Modelling Biochemical Signalling Pathways with Stochastic Process Algebra

Jane Hillston.
LFCS, University of Edinburgh

1st May 2007

Outline

- Introduction to Systems Biology
  - Motivation
- Challenges
- Stochastic Process Algebra for Biology
  - Stochastic Process Algebra
  - Abstraction, Modularity and Reasoning
  - Case Study
- Conclusions

Outline

- Introduction to Systems Biology
  - Motivation
- Challenges
- Stochastic Process Algebra for Biology
  - Stochastic Process Algebra
  - Abstraction, Modularity and Reasoning
  - Case Study
- Conclusions

Systems Biology

- Biological advances mean that much more is now known about the components of cells and the interactions between them.

Systems Biology

- Biological advances mean that much more is now known about the components of cells and the interactions between them.
- Systems biology aims to develop a better understanding of the processes involved.

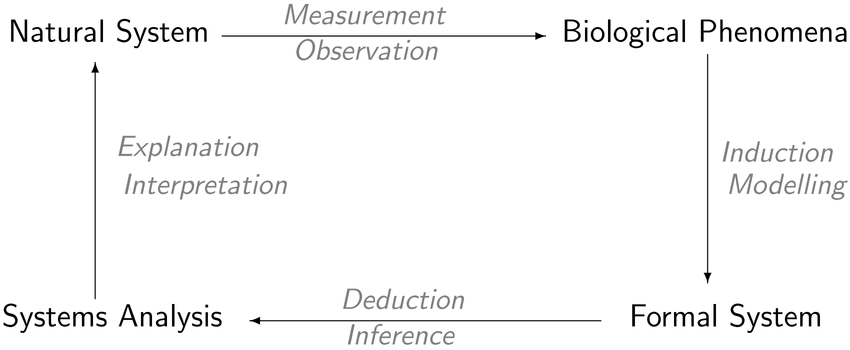
Systems Biology

- Biological advances mean that much more is now known about the components of cells and the interactions between them.
- Systems biology aims to develop a better understanding of the processes involved.
- It involves taking a systems theoretic view of biological processes — analysing inputs and outputs and the relationships between them.

Systems Biology

- Biological advances mean that much more is now known about the components of cells and the interactions between them.
- Systems biology aims to develop a better understanding of the processes involved.
- It involves taking a systems theoretic view of biological processes — analysing inputs and outputs and the relationships between them.
- A radical shift from earlier reductionist approaches, systems biology aims to provide a conceptual basis and a methodology for reasoning about biological phenomena.

Systems Biology Methodology



Natural System

Measurement

Observation

Biological Phenomena

Induction

Modelling

Formal System

Deduction

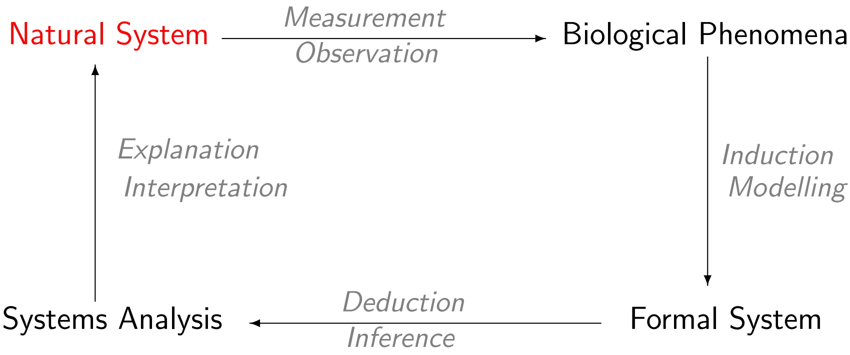
Inference

Systems Analysis

Explanation

Interpretation

Systems Biology Methodology



Natural System

Measurement

Observation

Biological Phenomena

Induction

Modelling

Formal System

Deduction

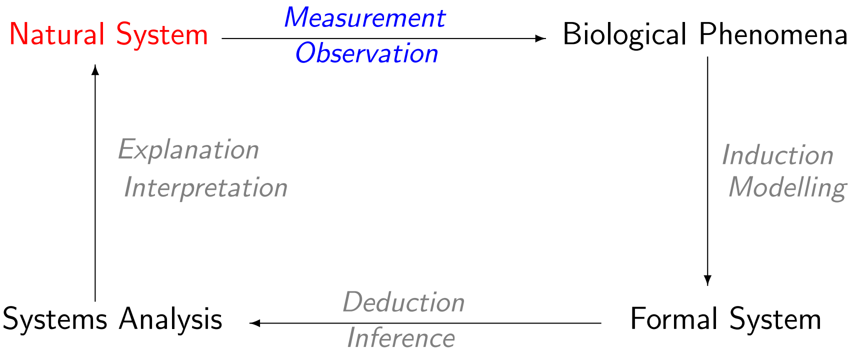
Inference

Systems Analysis

Explanation

Interpretation

Systems Biology Methodology



Natural System

Measurement

Observation

Biological Phenomena

Induction

Modelling

Formal System

Deduction

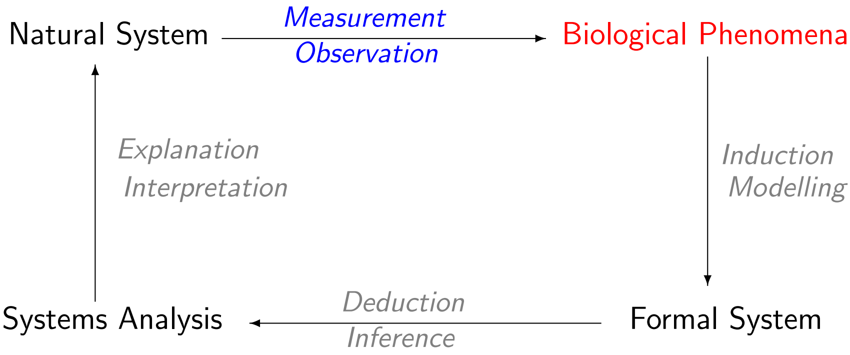
Inference

Systems Analysis

Explanation

Interpretation

Systems Biology Methodology



Natural System

Measurement

Observation

Biological Phenomena

Induction

Modelling

Formal System

Deduction

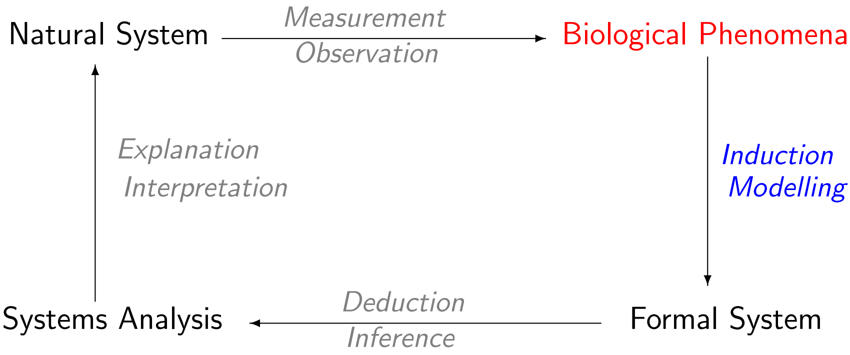
Inference

Systems Analysis

Explanation

Interpretation

Systems Biology Methodology



Natural System

Measurement
Observation

Biological Phenomena

Induction
Modelling

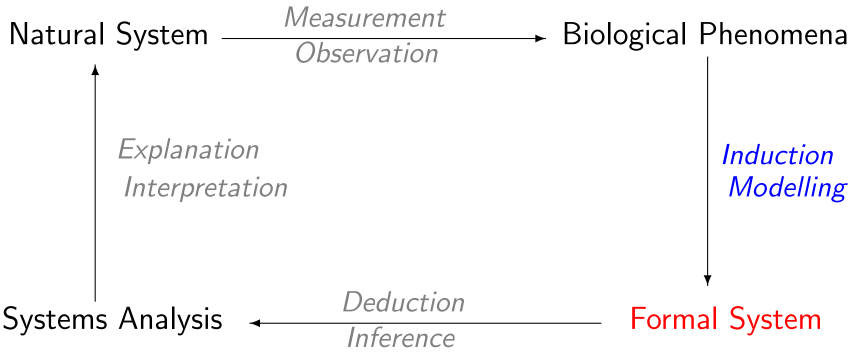
Formal System

Deduction
Inference

Systems Analysis

Explanation
Interpretation

Systems Biology Methodology



Natural System

Measurement

Observation

Biological Phenomena

Induction

Modelling

Formal System

Deduction

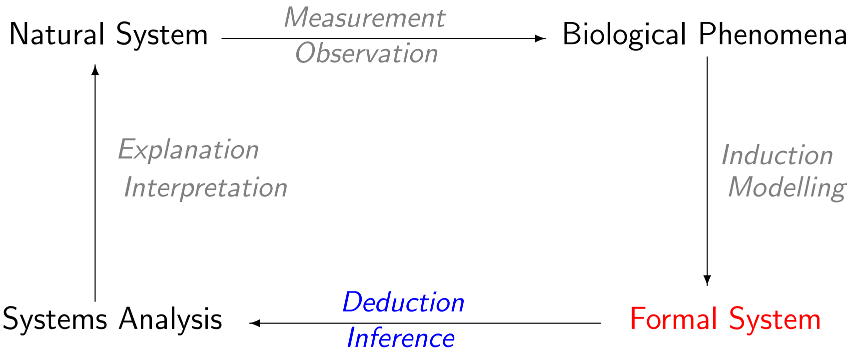
Inference

Systems Analysis

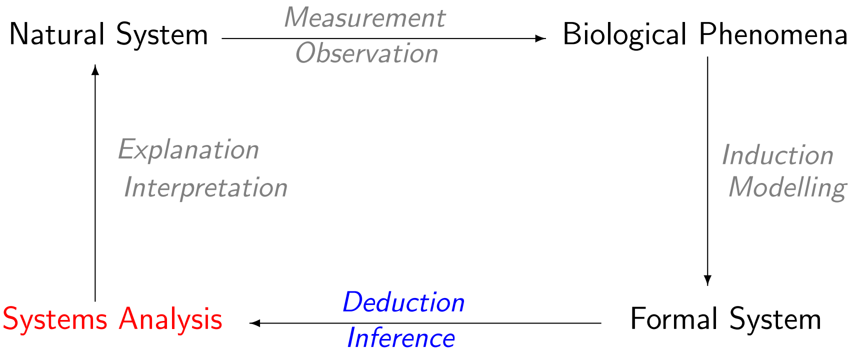
Explanation

Interpretation

Systems Biology Methodology



Systems Biology Methodology



Natural System

Measurement

Observation

Biological Phenomena

Induction

Modelling

Formal System

Deduction

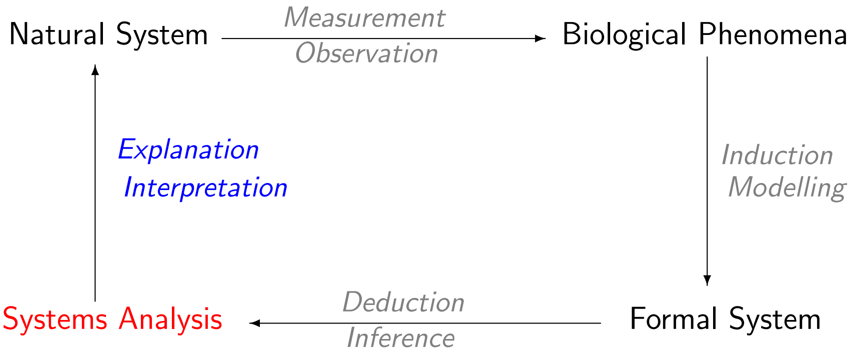
Inference

Systems Analysis

Explanation

Interpretation

Systems Biology Methodology



Natural System

Measurement

Observation

Biological Phenomena

Induction

Modelling

Formal System

Deduction

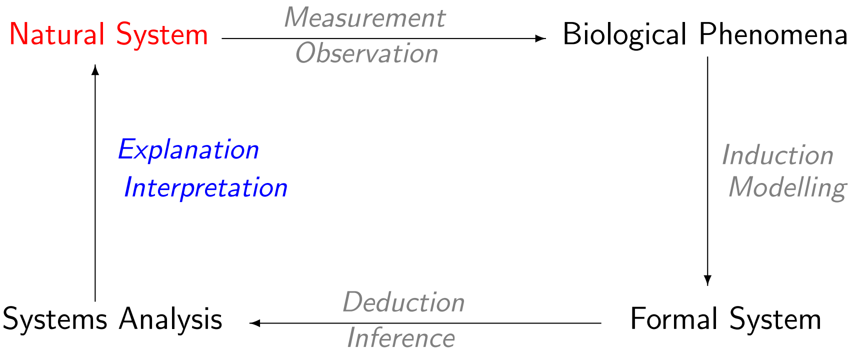
Inference

Systems Analysis

Explanation

Interpretation

Systems Biology Methodology



Natural System

Measurement
Observation

Biological Phenomena

Induction
Modelling

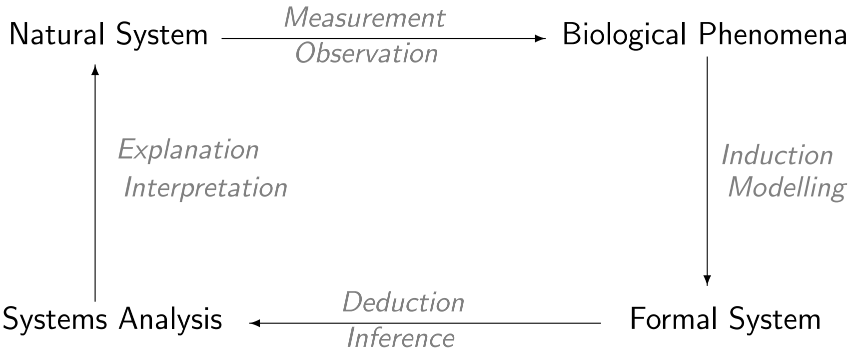
Formal System

Deduction
Inference

Systems Analysis

Explanation
Interpretation

Systems Biology Methodology



Natural System

Measurement

Observation

Biological Phenomena

Induction

Modelling

Formal System

Deduction

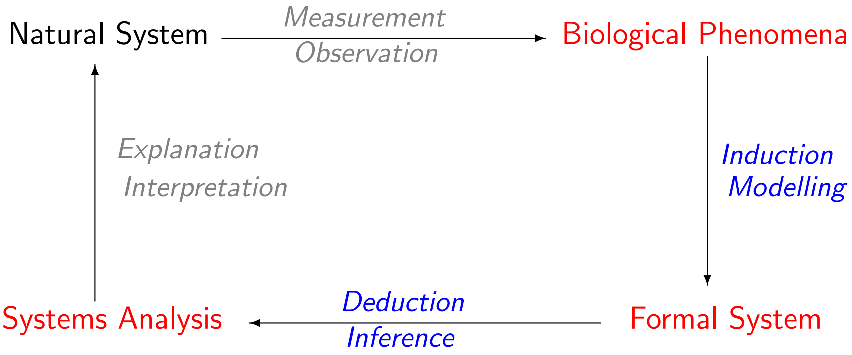
Inference

Systems Analysis

Explanation

Interpretation

Systems Biology Methodology


Natural System

Measurement

Observation

Biological Phenomena

Induction

Modelling

Formal System

Deduction

Inference

Systems Analysis

Explanation

Interpretation

Biochemical Signalling Pathways

At the intra-cellular level we can distinguish three distinct types of pathways or networks

Biochemical Signalling Pathways

At the intra-cellular level we can distinguish three distinct types of pathways or networks

Gene networks:	Genes control the production of proteins but are themselves regulated by the same or different proteins.

Biochemical Signalling Pathways

At the intra-cellular level we can distinguish three distinct types of pathways or networks

Gene networks: Genes control the production of proteins but are themselves regulated by the same or different proteins.

Signal transduction networks: External stimuli initiate messages that are carried through a cell via a cascade of biochemical reactions.

Biochemical Signalling Pathways

At the intra-cellular level we can distinguish three distinct types of pathways or networks

Gene networks:	Genes control the production of proteins but are themselves regulated by the same or different proteins.
Signal transduction networks:	External stimuli initiate messages that are carried through a cell via a cascade of biochemical reactions.
Metabolic pathways:	The survival of the cell depends on its ability to transform nutrients into energy.

Biochemical Signalling Pathways

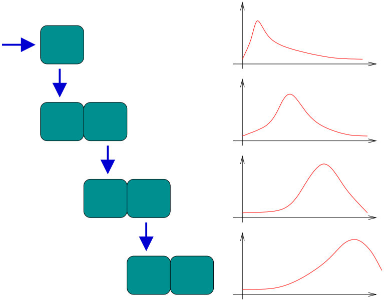
At the intra-cellular level we can distinguish three distinct types of pathways or networks

Gene networks:	Genes control the production of proteins but are themselves regulated by the same or different proteins.
Signal transduction networks:	External stimuli initiate messages that are carried through a cell via a cascade of biochemical reactions.
Metabolic pathways:	The survival of the cell depends on its ability to transform nutrients into energy.

But these distinctions are to some extent arbitrary as models may include elements of more than one pathway type.

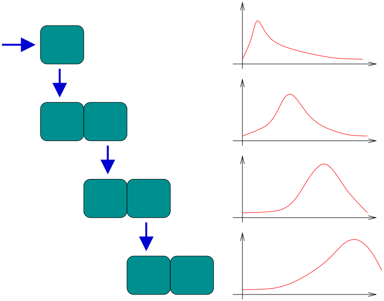
Signal transduction pathways

- All signalling is biochemical:



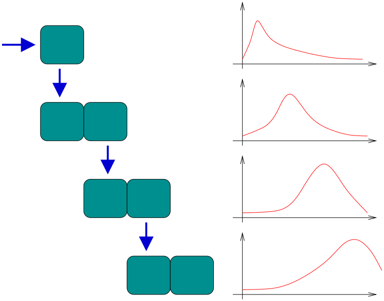
Signal transduction pathways

- All signalling is biochemical:
- Increasing protein concentration broadcasts the information about an event.



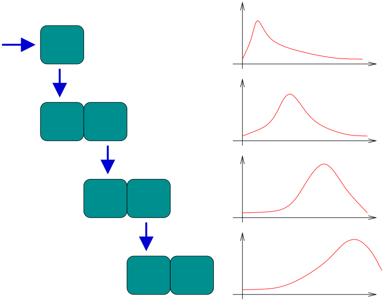
Signal transduction pathways

- All signalling is biochemical:
- Increasing protein concentration broadcasts the information about an event.
- The message is “received” by a concentration dependent response at the protein signal’s site of action.



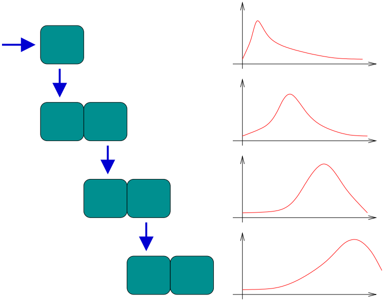
Signal transduction pathways

- All signalling is biochemical:
- Increasing protein concentration broadcasts the information about an event.
- The message is “received” by a concentration dependent response at the protein signal's site of action.
- This stimulates a response at the signalling protein's site of action.



Signal transduction pathways

- All signalling is biochemical:
- Increasing protein concentration broadcasts the information about an event.
- The message is “received” by a concentration dependent response at the protein signal’s site of action.
- This stimulates a response at the signalling protein’s site of action.
- Signals propagate through a series of protein accumulations.



Formal Systems

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

Formal Systems

There are two alternative approaches to contructing 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 contructing 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.

Systems Analysis

- In biochemical signalling pathways the events of interests are
  - when reagent concentrations start to increase;
  - when concentrations pass certain thresholds;
  - when a peak of concentration is reached.

Systems Analysis

- In biochemical signalling pathways the events of interests are
  - when reagent concentrations start to increase;
  - when concentrations pass certain thresholds;
  - when a peak of concentration is reached.
- For example, delay from the activation of a gene promoter until reaching an effective level to control the next promoter in a pathway depends on the rate of protein accumulation.

Systems Analysis

- In biochemical signalling pathways the events of interests are
  - when reagent concentrations start to increase;
  - when concentrations pass certain thresholds;
  - when a peak of concentration is reached.
- For example, delay from the activation of a gene promoter until reaching an effective level to control the next promoter in a pathway depends on the rate of protein accumulation.
- These are the data that can be collected from wet lab experiments.

Systems Analysis

- In biochemical signalling pathways the events of interests are
  - when reagent concentrations start to increase;
  - when concentrations pass certain thresholds;
  - when a peak of concentration is reached.
- For example, delay from the activation of a gene promoter until reaching an effective level to control the next promoter in a pathway depends on the rate of protein accumulation.
- These are the data that can be collected from wet lab experiments.
- The accumulation of protein is a stochastic process affected by several factors in the cell (temperature, pH, etc.).

Systems Analysis

- In biochemical signalling pathways the events of interests are
  - when reagent concentrations start to increase;
  - when concentrations pass certain thresholds;
  - when a peak of concentration is reached.
- For example, delay from the activation of a gene promoter until reaching an effective level to control the next promoter in a pathway depends on the rate of protein accumulation.
- These are the data that can be collected from wet lab experiments.
- The accumulation of protein is a stochastic process affected by several factors in the cell (temperature, pH, etc.).
- Thus it is more realistic to talk about a distribution rather than a deterministic time.

Outline

Individual vs. Population behaviour

- Biochemistry is concerned with the reactions between individual molecules and so it is often more natural to model at this level.

Individual vs. Population behaviour

- Biochemistry is concerned with the reactions between individual molecules and so it is often more natural to model at this level.
- Experimental data is usually more readily available in terms of populations cf. average reaction rates vs. the forces at play on an individual molecule in a particular physical context.

Individual vs. Population behaviour

- Biochemistry is concerned with the reactions between individual molecules and so it is often more natural to model at this level.
- Experimental data is usually more readily available in terms of populations cf. average reaction rates vs. the forces at play on an individual molecule in a particular physical context.
- These two views should be regarded as alternatives, each being appropriate for some models. The challenge then becomes when to use which approach.

Individual vs. Population behaviour

- Biochemistry is concerned with the reactions between individual molecules and so it is often more natural to model at this level.
- Experimental data is usually more readily available in terms of populations cf. average reaction rates vs. the forces at play on an individual molecule in a particular physical context.
- These two views should be regarded as alternatives, each being appropriate for some models. The challenge then becomes when to use which approach.
- Note that given a large enough number of molecules, an "individuals" model will (in many circumstances) be indistinguishable from the a "population" level model.

Noise vs. Determinism

- With perfect knowledge the behaviour of a biochemical reaction would be deterministic.

Noise vs. Determinism

- With perfect knowledge the behaviour of a biochemical reaction would be deterministic.
- However, in general, we do not have the requisite knowledge of thermodynamic forces, exact relative positions, temperature, velocity etc.

Noise vs. Determinism

- With perfect knowledge the behaviour of a biochemical reaction would be deterministic.
- However, in general, we do not have the requisite knowledge of thermodynamic forces, exact relative positions, temperature, velocity etc.
- Thus a reaction appears to display stochastic behaviour.

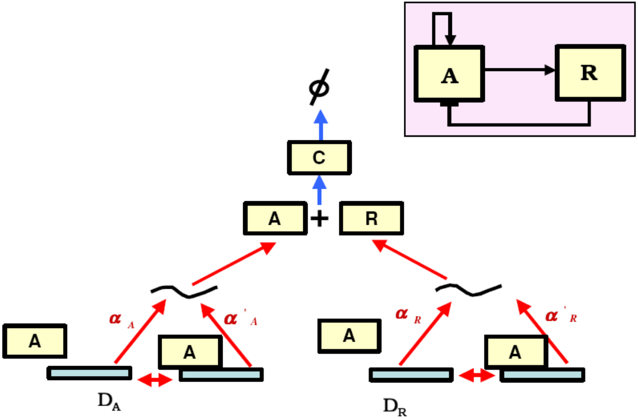
Noise vs. Determinism

- With perfect knowledge the behaviour of a biochemical reaction would be deterministic.
- However, in general, we do not have the requisite knowledge of thermodynamic forces, exact relative positions, temperature, velocity etc.
- Thus a reaction appears to display stochastic behaviour.
- When a large number of such reactions occur, the randomness of the individual reactions can cancel each other out and the apparent behaviour exhibits less variability.

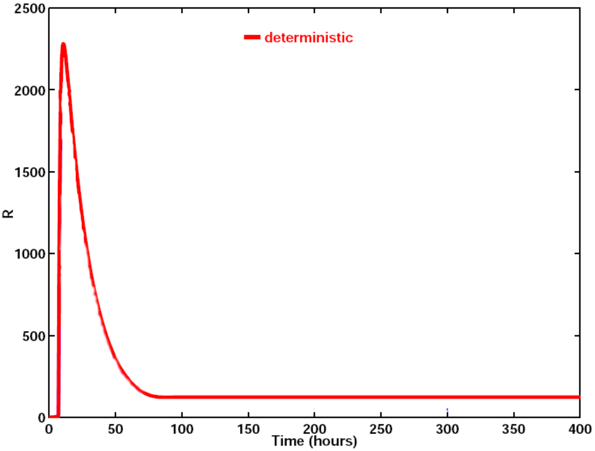
Noise vs. Determinism

- With perfect knowledge the behaviour of a biochemical reaction would be deterministic.
- However, in general, we do not have the requisite knowledge of thermodynamic forces, exact relative positions, temperature, velocity etc.
- Thus a reaction appears to display stochastic behaviour.
- When a large number of such reactions occur, the randomness of the individual reactions can cancel each other out and the apparent behaviour exhibits less variability.
- However, in some systems the variability in the stochastic behaviour plays a crucial role in the dynamics of the system.

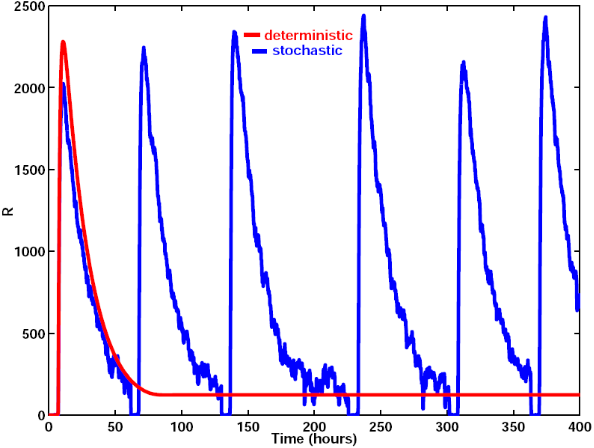
Circadian clock (cartoon)



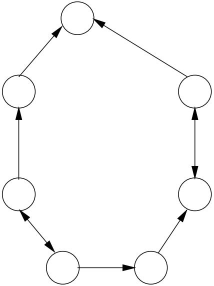
Circadian clock (deterministically . . .)



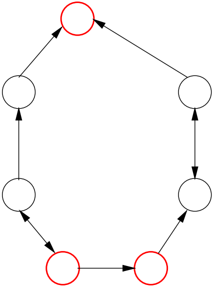
Circadian clock (. . . and stochastically)



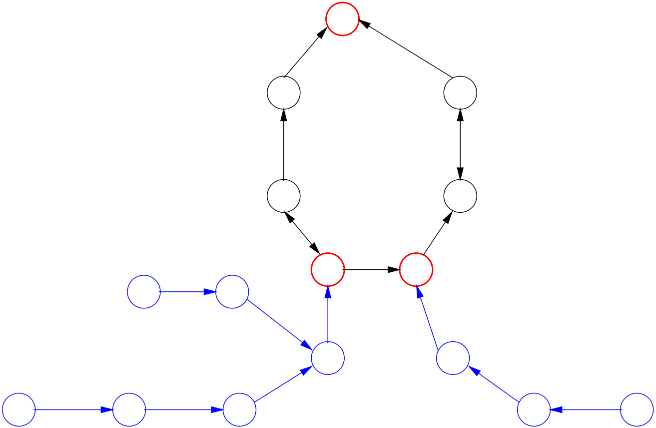
The problem of *Infinite Regress*



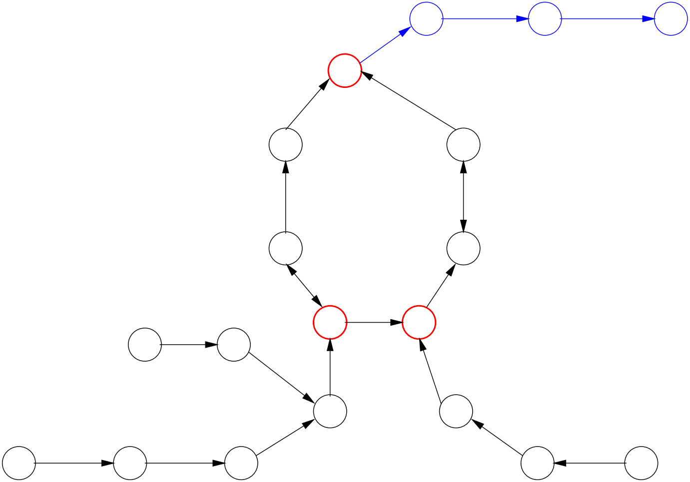
The problem of *Infinite Regress*



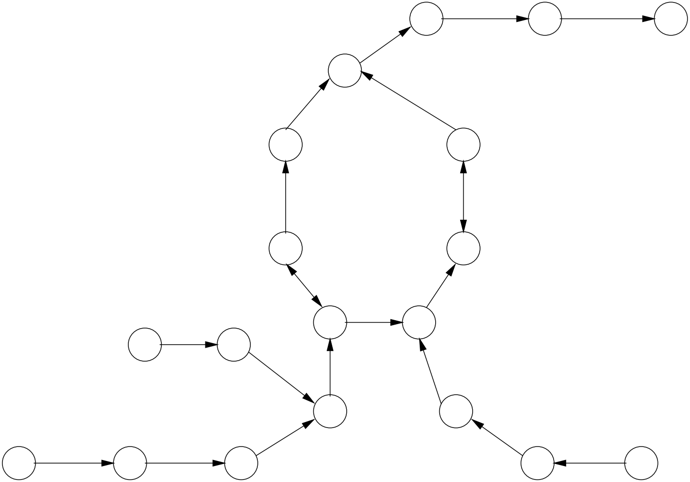
The problem of *Infinite Regress*



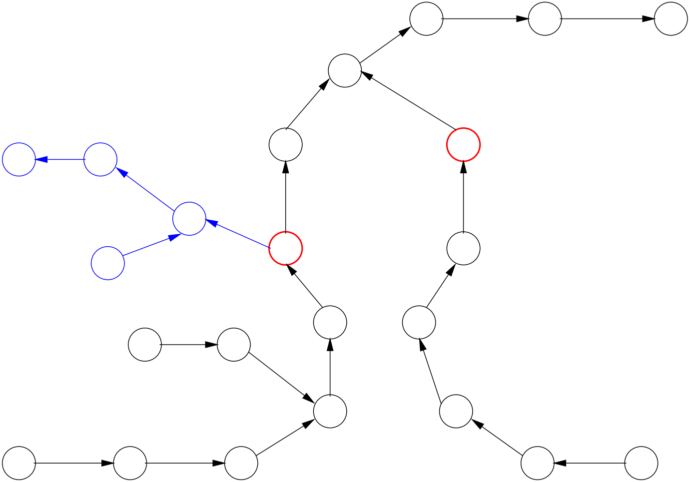
The problem of *Infinite Regress*



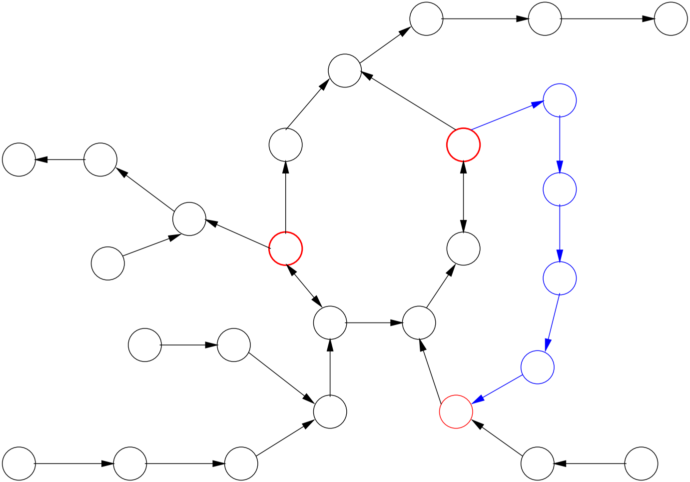
The problem of *Infinite Regress*



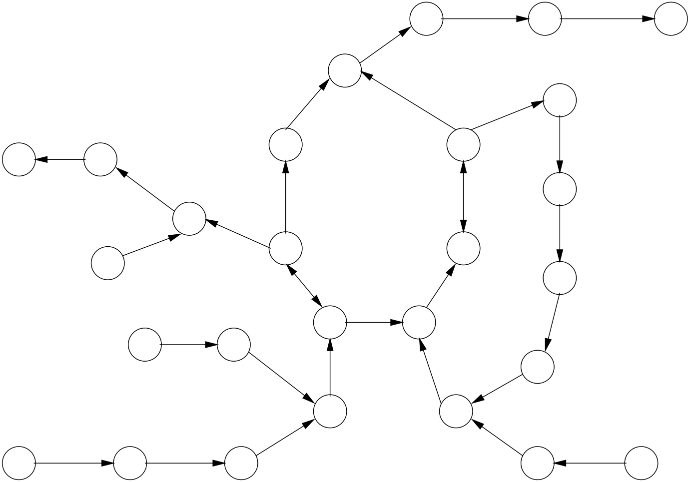
The problem of *Infinite Regress*



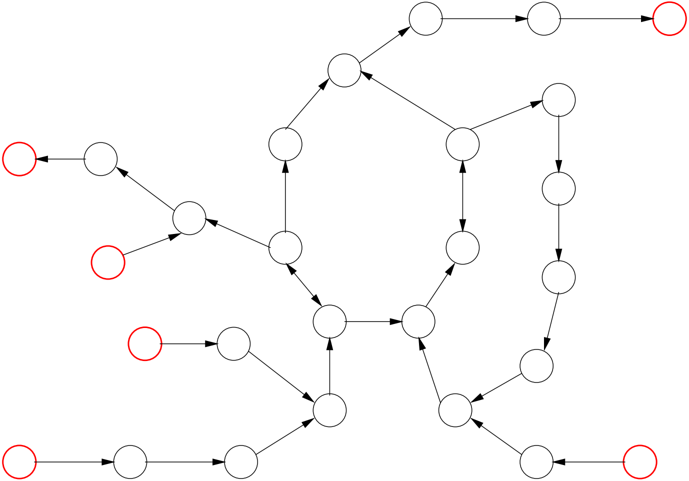
The problem of *Infinite Regress*



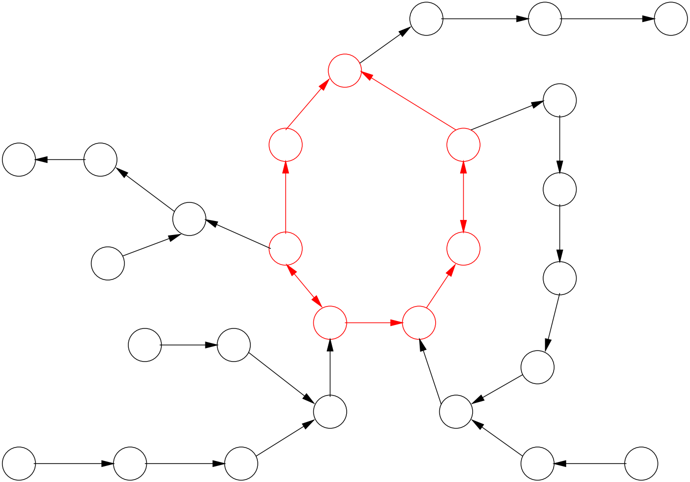
The problem of *Infinite Regress*



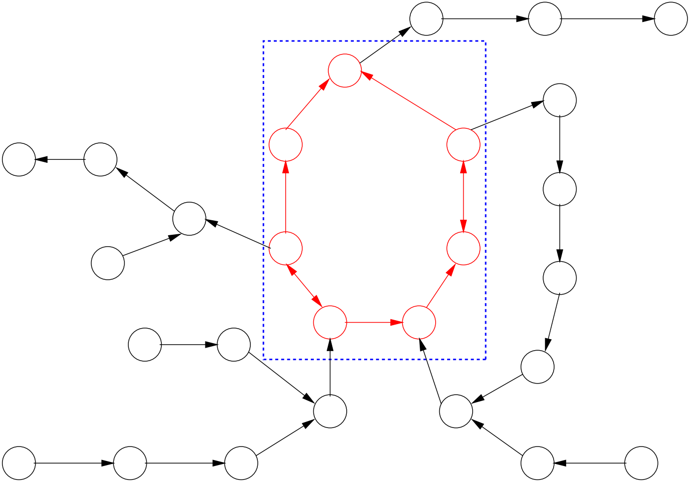
The problem of *Infinite Regress*



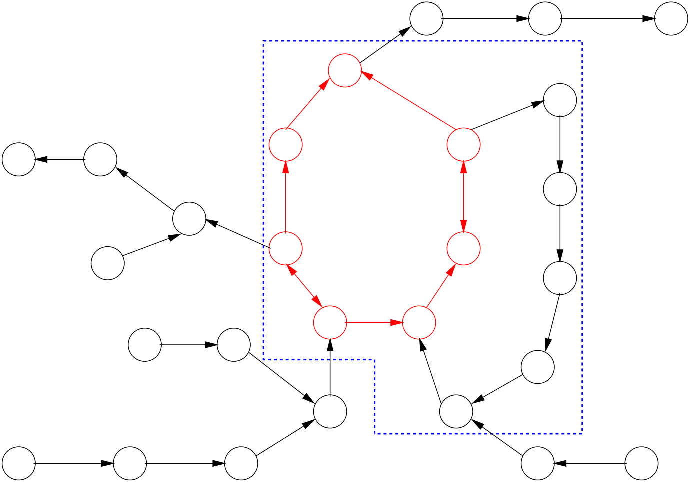
The problem of *Infinite Regress*



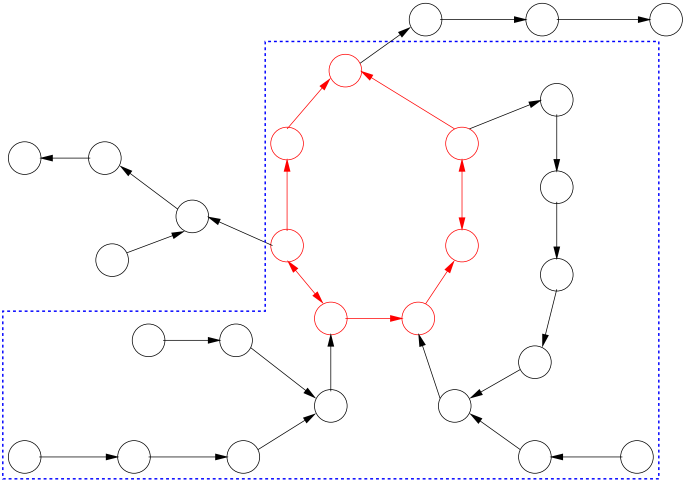
The problem of *Infinite Regress*



The problem of *Infinite Regress*



The problem of *Infinite Regress*



Modularity vs. Infinite Regress

As computer scientists we are firm believers in modularity and compositionality. When it comes to biochemical pathways opinion amongst biologists is divided about whether is makes sense to take a modular view of cellular pathways.

Modularity vs. Infinite Regress

As computer scientists we are firm believers in modularity and compositionality. When it comes to biochemical pathways opinion amongst biologists is divided about whether is makes sense to take a modular view of cellular pathways.

Some biologists (e.g. Leibler) argue that there is modularity, naturally occuring, where they define a module relative to a biological function.

Modularity vs. Infinite Regress

As computer scientists we are firm believers in modularity and compositionality. When it comes to biochemical pathways opinion amongst biologists is divided about whether is makes sense to take a modular view of cellular pathways.

Some biologists (e.g. Leibler) argue that there is modularity, naturally occuring, where they define a module relative to a biological function.

Others such as Cornish-Bowden are much more skeptical and cite the problem of infinite regress as being insurmountable.

Problems with Data

There is a fundamental challenge when modelling cellular pathways that little is known about some aspects of cellular processes.

Problems with Data

There is a fundamental challenge when modelling cellular pathways that little is known about some aspects of cellular processes.

In some cases this is because no experimental data is available, or that the experimental data that is available is inconsistent.

Problems with Data

There is a fundamental challenge when modelling cellular pathways that little is known about some aspects of cellular processes.

In some cases this is because no experimental data is available, or that the experimental data that is available is inconsistent.

In other cases the data is unknowable because experimental techniques do not yet exist to collect the data, or those that do involve modification to the system.

Problems with Data

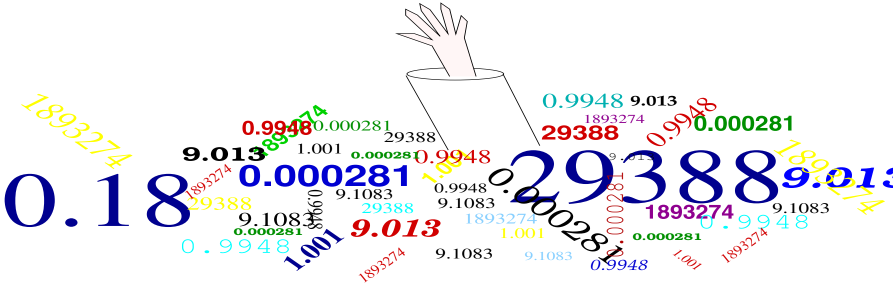
There is a fundamental challenge when modelling cellular pathways that little is known about some aspects of cellular processes.

In some cases this is because no experimental data is available, or that the experimental data that is available is inconsistent.

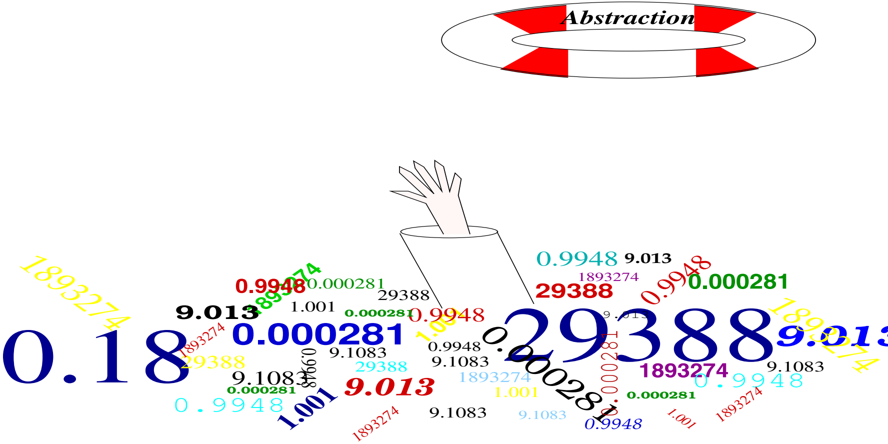
In other cases the data is unknowable because experimental techniques do not yet exist to collect the data, or those that do involve modification to the system.

Even when data exists the quality is often poor.

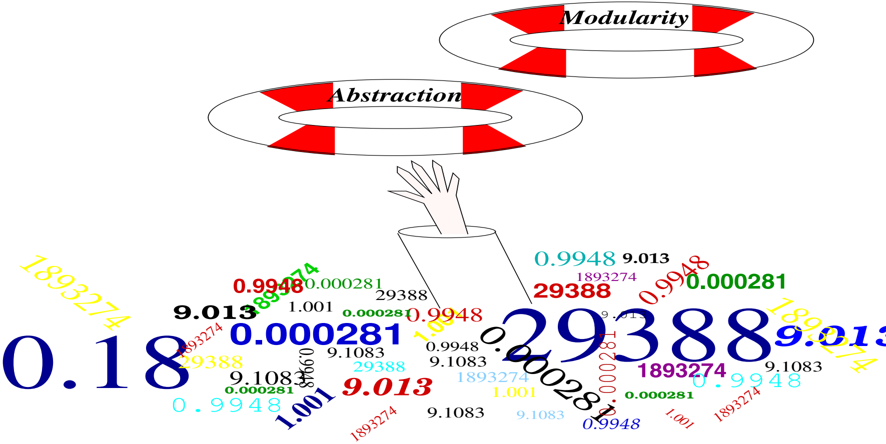
Computational Thinking to the rescue



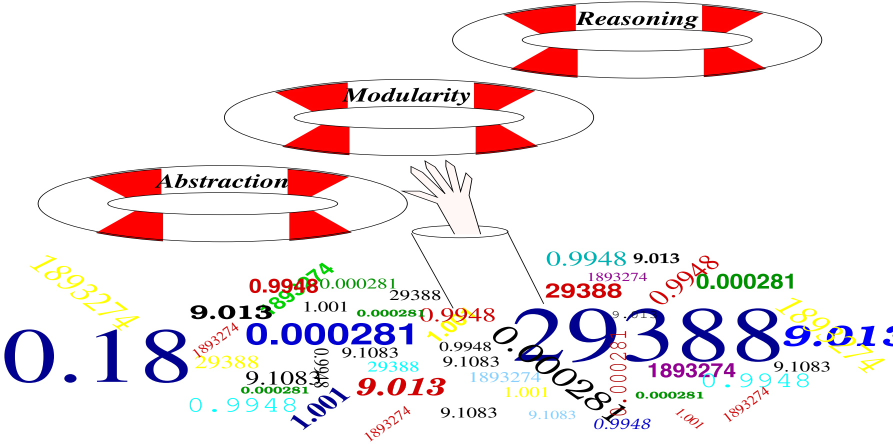
Computational Thinking to the rescue



Computational Thinking to the rescue



Computational Thinking to the rescue



Outline

Introduction to Systems Biology
  Motivation

Challenges

Stochastic Process Algebra for Biology
  Stochastic Process Algebra
  Abstraction, Modularity and Reasoning
  Case Study

Conclusions

Using Stochastic Process Algebras

Process algebras have several attractive features which could be useful for modelling and understanding biological systems:

Using Stochastic Process Algebras

Process algebras have several attractive features which could be useful for modelling and understanding biological systems:

- Process algebraic formulations are compositional and make interactions/constraints explicit.

Using Stochastic Process Algebras

Process algebras have several attractive features which could be useful for modelling and understanding biological systems:

- Process algebraic formulations are compositional and make interactions/constraints explicit.
- Equivalence relations allow formal comparison of high-level descriptions.

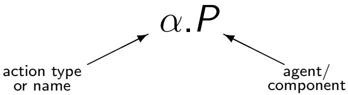
Using Stochastic Process Algebras

Process algebras have several attractive features which could be useful for modelling and understanding biological systems:

- Process algebraic formulations are compositional and make interactions/constraints explicit.
- Equivalence relations allow formal comparison of high-level descriptions.
- There are well-established techniques for reasoning about the behaviours and properties of models, supported by software. These include qualitative and quantitative analysis, and model checking.

Process Algebra

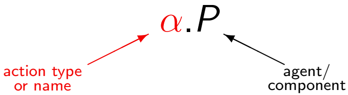
- Models consist of agents which engage in actions.



$$\alpha.P$$

Process Algebra

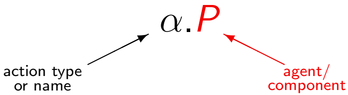
- Models consist of agents which engage in actions.



$$\alpha.P$$

Process Algebra

- Models consist of agents which engage in actions.



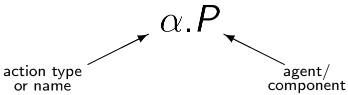
$$\alpha.P$$

action type
or name

agent/
component

Process Algebra

- Models consist of agents which engage in actions.



$$\alpha.P$$

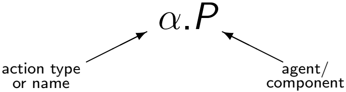
action type or name

agent/ component

- The structured operational (interleaving) semantics of the language is used to generate a labelled transition system.

Process Algebra

- Models consist of agents which engage in actions.



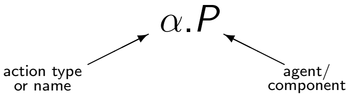
$$\alpha.P$$

- The structured operational (interleaving) semantics of the language is used to generate a labelled transition system.

Process algebra model

Process Algebra

- Models consist of agents which engage in actions.



$\alpha.P$

action type or name

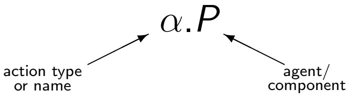
agent/ component

- The structured operational (interleaving) semantics of the language is used to generate a labelled transition system.

Process algebra model $\xrightarrow{\text{SOS rules}}$

Process Algebra

- Models consist of agents which engage in actions.



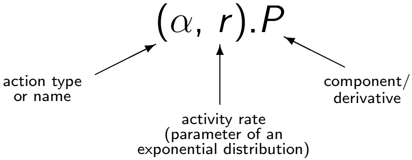
$$\alpha.P$$

- The structured operational (interleaving) semantics of the language is used to generate a labelled transition system.

$$\text{Process algebra model} \xrightarrow{\text{SOS rules}} \text{Labelled transition system}$$

Stochastic Process Algebra

- Models are constructed from components which engage in activities.



$$(\alpha, r).P$$

action type
or name

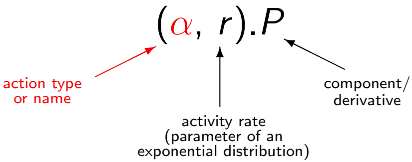
activity rate
(parameter of an
exponential distribution)

component/
derivative

The language may be used to generate a Markov Process (CTMC).

Stochastic Process Algebra

- Models are constructed from components which engage in activities.

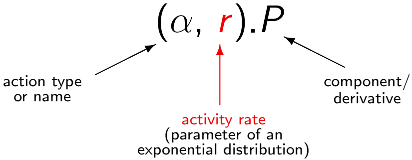


$$(\alpha, r).P$$

The language may be used to generate a Markov Process (CTMC).

Stochastic Process Algebra

- Models are constructed from components which engage in activities.



$$(\alpha, r).P$$

action type or name

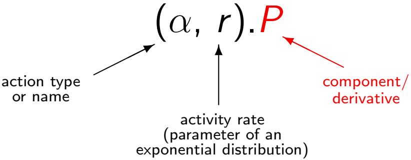
activity rate (parameter of an exponential distribution)

component/ derivative

The language may be used to generate a Markov Process (CTMC).

Stochastic Process Algebra

- Models are constructed from components which engage in activities.

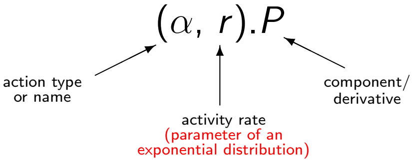


$$(\alpha, r).P$$

The language may be used to generate a Markov Process (CTMC).

Stochastic Process Algebra

- Models are constructed from components which engage in activities.



$$(\alpha, r).P$$

action type
or name

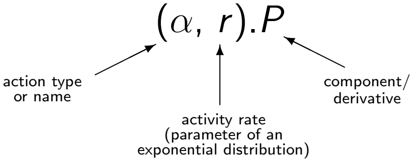
activity rate
(parameter of an
exponential distribution)

component/
derivative

The language may be used to generate a Markov Process (CTMC).

Stochastic Process Algebra

- Models are constructed from components which engage in activities.



$$(\alpha, r).P$$

action type
or name

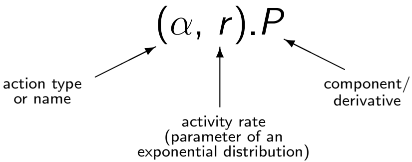
activity rate
(parameter of an
exponential distribution)

component/
derivative

The language may be used to generate a Markov Process (CTMC).

Stochastic Process Algebra

- Models are constructed from components which engage in activities.



$$(\alpha, r).P$$

action type
or name

activity rate
(parameter of an
exponential distribution)

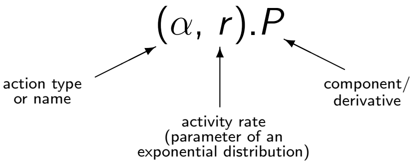
component/
derivative

The language may be used to generate a Markov Process (CTMC).

SPA
MODEL

Stochastic Process Algebra

- Models are constructed from components which engage in activities.

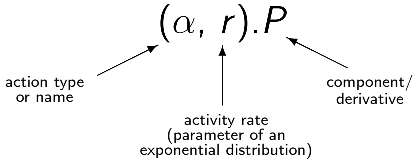


$$(\alpha, r).P$$

The language may be used to generate a Markov Process (CTMC).

Stochastic Process Algebra

- Models are constructed from components which engage in activities.



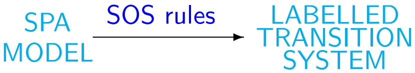
$$(\alpha, r).P$$

action type or name

activity rate (parameter of an exponential distribution)

component/ derivative

The language may be used to generate a Markov Process (CTMC).



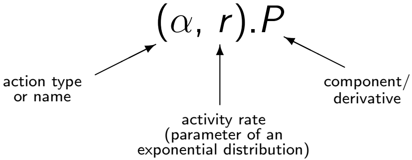
SPA MODEL

SOS rules

LABELLED TRANSITION SYSTEM

Stochastic Process Algebra

- Models are constructed from components which engage in activities.

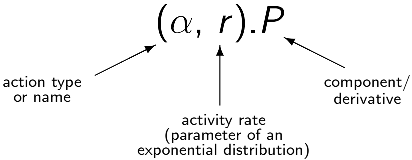


$$(\alpha, r).P$$

The language may be used to generate a Markov Process (CTMC).

Stochastic Process Algebra

- Models are constructed from components which engage in activities.



$$(\alpha, r).P$$

The language may be used to generate a Markov Process (CTMC).

Molecular processes as concurrent computations

Concurrency	**Molecular Biology**	**Metabolism**	**Signal Transduction**
Concurrent computational processes	Molecules	Enzymes and metabolites	Interacting proteins
Synchronous communication	Molecular interaction	Binding and catalysis	Binding and catalysis
Transition or mobility	Biochemical modification or relocation	Metabolite synthesis	Protein binding, modification or sequestration

[Regev *et al* 2000]

Molecular processes as concurrent computations

Concurrency	**Molecular Biology**	**Metabolism**	**Signal Transduction**
Concurrent computational processes	Molecules	Enzymes and metabolites	Interacting proteins
Synchronous communication	Molecular interaction	Binding and catalysis	Binding and catalysis
Transition or mobility	Biochemical modification or relocation	Metabolite synthesis	Protein binding, modification or sequestration

[Regev *et al* 2000]

Mapping biological systems to process algebra

The work using the stochastic π-calculus and related calculi, for modelling biochemical signalling within cells maps a molecule in a pathway to a process in the process algebra description.

Mapping biological systems to process algebra

The work using the stochastic π-calculus and related calculi, for modelling biochemical signalling within cells maps a molecule in a pathway to a process in the process algebra description.

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

Mapping biological systems to process algebra

The work using the stochastic π-calculus and related calculi, for modelling biochemical signalling within cells maps a molecule in a pathway to a process in the process algebra description.

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

In the PEPA modelling we have been doing we have experimented with more abstract mappings between process algebra constructs and elements of signalling pathways.

Mapping biological systems to process algebra

The work using the stochastic π-calculus and related calculi, for modelling biochemical signalling within cells maps a molecule in a pathway to a process in the process algebra description.

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

In the PEPA modelling we have been doing we have experimented with more abstract mappings between process algebra constructs and elements of signalling pathways.

In our mapping we focus on species (c.f. a type rather than an instance, or a class rather than an object).

Motivations for Abstraction

Our motivations for seeking more abstraction in process algebra models for systems biology comes from both key aspects of modelling:

Motivations for Abstraction

Our motivations for seeking more abstraction in process algebra models for systems biology comes from both key aspects of modelling:

- The data that we have available to parameterise models is sometimes speculative rather than precise.

Motivations for Abstraction

Our motivations for seeking more abstraction in process algebra models for systems biology comes from both key aspects of modelling:

- The data that we have available to parameterise models is sometimes speculative rather than precise. This suggests that we should use semiquantitative models rather than quantitative ones.

Motivations for Abstraction

Our motivations for seeking more abstraction in process algebra models for systems biology comes from both key aspects of modelling:

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

Motivations for Abstraction

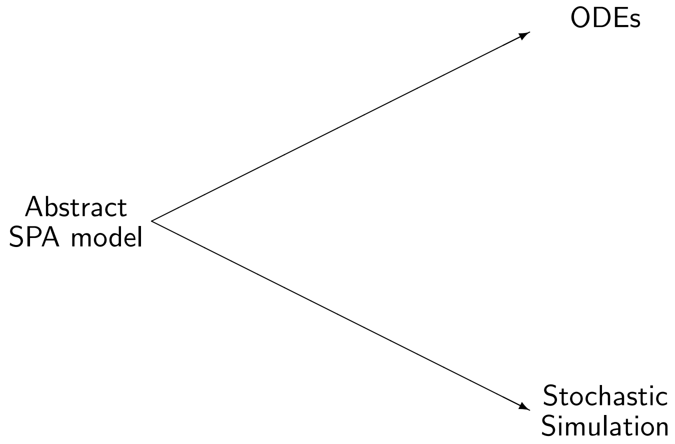
Our motivations for seeking more abstraction in process algebra models for systems biology comes from both key aspects of modelling:

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

Reagent-centric modelling

Reagent role	*Impact on reagent*	*Impact on reaction rate*
Producer	decreases concentration	has a positive impact, i.e. proportional to current concentration
Product	increases concentration	has no impact on the rate, except at saturation
Enzyme	concentration unchanged	has a positive impact, i.e. proportional to current concentration
Inhibitor	concentration unchanged	has a negative impact, i.e. inversely proportional to current concentration

Alternative Representations

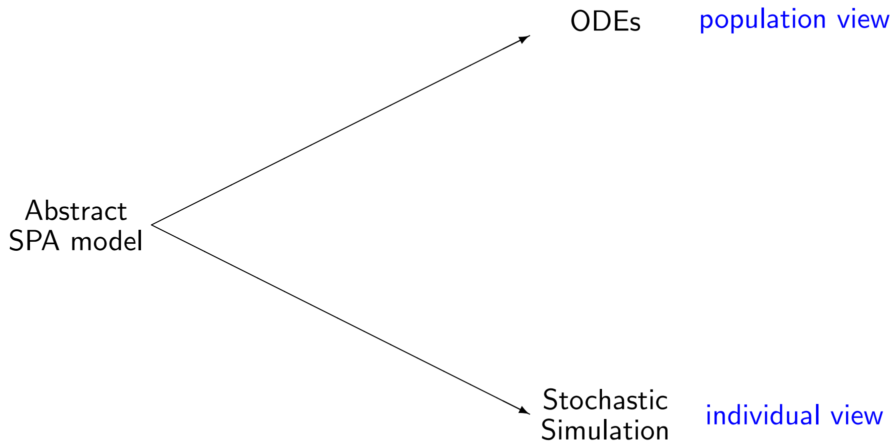


Abstract
SPA model

ODEs

Stochastic
Simulation

Alternative Representations



ODEs population view

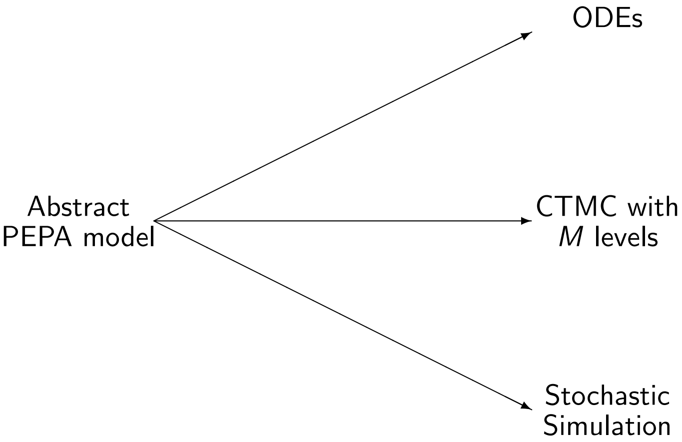
Abstract SPA model

Stochastic Simulation individual view

Discretising the population view

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



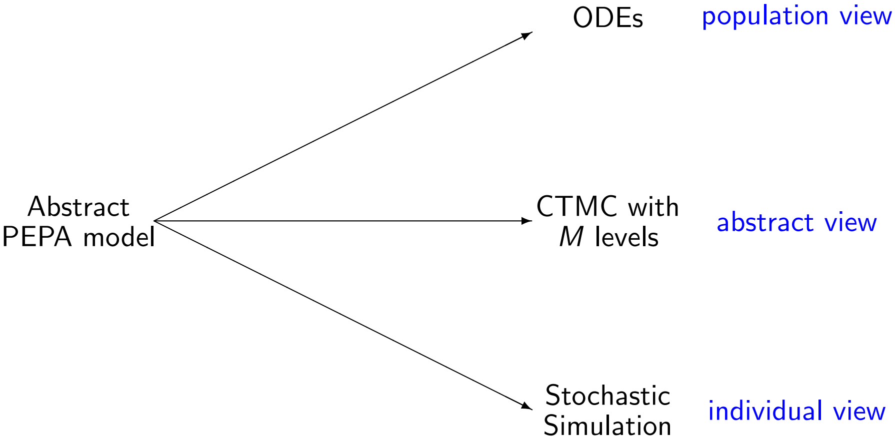
Abstract
PEPA model

ODEs

CTMC with
M levels

Stochastic
Simulation

Alternative Representations



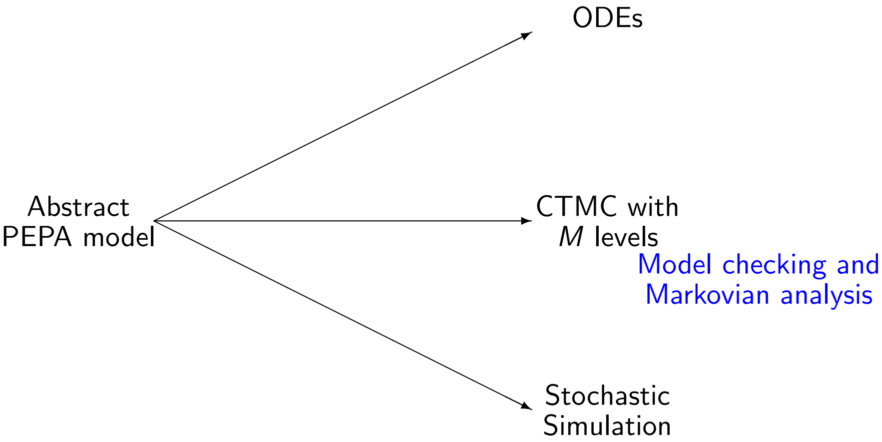
Abstract PEPA model

ODEs — population view

CTMC with M levels — abstract view

Stochastic Simulation — individual view

Alternative Representations



Abstract
PEPA model

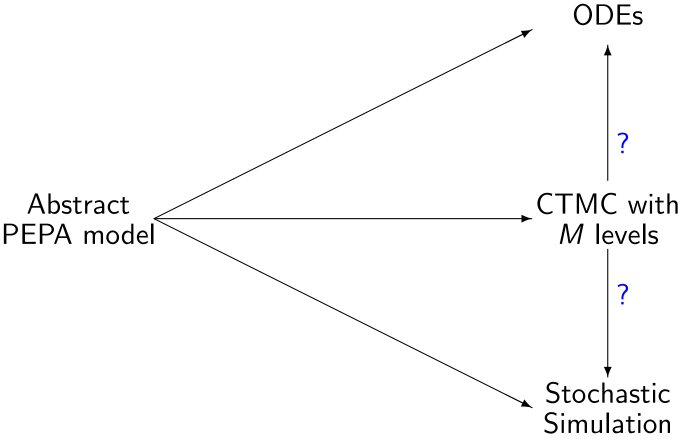
ODEs

CTMC with
M levels

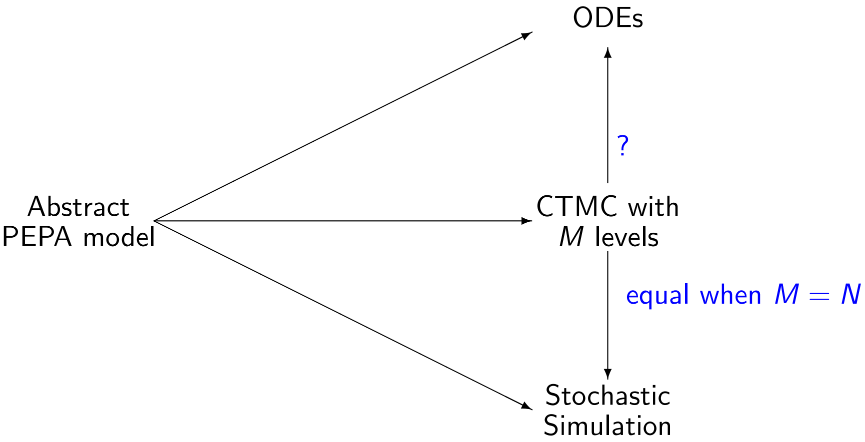
Model checking and
Markovian analysis

Stochastic
Simulation

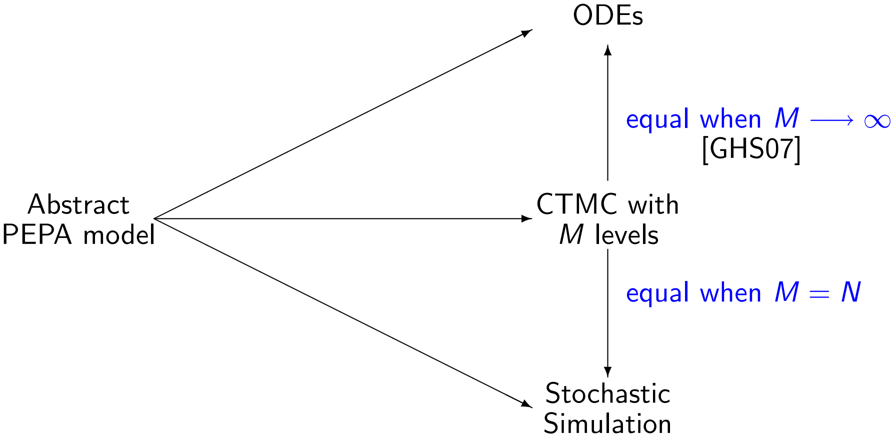
Alternative Representations



Alternative Representations



Alternative Representations



ODEs

equal when $M \longrightarrow \infty$ [GHS07]

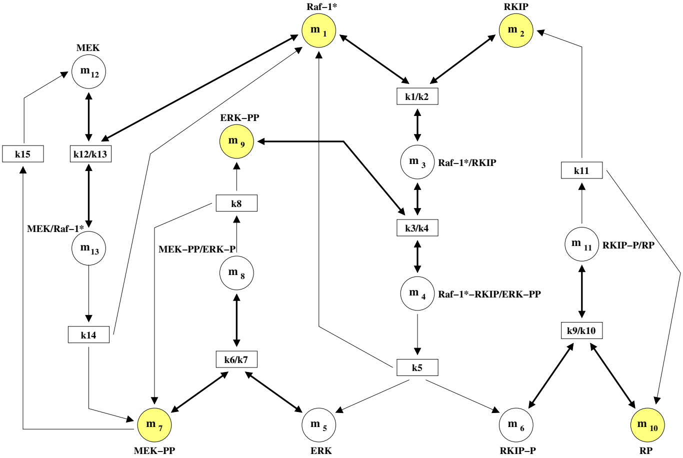
Abstract PEPA model

CTMC with M levels

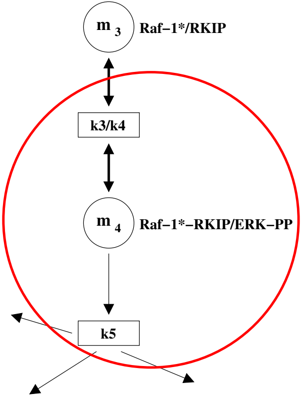
equal when $M = N$

Stochastic Simulation

Example: The Ras/Raf-1/MEK/ERK pathway



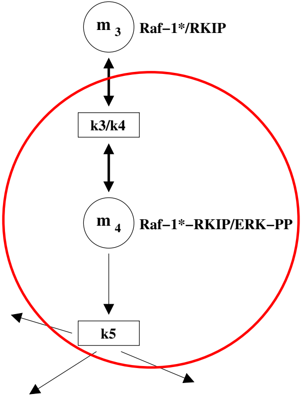
PEPA components of the reagent-centric model



$$\begin{aligned}\text{Raf-1}^*/\text{RKIP}/\text{ERK-PP}_\mathrm{H} &\overset{\mathit{def}}{=} (\mathit{k5product}, k_5).\text{Raf-1}^*/\text{RKIP}/\text{ERK-PP}_\mathrm{L} \\ &\quad + (\mathit{k4react}, k_4).\text{Raf-1}^*/\text{RKIP}/\text{ERK-PP}_\mathrm{L}\end{aligned}$$

$$\text{Raf-1}^*/\text{RKIP}/\text{ERK-PP}_\mathrm{L} \overset{\mathit{def}}{=} (\mathit{k3react}, k_3).\text{Raf-1}^*/\text{RKIP}/\text{ERK-PP}_\mathrm{H}$$

PEPA components of the reagent-centric model



$$
\begin{aligned}
\mathrm{Raf\text{-}1^*/RKIP/ERK\text{-}PP_H} &\overset{\mathit{def}}{=} \\
&(\mathit{k5product}, k_5).\mathrm{Raf\text{-}1^*/RKIP/ERK\text{-}PP_L} \\
&\quad + (\mathit{k4react}, k_4).\mathrm{Raf\text{-}1^*/RKIP/ERK\text{-}PP_L}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Raf\text{-}1^*/RKIP/ERK\text{-}PP_L} &\overset{\mathit{def}}{=} \\
&(\mathit{k3react}, k_3).\mathrm{Raf\text{-}1^*/RKIP/ERK\text{-}PP_H}
\end{aligned}
$$

Each reagent gives rise to a pair of PEPA definitions, one for high concentration and one for low concentration.

Commentary on the model

- I have shown the model with only high and low levels of concentration. In general we would discretise the concentration more coarsely with say 6 or 7 levels. As we add levels we are capturing the concentration at finer levels of granularity.

Commentary on the model

- I have shown the model with only high and low levels of concentration. In general we would discretise the concentration more coarsely with say 6 or 7 levels. As we add levels we are capturing the concentration at finer levels of granularity.
- The two levels (high/low) are sufficient to generate the ODEs or stochastic simulation.

Commentary on the model

- I have shown the model with only high and low levels of concentration. In general we would discretise the concentration more coarsely with say 6 or 7 levels. As we add levels we are capturing the concentration at finer levels of granularity.
- The two levels (high/low) are sufficient to generate the ODEs or stochastic simulation.
- For this model stochastic simulation and ODE analysis coincide.

Commentary on the model

- I have shown the model with only high and low levels of concentration. In general we would discretise the concentration more coarsely with say 6 or 7 levels. As we add levels we are capturing the concentration at finer levels of granularity.
- The two levels (high/low) are sufficient to generate the ODEs or stochastic simulation.
- For this model stochastic simulation and ODE analysis coincide.
- We also considered an alternative model of the pathway with subpathways as components, and we were able to use the process algebra equivalence to show that our two models have the same behaviour.

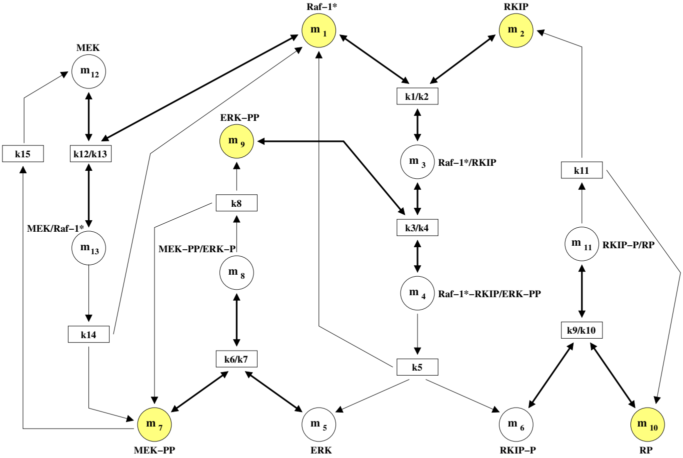
Reasoning and Model Checking

The original published model of the Ras/Raf-1/MEK/ERK pathway had a structure which allowed MEK to grow unboudedly which had not been detected in the ODE model.

Model checking using the PRISM probabilistic model checker allowed us to check properties such as whether one protein will exhibit a peak of concentration before another.

To do this we considered models with 6 levels of concentration, i.e. the range of possible values of concentration are split into 6 discrete levels — this shows good agreement with the ODE solution for transient behaviour.

Example: The Ras/Raf-1/MEK/ERK pathway



Markovian analysis

Numerical analysis of the CTMC can yield detailed information about the dynamic behaviour of the model.

Markovian analysis

Numerical analysis of the CTMC can yield 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.

Markovian analysis

Numerical analysis of the CTMC can yield 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.

Markovian analysis

Numerical analysis of the CTMC can yield detailed information about the dynamic behaviour of the model.

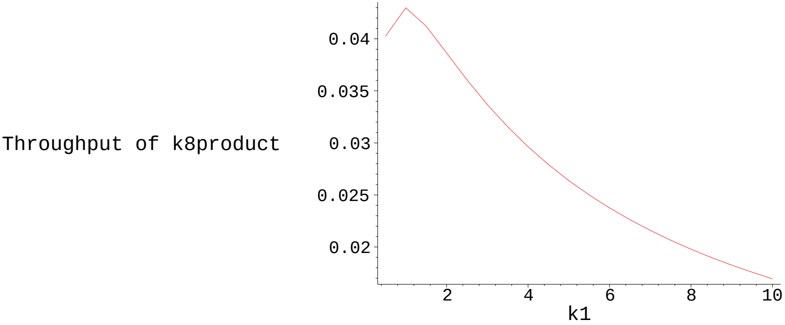
A steady state analysis provides statistics for average behaviour over a long run of the system, when the bias introduced by the initial state has been lost.

A transient analysis provides statistics relating to the evolution of the model over a fixed period. This will be dependent on the starting state.

Note, however, that a transient Markovian analysis is exact because it takes account of all possible evolutions, unlike a stochastic simulation which considers only one possible evolution in each run.

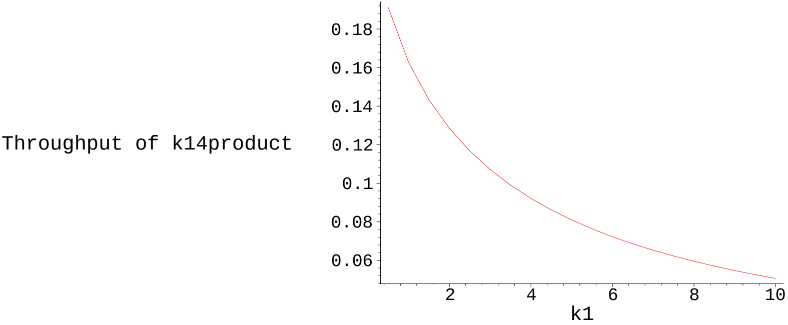
Quantified analysis – *k8product*

Approximating a variation in the initial concentration of RKIP by varying the rate constant k1, we can assess the impact on the production of ERK-PP.



Quantified analysis – *k14product*

Similarly we can assess the impact on the production of MEK-PP.



ODE analysis

Solving a system of ODEs will show how the concentrations of reagents vary over time.

ODE analysis

Solving a system of ODEs will show how the concentrations of reagents vary over time.

Solution is (relatively) fast and definitive but no variability is captured, unlike Markovian analyses (and real systems).

ODE analysis

Solving a system of ODEs will show how the concentrations of reagents vary over time.

Solution is (relatively) fast and definitive but no variability is captured, unlike Markovian analyses (and real systems).

There are advantages to be gained by using a process algebra model as an intermediary to the derivation of the ODEs.

ODE analysis

Solving a system of ODEs will show how the concentrations of reagents vary over time.

Solution is (relatively) fast and definitive but no variability is captured, unlike Markovian analyses (and real systems).

There are advantages to be gained by using a process algebra model as an intermediary to the derivation of the ODEs.

- The ODEs can be automatically generated from the descriptive process algebra model, thus reducing human error.

ODE analysis

Solving a system of ODEs will show how the concentrations of reagents vary over time.

Solution is (relatively) fast and definitive but no variability is captured, unlike Markovian analyses (and real systems).

There are advantages to be gained by using a process algebra model as an intermediary to the derivation of the ODEs.

- The ODEs can be automatically generated from the descriptive process algebra model, thus reducing human error.
- We can derive properties of the process algebra model, eg. freedom from deadlock, before numerical analysis is carried out.

ODE analysis

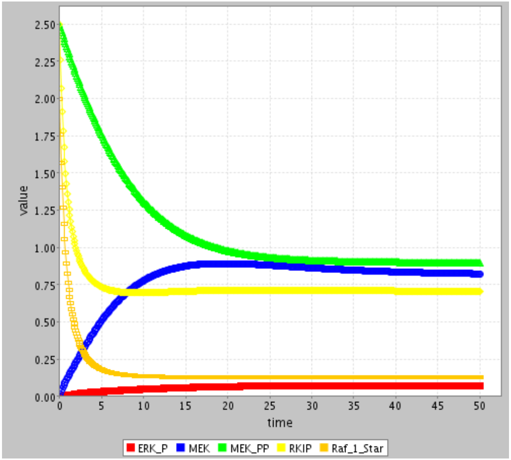
Solving a system of ODEs will show how the concentrations of reagents vary over time.

Solution is (relatively) fast and definitive but no variability is captured, unlike Markovian analyses (and real systems).

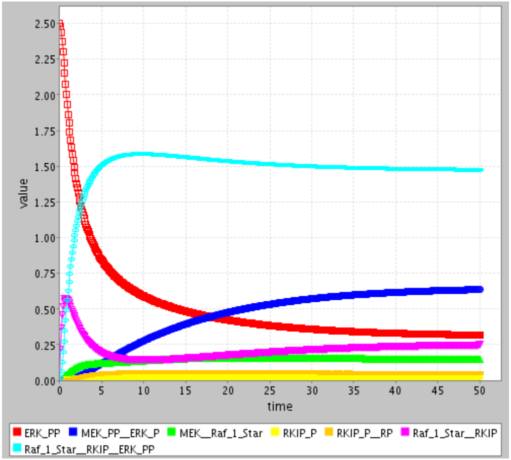
There are advantages to be gained by using a process algebra model as an intermediary to the derivation of the ODEs.

- The ODEs can be automatically generated from the descriptive process algebra model, thus reducing human error.
- We can derive properties of the process algebra model, eg. freedom from deadlock, before numerical analysis is carried out.
- The algebraic formulation of the model emphasises interactions between the biochemical entities.

ODE Analysis of the MAPK example



ODE Analysis of the MAPK example



Outline

- Introduction to Systems Biology
  - Motivation
- Challenges
- Stochastic Process Algebra for Biology
  - Stochastic Process Algebra
  - Abstraction, Modularity and Reasoning
  - Case Study
- Conclusions

Conclusions

- Ultimately we want to understand the functioning of cells as useful levels of abstraction, and to predict unknown behaviour.

Conclusions

- Ultimately we want to understand the functioning of cells as useful levels of abstraction, and to predict unknown behaviour.
- It remains an open and challenging problem to define a set of basic and general primitives for modelling biological systems, inspired by biological processes.

Conclusions

- Ultimately we want to understand the functioning of cells as useful levels of abstraction, and to predict unknown behaviour.
- It remains an open and challenging problem to define a set of basic and general primitives for modelling biological systems, inspired by biological processes.
- Achieving this goal is anticipated to have two broad benefits:

Conclusions

- Ultimately we want to understand the functioning of cells as useful levels of abstraction, and to predict unknown behaviour.
- It remains an open and challenging problem to define a set of basic and general primitives for modelling biological systems, inspired by biological processes.
- Achieving this goal is anticipated to have two broad benefits:
  - Better models and simulations of living phenomena

Conclusions

- Ultimately we want to understand the functioning of cells as useful levels of abstraction, and to predict unknown behaviour.
- It remains an open and challenging problem to define a set of basic and general primitives for modelling biological systems, inspired by biological processes.
- Achieving this goal is anticipated to have two broad benefits:
  - Better models and simulations of living phenomena
  - New models of computations that are biologically inspired.

Thank You!

Thank You!

Collaborators: Muffy Calder, Federica Ciocchetta, Adam Duguid, Nil Geisweiller, Stephen Gilmore and Marco Stenico.

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