BME & PN,
A Case Study - Signalling Cascade

Monika Heiner
@b-tu.de

joint work with David Gilbert, Robin Donaldson
University of Glasgow/Brunel University London

Ca' Foscari University Venezia, July, 2016

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

Semantics : Continuous Time Markov Chain / CSL, PLTLc

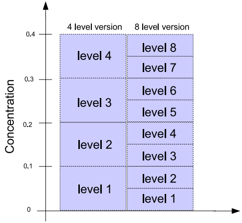
Definition :
A **biochemically interpreted continuous Petri net** is a quintuple $\mathcal{CPN}_{Bio} = (P, T, f, v, m_0)$, where

- 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
  - $H := \bigcup_{t \in T} \{ h_t \mid h_t : \mathbb{R}^{|{}^\bullet t|} \rightarrow \mathbb{R}^+ \}$
  - $v(t) = h_t$ 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*

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

- *concentration levels semantics*

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

Temporal logics to formally express properties

Various logics each with different expressivity :

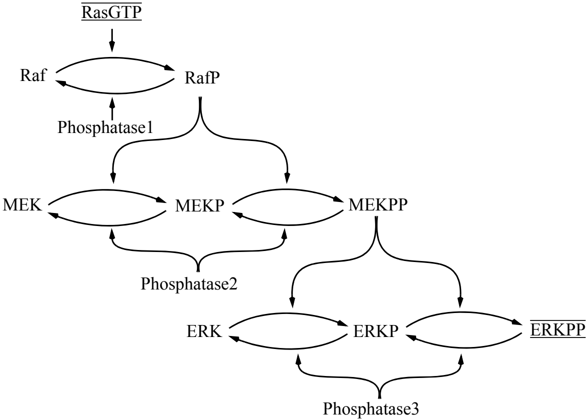
- *Branching-time logics* consider all branching time lines
  - Computational Tree Logic (CTL)
  - Continuous Stochastic Logic (CSL)

  "There is a possibility that I will stay hungry forever."
  "There is a possibility that eventually I am no longer hungry."

- *Linear-time logics* consider separately all single time lines
  - Linear-time Temporal Logics (LTL, LTLc, PLTLc)

  "I am hungry."
  "I am always hungry."
  "I will eventually be hungry."
  "I will be hungry until I eat something."

- *. . . a typical signalling cascade*



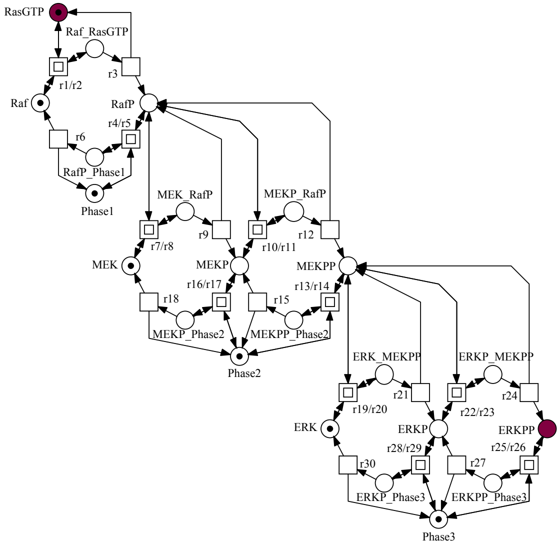
modelled in [Levchenko et al. 2000] like this . . .

Running Case Study - Origin

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

$$
\begin{aligned}
d\mathrm{Raf}/dt &= \mathrm{k}_2 * \mathrm{Raf\_RasGTP} + \mathrm{k}_6 * \mathrm{RafP\_Phase1} - \mathrm{k}_1 * \mathrm{Raf} * \mathrm{RasGTP} \\
d\mathrm{RasGTP}/dt &= \mathrm{k}_2 * \mathrm{Raf\_RasGTP} + \mathrm{k}_3 * \mathrm{Raf\_RasGTP} - \mathrm{k}_1 * \mathrm{Raf} * \mathrm{RasGTP} \\
d\mathrm{Raf\_RasGTP}/dt &= \mathrm{k}_1 * \mathrm{Raf} * \mathrm{RasGTP} - \mathrm{k}_2 * \mathrm{Raf\_RasGTP} - \mathrm{k}_3 * \mathrm{Raf\_RasGTP} \\
d\mathrm{RafP}/dt &= \mathrm{k}_3 * \mathrm{Raf\_RasGTP} + \mathrm{k}_{12} * \mathrm{MEKP\_RafP} + \mathrm{k}_9 * \mathrm{MEK\_RafP} + \\
&\quad \mathrm{k}_5 * \mathrm{RafP\_Phase1} + \mathrm{k}_8 * \mathrm{MEK\_RafP} + \mathrm{k}_{11} * \mathrm{MEKP\_RafP} - \\
&\quad \mathrm{k}_7 * \mathrm{RafP} * \mathrm{MEK} - \mathrm{k}_{10} * \mathrm{MEKP} * \mathrm{RafP} - \mathrm{k}_4 * \mathrm{Phase1} * \mathrm{RafP} \\
d\mathrm{RafP\_Phase1}/dt &= \mathrm{k}_4 * \mathrm{Phase1} * \mathrm{RafP} - \mathrm{k}_5 * \mathrm{RafP\_Phase1} - \mathrm{k}_6 * \mathrm{RafP\_Phase1} \\
d\mathrm{MEK\_RafP}/dt &= \mathrm{k}_7 * \mathrm{RafP} * \mathrm{MEK} - \mathrm{k}_8 * \mathrm{MEK\_RafP} - \mathrm{k}_9 * \mathrm{MEK\_RafP} \\
d\mathrm{MEKP\_RafP}/dt &= \mathrm{k}_{10} * \mathrm{MEKP} * \mathrm{RafP} - \mathrm{k}_{11} * \mathrm{MEKP\_RafP} - \mathrm{k}_{12} * \mathrm{MEKP\_RafP} \\
d\mathrm{MEKP\_Phase2}/dt &= \mathrm{k}_{16} * \mathrm{Phase2} * \mathrm{MEKP} - \mathrm{k}_{18} * \mathrm{MEKP\_Phase2} - \mathrm{k}_{17} * \mathrm{MEKP\_Phase2} \\
d\mathrm{MEKPP\_Phase2}/dt &= \mathrm{k}_{13} * \mathrm{MEKPP} * \mathrm{Phase2} - \mathrm{k}_{15} * \mathrm{MEKPP\_Phase2} - \mathrm{k}_{14} * \mathrm{MEKPP\_Phase} \\
d\mathrm{ERK}/dt &= \mathrm{k}_{20} * \mathrm{ERK\_MEKPP} + \mathrm{k}_{30} * \mathrm{ERKP\_Phase3} - \mathrm{k}_{19} * \mathrm{MEKPP} * \mathrm{ERK} \\
d\mathrm{ERK\_MEKPP}/dt &= \mathrm{k}_{19} * \mathrm{MEKPP} * \mathrm{ERK} - \mathrm{k}_{20} * \mathrm{ERK\_MEKPP} - \mathrm{k}_{21} * \mathrm{ERK\_MEKPP} \\
d\mathrm{ERKP\_MEKPP}/dt &= \mathrm{k}_{22} * \mathrm{MEKPP} * \mathrm{ERKP} - \mathrm{k}_{24} * \mathrm{ERKP\_MEKPP} - \mathrm{k}_{23} * \mathrm{ERKP\_MEKPP} \\
\mathit{etcetera} &= \ldots
\end{aligned}
$$

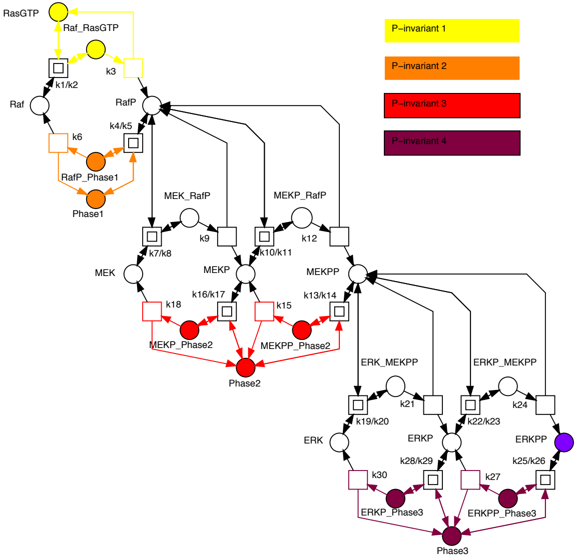
Running Case Study



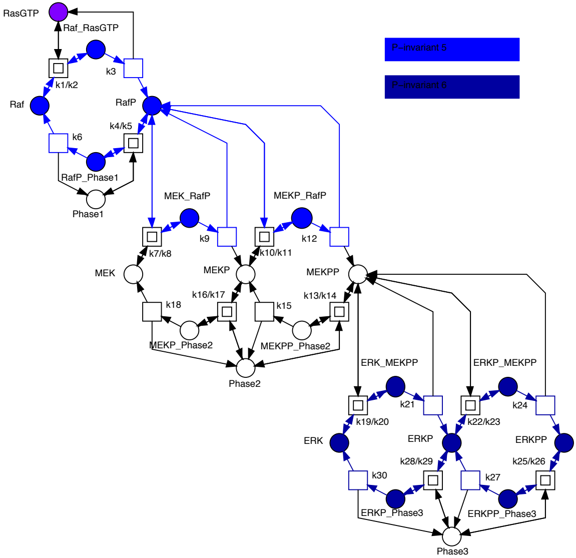
Qualitative Analysis

- **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

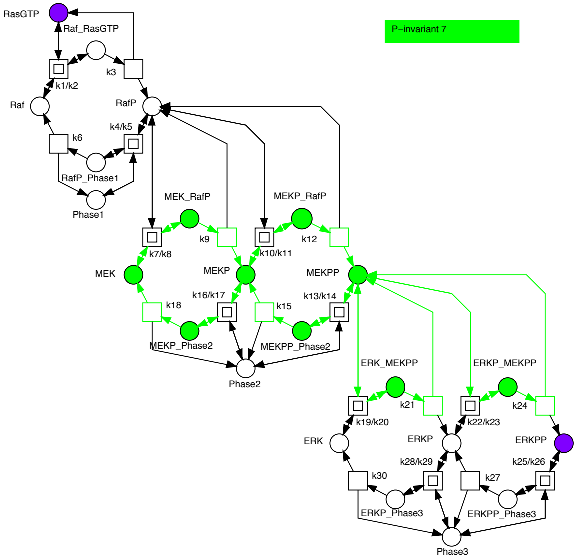
Running Case Study - P-invariants



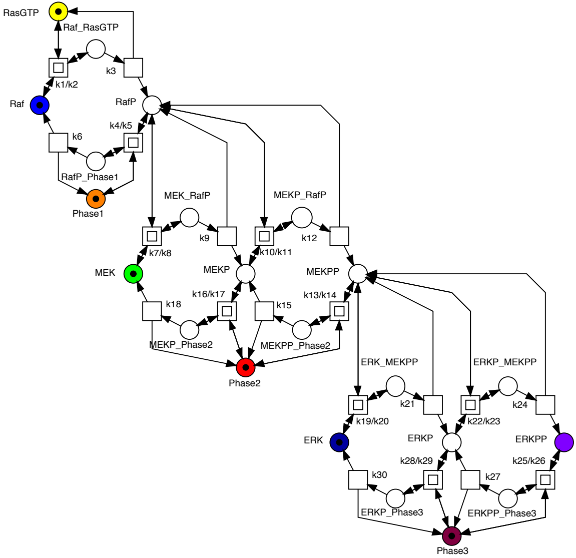
Running Case Study - P-invariants



Running Case Study - P-invariants



Running Case Study - initial marking



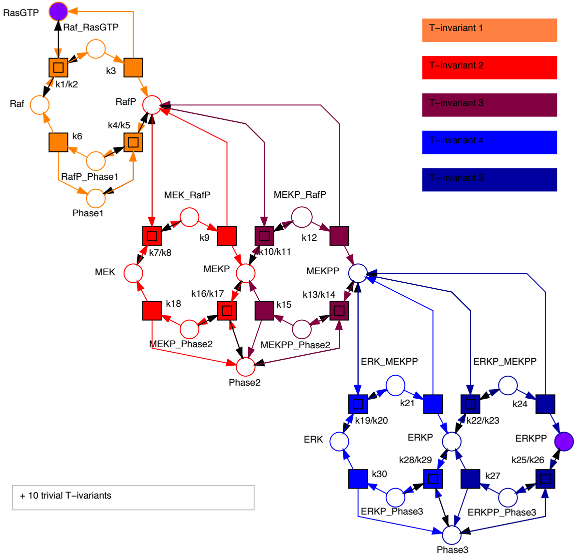
Running Case Study - general properties

- *state space*

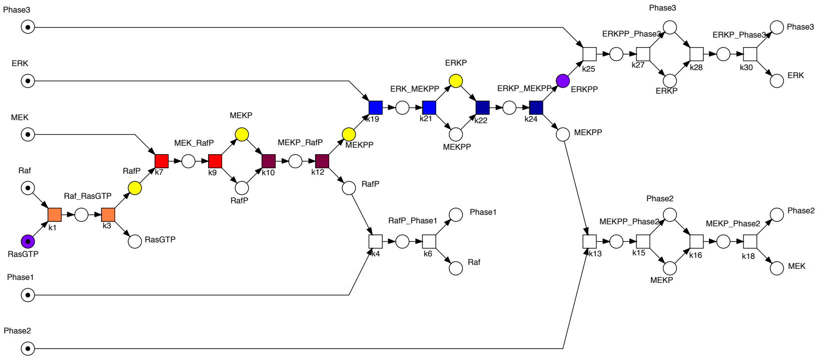
levels	reachability graph number of states	IDD data structure number of nodes
1	118	52
4	$2.4 \cdot 10^4$	115
8	$6.1 \cdot 10^6$	269
80	$5.6 \cdot 10^{18}$	13,472
120	$1.7 \cdot 10^{21}$	29,347

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

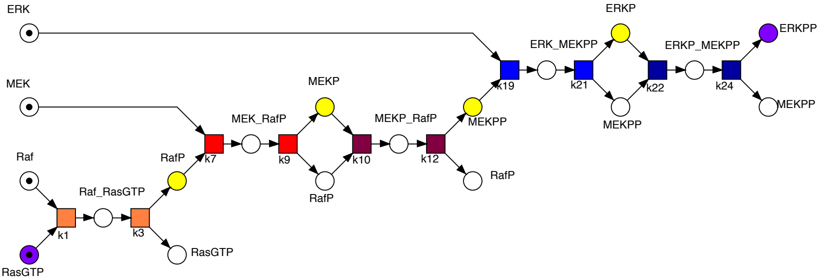
Running Case Study - T-invariants



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



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

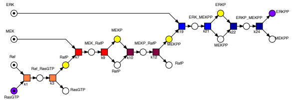


There is a path ...

- $EX\phi$
  - if there is a state reachable by one step where ϕ holds.
- $EF\phi$
  - if there is a path where ϕ holds finally, i.e., at some point.
- $EG\phi$
  - if there is a path where ϕ holds globally, i.e., forever.
- $E(\phi_1 U \phi_2)$
  - if there is a path where ϕ_1 holds until ϕ_2 holds.

For all path ...

- $AX\phi$
  - if ϕ holds for all states which are reachable by one step.
- $AF\phi$
  - if ϕ holds finally (at some point) for all paths.
- $AG\phi$
  - if ϕ holds globally (i.e. for ever) for all paths.
- $A(\phi_1 U \phi_2)$
  - if ϕ_1 holds until ϕ_2 holds for all paths.



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.

$$\begin{aligned} \neg\, [\, & \mathbf{E}\,(\neg\, \text{RafP}\ \mathbf{U}\ \text{MEKP}) \vee \\ & \mathbf{E}\,(\neg\, \text{MEKP}\ \mathbf{U}\ \text{MEKPP}) \vee \\ & \mathbf{E}\,(\neg\, \text{MEKPP}\ \mathbf{U}\ \text{ERKP}) \vee \\ & \mathbf{E}\,(\neg\, \text{ERKP}\ \mathbf{U}\ \text{ERKPP})\ ] \end{aligned}$$

Stochastic Model Checking - Preparation

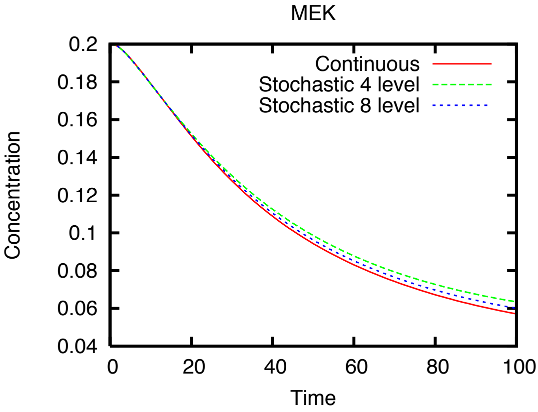
- *isomorphy of reachability graph and CTMC,* thus all qualitative properties still valid

- *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

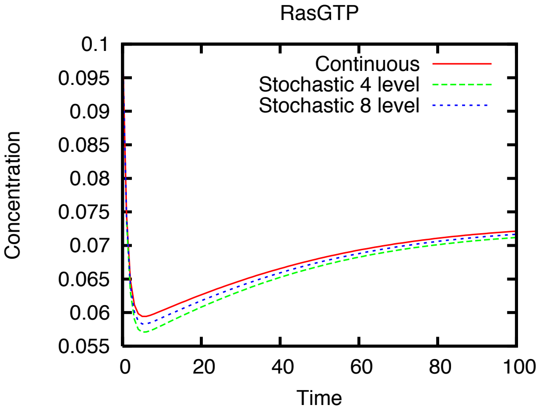
- *equivalence check*

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

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



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



Replaces the path quantifiers (E, A) in CTL by the probability operator $P_{\trianglelefteq x}$, where $\trianglelefteq x$ specifies the probability x of the formula.

- $P_{=?}[\, X\phi \,]$
  - prob there is a state reachable by one step where ϕ holds.
- $P_{=?}[\, F\phi \,]$
  - prob there is a path where ϕ holds finally, i.e., at some point.
- $P_{=?}[\, G\phi \,]$
  - prob there is a path where ϕ holds globally, i.e., forever.
- $P_{=?}[\, \phi_1 U \phi_2 \,]$
  - prob there is a path where ϕ_1 holds until ϕ_2 holds.

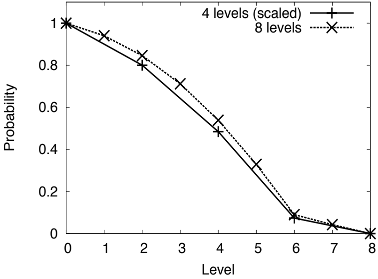
Syntactic sugar

- $\phi_1\{\phi_2\}$ - ϕ_1 happens from the first time ϕ_2 happens, where no temporal operators in ϕ_2.

property S1 :

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

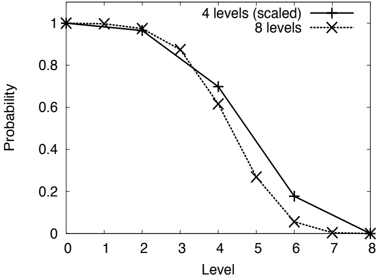
$$\mathbf{P}_{=?}\ [\ (\ \mathrm{RafP}\ =\ \mathrm{L}\ )\ \mathbf{U}^{<=100}\ (\ \mathrm{RafP} > \mathrm{L}\ )\ \{\ \mathrm{RafP} = \mathrm{L}\ \}\ ]$$



property S2 :

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

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



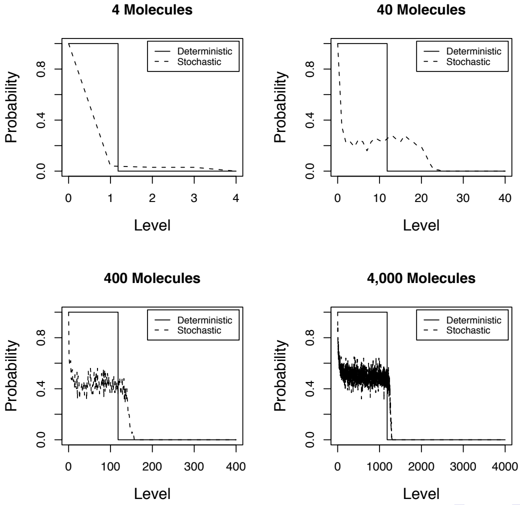
Stochastic Model Checking - Simulative Approach

Example figures for MC2 model checking of property S1 at varying number of levels/molecules.

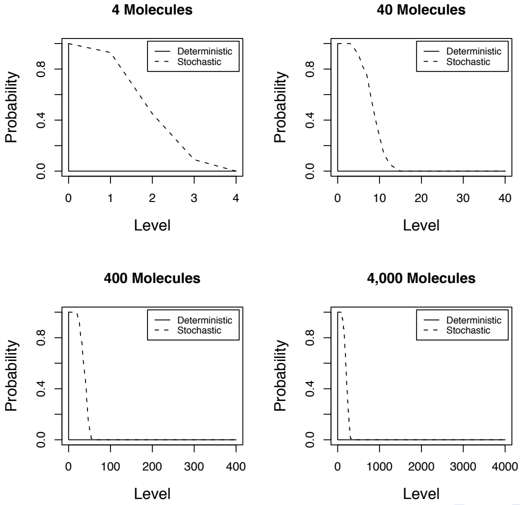
Levels	MC Time[a]	Simulation Output Size
4	10 s [b]	750 KB
8	15 s [b]	1.5 MB
40	1.5 minutes [b]	7.5 MB
400	1 minute [c]	80 MB
4,000	30 minutes [c]	900 MB

[a] Both Gillespie simulation and MC2 checking.
[b] Computation on a standard workstation.
[c] Distributed computation on a computer cluster comprising 45 Sun X2200 servers each with 2 dual core processors (180 CPU cores).

- *S1 at varying number of molecules shows progression towards deterministic behaviour as number of molecules increases.*

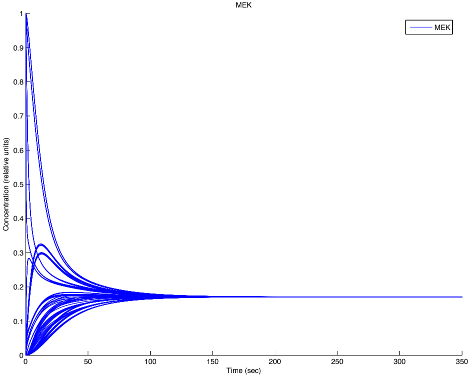


- *S2 at varying number of molecules shows progression towards deterministic behaviour as number of molecules increases.*



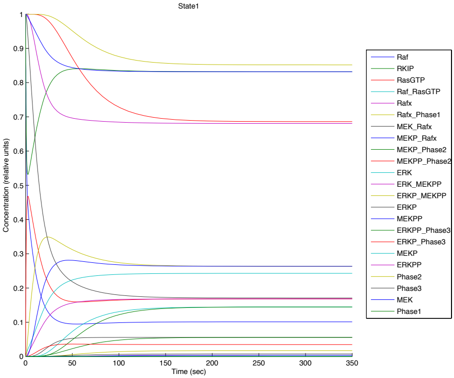
Continuous Model Checking - Preparation

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

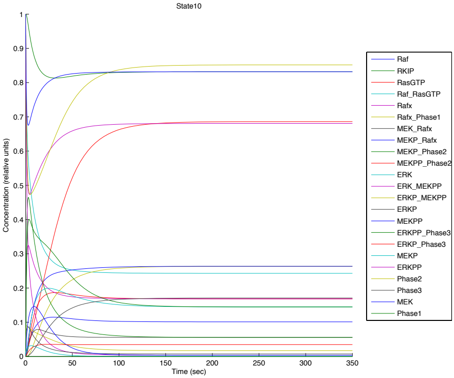


Continuous Model Checking - Preparation

- *steady state analysis for state 1 :*



- *steady state analysis for state 10 :*



For all single path ...

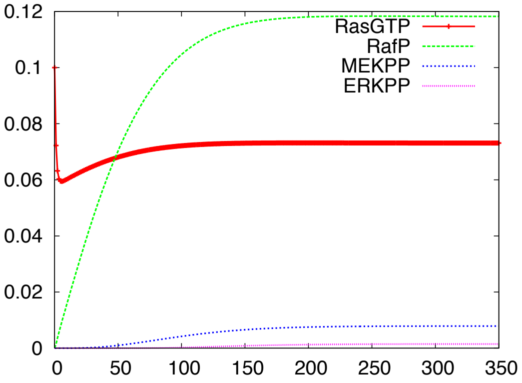
- $X\phi$
  - ϕ happens in the next time point.
- $F\phi$
  - ϕ happens at some time.
- $G\phi$
  - ϕ always happens.
- $A(\phi_1 U \phi_2)$
  - ϕ_1 happens until ϕ_2 happens.

Syntactic sugar

- $\phi_1\{\phi_2\}$ - ϕ_1 happens from the first time ϕ_2 happens, where no temporal operators in ϕ_2.

Continuous Model Checking (LTLc)

- *transient analysis for RasGTP, RafP, MEKPP, ERKPP :*



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.*

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

property C2 :

if the concentration of RafP is at a significant concentration level and that of ERKPP is close to zero, then both species remain in these states until the concentration of MEKPP becomes significant ; i.e. *MEKPP is the second species to react.*

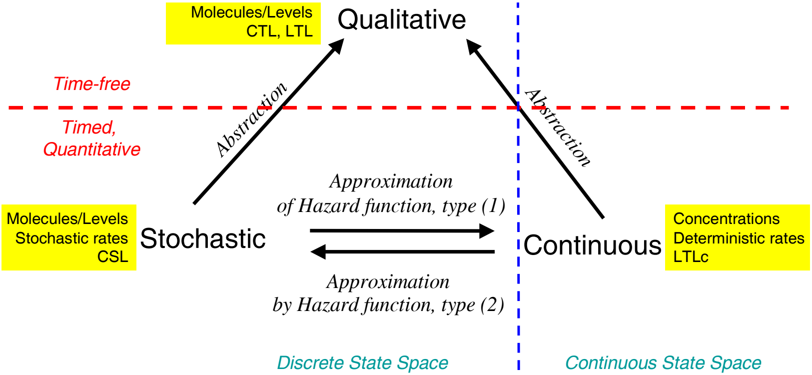
$$((\text{RafP} > 0.06) \wedge (\text{ERKPP} < 0.0002)) \Rightarrow ((\text{RafP} > 0.06) \wedge (\text{ERKPP} < 0.0002)) \; \mathbf{U} \; (\text{MEKPP} > 0.004)$$

property C3 :

if the concentrations of RafP and MEKPP are significant, they remain so, until the concentration of ERKPP becomes significant ; i.e. *ERKPP is the third species to react.*

$$((\text{RafP} > 0.06) \wedge (\text{MEKPP} > 0.004)) \Rightarrow$$
$$((\text{RafP} > 0.06) \wedge (\text{MEKPP} > 0.004)) \; \mathbf{U} \; (\text{ERKPP} > 0.0005)$$

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

Bionetworks, Validation

- *validation criterion 1*
  - all expected structural properties hold
  - all expected general behavioural properties hold
- *validation criterion 2*
  - CPI
  - 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 yield TRUE

Tools

- *model construction, animation, simulation*
  - Snoopy
- *qualitative analysis*
  - Charlie
  - Marcie
- *stochastic analysis*
  - analytical model checking : Marcie/CSL *(previously PRISM)*
  - simulative model checking : Marcie/PLTLc, MC2/PLTLc
- *continuous analysis*
  - model checking : MC2(LTLc) *(previously BioCham)*

Finally

This is an updated version of

M Heiner, D Gilbert, R Donaldson :
Petri Nets for Systems and Synthetic Biology ;
SFM 2008, Bertinoro, Springer LNCS 5016, pp. 215-264, 2008.

all data files and analysis results available at
www-dssz.informatik.tu-cottbus.de/examples/levchenko