Università di Bologna

Fourth day

Abstract interpretation of protein-protein interactions networks

Jérôme Feret

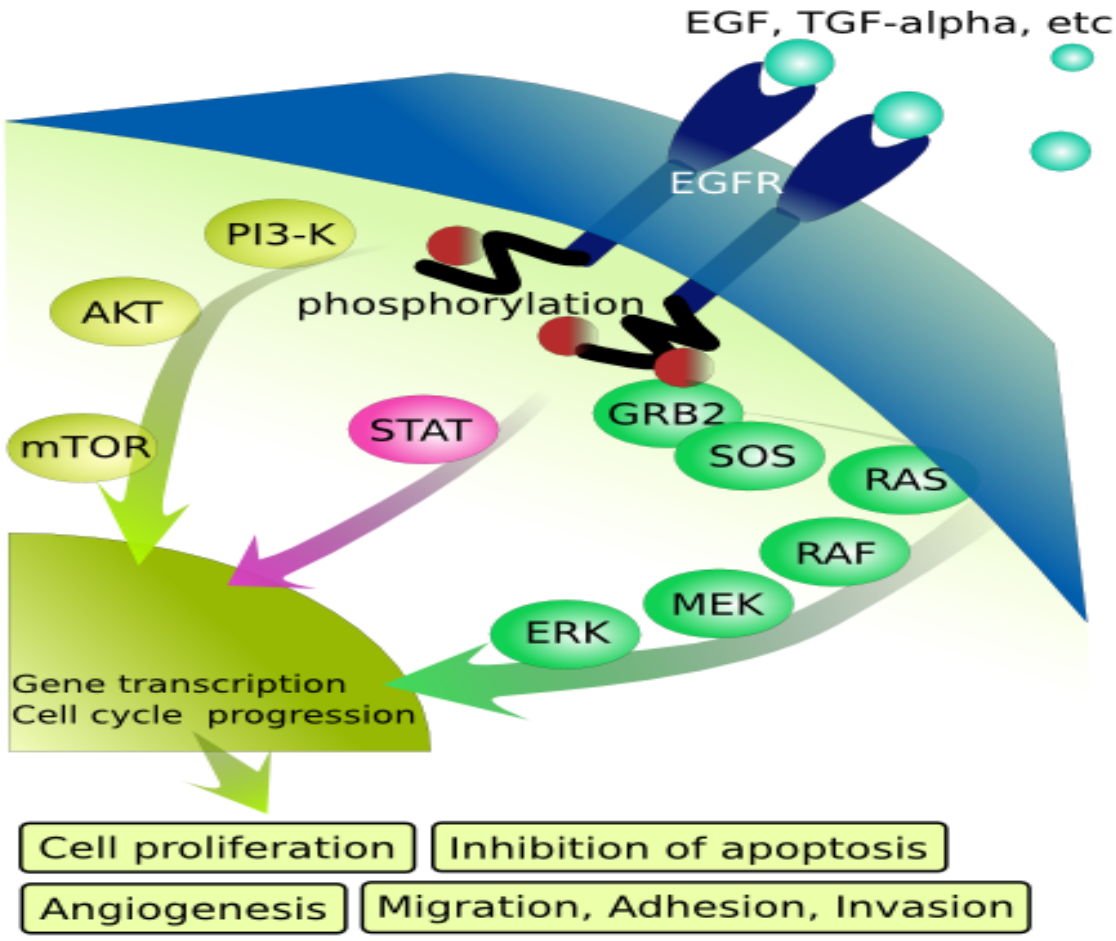
Laboratoire d'Informatique de l'École Normale Supérieure

INRIA, ÉNS, CNRS

http://www.di.ens.fr/∼feret

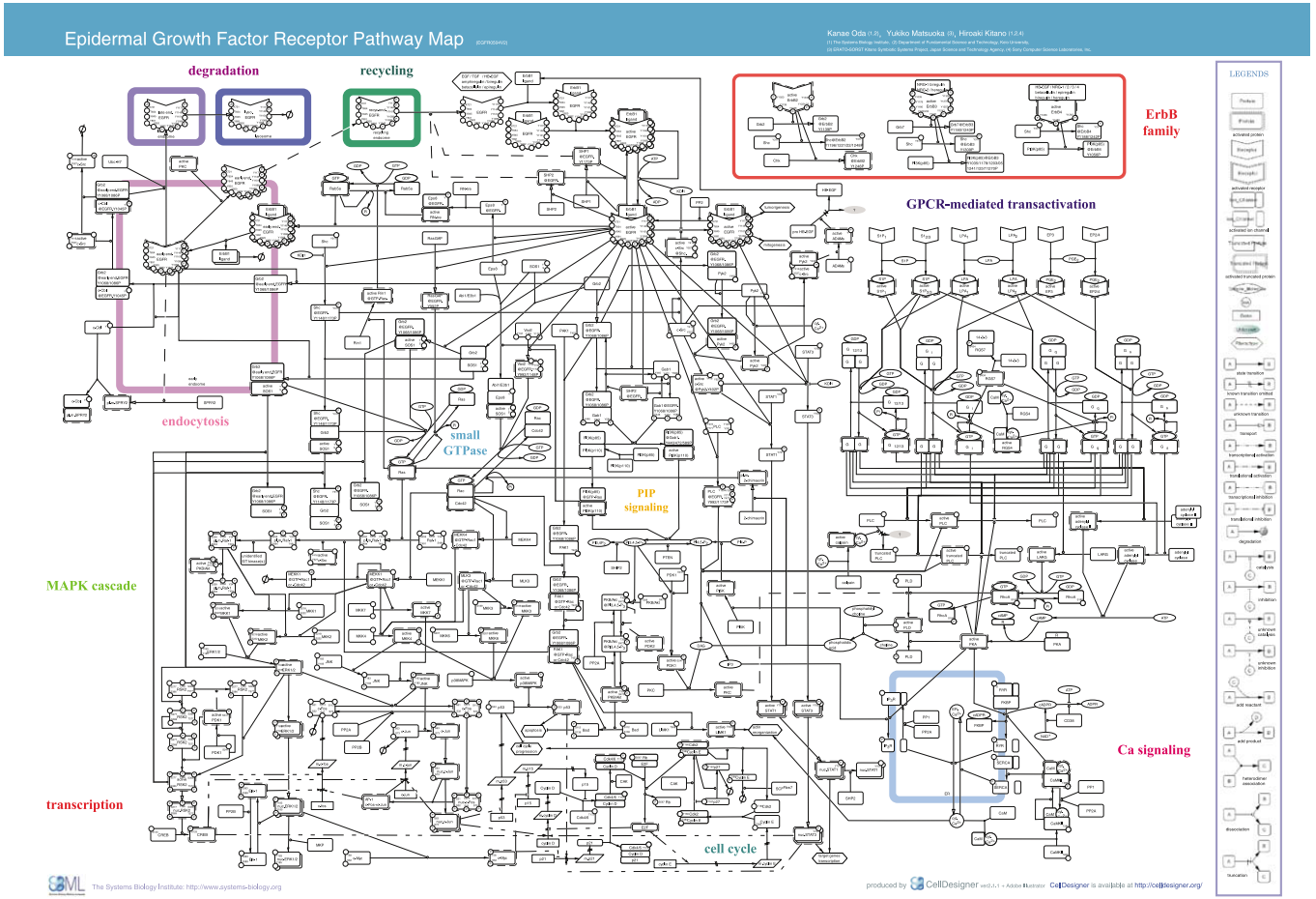
Thursday, April the 21st

Signalling Pathways



Eikuch, 2007

Pathway maps



Oda, Matsuoka, Funahashi, Kitano, Molecular Systems Biology, 2005

Differential models

$$
\begin{cases}
\frac{dx_1}{dt} = -k_1 \cdot x_1 \cdot x_2 + k_{-1} \cdot x_3 \\
\frac{dx_2}{dt} = -k_1 \cdot x_1 \cdot x_2 + k_{-1} \cdot x_3 \\
\frac{dx_3}{dt} = k_1 \cdot x_1 \cdot x_2 - k_{-1} \cdot x_3 + 2 \cdot k_2 \cdot x_3 \cdot x_3 - k_{-2} \cdot x_4) \\
\frac{dx_4}{dt} = k_2 \cdot x_3^2 - k_2 \cdot x_4 + \frac{v_4 \cdot x_5}{p_4 + x_5} - (k_3 \cdot x_4 - k_{-3} \cdot x_5) \\
\frac{dx_5}{dt} = \cdots \\
\quad \vdots \\
\frac{dx_n}{dt} = -k_1 \cdot x_1 \cdot c_2 + k_{-1} \cdot x_3
\end{cases}
$$

– do not describe the structure of molecules;
– combinatorial explosion: forces choices that are not principled;
– a nightmare to modify.

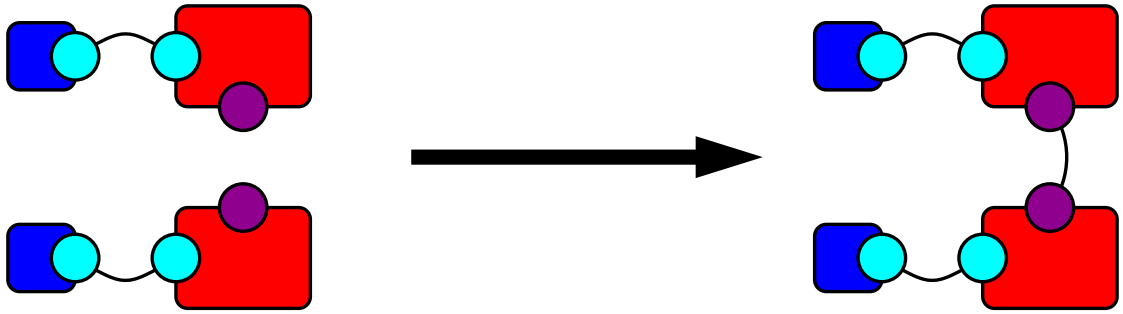
A gap between two worlds

Two levels of description:

1. Databases of proteins interactions in natural language
   + documented and detailed description
   + transparent description
   − cannot be interpreted
2. ODE-based models
   + can be integrated
   − opaque modelling process, models can hardly be modified
   − there are also some scalability issues.

Rule-based approach

We use site graph rewrite systems

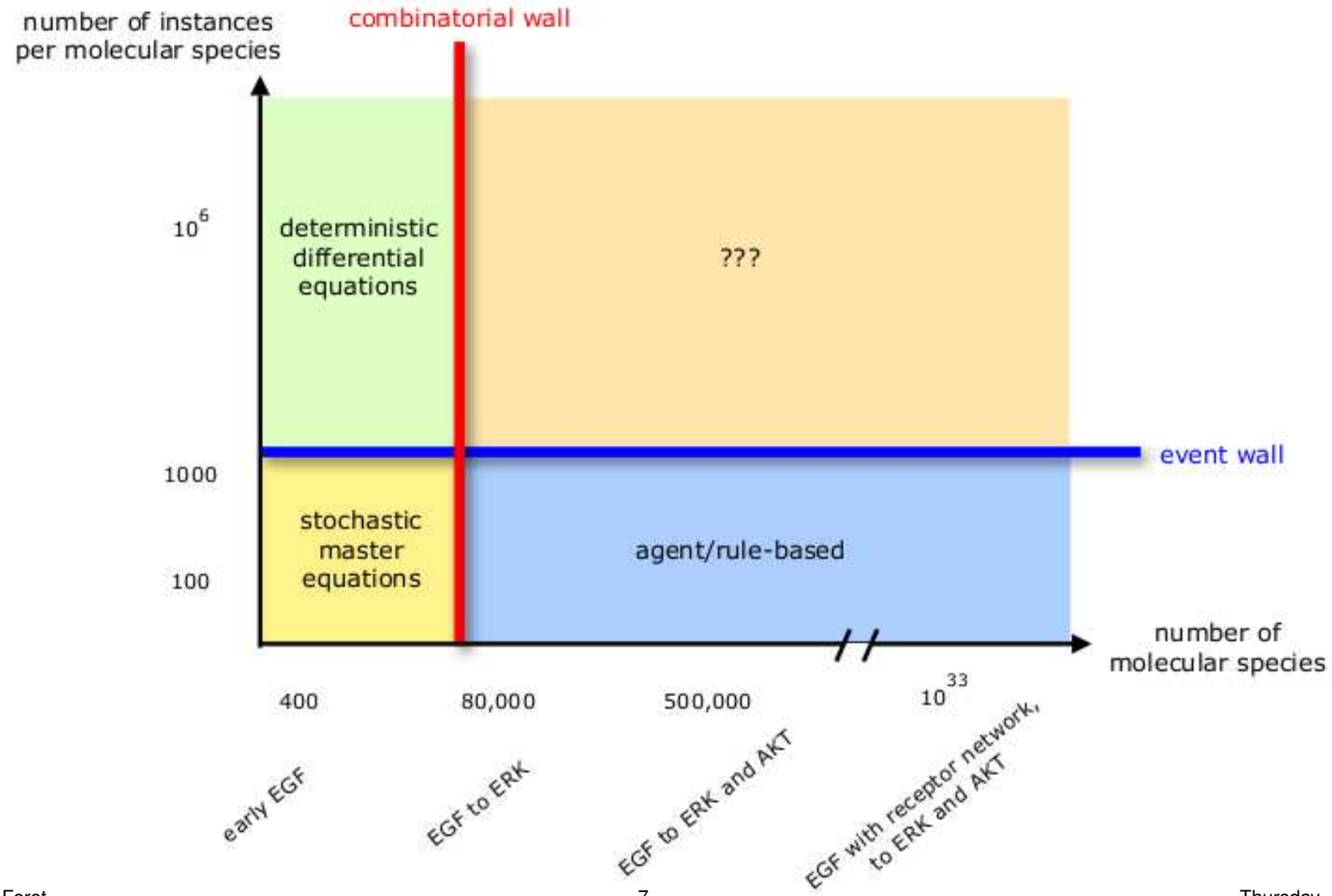


1. The description level matches with both
   - the observation level
   - and the intervention level

   of the biologist.

   We can tune the model easily.
2. Model description is very compact.
3. Quantitative semantics can be defined.

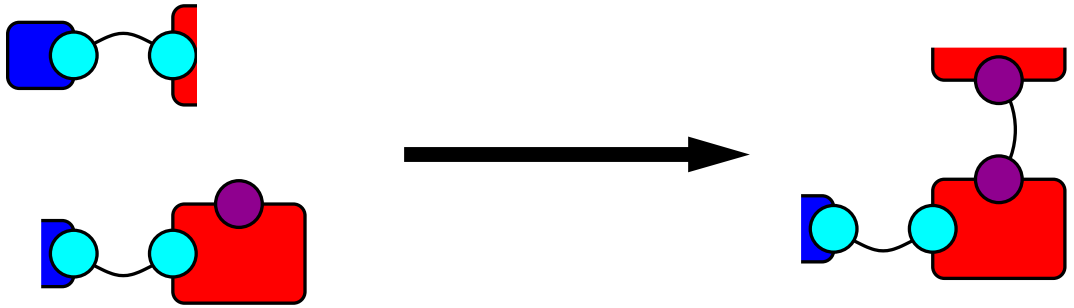
Complexity walls



Static analysis of reachable species (I/II)

Semi-fluid medium: the notion of individual is meaningless.

Design a static analysis to approximate the set of reachable species [VMCAI'08] which focuses on the relationships between the states of the sites of each agent:



This analysis is efficient, suitable to our problem, and accurate.

Static analysis of reachable species (II/II)

Applications:

1. check the consistency of a model [ICCMSE'07]
2. compute the properties to allow fast simulation [APLAS'07]
3. simplify models,
4. compute independent fragments of chemical species [PNAS'09, LICS'10,Chaos'10]

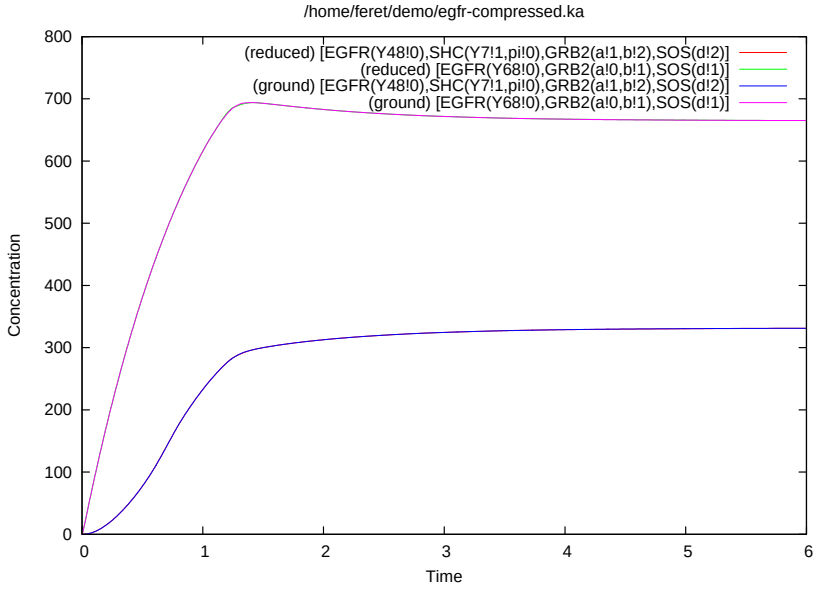
The analysis is complete (no false positif) for a significatif kernel of Kappa [VMCAI'08].

Model reduction

The ground differential system uses one variable per chemical species;

We directly compute its exact projection over independent fragments of chemical species.

With a small model, 356 chemical species are reduced into 38 fragments:



On a bigger model, 10^{19} chemical species are reduced into 180 000 fragments. [PNAS'09,LICS'10,Chaos'10]

Università di Bologna

Fourth day

Reachability Analysis of Rule-based Models

[ICCMSE'07,VMCAI'08]

Jérôme Feret

Laboratoire d'Informatique de l'École Normale Supérieure

INRIA, ÉNS, CNRS

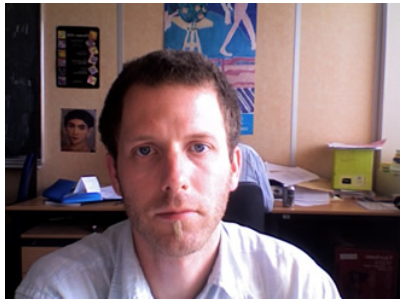
http://www.di.ens.fr/~feret

Thursday, April the 21st

Joint-work with...

Walter Fontana
Harvard Medical School

Vincent Danos
Edinburgh



Russ Harmer
Harvard Medical School

Jean Krivine
Paris VII

Overview

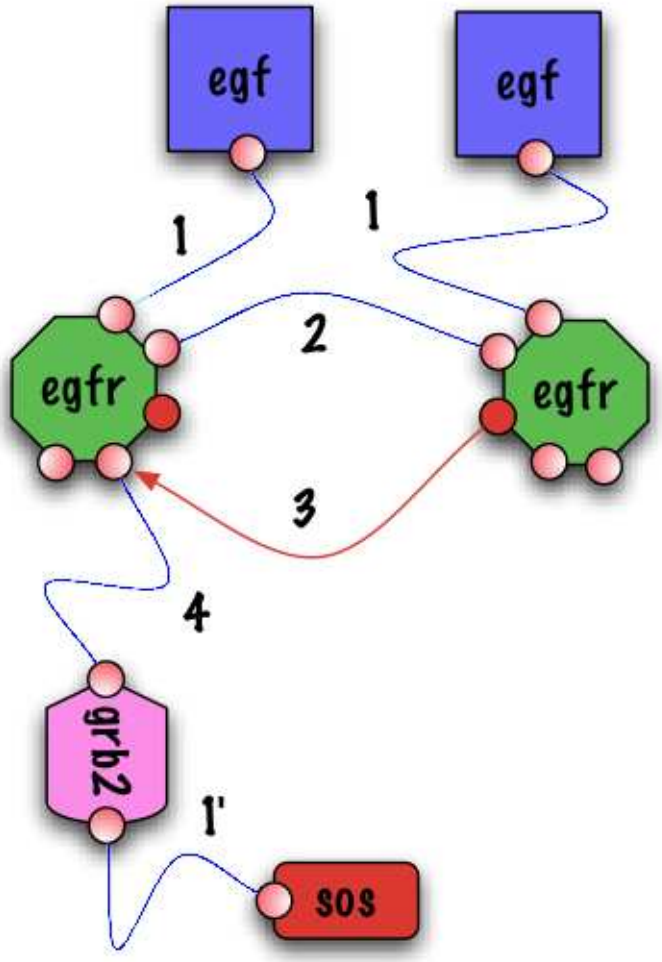
1. Introduction
2. Kappa language
3. Local views
4. Local set of chemical species
5. Local rule systems
6. Decontextualization
7. Conclusion

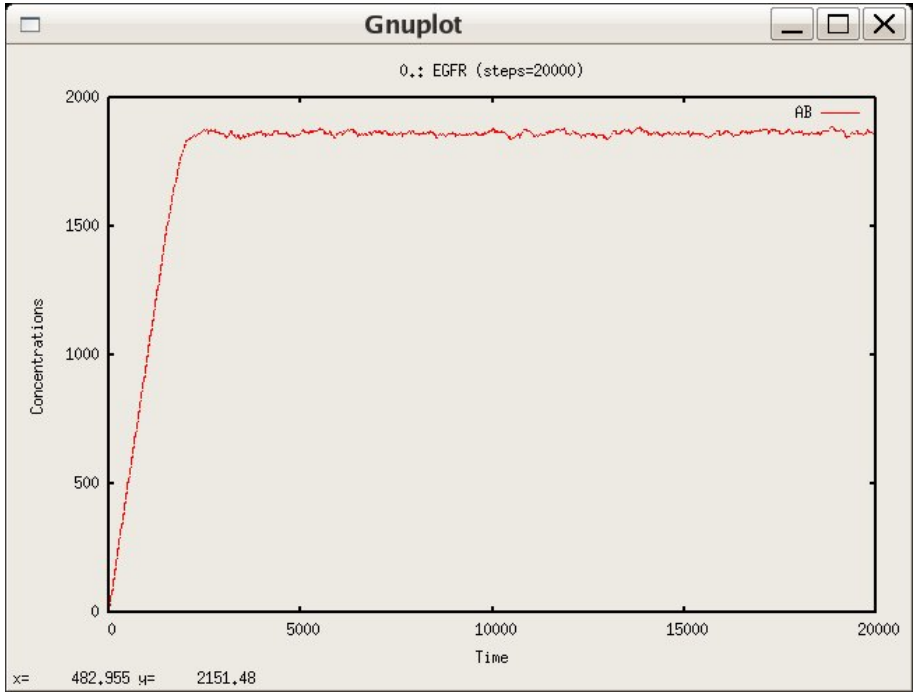
Modeling signaling pathway

- Signaling pathway:
  - A cell measures (i.e. checks thresholds, integrates, compares) the concentration of some proteins in order to make decisions.
  - Many proteins (enzymes, receptors, transport molecules) are involved.
  - They interact by binding with each other and activating each other.
- rule-based models:
  - A site graph-based rewrite language.
  - Description level matches with biologists' observation and manipulation level.
- Static analysis:

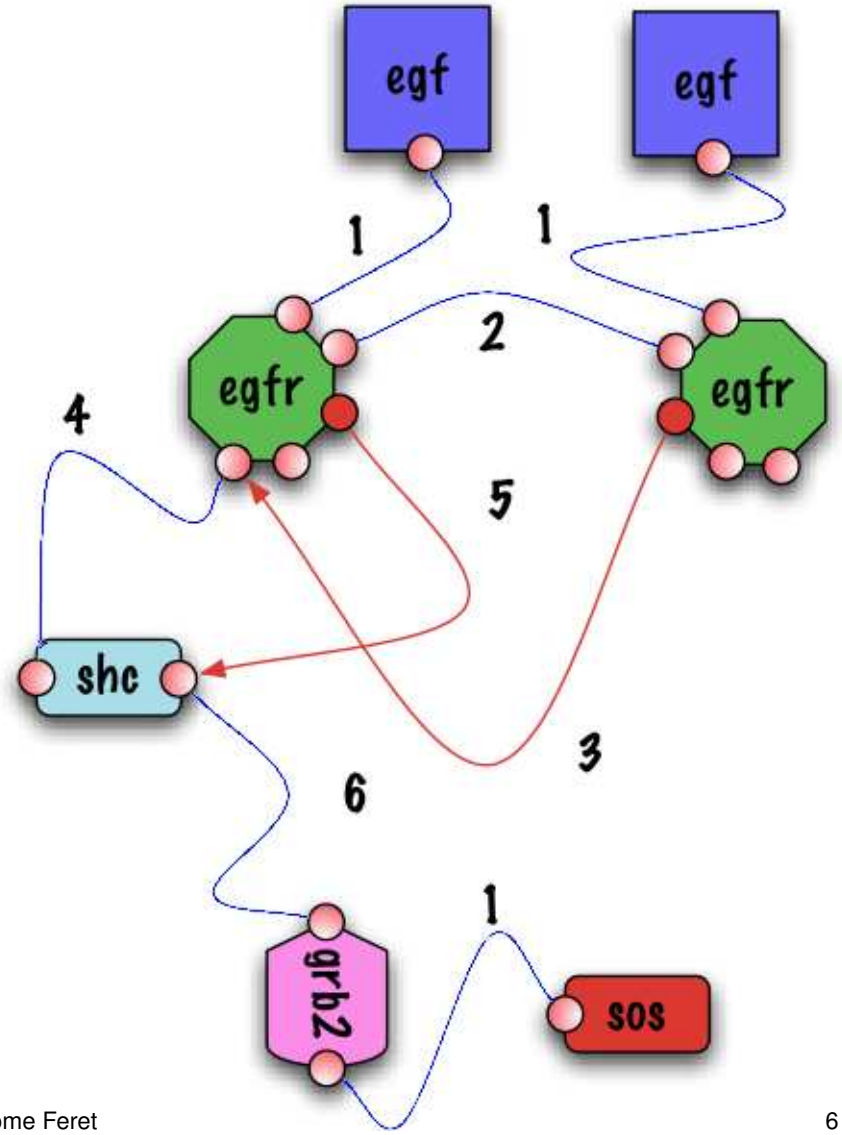
  We propose some static analysis tools in order to:
  - help the design of rule-based models;
  - compute (*abstract*) the properties of rule-based models.

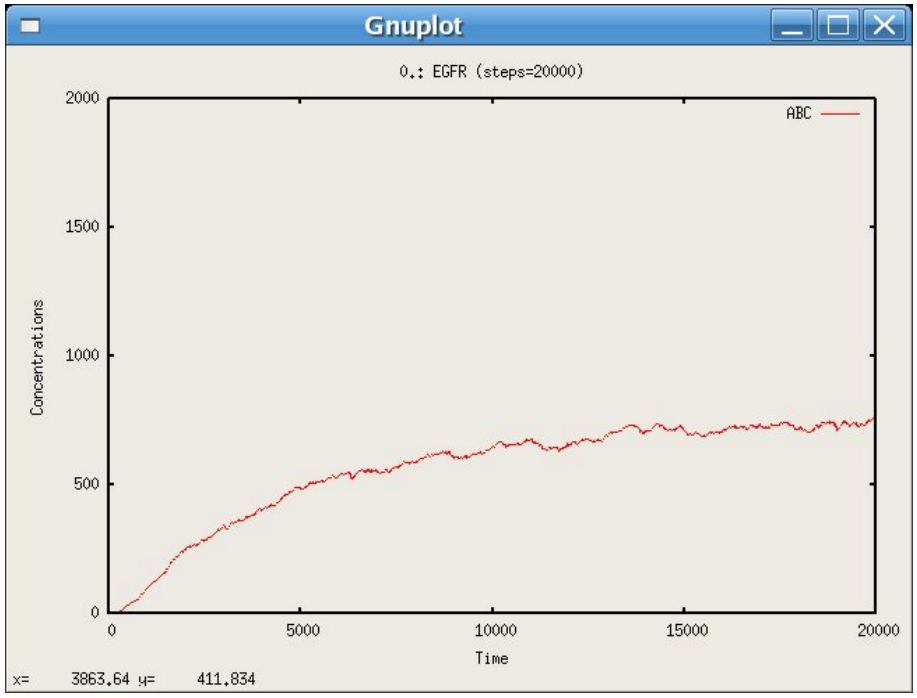
A single story





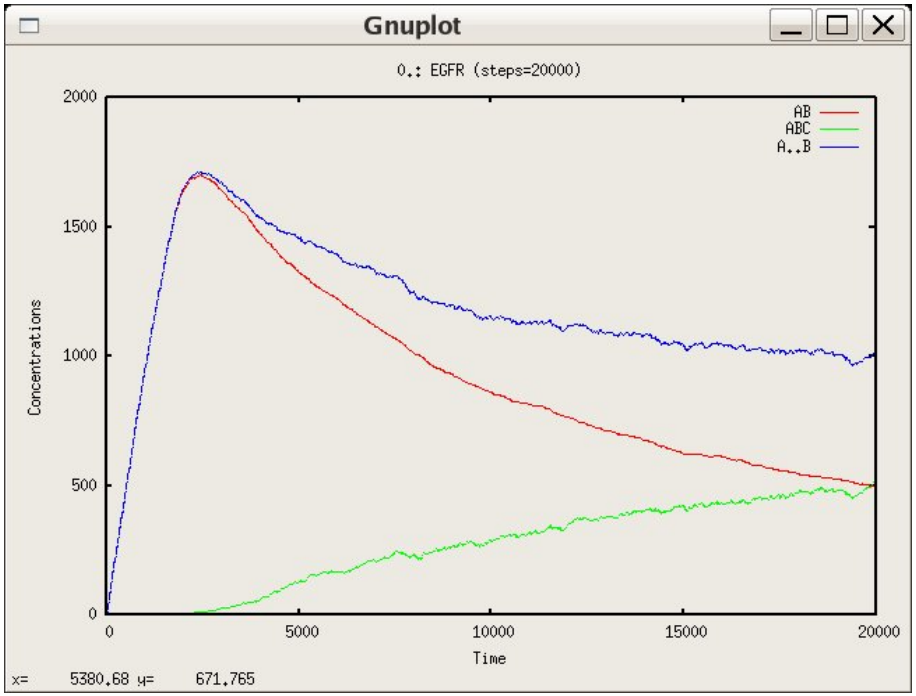
A concurrent story





Overshoot

When we combine the two stories. . .

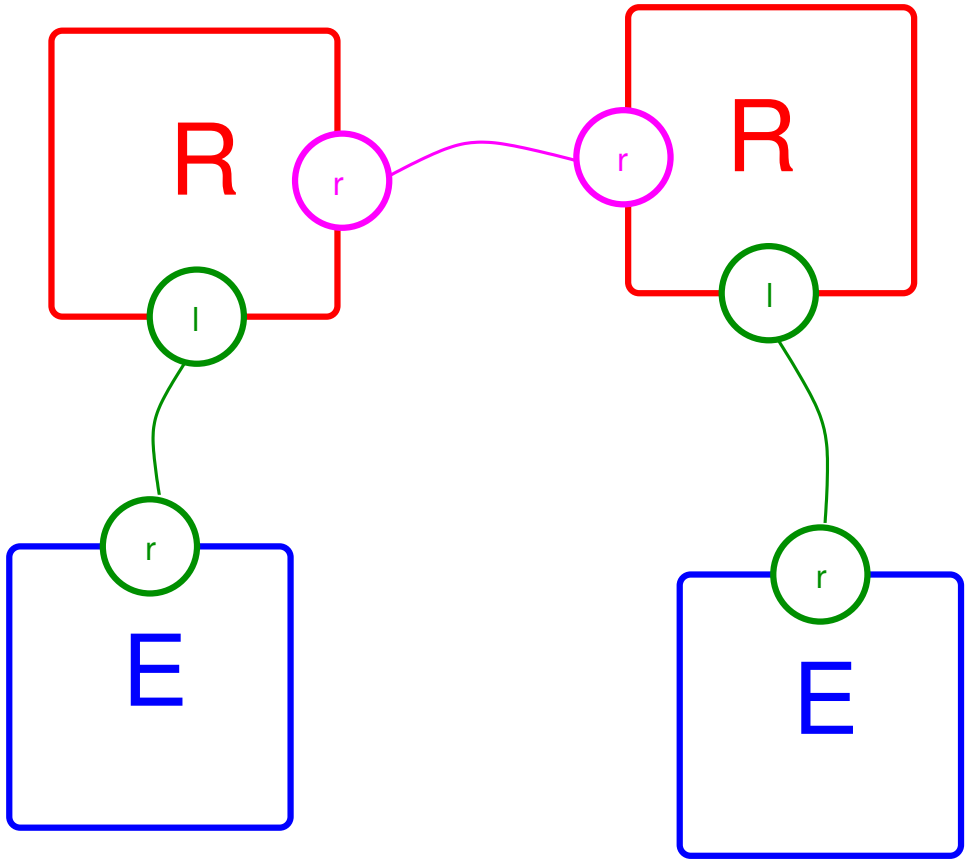


. . . we get an overshoot.

Overview

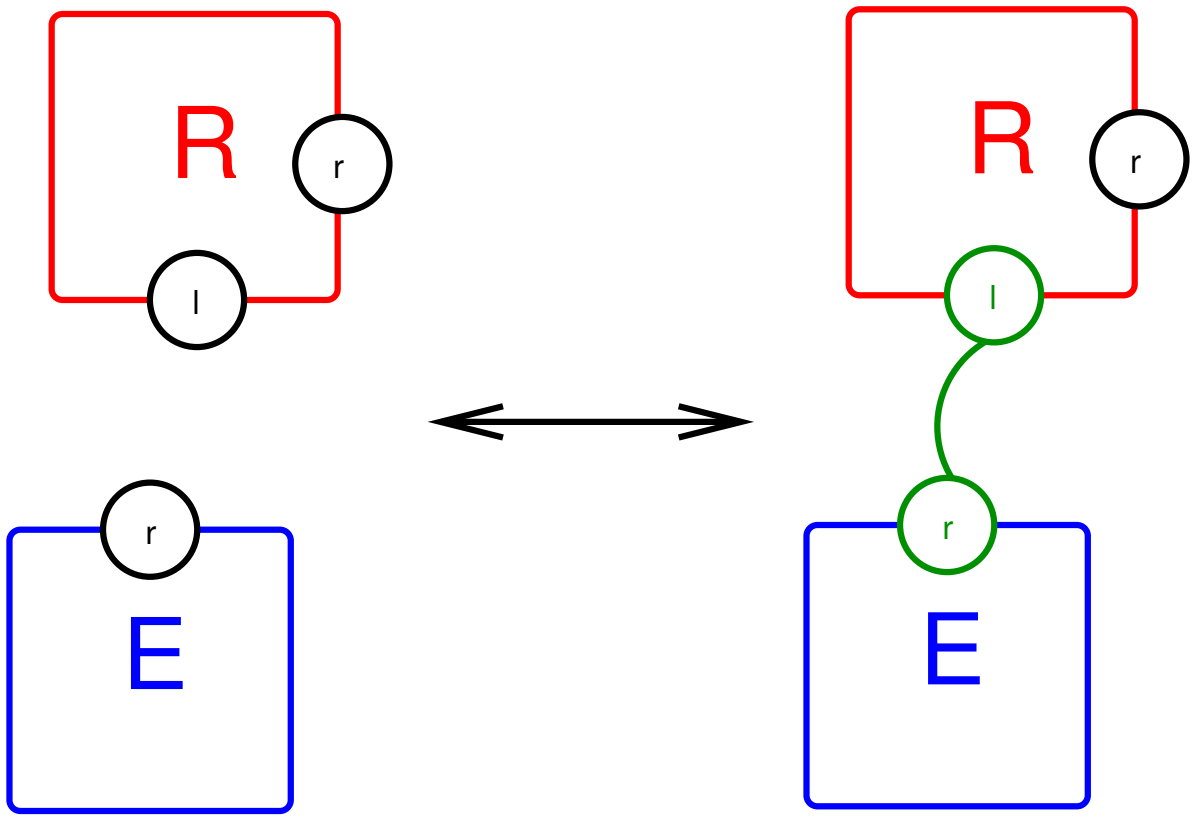
1. Introduction
2. Kappa language
3. Local views
4. Local set of chemical species
5. Local rule systems
6. Decontextualization
7. Conclusion

A chemical species



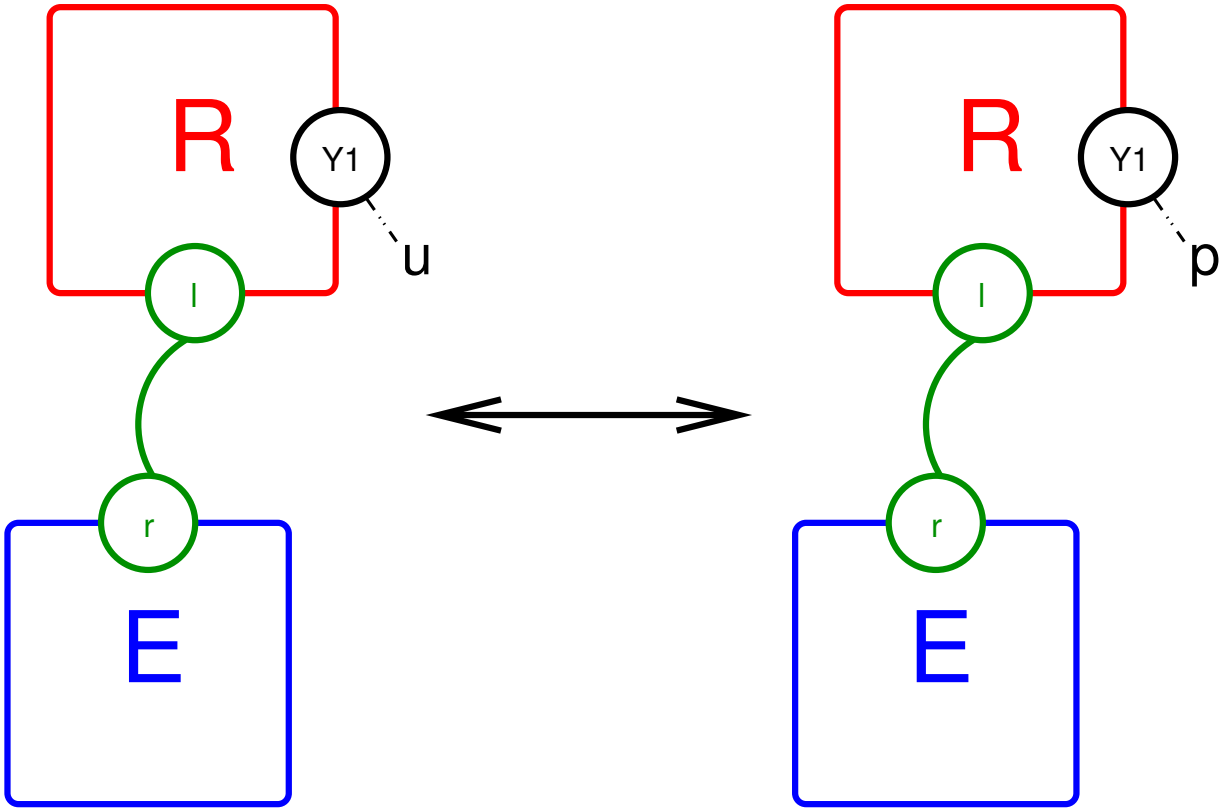
E(r!1), R(l!1,r!2), R(r!2,l!3), E(r!3)

A Unbinding/Binding Rule



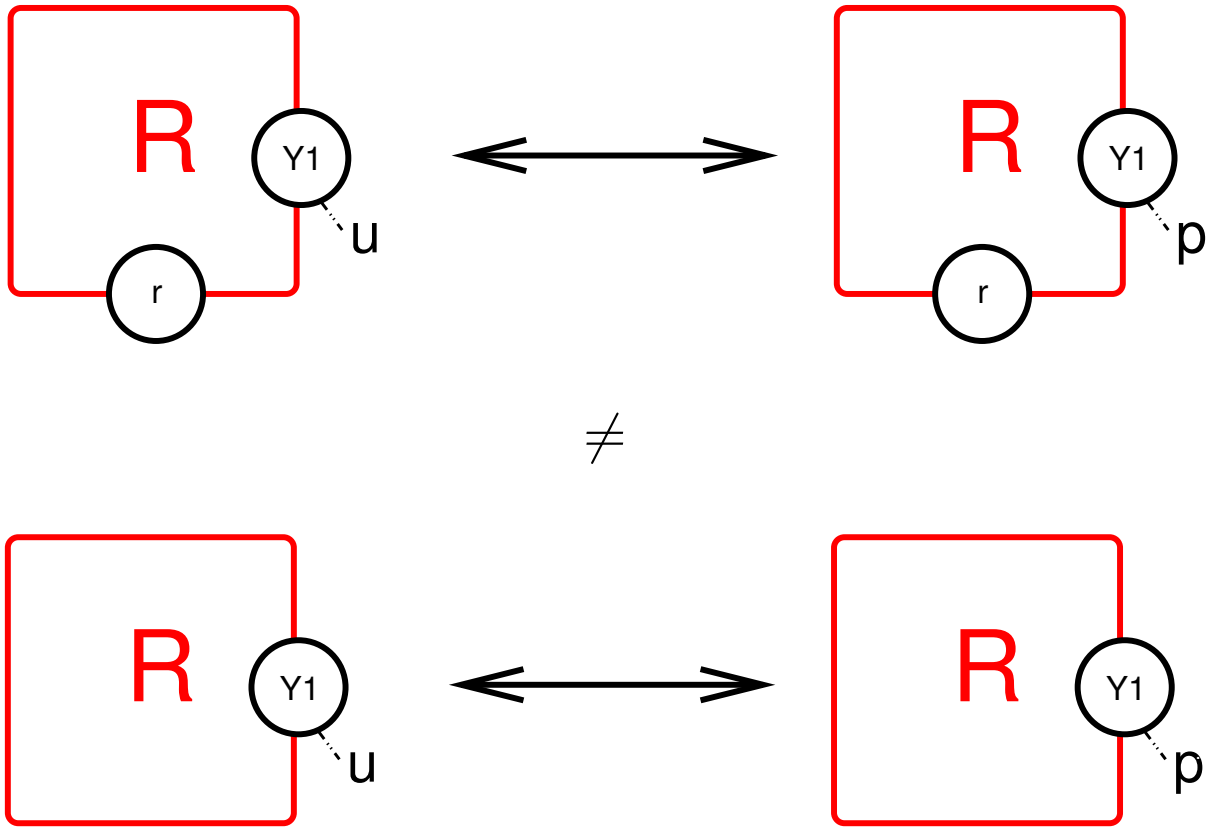
$$E(r), R(l,r) \longleftrightarrow E(r!1), R(l!1,r)$$

Internal state

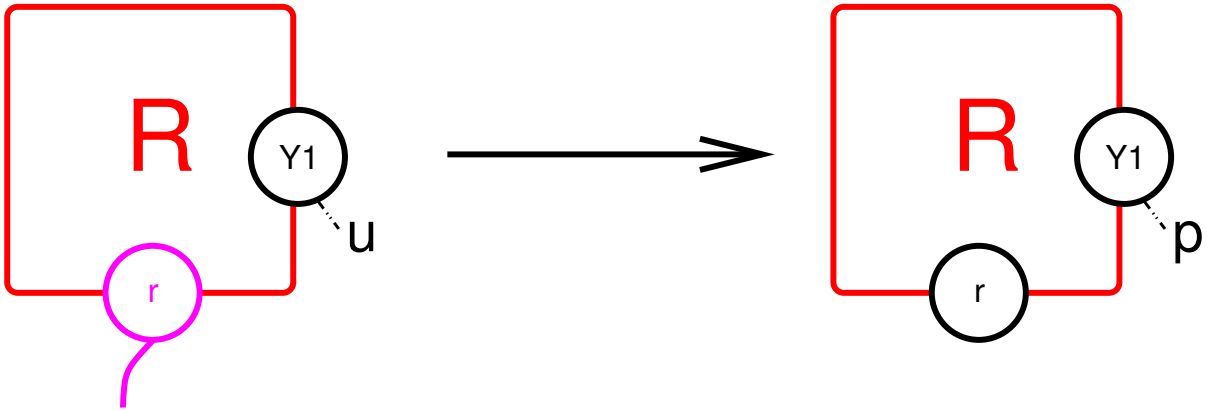


R(Y1~u,l!1), E(r!1) ⟷ R(Y1~p,l!1), E(r!1)

Don’t care, Don’t write

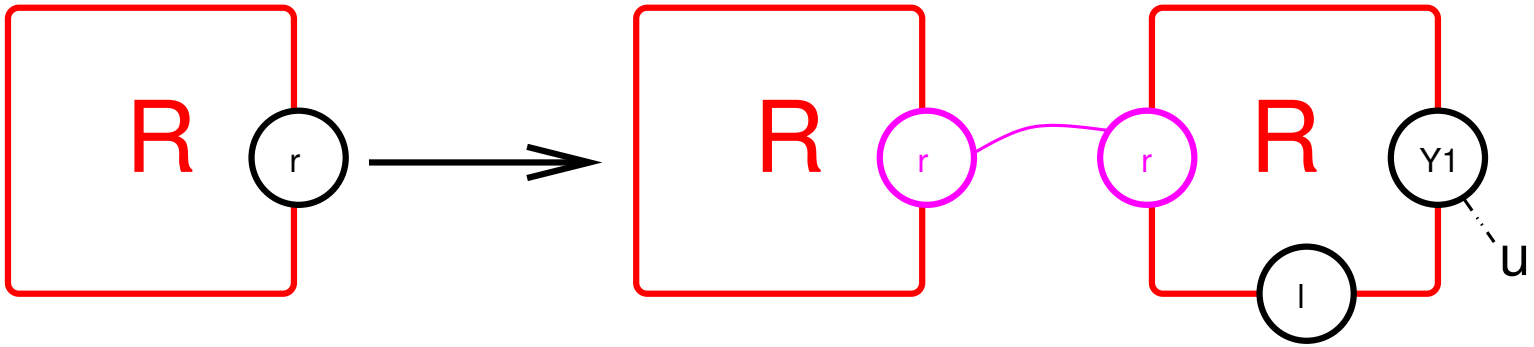


A contextual rule

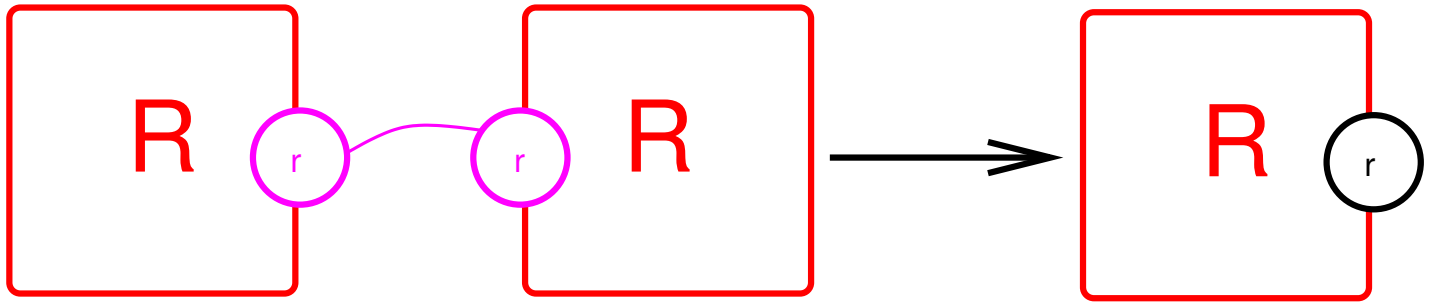


$$R(Y1{\sim}u,r!_) \rightarrow R(Y1{\sim}p,r)$$

Creation/Suppression



$R(r) \rightarrow R(r!1), R(r!1,l,Y1)$



$R(r!1), R(r!1) \rightarrow R(r)$

Early EGF example

egf rules 1

protein shorthands: E:=egf, R:=egfr, So:=Sos,Sh:=Sh,G:=grb2
site abbreviations & fusions: Y68:=Y1068, Y48:=Y1148/73, Y7:=Y317, π:=PTB/SH2

- Ligand-receptor binding, receptor dimerisation, rtk x-phosph, & de-phosph
  - 01: R(l,r), E(r) <-> R(l^1,r), E(r^1)
  - 02: R(l^1,r), R(l^2,r) <-> R(l^1,r^3), R(l^2,r^3)
  - 03: R(r^1,Y68) -> R(r^1,$Y68^p$)
  - R($Y68^p$) -> R(Y68)
  - 04: R(r^1,Y48) -> R(r^1,$Y48^p$)
  - R($Y48^p$) -> R(Y48)

receptor type: R(l,r,Y68,Y48)

- Sh x-phosph & de-phosph
  - 14: R(r^2,$Y48^{p1}$), Sh(π^1,Y7) -> R(r^2,$Y48^{p1}$), Sh(π^1,$Y7^p$)
  - ??: Sh(π^1,$Y7^p$) -> Sh(π^1,Y7)
  - 16: Sh(π,$Y7^p$) -> Sh(π,Y7)

refined from
Sh($Y7^p$)-> Sh(Y7)

- Y68-G binding
  - 09: R($Y68^p$), G(a,b) <-> R($Y68^{p1}$)+G(a^1,b)
  - 11: R($Y68^p$), G(a,b^2) <-> R($Y68^{p1}$)+G(a^1,b^2)

refined from
R($Y68^p$)+G(a)<->R($Y68^{p1}$)+G(a^1)

Early EGF example

egf rules 2

refined from
So(d)+G(b)<->So(d^1)+G(b^1)

interface note: highlight the interacting parts

- G-So binding
  - 10: R(Y68^{p1}), G(a^1,b), So(d) <-> R(Y68^{p1}), G(a^1,b^2), So(d^2)
  - 12: G(a,b), So(d) <-> G(a,b^1), So(d^1)
  - 22: Sh(π,Y7^{p2}), G(a^2,b), So(d) <-> Sh(π,Y7^{p2}), G(a^2,b^1), S(d^1)
  - 19: Sh(π^1,Y7^{p2}), G(a^2,b), So(d) <-> Sh(π^1,Y7^{p2}), G(a^2,b^1), S(d^1)
- Y48-Sh binding
  - 13: R(Y48^{p}), Sh(π,Y7) <-> R(Y48^{p1}), Sh(π^1,Y7)
  - 15: R(Y48^{p}), Sh(π,Y7^{p}) <-> R(Y48^{p1}), Sh(π^1,Y7^{p})
  - 18: R(Y48^{p}), Sh(π,Y7^{p1}), G(a^1,b) <-> R(Y48^{p2}), Sh(π^2,Y7^{p1}), G(a^1,b)
  - 20: R(Y48^{p}), Sh(π,Y7^{p1}), G(a^1,b^3), S(d^3) <-> R(Y48^{p2}), Sh(π^2,Y7^{p1}), G(a^1,b^3), S(d^3)

refined from
R(Y48^{p})+Sh(π)<->R(Y48^{p1})+Sh(π^1)

why not simply G(b^3)??

- Sh-G binding
  - 17: R(Y48^{p1}), Sh(π^1,Y7^{p}), G(a,b) <-> R(Y48^{p1}), Sh(π^1,Y7^{p2}), G(a^2,b)
  - 21: Sh(π,Y7^{p}), G(a,b) <-> Sh(π,Y7^{p1}), G(a^1,b)
  - 23: Sh(π,Y7^{p}), G(a,b^2) <-> Sh(π,Y7^{p1}), G(a^1,b^2)
  - 24: R(Y48^{p1}), Sh(π^1,Y7^{p}), G(a,b^3), S(d^3) <-> R(Y48^{p1}), Sh(π^1,Y7^{p2}), G(a^2, b^3), S(d^3)

refined from
Sh(π), G(a)<->Sh(π^1), G(a^1)

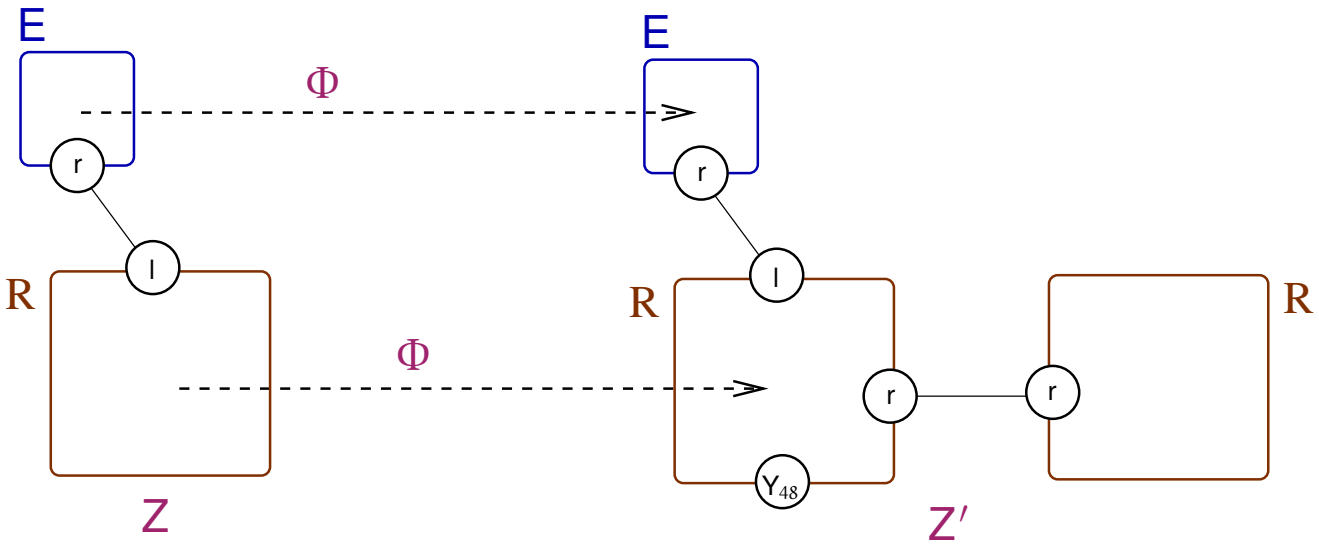
Properties of interest

1. Show the absence of modelling errors:
   - detect dead rules;
   - detect overlapping rules;
   - detect non exhaustive interactions;
   - detect rules with ambiguous molecularity.
2. Get idiomatic description of the networks:
   - capture causality;
   - capture potential interactions;
   - capture relationships between site states.
     (simplify rules)
3. Allow fast simulation:
   - capture accurate approximation of the wake-up relation.

Overview

1. Introduction
2. Kappa language
3. Local views
4. Local set of chemical species
5. Local rule systems
6. Decontextualization
7. Conclusion

Embedding



We write $Z \lhd_\Phi Z'$ iff:

- Φ is a site-graph morphism:
  - i is less specific than $\Phi(i)$,
  - if there is a link between (i, s) and (i', s'), then there is a link between $(\Phi(i), s)$ and $(\Phi(i'), s')$.
- Φ is an into map (injective):
  - $\Phi(i) = \Phi(i')$ implies that $i = i'$.

Set of reachable chemical species

Let $\mathcal{R} = \{R_i\}$ be a set of rules.
Let *Species* be the set of all chemical species ($C, c_1, c'_1, \ldots, c_k, c'_k, \ldots \in$ *Species*).
Let $Species_0$ be the set of initial .
We write:

$$c_1, \ldots, c_m \rightarrow_{R_k} c'_1, \ldots, c'_n$$

whenever:

1. there is an embedding of the lhs of R_k in the solution $c_1, \ldots, c_m$;
2. the (embedding/rule) produces the solution $c'_1, \ldots, c'_n$.

We are interested in $Species_\omega$ the set of all chemical species that can be constructed in one or several applications of rules in $\mathcal{R}$ starting from the set $Species_0$ of initial chemical species.

(We do not care about the number of occurrences of each chemical species).

Inductive definition

We define the mapping $\mathbb{F}$ as follows:

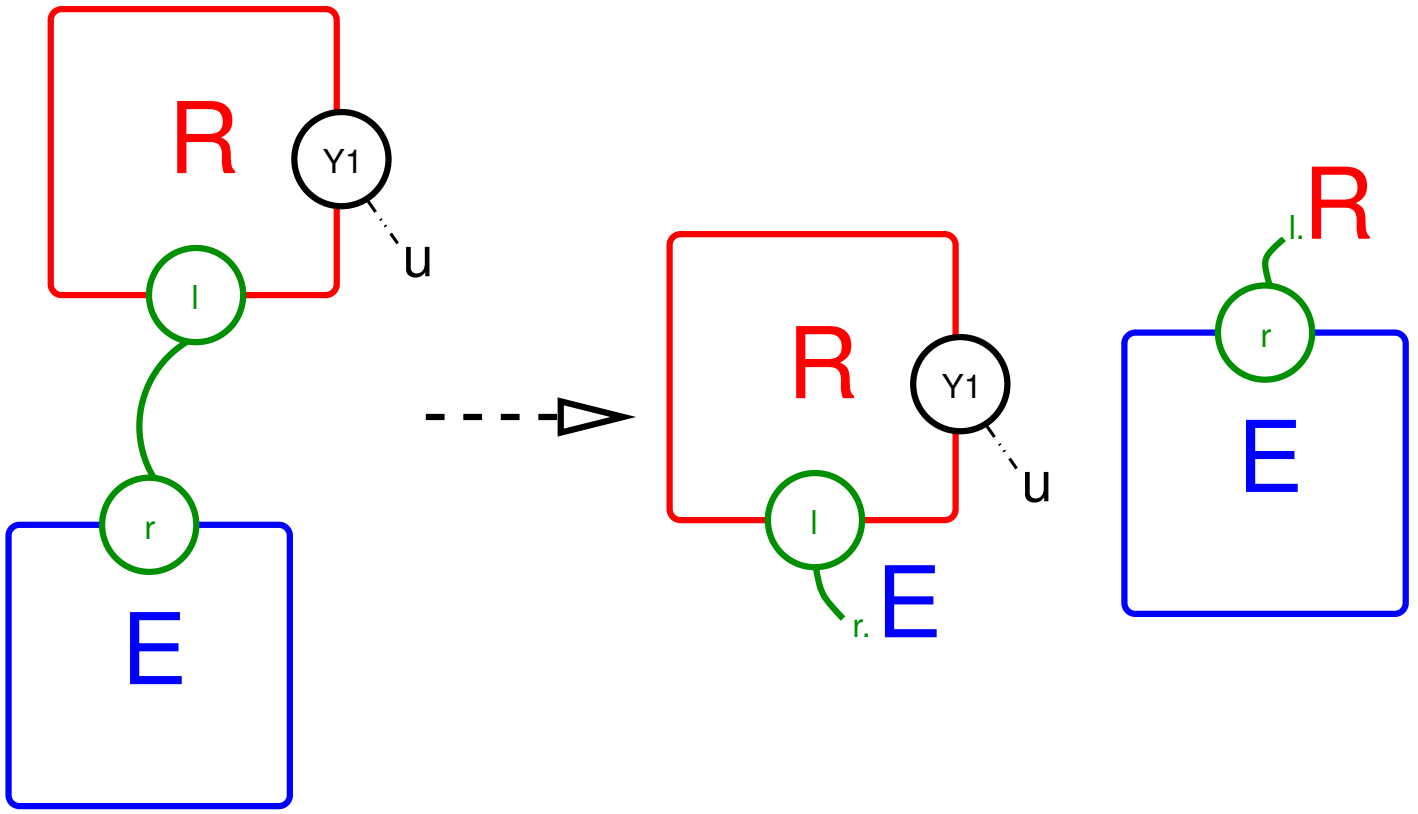
$$\mathbb{F} : \begin{cases} \wp(\mathit{Species}) & \to \wp(\mathit{Species}) \\ X & \mapsto X \cup \left\{ c'_j \;\middle|\; \begin{array}{l} \exists R_k \in \mathcal{R}, c_1, \ldots, c_m \in X, \\ c_1, \ldots, c_m \to_{R_k} c'_1, \ldots, c'_n \end{array} \right\}. \end{cases}$$

The set $\wp(\mathit{Species})$ is a complete lattice.
The mapping $\mathbb{F}$ is an extensive $\cup$-complete morphism.

We define the set of reachable chemical species as follows:

$$\mathit{Species}_\omega = \bigcup \left\{ \mathbb{F}^n(\mathit{Species}_0) \;\middle|\; n \in \mathbb{N} \right\}.$$

Local views



$\alpha(\{R(Y1\sim u,l!1), E(r!1)\}) = \{R(Y1\sim u,l!r.E); E(r!l.R)\}.$

Galois connexion

Let $\mathit{Local_view}$ be the set of all local views.

Let $\alpha \in \wp(\mathit{Species}) \to \wp(\mathit{Local_view})$ be the function that maps any set of complexes into the set of their local views.

The set $\wp(\mathit{Local_view})$ is a complete lattice.
The function α is a $\cup$-complete morphism.

Thus, it defines a Galois connexion:

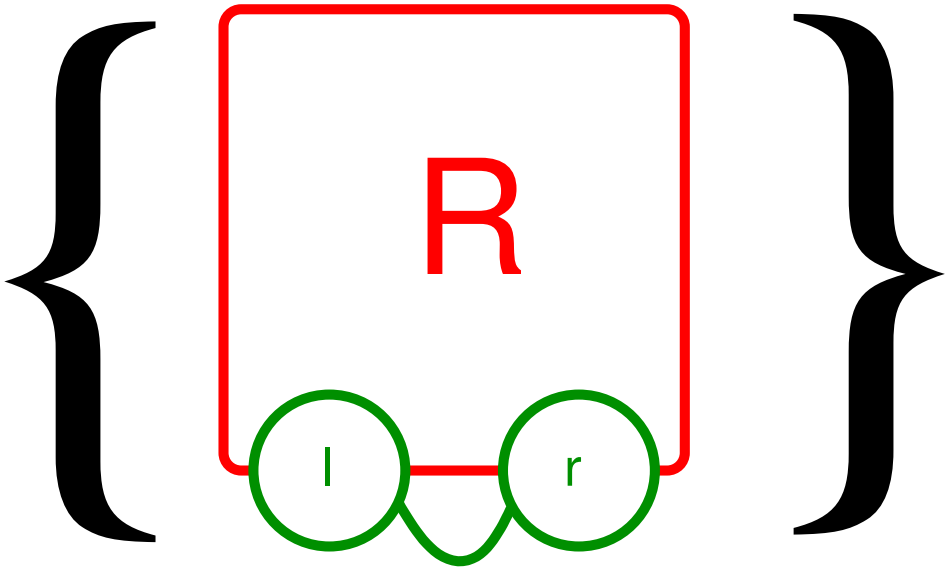
$$\wp(\mathit{Species}) \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \wp(\mathit{Local_view}).$$

(The function γ maps a set of local views into the set of complexes that can be built with these local views).

$\gamma \circ \alpha$

$\gamma \circ \alpha$ is an upper closure operator: it abstracts away some information.

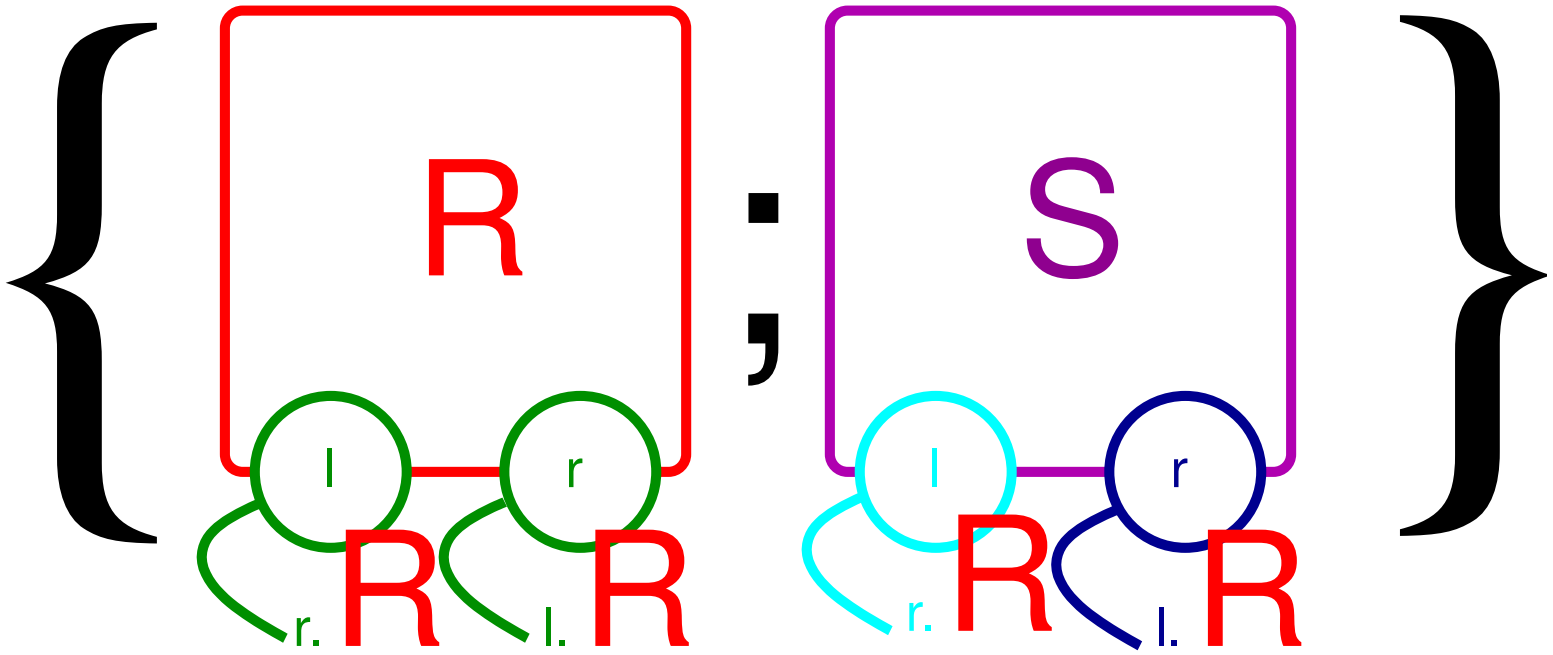
Guess the image of the following set of chemical species ?



$\alpha \circ \gamma$

$\alpha \circ \gamma$ is a lower closure operator: it simplifies (or reduces) constraints.

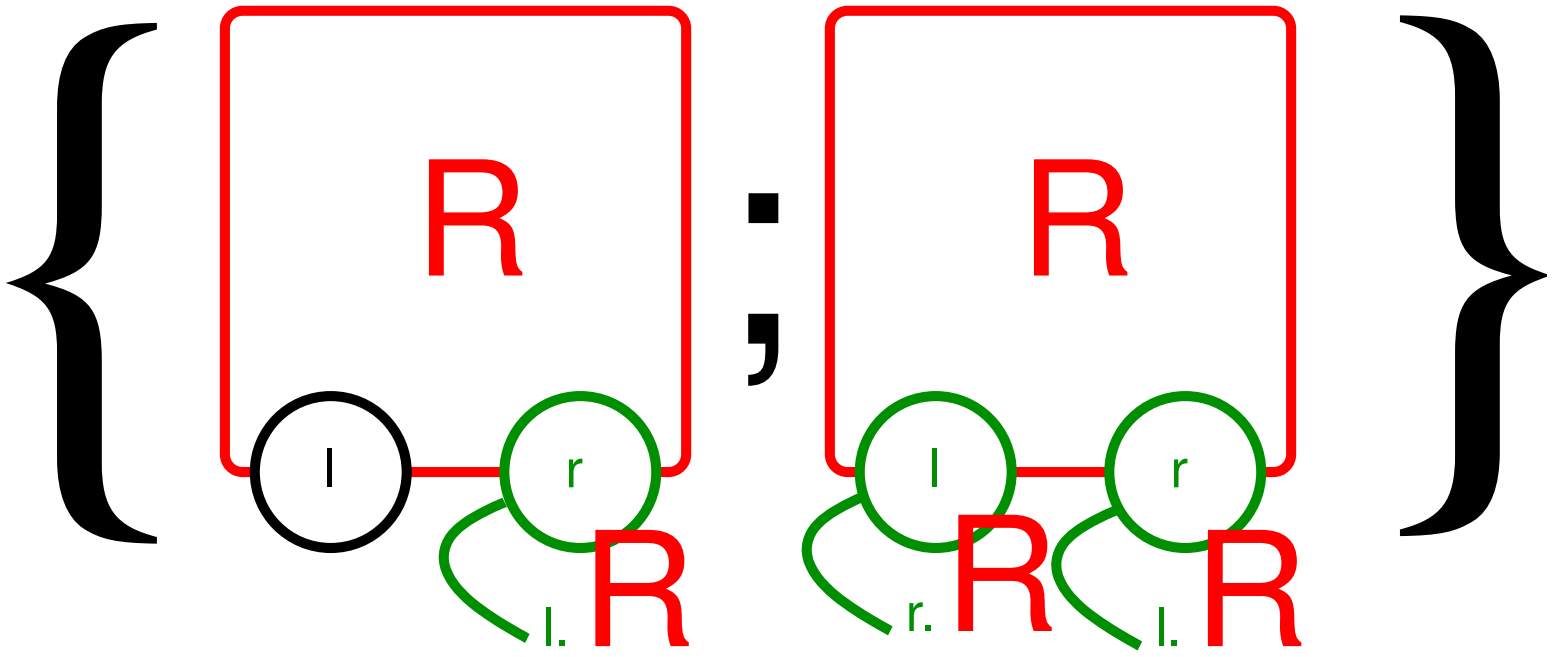
Guess the image of the following set of local views ?



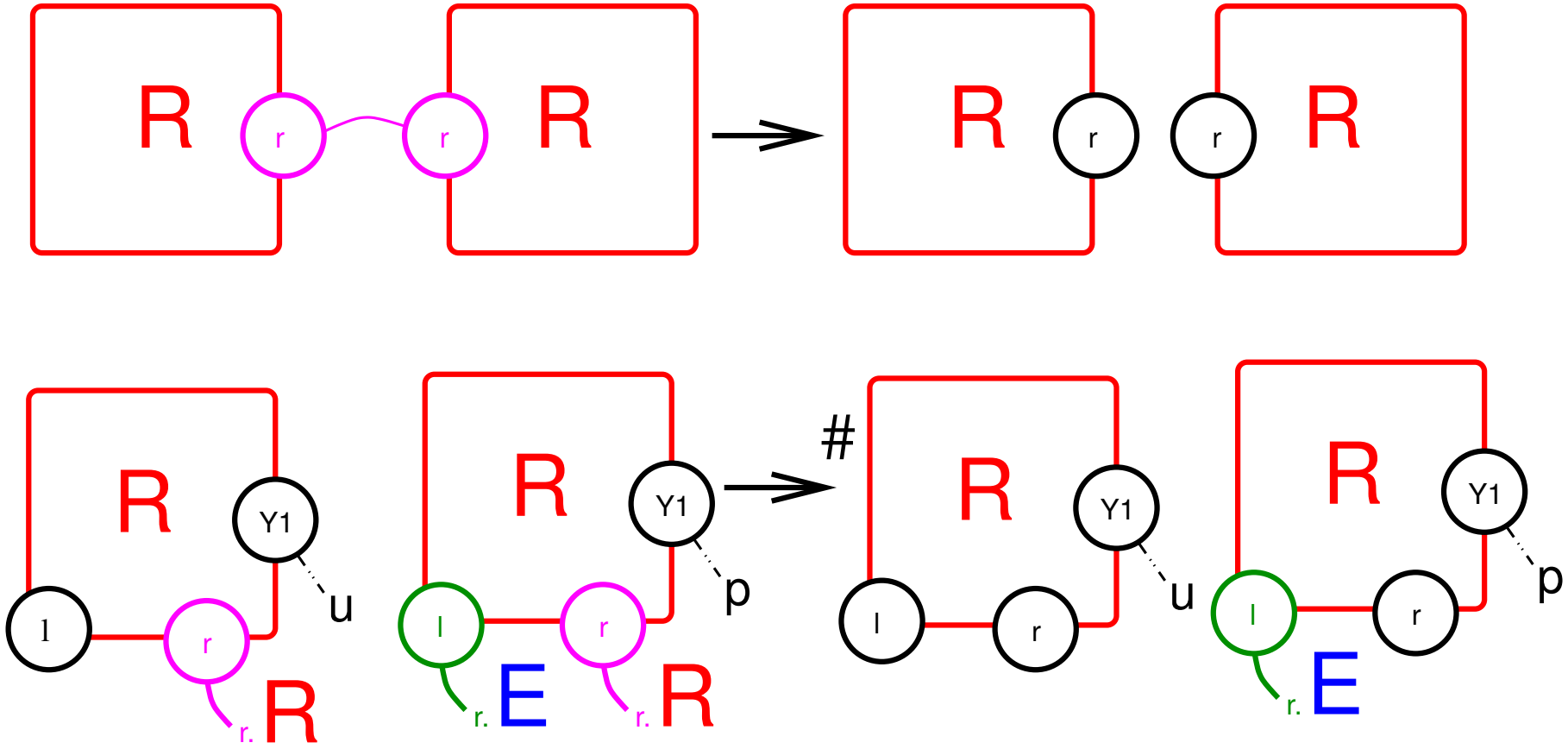
One more question

$\alpha \circ \gamma$ is a lower closure operator: it simplifies (or reduces) constraints.

Guess the image of the following set of local views ?



Abstract rules



Abstract counterpart to $\mathbb{F}$

We define $\mathbb{F}^\sharp$ as:

$$\mathbb{F}^\sharp : \begin{cases} \wp(\mathit{Local_view}) & \to \wp(\mathit{Local_view}) \\ X & \mapsto X \cup \left\{ \mathit{lv}'_j \;\middle|\; \begin{array}{l} \exists R_k \in \mathcal{R}, \mathit{lv}_1, \ldots, \mathit{lv}_m \in X, \\ \mathit{lv}_1, \ldots, \mathit{lv}_m \to^\sharp_{R_k} \mathit{lv}'_1, \ldots, \mathit{lv}'_n \end{array} \right\}. \end{cases}$$

We have:

- $\mathbb{F}^\sharp$ is extensive;
- $\mathbb{F}^\sharp$ is monotonic;
- $\mathbb{F} \circ \gamma \mathrel{\dot\subseteq} \gamma \circ \mathbb{F}^\sharp$;
- $\mathbb{F}^\sharp \circ \alpha = \alpha \circ \mathbb{F} \circ \gamma \circ \alpha$ (we will see later why).

Soundness

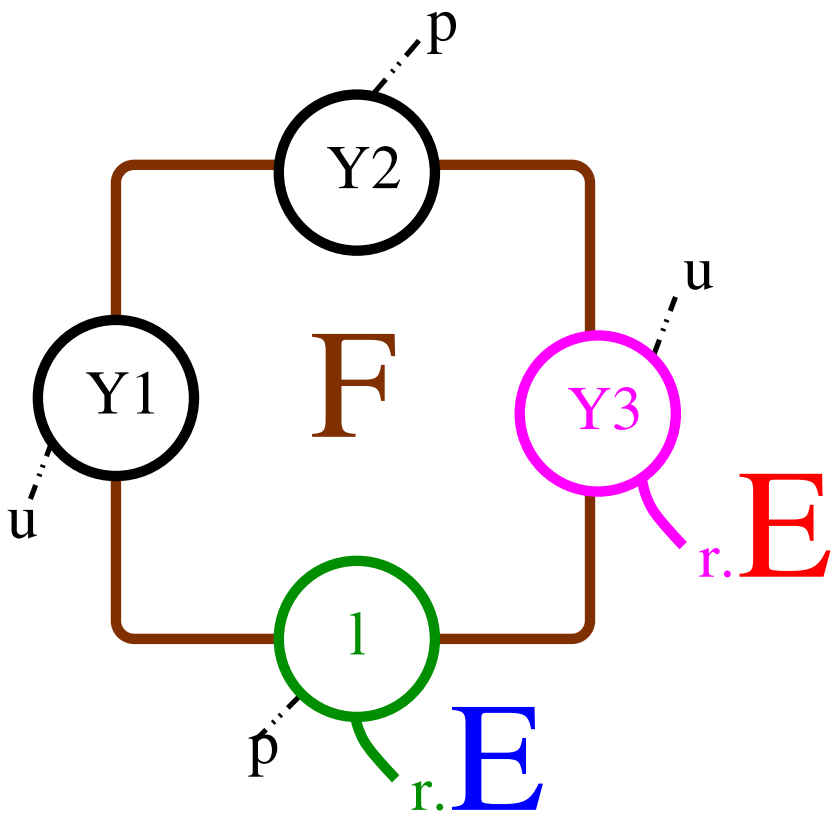
Theorem 1 Let:

1. $(D, \subseteq, \cup)$ and $(D^\sharp, \sqsubseteq, \sqcup)$ be chain-complete partial orders;
2. $D \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} D^\sharp$ be a Galois connexion;
3. $\mathbb{F} \in D \to D$ and $\mathbb{F}^\sharp \in D^\sharp \to D^\sharp$ be monotonic mappings such that: $\mathbb{F} \circ \gamma \dot{\subseteq} \gamma \circ \mathbb{F}^\sharp$;
4. $x_0 \in D$ be an element such that: $x_0 \subseteq \mathbb{F}(x_0)$;

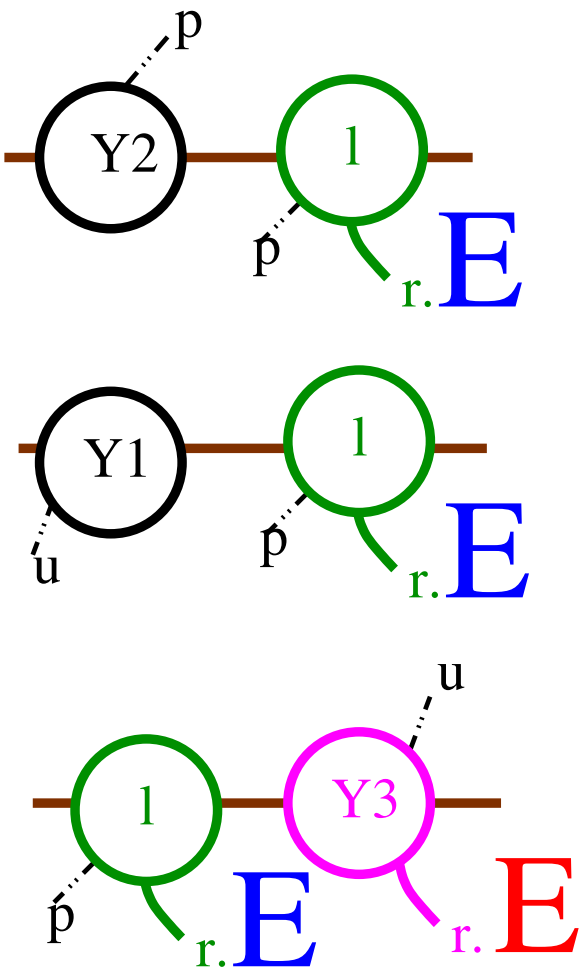
Then:

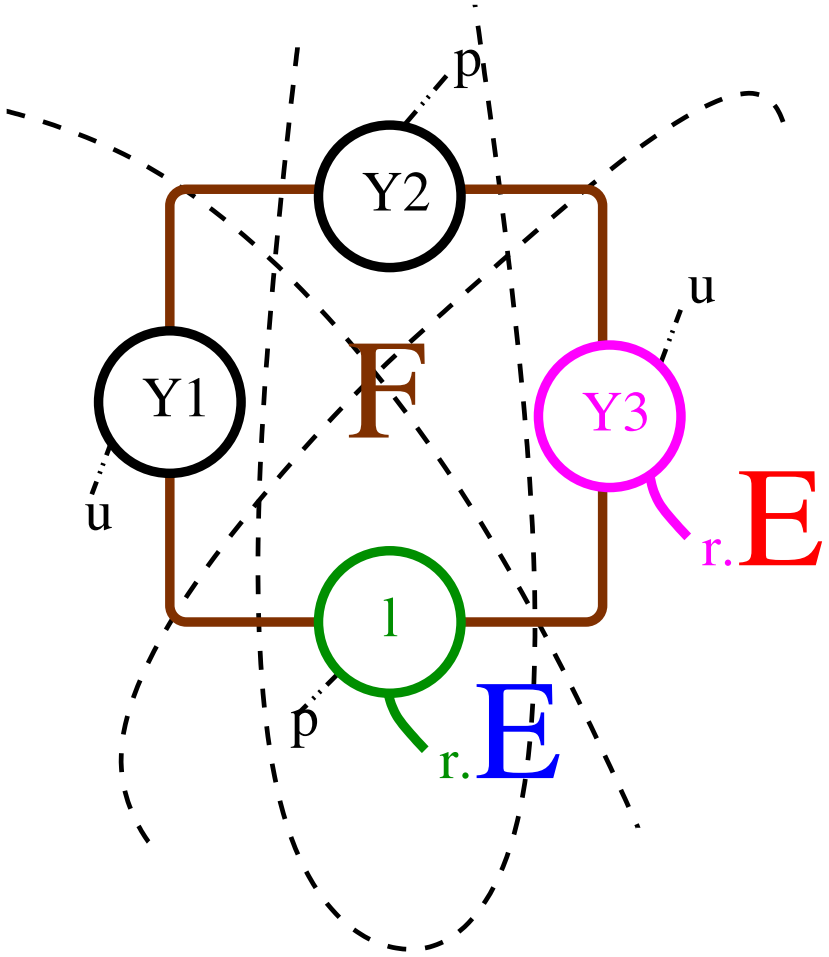
1. both $\mathit{lfp}_{x_0}\mathbb{F}$ and $\mathit{lfp}_{\alpha(x_0)}\mathbb{F}^\sharp$ exist,
2. $\mathit{lfp}_{x_0}\mathbb{F} \subseteq \gamma(\mathit{lfp}_{\alpha(x_0)}\mathbb{F}^\sharp)$.

Combinatorial blow up



Avoiding combinatorial blow up





Packing

- packing strategy: a covering P of the set of sites.
- template: a class $\mathrm{p} \in \mathrm{P}$ in the covering.
- projection: for any $\mathrm{p} \in \mathrm{P}$, the function $\Pi^{\mathrm{P}}_{\mathrm{p}}$ restricts a local view to p.
- sub-local-view: $\mathit{sub_local_view}_{\mathrm{P}} = \{\Pi^{\mathrm{P}}_{\mathrm{p}}(\mathit{lv}) \mid \mathit{lv} \in \mathit{Local_view},\ \mathrm{p} \in \mathrm{P}\}$.
- abstraction: the following function

$$\alpha_{\mathrm{P}} : \begin{cases} \wp(\mathit{Local_view}) & \to \wp(\mathit{sub_local_view}_{\mathrm{P}}) \\ \mathrm{X} & \mapsto \{\Pi_{\mathrm{p}}(\mathit{lv}) \mid \mathit{lv} \in \mathrm{X},\ \mathrm{p} \in \mathrm{P}\}. \end{cases}$$

  is a $\cup$-complete morphism.

  It defines a Galois connexion $\wp(\mathit{Local_view}) \underset{\alpha_{\mathrm{P}}}{\overset{\gamma_{\mathrm{P}}}{\leftrightarrows}} \wp(\mathit{sub_local_view}_{\mathrm{P}})$.

- counterpart: $\mathbb{F}^{\sharp}_{\mathrm{P}} = \alpha_{\mathrm{P}} \circ \mathbb{F}^{\sharp} \circ \gamma_{\mathrm{P}}$.

Automatic packing

Let P be a covering such that:

1. each site is at least in one class of P;
2. whenever a rule tests two sites a and b,
   then, for any $\mathrm{p} \in \mathrm{P}$, $\mathrm{a} \in \mathrm{p} \iff \mathrm{b} \in \mathrm{p}$;
3. whenever a rule modifies two sites a and b,
   then, for any $\mathrm{p} \in \mathrm{P}$, $\mathrm{a} \in \mathrm{p} \iff \mathrm{b} \in \mathrm{p}$;
4. whenever a rule modifies a site a and tests a site b,
   then for any $\mathrm{p} \in \mathrm{P}$, $\mathrm{a} \in \mathrm{p} \implies \mathrm{b} \in \mathrm{p}$;

Then:

$$\mathbb{F}^\sharp(\gamma_{\mathrm{P}}(\wp(\mathit{sub_local_view}_{\mathrm{P}}))) \subseteq \gamma_{\mathrm{P}}(\wp(\mathit{sub_local_view}_{\mathrm{P}}))$$

Relative completeness

Theorem 2 We suppose that:

1. $(\mathrm{D}, \subseteq, \cup)$ and $(\mathrm{D}^\sharp, \sqsubseteq, \sqcup)$ are chain-complete partial orders;
2. $(\mathrm{D}, \subseteq) \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} (\mathrm{D}^\sharp, \sqsubseteq)$ is a Galois connexion;
3. $\gamma(\mathrm{D}^\sharp)$ is closed by $\cup$;
4. $\mathbb{F} : \mathrm{D} \to \mathrm{D}$ is a monotonic map;
5. x_0 is a concrete element such that $x_0 \subseteq \mathbb{F}(x_0)$;
6. $\alpha \circ \mathbb{F} \circ \gamma = \mathbb{F}^\sharp$;
7. $x_0 \in \gamma(\mathrm{D}^\sharp)$;
8. $\mathbb{F}(\gamma(\mathrm{D}^\sharp)) \subseteq \gamma(\mathrm{D}^\sharp)$.

Then:

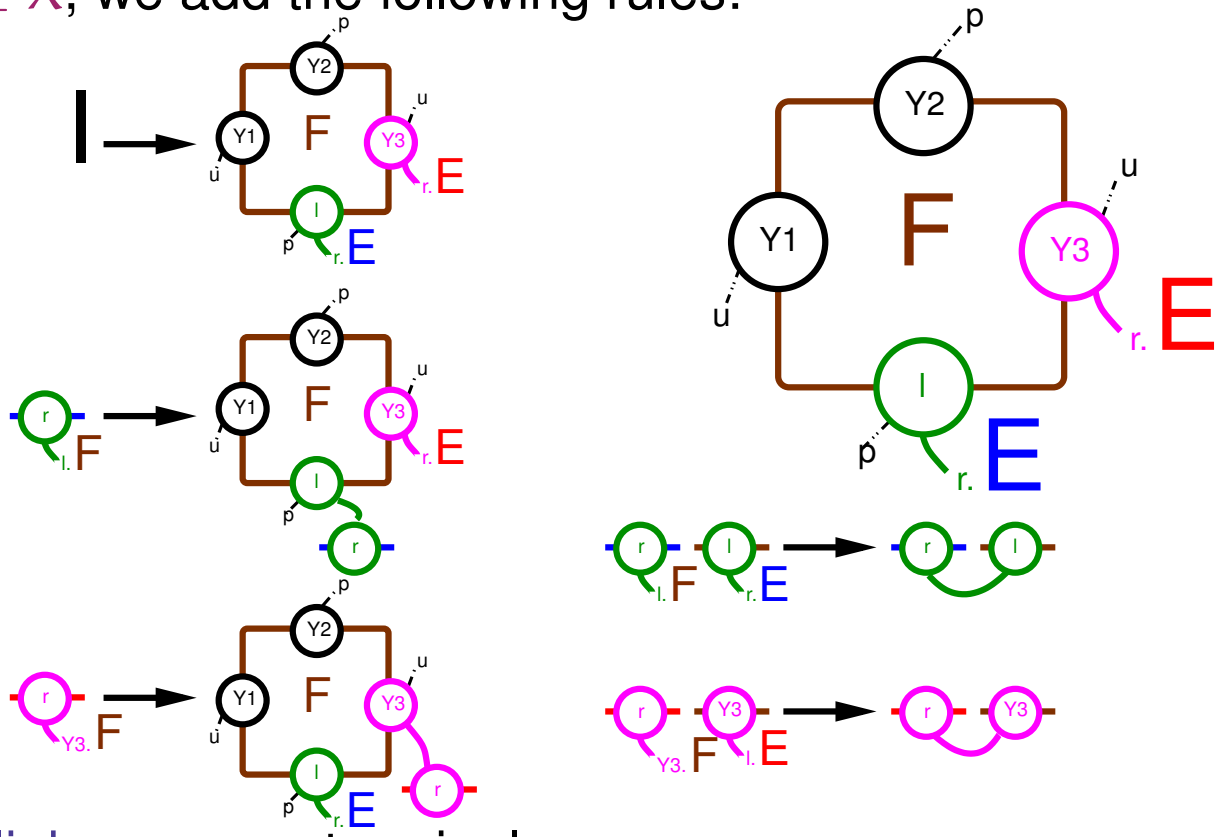
- $\mathit{lfp}_{x_0}\mathbb{F}$ and $\mathit{lfp}_{\alpha(x_0)}\mathbb{F}^\sharp$ exist;
- $\mathit{lfp}_{x_0}\mathbb{F} = \gamma(\mathit{lfp}_{\alpha(x_0)}\mathbb{F}^\sharp)$.

Overview

1. Introduction
2. Kappa language
3. Local views
4. Local set of chemical species
5. Local rule systems
6. Decontextualization
7. Conclusion

Concretization

For any $X \in \wp(Local_view)$, $\gamma(X)$ is given by a rewrite system:
For any $lv \in X$, we add the following rules:



I and semi-links are non-terminal.
I is the initial symbol.

Pumping lemma

- We use this rewrite system to enumerate the chemical species of $\gamma(X)$.
- There are two cases:
  1. either there is a finite number of rewrite sequences;
  2. or we encounter cyclic derivations
     i.e. an open chemical species with a cycle of the following form:

     R.l-r.E ... R.l-r.E

     can be built.
- We only enumerate chemical species that are reached through an acyclic rewriting computation.
- It turns out that: if $X \in \alpha(\wp(\mathit{Species}))$ then each rewrite sequence is the prefix of a terminating rewrite sequence.
  (So there is an unbounded number of species if, and only if, there is an unbounded number of rewrite sequences.)

Examples

1. Make the demo for egf
2. Make the demo for fgf
3. Make the demo for Global invariants

Counting chemical species

Given a set of local views X, we can easily count the number of species in $\gamma(X)$ by using the following lemmas:

Lemma 1 (rigidity) An embedding between two connected components is fully characterized by the image of one agent.

Lemma 2 (automorphism) If $\gamma(X)$ is finite, then for any $C \in \gamma(X)$:

- C has at most two automorphisms;
- if C has two automorphisms, then C has a bond of the form $R.r - r.R$. Moreover one automorphism swaps the two R of this bond.

Lemma 3 (Euler) If a chemical species has no cycle, then it has an agent with only one bound site.

sketch the algorithm

Which information is abstracted away ?

Our analysis is exact (no false positive):

- for EGF cascade (356 chemical species);
- for FGF cascade (79080 chemical species);

We know how to build systems with false positives. . .
. . .but they seem to be biologically meaningless.

This raises the following issues:

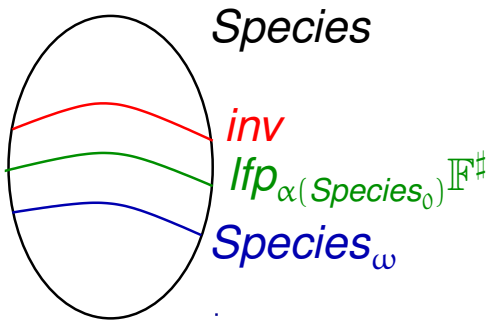
- Can we characterize which information is abstracted away ?
- Which is the form of the systems, for which we have no false positive ?
- Do we learn something about the biological systems that we describe ?

Which information is abstracted away ?

Theorem 3 We suppose that:

1. $(D, \subseteq)$ be a partial order;
2. $(D^\sharp, \sqsubseteq, \sqcup)$ be chain-complete partial order;
3. $D \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} D^\sharp$ be a Galois connexion;
4. $\mathbb{F} \in D \to D$ and $\mathbb{F}^\sharp \in D^\sharp \to D^\sharp$ are monotonic;
5. $\mathbb{F} \circ \gamma \dot{\subseteq} \gamma \circ \mathbb{F}^\sharp$;
6. $x_0, \mathit{inv} \in D$ such that:
   - $x_0 \subseteq \mathbb{F}(x_0) \subseteq \mathbb{F}(\mathit{inv}) \subseteq \mathit{inv}$,
   - $\mathit{inv} = \gamma(\alpha(\mathit{inv}))$,
   - and $\alpha(\mathbb{F}(\gamma(\alpha(\mathit{inv})))) = \mathbb{F}^\sharp(\alpha(\mathit{inv}))$;

Then, $\mathit{lfp}_{\alpha(x_0)}\mathbb{F}^\sharp$ exists and $\gamma(\mathit{lfp}_{\alpha(x_0)}\mathbb{F}^\sharp) \subseteq \mathit{inv}$.



When is there no false positive ?

Theorem 4 We suppose that:

1. $(D, \subseteq, \cup)$ and $(D^\sharp, \sqsubseteq, \sqcup)$ are chain-complete partial orders;
2. $(D, \subseteq) \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} (D^\sharp, \sqsubseteq)$ is a Galois connexion;
3. $\mathbb{F} : D \to D$ is a monotonic map;
4. x_0 is a concrete element such that $x_0 \subseteq \mathbb{F}(x_0)$;
5. $\mathbb{F} \circ \gamma \dot{\subseteq} \gamma \circ \mathbb{F}^\sharp$;
6. $\mathbb{F}^\sharp \circ \alpha = \alpha \circ \mathbb{F} \circ \gamma \circ \alpha$.

Then:

- $\mathit{lfp}_{x_0}\mathbb{F}$ and $\mathit{lfp}_{\alpha(x_0)}\mathbb{F}^\sharp$ exist;
- $\mathit{lfp}_{x_0}\mathbb{F} \in \gamma(D^\sharp) \iff \mathit{lfp}_{x_0}\mathbb{F} = \gamma(\mathit{lfp}_{\alpha(x_0)}\mathbb{F}^\sharp)$.

Local set of chemical species

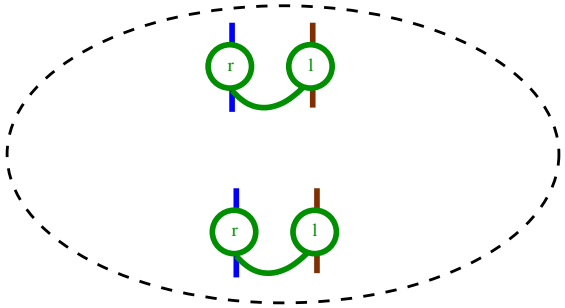
Definition 1 We say that a set $X \in \wp(Species)$ of chemical species is local if and only if $X \in \gamma(\wp(Local_view))$.

(ie. a set X is local if and only if X is exactly the set of all the species that are generated by a given set of local views.)

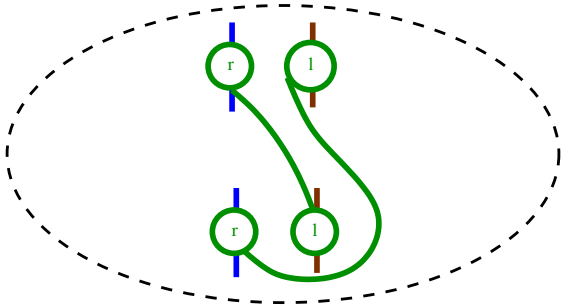
Swapping relation

We define the binary relation $\overset{\text{SWAP}}{\sim}$ among tuples *Species** of chemical species.
We say that $(C_1, \ldots, C_m) \overset{\text{SWAP}}{\sim} (D_1, \ldots, D_n)$ if and only if:

$(C_1, \ldots, C_m)$ matches with



while $(D_1, \ldots, D_n)$ matches with



Swapping closure

Theorem 5 Let $X \in \wp(\mathit{Species})$ be a set of chemical species.
The two following assertions are equivalent:

1. the set $X \in \wp(\mathit{Species})$ is local;
2. for any tuples $(C_i), (D_j) \in \mathit{Species}^*$ such that:
   - $(C_i) \in X^*$,
   - and $(C_i) \overset{\text{SWAP}}{\sim} (D_j)$;

   we have $(D_j) \in X^*$.

Consequences

Let $X \in \alpha(\wp(\mathit{Species}))$ be a set of chemical species.

1. Each open chemical species C built with the local views in X is a sub-chemical species of a close chemical species C' in $\gamma(X)$.

2. When considering the rewrite system that computes $\gamma(X)$, any partial rewriting sequence can be completed in a successful one.

3. We have $\mathbb{F}^\sharp \circ \alpha = \alpha \circ \mathbb{F} \circ \gamma \circ \alpha$.

Proof (easier implication way)

If:

- $X = \gamma(\alpha(X))$,
- $(C_i) \in X^*$,
- and $(C_i) \overset{\text{SWAP}}{\sim} (D_j)$;

Then:

we have $\alpha(\{C_i\}) = \alpha(\{D_j\})$	(because $(C_i) \overset{\text{SWAP}}{\sim} (D_j)$)
and $\alpha(\{C_i\}) \subseteq \alpha(X)$	(because $(C_i) \in X^*$ and by monotonicity);
so $\alpha(\{D_j\}) \subseteq \alpha(X)$;	
so $\{D_j\} \subseteq \gamma(\alpha(X))$	(by def. of Galois connexions);
so $\{D_j\} \subseteq X$	(since $X = \gamma(\alpha(X))$);
so $(D_j) \in X^*$.	

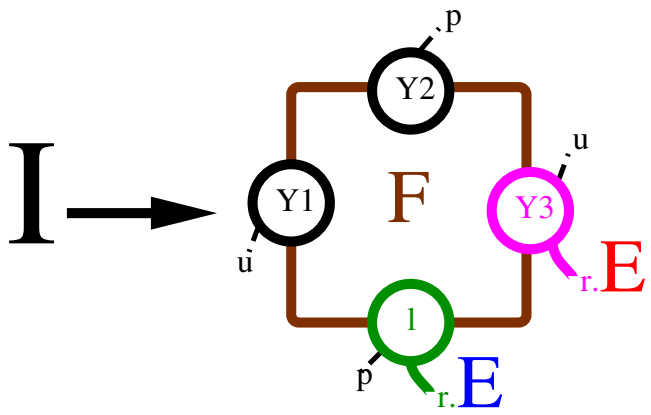
Proof (more difficult implication way)

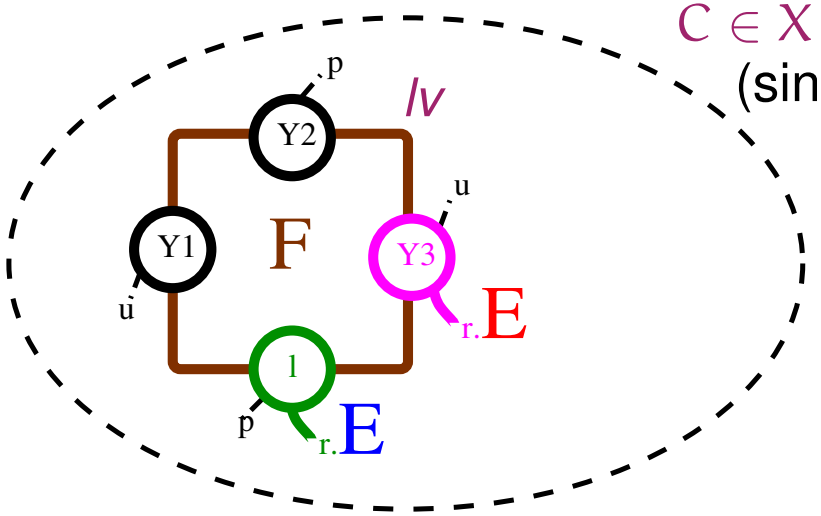
We suppose that X is close with respect to $\overset{\text{SWAP}}{\sim}$.
We want to prove that $\gamma(\alpha(X)) \subseteq X$.

We prove, by induction, that any open partial species that can be built using the rewrite system (associated with $\alpha(X)$) can be embedded in a partial species in X:

- By def. of α, this is satisfied for any local view in $\alpha(X)$;
- This remains satisfied after unfolding a semi-link with a local view;
- This remains satisfied after binding two semi-links.

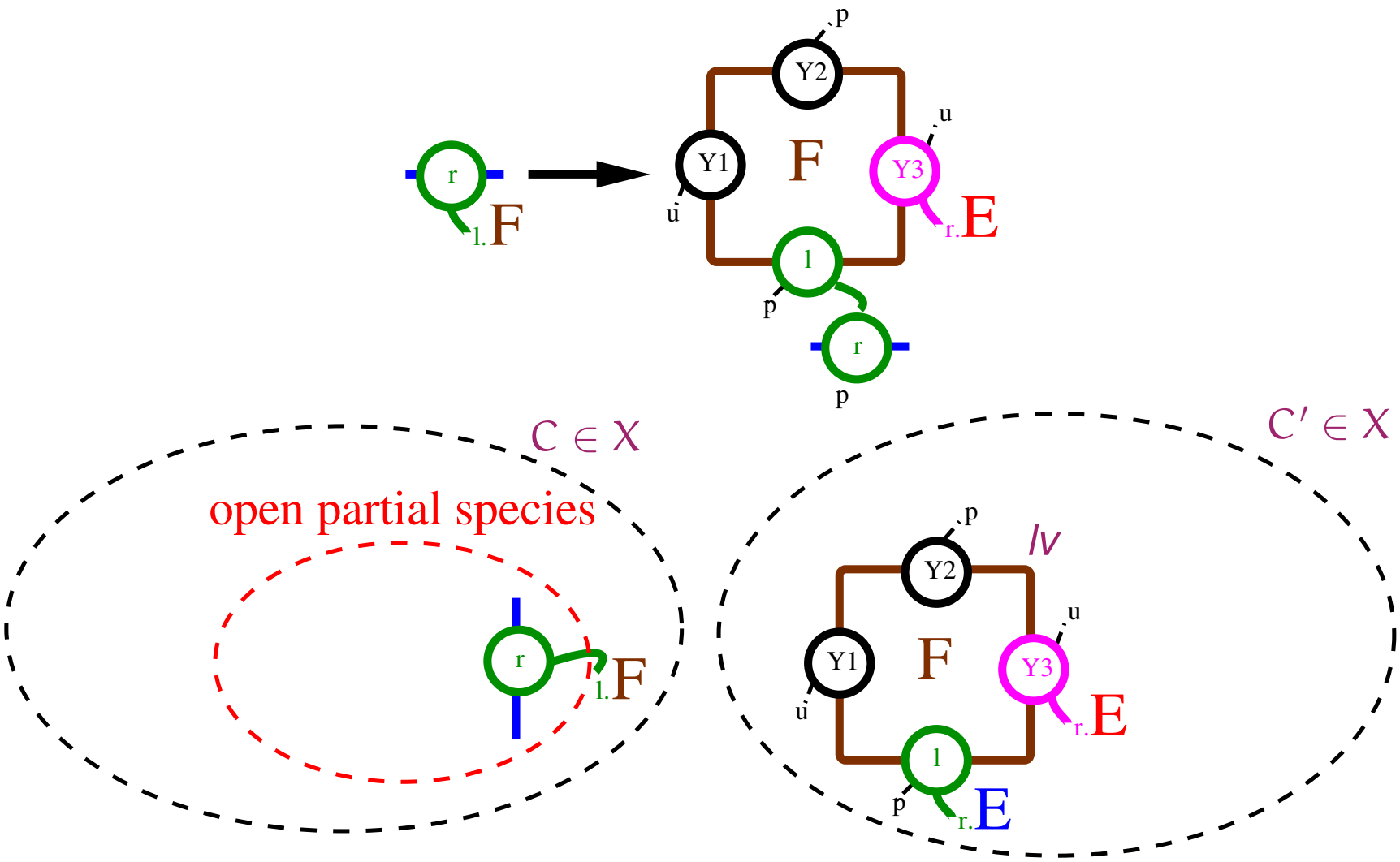
Initialization



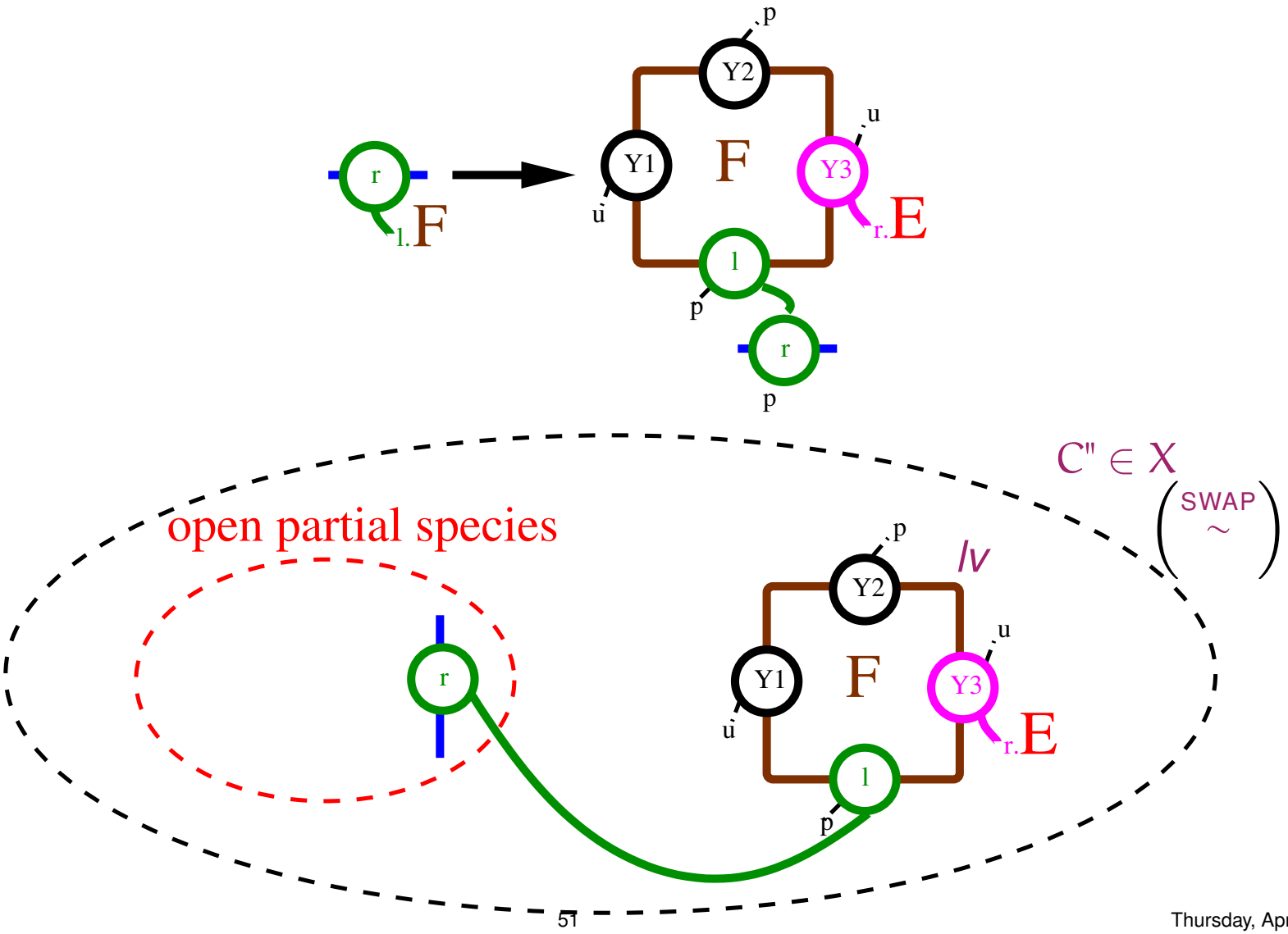


$C \in X$
(since $lv \in \alpha(X)$)

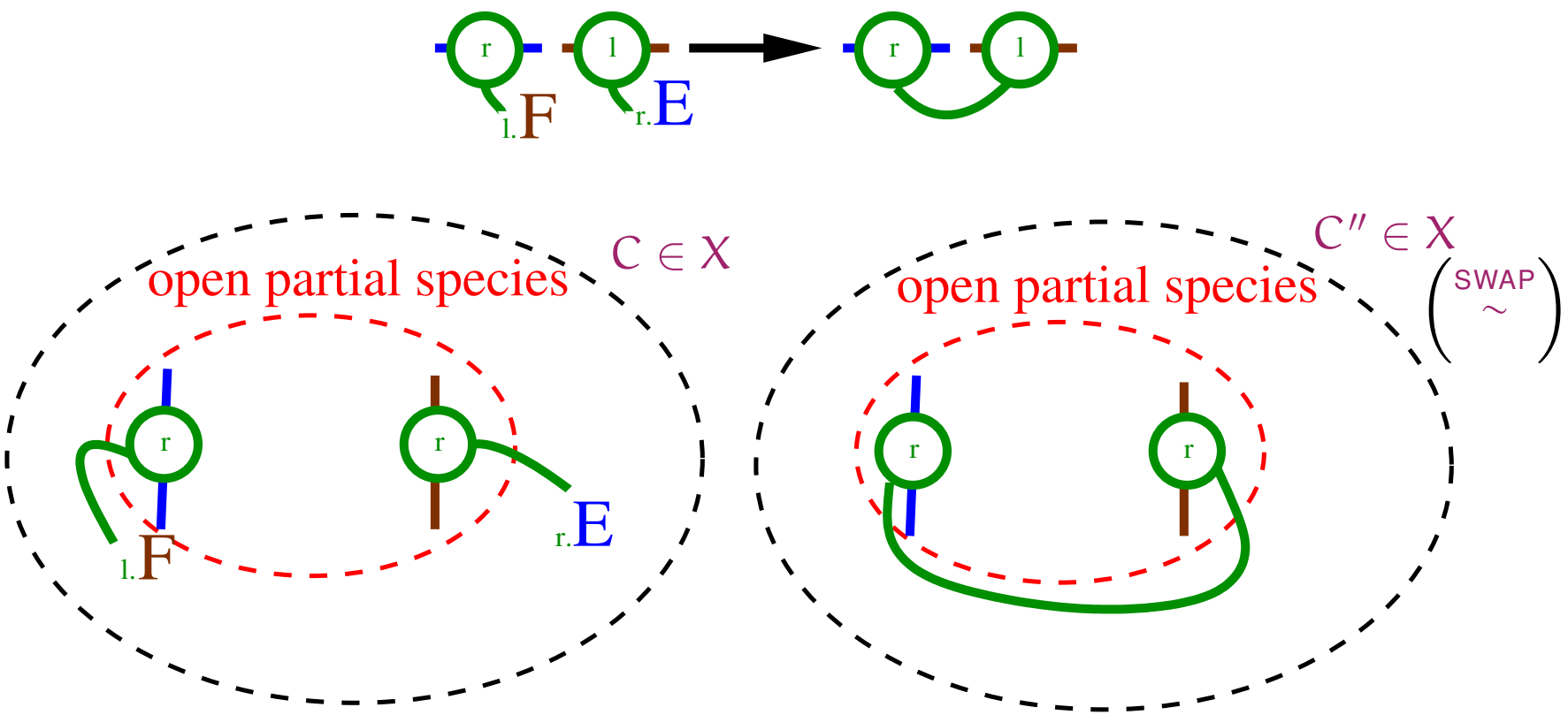
Unfolding a semi-link



Unfolding a semi-link



Binding two semi-links



Consequences

Let $\mathrm{X} \in \alpha(\wp(\mathit{Species}))$ be a set of partial species.

1. Each open partial species C built with the local views in X is a sub-partial species of a close partial species C' in $\gamma(\mathrm{X})$.

2. When considering the rewrite system that computes $\gamma(\mathrm{X})$, any partial rewriting sequence can be completed in a successful one.

3. We have $\mathbb{F}^\sharp \circ \alpha = \alpha \circ \mathbb{F} \circ \gamma \circ \alpha$.

Overview

1. Introduction
2. Kappa language
3. Local views
4. Local set of chemical species
5. Local rule systems
6. Decontextualization
7. Conclusion

Outline

We have proved that:

- if the set $\mathit{Species}_\omega$ of reachable chemical species is close with respect swapping $\overset{\text{SWAP}}{\sim}$,
- then the reachability analysis is exact (i.e. $\mathit{Species}_\omega = \gamma(\mathit{lfp}_{\alpha(\mathit{Species}_0)}\mathbb{F}^\sharp)$).

Now we give some sufficient conditions that ensure this property.

Sufficient conditions

Whenever the following assumptions:

1. initial agents are not bound;
2. rules are atomic;
3. rules are local:
   - only agents that interact are tested,
   - no cyclic patterns (neither in lhs, nor in rhs);
4. binding rules do not interfere i.e. if both:
   - A(a~m,S),B(b~n,T) $\rightarrow$ A(a~m!1,S),B(b~n!1,T)
   - and A(a~m',S'),B(b~n',T') $\rightarrow$ A(a~m'!1,S'),B(b~n'!1,T'),

   then:
   - A(a~m,S),B(b~n',T') $\rightarrow$ A(a~m!1,S),B(b~n'!1,T');
5. chemical species in $\gamma(\alpha(Species_{\omega}))$ are acyclic,

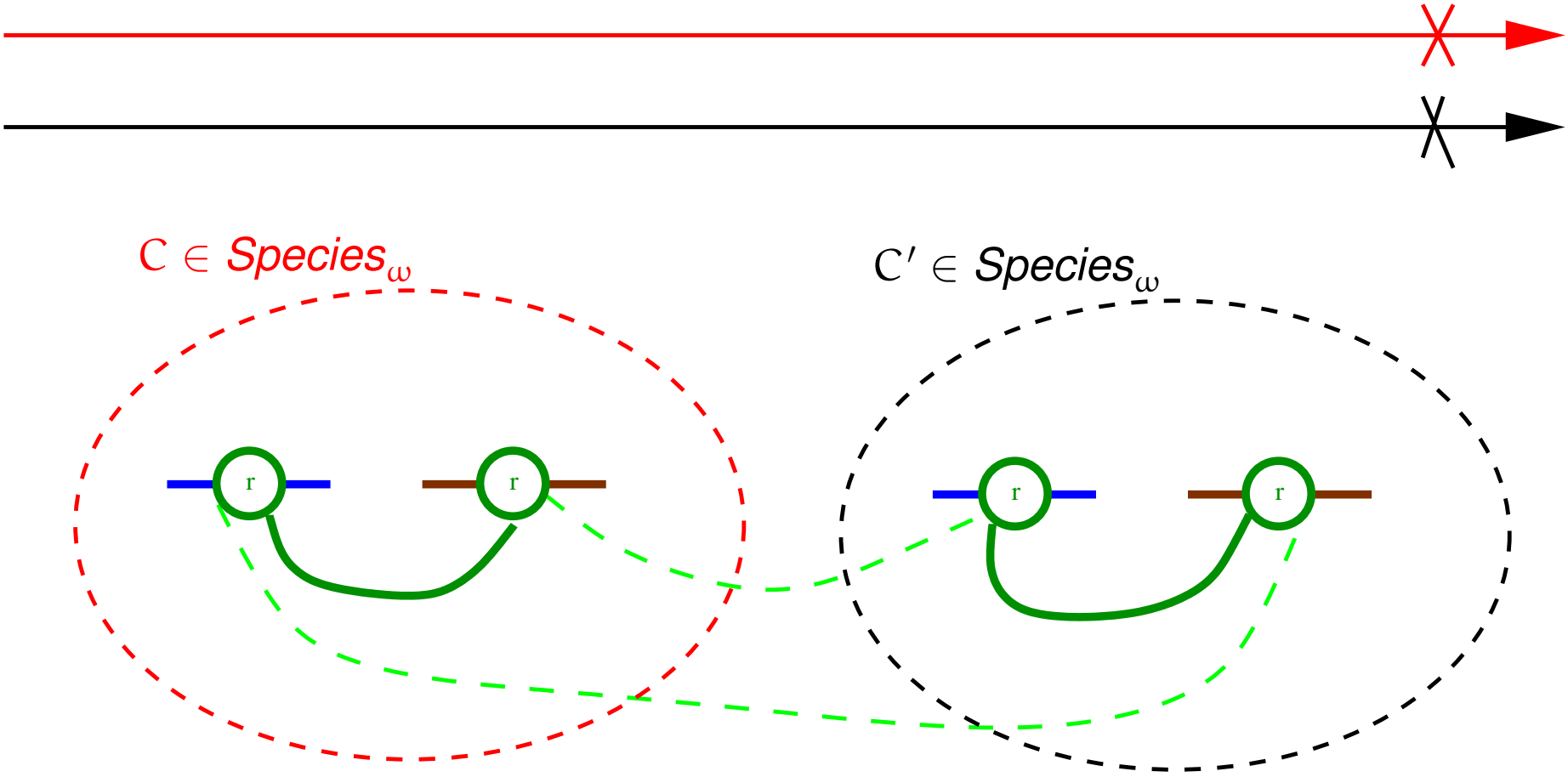
are satisfied, the set of reachable chemical species is local.

Proof outline

We sketch a proof in order to discover sufficient conditions that ensure this property:

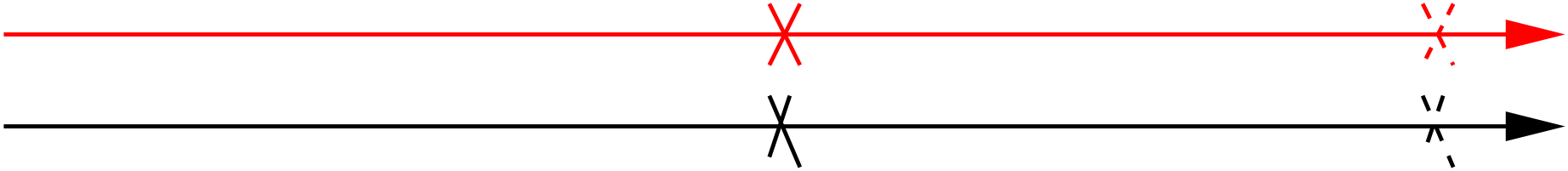
- We consider tuples of partial chemical species in which the same kind of links occur twice.
- We want to swap these links.
- We introduce the history of their computation.
- There are several cases. . .

First case (I/V)

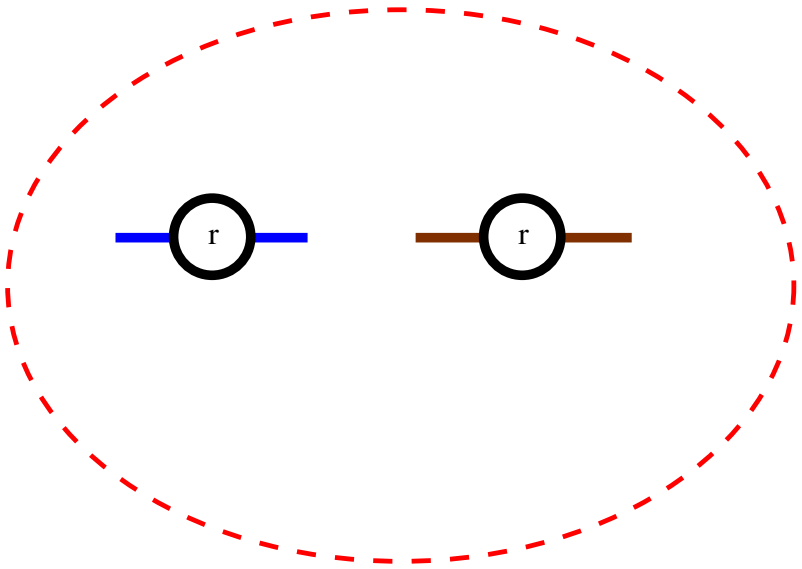


First case (II/V)

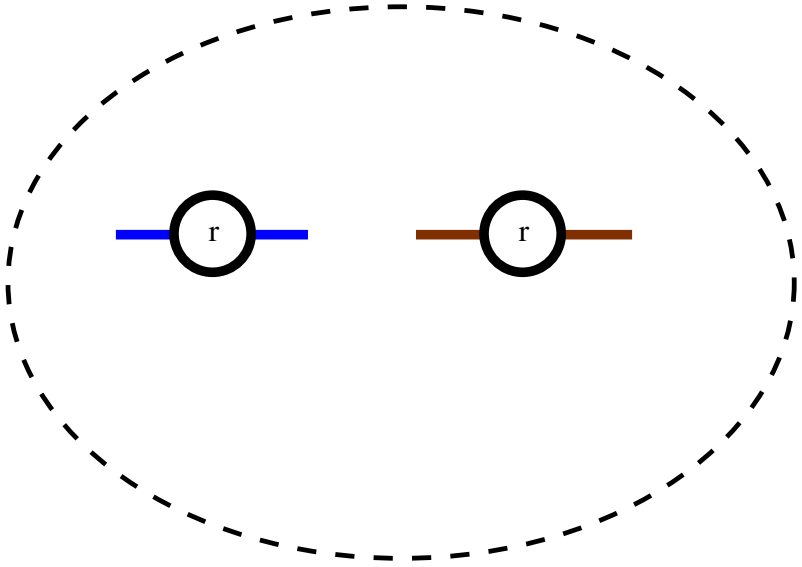
just before the links are made



$C \in Species^*_\omega$

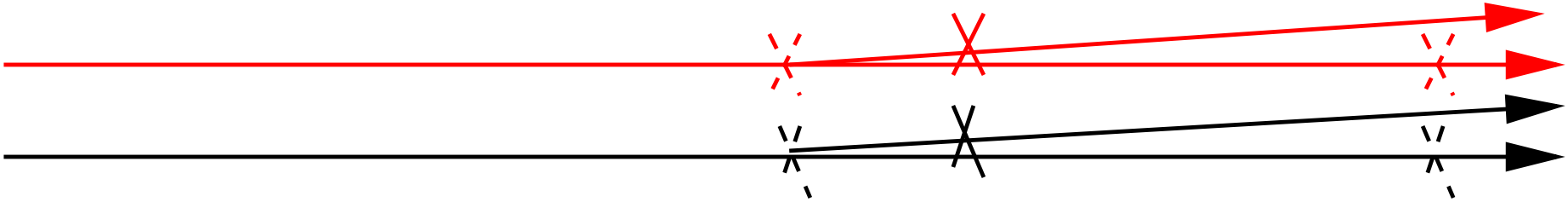


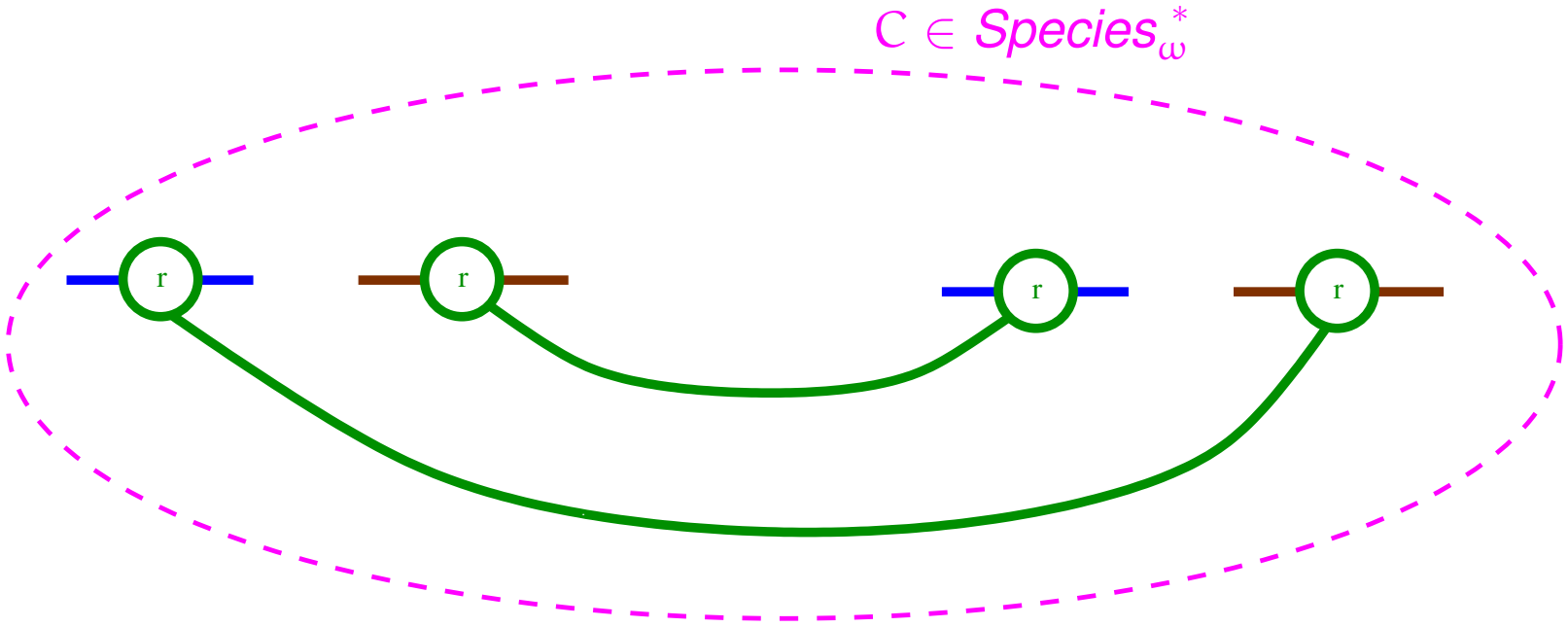
$C' \in Species^*_\omega$



First case (III/V)

we suppose we can swap the links

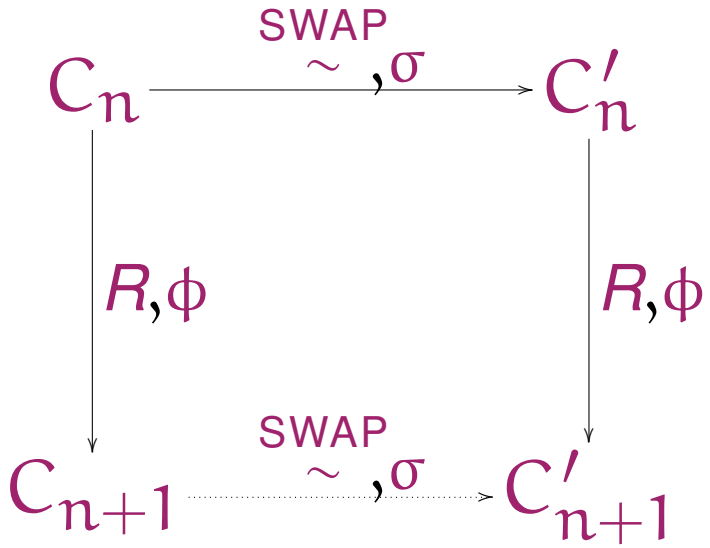




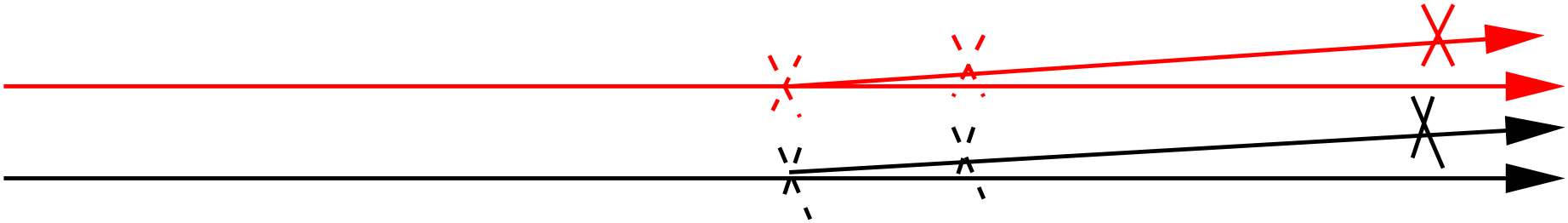
First case (IV/V)

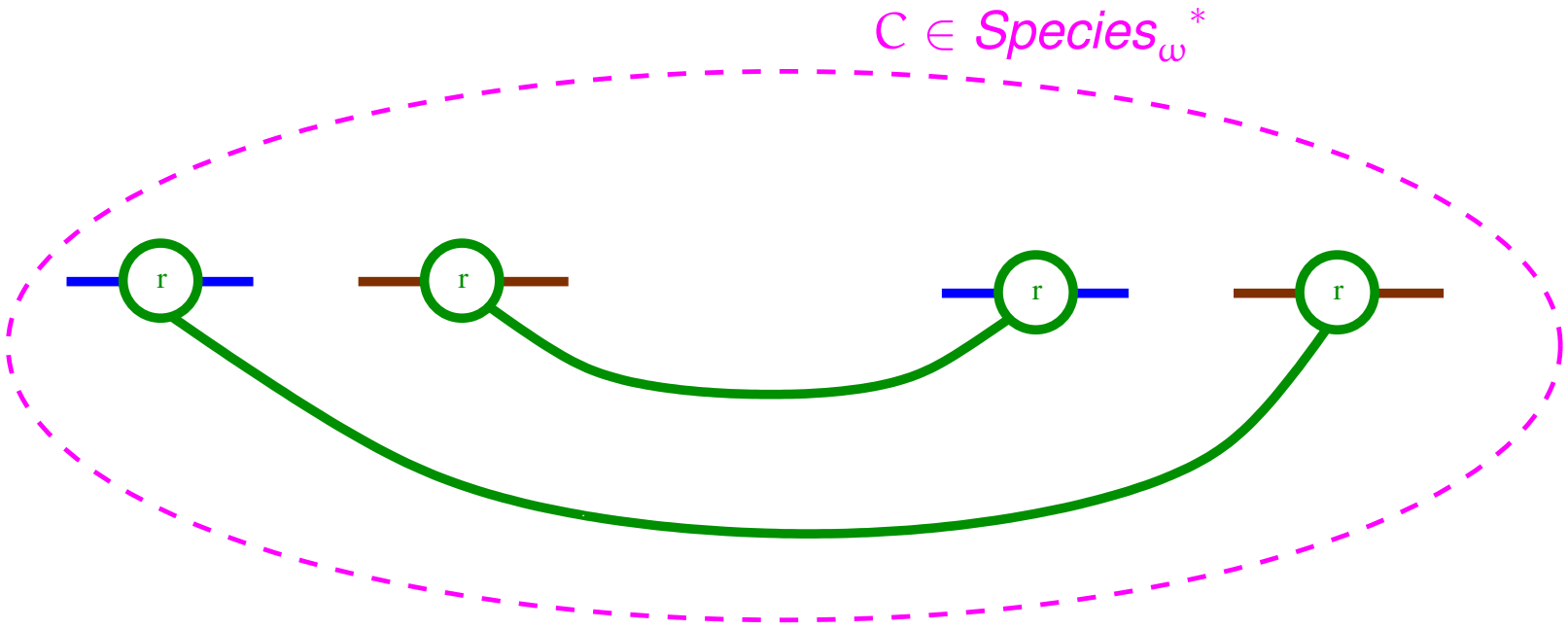
Then, we ensure that further computation steps:

- are always possible;
- have the same effect on local views;
- commute with the swapping relation $\overset{\text{SWAP}}{\sim}$.

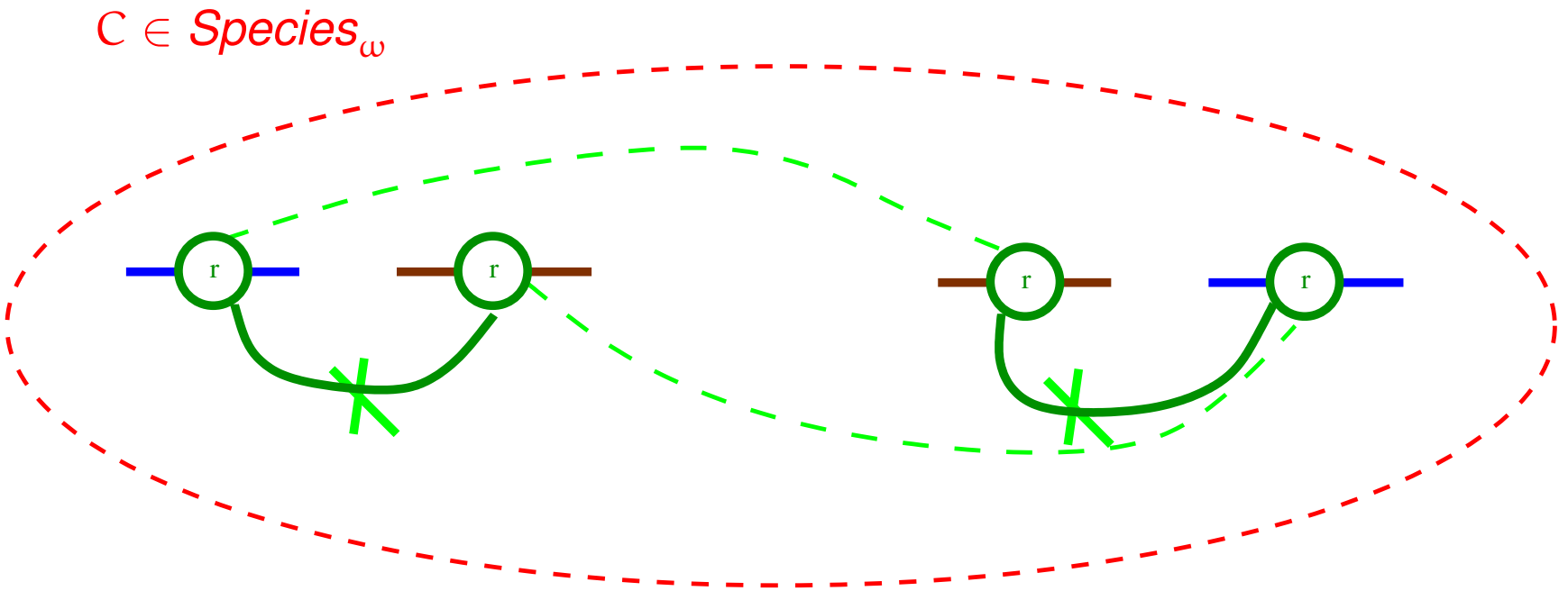


First case (V/V)



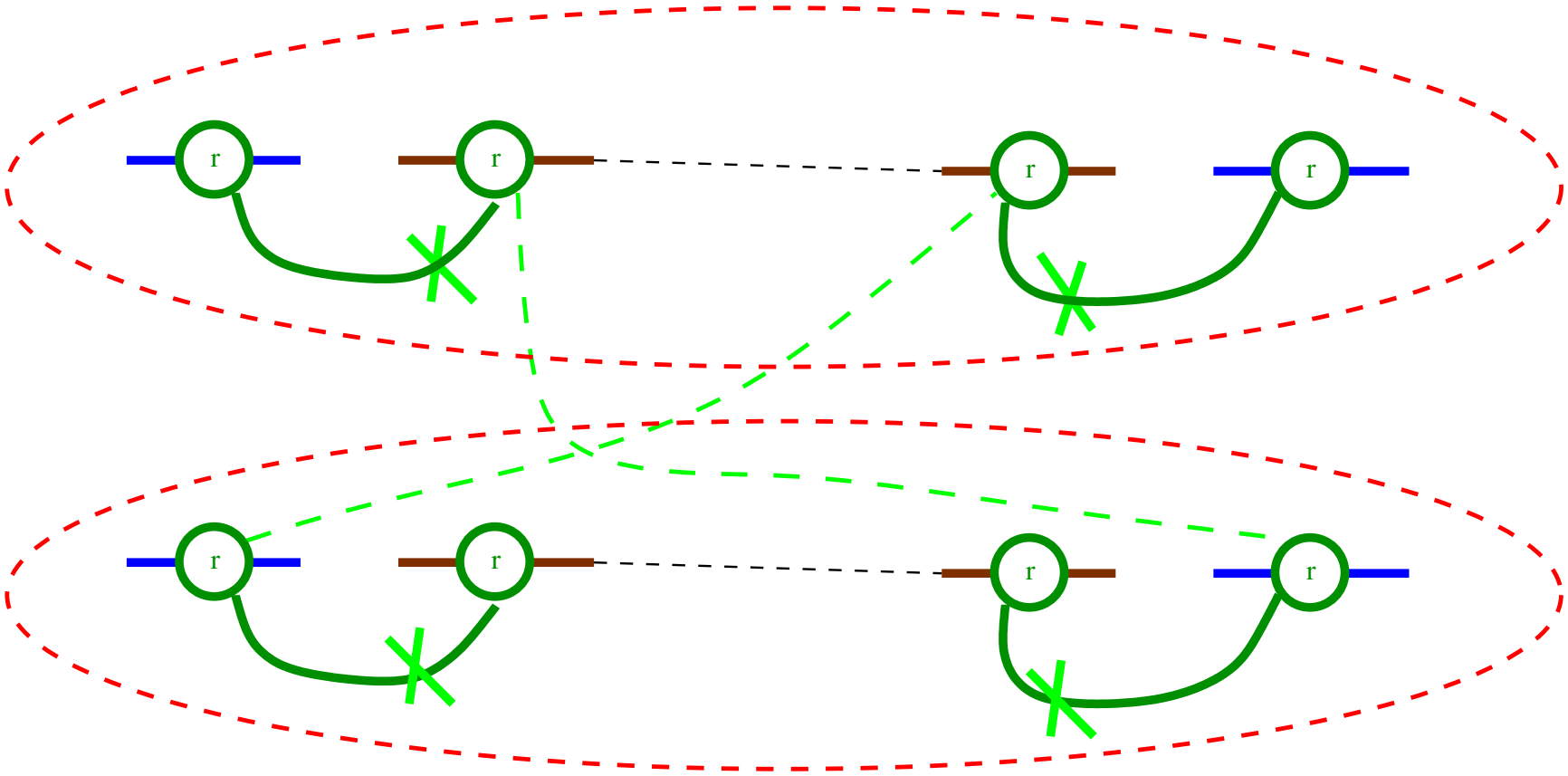


Second case (I/II)



we assume that the chemical species C is acyclic

Second case (II/II)



Sufficient conditions

Whenever the following assumptions:

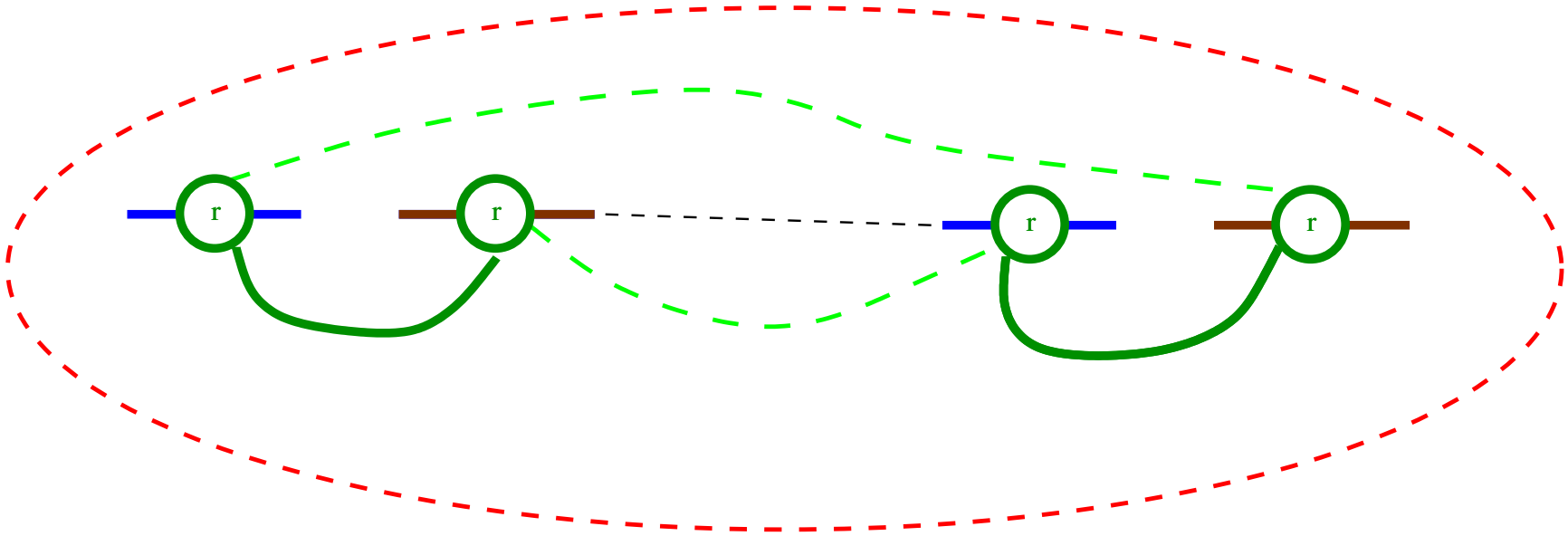
1. initial agents are not bound;
2. rules are atomic;
3. rules are local:
   - only agents that interact are tested,
   - no cyclic patterns (neither in lhs, nor in rhs);
4. binding rules do not interfere i.e. if both:
   - A(a~m,S),B(b~n,T) $\rightarrow$ A(a~m!1,S),B(b~n!1,T)
   - and A(a~m',S'),B(b~n',T') $\rightarrow$ A(a~m'!1,S'),B(b~n'!1,T'),

   then:
   - A(a~m,S),B(b~n',T') $\rightarrow$ A(a~m!1,S),B(b~n'!1,T');
5. chemical species in $\gamma(\alpha(Species_\omega))$ are acyclic,

are satisfied, the set of reachable chemical species is local.

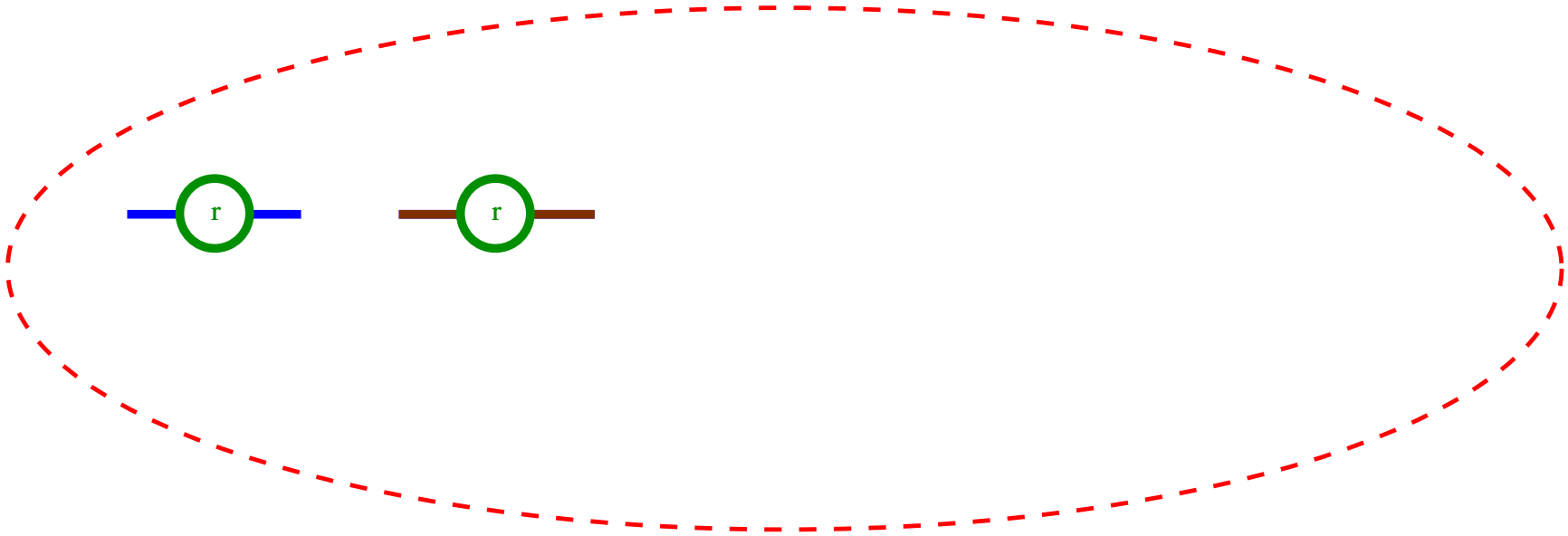
Third case (I/III)

$\mathcal{C} \in Species_\omega$

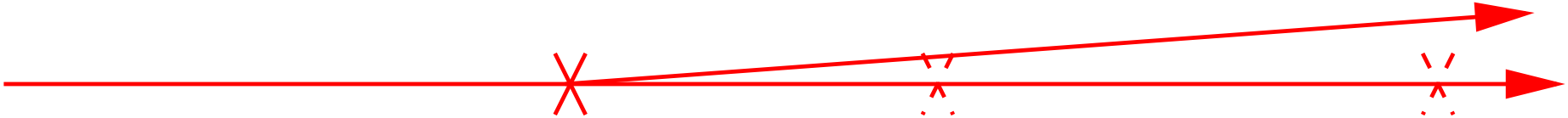


Third case (II/III)

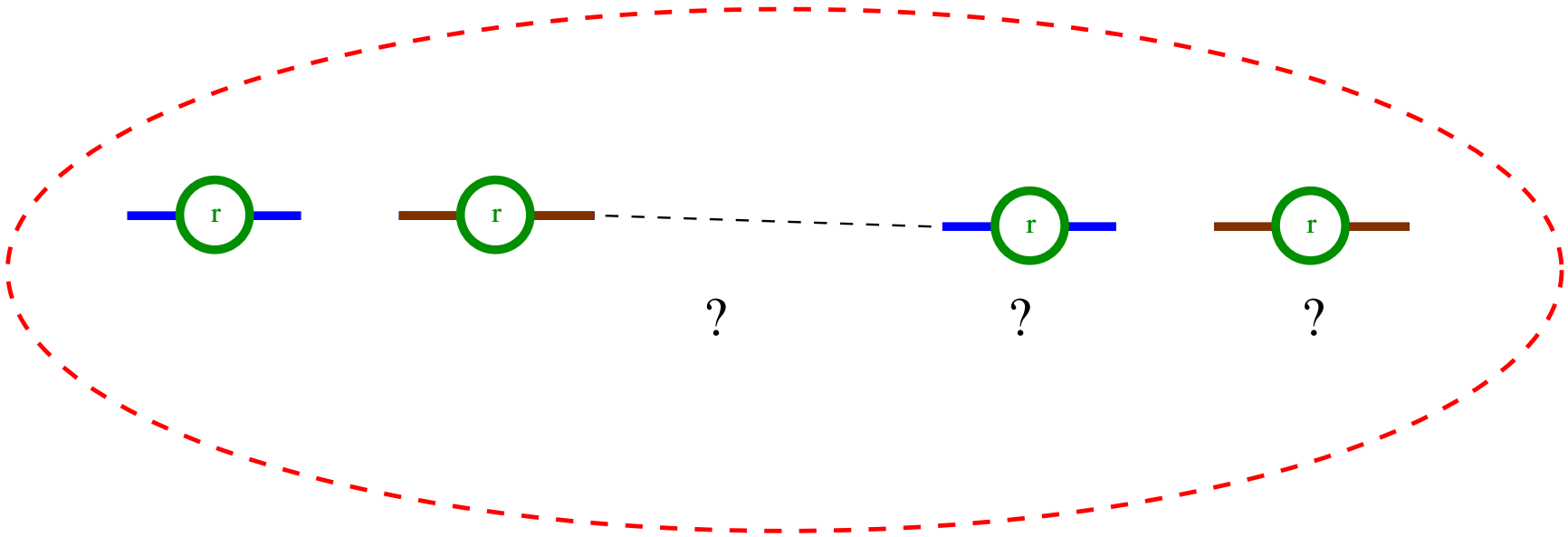
$\mathcal{C} \in Species^*_\omega$



Third case (II/III)



$C \in Species^*_\omega$



Dangerous sites

A site is dangerous if it may occur in cycles within a partial species in the set $(\gamma(\alpha(Species_\omega))$.

We would weaken the fifth requirement into:

- The binding state of a dangerous site is never tested, unless for binding or unbinding this site.
- When we bind dangerous sites, we only test that these sites are free.

Then, we prove that:

1. we can build any partial species with free dangerous sites,
2. then, we can bind them as much as we like.

Non local systems

$$Species_0 \triangleq \mathrm{R(a\sim u)}$$

$$Rules \triangleq \left\{ \begin{array}{lcl} \mathrm{R(a\sim u)} & \leftrightarrow & \mathrm{R(a\sim p)} \\ \mathrm{R(a\sim u),R(a\sim u)} & \rightarrow & \mathrm{R(a\sim u!1),R(a\sim u!1)} \\ \mathrm{R(a\sim p),R(a\sim u)} & \rightarrow & \mathrm{R(a\sim p!1),R(a\sim p!1)} \\ \mathrm{R(a\sim p),R(a\sim p)} & \rightarrow & \mathrm{R(a\sim p!1),R(a\sim p!1)} \end{array} \right\}$$

$\mathrm{R(a\sim u!1),R(a\sim u!1)} \in Species_\omega$

$\mathrm{R(a\sim p!1),R(a\sim p!1)} \in Species_\omega$

But $\mathrm{R(a\sim u!1),R(a\sim p!1)} \notin Species_\omega$.

Non local systems

$$Species_0 \triangleq \text{A(a}\sim\text{u),B(a}\sim\text{u)}$$

$$Rules \triangleq \left\{ \begin{array}{l} \text{A(a}\sim\text{u),B(a}\sim\text{u)} \rightarrow \text{A(a}\sim\text{u!1),B(a}\sim\text{u!1)} \\ \text{A(a}\sim\text{u!1),B(a}\sim\text{u!1)} \rightarrow \text{A(a}\sim\text{p!1),B(a}\sim\text{u!1)} \\ \text{A(a}\sim\text{u!1),B(a}\sim\text{u!1)} \rightarrow \text{A(a}\sim\text{u!1),B(a}\sim\text{p!1)} \end{array} \right\}$$

$\text{A(a}\sim\text{u!1),B(a}\sim\text{p!1)} \in Species_\omega$

$\text{A(a}\sim\text{p!1),B(a}\sim\text{u!1)} \in Species_\omega$

But $\text{A(a}\sim\text{p!1),B(a}\sim\text{p!1)} \notin Species_\omega$.

Non local systems

$$\mathit{Species}_0 \triangleq \mathrm{A(a{\sim}u)}$$

$$\mathit{Rules} \triangleq \left\{ \begin{array}{l} \mathrm{A(a{\sim}u)} \leftrightarrow \mathrm{A(a{\sim}p)} \\ \mathrm{A(a{\sim}u),A(a{\sim}p)} \rightarrow \mathrm{A(a{\sim}u!1),A(a{\sim}p!1)} \end{array} \right\}$$

$\mathrm{A(a{\sim}u!1),A(a{\sim}p!1)} \in \mathit{Species}_\omega$
But $\mathrm{A(a{\sim}p!1),A(a{\sim}p!1)} \notin \mathit{Species}_\omega$.

Non local systems

$$\mathit{Species}_0 \triangleq \text{R(a,b)}$$
$$\mathit{Rules} \triangleq \{\ \text{R(a,b),R(a)} \rightarrow \text{R(a,b!1),R(a!1)}\}$$

R(a,b!2),R(a!2,b!1),R(a!1,b)$\in \mathit{Species}_\omega$
But R(a!1,b!1) $\notin \mathit{Species}_\omega$.

Overview

1. Introduction
2. Kappa language
3. Local views
4. Local set of chemical species
5. Local rule systems
6. Decontextualization
7. Conclusion

Outline

- we have a syntactic criterion in order to ensure that the set of reachable chemical species of a kappa system is local ;
- we now design program transformations to help systems satisfying this criterion ;
  1. decontextualization
     - is fully automatic;
     - preserves the transition system;
     - simplifies rules thanks to reachability analysis.
  2. conjugation
     - manual;
     - preserves the set of reachable chemical species;
     - uses backtrack to add new rules.

Example

Initial rule:

R(l!2,r),R(l!1,r),E(r!1),E(r!2) $\rightarrow$ R(l!3,r!1),R(l!2,r!1),E(r!2),E(r!3)

Decontextualized rule:

R(l!_,r),R(l!_,r) $\rightarrow$ R(l!_,r!1),R(l!_,r!1)

We can remove redundant tests.

Example

Initial rules:

$$\begin{aligned}
\text{Sh(Y7}\sim\text{p!2,pi!1),G(a!2,b),R(Y48}\sim\text{p!1)} &\rightarrow \text{Sh(Y7}\sim\text{p,pi!1),G(a,b),R(Y48}\sim\text{p!1)}\\
\text{Sh(Y7}\sim\text{p!3,pi!1),G(a!3,b!2),So(d!2),R(Y48}\sim\text{p!1)} &\rightarrow \text{Sh(Y7}\sim\text{p,pi!1),G(a,b!2),So(d!2),R(Y48}\sim\text{p!1)}\\
\text{Sh(Y7}\sim\text{p!1,pi),G(a!1,b)} &\rightarrow \text{Sh(Y7}\sim\text{p,pi),G(a,b)}\\
\text{Sh(Y7}\sim\text{p!1,pi),G(a!1,b!_)} &\rightarrow \text{Sh(Y7}\sim\text{p,pi),G(a,b!_)}
\end{aligned}$$

Decontextualized rule:

$$\text{Sh(Y7!1),G(a!1)} \rightarrow \text{Sh(Y7),G(a)}$$

We can remove exhaustive enumerations.

How does it work ?

To remove a test, we prove that:

- this test is satisfied whenever the other tests are satisfied;
- or each complex that passes all tests but this one also matches with the left hand side of another rule that performs the same action.

More formally

More formally:

- Each rule R is associated with the set $S(\mathrm{R})$ of open chemical species that can match its lhs;
- Rules are gathered in equivalence classes according to the actions they perform;
- For each class $[\mathrm{R}]$, we compute:

$$\mathcal{G}([\mathrm{R}]) = \cup\{S(\mathrm{R}') \mid \mathrm{R}' \in [\mathrm{R}]\}.$$

- For each class $[\mathrm{R}]$, *Reach*$([\mathrm{R}])$ is an over approximation of the set of open chemical species that may match the lhs of a rule $\mathrm{R}' \in [\mathrm{R}]$.

A rule R may be decontextualized in a rule R' if:

$$S(\mathrm{R}') \cap \mathit{Reach}([\mathrm{R}]) \subseteq \mathcal{G}([\mathrm{R}]).$$

Decontextualization is more efficient, if the reachability analysis is accurate.

An undecontextualizable rule

Initial rule:

Sh(Y7~u,pi!1),R(Y48~p!1,r!_) -> Sh(Y7~p,pi!1),R(Y48~p!1,r!_)

Decontextualized rule:

Sh(Y7~u,pi!1),R(Y48!1,r!_) -> Sh(Y7~p,pi!1),R(Y48!1,r!_)

Conjugation

If a rule $\mathcal{R}'$ is equivalent to a rule in the transitive closure of the system.
Then it may be included in the system without modifying reachable states.
To remove the context $\mathcal{C}$ of a rule, we try to apply it for another context $\mathcal{C}'$ by:

1. removing the context $\mathcal{C}'$ (backtrack) ;
2. building the context $\mathcal{C}$;
3. applying the initial rule ;
4. removing the context $\mathcal{C}$ (backtrack) ;
5. building the context $\mathcal{C}'$.

This is proved manually.

Overview

1. Introduction
2. Kappa language
3. Local views
4. Local set of chemical species
5. Local rule systems
6. Decontextualization
7. Conclusion

Conclusion

- A scalable static analysis to abstract the reachable chemical species.
- A class of models for which the abstraction is complete.
- Many applications:
  - idiomatic description of reachable chemical species;
  - dead rule detection;
  - rule decontextualization;
  - computer-driven kinetic refinement.
- It can also help simulation algorithms:
  - wake up/inhibition map (agent-based simulation);
  - flat rule system generation (for bounded set of chemical species);
  - on the fly flat rule generation (for large/unbounded set)

Università di Bologna

Fourth day

Some notions of information flow

Jérôme Feret

Laboratoire d'Informatique de l'École Normale Supérieure

INRIA, ÉNS, CNRS

http://www.di.ens.fr/~feret

Thursday, April the 21st

Syntax

Let $\mathcal{V} \triangleq \{V, V_1, V_2, \ldots\}$ be a finite set of variables.

Let $\mathbb{Z} \triangleq \{z, \ldots\}$ be the set of relative numbers.

Expressions are polynomial of variables $\mathcal{V}$.

$$\mathrm{E} ::= z \mid \mathrm{V} \mid \mathrm{E} + \mathrm{E} \mid \mathrm{E} \times \mathrm{E}$$

Programs are given by the following grammar:

$$\begin{array}{rcl} \mathrm{P} & :== & \mathsf{skip} \\ & \mid & \mathrm{P};\mathrm{P} \\ & \mid & \mathrm{V} := \mathrm{E} \\ & \mid & \mathsf{if}\ (\mathrm{V} \geq 0)\ \{\mathrm{P}\}\ \mathsf{else}\ \{\mathrm{P}\} \\ & \mid & \mathsf{while}\ (\mathrm{V} \geq 0)\ \{\mathrm{P}\} \end{array}$$

Semantics

We define the semantics $[\![P]\!] \in ((\mathcal{V} \to \mathbb{Z}) \cup \Omega) \to ((\mathcal{V} \to \mathbb{Z}) \cup \Omega)$ of a program P:

- $[\![\text{skip}]\!](\rho) = \rho$,
- $[\![P_1;P_2]\!](\rho) = \begin{cases} \Omega & \text{if } [\![P_1]\!](\rho) = \Omega \\ [\![P_2]\!]([\![P_1]\!](\rho)) & \text{otherwise} \end{cases}$
- $[\![V := E]\!](\rho) = \begin{cases} \Omega & \text{if } \rho = \Omega \\ \rho\,[V \mapsto \overline{\rho}(E)] & \text{otherwise} \end{cases}$
- $[\![\text{if } (V \geq 0)\ \{P_1\} \text{ else } \{P_2\}]\!](\rho) = \begin{cases} \Omega & \text{if } \rho = \Omega \\ [\![P_1]\!](\rho) & \text{if } \rho(V) \geq 0 \\ [\![P_2]\!](\rho) & \text{otherwise} \end{cases}$
- $[\![\text{while } (V \geq 0)\ \{P\}]\!](\rho) = \begin{cases} \Omega & \text{if } \rho = \Omega \\ \Omega & \text{if } \{\rho' \in \mathit{Inv} \mid \rho'(V) < 0\} = \emptyset \\ \rho' & \text{if } \rho' = \{\rho' \in \mathit{Inv} \mid \rho'(V) < 0\} \end{cases}$

  where $\mathit{Inv} = \mathit{lfp}\,(X \mapsto \{\rho\} \cup \{\rho'' \mid \exists \rho' \in X,\ \rho'(V) \geq 0 \text{ and } \rho'' \in [\![P]\!](\rho')\})$.

Flow of information

Given a program P, we say that the variable V_1 flows into the variable V_2 if, and only if, the final value of V_2 depends on the initial value of V_1, which is written $V_1 \Rightarrow_P V_2$.

More formally,
$V_1 \Rightarrow_P V_2$ if and only if there exists $\rho \in \mathcal{V} \to \mathbb{Z}$, $z, z' \in \mathbb{Z}$ such that one of the following three assertions is satisfied:

1. $[\![P]\!](\rho[V_1 \mapsto z]) \neq \Omega$, $[\![P]\!](\rho[V_1 \mapsto z']) \neq \Omega$,
   and $[\![P]\!](\rho[V_1 \mapsto z])(V_2) \neq [\![P]\!](\rho[V_1 \mapsto z'])(V_2)$;
2. $[\![P]\!](\rho[V_1 \mapsto z]) = \Omega$ and $[\![P]\!](\rho[V_1 \mapsto z']) \neq \Omega$;
3. $[\![P]\!](\rho[V_1 \mapsto z]) \neq \Omega$ and $[\![P]\!](\rho[V_1 \mapsto z']) = \Omega$.

Syntactic approximation (tentative)

Let P be a program.

We define the following binary relation $\rightarrow_{\mathrm{P}}$ among variables in $\mathcal{V}$:
$V_1 \rightarrow_{\mathrm{P}} V_2$ if and only if there is an assignement in P of the form

$$V_2 := \mathrm{E}$$

such that V_1 occurs in E.

Does $V_1 \Rightarrow_{\mathrm{P}} V_2$ imply that $V_1 \rightarrow_{\mathrm{P}}^{*} V_2$?

Counter-example

We consider the following progrem P:

$$\begin{array}{rl} \mathrm{P} ::= & \text{if } (V_1 \geq 0) \\ & \quad \{V_2 := 0\} \\ & \text{else} \\ & \quad \{V_2 := 1\} \end{array}$$

For any $\rho \in \mathcal{V} \to \mathbb{Z}$,
we have $[\![\mathrm{P}]\!](\rho[V_1 \mapsto 0])(V_2) = 0$;
but, $[\![\mathrm{P}]\!](\rho[V_1 \mapsto 1])(V_2) = 1$;
so $V_1 \Rightarrow_{\mathrm{P}} V_2$;
But $V_1 \not\twoheadrightarrow_{\mathrm{P}}^{*} V_2$.

Syntactic approximation (tentative)

For each program points p in P,
we denote by *test*(p) the set of variables which occurs in the guard of the test and while loop the scope of which contains the program point p.

We define the following binary relation $\rightarrow$ among variables in $\mathcal{V}$:
$V_1 \rightarrow_P V_2$ if and only if there is an assignement in P of the form $V_2 := E$ at program point p such that:

1. either V_1 occurs in E;
2. or $V_1 \in$ *test*(p).

Does $V_1 \Rightarrow_P V_2$ imply that $V_1 \rightarrow_P^* V_2$?

Counter-example

We consider the following progrem P:

$$\mathrm{P} ::= \quad \text{while } (V_1 \geq 0) \ \{\text{skip}\}$$

For any $\rho \in \mathcal{V} \to \mathbb{Z}$,
we have $[\![\mathrm{P}]\!](\rho[V_1 \mapsto -1]) \neq \Omega$;
but, $[\![\mathrm{P}]\!](\rho[V_1 \mapsto 0]) = \Omega$;
so $V_1 \Rightarrow_{\mathrm{P}} V_2$;
But $V_1 \not\twoheadrightarrow^*_{\mathrm{P}} V_2$.

Approximation of the information flow

So as to get a sound approximation of the information flow,
we have to consider that a variable that is tested in the guard of a loop may flow in any variable.

We define the following binary relation $\rightarrow_P$ among variables in $\mathcal{V}$:
$V_1 \rightarrow V_2$ if and only if there is an assignement in P of the form $V_2 := E$ at program point p such that:

1. either V_1 occurs in E;
2. or V_1 is tested in the guard of a loop;
3. or $V_1 \in \mathit{test}(p)$.

Theorem 1 If $V_1 \Rightarrow_P V_2$, then $V_1 \rightarrow_P^* V_2$?

Limitations

The approximation is highly syntax-oriented.

- It is context-insensitive;
- It is very rough in the case of while loop,

  $\Longrightarrow$ we could show statically that some loops always terminate to avoid fictitious dependencies;
- we could detect some invariants to avoid fictitious dependencies.

Other forms of attacks could be modeled in the semantics: an attacker could observe:

- computation time;
- memory assumption;
- heating.

(attacks cannot be exhaustively specified).

Università di Bologna

Fourth day

Automatic reduction of differential semantics for protein-protein interaction networks, by abstract interpretation

[PNAS'09,LICS'10,MFPS'10,Chaos'10]

Jérôme Feret

Laboratoire d'Informatique de l'École Normale Supérieure

INRIA, ÉNS, CNRS

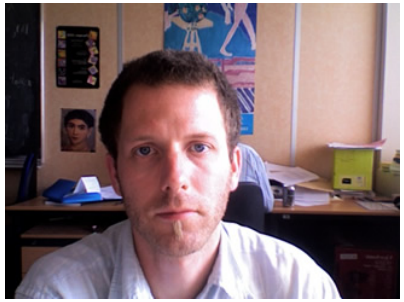
Thursday, April the 21st

Joint-work with...

Walter Fontana
Harvard Medical School

Vincent Danos
Edinburgh

Ferdinanda Camporesi
Bologna / ÉNS



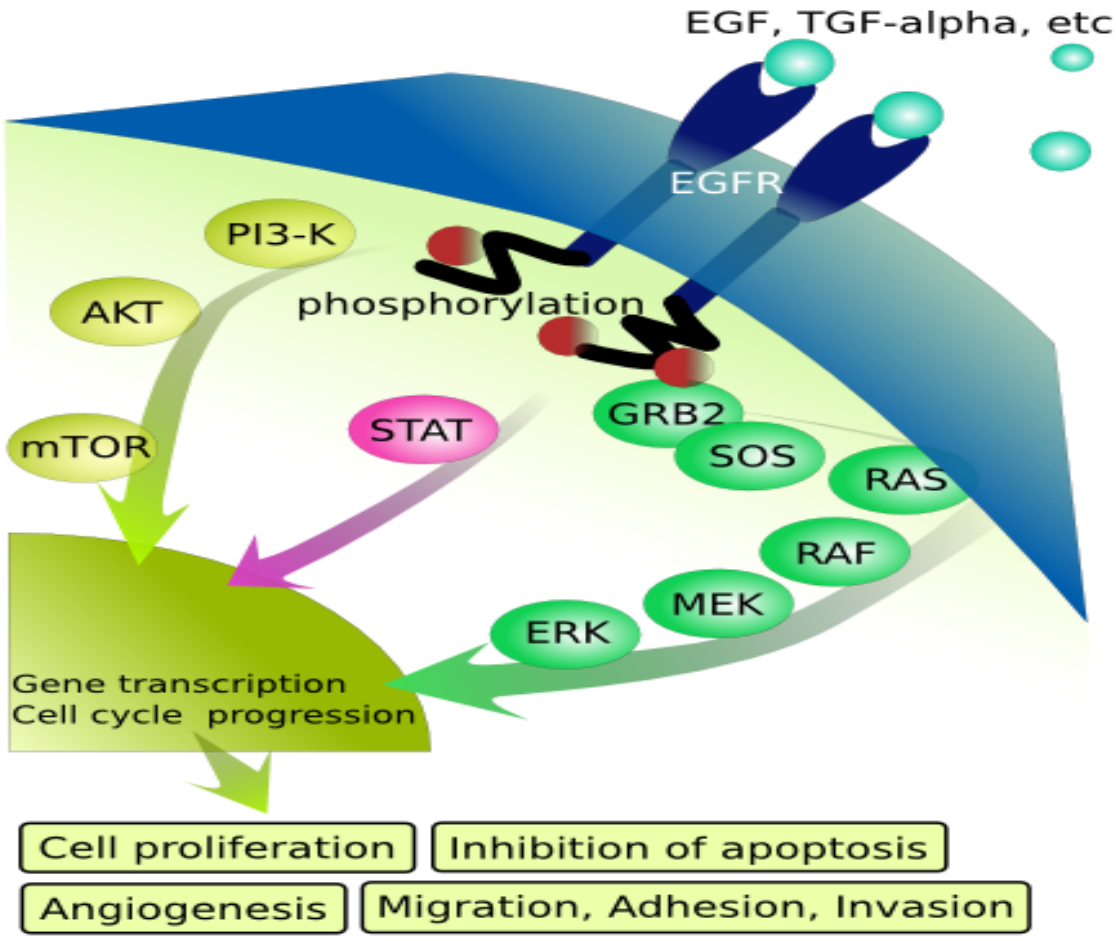
Russ Harmer
Harvard Medical School

Jean Krivine
Paris VII

Overview

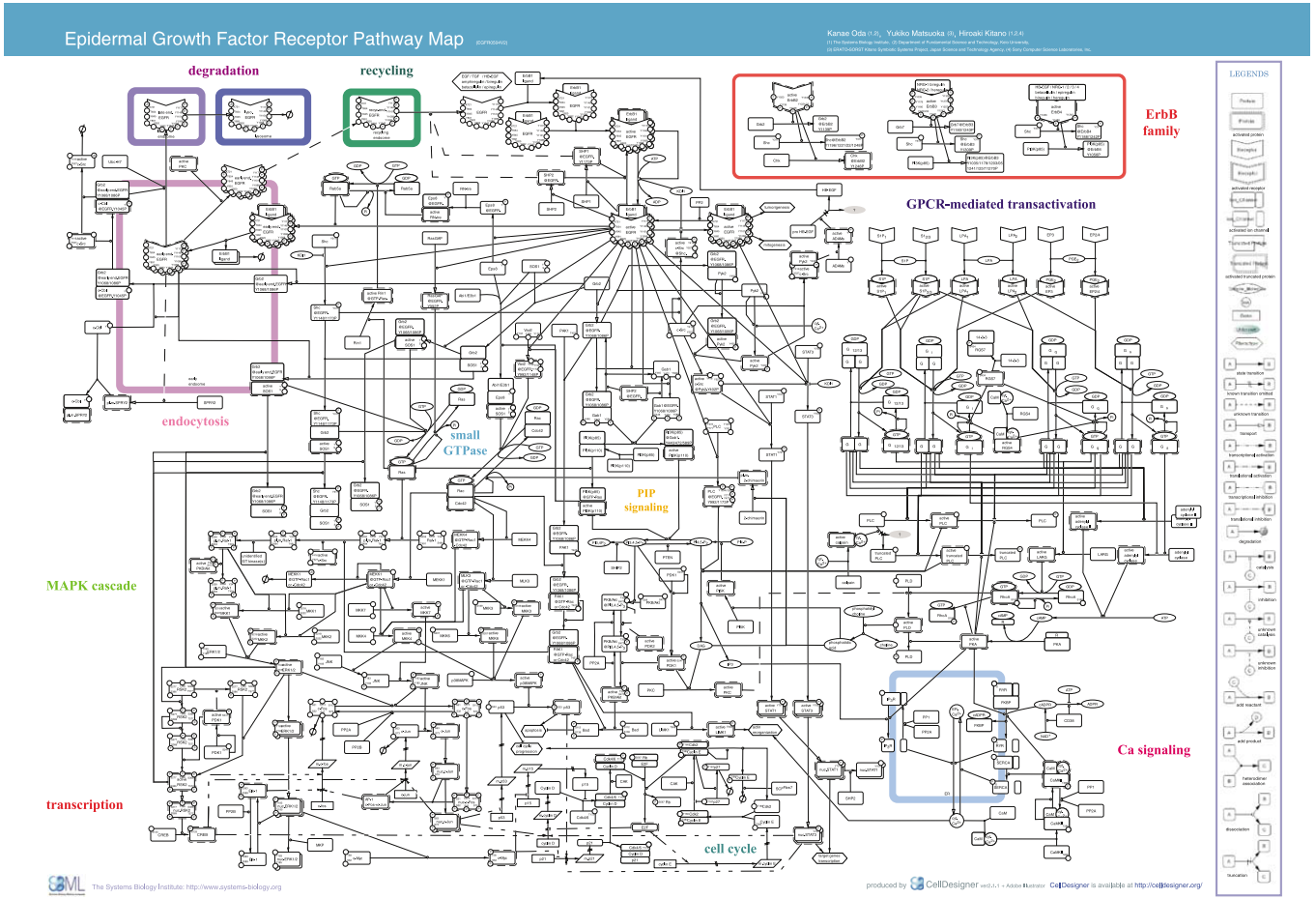
1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

Signalling Pathways



Eikuch, 2007

Pathway maps



Oda, Matsuoka, Funahashi, Kitano, Molecular Systems Biology, 2005

Differential models

$$
\begin{cases}
\frac{dx_1}{dt} = -k_1 \cdot x_1 \cdot x_2 + k_{-1} \cdot x_3 \\
\frac{dx_2}{dt} = -k_1 \cdot x_1 \cdot x_2 + k_{-1} \cdot x_3 \\
\frac{dx_3}{dt} = k_1 \cdot x_1 \cdot x_2 - k_{-1} \cdot x_3 + 2 \cdot k_2 \cdot x_3 \cdot x_3 - k_{-2} \cdot x_4) \\
\frac{dx_4}{dt} = k_2 \cdot x_3^2 - k_2 \cdot x_4 + \frac{v_4 \cdot x_5}{p_4 + x_5} - (k_3 \cdot x_4 - k_{-3} \cdot x_5) \\
\frac{dx_5}{dt} = \dots \\
\quad \vdots \\
\frac{dx_n}{dt} = -k_1 \cdot x_1 \cdot c_2 + k_{-1} \cdot x_3
\end{cases}
$$

– do not describe the structure of molecules;
– combinatorial explosion: forces choices that are not principled;
– a nightmare to modify.

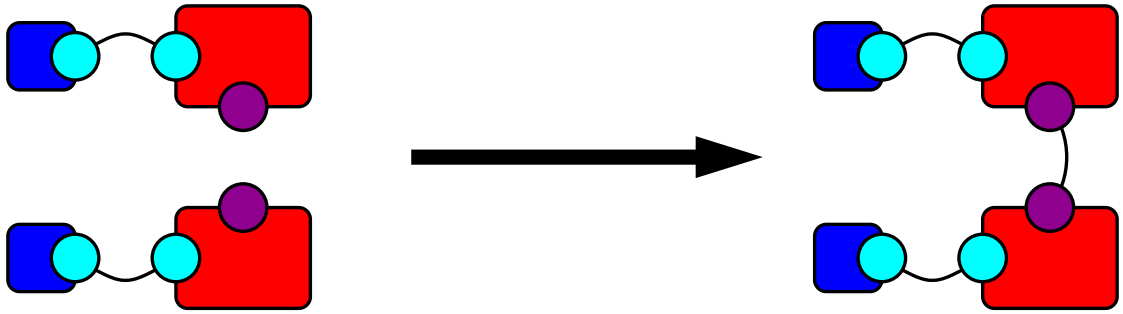
A gap between two worlds

Two levels of description:

1. Databases of proteins interactions in natural language
   + documented and detailed description
   + transparent description
   – cannot be interpreted
2. ODE-based models
   + can be integrated
   – opaque modelling process, models can hardly be modified
   – there are also some scalability issues.

Rule-based approach

We use site graph rewrite systems

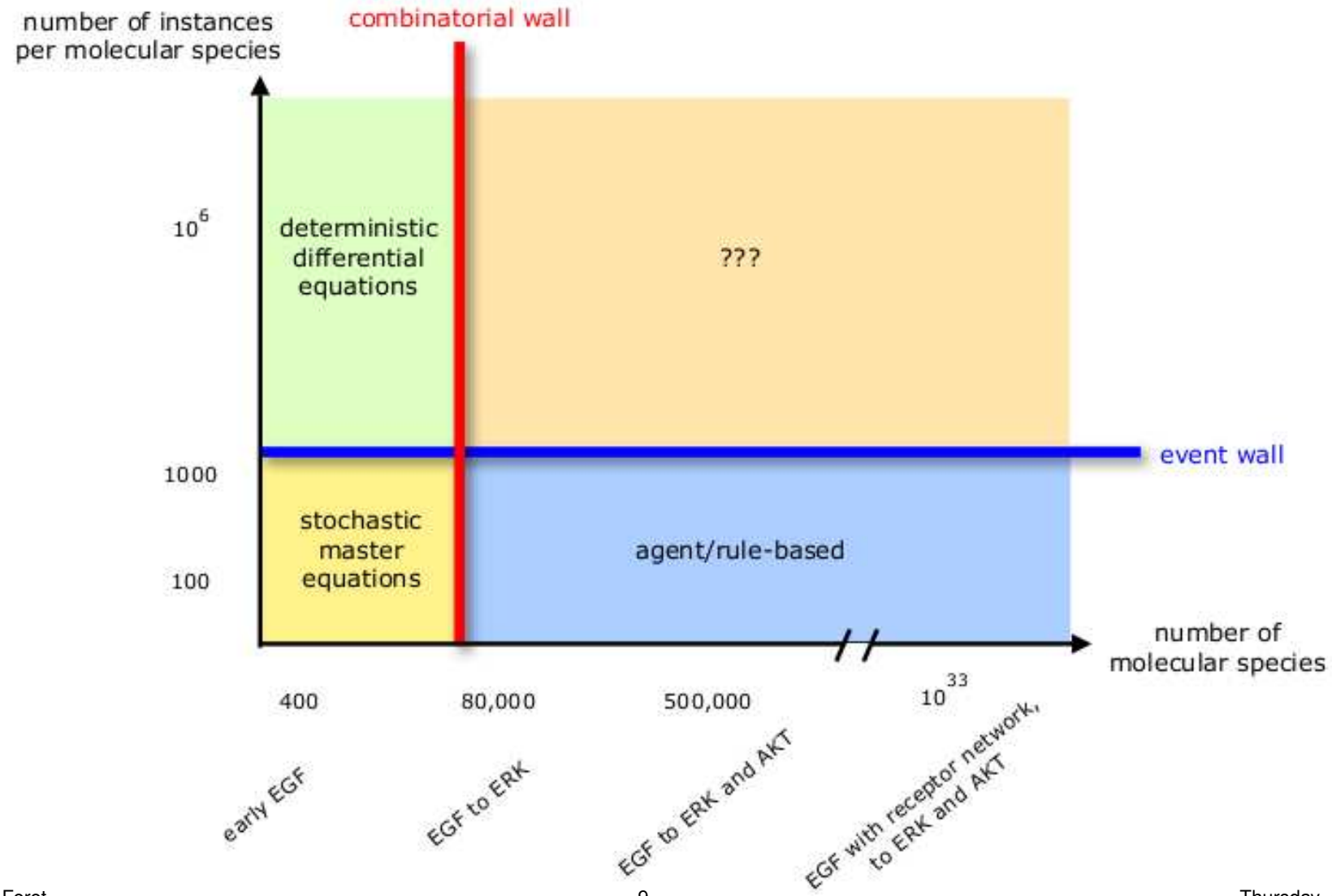


1. The description level matches with both
   - the observation level
   - and the intervention level

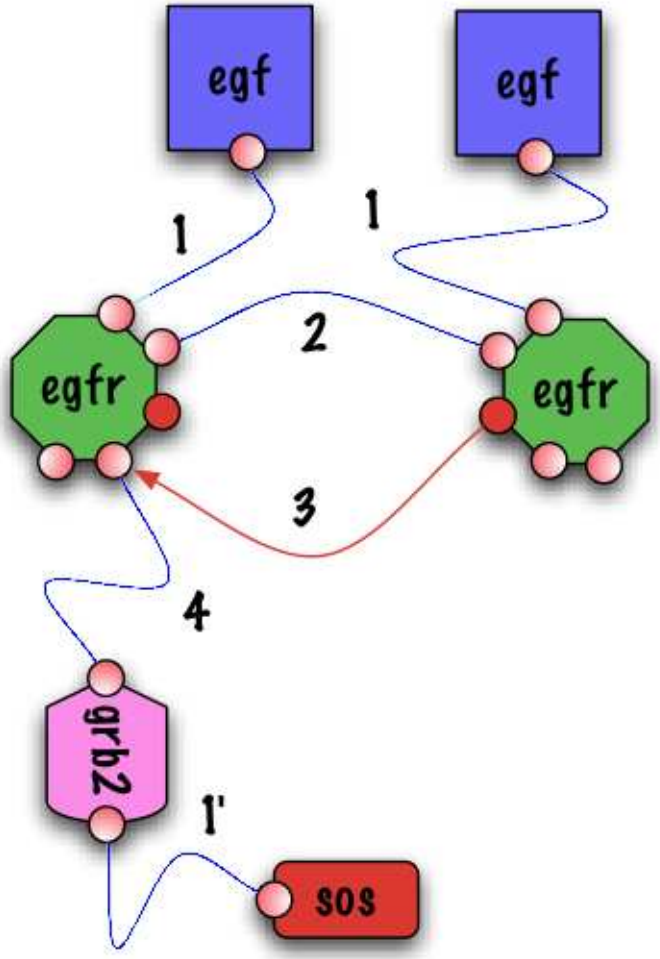
   of the biologist.

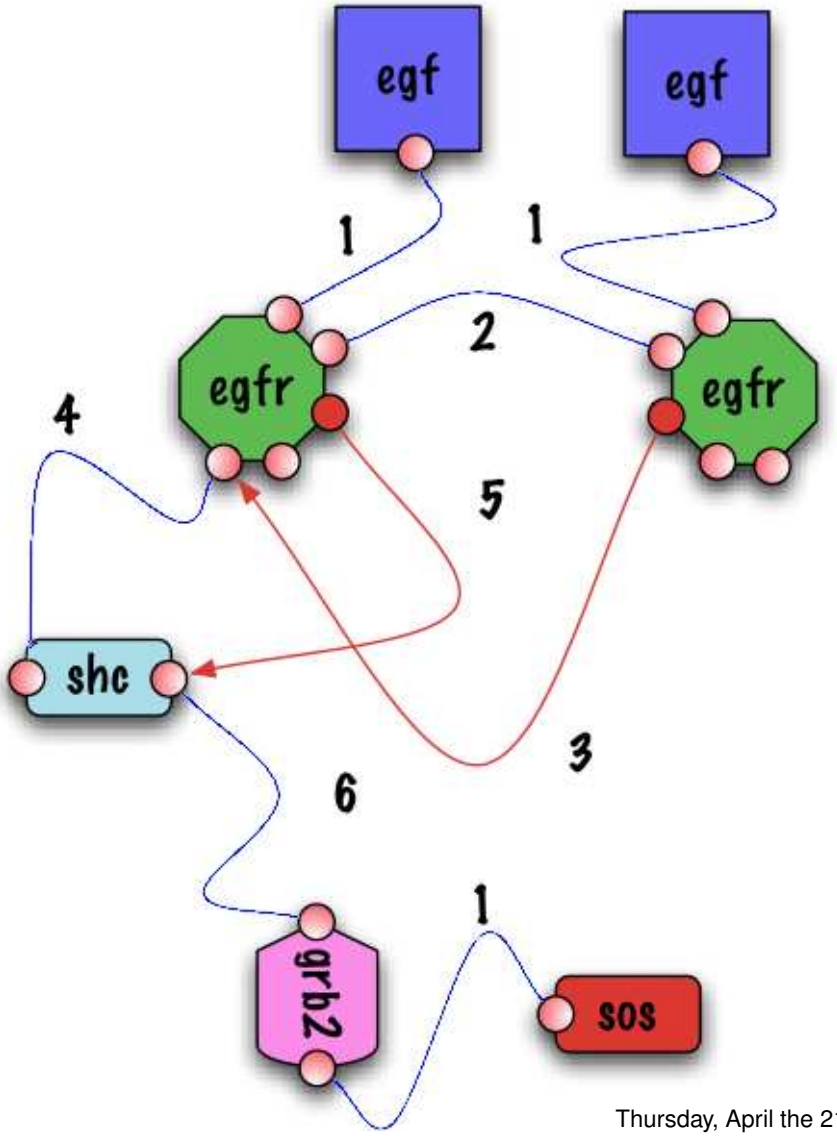
   We can tune the model easily.
2. Model description is very compact.
3. Quantitative semantics can be defined.

Complexity walls



A breach in the wall(s) ?

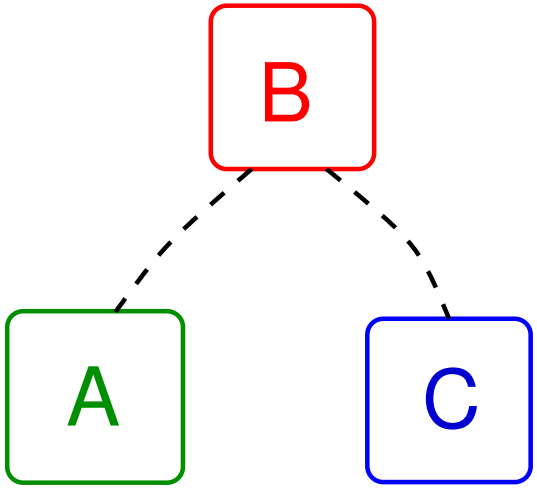




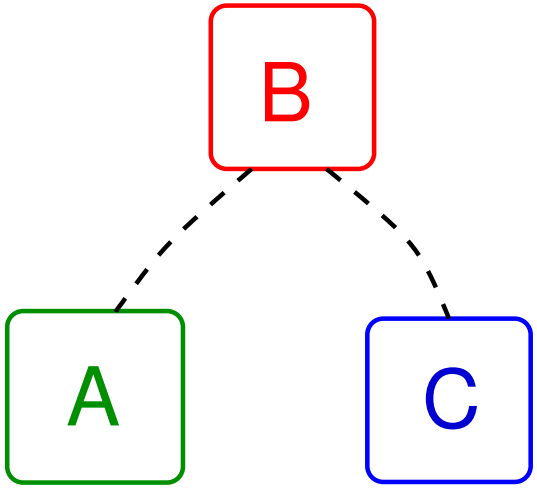
Overview

1. Context and motivations
2. Handmade ODEs
   (a) Independent subsystems
   (b) Self-consistent subsystems
   (c) Symmetric sites
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

Case study 1: A simple adapter

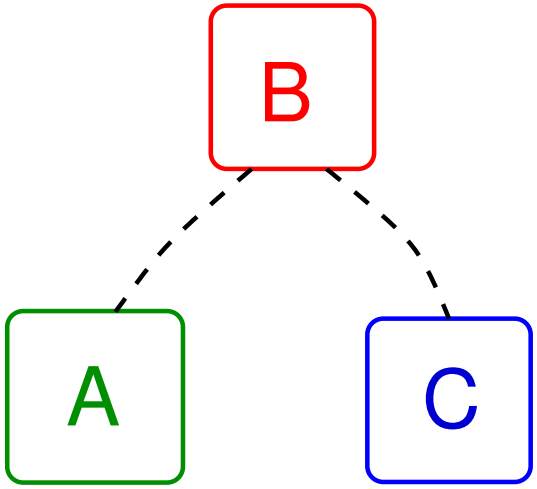


Case study 1: A simple adapter



$$A\,,\,\emptyset B\emptyset \longleftrightarrow AB\emptyset \qquad k^{AB}, k_d^{AB}$$
$$A\,,\,\emptyset BC \longleftrightarrow ABC \qquad k^{AB}, k_d^{AB}$$
$$\emptyset B\emptyset\,,\,C \longleftrightarrow \emptyset BC \qquad k^{BC}, k_d^{BC}$$
$$AB\emptyset\,,\,C \longleftrightarrow ABC \qquad k^{BC}, k_d^{BC}$$

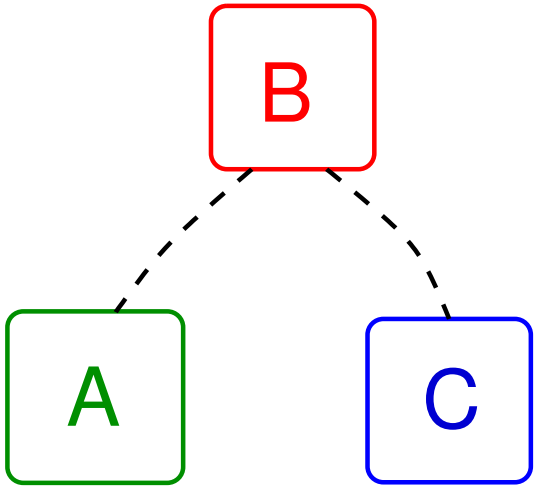
Case study 1: A simple adapter



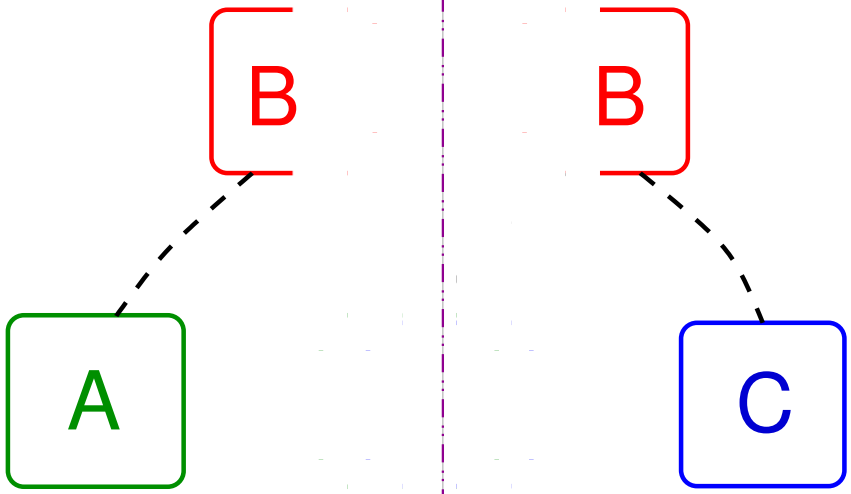
$$
\begin{aligned}
A\,,\,\emptyset B\emptyset &\longleftrightarrow AB\emptyset & k^{AB},k_d^{AB}\\
A\,,\,\emptyset BC &\longleftrightarrow ABC & k^{AB},k_d^{AB}\\
\emptyset B\emptyset\,,\,C &\longleftrightarrow \emptyset BC & k^{BC},k_d^{BC}\\
AB\emptyset\,,\,C &\longleftrightarrow ABC & k^{BC},k_d^{BC}
\end{aligned}
$$

$$
\begin{cases}
\frac{d[A]}{dt} = k_d^{AB}\cdot([AB\emptyset]+[ABC]) - [A]\cdot k^{AB}\cdot([\emptyset B\emptyset]+[\emptyset BC])\\
\frac{d[C]}{dt} = k_d^{BC}\cdot([\emptyset BC]+[ABC]) - [C]\cdot k^{BC}\cdot([\emptyset B\emptyset]+[AB\emptyset])\\
\frac{d[\emptyset B\emptyset]}{dt} = k_d^{AB}\cdot[AB\emptyset] + k_d^{BC}\cdot[\emptyset BC] - [\emptyset B\emptyset]\cdot([A]\cdot k^{AB}+[C]\cdot k^{BC})\\
\frac{d[AB\emptyset]}{dt} = [A]\cdot k^{AB}\cdot[\emptyset B\emptyset] + k_d^{BC}\cdot[ABC] - [AB\emptyset]\cdot(k_d^{AB}+[C]\cdot k^{BC})\\
\frac{d[\emptyset BC]}{dt} = k_d^{AB}\cdot[ABC] + [C]\cdot k^{BC}\cdot[\emptyset B\emptyset] - [\emptyset BC]\cdot(k_d^{BC}+[A]\cdot k^{AB})\\
\frac{d[ABC]}{dt} = [A]\cdot k^{AB}\cdot[\emptyset BC] + [C]\cdot k^{BC}\cdot[AB\emptyset] - [ABC]\cdot(k_d^{AB}+k_d^{BC})
\end{cases}
$$

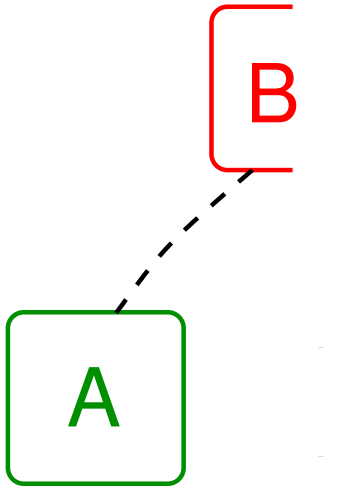
Case study 1: Two subsystems



Case study 1: Two subsystems

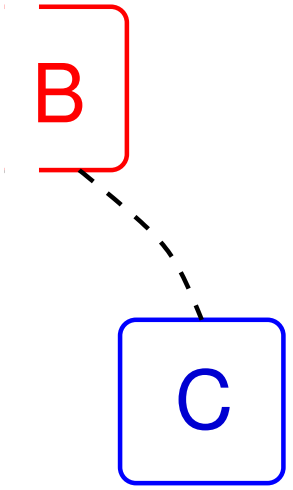


Case study 1: Two subsystems



$$[AB?] \stackrel{\Delta}{=} [AB\emptyset] + [ABC]$$
$$[\emptyset B?] \stackrel{\Delta}{=} [\emptyset B\emptyset] + [\emptyset BC]$$

$$\begin{cases} \frac{d[A]}{dt} = k_d^{AB} \cdot [AB?] - [A] \cdot k^{AB} \cdot [\emptyset B?] \\ \frac{d[AB?]}{dt} = [A] \cdot k^{AB} \cdot [\emptyset B?] - k_d^{AB} \cdot [AB?] \\ \frac{d[\emptyset B?]}{dt} = k_d^{AB} \cdot [AB?] - [A] \cdot k^{AB} \cdot [\emptyset B?] \end{cases}$$



$$[?BC] \stackrel{\Delta}{=} [\emptyset BC] + [ABC]$$
$$[?B\emptyset] \stackrel{\Delta}{=} [\emptyset B\emptyset] + [AB\emptyset]$$

$$\begin{cases} \frac{d[C]}{dt} = k_d^{BC} \cdot [?BC] - [C] \cdot k^{BC} \cdot [?B\emptyset] \\ \frac{d[?BC]}{dt} = [C] \cdot k^{BC} \cdot [?B\emptyset] - k_d^{BC} \cdot [?BC] \\ \frac{d[?B\emptyset]}{dt} = k_d^{BC} \cdot [?BC] - [C] \cdot k^{BC} \cdot [?B\emptyset] \end{cases}$$

Case study 1: Dependence index

We introduce:

$$[?B?] \stackrel{\Delta}{=} [?B\emptyset] + [?BC].$$

The binding with A and with C would be independent if, and only if:

$$\frac{[ABC]}{[?BC]} = \frac{[AB?]}{[?B?]}.$$

Thus we define the dependence index as follows:

$$X \stackrel{\Delta}{=} [ABC]\cdot[?B?] - [AB?]\cdot[?BC].$$

We have (after a short computation):

$$\frac{dX}{dt} = -X\cdot\left([A]\cdot k^{AB} + k_d^{AB} + [C]\cdot k^{BC} + k_d^{BC}\right)$$

So the property:

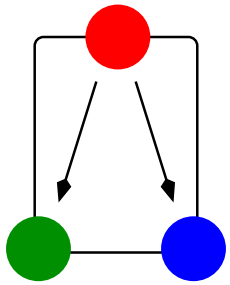
$$[ABC] = \frac{[AB?]\cdot[?BC]}{[?B?]}$$

is an invariant (i.e. if it holds at time t, it holds at any time $t' \geq t$).

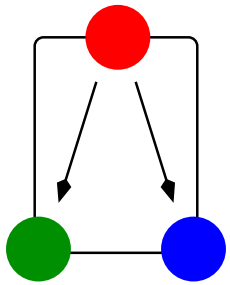
Overview

1. Context and motivations
2. Handmade ODEs
   (a) Independent subsystems
   (b) Self-consistent subsystems
   (c) Symmetric sites
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

Case study 2: A system with a switch

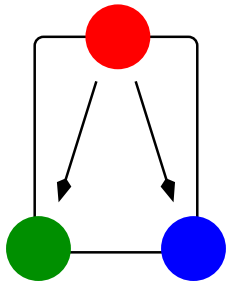


Case study 2: A system with a switch



$$
\begin{array}{lcll}
(u,u,u) & \longrightarrow & (u,p,u) & k^c \\
(u,p,u) & \longrightarrow & (p,p,u) & k^l \\
(u,p,p) & \longrightarrow & (p,p,p) & k^l \\
(u,p,u) & \longrightarrow & (u,p,p) & k^r \\
(p,p,u) & \longrightarrow & (p,p,p) & k^r
\end{array}
$$

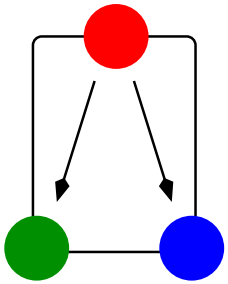
Case study 2: A system with a switch



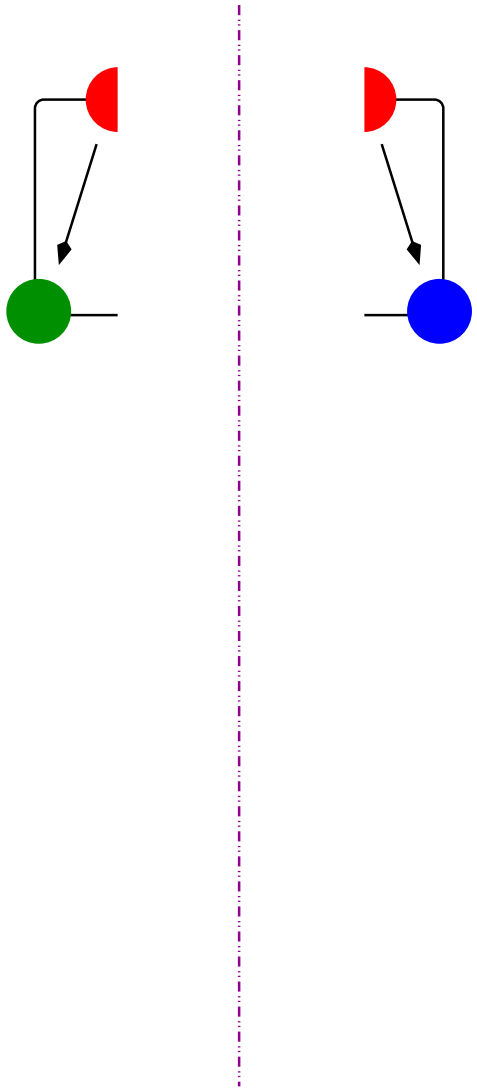
$$
\begin{aligned}
(u,u,u) &\longrightarrow (u,p,u) && k^c \\
(u,p,u) &\longrightarrow (p,p,u) && k^l \\
(u,p,p) &\longrightarrow (p,p,p) && k^l \\
(u,p,u) &\longrightarrow (u,p,p) && k^r \\
(p,p,u) &\longrightarrow (p,p,p) && k^r
\end{aligned}
$$

$$
\begin{cases}
\frac{d[(u,u,u)]}{dt} = -k^c \cdot [(u,u,u)] \\
\frac{d[(u,p,u)]}{dt} = -k^l \cdot [(u,p,u)] + k^c \cdot [(u,u,u)] - k^r \cdot [(u,p,u)] \\
\frac{d[(u,p,p)]}{dt} = -k^l \cdot [(u,p,p)] + k^r \cdot [(u,p,u)] \\
\frac{d[(p,p,u)]}{dt} = k^l \cdot [(u,p,u)] - k^r \cdot [(p,p,u)] \\
\frac{d[(p,p,p)]}{dt} = k^l \cdot [(u,p,p)] + k^r \cdot [(p,p,u)]
\end{cases}
$$

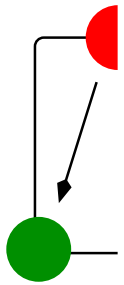
Case study 2: Two subsystems

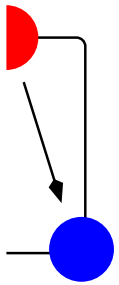


Case study 2: Two subsystems



Case study 2: Two subsystems





$$[(u,p,?)] \stackrel{\Delta}{=} [(u,p,u)] + [(u,p,p)]$$
$$[(p,p,?)] \stackrel{\Delta}{=} [(p,p,u)] + [(p,p,p)]$$

$$\begin{cases} \frac{d[(u,u,u)]}{dt} = -k^c \cdot [(u,u,u)] \\ \frac{d[(u,p,?)]}{dt} = -k^l \cdot [(u,p,?)] + k^c \cdot [(u,u,u)] \\ \frac{d[(p,p,?)]}{dt} = k^l \cdot [(u,p,?)] \end{cases}$$

$$[(?,p,u)] \stackrel{\Delta}{=} [(u,p,u)] + [(p,p,u)]$$
$$[(?,p,p)] \stackrel{\Delta}{=} [(u,p,p)] + [(p,p,p)]$$

$$\begin{cases} \frac{d[(u,u,u)]}{dt} = -k^c \cdot [(u,u,u)] \\ \frac{d[(?,p,u)]}{dt} = -k^r \cdot [(?,p,u)] + k^c \cdot [(u,u,u)] \\ \frac{d[(?,p,p)]}{dt} = k^r \cdot [(?,p,u)] \end{cases}$$

Case study 2: Dependence index

We introduce:

$$[(?,p,?)] \stackrel{\Delta}{=} [(?,p,u)] + [(?,p,p)]$$

The states of left site and right site would be independent if, and only if:

$$\frac{[(p,p,p)]}{[(p,p,?)]} = \frac{[(?,p,p)]}{[(?,p,?)]}.$$

Thus we define the dependence index as follows:

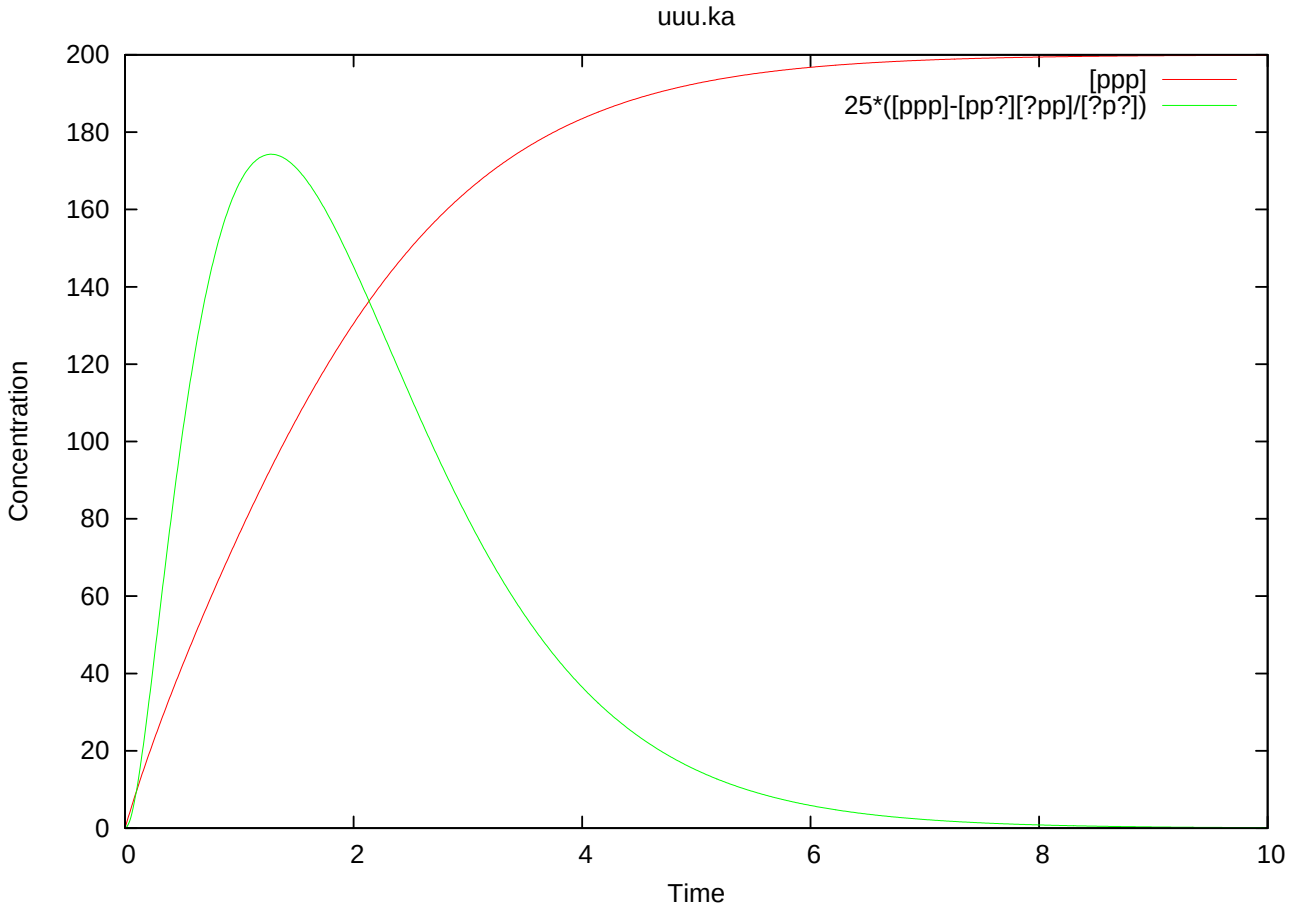
$$X \stackrel{\Delta}{=} [(p,p,p)]\cdot[(?,p,?)] - [(?,p,p)]\cdot[(p,p,?)].$$

We have (after a short computation):

$$\frac{dX}{dt} = -X \cdot \left(k^l + k^r\right) + k^c\cdot[(p,p,p)]\cdot[(u,u,u)].$$

As a consequence, the property $X = 0$ is not an invariant.
We can split the system into two subsystems,
but we cannot recombine both subsystems without errors.

Case study 2: Erroneous recombination



Concentrations evolution with respect to time ($[(u,u,u)](0) = 200$).

$[(p,p,p)]$ and $25 \cdot \left([(p,p,p)] - \frac{[(p,p,?)] \cdot [(?,p,p)]}{[(?,p,?)]}\right)$

Conclusion

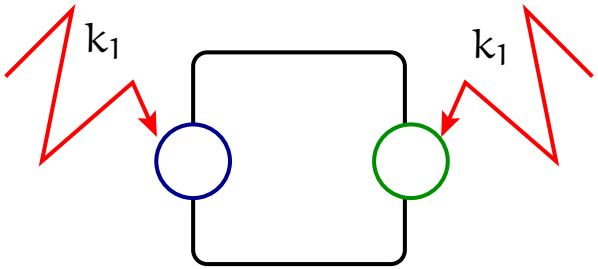
1. Independence:
   + the transformation is invertible:
     we can recover the concentration of any species;
   – it is a strong property
     which is hard to prove,
     which is hardly ever satisfied.
2. Self-consistency:
   – some information is abstracted away
     we cannot recover the concentration of any species;
   + it is a weak property
     which is easy to ensure,
     which is easy to propagate;
   + it captures the essence of the kinetics of systems.

We are going to track the correlations that are observed by the system.

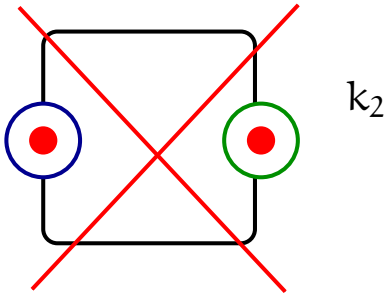
Overview

1. Context and motivations
2. Handmade ODEs
   (a) Independent subsystems
   (b) Self-consistent subsystems
   (c) Symmetric sites
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

A model with symmetries

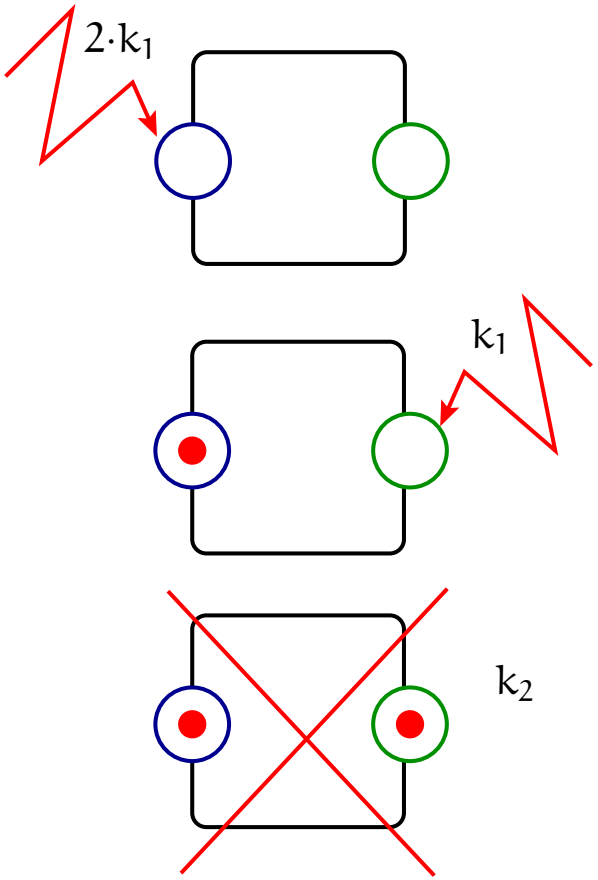


$$P \longrightarrow {}^{\star}P \quad k_1 \qquad P^{\star} \longrightarrow {}^{\star}P^{\star} \quad k_1$$
$$P \longrightarrow P^{\star} \quad k_1 \qquad {}^{\star}P \longrightarrow {}^{\star}P^{\star} \quad k_1$$



$${}^{\star}P^{\star} \longrightarrow \emptyset \quad k_2$$

Reduced model



$$P \longrightarrow {}^{\star}P \quad 2 \cdot k_1$$

$${}^{\star}P \longrightarrow {}^{\star}P^{\star} \quad k_1$$

$${}^{\star}P^{\star} \longrightarrow \emptyset \quad k_2$$

Exponential reduction!!!

Overview

1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
   (a) Concrete semantics
   (b) Abstraction
   (c) Bisimulation
   (d) Combination
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

Continuous differential semantics

Let $\mathcal{V}$, be a finite set of variables;
and $\mathbb{F}$, be a $\mathcal{C}^\infty$ mapping from $\mathcal{V} \to \mathbb{R}^+$ into $\mathcal{V} \to \mathbb{R}$,
as for instance,

- $\mathcal{V} \stackrel{\Delta}{=} \{[(u,u,u)], [(u,p,u)], [(p,p,u)], [(u,p,p)], [(p,p,p)]\}$,
- $$\mathbb{F}(\rho) \stackrel{\Delta}{=} \begin{cases} [(u,u,u)] \mapsto -k^c \cdot \rho([(u,u,u)]) \\ [(u,p,u)] \mapsto -k^l \cdot \rho([(u,p,u)]) + k^c \cdot \rho([(u,u,u)]) - k^r \cdot \rho([(u,p,u)]) \\ [(u,p,p)] \mapsto -k^l \cdot \rho([(u,p,p)]) + k^r \cdot \rho([(u,p,u)]) \\ [(p,p,u)] \mapsto k^l \cdot \rho([(u,p,u)]) - k^r \cdot \rho([(p,p,u)]) \\ [(p,p,p)] \mapsto k^l \cdot \rho([(u,p,p)]) + k^r \cdot \rho([(p,p,u)]). \end{cases}$$

The continuous semantics maps each initial state $X_0 \in \mathcal{V} \to \mathbb{R}^+$ to the maximal solution $X_{X_0} \in [0, T^{max}_{X_0}[\to (\mathcal{V} \to \mathbb{R}^+)$ which satisfies:

$$X_{X_0}(T) = X_0 + \int_{t=0}^{T} \mathbb{F}(X_{X_0}(t)) \cdot dt.$$

Overview

1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
   (a) Concrete semantics
   (b) Abstraction
   (c) Bisimulation
   (d) Combination
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

Abstraction

An abstraction $(\mathcal{V}^\sharp, \psi, \mathbb{F}^\sharp)$ is given by:

- $\mathcal{V}^\sharp$: a finite set of observables,
- ψ: a mapping from $\mathcal{V} \to \mathbb{R}$ into $\mathcal{V}^\sharp \to \mathbb{R}$,
- $\mathbb{F}^\sharp$: a $\mathcal{C}^\infty$ mapping from $\mathcal{V}^\sharp \to \mathbb{R}^+$ into $\mathcal{V}^\sharp \to \mathbb{R}$;

such that:

- ψ is linear with positive coefficients,
  and for any sequence $(x_n) \in (\mathcal{V} \to \mathbb{R}^+)^{\mathbb{N}}$ such that $(||x_n||)$ diverges towards $+\infty$, then $(||\psi(x_n)||^\sharp)$ diverges as well
  (for arbitrary norms $||\cdot||$ and $||\cdot||^\sharp$),
- $\mathbb{F}^\sharp$ is ψ-complete, i.e. the following diagram commutes:

$$\begin{array}{ccc} (\mathcal{V} \to \mathbb{R}^+) & \xrightarrow{\mathbb{F}} & (\mathcal{V} \to \mathbb{R}) \\ \psi \downarrow \ell^* & & \downarrow \ell^* \psi \\ (\mathcal{V}^\sharp \to \mathbb{R}^+) & \xrightarrow{\mathbb{F}^\sharp} & (\mathcal{V}^\sharp \to \mathbb{R}) \end{array}$$

i.e. $\psi \circ \mathbb{F} = \mathbb{F}^\sharp \circ \psi$.

Abstraction example

- $\mathcal{V} \stackrel{\Delta}{=} \{[(u,u,u)], [(u,p,u)], [(p,p,u)], [(u,p,p)], [(p,p,p)]\}$
- $\mathbb{F}(\rho) \stackrel{\Delta}{=} \begin{cases} [(u,u,u)] \mapsto -k^c \cdot \rho([(u,u,u)]) \\ [(u,p,u)] \mapsto -k^l \cdot \rho([(u,p,u)]) + k^c \cdot \rho([(u,u,u)]) - k^r \cdot \rho([(u,p,u)]) \\ [(u,p,p)] \mapsto -k^l \cdot \rho([(u,p,p)]) + k^r \cdot \rho([(u,p,u)]) \\ \dots \end{cases}$
- $\mathcal{V}^\sharp \stackrel{\Delta}{=} \{[(u,u,u)], [(?,p,u)], [(?,p,p)], [(u,p,?)], [(p,p,?)]\}$
- $\psi(\rho) \stackrel{\Delta}{=} \begin{cases} [(u,u,u)] \mapsto \rho([(u,u,u)]) \\ [(?,p,u)] \mapsto \rho([(u,p,u)]) + \rho([(p,p,u)]) \\ [(?,p,p)] \mapsto \rho([(u,p,p)]) + \rho([(p,p,p)]) \\ \dots \end{cases}$
- $\mathbb{F}^\sharp(\rho^\sharp) \stackrel{\Delta}{=} \begin{cases} [(u,u,u)] \mapsto -k^c \cdot \rho^\sharp([(u,u,u)]) \\ [(?,p,u)] \mapsto -k^r \cdot \rho^\sharp([(?,p,u)]) + k^c \cdot \rho^\sharp([(u,u,u)]) \\ [(?,p,p)] \mapsto k^r \cdot \rho^\sharp([(?,p,u)]) \\ \dots \end{cases}$

(Completeness can be checked analytically.)

Abstract continuous trajectories

Let $(\mathcal{V}, \mathbb{F})$ be a concrete system;
Let $(\mathcal{V}^\sharp, \psi, \mathbb{F}^\sharp)$ be an abstraction of the concrete system $(\mathcal{V}, \mathbb{F})$;
Let $X_0 \in \mathcal{V} \to \mathbb{R}^+$ be an initial (concrete) state.
We know that the following system:

$$Y_{\psi(X_0)}(T) = \psi(X_0) + \int_{t=0}^{T} \mathbb{F}^\sharp\left(Y_{\psi(X_0)}(t)\right) \cdot dt$$

has a unique maximal solution $Y_{\psi(X_0)}$ such that $Y_{\psi(X_0)} = \psi(X_0)$.

Theorem 1 Moreover, this solution is the projection of the maximal solution X_{X_0} of the system

$$X_{X_0}(T) = X_0 + \int_{t=0}^{T} \mathbb{F}\left(X_{X_0}(t)\right) \cdot dt,$$

which satisfies $X_{X_0}(0) = X_0$.
(ie $Y_{\psi(X_0)} = \psi(X_{X_0})$)

Abstract continuous trajectories
Proof sketch

Given an abstraction $(\mathcal{V}^\sharp, \psi, \mathbb{F}^\sharp)$, we have:

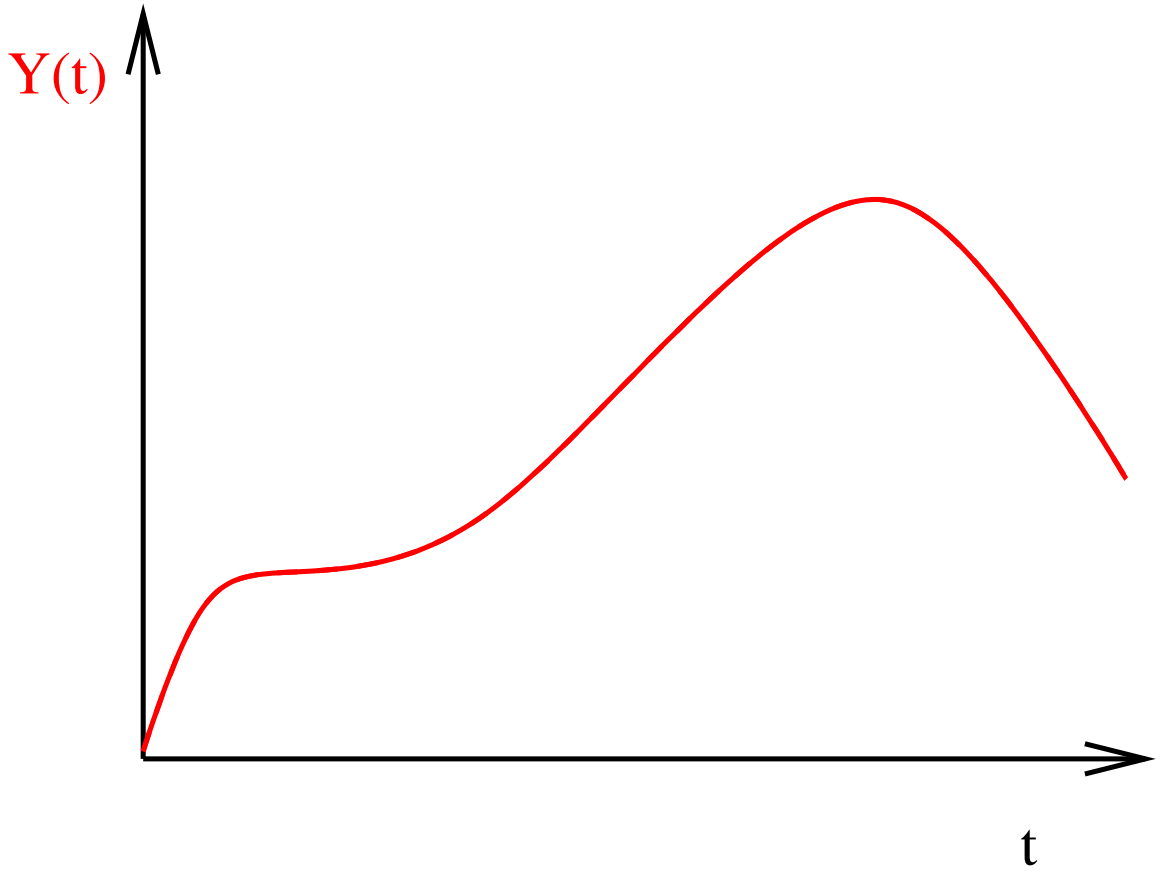
$$
\begin{aligned}
X_{X_0}(T) &= X_0 + \int_{t=0}^{T} \mathbb{F}\left(X_{X_0}(t)\right) \cdot dt \\
\psi\left(X_{X_0}(T)\right) &= \psi\left(X_0 + \int_{t=0}^{T} \mathbb{F}\left(X_{X_0}(t)\right) \cdot dt\right) \\
\psi\left(X_{X_0}(T)\right) &= \psi(X_0) + \int_{t=0}^{T} [\psi \circ \mathbb{F}]\left(X_{X_0}(t)\right) \cdot dt \quad (\psi \text{ is linear}) \\
\psi\left(X_{X_0}(T)\right) &= \psi(X_0) + \int_{t=0}^{T} \mathbb{F}^\sharp\left(\psi\left(X_{X_0}(t)\right)\right) \cdot dt \quad (\mathbb{F}^\sharp \text{ is } \psi\text{-complete})
\end{aligned}
$$

We set $Y_0 \stackrel{\Delta}{=} \psi(X_0)$ and $Y_{Y_0} \stackrel{\Delta}{=} \psi \circ X_{X_0}$.
Then we have:

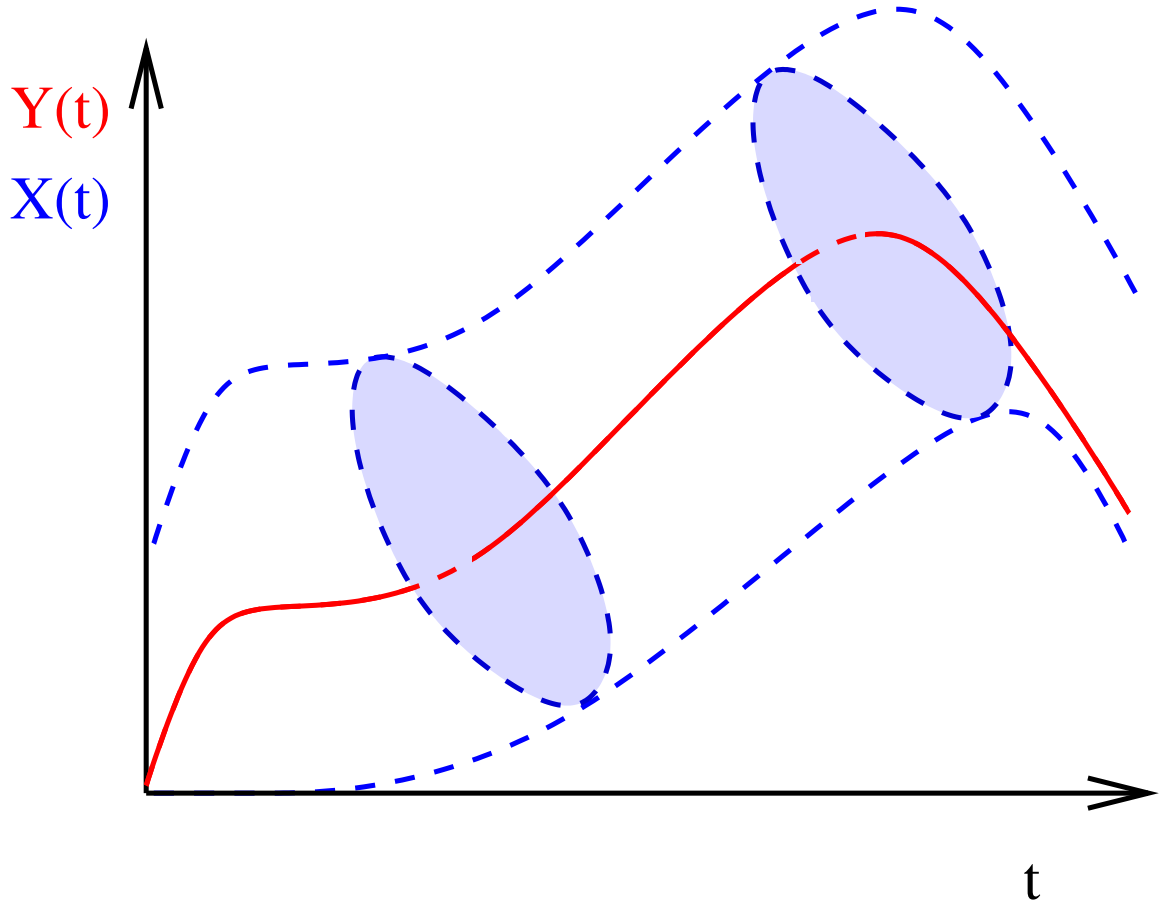
$$\boxed{Y_{Y_0}(T) = Y_0 + \int_{t=0}^{T} \mathbb{F}^\sharp\left(Y_{Y_0}(t)\right) \cdot dt}$$

The assumption about $\|\cdot\|$, $\|\cdot\|^\sharp$, and ψ ensures that $\psi \circ X_{X_0}$ is a maximal solution.

Fluid trajectories



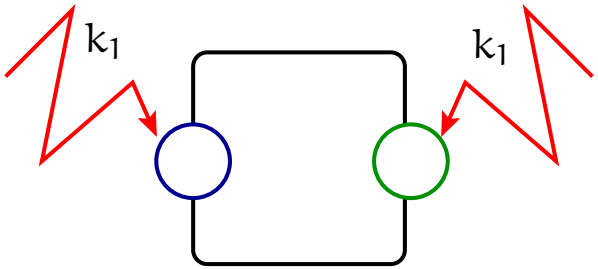
Fluid trajectories



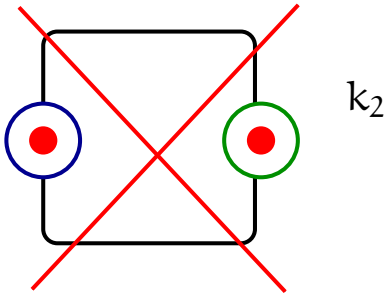
Overview

1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
   (a) Concrete semantics
   (b) Abstraction
   (c) Bisimulation
   (d) Combination
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

A model with symmetries

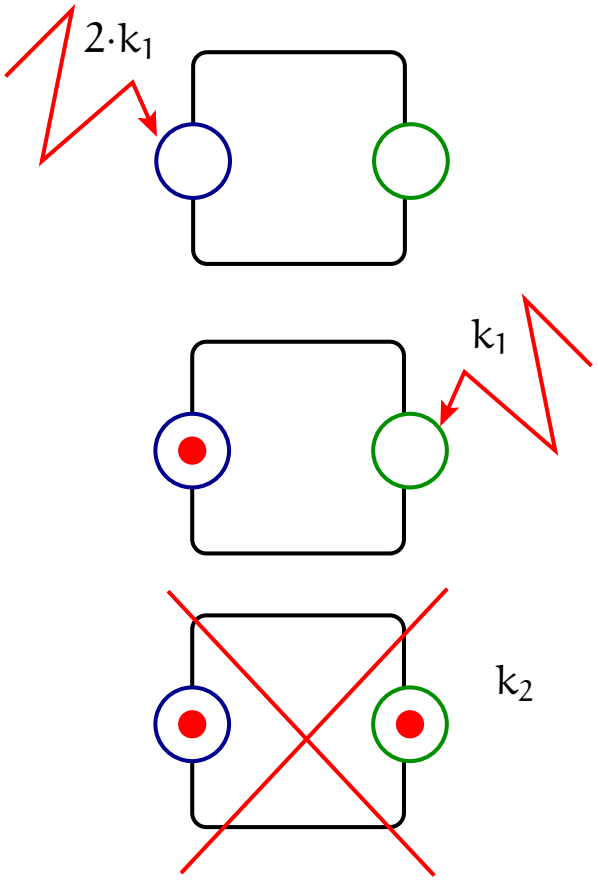


$$P \longrightarrow {}^{\star}P \quad k_1 \qquad P^{\star} \longrightarrow {}^{\star}P^{\star} \quad k_1$$
$$P \longrightarrow P^{\star} \quad k_1 \qquad {}^{\star}P \longrightarrow {}^{\star}P^{\star} \quad k_1$$



$${}^{\star}P^{\star} \longrightarrow \emptyset \quad k_2$$

Reduced model



$$P \longrightarrow {}^{\star}P \quad 2 \cdot k_1$$

$${}^{\star}P \longrightarrow {}^{\star}P^{\star} \quad k_1$$

$${}^{\star}P^{\star} \longrightarrow \emptyset \quad k_2$$

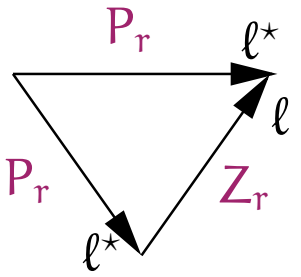
Exponential reduction!!!

Pair of projections induced by an equivalence relation among variables

Let r be a idempotent mapping from $\mathcal{V}$ to $\mathcal{V}$.
We define two linear projections $P_r, Z_r \in (\mathcal{V} \to \mathbb{R}^+) \mapsto (\mathcal{V} \to \mathbb{R}^+)$ by:

- $P_r(\rho)(V) = \begin{cases} \sum\{\rho(V') \mid r(V') = r(V)\} & \text{when } V = r(V) \\ 0 & \text{when } V \neq r(V); \end{cases}$
- $Z_r(\rho) = \begin{cases} V \mapsto \rho(V) & \text{when } V = r(V) \\ V \mapsto 0 & \text{when } V \neq r(V). \end{cases}$

We notice that the following diagram commutes:



Induced bisimulation

The mapping r induces a bisimulation,

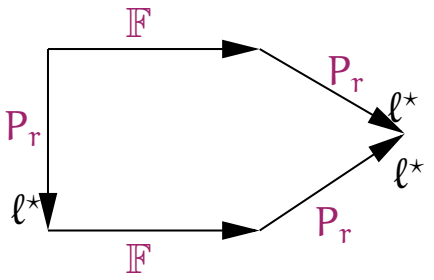
$\overset{\Delta}{\Longleftrightarrow}$

for any $\sigma, \sigma' \in \mathcal{V} \to \mathbb{R}^+$, $\mathrm{P_r}(\sigma) = \mathrm{P_r}(\sigma') \implies \mathrm{P_r}(\mathbb{F}(\sigma)) = \mathrm{P_r}(\mathbb{F}(\sigma'))$.

Indeed the mapping r induces a bisimulation,

$\Longleftrightarrow$

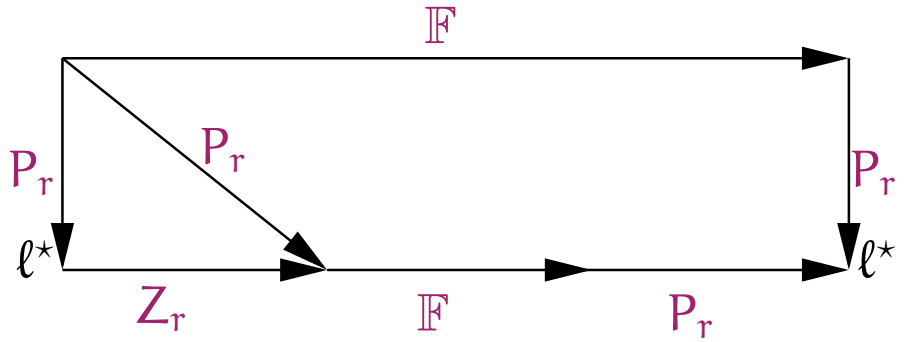
for any $\sigma \in \mathcal{V} \to \mathbb{R}^+$, $\mathrm{P_r}(\mathbb{F}(\sigma)) = \mathrm{P_r}(\mathbb{F}(\mathrm{P_r}(\sigma)))$.



Induced abstraction

Under these assumptions $(r(\mathcal{V}), P_r, P_r \circ \mathbb{F} \circ Z_r)$ is an abstraction of $(\mathcal{V}, \mathbb{F})$:

As proved in the following commutative diagram:



Overview

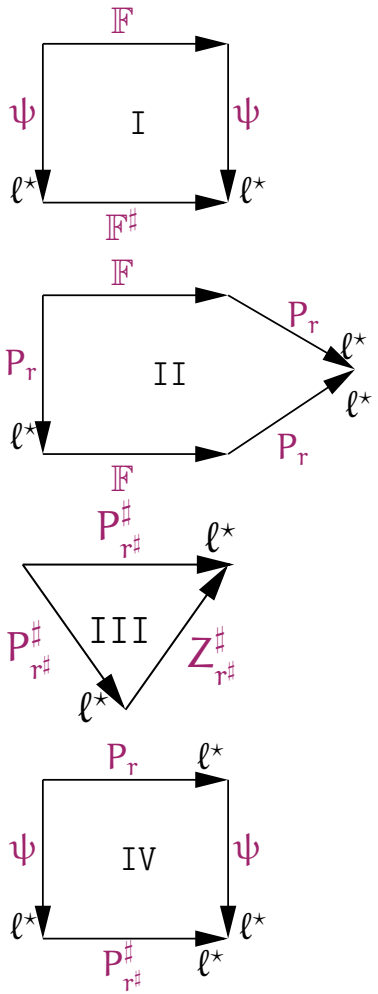
1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
   (a) Concrete semantics
   (b) Abstraction
   (c) Bisimulation
   (d) Combination
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

Abstract projection

We assume that we are given:

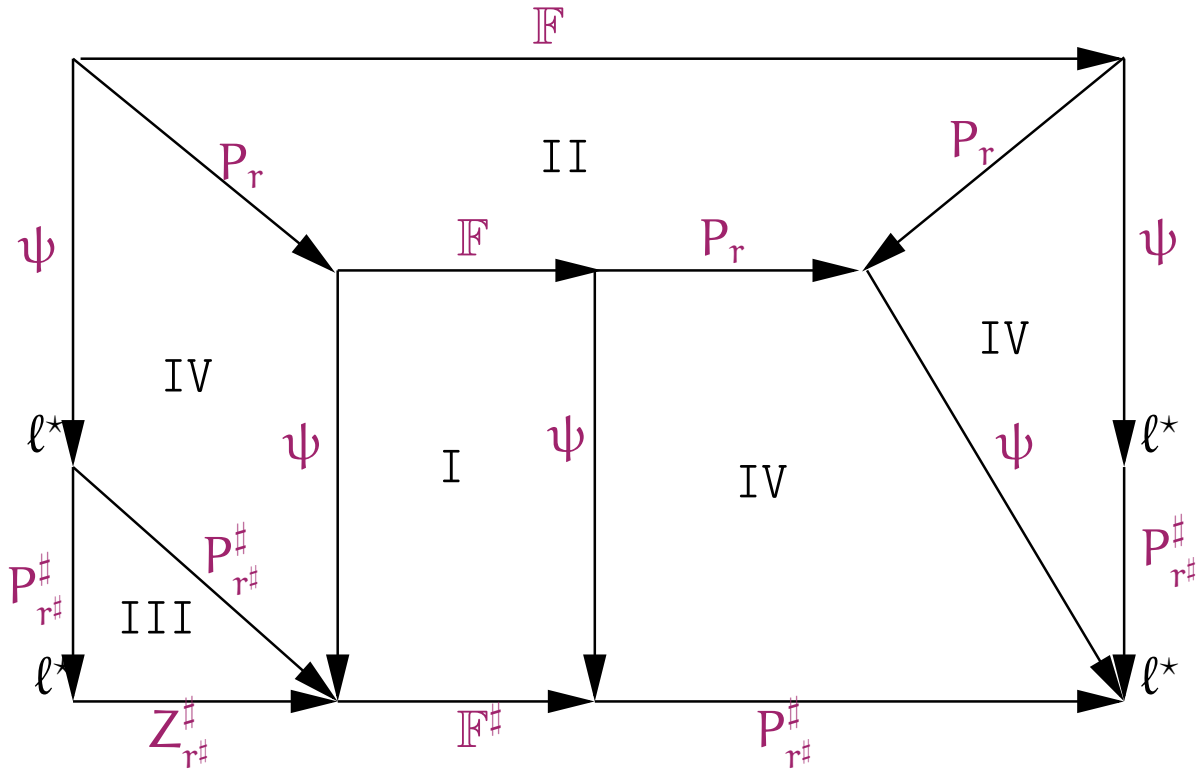
- a concrete system $(\mathcal{V}, \mathbb{F})$;
- an abstraction $(\mathcal{V}^\sharp, \psi, \mathbb{F}^\sharp)$ of $(\mathcal{V}, \mathbb{F})$ (I);
- a binary relation r over $\mathcal{V}$ which induces a bisimulation (II);
- (P_r, Z_r), the pair of projections induced by r;
- a binary relation $r^\sharp$ over $\mathcal{V}^\sharp$;
- $(P^\sharp_{r^\sharp}, Z^\sharp_{r^\sharp})$, the pair of projections over $\mathcal{V}^\sharp$ induced by $r^\sharp$ (III);

such that: $\psi \circ P_r = P^\sharp_{r^\sharp} \circ \psi$ (IV).



Combination of abstractions

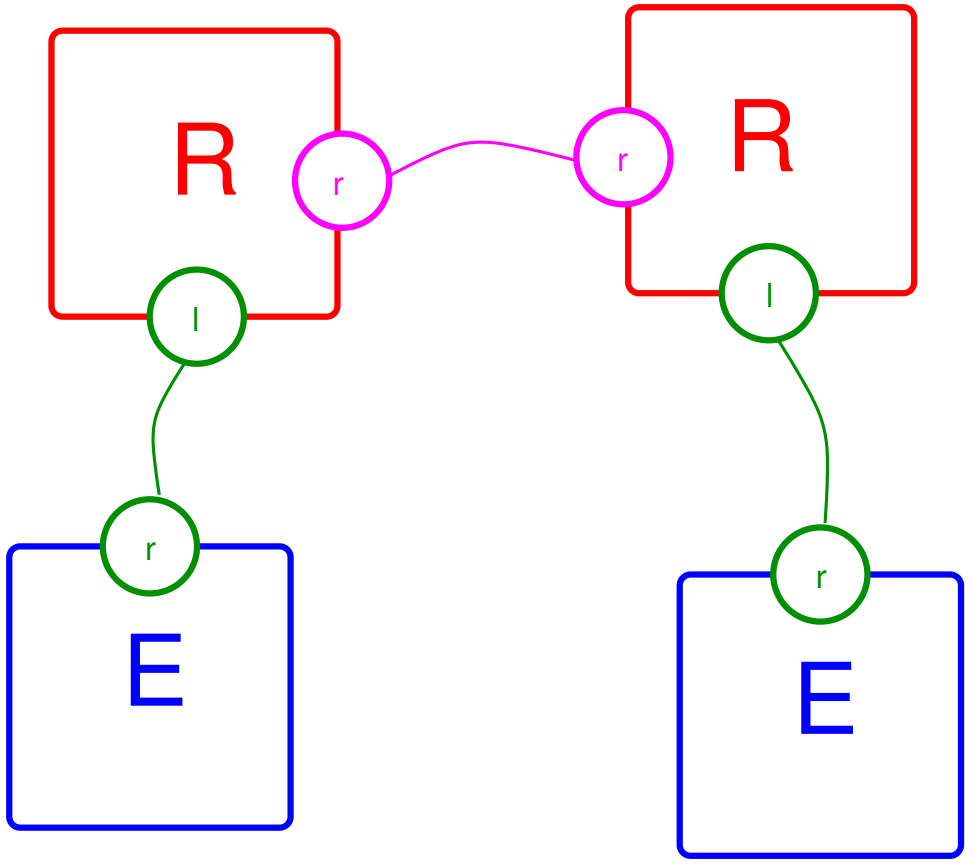
Under these assumptions, $(r^\sharp(\mathcal{V}^\sharp), P^\sharp_{r^\sharp} \circ \psi, P^\sharp_{r^\sharp} \circ \mathbb{F}^\sharp \circ Z^\sharp_{r^\sharp})$ is an abstraction of $(\mathcal{V}, \mathbb{F})$,
as proved in the following commutative diagram:



Overview

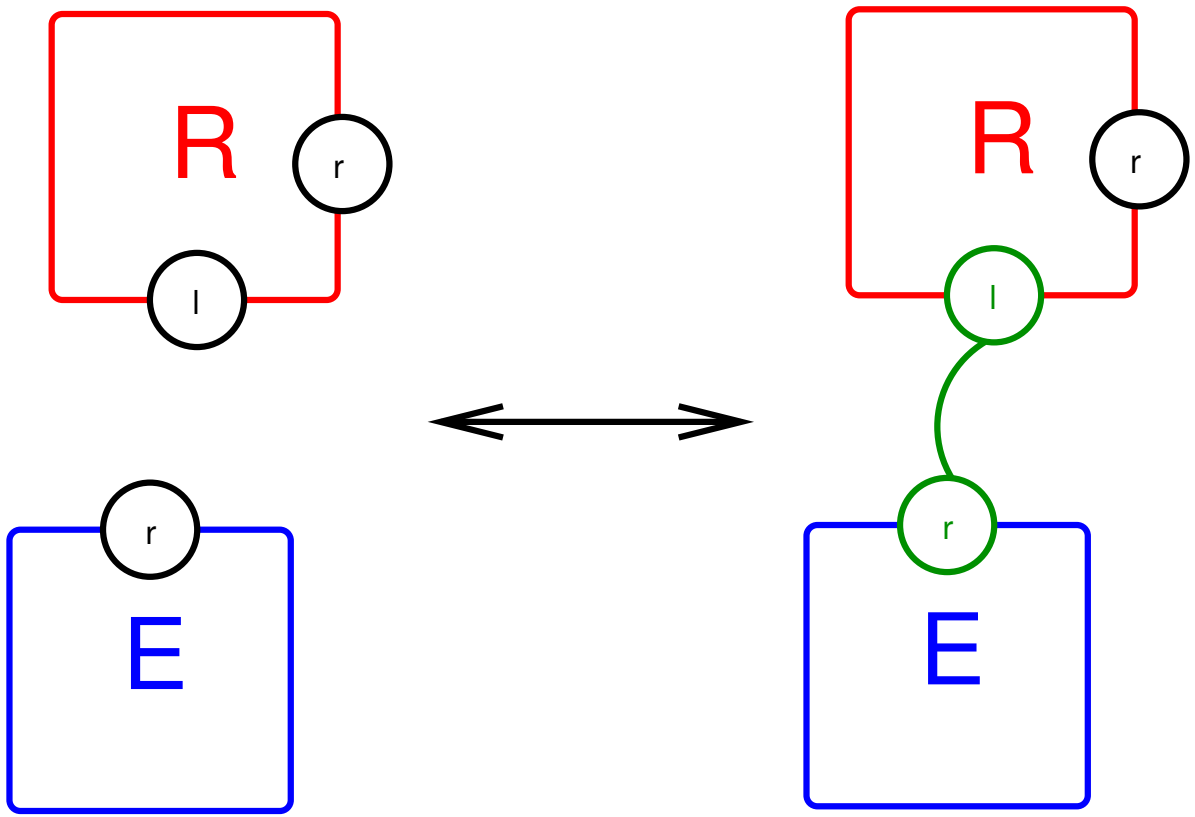
1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

A species



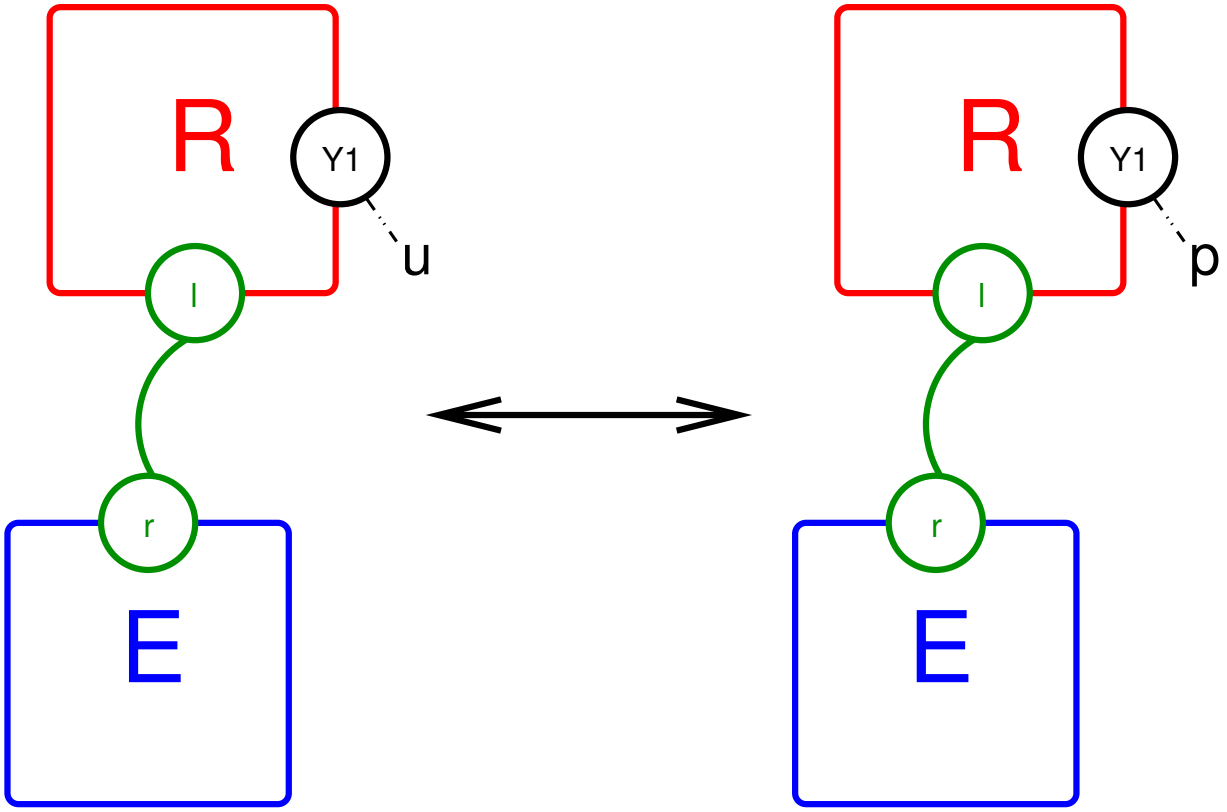
E(r!1), R(l!1,r!2), R(r!2,l!3), E(r!3)

A Unbinding/Binding Rule



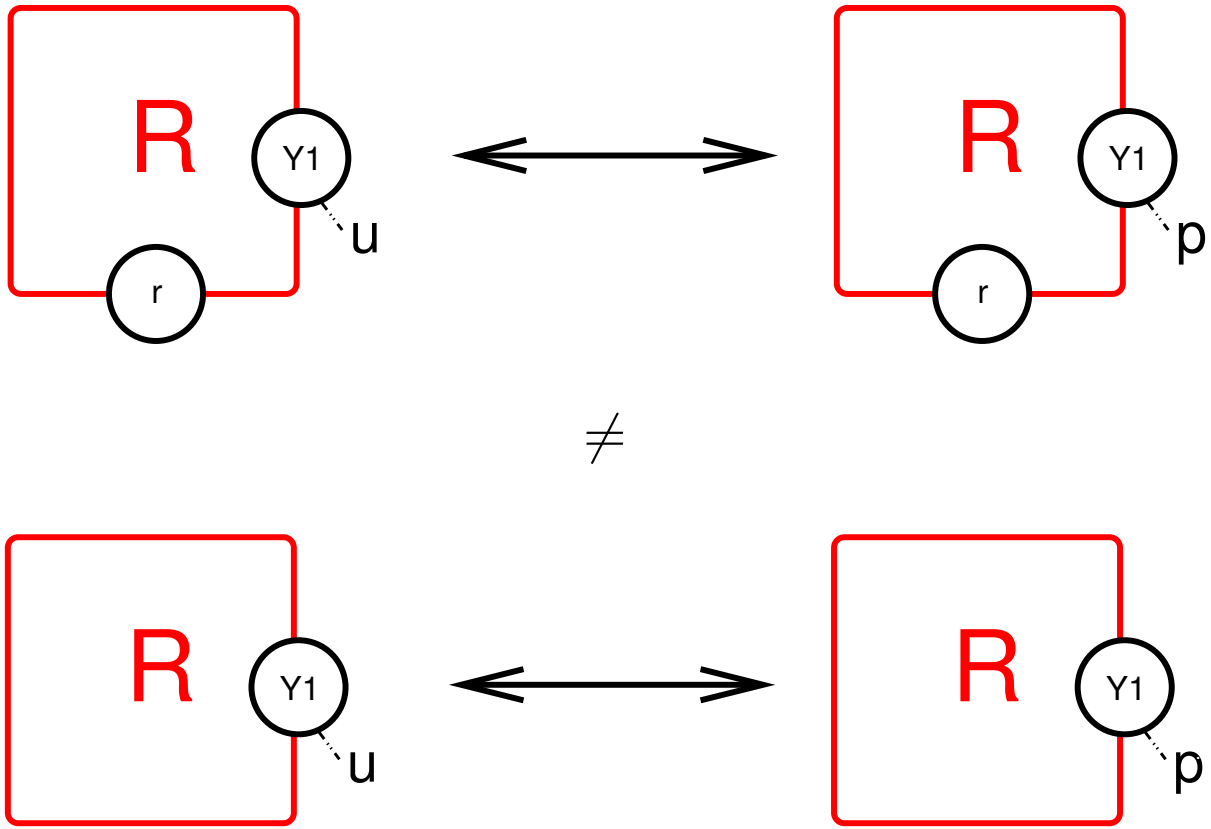
$$E(r),\ R(l,r) \longleftrightarrow E(r!1),\ R(l!1,r)$$

Internal state

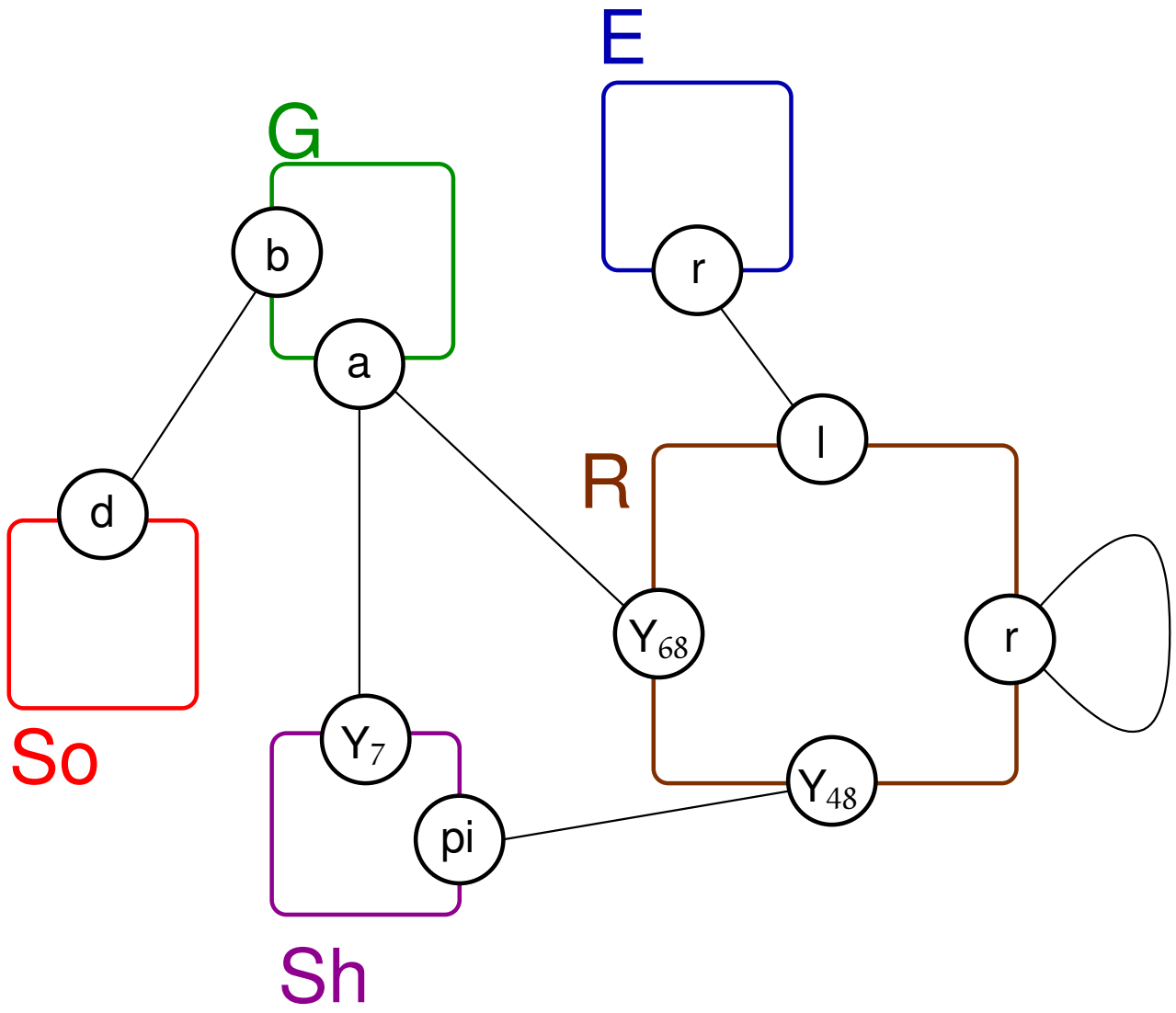


R(Y1~u,l!1), E(r!1) ⟷ R(Y1~p,l!1), E(r!1)

Don’t care, Don’t write



Contact map



Overview

1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

Requirements

1. Reachable species

   A set $\mathcal{R}$ of connected site-graphs such that:

   - $\mathcal{R}$ is finite;
   - $\mathcal{R}$ is closed with respect to rule application: i.e. applying a rule with a tuple of site-graphs in $\mathcal{R}$ gives a tuple of site-graphs in $\mathcal{R}$;

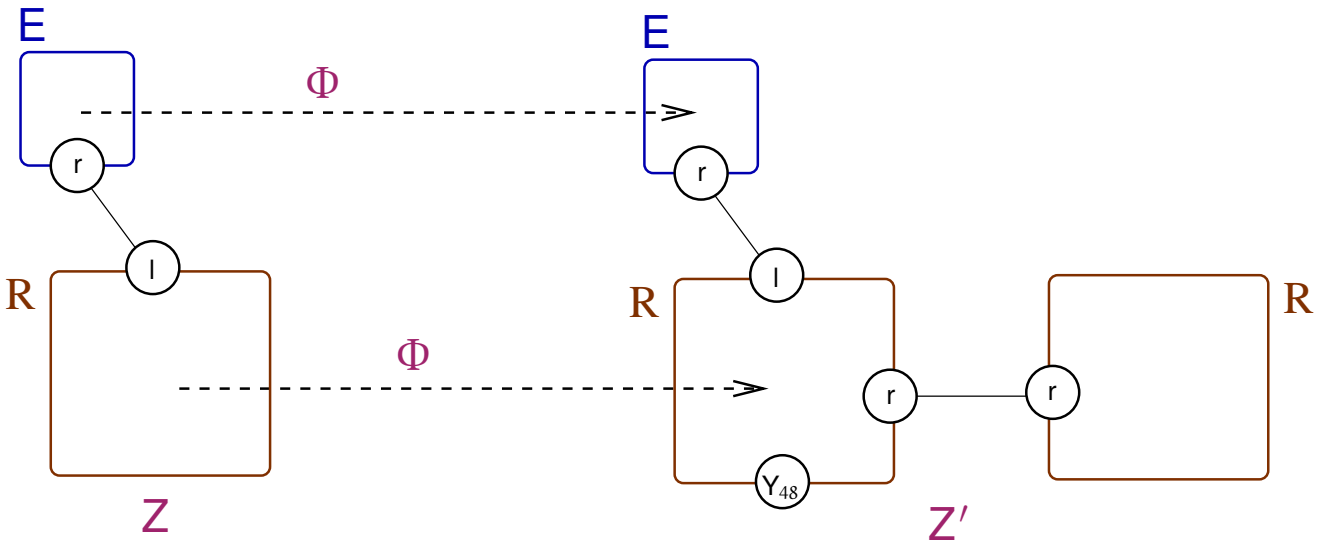
2. Rules are associated with kinetic factors

   - the unit depends on the arity of the rule as follows:

$$\left(\frac{L}{mol}\right)^{arity-1} \cdot s^{-1}$$

   where *arity* is the number of connected components in the lhs.

Embedding



We write $Z \lhd_\Phi Z'$ iff:

- Φ is a site-graph morphism:
  - i is less specific than $\Phi(i)$,
  - if there is a link between (i, s) and (i', s'), then there is a link between $(\Phi(i), s)$ and $(\Phi(i'), s')$.
- Φ is an into map (injective):
  - $\Phi(i) = \Phi(i')$ implies that $i = i'$.

Differential system

Let us consider a rule *rule*:

$$\mathit{lhs} \rightarrow \mathit{rhs} \quad k.$$

1. We write *lhs* as a multi-set $\{C_i\}$ of non empty connected components.
2. A ground instantiation of the rule *rule* is defined by a tuple $(r_i, \Phi i)$ such that $\forall i, r_i \in \mathcal{R}$ and $C_i \lhd_{\Phi i} r_i$.
3. The ground instantiation can be written as follows:

$$r_1, \ldots, r_m \rightarrow p_1, \ldots, p_n \quad k.$$

4. The activity of a ground instantiation is defined as:

$$\mathit{act}_{(r_i,\Phi i)} = \frac{k \cdot \prod [r_i]}{\sharp\{\Phi \mid \mathit{lhs} \lhd_\Phi \mathit{lhs}\}}.$$

5. Each ground instantiation induces the following contributions:

$$\frac{d[r_i]}{dt} \stackrel{+}{=} -\mathit{act}_{(r_i,\Phi i)}, \qquad \frac{d[p_i]}{dt} \stackrel{+}{=} \mathit{act}_{(r_i,\Phi i)}.$$

Overview

1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
   (a) Fragments
   (b) Soundness criteria
   (c) Abstract counterpart
   (d) Symmetries between sites
7. Conclusion

Abstract domain

We are looking for suitable pair $(\mathcal{V}^\sharp, \psi)$ (such that $\mathbb{F}^\sharp$ exists)

The set of linear variable replacements is too big to be explored.

We introduce a specific shape on $(\mathcal{V}^\sharp, \psi)$ so as:

- restrict the exploration;
- drive the intuition;
- having efficient way to find suitable abstractions $(\mathcal{V}^\sharp, \psi)$ and to compute $\mathbb{F}^\sharp$.

Our choice might be not optimal, but we can live with that.

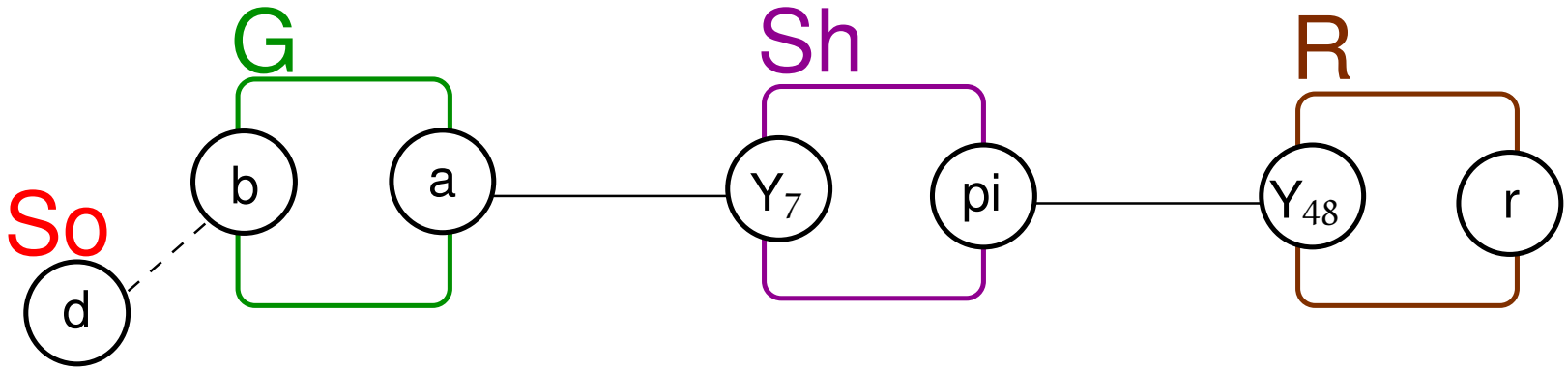
Partial species

Fragments are well-chosen partial species.

A partial species $X \in \mathcal{P}$ is a connected site-graph such that:

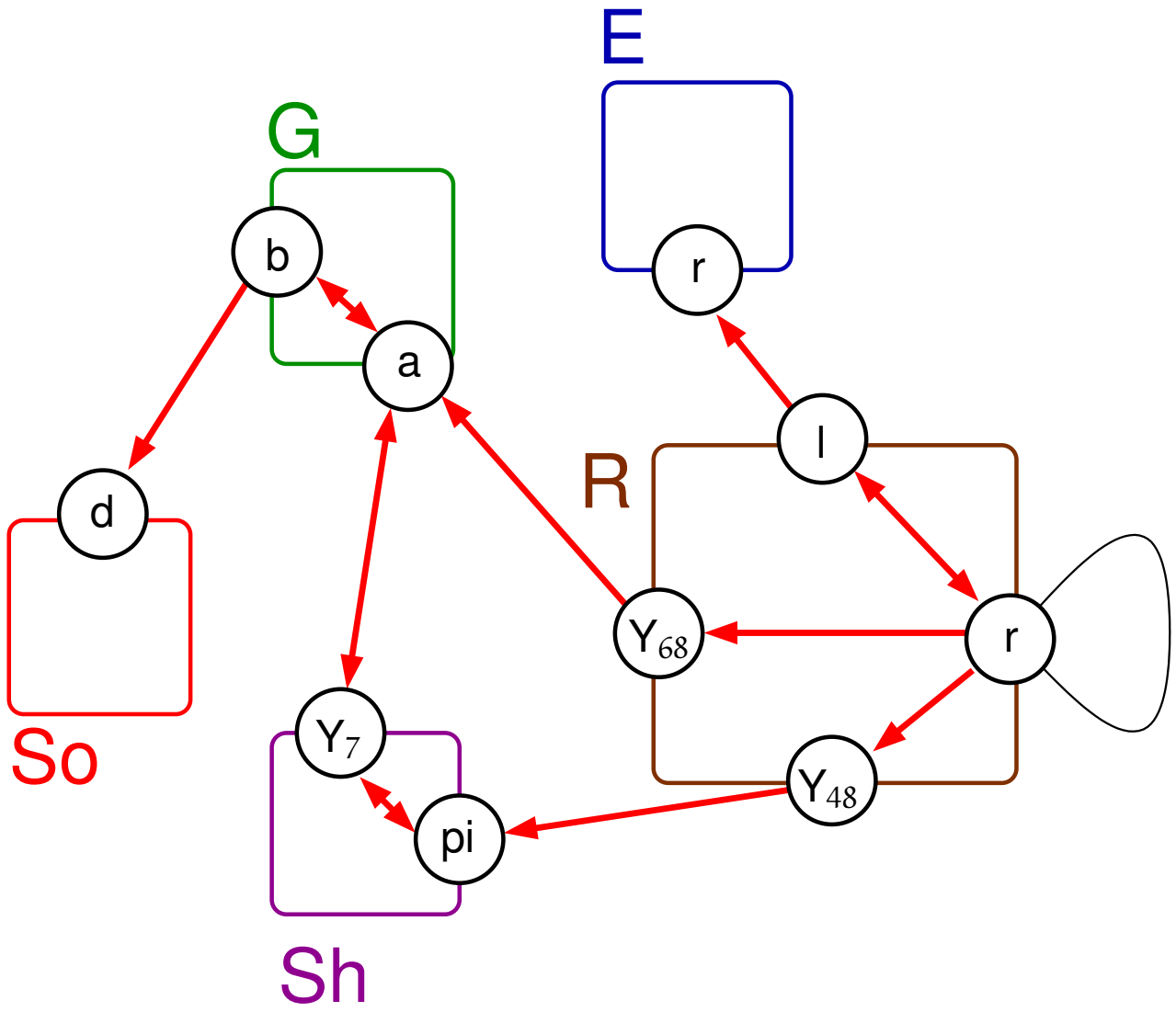
- the set of the sites of each node of type A is a subset of the set of the sites of A;
- sites are free, bound to an other site, or tagged with a binding type.

For instance:



G(b!d.So,a!1),Sh(Y_7!1,pi!2),R(Y_{48}!2,r)

Annotated contact map



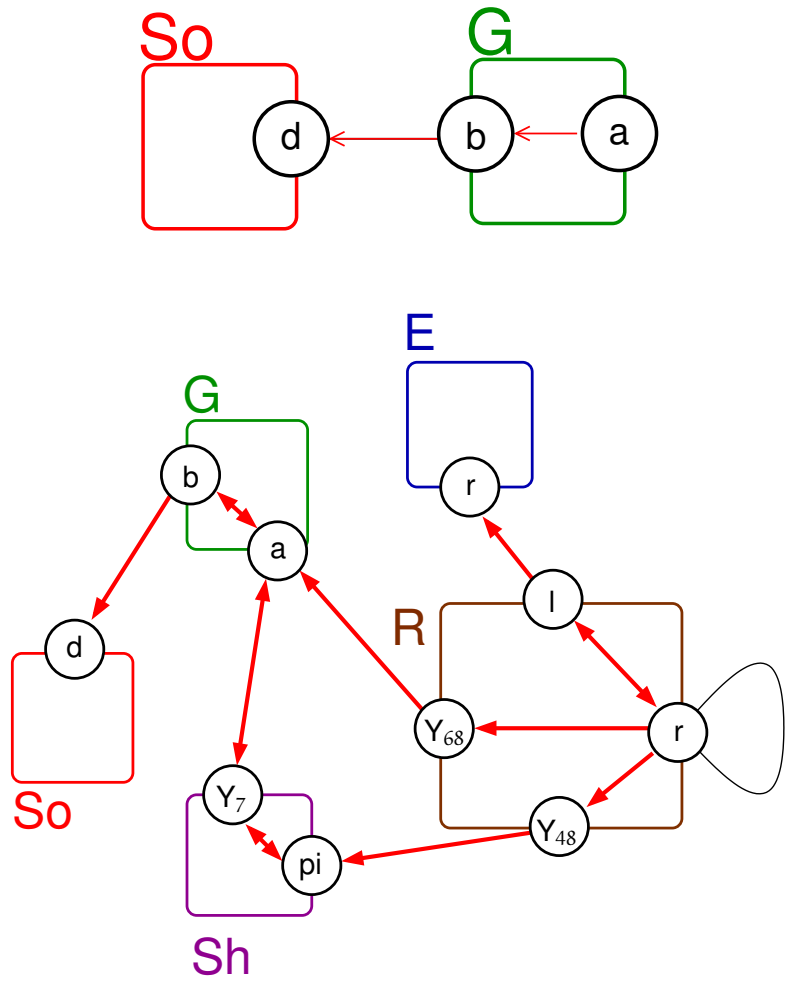
Fragments and prefragments

A prefragment is a partial species for which there exists a binary relations $\rightarrow$ between sites such that:

- Tank: there is a site that can be reached by each site via the reflexive and transitive closure of $\rightarrow$;
- Compatibility: any edge $\rightarrow$ can be projected to an edge in the annotated contact map.

A fragment is a prefragment F such that:

- Parsimoniousness: for any prefragment F' such that $\mathrm{F} \lhd \mathrm{F}'$, we also have $\mathrm{F}' \lhd \mathrm{F}$.



Inductive definition

We can provide an alternative inductive definition.

We use anchored partial species.
(ie a partial species with a site that is tagged).

Let us introduce $\mathcal{F}_I^\star$ the set of any partial species with only one site (this site is identified).

Then we call a fragment any partial species F such that there exists a anchorage $F^\star$ of F

1. transitive closure: $\exists X^\star \in \mathcal{F}_I^\star, X^\star \xrightarrow{\text{grow}}{}^\star F^\star$;
2. maximal: $\nexists (X^\star, G^\star), F^\star \xrightarrow{\text{grow}}{}^\star G^\star \xrightarrow{\text{grow}} X^\star$ such that $F = G$ and $F \neq X$.

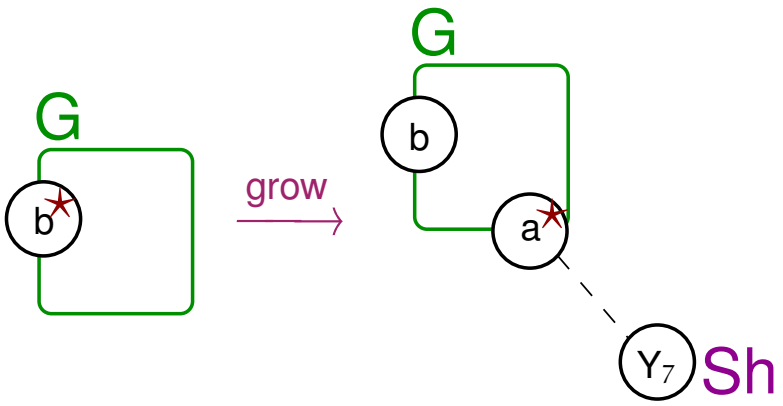
(The relation $\xrightarrow{\text{grow}}$ is defined in the next slides)

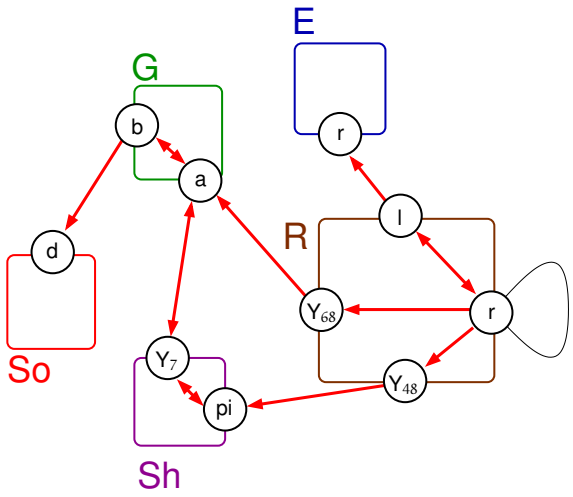
Forward growing: fresh site

Whenever:

1. an agent of type A documents the site x but not a site y,
2. the site x is identified;
3. $x.A \rightarrow y.A$ in the annotated contact map;

then, the site y can be added with an arbitrary internal state, it can be either free, or associated with an arbitrary binding type. The anchor is moved to the newsite y.



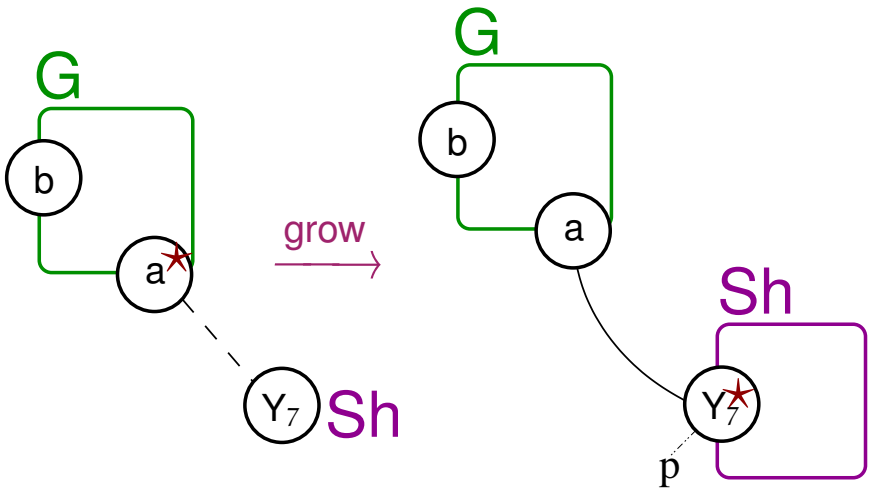


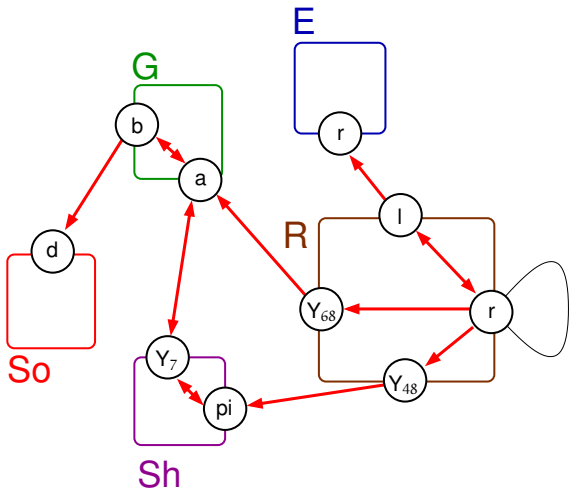
Forward growing: fresh agent

Whenever:

1. the site x of an agent of type A is associated with a binding type $y.B$;
2. the site x is identified;
3. $x.A \rightarrow y.B$ in the annotated contact map;

then, the binding type $x.A$ can be replaced with a bond to the site y of a fresh agent of type B, with an arbitrary internal state. The anchor is moved to $y.B$.



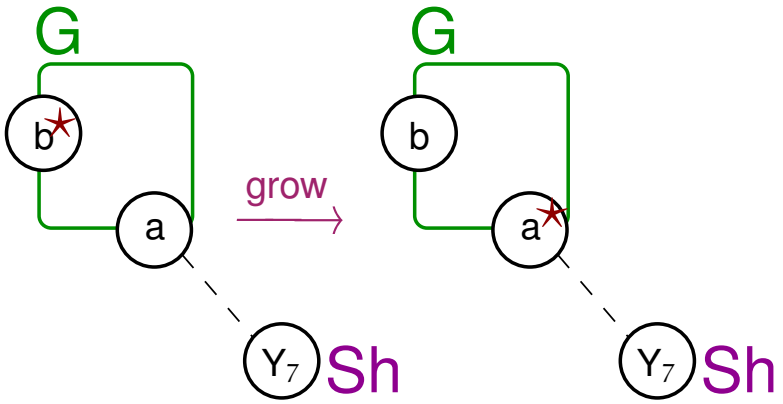


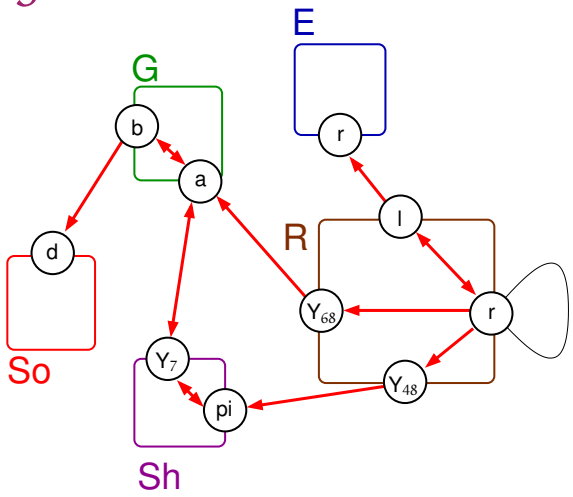
Silent move: sites

Whenever:

1. an agent of type A documents the site x and the site y,
2. the site x is identified;
3. $x.A \rightarrow y.A$ in the annotated contact map;

then, the anchor can be moved to the site y.



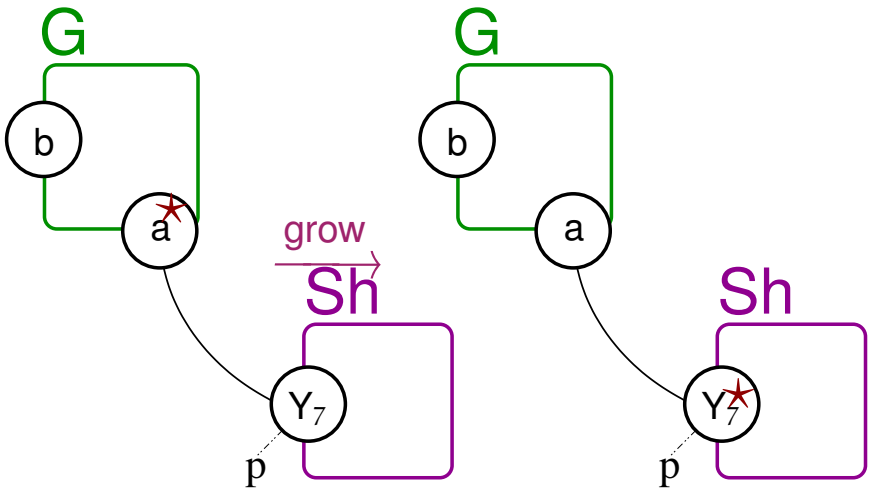


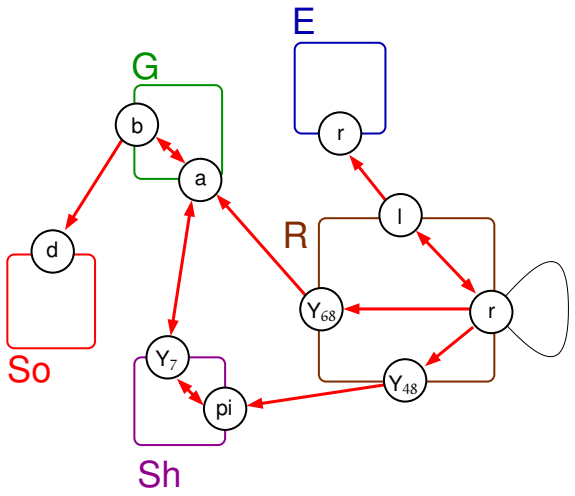
Silent move: bonds

Whenever:

1. the site x of an agent of type A and the site y of an agent of type B are bound together;
2. the site x is identified;
3. $x.A \rightarrow y.B$ in the annotated contact map;

then, we can move the anchor from $x.A$ to $y.B$.



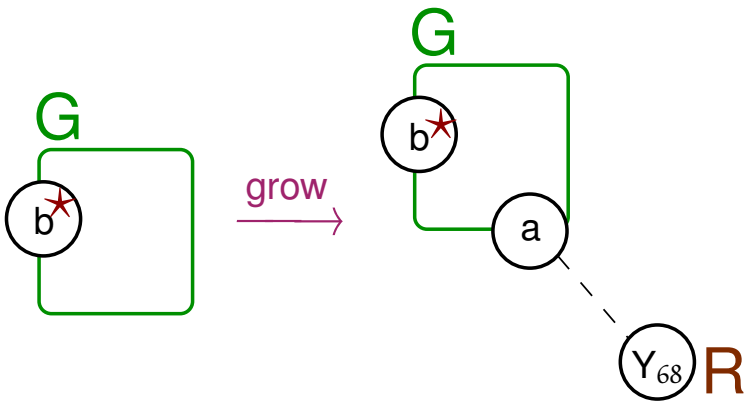


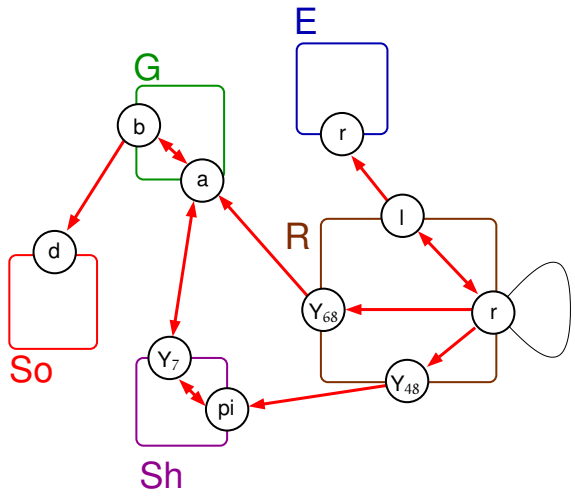
Backward growing: fresh site

Whenever:

1. an agent of type A documents a site x but not a site y;
2. $y.A \rightarrow x.A$ in the annotated contact map;

then, the site y can be added with an arbitrary internal state, it can be either free, or associated with an arbitrary binding type.



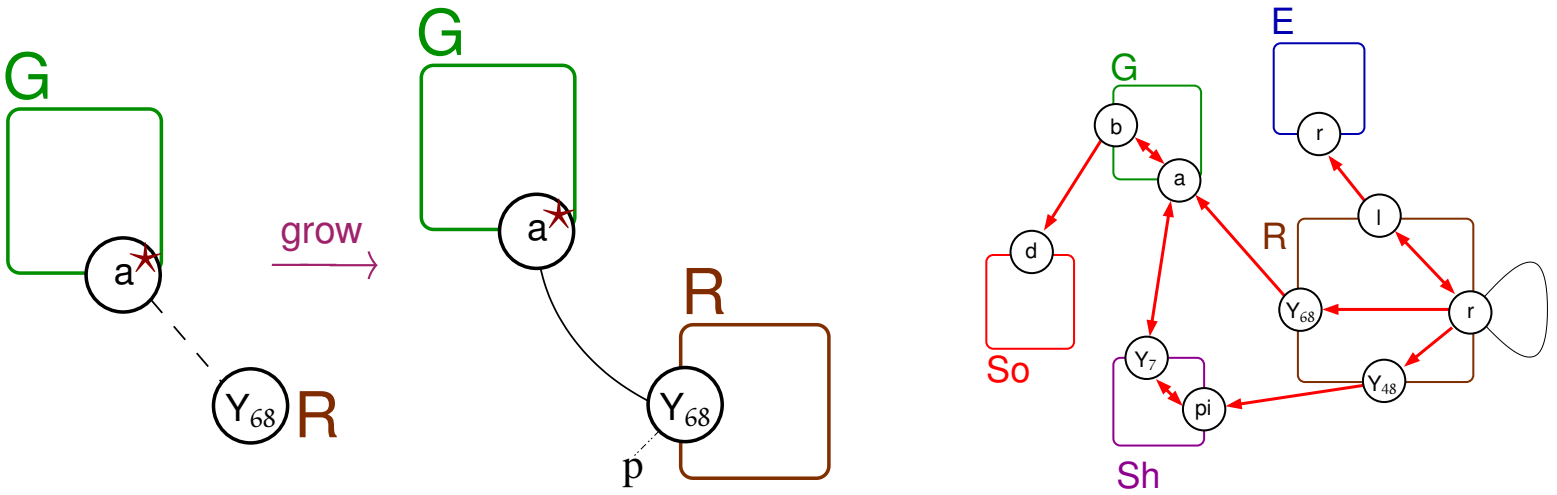


Backward growing: fresh agent

Whenever:

1. the site x of an agent of type A is associated with a binding type $y.B$;
2. $y.B \rightarrow x.A$ in the annotated contact map;

then, the binding type $x.A$ can be replaced with a bond to the site y of a fresh agent of type B, with an arbitrary internal state.

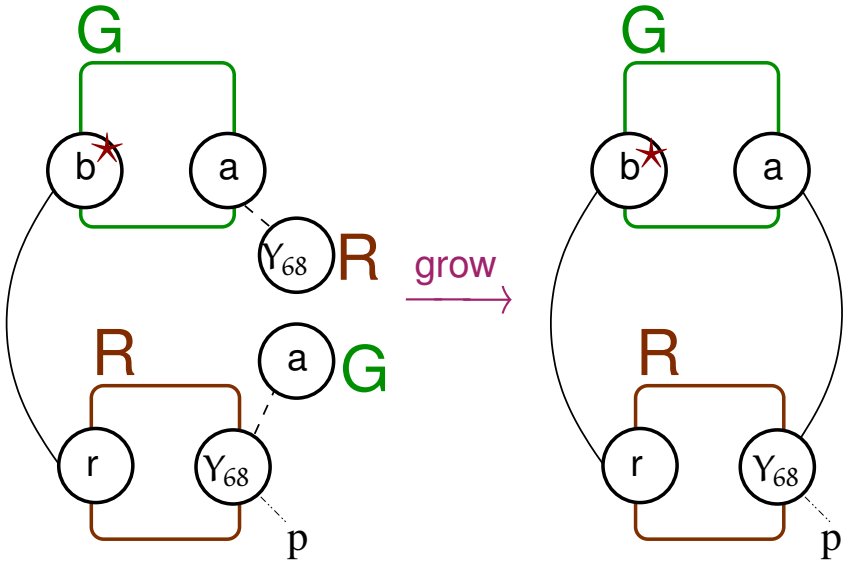


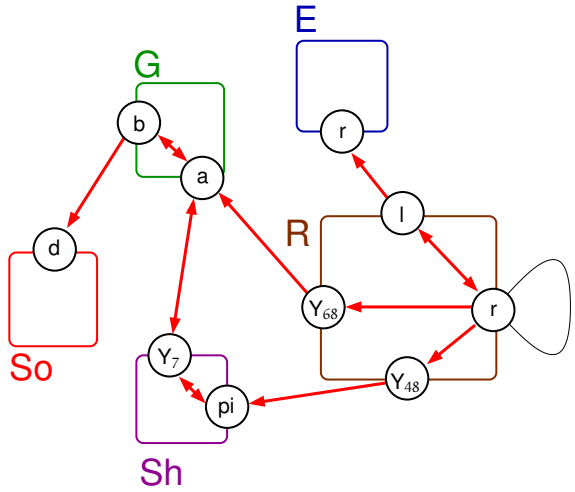
Backward growing: internal link

Whenever:

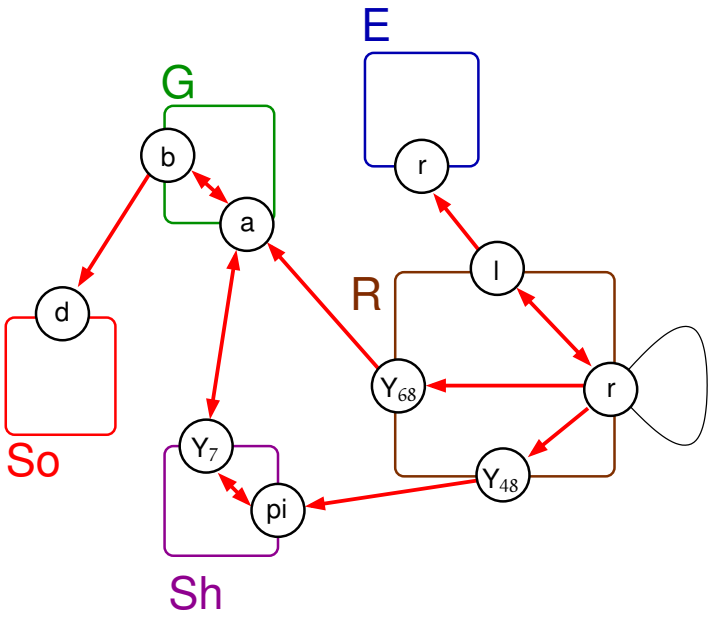
1. the site x of an agent of type A is associated with a binding type $y.B$;
2. the site y of an agent of type B is associated with a binding type $x.A$,
3. $x.A \rightarrow y.B$ in the annotated contact map;

then, both binding types can be replaced with a bond between the two sites.

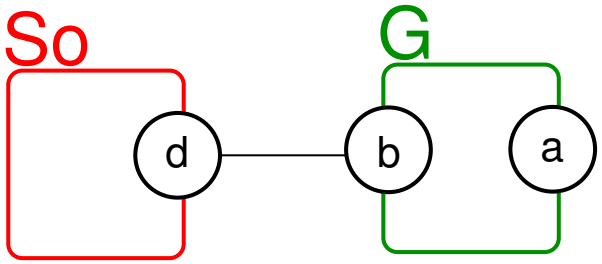


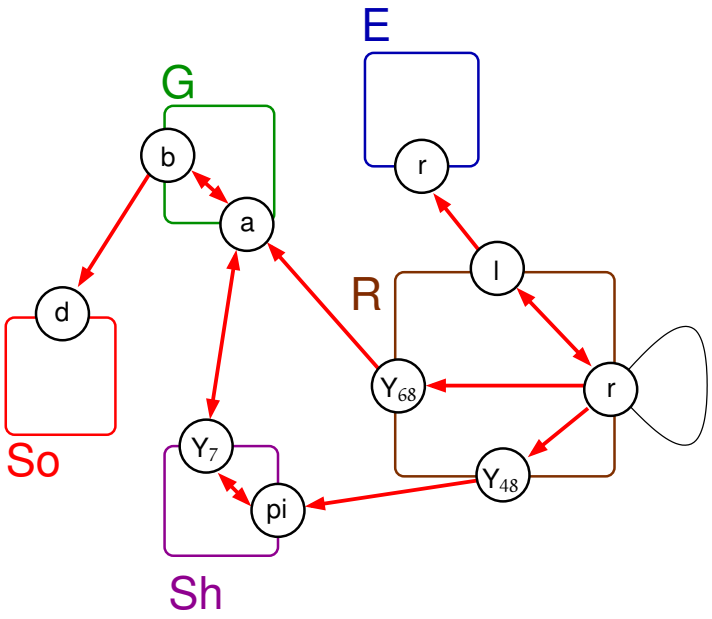


Are they fragments ?

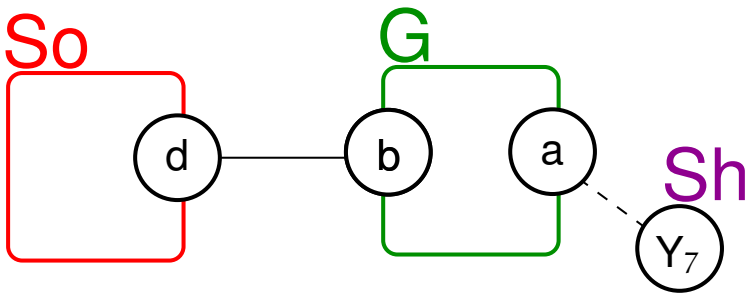


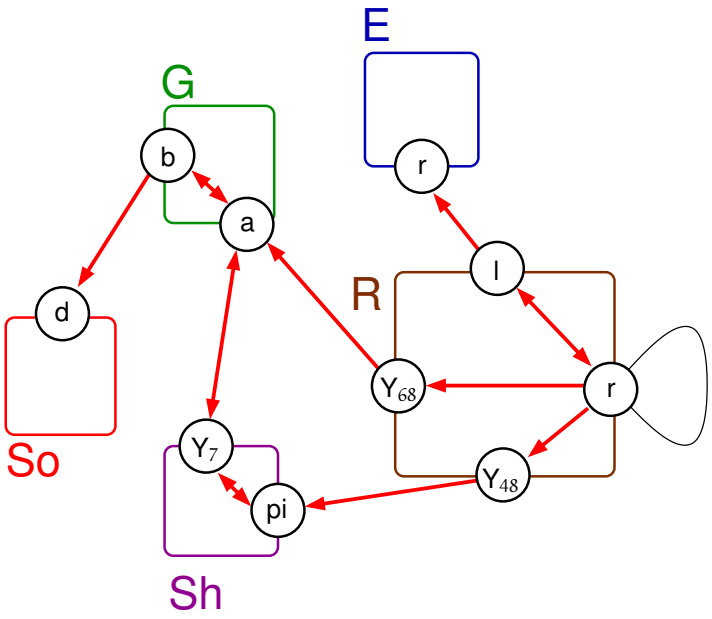
Are they fragments ?



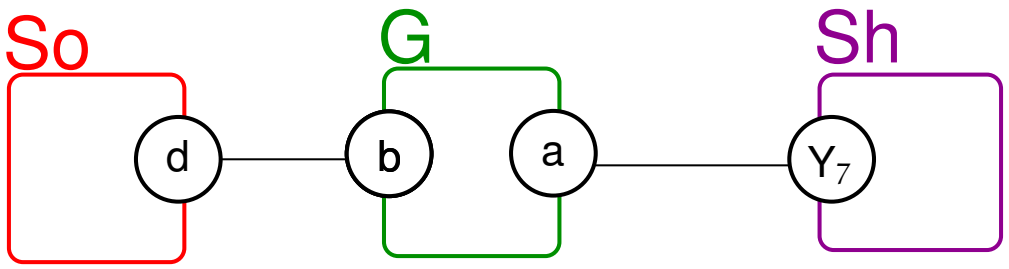


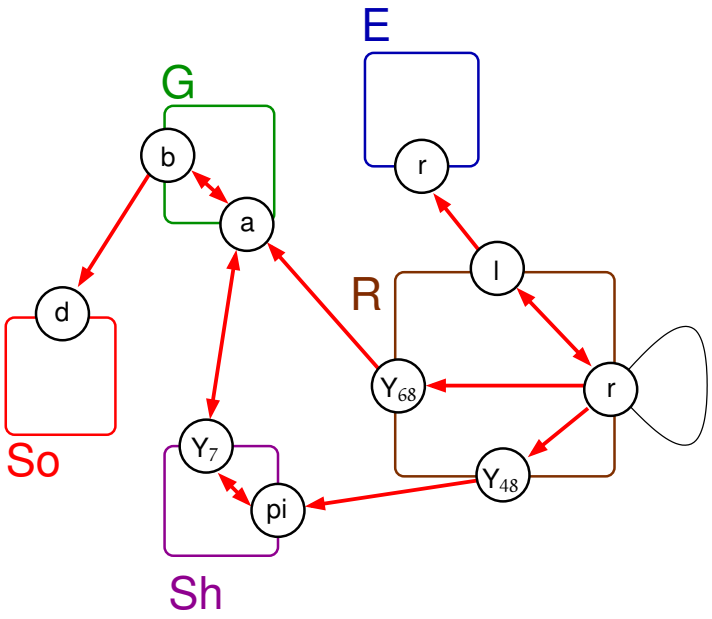
Are they fragments ?



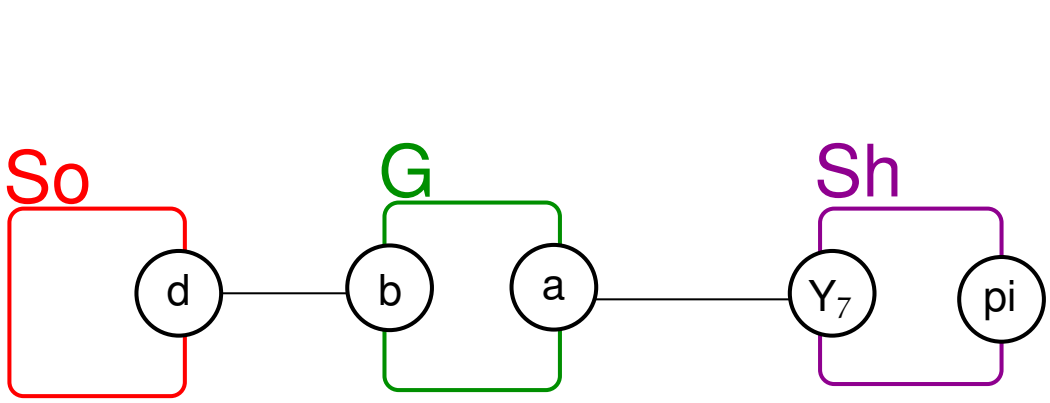


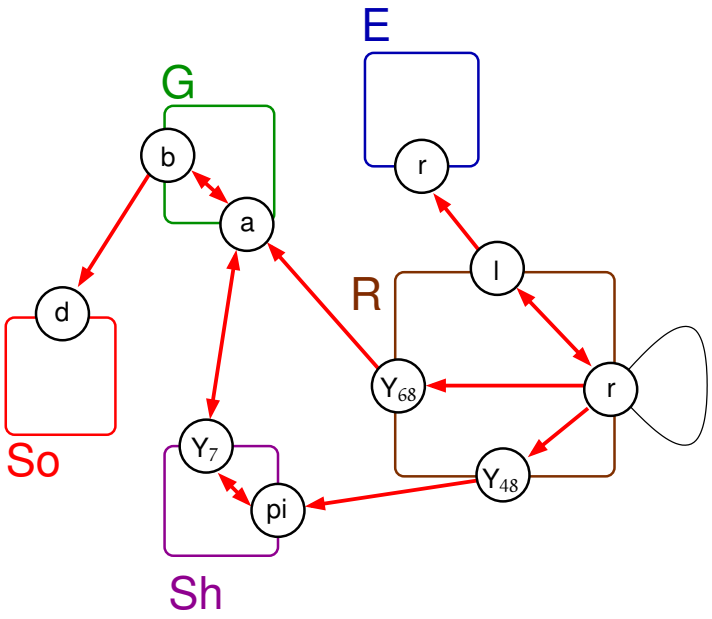
Are they fragments ?



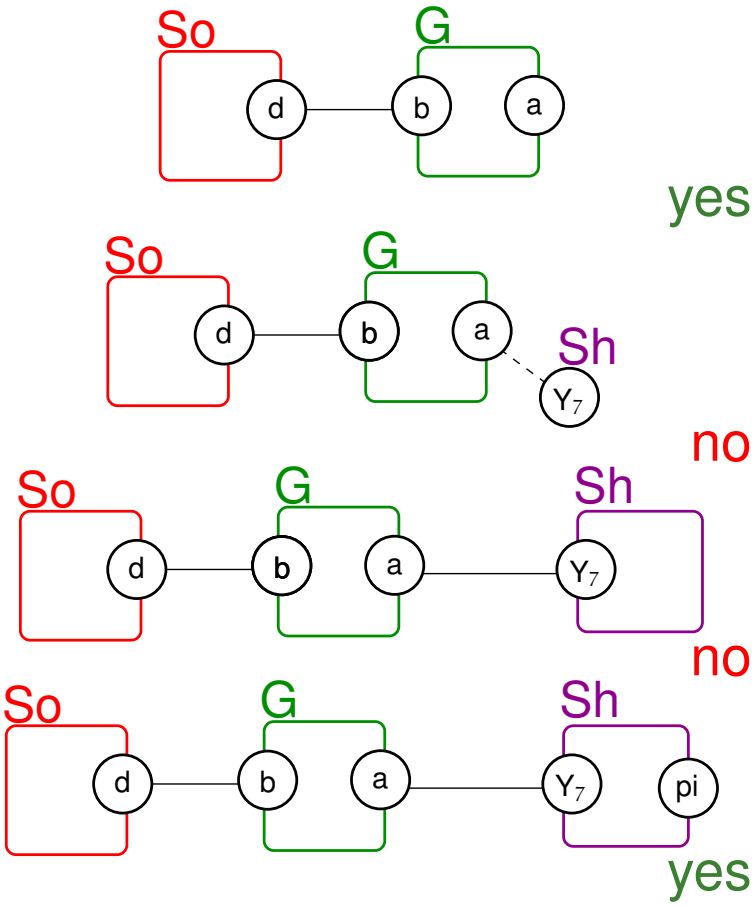


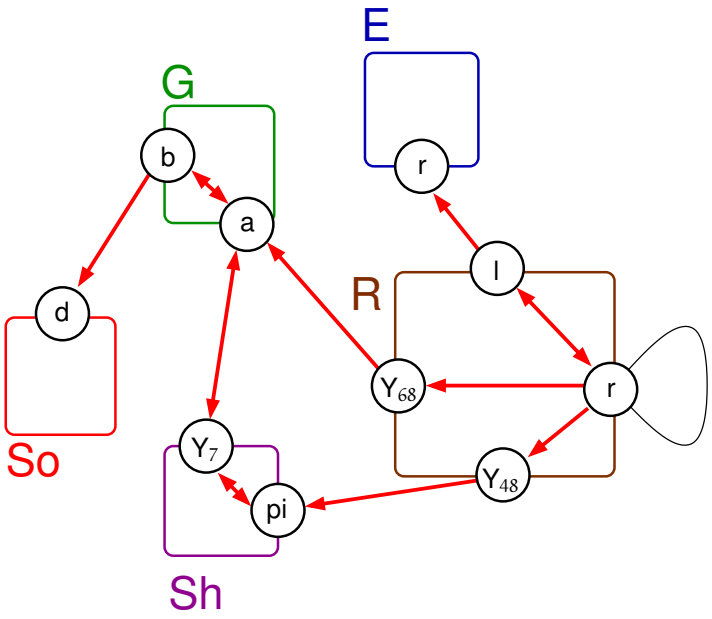
Are they fragments ?





Are they fragments ?





Basic properties

1. We call a sub-fragment any partial species which can be embedded into a fragment.

   Property 1 (sub-fragment) The concentration of any sub-fragment can be expressed as a linear combination of the concentration of some fragments.

2. We consider two norms $||\cdot||$ on $\mathcal{V} \to \mathbb{R}^+$ and $||\cdot||^\sharp$ on $\mathcal{V}^\sharp \to \mathbb{R}^+$.

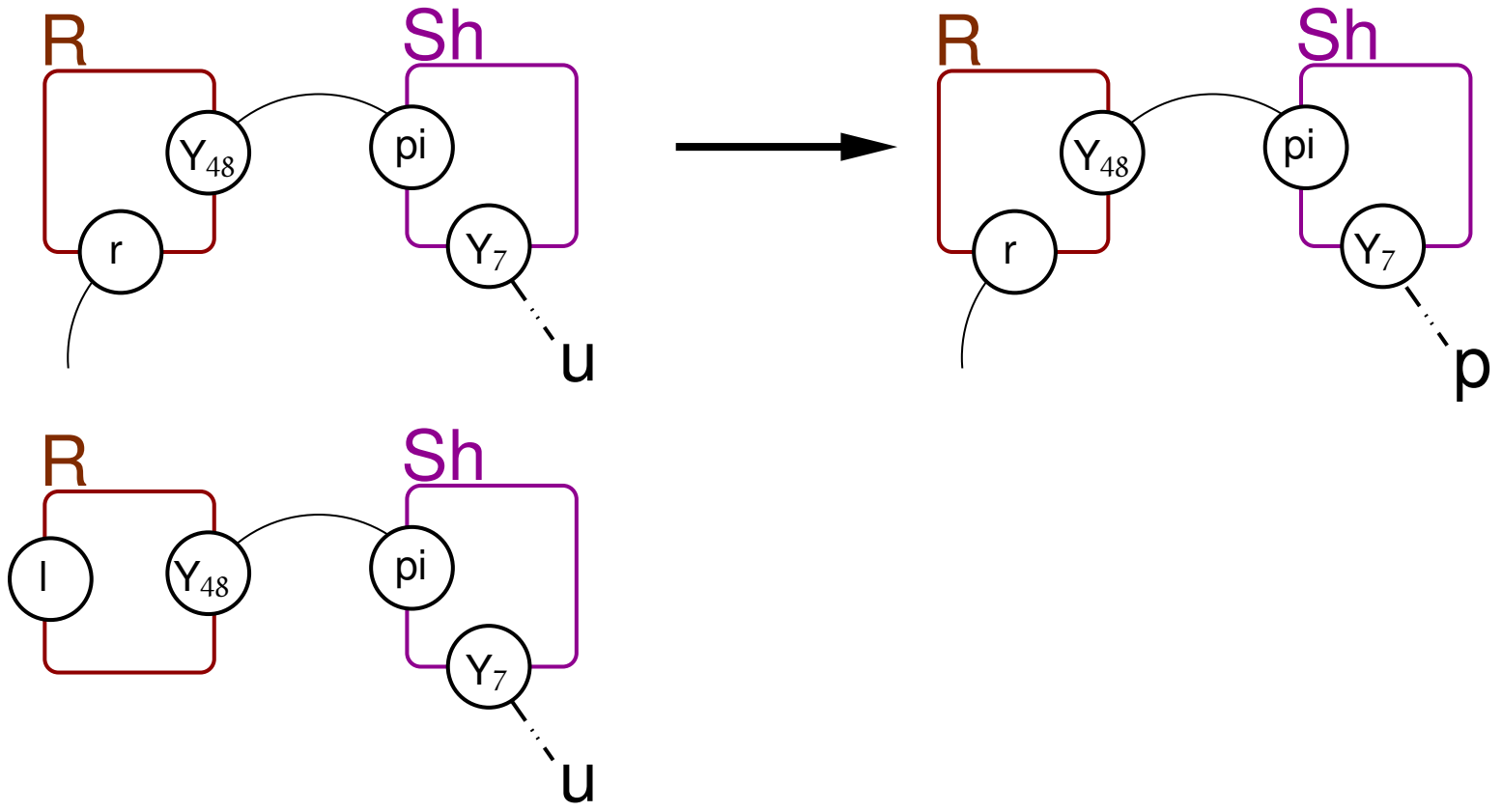
   Property 2 (non-degenerescence) Given a sequence of valuations $(x_n)_{n\in\mathbb{N}} \in (\mathcal{V} \to \mathbb{R}^+)^{\mathbb{N}}$ such that $||x_n||$ diverges toward $+\infty$, then $||\phi(x_n)||^\sharp$ diverges toward $+\infty$ as well.

Which other properties do we need so that the function $\mathbb{F}^\sharp$ can be defined ?

Overview

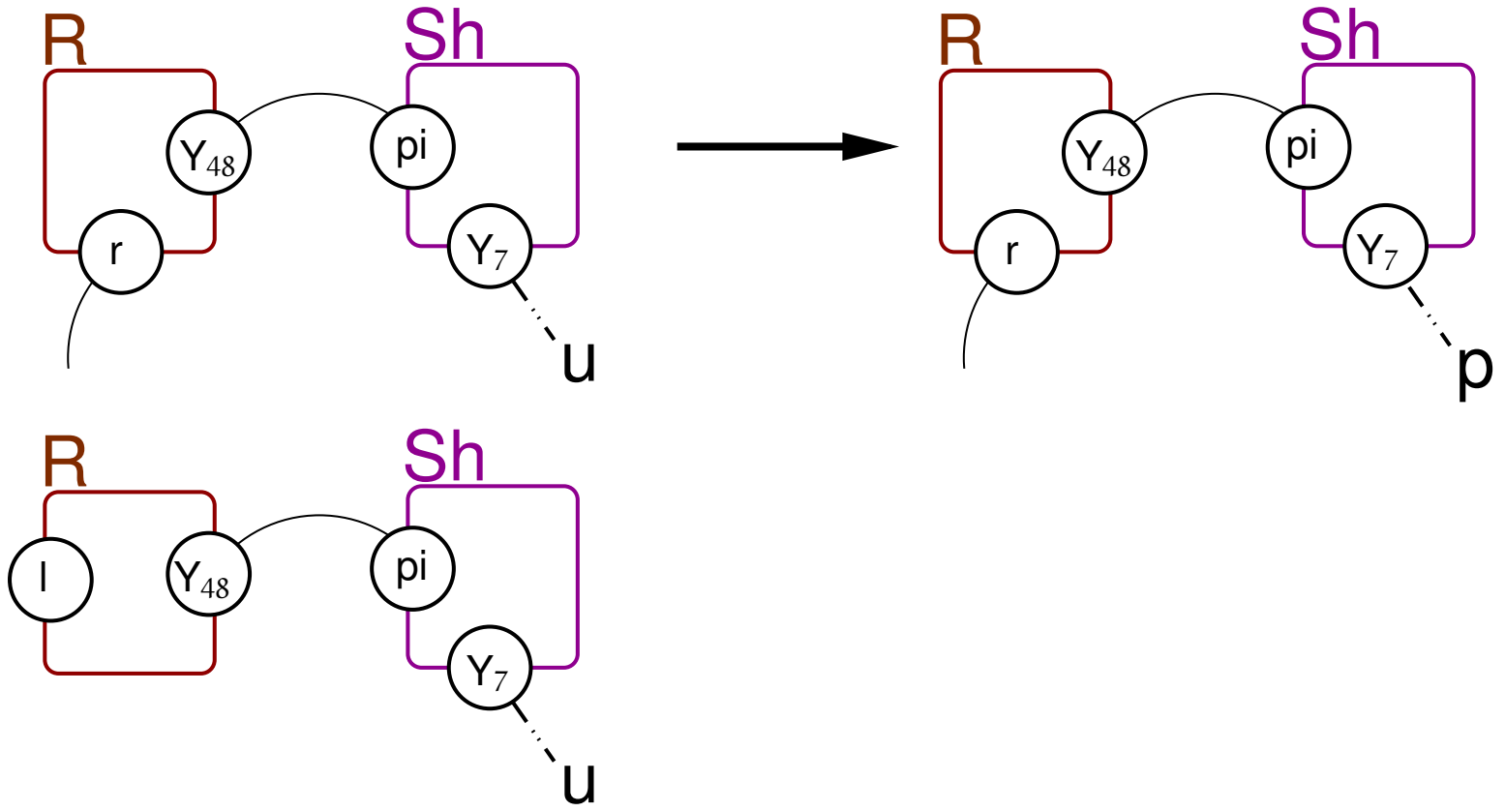
1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
   (a) Fragments
   (b) Soundness criteria
   (c) Abstract counterpart
   (d) Symmetries between sites
7. Conclusion

Fragments consumption



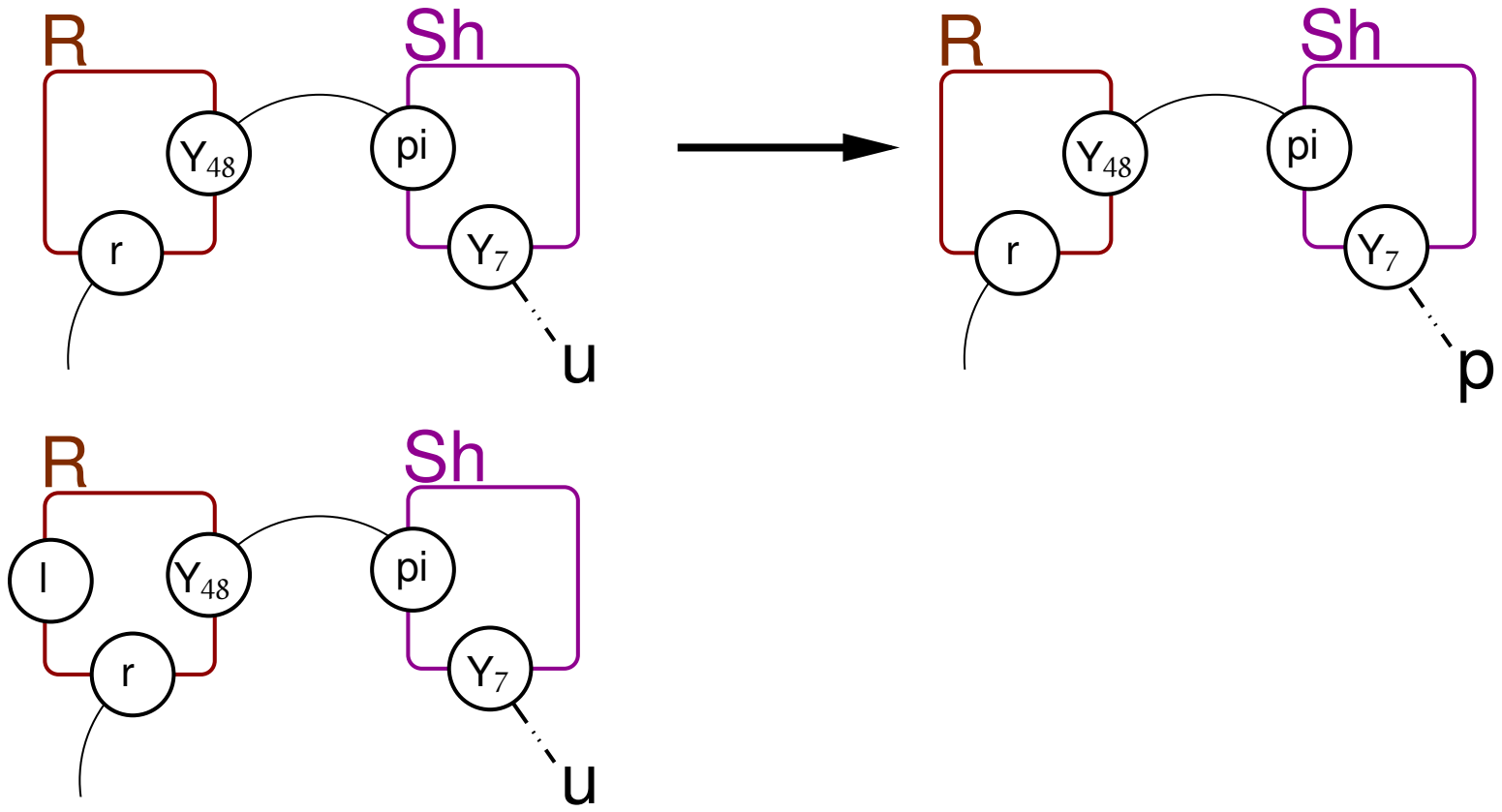
Can we express the amount (per time unit) of this fragment (bellow) concentration that is consumed by this rule (above)?

Fragments consumption



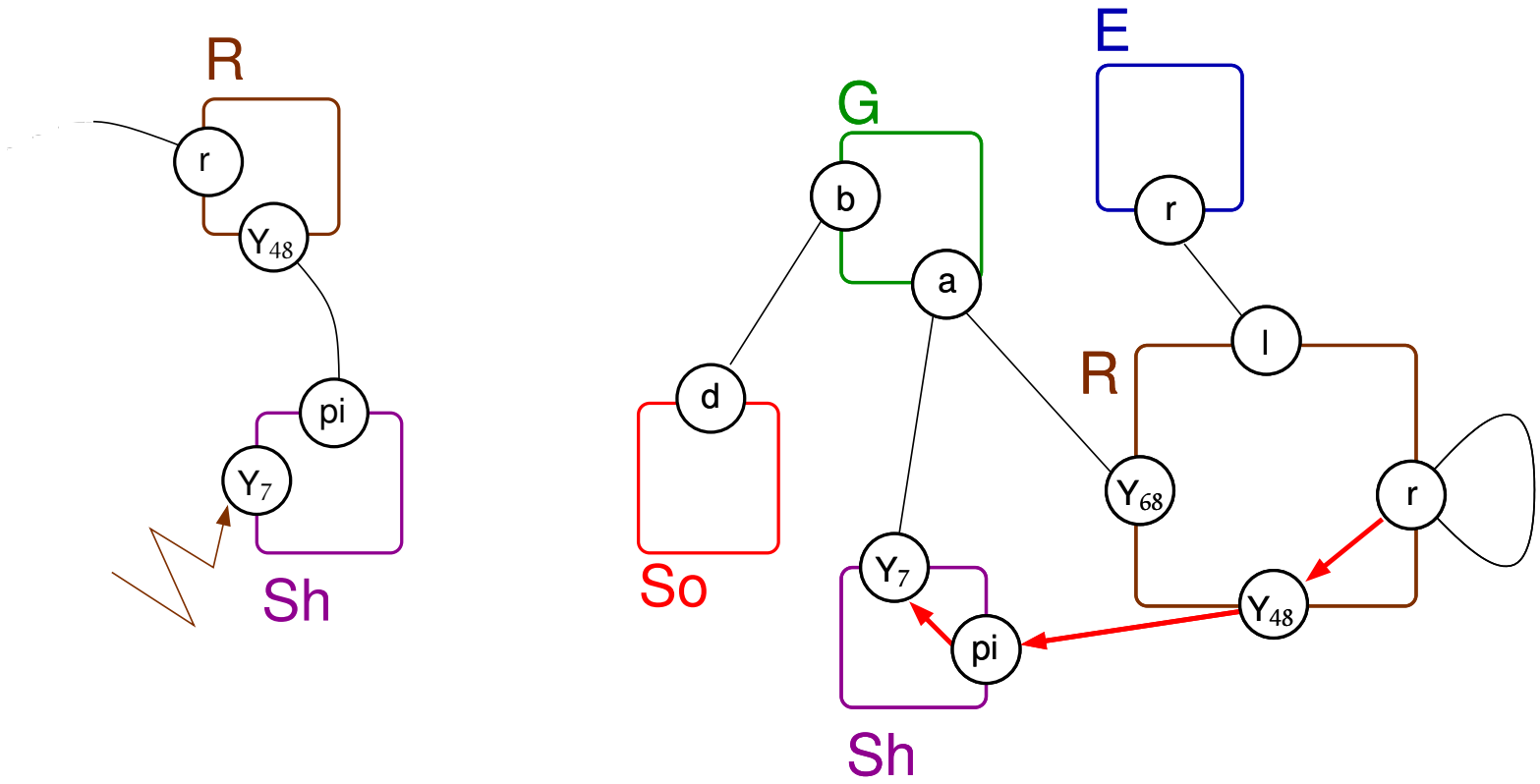
No, because we have abstracted away the correlation between the state of the site `r` and the state of the site `l`.

Fragments consumption
Proper intersection



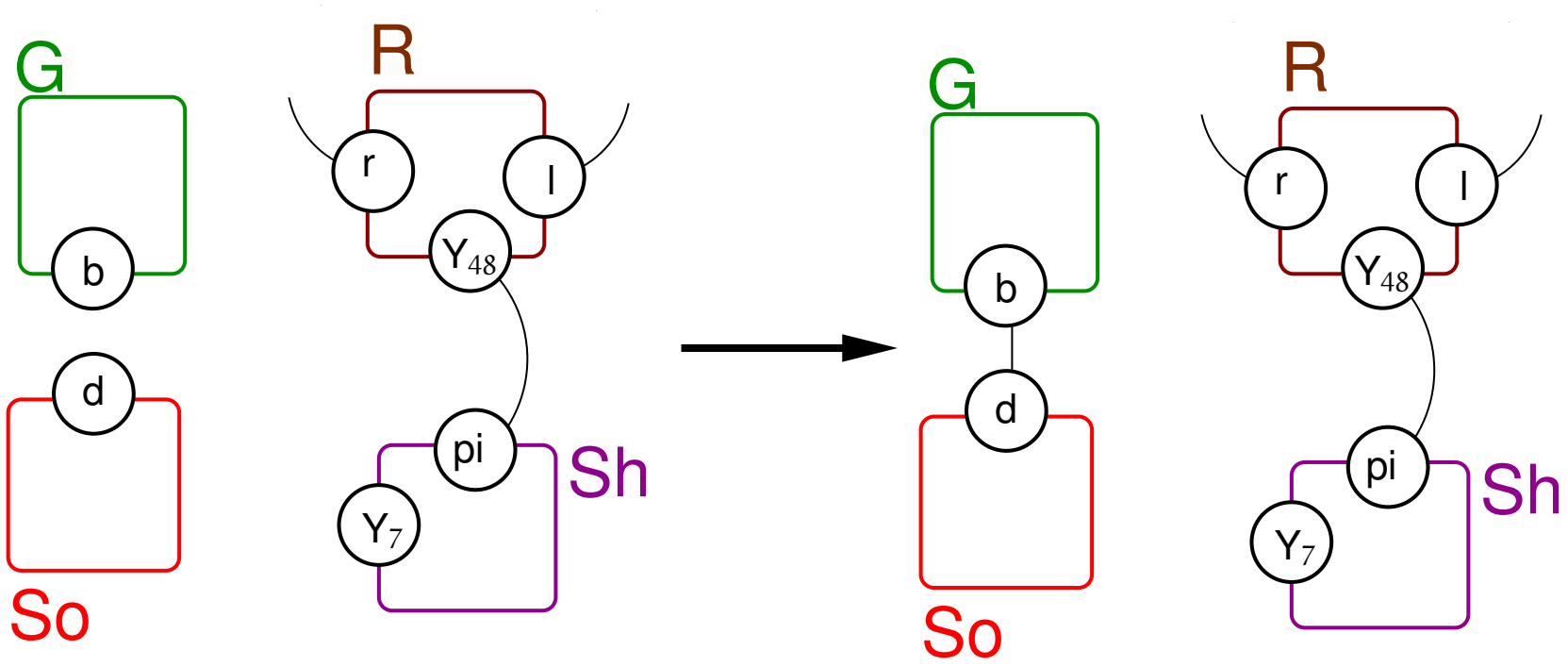
Whenever a fragment intersects a connected component of a lhs on a modified site, then the connected component must be embedded in the fragment!

Fragment consumption
Syntactic criteria



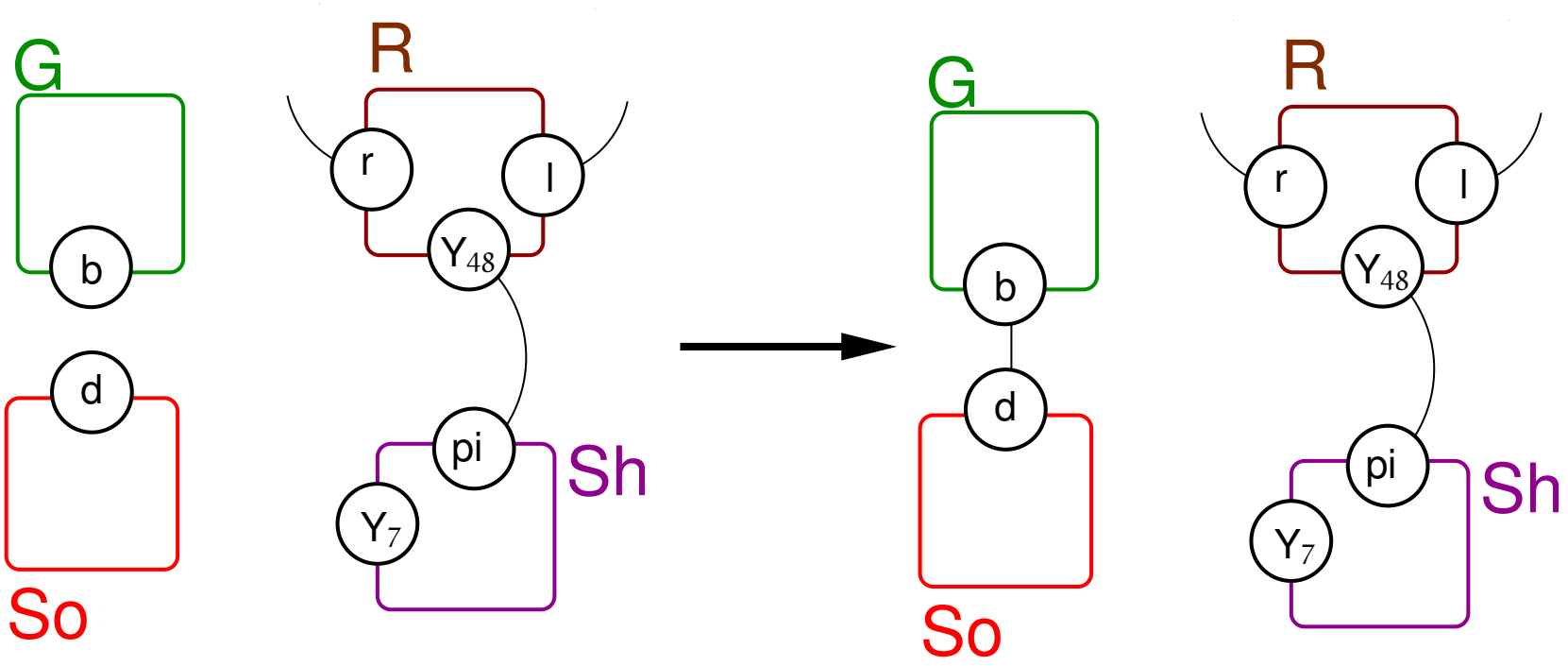
We reflect, in the annotated contact map, each path that stems from a tested site to a modified site (in the lhs of a rule).

Connected components



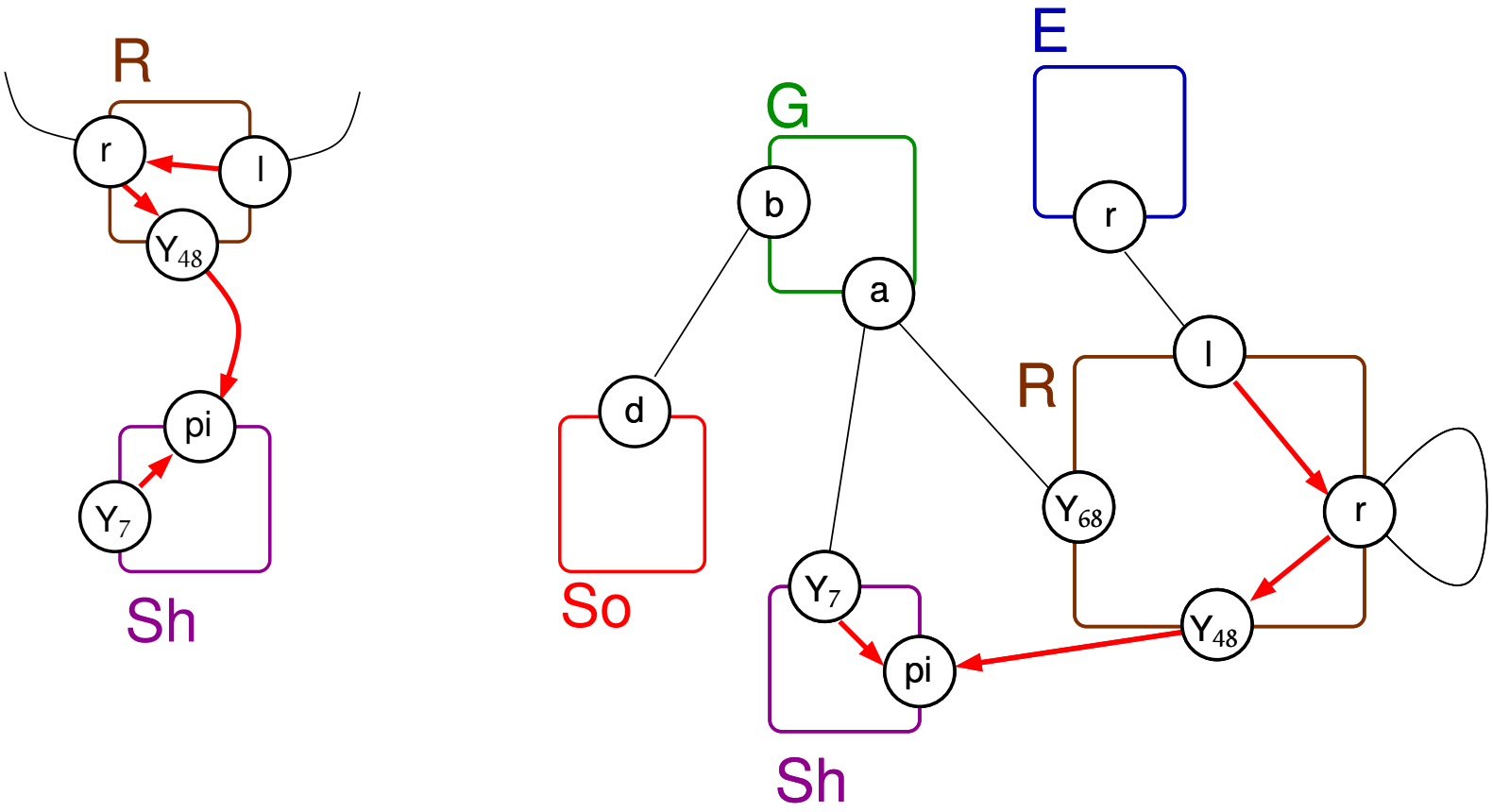
We need to express the “concentration” of any connected component of a lhs with respect to the “concentration” of fragments.

Connected components
Sub-fragment



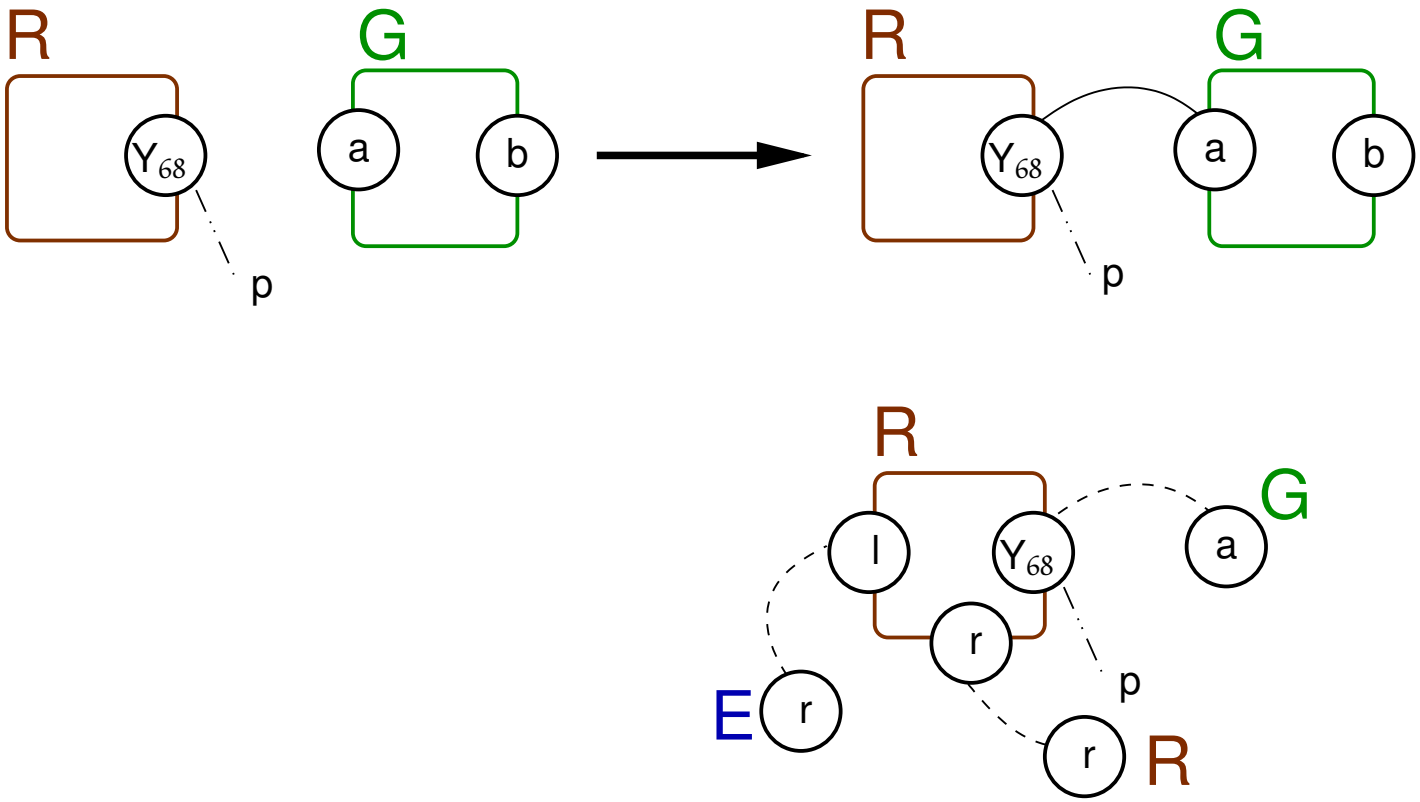
Each connected component of a lhs must be a sub-fragment.

Connected components
Syntactic criteria



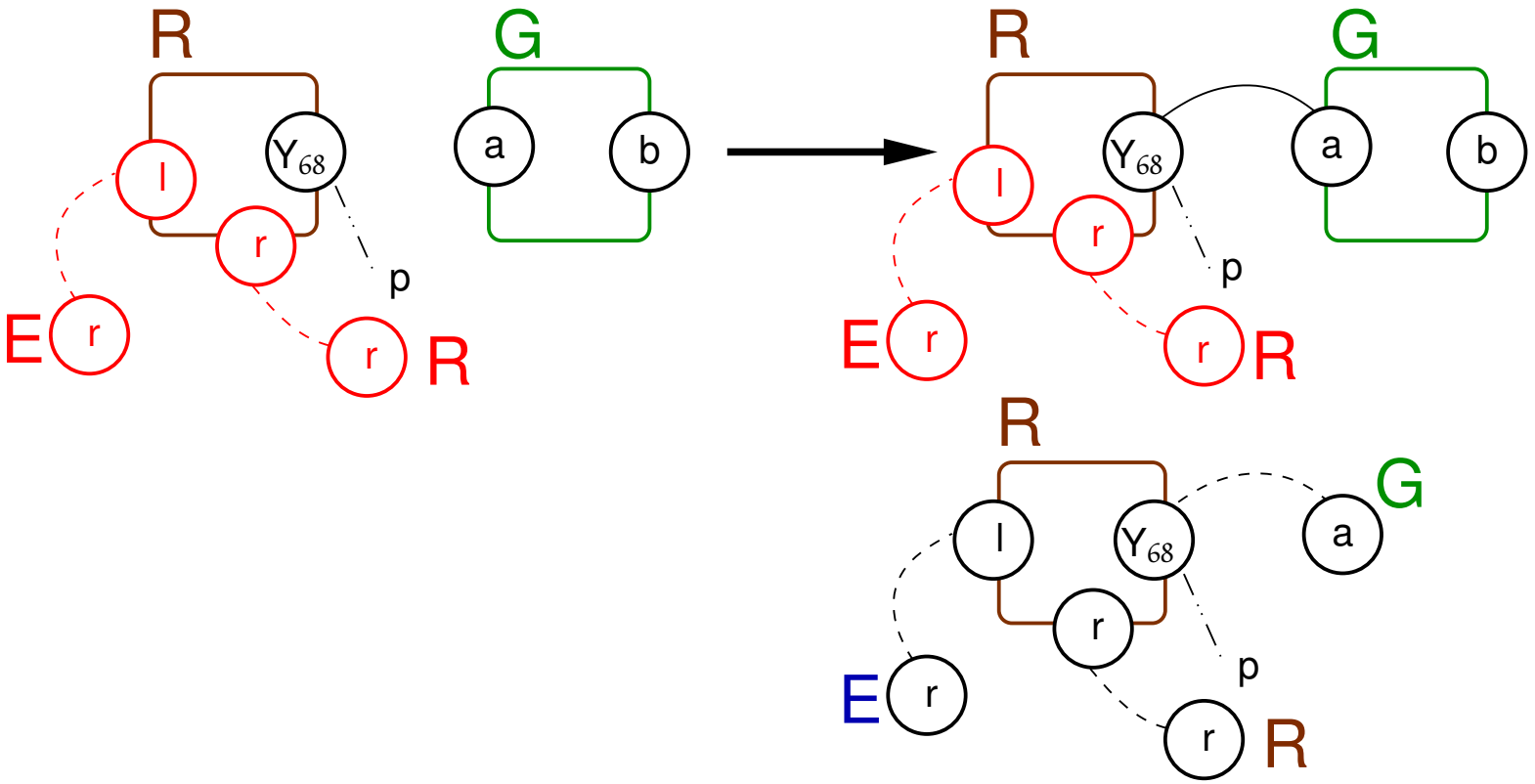
For each connected component of a lhs, there must exists a spanning tree, which reflects it-self in the annotated contact map.

Fragments production



Can we express the amount (per time unit) of this fragment (bellow) concentration that is produced by the rule (above)?

Fragments production
Proper intersection (bis)



Yes, if the connected components of the lhs of the refinement are sub-fragments.
This is already satisfied thans to the previous syntactic criteria.

Fragment properties

If:

- an annotated contact map satisfies the syntactic criteria,
- fragments are defined by this annotated contact map,
- we know the concentration of fragments;

then:

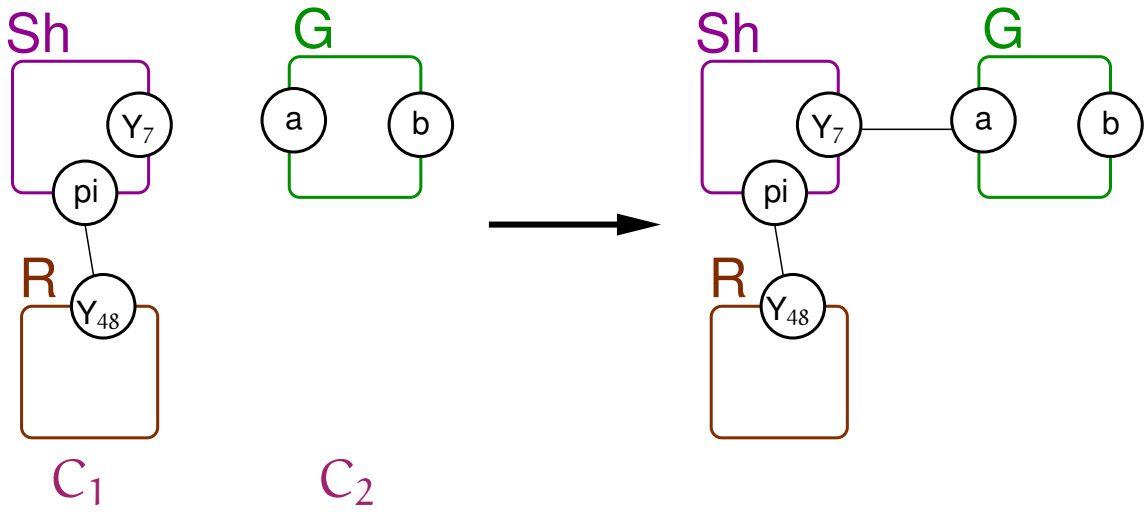
- we can express the concentration of any connected component occuring in lhss,
- we can express fragment proper consumption,
- we can express fragment proper production (eg. see the LICS'2010 paper),
- WE HAVE A CONSTRUCTIVE DEFINITION FOR $\mathbb{F}^\sharp$.

Overview

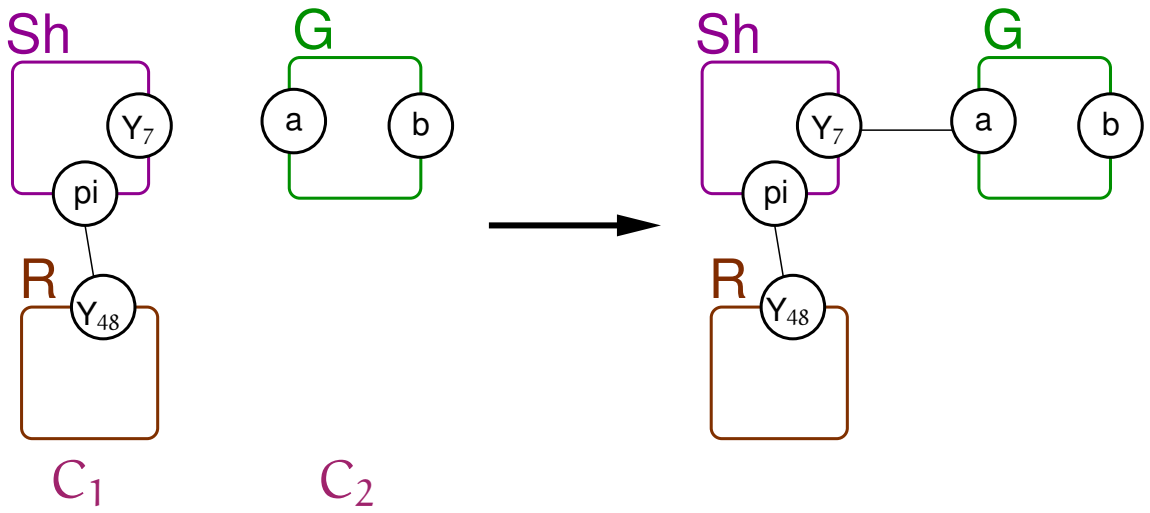
1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
   (a) Fragments
   (b) Soundness criteria
   (c) Abstract counterpart
   (d) Symmetries between sites
7. Conclusion

A binding rule

Let us abstract the contribution of a binding rule:



A binding rule: reactants

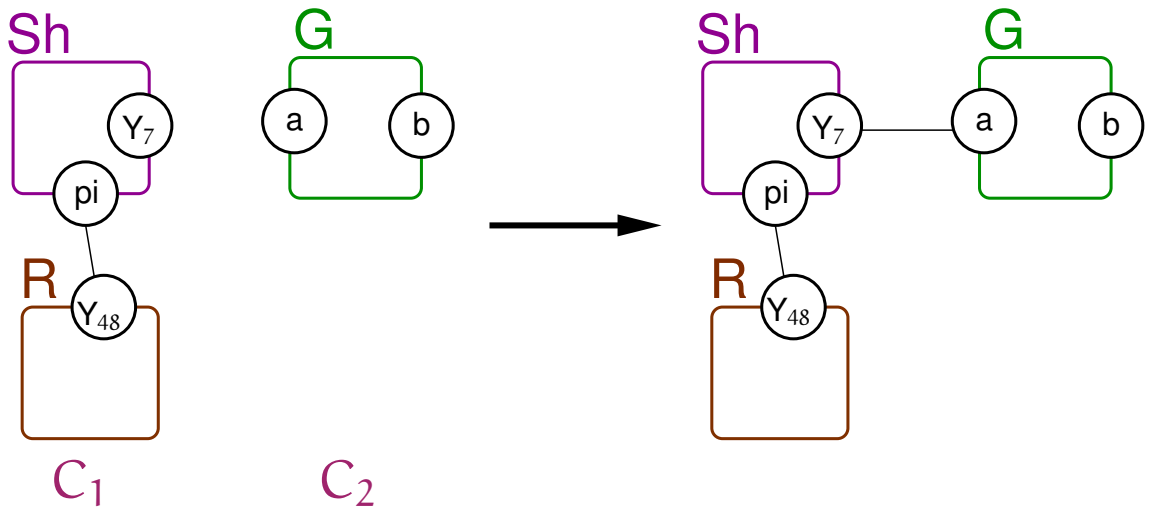


k

For any (F, Φ) such that $C_i \lhd_\Phi F$,

$$\frac{d[F]}{dt} \stackrel{+}{=} -\frac{k \cdot [F] \cdot [C_{3-i}]}{\sharp\{\Phi' \mid C_1, C_2 \lhd_{\Phi'} C_1, C_2\}}.$$

Binding rules: products

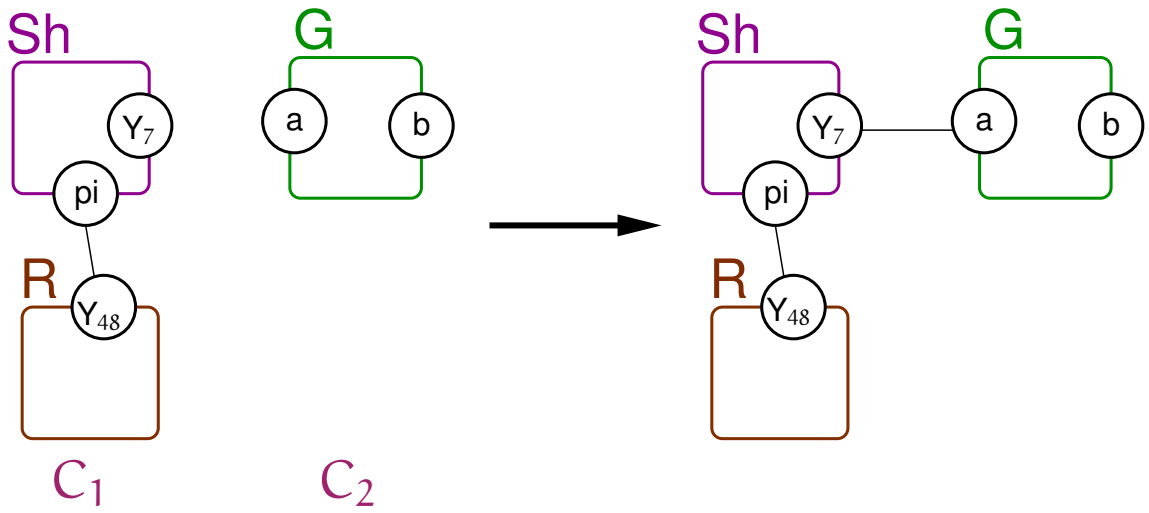


k

If the edge is solid, for any (F_1, Φ_1) and (F_2, Φ_2), such that $C_1 \lhd_{\Phi_1} F_1$ and $C_2 \lhd_{\Phi_2} F_2$,

$$\frac{d[F_1 - F_2]}{dt} = \frac{k \cdot [F_1] \cdot [F_2]}{\sharp\{\Phi' \mid C_1, C_2 \lhd_{\Phi'} C_1, C_2\}}$$

Binding rules: products



k

If the edge is dotted, for any (F, Φ) such that $C_i \lhd_\Phi F$,

$$\frac{d[F—]}{dt} \stackrel{+}{=} \frac{k \cdot [F] \cdot [C_{3-i}]}{\sharp\{\Phi' \mid C_1, C_2 \lhd_{\Phi'} C_1, C_2\}}.$$

Overview

1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
   (a) Fragments
   (b) Soundness criteria
   (c) Abstract counterpart
   (d) Symmetries between sites
7. Conclusion

Symmetries among sites

Let $\mathcal{R}$ be a set of rules.

Two sites x_1 and x_2 are symmetric in the agent A in the set of rules $\mathcal{R}$,
$\overset{\Delta}{\Longleftrightarrow}$
$\mathcal{R}$ is preserved (modulo $\equiv$) if we replace each rule with all the combinations of rules which can be obtained by replacing (independently) each occurrence of x_1 and x_2 with x_1 or x_2 (and dividing kinetic rate by the number of combinations, and taking care of gain/loss of automorphisms).

Example I

$$
\begin{array}{lcll}
A\left(x_{u}\right) & \longrightarrow & A\left(x_{p}\right) & k_{1} \\
A\left(y_{u}\right) & \longrightarrow & A\left(y_{p}\right) & k_{1} \\
A\left(x_{p}, y_{p}\right) & \longrightarrow & & k_{2}
\end{array}
$$

Example I

$$
\begin{array}{lcll}
A\left(x_u\right) & \longrightarrow & A\left(x_p\right) & k_1 \\
A\left(y_u\right) & \longrightarrow & A\left(y_p\right) & k_1 \\
A\left(x_p, y_p\right) & \longrightarrow & & k_2
\end{array}
$$

We get:

$$
\begin{array}{lcll}
A\left(x_u\right) & \longrightarrow & A\left(x_p\right) & \frac{k_1}{2} + \frac{k_1}{2} \\
A\left(y_u\right) & \longrightarrow & A\left(y_p\right) & \frac{k_1}{2} + \frac{k_1}{2} \\
A\left(x_p, y_p\right) & \longrightarrow & & \frac{k_2}{2} + \frac{k_2}{2}
\end{array}
$$

So, x and y are symmetric in A!

Example II

$$
\begin{array}{lcll}
A\left(x_u\right) & \longrightarrow & A\left(x_p\right) & k_1 \\
A\left(y_u\right) & \longrightarrow & A\left(y_p\right) & k_2 \\
A\left(x_p, y_p\right) & \longrightarrow & & k_3
\end{array}
$$

Example II

$$
\begin{array}{lcll}
A\left(x_u\right) & \longrightarrow & A\left(x_p\right) & k_1 \\
A\left(y_u\right) & \longrightarrow & A\left(y_p\right) & k_2 \\
A\left(x_p, y_p\right) & \longrightarrow & & k_3
\end{array}
$$

We get:

$$
\begin{array}{lcll}
A\left(x_u\right) & \longrightarrow & A\left(x_p\right) & \frac{k_1}{2} + \frac{k_2}{2} \\
A\left(y_u\right) & \longrightarrow & A\left(y_p\right) & \frac{k_1}{2} + \frac{k_2}{2} \\
A\left(x_p, y_p\right) & \longrightarrow & & \frac{k_3}{2} + \frac{k_3}{2}
\end{array}
$$

So, x and y are symmetric in A, if and only if $k_1 = k_2$!

Example III

$$
\begin{array}{llll}
A\left(x\right) , A\left(x\right) & \longrightarrow & A\left(x^1\right) , A\left(x^1\right) & \mathrm{k} \\
A\left(y\right) , A\left(y\right) & \longrightarrow & A\left(y^1\right) , A\left(y^1\right) & \mathrm{k}
\end{array}
$$

Example III

$$
\begin{array}{llll}
A(x) \ , A(x) & \longrightarrow & A(x^1) \ , A(x^1) & \mathrm{k} \\
A(y) \ , A(y) & \longrightarrow & A(y^1) \ , A(y^1) & \mathrm{k}
\end{array}
$$

We get:

$$
\begin{array}{llll}
A(x) \ , A(x) & \longrightarrow & A(x^1) \ , A(x^1) & \frac{\mathrm{k}}{2} \\
A(y) \ , A(y) & \longrightarrow & A(y^1) \ , A(y^1) & \frac{\mathrm{k}}{2} \\
A(x) \ , A(y) & \longrightarrow & A(x^1) \ , A(y^1) & \frac{\mathrm{k}}{2}
\end{array}
$$

So, x and y are symmetric in A, if and only if $\mathrm{k} = 0$!

Example IV

$$
\begin{array}{llll}
A(x) \ , A(x) & \longrightarrow & A(x^1) \ , A(x^1) & k_1 \\
A(y) \ , A(y) & \longrightarrow & A(y^1) \ , A(y^1) & k_2 \\
A(x) \ , A(y) & \longrightarrow & A(x^1) \ , A(y^1) & k_3
\end{array}
$$

Example IV

$$
\begin{array}{llll}
A(x) \ , A(x) & \longrightarrow & A(x^1) \ , A(x^1) & k_1 \\
A(y) \ , A(y) & \longrightarrow & A(y^1) \ , A(y^1) & k_2 \\
A(x) \ , A(y) & \longrightarrow & A(x^1) \ , A(y^1) & k_3
\end{array}
$$

We get:

$$
\begin{array}{llll}
A(x) \ , A(x) & \longrightarrow & A(x^1) \ , A(x^1) & \frac{k_1}{4} + \frac{k_2}{4} + \frac{k_3}{2} \\
A(y) \ , A(y) & \longrightarrow & A(y^1) \ , A(y^1) & \frac{k_1}{4} + \frac{k_2}{4} + \frac{k_3}{2} \\
A(x) \ , A(y) & \longrightarrow & A(x^1) \ , A(y^1) & \frac{k_1}{4} + \frac{k_2}{4} + \frac{k_3}{2}
\end{array}
$$

So, x and y are symmetric in A, if and only if $k_1 = k_2 = k_3$!

Symmetries among sites

- We consider a family of triples $(x_i, y_i, A_i)_{i \in I}$ such that, for each $i \in I$:
  - x_i and y_i are symmetric in the agent A_i;
  - x_i and y_i are connected in both directions in the annotated contact map;
- We define $\sim_{ag}$ over agents (with interfaces) by $A_i(\sigma[x_i, y_i]) \sim_{ag} A_i(\sigma[y_i, x_i])$.
- We define $\sim_{pattern}$ over expressions by:

$$\frac{A_i \sim_{ag} A'_i, 1 \leq i \leq k}{A_1, \ldots, A_k \sim_{pattern} A'_1, \ldots, A'_k}.$$

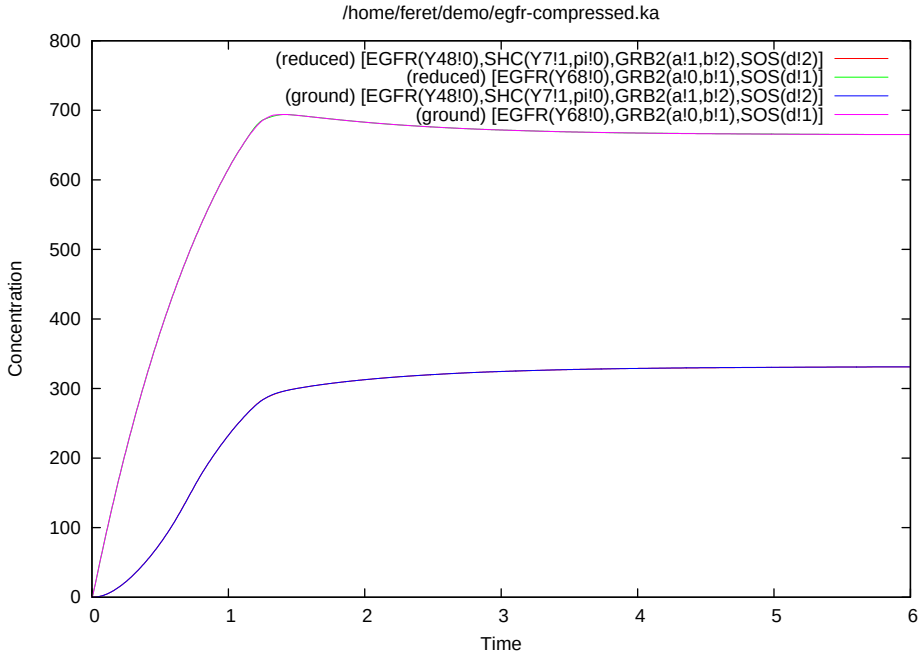
- Then, it is (quite) easy to build $r \in \mathcal{V} \to \mathcal{V}$ and $r^\sharp \in \mathcal{V}^\sharp \to \mathcal{V}^\sharp$, such that:
  1. for any $X \in \mathcal{V}$, $r(X) \sim_{pattern} X$,
  2. for any $F \in \mathcal{V}^\sharp$, $r^\sharp(F) \sim_{pattern} F$,
  3. and $\psi \circ P_r = P^\sharp_{r^\sharp} \circ \psi$.

Overview

1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

Experimental results

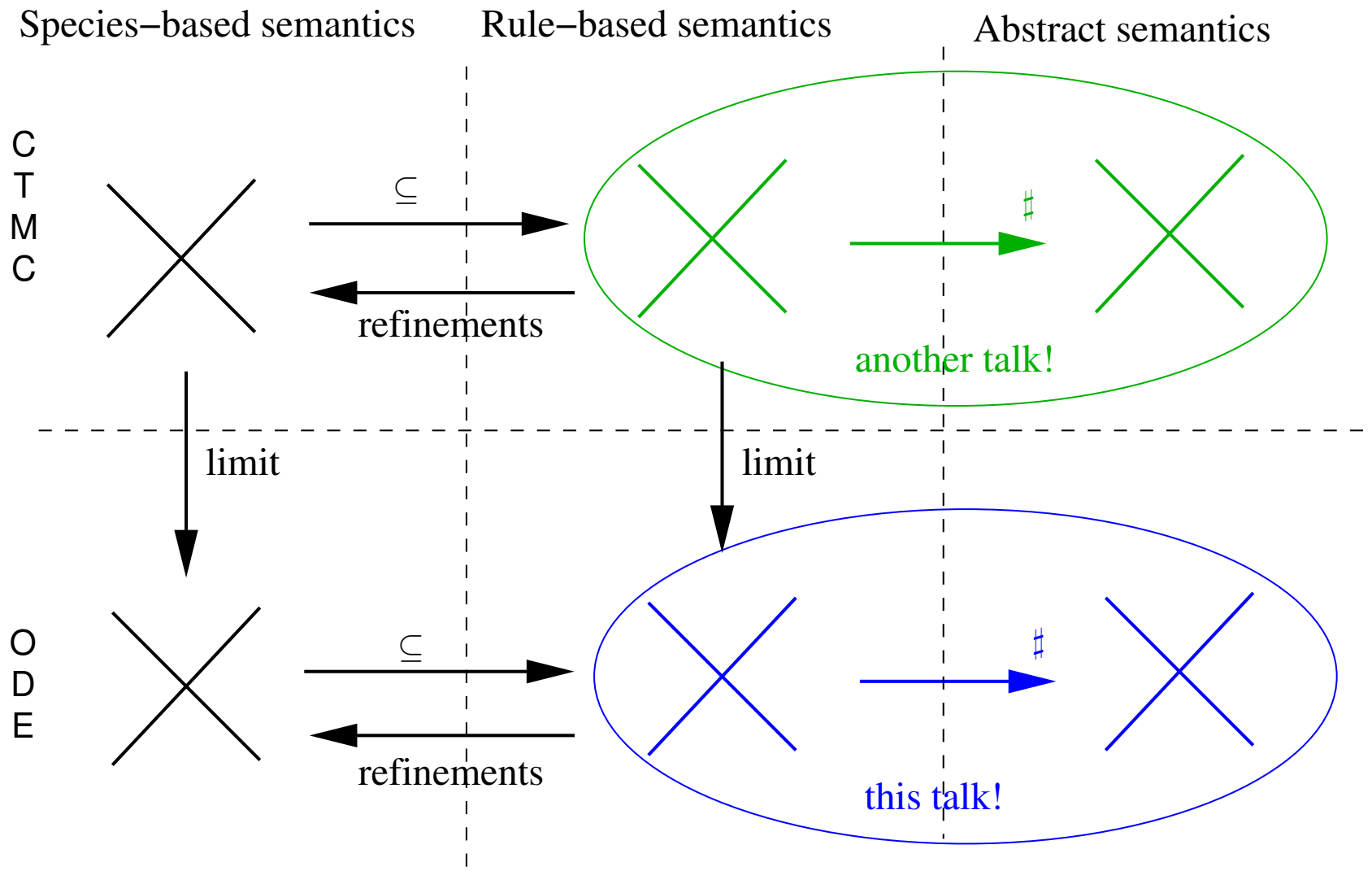
On early egfr, 356 species are simplified into 38 fragments:



Superposition of the ground and the abstract differential semantics.

On a bigger example, $\approx 2 \cdot 10^{19}$ species are simplified into $\approx 2 \cdot 10^{5}$ fragments.

Related issues I: Semantics comparisons



Related issues II: Semantics approximations

1. ODE approximations:
   - Concrete definition of the control flow and hierarchy of abstractions. (A notion of control flow which would be invariant by:
     - neutral rule refinement;
     - compilation of a Kappa system into a Kappa system with only one agent type.

   Joint work with Ferdinanda Camporesi (Bologna)
2. Stochastic semantics approximations:
   - Can we design abstraction ?
   - Find the adequate soundness criteria.

   Joint work with Thomas Henzinger (IST-Vienna), Heinz Koeppl (ETH-Zurich), Tatjana Petrov (EPFL)

Università di Bologna

Fourth day

Model reduction of stochastic rules-based models

[CS2Bio'10,MFPS'10,MeCBIC'10,ICNAAM'10]

Jérôme Feret

Laboratoire d'Informatique de l'École Normale Supérieure

INRIA, ÉNS, CNRS

Thursday, April the 21st

Joint-work with...

Ferdinanda Camporesi
Bologna / ÉNS

Thomas Henzinger
IST Austria

Heinz Koeppl
ETH Zürich

Tatjana Petrov
EPFL

Overview

1. Introduction
2. Examples of information flow
3. Symmetric sites
4. Stochastic semantics
5. Lumpability
6. Bisimulations
7. Hierarchy of semantics
8. Conclusion

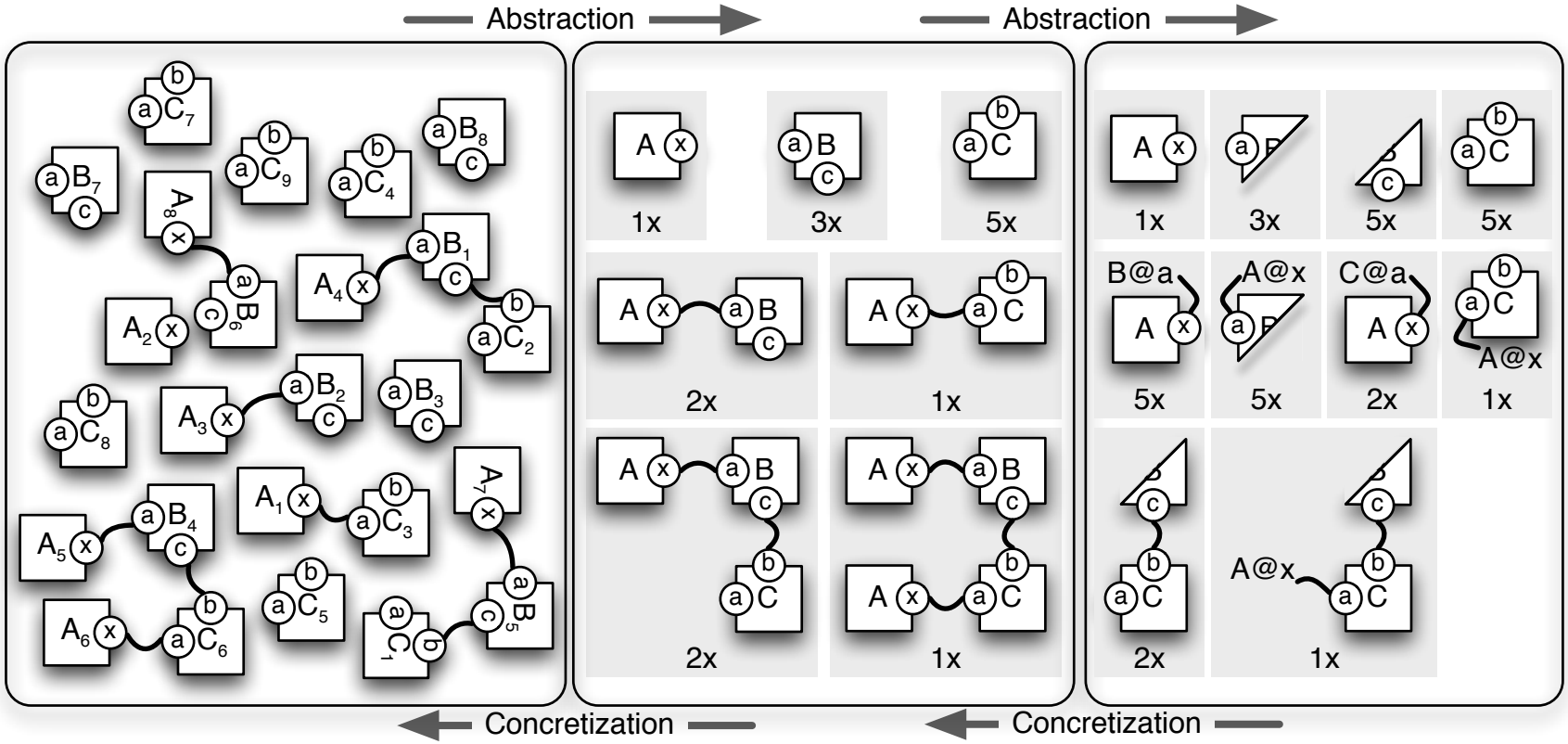
ODE fragments

In the ODE semantics, using the flow of information backward, we can detect which correlations are not relevant for the system, and deduce a small set of portions of chemical species (called fragments) the behavior of the concentration of which can be described in a self-consistent way.

(ie. the trajectory of the reduced model are the exact projection of the trajectory of the initial model).

Can we do the same for the stochastic semantics?

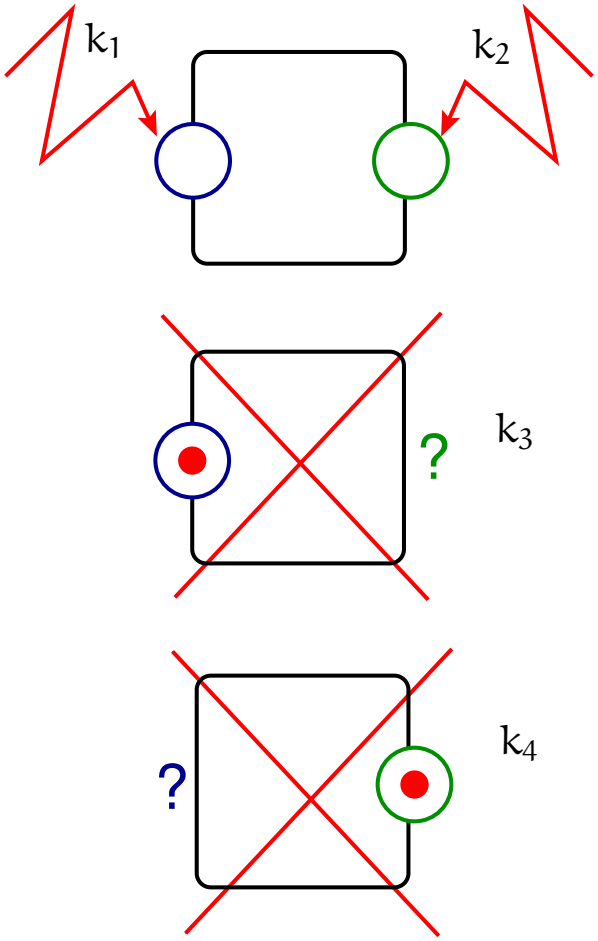
Stochastic fragments ?



Overview

1. Introduction
2. Examples of information flow
3. Symmetric sites
4. Stochastic semantics
5. Lumpability
6. Bisimulations
7. Hierarchy of semantics
8. Conclusion

A model with ubiquitination



$$P \xrightarrow{k_1} {}^{\star}P \qquad P^{\star} \xrightarrow{k_1} {}^{\star}P^{\star}$$

$$P \xrightarrow{k_2} P^{\star} \qquad {}^{\star}P \xrightarrow{k_2} {}^{\star}P^{\star}$$

$${}^{\star}P \xrightarrow{k_3} \emptyset$$

$${}^{\star}P^{\star} \xrightarrow{k_3} \emptyset$$

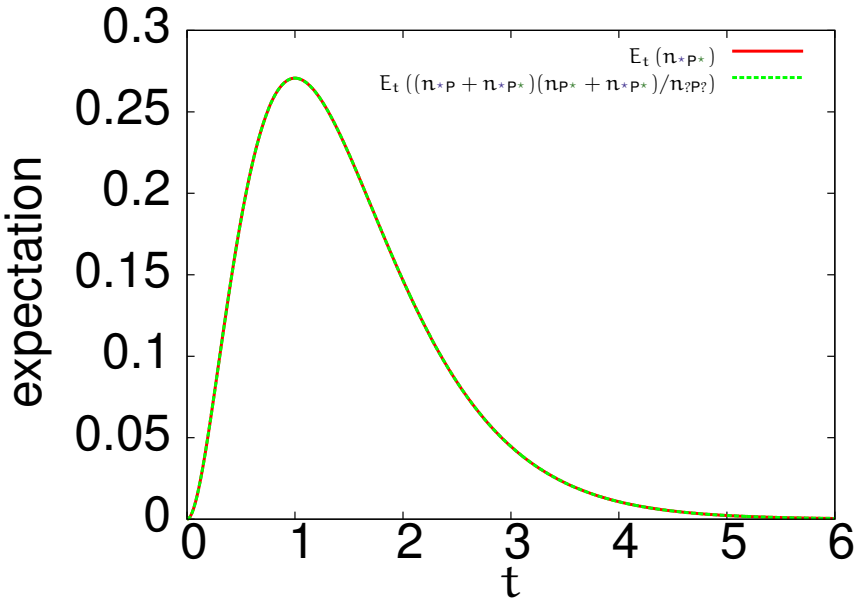
$$P^{\star} \xrightarrow{k_4} \emptyset$$

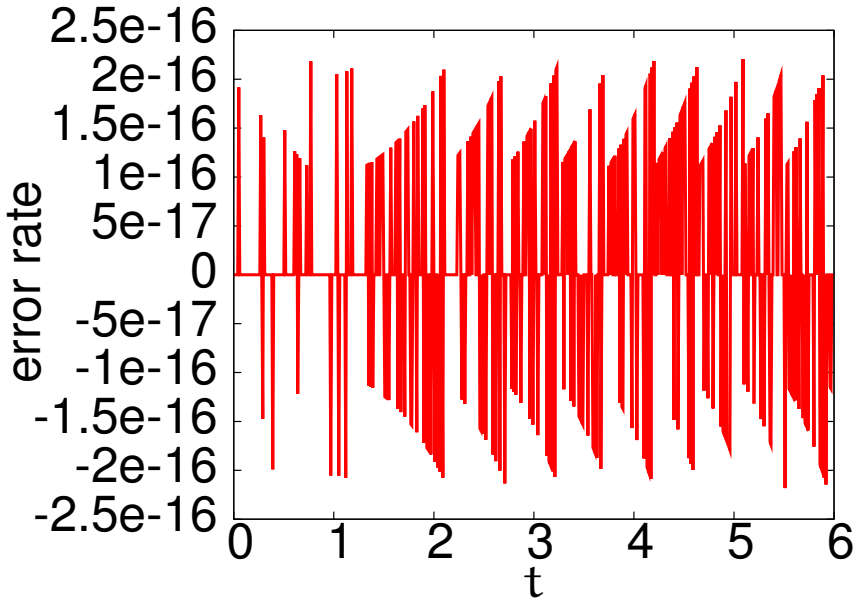
$${}^{\star}P^{\star} \xrightarrow{k_4} \emptyset$$

Statistical independence

We check numerically that:

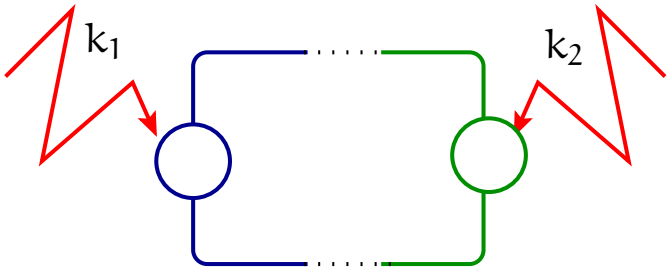
$$\mathrm{E}_t\left(n_{\star P\star}\right) = \mathrm{E}_t\left(\frac{(n_{\star P} + n_{\star P\star})(n_{P\star} + n_{\star P\star})}{n_P + n_{P\star} + n_{\star P} + n_{\star P\star}}\right).$$





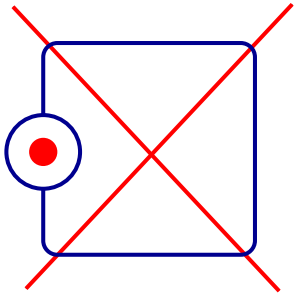
with $k_1 = k_2 = k_3 = k_4 = 1$
and two instances of P at time $t = 0$.

Reduced model



$$P \xrightarrow{k_1} {}^{\star}P$$

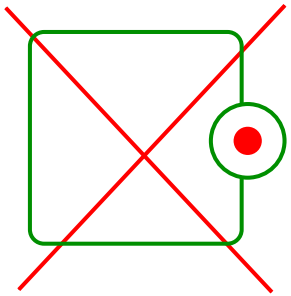
$$P \xrightarrow{k_2} P^{\star}$$



k_3

$${}^{\star}P \xrightarrow{k_3} \emptyset$$

+ side effect: remove one P

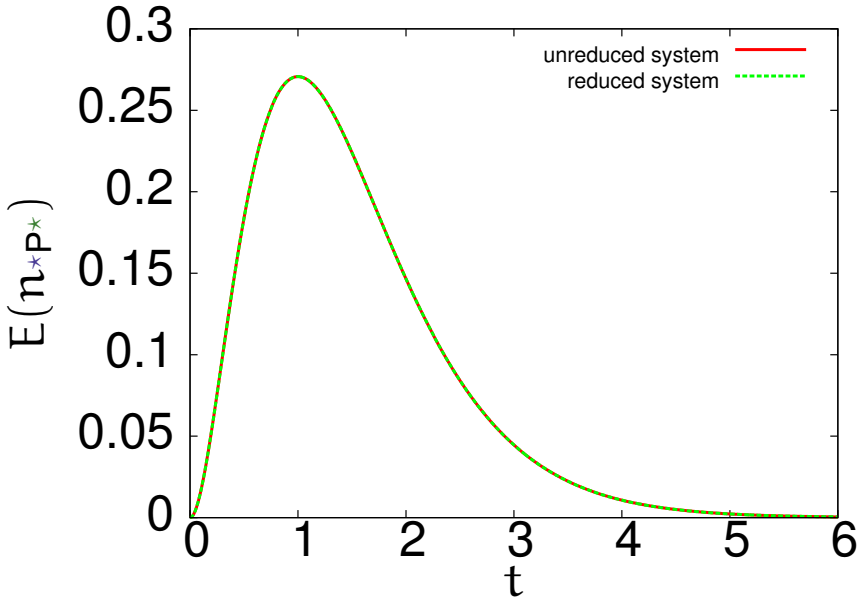


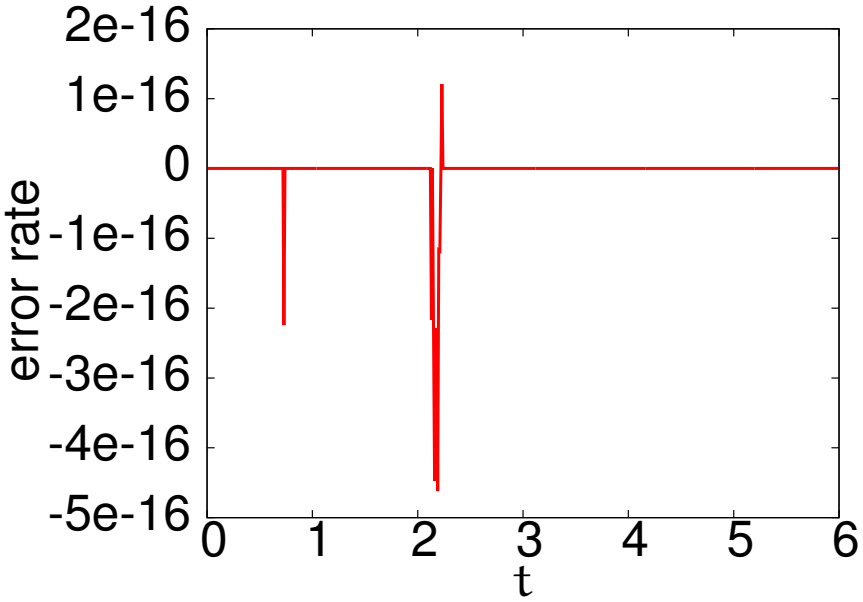
k_4

$$P^{\star} \xrightarrow{k_4} \emptyset$$

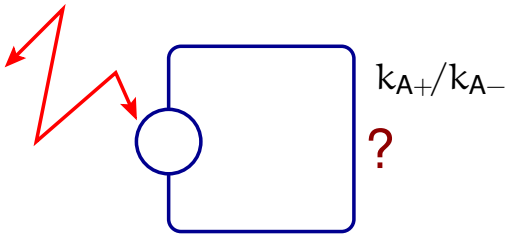
+ side effect: remove one P

Comparison between the two models

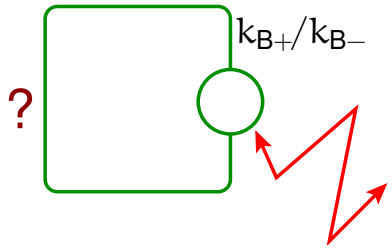




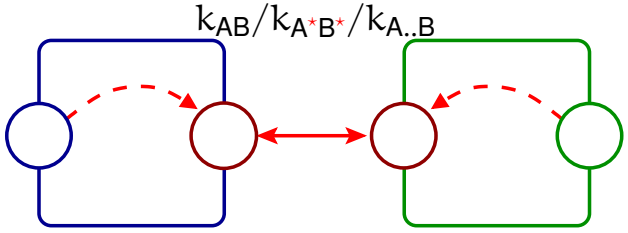
Coupled semi-reactions



$$A \underset{k_{A-}}{\overset{k_{A+}}{\rightleftharpoons}} A^\star, \quad AB \underset{k_{A-}}{\overset{k_{A+}}{\rightleftharpoons}} A^\star B, \quad AB^\star \underset{k_{A-}}{\overset{k_{A+}}{\rightleftharpoons}} A^\star B^\star$$



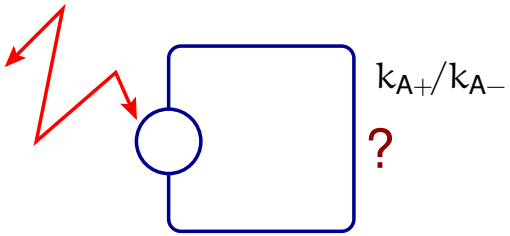
$$B \underset{k_{B-}}{\overset{k_{B+}}{\rightleftharpoons}} B^\star, \quad AB \underset{k_{B-}}{\overset{k_{B+}}{\rightleftharpoons}} AB^\star, \quad A^\star B \underset{k_{B-}}{\overset{k_{B+}}{\rightleftharpoons}} A^\star B^\star$$

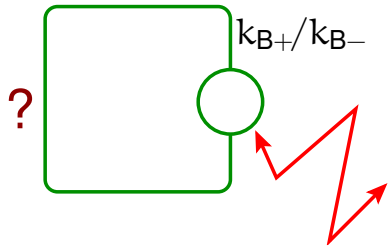


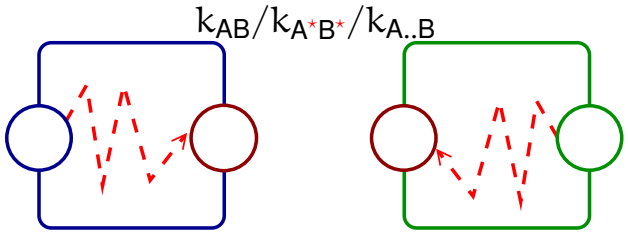
$$A + B \underset{k_{A..B}}{\overset{k_{AB}}{\rightleftharpoons}} AB, \quad A^\star + B \underset{k_{A..B}}{\overset{k_{AB}}{\rightleftharpoons}} A^\star B,$$

$$A + B^\star \underset{k_{A..B}}{\overset{k_{AB}}{\rightleftharpoons}} AB^\star, \quad A^\star + B^\star \underset{k_{A..B}}{\overset{k_{A^\star B^\star}}{\rightleftharpoons}} A^\star B^\star$$

Reduced model







$$A \underset{k_{A-}}{\overset{k_{A+}}{\rightleftharpoons}} A^\star, \quad AB^\diamond \underset{k_{A-}}{\overset{k_{A+}}{\rightleftharpoons}} A^\star B^\diamond,$$

$$B \underset{k_{B-}}{\overset{k_{B+}}{\rightleftharpoons}} B^\star, \quad A^\diamond B \underset{k_{B-}}{\overset{k_{B+}}{\rightleftharpoons}} A^\diamond B^\star,$$

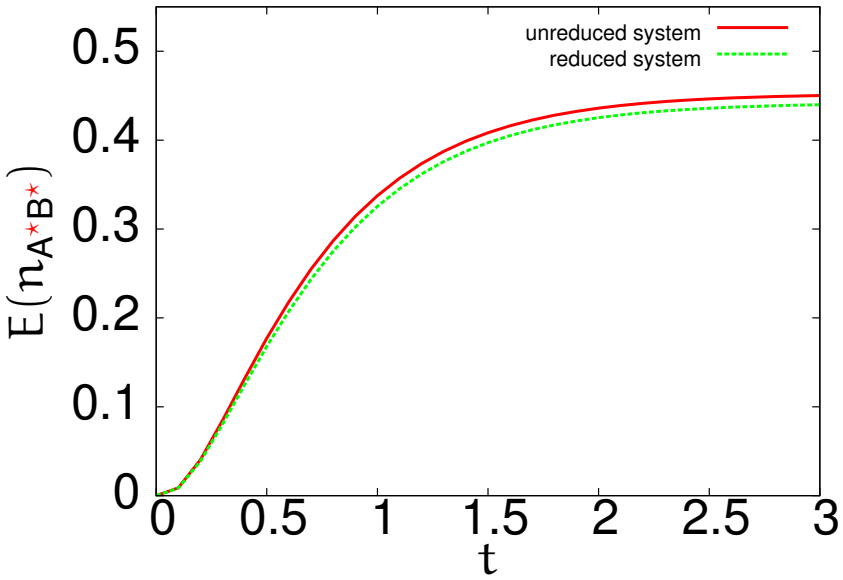
$$A + B \underset{k_{A..B}/(n_{A^\diamond B}+n_{A^\diamond B^\star})}{\overset{k_{AB}}{\rightleftharpoons}} AB^\diamond + A^\diamond B,$$

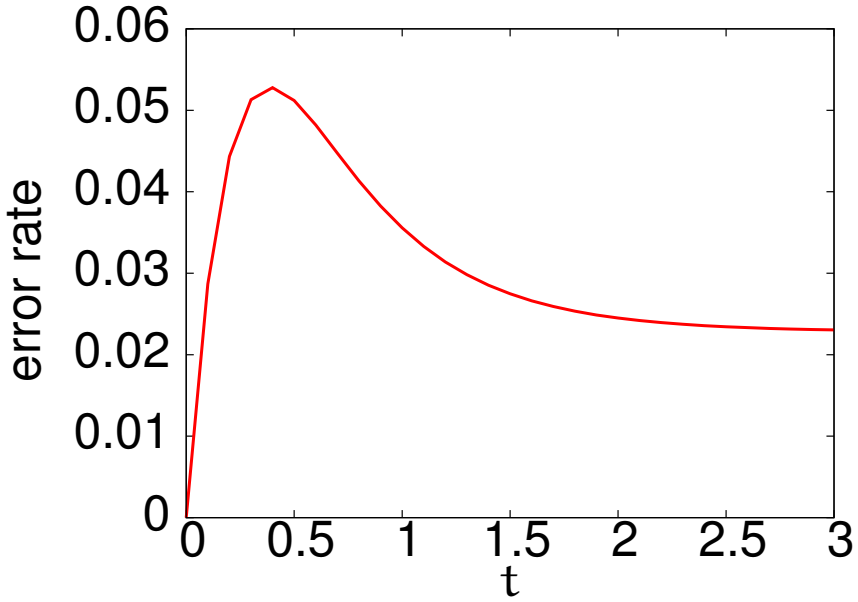
$$A^\star + B \underset{k_{A..B}/(n_{A^\diamond B}+n_{A^\diamond B^\star})}{\overset{k_{AB}}{\rightleftharpoons}} A^\star B^\diamond + A^\diamond B,$$

$$A + B^\star \underset{k_{A..B}/(n_{A^\diamond B}+n_{A^\diamond B^\star})}{\overset{k_{AB}}{\rightleftharpoons}} AB^\diamond + A^\diamond B^\star,$$

$$A^\star + B^\star \underset{k_{A..B}/(n_{A^\diamond B}+n_{A^\diamond B^\star})}{\overset{k_{A^\star B^\star}}{\rightleftharpoons}} A^\star B^\diamond + A^\diamond B^\star$$

Comparison between the two models

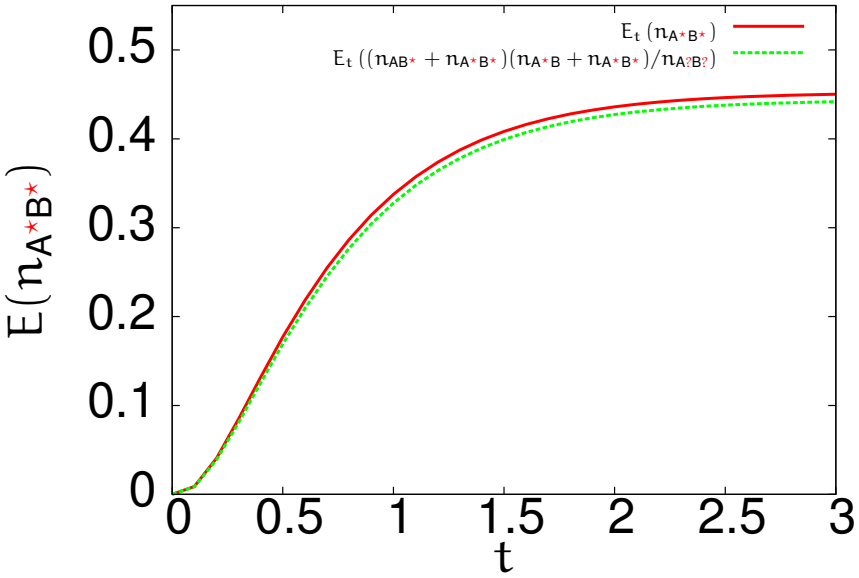


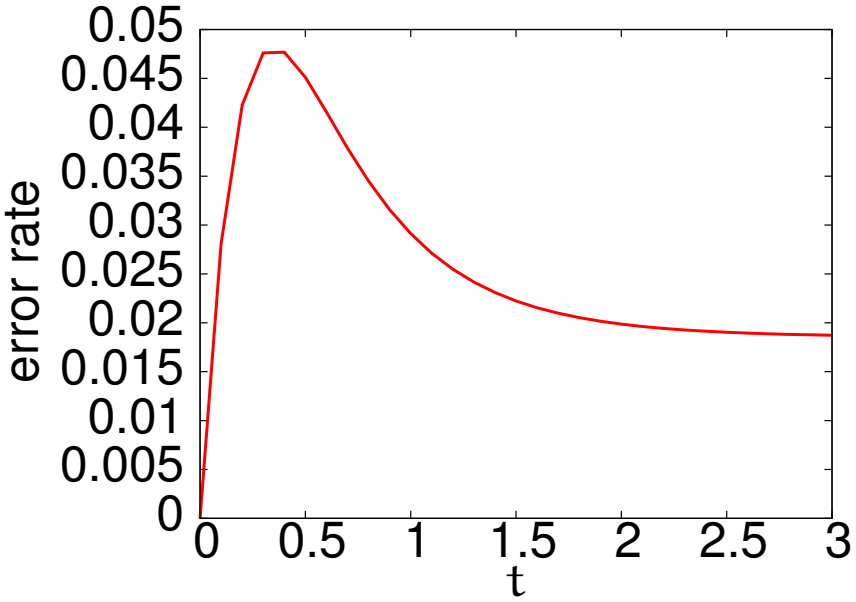


with $k_{A+} = k_{A-} = k_{B+} = k_{B-} = k_{AB} = k_{A..B} = 1$, $k_{A^\star B^\star} = 10$,
and two instances of A and B at time $t = 0$.

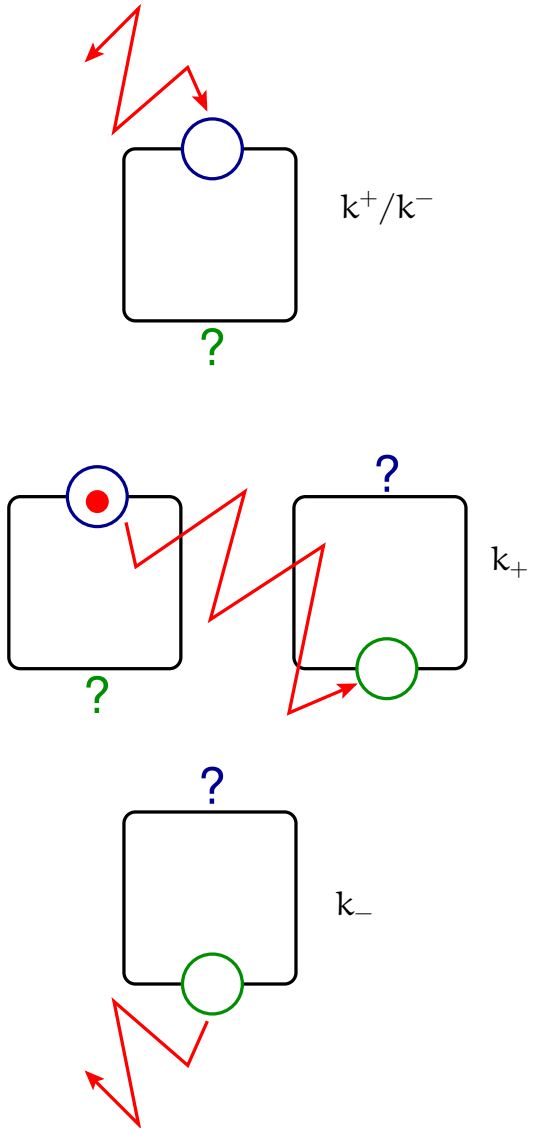
Although the reduction is correct in the ODE semantics.

Degree of correlation (in the unreduced model)





Distant control



$$A \underset{k^-}{\overset{k^+}{\rightleftharpoons}} A^\star$$

$$A_\star \underset{k^-}{\overset{k^+}{\rightleftharpoons}} A^\star_\star$$

$$A + A^\star \xrightarrow{k_+} A_\star + A^\star$$

$$A^\star + A^\star \xrightarrow{k_+} A^\star_\star + A^\star$$

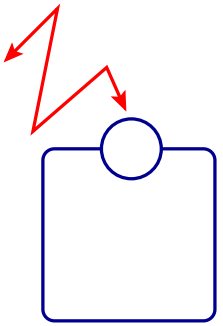
$$A + A^\star_\star \xrightarrow{k_+} A_\star + A^\star_\star$$

$$A^\star + A^\star_\star \xrightarrow{k_+} A^\star_\star + A^\star_\star$$

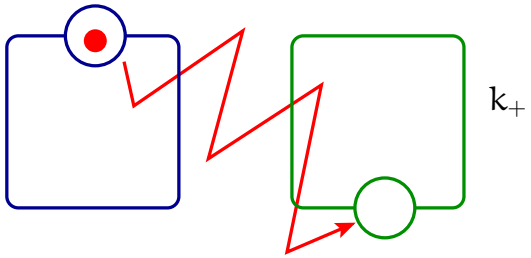
$$A^\star_\star \xrightarrow{k_-} A^\star$$

$$A_\star \xrightarrow{k_-} A$$

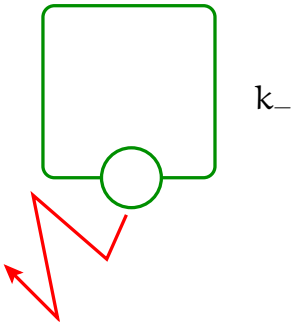
Reduced model



$$A \underset{k^-}{\overset{k^+}{\rightleftharpoons}} A^\star$$

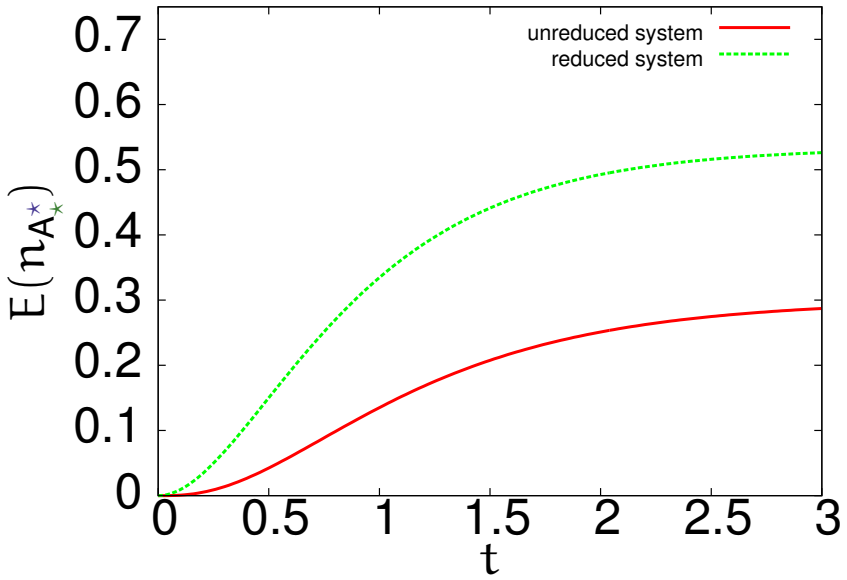


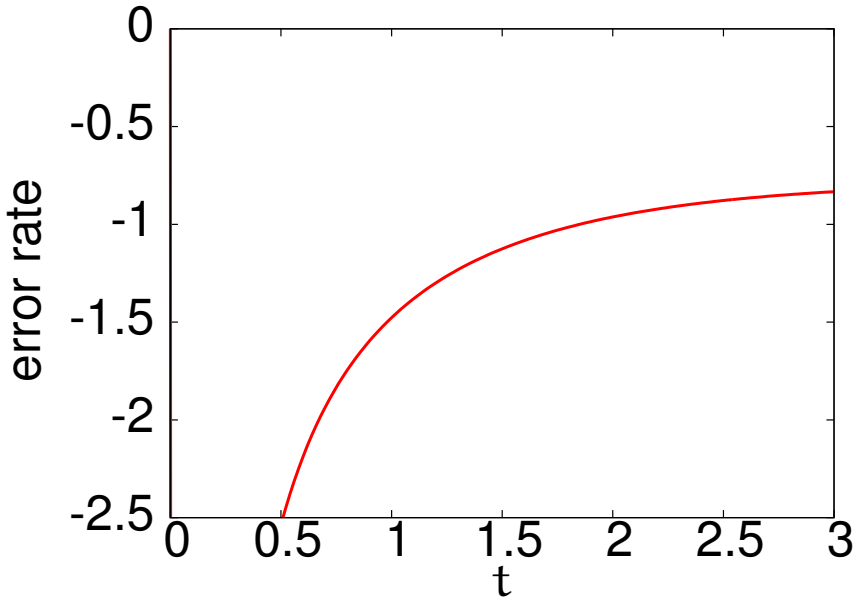
$$A + A^\star \xrightarrow{k_+} A_\star + A^\star$$



$$A_\star \xrightarrow{k_-} A$$

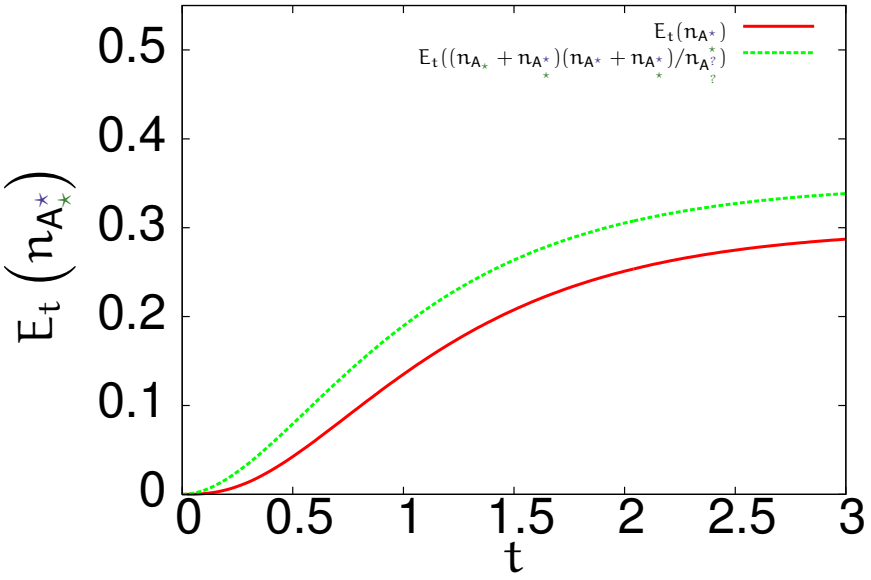
Comparison between the two models

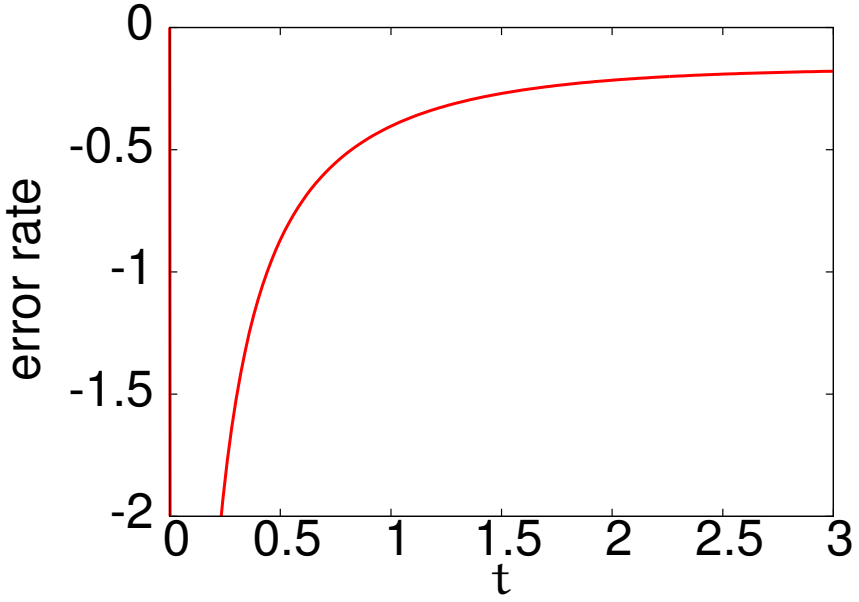




with $k^+ = k^- = k_+ = k_- = 1$,
and two instances of A at time $t = 0$.

Degree of correlation (in the unreduced model)

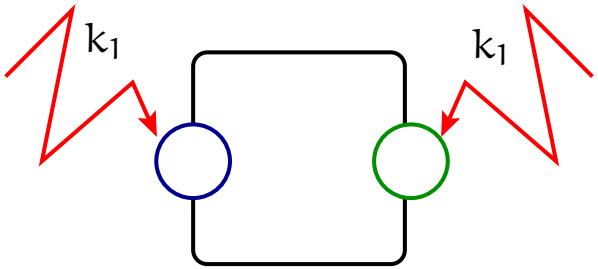




Overview

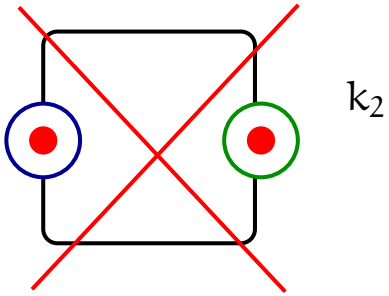
1. Introduction
2. Examples of information flow
3. Symmetric sites
4. Stochastic semantics
5. Lumpability
6. Bisimulations
7. Hierarchy of semantics
8. Conclusion

A model with symmetries



$$P \xrightarrow{k_1} {}^\star P \qquad P^\star \xrightarrow{k_1} {}^\star P^\star$$

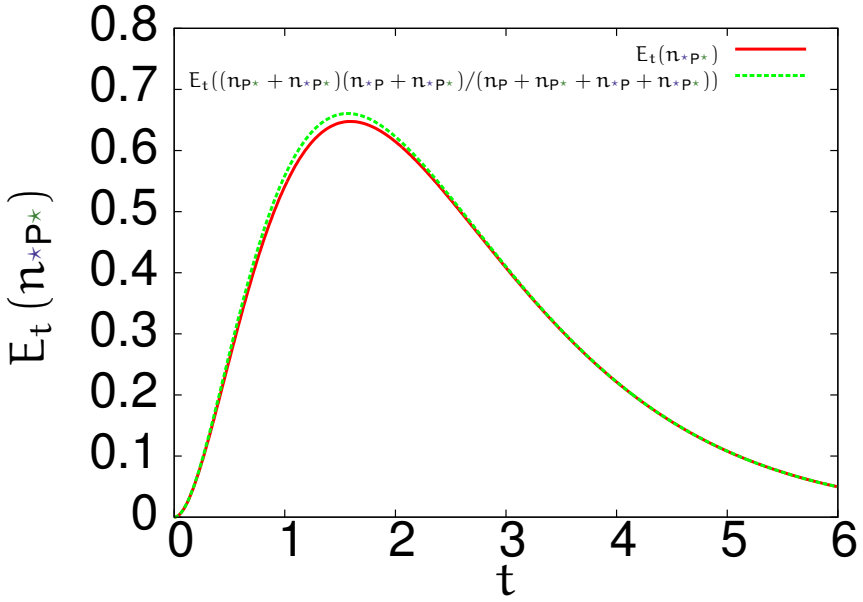
$$P \xrightarrow{k_1} P^\star \qquad {}^\star P \xrightarrow{k_1} {}^\star P^\star$$

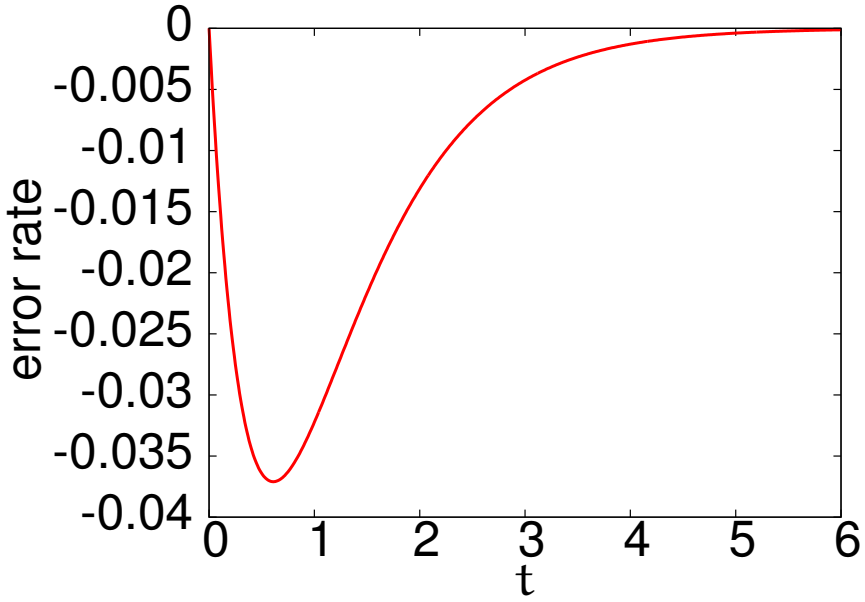


$${}^\star P^\star \xrightarrow{k_2} \emptyset$$

Degree of correlation (in the unreduced model)

$$E_t\left(n_{\star P^\star}\right) = E_t\left(\frac{(n_{\star P} + n_{\star P^\star})(n_{P^\star} + n_{\star P^\star})}{n_P + n_{P^\star} + n_{\star P} + n_{\star P^\star}}\right).$$



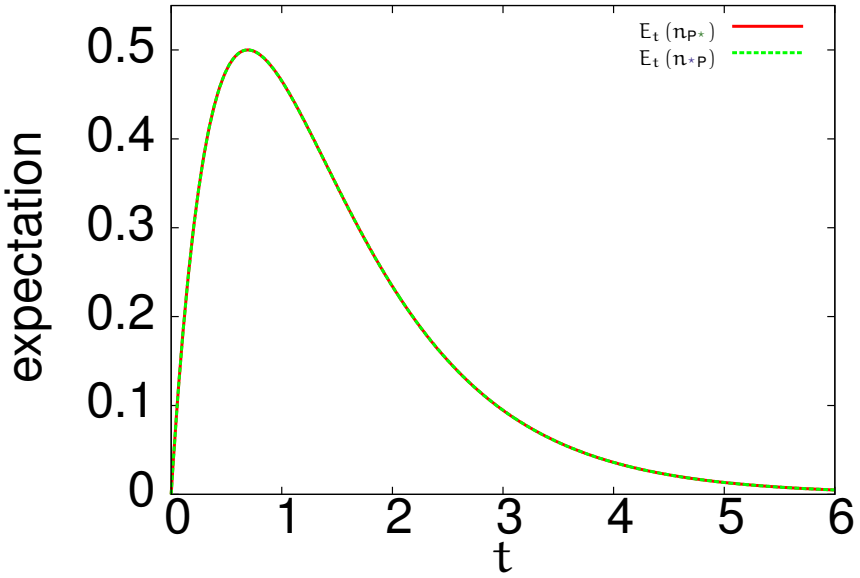


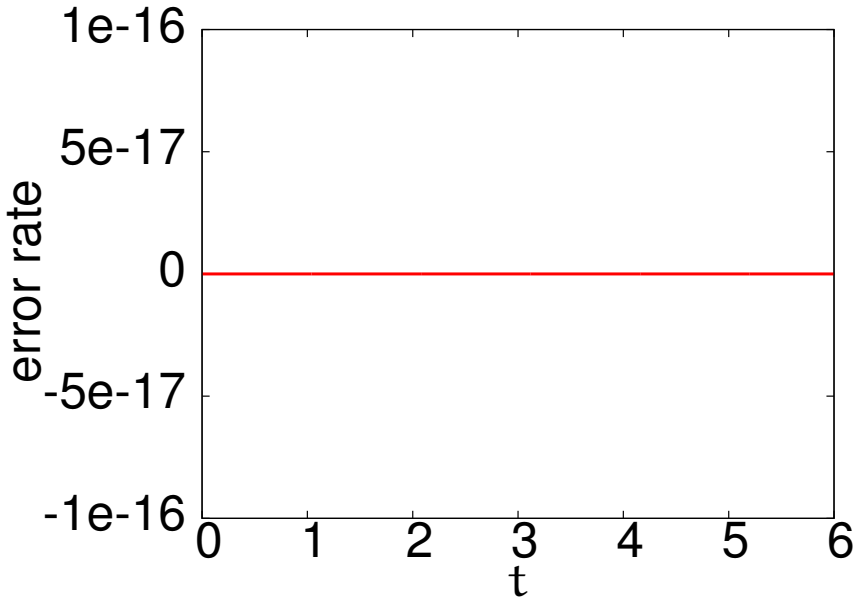
with $k_1 = k_2 = 1$
and two instances of P at time $t = 0$.

Equivalent chemical species

We check numerically that:

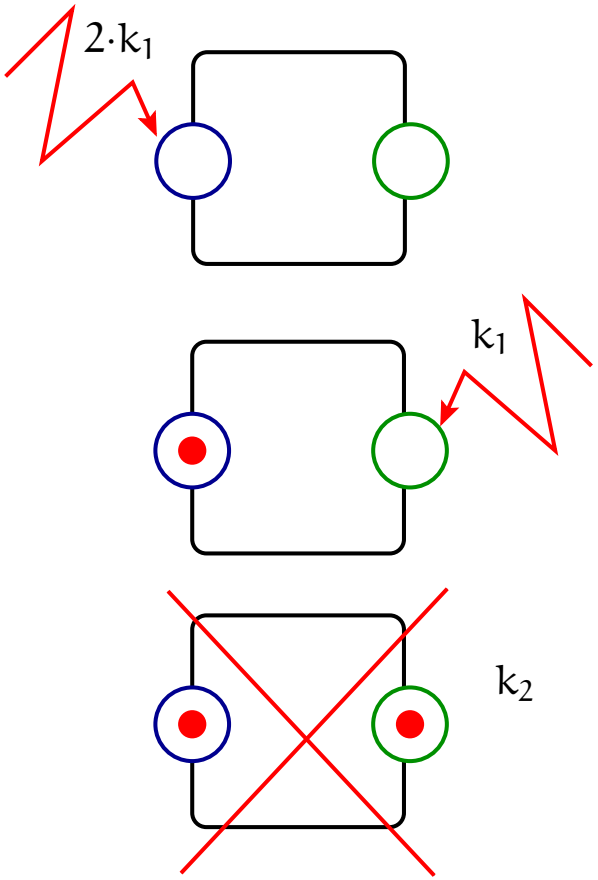
$$\mathrm{E}_t\left(n_{P^\star}\right) = \mathrm{E}_t\left(n_{\star P}\right).$$





with $k_1 = k_2 = 1$
and two instances of P at time $t = 0$.

Reduced model



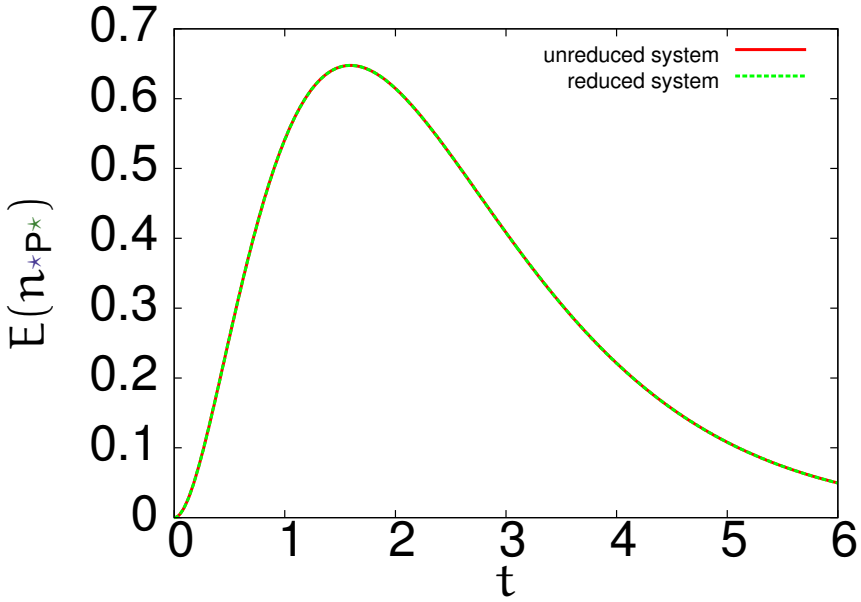
$$P \xrightarrow{2 \cdot k_1} {}^{\star}P$$

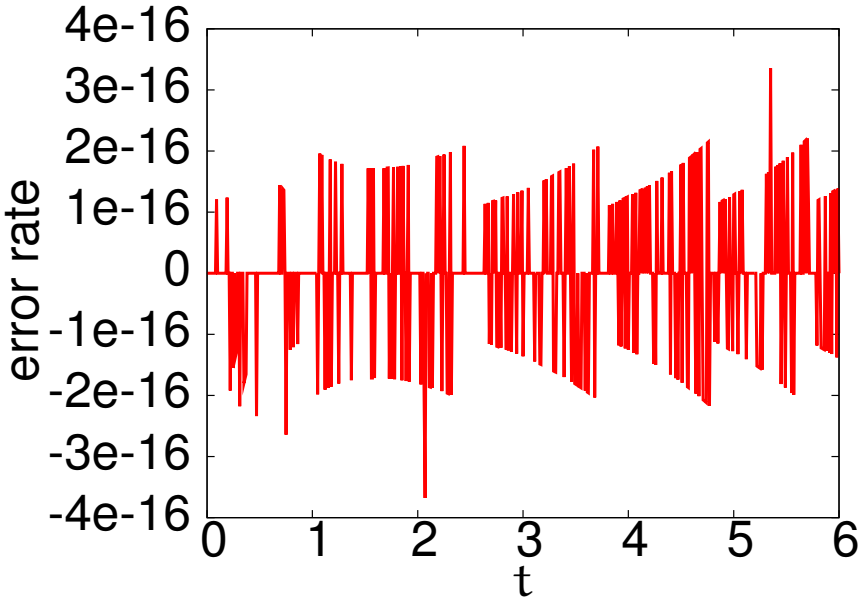
$${}^{\star}P \xrightarrow{k_1} {}^{\star}P^{\star}$$

$${}^{\star}P^{\star} \xrightarrow{k_2} \emptyset$$

Exponential reduction!!!

Comparison between the two models





with $k_1 = k_2 = 1$
and two instances of P at time $t = 0$.

Overview

1. Introduction
2. Examples of information flow
3. Symmetric sites
4. Stochastic semantics
5. Lumpability
6. Bisimulations
7. Hierarchy of semantics
8. Conclusion

Weighted Labelled Transition Systems

A weighted-labelled transition system $\mathcal{W}$ is given by:

- $\mathcal{Q}$, a countable set of states;
- $\mathcal{L}$, a set of labels;
- $w : \mathcal{Q} \times \mathcal{L} \times \mathcal{Q} \to \mathbb{R}_0^+$, a weight function;
- $\pi_0 : \mathcal{Q} \to [0, 1]$, an initial probability distribution.

We also assume that:

- the system is finitely branching, i.e.:
  - the set $\{q \in \mathcal{Q} \mid \pi_0(q) > 0\}$ is finite
  - and, for any $q \in \mathcal{Q}$, the set $\{l, q' \in \mathcal{L} \times \mathcal{Q} \mid w(q, l, q') > 0\}$ is finite.
- the system is deterministic:
  if $w(q, \lambda, q_1) > 0$ and $w(q, \lambda, q_2) > 0$, then: $q_1 = q_2$.

Trace distribution

A cylinder set of traces is defined as:

$$\tau \stackrel{\Delta}{=} q_0 \stackrel{\lambda_1, I_1}{\rightarrow} q_1 \ldots q_{k-1} \stackrel{\lambda_k, I_k}{\rightarrow} q_k$$

where:

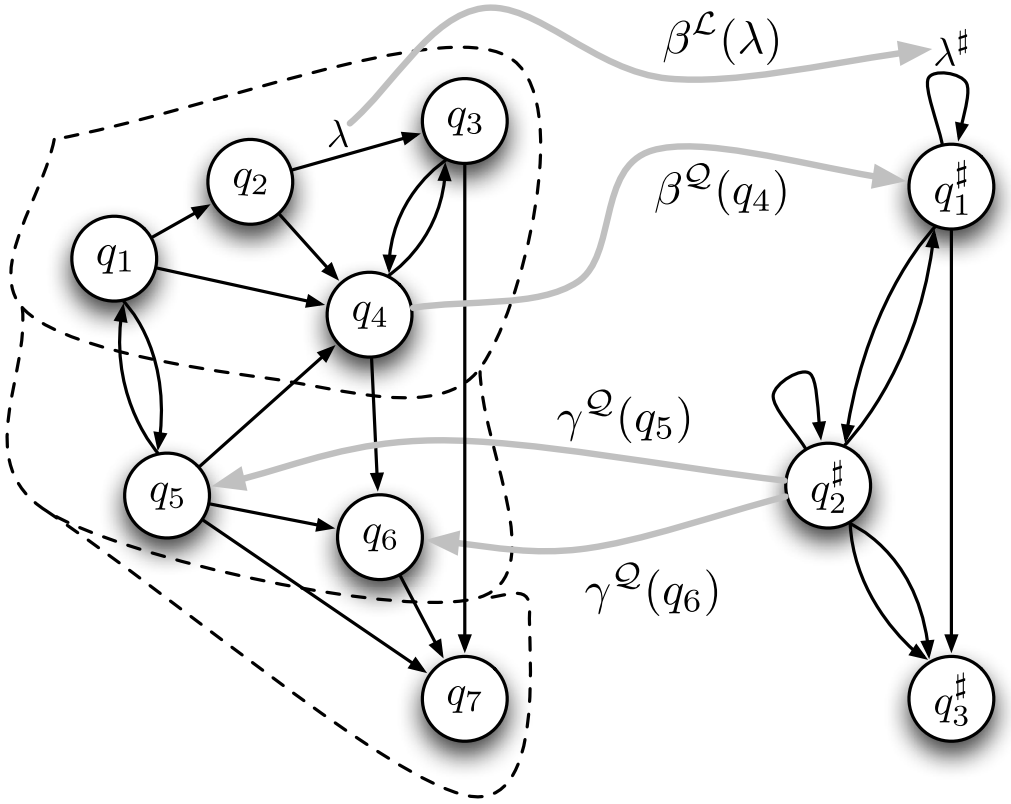
- $(q_i)_{0 \leq i \leq k} \in \mathcal{Q}^{k+1}$ and $(\lambda_i)_{1 \leq i \leq k} \in \mathcal{L}^k$,
- $(I_i)_{1 \leq i \leq k}$ is a family of open intervals in $\mathbb{R}_0^+$.

The probability of a cylinder set of traces is defined as follows:

$$\mathcal{P}r(\tau) \stackrel{\Delta}{=} \pi_0(q_0) \prod_{i=1}^{k} \frac{w(q_{i-1}, l_i, q_i)}{a(q_{i-1})} \left(e^{-a(q_{i-1}) \cdot \mathsf{inf}(I_i)} - e^{-a(q_{i-1}) \cdot \mathsf{sup}(I_i)} \right),$$

where $a(q) \stackrel{\Delta}{=} \sum_{\lambda, q'} w(q, \lambda, q')$.

Abstraction between WLTS



Soundness

Given:

- two WLTS $\mathcal{S} \triangleq (\mathcal{Q}, \mathcal{L}, \rightarrow, w, \mathcal{I}, \pi_0)$ and $\mathcal{S}^\sharp \triangleq (\mathcal{Q}^\sharp, \mathcal{L}^\sharp, \rightsquigarrow, w^\sharp, \mathcal{I}^\sharp, \pi_0^\sharp)$,
- two abstraction functions $\beta^{\mathcal{Q}} : \mathcal{Q} \rightarrow \mathcal{Q}^\sharp$ and $\beta^{\mathcal{L}} : \mathcal{L} \rightarrow \mathcal{L}^\sharp$,

$\mathcal{S}^\sharp$ is a sound abstraction of $\mathcal{S}$, if and only if, for any cylinder set τ of traces of $\mathcal{S}$, we have:

$$\mathcal{P}r(\beta^{\mathbb{T}}(\tau)) = \sum\nolimits_{\tau'} (\mathcal{P}r(\tau') \mid \beta^{\mathbb{T}}(\tau) = \beta^{\mathbb{T}}(\tau')),$$

where,

$$\beta^{\mathbb{T}}(q_0 \overset{\lambda_1, I_1}{\rightarrow} q_1 \dots q_{k-1} \overset{\lambda_k, I_k}{\rightarrow} q_k) \triangleq \beta^{\mathcal{Q}}(q_0) \overset{\beta^{\mathcal{L}}(\lambda_1), I_1}{\rightarrow} \beta^{\mathcal{Q}}(q_1) \dots \beta^{\mathcal{Q}}(q_{k-1}) \overset{\beta^{\mathcal{L}}(\lambda_k), I_k}{\rightarrow} \beta^{\mathcal{Q}}(q_k).$$

Completeness

Given:

- two WLTS $\mathcal{S} \stackrel{\Delta}{=} (\mathcal{Q}, \mathcal{L}, \rightarrow, w, \mathcal{I}, \pi_0)$ and $\mathcal{S}^\sharp \stackrel{\Delta}{=} (\mathcal{Q}^\sharp, \mathcal{L}^\sharp, \rightsquigarrow, w^\sharp, \mathcal{I}^\sharp, \pi_0^\sharp)$,
- two abstraction functions $\beta^{\mathcal{Q}} : \mathcal{Q} \rightarrow \mathcal{Q}^\sharp$ and $\beta^{\mathcal{L}} : \mathcal{L} \rightarrow \mathcal{L}^\sharp$,
- a concretization function $\gamma^{\mathcal{Q}} : \mathcal{Q} \rightarrow \mathbb{R}^+$,

$\mathcal{S}^\sharp$ is a sound and complete abstraction of $\mathcal{S}$, if and only if,

1. it is a sound abstraction;
2. for any cylinder set $\tau^\sharp$ of abstract traces of $\mathcal{S}^\sharp$ which ends in the abstract state $q_k^\sharp$, we have:

$$\gamma^{\mathcal{Q}}(s) = \mathcal{P}r(q_k = s \mid \tau \text{ such that } \beta^{\mathbb{T}}(\tau) \in \tau^\sharp) \times \sum \{\gamma^{\mathcal{Q}}(s') \mid \beta^{\mathcal{Q}}(s') = q_k^\sharp\}.$$

Overview

1. Introduction
2. Examples of information flow
3. Symmetric sites
4. Stochastic semantics
5. Lumpability
6. Bisimulations
7. Hierarchy of semantics
8. Conclusion

Markovian Property

We consider a stochastic process:

- $\mathbb{T} = \mathbb{R}_0^+$: time range;
- $\mathcal{Q}$: a countable set of states;
- $(\mathcal{X}_t)_{t \in \mathbb{T}}$: a family of random variables over $\mathcal{Q}$;

We say that $(\mathcal{X}_t)$ satisfies the Markovian property,
if, for any family $(s_t)_{t \in \mathbb{T}}$ of states indexed over $\mathbb{T}$, and any time $t_1 < t_2$,
we have:

$$\mathcal{P}r(X_{t_2} = s_{t_2} \mid X_{t_1} = s_{t_1}) = \mathcal{P}r(X_{t_2} = s_{t_2} \mid X_t = s_t, \forall t < t_1).$$

Lumpability property

Given:

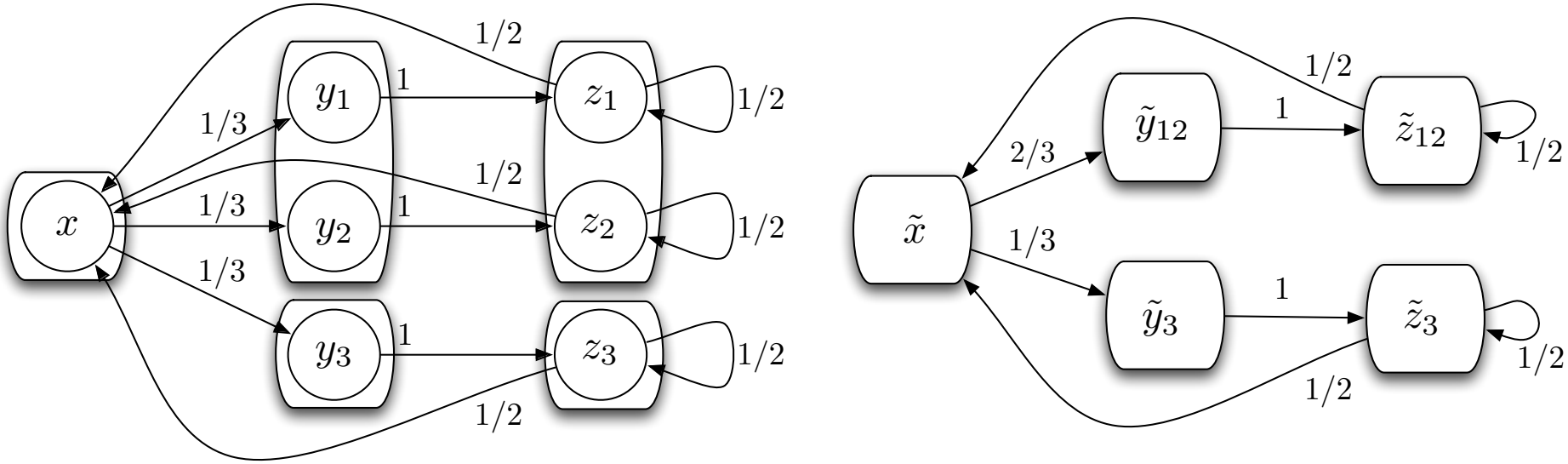
- a stochastic process $(\mathcal{X}_t)$ which satisfies the Markovian property,
- an initial distribution $\pi_0 : \mathcal{Q} \to [0, 1]$,
- an equivalence relation $\sim$ over $\mathcal{Q}$,

we define the lumped process $(\mathcal{Y}_t)$ on the state space $\mathcal{Q}_{/\sim}$ as:

$$\mathcal{P}r(\mathcal{Y}_t = [x_t]_{/\sim} \mid \mathcal{Y}_0 = [s_0]_{/\sim}) \stackrel{\Delta}{=} \mathcal{P}r(\mathcal{X}_t \in [s_t]_{/\sim} \mid \mathcal{X}_0 \in [s_0]_{/\sim}).$$

We say that $(\mathcal{X})_t$ is $\sim$-lumpable with respect to π_0 if and only if, the stochastic process $(\mathcal{Y}_t)$ satisfies the Markovian property as well.

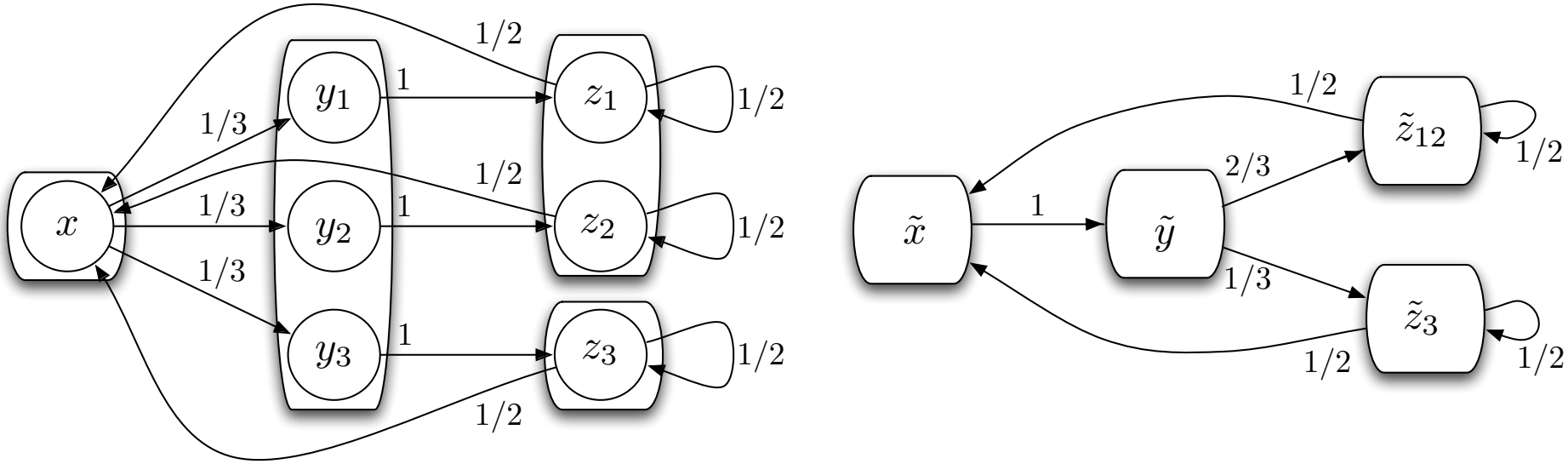
Strong lumpability



A stochastic process is $\sim$-strongly lumpable, if:

it is $\sim$-lumpable with respect to any initial distribution.

Weak lumpability



A stochastic process $(\mathcal{X}_t)$ is $\sim$-weakly lumpable, if:
there exists an initial distribution with respect to which $(\mathcal{X}_t)$ is $\sim$-lumpable.

Overview

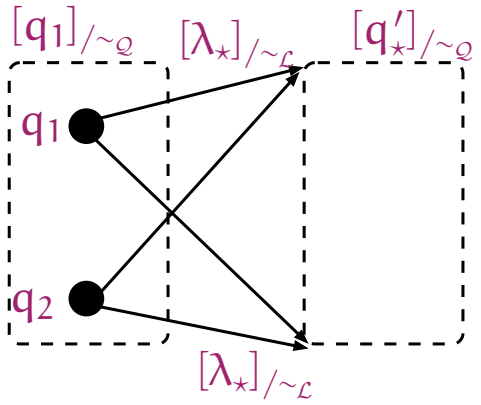
1. Introduction
2. Examples of information flow
3. Symmetric sites
4. Stochastic semantics
5. Lumpability
6. Bisimulations
7. Hierarchy of semantics
8. Conclusion

Forward bisimulation

Let $\sim_{\mathcal{Q}}$ be an equivalence relation over $\mathcal{Q}$ and $\sim_{\mathcal{L}}$ be an equivalence relation over $\mathcal{L}$.

We say that $(\sim_{\mathcal{Q}}, \sim_{\mathcal{L}})$ is a forward bisimulation, if and only if, for any $q_1, q_2 \in \mathcal{Q}$ such that $q_1 \sim_{\mathcal{Q}} q_2$:

- $a(q_1) = a(q_2)$;
- and for any $\lambda_\star \in \mathcal{L}$, $q'_\star \in \mathcal{Q}$, $\mathsf{fwd}(q_1, [\lambda_\star]_{/\sim_{\mathcal{L}}}, [q'_\star]_{/\sim_{\mathcal{Q}}}) = \mathsf{fwd}(q_2, [\lambda_\star]_{/\sim_{\mathcal{L}}}, [q'_\star]_{/\sim_{\mathcal{Q}}})$

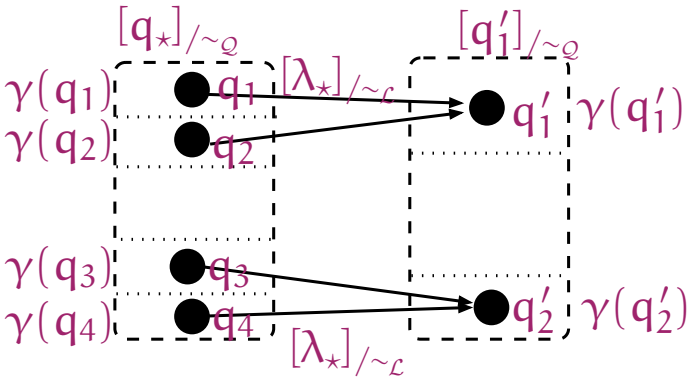


where: $\mathsf{fwd}(q, [\lambda_\star]_{/\sim_{\mathcal{L}}}, [q'_\star]_{/\sim_{\mathcal{Q}}}) = \sum_{\lambda', q'} (w(q, \lambda', q') \mid \lambda' \sim_{\mathcal{L}} \lambda_\star,\ q' \sim_{\mathcal{Q}} q'_\star)$.

Backward bisimulation

Let $\sim_{\mathcal{Q}}$ be an equivalence relation over $\mathcal{Q}$ and $\sim_{\mathcal{L}}$ be an equivalence relation over $\mathcal{L}$.

We say that $(\sim_{\mathcal{Q}}, \sim_{\mathcal{L}})$ is a backward bisimulation, if and only if, there exists $\gamma : \mathcal{Q} \to \mathbb{R}^+$, such that: for any $q_1', q_2' \in \mathcal{Q}$ which satisfies $q_1' \sim_{\mathcal{Q}} q_2'$:



- $a(q_1') = a(q_2')$;
- and for any $\lambda_\star \in \mathcal{L}$, $q_\star \in \mathcal{Q}$,
  $\mathrm{bwd}([q_\star]_{/\sim_{\mathcal{Q}}}, [\lambda_\star]_{\sim_{/\mathcal{L}}}, q_1') = \mathrm{bwd}([q_\star]_{/\sim_{\mathcal{Q}}}, [\lambda_\star]_{\sim_{/\mathcal{L}}}, q_2')$

  where: $\mathrm{bwd}([q_\star]_{/\sim_{\mathcal{Q}}}, [\lambda_\star]_{\sim_{/\mathcal{L}}}, q') = \sum_{q,\lambda'} \left(\frac{\gamma(q)}{\gamma(q')} w(q, \lambda', q') \,|\, q \sim_{\mathcal{Q}} q_\star, \lambda' \sim_{\mathcal{L}} \lambda_\star \right)$.

Logical implications

- if $(\sim_{\mathcal{Q}}, \sim_{\mathcal{L}})$ is a forward bisimulation, then the process is $\sim_{\mathcal{Q}}$-strongly lumpable,

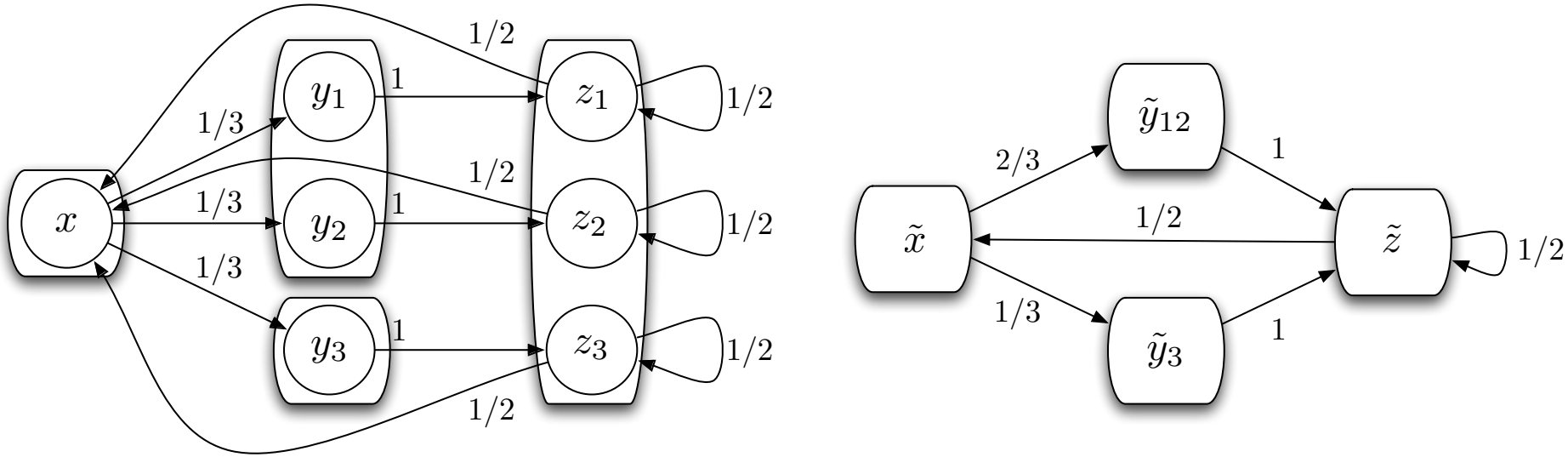
  moreover, it induces a sound abstraction;
- if $(\sim_{\mathcal{Q}}, \sim_{\mathcal{L}})$ is a backward bisimulation, then the process is $\sim_{\mathcal{Q}}$-weakly lumpable, for the initial distributions which satisfy:

$$q \sim_{\mathcal{Q}} q' \Rightarrow [\pi_0(q) \cdot \gamma(q') = \pi_0(q') \cdot \gamma(q)];$$

  it induces a sound and complete abstraction for these initial distributions.;
- there exist forward bisimulations which are not backward bisimulations;
- there exist backward bisimulations which are not forward bisimulations.

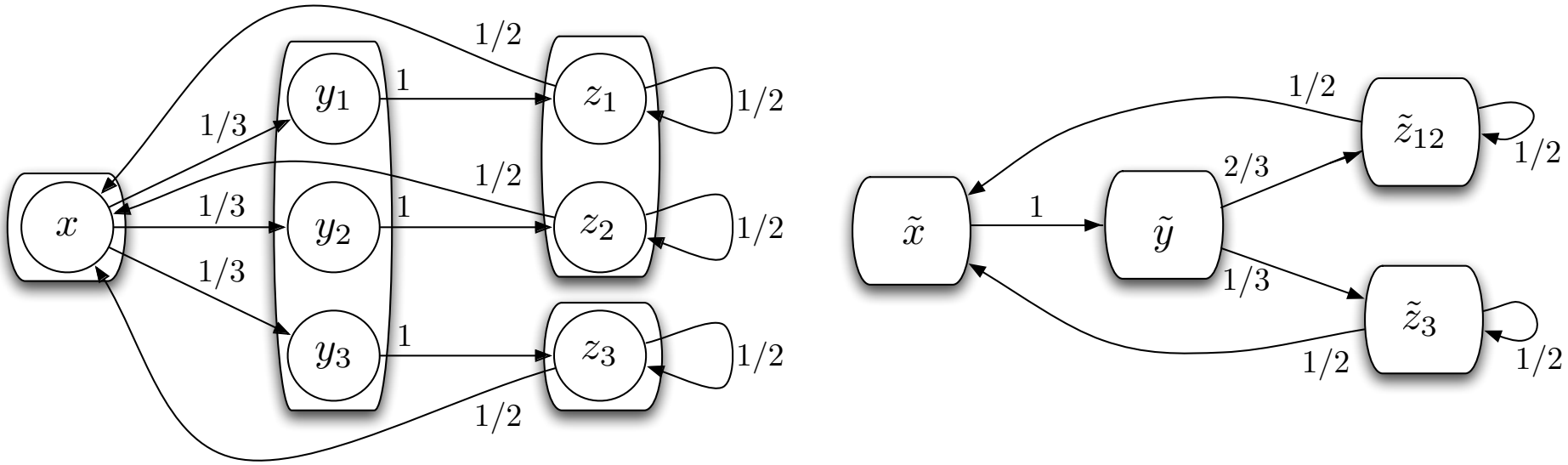
Counter-example I

A forward bisimulation which is not a backward bisimulation:



Counter-example II

A backward bisimulation which is not a forward bisimulation:



Uniform backward bisimulation

Given $q_\star, q' \in \mathcal{Q}$ and $\lambda_\star \in \mathcal{L}$, we denote:

$$\text{pred}([q_\star]_{/\sim_\mathcal{Q}}, [\lambda_\star]_{\sim_{/\mathcal{L}}}, q') \stackrel{\Delta}{=} \{(q, \lambda) \mid w(q, \lambda, q') > 0, q \sim_\mathcal{Q} q_\star,\ \lambda \sim_\mathcal{L} \lambda_\star\}.$$

If,

- $q_1 \sim_\mathcal{Q} q_2 \implies a(q_1) = a(q_2)$;
- for any $q_1', q_2' \in \mathcal{Q}$, such that $q_1' \sim_\mathcal{Q} q_2'$, and any $q_\star \in \mathcal{Q}$ and $\lambda_\star \in \mathcal{L}$, there is a 1-to-1 mapping between $\text{pred}([q_\star]_{/\sim_\mathcal{Q}}, [\lambda_\star]_{\sim_{/\mathcal{L}}}, q_1')$ and $\text{pred}([q_\star]_{/\sim_\mathcal{Q}}, [\lambda_\star]_{\sim_{/\mathcal{L}}}, q_2')$ which is compatible with w,

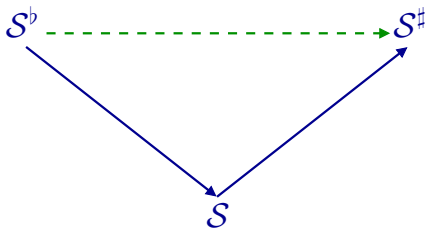
then:

- $(\sim_\mathcal{Q}, \sim_\mathcal{L})$ is a backward bisimulation (with $\gamma(q) = 1,\ \forall q \in \mathcal{Q}$).

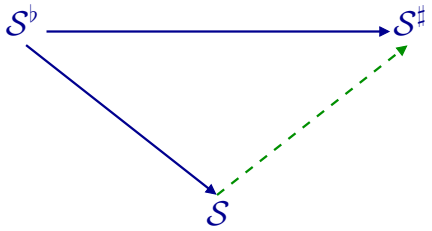
Abstraction algebra

(Sound) abstractions can be:

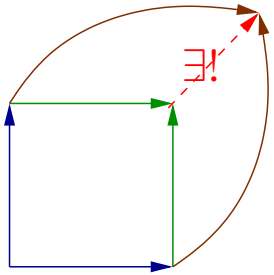
- composed:



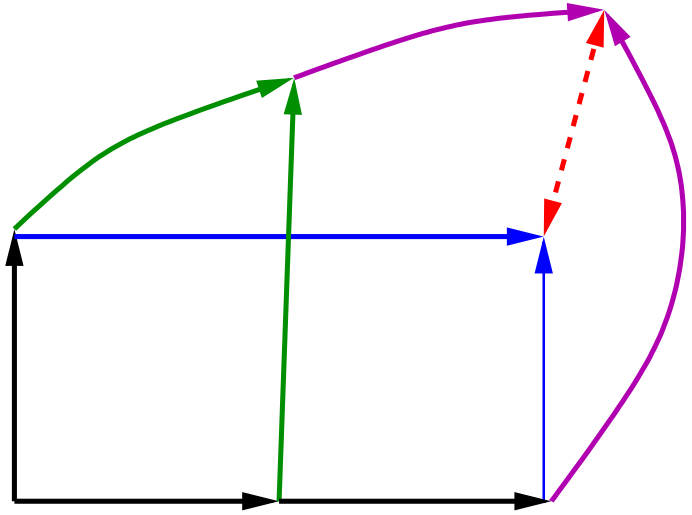
- factored:



- combined with a symmetric product (c.f. lub or pushout):

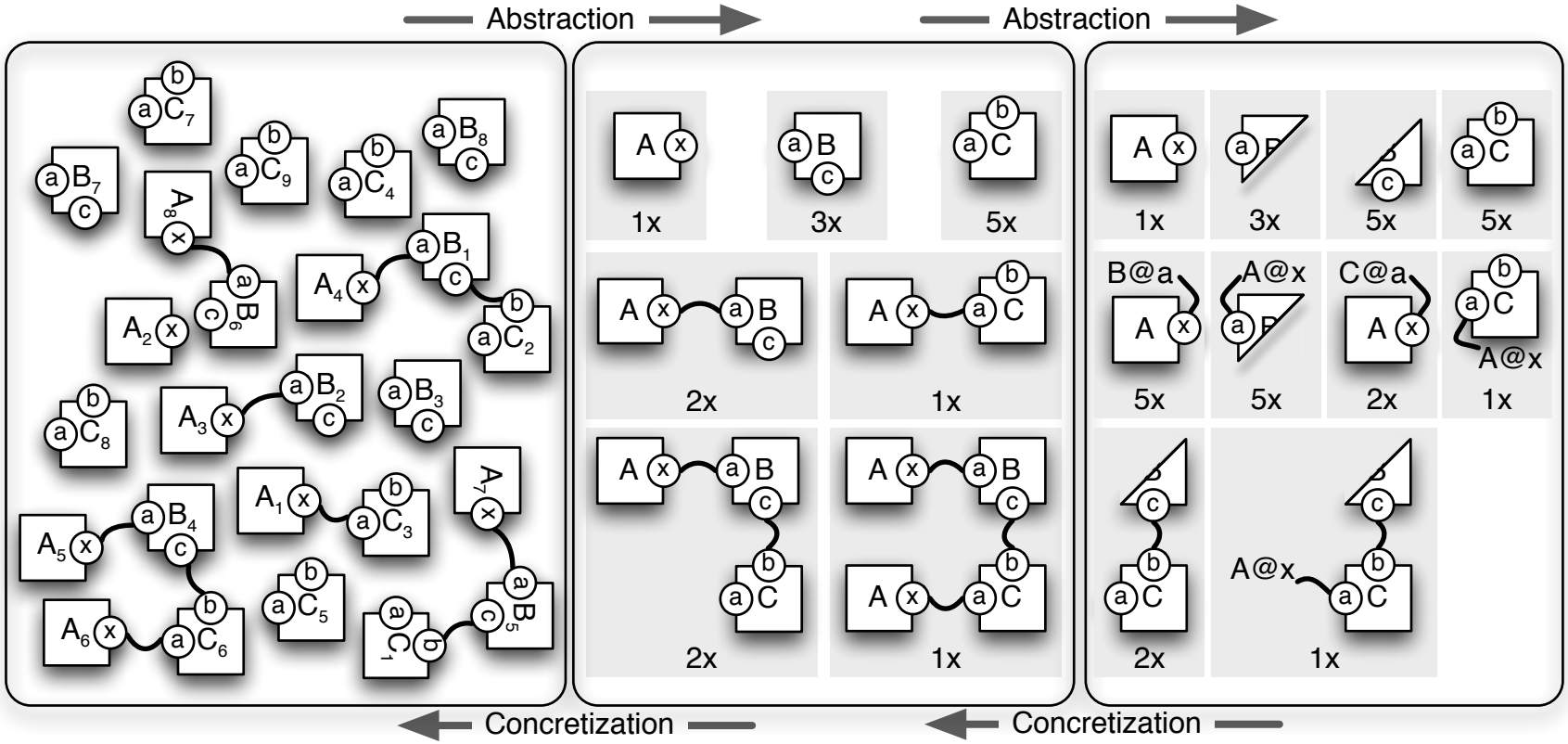


Compatibility between composition and pushout



Overview

1. Introduction
2. Examples of information flow
3. Symmetric sites
4. Stochastic semantics
5. Lumpability
6. Bisimulations
7. Hierarchy of semantics
8. Conclusion


Abstraction
Abstraction
1x
3x
5x
2x
1x
2x
1x
1x
3x
5x
5x
B@a
A@x
C@a
5x
5x
2x
1x
2x
A@x
1x
Concretization
Concretization

From individuals to population

- Individual semantics:

  In the individual semantics, each agent is tagged with a unique identifier which can be tracked along the trace;

- Population semantics:

  In the population semantics, the state of the system is seen up to injective substitution of agent identifier;

  equivalently, the state of the system is a multi-set of chemical species.

Fragments

An annotated contact map is valid with respect to the stochastic semantics, if:

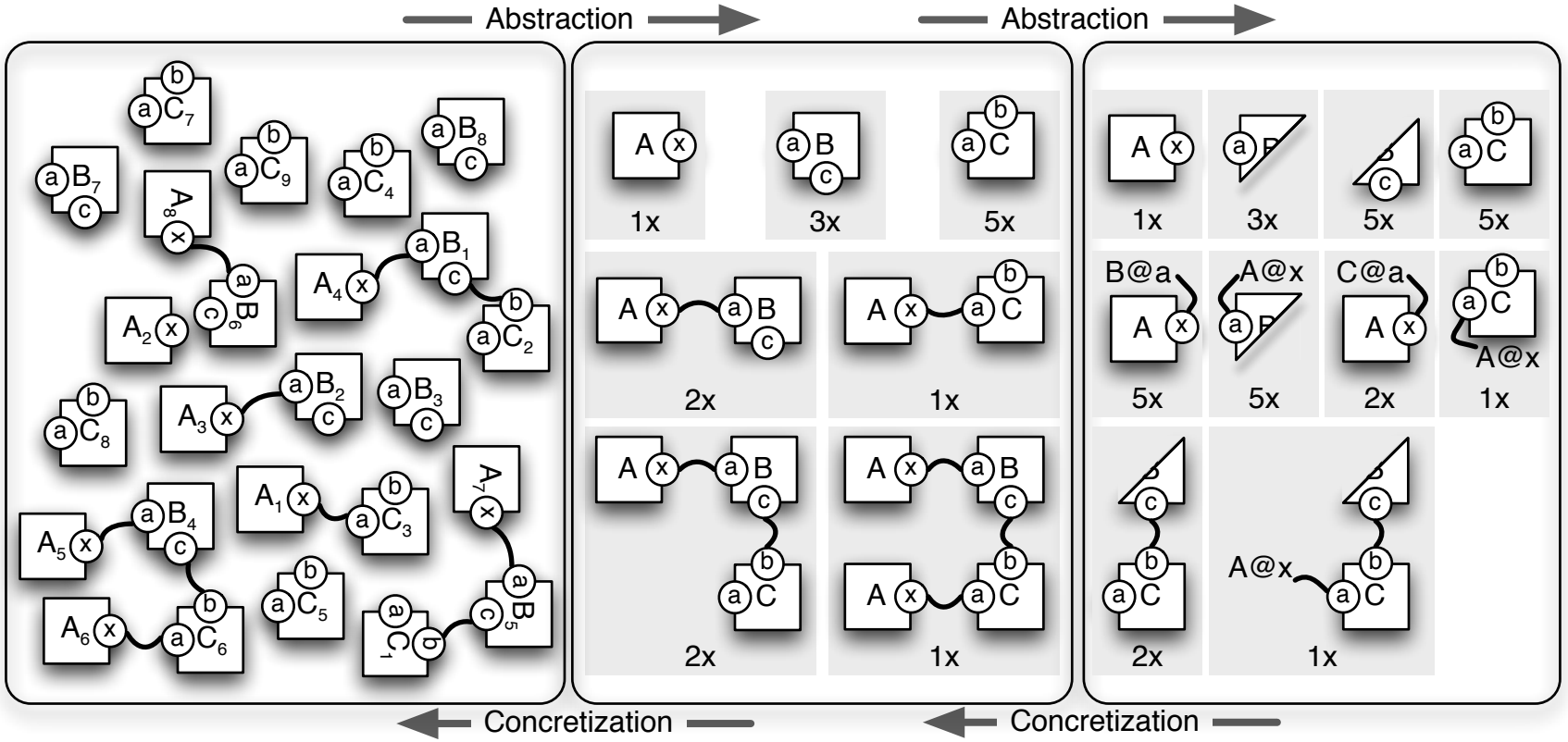
- Whenever the site x and y both occurs in the same or in distinct agent of type A in a rule, then, there should be a bidirectional edge between the site x and the y of A.
- Whenever there is a bond between two sites, each of which either carries an internal state of, is connected to some other sites of its agent, then the bond if oriented in both directions.

From population to fragments

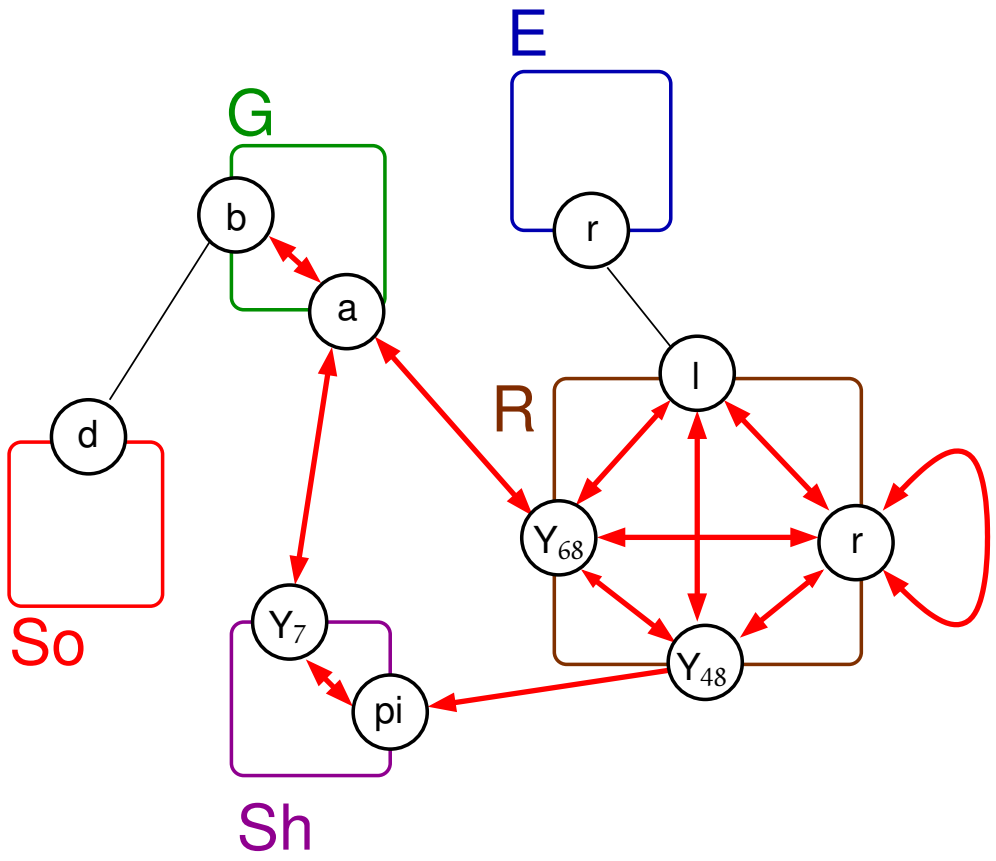
- Population of fragments:

1. In the annotated contact, each agent is fitted with a binary equivalence over its sites. We split the interface of agents into equivalence classes of sites. Then we abstract away which subagents belong to the same agent.

2. Whenever an edge is not oriented in the annotated contact map, we cut each instance of this bond into two half bonds, and abstract away which partners are bond together.


Abstraction
Abstraction
1x
3x
5x
2x
1x
2x
1x
1x
3x
5x
5x
B@a
A@x
C@a
5x
5x
2x
1x
2x
1x
A@x
Concretization
Concretization

Example



Symmetries among sites

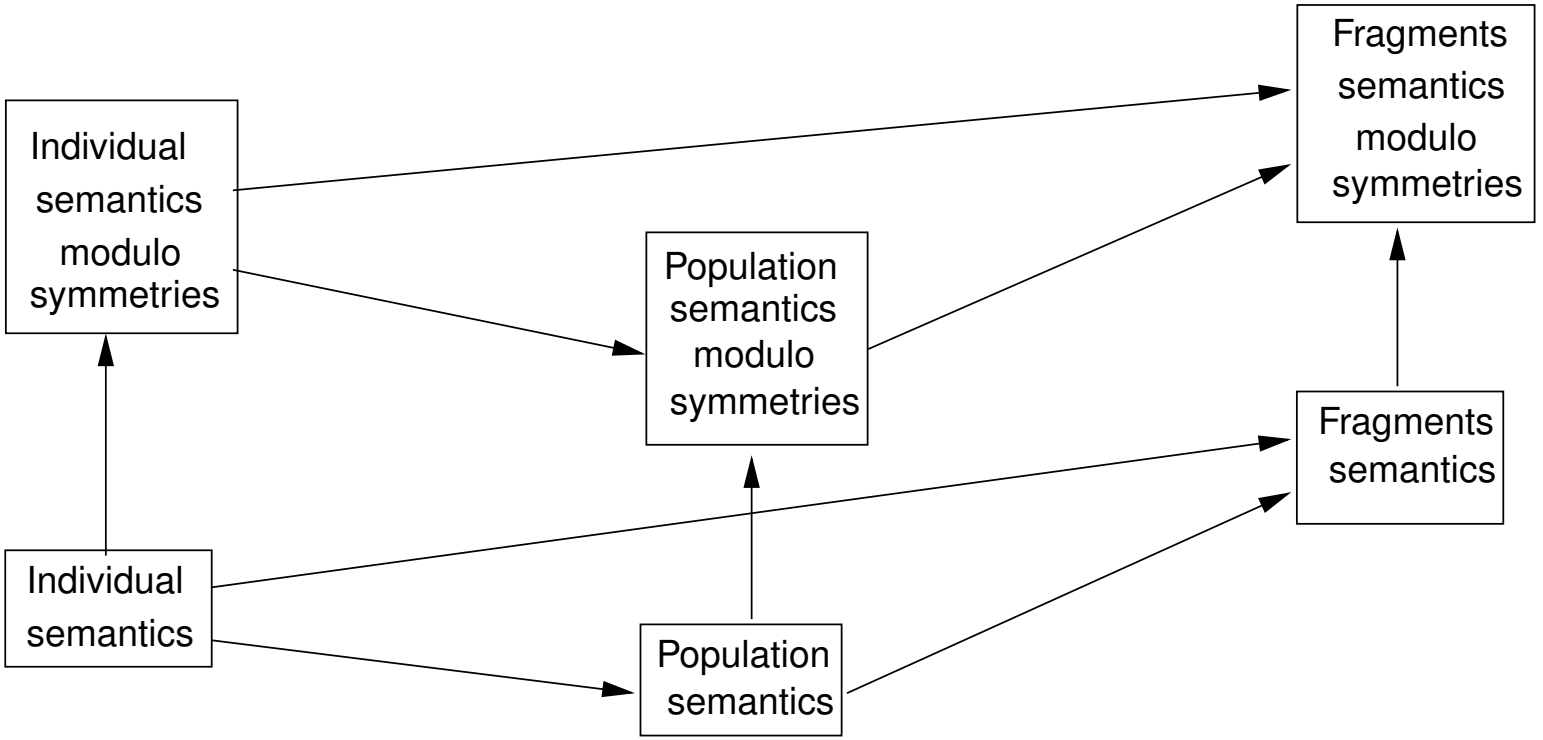
Let $\mathcal{R}$ be a set of rules and $\mathcal{M}_0$ be an initial mixture.

Two sites x_1 and x_2 are symmetric in the agent A in the set of rules $\mathcal{R}$ and the initial mixture $\mathcal{M}_0$

$\overset{\Delta}{\Longleftrightarrow}$

- $\mathcal{R}$ is preserved (modulo $\equiv$) if we replace each rule with all the combinations of rules which can be obtained by replacing (independently) each occurrence of x_1 and x_2 with x_1 or x_2 (and dividing the kinetic rate by the number of combinations, and taking care of gain/loss of automorphisms).
- each agent of type A_i in $\mathcal{M}_0$ has their sites x_1 and x_2 free, with the same internal state.

Hierarchy of semantics



Overview

1. Introduction
2. Examples of information flow
3. Symmetric sites
4. Stochastic semantics
5. Lumpability
6. Bisimulations
7. Hierarchy of semantics
8. Conclusion

Conclusion

- A framework for reducing stochastic rule-based models.
  - We use:
    - the sites the state of which are uncorrelated;
    - the sites having the same capabilities of interactions.
  - Algebraic operators combine these abstractions.

- We use backward bisimulations in order to prove statistical invariants, we use them to reduce the dimension of the continuous-time Markov chains.

Future works

- Investigate the use of hybrid bisimulation.

- Propose approximated simulation algorithms to approximate different scale rate reactions.
  - hybrid systems,
  - tau-leaping,
  - ...

Announcements

- Implementation is freely distributed on http://www.kappalanguage.org.
- Call for candidates (Internships, Post-doc positions...):

  ANR-Chair of Excellence: AbstractCell
  http://www.di.ens.fr/~feret/abstractcell

- Call for papers/participation:

  Second Workshop on Static Analysis and Systems Biology
  (SASB 2011)
  (co-chaired with Andre Levchenko)
  13th Sept 2011, Venice
  http://www.di.ens.fr/sasb2011