SPA for quantitative analysis: Lecture 6 — Modelling Biological Processes

Jane Hillston

LFCS, School of Informatics
The University of Edinburgh
Scotland

7th March 2013

Outline

Outline

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.
- Formalisms from theoretical computer science have found a new role in developing models for systems biology, allowing biologists to test hypotheses and prioritise experiments.

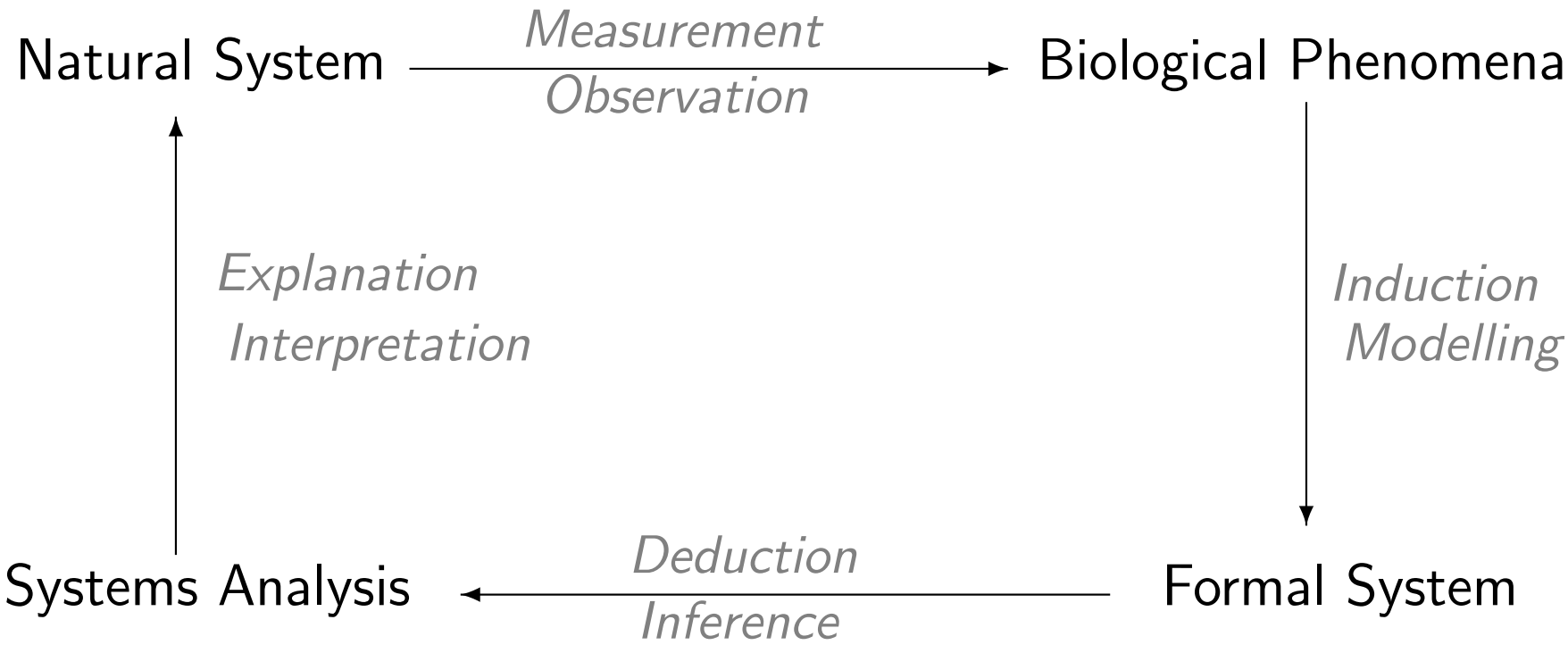
Systems Biology

- 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.
- Formalisms from theoretical computer science have found a new role in developing models for systems biology, allowing biologists to test hypotheses and prioritise experiments.

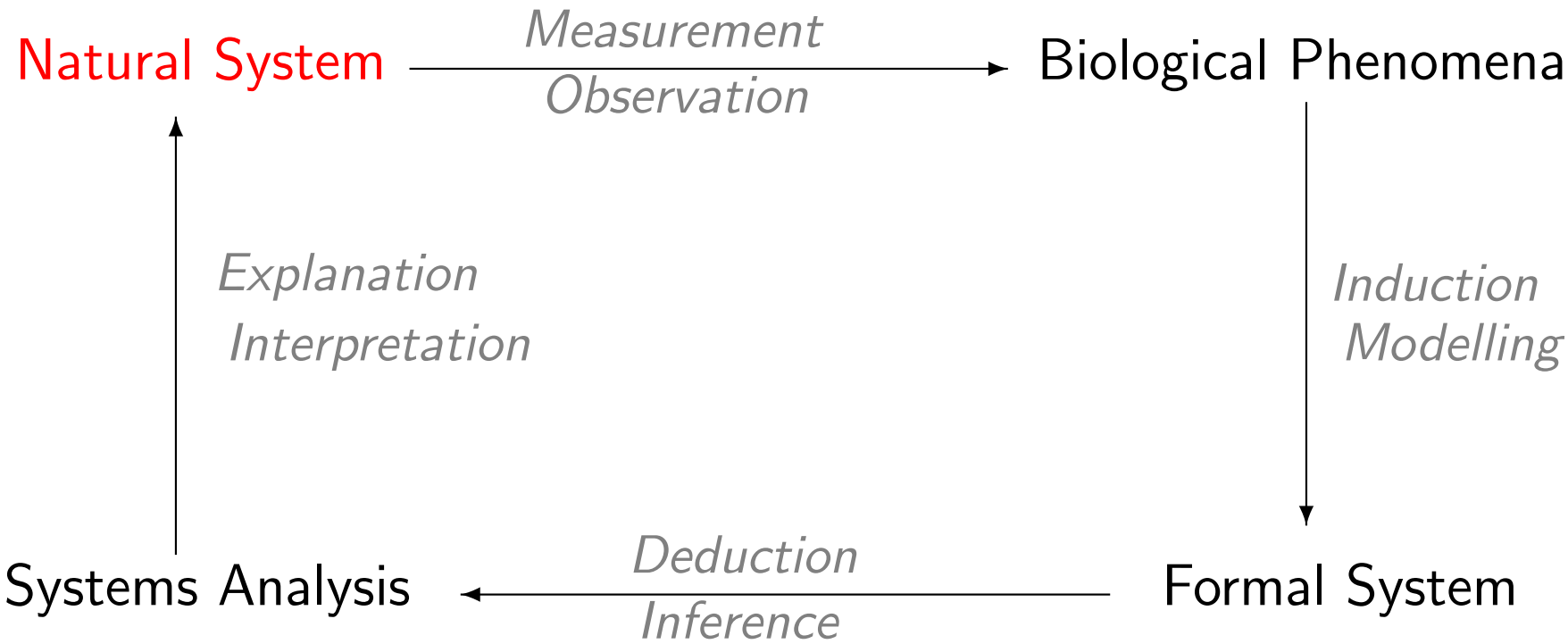
Systems Biology

- 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.
- Formalisms from theoretical computer science have found a new role in developing models for systems biology, allowing biologists to test hypotheses and prioritise experiments.

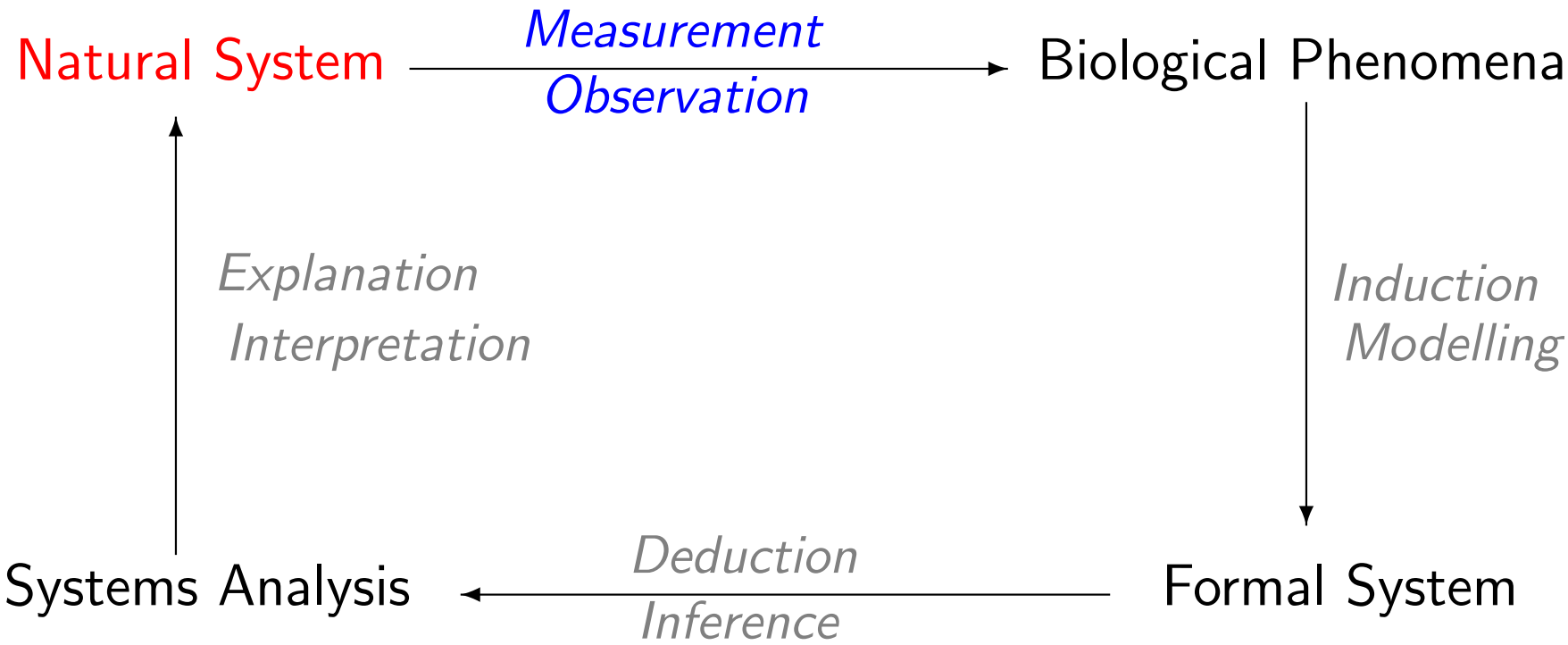
Systems Biology Methodology



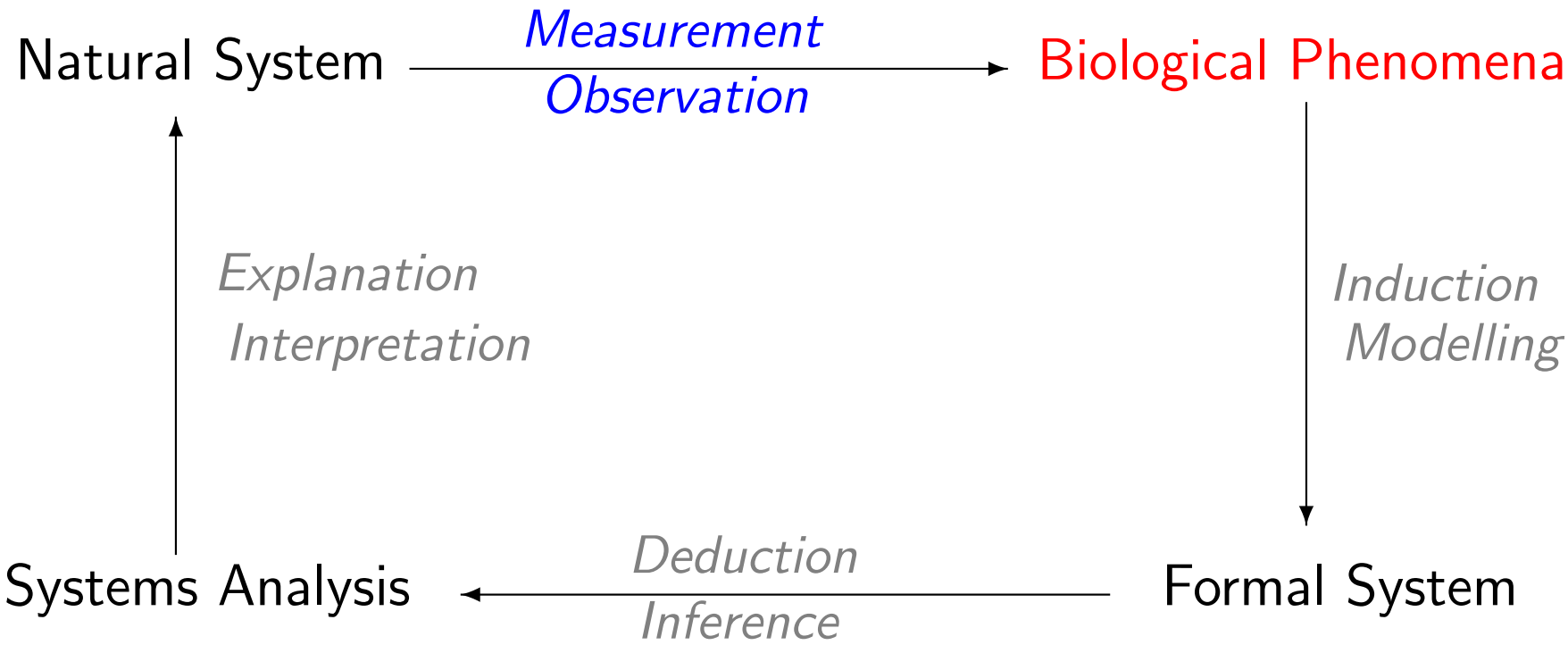
Systems Biology Methodology



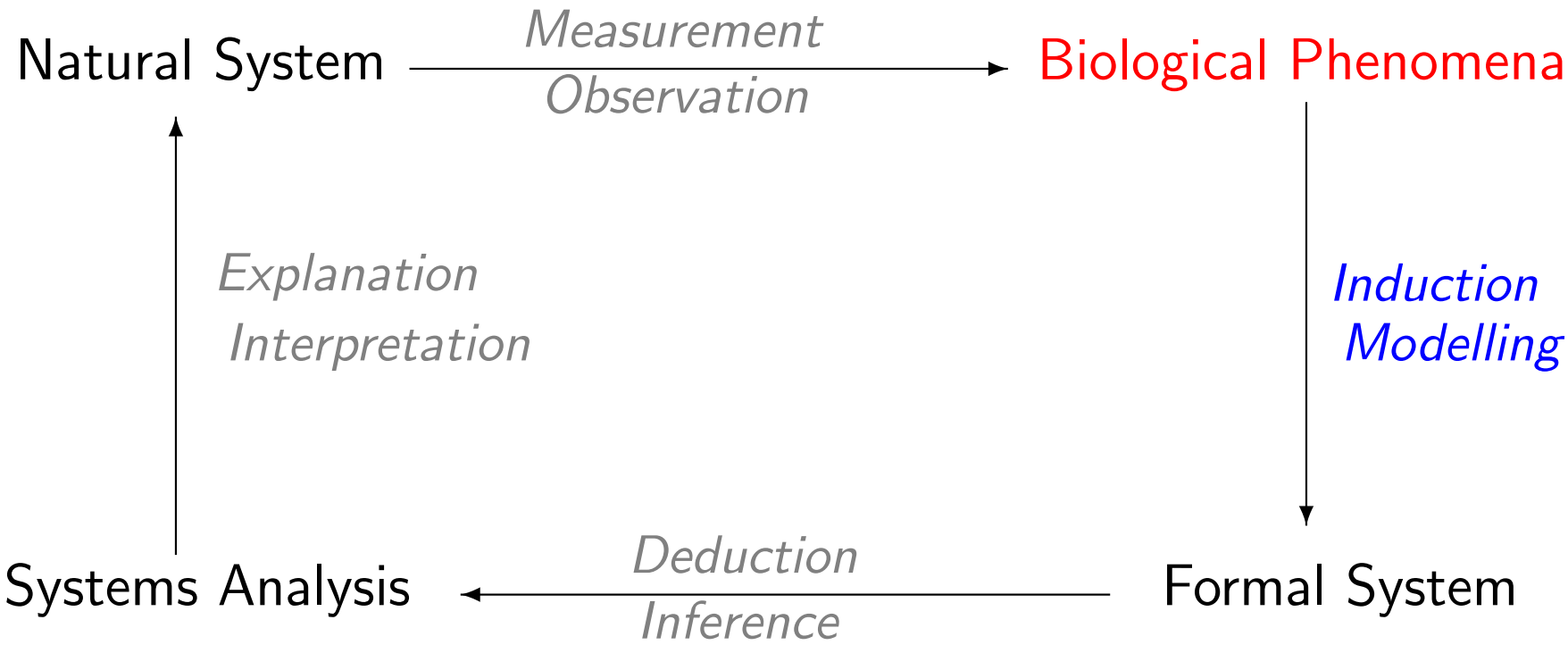
Systems Biology Methodology



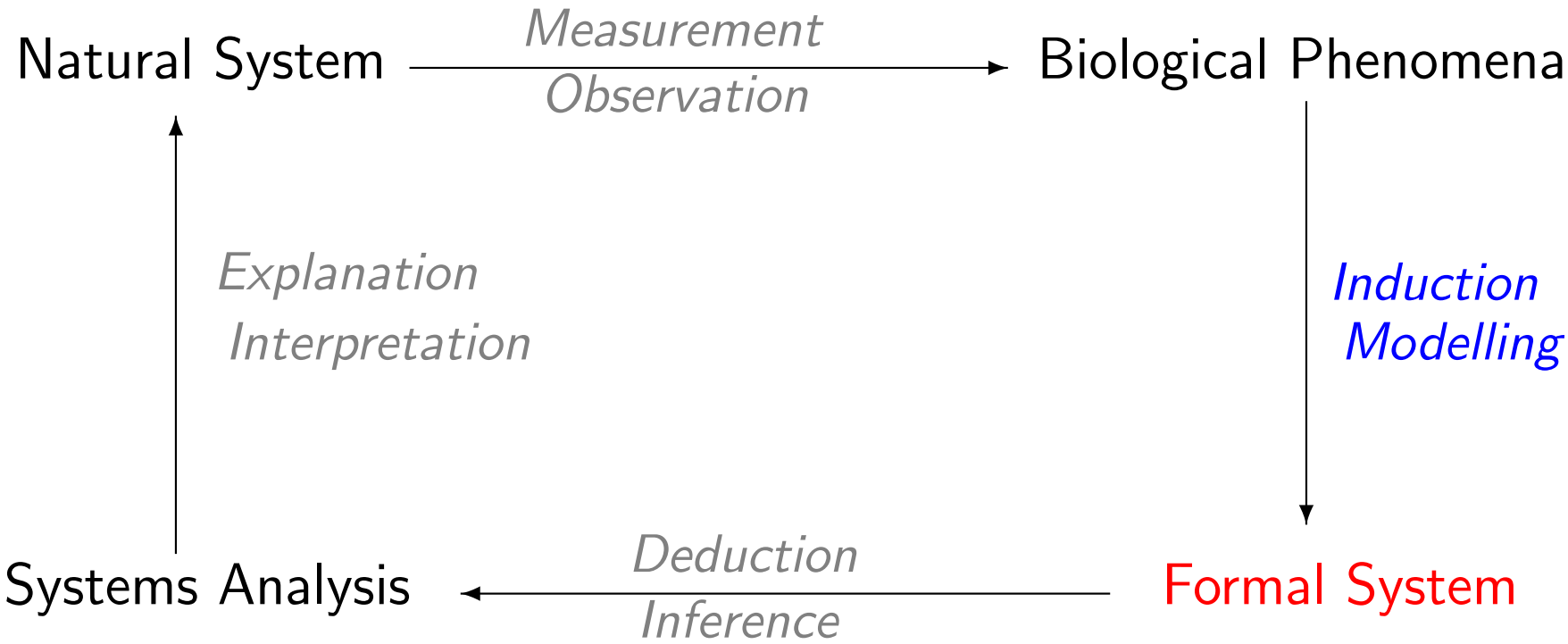
Systems Biology Methodology



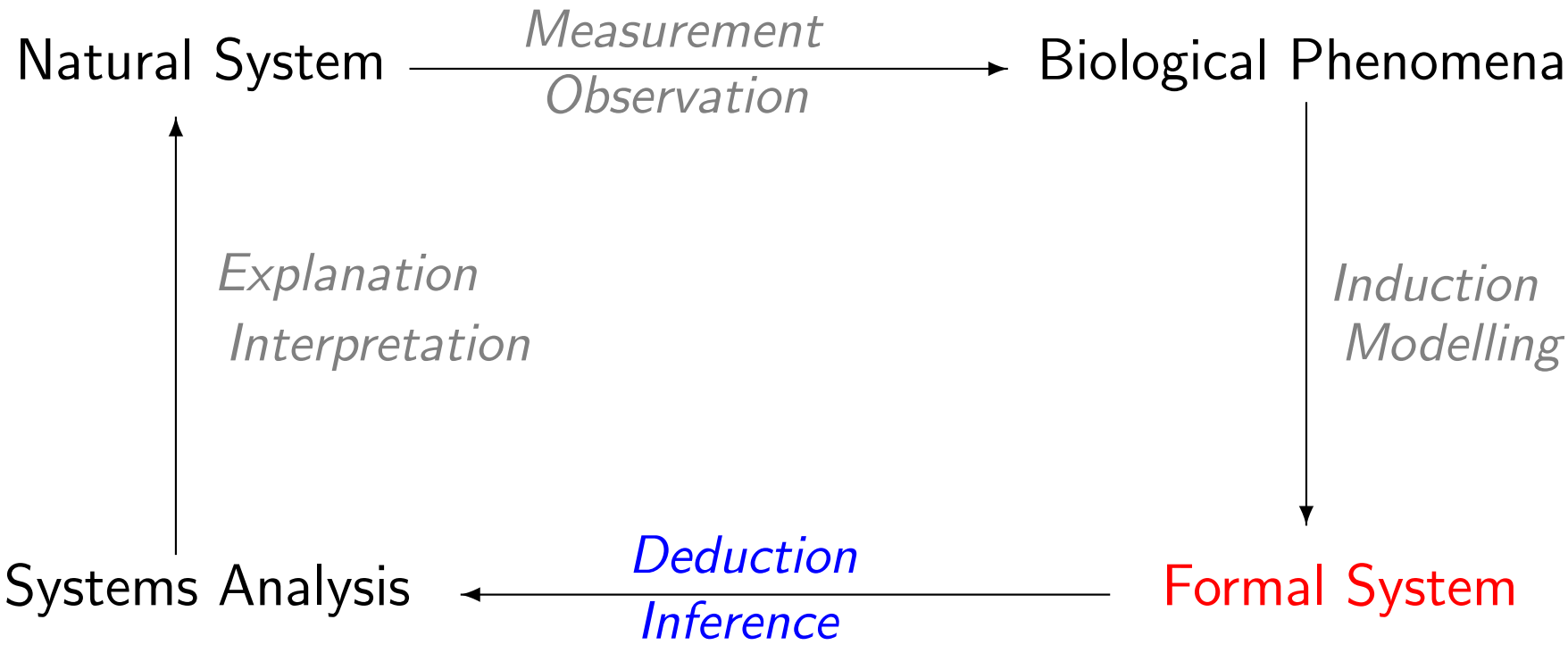
Systems Biology Methodology



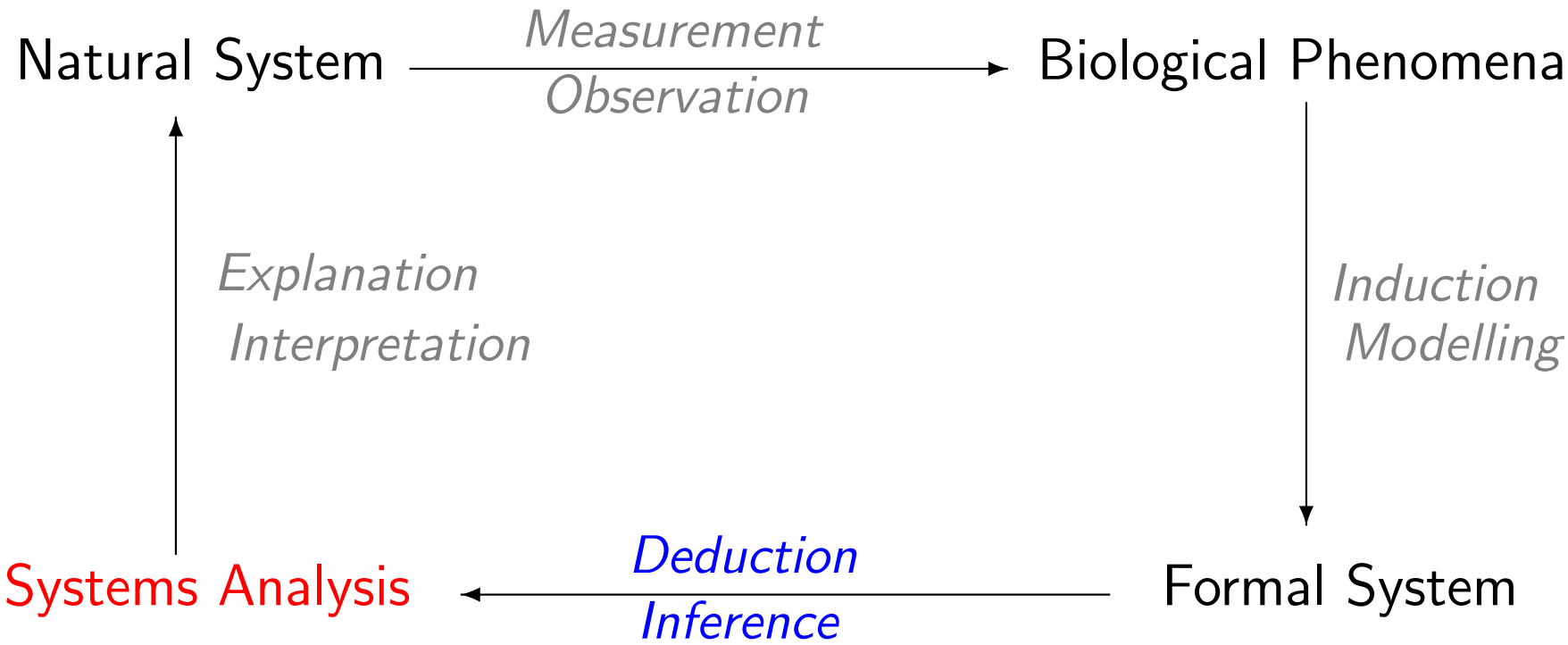
Systems Biology Methodology



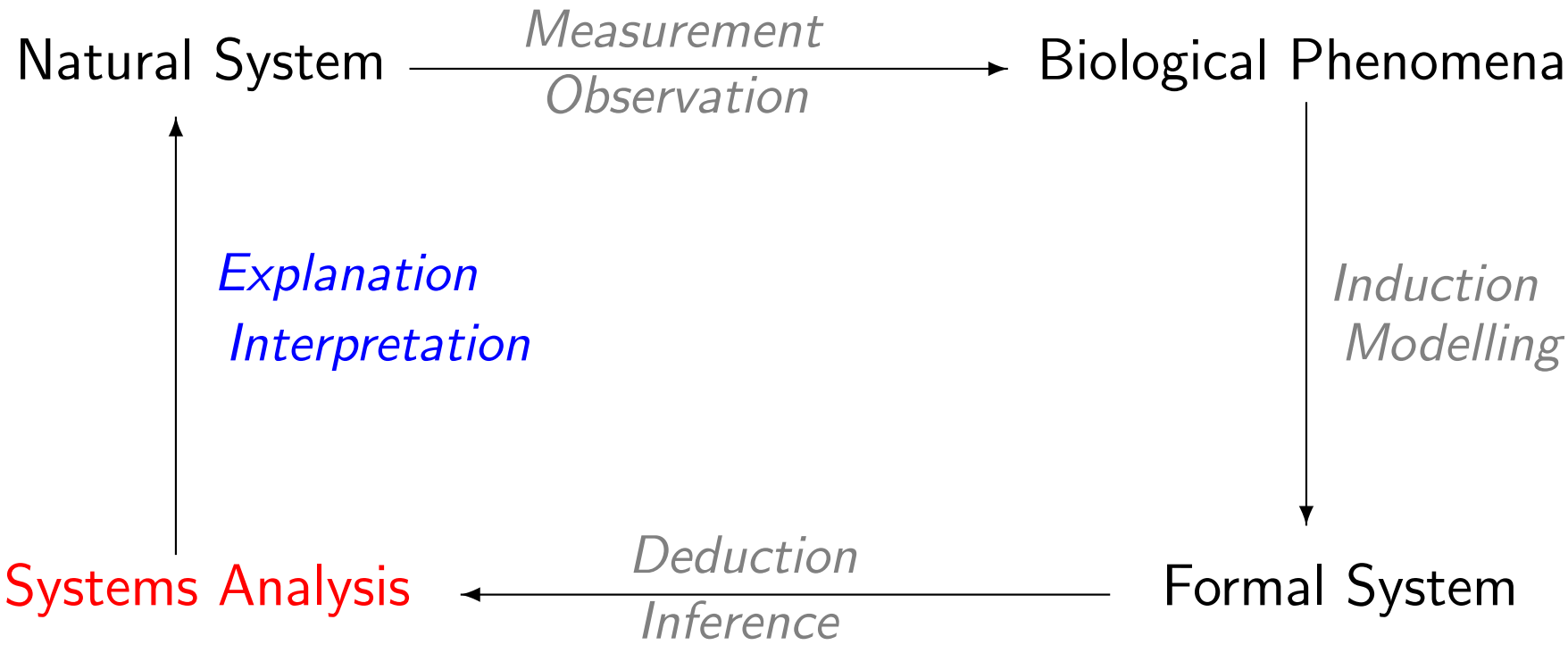
Systems Biology Methodology



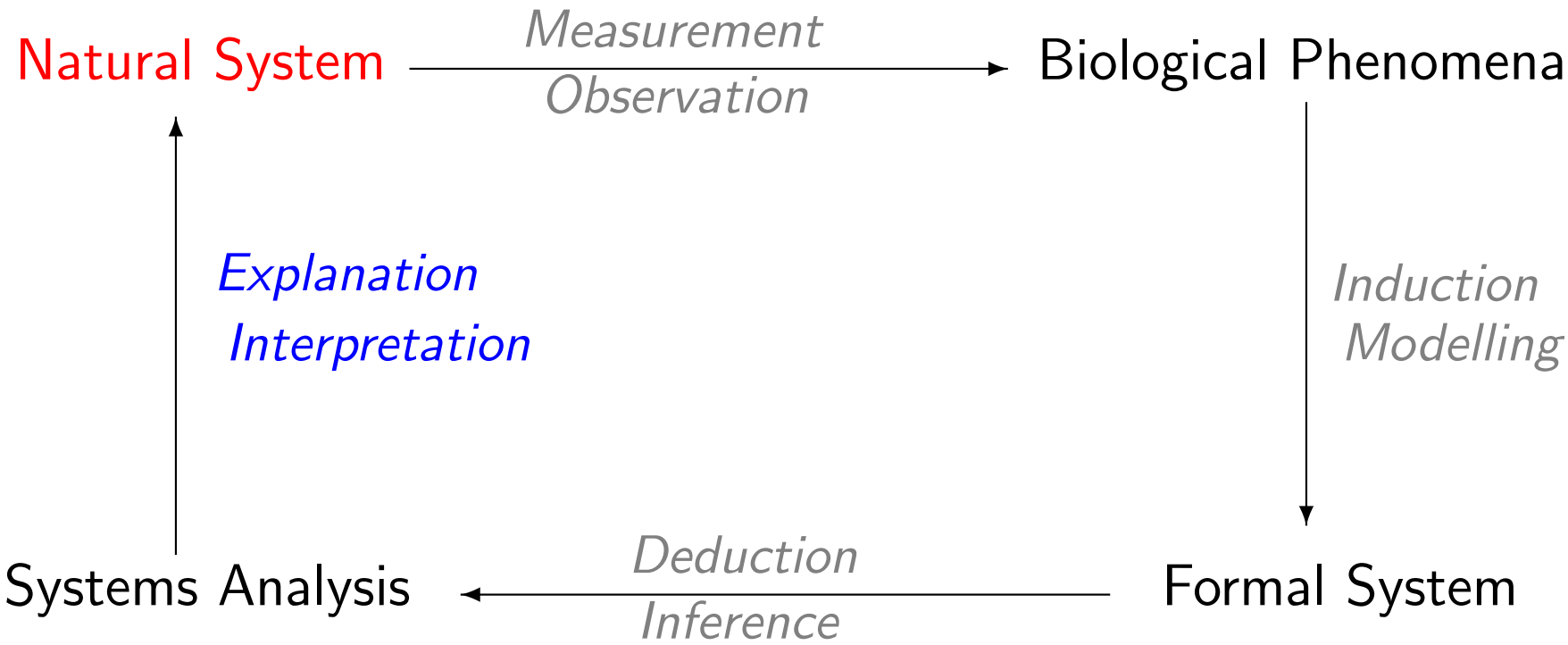
Systems Biology Methodology



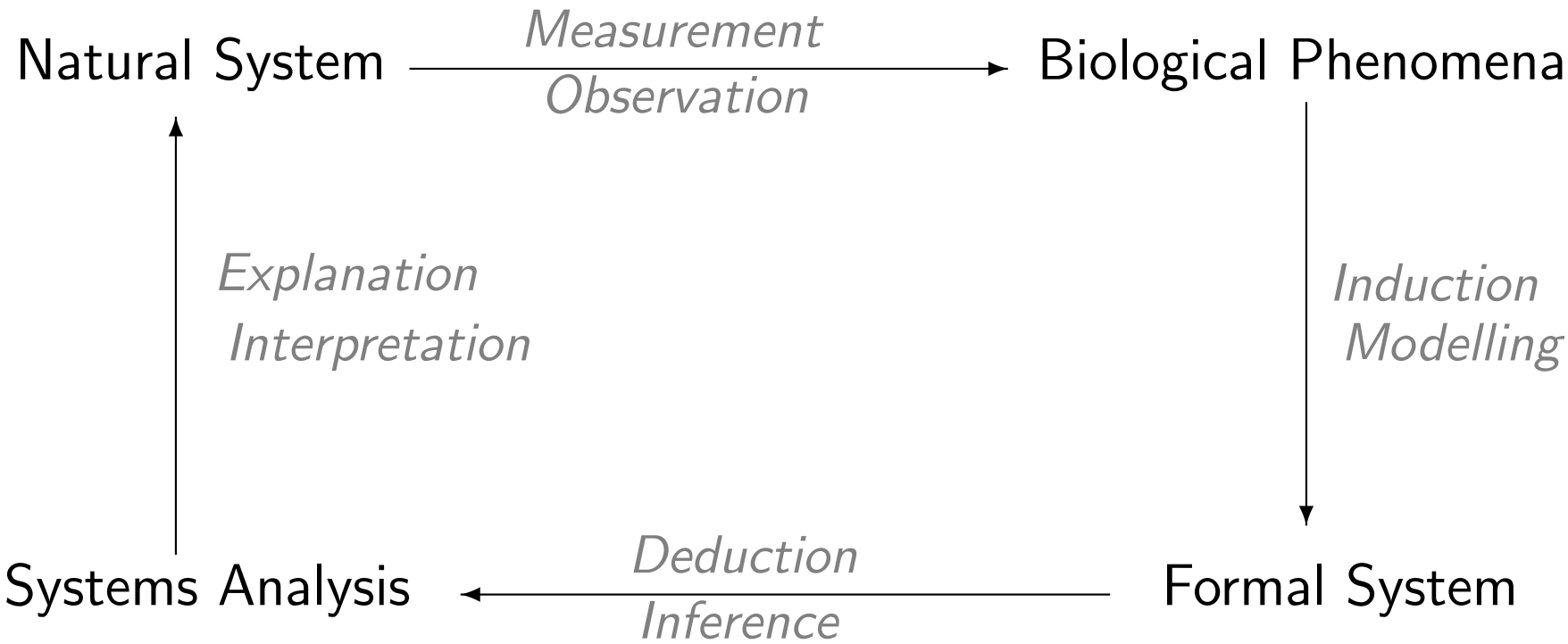
Systems Biology Methodology



Systems Biology Methodology



Systems Biology Methodology



Networks in cells

We can distinguish three distinct types of links or networks in cells

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.

Networks in cells

We can distinguish three distinct types of links or networks in cells

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.

Networks in cells

We can distinguish three distinct types of links or networks in cells

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.

Networks in cells

We can distinguish three distinct types of links or networks in cells

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

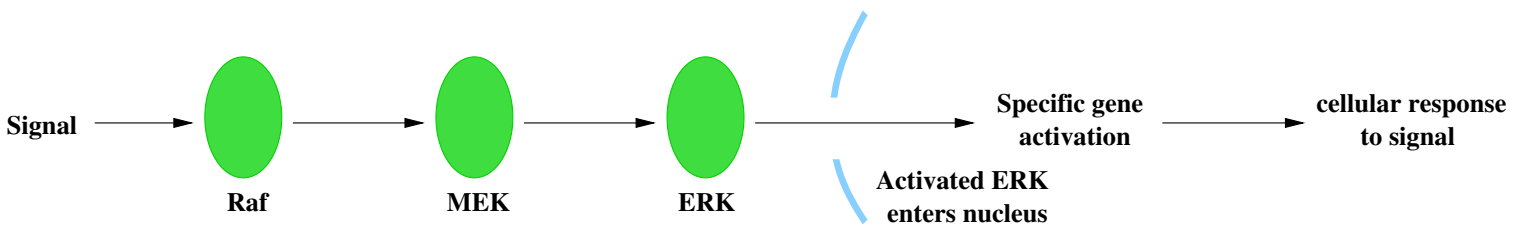
Signal transduction networks: External stimuli initiate messages that are carried through a cell via a cascade of biochemical reactions.

Metabolic pathways: The survival of the cell depends on its ability to transform nutrients into energy.

Extracellular signalling

Extracellular signalling — communication between cells.

- Signalling molecules released by one cell migrate to another;
- These molecules enter the cell and instigate a pathway, or series of reactions, which carries the information from the membrane to the nucleus;
- For example, the Ras/Raf-1/MEK/ERK pathway conveys differentiation signals to the nucleus of a cell.



Cell signalling

- All signalling is biochemical:
- Increasing protein concentration broadcasts the information about an event; for example, that a gene promoter is “on”.
- 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.

Cell signalling

- All signalling is biochemical:
- Increasing protein concentration broadcasts the information about an event; for example, that a gene promoter is “on”.
- 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.

Cell signalling

- All signalling is biochemical:
- Increasing protein concentration broadcasts the information about an event; for example, that a gene promoter is “on”.
- 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.

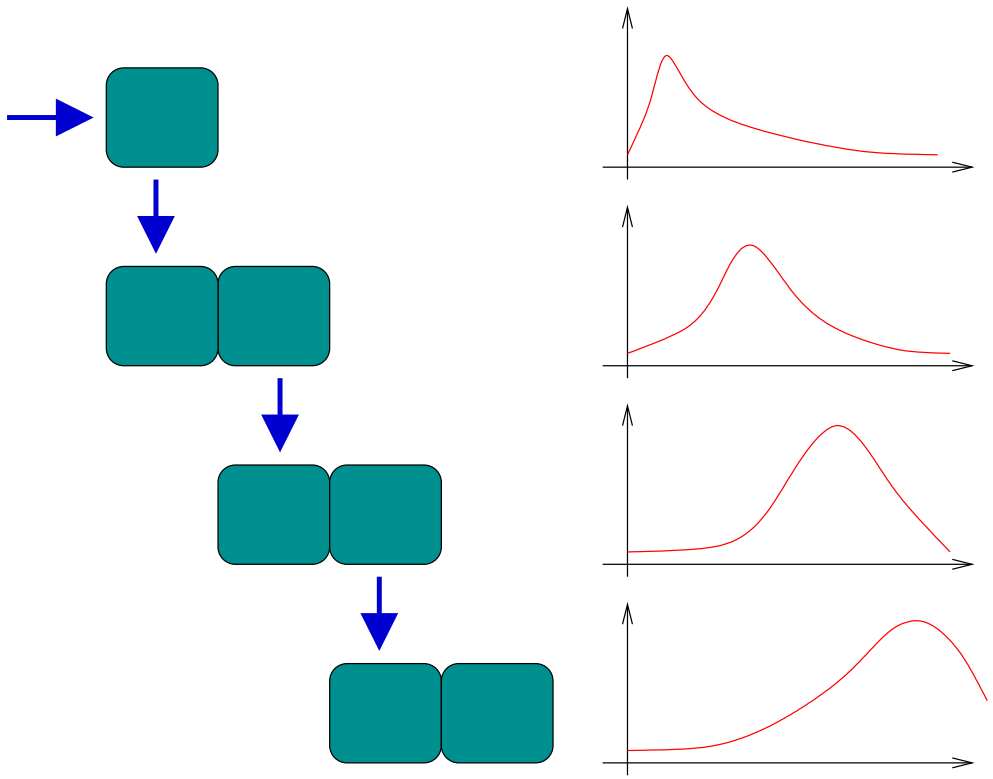
Cell signalling

- All signalling is biochemical:
- Increasing protein concentration broadcasts the information about an event; for example, that a gene promoter is “on”.
- 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.

Cell signalling

- All signalling is biochemical:
- Increasing protein concentration broadcasts the information about an event; for example, that a gene promoter is “on”.
- 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.

Signal transduction pathways



A series of biochemical reactions serve to pass a message from the cell membrane to the nucleus.

Gene expression pathways

- Genetic activity is controlled by molecular signals that determine when and how often a given gene is transcribed.
- The product encoded by one gene often regulates the expression of other genes.
- For appropriate combinations of input signals transcription is initiated and protein product accumulates when production exceeds degradation.
- Links are established between genes when the product of one regulates the expression of another.
- Thus networks of interaction can be deduced and these may be quite complex.

Gene expression pathways

- Genetic activity is controlled by molecular signals that determine when and how often a given gene is transcribed.
- The product encoded by one gene often regulates the expression of other genes.
- For appropriate combinations of input signals transcription is initiated and protein product accumulates when production exceeds degradation.
- Links are established between genes when the product of one regulates the expression of another.
- Thus networks of interaction can be deduced and these may be quite complex.

Gene expression pathways

- Genetic activity is controlled by molecular signals that determine when and how often a given gene is transcribed.
- The product encoded by one gene often regulates the expression of other genes.
- For appropriate combinations of input signals transcription is initiated and protein product accumulates when production exceeds degradation.
- Links are established between genes when the product of one regulates the expression of another.
- Thus networks of interaction can be deduced and these may be quite complex.

Gene expression pathways

- Genetic activity is controlled by molecular signals that determine when and how often a given gene is transcribed.
- The product encoded by one gene often regulates the expression of other genes.
- For appropriate combinations of input signals transcription is initiated and protein product accumulates when production exceeds degradation.
- Links are established between genes when the product of one regulates the expression of another.
- Thus networks of interaction can be deduced and these may be quite complex.

Gene expression pathways

- Genetic activity is controlled by molecular signals that determine when and how often a given gene is transcribed.
- The product encoded by one gene often regulates the expression of other genes.
- For appropriate combinations of input signals transcription is initiated and protein product accumulates when production exceeds degradation.
- Links are established between genes when the product of one regulates the expression of another.
- Thus networks of interaction can be deduced and these may be quite complex.

Dynamic issues

- In biochemical regulatory networks, the delay between events are determined by the delay while signal molecule concentrations accumulate or decline sufficiently.
- All reactions rely on the random process of molecules colliding with the cell.
- Thus the accumulation of protein is a stochastic process affected by several factors in the cell (temperature, pH, etc.).
- Thus the "reaction time" is a distribution rather than a deterministic time, and this can be shown to be an exponential distribution for basis molecular reactions.

Dynamic issues

- In biochemical regulatory networks, the delay between events are determined by the delay while signal molecule concentrations accumulate or decline sufficiently.
- All reactions rely on the random process of molecules colliding with the cell.
- Thus the accumulation of protein is a stochastic process affected by several factors in the cell (temperature, pH, etc.).
- Thus the "reaction time" is a distribution rather than a deterministic time, and this can be shown to be an exponential distribution for basis molecular reactions.

Dynamic issues

- In biochemical regulatory networks, the delay between events are determined by the delay while signal molecule concentrations accumulate or decline sufficiently.
- All reactions rely on the random process of molecules colliding with the cell.
- Thus the accumulation of protein is a stochastic process affected by several factors in the cell (temperature, pH, etc.).
- Thus the "reaction time" is a distribution rather than a deterministic time, and this can be shown to be an exponential distribution for basis molecular reactions.

Dynamic issues

- In biochemical regulatory networks, the delay between events are determined by the delay while signal molecule concentrations accumulate or decline sufficiently.
- All reactions rely on the random process of molecules colliding with the cell.
- Thus the accumulation of protein is a stochastic process affected by several factors in the cell (temperature, pH, etc.).
- Thus the "reaction time" is a distribution rather than a deterministic time, and this can be shown to be an exponential distribution for basis molecular reactions.

Stochastic behaviour

- The stochastic reaction rate of a chemical reaction is a function of only those molecular species involved as reactants or catalysts, and a stochastic rate constant c.
- The stochastic rate constant takes into account volume, temperature, pH and other environmental factors.
- The stoichiometry of the reaction — how many molecules of each reactant species are required — also has an impact.
- Commonly the law of mass action is used to determine the effective rate of a reaction: this is the product of the stochastic rate constant and the amount of each of the reactants.

Using Stochastic Process Algebras

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

- Process algebraic formulations are compositional and make interactions/constraints explicit.
- Structure can also be apparent.
- 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.

Using Stochastic Process Algebras

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

- Process algebraic formulations are compositional and make interactions/constraints explicit.
- Structure can also be apparent.
- 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.

Using Stochastic Process Algebras

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

- Process algebraic formulations are compositional and make interactions/constraints explicit.
- Structure can also be apparent.
- 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.

Using Stochastic Process Algebras

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

- Process algebraic formulations are compositional and make interactions/constraints explicit.
- Structure can also be apparent.
- 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.

Using Stochastic Process Algebras

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

- Process algebraic formulations are compositional and make interactions/constraints explicit.
- Structure can also be apparent.
- 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.

Molecular processes as concurrent computations

A correspondence between cellular processes and process alebras was first highlighted by Regev and her co-authors in the early 2000s.

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

A correspondence between cellular processes and process alebras was first highlighted by Regev and her co-authors in the early 2000s.

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]

Calculi for Systems Biology

Several different process caluli have been developed or adapted for application in systems biology. Each of them has different properties able to render different aspects of biological phenomena. They may be divided into two main categories:

- Calculi defined originally in computer science and then applied in biology, such as the biochemical stochastic π-calculus, SCCP, CCS-R and PEPA;
- Calculi defined specifically by observing biological structures and phenomena, such as BioAmbients, Brane Calculi, Beta-binders, BlenX and Bio-PEPA.

Calculi for Systems Biology

Several different process caluli have been developed or adapted for application in systems biology. Each of them has different properties able to render different aspects of biological phenomena. They may be divided into two main categories:

- Calculi defined originally in computer science and then applied in biology, such as the biochemical stochastic π-calculus, SCCP, CCS-R and PEPA;
- Calculi defined specifically by observing biological structures and phenomena, such as BioAmbients, Brane Calculi, Beta-binders, BlenX and Bio-PEPA.

Calculi for Systems Biology

Several different process caluli have been developed or adapted for application in systems biology. Each of them has different properties able to render different aspects of biological phenomena. They may be divided into two main categories:

- Calculi defined originally in computer science and then applied in biology, such as the biochemical stochastic π-calculus, SCCP, CCS-R and PEPA;
- Calculi defined specifically by observing biological structures and phenomena, such as BioAmbients, Brane Calculi, Beta-binders, BlenX and Bio-PEPA.

Calculi for Systems Biology

Several different process caluli have been developed or adapted for application in systems biology. Each of them has different properties able to render different aspects of biological phenomena. They may be divided into two main categories:

- Calculi defined originally in computer science and then applied in biology, such as the biochemical stochastic π-calculus, SCCP, CCS-R and PEPA;
- Calculi defined specifically by observing biological structures and phenomena, such as BioAmbients, Brane Calculi, Beta-binders, BlenX and Bio-PEPA.

Stochastic π-calculus

- The stochastic π-calculus has been used to model and analyse a wide variety of biological systems.
- Examples include metabolic pathways, gene transcription and signal transduction.
- Analysis is mainly based on stochastic simulation (Gillespie's algorithm).
- Two tools: BioSPI and SPIM which implement slightly different versions of the language.
- There has also been some work on a graphical notation associated with the SPIM tool.

Stochastic π-calculus

- The stochastic π-calculus has been used to model and analyse a wide variety of biological systems.
- Examples include metabolic pathways, gene transcription and signal transduction.
- Analysis is mainly based on stochastic simulation (Gillespie's algorithm).
- Two tools: BioSPI and SPIM which implement slightly different versions of the language.
- There has also been some work on a graphical notation associated with the SPIM tool.

Stochastic π-calculus

- The stochastic π-calculus has been used to model and analyse a wide variety of biological systems.
- Examples include metabolic pathways, gene transcription and signal transduction.
- Analysis is mainly based on stochastic simulation (Gillespie's algorithm).
- Two tools: BioSPI and SPIM which implement slightly different versions of the language.
- There has also been some work on a graphical notation associated with the SPIM tool.

Stochastic π-calculus

- The stochastic π-calculus has been used to model and analyse a wide variety of biological systems.
- Examples include metabolic pathways, gene transcription and signal transduction.
- Analysis is mainly based on stochastic simulation (Gillespie's algorithm).
- Two tools: BioSPI and SPIM which implement slightly different versions of the language.
- There has also been some work on a graphical notation associated with the SPIM tool.

Stochastic π-calculus

- The stochastic π-calculus has been used to model and analyse a wide variety of biological systems.
- Examples include metabolic pathways, gene transcription and signal transduction.
- Analysis is mainly based on stochastic simulation (Gillespie's algorithm).
- Two tools: BioSPI and SPIM which implement slightly different versions of the language.
- There has also been some work on a graphical notation associated with the SPIM tool.

Stochastic π-calculus

Stochastic π-calculus [*Priami, 1995*] extends the π-calculus with exponentially-distributed rates.

Stochastic π-calculus

Stochastic π-calculus [*Priami, 1995*] extends the π-calculus with exponentially-distributed rates.

0	Nil

Stochastic π-calculus

Stochastic π-calculus [*Priami, 1995*] extends the π-calculus with exponentially-distributed rates.

0	Nil
$(\pi, r).P$	Prefix

Stochastic π-calculus

Stochastic π-calculus [*Priami, 1995*] extends the π-calculus with exponentially-distributed rates.

0	Nil
$(\pi, r).P$	Prefix
$(\nu n)P$	New

Stochastic π-calculus

Stochastic π-calculus [*Priami, 1995*] extends the π-calculus with exponentially-distributed rates.

0	Nil
$(\pi, r).P$	Prefix
$(\nu n)P$	New
$[x = y]P$	Matching

Stochastic π-calculus

Stochastic π-calculus [*Priami, 1995*] extends the π-calculus with exponentially-distributed rates.

0	Nil
$(\pi, r).P$	Prefix
$(\nu n)P$	New
$[x = y]P$	Matching
$P_1 \mid P_2$	Parallel

Stochastic π-calculus

Stochastic π-calculus [*Priami, 1995*] extends the π-calculus with exponentially-distributed rates.

0	Nil
$(\pi, r).P$	Prefix
$(\nu n)P$	New
$[x = y]P$	Matching
$P_1 \mid P_2$	Parallel
$P_1 + P_2$	Choice

Stochastic π-calculus

Stochastic π-calculus [*Priami, 1995*] extends the π-calculus with exponentially-distributed rates.

0	Nil
$(\pi, r).P$	Prefix
$(\nu n)P$	New
$[x = y]P$	Matching
$P_1 \mid P_2$	Parallel
$P_1 + P_2$	Choice
$!P$	Replication

Stochastic π-calculus

Stochastic π-calculus [*Priami, 1995*] extends the π-calculus with exponentially-distributed rates.

0	Nil
$(\pi, r).P$	Prefix
$(\nu n)P$	New
$[x = y]P$	Matching
$P_1 \mid P_2$	Parallel
$P_1 + P_2$	Choice
$!P$	Replication

where π is either $x(y)$ (input), $\overline{x}y$ (output) or τ (silent).

Example: The VICE project

- The aim was to construct a minimal cell in silico in order to track the dynamics of a complete metabolome.
- Thus a VIrtual CEll was defined as a stochastic π-calculus model, which seems to behave as a simplified prokaryote.
- Started from a published minimal gene set which eliminated duplicated genes and other redundancies from the smallest known bacterial genomes. This was further reduced to 180 different genes.
- Experimental results were in accordance with those available from in vivo experiments.
- Some extensions to the stochastic π-calculus were needed.

Example: The VICE project

- The aim was to construct a minimal cell in silico in order to track the dynamics of a complete metabolome.
- Thus a VIrtual CEll was defined as a stochastic π-calculus model, which seems to behave as a simplified prokaryote.
- Started from a published minimal gene set which eliminated duplicated genes and other redundancies from the smallest known bacterial genomes. This was further reduced to 180 different genes.
- Experimental results were in accordance with those available from in vivo experiments.
- Some extensions to the stochastic π-calculus were needed.

Example: The VICE project

- The aim was to construct a minimal cell in silico in order to track the dynamics of a complete metabolome.
- Thus a VIrtual CEll was defined as a stochastic π-calculus model, which seems to behave as a simplified prokaryote.
- Started from a published minimal gene set which eliminated duplicated genes and other redundancies from the smallest known bacterial genomes. This was further reduced to 180 different genes.
- Experimental results were in accordance with those available from in vivo experiments.
- Some extensions to the stochastic π-calculus were needed.

Example: The VICE project

- The aim was to construct a minimal cell in silico in order to track the dynamics of a complete metabolome.
- Thus a VIrtual CEll was defined as a stochastic π-calculus model, which seems to behave as a simplified prokaryote.
- Started from a published minimal gene set which eliminated duplicated genes and other redundancies from the smallest known bacterial genomes. This was further reduced to 180 different genes.
- Experimental results were in accordance with those available from in vivo experiments.
- Some extensions to the stochastic π-calculus were needed.

Example: The VICE project

- The aim was to construct a minimal cell in silico in order to track the dynamics of a complete metabolome.
- Thus a VIrtual CEll was defined as a stochastic π-calculus model, which seems to behave as a simplified prokaryote.
- Started from a published minimal gene set which eliminated duplicated genes and other redundancies from the smallest known bacterial genomes. This was further reduced to 180 different genes.
- Experimental results were in accordance with those available from in vivo experiments.
- Some extensions to the stochastic π-calculus were needed.

Outline

Process Algebras for Systems Biology

The process algebra of choice for Regev and her co-workers was the stochastic π-calculus.

This work was hugely influential and many people followed their lead in applying the stochastic π-calculus to model intracellular processes.

However the style of modelling in the π-calculus is not always ideal for representing biochemical processes.

Nevertheless a certain "flavour" of the π-calculus still predominately influences many of the process algebras for systems biology.

Process Algebras for Systems Biology

The process algebra of choice for Regev and her co-workers was the stochastic π-calculus.

This work was hugely influential and many people followed their lead in applying the stochastic π-calculus to model intracellular processes.

However the style of modelling in the π-calculus is not always ideal for representing biochemical processes.

Nevertheless a certain "flavour" of the π-calculus still predominately influences many of the process algebras for systems biology.

Process Algebras for Systems Biology

The process algebra of choice for Regev and her co-workers was the stochastic π-calculus.

This work was hugely influential and many people followed their lead in applying the stochastic π-calculus to model intracellular processes.

However the style of modelling in the π-calculus is not always ideal for representing biochemical processes.

Nevertheless a certain "flavour" of the π-calculus still predominately influences many of the process algebras for systems biology.

Process Algebras for Systems Biology

The process algebra of choice for Regev and her co-workers was the stochastic π-calculus.

This work was hugely influential and many people followed their lead in applying the stochastic π-calculus to model intracellular processes.

However the style of modelling in the π-calculus is not always ideal for representing biochemical processes.

Nevertheless a certain "flavour" of the π-calculus still predominately influences many of the process algebras for systems biology.

Issues with the π-calculus

- The restriction to always consider actions as occurring in conjugate pairs does not always match well with biochemical reactions where you might want more than two molecules to be involved.

- The molecule-as-processes abstraction can lead to problems pragmatically. By focussing on the individual molecules the calculus forces the modeller into an individuals-based interpretation of the model. This means that simulation is often the only feasible interpretation.

Issues with the π-calculus

- The restriction to always consider actions as occurring in conjugate pairs does not always match well with biochemical reactions where you might want more than two molecules to be involved.

- The molecule-as-processes abstraction can lead to problems pragmatically. By focussing on the individual molecules the calculus forces the modeller into an individuals-based interpretation of the model. This means that simulation is often the only feasible interpretation.

Bio-PEPA: motivations

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

We also wanted to be able to capture more of the biological features expressed in the models such as those found in the BioModels database.

Bio-PEPA: motivations

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

We also wanted to be able to capture more of the biological features expressed in the models such as those found in the BioModels database.

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.

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.

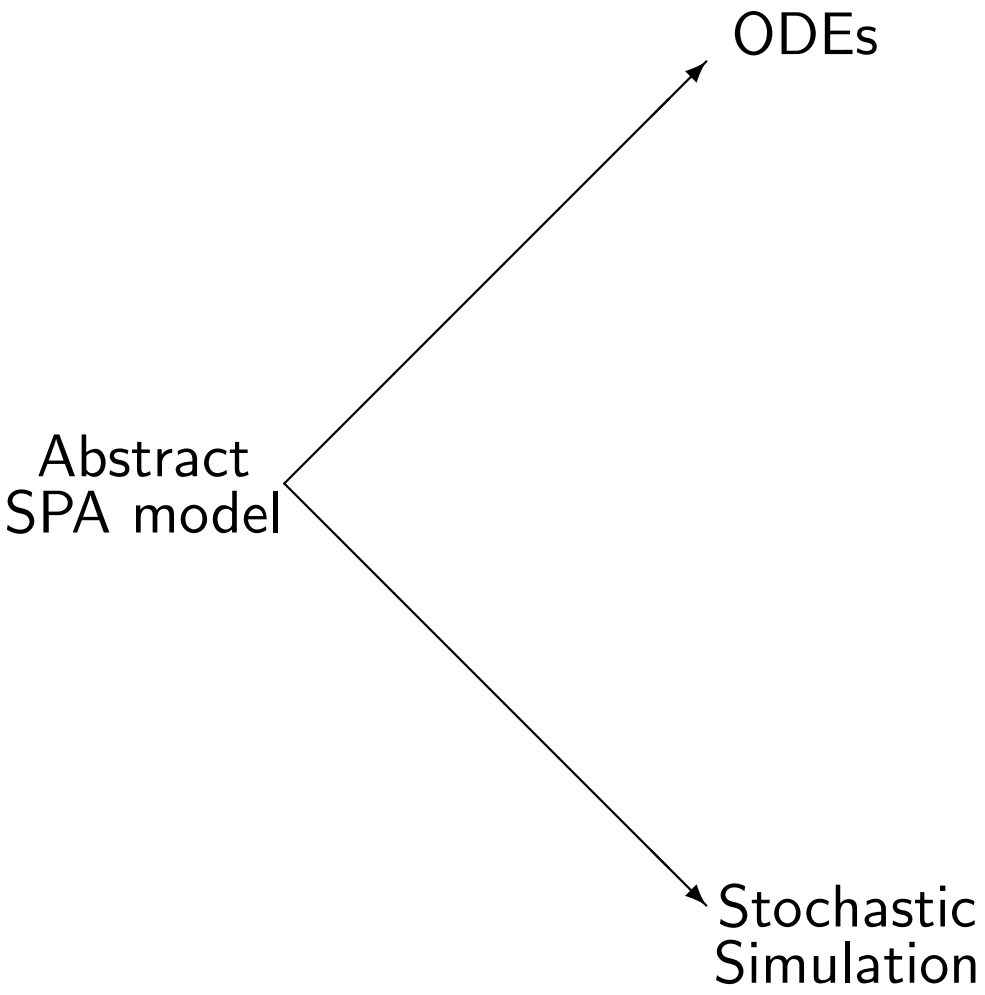
Motivations for Abstraction

Our motivations for seeking more abstraction:

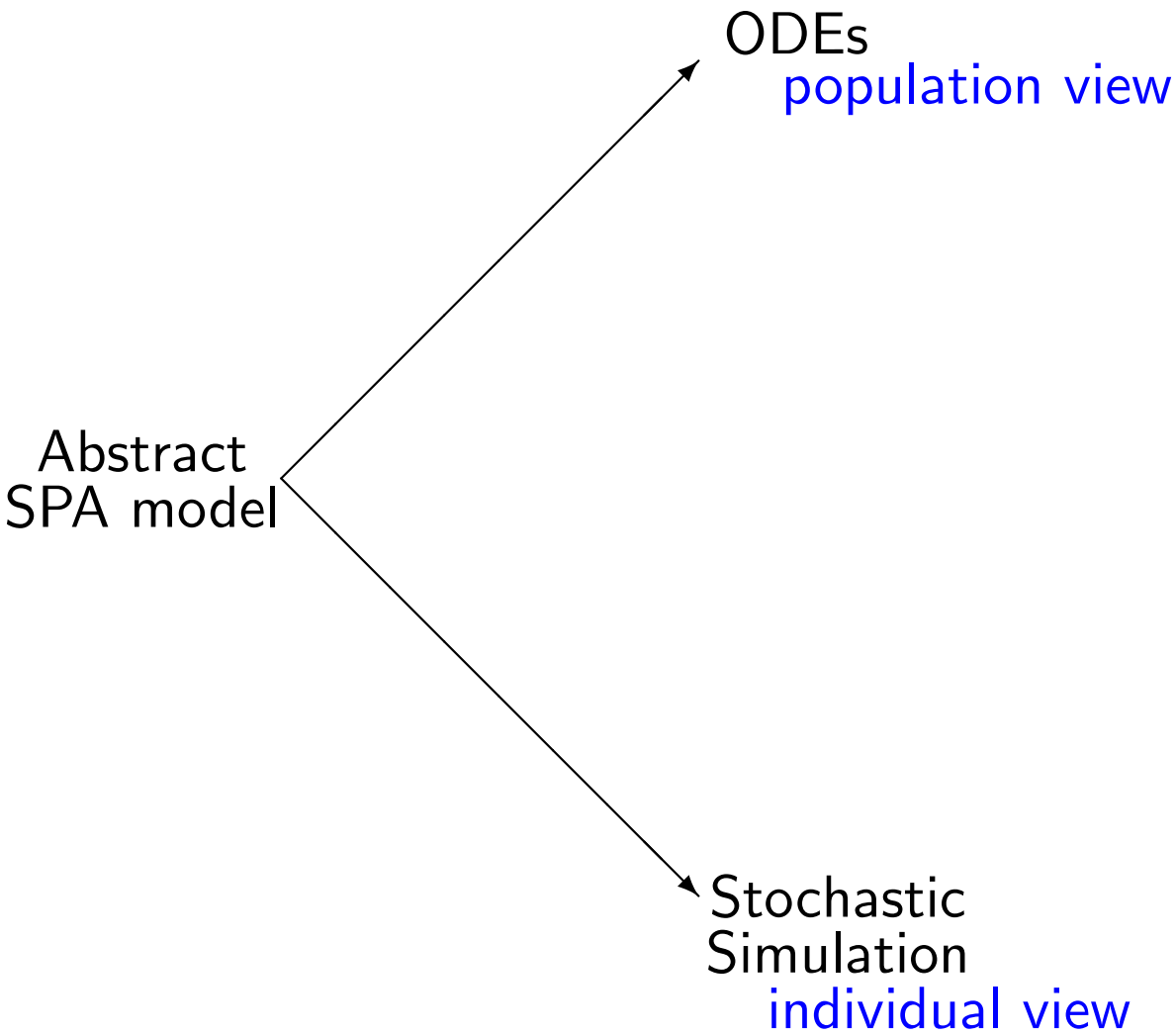
- Process algebra-based analyses such as comparing models (e.g. for equivalence or simulation) and model checking are only possible if the state space is not prohibitively large.
- The data that we have available to parameterise models is sometimes speculative rather than precise.

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

Alternative Representations



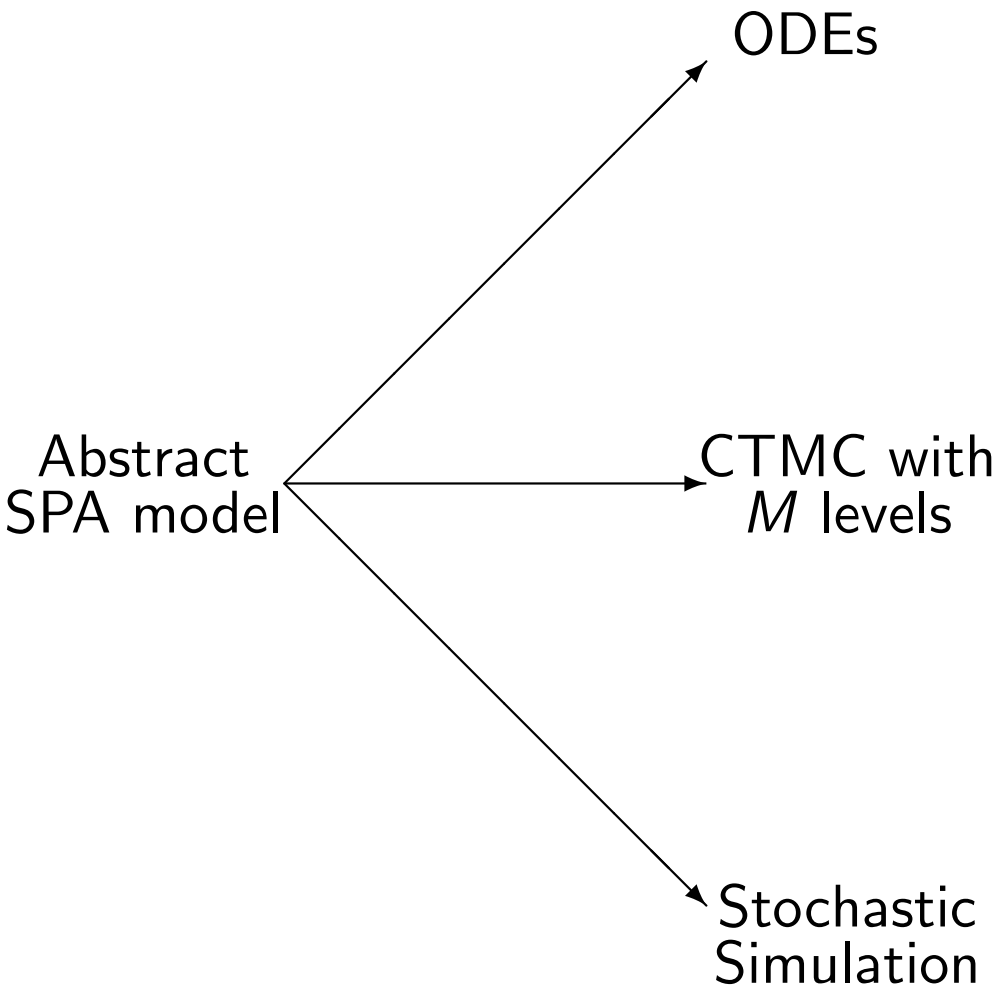
Alternative Representations



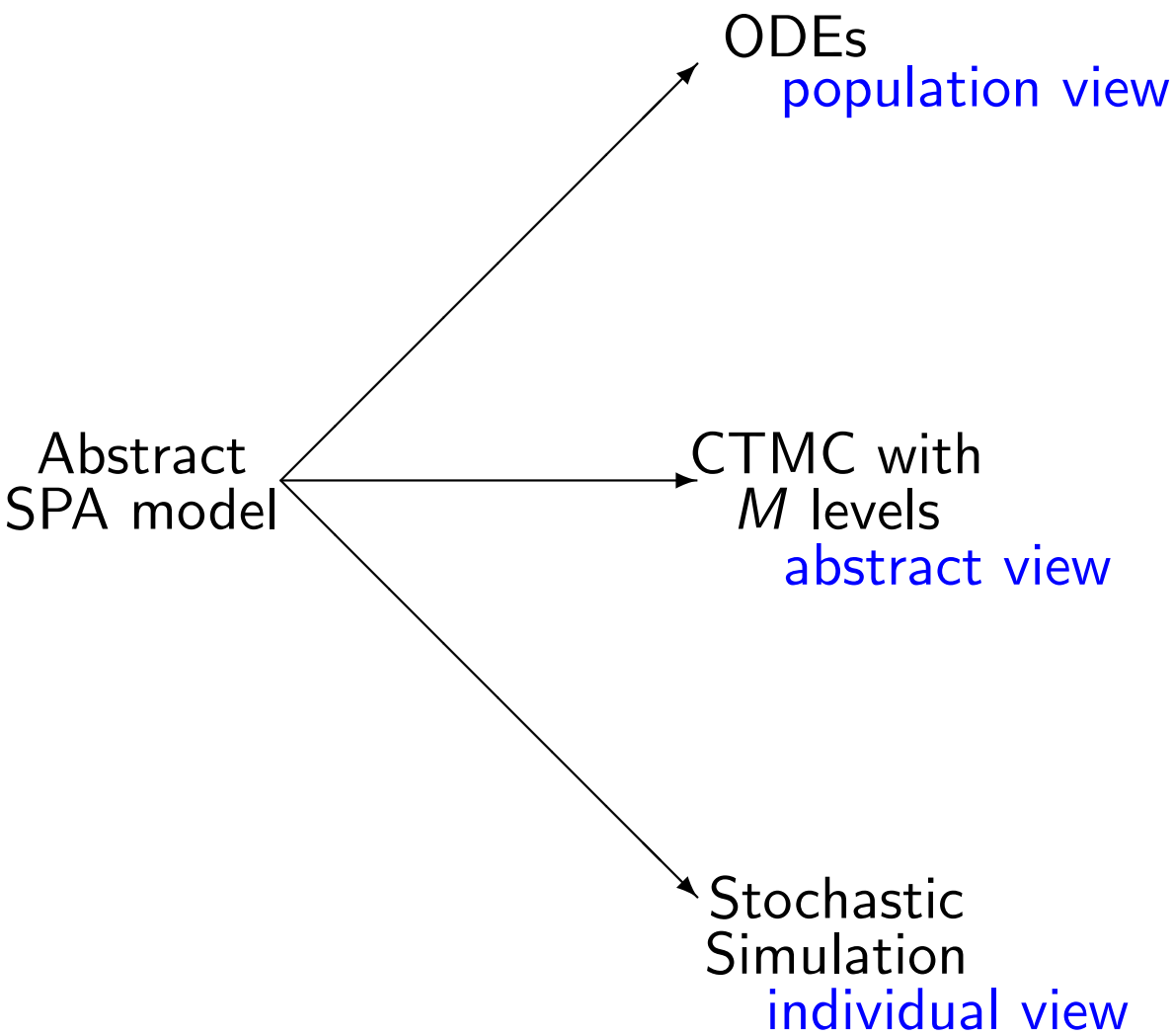
Discretising the population view

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



Alternative models

- The ODE model can be regarded as an approximation of a CTMC in which the number of molecules is large enough that the randomness averages out and the system is essentially deterministic.
- The full molecular view can be used for detailed study of the stochastic behaviour, usually via stochastic simulation
- In models with levels, each level of granularity gives rise to a CTMC, and the behaviour of this sequence of Markov processes converges to the behaviour of the system of ODEs.
- Some analyses which can be carried out via numerical solution of the CTMC are not readily available from ODEs or stochastic simulation.

Alternative models

- The ODE model can be regarded as an approximation of a CTMC in which the number of molecules is large enough that the randomness averages out and the system is essentially deterministic.
- The full molecular view can be used for detailed study of the stochastic behaviour, usually via stochastic simulation
- In models with levels, each level of granularity gives rise to a CTMC, and the behaviour of this sequence of Markov processes converges to the behaviour of the system of ODEs.
- Some analyses which can be carried out via numerical solution of the CTMC are not readily available from ODEs or stochastic simulation.

Alternative models

- The ODE model can be regarded as an approximation of a CTMC in which the number of molecules is large enough that the randomness averages out and the system is essentially deterministic.
- The full molecular view can be used for detailed study of the stochastic behaviour, usually via stochastic simulation
- In models with levels, each level of granularity gives rise to a CTMC, and the behaviour of this sequence of Markov processes converges to the behaviour of the system of ODEs.
- Some analyses which can be carried out via numerical solution of the CTMC are not readily available from ODEs or stochastic simulation.

Alternative models

- The ODE model can be regarded as an approximation of a CTMC in which the number of molecules is large enough that the randomness averages out and the system is essentially deterministic.
- The full molecular view can be used for detailed study of the stochastic behaviour, usually via stochastic simulation
- In models with levels, each level of granularity gives rise to a CTMC, and the behaviour of this sequence of Markov processes converges to the behaviour of the system of ODEs.
- Some analyses which can be carried out via numerical solution of the CTMC are not readily available from ODEs or stochastic simulation.

Modelling biological features

There are some features of biochemical reaction systems which are not readily captured by π-calculus-based stochastic process algebras.

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.

Modelling biological features

There are some features of biochemical reaction systems which are not readily captured by π-calculus-based stochastic process algebras.

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.

Modelling biological features

There are some features of biochemical reaction systems which are not readily captured by π-calculus-based stochastic process algebras.

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.

Modelling biological features

There are some features of biochemical reaction systems which are not readily captured by π-calculus-based stochastic process algebras.

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.

Modelling biological features

There are some features of biochemical reaction systems which are not readily captured by π-calculus-based stochastic process algebras.

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.

Illustration

Consider a conversion of a substrate S, with stoichiometry 2, to a product P, under the influence of an enzyme E, i.e.

$$2 \times S \xrightarrow{E} P$$

In the stochastic π-calculus (for example) this must modelled as a sequence of unary and binary reactions:

- $S + S \longrightarrow 2S$
- $2S + E \longrightarrow 2S : E$
- $2S : E \longrightarrow P : E$
- $P : E \longrightarrow P + E$

Illustration

Consider a conversion of a substrate S, with stoichiometry 2, to a product P, under the influence of an enzyme E, i.e.

$$2 \times S \xrightarrow{E} P$$

In the stochastic π-calculus (for example) this must modelled as a sequence of unary and binary reactions:

- $S + S \longrightarrow 2S$
- $2S + E \longrightarrow 2S : E$
- $2S : E \longrightarrow P : E$
- $P : E \longrightarrow P + E$

Illustration cont.

The problems with this are various:

- Rates must be found for all the intermediate steps.
- Alternate intermediate states are possible and it may not be known which is the appropriate one.
- The number of "states" of the system is significantly increased which has implications for computational efficiency/tractability.

The use of multiway synchronisation, and the reagent-centric style of modelling, avoids these problems

Illustration cont.

The problems with this are various:

- Rates must be found for all the intermediate steps.
- Alternate intermediate states are possible and it may not be known which is the appropriate one.
- The number of "states" of the system is significantly increased which has implications for computational efficiency/tractability.

The use of multiway synchronisation, and the reagent-centric style of modelling, avoids these problems

Illustration cont.

The problems with this are various:

- Rates must be found for all the intermediate steps.
- Alternate intermediate states are possible and it may not be known which is the appropriate one.
- The number of "states" of the system is significantly increased which has implications for computational efficiency/tractability.

The use of multiway synchronisation, and the reagent-centric style of modelling, avoids these problems

Illustration cont.

The problems with this are various:

- Rates must be found for all the intermediate steps.
- Alternate intermediate states are possible and it may not be known which is the appropriate one.
- The number of "states" of the system is significantly increased which has implications for computational efficiency/tractability.

The use of multiway synchronisation, and the reagent-centric style of modelling, avoids these problems

Illustration cont.

The problems with this are various:

- Rates must be found for all the intermediate steps.
- Alternate intermediate states are possible and it may not be known which is the appropriate one.
- The number of "states" of the system is significantly increased which has implications for computational efficiency/tractability.

The use of multiway synchronisation, and the reagent-centric style of modelling, avoids these problems

Illustration cont.

The reaction $2S \xrightarrow{E} P$ represents the enzymatic reaction from the substrate S, with stoichiometry 2, to the product P with enzyme E.

In Bio-PEPA this is described as:

$$\begin{aligned} S &\stackrel{def}{=} (\alpha, 2)\downarrow S \\ E &\stackrel{def}{=} (\alpha, 1) \oplus E \\ P &\stackrel{def}{=} (\alpha, 1)\uparrow P \end{aligned}$$

$$(S(l_{S0}) \underset{\{\alpha\}}{\bowtie} E(l_{E0})) \underset{\{\alpha\}}{\bowtie} P(l_{P0})$$

The dynamics is described by the law $f_\alpha = \frac{v \times E \times S^2}{(K + S^2)}$.

Illustration cont.

The reaction $2S \xrightarrow{E} P$ represents the enzymatic reaction from the substrate S, with stoichiometry 2, to the product P with enzyme E.

In Bio-PEPA this is described as:

$$\begin{aligned} S &\stackrel{def}{=} (\alpha, 2)\downarrow S \\ E &\stackrel{def}{=} (\alpha, 1) \oplus E \\ P &\stackrel{def}{=} (\alpha, 1)\uparrow P \end{aligned}$$

$$(S(l_{S0}) \underset{\{\alpha\}}{\bowtie} E(l_{E0})) \underset{\{\alpha\}}{\bowtie} P(l_{P0})$$

The dynamics is described by the law $f_\alpha = \frac{v \times E \times S^2}{(K + S^2)}$.

Illustration cont.

The reaction $2S \xrightarrow{E} P$ represents the enzymatic reaction from the substrate S, with stoichiometry 2, to the product P with enzyme E.

In Bio-PEPA this is described as:

$$\begin{aligned} S &\stackrel{def}{=} (\alpha, 2)\downarrow S \\ E &\stackrel{def}{=} (\alpha, 1) \oplus E \\ P &\stackrel{def}{=} (\alpha, 1)\uparrow P \end{aligned}$$

$$(S(l_{S0}) \underset{\{\alpha\}}{\bowtie} E(l_{E0})) \underset{\{\alpha\}}{\bowtie} P(l_{P0})$$

The dynamics is described by the law $f_\alpha = \frac{v \times E \times S^2}{(K + S^2)}$.

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.

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.

Bio-PEPA

In Bio-PEPA:

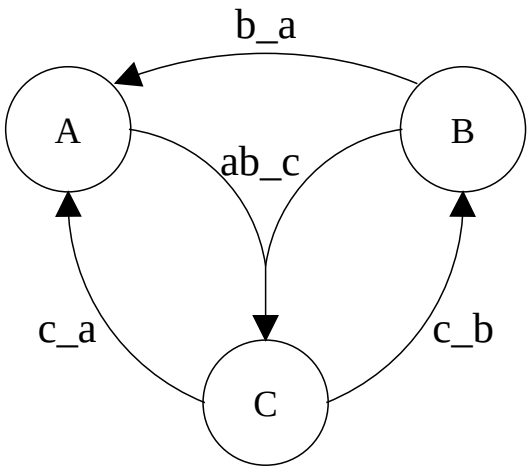
- Unique rates are associated with each reaction (action) type, separately from the specification of the logical behaviour. These rates may be specified by functions.
- The representation of an action within a component (species) records the stoichiometry of that entity with respect to that reaction.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

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.

Bio-PEPA reagent-centric example



$$
\begin{aligned}
A &\stackrel{def}{=} (ab\_c, 1)\downarrow A + (b\_a, 1)\uparrow A \\
&\quad + (c\_a, 1)\uparrow A \\
B &\stackrel{def}{=} (ab\_c, 1)\downarrow B + (b\_a, 1)\downarrow B \\
&\quad + (c\_b, 1)\uparrow B \\
C &\stackrel{def}{=} (c\_a, 1)\downarrow C + (c\_b, 1)\downarrow C \\
&\quad + (ab\_c, 1)\uparrow C
\end{aligned}
$$

$$
\left(A(l_{A0}) \underset{\{ab\_c, b\_a\}}{\bowtie} B(l_{B0})\right) \underset{\{ab\_c, c\_a, c\_b\}}{\bowtie} C(l_{C0})
$$

Reagent-centric view

Role	*Impact on reaction rate*	*Impact on reagent*
Reactant	positive impact, e.g. proportional to current concentration	decreases level
Product	no impact, except at saturation	increases level
Enzyme	positive impact, e.g. proportional to current concentration	level unchanged
Inhibitor	negative impact, e.g. inversely proportional to current concentration	level unchanged

Outline

The syntax

Sequential component (species component)

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

Model component

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

The parameter I 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 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

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

The parameter I 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 syntax

Sequential component (species component)

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

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 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

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

The parameter I 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 syntax

Sequential component (species component)

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

Model component

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

The parameter I 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 syntax

Sequential component (species component)

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

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 syntax

Sequential component (species component)

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

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 syntax

Sequential component (species component)

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

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 syntax

Sequential component (species component)

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

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 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.

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.

Outline

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

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

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

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

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

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

Semantics: prefix rules

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

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

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

Semantics: prefix rules

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

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

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

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

Semantics: constant and choice rules

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

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

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

Semantics: constant and choice rules

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

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

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

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

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

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

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

Semantics: cooperation rules

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

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

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

Semantics: cooperation rules

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

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

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

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 $r \in \mathbb{R}^+$.

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 $f_{\alpha_j}(\mathcal{V}, \mathcal{N}, \mathcal{K})/h$.

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 $f_{\alpha_j}(\mathcal{V}, \mathcal{N}, \mathcal{K})/h$.

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 $f_{\alpha_j}(\mathcal{V}, \mathcal{N}, \mathcal{K})/h$.

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 $f_{\alpha_j}(\mathcal{V}, \mathcal{N}, \mathcal{K})/h$.

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 $f_{\alpha_j}(\mathcal{V}, \mathcal{N}, \mathcal{K})/h$.

Example: Michaelis-Menten

The reaction $S \xrightarrow{E} P$ represents the enzymatic reaction from the substrate S to the product P with enzyme E.

The dynamics is described by the law $\frac{v \times E \times S}{(K+S)}$.

$$
\begin{array}{lcl}
S & \stackrel{def}{=} & (\alpha, 1)\downarrow S \\
E & \stackrel{def}{=} & (\alpha, 1) \oplus E \\
P & \stackrel{def}{=} & (\alpha, 1)\uparrow P
\end{array}
$$

$$
(S(l_{S0}) \underset{\{\alpha\}}{\bowtie} E(l_{E0})) \underset{\{\alpha\}}{\bowtie} P(l_{P0})
$$

Example: Michaelis-Menten

The reaction $S \xrightarrow{E} P$ represents the enzymatic reaction from the substrate S to the product P with enzyme E.

The dynamics is described by the law $\frac{v \times E \times S}{(K+S)}$.

$$
\begin{aligned}
S &\stackrel{def}{=} (\alpha, 1)\downarrow S \\
E &\stackrel{def}{=} (\alpha, 1) \oplus E \\
P &\stackrel{def}{=} (\alpha, 1)\uparrow P
\end{aligned}
$$

$$(S(l_{S0}) \underset{\{\alpha\}}{\bowtie} E(l_{E0})) \underset{\{\alpha\}}{\bowtie} P(l_{P0})$$

Example: Michaelis-Menten

The reaction $S \xrightarrow{E} P$ represents the enzymatic reaction from the substrate S to the product P with enzyme E.

The dynamics is described by the law $\frac{v \times E \times S}{(K+S)}$.

$$
\begin{array}{lcl}
S & \stackrel{def}{=} & (\alpha, 1)\downarrow S \\
E & \stackrel{def}{=} & (\alpha, 1) \oplus E \\
P & \stackrel{def}{=} & (\alpha, 1)\uparrow P
\end{array}
$$

$$
(S(l_{S0}) \underset{\{\alpha\}}{\bowtie} E(l_{E0})) \underset{\{\alpha\}}{\bowtie} P(l_{P0})
$$

Outline

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 they can be used in conjunction.

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 they can be used in conjunction.

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 they can be used in conjunction.

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 they can be used in conjunction.

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 they can be used in conjunction.

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 they can be used in conjunction.

The abstraction

When modelling in Bio-PEPA we use the following abstraction:

- Each species i is described by a Bio-PEPA component C_i.
- Each reaction j is associated with an action type α_j and its dynamics is described by a specific function f_{α_j}.

Given a reaction j, all the species/components cooperate on the action type α_j and consequently, reactants decrease their levels, while products increase them.

The species components are then composed together to describe the behaviour of the system.

The abstraction

When modelling in Bio-PEPA we use the following abstraction:

- Each species i is described by a Bio-PEPA component C_i.
- Each reaction j is associated with an action type α_j and its dynamics is described by a specific function f_{α_j}.

  Given a reaction j, all the species/components cooperate on the action type α_j and consequently, reactants decrease their levels, while products increase them.

The species components are then composed together to describe the behaviour of the system.

The abstraction

When modelling in Bio-PEPA we use the following abstraction:

- Each species i is described by a Bio-PEPA component C_i.
- Each reaction j is associated with an action type α_j and its dynamics is described by a specific function f_{α_j}.

  Given a reaction j, all the species/components cooperate on the action type α_j and consequently, reactants decrease their levels, while products increase them.

The species components are then composed together to describe the behaviour of the system.

The abstraction

When modelling in Bio-PEPA we use the following abstraction:

- Each species i is described by a Bio-PEPA component C_i.
- Each reaction j is associated with an action type α_j and its dynamics is described by a specific function f_{α_j}.

  Given a reaction j, all the species/components cooperate on the action type α_j and consequently, reactants decrease their levels, while products increase them.

The species components are then composed together to describe the behaviour of the system.

State representation

- The state of the system at any time consists of the local states of each of its sequential/species components.
- 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.
- A component varying its state corresponds to it varying its quantity or level.
- This is captured by an integer parameter associated with the species and the effect of a reaction is to vary that parameter by a number of levels corresponding to the stoichiometry of this species in the reaction as we saw in the formal semantics.

State representation

- The state of the system at any time consists of the local states of each of its sequential/species components.
- 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.
- A component varying its state corresponds to it varying its quantity or level.
- This is captured by an integer parameter associated with the species and the effect of a reaction is to vary that parameter by a number of levels corresponding to the stoichiometry of this species in the reaction as we saw in the formal semantics.

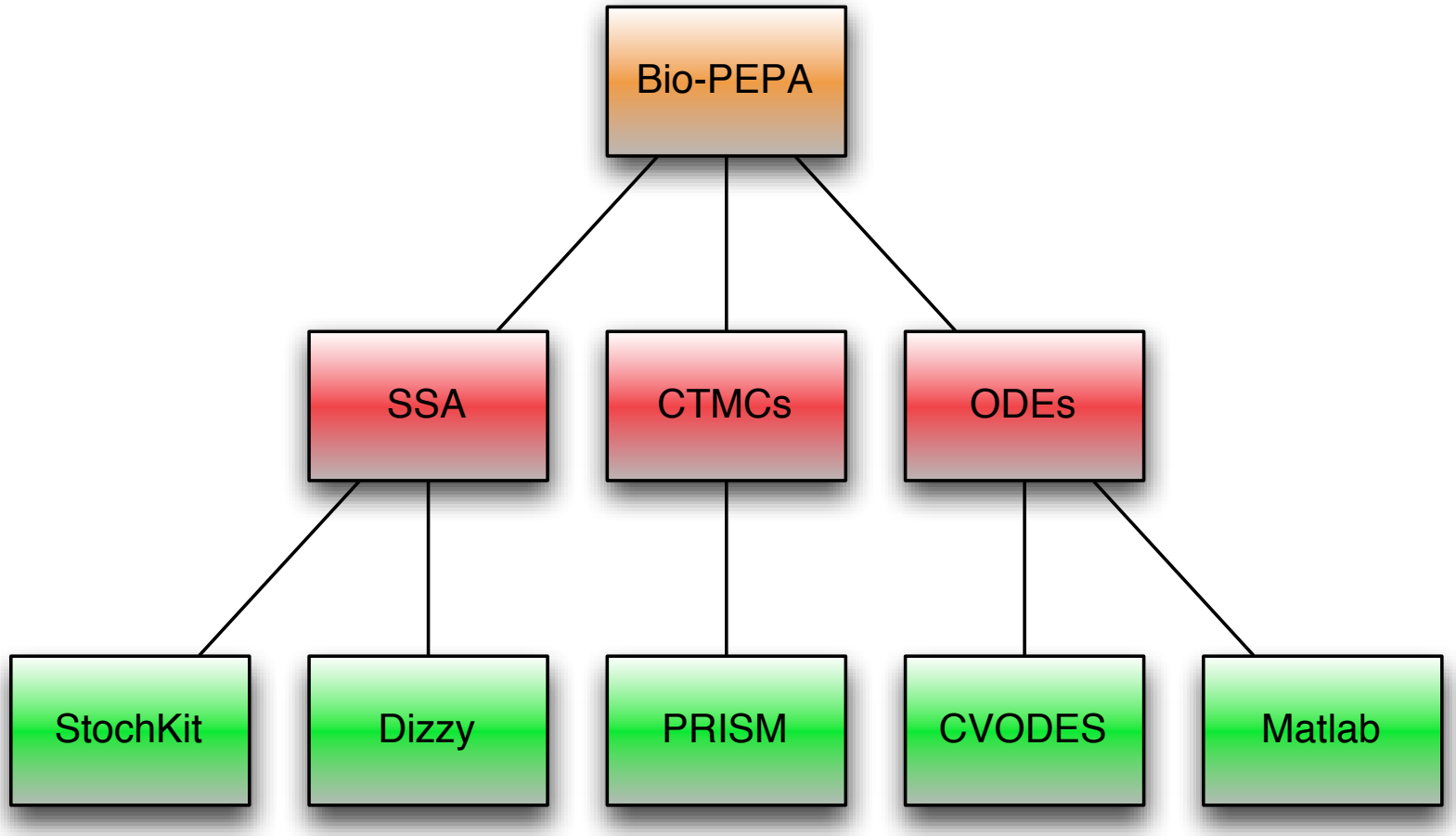
State representation

- The state of the system at any time consists of the local states of each of its sequential/species components.
- 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.
- A component varying its state corresponds to it varying its quantity or level.
- This is captured by an integer parameter associated with the species and the effect of a reaction is to vary that parameter by a number of levels corresponding to the stoichiometry of this species in the reaction as we saw in the formal semantics.

State representation

- The state of the system at any time consists of the local states of each of its sequential/species components.
- 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.
- A component varying its state corresponds to it varying its quantity or level.
- This is captured by an integer parameter associated with the species and the effect of a reaction is to vary that parameter by a number of levels corresponding to the stoichiometry of this species in the reaction as we saw in the formal semantics.

Analysis with Bio-PEPA



Outline

CTMC

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

In the biological context transient analysis is appropriate much more frequently than steady state analysis.

If molecule counts are kept in states then either analysis rapidly becomes infeasible.

Bio-PEPA models can also be simulated using the Gillespie Stochastic Simulation Algorithm for CTMCs that we considered for PEPA.

CTMC

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

In the biological context transient analysis is appropriate much more frequently than steady state analysis.

If molecule counts are kept in states then either analysis rapidly becomes infeasible.

Bio-PEPA models can also be simulated using the Gillespie Stochastic Simulation Algorithm for CTMCs that we considered for PEPA.

CTMC

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

In the biological context transient analysis is appropriate much more frequently than steady state analysis.

If molecule counts are kept in states then either analysis rapidly becomes infeasible.

Bio-PEPA models can also be simulated using the Gillespie Stochastic Simulation Algorithm for CTMCs that we considered for PEPA.

CTMC

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

In the biological context transient analysis is appropriate much more frequently than steady state analysis.

If molecule counts are kept in states then either analysis rapidly becomes infeasible.

Bio-PEPA models can also be simulated using the Gillespie Stochastic Simulation Algorithm for CTMCs that we considered for PEPA.

CTMC

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

In the biological context transient analysis is appropriate much more frequently than steady state analysis.

If molecule counts are kept in states then either analysis rapidly becomes infeasible.

Bio-PEPA models can also be simulated using the Gillespie Stochastic Simulation Algorithm for CTMCs that we considered for PEPA.

CTMC

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

In the biological context transient analysis is appropriate much more frequently than steady state analysis.

If molecule counts are kept in states then either analysis rapidly becomes infeasible.

Bio-PEPA models can also be simulated using the Gillespie Stochastic Simulation Algorithm for CTMCs that we considered for PEPA.

CTMC with levels

- Models are based on discrete levels of concentration within a species.
- The granularity of the system is defined in terms of the step size h of the concentration intervals.
- We define the same step size h for all the species. This follows from the law of conservation of mass.
- If l_i is the concentration level for the species i, the concentration is taken to be $x_i = l_i \times h$.

CTMC with levels

- Models are based on discrete levels of concentration within a species.
- The granularity of the system is defined in terms of the step size h of the concentration intervals.
- We define the same step size h for all the species. This follows from the law of conservation of mass.
- If l_i is the concentration level for the species i, the concentration is taken to be $x_i = l_i \times h$.

CTMC with levels

- Models are based on discrete levels of concentration within a species.
- The granularity of the system is defined in terms of the step size h of the concentration intervals.
- We define the same step size h for all the species. This follows from the law of conservation of mass.
- If l_i is the concentration level for the species i, the concentration is taken to be $x_i = l_i \times h$.

CTMC with levels

- Models are based on discrete levels of concentration within a species.
- The granularity of the system is defined in terms of the step size h of the concentration intervals.
- We define the same step size h for all the species. This follows from the law of conservation of mass.
- If l_i is the concentration level for the species i, the concentration is taken to be $x_i = l_i \times h$.

CTMC with levels

- The rate of a transition is set to be consistent with the granularity.
- The granularity must be specified by the modeller as the expected range of concentration values and the number of levels considered.
- 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].

CTMC with levels

- The rate of a transition is set to be consistent with the granularity.
- The granularity must be specified by the modeller as the expected range of concentration values and the number of levels considered.
- 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].

CTMC with levels

- The rate of a transition is set to be consistent with the granularity.
- The granularity must be specified by the modeller as the expected range of concentration values and the number of levels considered.
- 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].

CTMC with levels

- The rate of a transition is set to be consistent with the granularity.
- The granularity must be specified by the modeller as the expected range of concentration values and the number of levels considered.
- 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].

ODE system

The derivation of the ODEs from the Bio-PEPA is straightforward, based on the definitions of the species components.

1. definition of the $(N \times M)$ stoichiometry matrix D, where N is the number of species and M is the number of reactions;
2. definition of the kinetic law vector v_{KL} containing the kinetic law of each reaction;
3. association of the variable x_i with each component C_i and definition of the vector $\overline{x}$.

The ODE system is then obtained as

$$\frac{d\overline{x}}{dt} = D \times v_{KL}$$

ODE system

The derivation of the ODEs from the Bio-PEPA is straightforward, based on the definitions of the species components.

1. definition of the $(N \times M)$ stoichiometry matrix D, where N is the number of species and M is the number of reactions;
2. definition of the kinetic law vector v_{KL} containing the kinetic law of each reaction;
3. association of the variable x_i with each component C_i and definition of the vector $\overline{x}$.

The ODE system is then obtained as

$$\frac{d\overline{x}}{dt} = D \times v_{KL}$$

ODE system

The derivation of the ODEs from the Bio-PEPA is straightforward, based on the definitions of the species components.

1. definition of the $(N \times M)$ stoichiometry matrix D, where N is the number of species and M is the number of reactions;
2. definition of the kinetic law vector v_{KL} containing the kinetic law of each reaction;
3. association of the variable x_i with each component C_i and definition of the vector $\overline{x}$.

The ODE system is then obtained as

$$\frac{d\overline{x}}{dt} = D \times v_{KL}$$

ODE system

The derivation of the ODEs from the Bio-PEPA is straightforward, based on the definitions of the species components.

1. definition of the $(N \times M)$ stoichiometry matrix D, where N is the number of species and M is the number of reactions;
2. definition of the kinetic law vector v_{KL} containing the kinetic law of each reaction;
3. association of the variable x_i with each component C_i and definition of the vector $\overline{x}$.

The ODE system is then obtained as

$$\frac{d\overline{x}}{dt} = D \times v_{KL}$$

ODE system

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.
- The process algebra model allow us to derive properties of the model, such as freedom from deadlock, before numerical analysis is carried out.
- The algebraic formulation of the model emphasises interactions between the biochemical entities.

ODE system

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.
- The process algebra model allow us to derive properties of the model, such as freedom from deadlock, before numerical analysis is carried out.
- The algebraic formulation of the model emphasises interactions between the biochemical entities.

ODE system

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.
- The process algebra model allow us to derive properties of the model, such as freedom from deadlock, before numerical analysis is carried out.
- The algebraic formulation of the model emphasises interactions between the biochemical entities.

ODE system

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.
- The process algebra model allow us to derive properties of the model, such as freedom from deadlock, before numerical analysis is carried out.
- The algebraic formulation of the model emphasises interactions between the biochemical entities.

PRISM model and model checking

- Analysing models of biological processes via stochastic 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

- Analysing models of biological processes via stochastic 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

- Analysing models of biological processes via stochastic 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

- Analysing models of biological processes via stochastic 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

- Stochastic 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.
- In practice, it is more straightforward to directly map to the input language of the tool, the language of 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.

PRISM model and model checking

- Stochastic 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.
- In practice, it is more straightforward to directly map to the input language of the tool, the language of 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.

PRISM model and model checking

- Stochastic 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.
- In practice, it is more straightforward to directly map to the input language of the tool, the language of 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.

PRISM model and model checking

- Stochastic 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.
- In practice, it is more straightforward to directly map to the input language of the tool, the language of 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.

Outline

Goldbeter's model [Goldbeter 91]

- Goldbeter's model describes the activity of the cyclin in the cell cycle.
- The cyclin promotes the activation of a cdk (cdc2) which in turn activates a cyclin protease.
- This protease promotes cyclin degradation.
- This leads to a negative feedback loop.
- In the model most of the kinetic laws are of kind Michaelis-Menten and this can be reflected in the Bio-PEPA model

Goldbeter's model [Goldbeter 91]

- Goldbeter's model describes the activity of the cyclin in the cell cycle.
- The cyclin promotes the activation of a cdk (cdc2) which in turn activates a cyclin protease.
- This protease promotes cyclin degradation.
- This leads to a negative feedback loop.
- In the model most of the kinetic laws are of kind Michaelis-Menten and this can be reflected in the Bio-PEPA model

Goldbeter's model [Goldbeter 91]

- Goldbeter's model describes the activity of the cyclin in the cell cycle.
- The cyclin promotes the activation of a cdk (cdc2) which in turn activates a cyclin protease.
- This protease promotes cyclin degradation.
- This leads to a negative feedback loop.
- In the model most of the kinetic laws are of kind Michaelis-Menten and this can be reflected in the Bio-PEPA model

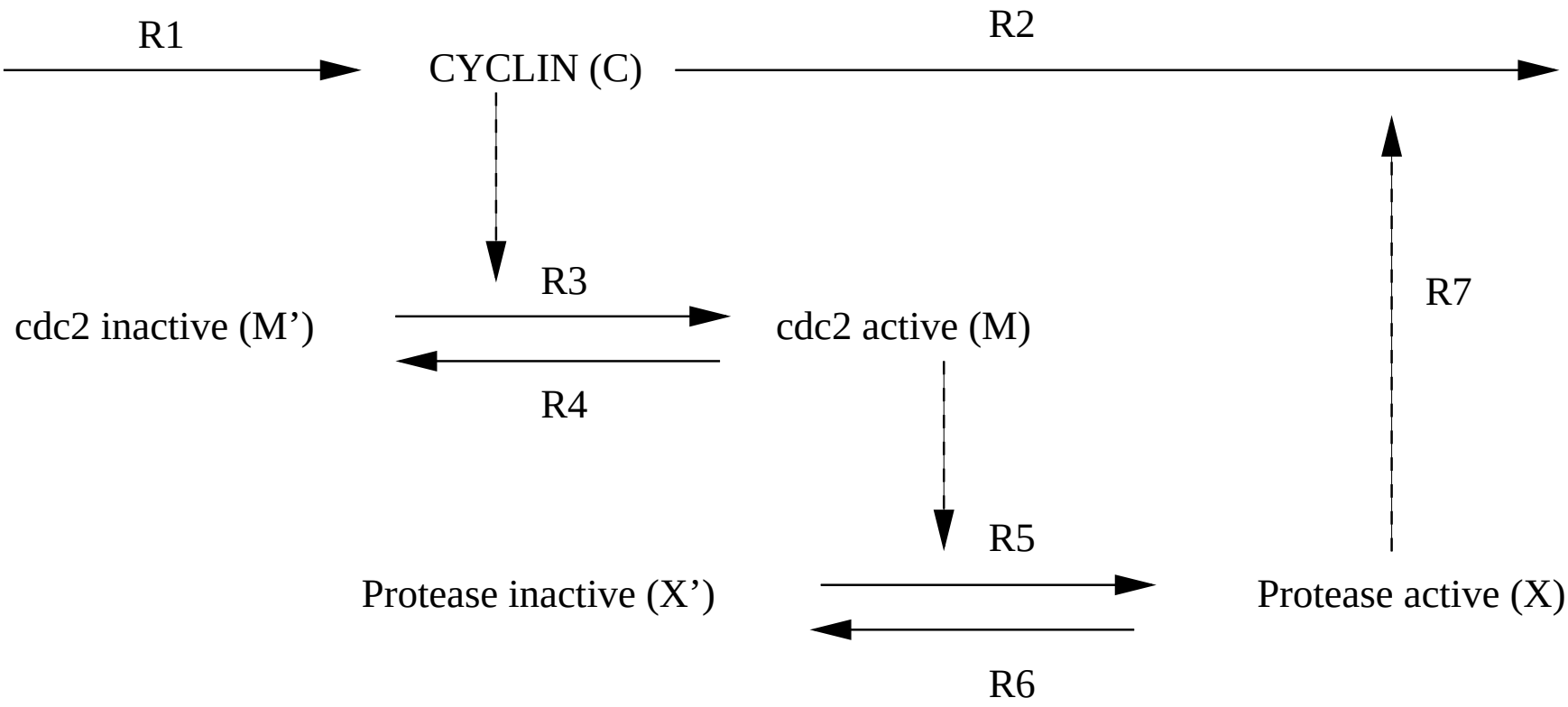
Goldbeter's model [Goldbeter 91]

- Goldbeter's model describes the activity of the cyclin in the cell cycle.
- The cyclin promotes the activation of a cdk (cdc2) which in turn activates a cyclin protease.
- This protease promotes cyclin degradation.
- This leads to a negative feedback loop.
- In the model most of the kinetic laws are of kind Michaelis-Menten and this can be reflected in the Bio-PEPA model

Goldbeter's model [Goldbeter 91]

- Goldbeter's model describes the activity of the cyclin in the cell cycle.
- The cyclin promotes the activation of a cdk (cdc2) which in turn activates a cyclin protease.
- This protease promotes cyclin degradation.
- This leads to a negative feedback loop.
- In the model most of the kinetic laws are of kind Michaelis-Menten and this can be reflected in the Bio-PEPA model

The biological model



The biological model (2)

There are three different biological species involved:

- cyclin, the protein protagonist of the cycle, represented as C;
- cdc2 kinase, in both active and inactive form. The variables used to represent them are M and M', respectively;
- cyclin protease, in both active and inactive form. The variable are X and X'.

The biological model (2)

There are three different biological species involved:

- cyclin, the protein protagonist of the cycle, represented as C;
- cdc2 kinase, in both active and inactive form. The variables used to represent them are M and M', respectively;
- cyclin protease, in both active and inactive form. The variable are X and X'.

The biological model (2)

There are three different biological species involved:

- cyclin, the protein protagonist of the cycle, represented as C;
- cdc2 kinase, in both active and inactive form. The variables used to represent them are M and M', respectively;
- cyclin protease, in both active and inactive form. The variable are X and X'.

The biological model (2)

There are three different biological species involved:

- cyclin, the protein protagonist of the cycle, represented as C;
- cdc2 kinase, in both active and inactive form. The variables used to represent them are M and M', respectively;
- cyclin protease, in both active and inactive form. The variable are X and X'.

Reactions

id	reaction	react.	prod.	mod.	kinetic laws
α_1	creation of cyclin	-	C	-	v_i
α_2	degradation of cyclin	C	-	-	$kd \times C$
α_3	activation of *cdc*2 kinase	M'	M	C	$\frac{C \times V_{M1}}{(K_c + C)} \frac{M'}{(K_1 + M')}$
α_4	deactivation of *cdc*2 kinase	M	M'	-	$\frac{M \times V_2}{(K_2 + M)}$
α_5	activation of cyclin protease	X'	X	M	$\frac{X' \times M \times V_{M3}}{(K_3 + X')}$
α_6	deactivation of cyclin protease	X	X'	-	$\frac{X \times V_4}{K_4 + X}$
α_7	X triggered degradation of cyclin	C	-	X	$\frac{C \times v_d \times X}{C + K_d}$

α_1, α_2 have mass-action kinetics; others are Michaelis-Menten.

Translation into Bio-PEPA

Definition of the set $\mathcal{N}$:

$$\mathcal{N} = [C : h_C, N_c;\ M : h_M, N_M;\ M' : h_{M'}, N_{M'};\ X : h_X, N_X, ;\ X' : h_{X'}, N_{X'}]$$

Definition of functional rates ($\mathcal{F}$):

$$f_{\alpha_1} = fMA(v_i); \qquad f_{\alpha_2} = fMA(k_d);$$
$$f_{\alpha_4} = fMM(V_2, K_2); \qquad f_{\alpha_5} = fMM(V_3, K_3);$$
$$f_{\alpha_6} = fMM(V_4, K_4); \qquad f_{\alpha_7} = fMM(V_d, K_d);$$

$$f_{\alpha_3} = \frac{v_1 \times C}{K_c + C} \frac{M'}{K1 + M'}$$

Translation into Bio-PEPA

Definition of the set $\mathcal{N}$:

$$\mathcal{N} = [C : h_C, N_c;\ M : h_M, N_M;\ M' : h_{M'}, N_{M'};\ X : h_X, N_X, ;\ X' : h_{X'}, N_{X'}]$$

Definition of functional rates ($\mathcal{F}$):

$$\begin{array}{lcllcl} f_{\alpha_1} & = & fMA(v_i); & f_{\alpha_2} & = & fMA(k_d); \\ f_{\alpha_4} & = & fMM(V_2, K_2); & f_{\alpha_5} & = & fMM(V_3, K_3); \\ f_{\alpha_6} & = & fMM(V_4, K_4); & f_{\alpha_7} & = & fMM(V_d, K_d); \end{array}$$

$$f_{\alpha_3} = \frac{v_1 \times C}{K_c + C} \frac{M'}{K1 + M'}$$

The Bio-PEPA model

Definition of species components (*Comp*):

$$
\begin{aligned}
C &\stackrel{def}{=} (\alpha_1, 1)\uparrow C + (\alpha_2, 1)\downarrow C + (\alpha_7, 1)\downarrow C + (\alpha_3, 1) \oplus C \\
M' &\stackrel{def}{=} (\alpha_4, 1)\uparrow M' + (\alpha_3, 1)\downarrow M' \\
M &\stackrel{def}{=} (\alpha_3, 1)\uparrow M + (\alpha_4, 1)\downarrow M + (\alpha_5, 1) \oplus M \\
X' &\stackrel{def}{=} (\alpha_6, 1)\uparrow X' + (\alpha_5, 1)\downarrow X' \\
X &\stackrel{def}{=} (\alpha_5, 1)\uparrow X + (\alpha_6, 1)\downarrow X + (\alpha_7, 1) \oplus X
\end{aligned}
$$

Definition of the model component (*P*):

$$
C(l_{0C}) \underset{\{\alpha_3\}}{\bowtie} M(l_{0M}) \underset{\{\alpha_3,\alpha_4\}}{\bowtie} M'(l_{0M'}) \underset{\{\alpha_5,\alpha_7\}}{\bowtie} X(l_{0X}) \underset{\{\alpha_5,\alpha_6\}}{\bowtie} X'(l_{0X'})
$$

The Bio-PEPA model

Definition of species components (*Comp*):

$$
\begin{aligned}
C &\stackrel{def}{=} (\alpha_1, 1)\uparrow C + (\alpha_2, 1)\downarrow C + (\alpha_7, 1)\downarrow C + (\alpha_3, 1) \oplus C \\
M' &\stackrel{def}{=} (\alpha_4, 1)\uparrow M' + (\alpha_3, 1)\downarrow M' \\
M &\stackrel{def}{=} (\alpha_3, 1)\uparrow M + (\alpha_4, 1)\downarrow M + (\alpha_5, 1) \oplus M \\
X' &\stackrel{def}{=} (\alpha_6, 1)\uparrow X' + (\alpha_5, 1)\downarrow X' \\
X &\stackrel{def}{=} (\alpha_5, 1)\uparrow X + (\alpha_6, 1)\downarrow X + (\alpha_7, 1) \oplus X
\end{aligned}
$$

Definition of the model component (*P*):

$$
C(l_{0C}) \underset{\{\alpha_3\}}{\bowtie} M(l_{0M}) \underset{\{\alpha_3,\alpha_4\}}{\bowtie} M'(l_{0M'}) \underset{\{\alpha_5,\alpha_7\}}{\bowtie} X(l_{0X}) \underset{\{\alpha_5,\alpha_6\}}{\bowtie} X'(l_{0X'})
$$

ODEs

The stoichiometry matrix D:

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	
C	$+1$	0	0	0	0	0	-1	x_C
M'	0	0	-1	$+1$	0	0	0	$x_{M'}$
M	0	0	$+1$	-1	0	0	0	x_M
X'	0	0	0	0	-1	$+1$	0	$x_{X'}$
X	0	0	0	0	$+1$	-1	0	x_X

The vector that contains the kinetic laws is:

$$w = \left(v_i \times 1, k_d \times x_C, \frac{V_{M1} \times x_C}{K_c + x_C} \frac{x_{M'}}{(K_1 + x_{M'})}, \frac{V_2 \times x_M}{(K_2 + x_M)}, \right.$$
$$\left. \frac{V_{M3} \times x_M \times x_{X'}}{(K_3 + x_{X'})}, \frac{V_4 \times x_X}{(K_4 + x_X)}, \frac{v_d \times x_C \times x_X}{(K_d + x_C)} \right)$$

ODEs

The stoichiometry matrix D:

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	
C	$+1$	0	0	0	0	0	-1	x_C
M'	0	0	-1	$+1$	0	0	0	$x_{M'}$
M	0	0	$+1$	-1	0	0	0	x_M
X'	0	0	0	0	-1	$+1$	0	$x_{X'}$
X	0	0	0	0	$+1$	-1	0	x_X

The vector that contains the kinetic laws is:

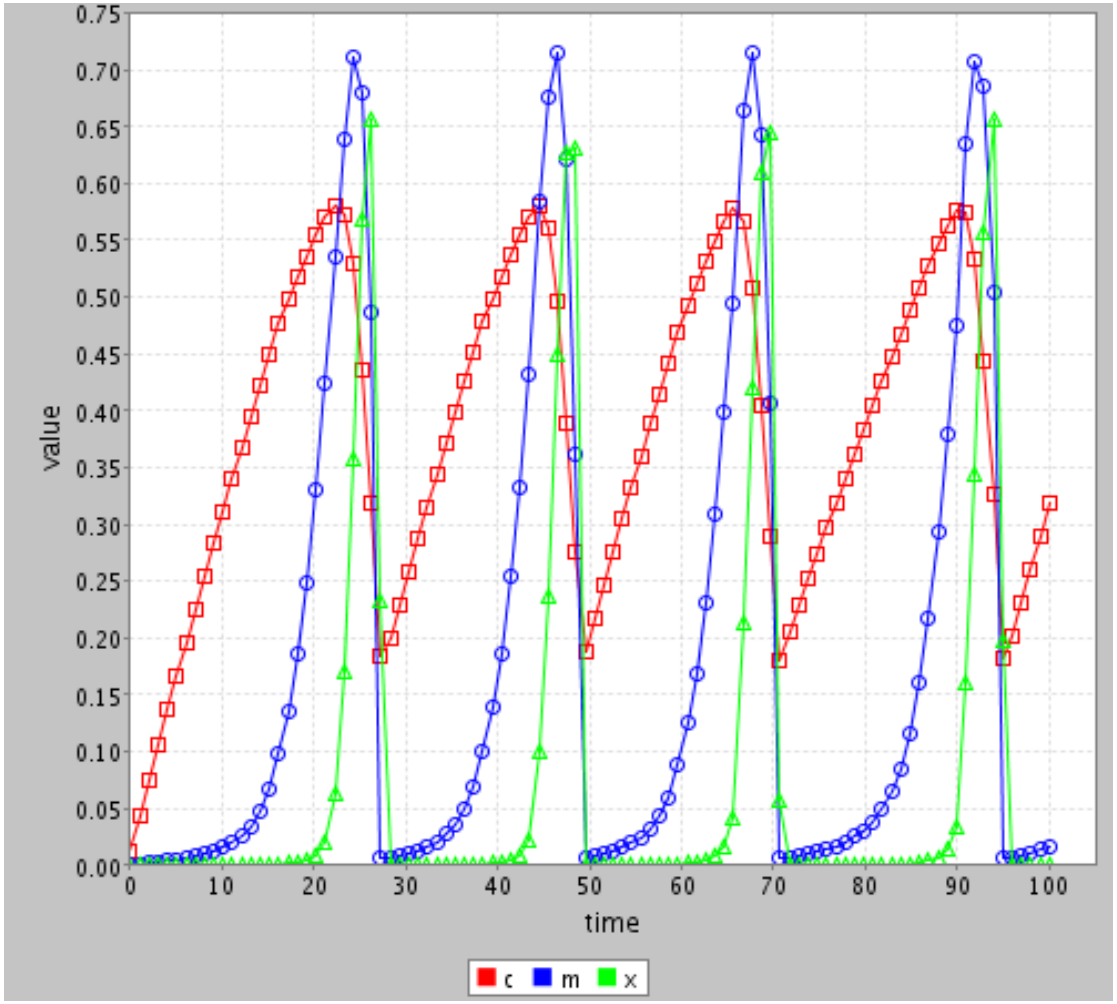
$$w = \left(v_i \times 1, k_d \times x_C, \frac{V_{M1} \times x_C}{K_c + x_C} \frac{x_{M'}}{(K_1 + x_{M'})}, \frac{V_2 \times x_M}{(K_2 + x_M)}, \right.$$
$$\left. \frac{V_{M3} \times x_M \times x_{X'}}{(K_3 + x_{X'})}, \frac{V_4 \times x_X}{(K_4 + x_X)}, \frac{v_d \times x_C \times x_X}{(K_d + x_C)} \right)$$

ODEs (2)

The system of ODEs is obtained as $\frac{d\bar{x}}{dt} = D \times w$, where $\bar{x}^T =: (x_C, x_{M'}, x_M, x_{X'}, x_X)$ is the vector of the species variables:

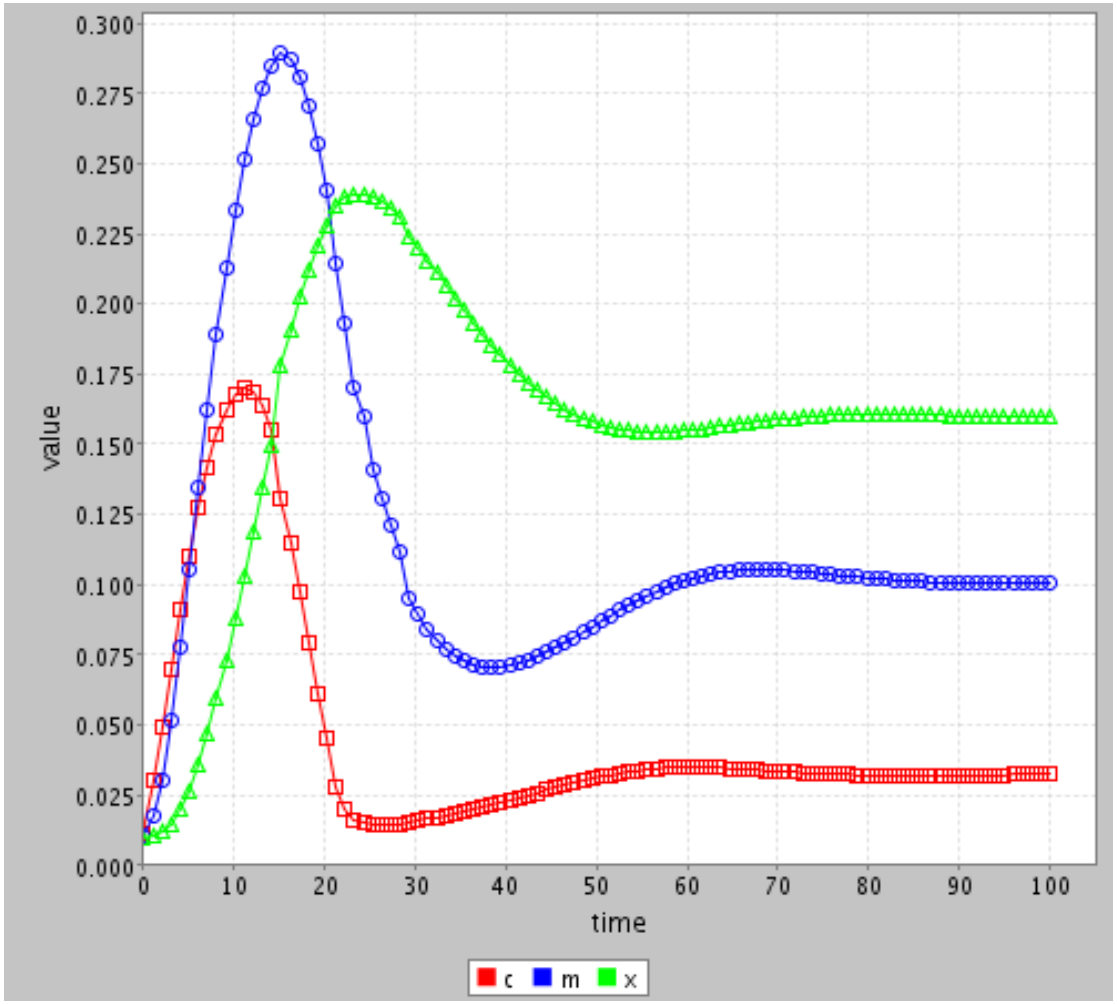
$$
\begin{aligned}
\frac{dx_C}{dt} &= v_i \times 1 - k_d \times x_C - \frac{v_d \times x_C \times x_X}{(K_d + x_C)} \\
\frac{dx_{M'}}{dt} &= -\frac{V_{M1} \times x_C}{K_c + x_C} \frac{x_{M'}}{(K_1 + x_{M'})} + \frac{V_2 \times x_M}{(K_2 + x_M)} \\
\frac{dx_M}{dt} &= +\frac{V_{M1} \times x_C}{K_c + x_C} \frac{x_{M'}}{(K_1 + x_{M'})} - \frac{V_2 \times x_M}{(K_2 + x_M)} \\
\frac{dx_{X'}}{dt} &= -\frac{V_{M3} \times x_M \times x_{X'}}{(K_3 + x_{X'})} + \frac{V_4 \times x_X}{(K_4 + x_X)} \\
\frac{dx_X}{dt} &= \frac{V_{M3} \times x_M \times x_{X'}}{(K_3 + x_{X'})} - \frac{V_4 \times x_X}{(K_4 + x_X)}
\end{aligned}
$$

ODE results



$$K_1 = K_2 = K_3 = K_4 = 0.02\mu M$$

ODE results



$$K_1 = K_2 = K_3 = K_4 = 40\mu M$$

PRISM model

Each species is represented as a PRISM module.

For example, the cyclin is represented as:

module c
$c : [0..Nc]$ **init** 0;
$[a1]c < Nc \rightarrow (c' = c + 1)$;
$[a2]c > 0 \rightarrow (c' = c - 1)$;
$[a3]c > 0 \rightarrow (c' = c)$;
$[a7]c > 0 \rightarrow (c' = c - 1)$;
endmodule

We assume that there are 12 levels of C and 20 levels of the other species. This results in 5733 and 31744 transitions.

PRISM model

Each species is represented as a PRISM module.
For example, the cyclin is represented as:

module c
$c : [0..Nc]$ **init** 0;
$[a1]c < Nc \rightarrow (c' = c + 1)$;
$[a2]c > 0 \rightarrow (c' = c - 1)$;
$[a3]c > 0 \rightarrow (c' = c)$;
$[a7]c > 0 \rightarrow (c' = c - 1)$;
endmodule

We assume that there are 12 levels of C and 20 levels of the other species. This results in 5733 and 31744 transitions.

PRISM model

Each species is represented as a PRISM module.
For example, the cyclin is represented as:

module c
$c : [0..Nc]$ **init** 0;
$[a1]c < Nc \rightarrow (c' = c + 1)$;
$[a2]c > 0 \rightarrow (c' = c - 1)$;
$[a3]c > 0 \rightarrow (c' = c)$;
$[a7]c > 0 \rightarrow (c' = c - 1)$;
endmodule

We assume that there are 12 levels of C and 20 levels of the other species. This results in 5733 and 31744 transitions.

PRISM model (2)

An additional (dummy) module is needed to capture the kinetic rates.

module Functional_rates
dummy: bool **init** true;
$[a1]dummy = true \rightarrow r1 : (dummy' = dummy);$
$[a2]dummy = true \rightarrow r2 : (dummy' = dummy);$
$[a3]dummy = true \rightarrow r3 : (dummy' = dummy);$
$[a4]dummy = true \rightarrow r4 : (dummy' = dummy);$
$[a5]dummy = true \rightarrow r5 : (dummy' = dummy);$
$[a6]dummy = true \rightarrow r6 : (dummy' = dummy);$
$[a7]dummy = true \rightarrow r7 : (dummy' = dummy);$
endmodule

PRISM analysis

- **Probability that cyclin is exhausted in the cell.**

$$P =?[trueU[T, T]cyclin = 0]$$

- **Expected number of degradation reactions (both standard and by means of X)**

$$R\alpha_2 =?[C <= T] \quad \text{and} \quad R\alpha_7 =?[C <= T]$$

- **Probability that the level of active kinase (M) is greater than the level of inactive kinase (M') at time T**

$$P =?[trueU[T, T]M > M']$$

PRISM analysis

- **Probability that cyclin is exhausted in the cell.**

$$P =?[trueU[T, T]cyclin = 0]$$

- **Expected number of degradation reactions (both standard and by means of X)**

$$R\alpha_2 =?[C <= T] \quad \text{and} \quad R\alpha_7 =?[C <= T]$$

- **Probability that the level of active kinase (M) is greater than the level of inactive kinase (M') at time T**

$$P =?[trueU[T, T]M > M']$$

PRISM analysis

- **Probability that cyclin is exhausted in the cell.**

$$P =?[trueU[T, T]cyclin = 0]$$

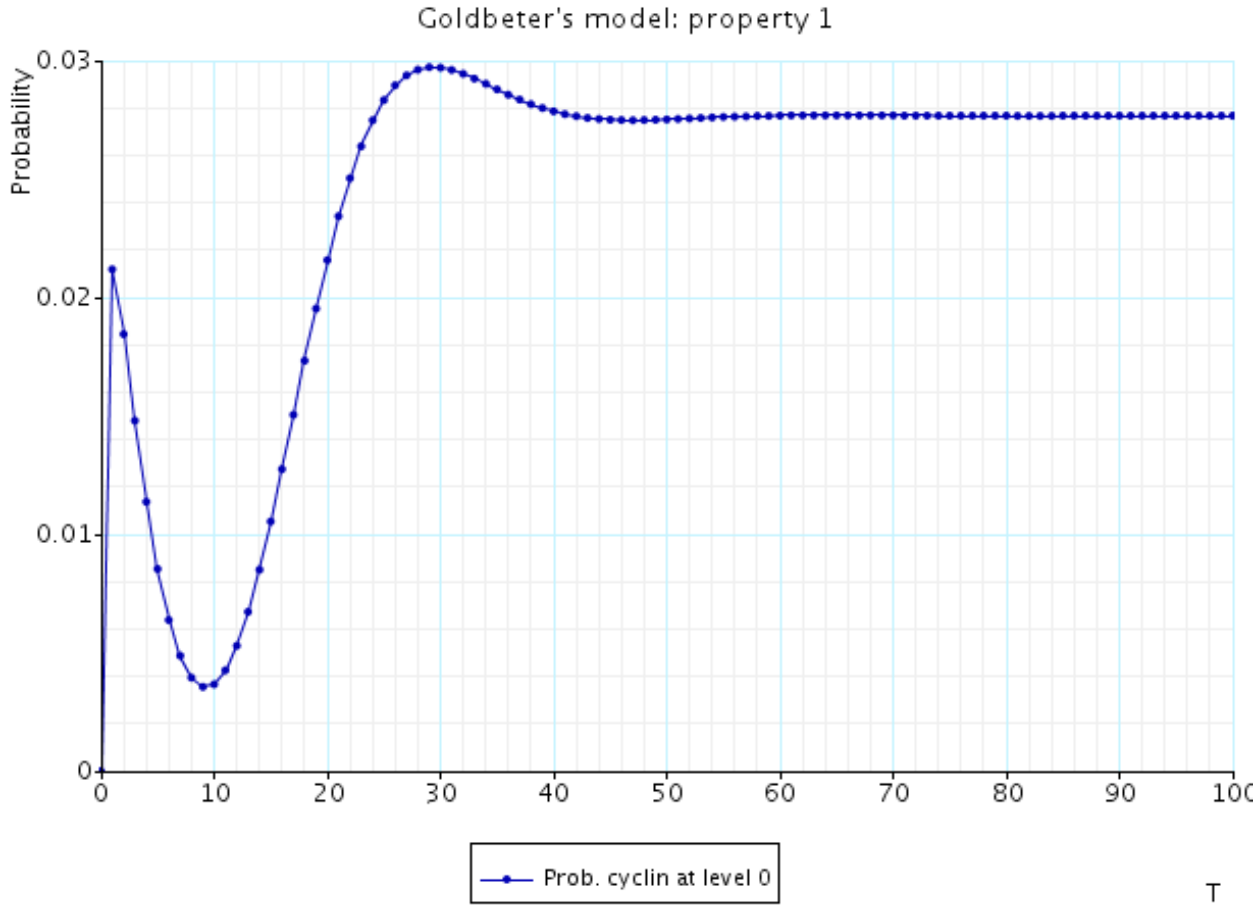
- **Expected number of degradation reactions (both standard and by means of X)**

$$R\alpha_2 =?[C <= T] \quad \text{and} \quad R\alpha_7 =?[C <= T]$$

- **Probability that the level of active kinase (M) is greater than the level of inactive kinase (M') at time T**

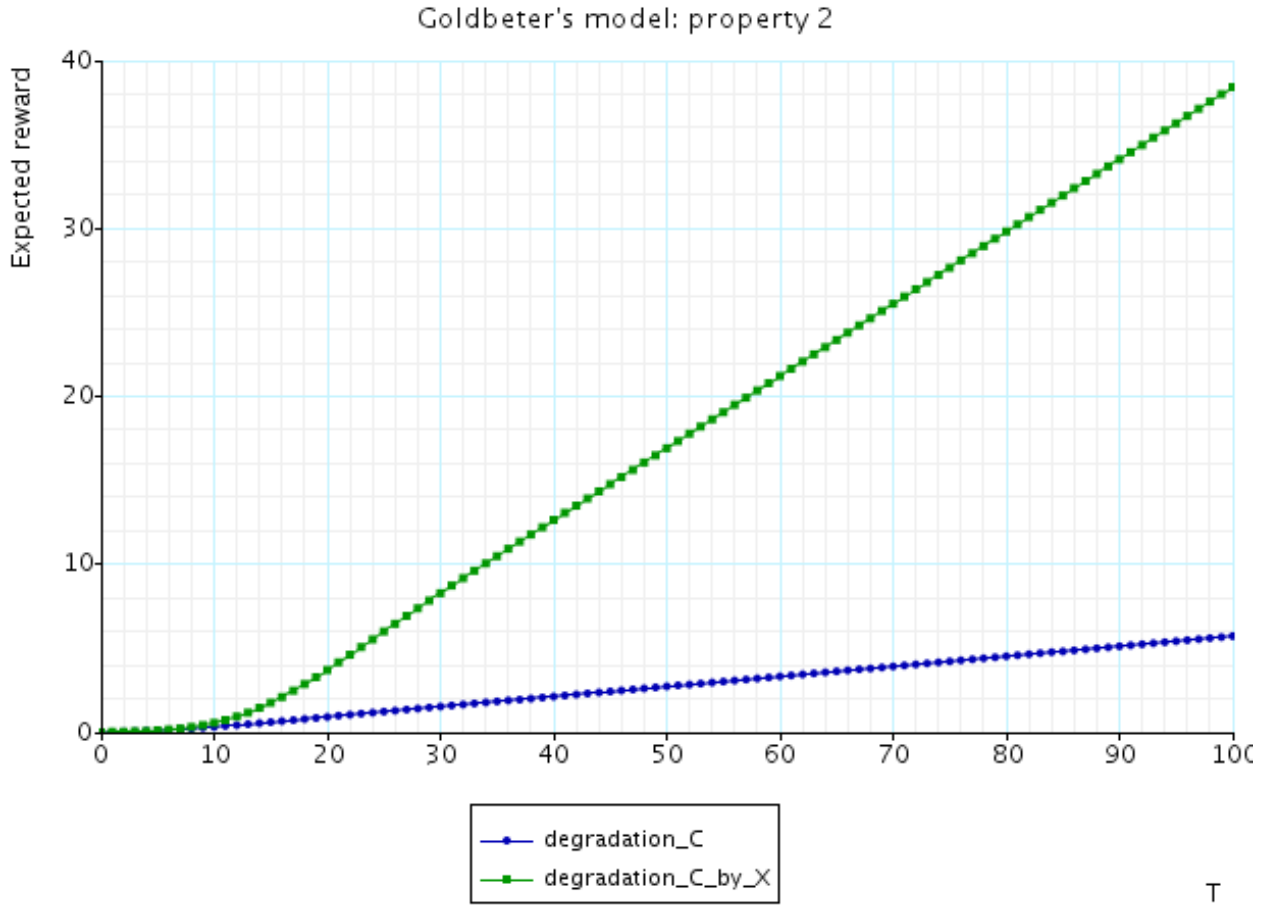
$$P =?[trueU[T, T]M > M']$$

PRISM results



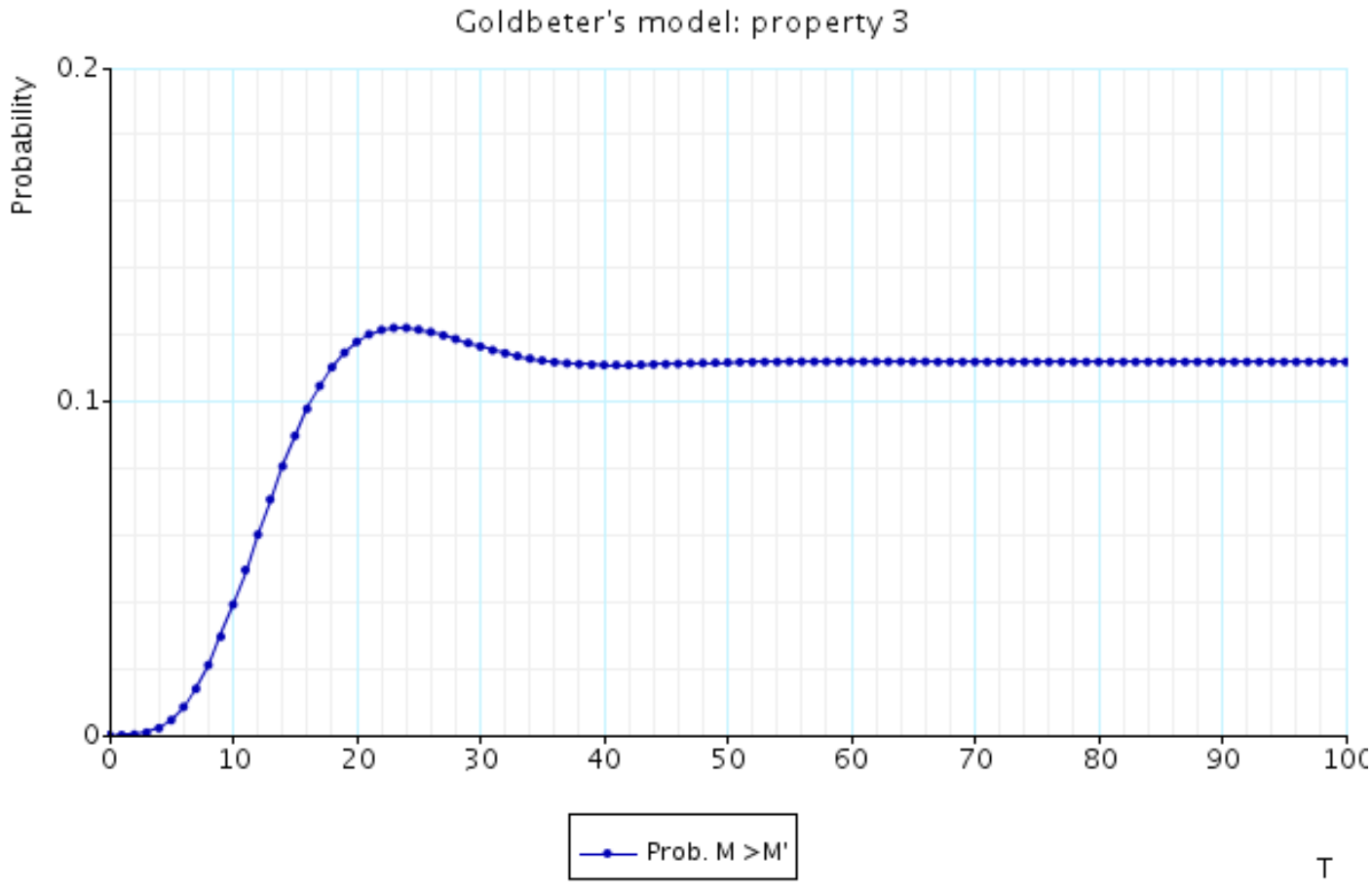
$$P =?[trueU[T, T]cyclin = 0]$$

PRISM results



$R\alpha_2 =?[C <= T]$ and $R\alpha_7 =?[C <= T]$

PRISM results



$$P =?[trueU[T, T]M > M']$$

References

- A. Regev and E. Shapiro, *Cells as computation*, in Nature, 419, pp. 343, 2001.
- C. Priami, A. Regev, W. Silverman and E. Shapiro, *Application of stochastic name-passing calculus to representation and simulation of molecular processes*, in Information Processing Letters, 80, pp. 25–31, 2001.
- D. Chiarugi, M. Curti, P. Degano and R. Marangoni, *VICE: A VIrtual CEll*, in *Proc. of 2nd Intl. Workshop on Computational Methods in Systems Biology* Paris, France, April 2004.
- F. Ciocchetta and J. Hillston, *Bio-PEPA: A framework for the modelling and analysis of biological systems*, in Theoretical Computer Science, 410(33-34), pp. 3065–3084, 2009.
- A. Goldbeter, *A Minimal Cascade Model for the Mitotic Oscillator Involving Cycling and CDC2 kinase*, in PNAS 8, pp. 9107–9111, 1991.