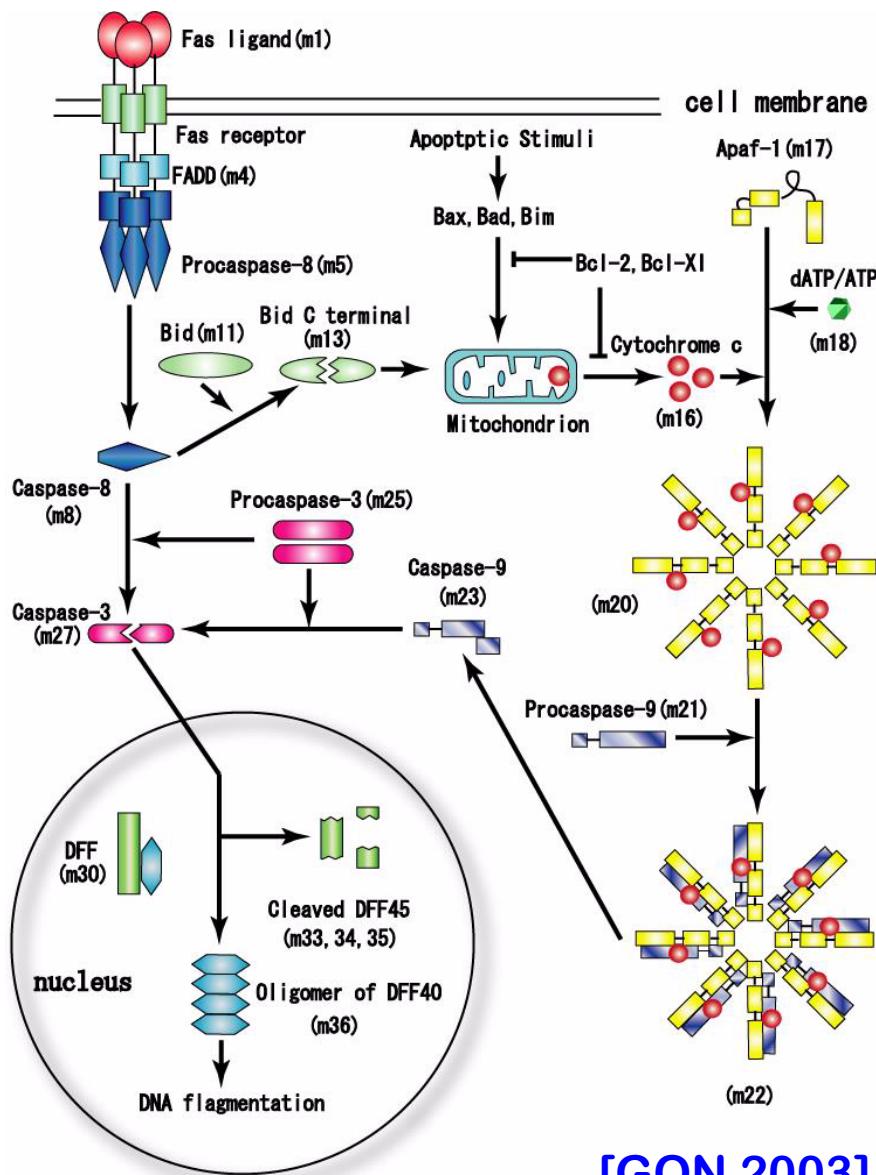


[KOCHE; JUNKER; HEINER 2005]

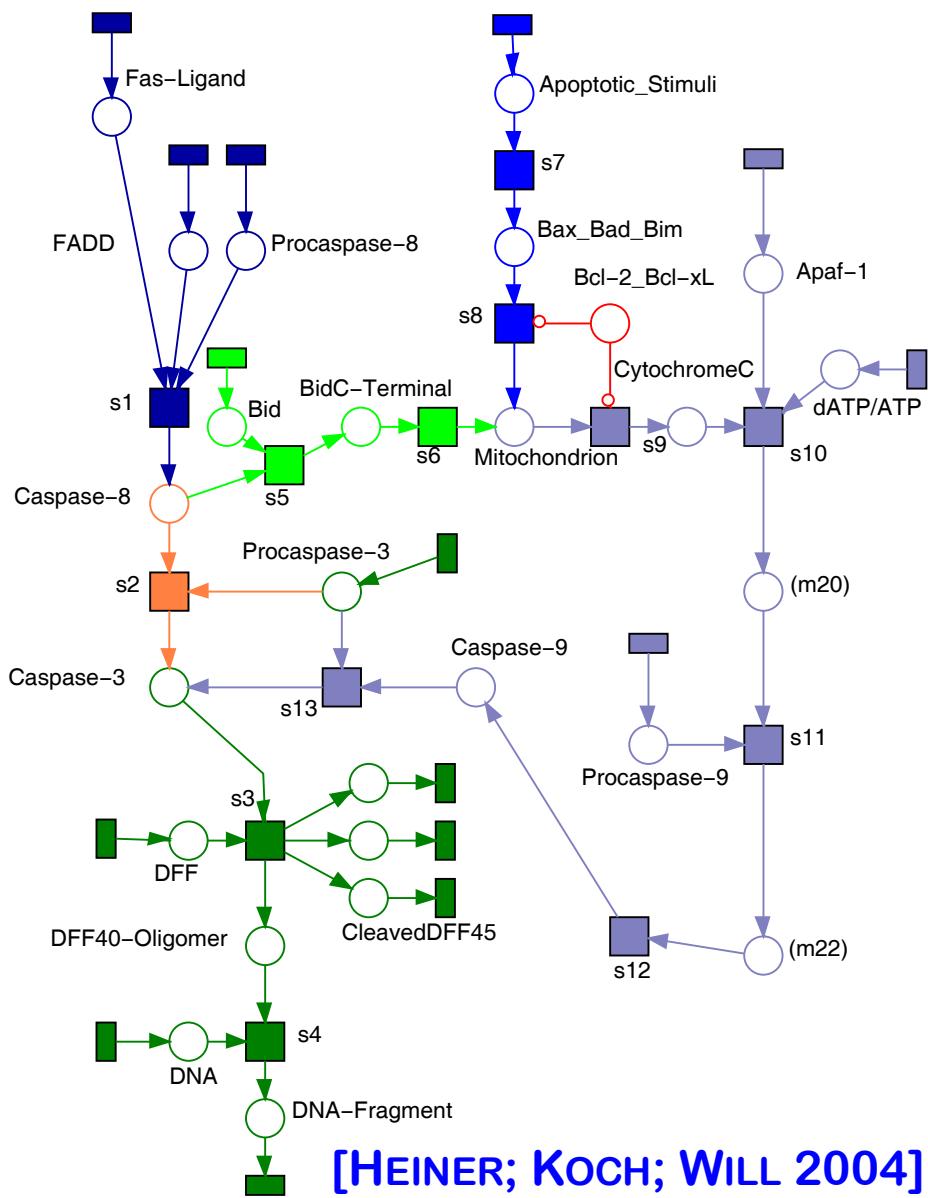


## Ex3: APOPTOSIS IN MAMMALIAN CELLS

PN & Systems Biology



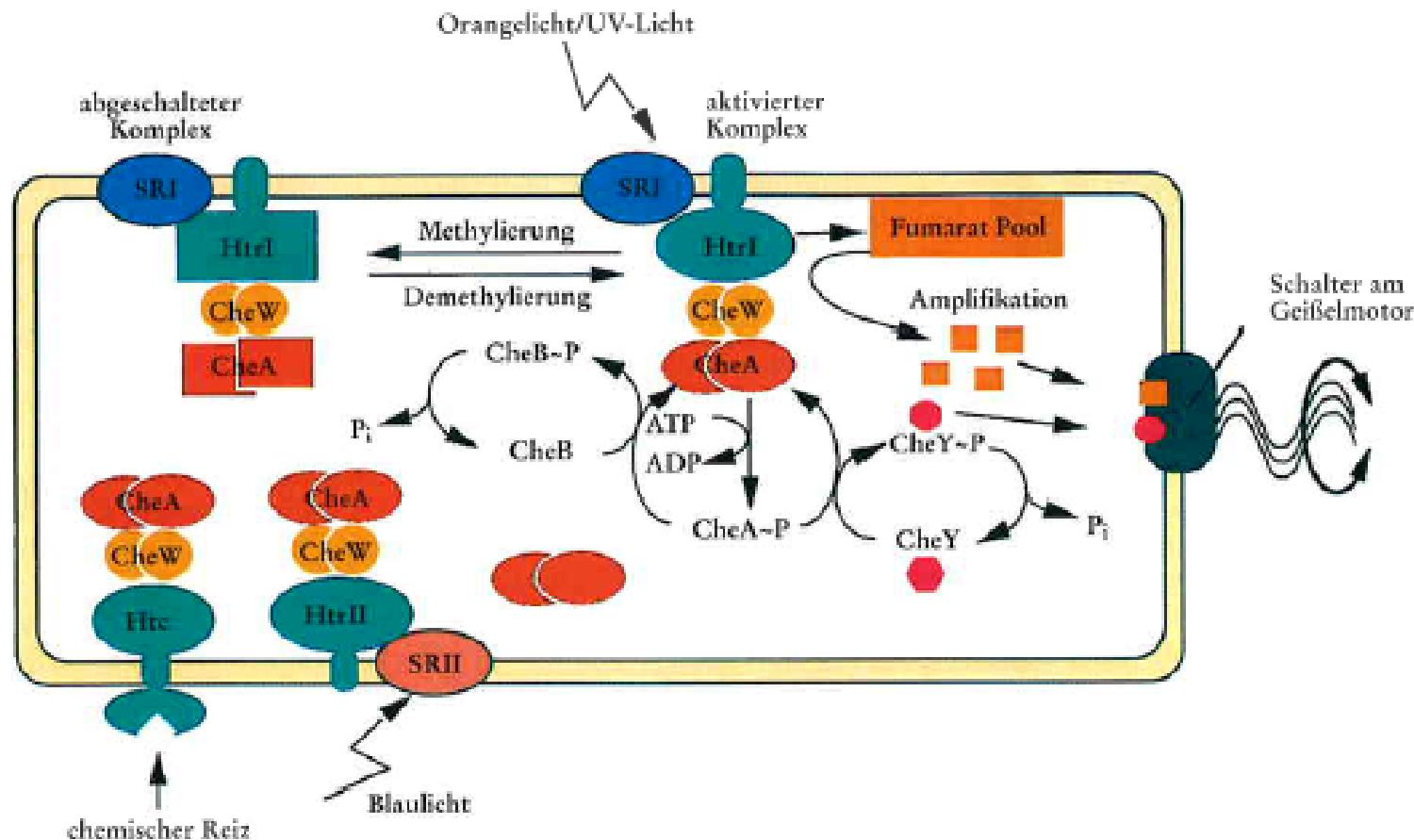
[GON 2003]



[HEINER; KOCH; WILL 2004]

## Ex4 - SWITCH CYCLE HALOBACTERIUM SALINARUM

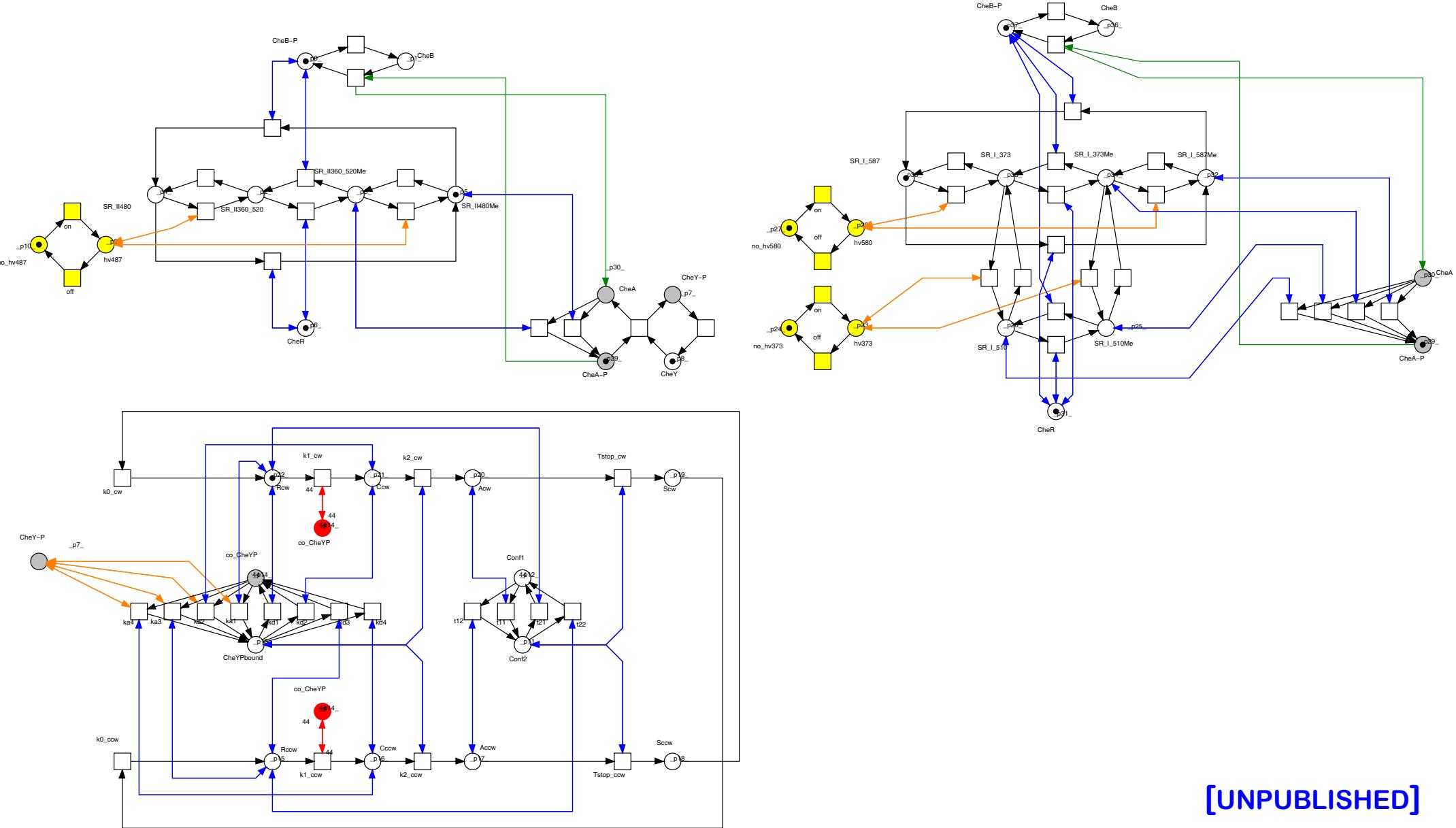
PN & Systems Biology



[Marwan; Oesterhelt 1999]

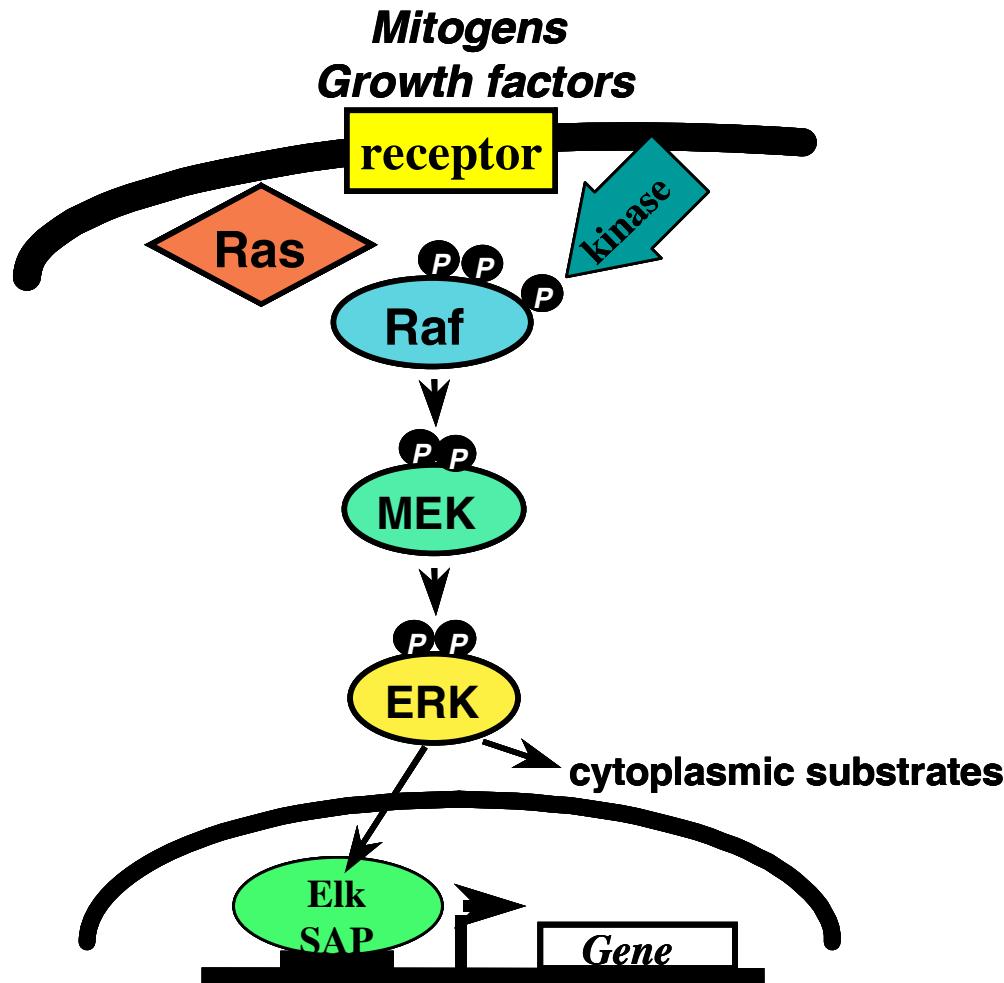
# Ex4 - SWITCH CYCLE HALOBACTERIUM SALINARUM

PN & Systems Biology



[UNPUBLISHED]

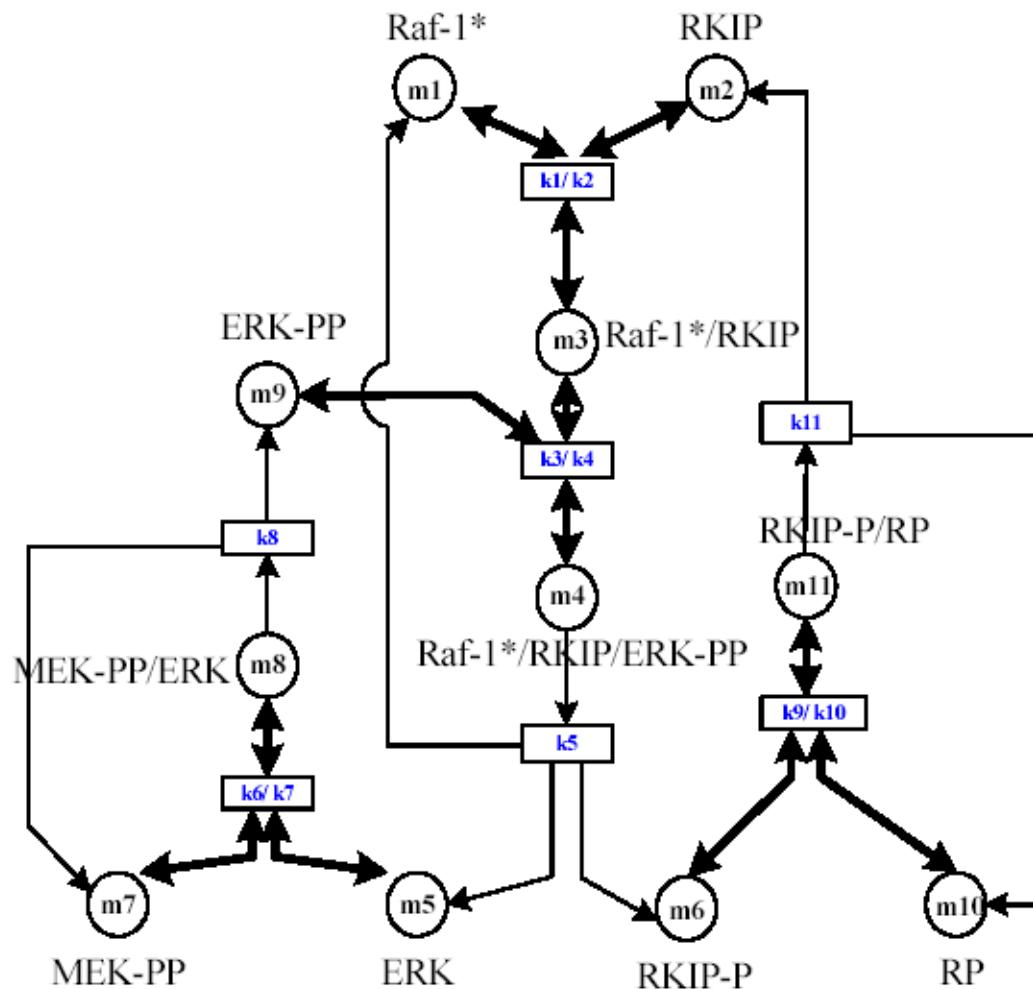
...one pathway...



## Ex5 - THE RKIP PATHWAY

PN & Systems Biology

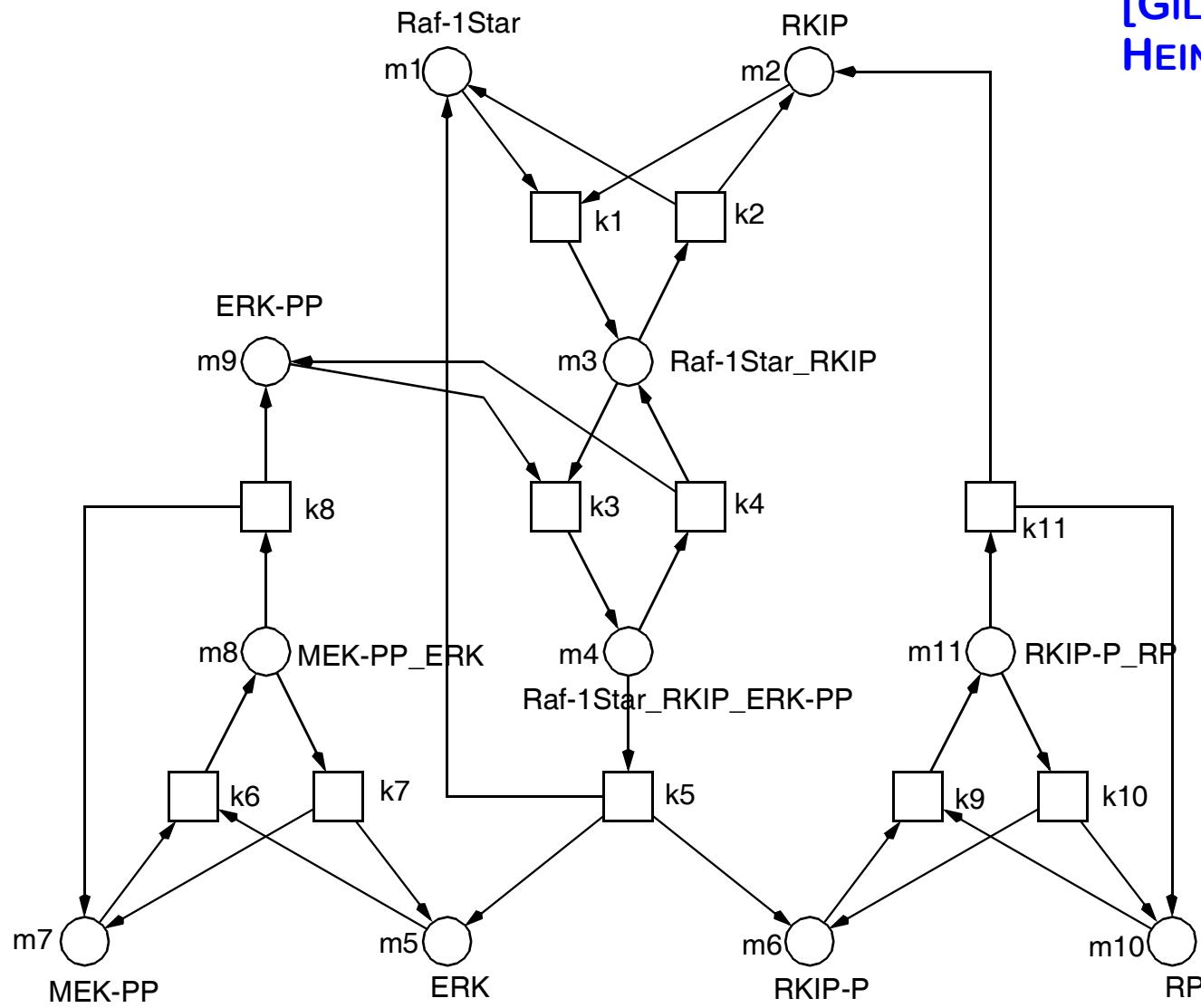
[Cho et al.,  
CMSB 2003]



## Ex5 - THE RKIP PATHWAY, PETRI NET

PN & Systems Biology

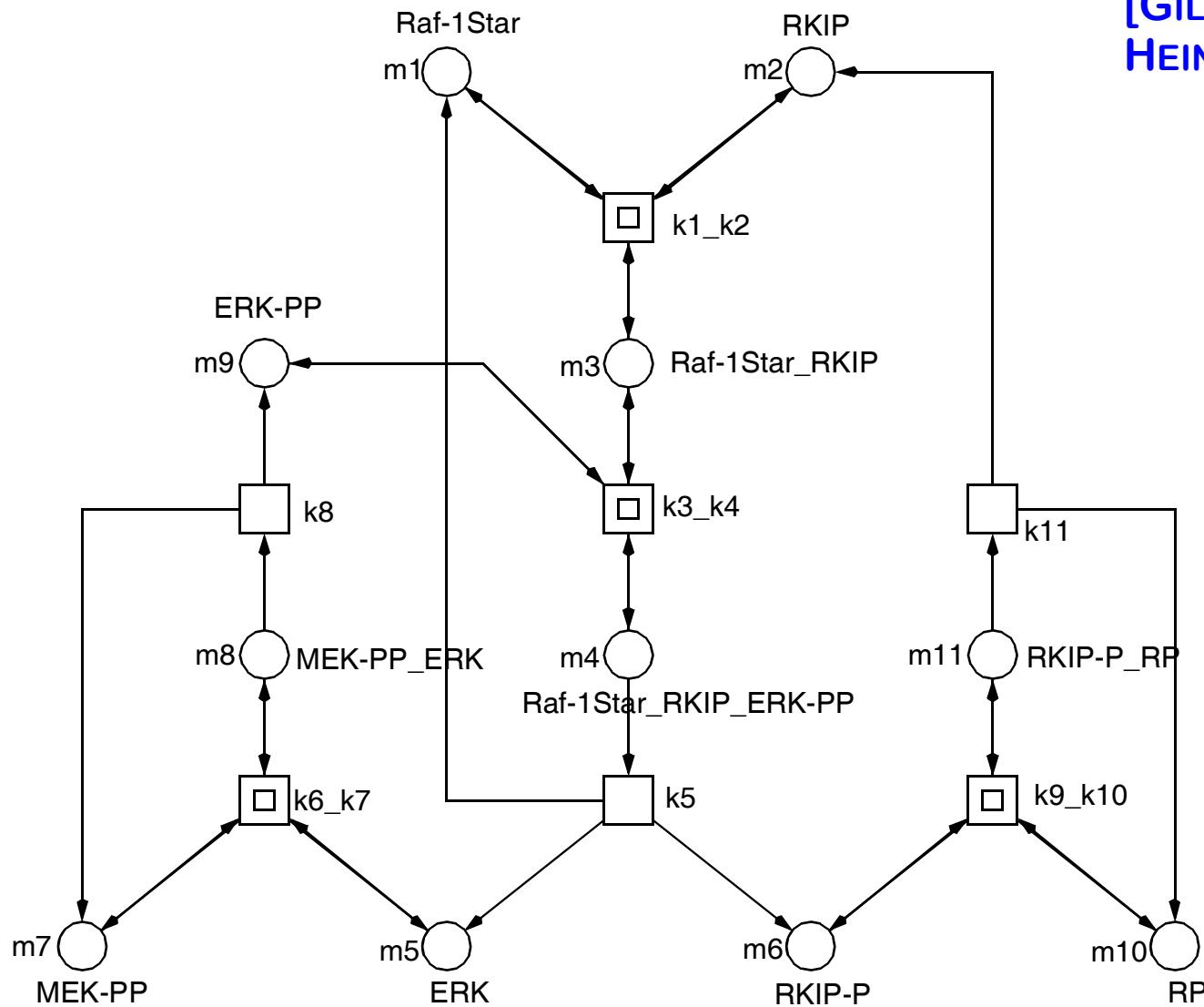
[GILBERT,  
HEINER 2006]



# Ex5 -THE RKIP PATHWAY, HIERARCHICAL PETRI NET

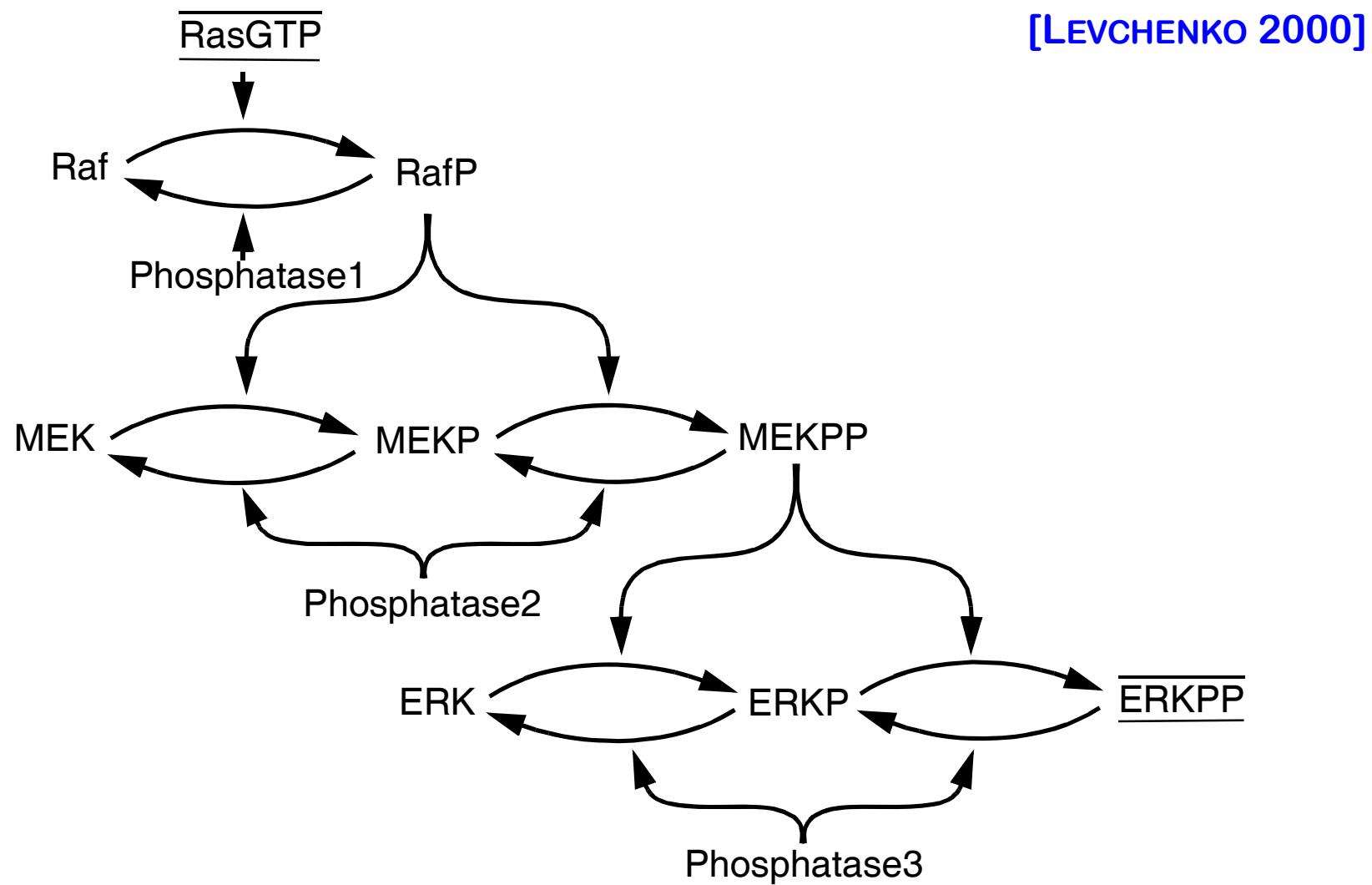
PN & Systems Biology

[GILBERT,  
HEINER 2006]



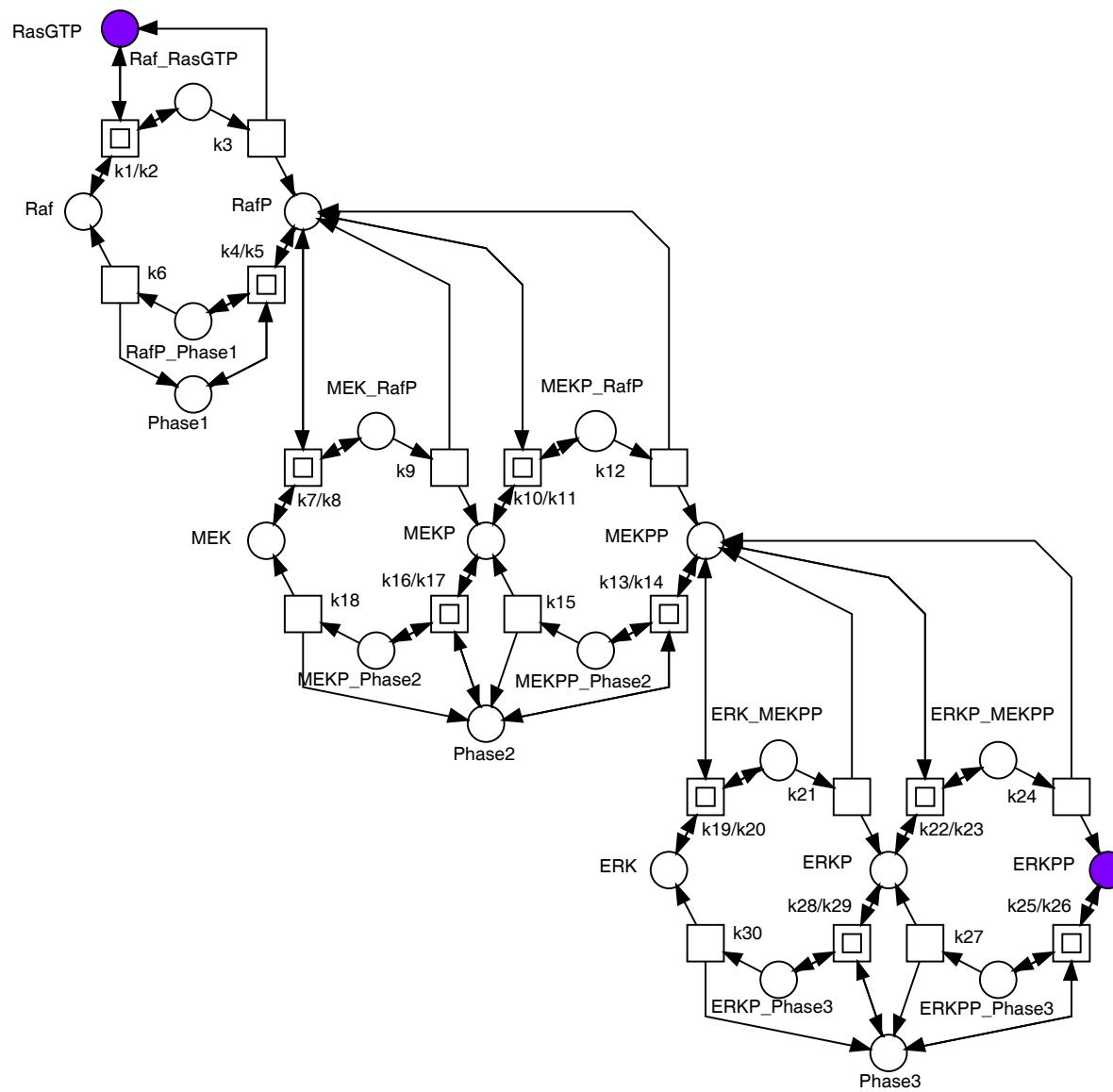
## Ex6 - SIGNALLING CASCADE

PN & Systems Biology



## Ex6 - SIGNALLING CASCADE

PN & Systems Biology



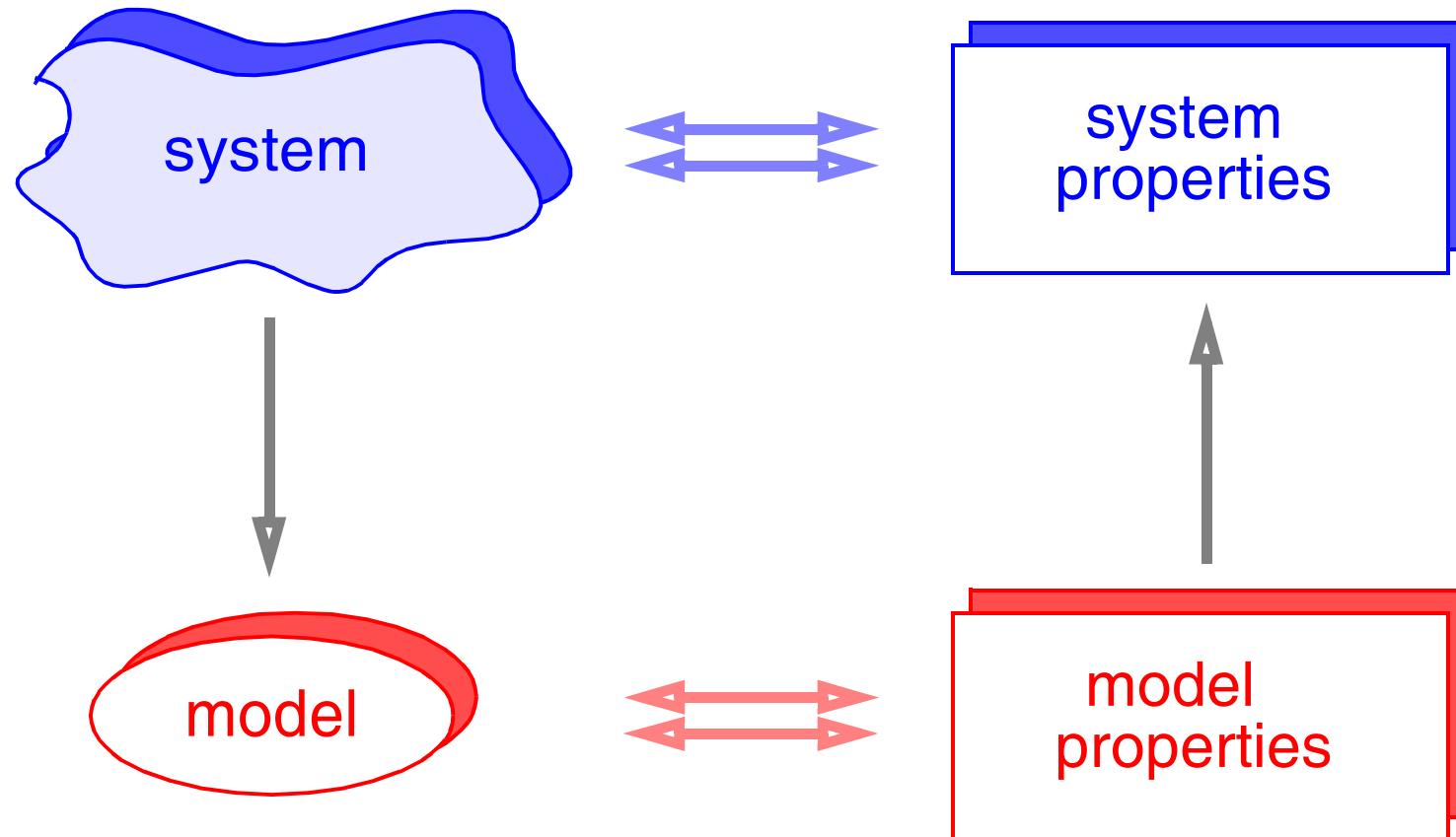
[GILBERT,  
HEINER,  
LEHRACK  
2007]

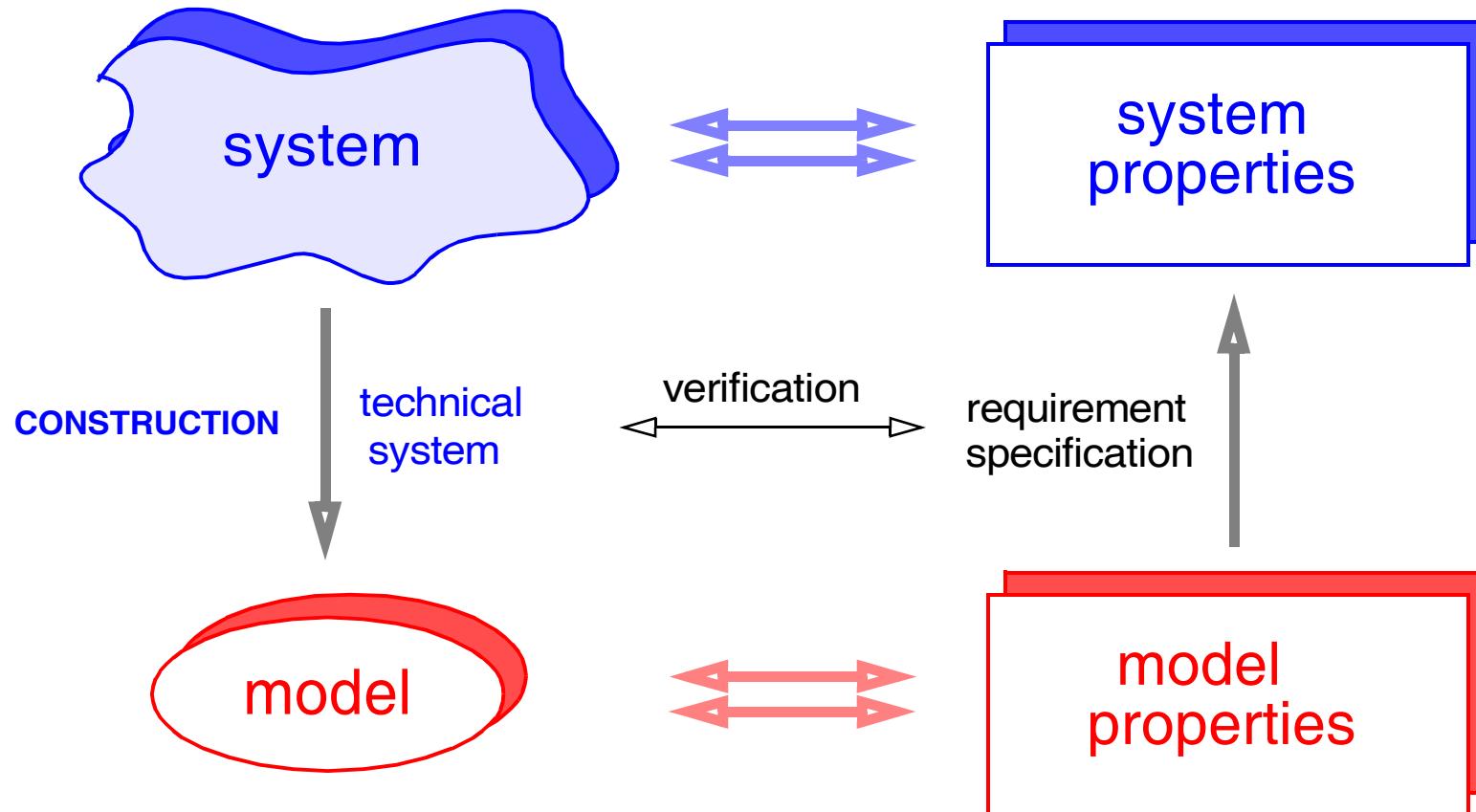
[HEINER,  
GILBERT,  
DONALDSON  
2008]

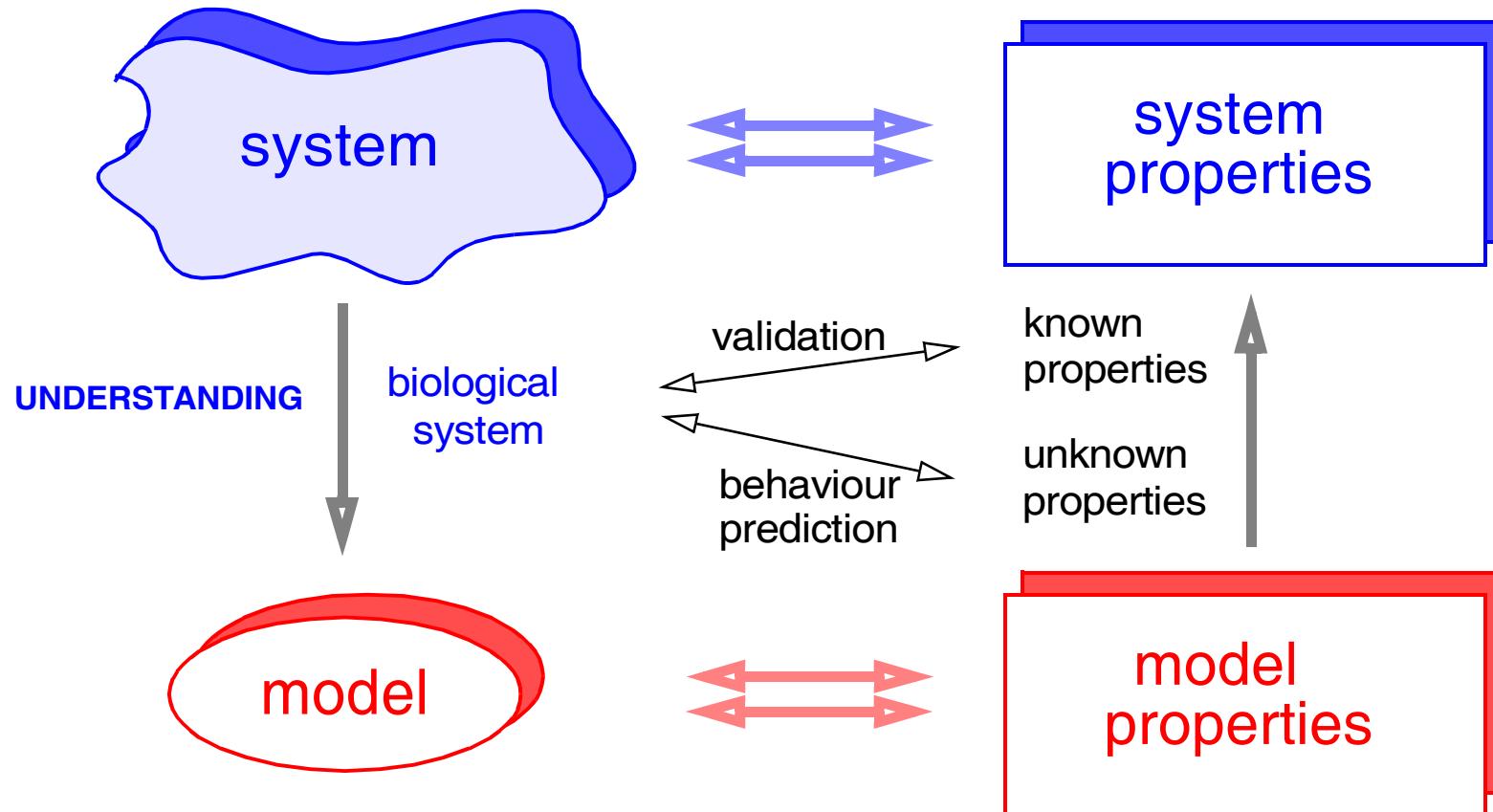
**END OF  
PART I**

# **BIO PETRI NETS, PART II**

## **MODEL-BASED SYSTEM ANALYSIS**







# QUALITATIVE ANALYSES

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

->  $(0, 1, k, oo)$

-> **BOUNDEDNESS**

□ How often can a transition fire ?

-> *(0-times, n-times, oo-times)*

-> **LIVENESS**

□ How often can a system state be reached ?

-> *never*

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

-> *n-times*

-> **REPRODUCIBLE**

-> *always reachable again*

-> **REVERSIBLE (HOME STATE)**

-> *reversible initial state*

-> **REVERSIBILITY**

## GENERAL BEHAVIOURAL PROPERTIES

-> *orthogonal*

-> *general decidable*

# MODEL ANIMATION (?)

# DYNAMIC ANALYSES

# DYNAMIC ANALYSES

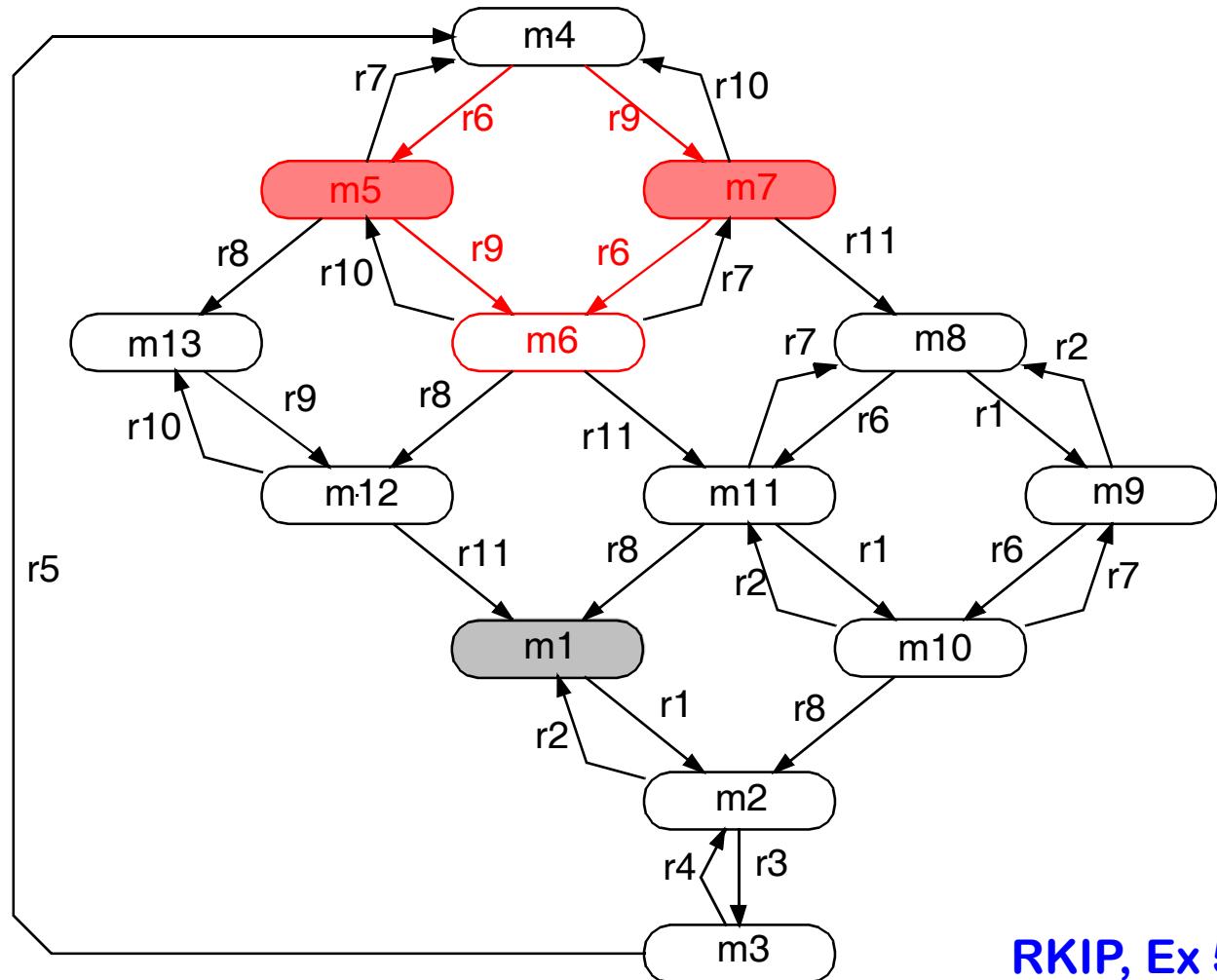
reachability / occurrence graph,  
(labelled) state transition graph

CTMC, Kripke structure

# REACHABILITY GRAPH CONSTRUCTION

PN & Systems Biology

- ❑ simple algorithm
- ❑ nodes : system states
- ❑ arcs : the (single) firing transition
- ❑ single step firing rule



- **interleaving semantics**

- > *(sequential) finite automaton*
- > *concurrency == enumerating all interleaving sequences*

- **boundedness**

- > *finite graph*

- **reversibility**

- > *one Strongly Connected Component (SCC)*

- **liveness**

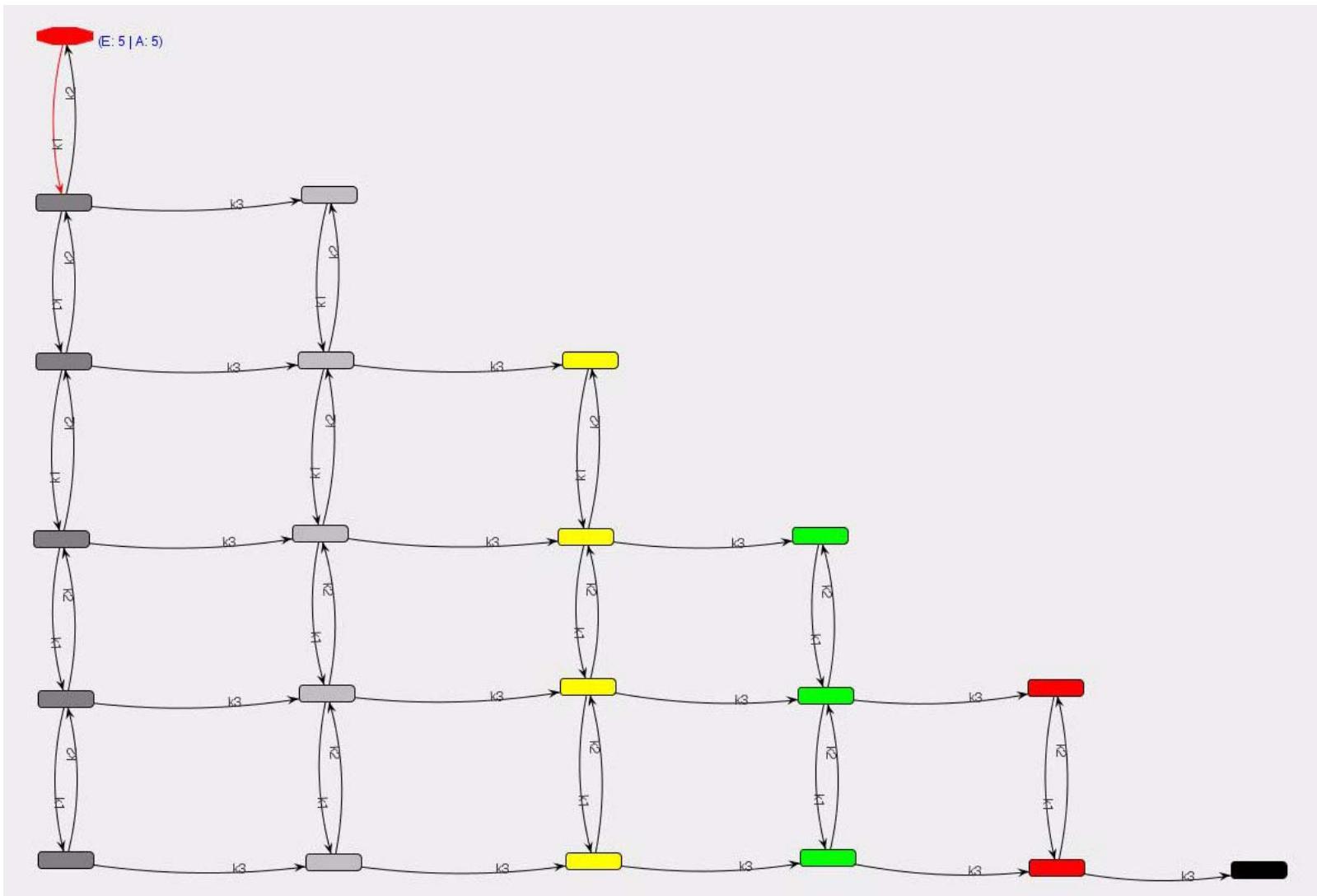
- > *every transition contained in all terminal SCC*

- **dead states**

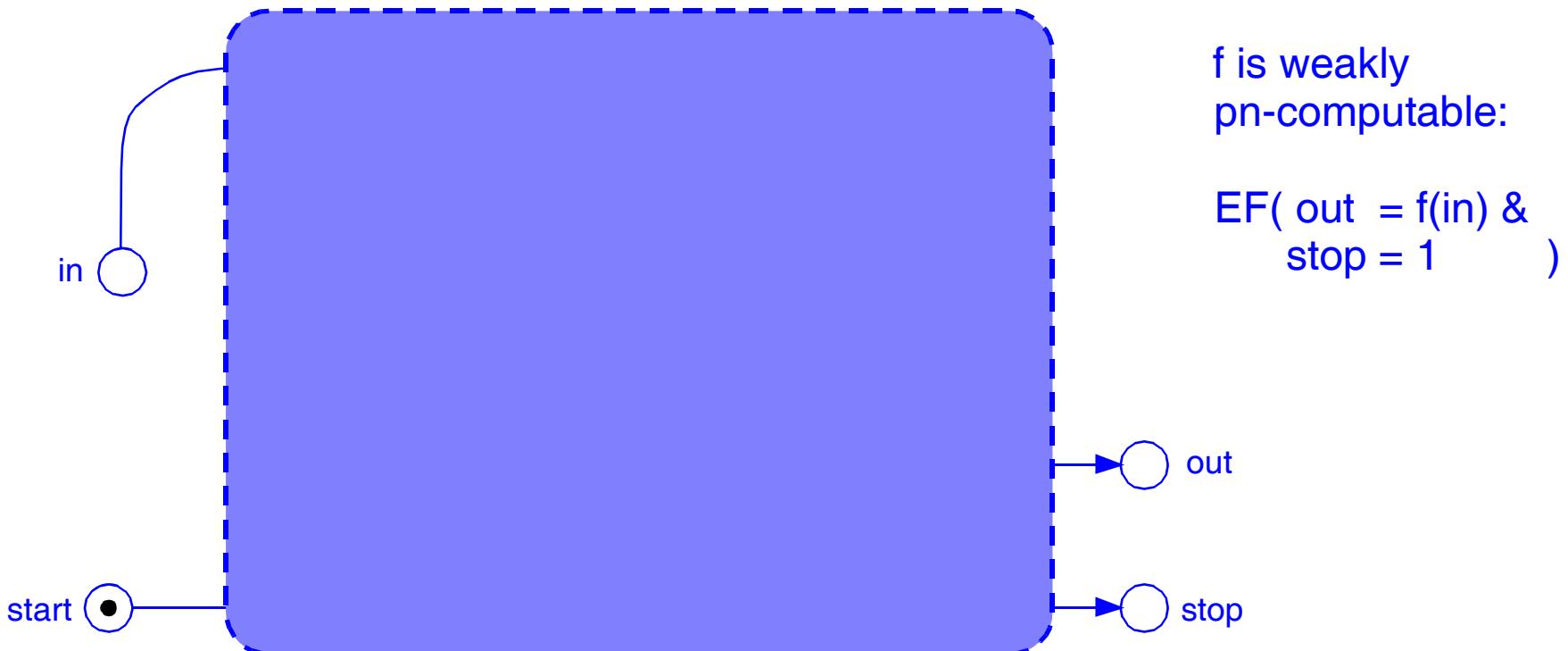
- > *terminal nodes*

# REACHABILITY GRAPH, EX: MA1, 5 TOKENS

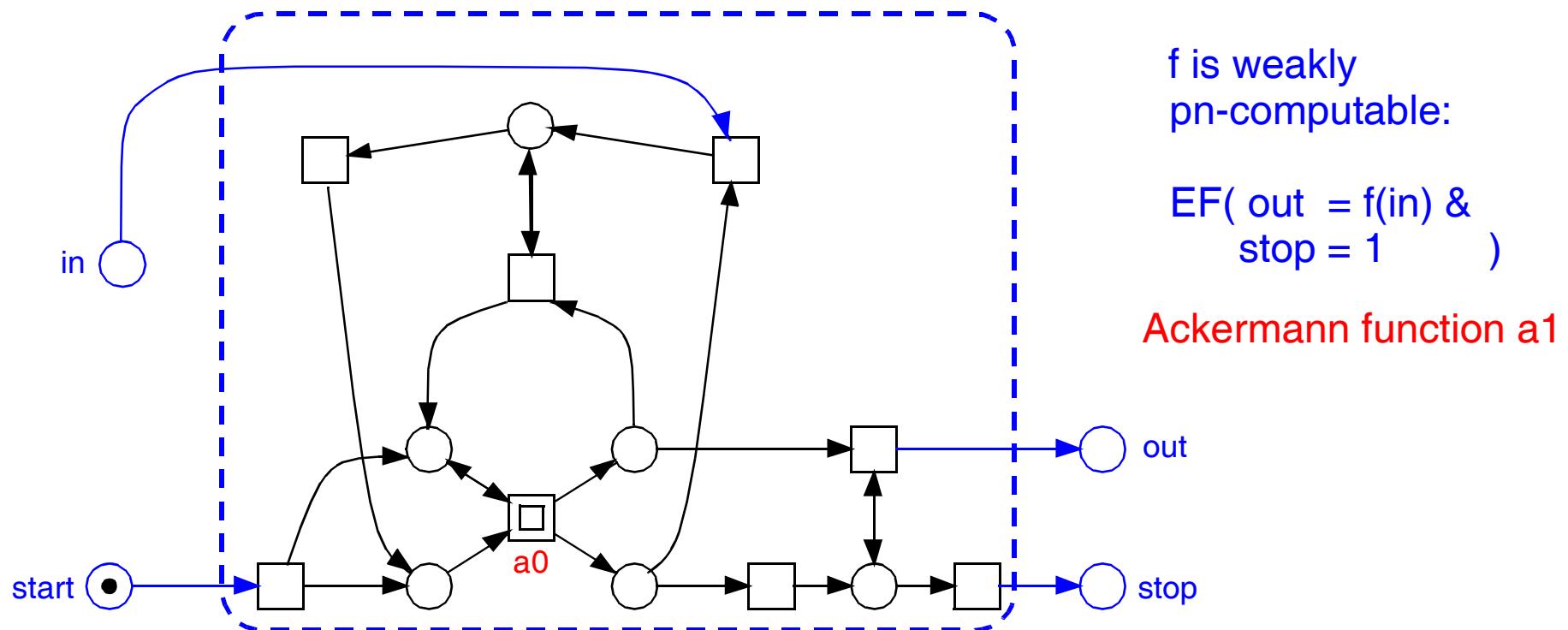
PN & Systems Biology



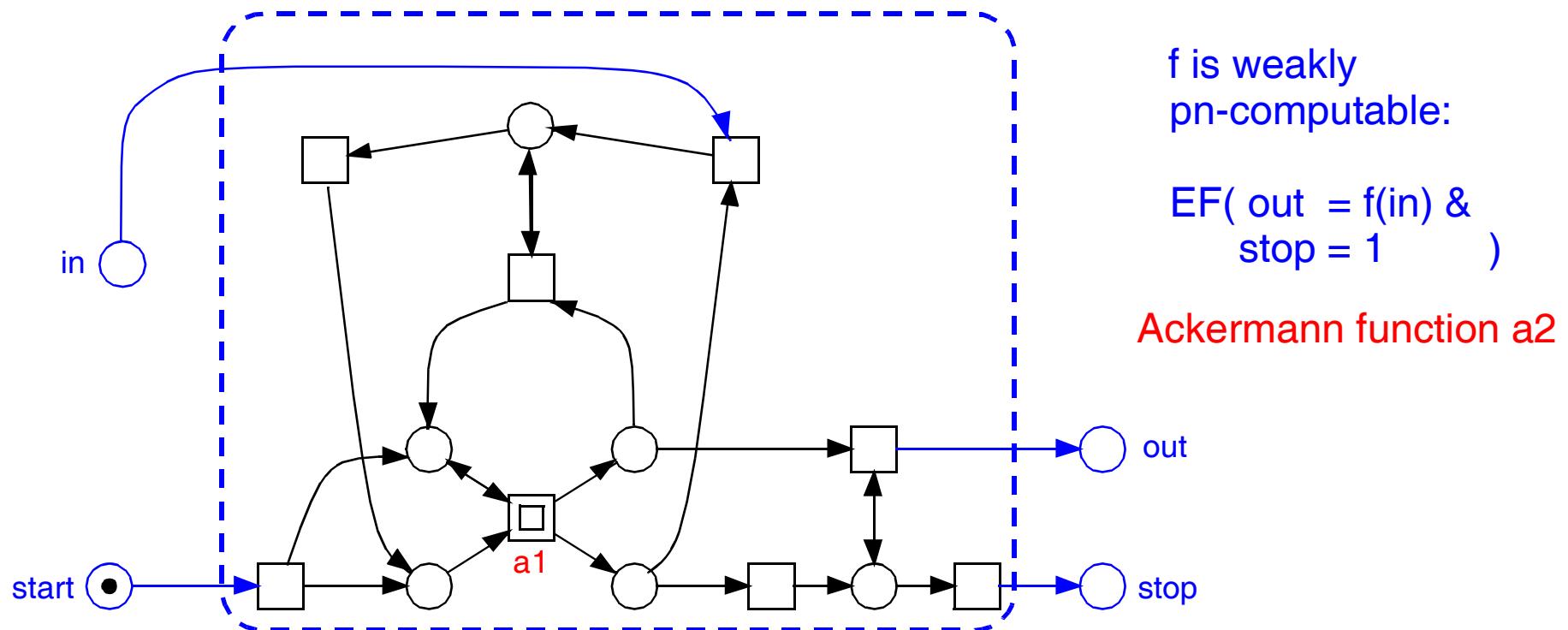
- infinite for unbounded nets
- worst-case for finite state spaces [Pries, Wimmel 2003]  
... *cannot be bounded by a primitive recursive function ...*
- proof -> Petri net computer for a function  $f: \mathbb{N}_0^m \rightarrow \mathbb{N}_0$



- infinite for unbounded nets
- worst-case for finite state spaces [Priese, Wimmel 2003]  
... *cannot be bounded by a primitive recursive function ...*
- proof -> Petri net computer for a function  $f: \mathbb{N}_0^m \rightarrow \mathbb{N}_0$

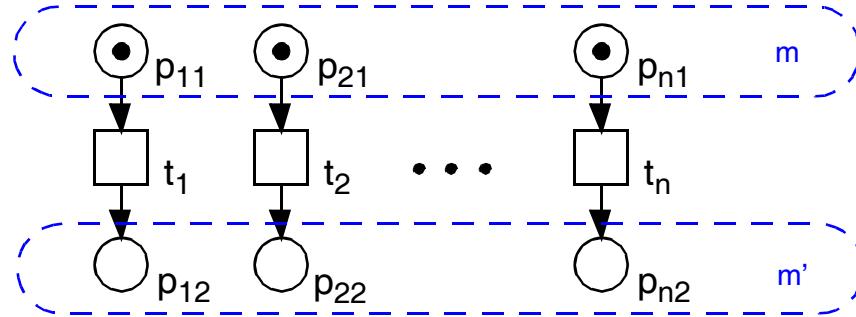


- infinite for unbounded nets
- worst-case for finite state spaces [Priese, Wimmel 2003]  
... *cannot be bounded by a primitive recursive function ...*
- proof -> Petri net computer for a function  $f: \mathbb{N}_0^m \rightarrow \mathbb{N}_0$



# STATE SPACE COMPLEXITY, CAUSES

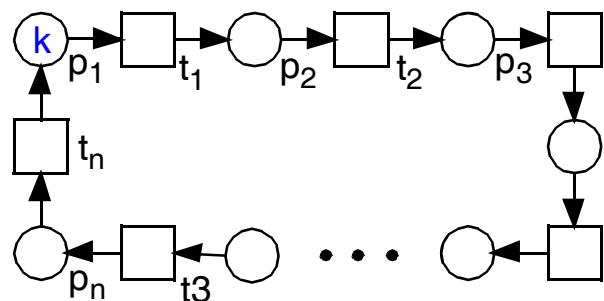
PN & Systems Biology



$n!$  interleaving sequences

$m \rightarrow m'$

$2^n - 2$  intermediate states



$\frac{(n + k - 1)!}{(n - 1)! k!}$  states

(combination with repetition)

- 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,  
i.e. 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)
  
  - => expressed in temporal logics (CTL / LTL),  
as very flexible & powerful query language

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

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

-> *BOUNDEDNESS*

- How often can a transition fire ?

-> *(0-times, n-times, oo-times)*

-> *LIVENESS*

- How often can a system state be reached ?

-> *never*

-> *UNREACHABLE* -> *SAFETY PROPERTIES*

-> *n-times*

-> *REPRODUCIBLE*

-> *always reachable again*

-> *REVERSIBLE (HOME STATE)*

-> *reversible initial state*

-> *REVERSIBILITY*

- Are there behaviourally invariant subnet structures ?

-> *token conservation*

-> *P - INVARIANTS*

-> *token distribution reproduction*

-> *T - INVARIANTS*

- ... and many more -> temporal logics (CTL, LTL)



## □ Petri net theory

- > INA (HU Berlin)
- > TINA (LAAS/CNRS)
- > Charlie

## □ model checking

		CTL	LTL
-> reachability graph	->	INA, Charlie	Charlie
		PROD, MARIA	PROD, MARIA
-> lazy state spaces			
- stubborn set reduction	->	LoLA	PROD (LTL\X)
- symmetry reduction	->	LoLA	
-> compressed state spaces (BDD, NDD, ... , IDD)	->	bdd-CTL, SMART idd-CTL	bdd-LTL idd-LTL
-> Kronecker algebra	->	[Kemper]	
<hr/>			
-> prefix	->	PEP ( $CTL_0$ )	QQ (LTL\X)
-> process automata	->	[pd]	

Petri net theory

- > INA (HU Berlin)
- > TINA (LAAS/CNRS)
- > Charlie

 model checking

- > reachability graph
- > lazy state spaces
  - stubborn set reduction
  - symmetry reduction
- > compressed state spaces (BDD, NDE, ..., PDD)
- > Kronecker algebra
- > refix
- > process automata

-&gt;

CTL

INA, Charlie

PROD, MARIA

-&gt;

LoLA  
LoLA

-&gt;

bdd-CTL, SMART  
idd-CTL

-&gt;

[Kemper]

-&gt;

PEP ( $CTL_0$ )  
[pd]

-&gt;

LTL

Charlie

PROD, MARIA

PROD (LTL\X)

bdd-LTL  
idd-LTL

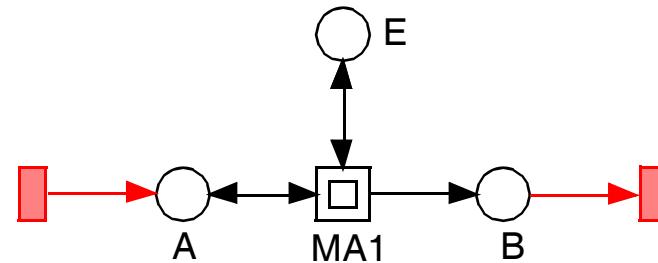
QQ (LTL\X)

**TOOLBOX**

# STATIC ANALYSES

□ **boundary nodes**

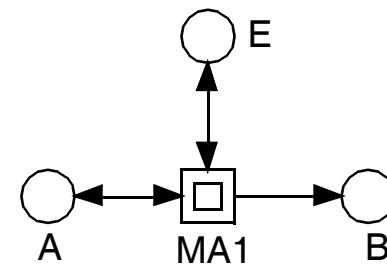
- > *input transitions* -> *not BND*
- > *input places* -> *not LIVE*
- > *LIVE & BND* -> *no boundary nodes*



□ **conservative -> BND**

□ **Deadlock-Trap Property (DTP)**

- > *no structural deadlock* -> *live*
- > *ORD & DTP* -> *no dead states* (Ex6)
- > *ORD & ES & DTP* -> *LIVE* (Ex5)
- > *ORD & EFC & DTP* <-> *LIVE*



## INCIDENCE MATRIX C

PN & Systems Biology

- a representation of the net structure

=> stoichiometric matrix

P \ T	t1	...	tj	...	tm
p1					
pi			cij		
:			$\Delta t j$		
pn					

$$c_{ij} = (p_i, t_j) = F(t_j, p_i) - F(p_i, t_j) = \Delta t_j(p_i)$$

$$\Delta t_j = \Delta t_j(*)$$

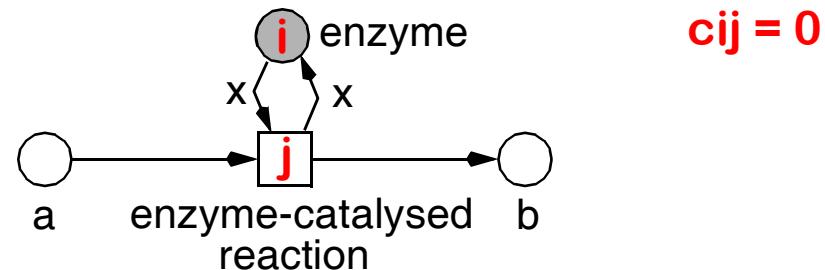
- matrix entry  $c_{ij}$ :

token change in place  $p_i$  by firing of transition  $t_j$

- matrix column  $\Delta t_j$ :

vector describing the change of the whole marking by firing of  $t_j$

- side-conditions are neglected

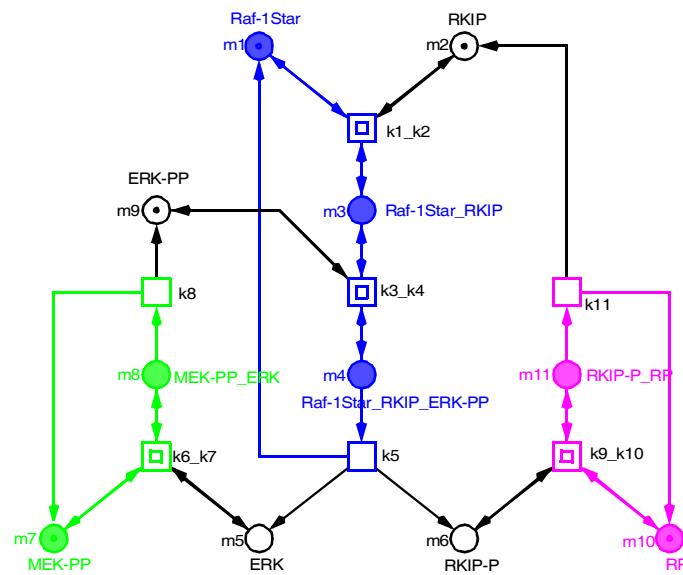


- Lautenbach, 1973
- P-invariants
  - > *integer solutions  $y$  of*  $yC = 0, y \neq 0, y \geq 0$
  - > *multisets of places*
- minimal P-invariants
  - > *there is no P-invariant with a smaller support*
  - > *gcd of all entries is 1*
  - > *sets of places*
- any P-invariant is a non-negative linear combination of minimal ones
  - > *multiplication with a positive integer*
  - > *addition*
  - > *Division by gcd*
  - >  $ky = \sum_i a_i y_i$
- Covered by P-Invariants (CPI)
  - > *each place belongs to a P-invariant*
  - > *CPI => BND (sufficient condition)*

- the firing of any transition has no influence on the weighted sum of tokens on the P-invariant's places
  - > *for all t: the effect of the arcs, removing tokens from a P-invariant's places is equal to the effect of the arcs, adding tokens to a P-invariant's places*
- set of places with
  - > *a constant weighted sum of tokens for all markings m reachable from  $m_0$*   
 $ym = ym_0$
  - > *token / compound preservation,*
  - > *moieties*
  - > *a place belonging to a P-invariant is bounded*
- a P-invariant defines a subnet
  - > *the P-invariant's places (the support),  
+ all their pre- and post-transitions  
+ the arcs in between*
  - > *pre-sets of supports = post-sets of supports*      -> **self-contained**

# THE RKIP PATHWAY, P-INVARIANTS

PN & Systems Biology



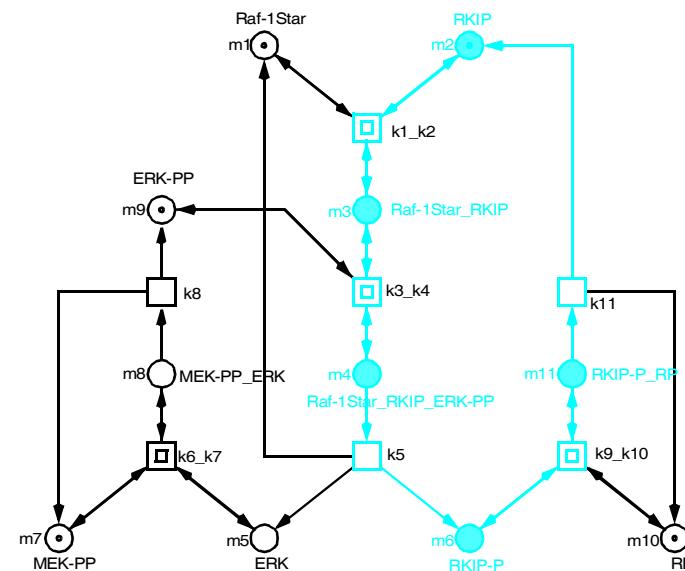
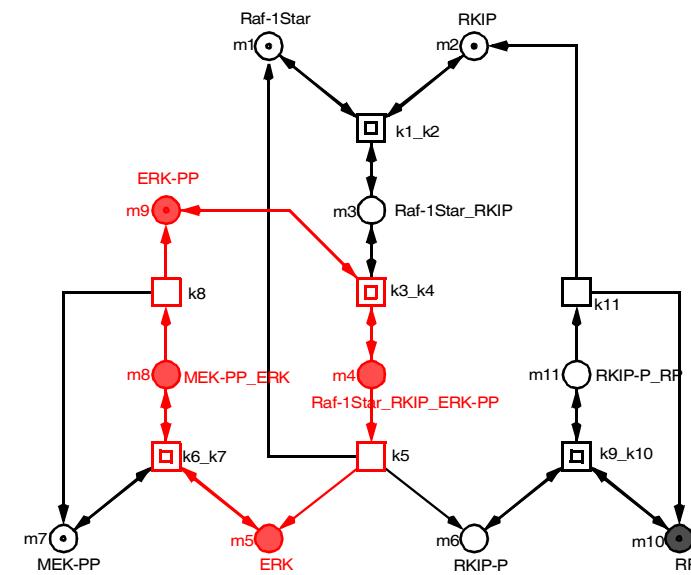
P-INV1: MEK

P-INV2: RAF-1STAR

P-INV3: RP

P-INV4: ERK

P-INV5: RKIP



- Lautenbach, 1973
- T-invariants
  - > *integer solutions  $x$  of*  $Cx = 0, x \neq 0, x \geq 0$
- minimal T-invariants
  - > *there is no T-invariant with a smaller support*
  - > *gcd of all entries is 1*
- any T-invariant is a non-negative linear combination of minimal ones
  - > *multiplication with a positive integer*
  - > *addition*
  - > *Division by gcd*
- Covered by T-Invariants (CTI)
  - > *each transition belongs to a T-invariant*
  - > *BND & LIVE => CTI (necessary condition)*

-> Schuster, 1993

-> *multisets of transitions*

-> *Parikh vector*

-> *sets of transitions*

$$kx = \sum_i a_i x_i$$

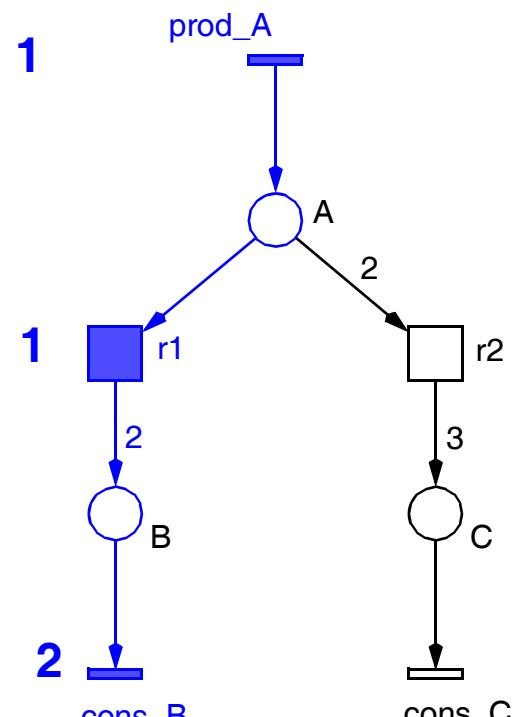
- T-invariants = (multi-) sets of transitions = Parikh vector
  - > zero effect on marking
  - > reproducing a marking / system state
- two interpretations
  1. partially ordered transition sequence
    - of transitions occurring one after the other
    - > substance / signal flow
  2. relative transition firing rates
    - of transitions occurring permanently & concurrently
    - > steady state behaviour
- a minimal T-invariant defines a connected subnet
  - > the T-invariant's transitions (the support),
    - + all their pre- and post-places
    - + the arcs in between
  - > pre-set of support = post-set of support

# T-INVARIANTS, Ex1

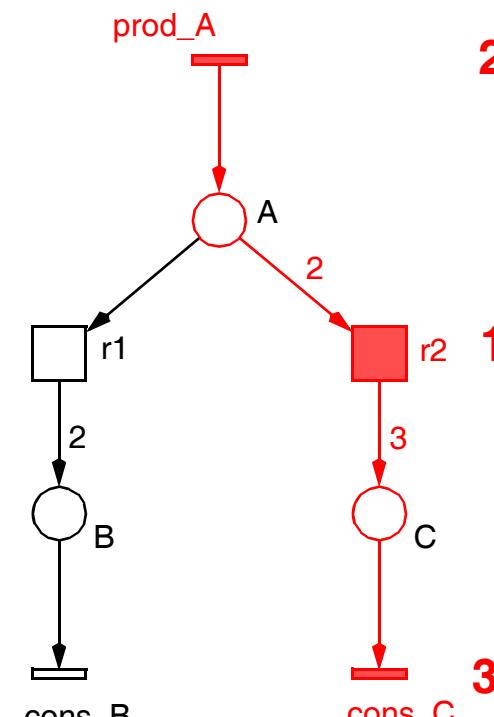
PN & Systems Biology

$r1: A \rightarrow 2B$

$r2: 2A \rightarrow 3C$



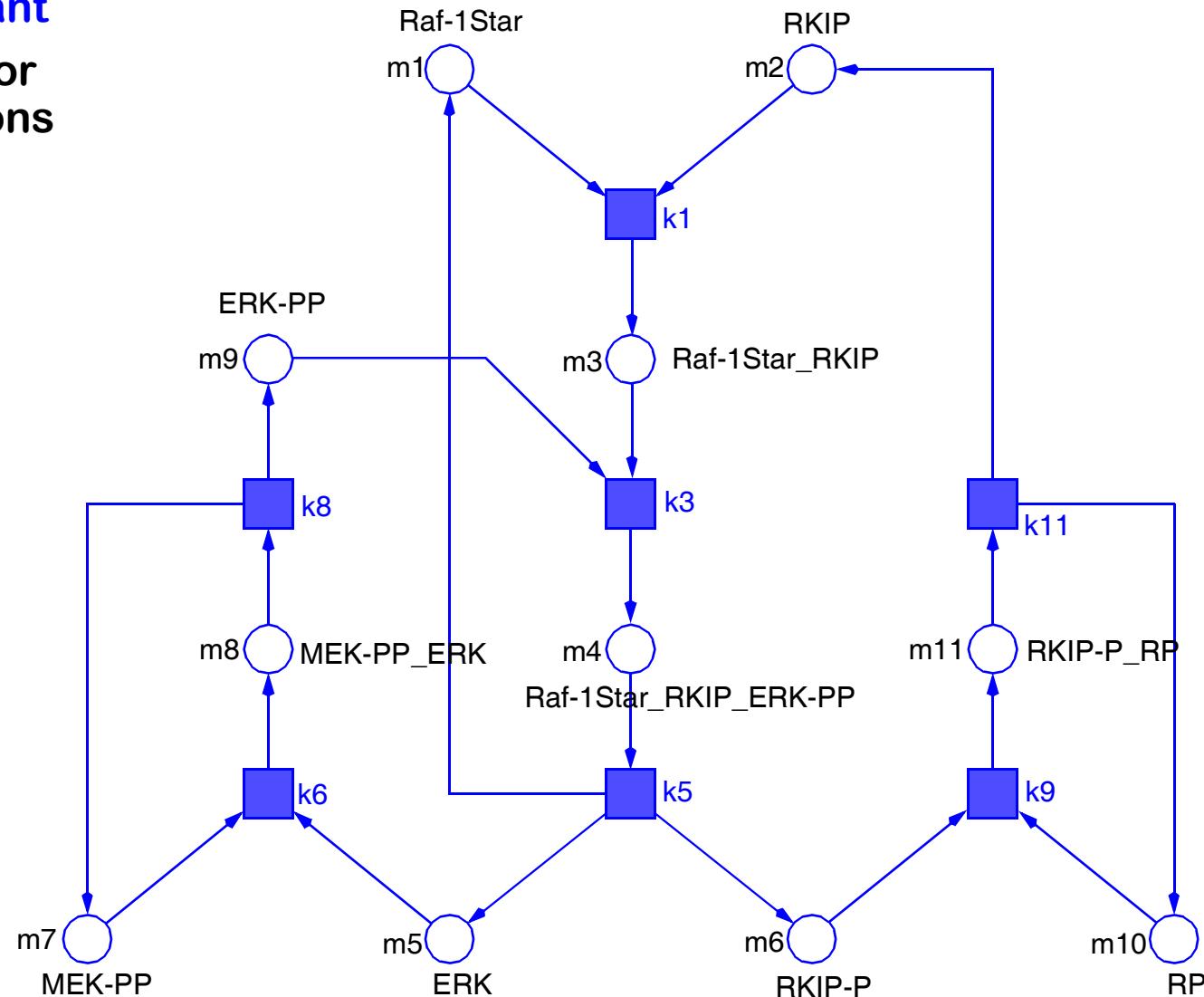
T-INVARIANT 1



T-INVARIANT 2

-> non-trivial T-invariant

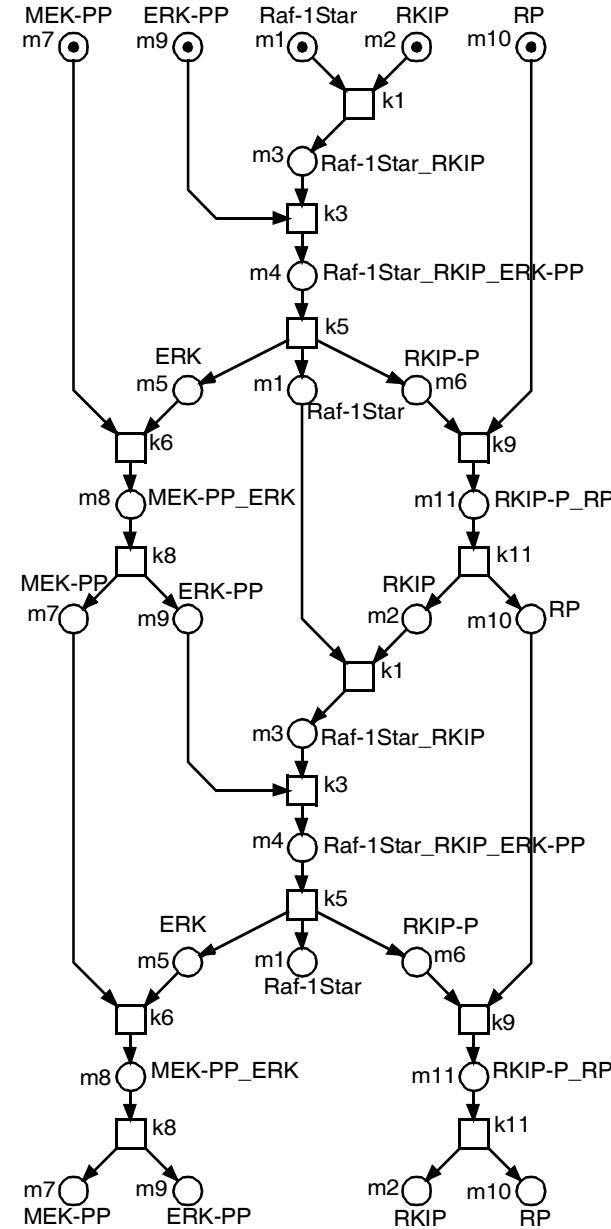
+ four trivial ones for reversible reactions



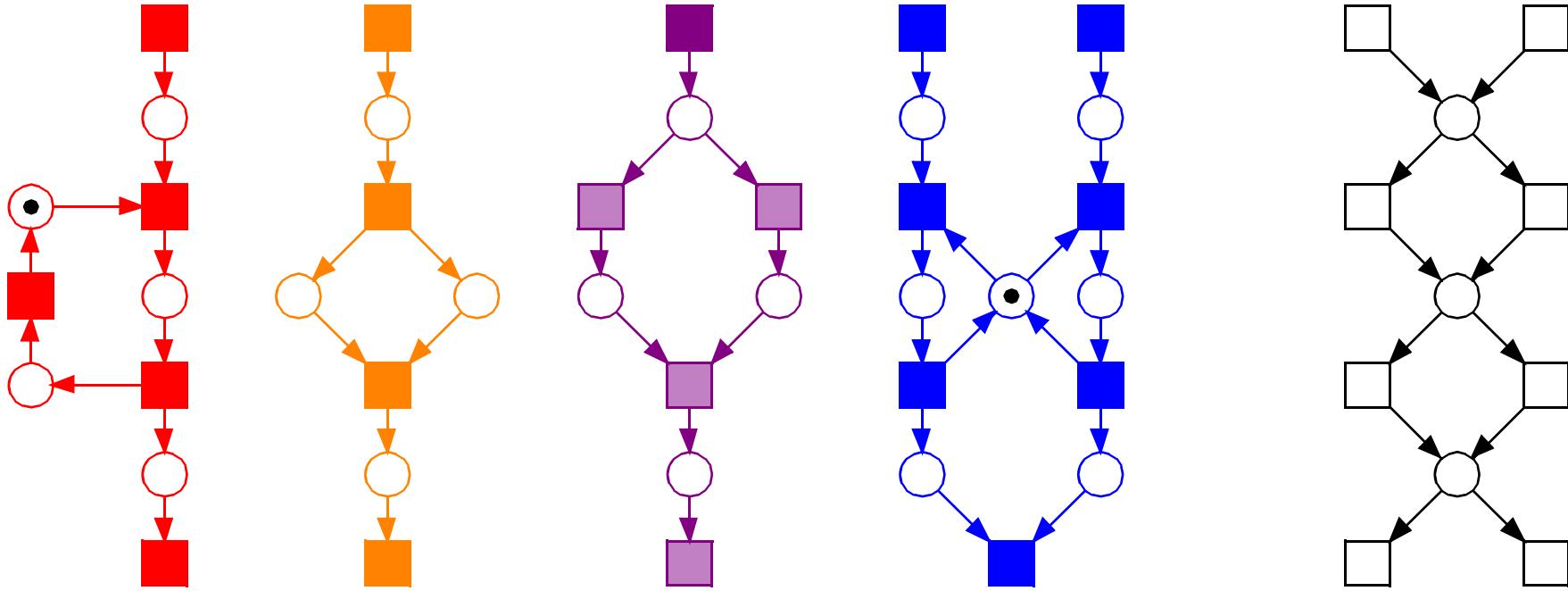
## NON-TRIVIAL T-INVARIANT, RUN

PN & Systems Biology

- realizability check under the constructed marking
- T-invariant's unfolding to describe its behaviour  
-> **partial order structure**
- labelled condition / event net  
-> *events (boxes)*  
- *transition occurrences*  
-> *conditions (circles)*  
- *involved compounds*
- **occurrence net**  
-> *acyclic*  
-> *no backward branching conditions*  
-> *infinite*



- T-invariants may contain any structure



- T-invariants generally overlap  
-> combinatorial effect brings *explosion* in the number of min. T-invariants ( $2^4$ )

- each P-invariant gets at least one token
  - > *P-invariants are structural deadlocks and traps*
- in signal transduction
  - > *exactly 1 token, corresponding to species conservation*
  - > *token in least active state*
- all (non-trivial) T-invariants get feasible
  - > *to make the net live*
- minimal marking
  - > *minimization of the state space*

-> **UNIQUE INITIAL MARKING** <-

## □ validation criterion 1

- > *all expected structural properties hold*
- > *all expected general behavioural properties hold*

## □ validation criterion 2

- > *CPI (if closed model)*
- > *no minimal P-invariant without biological interpretation*

## □ validation criterion 3

- > *CTI*
- > *no minimal T-invariant without biological interpretation*
- > *no known biological behaviour without corresponding T-invariant*

## □ validation criterion 4

- > *all expected special behavioural properties hold*
- > *temporal-logic properties -> TRUE*

- ❑ construction of initial marking
- ❑ subnetwork identification
  - > *P-invariants: token preserving modules (mass conservation)*
  - > *T-invariants: state repeating modules (elementary modes)*
- ❑ network validation
  - > *structure (topology)*
  - > *initial conditions*
- ❑ choice of stochastic analysis techniques

**NOW WE ARE READY  
FOR SOPHISTICATED  
QUANTITATIVE ANALYSES !**

## ❑ transitions

- > *exponentially distributed waiting time*
- > *state-dependent propensity (hazard) function*

## ❑ semantics

- > *Continuous Time Markov Chain (CTMC)*

## ❑ CTMC ~ reachability graph + transition rates

- > *all qualitative properties are preserved*
- > *reversibility* -> *ergodicity*

## ❑ (sufficiently) finite CTMC

- > *analytical = exact CSL model checking*

## ❑ (practically) infinite CTMC

- > *simulative = approximative PLTLC model checking*
- > *approximation: finite number of finite simulation traces*

- continuous places carry continuous tokens

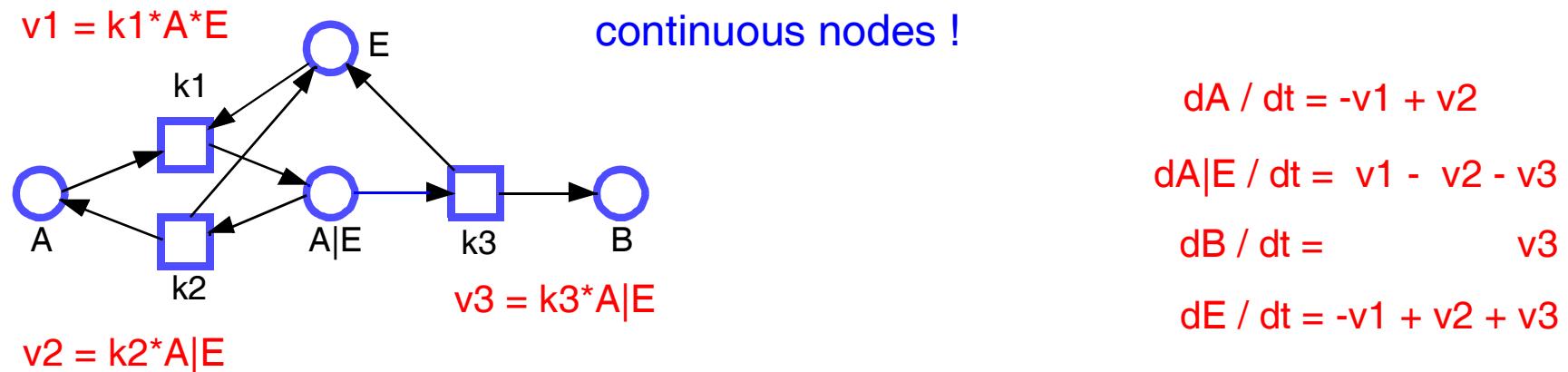
> real numbers -> compound concentrations

- continuous transitions

-> continuous firing / fluxes (if any)

-> state-dependent rate functions

- continuous Petri nets = ODEs

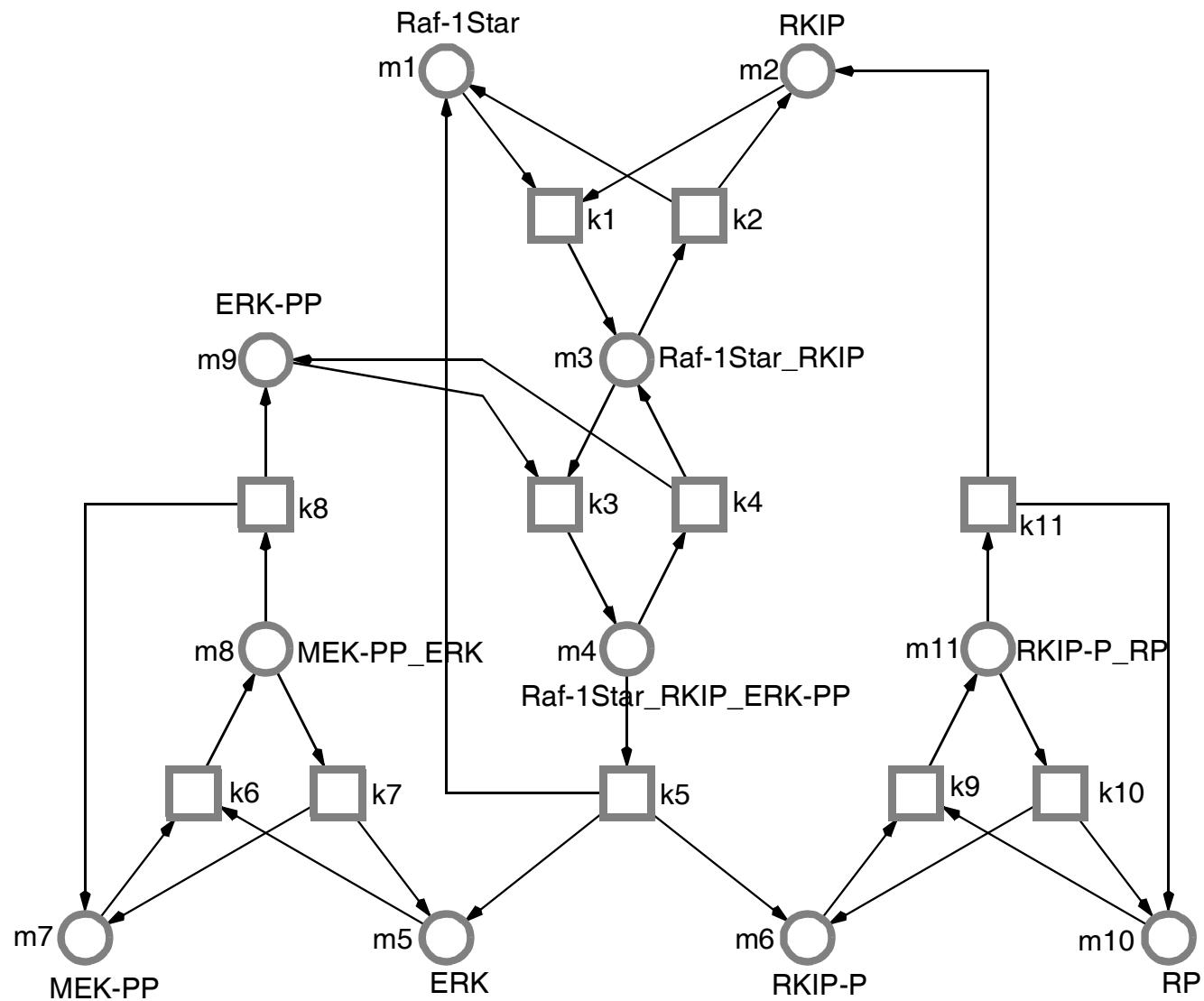


- analysis

-> all standard ODEs techniques + LTLc model checking

# THE RKIP PATHWAY, CONTINUOUS PETRI NET

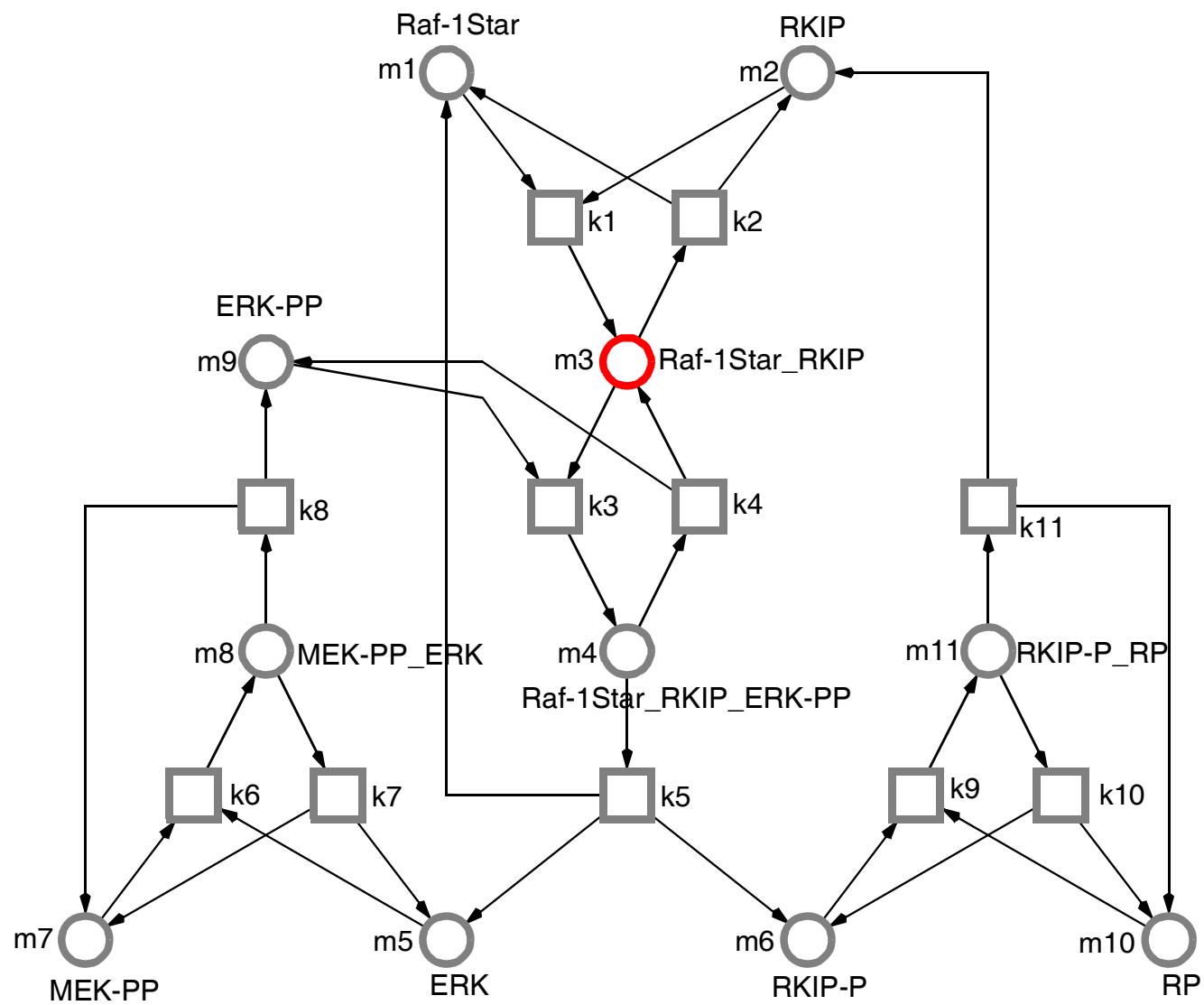
PN & Systems Biology



# THE RKIP PATHWAY, CONTINUOUS PETRI NET

PN & Systems Biology

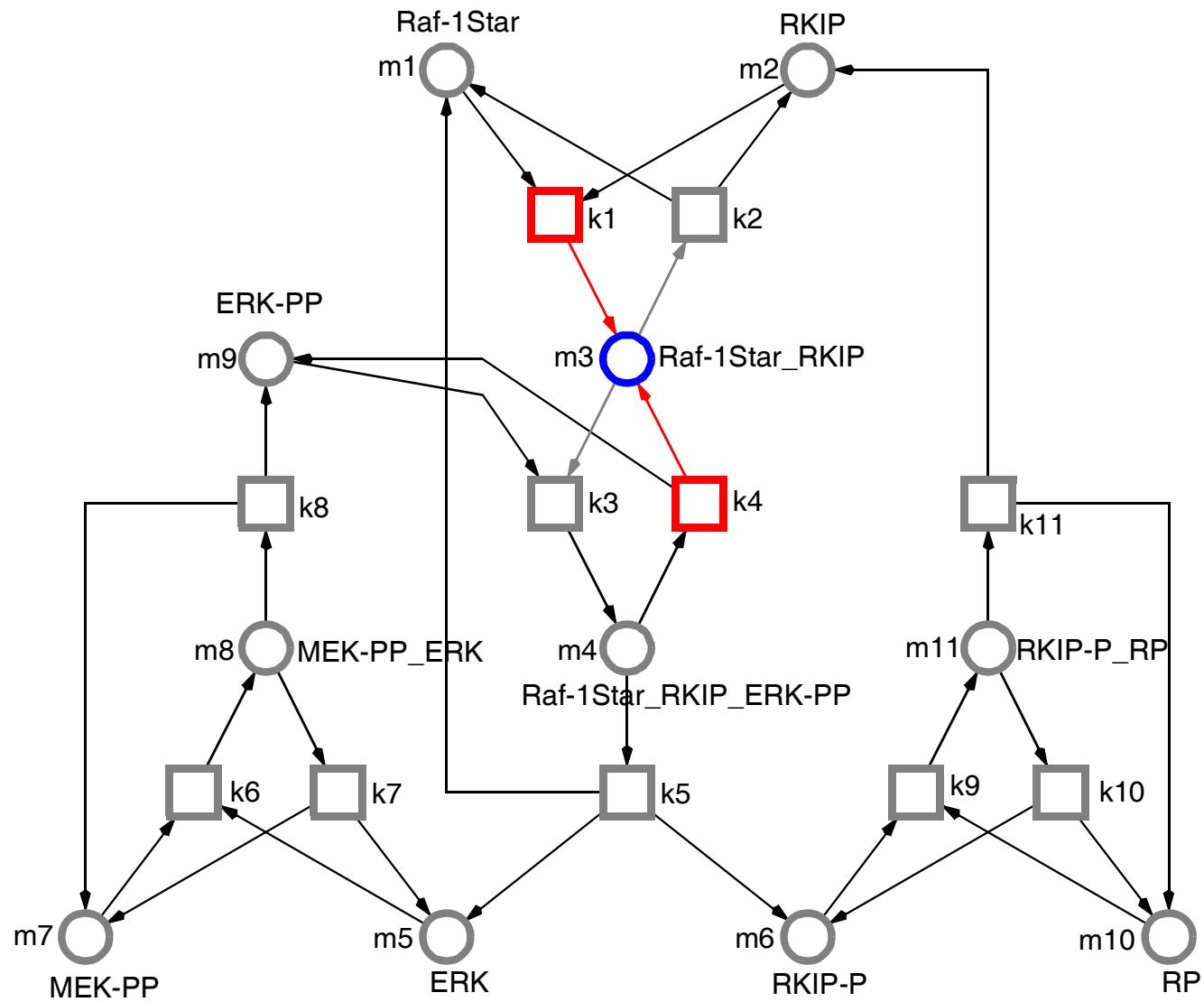
$$\frac{dm_3}{dt} =$$



# THE RKIP PATHWAY, CONTINUOUS PETRI NET

PN & Systems Biology

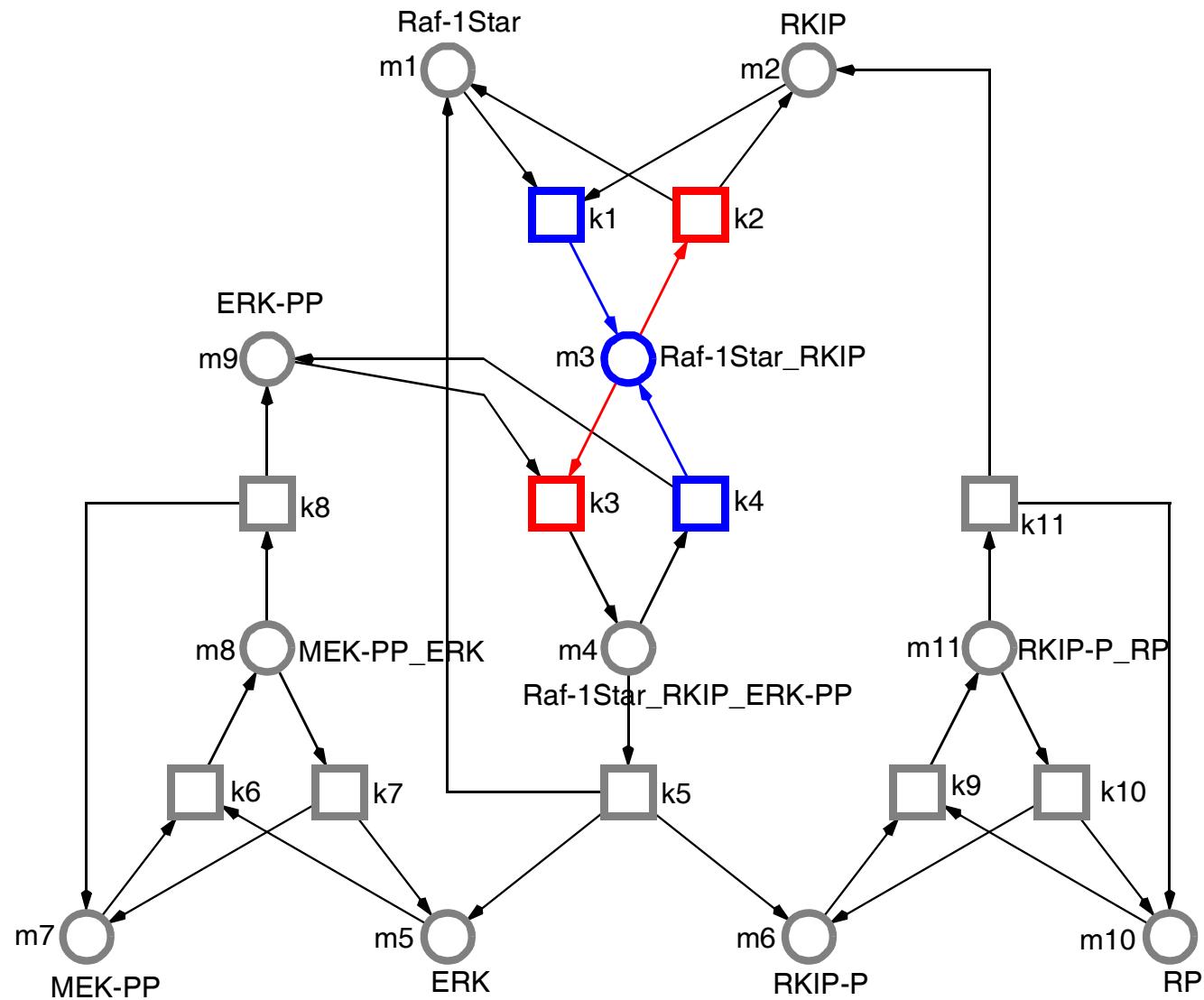
$$\frac{dm_3}{dt} = +r_1 \\ +r_4$$



# THE RKIP PATHWAY, CONTINUOUS PETRI NET

PN & Systems Biology

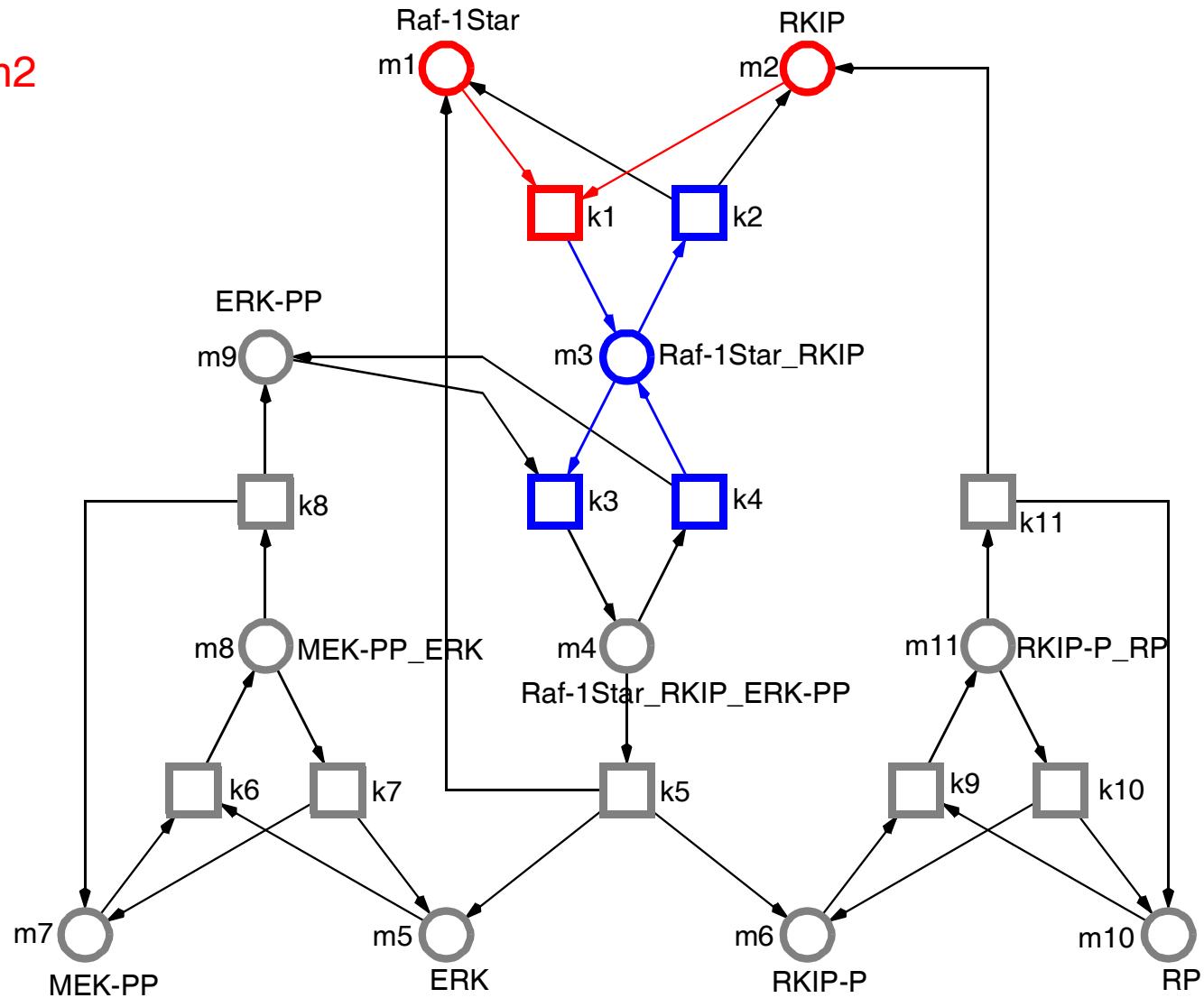
$$\frac{dm_3}{dt} = + r_1 \\ + r_4 \\ - r_2 \\ - r_3$$



# THE RKIP PATHWAY, CONTINUOUS PETRI NET

PN & Systems Biology

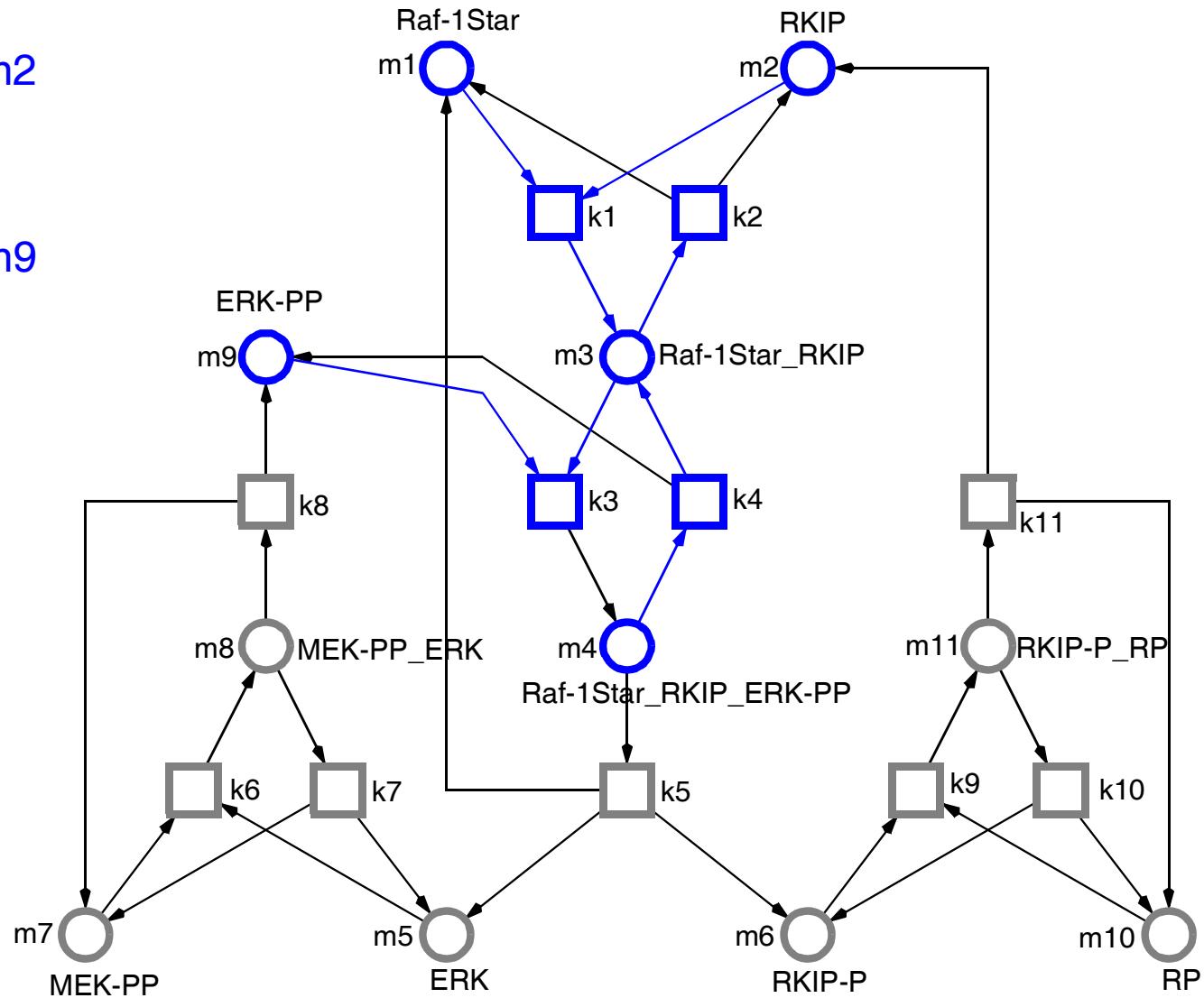
$$\frac{dm_3}{dt} = + k_1 * m_1 * m_2 \\ + r_4 \\ - r_2 \\ - r_3$$



# THE RKIP PATHWAY, CONTINUOUS PETRI NET

PN & Systems Biology

$$\frac{dm_3}{dt} = + k_1 * m_1 * m_2 \\ + k_4 * m_4 \\ - k_2 * m_3 \\ - k_3 * m_3 * m_9$$



**THE QUALITATIVE MODEL  
BECOMES  
THE STRUCTURED DESCRIPTION  
OF THE QUANTITATIVE MODELS !**

# Bio Petri Nets, Part III A Case Study

Monika Heiner  
@tu-cottbus.de

*joint work with David Gilbert, Robin Donaldson  
Bioinformatics Research Centre, University of Glasgow*

Bertinoro, June, 2008

## Definition :

A **place/transition Petri net** is a quadruple

$$\mathcal{PN} = (P, T, f, m_0), \text{ where}$$

- $P, T$  - finite, non empty, disjoint sets (**places, transitions**)
- $f : ((P \times T) \cup (T \times P)) \rightarrow \mathbb{N}_0$  (**weighted directed arcs**)
- $m_0 : P \rightarrow \mathbb{N}_0$  (**initial marking**)

**Interleaving Semantics :** reachability graph / CTL, LTL

## Definition :

A biochemically interpreted stochastic Petri net is a quintuple  $\mathcal{SPN}_{Bio} = (P, T, f, v, m_0)$ , where

- $P, T$  - finite, non empty, disjoint sets (**places, transitions**)
- $f : ((P \times T) \cup (T \times P)) \rightarrow \mathbb{N}_0$  (**weighted directed arcs**)
- $m_0 : P \rightarrow \mathbb{N}_0$  (**initial marking**)
- $v : T \rightarrow H$  (**stochastic firing rate functions**) with
  - $H := \bigcup_{t \in T} \left\{ h_t \mid h_t : \mathbb{N}_0^{|P|} \rightarrow \mathbb{R}^+ \right\}$
  - $v(t) = h_t$  for all transitions  $t \in T$

**Semantics** : Continuous Time Markov Chain / CSL, PLTLc

## Definition :

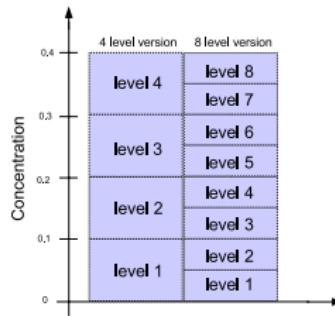
A biochemically interpreted continuous Petri net is a quintuple  $\mathcal{CPN}_{Bio} = (P, T, f, v, m_0)$ , where

- $P, T$  - finite, non empty, disjoint sets (**places, transitions**)
- $f : ((P \times T) \cup (T \times P)) \rightarrow \mathbb{R}_0^+$  (**weighted directed arcs**)
- $m_0 : P \rightarrow \mathbb{R}_0^+$  (**initial marking**)
- $v : T \rightarrow H$  (**continuous firing rate functions**) with
  - $H := \bigcup_{t \in T} \{h_t \mid h_t : \mathbb{R}^{|P|} \rightarrow \mathbb{R}^+\}$
  - $v(t) = h_t$  for all transitions  $t \in T$

**Semantics** : ODEs / LTLc

## Interpretation of tokens :

- *tokens = molecules, moles*
- *tokens = concentration levels*



Specialised stochastic firing rate function, two examples :

- *molecules semantics*

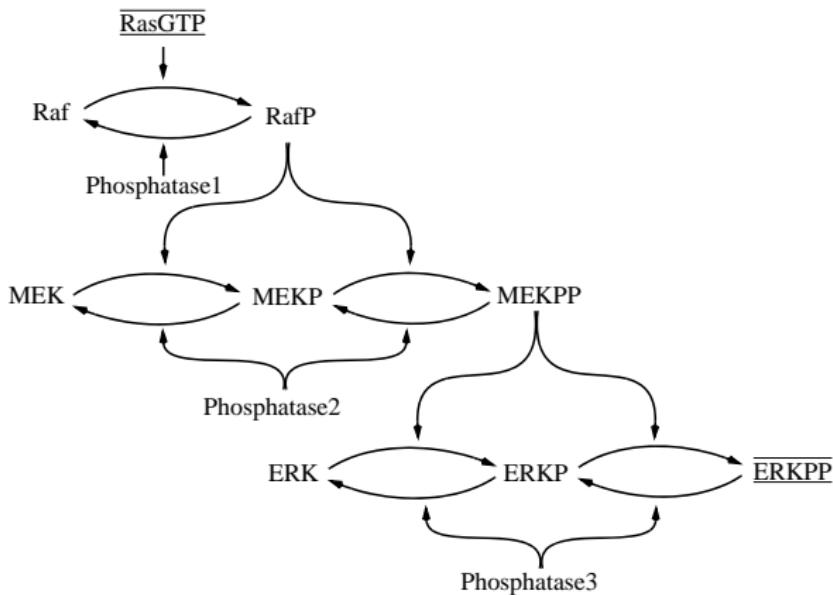
$$h_t := \textcolor{red}{c_t} \cdot \prod_{p \in \bullet t} \binom{m(p)}{f(p, t)} \quad (1)$$

- *concentration levels semantics*

$$h_t := \textcolor{red}{k_t} \cdot N \cdot \prod_{p \in \bullet t} \left( \frac{m(p)}{N} \right) \quad (2)$$

# Running Case Study

- ... a typical signalling cascade



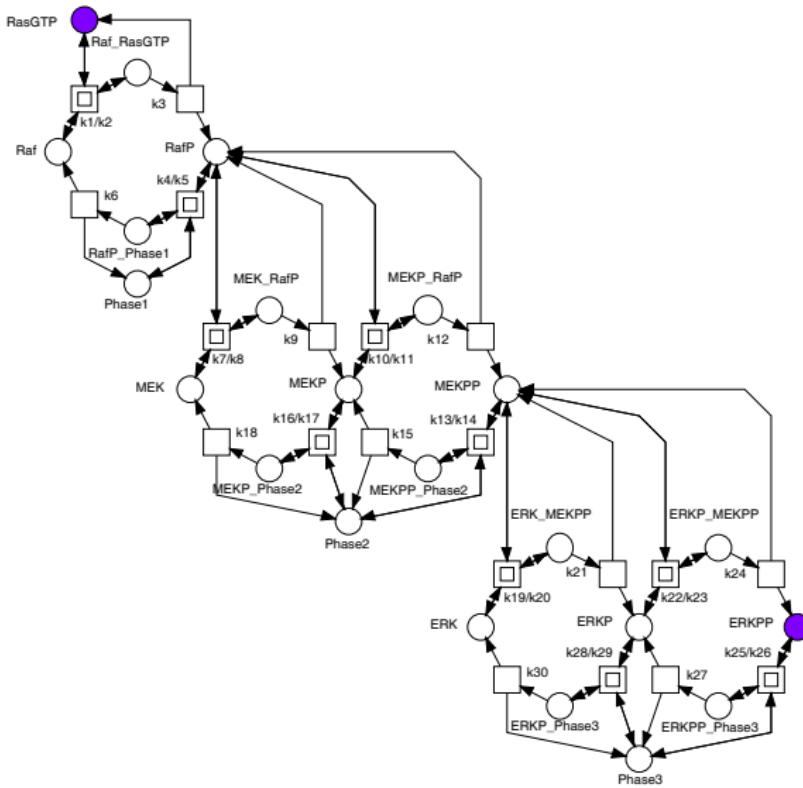
modelled in [Levchenko et al. 2000] like this ...

# Running Case Study - Origin

[Levchenko et al. 2000], *Supplemental Material : ODEs*

$$\begin{aligned} \frac{dRaf}{dt} &= k_2 * Raf\_RasGTP + k_6 * RafP\_Phase1 - k_1 * Raf * RasGTP \\ \frac{dRasGTP}{dt} &= k_2 * Raf\_RasGTP + k_3 * Raf\_RasGTP - k_1 * Raf * RasGTP \\ \frac{dRaf\_RasGTP}{dt} &= k_1 * Raf * RasGTP - k_2 * Raf\_RasGTP - k_3 * Raf\_RasGTP \\ \frac{dRafP}{dt} &= k_3 * Raf\_RasGTP + k_{12} * MEKP\_RafP + k_9 * MEK\_RafP + \\ &\quad k_5 * RafP\_Phase1 + k_8 * MEK\_RafP + k_{11} * MEKP\_RafP - \\ &\quad k_7 * RafP * MEK - k_{10} * MEKP * RafP - k_4 * Phase1 * RafP \\ \frac{dRafP\_Phase1}{dt} &= k_4 * Phase1 * RafP - k_5 * RafP\_Phase1 - k_6 * RafP\_Phase1 \\ \frac{dMEK\_RafP}{dt} &= k_7 * RafP * MEK - k_8 * MEK\_RafP - k_9 * MEK\_RafP \\ \frac{dMEKP\_RafP}{dt} &= k_{10} * MEKP * RafP - k_{11} * MEKP\_RafP - k_{12} * MEKP\_RafP \\ \frac{dMEKP\_Phase2}{dt} &= k_{16} * Phase2 * MEKP - k_{18} * MEKP\_Phase2 - k_{17} * MEKP\_Phase2 \\ \frac{dMEKPP\_Phase2}{dt} &= k_{13} * MEKPP * Phase2 - k_{15} * MEKPP\_Phase2 - k_{14} * MEKPP\_Phase2 \\ \frac{dERK}{dt} &= k_{20} * ERK\_MEKPP + k_{30} * ERKP\_Phase3 - k_{19} * MEKPP * ERK \\ \frac{dERK\_MEKPP}{dt} &= k_{19} * MEKPP * ERK - k_{20} * ERK\_MEKPP - k_{21} * ERK\_MEKPP \\ \frac{dERKP\_MEKPP}{dt} &= k_{22} * MEKPP * ERKP - k_{24} * ERKP\_MEKPP - k_{23} * ERKP\_MEKPP \\ \text{etcetera} &= \dots \end{aligned}$$

# Running Case Study



- initial marking construction

P-invariants

- subnetwork identification

- P-invariants : token preserving modules (*mass conservation*)
- T-invariants : state repeating modules (*elementary modes*)

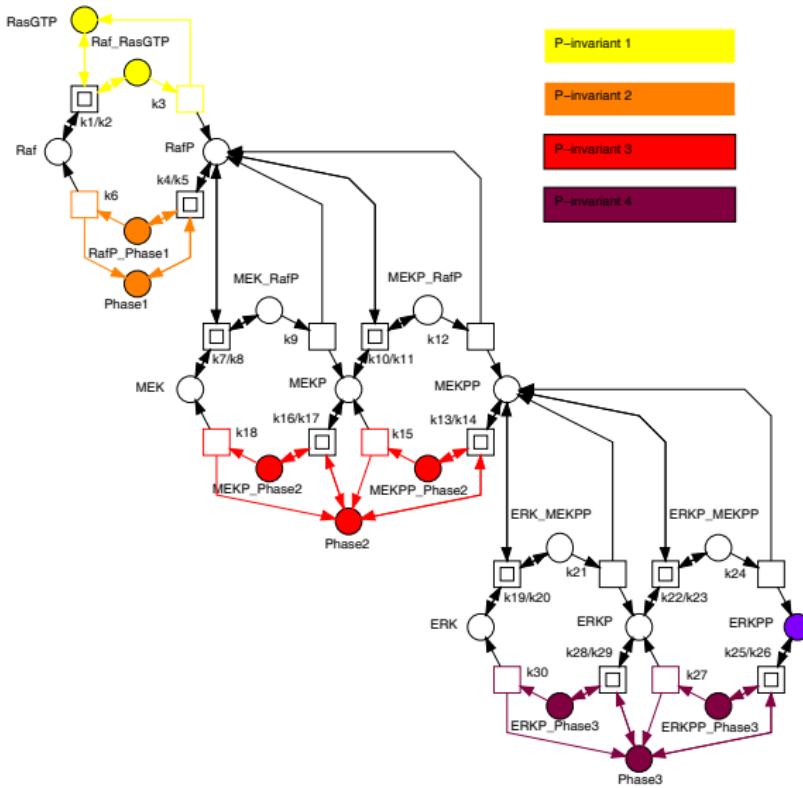
- general behavioural properties

- *boundedness* : every place gets finite token number only
- *liveness* : every transition may happen forever
- *reversibility* : every state may be reached forever

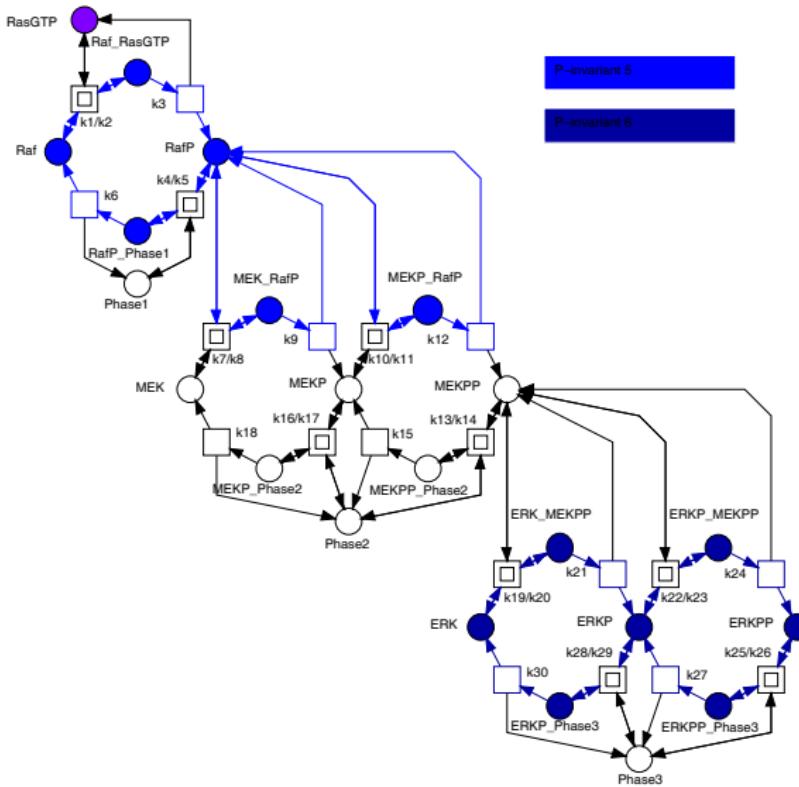
- special behavioural properties

CTL / LTL model checking

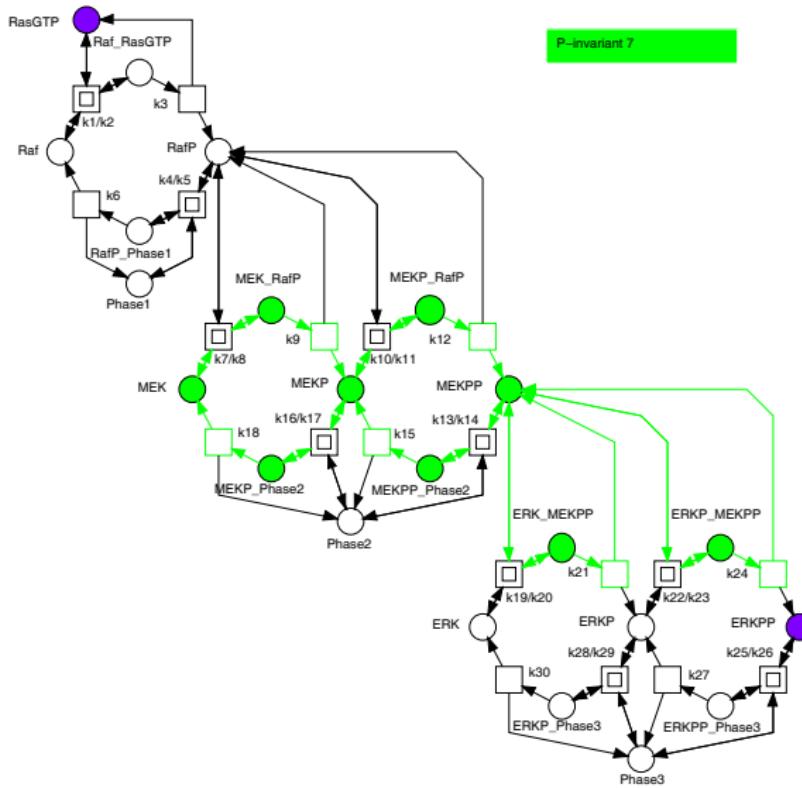
# Running Case Study - P-invariants



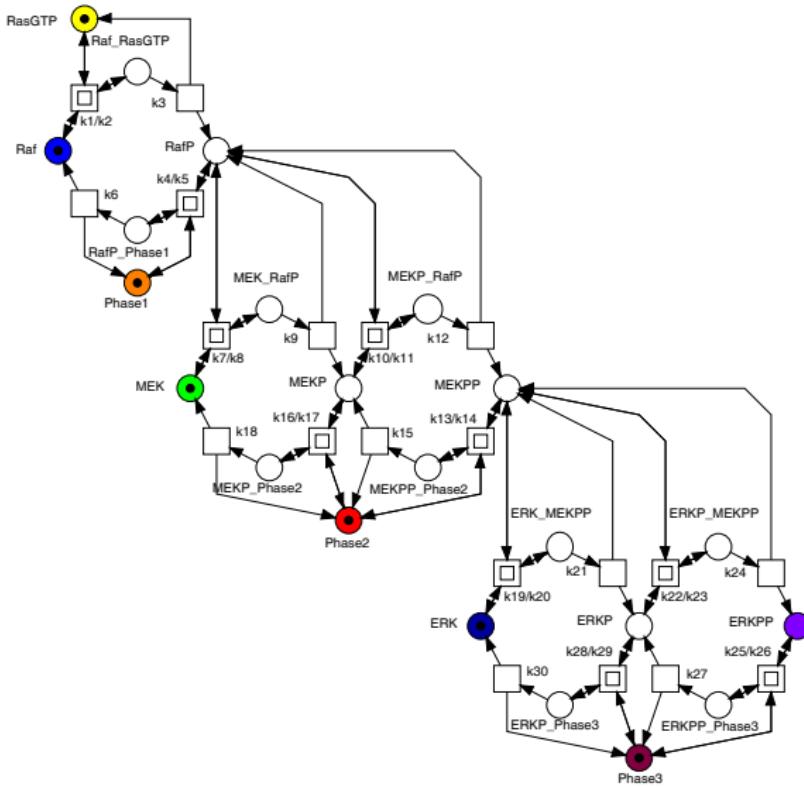
# Running Case Study - P-invariants



# Running Case Study - P-invariants



# Running Case Study - initial marking



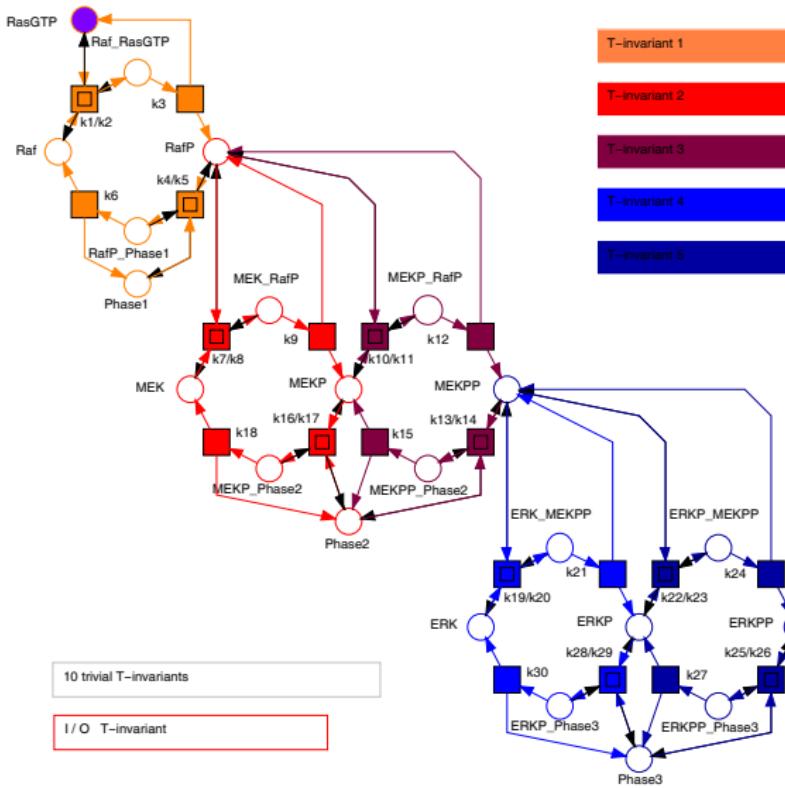
# Running Case Study - general properties

- *state space*

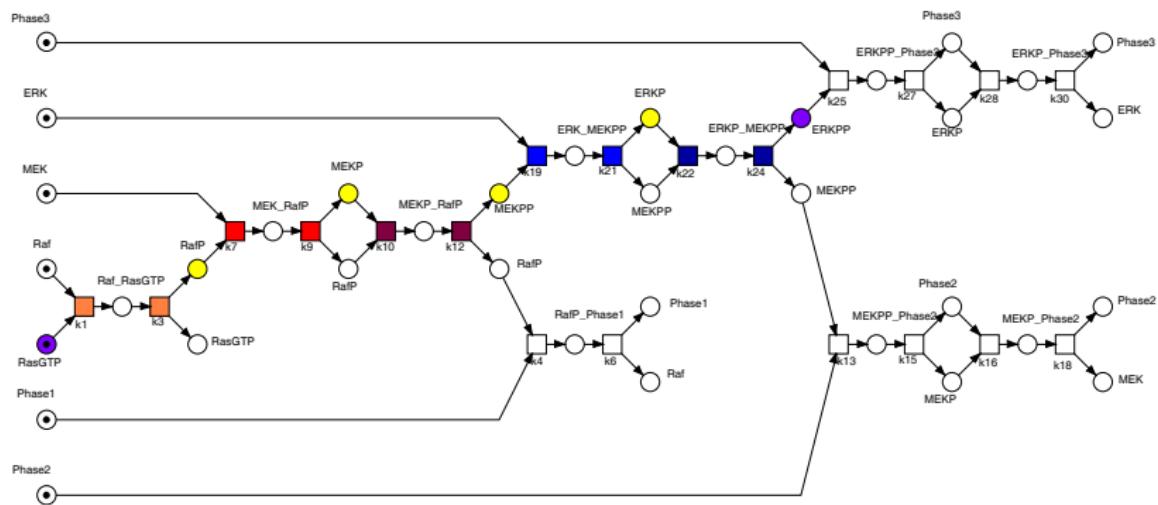
levels	reachability graph number of states	IDD data structure number of nodes
1	118	52
4	$2.4 \cdot 10^4$	115
8	$6.1 \cdot 10^6$	269
80	$5.6 \cdot 10^{18}$	13,472
120	$1.7 \cdot 10^{21}$	29,347

- Covered by P-invariants (CPI)  $\Rightarrow$  **bounded**
- Deadlock-Trap Property (DTP) holds  $\Rightarrow$  **no dead states**
- reachability graph
  - strongly connected  $\Rightarrow$  **reversible**
  - contains every transition (reaction)  $\Rightarrow$  **live**

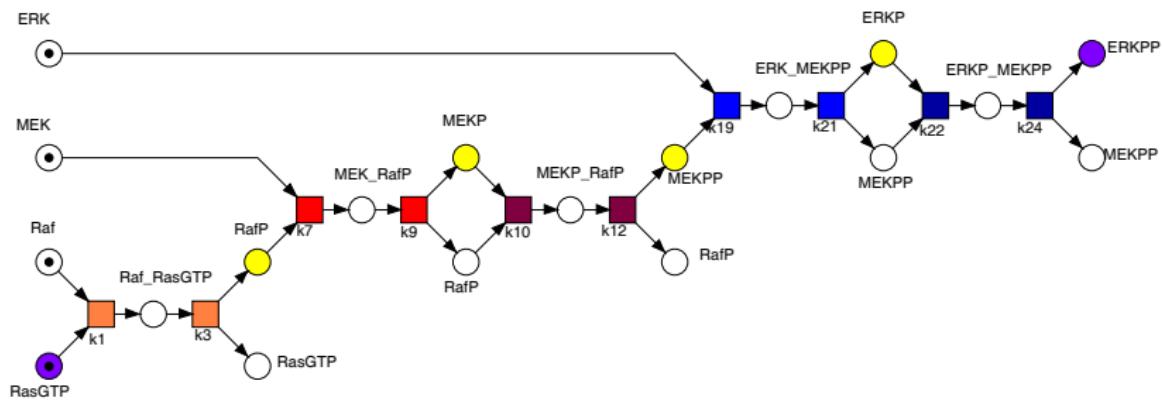
# Running Case Study - T-invariants

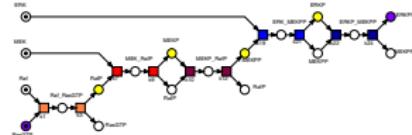


# Running Case Study - partial order run of I/O T-invariant



# Running Case Study - partial order run of I/O T-invariant





## property Q1 :

The signal sequence predicted by the partial order run of the I/O T-invariant is the only possible one;  
i.e., starting at the initial state, it is necessary to pass through RafP, MEKP, MEKPP and ERKP in order to reach ERKPP.

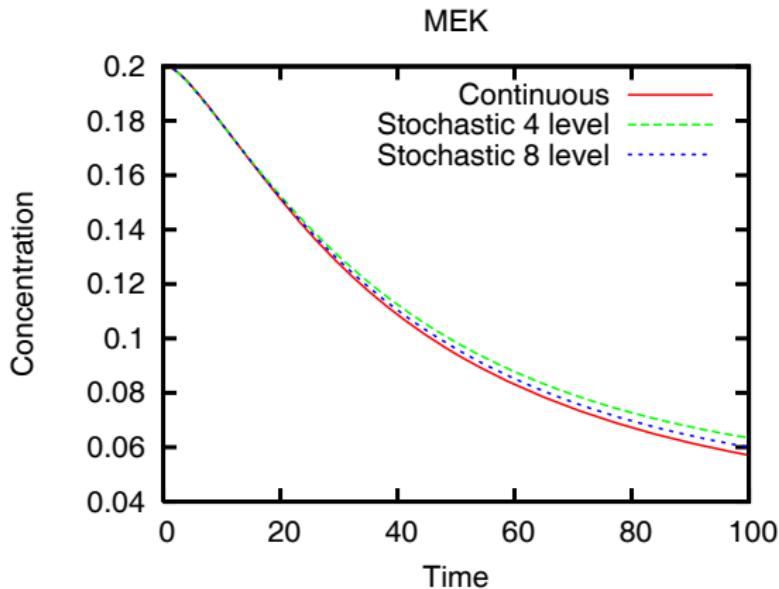
$$\neg [ \mathbf{E} (\neg \text{RafP} \mathbf{U} \text{MEKP}) \vee \\ \mathbf{E} (\neg \text{MEKP} \mathbf{U} \text{MEKPP}) \vee \\ \mathbf{E} (\neg \text{MEKPP} \mathbf{U} \text{ERKP}) \vee \\ \mathbf{E} (\neg \text{ERKP} \mathbf{U} \text{ERKPP}) ]$$

- *isomorphy of reachability graph and CTMC,*  
thus all qualitative properties still valid
- *How many levels needed for quantitative evaluation ?*
  - state space(1 levels) = 118 (Boolean interpretation)
  - state space(4 levels) = 24,065
  - state space(8 levels) = 6,110,643
- *equivalence check*

$$C_{RafP}(t) = \frac{0.1}{s} \cdot \underbrace{\sum_{i=1}^{4s} (i \cdot P(L_{RafP}(t) = i))}_{\text{expected value of } L_{RafP}(t)}$$

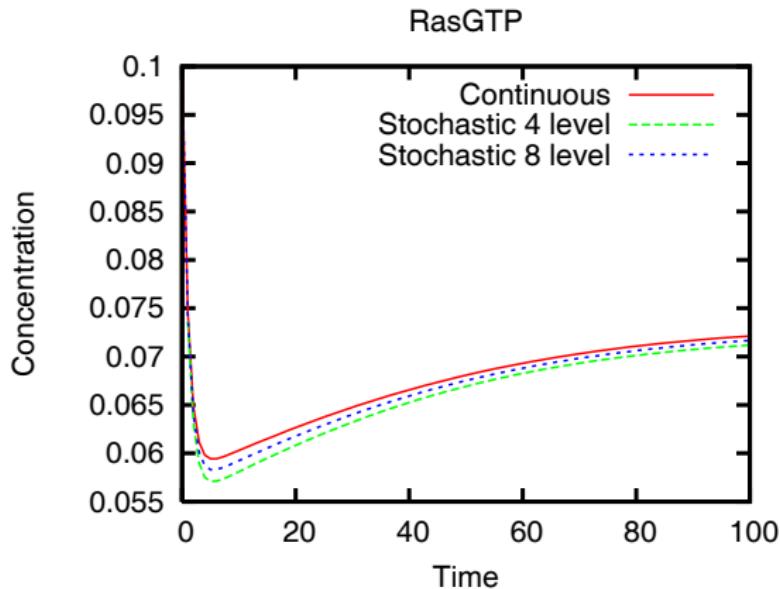
# Stochastic Model Checking - Preparation

- equivalence check, results, e.g. for MEK :



# Stochastic Model Checking - Preparation

- equivalence check, results, e.g. for RasGTP :

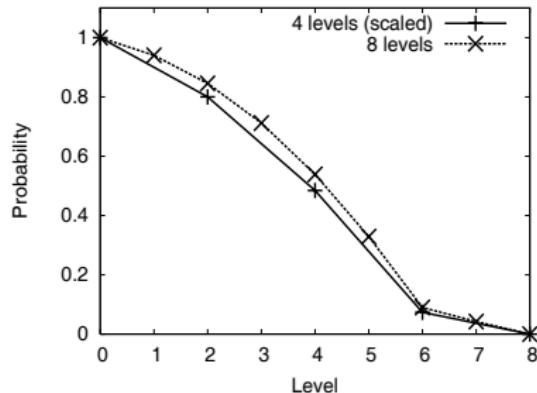


# Stochastic Model Checking (CSL)

## property S1 :

What is the probability of the concentration of RafP increasing, when starting in a state where the level is already at L ?

$$P_{=?} [ ( \text{RafP} = L ) \mathbf{U}^{<=100} ( \text{RafP} > L ) \{ \text{RafP} = L \} ]$$

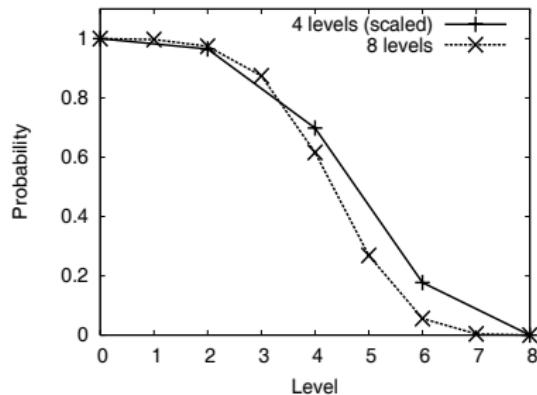


# Stochastic Model Checking (CSL)

**property S2 :**

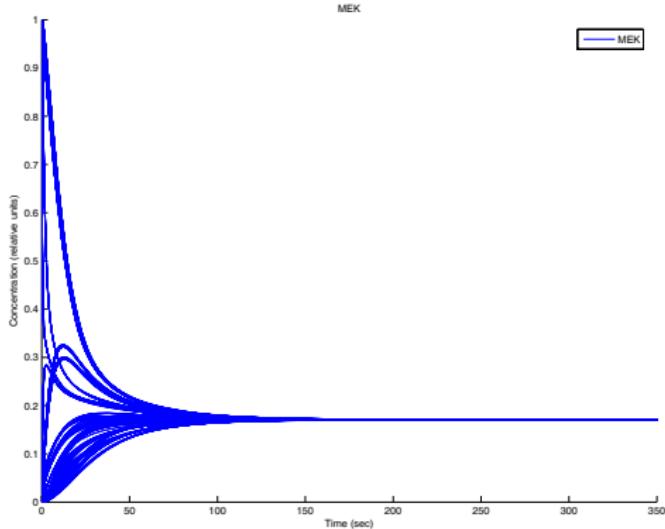
What is the probability that RafP is the first species to react?

$$\mathbf{P}_{=?} [ (( \text{MEKPP} = 0 ) \wedge ( \text{ERKPP} = 0 )) \mathbf{U}^{<=100} ( \text{RafP} > \text{L} ) \\ \{ ( \text{MEKPP} = 0 ) \wedge ( \text{ERKPP} = 0 ) \wedge ( \text{RafP} = 0 ) \} ]$$



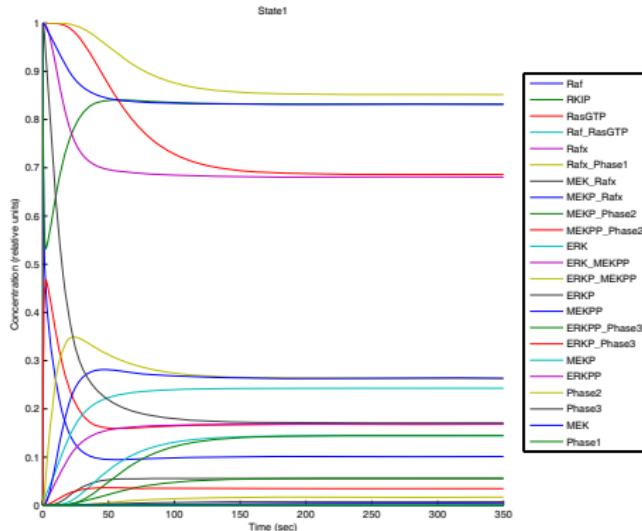
# Continuous Model Checking - Preparation

- steady state analysis, results for all 118 'good' states, e.g. for MEK :



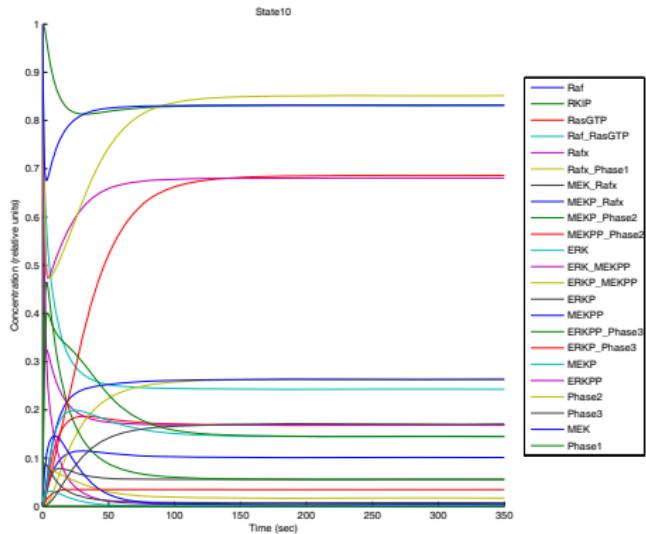
# Continuous Model Checking - Preparation

- steady state analysis for state 1 :



# Continuous Model Checking - Preparation

- steady state analysis for state 10 :

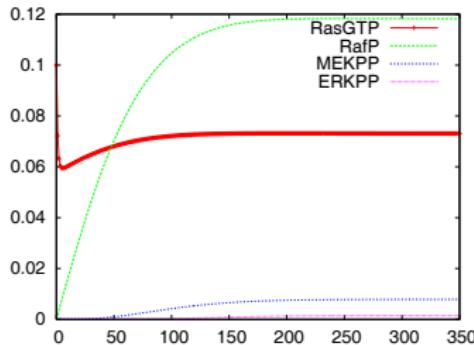


# Continuous Model Checking (LTLc)

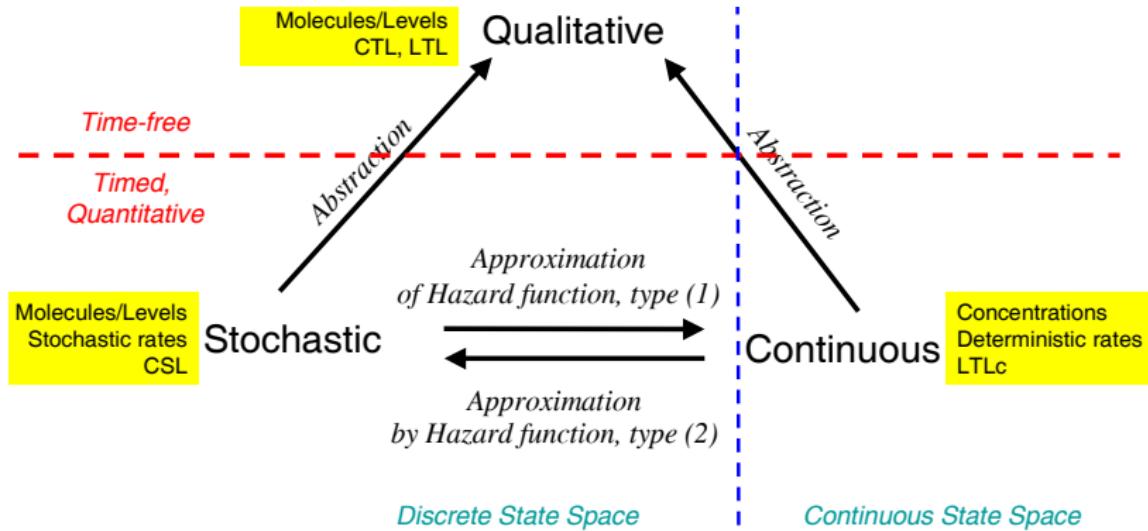
## property C1 :

The concentration of RafP rises to a significant level, while the concentrations of MEKPP and ERKPP remain close to zero ; i.e. RafP is really the first species to react.

$$((\text{MEKPP} < 0.001) \wedge (\text{ERKPP} < 0.0002)) \text{ U } (\text{RafP} > 0.06)$$



# Framework



- *model construction, animation, simulation*
  - Snoopy (*Cottbus*)
- *qualitative analysis*
  - Charlie (*Cottbus*), INA
  - BDD-CTL model checker (Boolean semantics) (*Cottbus*)
  - IDD-CTL model checker (integer semantics) (*Cottbus*)
- *stochastic analysis*
  - analytical model checking : PRISM/CSL
  - simulative model checking : MC2(PLTLc) (*Glasgow*)
- *continuous analysis*
  - MATLAB
  - BioNessie (*Glasgow*)
  - LTLc model checking : MC2(PLTLc) (*Glasgow*), BioCham

# end of part III

- all data files and analysis results available at  
[www-dssz.informatik.tu-cottbus.de/examples/levchenko](http://www-dssz.informatik.tu-cottbus.de/examples/levchenko)
- *laptop demonstration available*

# **BIO PETRI NETS, PART IV**

## **SUMMARY**

- Carl Adam Petri, 1962, PhD University of Technology Darmstadt  
-> *basic ideas introduced*
- early 1970's  
-> *first papers contributing to Petri net theory*
- Petri, 1976  
-> *application to chemical networks mentioned*
- early 1980's  
-> *first monographs on Petri net theory*
- Reddy, 1993  
-> *first paper on bio application*
- late 1990's  
-> *increasing interest for modelling and analysis of bio networks*



C. A. PETRI, NOVEMBER 2006



### ❑ representation of bio networks by Petri nets

- > *partial order representation*
- > *formal semantics*
- > *unifying view*
- > *better comprehension*
- > *sound analysis techniques*
- > *various abstraction levels*

### ❑ purposes

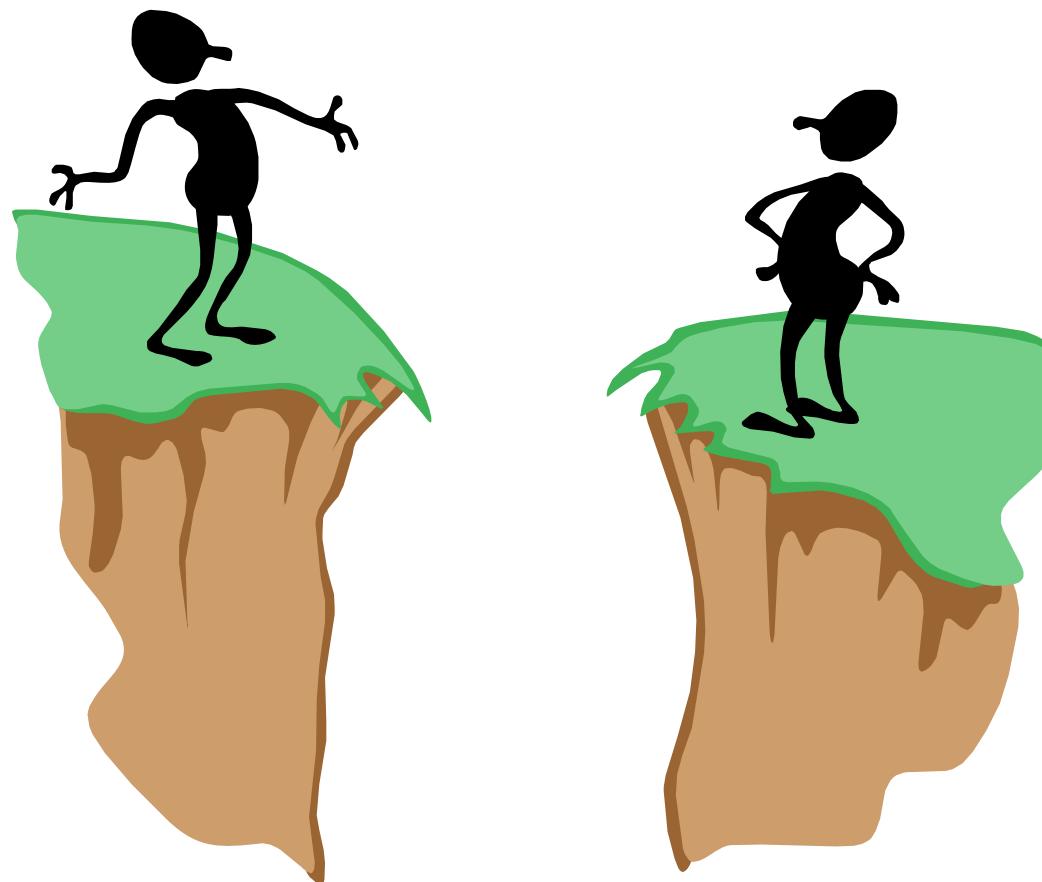
- > *animation*
- > *model validation against consistency criteria*
- > *qualitative / quantitative behaviour prediction*
- > *to experience the model*
- > *to increase confidence*
- > *experiment design, new insights*

### ❑ step-wise model development

- > *qualitative model*
- > *discrete quantitative model*
- > *continuous quantitative model*
- > *discrete Petri nets*
- > *stochastic Petri nets*
- > *continuous Petri nets = ODEs*

- **increasing level number = increasing accuracy**  
**BUT,** monotonous liveness holds for substructures only !
- **unbounded qualitative model + time = bounded model**  
**BUT,** that's not always the case !  
-> (structural) criteria for time-dependent boundedness ?
- **continuous behaviour = averaged stochastic behaviour**  
**BUT,** that's not always the case !  
-> stochastic and continuous behaviour may differ; why ? when ?
- **sharing structure = sharing properties**  
**BUT,** to which extend ?  
-> relation: qualitative & continuous behaviour ?

- M Heiner, D Gilbert, R Donaldson:  
**Petri Nets for Systems and Synthetic Biology**  
Springer LNCS 5016, pp. 215-264, 2008.
  
- R Breitling, D Gilbert, M Heiner, R Orton:  
A structured approach for the engineering of biochemical network models,  
illustrated for signalling pathways;  
Journal **Briefings in Bioinformatics**, accepted April 2008.
  
- D Gilbert, M Heiner, S Rosser, R Fulton, X Gu, M Trybilo:  
A Case Study in Model-driven Synthetic Biology;  
IFIP WCC 2008, 2nd IFIP Conference on Biologically Inspired Collaborative  
Computing (BICC 2008), Milano, Sept. 2008, to appear.



**THANKS !**  
**[HTTP://WWW-DSSZ.INFORMATIK.TU-COTTBUS.DE](http://WWW-DSSZ.INFORMATIK.TU-COTTBUS.DE)**