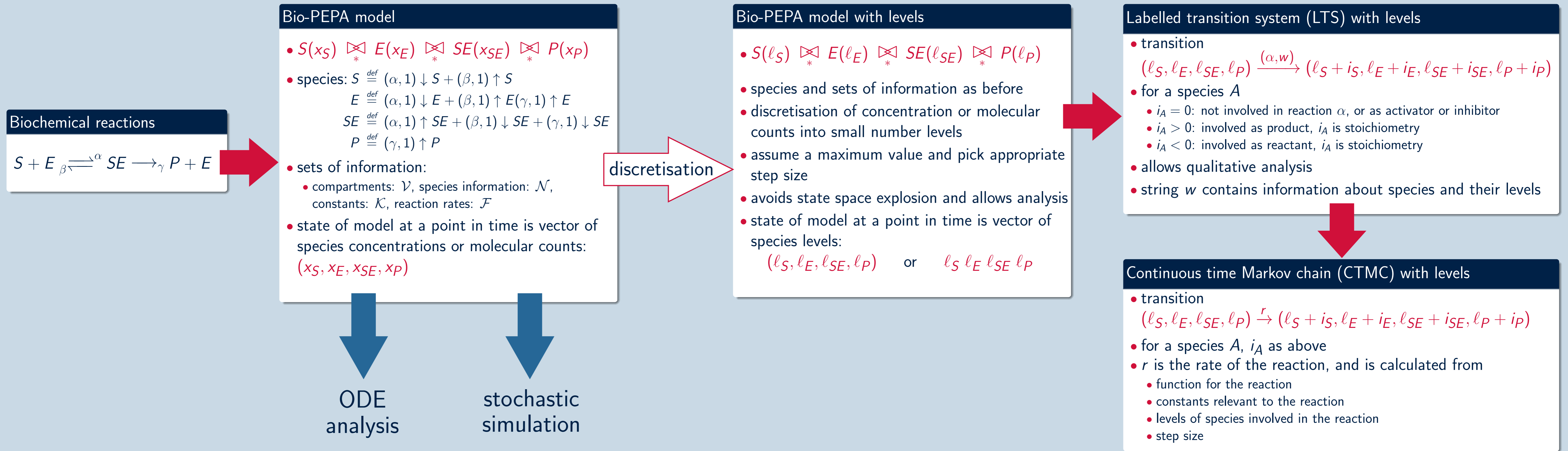# Using semantic equivalences to model biological behaviour

Vashti Galpin*, Federica Ciocchetta[‡], Jane Hillston*[†]

*Laboratory for Foundations of Computer Science, School of Informatics, University of Edinburgh; [‡]COSBI, Trento, Italy; [†]CSBE, University of Edinburgh
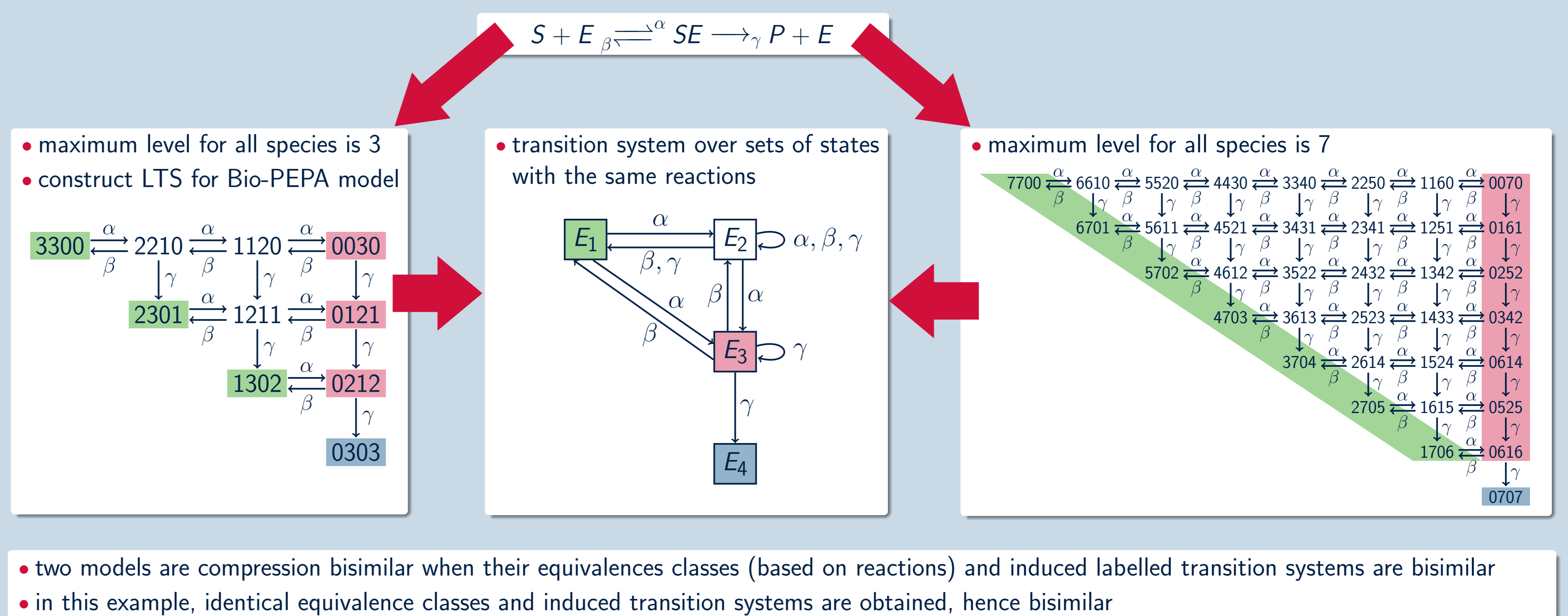
Process algebra modelling of biological systems

Biochemical reactions

$$S + E \; {}_\beta\!\rightleftharpoons^\alpha SE \longrightarrow_\gamma P + E$$

Bio-PEPA model

- $S(x_S) \bowtie_* E(x_E) \bowtie_* SE(x_{SE}) \bowtie_* P(x_P)$
- species:
$$S \stackrel{def}{=} (\alpha, 1) \downarrow S + (\beta, 1) \uparrow S$$
$$E \stackrel{def}{=} (\alpha, 1) \downarrow E + (\beta, 1) \uparrow E(\gamma, 1) \uparrow E$$
$$SE \stackrel{def}{=} (\alpha, 1) \uparrow SE + (\beta, 1) \downarrow SE + (\gamma, 1) \downarrow SE$$
$$P \stackrel{def}{=} (\gamma, 1) \uparrow P$$
- sets of information:
  - compartments: $\mathcal{V}$, species information: $\mathcal{N}$, constants: $\mathcal{K}$, reaction rates: $\mathcal{F}$
- state of model at a point in time is vector of species concentrations or molecular counts:
$$(x_S, x_E, x_{SE}, x_P)$$

→ discretisation →

Bio-PEPA model with levels

- $S(\ell_S) \bowtie_* E(\ell_E) \bowtie_* SE(\ell_{SE}) \bowtie_* P(\ell_P)$
- species and sets of information as before
- discretisation of concentration or molecular counts into small number levels
- assume a maximum value and pick appropriate step size
- avoids state space explosion and allows analysis
- state of model at a point in time is vector of species levels:
$$(\ell_S, \ell_E, \ell_{SE}, \ell_P) \quad \text{or} \quad \ell_S \; \ell_E \; \ell_{SE} \; \ell_P$$

Labelled transition system (LTS) with levels

- transition
$$(\ell_S, \ell_E, \ell_{SE}, \ell_P) \xrightarrow{(\alpha,w)} (\ell_S + i_S, \ell_E + i_E, \ell_{SE} + i_{SE}, \ell_P + i_P)$$
- for a species A
  - $i_A = 0$: not involved in reaction α, or as activator or inhibitor
  - $i_A > 0$: involved as product, i_A is stoichiometry
  - $i_A < 0$: involved as reactant, i_A is stoichiometry
- allows qualitative analysis
- string w contains information about species and their levels

Continuous time Markov chain (CTMC) with levels

- transition
$$(\ell_S, \ell_E, \ell_{SE}, \ell_P) \xrightarrow{r} (\ell_S + i_S, \ell_E + i_E, \ell_{SE} + i_{SE}, \ell_P + i_P)$$
- for a species A, i_A as above
- r is the rate of the reaction, and is calculated from
  - function for the reaction
  - constants relevant to the reaction
  - levels of species involved in the reaction
  - step size

ODE analysis

stochastic simulation

Semantic equivalences

- relates Bio-PEPA models that have the same behaviour
- defined over LTS with levels or CTMC with levels
- various notions of same behaviour
  - behaviour exhibited by different discretisations of same Bio-PEPA model
  - behaviour seen in different abstractions of reaction
  - from computer science – bisimulation
- basic bisimulation definition for two models P and Q
  - for each state in P there is a state in Q
    - such that if P has a transition, $P \xrightarrow{\theta} P'$
    - then Q has a transition, $Q \xrightarrow{\phi} Q'$
    - with reaction info θ and ϕ related
    - and the states P' and Q' are related in the same way as P and Q
  - for each state in Q, the same holds
- types of equivalence
  - qualitative – does not consider rates of reactions
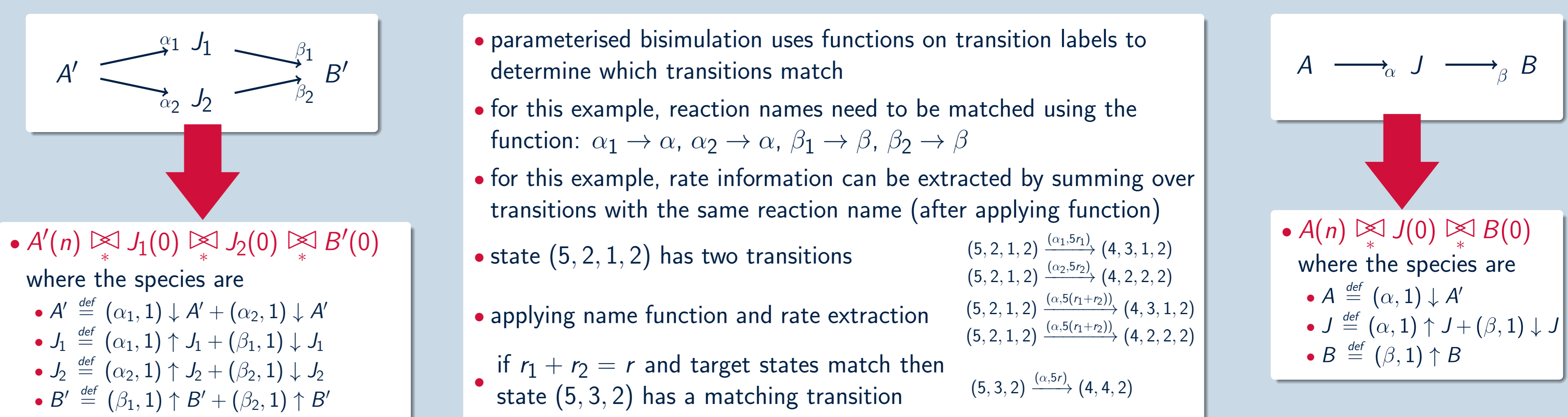 - quantitative – partly determined by rates of reactions

Equivalent behaviour based on two discretisations of same model: compression bisimulation
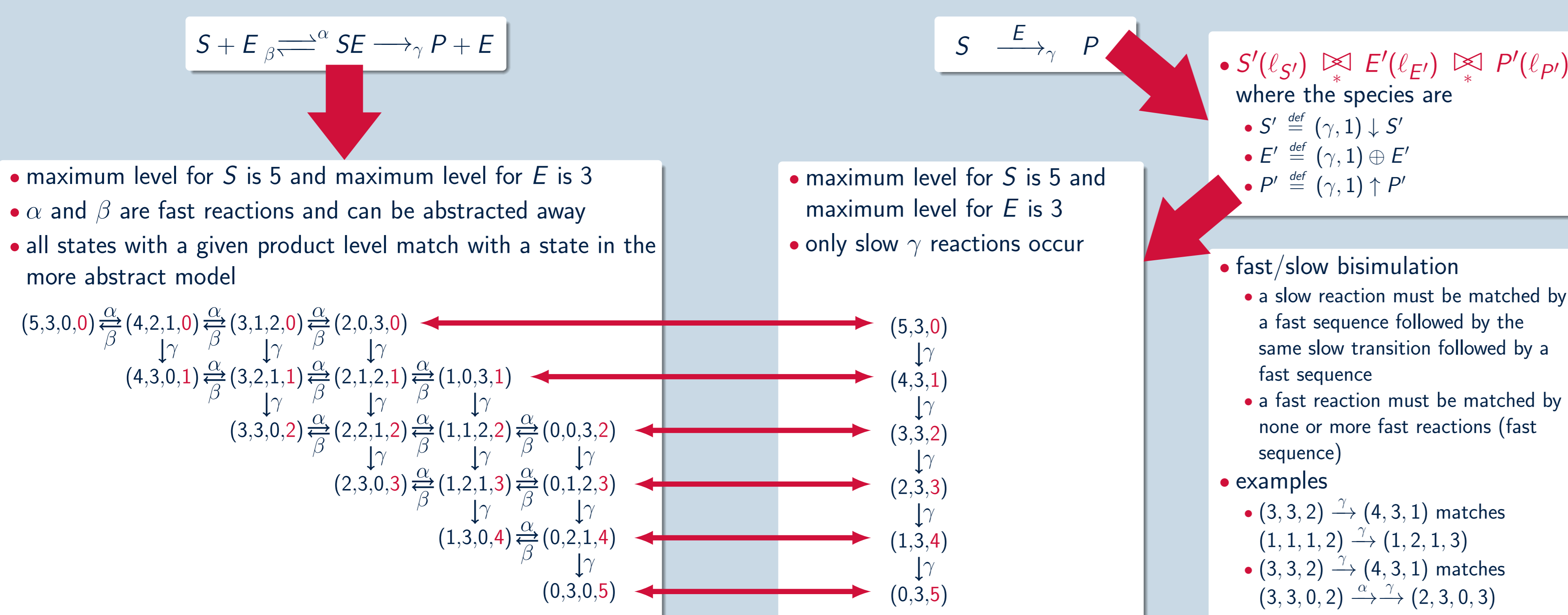
$$S + E \; {}_\beta\!\rightleftharpoons^\alpha SE \longrightarrow_\gamma P + E$$

- maximum level for all species is 3
- construct LTS for Bio-PEPA model

- transition system over sets of states with the same reactions

- maximum level for all species is 7

- two models are compression bisimilar when their equivalences classes (based on reactions) and induced labelled transition systems are bisimilar
- in this example, identical equivalence classes and induced transition systems are obtained, hence bisimilar

Use and application

- congruence
  - when a semantic equivalence interacts well with model language
  - consider $P \bowtie_* A$ which adds species A to system P
  - if P and Q have the same behaviour then $P \bowtie_* A$ has the same behaviour as $Q \bowtie_* A$
  - can consider congruence at system level and at species level
- can substitute smaller congruent systems into model
- reduce state space size and gives same behaviour
- further research
  - use equivalences developed on other models
  - investigate variants of parameterised bisimulation

Equivalent behaviour based on different abstractions of reactions: parameterised bisimulation

- $A'(n) \bowtie_* J_1(0) \bowtie_* J_2(0) \bowtie_* B'(0)$
  where the species are
- $A' \stackrel{def}{=} (\alpha_1, 1) \downarrow A' + (\alpha_2, 1) \downarrow A'$
- $J_1 \stackrel{def}{=} (\alpha_1, 1) \uparrow J_1 + (\beta_1, 1) \downarrow J_1$
- $J_2 \stackrel{def}{=} (\alpha_2, 1) \uparrow J_2 + (\beta_2, 1) \downarrow J_2$
- $B' \stackrel{def}{=} (\beta_1, 1) \uparrow B' + (\beta_2, 1) \uparrow B'$

- parameterised bisimulation uses functions on transition labels to determine which transitions match
- for this example, reaction names need to be matched using the function: $\alpha_1 \to \alpha$, $\alpha_2 \to \alpha$, $\beta_1 \to \beta$, $\beta_2 \to \beta$
- for this example, rate information can be extracted by summing over transitions with the same reaction name (after applying function)
- state $(5, 2, 1, 2)$ has two transitions
$$(5,2,1,2) \xrightarrow{(\alpha_1, 5r_1)} (4,3,1,2)$$
$$(5,2,1,2) \xrightarrow{(\alpha_2, 5r_2)} (4,2,2,2)$$
- applying name function and rate extraction
$$(5,2,1,2) \xrightarrow{(\alpha, 5(r_1+r_2))} (4,3,1,2)$$
$$(5,2,1,2) \xrightarrow{(\alpha, 5(r_1+r_2))} (4,2,2,2)$$
- if $r_1 + r_2 = r$ and target states match then state $(5, 3, 2)$ has a matching transition
$$(5,3,2) \xrightarrow{(\alpha, 5r)} (4,4,2)$$

$$A \longrightarrow_\alpha J \longrightarrow_\beta B$$

- $A(n) \bowtie_* J(0) \bowtie_* B(0)$
  where the species are
- $A \stackrel{def}{=} (\alpha, 1) \downarrow A'$
- $J \stackrel{def}{=} (\alpha, 1) \uparrow J + (\beta, 1) \downarrow J$
- $B \stackrel{def}{=} (\beta, 1) \uparrow B$

Equivalent behaviour based on different abstractions of reactions: fast/slow bisimulation

$$S + E \; {}_\beta\!\rightleftharpoons^\alpha SE \longrightarrow_\gamma P + E$$

$$S \xrightarrow{E}_\gamma P$$

- $S'(\ell_{S'}) \bowtie_* E'(\ell_{E'}) \bowtie_* P'(\ell_{P'})$
  where the species are
- $S' \stackrel{def}{=} (\gamma, 1) \downarrow S'$
- $E' \stackrel{def}{=} (\gamma, 1) \oplus E'$
- $P' \stackrel{def}{=} (\gamma, 1) \uparrow P'$

- maximum level for S is 5 and maximum level for E is 3
- α and β are fast reactions and can be abstracted away
- all states with a given product level match with a state in the more abstract model

- maximum level for S is 5 and maximum level for E is 3
- only slow γ reactions occur

- fast/slow bisimulation
  - a slow reaction must be matched by a fast sequence followed by the same slow transition followed by a fast sequence
  - a fast reaction must be matched by none or more fast reactions (fast sequence)
- examples
  - $(3, 3, 2) \xrightarrow{\gamma} (4, 3, 1)$ matches $(1, 1, 1, 2) \xrightarrow{\gamma} (1, 2, 1, 3)$
  - $(3, 3, 2) \xrightarrow{\gamma} (4, 3, 1)$ matches $(3, 3, 0, 2) \xrightarrow{\gamma} (2, 3, 0, 3)$

References and acknowledgements

- F. Ciocchetta and J. Hillston, Bio-PEPA: a framework for the modelling and analysis of biological systems, *Theoretical Computer Science*, 410:30653084, 2009.
- V. Galpin and J. Hillston, A semantic equivalence for Bio-PEPA based on discretisation of continuous values, *Theoretical Computer Science*, to appear, 2011.
- V. Galpin, Equivalences for a biological process algebra, submitted for publication
- F. Ciocchetta, V. Galpin and J. Hillston, An equivalence for abstracting from fast reactions, in preparation