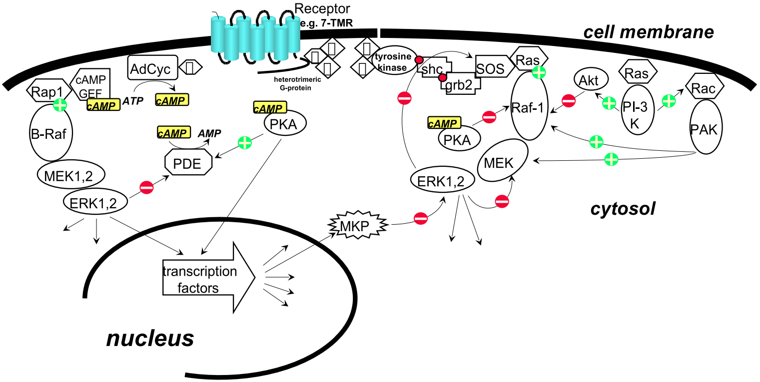
Petri Nets
for Systems and Synthetic Biology

Monika Heiner
@tu-cottbus.de

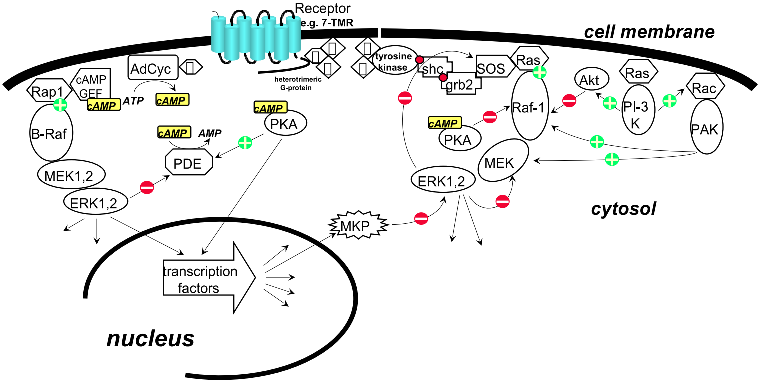
joint work with David Gilbert, Robin Donaldson
Bioinformatics Research Centre, University of Glasgow

Workshop on Computational Models for Cell Processes
Turku, May 27, 2008

- *. . . are networks of (bio-) chemical reactions*

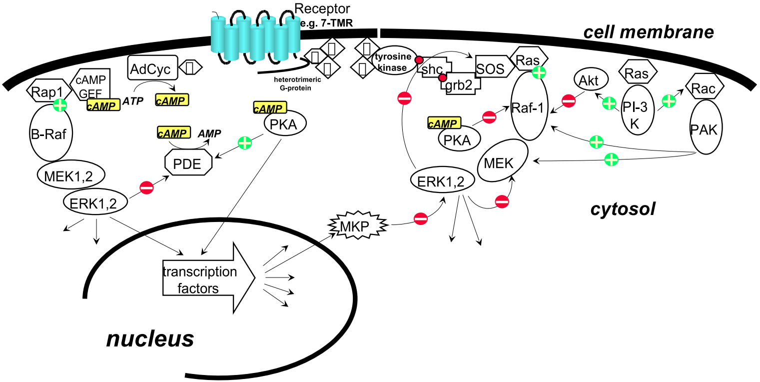


- . . . *are networks of (bio-) chemical reactions*



How to model this ?

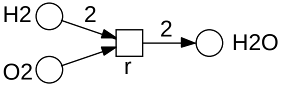
- ... *are networks of (bio-) chemical reactions*



How to model this ?

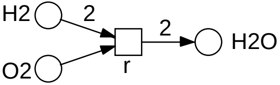
How to analyse this ?

- **bipartite - species & reactions** : r : $2H_2 + O_2 \rightarrow 2H_2O$

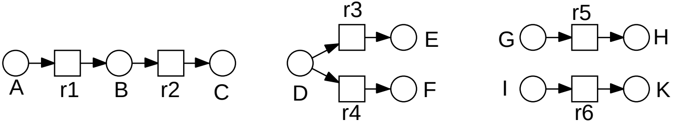


Biochemical Networks, Three Basic Properties

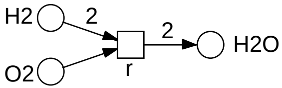
- bipartite - species & reactions : r : $2H_2 + O_2 \rightarrow 2H_2O$



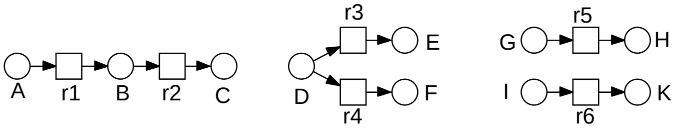
- reactions - sequential, alternative, concurrent



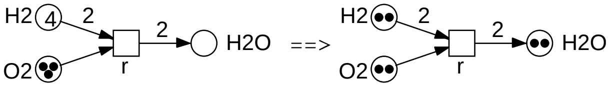
- **bipartite - species & reactions** : r : $2H_2 + O_2 \rightarrow 2H_2O$



- **reactions - sequential, alternative, concurrent**



- **behaviour - stochastic**

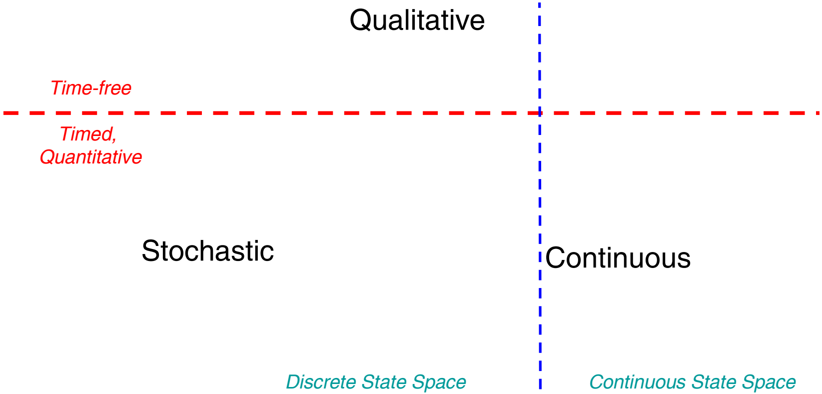


Framework

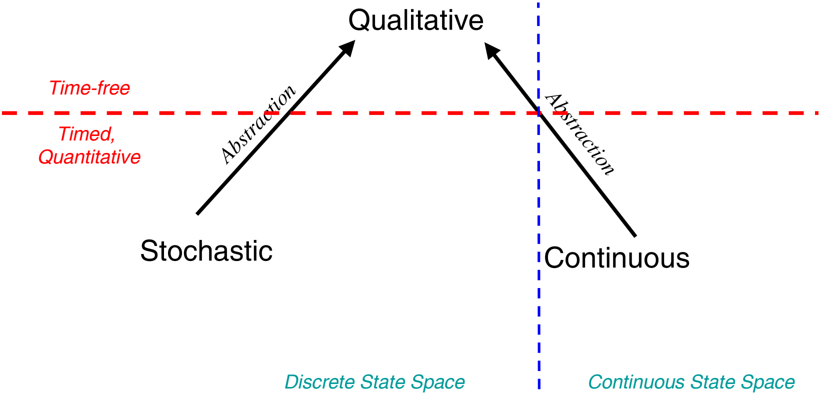
Qualitative

Stochastic

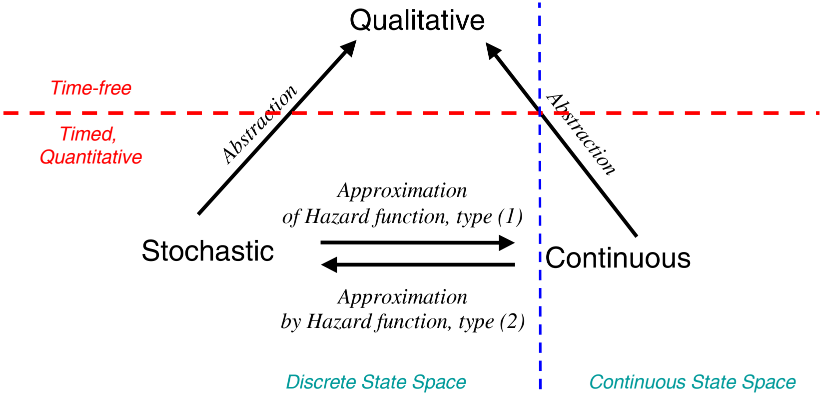
Continuous


Qualitative
Time-free
Timed,
Quantitative
Stochastic
Continuous
Discrete State Space
Continuous State Space

Framework


Qualitative
Time-free
Timed,
Quantitative
Abstraction
Abstraction
Stochastic
Continuous
Discrete State Space
Continuous State Space

Framework



Qualitative
Time-free
Timed,
Quantitative
Abstraction
Abstraction
Approximation
of Hazard function, type (1)
Stochastic
Continuous
Approximation
by Hazard function, type (2)
Discrete State Space
Continuous State Space

Definition :

A **place/transition Petri net** is a quadruple $\mathcal{PN} = (P, T, f, m_0)$, where

- P, T - finite, non empty, disjoint sets *(places, transitions)*

(Qualitative) Petri net

Definition :

A **place/transition Petri net** is a quadruple
$\mathcal{PN} = (P, T, f, m_0)$, 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)**

Definition :
A **place/transition Petri net** is a quadruple
$\mathcal{PN} = (P, T, f, m_0)$, where

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

Definition :

A **place/transition Petri net** is a quadruple
$\mathcal{PN} = (P, T, f, m_0)$, where

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

Interleaving Semantics : reachability graph / CTL, LTL

Stochastic Petri net

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)***
- $f : ((P \times T) \cup (T \times P)) \to \mathbb{N}_0$ ***(weighted directed arcs)***
- $m_0 : P \to \mathbb{N}_0$ ***(initial marking)***

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
  - $H := \bigcup_{t \in T} \left\{ h_t \mid h_t : \mathbb{N}_0^{|{}^\bullet t|} \rightarrow \mathbb{R}^+ \right\}$
  - $v(t) = h_t$ for all transitions $t \in T$

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

Definition :

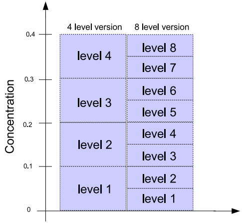
A **biochemically interpreted continuous Petri net** is a quintuple $\mathcal{CPN}_{Bio} = (P, T, f, v, m_0)$, where

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

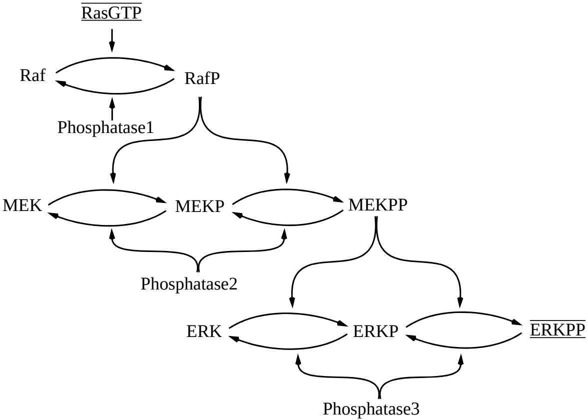
$$h_t := c_t \cdot \prod_{p \in {}^{\bullet}t} \binom{m(p)}{f(p,t)} \tag{1}$$

- *concentration levels semantics*

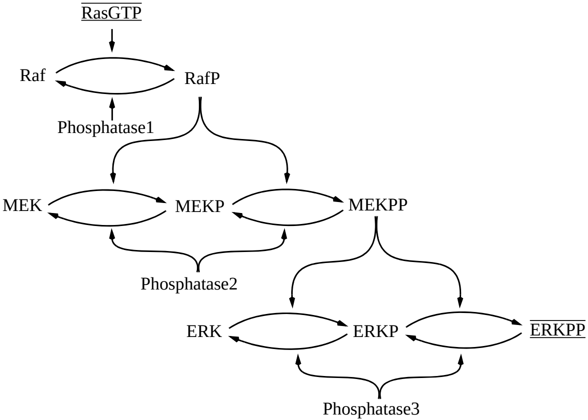
$$h_t := k_t \cdot N \cdot \prod_{p \in {}^{\bullet}t} \left(\frac{m(p)}{N}\right) \tag{2}$$

Running Case Study

- *. . . a typical signalling cascade*



- *. . . a typical signalling cascade*



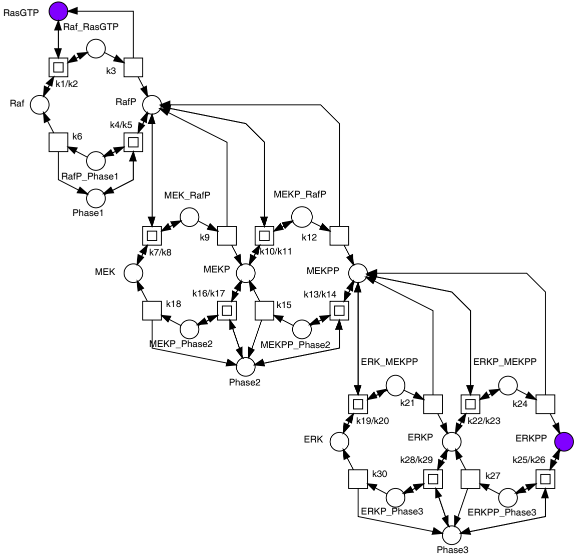
modelled in [Levchenko et al. 2000] like this . . .

Running Case Study - Origin

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

$$
\begin{aligned}
d\mathrm{Raf}/dt &= k_2 * \mathrm{Raf\_RasGTP} + k_6 * \mathrm{RafP\_Phase1} - k_1 * \mathrm{Raf} * \mathrm{RasGTP} \\
d\mathrm{RasGTP}/dt &= k_2 * \mathrm{Raf\_RasGTP} + k_3 * \mathrm{Raf\_RasGTP} - k_1 * \mathrm{Raf} * \mathrm{RasGTP} \\
d\mathrm{Raf\_RasGTP}/dt &= k_1 * \mathrm{Raf} * \mathrm{RasGTP} - k_2 * \mathrm{Raf\_RasGTP} - k_3 * \mathrm{Raf\_RasGTP} \\
d\mathrm{RafP}/dt &= k_3 * \mathrm{Raf\_RasGTP} + k_{12} * \mathrm{MEKP\_RafP} + k_9 * \mathrm{MEK\_RafP} + \\
&\quad k_5 * \mathrm{RafP\_Phase1} + k_8 * \mathrm{MEK\_RafP} + k_{11} * \mathrm{MEKP\_RafP} - \\
&\quad k_7 * \mathrm{RafP} * \mathrm{MEK} - k_{10} * \mathrm{MEKP} * \mathrm{RafP} - k_4 * \mathrm{Phase1} * \mathrm{RafP} \\
d\mathrm{RafP\_Phase1}/dt &= k_4 * \mathrm{Phase1} * \mathrm{RafP} - k_5 * \mathrm{RafP\_Phase1} - k_6 * \mathrm{RafP\_Phase1} \\
d\mathrm{MEK\_RafP}/dt &= k_7 * \mathrm{RafP} * \mathrm{MEK} - k_8 * \mathrm{MEK\_RafP} - k_9 * \mathrm{MEK\_RafP} \\
d\mathrm{MEKP\_RafP}/dt &= k_{10} * \mathrm{MEKP} * \mathrm{RafP} - k_{11} * \mathrm{MEKP\_RafP} - k_{12} * \mathrm{MEKP\_RafP} \\
d\mathrm{MEKP\_Phase2}/dt &= k_{16} * \mathrm{Phase2} * \mathrm{MEKP} - k_{18} * \mathrm{MEKP\_Phase2} - k_{17} * \mathrm{MEKP\_Phase2} \\
d\mathrm{MEKPP\_Phase2}/dt &= k_{13} * \mathrm{MEKPP} * \mathrm{Phase2} - k_{15} * \mathrm{MEKPP\_Phase2} - k_{14} * \mathrm{MEKPP\_Phase} \\
d\mathrm{ERK}/dt &= k_{20} * \mathrm{ERK\_MEKPP} + k_{30} * \mathrm{ERKP\_Phase3} - k_{19} * \mathrm{MEKPP} * \mathrm{ERK} \\
d\mathrm{ERK\_MEKPP}/dt &= k_{19} * \mathrm{MEKPP} * \mathrm{ERK} - k_{20} * \mathrm{ERK\_MEKPP} - k_{21} * \mathrm{ERK\_MEKPP} \\
d\mathrm{ERKP\_MEKPP}/dt &= k_{22} * \mathrm{MEKPP} * \mathrm{ERKP} - k_{24} * \mathrm{ERKP\_MEKPP} - k_{23} * \mathrm{ERKP\_MEKPP} \\
\mathit{etcetera} &= \ldots
\end{aligned}
$$

Running Case Study



- **initial marking construction**
  P-invariants

Qualitative Analysis

- **initial marking construction**
  P-invariants

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

Qualitative Analysis

- **initial marking construction**
  P-invariants

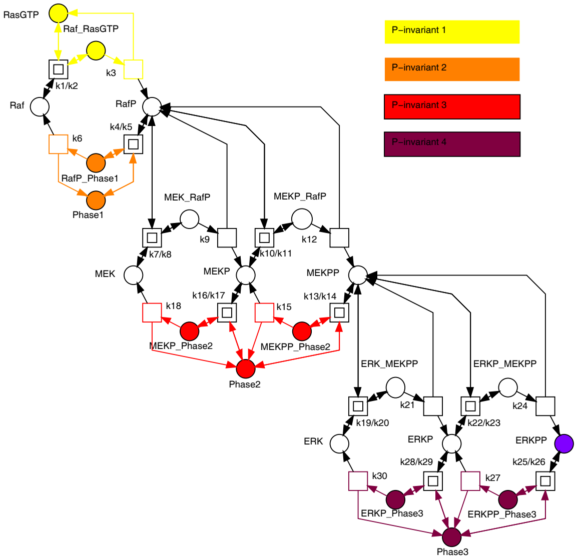
- **subnetwork identification**
  - P-invariants : token preserving modules (*mass conservation*)
  - T-invariants : state repeating modules (*elementary modes*)

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

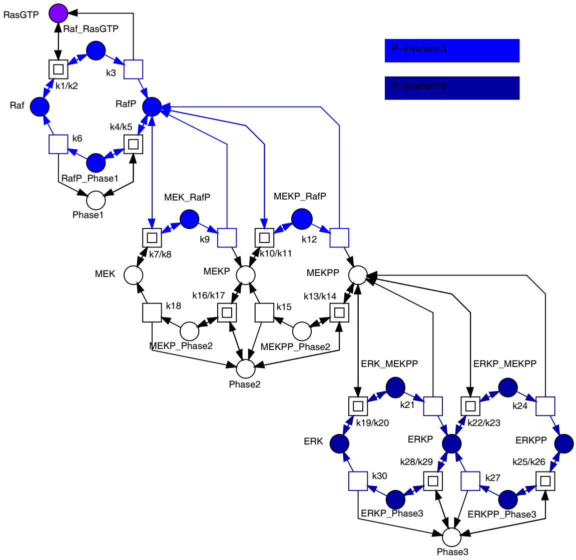
Qualitative Analysis

- **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

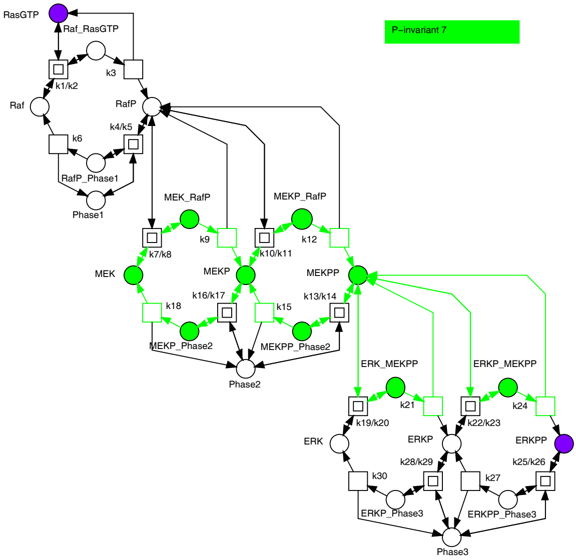
Running Case Study - P-invariants



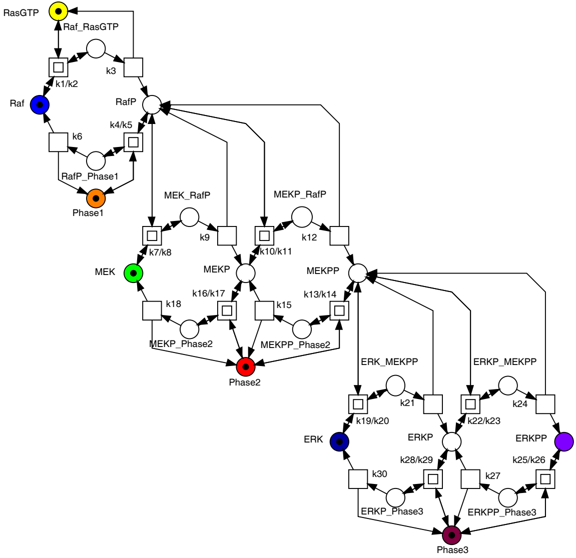
Running Case Study - P-invariants



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

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**

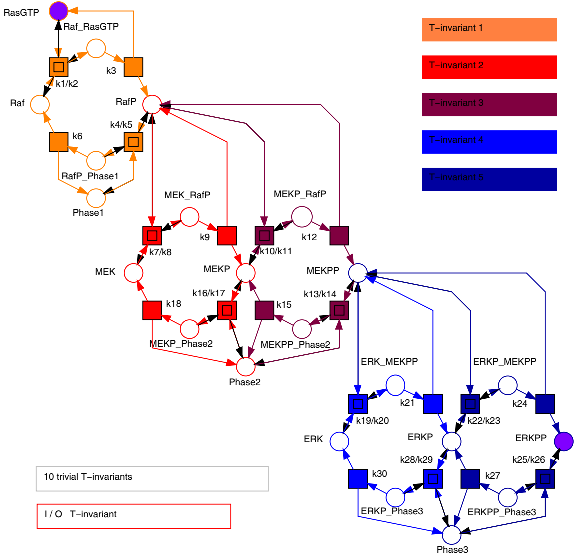
Running Case Study - general properties

- *state space*

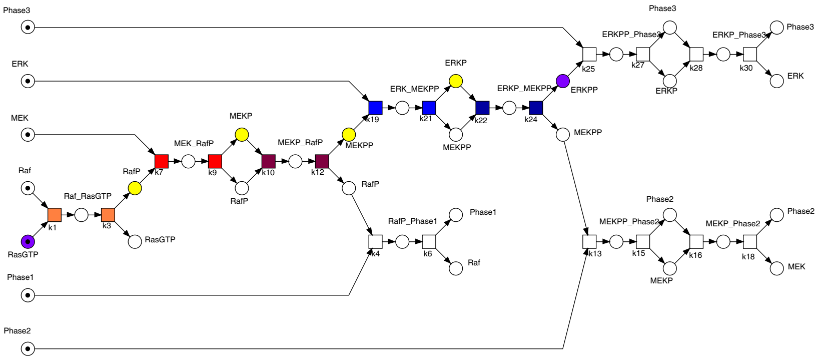
levels	reachability graph number of states	IDD data structure number of nodes
1	118	52
4	$2.4 \cdot 10^4$	115
8	$6.1 \cdot 10^6$	269
80	$5.6 \cdot 10^{18}$	13,472
120	$1.7 \cdot 10^{21}$	29,347

- 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**

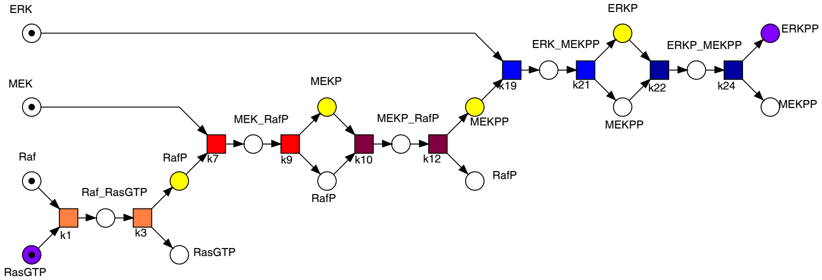
Running Case Study - T-invariants

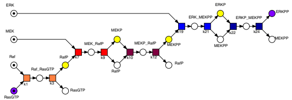


Running Case Study - partial order run of I/O T-invariant



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.

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

Stochastic Model Checking - Preparation

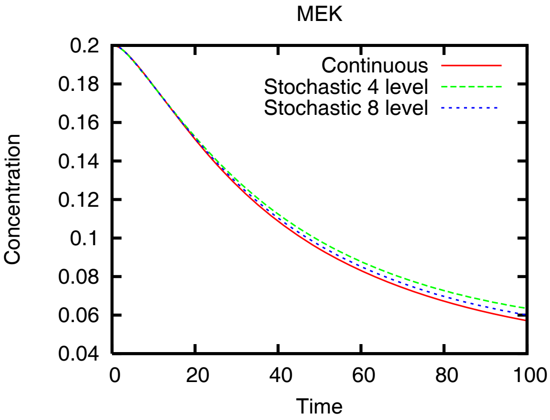
- *isomorphy of reachability graph and CTMC,*
  thus all qualitative properties still valid

- *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

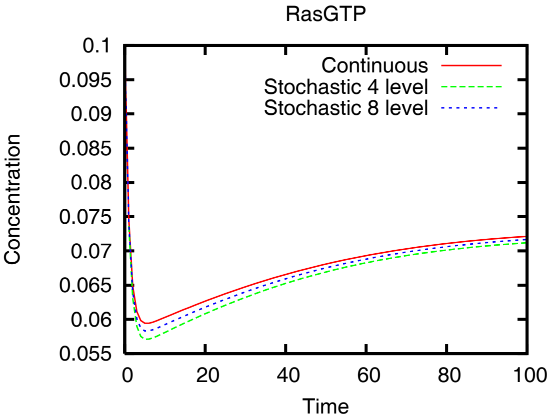
- *equivalence check*

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

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



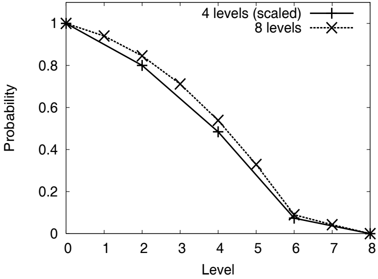
- *equivalence check, results, e.g. for RasGTP :*



property S1 :

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

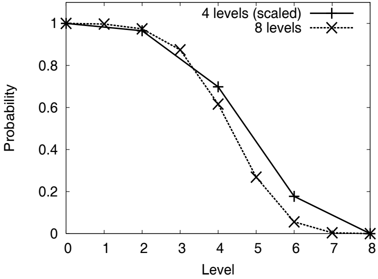
$$\mathbf{P}_{=?} \; [\; (\; \mathrm{RafP} \; = \; \mathrm{L} \;) \; \mathbf{U}^{<=100} \; (\; \mathrm{RafP} > \mathrm{L} \;) \{ \; \mathrm{RafP} = \mathrm{L} \; \} \;]$$



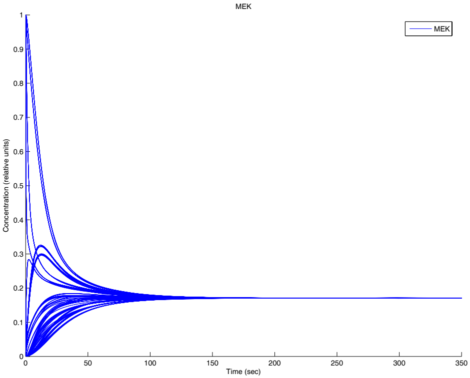
property S2 :

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

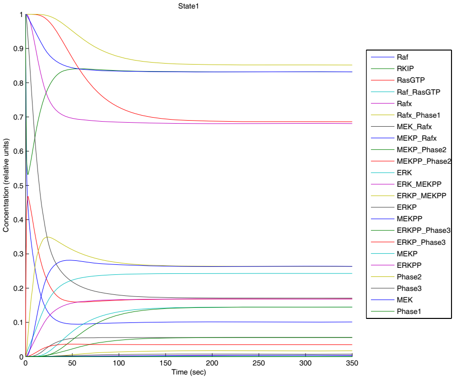
$$P_{=?}\,[\,(\,(\,MEKPP\ =\ 0\,)\wedge(\,ERKPP\ =\ 0\,)\,)\ \mathbf{U}^{<=100}\,(\,RafP > \ L\,)\\ \{\,(\,MEKPP\ =\ 0\,)\ \wedge\ (\,ERKPP\ =\ 0\,)\ \ \wedge\ (\,RafP =\ 0\,)\,\}\ ]$$



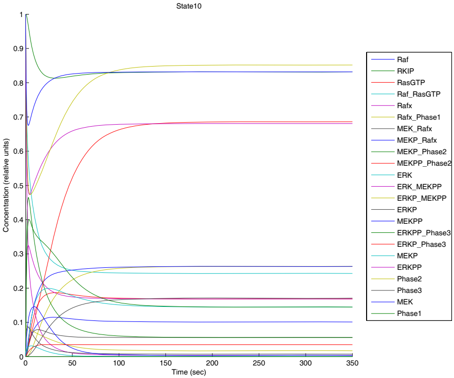
- *steady state analysis, results for all 118 'good' states, e.g. for MEK :*



- *steady state analysis for state 1 :*



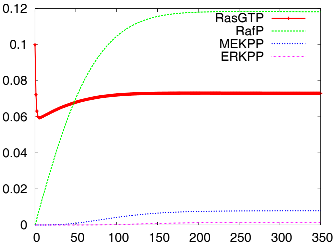
- *steady state analysis for state 10 :*



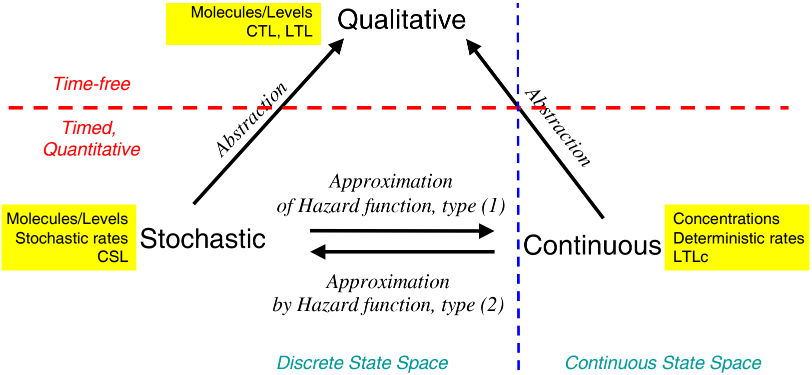
property C1 :

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.

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



Framework



Molecules/Levels
CTL, LTL
Qualitative
Time-free
Timed,
Quantitative
Abstraction
Abstraction
Approximation
of Hazard function, type (1)
Molecules/Levels
Stochastic rates
CSL
Stochastic
Continuous
Concentrations
Deterministic rates
LTLc
Approximation
by Hazard function, type (2)
Discrete State Space
Continuous State Space

Tools

- *model construction, animation, simulation*
  - Snoopy *(Cottbus)*

Tools

- *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)*

Tools

- *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)*

Tools

- *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

Summary

- *unifying framework*
  qualitative & stochastic & continuous paradigms

Summary

- *unifying framework*
  qualitative & stochastic & continuous paradigms
- *three models sharing structure*
  - qualitative Petri nets $\longrightarrow$ time-free analyses
  - stochastic Petri nets $\longrightarrow$ CTMC
  - continuous Petri nets $\longrightarrow$ ODEs

Summary

- *unifying framework*
  qualitative & stochastic & continuous paradigms
- *three models sharing structure*
  - qualitative Petri nets $\longrightarrow$ time-free analyses
  - stochastic Petri nets $\longrightarrow$ CTMC
  - continuous Petri nets $\longrightarrow$ ODEs
- *running case study*
  ERK signalling pathway

Summary

- *unifying framework*
  qualitative & stochastic & continuous paradigms
- *three models sharing structure*
  - qualitative Petri nets → time-free analyses
  - stochastic Petri nets → CTMC
  - continuous Petri nets → ODEs
- *running case study*
  ERK signalling pathway
- *focus - transient analysis*, esp. by
  - transition invariants & partial order run
  - qualitative & stochastic & continuous model checking

Summary

- *unifying framework*
  qualitative & stochastic & continuous paradigms
- *three models sharing structure*
  - qualitative Petri nets → time-free analyses
  - stochastic Petri nets → CTMC
  - continuous Petri nets → ODEs
- *running case study*
  ERK signalling pathway
- *focus - transient analysis*, esp. by
  - transition invariants & partial order run
  - qualitative & stochastic & continuous model checking
- **not bound to the Petri net perspective**

Some Open Problems

- *increasing level number = increasing accuracy*
  BUT, monotonous liveness holds for substructures (EFC) only !

Some Open Problems

- *increasing level number = increasing accuracy*
  BUT, monotonous liveness holds for substructures (EFC) only !
- *unbounded qualitative model + time = bounded model*
  BUT, that's not always the case !
  (structural) criteria for time-dependent boundedness ?

Some Open Problems

- *increasing level number = increasing accuracy*
  BUT, monotonous liveness holds for substructures (EFC) 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 ?

Some Open Problems

- *increasing level number = increasing accuracy*
  BUT, monotonous liveness holds for substructures (EFC) 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 properties & continuous behaviour ?

Thanks !

- all data files and analysis results available at www-dssz.informatik.tu-cottbus.de/examples/levchenko
- M Heiner, D Gilbert, R Donaldson : Petri Nets for Systems and Synthetic Biology ; SFM 2008, Springer LNCS 5016, pp. 215-264, 2008.
- *laptop demonstration available*