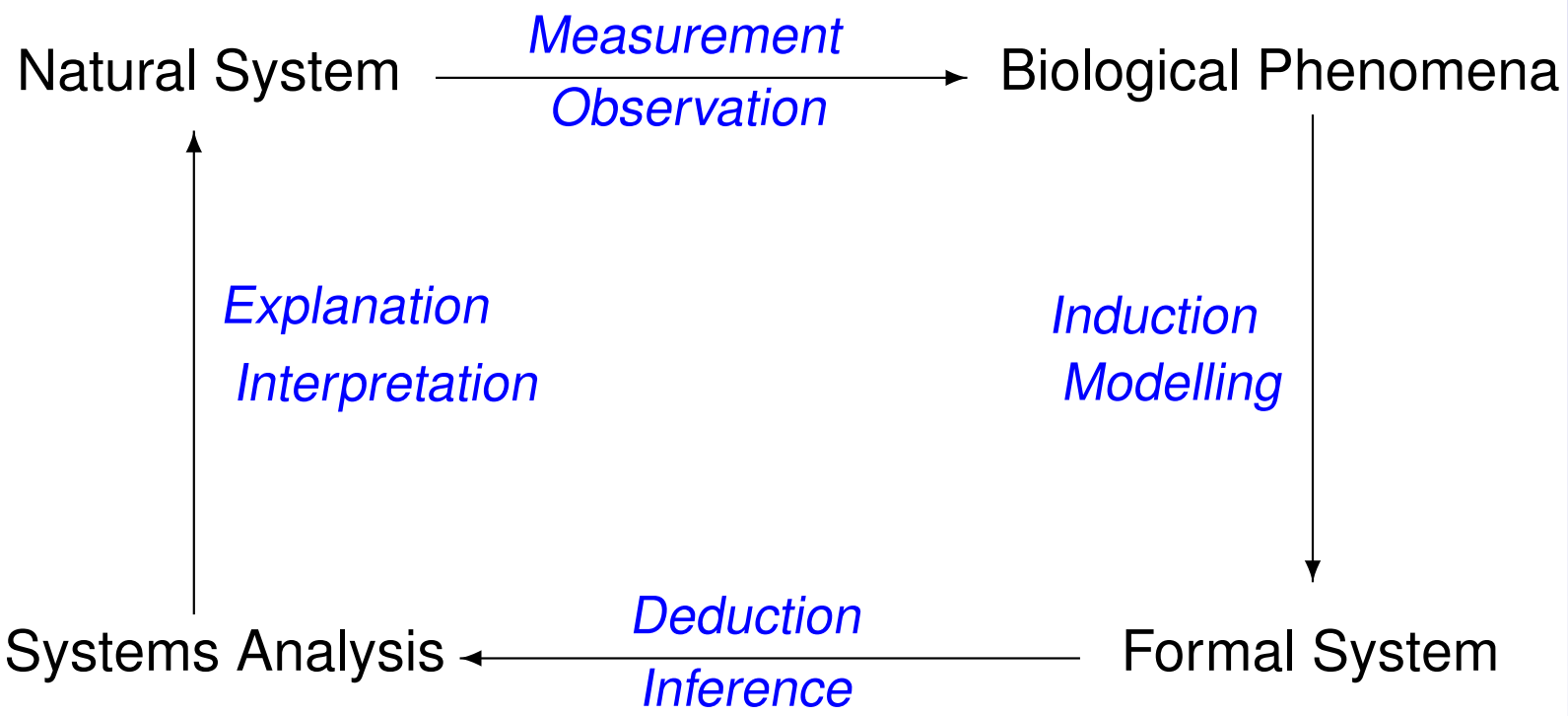
Bio-PEPA: A collective dynamics approach to systems biology

Jane Hillston.
LFCS and CSBE, University of Edinburgh

30th April 2010

Outline

Systems Biology Methodology



Jane Hillston.
University of Edinburgh.

Formal Systems

There are two alternative approaches to constructing dynamic models of biochemical pathways commonly used by biologists:

Formal Systems

There are two alternative approaches to constructing dynamic models of biochemical pathways commonly used by biologists:

- Ordinary Differential Equations:
  - continuous time,
  - continuous behaviour (concentrations),
  - deterministic.

Formal Systems

There are two alternative approaches to constructing dynamic models of biochemical pathways commonly used by biologists:

- Ordinary Differential Equations:
  - continuous time,
  - continuous behaviour (concentrations),
  - deterministic.
- Stochastic Simulation:
  - continuous time,
  - discrete behaviour (no. of molecules),
  - stochastic.

Current Hypothesis

Some of the techniques we have developed over the last thirty years for modelling complex software systems can be beneficially applied to the modelling aspects of systems biology.

Current Hypothesis

Some of the techniques we have developed over the last thirty years for modelling complex software systems can be beneficially applied to the modelling aspects of systems biology.

In particular formalisms which encompass support for

- Abstraction
- Modularity and
- Reasoning

have a key role to play.

Current Hypothesis

Some of the techniques we have developed over the last thirty years for modelling complex software systems can be beneficially applied to the modelling aspects of systems biology.

In particular formalisms which encompass support for

- Abstraction
- Modularity and
- Reasoning

have a key role to play.

Process algebras have mechanisms for each of these, and stochastic extensions which allow dynamic properties to be analysed.

Bio-PEPA: motivations

Work on the stochastic π-calculus and related calculi, is typically based on Regev and Shapiro's mapping: molecules as processes.

Bio-PEPA: motivations

Work on the stochastic π-calculus and related calculi, is typically based on Regev and Shapiro's mapping: molecules as processes.

This is an inherently individuals-based view of the system and analysis will generally then be via stochastic simulation.

Bio-PEPA: motivations

Work on the stochastic π-calculus and related calculi, is typically based on Regev and Shapiro's mapping: molecules as processes.

This is an inherently individuals-based view of the system and analysis will generally then be via stochastic simulation.

With Bio-PEPA we have been experimenting with more abstract mappings between elements of signalling pathways and process algebra constructs: species as processes.

Bio-PEPA: motivations

Work on the stochastic π-calculus and related calculi, is typically based on Regev and Shapiro's mapping: molecules as processes.

This is an inherently individuals-based view of the system and analysis will generally then be via stochastic simulation.

With Bio-PEPA we have been experimenting with more abstract mappings between elements of signalling pathways and process algebra constructs: species as processes.

Abstract models are more amenable to a collective dynamics view and integrated analysis..

Bio-PEPA: motivations

Work on the stochastic π-calculus and related calculi, is typically based on Regev and Shapiro's mapping: molecules as processes.

This is an inherently individuals-based view of the system and analysis will generally then be via stochastic simulation.

With Bio-PEPA we have been experimenting with more abstract mappings between elements of signalling pathways and process algebra constructs: species as processes.

Abstract models are more amenable to a collective dynamics view and integrated analysis..

We also wanted to be able to capture more biological features .

Motivations for Collective Dynamics

The scale of biological models is vast with tens or hundreds of thousands of molecules of each type often present within the cell.

Motivations for Collective Dynamics

The scale of biological models is vast with tens or hundreds of thousands of molecules of each type often present within the cell.

Many experimental techniques have poor resolution — they cannot distinguish individual cells never mind individual molecules.

Motivations for Collective Dynamics

The scale of biological models is vast with tens or hundreds of thousands of molecules of each type often present within the cell.

Many experimental techniques have poor resolution — they cannot distinguish individual cells never mind individual molecules.

Ordinary differential equations have been used in the context of biochemistry for decades and for many systems that level of abstraction is appropriate so the collective dynamics view is one which biologists are comfortable with.

Motivations for Abstraction

Our motivations for seeking more abstraction:

Motivations for Abstraction

Our motivations for seeking more abstraction:

- Process algebra-based analyses such as comparing models (e.g. for equivalence or simulation) and model checking are only possible if the state space is not prohibitively large.

Motivations for Abstraction

Our motivations for seeking more abstraction:

- Process algebra-based analyses such as comparing models (e.g. for equivalence or simulation) and model checking are only possible if the state space is not prohibitively large.
- The data that we have available to parameterise models is sometimes speculative rather than precise.

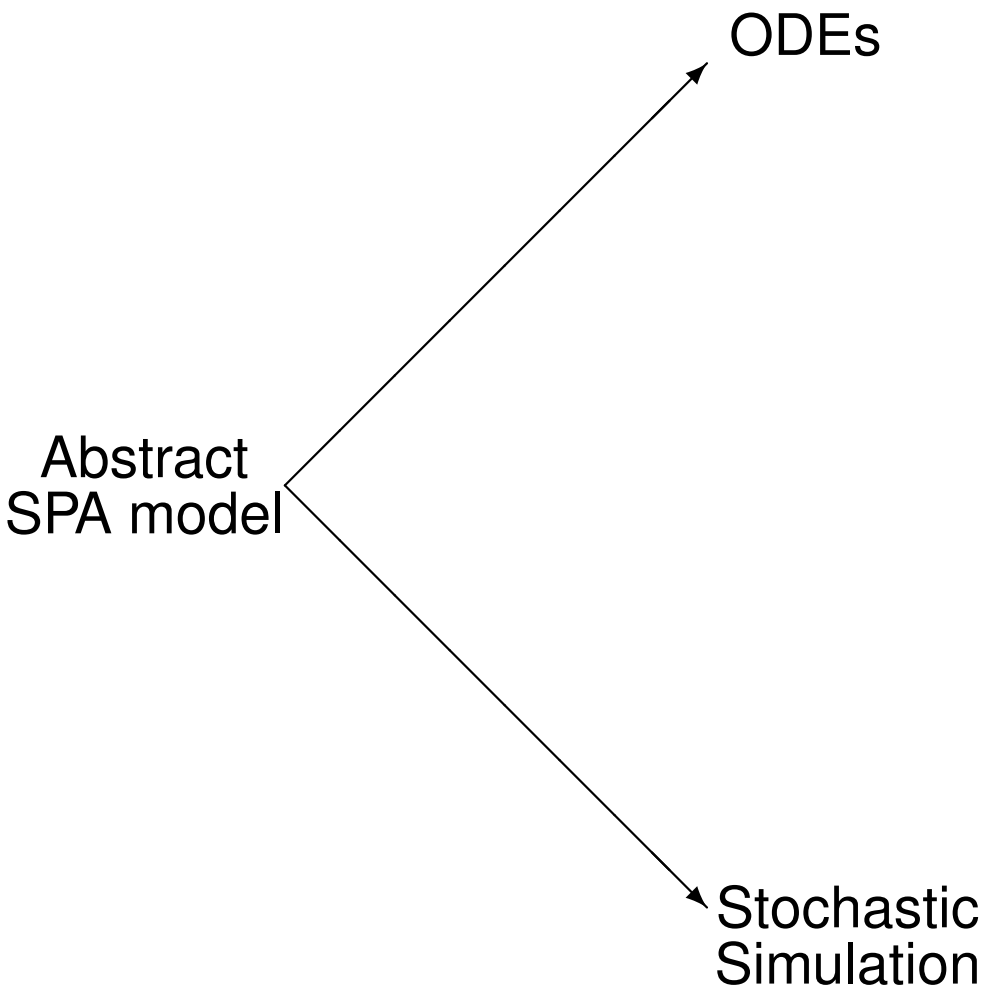
Motivations for Abstraction

Our motivations for seeking more abstraction:

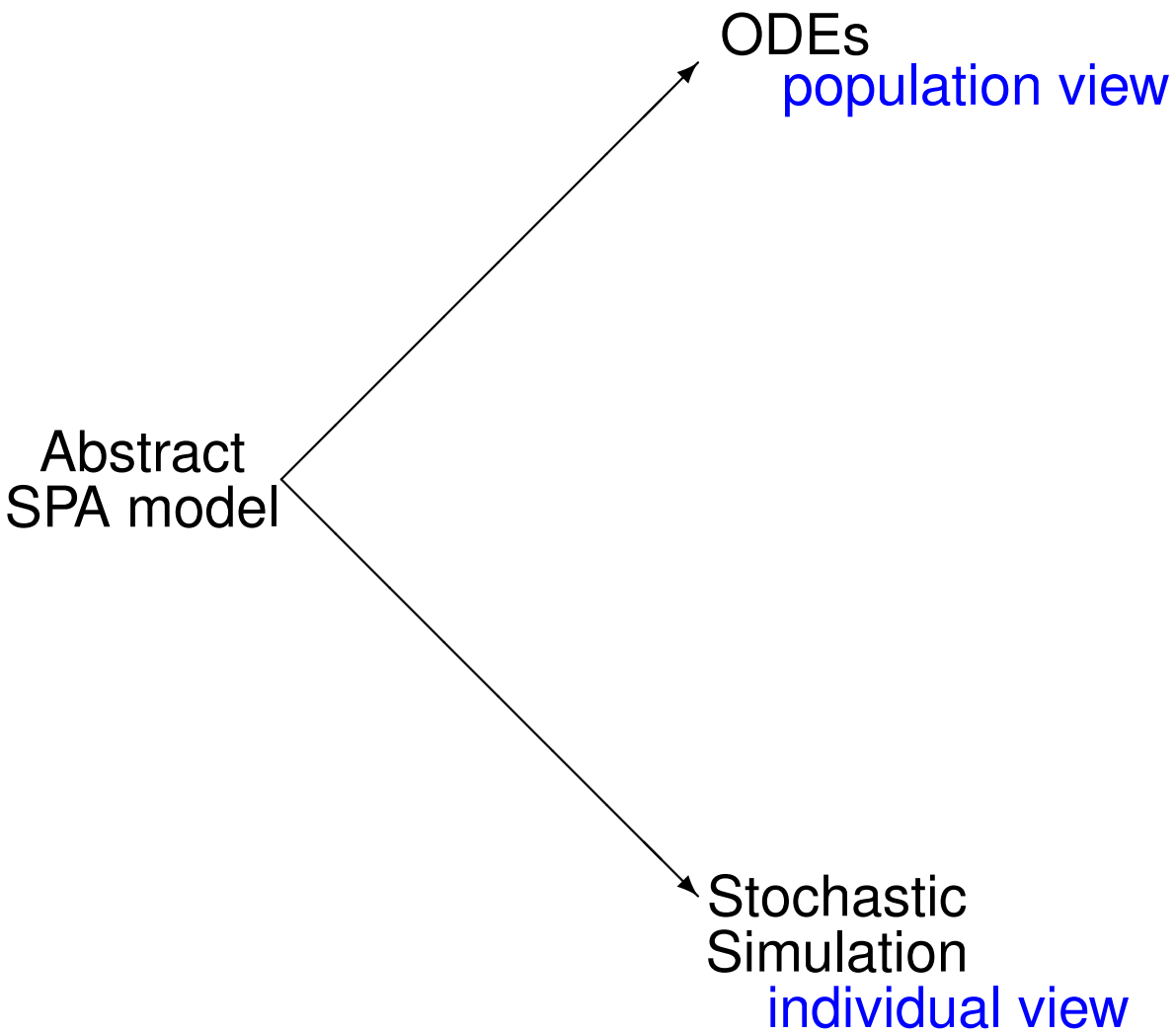
- Process algebra-based analyses such as comparing models (e.g. for equivalence or simulation) and model checking are only possible if the state space is not prohibitively large.
- The data that we have available to parameterise models is sometimes speculative rather than precise.

  This suggests that it can be useful to use semi-quantitative models rather than quantitative ones.

Alternative Representations



Alternative Representations

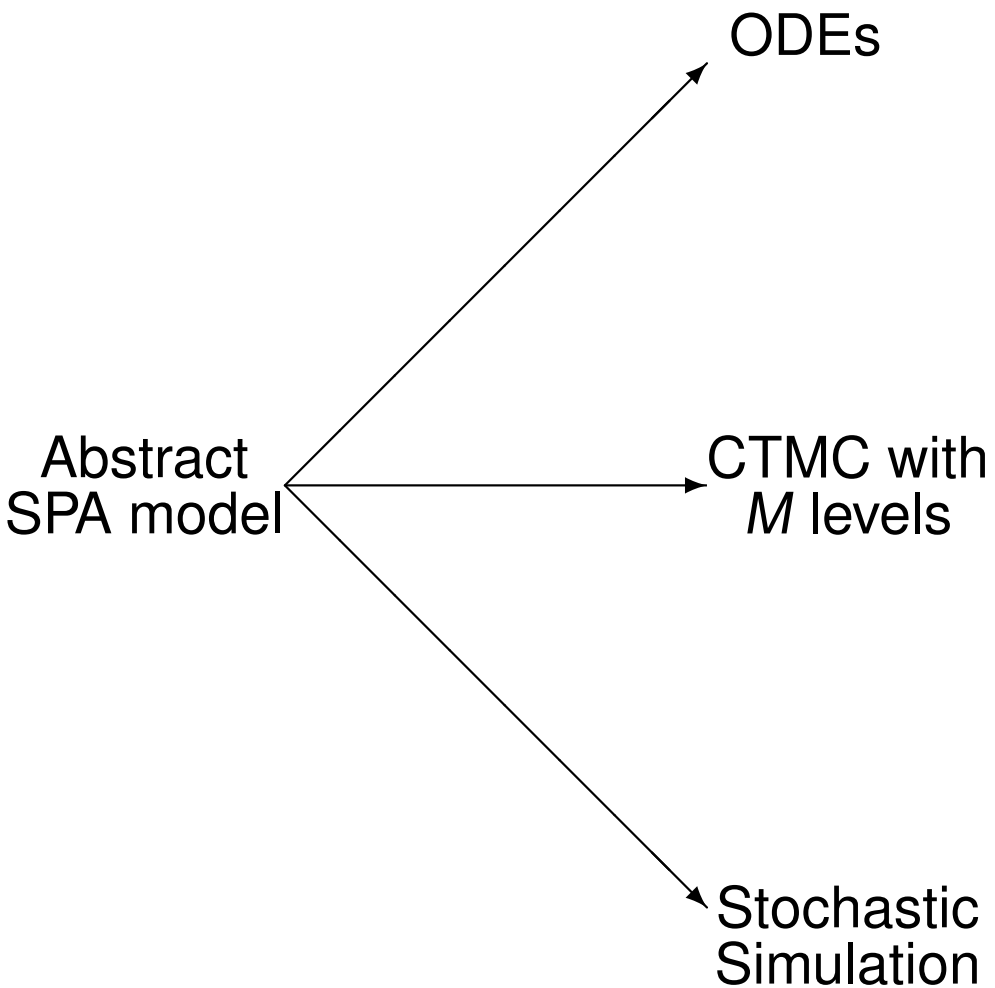


Discretising the population view

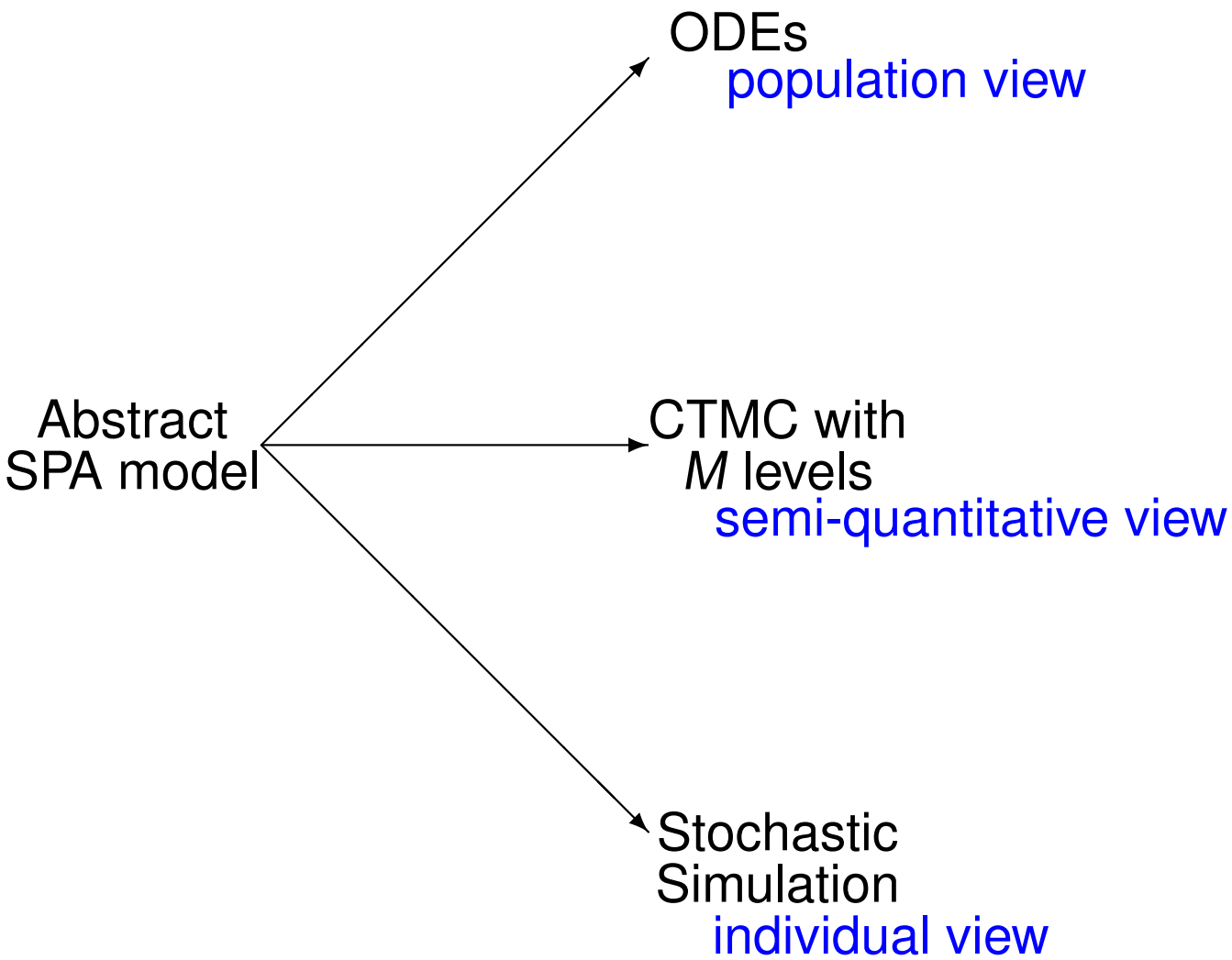
Jane Hillston.
University of Edinburgh.

We can discretise the continuous range of possible concentration values into a number of distinct states. These form the possible states of the component representing the reagent.

Alternative Representations



Alternative Representations



Jane Hillston.
University of Edinburgh.

Modelling biological features

There are some features of biochemical reaction systems which are not readily captured by many of the stochastic process algebras that are currently in use.

Modelling biological features

There are some features of biochemical reaction systems which are not readily captured by many of the stochastic process algebras that are currently in use.

Particular problems are encountered with:

Modelling biological features

There are some features of biochemical reaction systems which are not readily captured by many of the stochastic process algebras that are currently in use.

Particular problems are encountered with:

- stoichiometry — the multiplicity in which an entity participates in a reaction;

Modelling biological features

There are some features of biochemical reaction systems which are not readily captured by many of the stochastic process algebras that are currently in use.

Particular problems are encountered with:

- stoichiometry — the multiplicity in which an entity participates in a reaction;
- general kinetic laws — although mass action is widely used other kinetics are also commonly employed.

Modelling biological features

There are some features of biochemical reaction systems which are not readily captured by many of the stochastic process algebras that are currently in use.

Particular problems are encountered with:

- stoichiometry — the multiplicity in which an entity participates in a reaction;
- general kinetic laws — although mass action is widely used other kinetics are also commonly employed.
- multiway reactions — although thermodynamic arguments can be made that there are never more than two reagents involved in a reaction, in practice it is often useful to model at a more abstract level.

Bio-PEPA

In Bio-PEPA:

Bio-PEPA

In Bio-PEPA:

- Unique rates are associated with each reaction (action) type, separately from the specification of the logical behaviour. These rates may be specified by functions.

Bio-PEPA

In Bio-PEPA:

- Unique rates are associated with each reaction (action) type, separately from the specification of the logical behaviour. These rates may be specified by functions.
- The representation of an action within a component (species) records the stoichiometry of that entity with respect to that reaction. The role of the entity is also distinguished.

Bio-PEPA

In Bio-PEPA:

- Unique rates are associated with each reaction (action) type, separately from the specification of the logical behaviour. These rates may be specified by functions.
- The representation of an action within a component (species) records the stoichiometry of that entity with respect to that reaction. The role of the entity is also distinguished.
- The local states of components are quantitative rather than functional, i.e. distinct states of the species are represented as distinct components, not derivatives of a single component.

Outline

The syntax

Sequential component (species component)

$S ::= (\alpha, \kappa)\ \mathrm{op}\ S \mid S + S \mid C \qquad \text{where op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$

The syntax

Sequential component (species component)

$S ::= (\alpha, \kappa)\ \mathrm{op}\ S \mid S + S \mid C \qquad \text{where op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$

The syntax

Sequential component (species component)

$S ::= (\alpha, \kappa)\ \mathbf{op}\ S \mid S+S \mid C \qquad$ where $\mathbf{op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$

The syntax

Sequential component (species component)

$S ::= (\alpha, \kappa)\ \mathrm{op}\ S \mid S{+}S \mid C \qquad \text{where op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$

The syntax

Sequential component (species component)

$S ::= (\alpha, \kappa)\ \mathrm{op}\ S \mid S + S \mid C \qquad \text{where op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$

The syntax

Sequential component (species component)

$$S ::= (\alpha, \kappa) \text{ op } S \mid S + S \mid C \qquad \text{where op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$$

Model component

$$P ::= P \underset{\mathcal{L}}{\bowtie} P \mid S(l)$$

The syntax

Sequential component (species component)

$$S ::= (\alpha, \kappa) \text{ op } S \mid S + S \mid C \qquad \text{where op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$$

Model component

$$P ::= P \underset{\mathcal{L}}{\bowtie} P \mid S(l)$$

The syntax

Sequential component (species component)

$$S ::= (\alpha, \kappa) \text{ op } S \mid S + S \mid C \qquad \text{where op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$$

Model component

$$P ::= P \underset{\mathcal{L}}{\bowtie} P \mid S(l)$$

The syntax

Sequential component (species component)

$$S ::= (\alpha, \kappa) \text{ op } S \mid S+S \mid C \qquad \text{where op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$$

Model component

$$P ::= P \underset{\mathcal{L}}{\bowtie} P \mid S(l)$$

The parameter l is abstract, recording quantitative information about the species.

The syntax

Sequential component (species component)

$S ::= (\alpha, \kappa)\ \mathrm{op}\ S \mid S + S \mid C \qquad \text{where op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$

Model component

$$P ::= P \underset{\mathcal{L}}{\bowtie} P \mid S(l)$$

The parameter l is abstract, recording quantitative information about the species.

Depending on the interpretation, this quantity may be:

- number of molecules (SSA),
- concentration (ODE) or
- a level within a semi-quantitative model (CTMC).

The Bio-PEPA system

A Bio-PEPA system $\mathcal{P}$ is a 6-tuple $\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}_R, \mathit{Comp}, P \rangle$, where:

- $\mathcal{V}$ is the set of compartments;
- $\mathcal{N}$ is the set of quantities describing each species (step size, number of levels, location, ...);
- $\mathcal{K}$ is the set of parameter definitions;
- $\mathcal{F}_R$ is the set of functional rate definitions;
- *Comp* is the set of definitions of sequential components;
- P is the model component describing the system.

Semantics

The semantics of Bio-PEPA is given as a small-step operational semantics, intended for deriving the CTMC with levels.

Semantics

The semantics of Bio-PEPA is given as a small-step operational semantics, intended for deriving the CTMC with levels.

We define two relations over the processes:

1. capability relation, that supports the derivation of quantitative information;

Semantics

The semantics of Bio-PEPA is given as a small-step operational semantics, intended for deriving the CTMC with levels.

We define two relations over the processes:

1. capability relation, that supports the derivation of quantitative information;
2. stochastic relation, that gives the rates associated with each action.

Semantics: prefix rules

$$\texttt{prefixReac} \quad ((\alpha,\kappa)\downarrow S)(l) \xrightarrow{(\alpha,[S:\downarrow(l,\kappa)])}_c S(l-\kappa) \qquad \kappa \leq l \leq N$$

Semantics: prefix rules

$$\texttt{prefixReac} \quad ((\alpha,\kappa)\downarrow S)(l) \xrightarrow{(\alpha,[S:\downarrow(l,\kappa)])}_c S(l-\kappa) \qquad \kappa \leq l \leq N$$

$$\texttt{prefixProd} \quad ((\alpha,\kappa)\uparrow S)(l) \xrightarrow{(\alpha,[S:\uparrow(l,\kappa)])}_c S(l+\kappa) \qquad 0 \leq l \leq (N-\kappa)$$

Semantics: prefix rules

$$\texttt{prefixReac} \quad ((\alpha,\kappa)\downarrow S)(l) \xrightarrow{(\alpha,[S:\downarrow(l,\kappa)])}_c S(l-\kappa) \qquad \kappa \leq l \leq N$$

$$\texttt{prefixProd} \quad ((\alpha,\kappa)\uparrow S)(l) \xrightarrow{(\alpha,[S:\uparrow(l,\kappa)])}_c S(l+\kappa) \qquad 0 \leq l \leq (N-\kappa)$$

$$\texttt{prefixMod} \quad ((\alpha,\kappa)\, op\, S)(l) \xrightarrow{(\alpha,[S:op(l,\kappa)])}_c S(l) \qquad 0 < l \leq N$$

with $op = \odot, \oplus$, or $\ominus$

Semantics: constant and choice rules

$$\texttt{Choice1} \quad \frac{S_1(l) \xrightarrow{(\alpha,v)}_c S'_1(l')}{(S_1 + S_2)(l) \xrightarrow{(\alpha,v)}_c S'_1(l')}$$

Semantics: constant and choice rules

$$\texttt{Choice1} \quad \frac{S_1(l) \xrightarrow{(\alpha,v)}_c S'_1(l')}{(S_1 + S_2)(l) \xrightarrow{(\alpha,v)}_c S'_1(l')}$$

$$\texttt{Choice2} \quad \frac{S_2(l) \xrightarrow{(\alpha,v)}_c S'_2(l')}{(S_1 + S_2)(l) \xrightarrow{(\alpha,v)}_c S'_2(l')}$$

Semantics: constant and choice rules

$$\texttt{Choice1} \quad \frac{S_1(l) \xrightarrow{(\alpha,v)}_c S'_1(l')}{(S_1 + S_2)(l) \xrightarrow{(\alpha,v)}_c S'_1(l')}$$

$$\texttt{Choice2} \quad \frac{S_2(l) \xrightarrow{(\alpha,v)}_c S'_2(l')}{(S_1 + S_2)(l) \xrightarrow{(\alpha,v)}_c S'_2(l')}$$

$$\texttt{Constant} \quad \frac{S(l) \xrightarrow{(\alpha,S:[op(l,\kappa))]}_c S'(l')}{C(l) \xrightarrow{(\alpha,C:[op(l,\kappa))]}_c S'(l')} \quad \text{with } C \stackrel{def}{=} S$$

Semantics: cooperation rules

$$\text{coop1} \quad \frac{P_1 \xrightarrow{(\alpha,v)}_c P_1'}{P_1 \underset{\mathcal{L}}{\bowtie} P_2 \xrightarrow{(\alpha,v)}_c P_1' \underset{\mathcal{L}}{\bowtie} P_2} \quad \text{with } \alpha \notin \mathcal{L}$$

Semantics: cooperation rules

$$\text{coop1} \quad \frac{P_1 \xrightarrow{(\alpha,v)}_c P_1'}{P_1 \underset{\mathcal{L}}{\bowtie} P_2 \xrightarrow{(\alpha,v)}_c P_1' \underset{\mathcal{L}}{\bowtie} P_2} \quad \text{with } \alpha \notin \mathcal{L}$$

$$\text{coop2} \quad \frac{P_2 \xrightarrow{(\alpha,v)}_c P_2'}{P_1 \underset{\mathcal{L}}{\bowtie} P_2 \xrightarrow{(\alpha,v)}_c P_1 \underset{\mathcal{L}}{\bowtie} P_2'} \quad \text{with } \alpha \notin \mathcal{L}$$

Semantics: cooperation rules

$$\texttt{coop1} \quad \frac{P_1 \xrightarrow{(\alpha,v)}_c P_1'}{P_1 \underset{\mathcal{L}}{\bowtie} P_2 \xrightarrow{(\alpha,v)}_c P_1' \underset{\mathcal{L}}{\bowtie} P_2} \quad \text{with } \alpha \notin \mathcal{L}$$

$$\texttt{coop2} \quad \frac{P_2 \xrightarrow{(\alpha,v)}_c P_2'}{P_1 \underset{\mathcal{L}}{\bowtie} P_2 \xrightarrow{(\alpha,v)}_c P_1 \underset{\mathcal{L}}{\bowtie} P_2'} \quad \text{with } \alpha \notin \mathcal{L}$$

$$\texttt{coopFinal} \quad \frac{P_1 \xrightarrow{(\alpha,v_1)}_c P_1' \qquad P_2 \xrightarrow{(\alpha,v_2)}_c P_2'}{P_1 \underset{\mathcal{L}}{\bowtie} P_2 \xrightarrow{(\alpha,v_1::v_2)}_c P_1' \underset{\mathcal{L}}{\bowtie} P_2'} \quad \text{with } \alpha \in \mathcal{L}$$

Semantics: rates and transition system

In order to derive the rates we consider the stochastic relation $\rightarrow_s \subseteq \mathcal{P} \times \Gamma \times \mathcal{P}$, with $\gamma \in \Gamma := (\alpha, r)$ and $r \in \mathbb{R}^+$.

Semantics: rates and transition system

In order to derive the rates we consider the stochastic relation $\rightarrow_s \subseteq \mathcal{P} \times \Gamma \times \mathcal{P}$, with $\gamma \in \Gamma := (\alpha, r)$ and $r \in \mathbb{R}^+$.

The relation is defined in terms of the previous one:

Semantics: rates and transition system

In order to derive the rates we consider the stochastic relation $\rightarrow_s \subseteq \mathcal{P} \times \Gamma \times \mathcal{P}$, with $\gamma \in \Gamma := (\alpha, r)$ and $r \in \mathbb{R}^+$.

The relation is defined in terms of the previous one:

$$\frac{P \xrightarrow{(\alpha_j, v)}_c P'}{\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}_R, Comp, P \rangle \xrightarrow{(\alpha_j, r_{\alpha_j})}_s \langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}_R, Comp, P' \rangle}$$

Semantics: rates and transition system

In order to derive the rates we consider the stochastic relation $\rightarrow_s \subseteq \mathcal{P} \times \Gamma \times \mathcal{P}$, with $\gamma \in \Gamma := (\alpha, r)$ and $r \in \mathbb{R}^+$.

The relation is defined in terms of the previous one:

$$\frac{P \xrightarrow{(\alpha_j, v)}_c P'}{\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}_R, Comp, P \rangle \xrightarrow{(\alpha_j, r_{\alpha_j})}_s \langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}_R, Comp, P' \rangle}$$

r_{α_j} represents the parameter of an exponential distribution and the dynamic behaviour is determined by a race condition.

.

Semantics: rates and transition system

In order to derive the rates we consider the stochastic relation $\rightarrow_s \subseteq \mathcal{P} \times \Gamma \times \mathcal{P}$, with $\gamma \in \Gamma := (\alpha, r)$ and $r \in \mathbb{R}^+$.

The relation is defined in terms of the previous one:

$$\frac{P \xrightarrow{(\alpha_j, v)}_c P'}{\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}_R, Comp, P \rangle \xrightarrow{(\alpha_j, r_{\alpha_j})}_s \langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}_R, Comp, P' \rangle}$$

r_{α_j} represents the parameter of an exponential distribution and the dynamic behaviour is determined by a race condition.

The rate r_{α_j} is defined as $f_{\alpha_j}(\mathcal{V}, \mathcal{N}, \mathcal{K})/h$.

Analysis

A Bio-PEPA system is a formal, intermediate and compositional representation of the system.

Analysis

A Bio-PEPA system is a formal, intermediate and compositional representation of the system.

From it we can obtain

Analysis

A Bio-PEPA system is a formal, intermediate and compositional representation of the system.

From it we can obtain

- a CTMC (with levels)

Analysis

A Bio-PEPA system is a formal, intermediate and compositional representation of the system.

From it we can obtain

- a CTMC (with levels)
- a ODE system for simulation and other kinds of analysis

Analysis

A Bio-PEPA system is a formal, intermediate and compositional representation of the system.

From it we can obtain

- a CTMC (with levels)
- a ODE system for simulation and other kinds of analysis
- a Gillespie model for stochastic simulation

Analysis

A Bio-PEPA system is a formal, intermediate and compositional representation of the system.

From it we can obtain

- a CTMC (with levels)
- a ODE system for simulation and other kinds of analysis
- a Gillespie model for stochastic simulation
- a PRISM model for model checking

Analysis

A Bio-PEPA system is a formal, intermediate and compositional representation of the system.

From it we can obtain

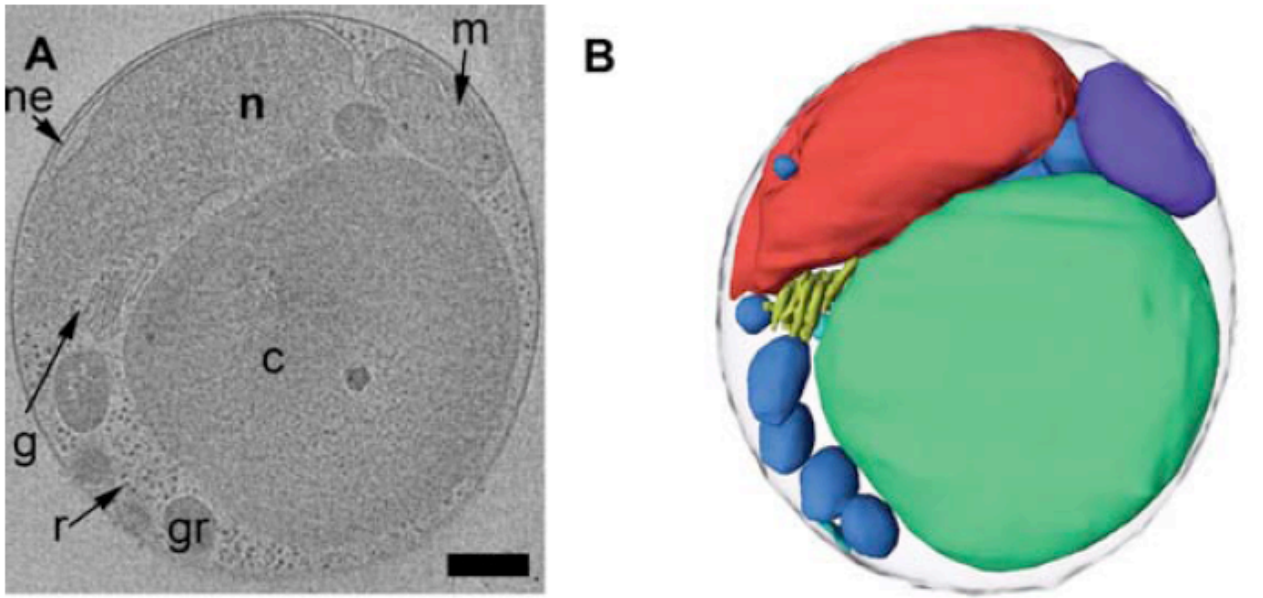
- a CTMC (with levels)
- a ODE system for simulation and other kinds of analysis
- a Gillespie model for stochastic simulation
- a PRISM model for model checking

Each of these kinds of analysis can be of help for studying different aspects of the biological model. Moreover we are exploring how they can be used in conjunction.

Outline

Circadian Rhythms in *Ostreococcus Tauri*

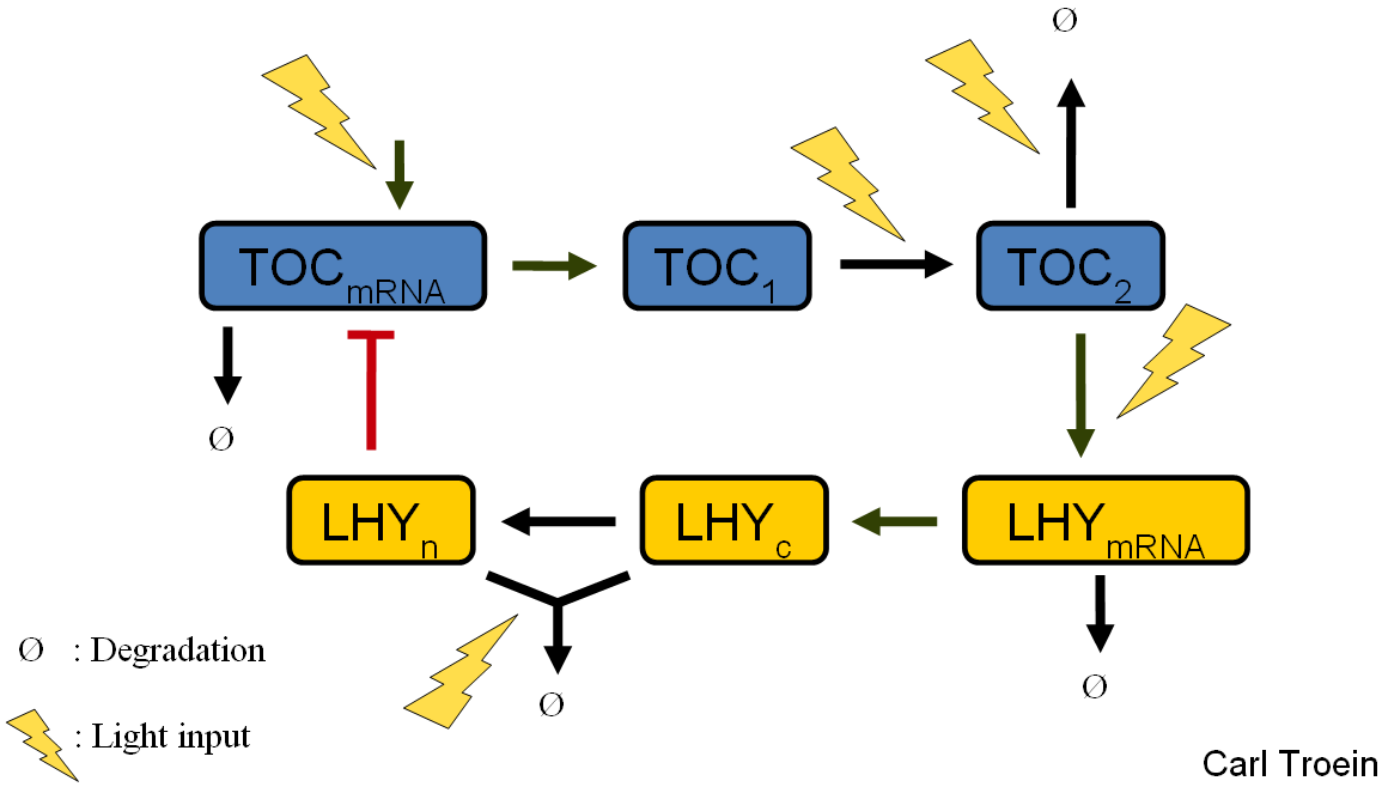
Jane Hillston.
University of Edinburgh.



The green alga *Ostreococcus Tauri* is an ideal model system for understanding the function of plant clocks.

In particular we use a Bio-PEPA model to study the variability and robustness of the clock's functional behaviour with respect to internal stochastic noise and environmental changes.

The biological model



- Feedback loop between the two main genes (TOC1 and LHY).
- Effect of light on several parts of the network.

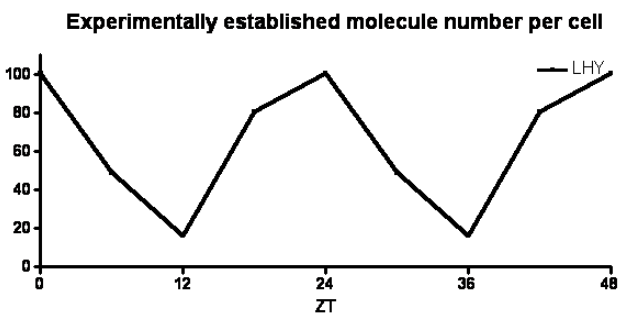
Modelling overview

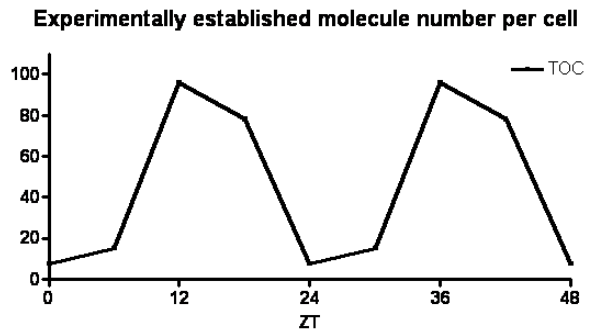
- A previous deterministic (ODE) model of the system had been developed by hand.
- We developed a Bio-PEPA model and validated that its continuous-deterministic interpretation coincided with the original model.
- Stochastic simulation was used with the discrete-stochastic interpretation of the Bio-PEPA model to investigate stochastic fluctuations, focusing on clock phase and sensitivity analysis.
- Model-checking was also used to give more detailed analysis of variability within the system, via specific questions about model behaviour.

Model and experimental settings

- Lab experiments in different light conditions and photoperiods
  - DD (24 hours dark)
  - LL (24 hours light)
  - LD 12:12 (12 hours light / 12 hours dark)
  - LD 6:18 (6 hours light / 18 hours dark)
  - LD 18:6 (18 hours light / 6 hours dark)

- Lab measurements on amounts of proteins





Modelling in Bio-PEPA

- Biological species ⇒ interacting species components
- Reactions ⇒ actions, with functional rates expressing the kinetic rate laws
- Light on/off mechanism for entrainment to day/night cycle ⇒ events switching between day-time and night-time rates

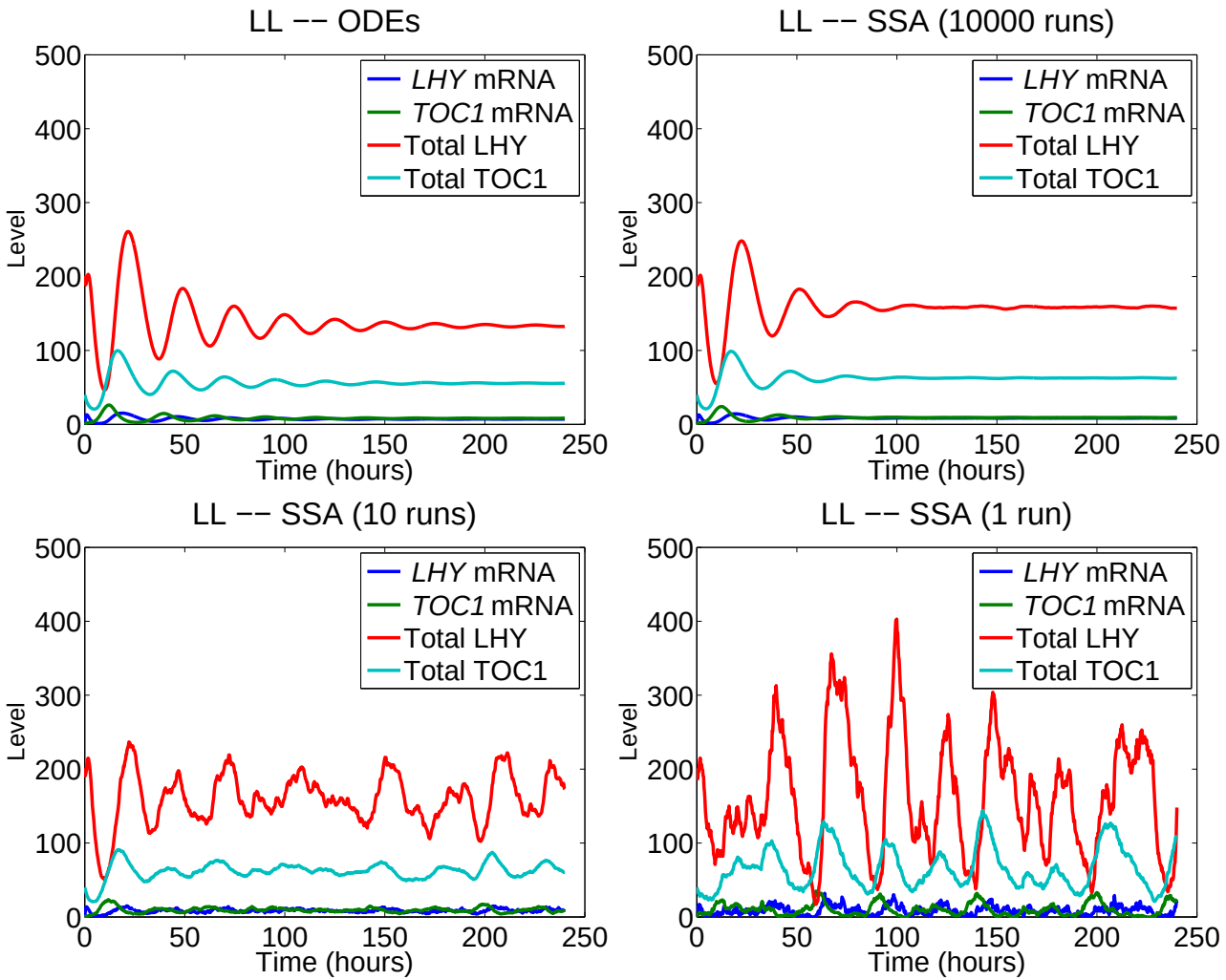
Extracts from the Bio-PEPA model

$$
\begin{aligned}
TOC1\_mRNA &= (transc_3, 1)\uparrow + (transl_5, 1) \oplus + (deg_7, 1)\downarrow \\
TOC1\_a &= (deg_4, 1)\downarrow + (conv_6, 1)\uparrow + (transc_8, 1) \oplus \\
TOC1\_i &= (transl_5, 1)\uparrow + (conv_6, 1)\downarrow \\
\\
LHY\_mRNA &= (transc_8, 1)\uparrow + (deg_9, 1)\downarrow + (transl_{10}, 1) \oplus \\
LHY\_c &= (transl_{10}, 1)\uparrow + (transp_{11}, 1)\downarrow + (deg_{12}, 1)\downarrow \\
LHY\_n &= (transc_3, 1) \ominus + (transp_{11}, 1)\uparrow + (deg_{13}, 1)\downarrow \\
\\
acc &= (prod_1, 1)\uparrow + (deg_2, 1)\downarrow + (transc_3, 1) \oplus
\end{aligned}
$$

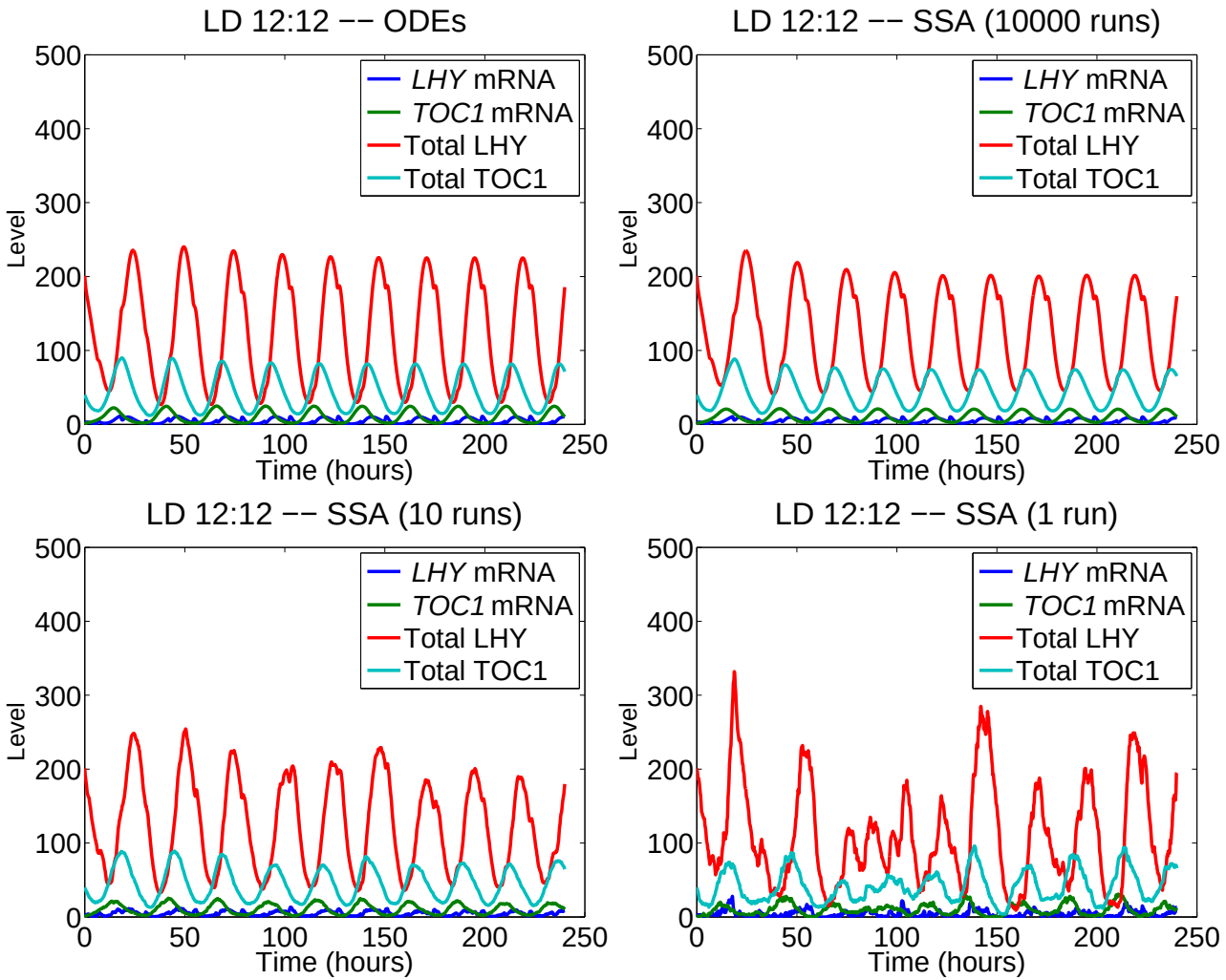
$$
\begin{aligned}
\text{total\_TOC1} &= \text{TOC1\_i} + \text{TOC1\_a} \\
\text{total\_LHY} &= \text{LHY\_c} + \text{LHY\_n}
\end{aligned}
$$

$t\_dawn = 6, \quad t\_dusk = 18$ (e.g. for LD 12:12 system)

Simulations – constant light (LL)



Simulations – light/dark (LD 12:12)



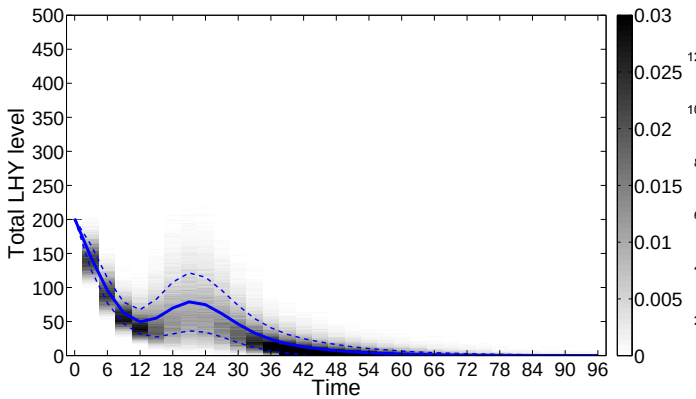
Model-checking

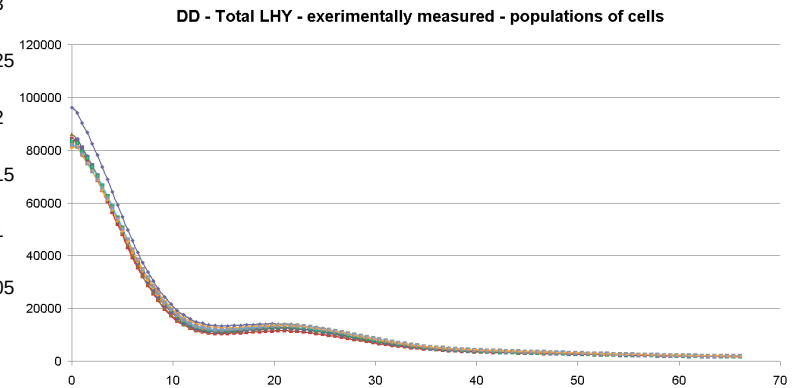
- Here we use statistical model-checking (discrete event simulation and sampling over multiple runs): approximate results.
- The underlying simulation model is enhanced with a representation of time so that rates of reactions can change appropriately at dawn and dusk.
- The model checker can then be used to investigate the probability distribution of the number of molecules of a species over time.

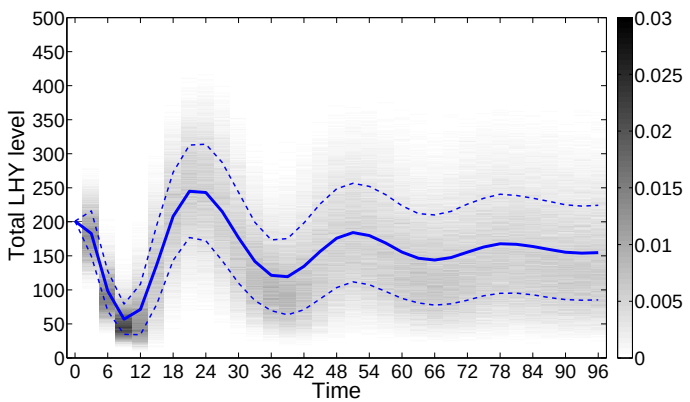
Model-checking – probability distribution of LHY value over 0-96 h

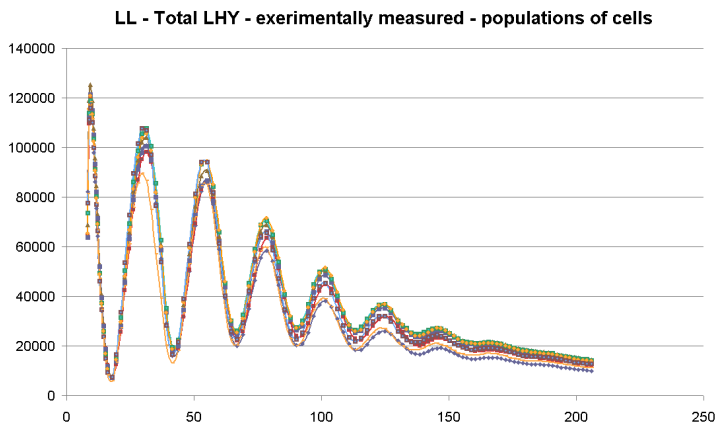
Jane Hillston.
University of Edinburgh.

$$[T, T](\_LHY\_c + \_LHY\_n = level), T = 0 : 3 : 96, level = 0 : 1 : 500$$









Probability that the total LHY stays below some threshold e in the long run

$$[96, 500](LHY\_c + LHY\_n \leq 0 + e)$$

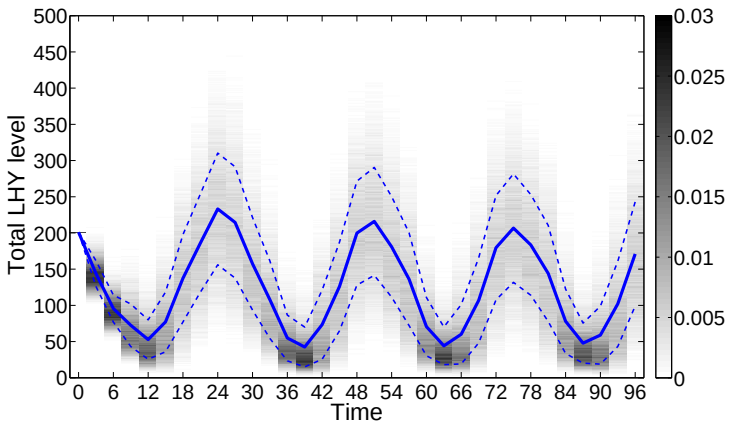
e	0	1	2	3	4
Prob	0.93	0.96	0.96	0.98	0.98

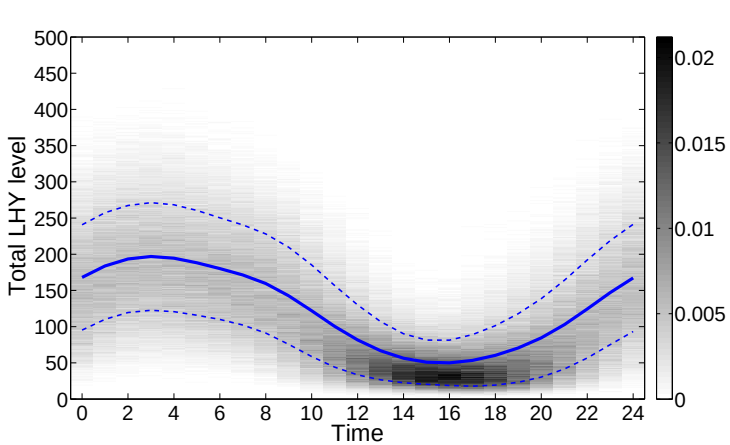
e	5	6	7	8	9	10
Prob	0.99	0.99	0.99	0.99	0.99	1.0

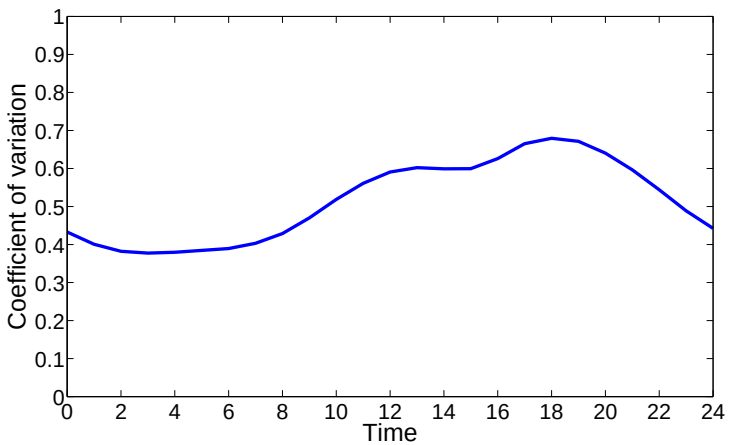
Probability distribution/coefficient of variation of LHY value – LD 12:12

Jane Hillston.
University of Edinburgh.

$$[T, T](\_LHY\_c + \_LHY\_n = level), T = 120 : 1 : 144, level = 0 : 1 : 500$$







Outline

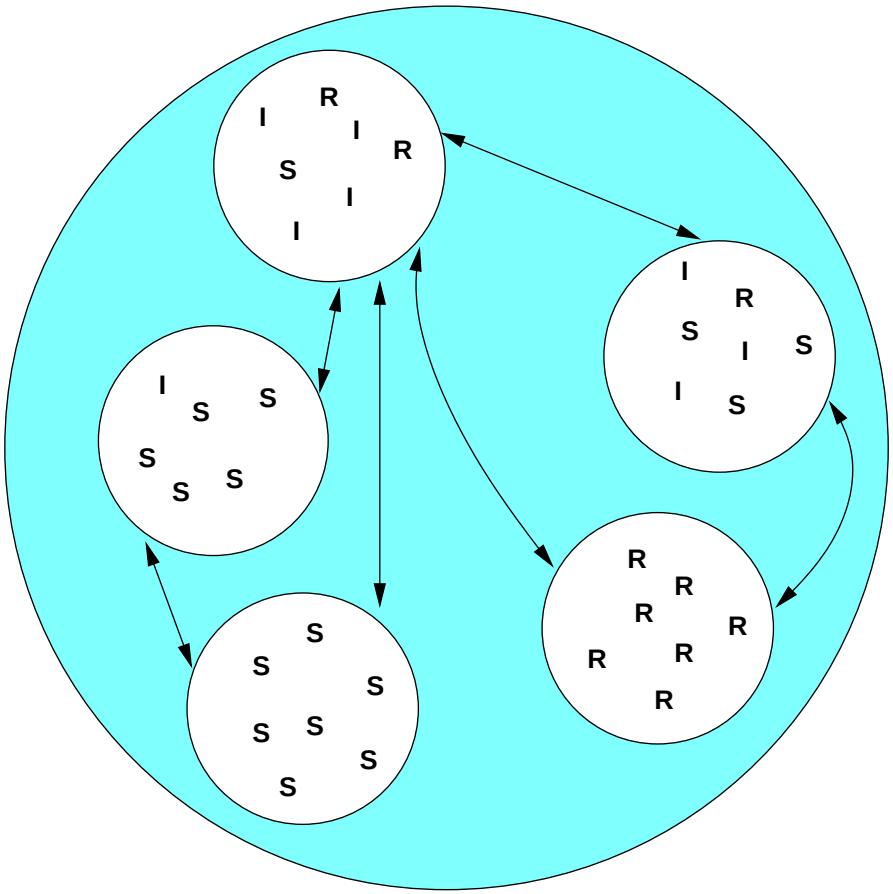
Epidemiological models

Epidemiological models study the spread of disease within a population.

For this purpose the population can be divided into classes: susceptibles (S), those who are infective (I) and those who have recovered (R) and are immune.

Refinements of this might include differentiating those which are symptomatic or asymptomatic, treated or untreated, or suffering from a drug-resistant form of the disease.

Spatial structures also impact infection spread



Jane Hillston.
University of Edinburgh.

Using Bio-PEPA for epidemiological models

- Each species will correspond to a subset of individuals (e.g. susceptible, infective or recovered individuals), possibly also differentiating locations.
- The role of the individual with respect to an action can be used to indicate that the species decreases, remains invariant or increases in an interaction.
- Interactions such as $I + S \rightarrow 2I$ are possible, where an entity is present on both sides of the interaction with different multiplicity. Note that this cannot be represented in Bio-PEPA.

Modifying Bio-PEPA for epidemiology

The key change is that two multiplicities can be associated with an action in a component, rather than the single stoichiometry used in biochemistry:

Modifying Bio-PEPA for epidemiology

The key change is that two multiplicities can be associated with an action in a component, rather than the single stoichiometry used in biochemistry:

A Bio-PEPA model for epidemiological system is described by the following syntax:

$$S ::= (\alpha, \kappa) \downarrow S \mid (\alpha, \kappa) \uparrow S \mid (\alpha, (\kappa_1, \kappa_2)) \odot S \mid S + S \mid C \mid S@L$$

$$P ::= P \underset{\mathcal{I}}{\bowtie} P \mid S(x)$$

Modifying Bio-PEPA for epidemiology

The key change is that two multiplicities can be associated with an action in a component, rather than the single stoichiometry used in biochemistry:

A Bio-PEPA model for epidemiological system is described by the following syntax:

$$S ::= (\alpha, \kappa) \downarrow S \mid (\alpha, \kappa) \uparrow S \mid (\alpha, (\kappa_1, \kappa_2)) \odot S \mid S + S \mid C \mid S@L$$

$$P ::= P \underset{\mathcal{I}}{\bowtie} P \mid S(x)$$

The abstraction

The translation of epidemiological models into Bio-PEPA is based on the following correspondences:

- Each subpopulation/patch is abstracted by a location.
- Each type of individual is represented by a species component, whose subterms describe its interaction capabilities.
- Each interaction is represented by an action type. The dynamics are described by a functional rate.
- The model component represents how the species interact and contains information about the initial state.

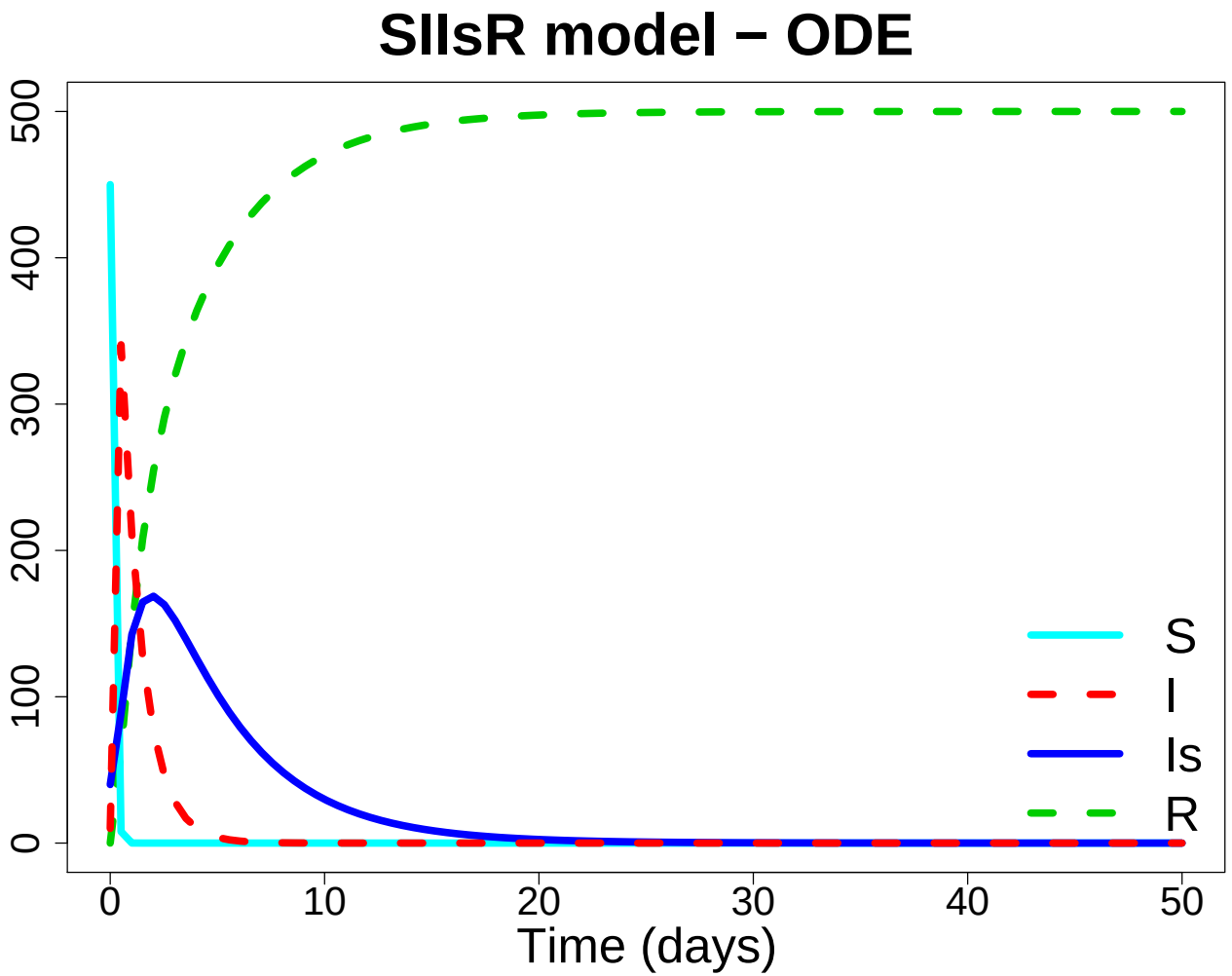
Simple model of H5N1 Avian Influenza: single location, no drug treatment

We distinguish between asymptomatic (I) and symptomatic (I_s) individuals:

$$
\begin{aligned}
S &\stackrel{def}{=} (contact1, 1)\downarrow S + (contact2, 1)\downarrow S \\
I &\stackrel{def}{=} (contact1, (1, 2)) \odot I + (contact2, 1)\uparrow I \\
&\quad + (recovery1, 1)\downarrow I + (symp, 1)\downarrow I \\
I_s &\stackrel{def}{=} (contact2, (1, 1)) \odot I_s + (recovery2, 1)\downarrow I_s \\
&\quad + (symp, 1)\uparrow I_s \\
R &\stackrel{def}{=} (recovery1, 1)\uparrow R + (recovery2, 1)\uparrow R
\end{aligned}
$$

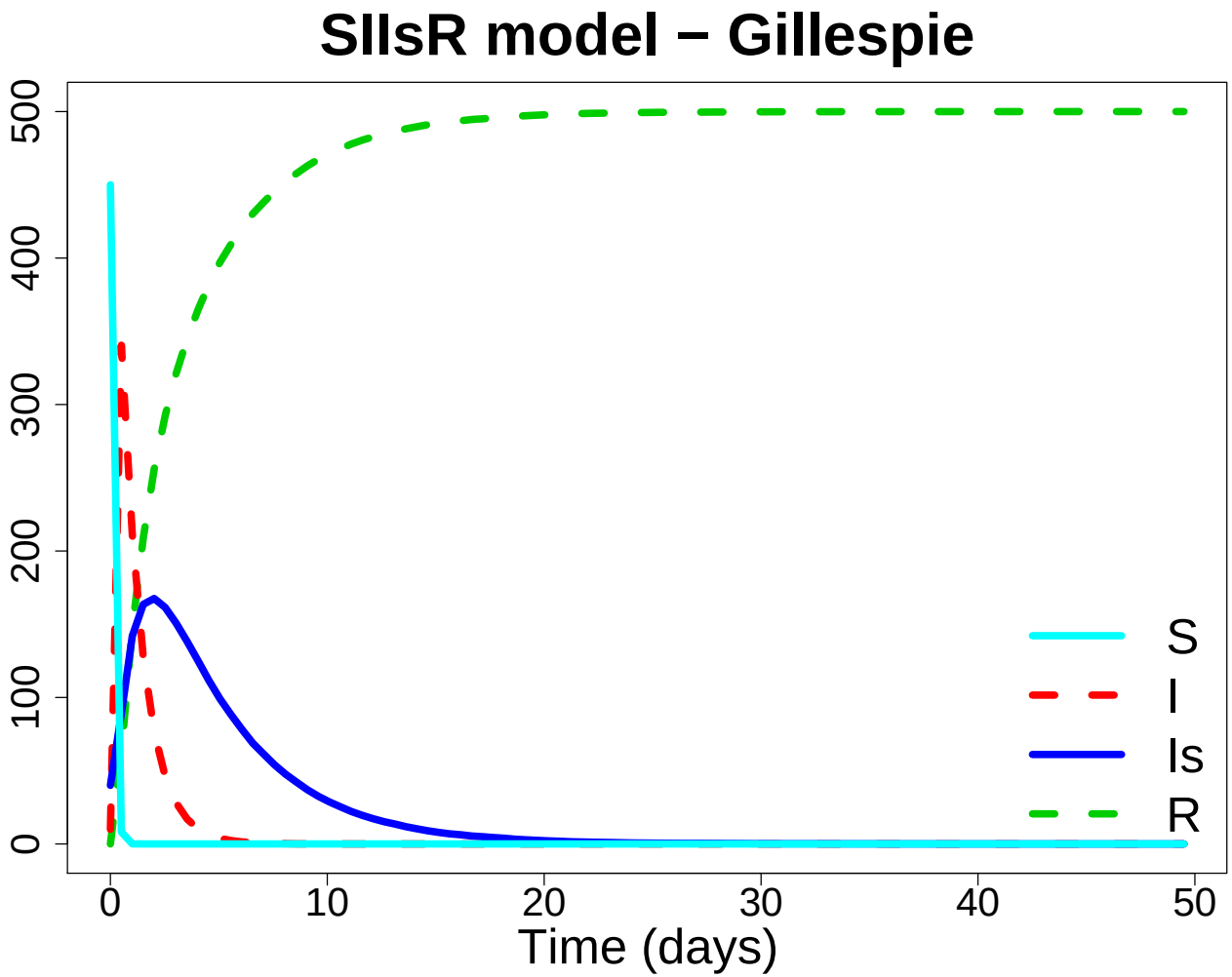
$$S(450) \underset{*}{\bowtie} I(10) \underset{*}{\bowtie} I_s(40) \underset{*}{\bowtie} R(0)$$

Simulation Results



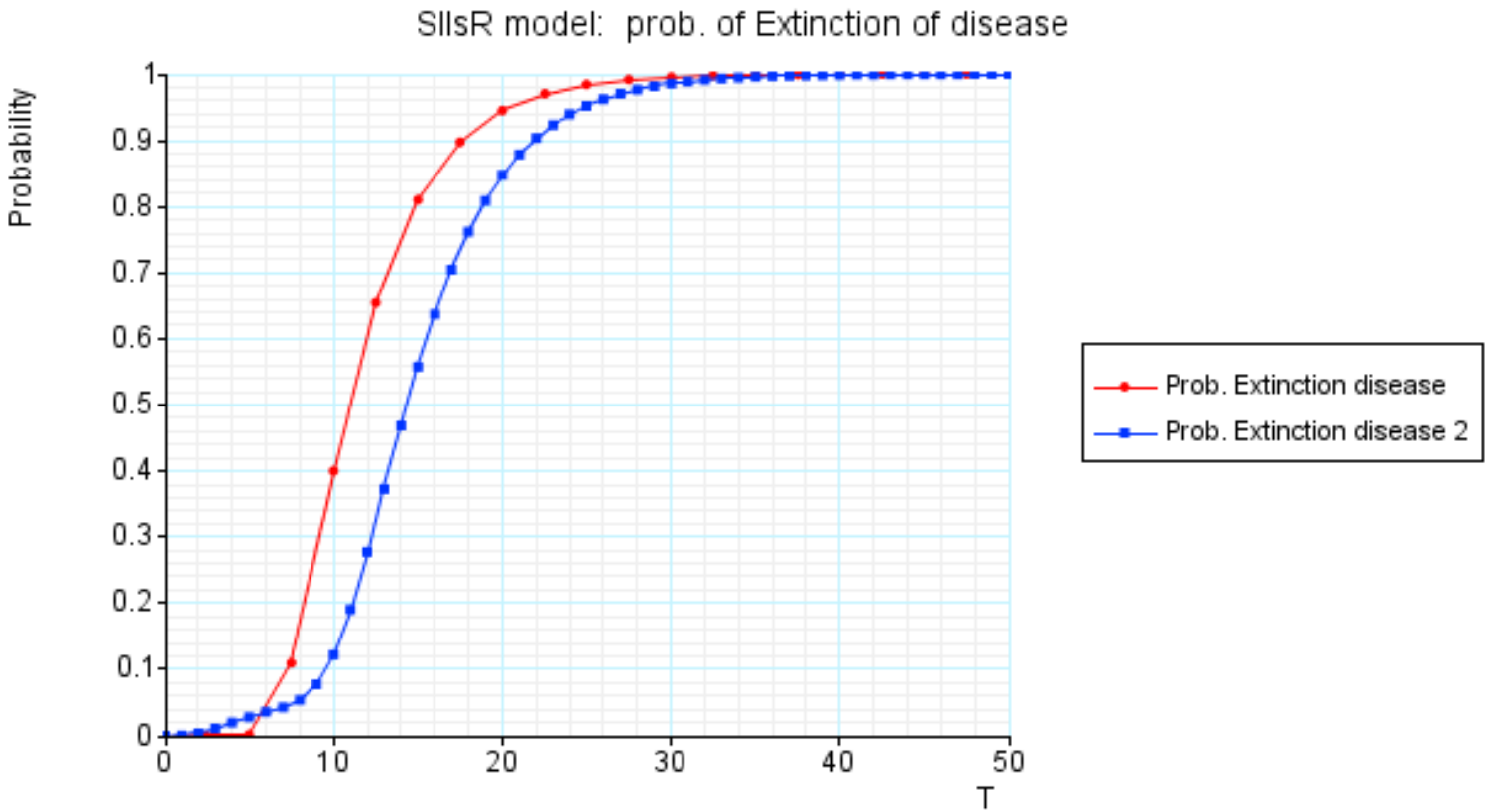
Jane Hillston.
University of Edinburgh.

Simulation Results



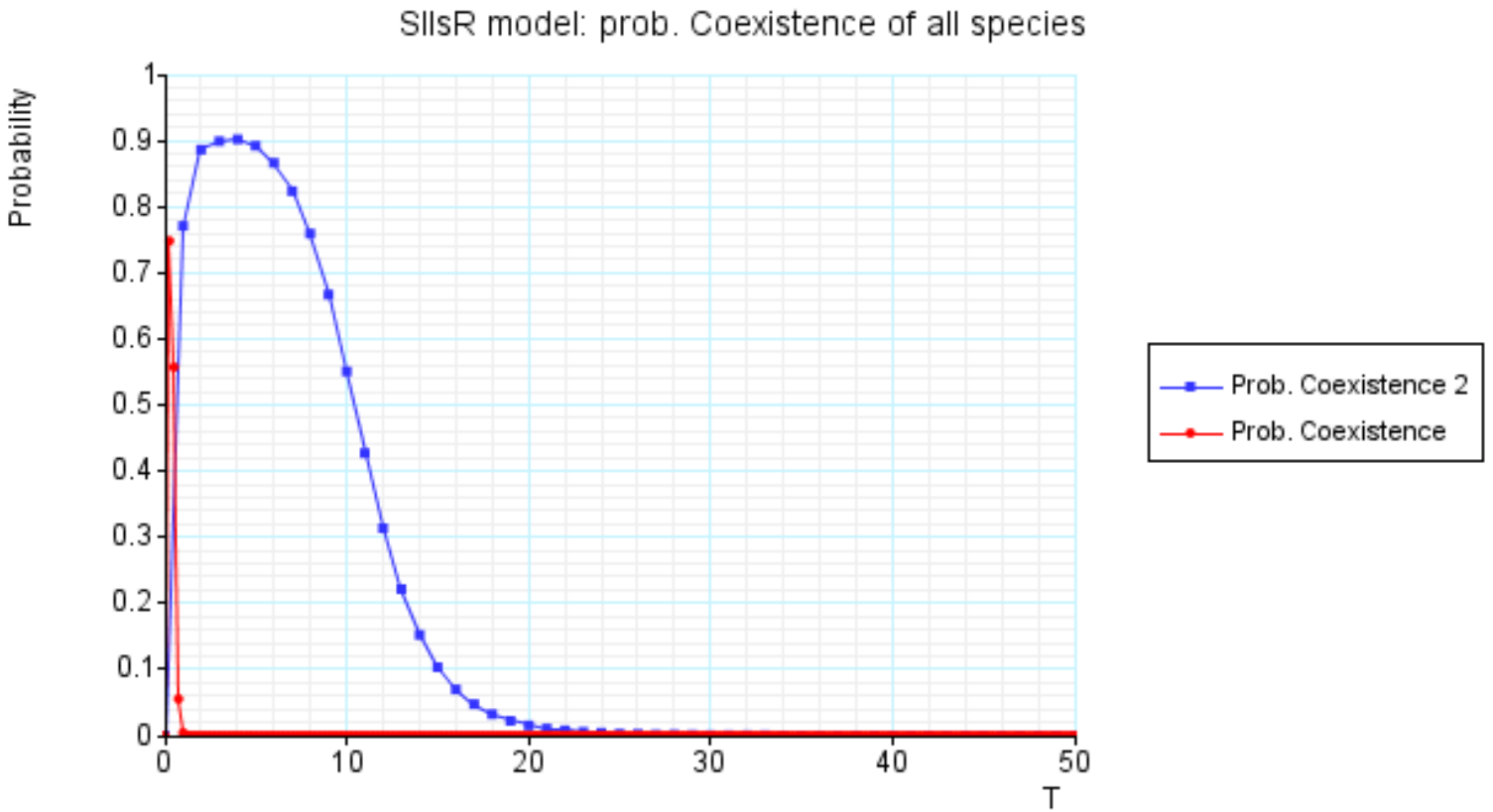
Jane Hillston.
University of Edinburgh.

PRISM Results — Varying contact rates



Jane Hillston.
University of Edinburgh.

PRISM Results — Varying contact rates



Jane Hillston.
University of Edinburgh.

Models with multiple locations

We investigated the impact of having multiple locations and the spatial arrangement of those locations.

Models with multiple locations

We investigated the impact of having multiple locations and the spatial arrangement of those locations.

Interactions between individuals are as in the previous model but constrained to only occur when they are in the same location.

Models with multiple locations

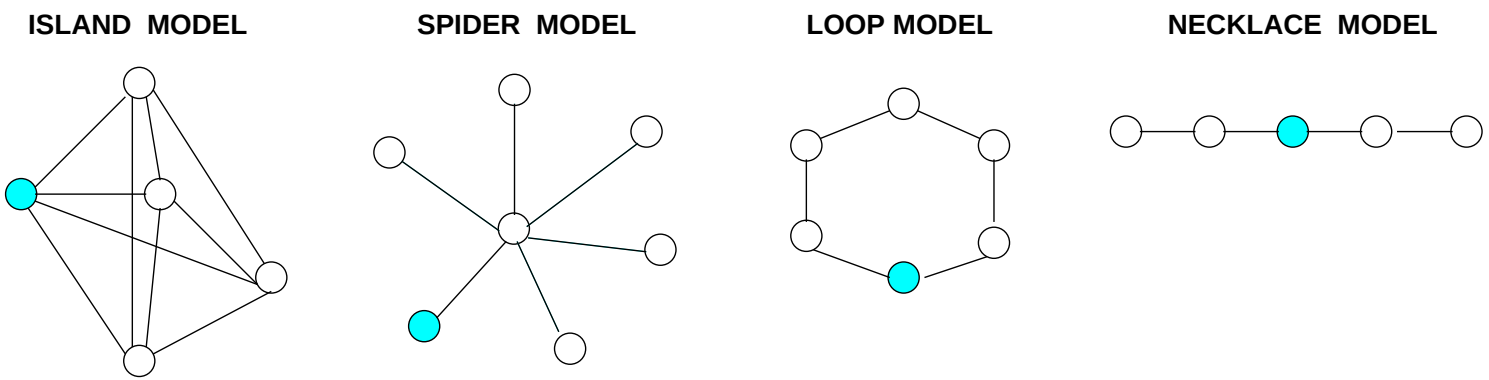
We investigated the impact of having multiple locations and the spatial arrangement of those locations.

Interactions between individuals are as in the previous model but constrained to only occur when they are in the same location.

Additionally there are migration actions, e.g:

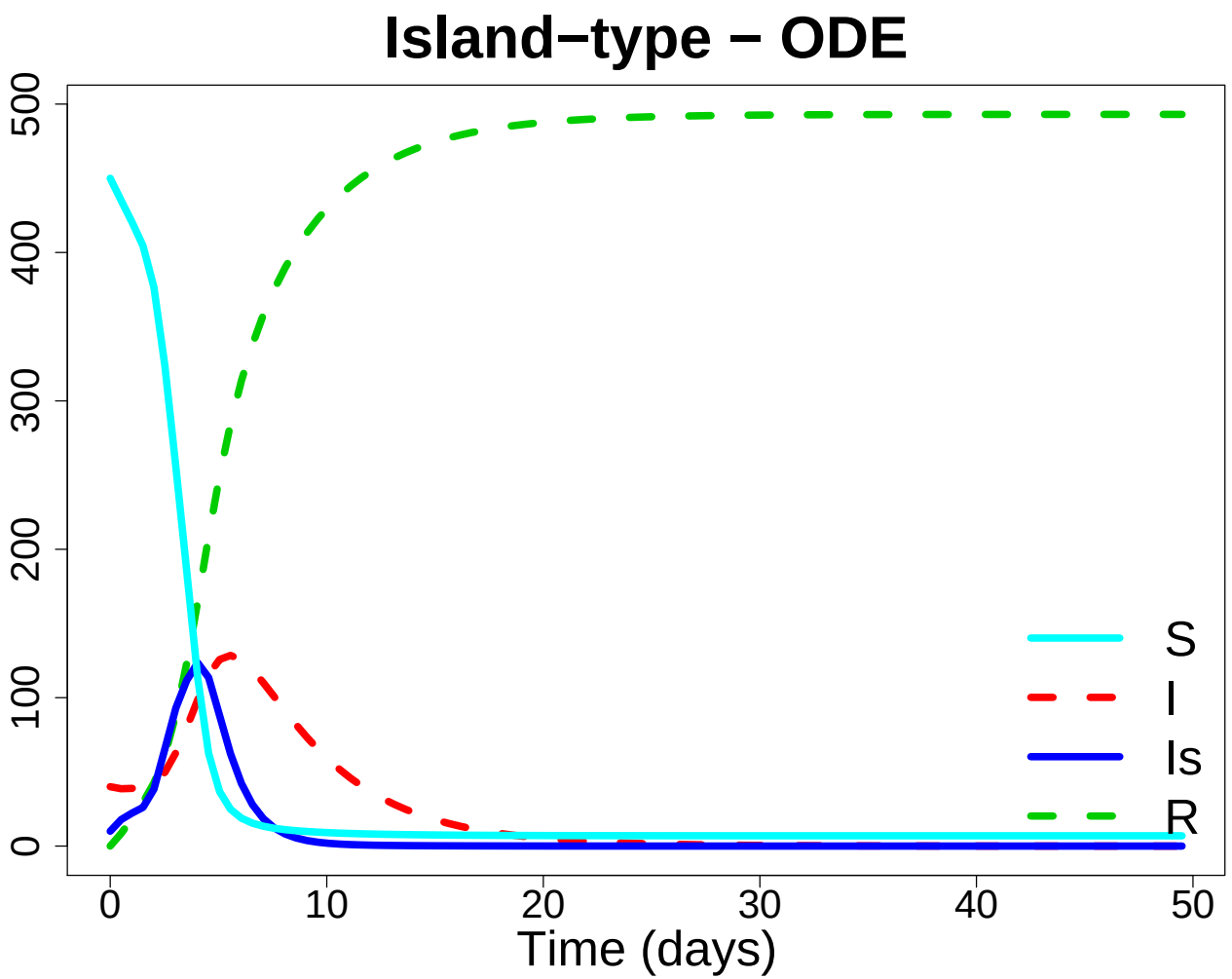
$$S \stackrel{def}{=} (contact1, 1)\downarrow S + (contact2, 1)\downarrow S + \sum_{M(i,j)\neq 0} (m_{ij,S}[location\_i \rightarrow location\_j], (1,1)) \odot S$$

Population structures



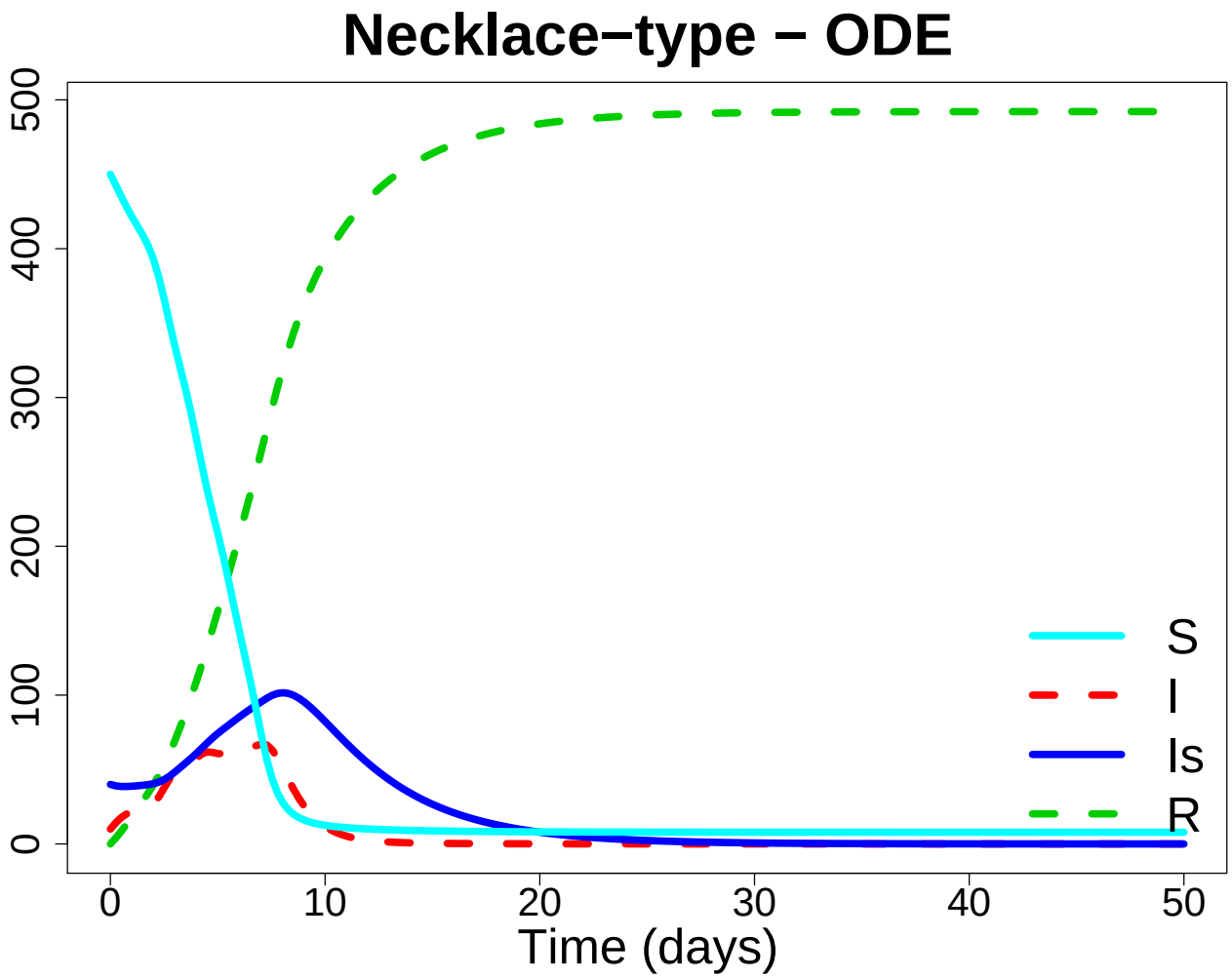
The connections between the patches indicate how individuals might move between patches, thus relaying infections through the metapopulation.

Comparing structures



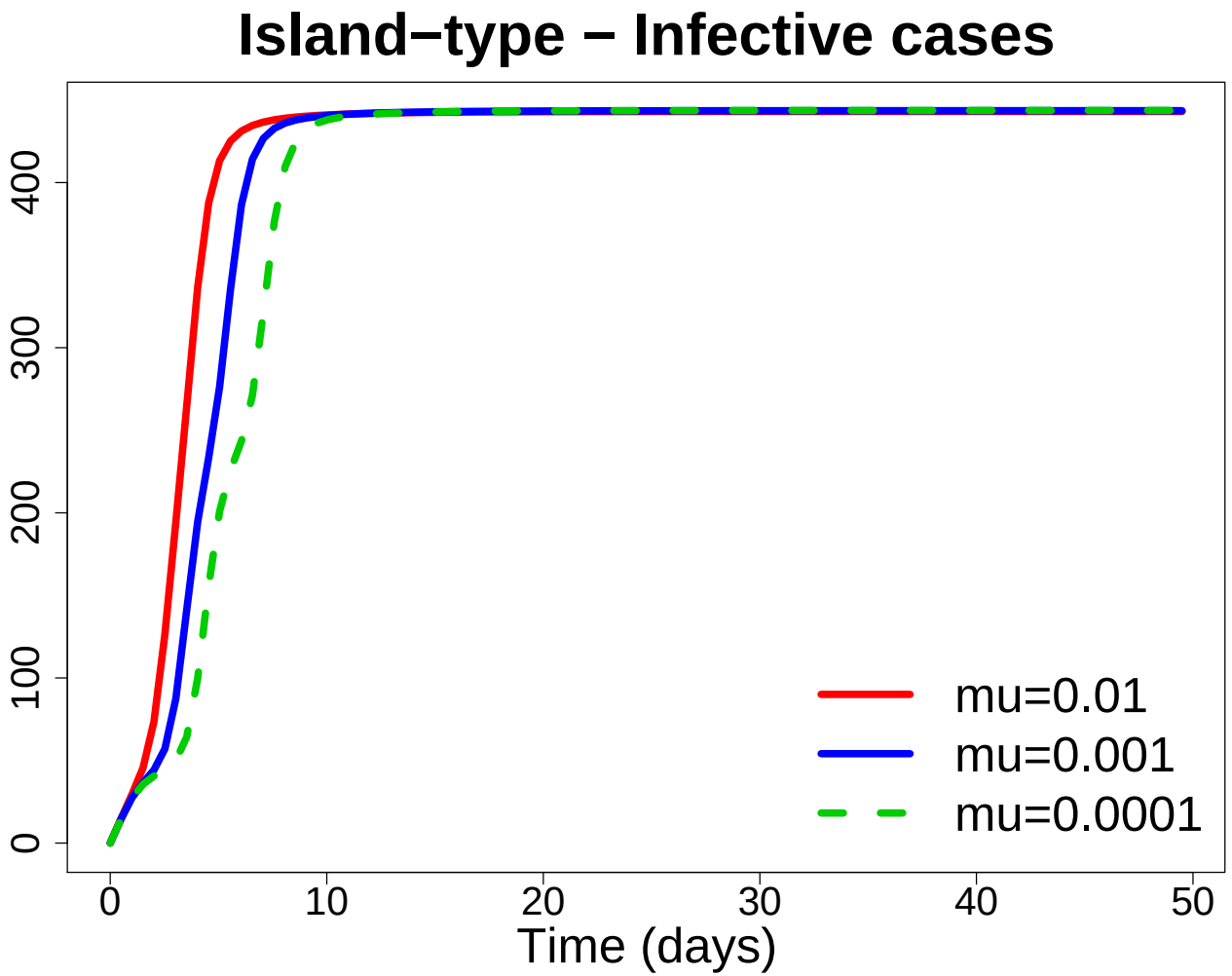
Jane Hillston.
University of Edinburgh.

Comparing structures



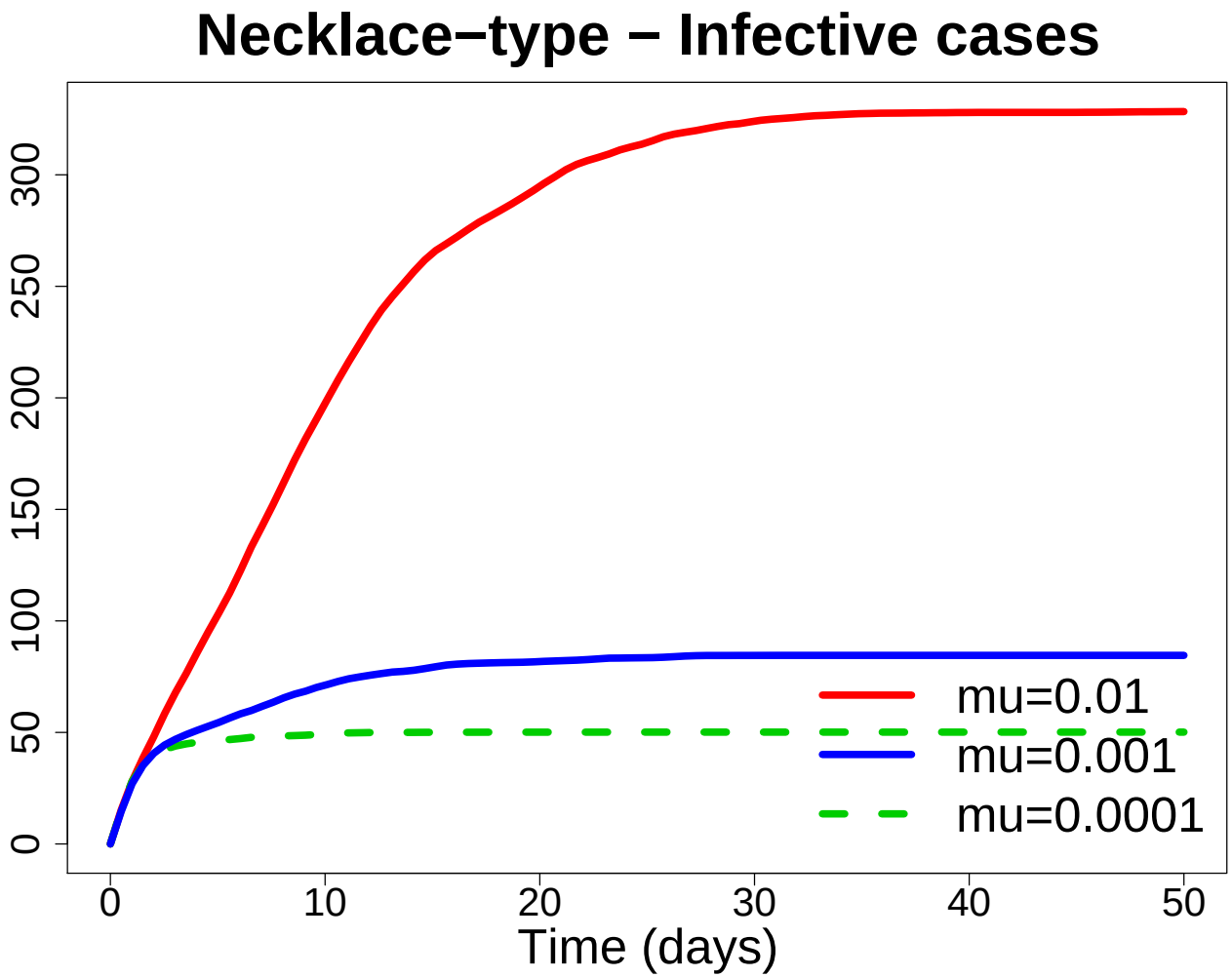
Jane Hillston.
University of Edinburgh.

The effect of the migration rate



Jane Hillston.
University of Edinburgh.

The effect of the migration rate

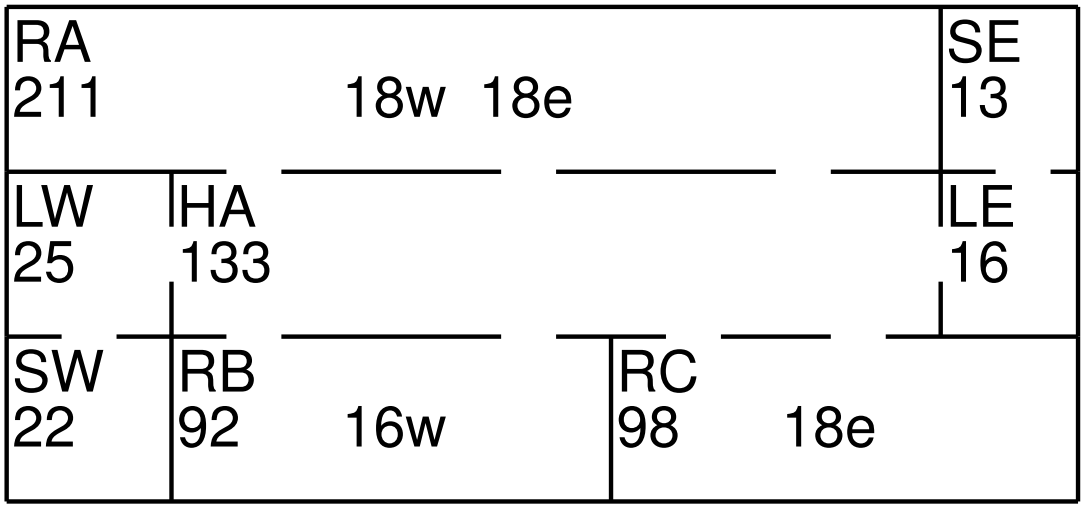


Jane Hillston.
University of Edinburgh.

Designing for human crowd dynamics

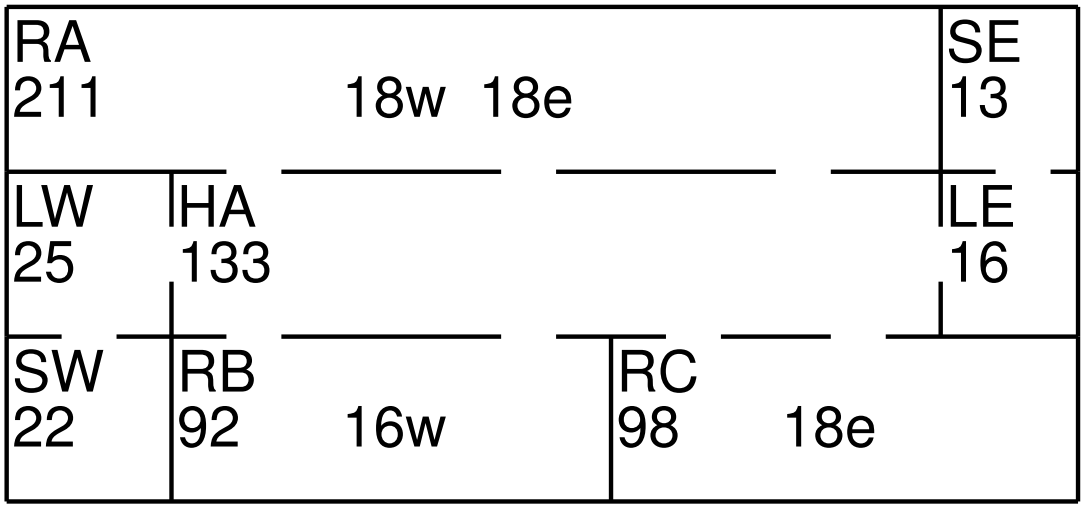
- Widespread take up of mobile and communicating computational devices is making ubiquitous systems a reality and creating new ways for us to interact with our environment.
- One application is to provide routing information to help people navigate through unfamiliar locations.
- In these case the dynamic behaviour of the system as a whole is important to ensure the satisfaction of the users.
- Emergency egress can be regarded as a particular case, when the location may be familiar but circumstances may alter the usual topology and make efficient movement particularly important

Example scenario



The layout of the building is described in terms of the arrangement of the rooms, hallways, landing and stairs. Each has a capacity and may have an initial occupancy.

Example scenario



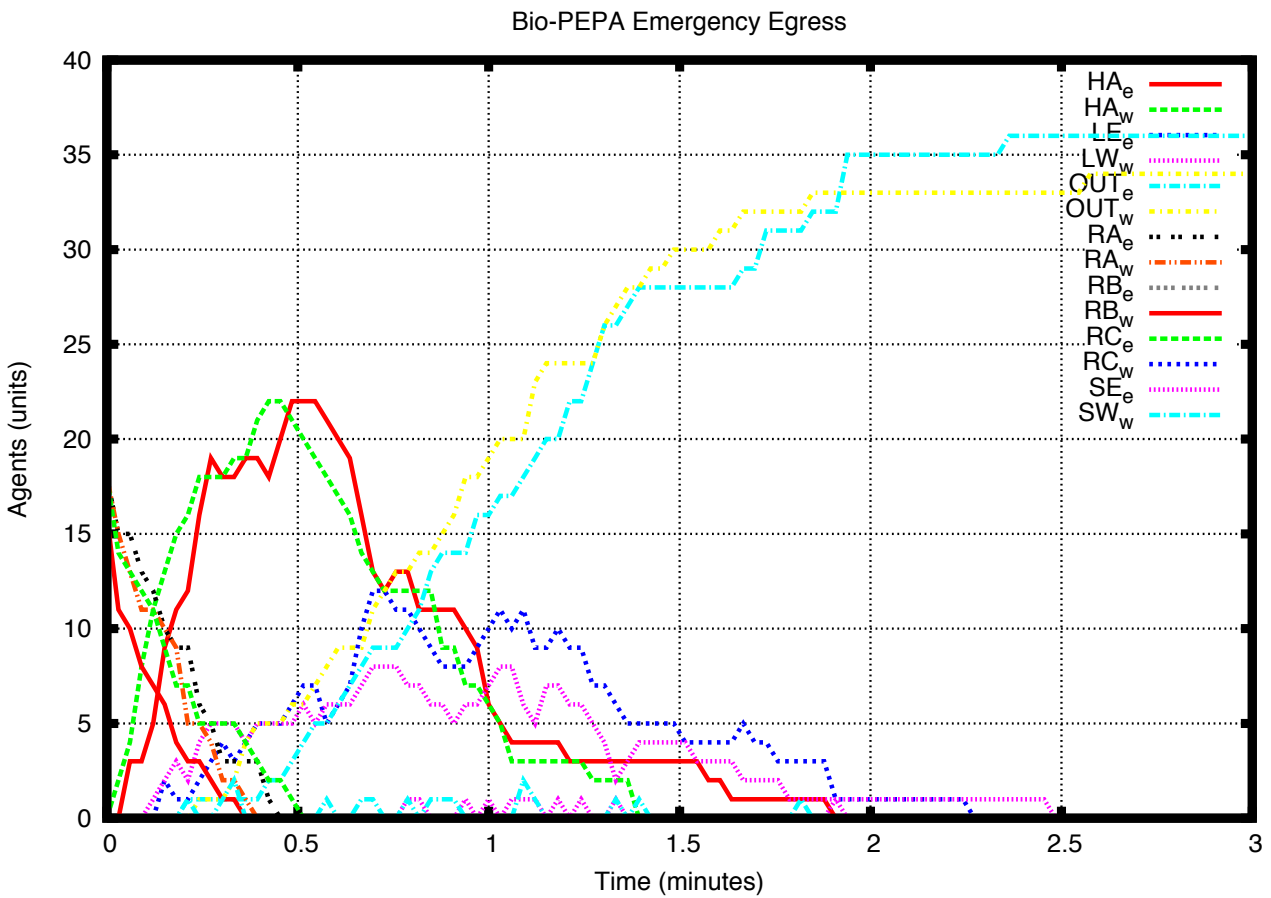
The layout of the building is described in terms of the arrangement of the rooms, hallways, landing and stairs. Each has a capacity and may have an initial occupancy.

Bio-PEPA species describe the behaviours of individuals, but also rooms and information dissemination.

Model specification

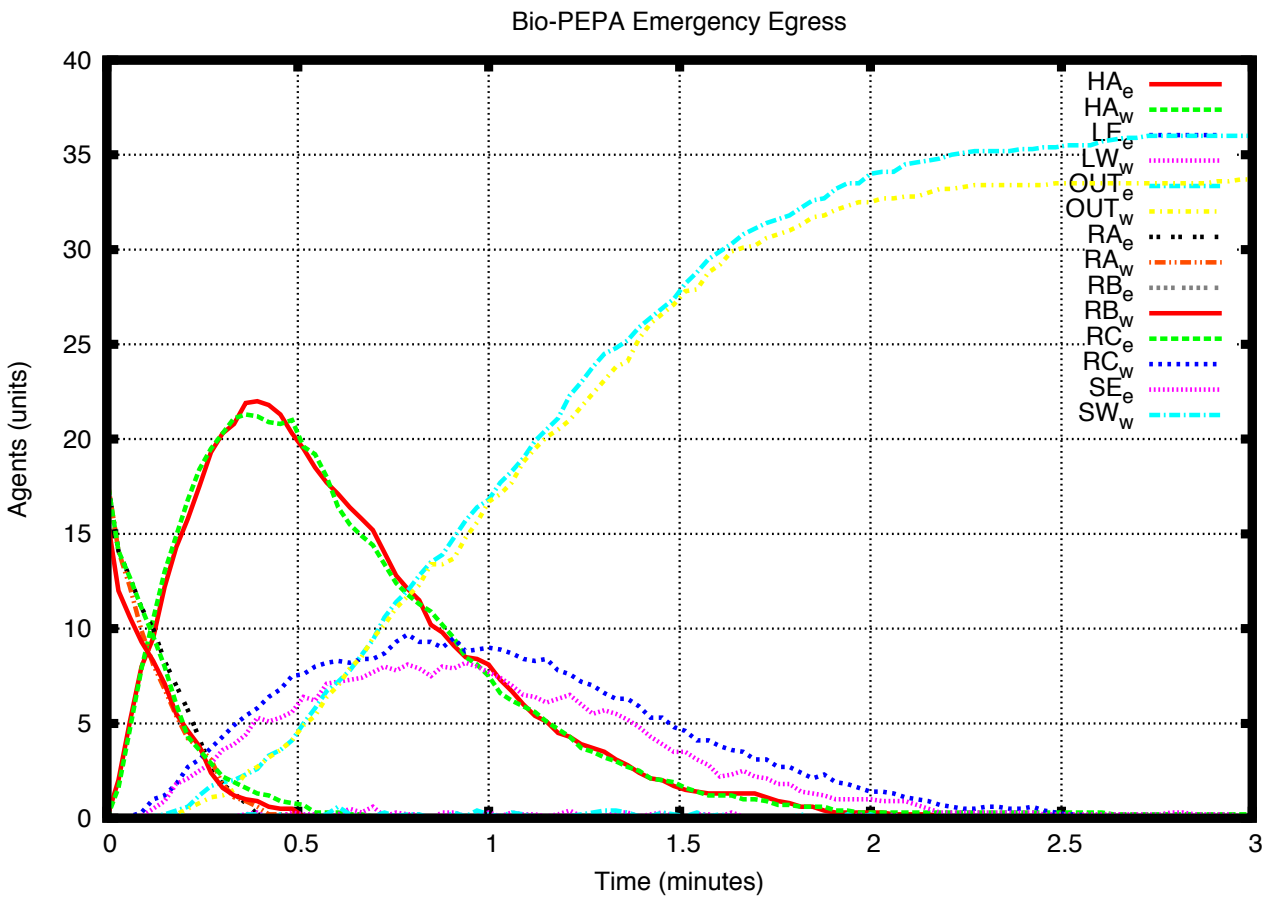
```
// BUILDING LAYOUT   (COMPARTMENTS)

location ra :       size = normal_room,  type= compartment;
location d1_ra_ha : size = normal_door,  type= compartment;
...
// PARAMETERS SET UP

to_ra_d1      =  6;                    // 3 + (60/7);
from_d1       =  door_exit_rate;
...
occupancy_d1 =  D1_ra_e@d1_ra_ha + D1_ra_w@d1_ra_ha + ... :
full_d1       =  H(capacity_d1 - occupancy_d1);
switch_d1     =  open_d1*full_d1;
...
// AGENT DYNAMICS    (FUNCTIONAL RATES)

//   From ra to ha through d1
kineticLawOf ra_e_in__d1_ra_ha  : fMA(to_ra_d1 * switch_d1 * ra_e_in_safe);
kineticLawOf ha_e_out__d1_ra_ha : fMA(from_d1 *  open_d1 * safeD1_ra_e * allowance_ha);
kineticLawOf ra_w_in__d1_ra_ha  : ...
kineticLawOf ha_w_out__d1_ra_ha : ...
//   ... and back
kineticLawOf ha_e_in__d1_ra_ha  : ...
...

// AGENT DEFINITIONS (SEQUENTIAL PROCESSES)

RA_e = (ra_e_in__d1_ra_ha, 1)     << RA_e@ra  +  ...
       (ra_e_out__d1_ra_ha, 1)    >> RA_e@ra  +  ...
RA_w =  ...
HA_e =  ...
...
D1_ra_e = (ra_e_in__d1_ra_ha, 1)  >> D1_ra_e@d1_ra_ha +
          (ha_e_out__d1_ra_ha, 1) << D1_ra_e@d1_ra_ha;
D1_ra_w = ...
D1_ha_e = ...
D1_ha_w = ...
...
// SYSTEM DEFINITION (MODEL COMPONENT)

building ::=     (RA_e@ra[18] <> RA_w@ra[18] <> HA_e@ha[0] <> ...)
```

Example results: room occupancy



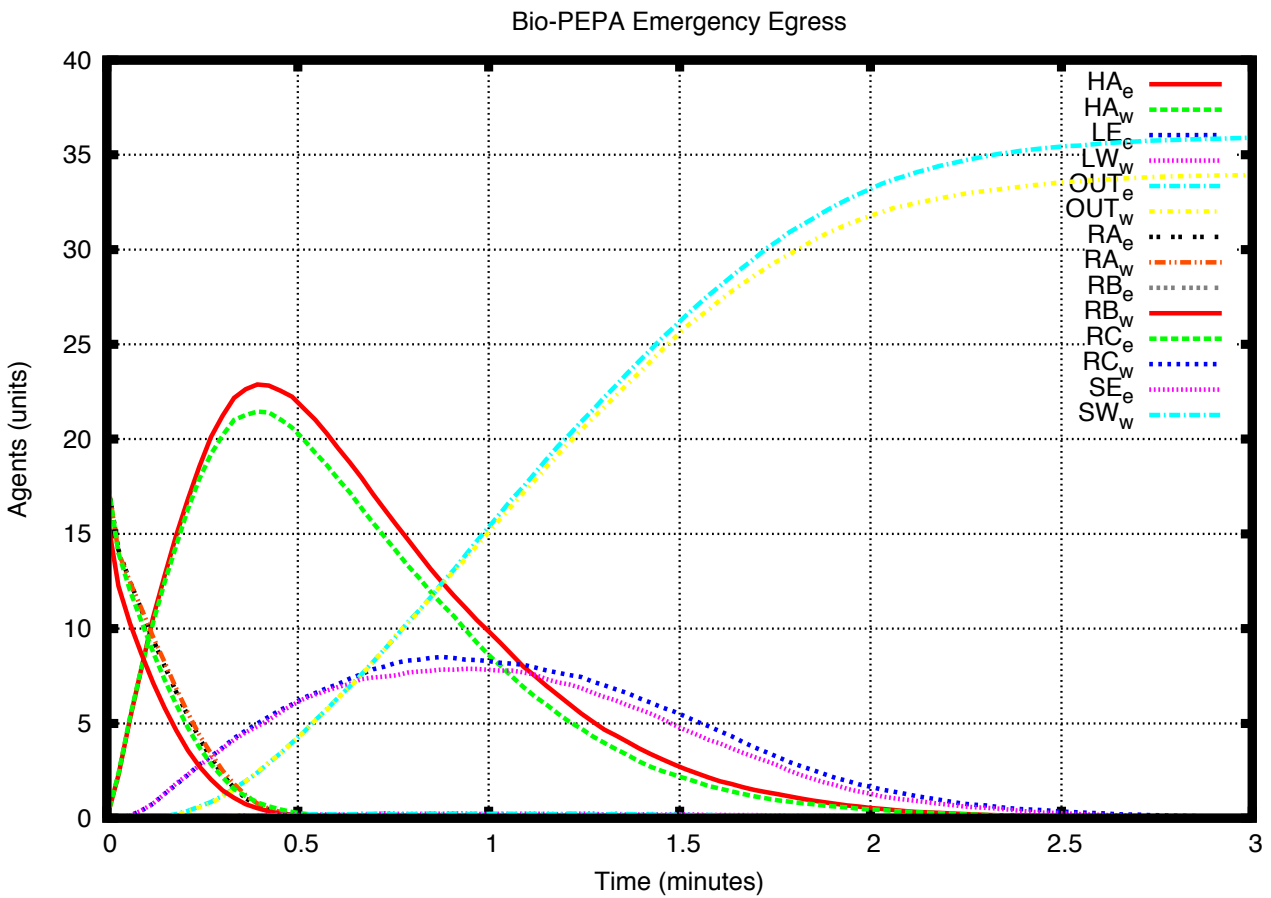
One stochastic simulation run

Example results: room occupancy



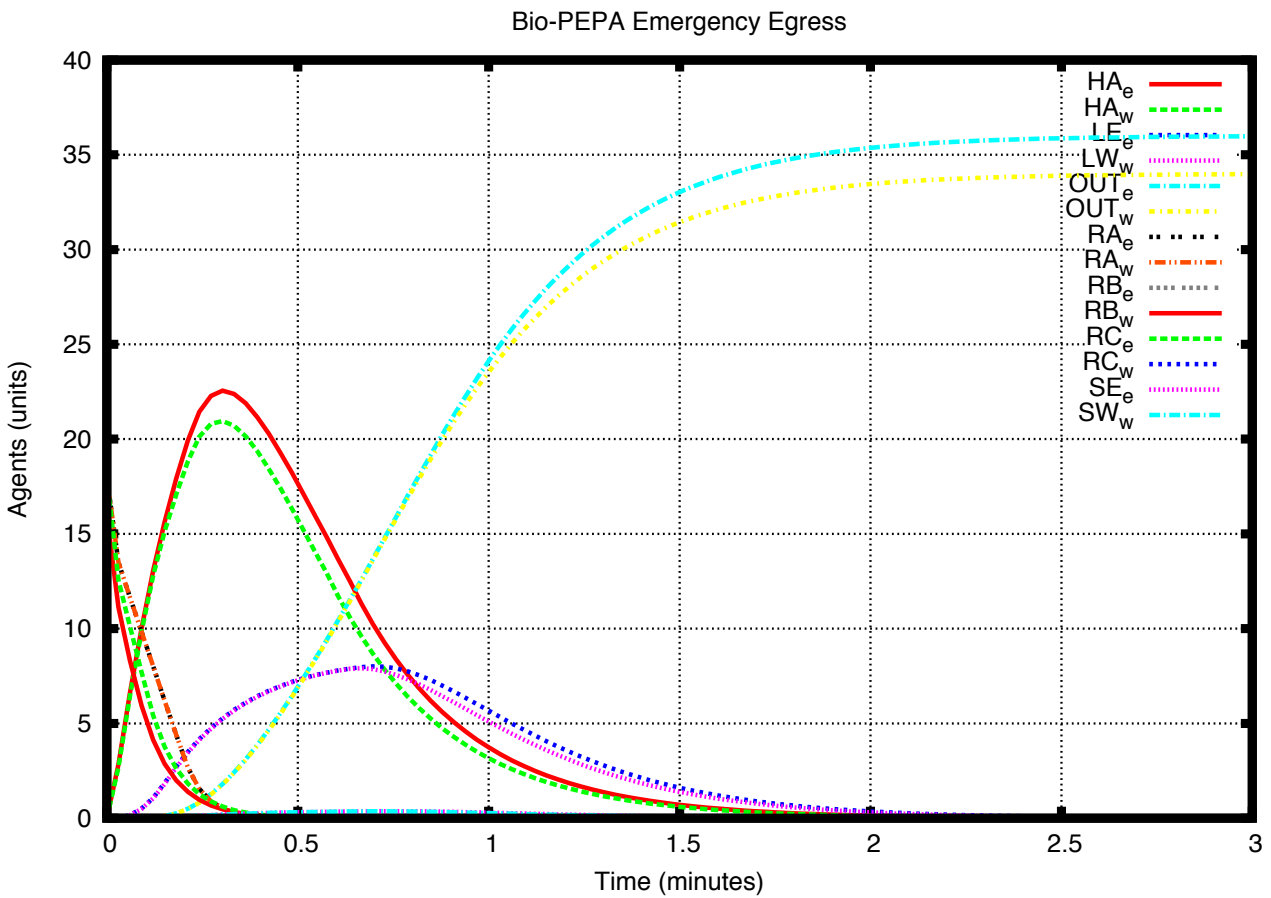
10 stochastic simulation runs

Example results: room occupancy



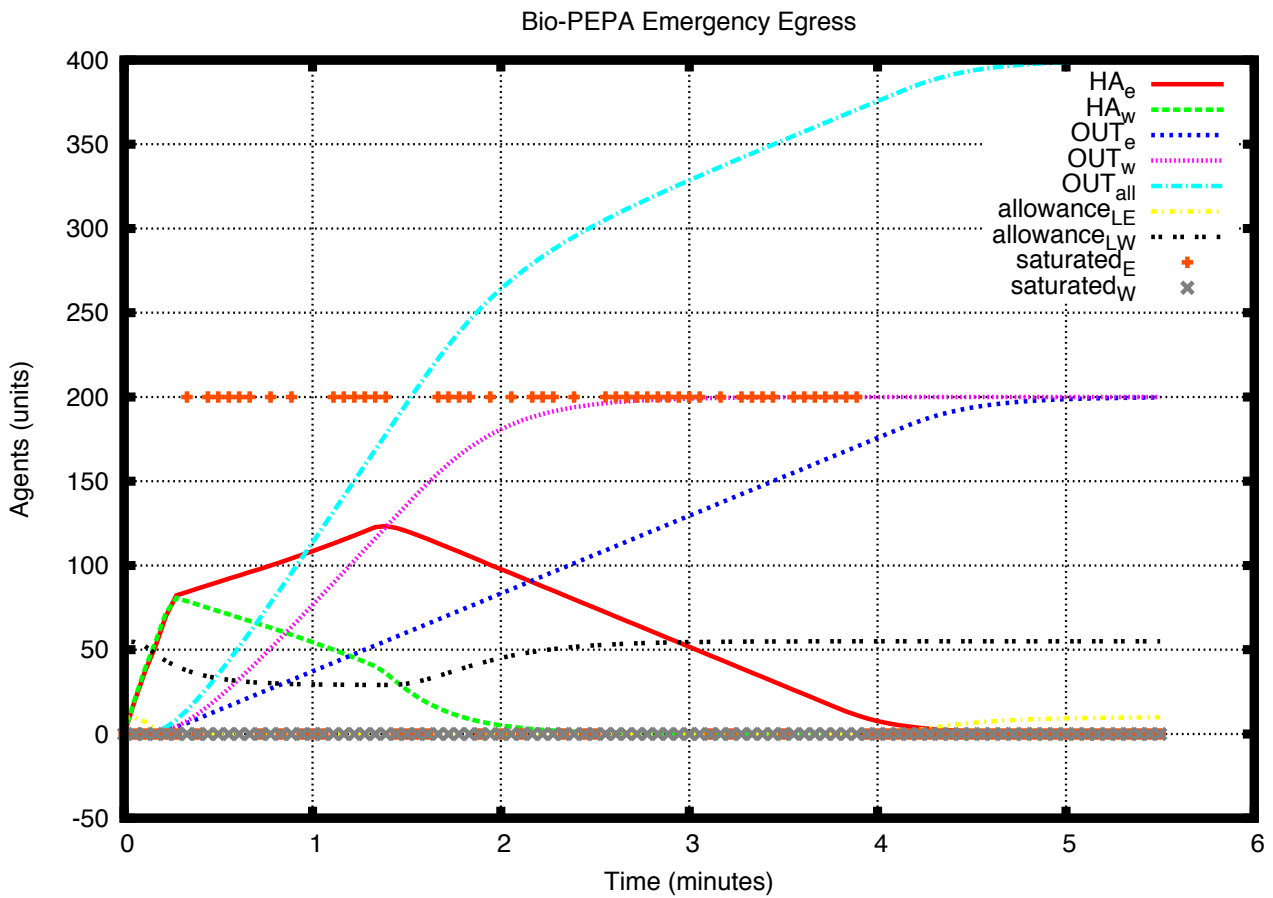
500 stochastic simulation runs

Example results: room occupancy



ODE numerical simulation

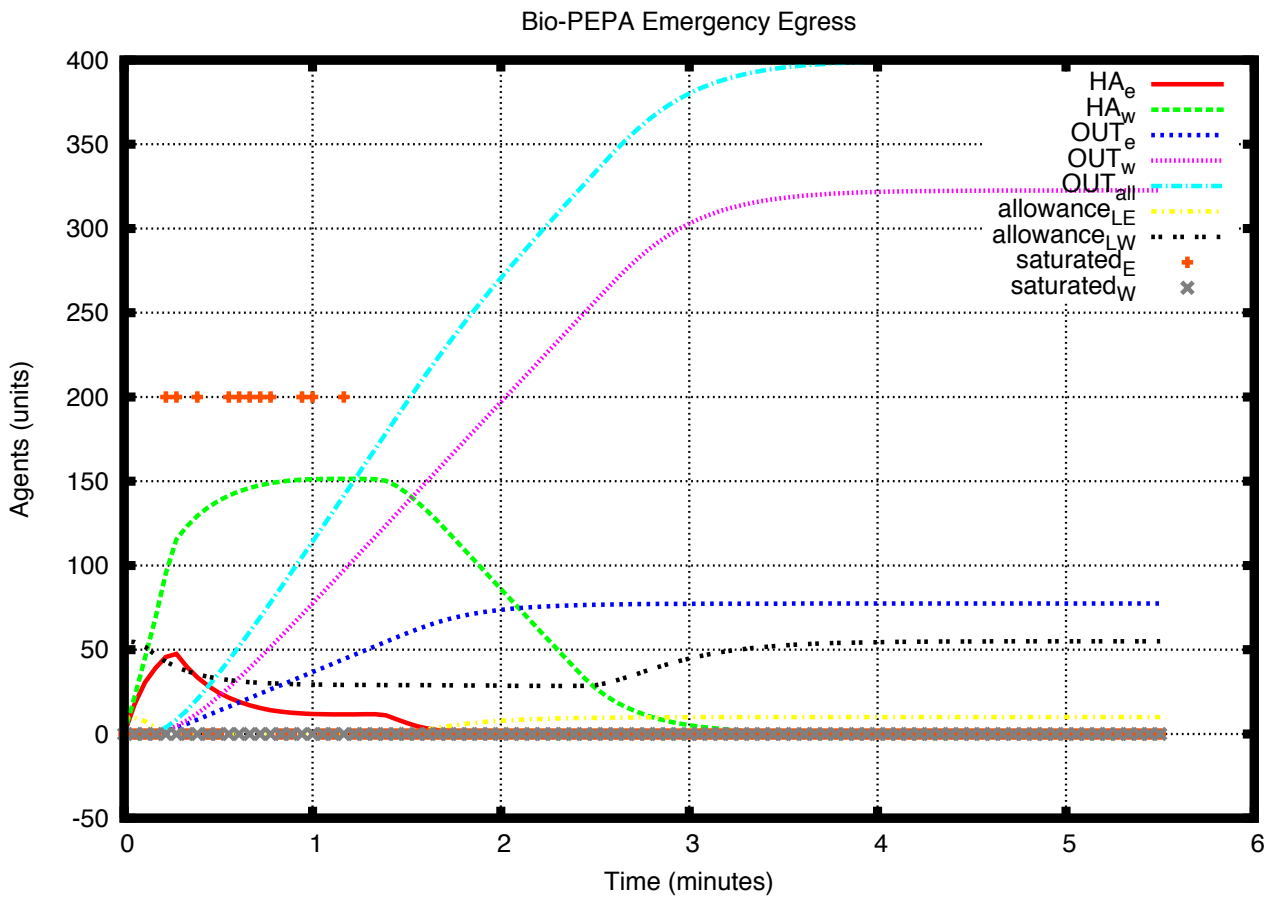
Example results: rerouting through mediation



Room occupancy over time without rerouting capability

Jane Hillston.
University of Edinburgh.

Example results: rerouting through mediation



Room occupancy over time with rerouting capability

Jane Hillston.
University of Edinburgh.

Outline

Conclusions

- Having formal representations of biological pathways eases the development of integrated analysis techniques.
- The compositionality of process algebras offers some benefits for model construction...
- ... however there is little work yet exploiting that compositional structure for the benefit of analysis.
- The semi-quantitative approach, based on CTMC with levels, offers a compomise between state space explosion and retaining some stochasticity in the model.
- Analysis based on the collective dynamics allows an efficient alternative to stochastic simulation which in many cases gives a reasonable approximation. Moreover being able to readily compare results of the two approaches can give added insight.

More Information and tool downloads

http://www.biopepa.org

Thank you!

People involved in the work:

- Ozgur Akman
- Andrea Bracciali
- Federica Ciocchetta
- Vashti Galpin
- Stephen Gilmore
- Maria Luisa Guerriero
- Diego Latella
- Mieke Massink
- Andrew Millar
- Carl Troein

Acknowledgements: Engineering and Physical Sciences Research Council (EPSRC) and Biotechnology and Biological Sciences Research Council (BBSRC).

CTMC with levels

- Analysis of the Markov process can yield quite detailed information about the dynamic behaviour of the model.

CTMC with levels

- Analysis of the Markov process can yield quite detailed information about the dynamic behaviour of the model.
- A steady state analysis provides statistics for average behaviour over a long run of the system, when the bias introduced by the initial state has been lost.

CTMC with levels

- Analysis of the Markov process can yield quite detailed information about the dynamic behaviour of the model.
- A steady state analysis provides statistics for average behaviour over a long run of the system, when the bias introduced by the initial state has been lost.
- A transient analysis provides statistics relating to the evolution of the model over a fixed period. This will be dependent on the starting state.

CTMC with levels

- Analysis of the Markov process can yield quite detailed information about the dynamic behaviour of the model.
- A steady state analysis provides statistics for average behaviour over a long run of the system, when the bias introduced by the initial state has been lost.
- A transient analysis provides statistics relating to the evolution of the model over a fixed period. This will be dependent on the starting state.

SPA
MODEL

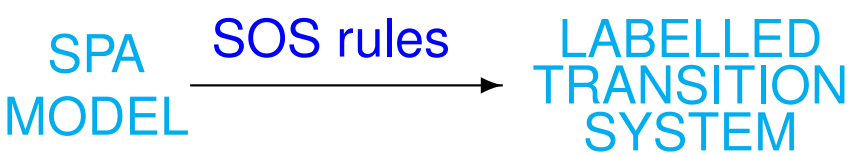
CTMC with levels

- Analysis of the Markov process can yield quite detailed information about the dynamic behaviour of the model.
- A steady state analysis provides statistics for average behaviour over a long run of the system, when the bias introduced by the initial state has been lost.
- A transient analysis provides statistics relating to the evolution of the model over a fixed period. This will be dependent on the starting state.

SPA MODEL $\xrightarrow{\text{SOS rules}}$

CTMC with levels

- Analysis of the Markov process can yield quite detailed information about the dynamic behaviour of the model.
- A steady state analysis provides statistics for average behaviour over a long run of the system, when the bias introduced by the initial state has been lost.
- A transient analysis provides statistics relating to the evolution of the model over a fixed period. This will be dependent on the starting state.



CTMC with levels

- Analysis of the Markov process can yield quite detailed information about the dynamic behaviour of the model.
- A steady state analysis provides statistics for average behaviour over a long run of the system, when the bias introduced by the initial state has been lost.
- A transient analysis provides statistics relating to the evolution of the model over a fixed period. This will be dependent on the starting state.

CTMC with levels

- Analysis of the Markov process can yield quite detailed information about the dynamic behaviour of the model.
- A steady state analysis provides statistics for average behaviour over a long run of the system, when the bias introduced by the initial state has been lost.
- A transient analysis provides statistics relating to the evolution of the model over a fixed period. This will be dependent on the starting state.

CTMC with levels

The granularity of the system is defined in terms of the step size h of the concentration intervals and the same step size is used for all the species, with few exceptions. (Law of conservation of mass).

CTMC with levels

The granularity of the system is defined in terms of the step size h of the concentration intervals and the same step size is used for all the species, with few exceptions. (Law of conservation of mass).

The granularity must be specified by the modeller based on the expected range of concentration values and the number of levels considered and transition rates are made consistent with the granularity.

CTMC with levels

The granularity of the system is defined in terms of the step size h of the concentration intervals and the same step size is used for all the species, with few exceptions. (Law of conservation of mass).

The granularity must be specified by the modeller based on the expected range of concentration values and the number of levels considered and transition rates are made consistent with the granularity.

The structure of the CTMC derived from Bio-PEPA, which we term the CTMC with levels, will depend on the granularity of the model.

CTMC with levels

The granularity of the system is defined in terms of the step size h of the concentration intervals and the same step size is used for all the species, with few exceptions. (Law of conservation of mass).

The granularity must be specified by the modeller based on the expected range of concentration values and the number of levels considered and transition rates are made consistent with the granularity.

The structure of the CTMC derived from Bio-PEPA, which we term the CTMC with levels, will depend on the granularity of the model.

As the granularity tends to zero the behaviour of this CTMC with levels tends to the behaviour of the ODEs [CDHC FBTC08].

PRISM model and model checking

- Analysing models of biological processes via probabilistic model-checking has considerable appeal.

PRISM model and model checking

- Analysing models of biological processes via probabilistic model-checking has considerable appeal.
- As with stochastic simulation the answers which are returned from model-checking give a thorough stochastic treatment to the small-scale phenomena.

PRISM model and model checking

- Analysing models of biological processes via probabilistic model-checking has considerable appeal.
- As with stochastic simulation the answers which are returned from model-checking give a thorough stochastic treatment to the small-scale phenomena.
- However, in contrast to a simulation run which generates just one trajectory, probabilistic model-checking gives a definitive answer so it is not necessary to re-run the analysis repeatedly and compute ensemble averages of the results.

PRISM model and model checking

- Analysing models of biological processes via probabilistic model-checking has considerable appeal.
- As with stochastic simulation the answers which are returned from model-checking give a thorough stochastic treatment to the small-scale phenomena.
- However, in contrast to a simulation run which generates just one trajectory, probabilistic model-checking gives a definitive answer so it is not necessary to re-run the analysis repeatedly and compute ensemble averages of the results.
- Building a reward structure over the model it is possible to express complex analysis questions.

PRISM model and model checking

- Probabilistic model checking in PRISM is based on a CTMC and the logic CSL.

PRISM model and model checking

- Probabilistic model checking in PRISM is based on a CTMC and the logic CSL.
- Formally the mapping from Bio-PEPA is based on the structured operational semantics, generating the underlying CTMC in the usual way.

PRISM model and model checking

- Probabilistic model checking in PRISM is based on a CTMC and the logic CSL.
- Formally the mapping from Bio-PEPA is based on the structured operational semantics, generating the underlying CTMC in the usual way.
- As with SSA, in practice it is more straightforward to directly map to the input language of the tool, as interacting reactive modules.

PRISM model and model checking

- Probabilistic model checking in PRISM is based on a CTMC and the logic CSL.
- Formally the mapping from Bio-PEPA is based on the structured operational semantics, generating the underlying CTMC in the usual way.
- As with SSA, in practice it is more straightforward to directly map to the input language of the tool, as interacting reactive modules.
- From a Bio-PEPA description one module is generated for each species component with an additional module to capture the functional rate information.

Integrated analyses [CHDC FBTC08] and [CGGH PASM08]

As well as offering alternative analyses we can use the different analysis tools in a complementary way. For example using stochastic simulation and probabilistic model checking in tandem.

Integrated analyses [CHDC FBTC08] and [CGGH PASM08]

As well as offering alternative analyses we can use the different analysis tools in a complementary way. For example using stochastic simulation and probabilistic model checking in tandem.

- The exact discrete-state representation of probabilistic model-checking means that its use is limited by state space explosion.

Integrated analyses [CHDC FBTC08] and [CGGH PASM08]

As well as offering alternative analyses we can use the different analysis tools in a complementary way. For example using stochastic simulation and probabilistic model checking in tandem.

- The exact discrete-state representation of probabilistic model-checking means that its use is limited by state space explosion.
- Moreover, the finite nature of the state representation used means that *a priori* bounds must be set (whether numbers of molecules or discrete levels for each species are used).

Integrated analyses [CHDC FBTC08] and [CGGH PASM08]

As well as offering alternative analyses we can use the different analysis tools in a complementary way. For example using stochastic simulation and probabilistic model checking in tandem.

- The exact discrete-state representation of probabilistic model-checking means that its use is limited by state space explosion.
- Moreover, the finite nature of the state representation used means that *a priori* bounds must be set (whether numbers of molecules or discrete levels for each species are used).
- We can use stochastic simulation to establish appropriate bounds to use for defining the PRISM state space.