Parameter Scanning by Parallel Model Checking with applications in Systems Biology

HIBI 2010
University of Twente

David Šafránek
with
Jiří Barnat, Luboš Brim, Martin Vejnár
Masaryk University Brno

1st October, 2010

Outline

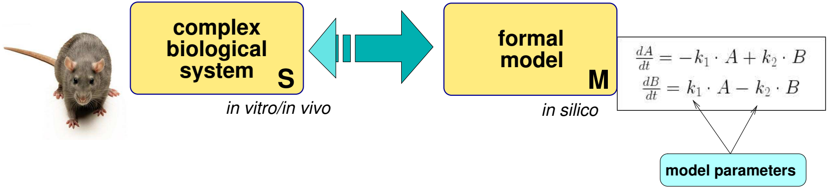
1. Motivation and Background
2. Method Description
3. Case Study

Outline

1. Motivation and Background
2. Method Description
3. Case Study

Motivation

Tuning the Biological Models



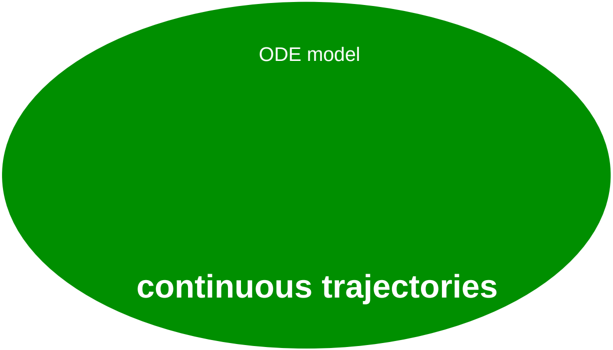
- scanning of the proper parameter values
  - the set of unknown model parameters χ
  - traditional methods of fitting models to *in vitro* data
  - **property-driven parameter scanning**

    $\Rightarrow$ find the parameter values exibiting the given dynamic phenomena

To learn the rat dancing...

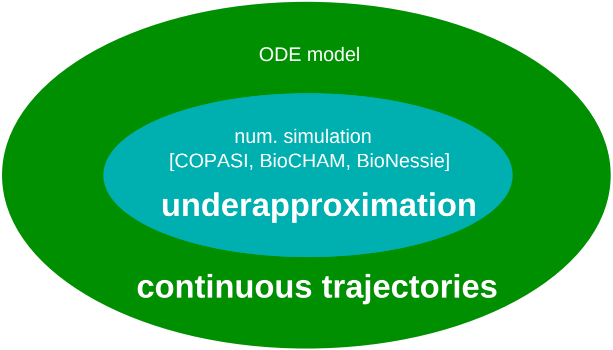
Modeling Approach

How to obtain a finite abstraction?



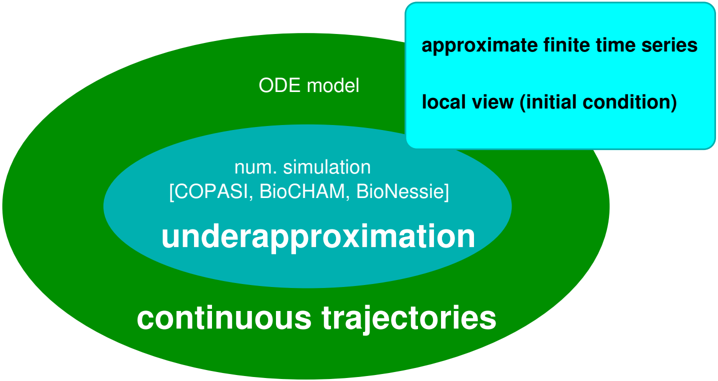
Modeling Approach

How to obtain a finite abstraction?



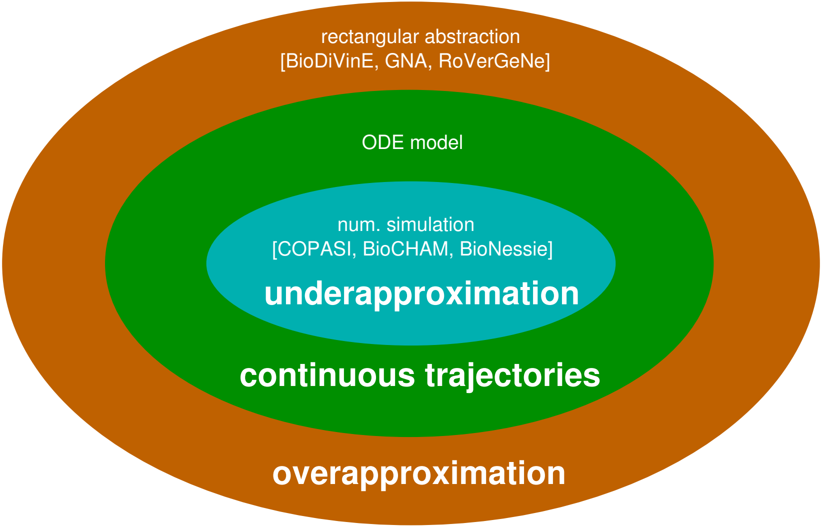
Modeling Approach

How to obtain a finite abstraction?



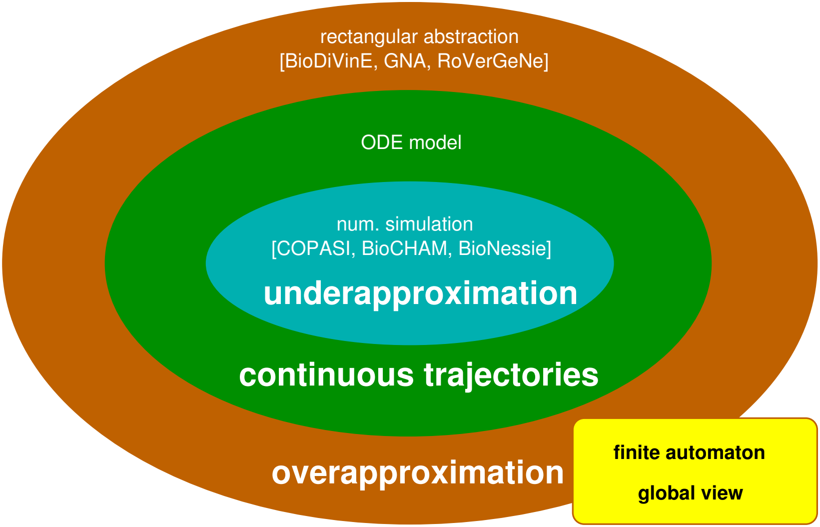
Modeling Approach

How to obtain a finite abstraction?

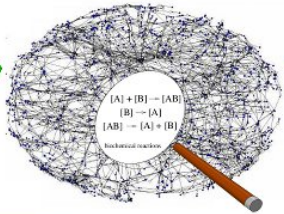


Modeling Approach

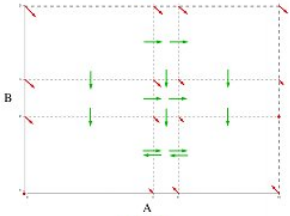
How to obtain a finite abstraction?

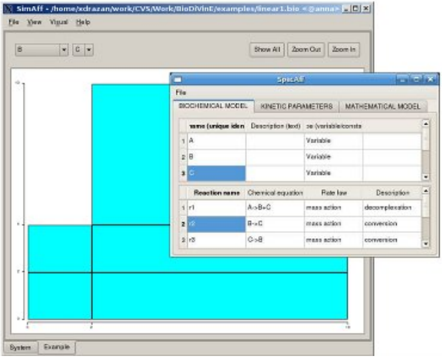


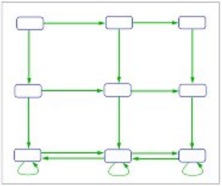
Modeling Approach



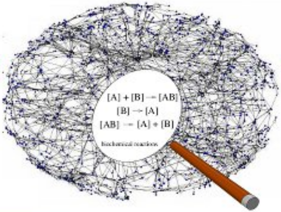
$$\frac{dA}{dt} = -k_3 \cdot [AB] + k_2 \cdot B - k_1 \cdot A \cdot B$$
$$\frac{dB}{dt} = k_3 \cdot [AB] - k_1 \cdot A \cdot B - k_2 \cdot B$$
$$\frac{dAB}{dt} = k_1 \cdot A \cdot B - k_3 \cdot [AB]$$



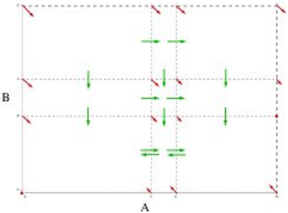


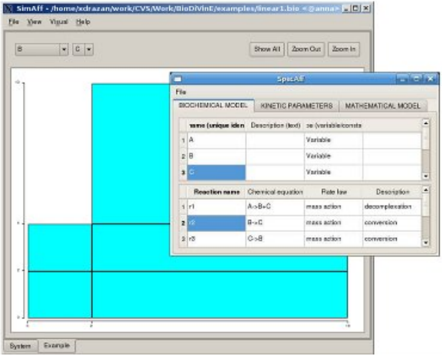


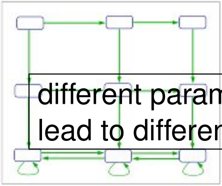
Modeling Approach



$$\frac{dA}{dt} = -k_3 \cdot [AB] + k_2 \cdot B - k_1 \cdot A \cdot B$$
$$\frac{dB}{dt} = k_3 \cdot [AB] - k_1 \cdot A \cdot B - k_2 \cdot B$$
$$\frac{dAB}{dt} = k_1 \cdot A \cdot B - k_3 \cdot [AB]$$







different parameter values lead to different transitions

Problem Definition

Robustness

Given a dynamic (temporal) property φ and a parameterized model $\mathcal{M}(\chi)$ check if $\mathcal{M}(p) \models \varphi$ **holds for all possible parameterizations** p (valuations of χ).

Parameter Scanning Problem

Given a dynamic (temporal) property φ and a parameterized model $\mathcal{M}(\chi)$ **find the maximal set** P **of parameterizations of** χ such that $\mathcal{M}(p) \models \varphi$ for all $p \in P$.

Problem Reduction

Robustness is reduced to Parameter Scanning Problem by taking the set of all possible parameterizations as P.

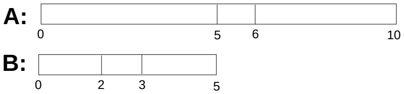
Related Work

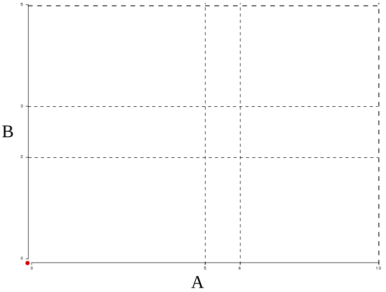
- BioCHAM [Fages et.al.]
  - based on numerical simulation
  - LTL with constraints interpreted over simulations
  - notion of “satisfaction degree” defined
  - analysis dependent on precise initial conditions
- RoVerGeNe [Batt et.al.]
  - model checking over a finite discrete abstraction
  - properties “decidable” globally
  - BDD representation of parameter space
  - w.c.s.: two model checking operations per each parameterization
  - computationaly hard (3 variable model hours-days of computing)
- our contribution
  - revisit the latter approach
  - new algorithm based on enumerative LTL model checking
  - parallel implementation

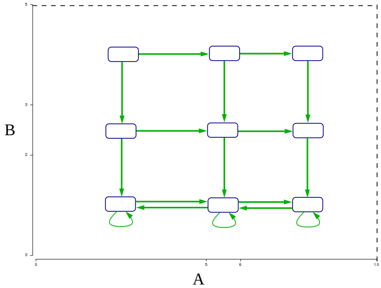
Rectangular Abstraction

$$A \overset{k_1}{\longleftarrow} \overset{k_2}{\longrightarrow} B$$

$$\frac{dA}{dt} = -k_1 \cdot A + k_2 \cdot B$$
$$\frac{dB}{dt} = k_1 \cdot A - k_2 \cdot B$$







Supported Classes of Models

Mass Action Kinetics

- biochemical reactions modeled by **mass action kinetics**
- the supported format of reactions:

$$\gamma_1 X_1 + \cdots + \gamma_m X_m \xrightarrow{k} \delta_1 Y_1 + \cdots + \delta_n Y_n, \quad \gamma_i \in \{0, 1\}$$

- resulting ODE system describes dynamics of each variable (chemical species):

$$\forall i \in \{1, ..., n\}. \frac{dY_i}{dt} = f(X_1, ..., X_m) = kX_1^{\gamma_1} X_2^{\gamma_2} \cdots X_m^{\gamma_m}$$

$$\forall i \in \{1, ..., m\}. \frac{dX_i}{dt} = f(X_1, ..., X_m) = -kX_1^{\gamma_1} X_2^{\gamma_2} \cdots X_m^{\gamma_m}$$

- $k \in \mathbb{R}^+$ is *kinetic parameter*

Supported Classes of Models

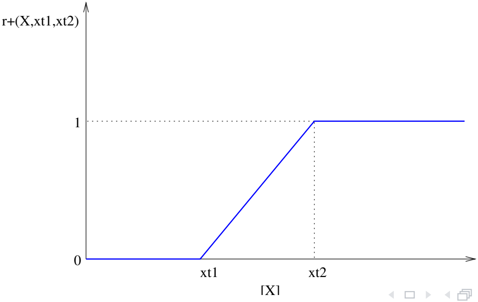
Regulatory Kinetics

- protein dynamics driven by protein-regulated transcription
- Hill kinetics approximated in terms of *ramp functions*

$$X \longrightarrow Y$$

$$\frac{dY}{dt} = kr^{+}(X, xt_1, xt_2)$$

- $k \in \mathbb{R}^+$ is *kinetic parameter*
- xt_1, xt_2 determine the partitioning of X



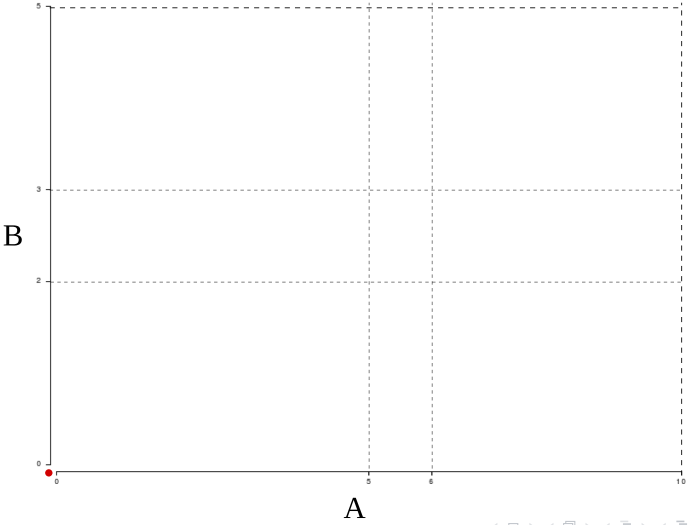
Supported Classes of Models

Mixing the Both

- we combine both kinetics aspects under one ODE formalism
- right-hand side of any ODE is a function $f(\boldsymbol{X}, \boldsymbol{p})$ where $\boldsymbol{p}$ is a vector of unknown parameters
  - (piece-wise) multi-affine in $\boldsymbol{X}$
  - affine in $\boldsymbol{p}$
- these properties enable us to (are necessary to):
  - make a discrete finite overapproximation of the system dynamics
  - finitely represent the *parameter space* – possible values of $\boldsymbol{p}$
- linear combinations of unknown parameters not considered
  $\Rightarrow$ at most one unknown parameter per equation
  $\Rightarrow$ simple manipulation with the parameter space

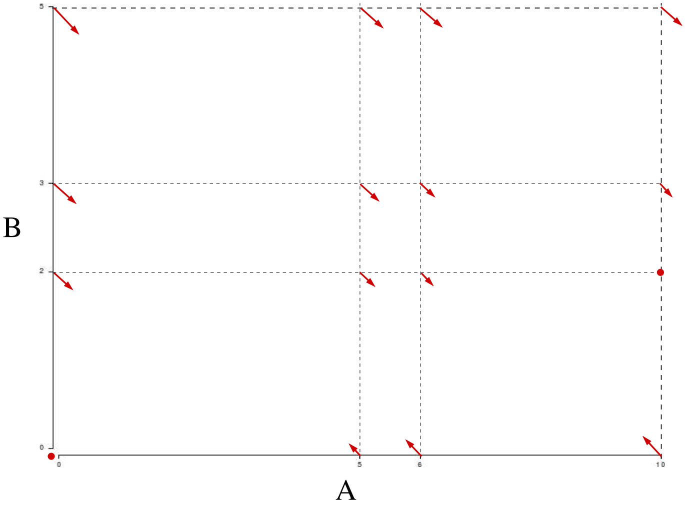
The Principle of Rectangular Abstraction

- approach of Belta, Habets, van Schuppen
- continuous solutions abstracted by a non-deterministic automaton



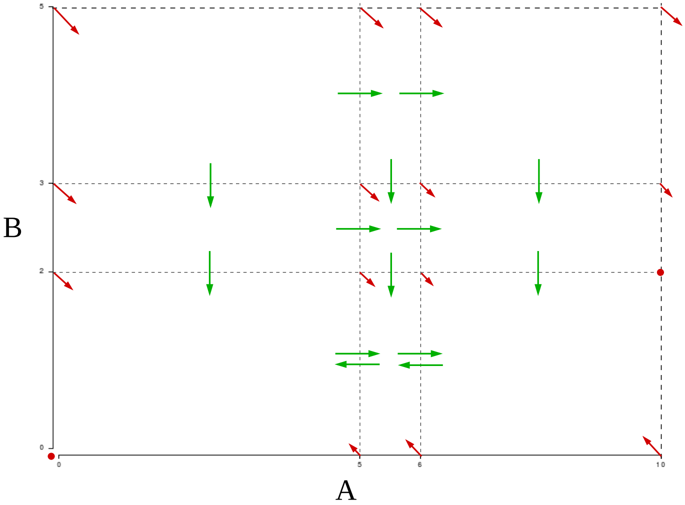
The Principle of Rectangular Abstraction

- approach of Belta, Habets, van Schuppen
- continuous solutions abstracted by a non-deterministic automaton



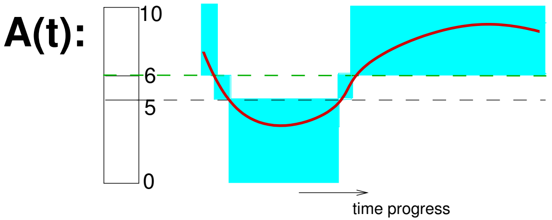
The Principle of Rectangular Abstraction

- approach of Belta, Habets, van Schuppen
- continuous solutions abstracted by a non-deterministic automaton



Analysed Properties

- reasoning about concentration levels (up-to the abstraction)
- reachability
  - *global*: regardless the initial state, *B* eventually falls below 2
  - *local*: if *B* initally below 2 then *B* does not exceed 2
- temporal properties
  - there is no initial state from which *A* falls below 6 before *A* exceeds 6



- properties can be formulated in Linear Temporal Logic (LTL)

Outline

1 Motivation and Background

2 Method Description

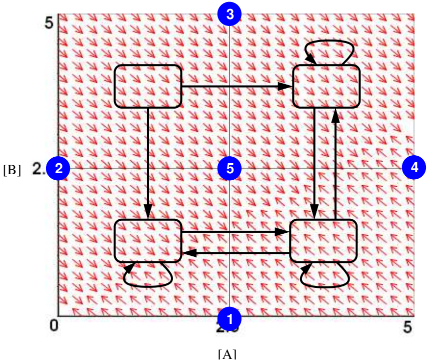
3 Case Study

Computing the Parameter Space

$$\frac{dA}{dt} = -k_1 \cdot A + k_2 \cdot B$$
$$\frac{dB}{dt} = k_1 \cdot A - k_2 \cdot B$$

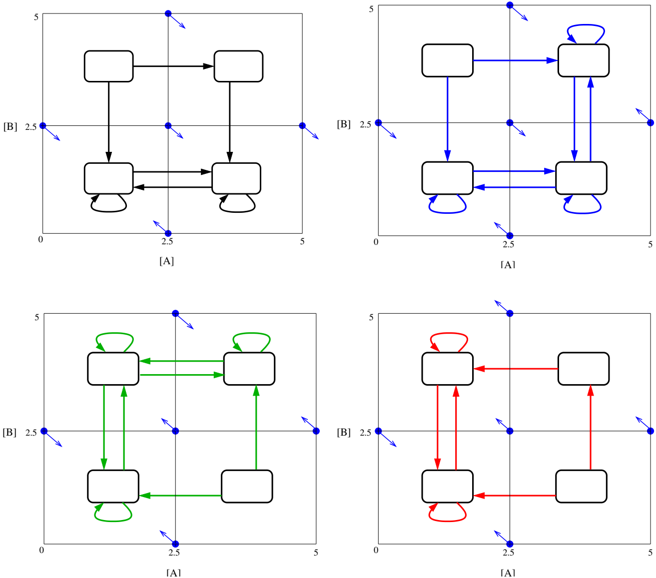
$k_2 = 0.8$

$k_1 = ?$



	value of k1: (0,0.4)	(0.4,0.8)	(0.8,1.6)	(1.6,max)
1	↖	↖	↖	↖
2	↘	↘	↘	↘
3	↘	↘	↘	↖
4	↘	↖	↖	↖
5	↘	↘	↖	↖

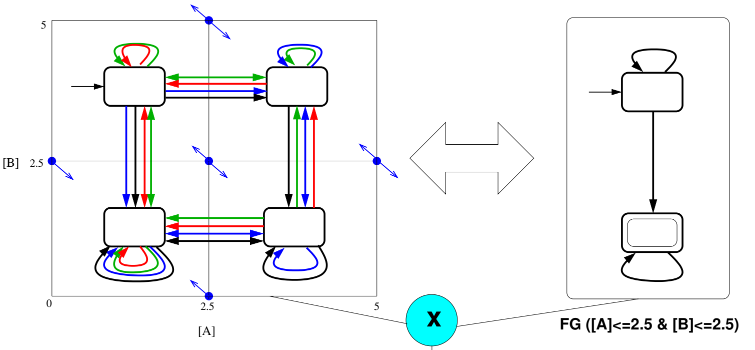
Effect of Parameters on Abstraction Automaton



Parameter Scanning by LTL Model Checking

parameterized Kripke structure of the model

never claim Buchi automaton



check if the product accepts an empty language

normally for each parameterization separately

we compute all parameterizations together

YES

property is robust

GF ([A]>2.5 | [B]>2.5)

NO

set of parameter values P violating the property

counterexample may be a false positive

inverse of P in entire parameter space
is the maximal set of valid parameters

may be underapproximated

Model Checking on Coloured Graphs

Idea

- represent each parameterization by a distinct colour
- assume all transitions for each parameterization adequately coloured
- find accepting cycles and get colours enabling accepting runs

Procedure

1. compute initial mapping of colours to states
   $\Rightarrow$ propagate colours through the entire graph (BFS reachability)
   $\Rightarrow$ states on accepting cycles know all colours by which they are reached
2. for each reachable accepting cycle aggregate (scan) the valid colours

Complexity Issues

- worst case: $O(|S| \cdot |E| \cdot |F| \cdot |\mathcal{P}|)$
  $|S|$...states, E...edges, F...accepting states, $\mathcal{P}$...colours
- in expected cases $|S|$ is reduced (levels of BFS)

Parallel Implementation

Problem

- *number of states exponential w.r.t. number of variables*
- *size of the parameter space exponential w.r.t. number of unknown parameters*
- *many computations performed on a single graph*

Solution

- *multi-core data-parallel implementation of colour mapping propagation*
- *states evenly distributed among threads by a hash-function*
- *each thread responsible for a unique partition of colour mapping*
- *threads communicate via a colour mapping update qeue*
  - *implemented as a set of lock-free qeues*
  - *one qeue per thread*
  - *threads synchronize on BFS levels*

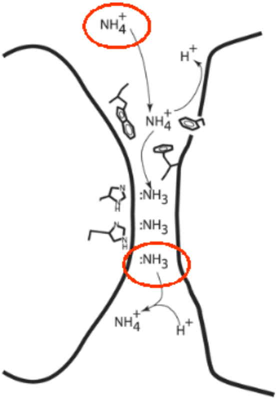
Outline

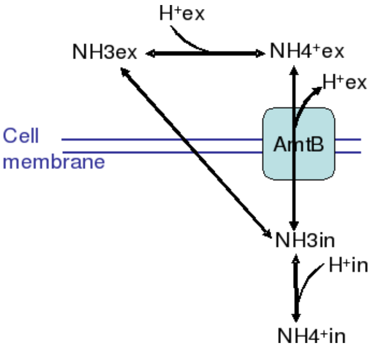
1 Motivation and Background

2 Method Description

3 Case Study

Case Study I – E. Coli Ammonium Transport





$$\frac{d[AmtB]}{dt} = -k_1[AmtB][NH_4ex] + k_2[AmtB:NH_4] + k_4[AmtB:NH_3]$$
$$\frac{d[AmtB:NH_3]}{dt} = k_3[AmtB:NH_4] - k_4[AmtB:NH_3]$$
$$\frac{d[AmtB:NH_4]}{dt} = k_1[AmtB][NH_4ex] - k_2[AmtB:NH_4] - k_3[AmtB:NH_4]$$
$$\frac{d[NH_3in]}{dt} = k_4[AmtB:NH_3] - k_6[NH_3in][H_{in}] + k_7[NH_4in] + k_9[NH_3ex]$$
$$\frac{d[NH_4in]}{dt} = k_6[NH_3in][H_{in}] - k_5[NH_4in] - k_7[NH_4in]$$

E. Coli Ammonium Transport: Model Settings

Settings

- mass action kinetics $\Rightarrow$ multi-affine ODE model
- abstraction – number of discrete concentration levels considered:

$AmtB$	$AmtB : NH_3$	$AmtB : NH_4$	NH_3in	NH_4in
7	9	3	8	26

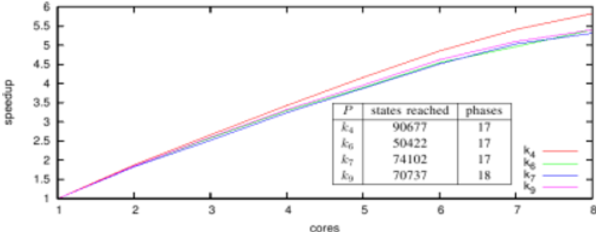
- initial conditions set to impose low external ammonium conditions

Experiments

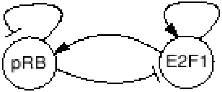
- find the maximal set of parameter values for the given uknown parameter ensuring the maximal reachable level of internal NH_3 is $1.1 \cdot 10^6$ *mol*
- the employed LTL property: $\mathbf{G}(NH_3in < 1.1 \cdot 10^6)$

E. Coli Ammonium Transport: Experiments

χ	intervals of validity	time
k_4	$(1 \cdot 10^{-12}, 2.7 \cdot 10^{6})$	30 s
k_6	$(5.2 \cdot 10^{6}, 1 \cdot 10^{12})$	22 s
k_7	$(1 \cdot 10^{-12}, 3.3 \cdot 10^{6})$	33 s
k_9	$(1 \cdot 10^{-12}, 2.7 \cdot 10^{6})$	20 s
$k_{1,6,10}$	see the paper	19 min



Case Study II – Genetic Regulation of G_1/S Transition



$$\frac{d[pRB]}{dt} = k_1 \varrho_1(pRB, E2F1) - \gamma_{pRB}[pRB]$$
$$\frac{d[E2F1]}{dt} = k_p + k_2 \varrho_2(pRB, E2F1) - \gamma_{E2F1}[E2F1]$$

- central module controlling G_1/S transition of mammalian cells
- bistability w.r.t. setting of γ_{pRB} parameter in the range $[0.01, 0.1]$
- liveness property $\mathbf{FG}[E2F1] > 8$
- many false-positive runs arise due to time-convergent behaviour introduced by abstraction
- in the paper we present a (non-universal) solution

Conclusions

- new algorithm introduced (model checking on coloured graphs)
- scalability achieved on a multi-core implementation
- full support of multi-affine ODE models
- case studies show practicability for common models
  - large-scale models still a challenge
  - for simple parameter scanning problems effective alternative to RoVerGene
- future work
  - reduction of false-positives
  - distributed implementation
  - computation of more complicated parameter spaces
  - other application of the idea of model checking on coloured graphs

Thank you for your attention!