MPRI

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

Friday, the 31rd of January, 2014

Joint-work with...

Walter Fontana
Harvard Medical School

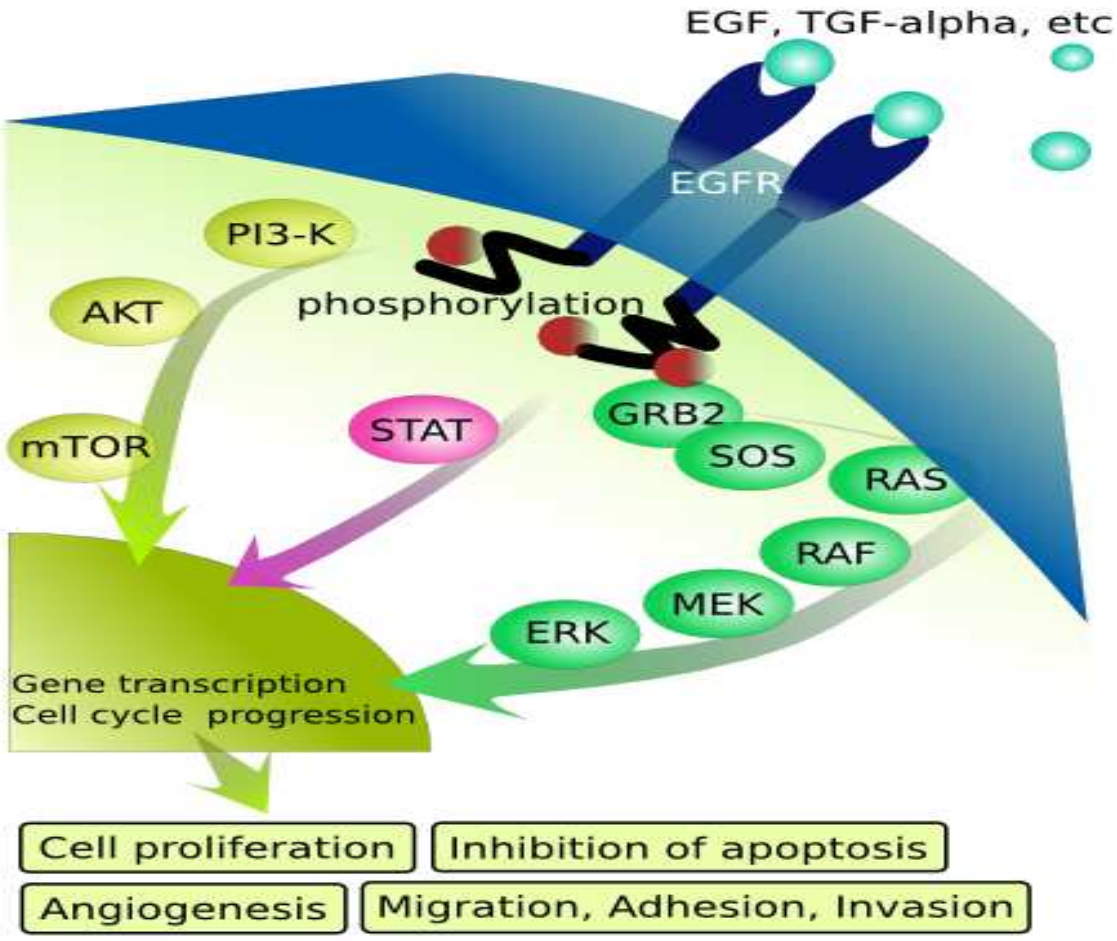
Vincent Danos
Edinburgh

Ferdinanda Camporesi
Bologna / ÉNS

Russ Harmer
Paris VII

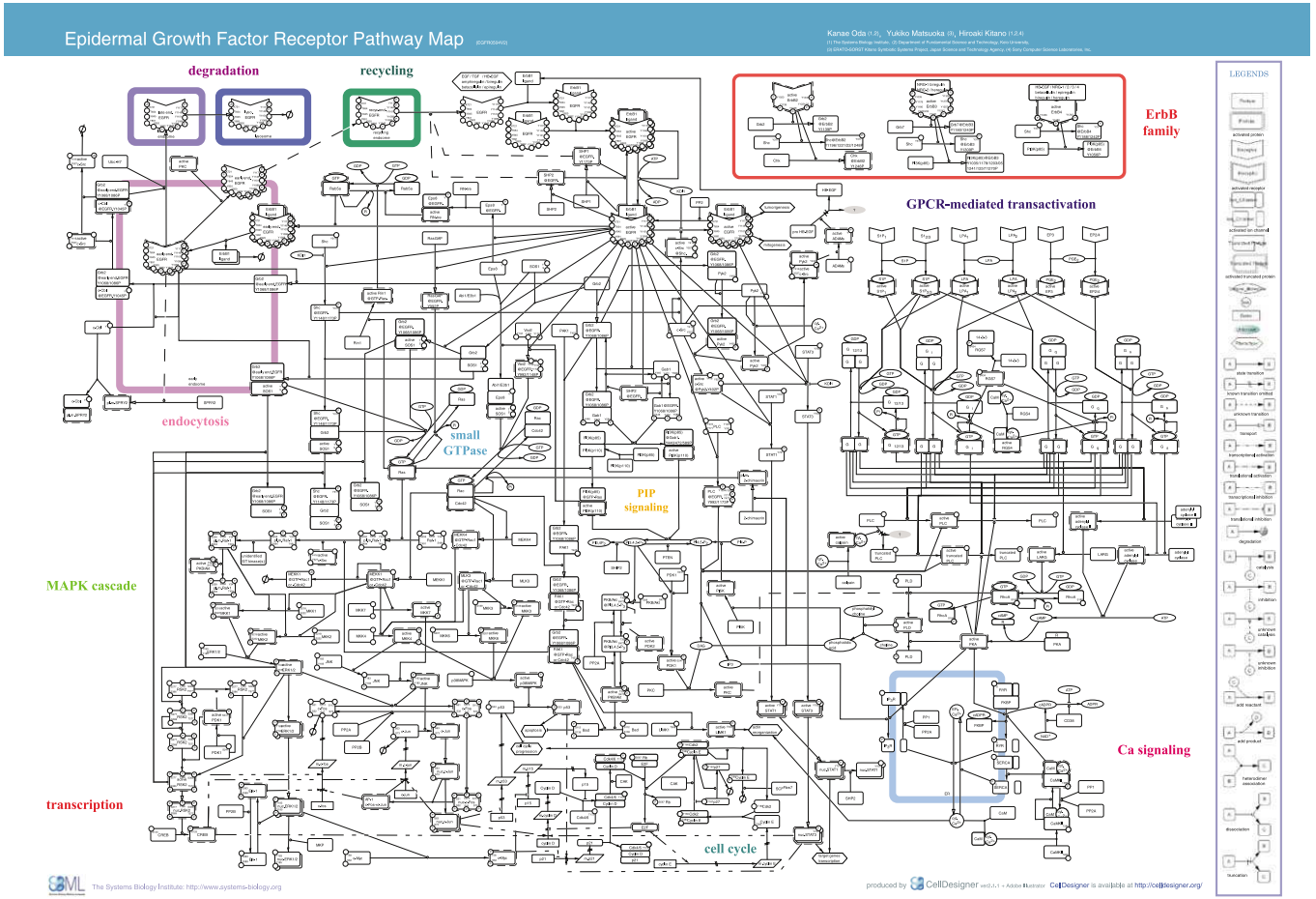
Jean Krivine
Paris VII

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.

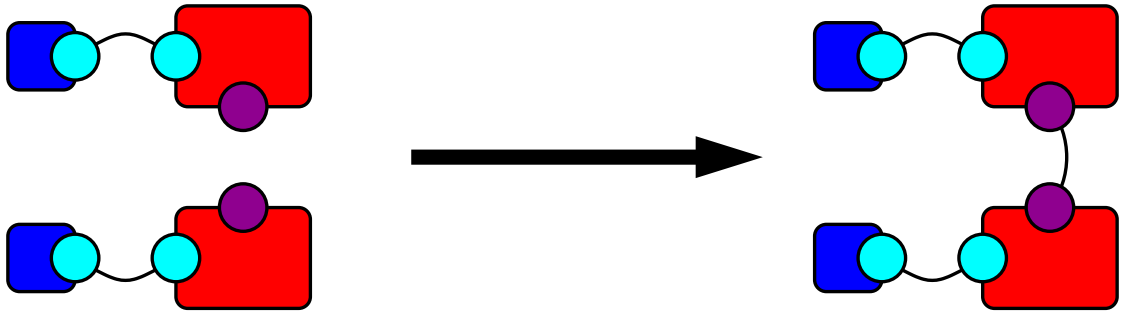
A gap between two worlds

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

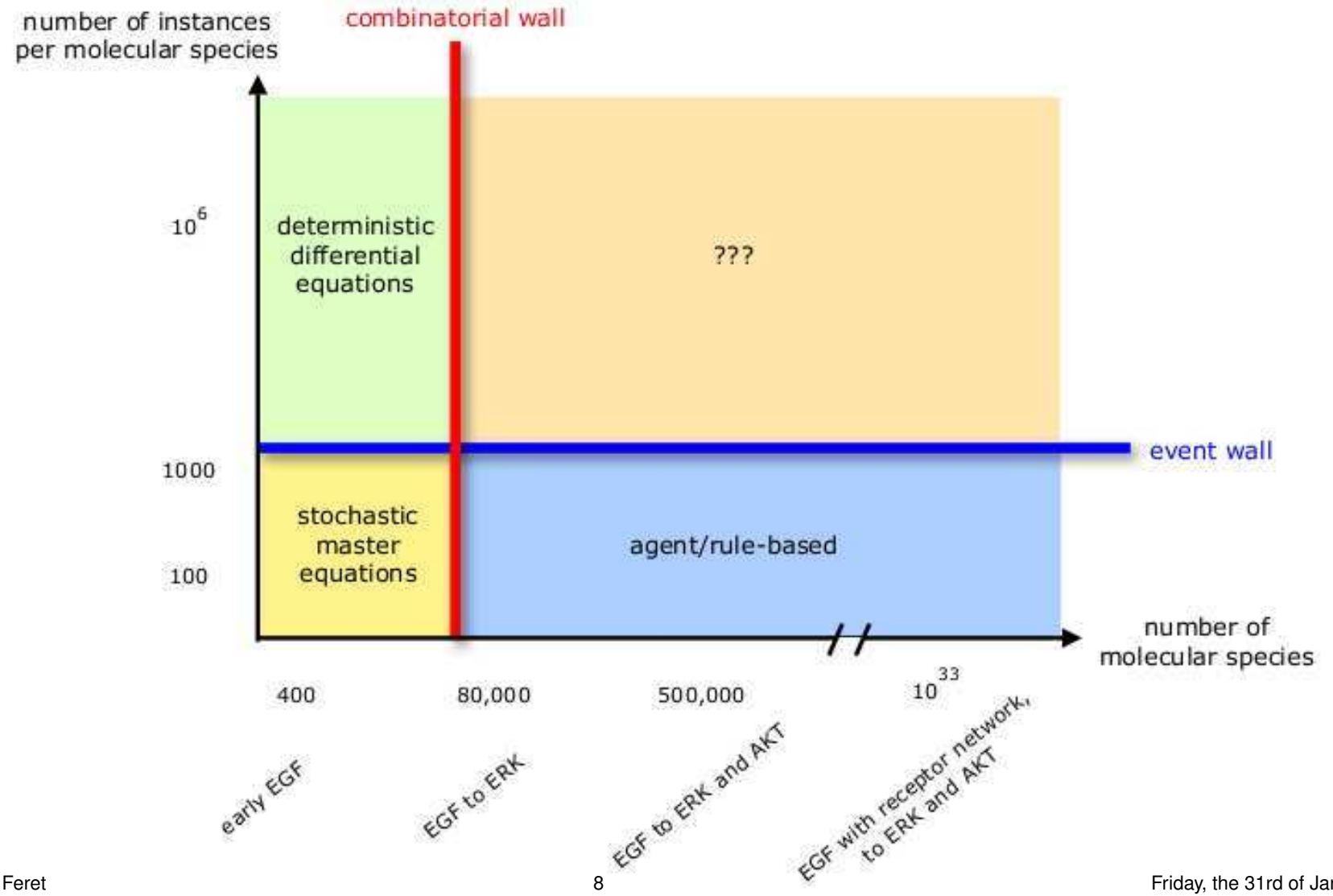


1. The description level matches with both
   - the observation level
   - and the intervention level

   of the biologist.

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

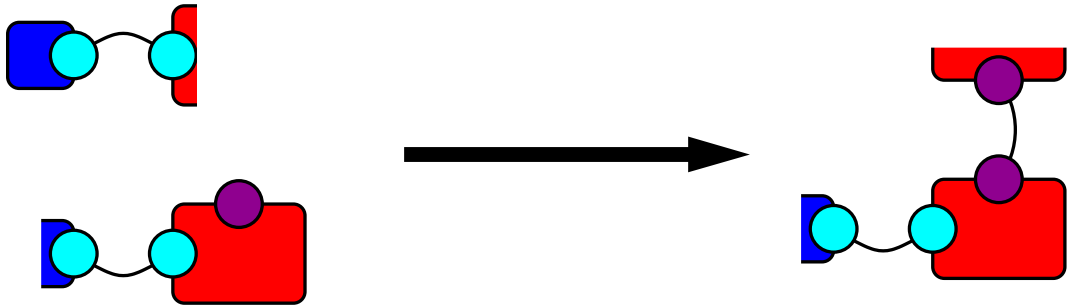
Complexity walls



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]

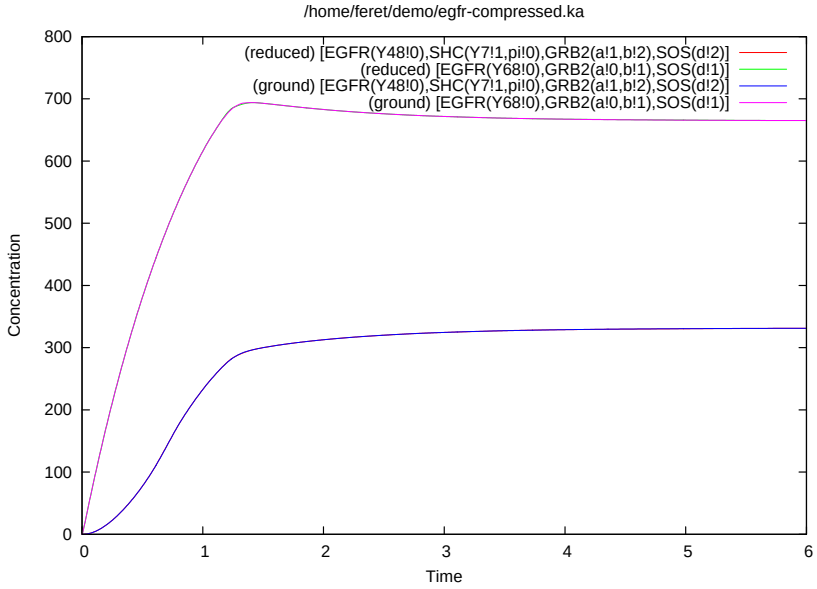
The analysis is complete (no false positif) for a significatif kernel of Kappa [VMCAI'08].

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]

MPRI

Reachability Analysis of Rule-based Models

[ICCMSE'07,VMCAI'08]

Jérôme Feret
Département d'Informatique de l'École Normale Supérieure
INRIA, ÉNS, CNRS
http://www.di.ens.fr/~feret

Friday, the 31rd of January, 2014

In this talk...

We illustrate the following concepts:

- Galois connections:
  - the upper closure operator $\gamma \circ \alpha$,
  - the lower closure operator $\alpha \circ \gamma$;
- soundness:
  - the abstraction forgets no behavior;
- completeness:
  - sufficient conditions that ensure the absence of false positive;

on an abstraction of the reachable connected components in a site-graph rewriting language.

Joint-work with...

Walter Fontana
Harvard Medical School

Vincent Danos
Edinburgh

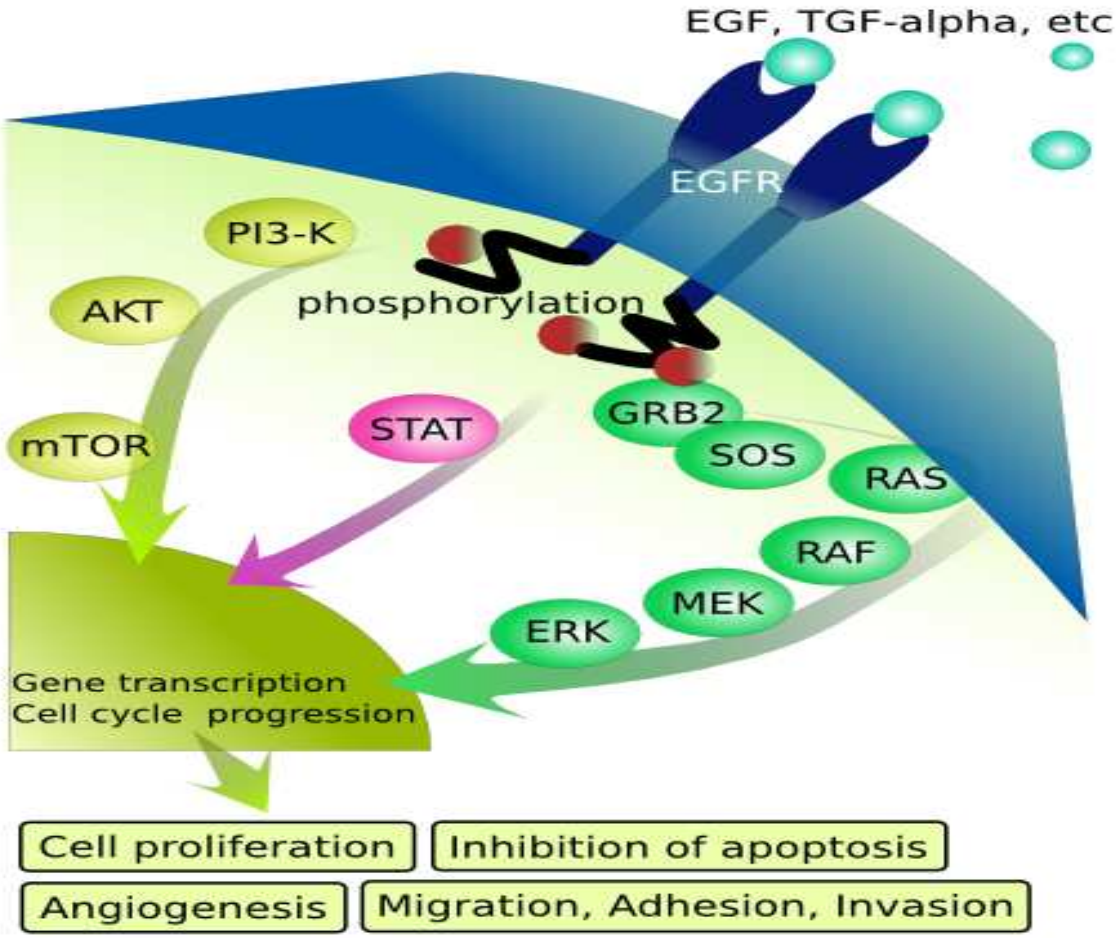
Russ Harmer
Paris VII

Jean Krivine
Paris VII

Overview

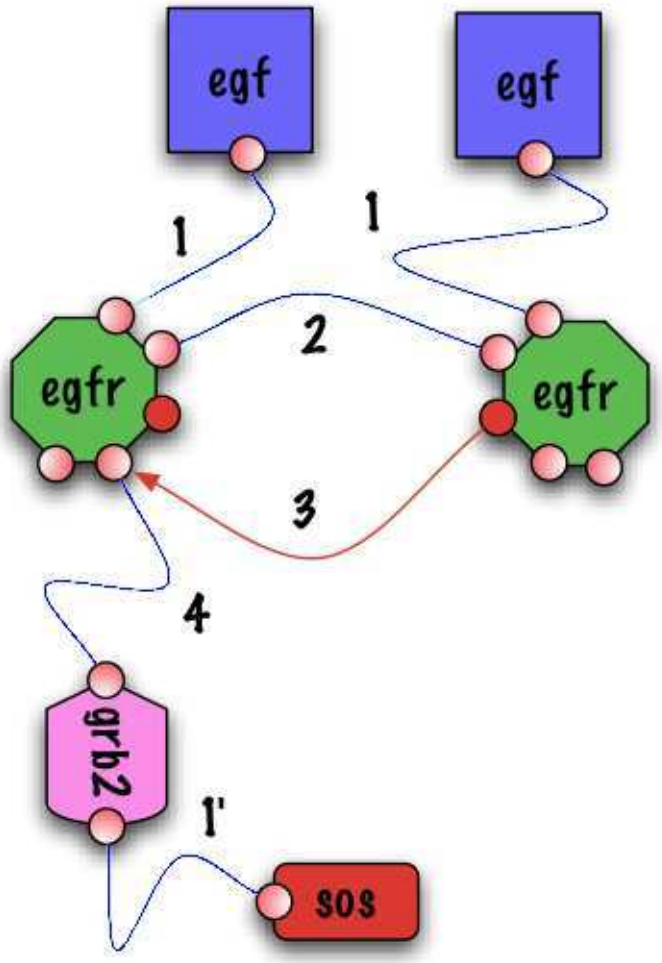
1. Introduction
2. Language: Kappa
3. Abstraction: Local views
4. Completeness: false positives?
5. Local fragment of Kappa
6. Decontextualization
7. Conclusion

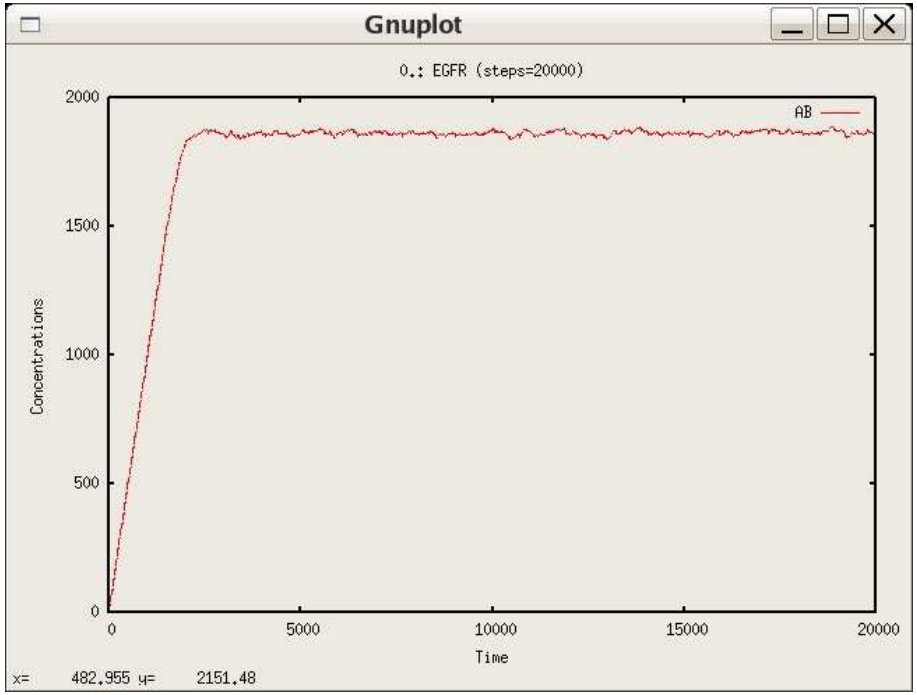
Signaling Pathways



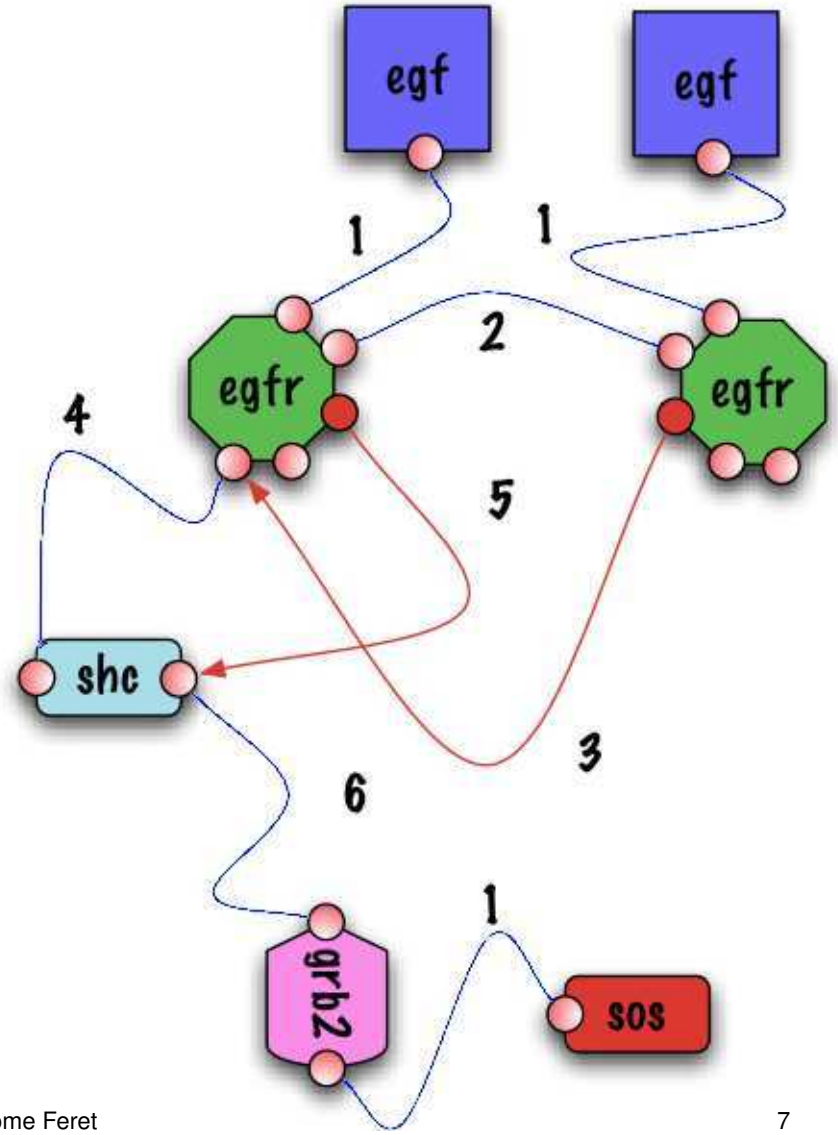
Eikuch, 2007

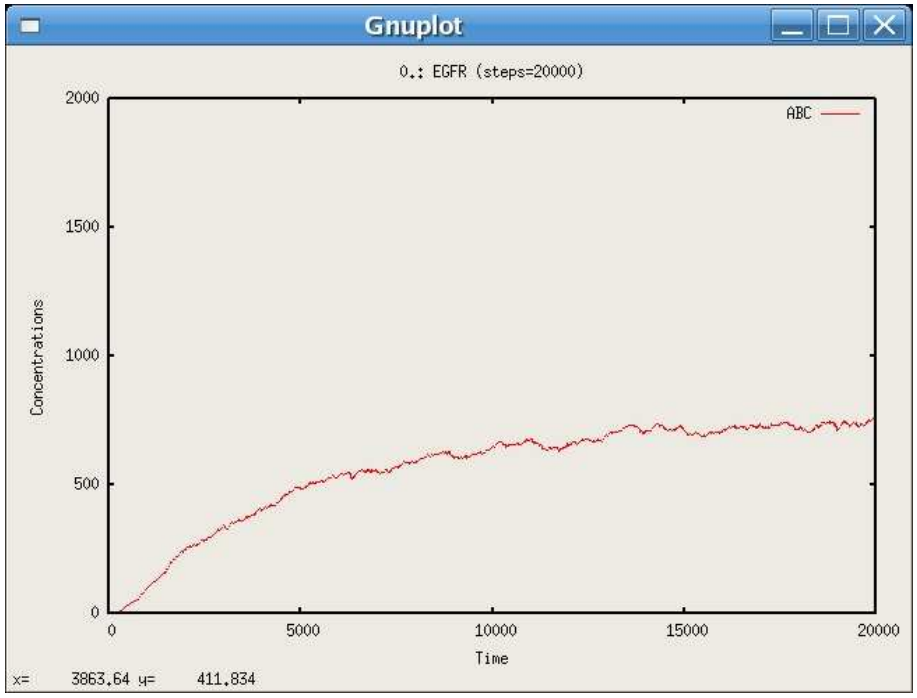
A single story





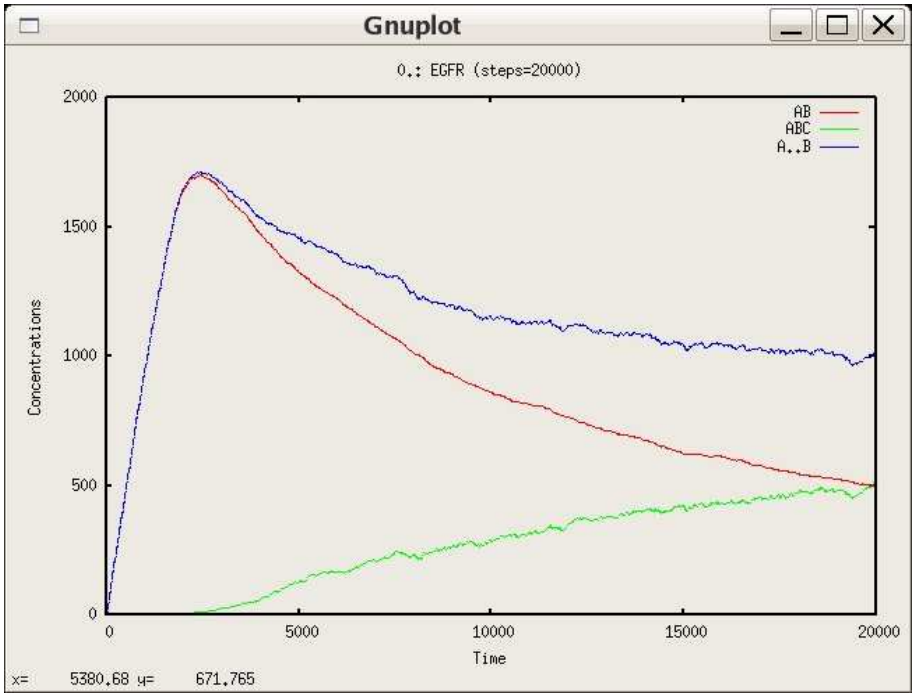
A concurrent story





Overshoot

When we combine the two stories. . .

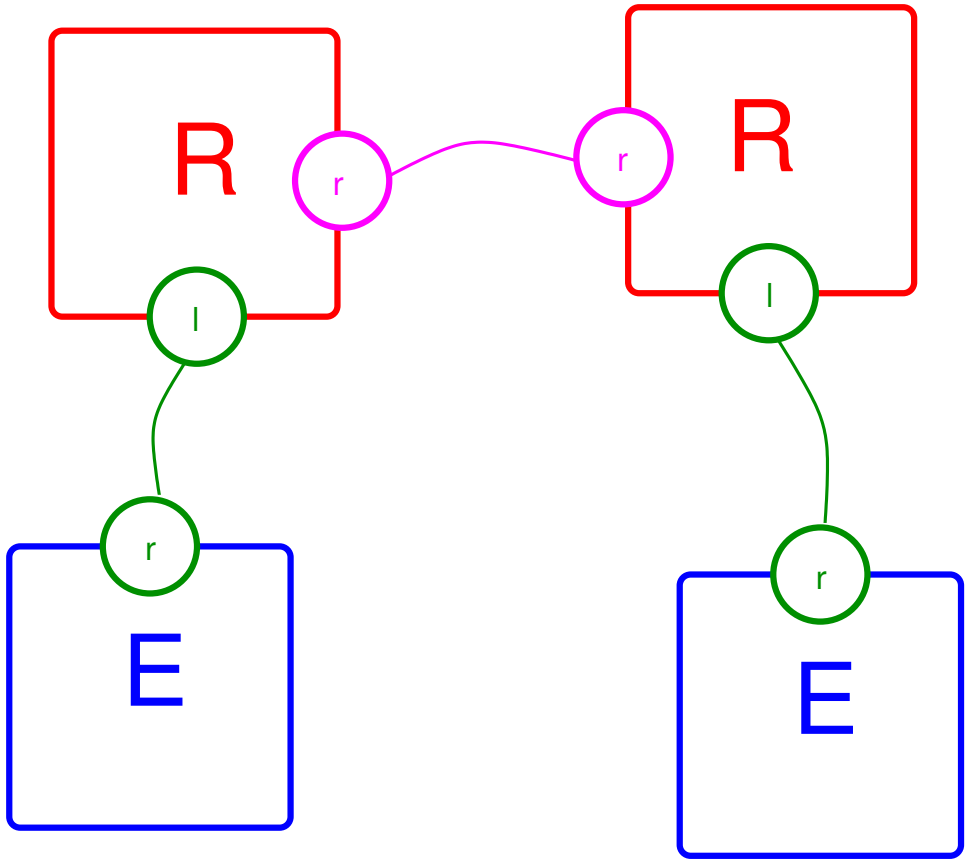


. . . we get an overshoot.

Overview

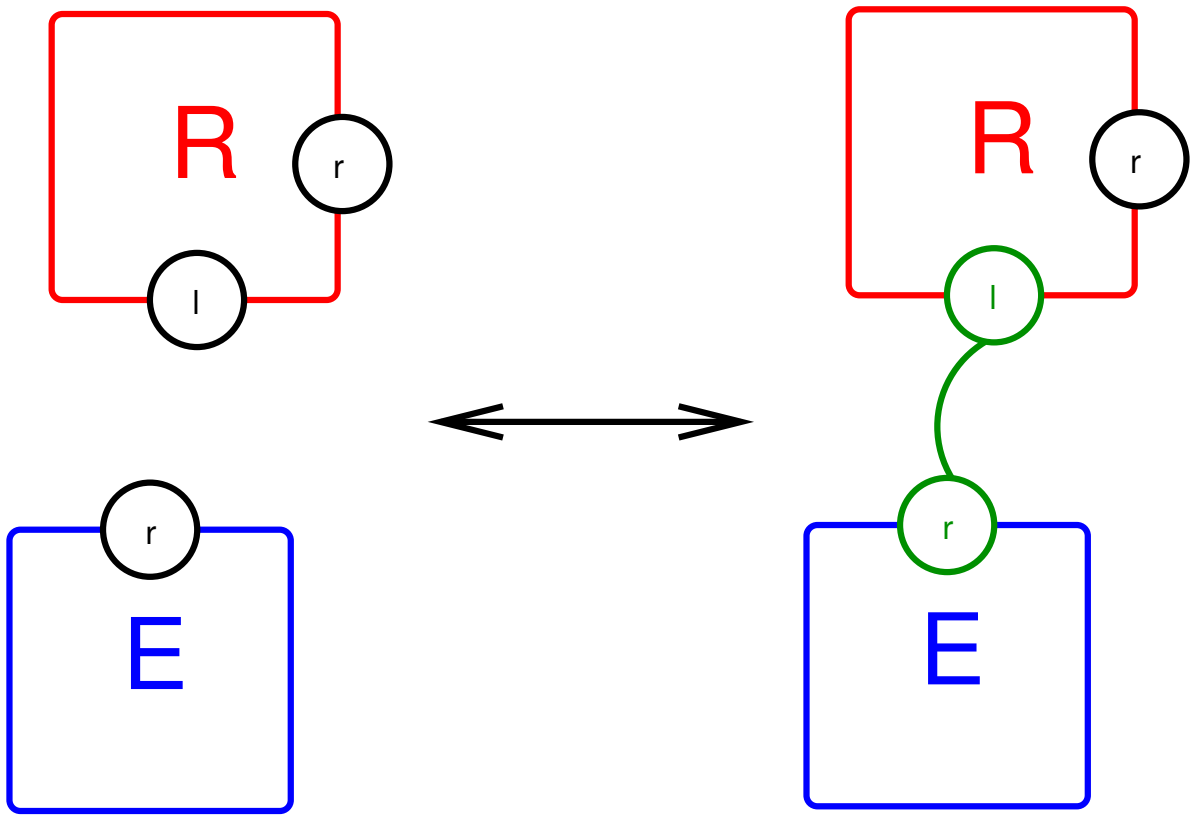
1. Introduction
2. Language: Kappa
3. Abstraction: Local views
4. Completeness: false positives?
5. Local fragment of Kappa
6. Decontextualization
7. Conclusion

A chemical species



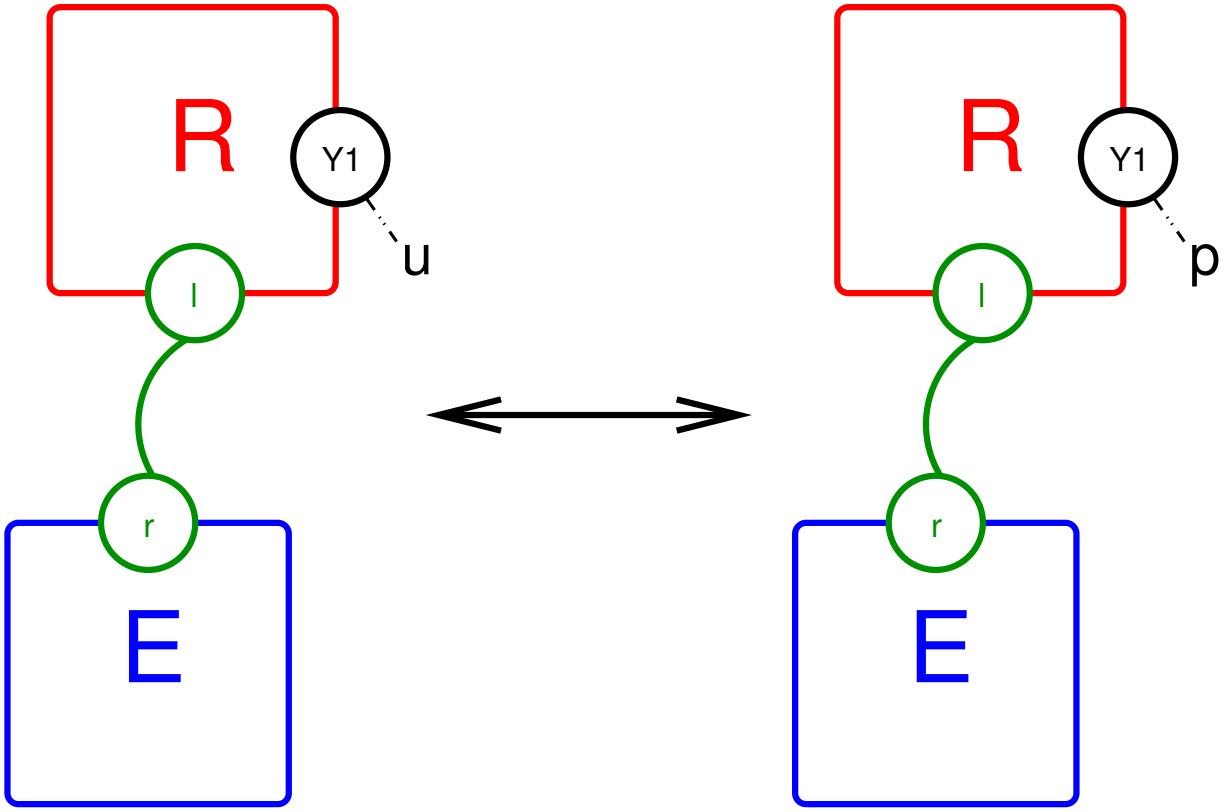
E(r!1), R(l!1,r!2), R(r!2,l!3), E(r!3)

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)

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)

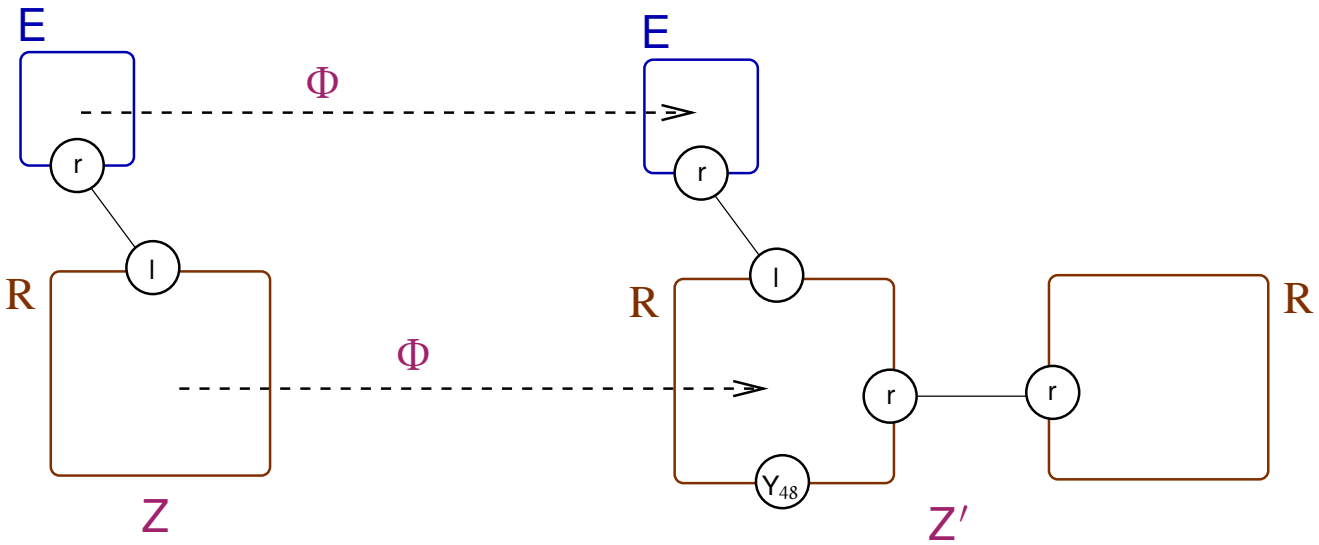
Properties of interest

1. Show the absence of modeling 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. Language: Kappa
3. Abstraction: Local views
4. Completeness: false positives?
5. Local fragment of Kappa
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:

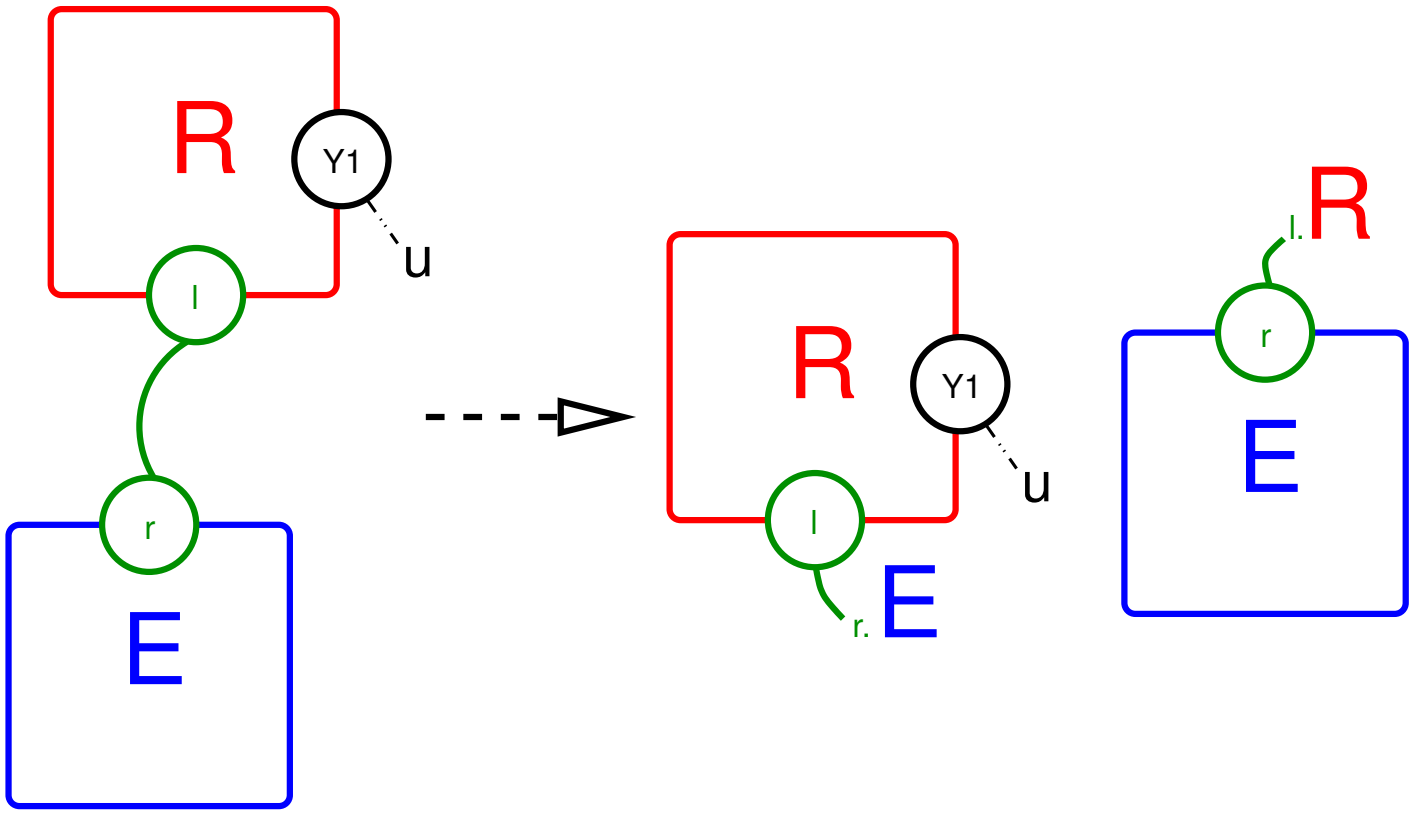
$$\mathbb{F} : \begin{cases} \wp(\mathit{Species}) & \to \wp(\mathit{Species}) \\ X & \mapsto X \cup \left\{ c'_j \;\middle|\; \begin{array}{l} \exists R_k \in \mathcal{R}, c_1, \ldots, c_m \in X, \\ c_1, \ldots, c_m \to_{R_k} c'_1, \ldots, c'_n \end{array} \right\}. \end{cases}$$

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 connection

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 chemical species 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 connection:

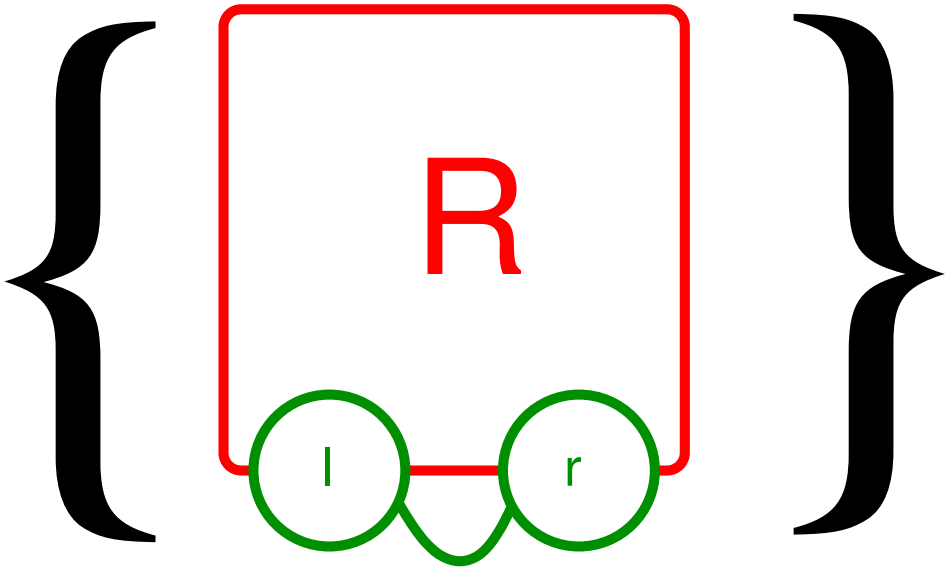
$$\wp(\mathit{Species}) \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \wp(\mathit{Local\_view}).$$

(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.

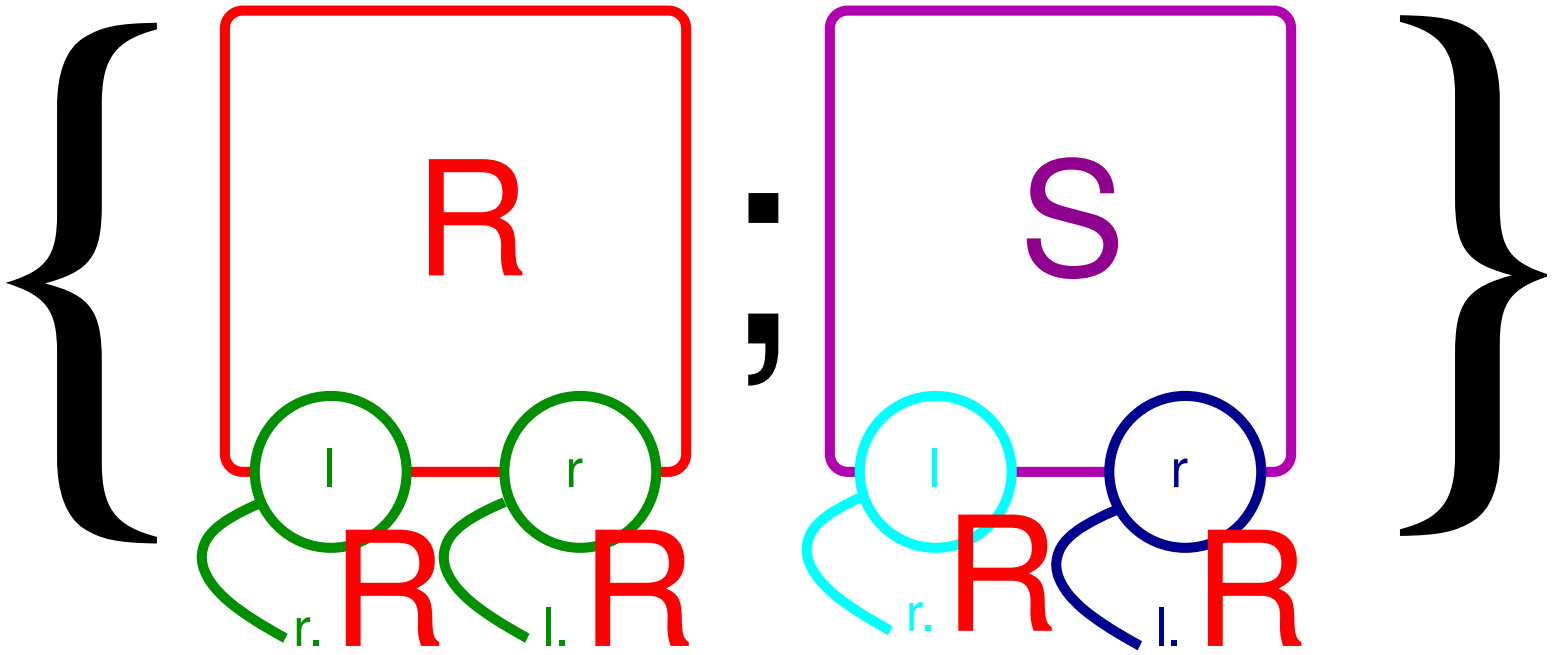
Guess the image of the following set of chemical species ?



$\alpha \circ \gamma$

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

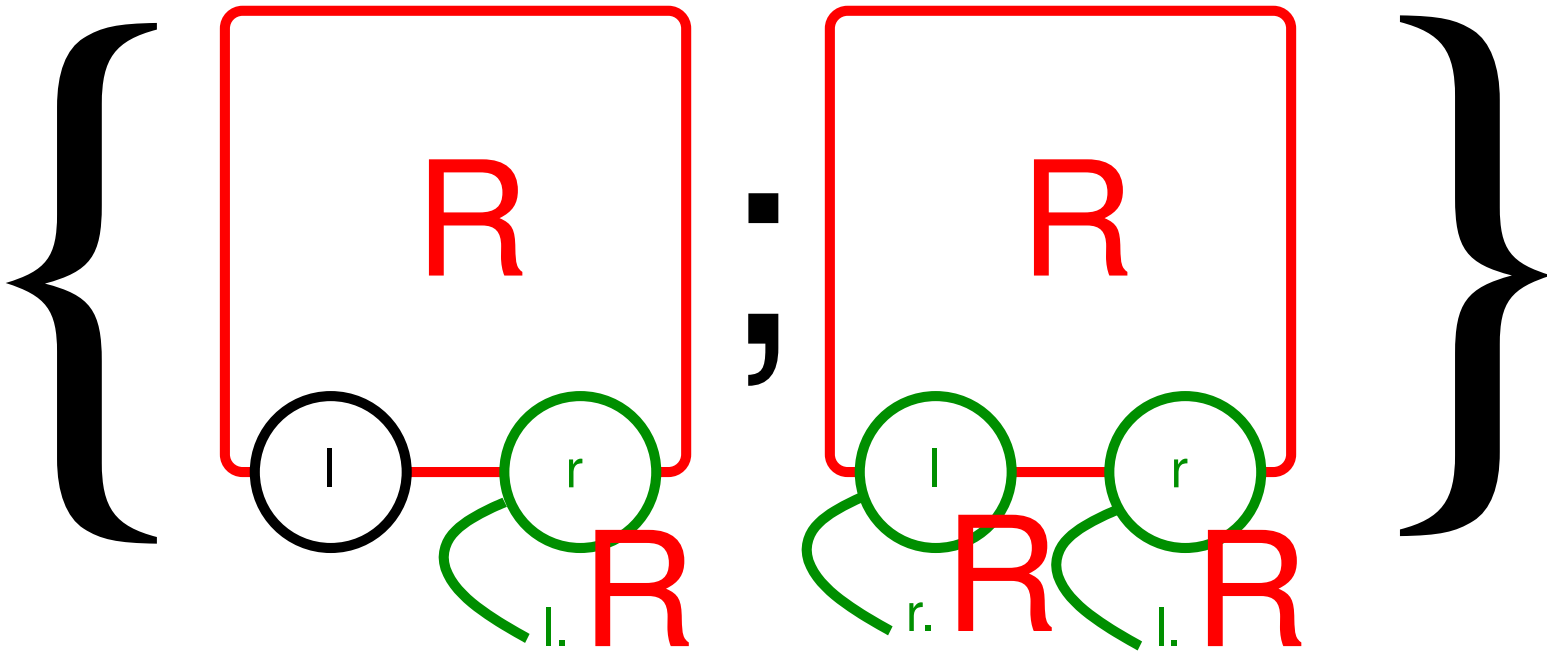
Guess the image of the following set of local views ?



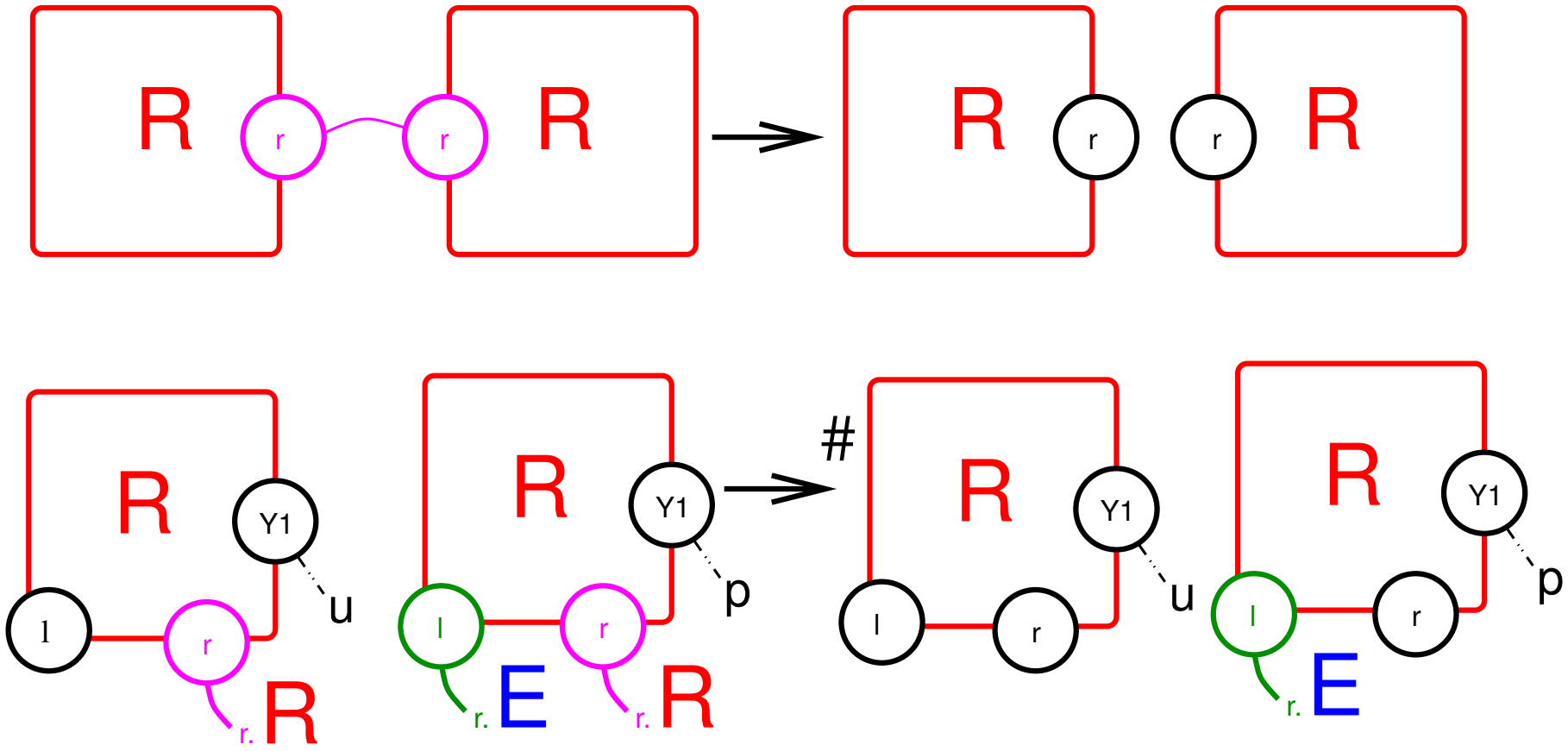
One more question

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

Guess the image of the following set of local views ?



Abstract reactions



Abstract counterpart to $\mathbb{F}$

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

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

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 connection;
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. $\mathit{lfp}_{X_0}\mathbb{F} \subseteq \gamma(\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp)$.

Overview

1. Introduction
2. Language: Kappa
3. Abstraction: Local views
4. Completeness: false positives?
5. Local fragment of Kappa
6. Decontextualization
7. Conclusion

Which information is abstracted away ?

Our analysis is exact (no false positive):

- for EGF cascade (356 chemical species);
- for FGF cascade (79080 chemical species);
- for SBF cascade (around 10^{19} chemical species).

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

This raises the following issues:

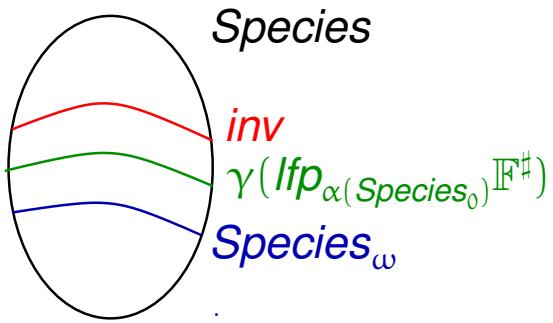
- Can we characterize which information is abstracted away ?
- Which is the form of the systems, for which we have no false positive ?
- Do we learn something about the biological systems that we describe ?

Which information is abstracted away ?

Theorem 2 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 connection;
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 $\alpha(\mathbb{F}(\mathit{inv})) = \mathbb{F}^\sharp(\alpha(\mathit{inv}))$;

Then, $\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp$ exists and $\gamma(\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp) \subseteq \mathit{inv}$.



Proof I/III

We have already seen (previous lectures) that:

1. $\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp$ exists;
2. there exists an ordinal δ such that $\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp = \mathbb{F}^{\sharp\delta}(\alpha(X_0))$.

Proof II/III

Let us show that $\gamma(\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp) \subseteq \mathit{inv}$.

Let us prove instead by induction over δ that $\mathbb{F}^{\sharp\delta}(\alpha(X_0)) \sqsubseteq \alpha(\mathit{inv})$.

- If $Y \in D^\sharp$ is an element such that $Y \sqsubseteq \alpha(\mathit{inv})$,
  $\mathbb{F}^\sharp(Y) \sqsubseteq \mathbb{F}^\sharp(\alpha(\mathit{inv}))$ ($\mathbb{F}^\sharp$ is mon)
  $\mathbb{F}^\sharp(\alpha(\mathit{inv})) = \alpha(\mathbb{F}(\mathit{inv}))$ (assumption)
  $\alpha(\mathbb{F}(\mathit{inv})) \sqsubseteq \alpha(\mathit{inv})$. ($\alpha$ is mon and inv is a post)

  Thus: $\mathbb{F}^\sharp(Y) \sqsubseteq \alpha(\mathit{inv})$
- If $Y_i \in D^{\sharp I}$ is a chain of elements such that $Y_i \sqsubseteq \alpha(\mathit{inv})$ for any $i \in I$, then, $\sqcup Y_i \sqsubseteq \alpha(\mathit{inv})$ (lub).

So: $\mathbb{F}^{\sharp\delta}(\alpha(X_0)) \sqsubseteq \alpha(\mathit{inv})$.

Proof III/III

We have:

$$\mathbb{F}^{\sharp\delta}(\alpha(X_0)) \sqsubseteq \alpha(\mathit{inv}).$$

Since γ is monotonic:

$$\gamma(\mathbb{F}^{\sharp\delta}(\alpha(X_0))) \subseteq \gamma(\alpha(\mathit{inv})).$$

But, by assumption, $\gamma(\alpha(\mathit{inv})) = \mathit{inv}$.
Thus,

$$\gamma(\mathbb{F}^{\sharp\delta}(\alpha(X_0))) \subseteq \mathit{inv}.$$

When is there no false positive ?

Theorem 3 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 connection;
3. $\mathbb{F} : \mathrm{D} \to \mathrm{D}$ is a monotonic map;
4. X_0 is a concrete element such that $\mathrm{X}_0 \subseteq \mathbb{F}(\mathrm{X}_0)$;
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:

- $\mathit{lfp}_{\mathrm{X}_0}\mathbb{F}$ and $\mathit{lfp}_{\alpha(\mathrm{X}_0)}\mathbb{F}^\sharp$ exist;
- $\mathit{lfp}_{\mathrm{X}_0}\mathbb{F} = \gamma(\alpha(\mathit{lfp}_{\mathrm{X}_0}\mathbb{F})) \Longleftrightarrow \mathit{lfp}_{\mathrm{X}_0}\mathbb{F} = \gamma(\mathit{lfp}_{\alpha(\mathrm{X}_0)}\mathbb{F}^\sharp)$.

Proof I/V

The (transfinite) sequence $(\mathbb{F}^{\sharp o}(\alpha(X_0)))$ is defined and increasing since:

1. $\mathbb{F}^\sharp \circ \alpha = \alpha \circ \mathbb{F} \circ \gamma \circ \alpha$, so:
   - $\mathbb{F}^\sharp \circ \alpha$ is monotonic;
   - for any $X \in D$,
     there exists $X' \in D$, such that $\mathbb{F}^\sharp(\alpha(X)) = \alpha(X')$;
2. Moreover, α is a $\cup/\sqcup$-complete morphism, so:
   - for any increasing chain $(\alpha(X_i))_{i \in I}$,
     there exists $X' \in D$, such that $\sqcup \alpha(X_i) = \alpha(X')$.

So there exists an ordinal δ such that:

$$\mathbb{F}^\sharp(\mathbb{F}^{\sharp\delta}(\alpha(X_0))) = \mathbb{F}^{\sharp\delta}(\alpha(X_0)).$$

Proof II/V

Let us show that $\mathbb{F}^{\sharp\delta}(\alpha(X_0))$ is a least fixpoint (greater than $\alpha(X_0)$):

Let Y be such that $\mathbb{F}^{\sharp}(Y) = Y$ and $\alpha(X_0) \sqsubseteq Y$.
Let us prove that $\mathbb{F}^{\sharp\delta}(\alpha(X_0)) \sqsubseteq Y$ by (transfinite) induction:

1. If $\alpha(X) \sqsubseteq Y$, we have: $\mathbb{F}^{\sharp}(\alpha(X)) \sqsubseteq Y$.

   Since:

   $\mathbb{F}^{\sharp}(\alpha(X)) = \alpha(\mathbb{F}(\gamma(\alpha(X))))$
   $\alpha(\mathbb{F}(\gamma(\alpha(X)))) \sqsubseteq \alpha(\mathbb{F}(\gamma(Y)))$ ($\alpha \circ \mathbb{F} \circ \gamma$ is mon)
   $\alpha(\mathbb{F}(\gamma(Y))) \sqsubseteq \alpha(\gamma(\mathbb{F}^{\sharp}(Y)))$ ($\mathbb{F} \circ \gamma \dot{\sqsubseteq} \gamma \circ \mathbb{F}^{\sharp}$ and α mon)
   $\alpha(\gamma(\mathbb{F}^{\sharp}(Y))) \sqsubseteq \mathbb{F}^{\sharp}(Y)$ ($\alpha \circ \gamma$ is reductive)
   $\mathbb{F}^{\sharp}(Y) = Y$ (fixpoint)

2. If $(X_i)_{i \in I} \in D^I$ is such that:
   - $(\alpha(X_i))_{i \in I}$ is a chain of elements;
   - $\alpha(X_i) \sqsubseteq Y$ for any $i \in I$.

   Then $\sqcup \alpha(X_i) \sqsubseteq Y$ ($\sqcup$ is a least upper bound)

Proof III/V

By induction, we have proved that:

1. $\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp$ exists;
2. $\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp = \mathbb{F}^{\sharp\delta}(\alpha(X_0))$.

Proof IV/V

Let us now show that:

$$\mathit{lfp}_{X_0}\mathbb{F} = \gamma(\alpha(\mathit{lfp}_{X_0}\mathbb{F})) \iff \mathit{lfp}_{X_0}\mathbb{F} = \gamma(\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp).$$

- Easy implication: $\Leftarrow$.
  If $\mathit{lfp}_{X_0}\mathbb{F} = \gamma(\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp)$,
  then $\mathit{lfp}_{X_0}\mathbb{F} = \gamma(\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp) = \gamma(\alpha(\gamma(\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp))) = \gamma(\alpha(\mathit{lfp}_{X_0}\mathbb{F}))$.

- Other implication:
  - (easy inclusion)
    We always have $\mathit{lfp}_{X_0}\mathbb{F} \subseteq \gamma(\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp)$.

Proof V/V

- other inclusion

  We assume that $\gamma(\alpha(\mathit{lfp}_{X_0}\mathbb{F})) = \mathit{lfp}_{X_0}\mathbb{F}$.

  We have to prove that: $\gamma(\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp) \subseteq \mathit{lfp}_{X_0}\mathbb{F}$.

  We have: $\mathbb{F}^\sharp(\alpha(\mathit{lfp}_{X_0}\mathbb{F})) = \alpha(\mathit{lfp}_{X_0}\mathbb{F})$.

  Since:

  $\mathbb{F}^\sharp(\alpha(\mathit{lfp}_{X_0}\mathbb{F})) = \alpha(\mathbb{F}(\gamma(\alpha(\mathit{lfp}_{X_0}\mathbb{F}))))$ $(\mathbb{F}^\sharp \circ \alpha = \alpha \circ \mathbb{F} \circ \gamma \circ \alpha)$

  $\alpha(\mathbb{F}(\gamma(\alpha(\mathit{lfp}_{X_0}\mathbb{F})))) = \alpha(\mathbb{F}(\mathit{lfp}_{X_0}\mathbb{F}))$ $(\gamma(\alpha(\mathit{lfp}_{X_0}\mathbb{F})) = \mathit{lfp}_{X_0}\mathbb{F})$

  $\alpha(\mathbb{F}(\mathit{lfp}_{X_0}\mathbb{F})) = \alpha(\mathit{lfp}_{X_0}\mathbb{F})$ (fixpoint)

  It follows that: $\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp \sqsubseteq \alpha(\mathit{lfp}_{X_0}\mathbb{F})$ (lfp and $\alpha(X_0) \sqsubseteq \alpha(\mathit{lfp}_{X_0}\mathbb{F})$)

  Then, since γ is mon:

$$\gamma(\mathit{lfp}_{\alpha(X_0)}\mathbb{F}^\sharp) \sqsubseteq \gamma(\alpha(\mathit{lfp}_{X_0}\mathbb{F})) = \mathit{lfp}_{X_0}\mathbb{F}.$$

When is there no false positive ?

Theorem 3 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 connection;
3. $\mathbb{F} : \mathrm{D} \to \mathrm{D}$ is a monotonic map;
4. X_0 is a concrete element such that $\mathrm{X}_0 \subseteq \mathbb{F}(\mathrm{X}_0)$;
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:

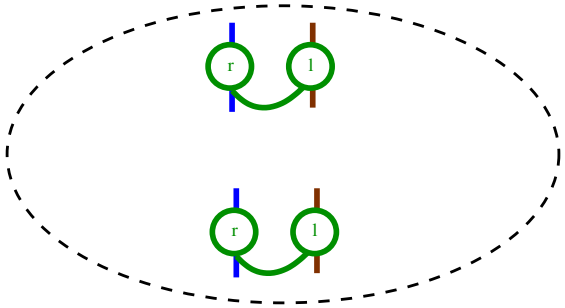
- $\mathit{lfp}_{\mathrm{X}_0}\mathbb{F}$ and $\mathit{lfp}_{\alpha(\mathrm{X}_0)}\mathbb{F}^\sharp$ exist;
- $\mathit{lfp}_{\mathrm{X}_0}\mathbb{F} = \gamma(\alpha(\mathit{lfp}_{\mathrm{X}_0}\mathbb{F})) \Longleftrightarrow \mathit{lfp}_{\mathrm{X}_0}\mathbb{F} = \gamma(\mathit{lfp}_{\alpha(\mathrm{X}_0)}\mathbb{F}^\sharp)$.

We need to understand under which assumptions $\mathit{lfp}_{\mathrm{X}_0}\mathbb{F} = \gamma(\alpha(\mathit{lfp}_{\mathrm{X}_0}\mathbb{F}))$.

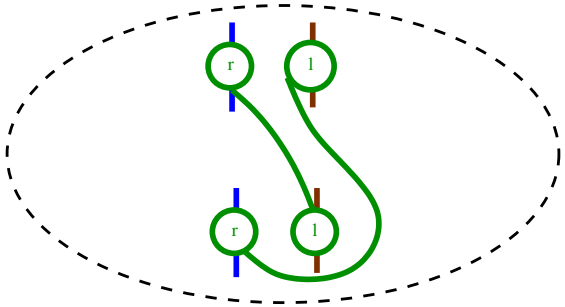
Swapping relation

We define the binary relation $\overset{\text{SWAP}}{\sim}$ among tuples $\mathit{Species}^*$ 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 4 Let $X \in \wp(\mathit{Species})$ be a set of chemical species.

The two following assertions are equivalent:

1. $X = \gamma(\alpha(X))$;
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 $(D_j) \in X^*$.

Proof (easier implication way)

If:

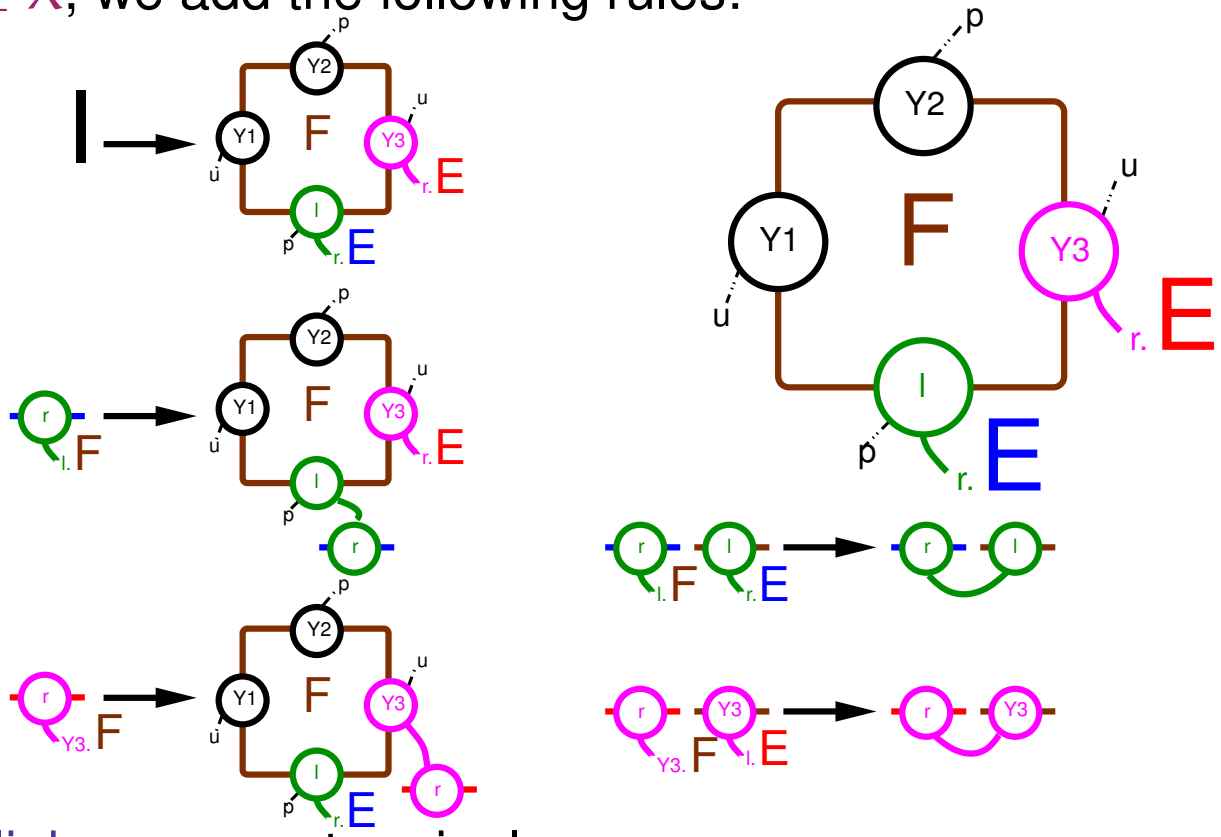
- $X = \gamma(\alpha(X))$,
- $(C_i)_{i \in I} \in X^*$,
- and $(C_i)_{i \in I} \overset{\text{SWAP}}{\sim} (D_j)_{j \in J}$;

Then:

we have $\alpha(\{C_i \mid i \in I\}) = \alpha(\{D_j \mid j \in J\})$	(because $(C_i) \overset{\text{SWAP}}{\sim} (D_j)$)
and $\alpha(\{C_i \mid i \in I\}) \subseteq \alpha(X)$	(because $(C_i) \in X^*$ and α mon);
so $\alpha(\{D_j \mid j \in J\}) \subseteq \alpha(X)$;	
so $\{D_j \mid j \in J\} \subseteq \gamma(\alpha(X))$	(by def. of Galois connections);
so $\{D_j \mid j \in J\} \subseteq X$	(since $X = \gamma(\alpha(X))$);
so $(D_j)_{j \in J} \in X^*$.	

Proof: more difficult implication way

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.

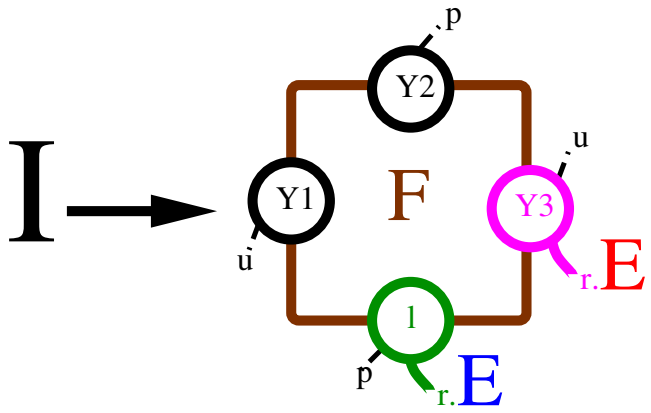
Proof (more difficult implication way)

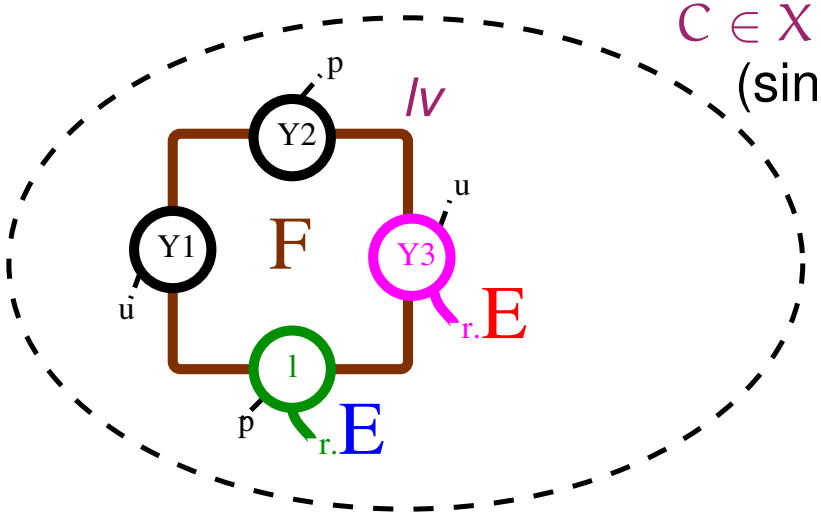
We suppose that X is close with respect to $\overset{\text{SWAP}}{\sim}$.
We want to prove that $\gamma(\alpha(X)) \subseteq X$.

We prove, by induction, that any open complex that can be built by gathering the views of $\alpha(X)$, can be embedded in a complex 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.

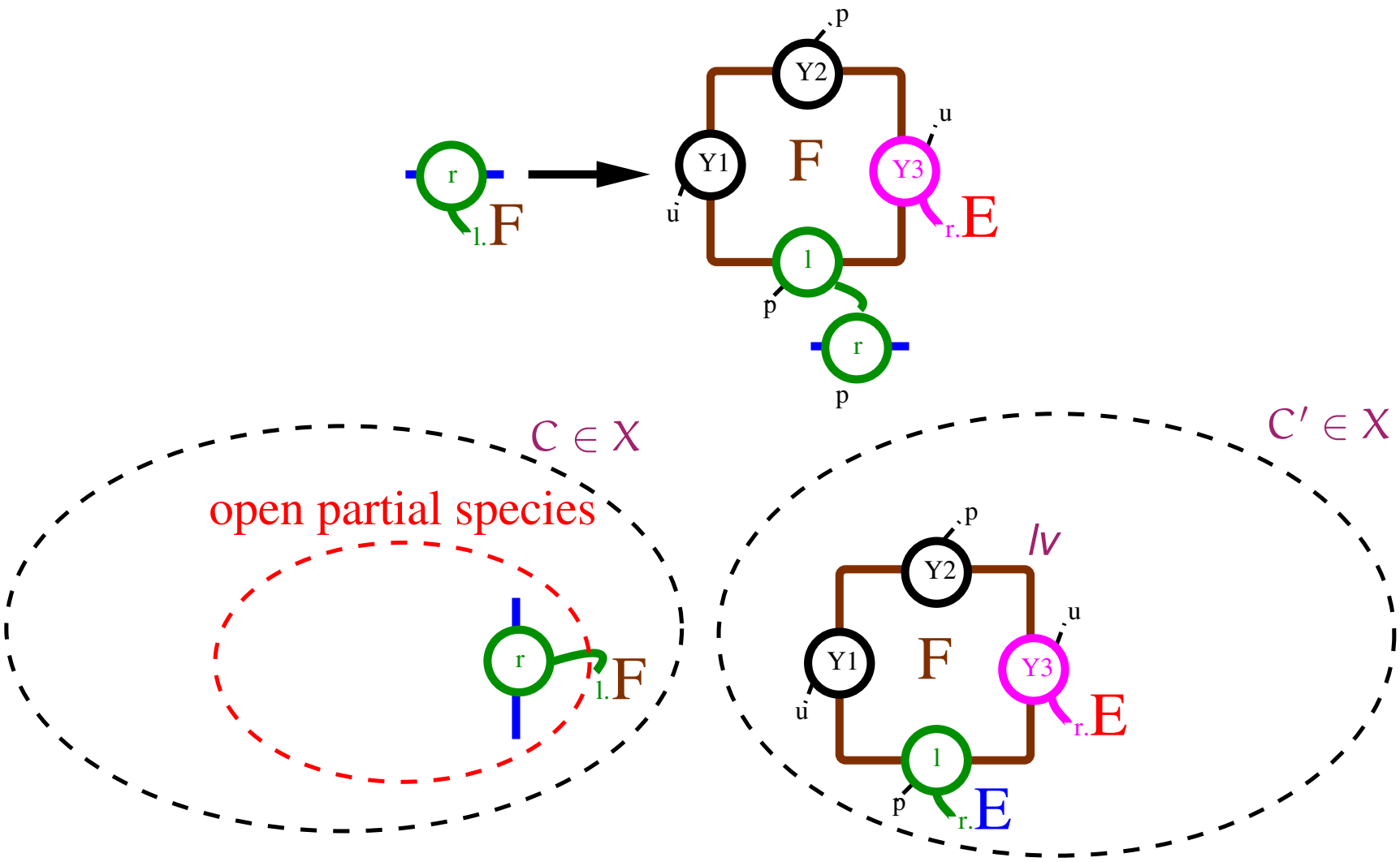
Initialization



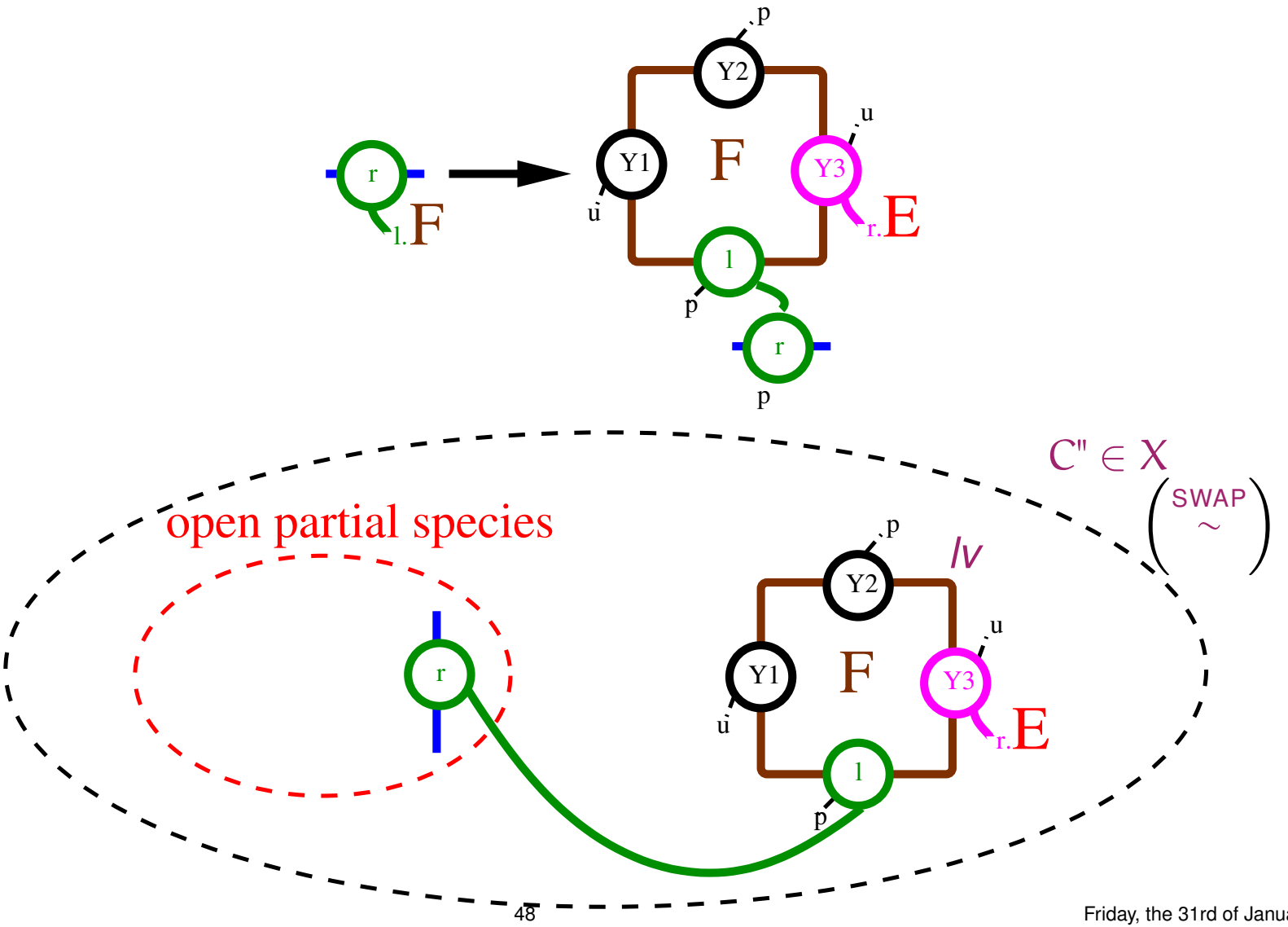


$C \in X$
(since $lv \in \alpha(X)$)

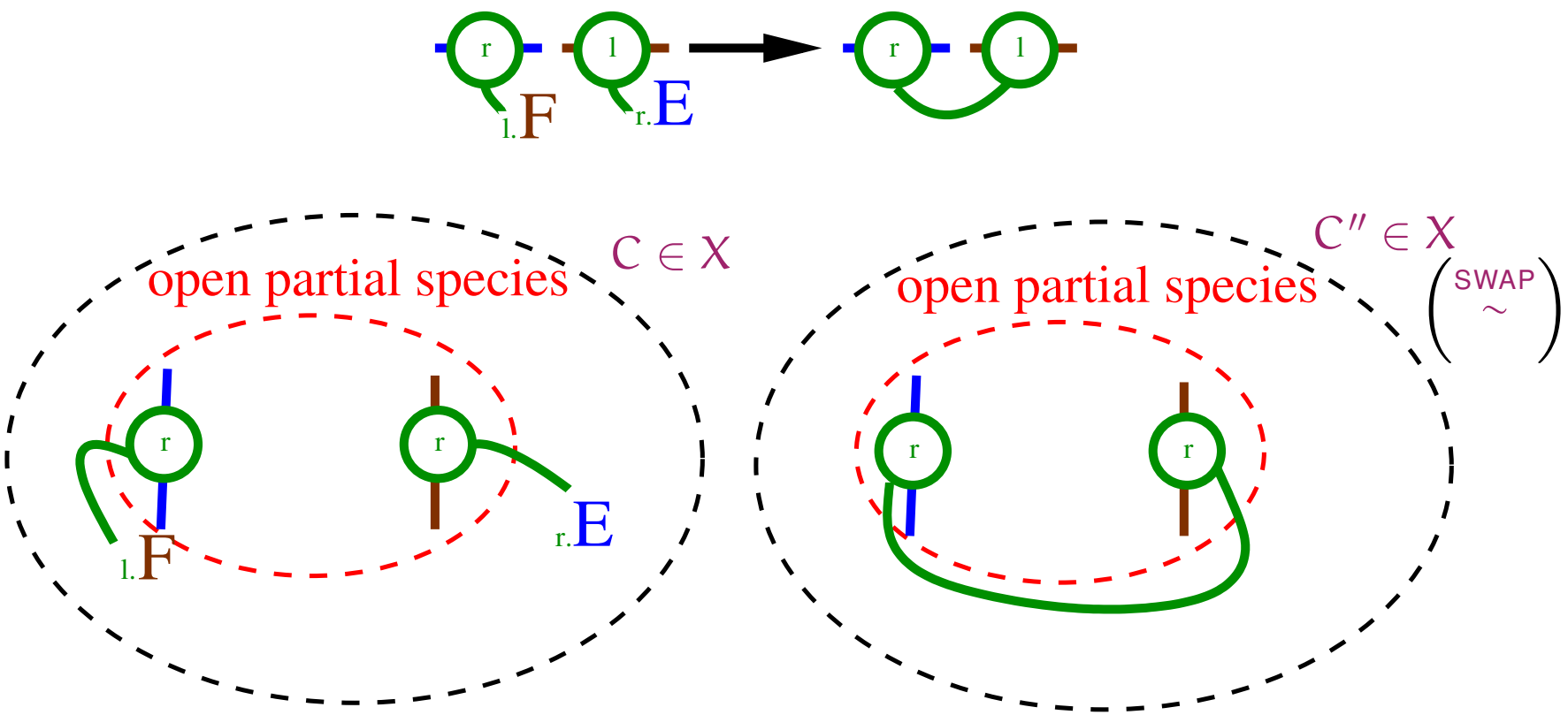
Unfolding a semi-link



Unfolding a semi-link



Binding two semi-links



Consequences

Let $Y \in \wp(\mathit{Local\_view})$ be a set of local views such that $\alpha(\gamma(Y)) = Y$.

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

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

   Thus:

   (a) $\gamma(Y)$ is finite if and only if the grammar has a finite set of prefixes (and the latter is decidable);

   (b) We have $\mathbb{F}^\sharp \circ \alpha = \alpha \circ \mathbb{F} \circ \gamma \circ \alpha$.

Overview

1. Introduction
2. Language: Kappa
3. Abstraction: Local views
4. Completeness: false positives?
5. Local fragment of Kappa
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,

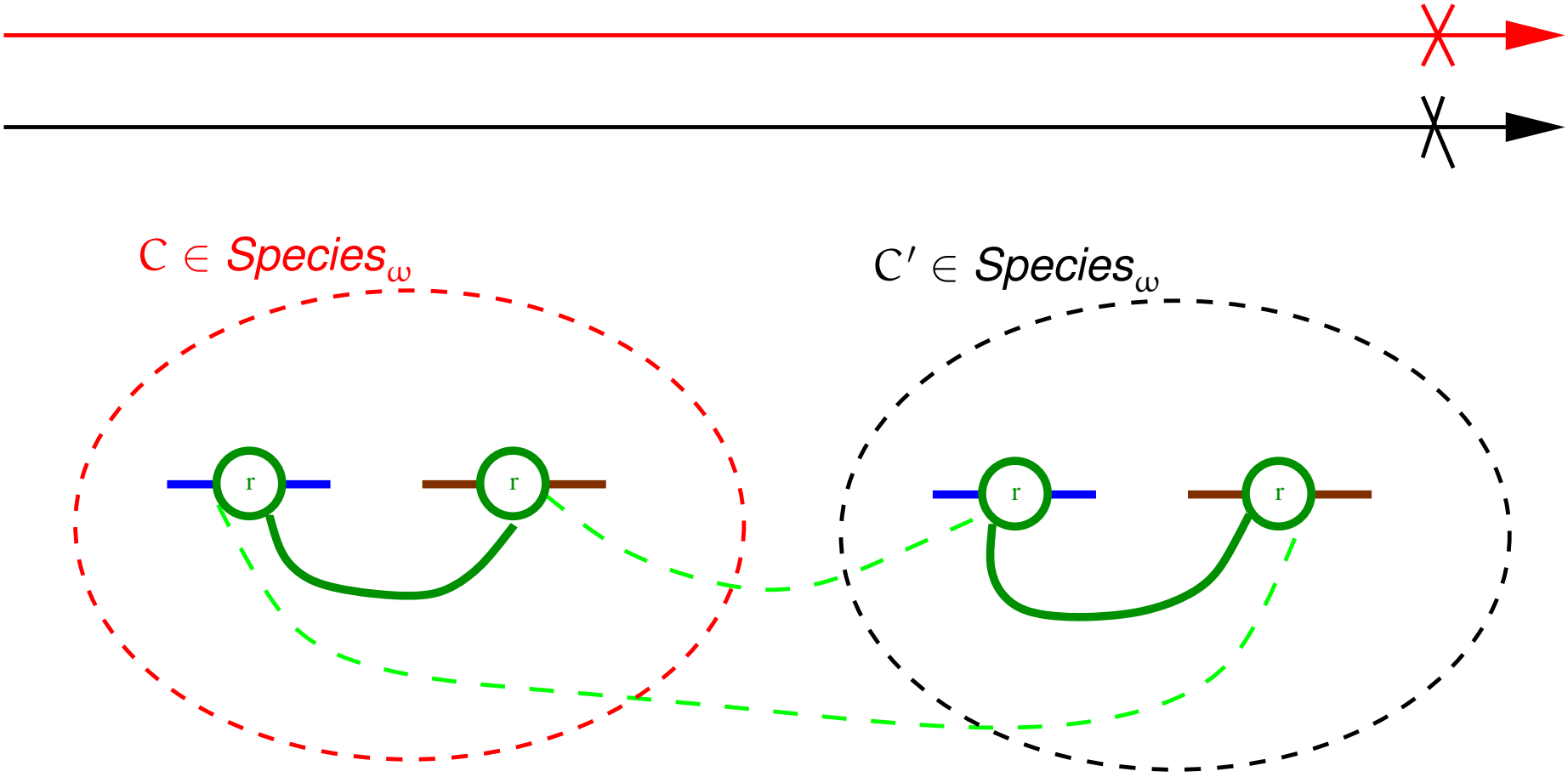
are satisfied, the set of reachable chemical species is local.

Proof outline

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

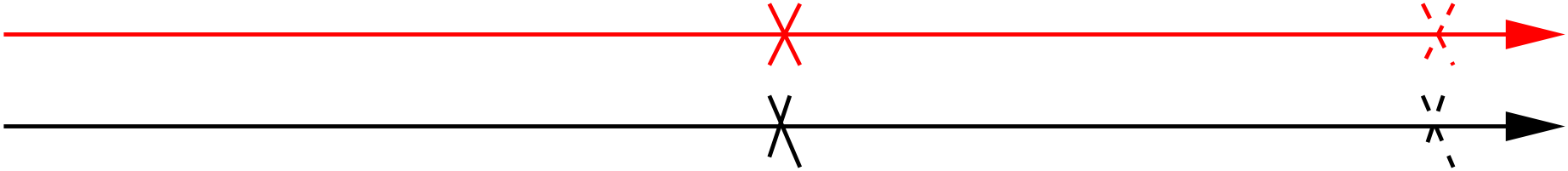
- We consider tuples of complexes in which the same kind of links occur twice.
- We want to swap these links.
- We introduce the history of their computation.
- There are several cases. . .

First case (I/V)

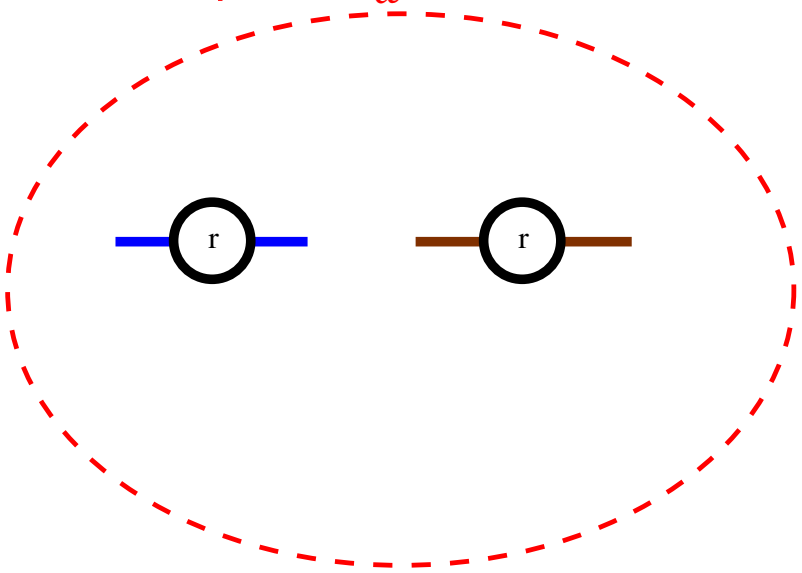


First case (II/V)

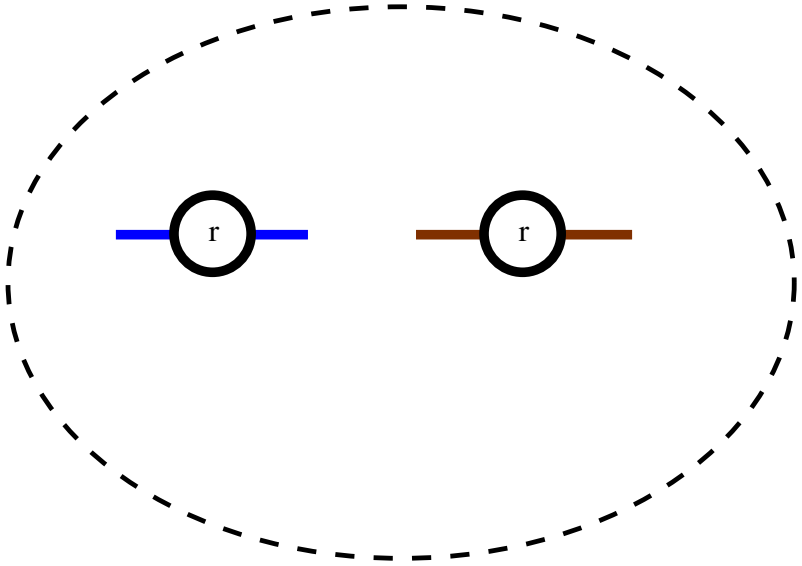
just before the links are made



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

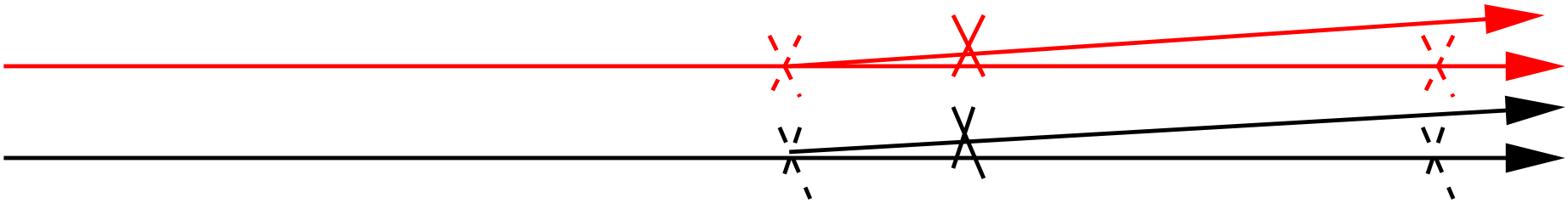


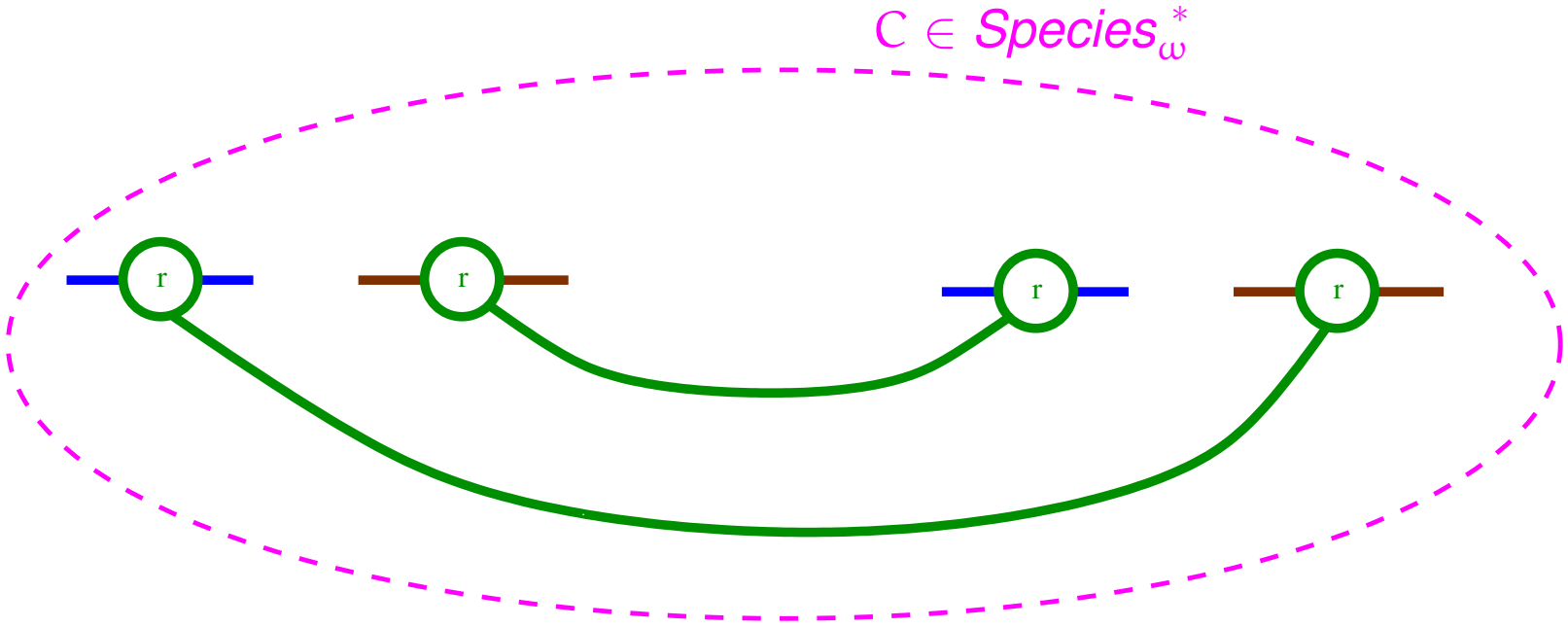
$\mathcal{C}' \in Species^*_\omega$



First case (III/V)

we suppose we can swap the links

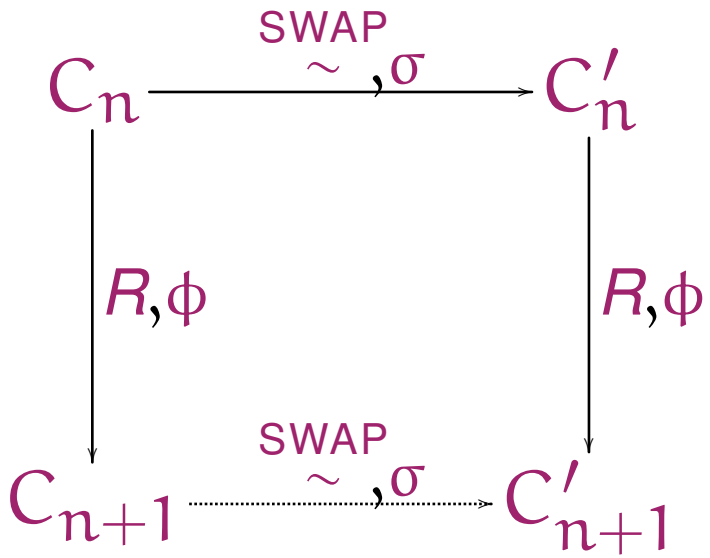




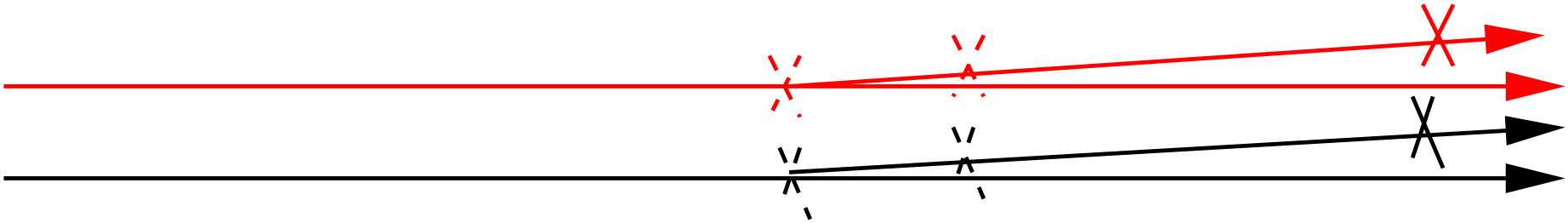
First case (IV/V)

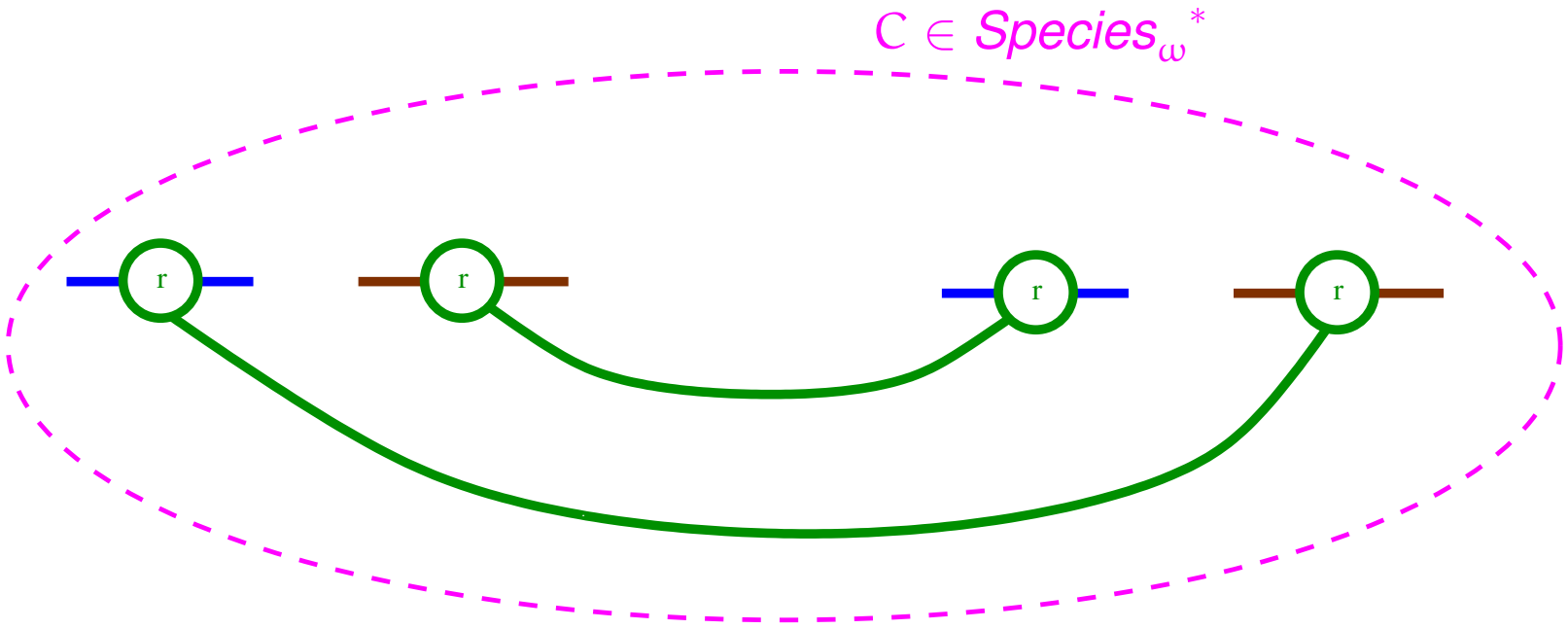
Then, we ensure that further computation steps:

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

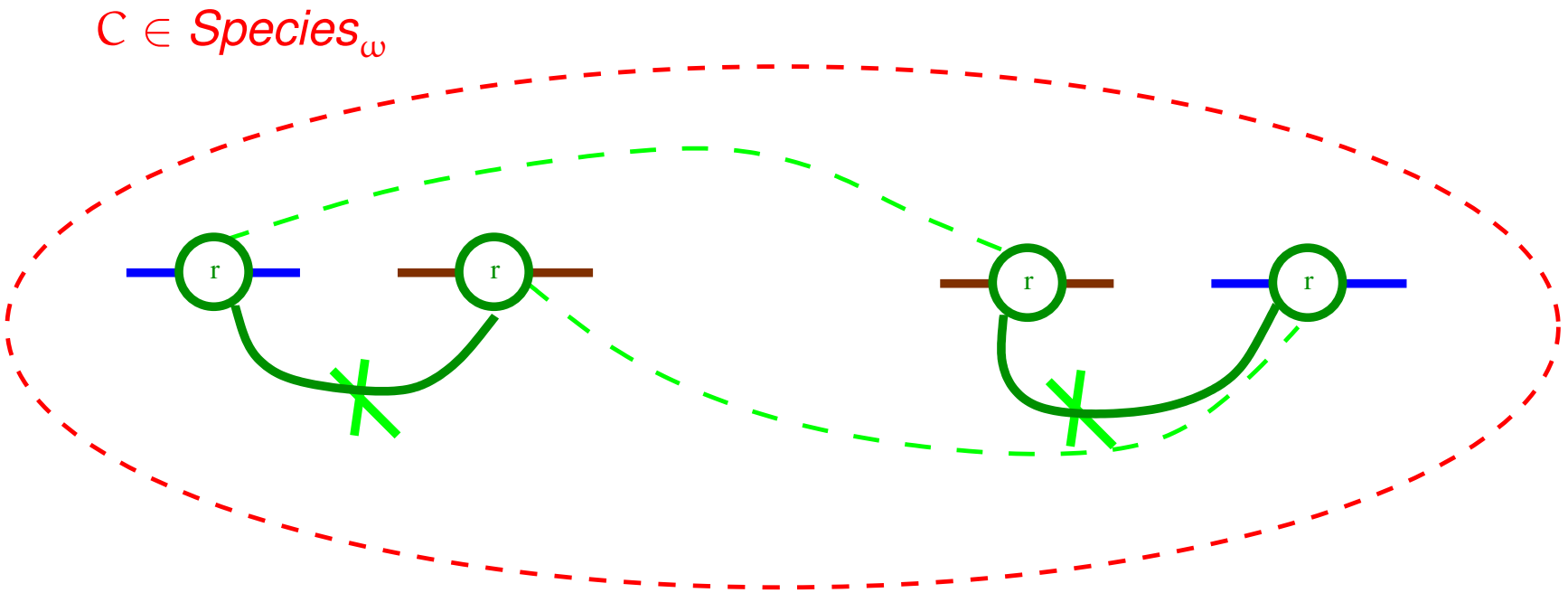


First case (V/V)



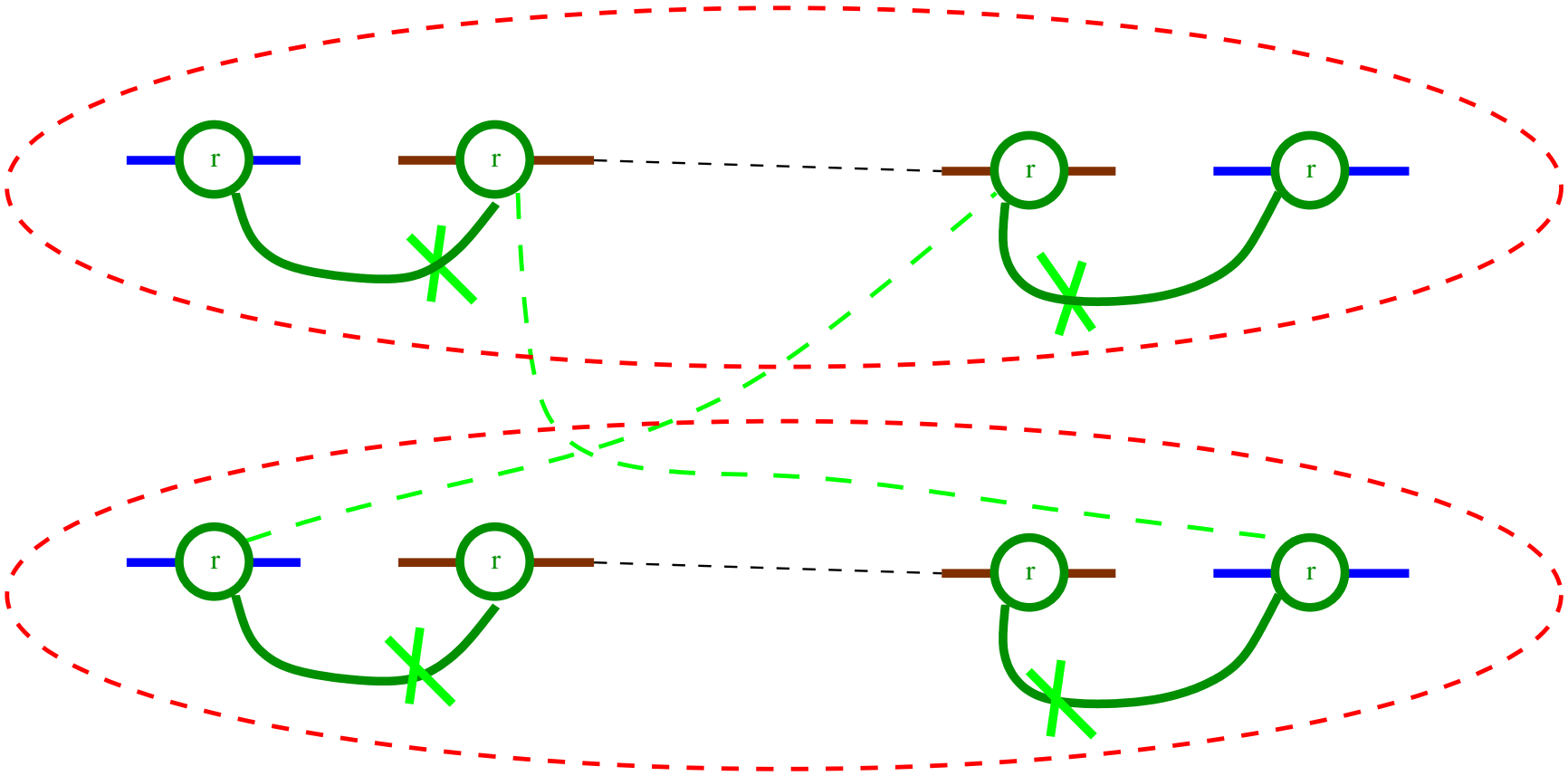


Second case (I/II)



we assume that the chemical species C is acyclic

Second case (II/II)



Sufficient conditions

Whenever the following assumptions:

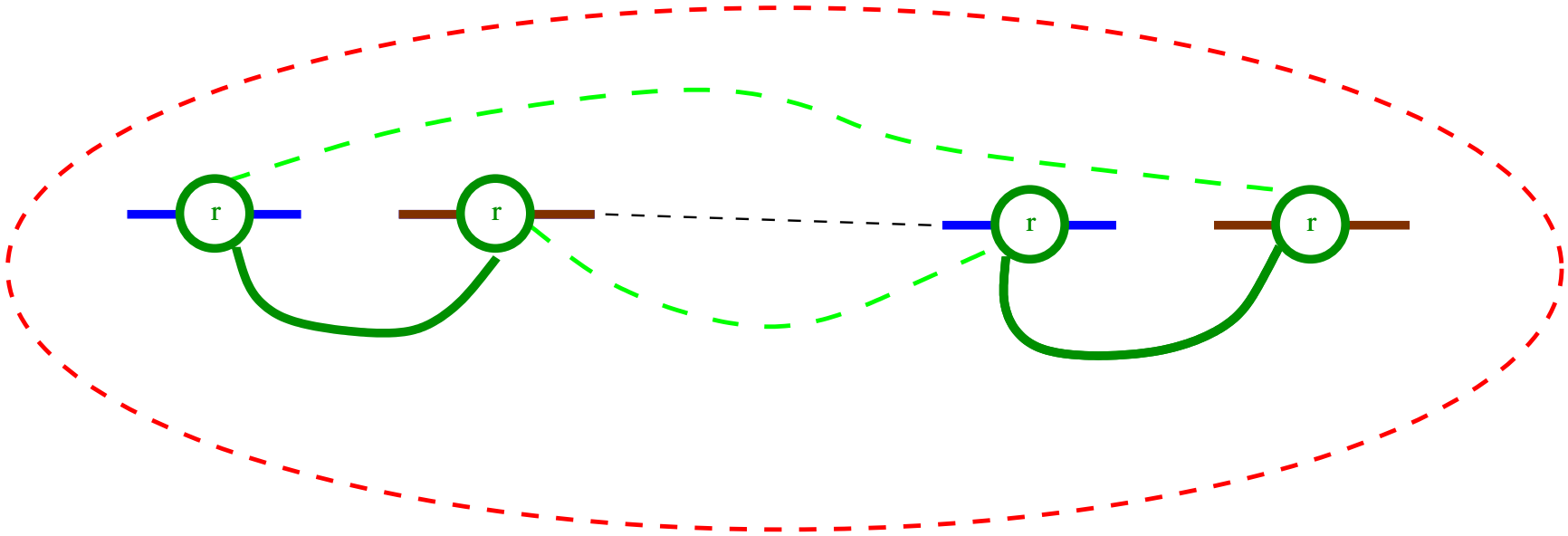
1. initial agents are not bound;
2. rules are atomic;
3. rules are local:
   - only agents that interact are tested,
   - no cyclic patterns (neither in lhs, nor in rhs);
4. binding rules do not interfere i.e. if both:
   - A(a~m,S),B(b~n,T) $\rightarrow$ A(a~m!1,S),B(b~n!1,T)
   - and A(a~m',S'),B(b~n',T')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

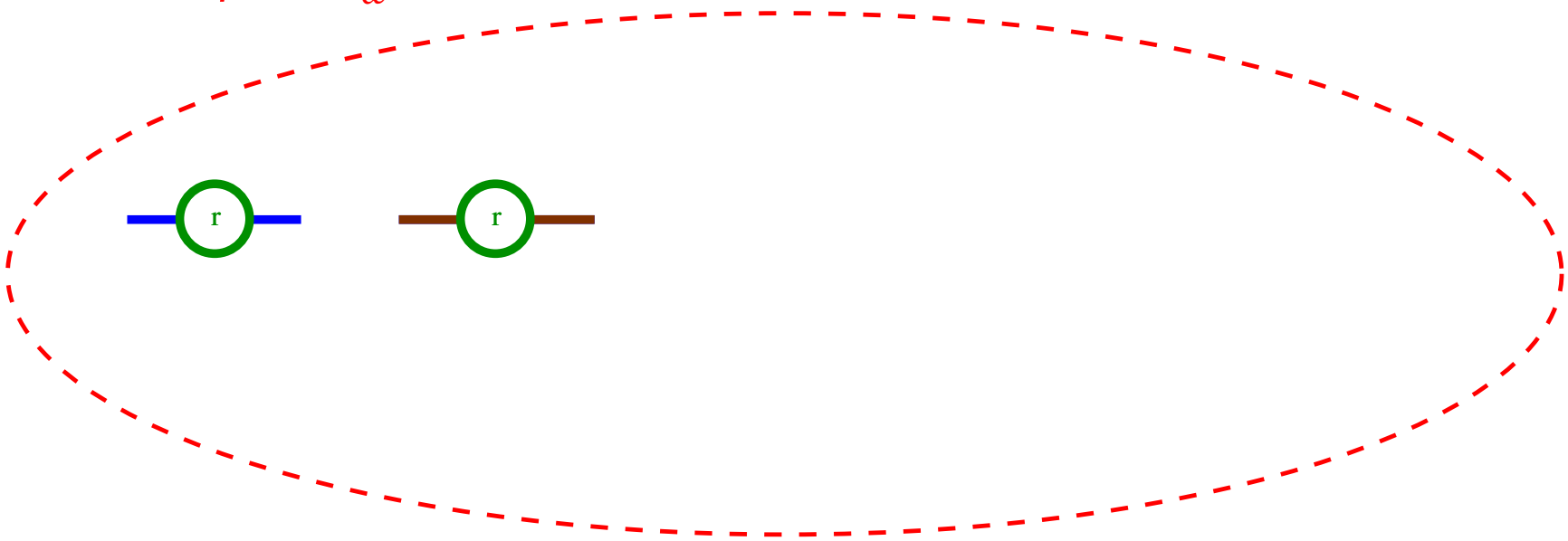
Third case (I/III)

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

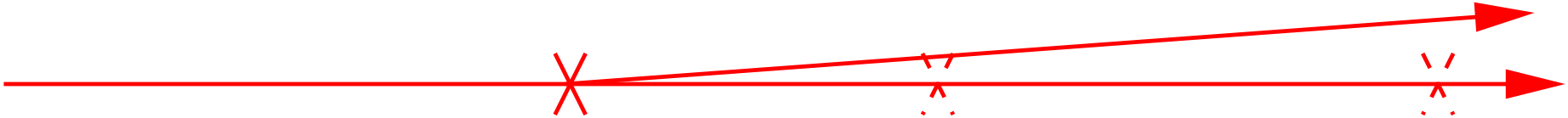


Third case (II/III)

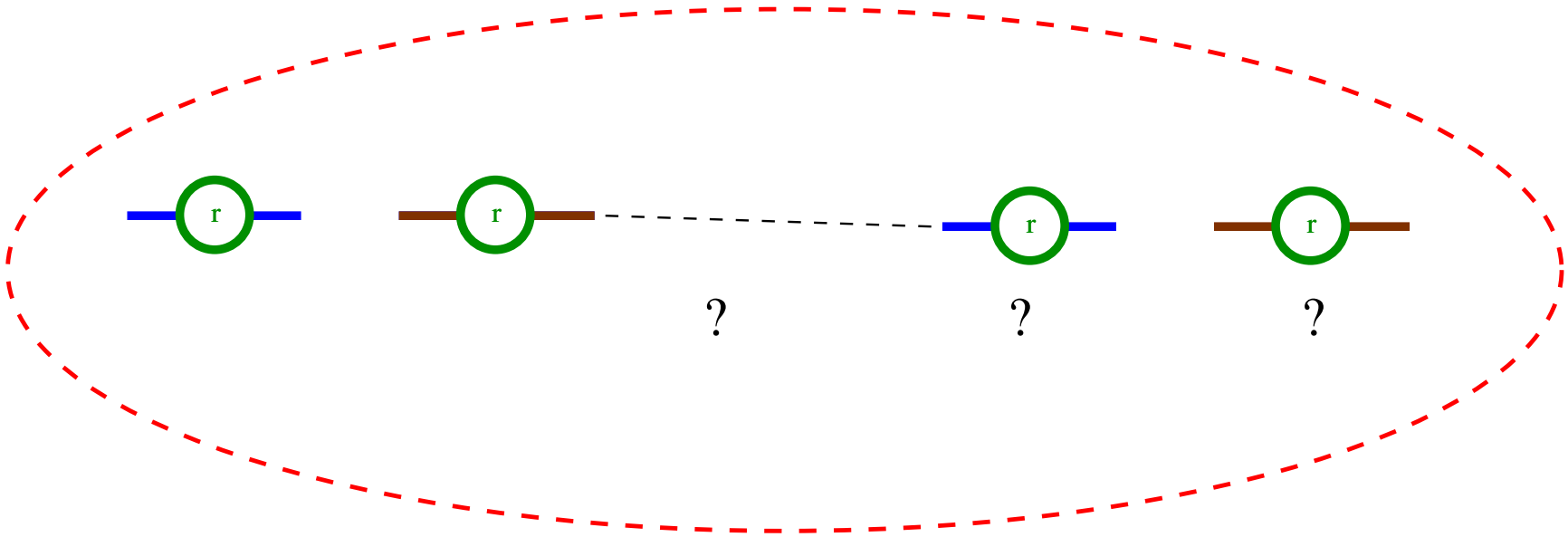
$\mathcal{C} \in Species^*_\omega$



Third case (II/III)



$\mathbb{C} \in Species_{\omega}^{*}$



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

$$\mathit{Species}_0 \triangleq \text{A(a}\sim\text{u),B(a}\sim\text{u)}$$

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

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

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

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

Non local systems

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

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

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

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. Language: Kappa
3. Abstraction: Local views
4. Completeness: false positives?
5. Local fragment of Kappa
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:

R2(l!2,r),R1(l!1,r),E2(r!1),E1(r!2) $\rightarrow$ R2(l!3,r!1),R1(l!2,r!1),E2(r!2),E1(r!3)

Decontextualized rule:

R2(l!_,r),R1(l!_,r) $\rightarrow$ R2(l!_,r!1),R1(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 $\mathcal{S}(\mathrm{R})$ 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:

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

- For each class $[\mathrm{R}]$, $\mathit{Reach}([\mathrm{R}])$ 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:

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

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 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 C of a rule, we try to apply it for another context C' by:

1. removing the context C' (backtrack) ;
2. building the context C ;
3. applying the initial rule ;
4. removing the context C (backtrack) ;
5. building the context C'.

This is proved manually.

Overview

1. Introduction
2. Language: Kappa
3. Abstraction: Local views
4. Completeness: false positives?
5. Local fragment of Kappa
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)

MPRI

Some notions of information flow

Jérôme Feret
Laboratoire d'Informatique de l'École Normale Supérieure
INRIA, ÉNS, CNRS
http://www.di.ens.fr/~feret

Friday, the 7th of February, 2014

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 $[\![P]\!] \in \mathcal{F}((\mathcal{V} \to \mathbb{Z}) \cup \Omega)$ 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 pf 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 $V_1 \not\rightarrow^*_{\mathrm{P}} V_2$.

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 atacker could observe:

- computation time;
- memory assumption;
- heating.

(attacks cannot be exhaustively specified).

Cours MPRI

Formal model reduction

Jérôme Feret
Laboratoire d'Informatique de l'École Normale Supérieure
INRIA, ÉNS, CNRS

Friday, the 7th of February, 2014

Joint-work with...

Walter Fontana
Harvard Medical School

Vincent Danos
Edinburgh

Ferdinanda Camporesi
Bologna / ÉNS

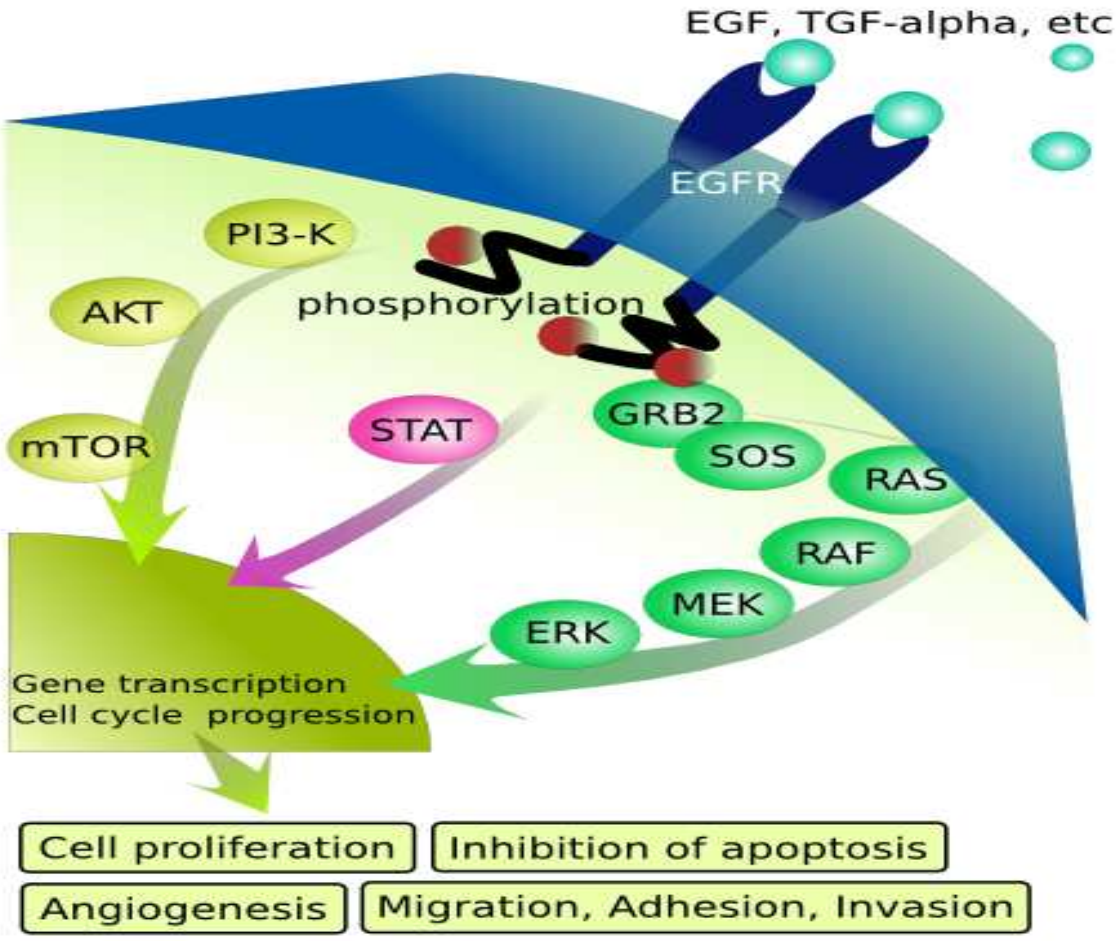
Russ Harmer
Paris VII

Jean Krivine
Paris VII

Overview

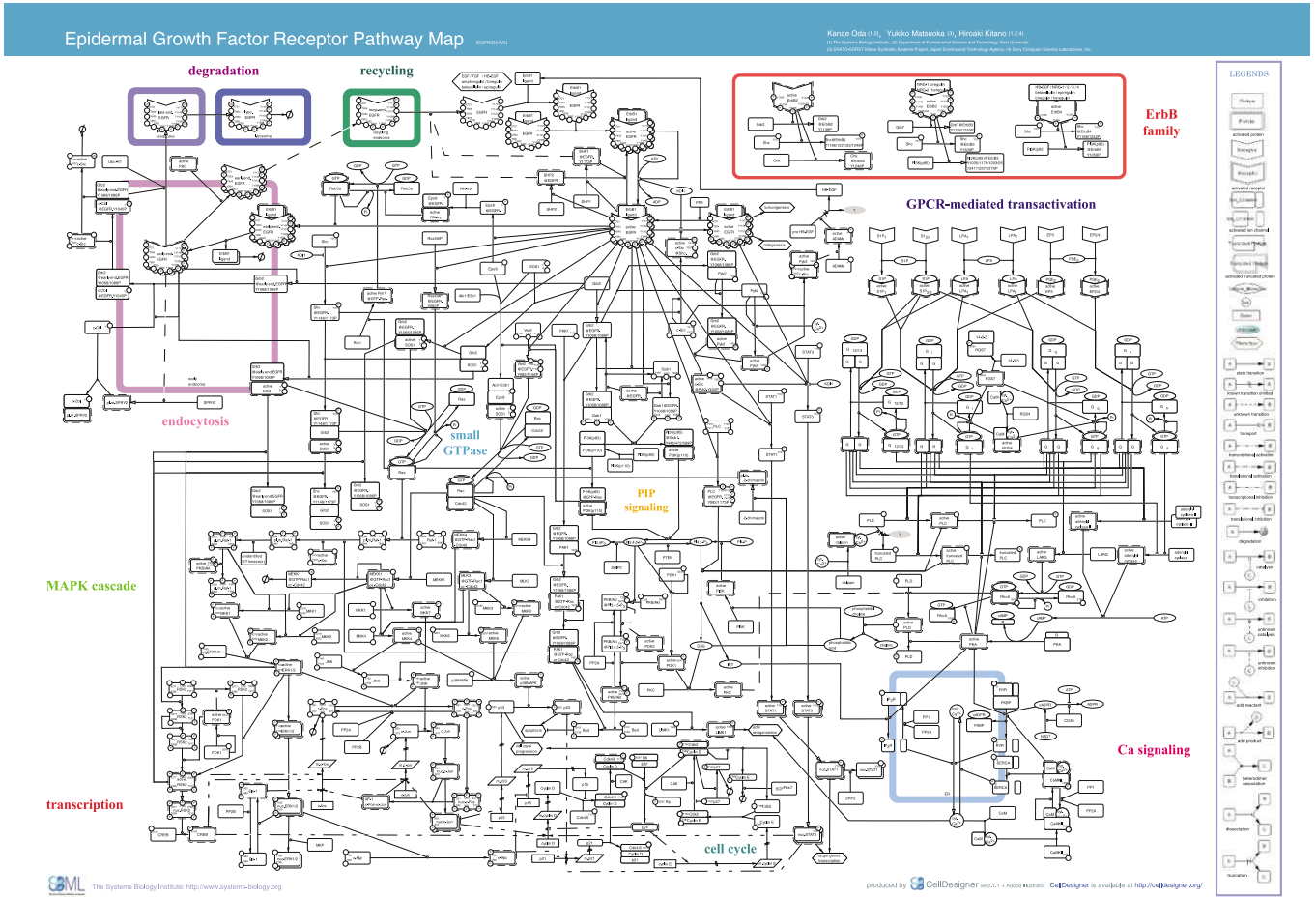
1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

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.

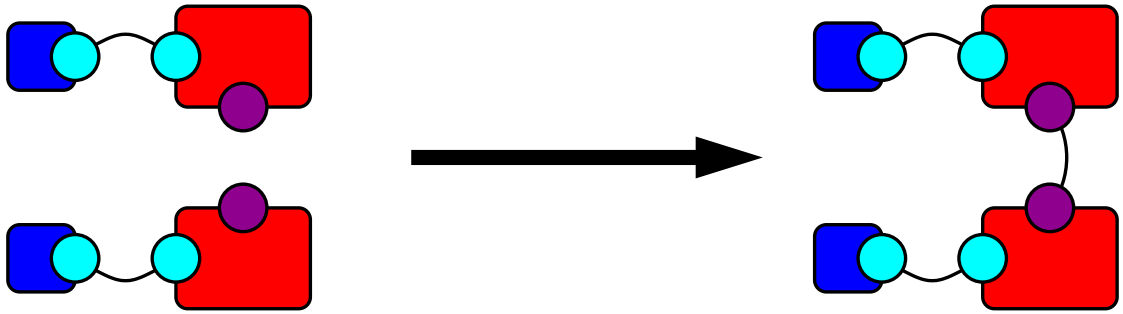
A gap between two worlds

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



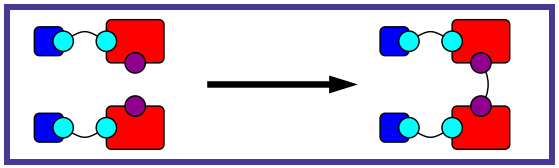
1. The description level matches with both
   - the observation level
   - and the intervention level

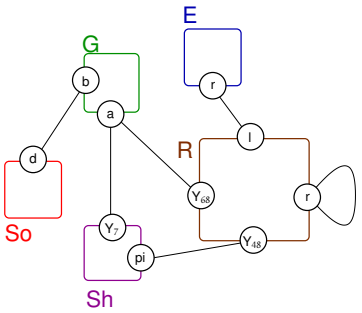
   of the biologist.

   We can tune the model easily.
2. Model description is very compact.

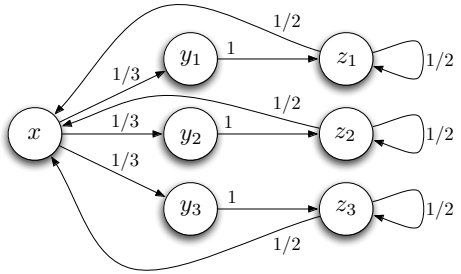
Semantics

Several semantics (qualititative and/or quantitative) can be defined.





Interaction map

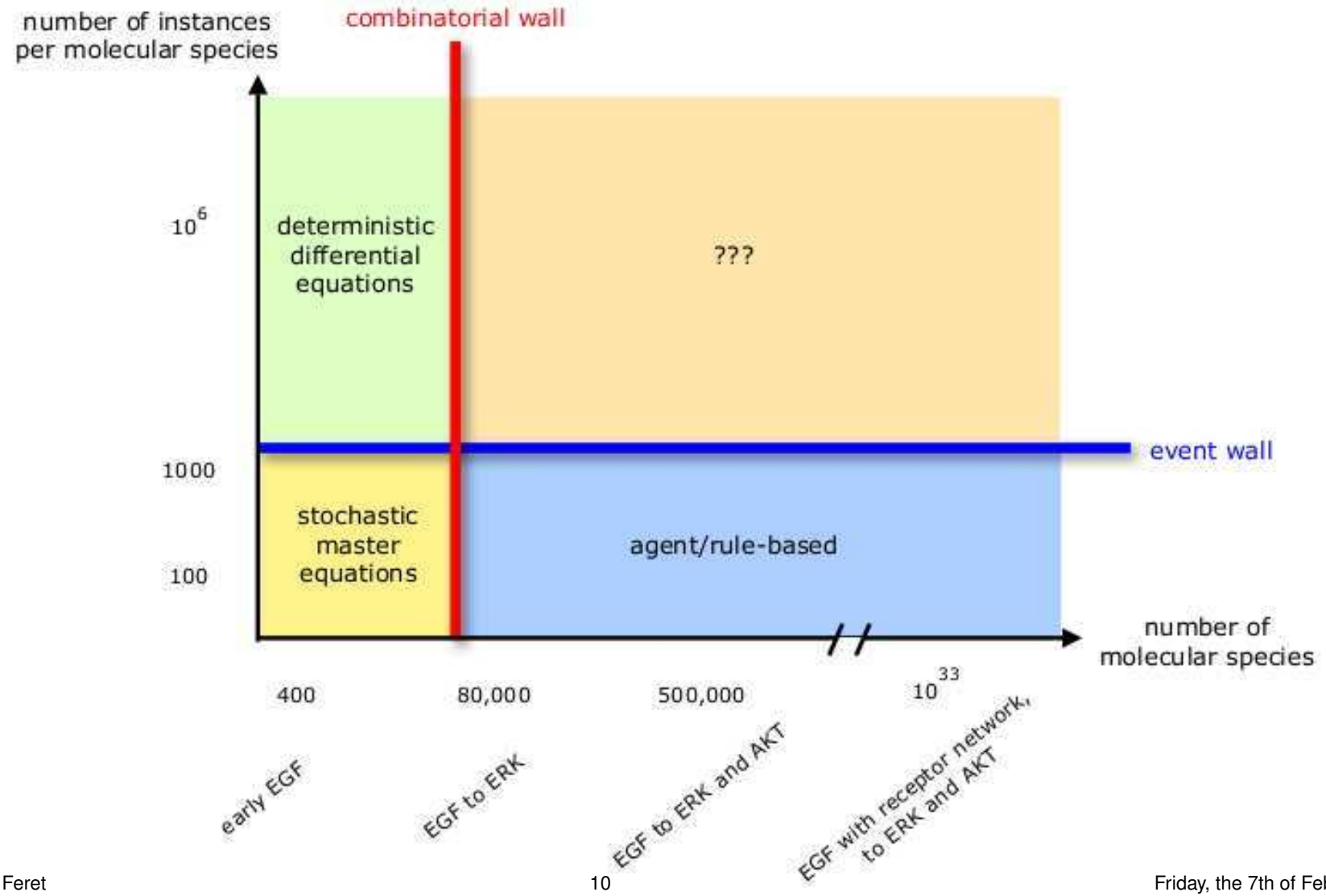


CTMC

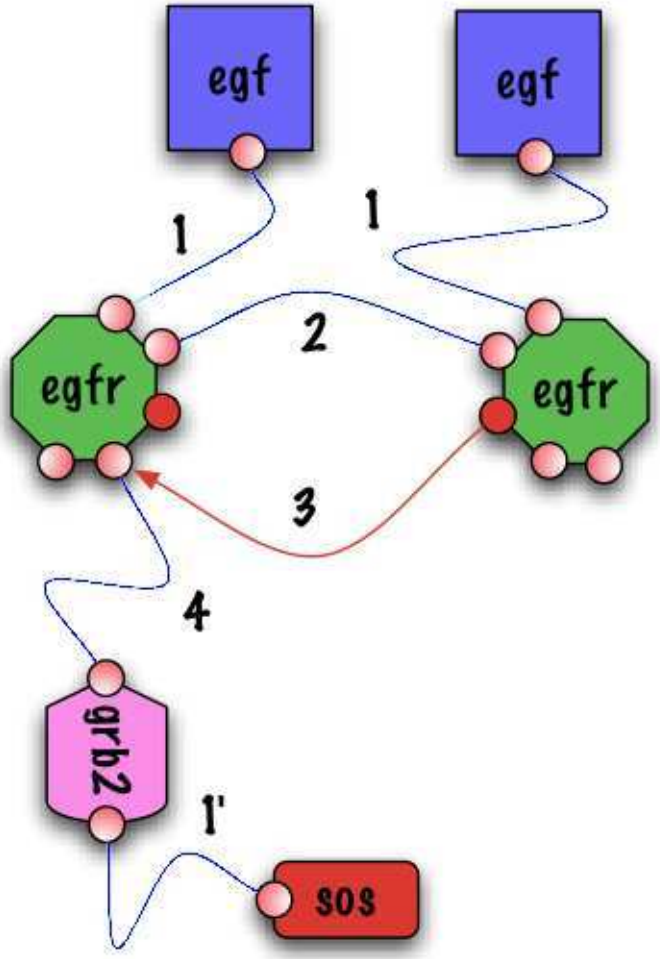
$$\begin{cases} \frac{dx_1}{dt} = -k_1 \cdot x_1 \cdot x_2 + k_{-1} \cdot x_3 \\ \frac{dx_2}{dt} = -k_1 \cdot x_1 \cdot x_2 + k_{-1} \cdot x_3 \\ \frac{dx_3}{dt} = k_1 \cdot x_1 \cdot x_2 - k_{-1} \cdot x_3 + 2 \cdot k_2 \cdot x_3 \cdot x_3 - k_{-2} \cdot x_4 \\ \frac{dx_4}{dt} = k_2 \cdot x_3^2 - k_2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

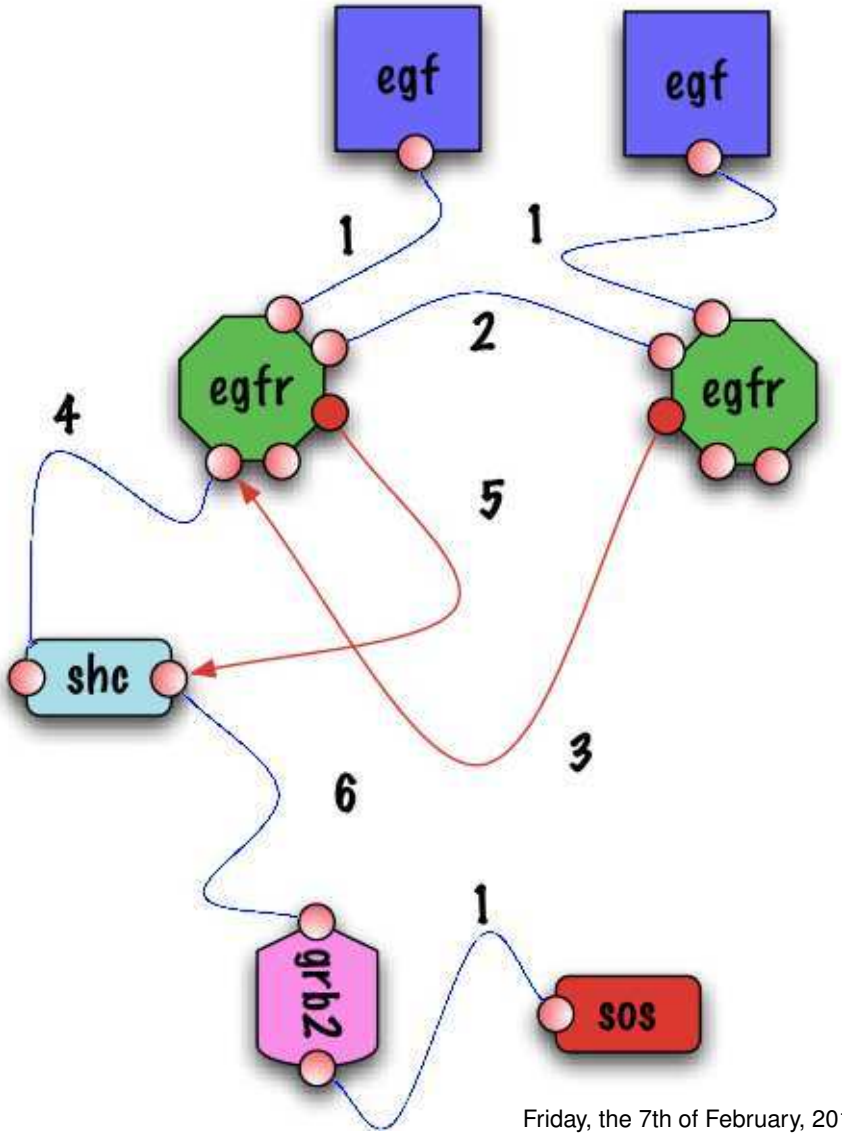
ODEs

Complexity walls



A breach in the wall(s) ?

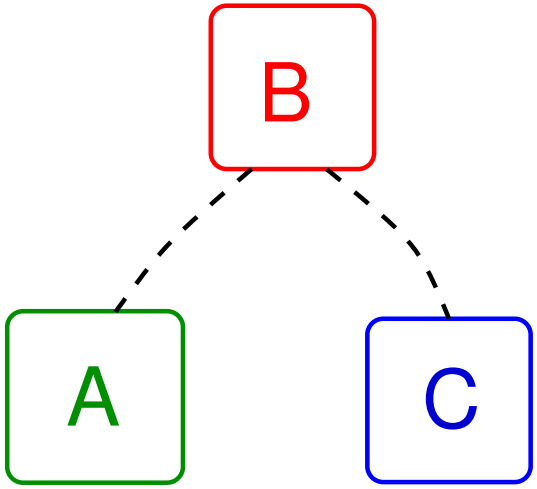




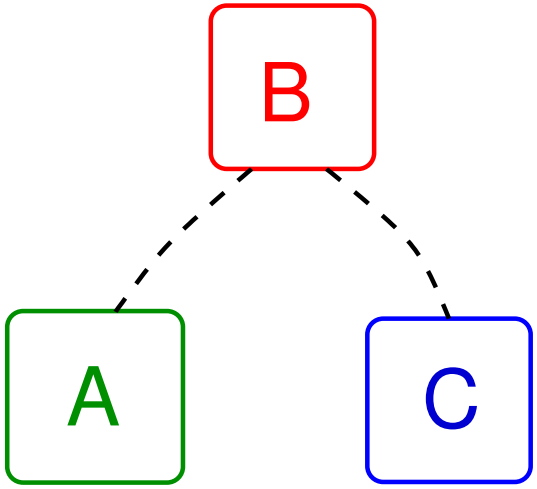
Overview

1. Context and motivations
2. Handmade ODEs
   (a) a simple adapter
   (b) a system with a switch
   (c) a system with symmetries
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

A simple adapter

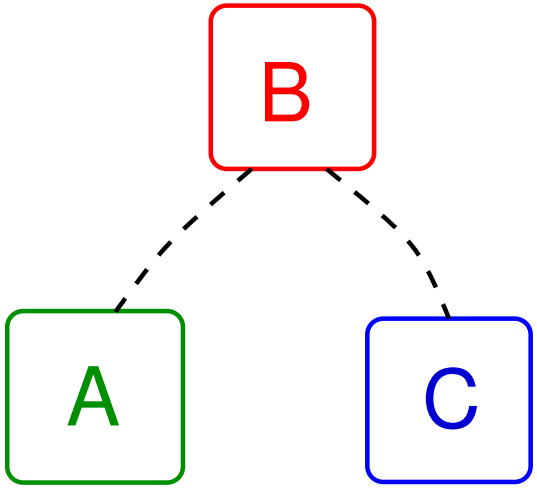


A simple adapter



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

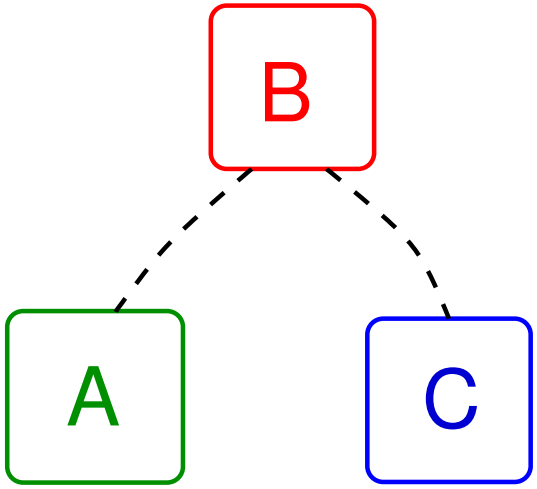
A simple adapter



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

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

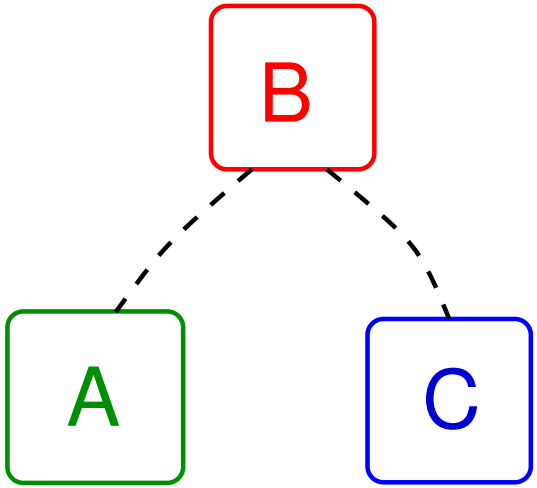
A simple adapter



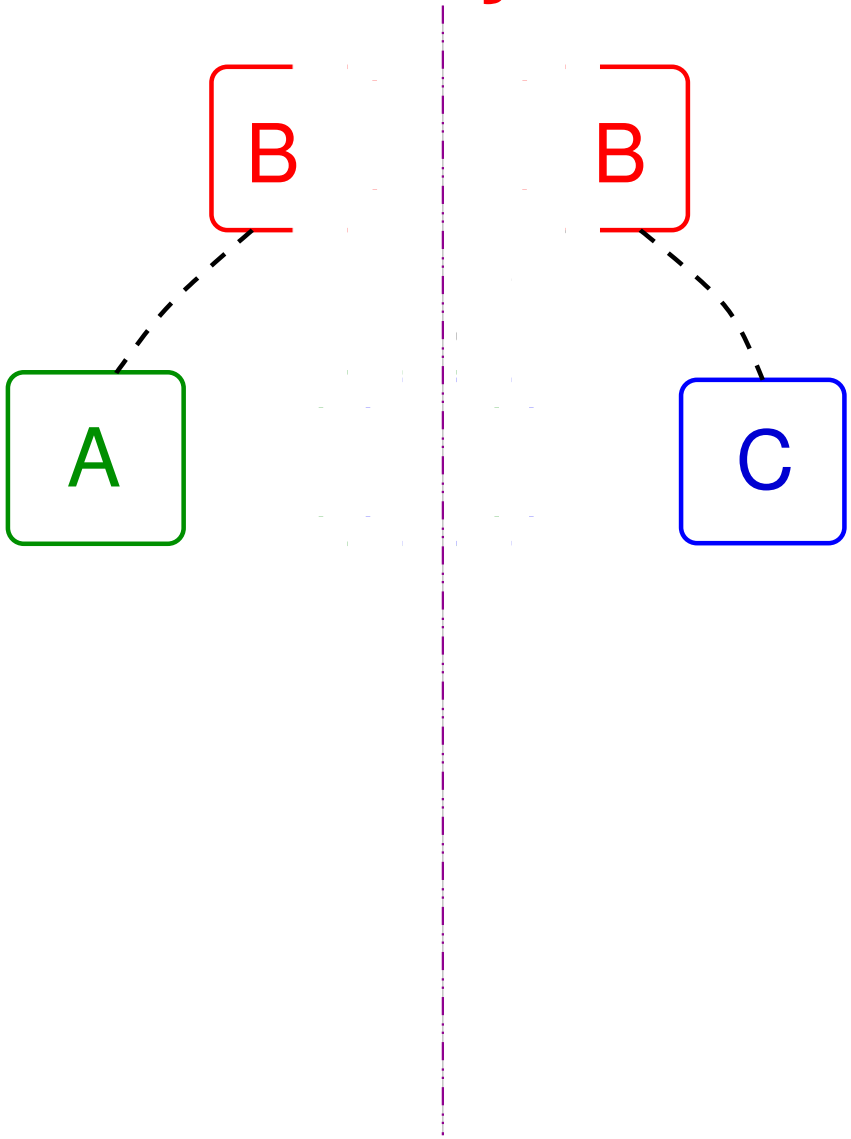
$$
\begin{array}{lll}
A\,,\,\emptyset B\emptyset & \longleftrightarrow & AB\emptyset \qquad k^{AB},k_d^{AB} \\
A\,,\,\emptyset BC & \longleftrightarrow & ABC \qquad k^{AB},k_d^{AB} \\
\emptyset B\emptyset\,,\,C & \longleftrightarrow & \emptyset BC \qquad k^{BC},k_d^{BC} \\
AB\emptyset\,,\,C & \longleftrightarrow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

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

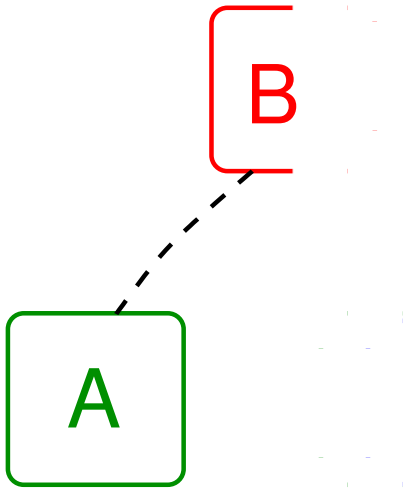
Two subsystems



Two subsystems



Two subsystems

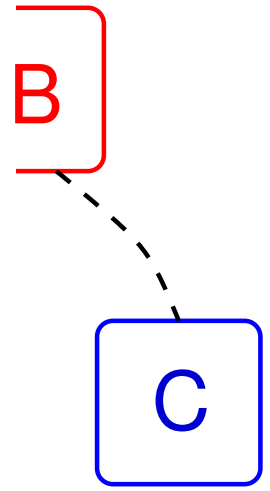


$$[A] = [A]$$

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

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

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



$$[C] = [C]$$

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

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

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

Dependence index

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

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

Thus we define the dependence index as follows:

$$X \stackrel{\Delta}{=} [\mathrm{ABC}]\cdot([\emptyset\mathrm{B}?] + [\mathrm{AB}?]) - [\mathrm{AB}?]\cdot[?\mathrm{BC}].$$

We have (after a short computation):

$$\frac{\mathrm{d}X}{\mathrm{d}t} = -X\cdot\left([\mathrm{A}]\cdot k^{\mathrm{AB}} + k_d^{\mathrm{AB}} + [\mathrm{C}]\cdot k^{\mathrm{BC}} + k_d^{\mathrm{BC}}\right).$$

So the property:

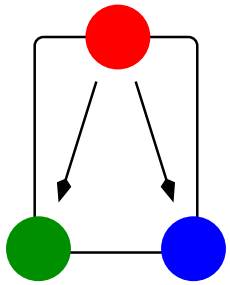
$$\frac{[\mathrm{ABC}]}{[?\mathrm{BC}]} = \frac{[\mathrm{AB}?]}{[\emptyset\mathrm{B}?] + [\mathrm{AB}?]}.$$

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

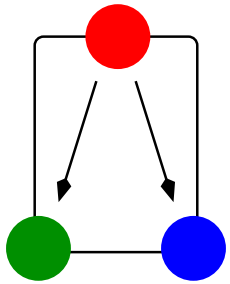
Overview

1. Context and motivations
2. Handmade ODEs
   (a) a simple adapter
   (b) a system with a switch
   (c) a system with symmetries
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

A system with a switch

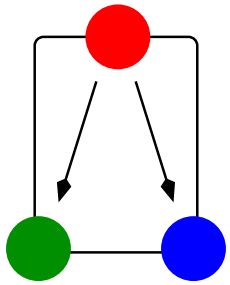


A system with a switch



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

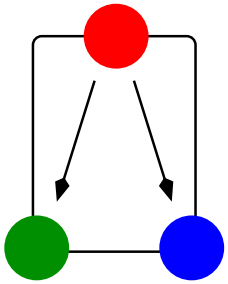
A system with a switch



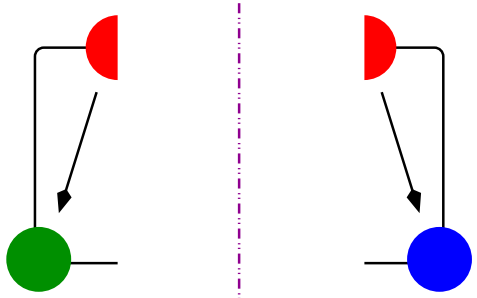
$$
\begin{aligned}
(u,u,u) &\longrightarrow (u,p,u) \quad k^c \\
(u,p,u) &\longrightarrow (p,p,u) \quad k^l \\
(u,p,p) &\longrightarrow (p,p,p) \quad k^l \\
(u,p,u) &\longrightarrow (u,p,p) \quad k^r \\
(p,p,u) &\longrightarrow (p,p,p) \quad k^r
\end{aligned}
$$

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

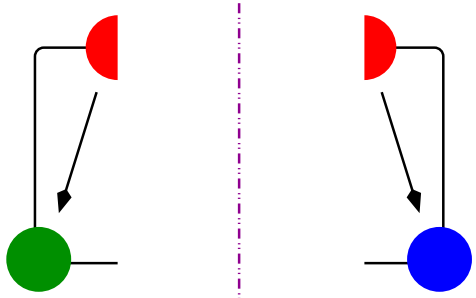
Two subsystems



Two subsystems



Two subsystems



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

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

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

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

Dependence index

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

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

Thus we define the dependence index as follows:

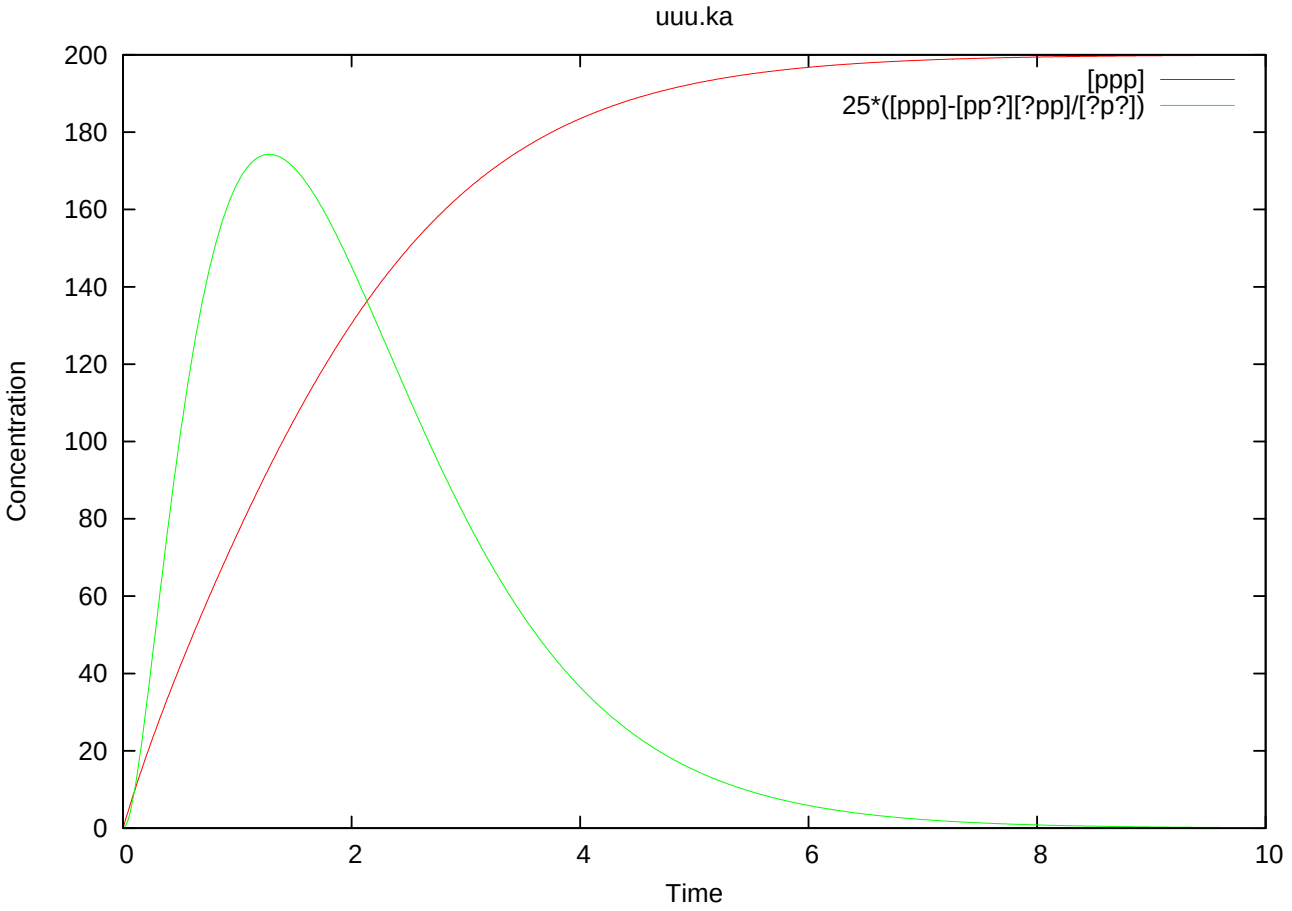
$$X \stackrel{\Delta}{=} [(p,p,p)]\cdot([(?,p,u)]+[(?,p,p)])-[(?,p,p)]\cdot[(p,p,?)].$$

We have:

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

So the property $(X=0)$ is not an invariant.

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

We can use the absence of flow of information to cut chemical species into self-consistent fragments of chemical species:

- some information is abstracted away:
  we cannot recover the concentration of any species;

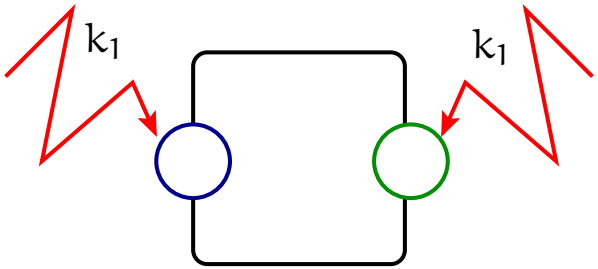
+ flow of information is easy to abstract;

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

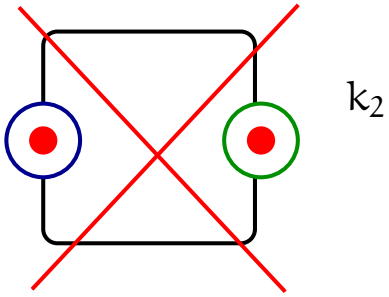
Overview

1. Context and motivations
2. Handmade ODEs
   (a) a simple adapter
   (b) a system with a switch
   (c) a system with symmetries
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

A model with symmetries

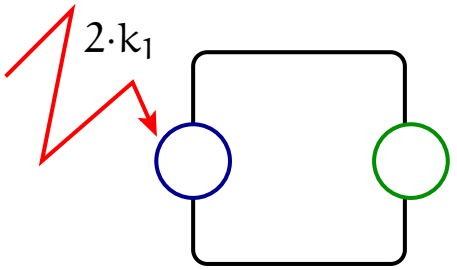


$$P \longrightarrow {}^{\star}P \quad k_1 \qquad P^{\star} \longrightarrow {}^{\star}P^{\star} \quad k_1$$
$$P \longrightarrow P^{\star} \quad k_1 \qquad {}^{\star}P \longrightarrow {}^{\star}P^{\star} \quad k_1$$

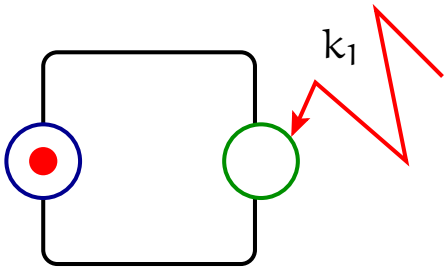


$${}^{\star}P^{\star} \longrightarrow \emptyset \quad k_2$$

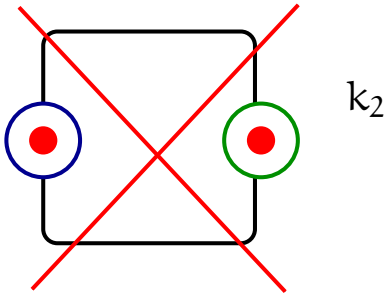
Reduced model



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



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



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

Differential equations

- Initial system:

$$\frac{d}{dt}\begin{bmatrix} P \\ {}^{\star}P \\ P^{\star} \\ {}^{\star}P^{\star} \end{bmatrix} = \begin{bmatrix} -2\cdot k_1 & 0 & 0 & 0 \\ k_1 & -k_1 & 0 & 0 \\ k_1 & 0 & -k_1 & 0 \\ 0 & k_1 & k_1 & -k_2 \end{bmatrix} \cdot \begin{bmatrix} P \\ {}^{\star}P \\ P^{\star} \\ {}^{\star}P^{\star} \end{bmatrix}$$

- Reduced system:

$$\frac{d}{dt}\begin{bmatrix} P \\ {}^{\star}P + P^{\star} \\ 0 \\ {}^{\star}P^{\star} \end{bmatrix} = \begin{bmatrix} -2\cdot k_1 & 0 & 0 & 0 \\ 2\cdot k_1 & -k_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & k_1 & 0 & -k_2 \end{bmatrix} \cdot \begin{bmatrix} P \\ {}^{\star}P + P^{\star} \\ 0 \\ {}^{\star}P^{\star} \end{bmatrix}$$

Invariant

We wonder whether or not:

$$[{}^{\star}P] = [P^{\star}],$$

Thus we define the difference X as follows:

$$X \stackrel{\Delta}{=} [{}^{\star}P] - [P^{\star}].$$

We have:

$$\frac{dX}{dt} = -k_1 \cdot X.$$

So the property $(X = 0)$ is an invariant.

Thus, if $[{}^{\star}P] = [P^{\star}]$ at time $t = 0$, then $[{}^{\star}P] = [P^{\star}]$ forever.

Conclusion

We can abstract away the distinction between chemical species which are equivalent up to symmetries (with respect to the reactions).

1. If the symmetries are satisfied in the initial state:
   + the abstraction is invertible:
     we can recover the concentration of any species,
     (thanks to the invariants).
2. Otherwise:
   − some information is abstracted away:
     we cannot recover the concentration of any species;
   + the system converges to a state which satisfies the symmetries.

Overview

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

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 differential 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
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,
- 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

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

(Completeness can be checked analytically.)

Abstract differential semantics

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:

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

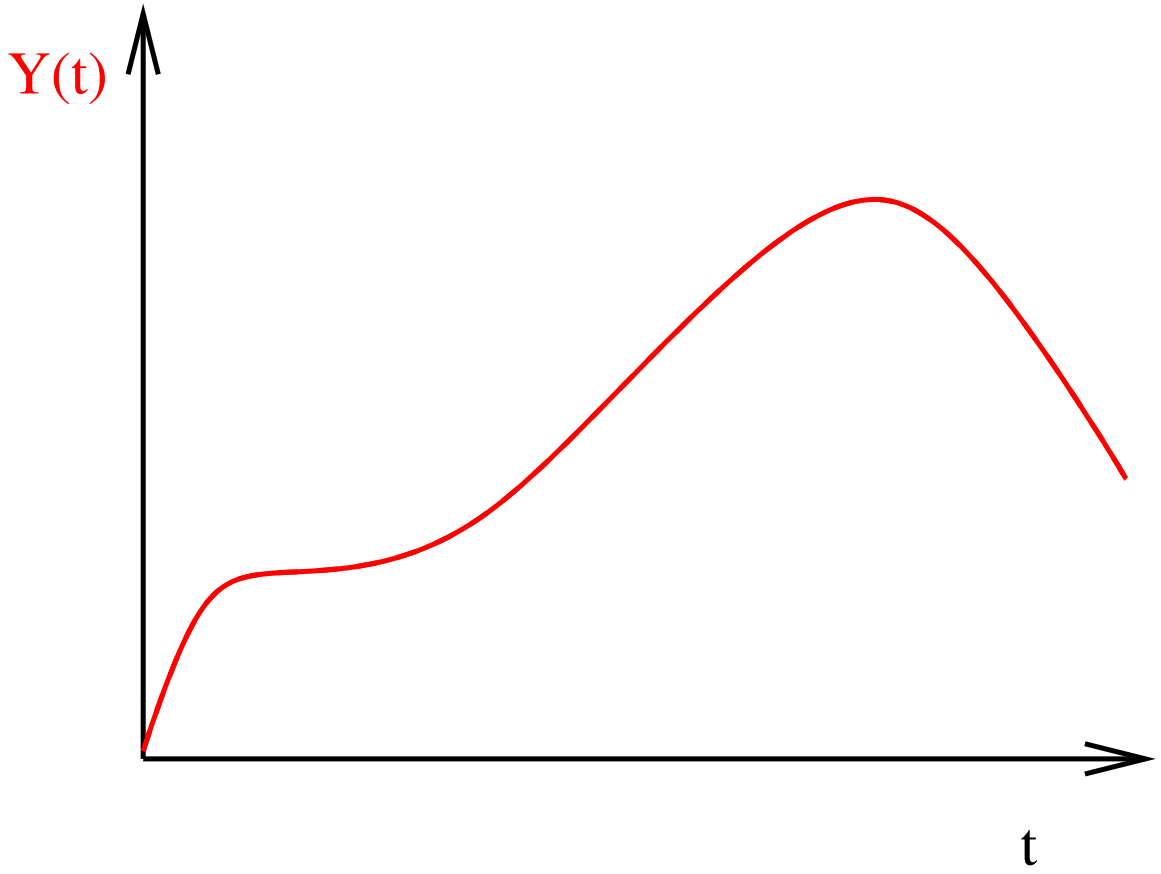
has a unique maximal solution $Y_{\psi(X_0)}$ such that $Y_{\psi(X_0)} = \psi(X_0)$.

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

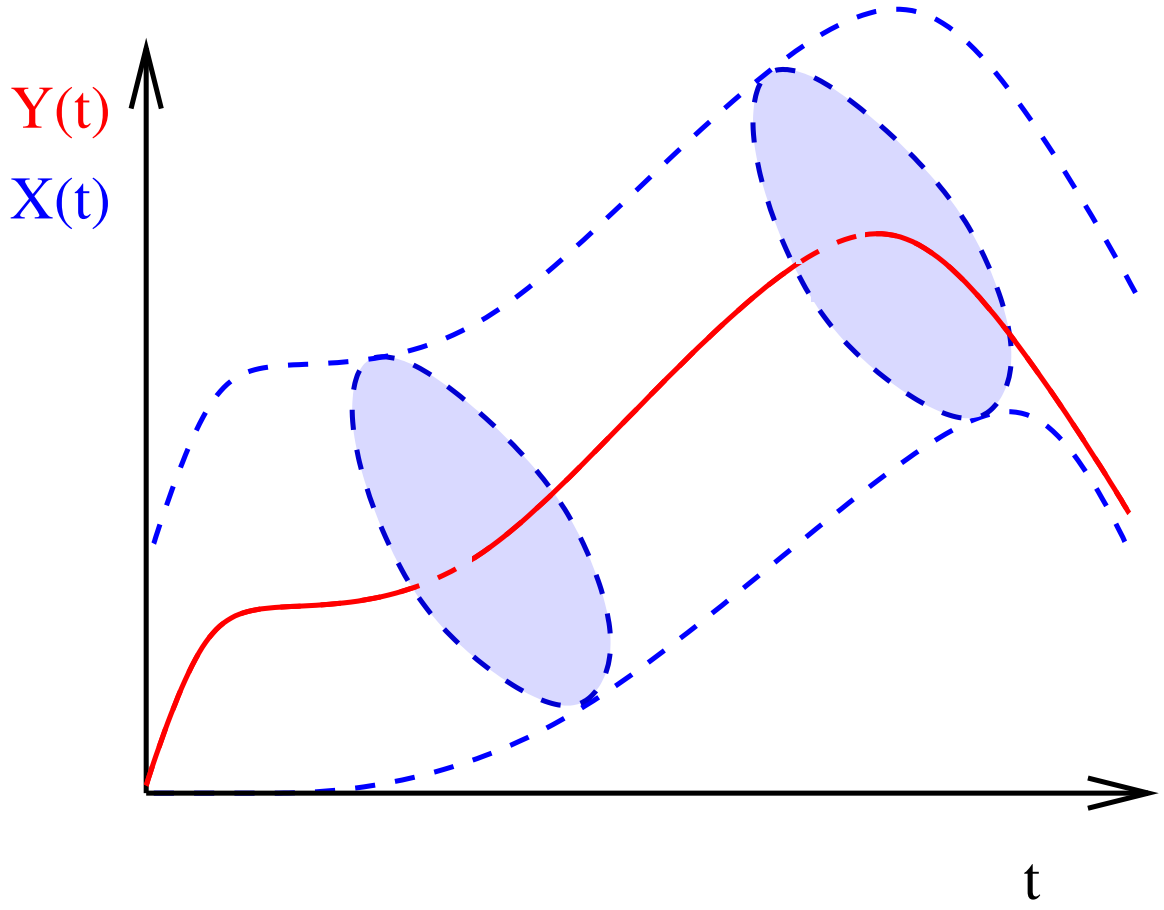
$$X_{X_0}(T) = X_0 + \int_{t=0}^{T} \mathbb{F}\left(X_{X_0}(t)\right) \cdot dt.$$

(i.e. $Y_{\psi(X_0)} = \psi(X_{X_0})$)

Fluid trajectories



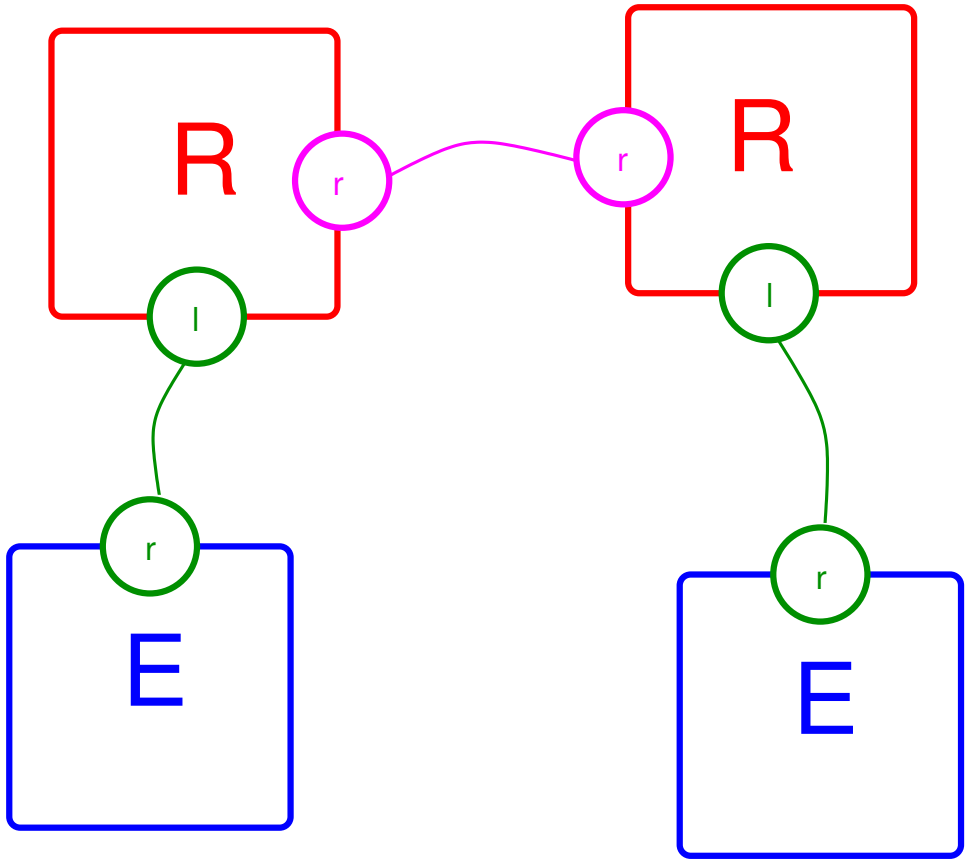
Fluid trajectories



Overview

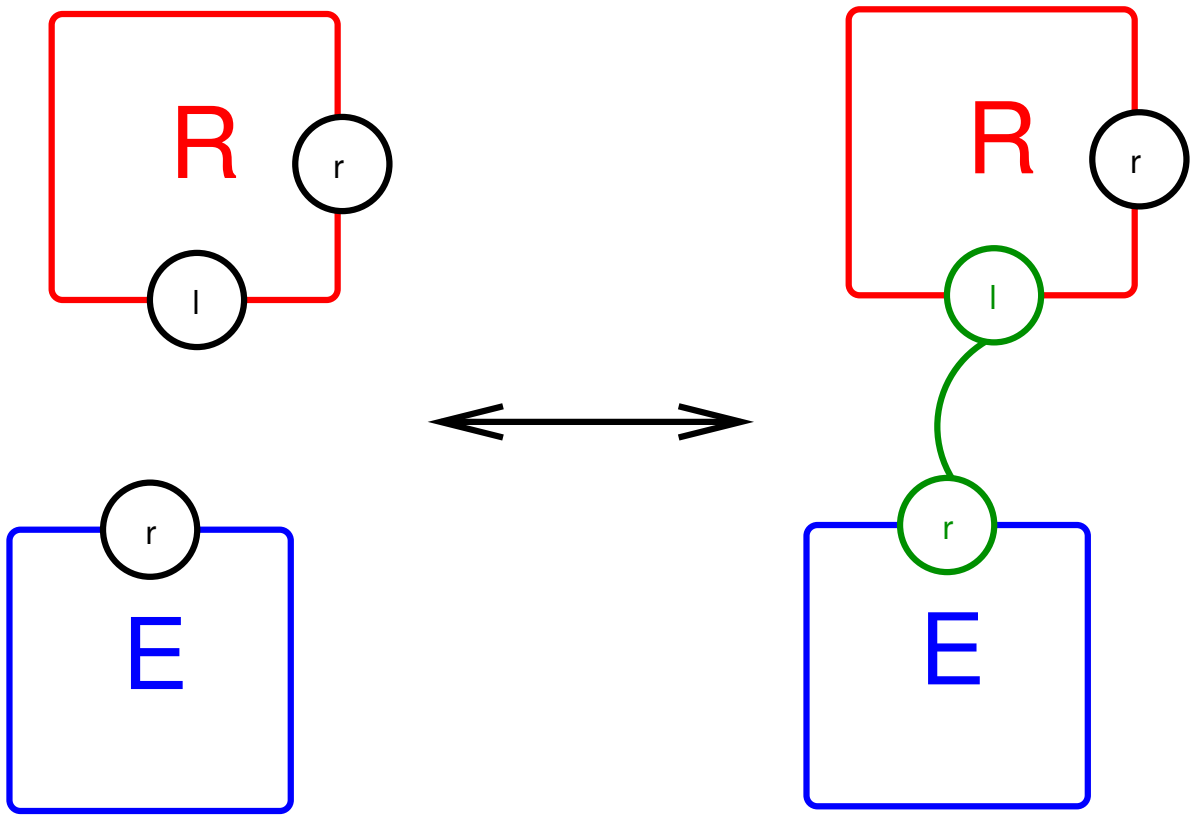
1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

A species



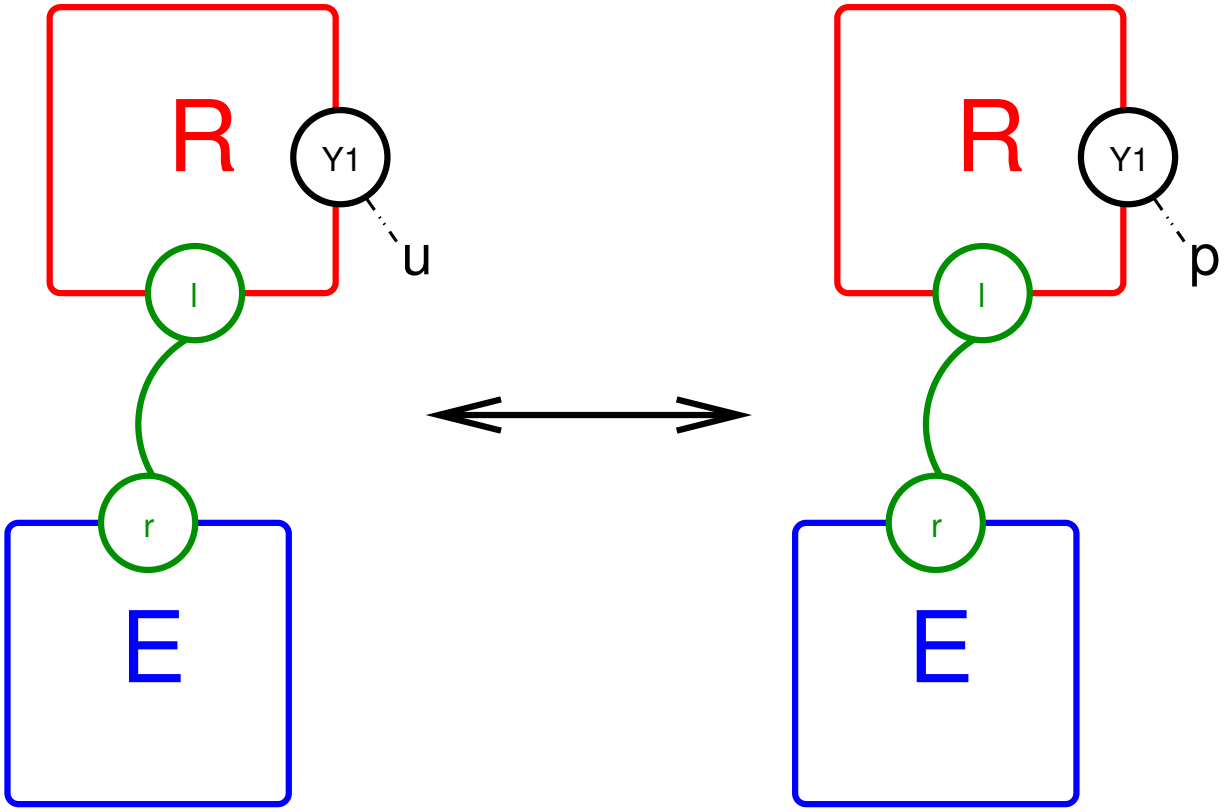
E(r!1), R(l!1,r!2), R(r!2,l!3), E(r!3)

A Unbinding/Binding Rule



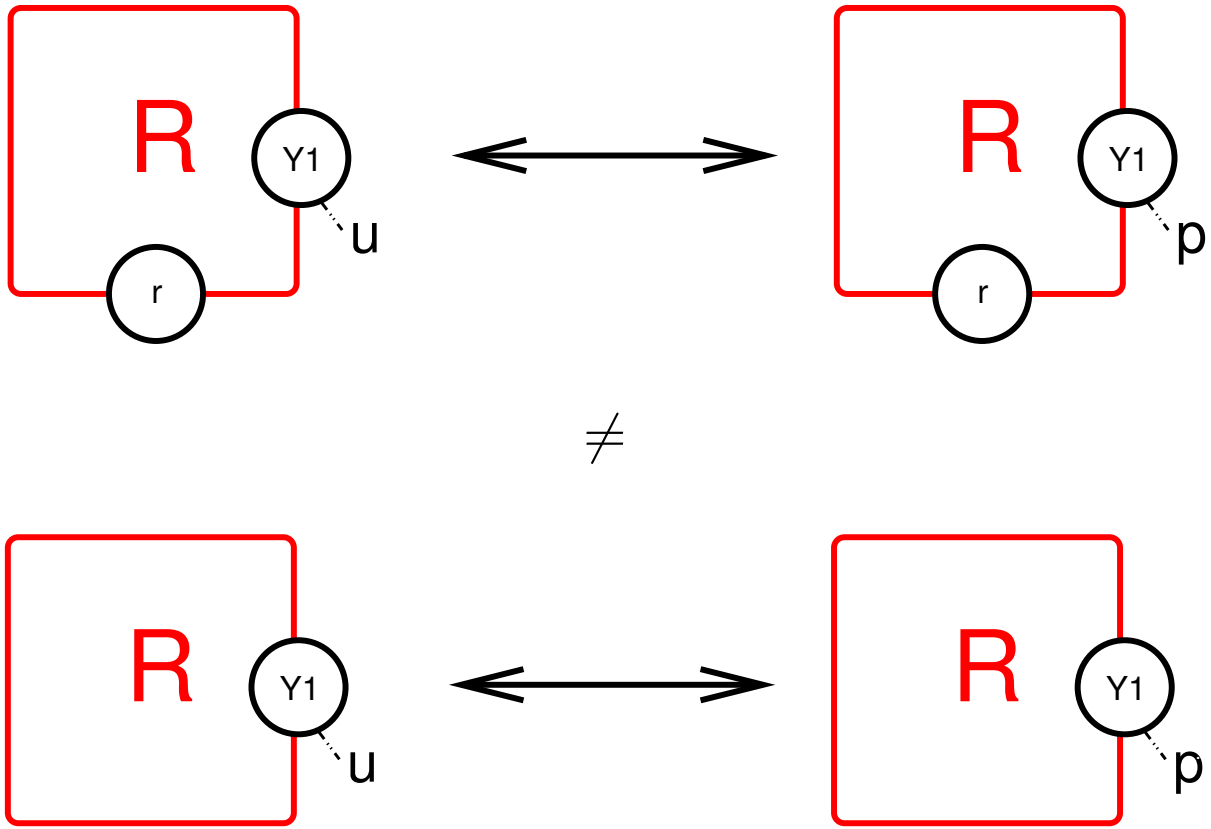
E(r), R(l,r) ⟷ E(r!1), R(l!1,r)

Internal state

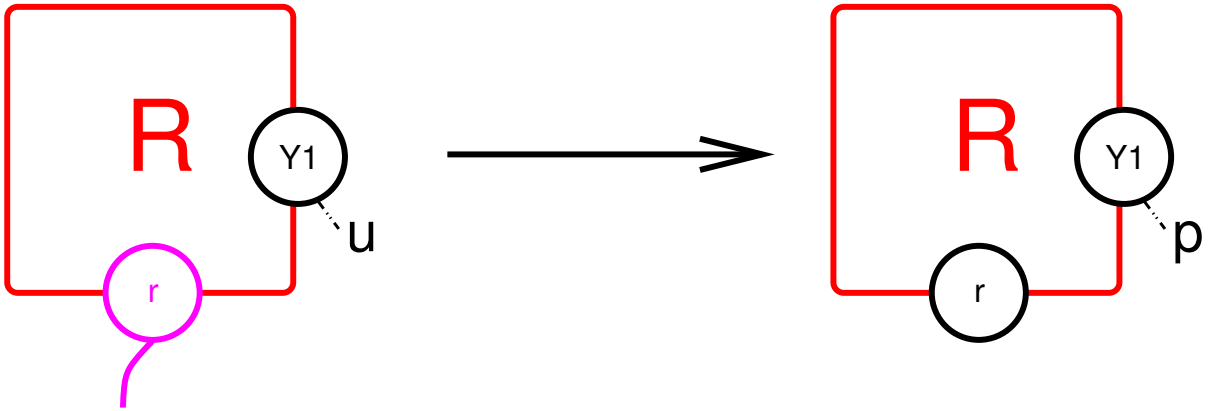


R(Y1~u,l!1), E(r!1) ⟷ R(Y1~p,l!1), E(r!1)

Don’t care, Don’t write

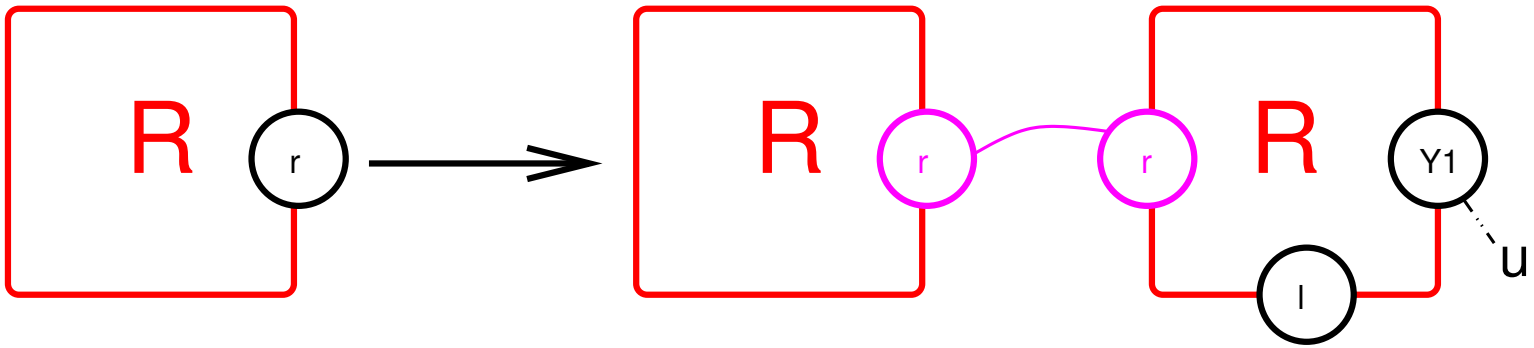


A contextual rule

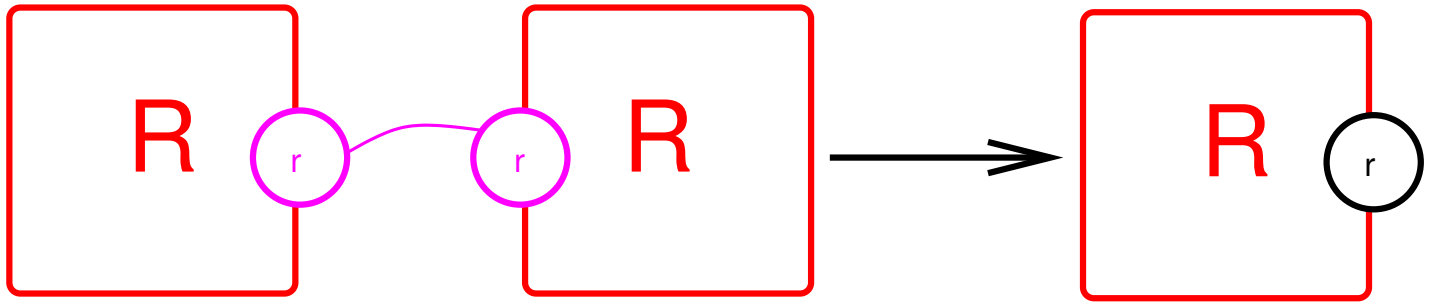


$$R(Y1{\sim}u,r!\_) \rightarrow R(Y1{\sim}p,r)$$

Creation/Suppression



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

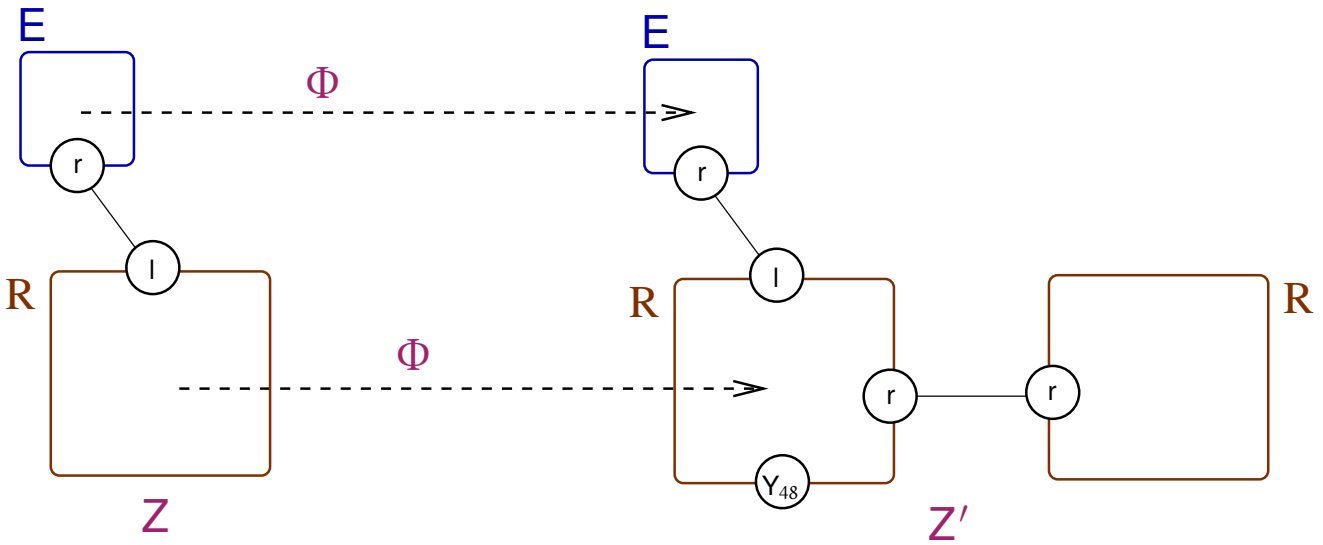


$R(r!1), R(r!1) \rightarrow R(r)$

Overview

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

Embedding



We say that Φ is an embedding between Z and Z' iff:

- Φ is a site-graph morphism:
  - i is less specific than $\Phi(i)$,
  - there is a link between (i, s) and (i', s'),
    if and only if 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

Each rule *rule*: *lhs* $\rightarrow$ *rhs* is associated with a rate constant k.

Such a rule is seen as a generic representation of a set of chemical reactions:

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

For each such reaction, we get the following contribution:

$$\frac{d[r_i]}{dt} \stackrel{-}{=} \frac{k \cdot \prod [r_i]}{\text{SYM}(\textit{lhs})} \quad \text{and} \quad \frac{d[p_i]}{dt} \stackrel{+}{=} \frac{k \cdot \prod [r_i]}{\text{SYM}(\textit{lhs})}.$$

where $\text{SYM}(E)$ is the number of automorphisms in E.

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
7. Conclusion

Abstract domain

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

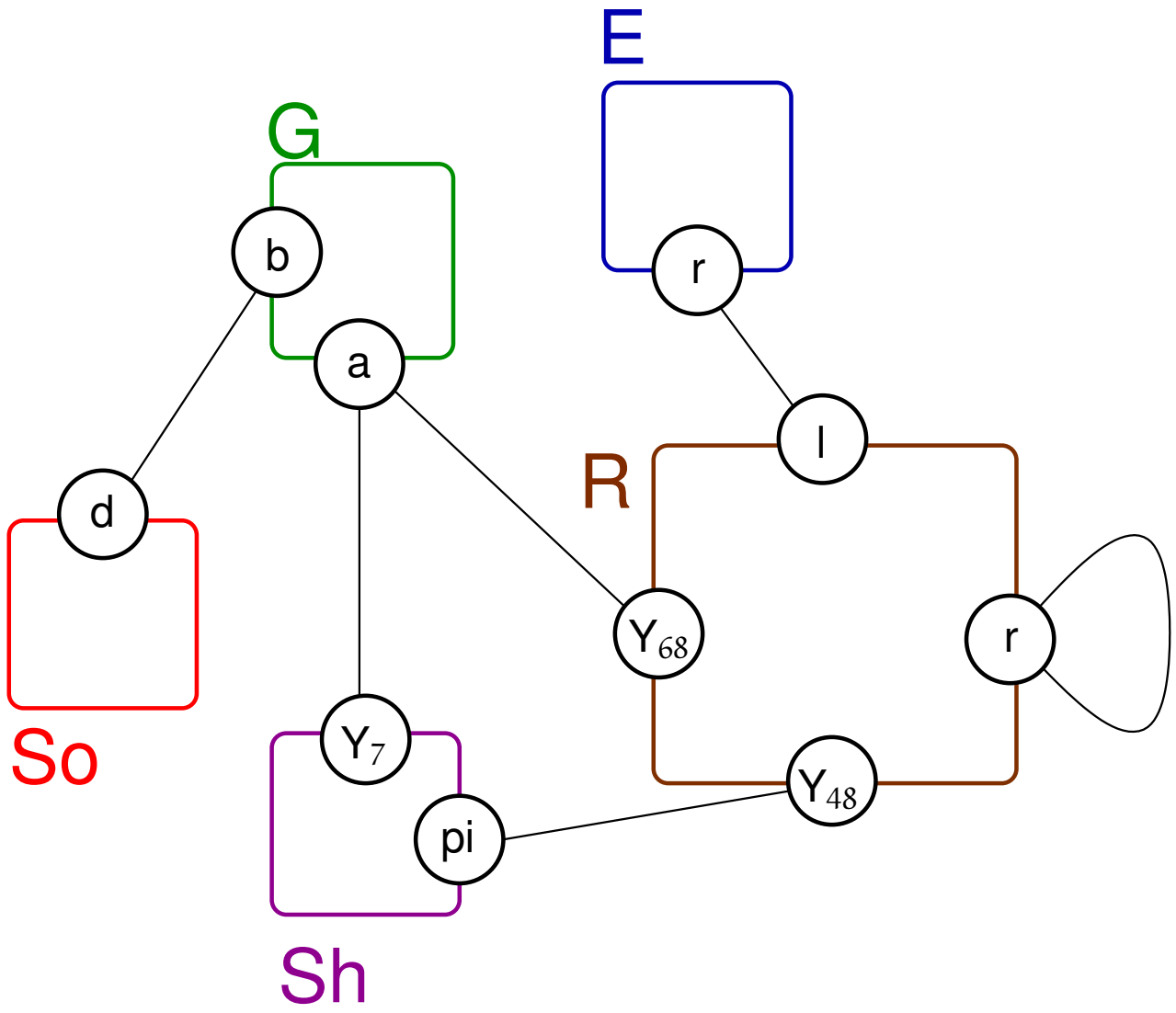
The set of linear variable replacements is too big to be explored.

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

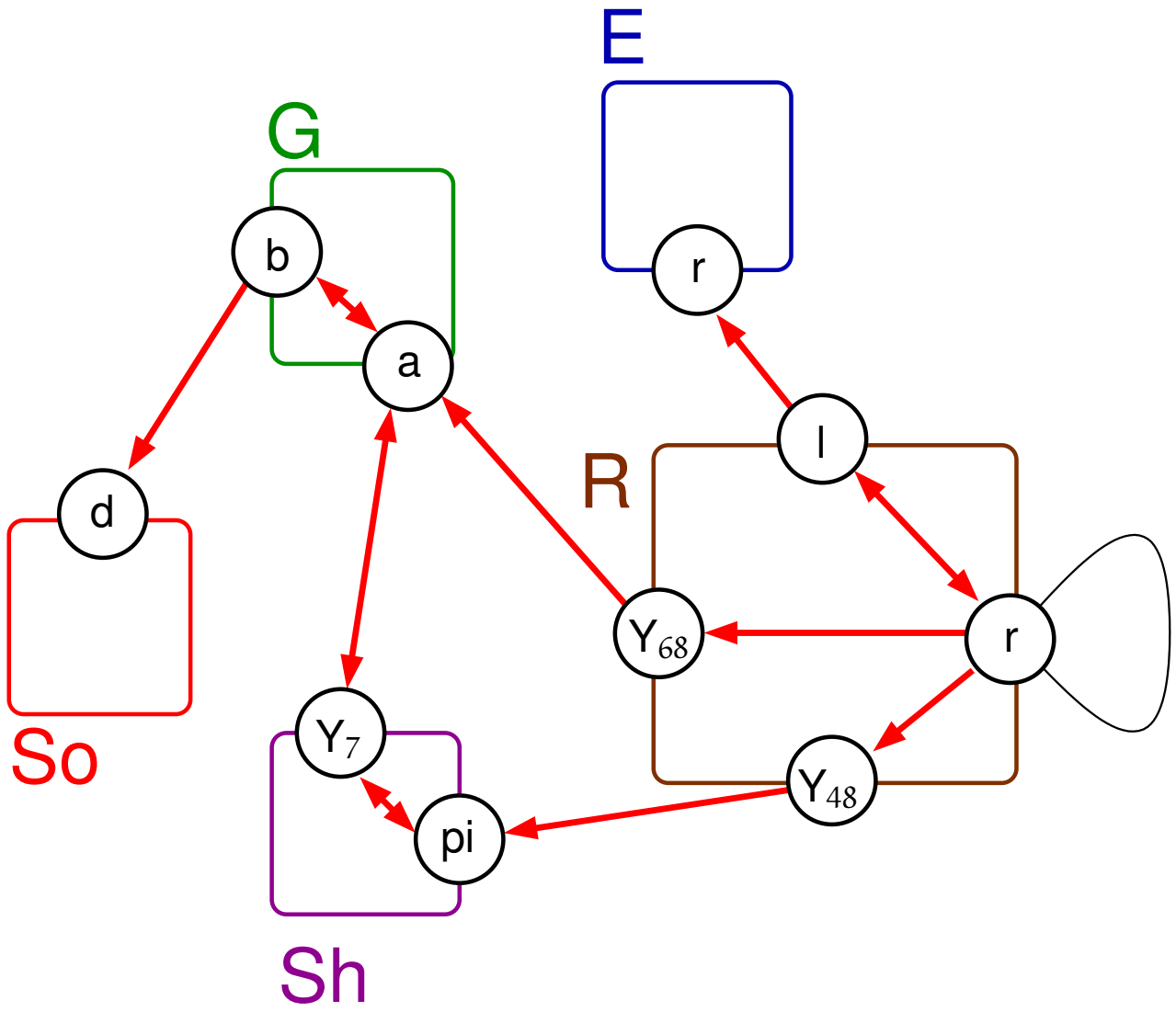
- restrict the exploration;
- drive the intuition (by using the flow of information);
- having efficient way to find suitable abstractions $(\mathcal{V}^\sharp, \psi)$ and to compute $\mathbb{F}^\sharp$.

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

Contact map



Annotated contact map



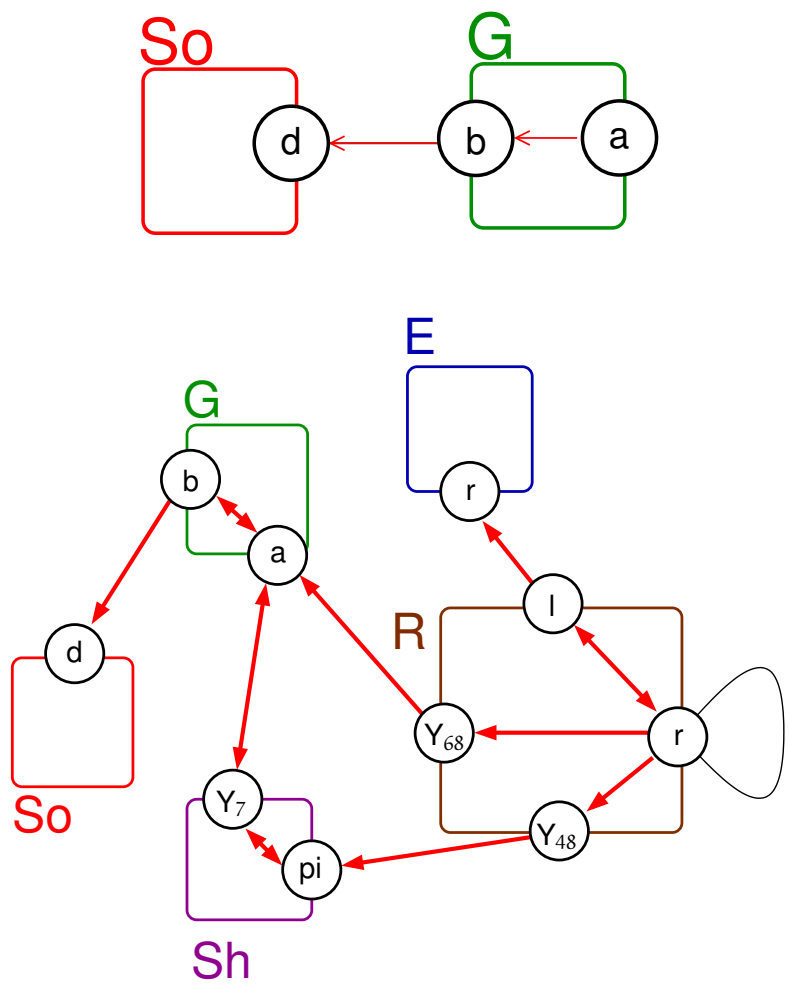
Fragments and prefragments

A prefragment is a connected site graph for which there exists a binary relations $\rightarrow$ between sites such that:

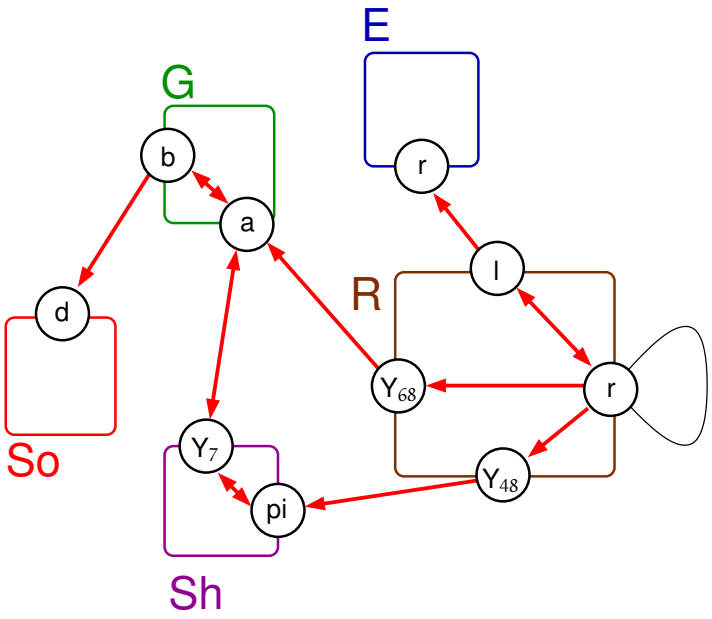
- Directed preorder: for any pair of sites x and y there is a site z such that: $x \rightarrow^\star z$ and $y \rightarrow^\star z$.
- Compatibility: any edge $\rightarrow$ can be projected to an edge in the annotated contact map.

A fragment is a prefragment F such that:

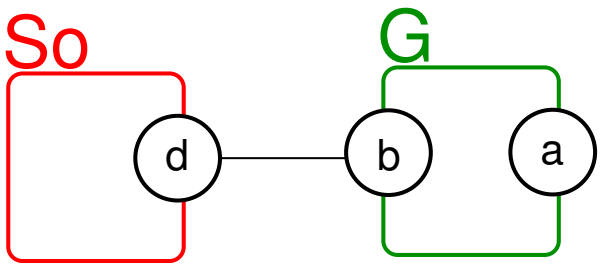
- Parsimoniousness: for any prefragment F' such that F embeds in F', F' also embeds into F.

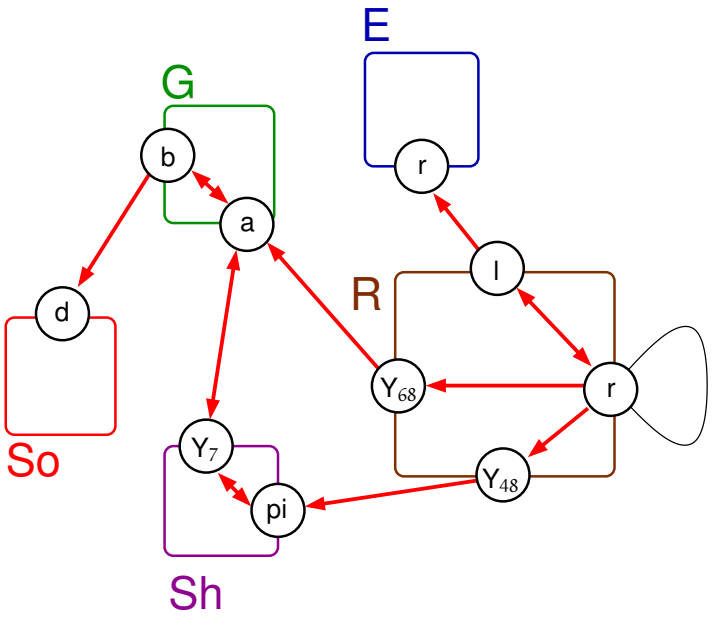


Are they fragments ?

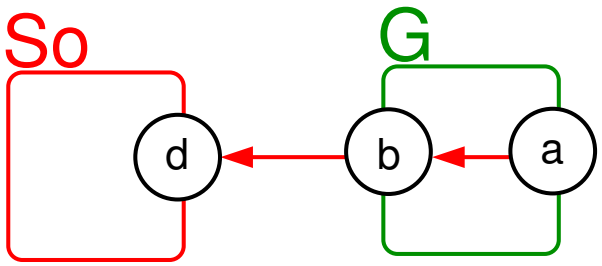


Are they fragments ?

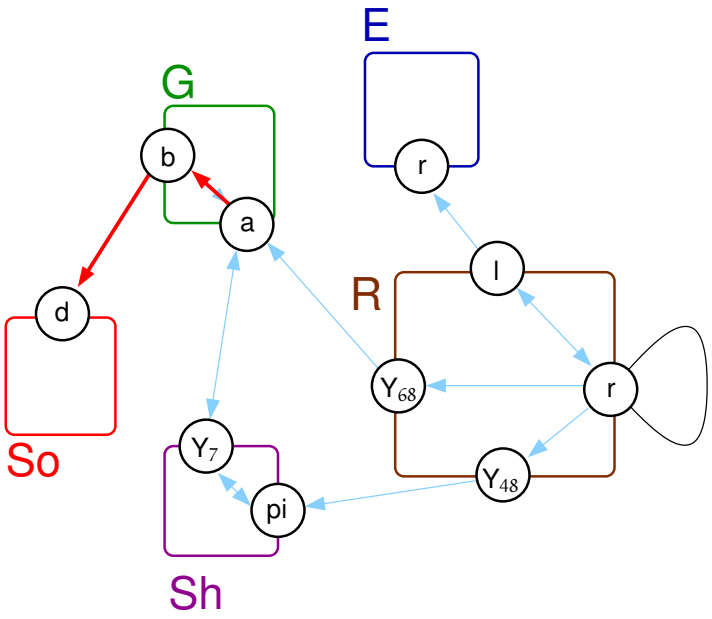




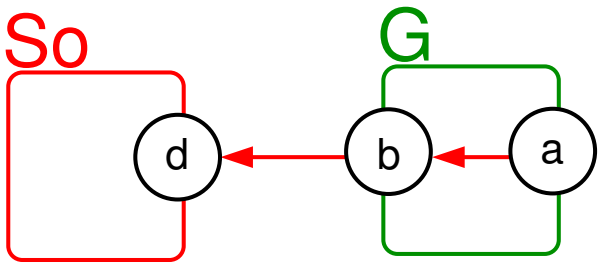
Are they fragments ?



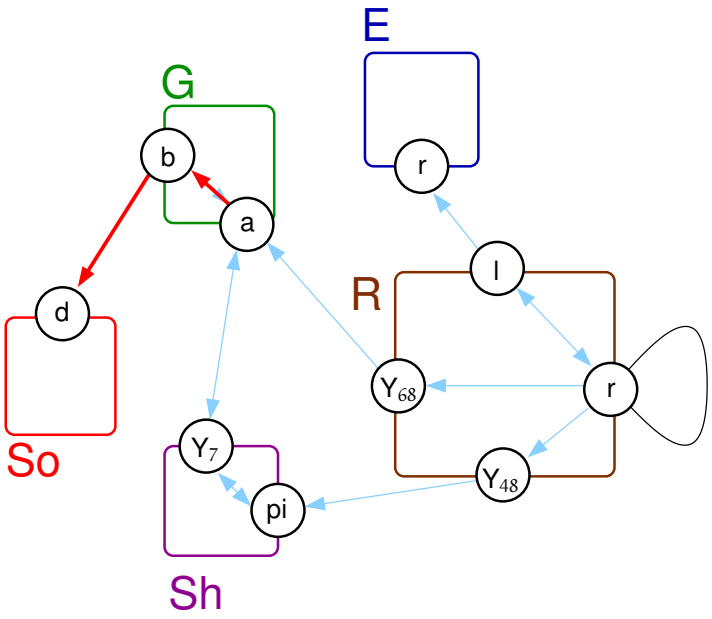
Thus, it is a prefragment.



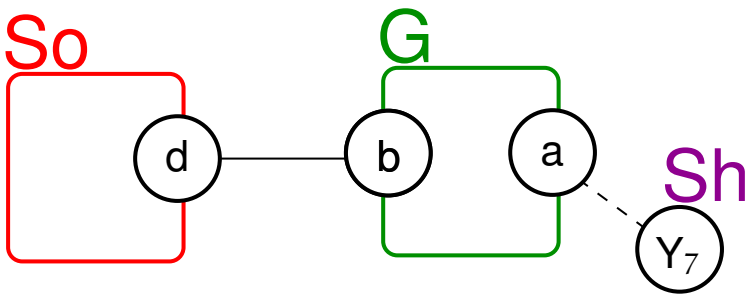
Are they fragments ?

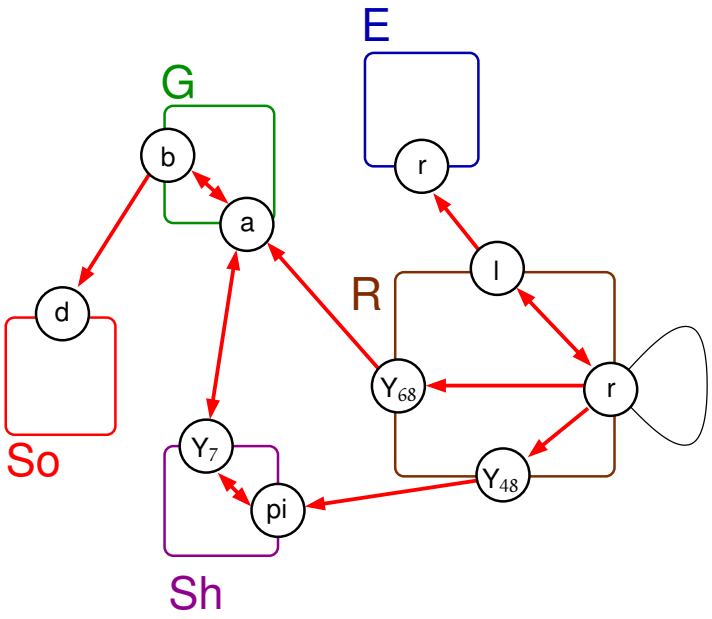


It is maximally specified.
Thus it is a fragment.

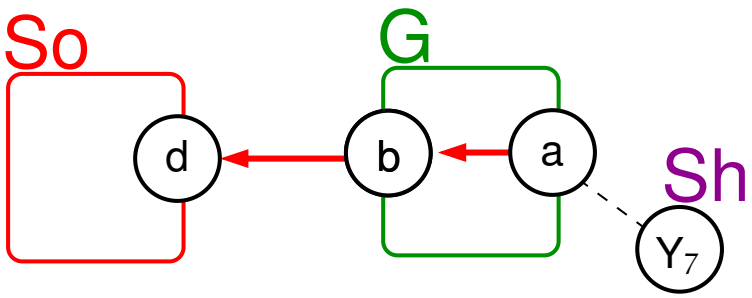


Are they fragments ?

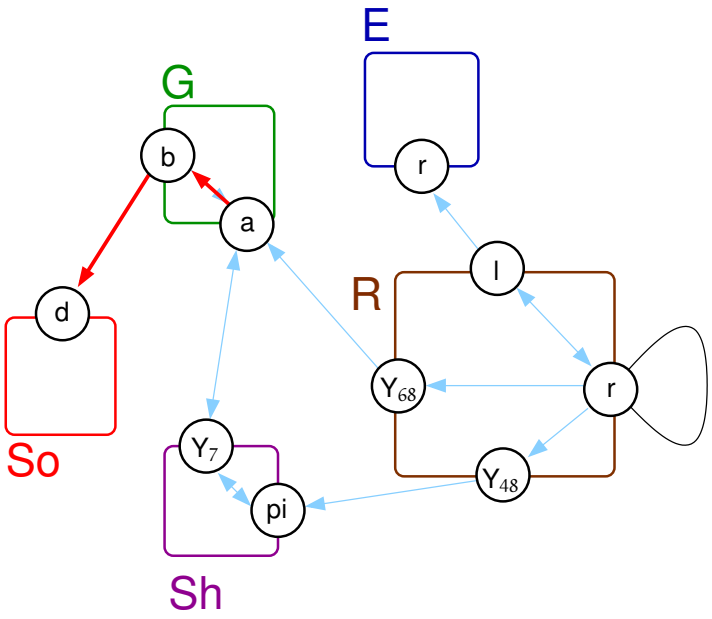




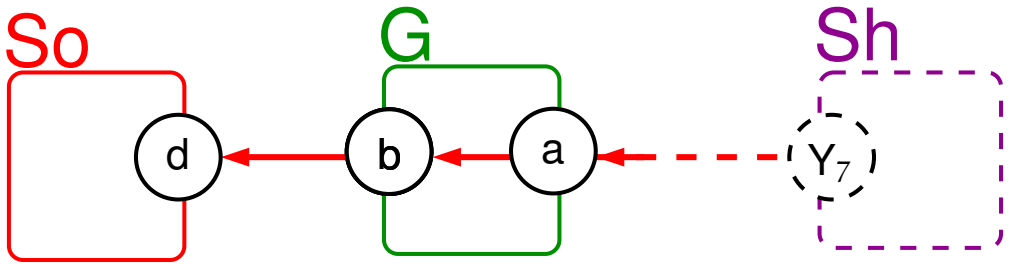
Are they fragments ?



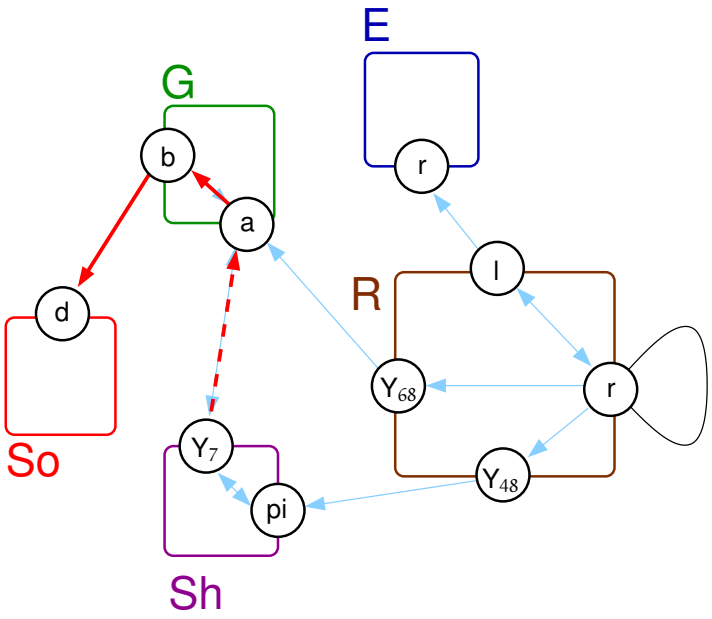
Thus, it is a prefragment.



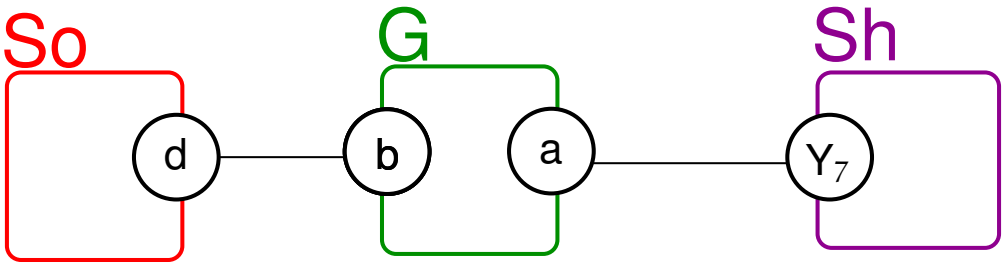
Are they fragments ?

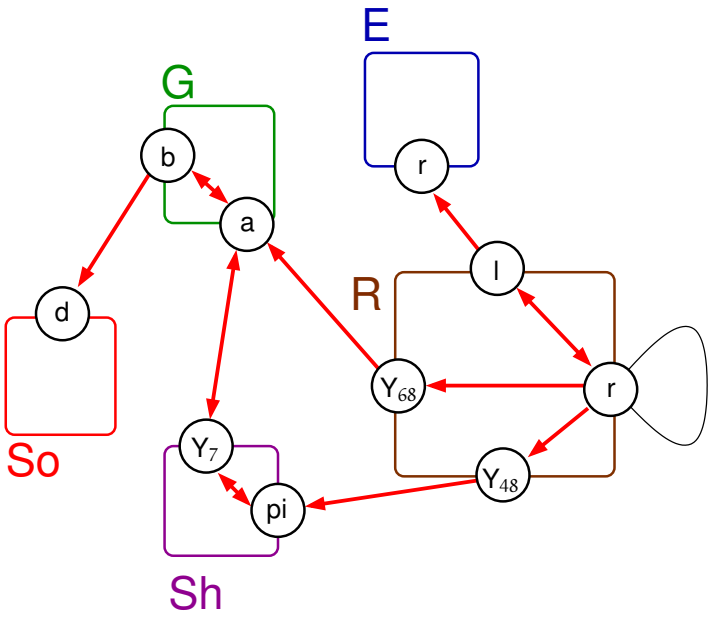


It can be refined into another prefragment.
Thus, it is not a fragment.

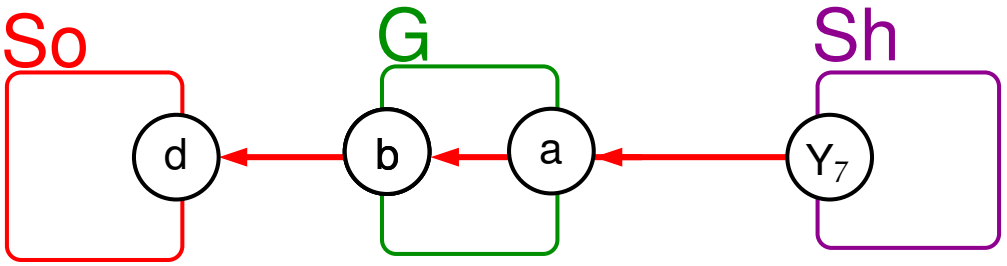


Are they fragments ?

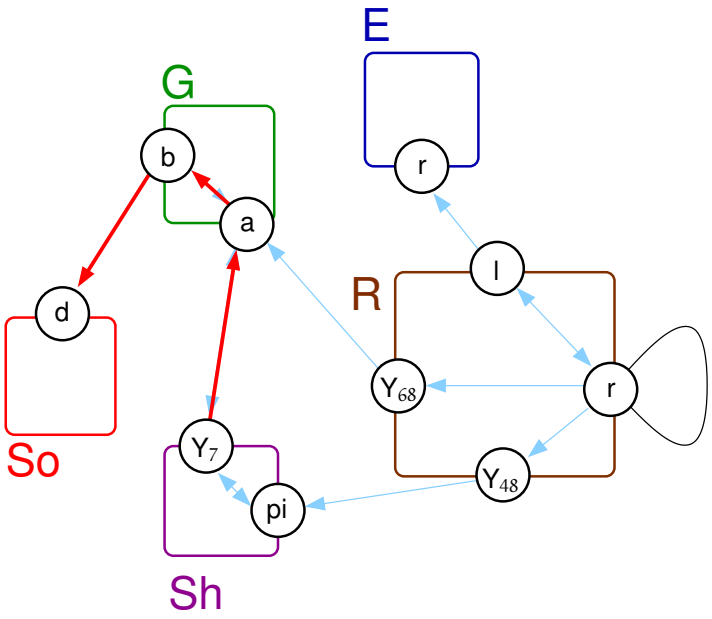




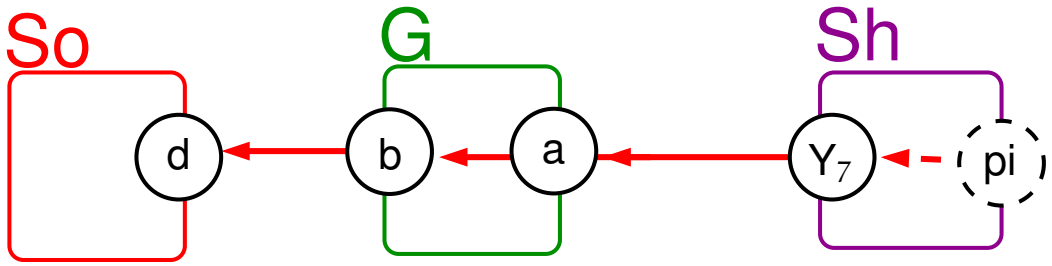
Are they fragments ?



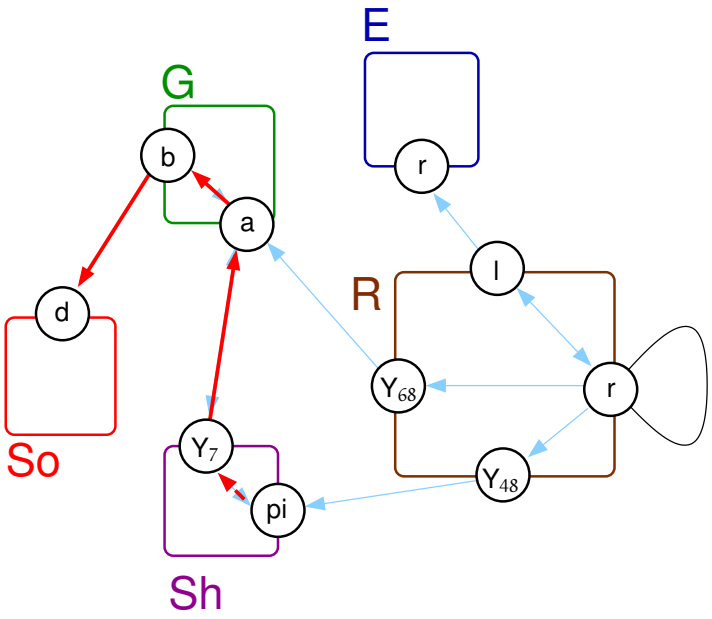
Thus, it is a prefragment.



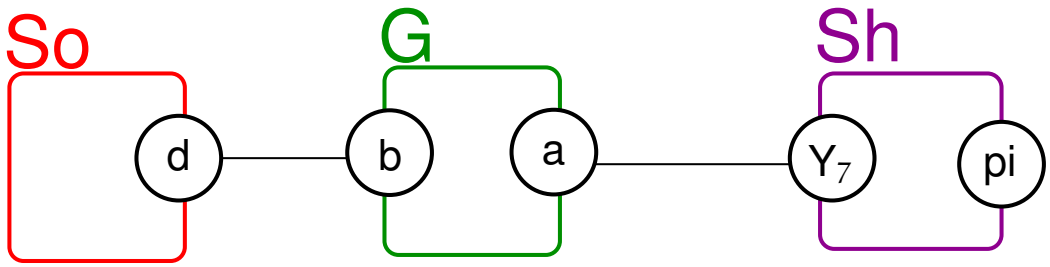
Are they fragments ?

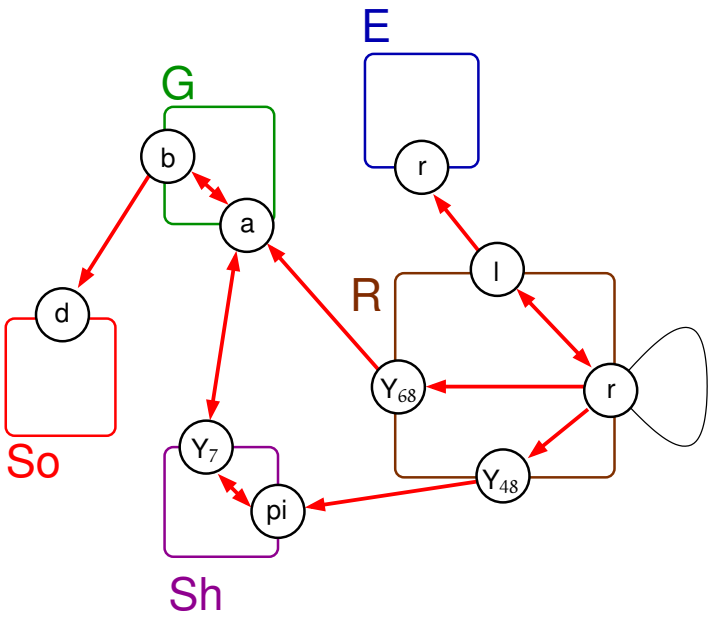


It can be refined into another prefragment.
Thus, it is not a fragment.

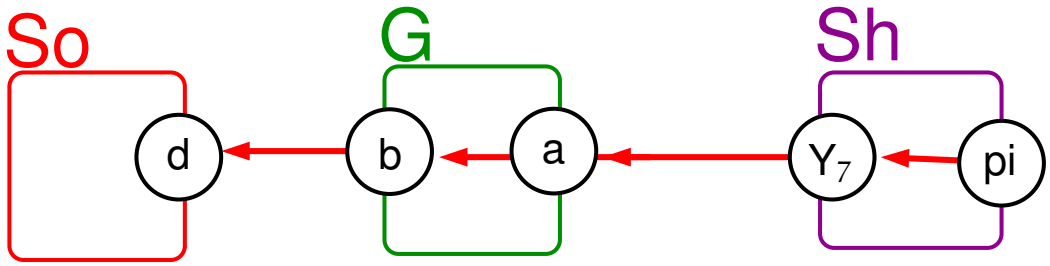


Are they fragments ?

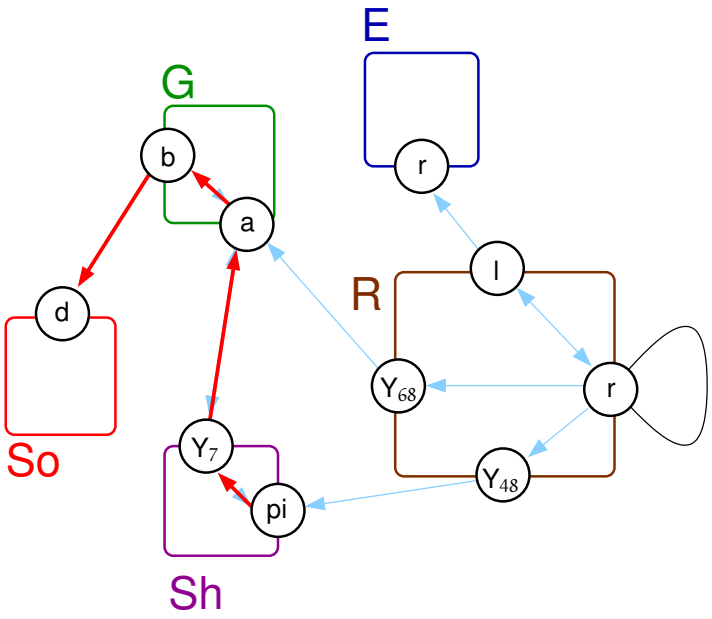




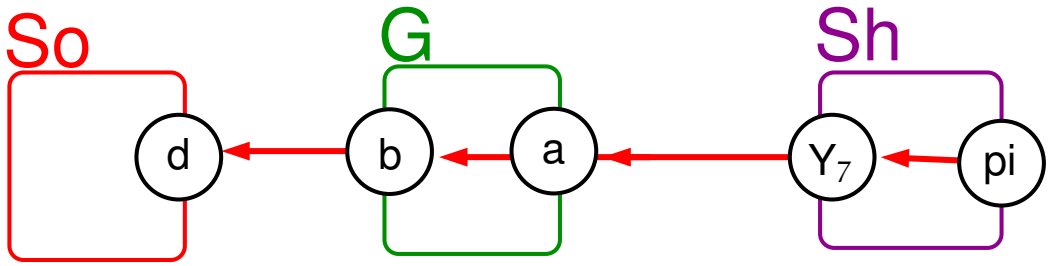
Are they fragments ?



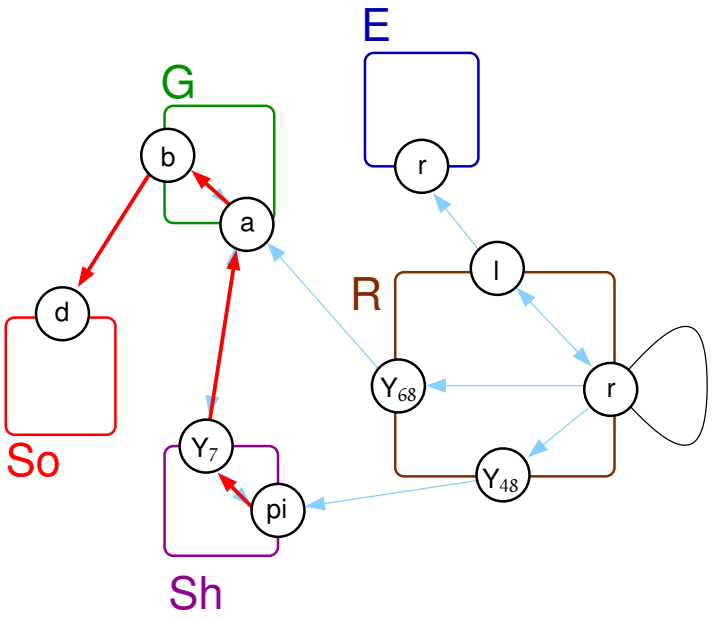
Thus, it is a prefragment.



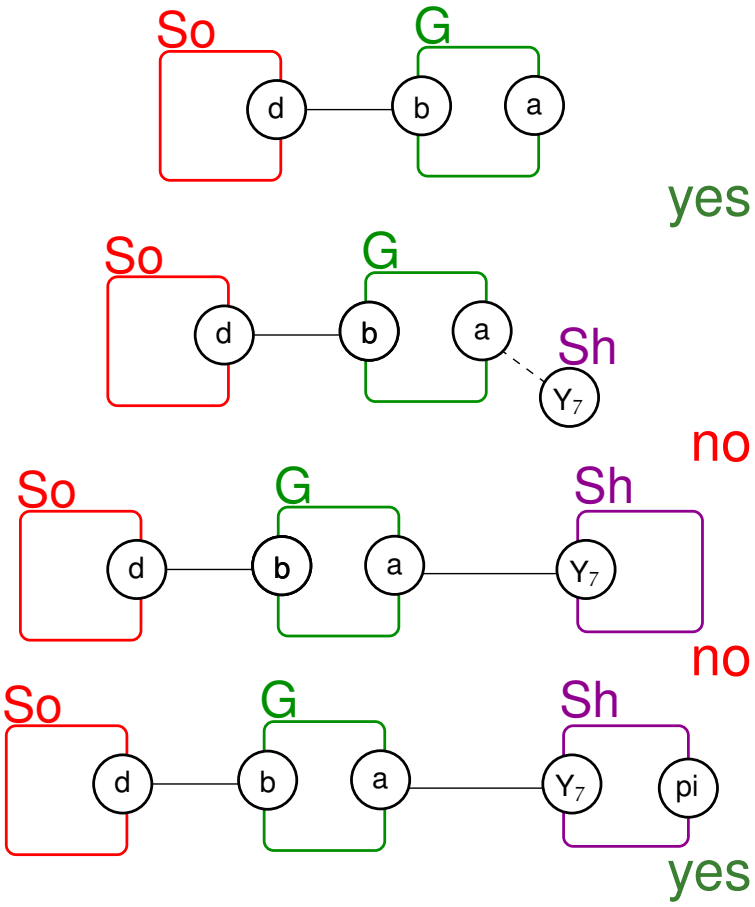
Are they fragments ?

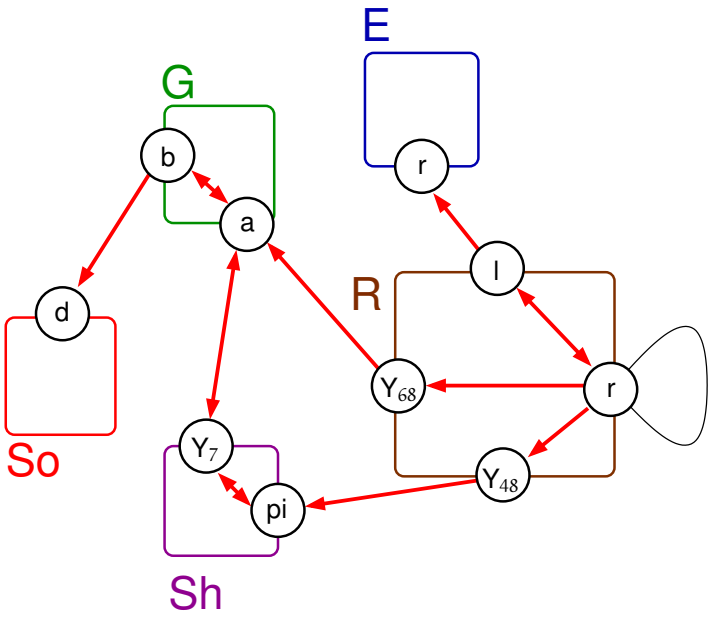


It is maximally specified.
Thus it is a fragment.

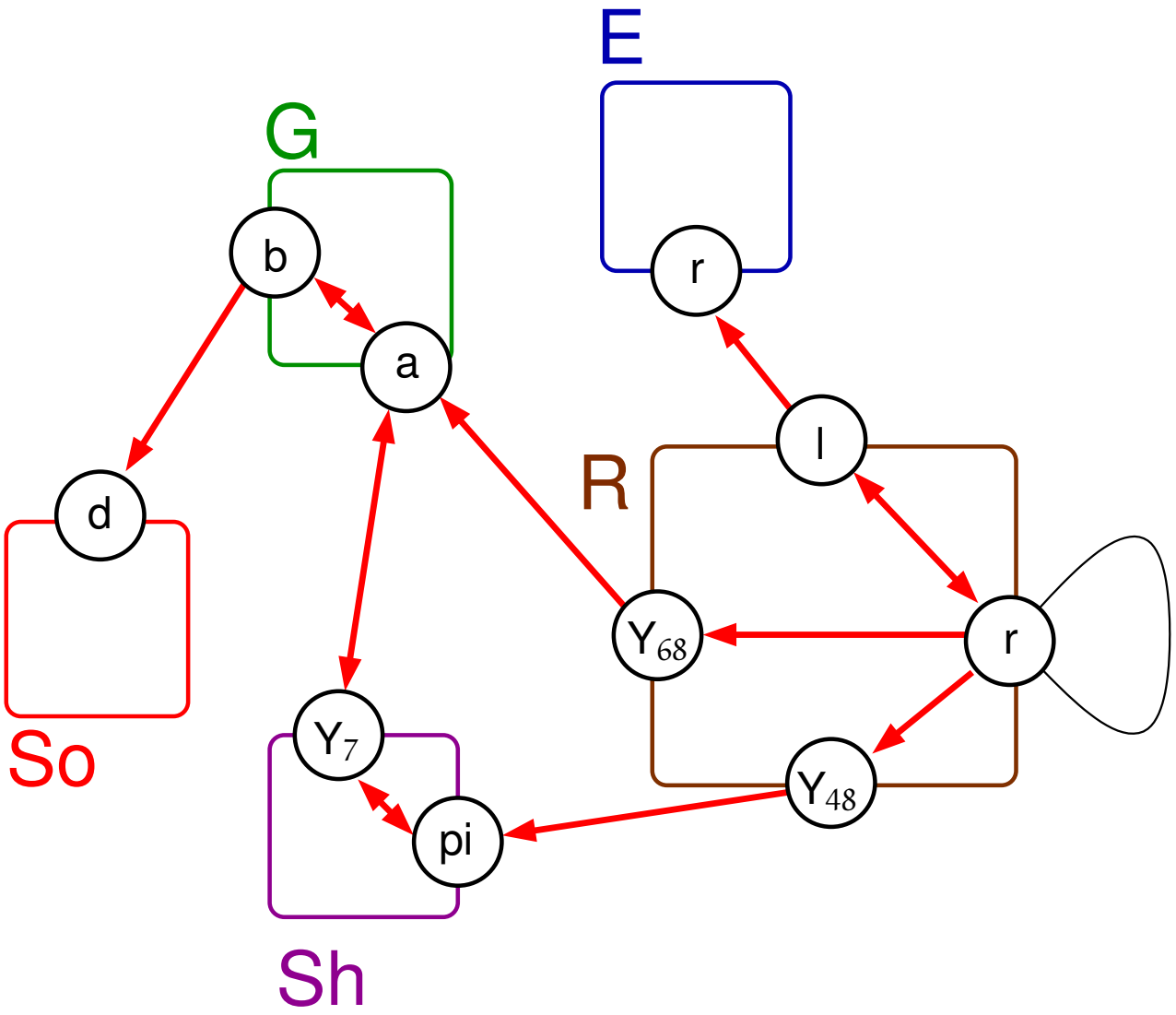


Are they fragments ?

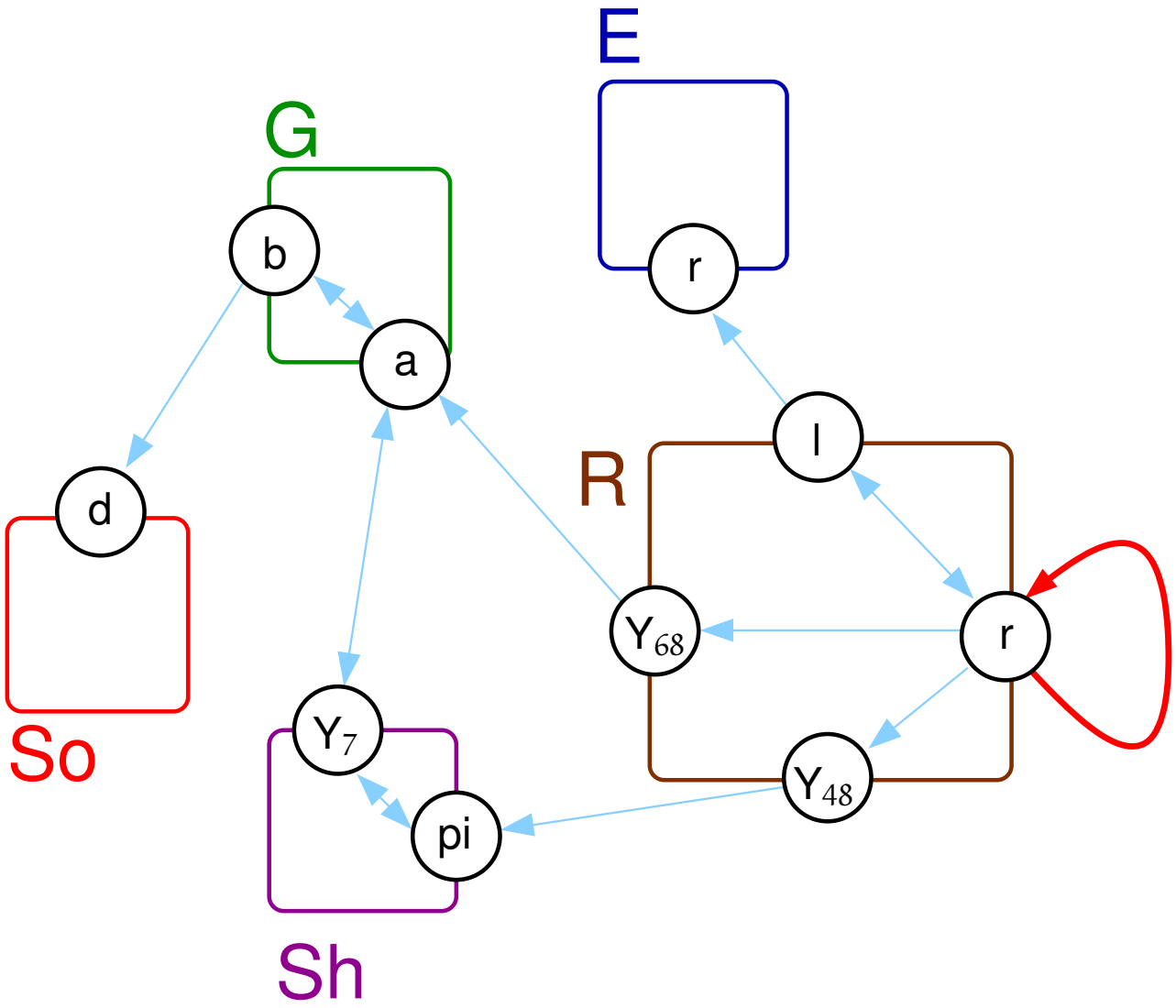




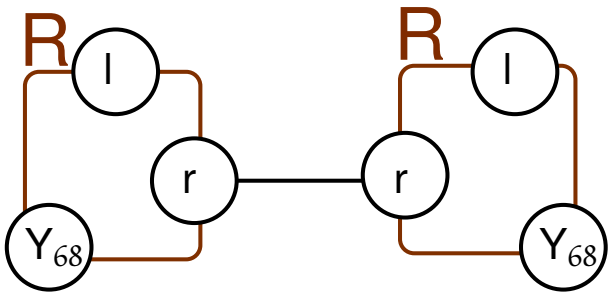
Annotated contact map

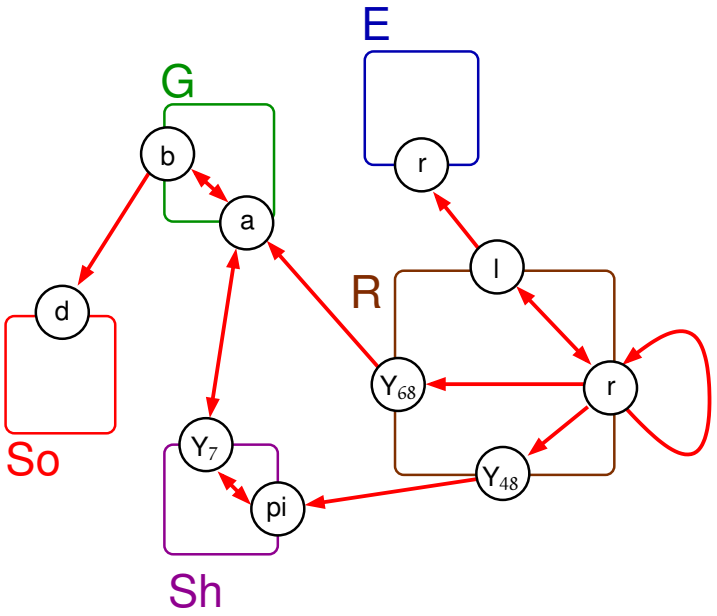


What if we were adding this flow ?

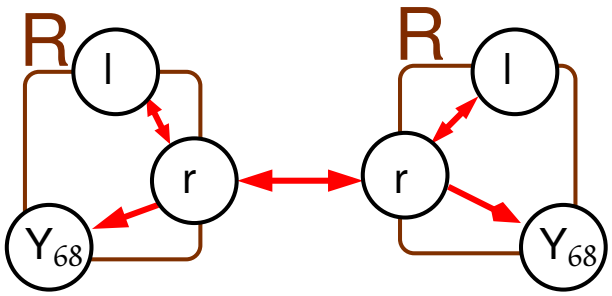


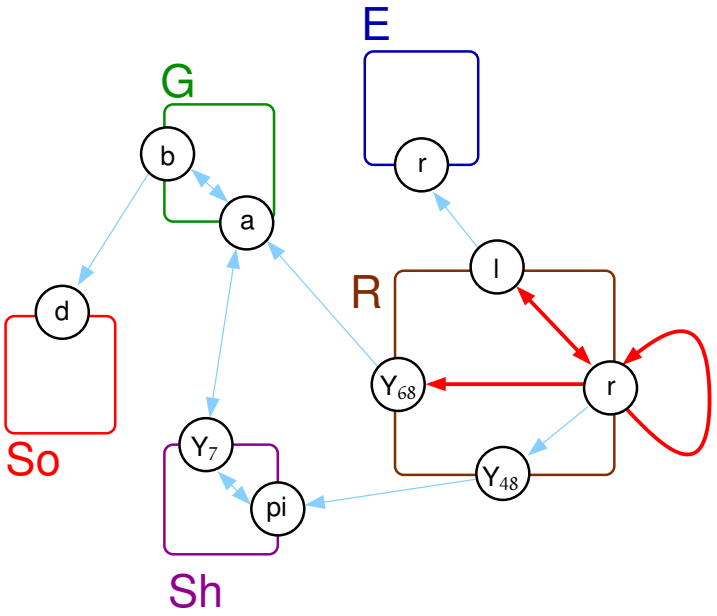
Are they fragments ?
stage 2





Are they fragments ?
stage 2

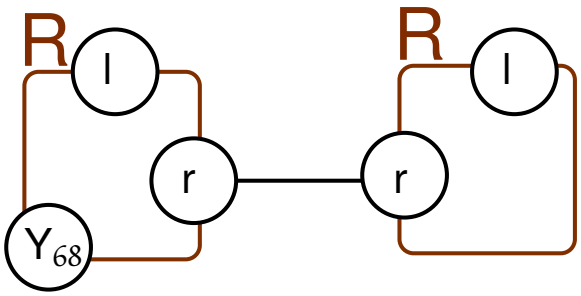


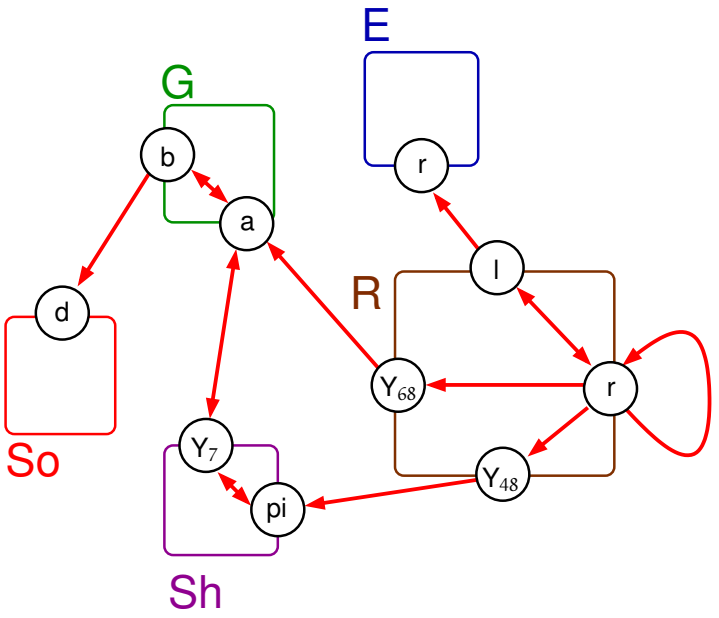


There is no way to make a path from the first Y_{68} and the second one or to make a path from the second one to the first one.

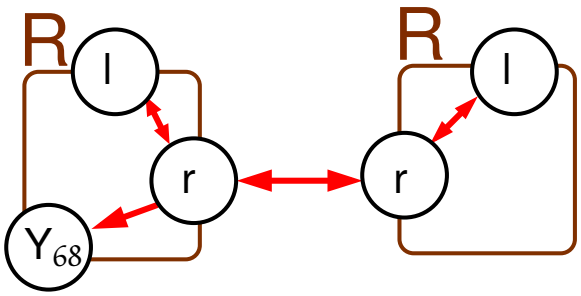
Thus it is not even a prefragment.

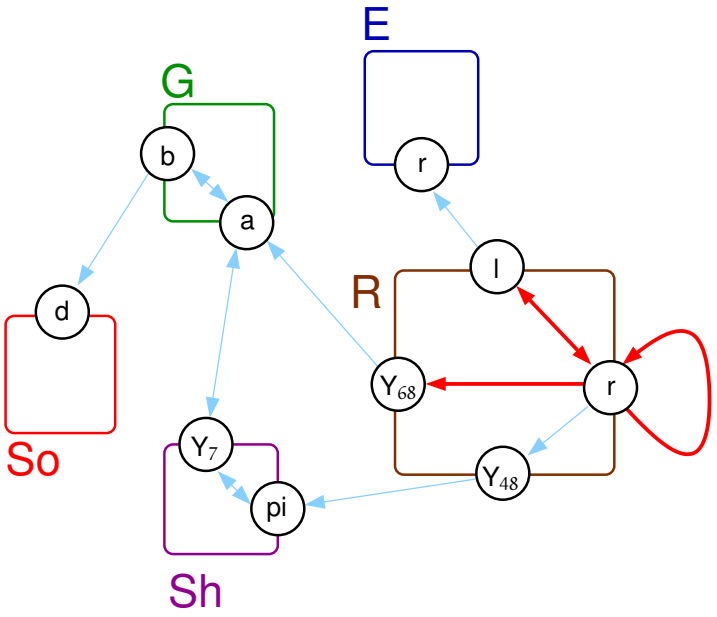
Are they fragments ?
stage 2





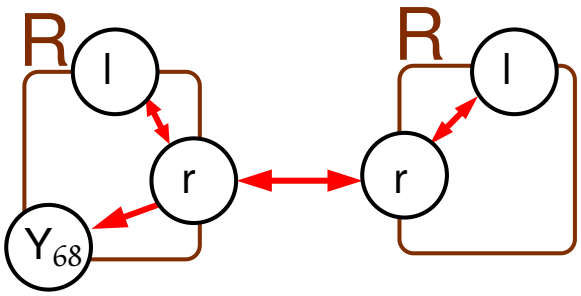
Are they fragments ?
stage 2





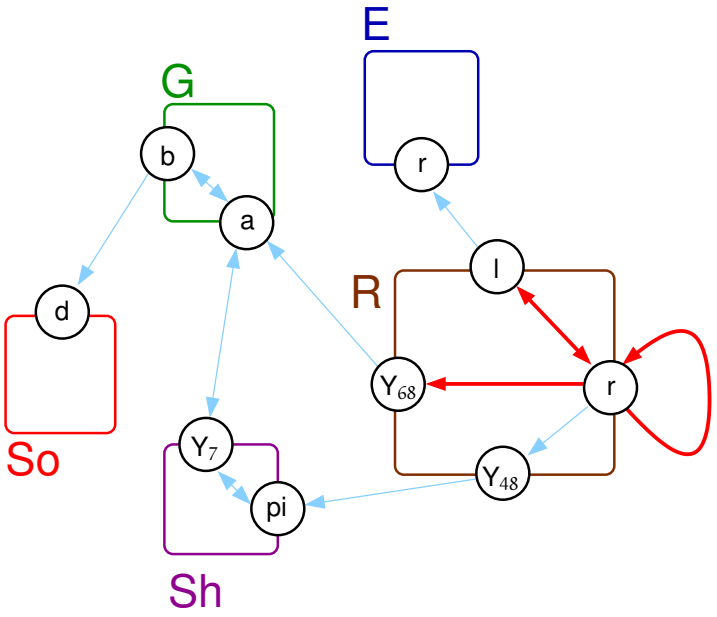
Thus it is a prefragment.

Are they fragments ?
stage 2



There is no way to refine it, while preserving the directedness.

Thus it is a fragment.



Orthogonal refinement

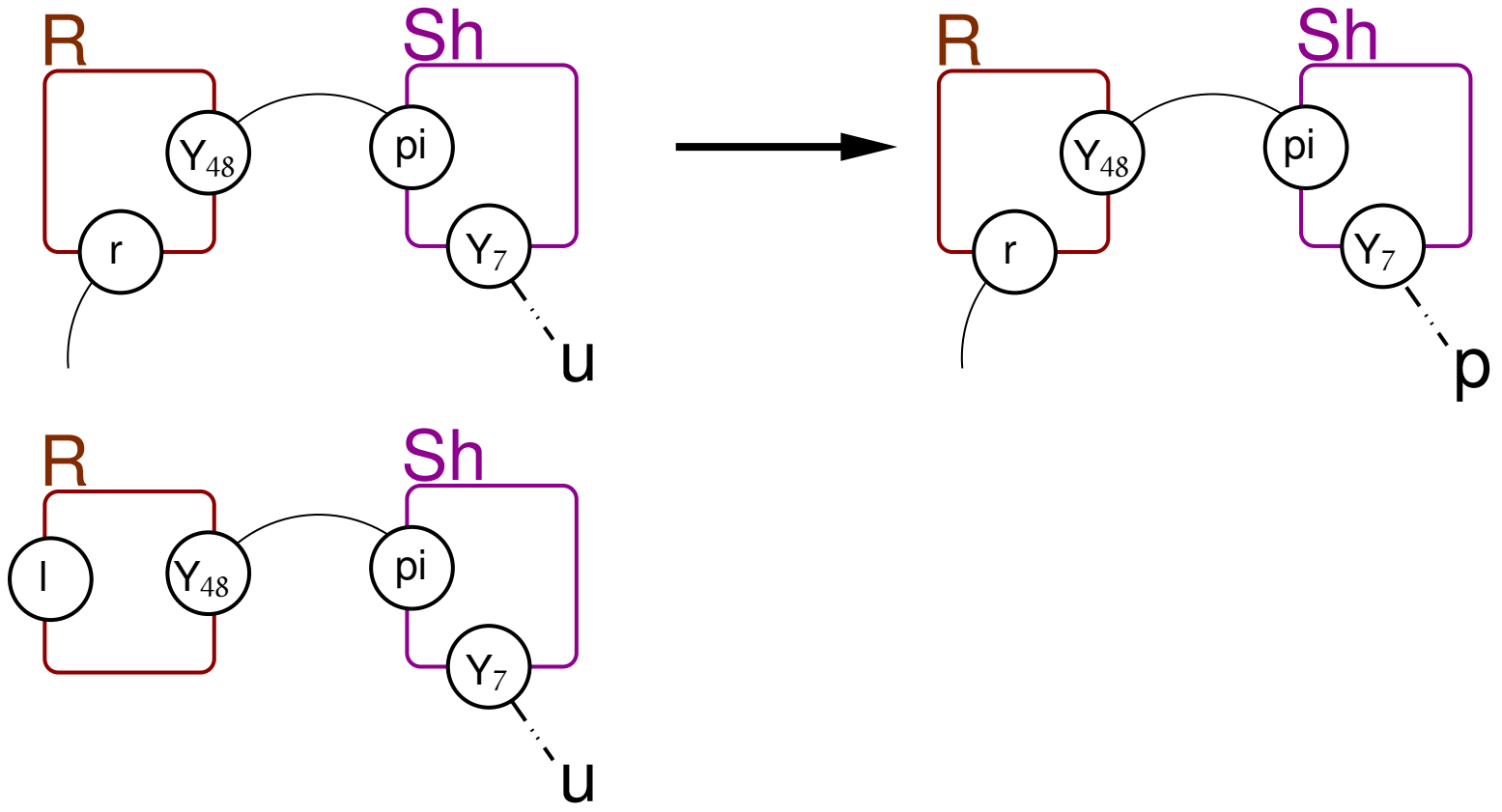
Property 1 (prefragment) The concentration of any prefragment can be expressed as a linear combination of the concentration of some fragments.

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

Overview

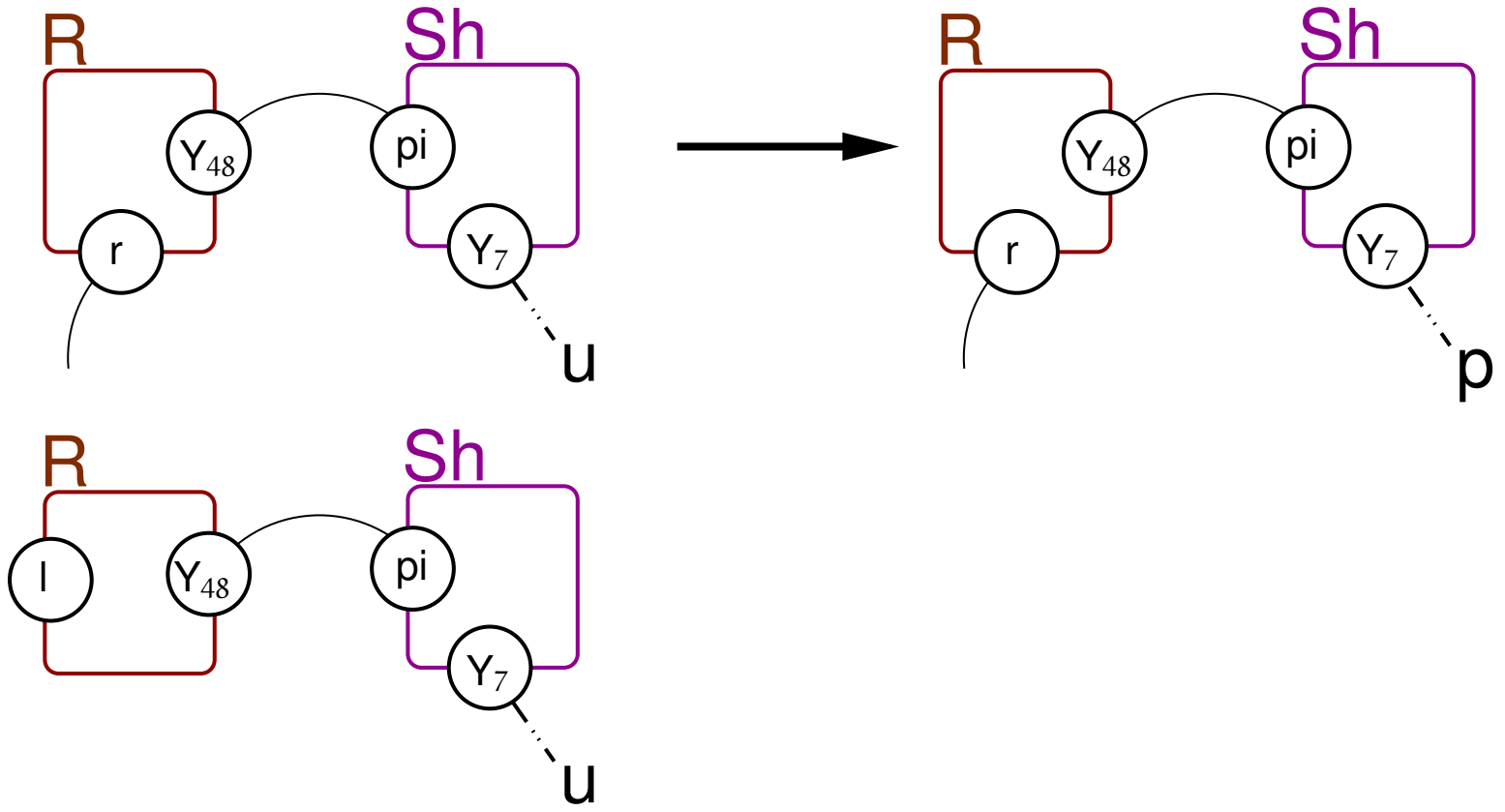
1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
   (a) Fragments
   (b) Soundness criteria
7. Conclusion

Fragments consumption



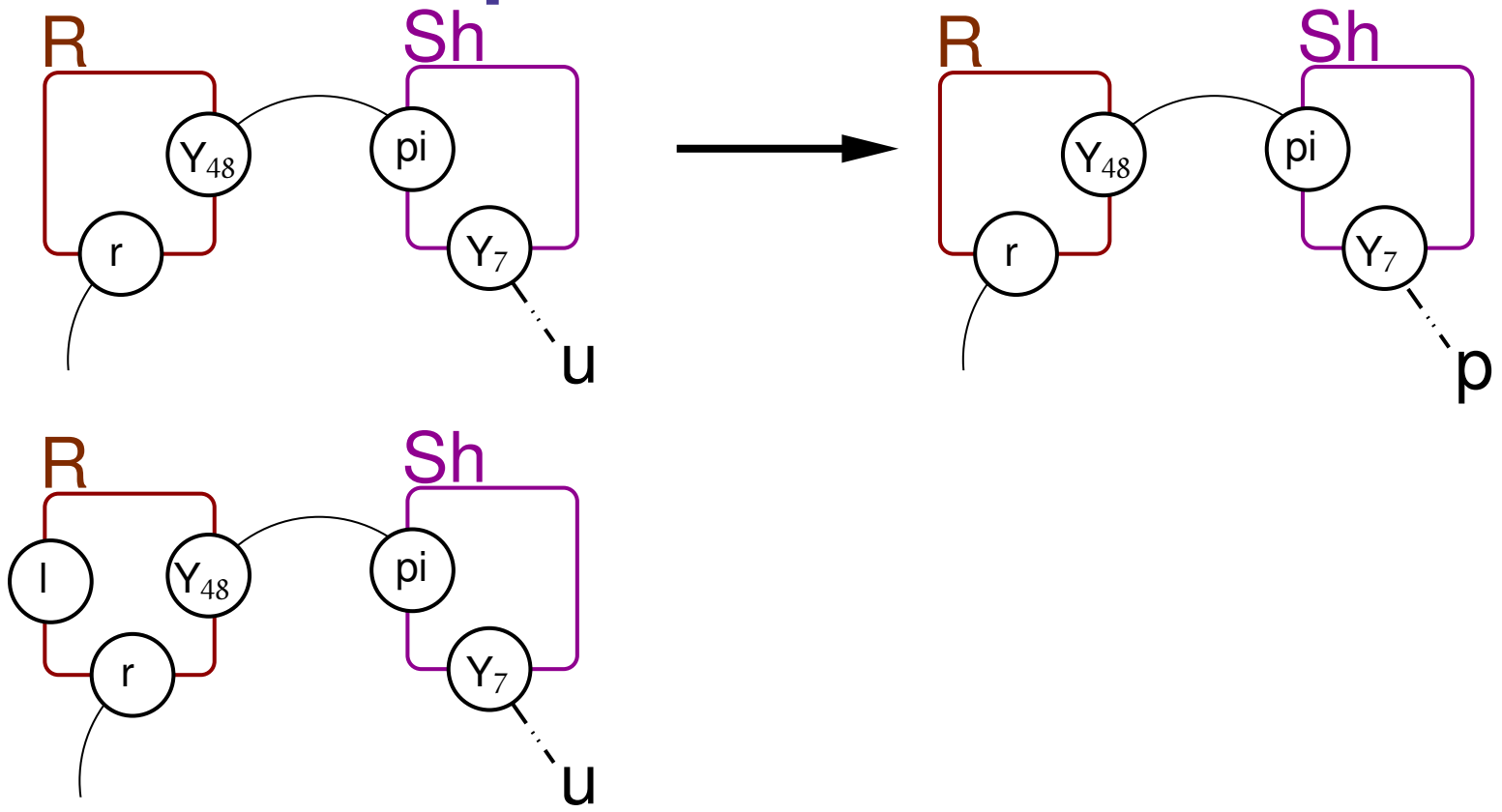
Can we express the amount (per time unit) of this fragment (bellow) concentration that is consumed by this rule (above)?

Fragments consumption



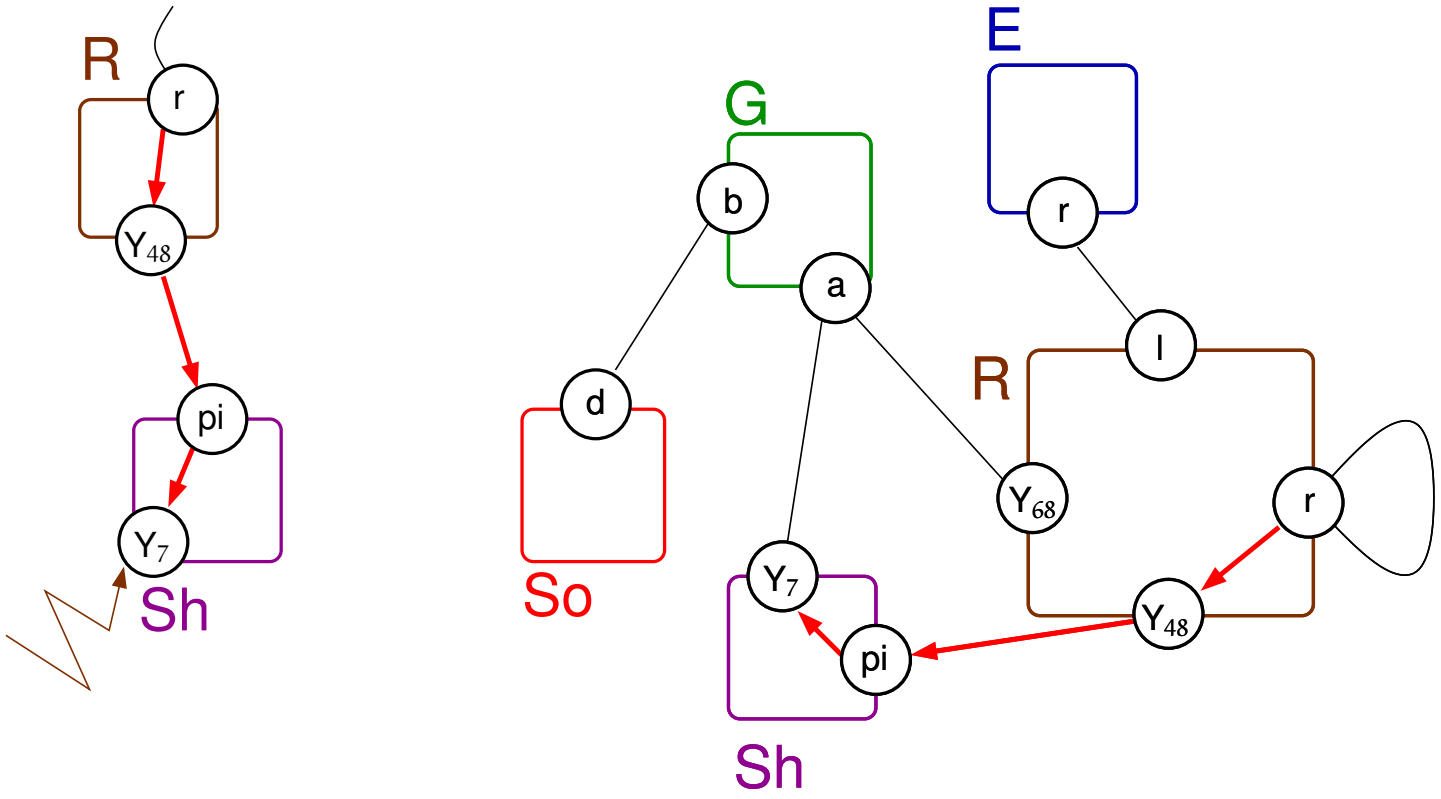
No, because we have abstracted away the correlation between the state of the site `r` and the state of the site `l`.

Fragments consumption
Proper intersection



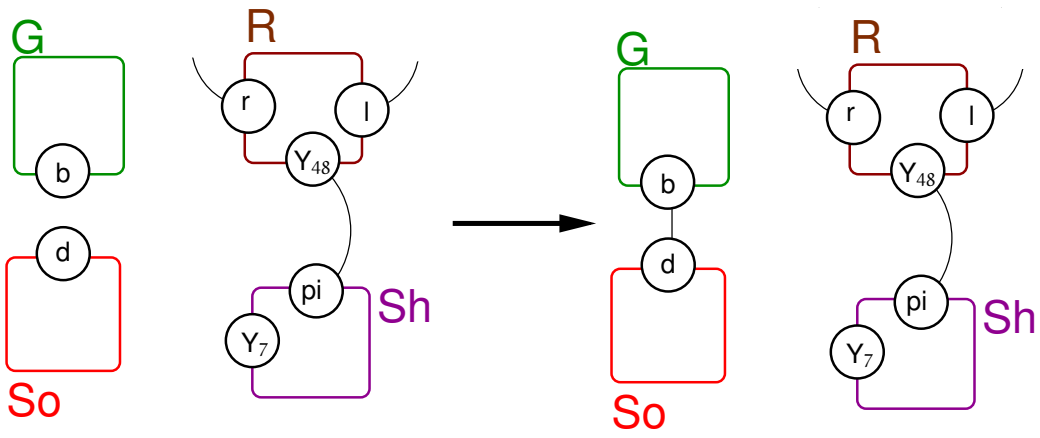
Whenever a fragment intersects a connected component of a lhs on a modified site, then the connected component is indeed embedded in the fragment!

Fragment consumption
Syntactic criteria



We reflect, in the annotated contact map, each path that stems from a site that is tested to a site that is modified.

Fragment consumption



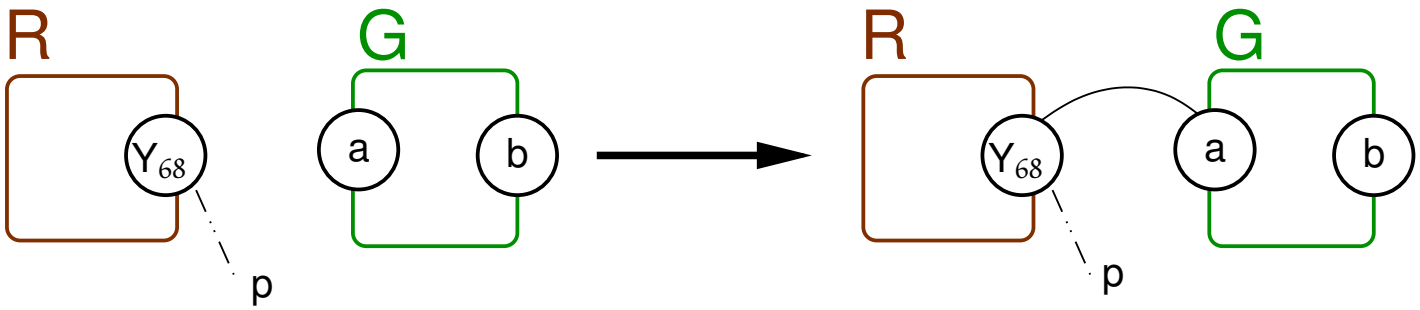
For any rule:

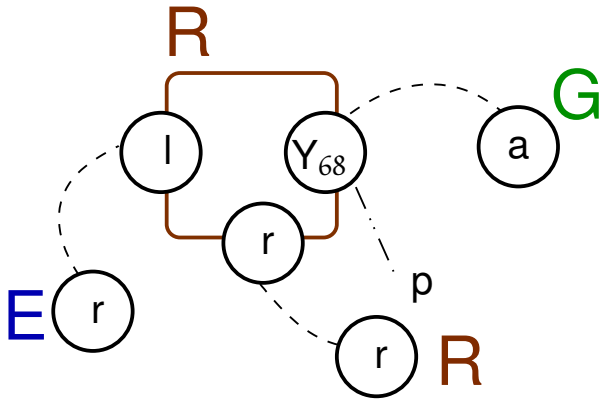
$$\textit{rule}: \; C_1, \ldots, C_n \rightarrow \textit{rhs} \quad k$$

and any embedding between a modified connected component C_k and a fragment F, we get:

$$\frac{d[F]}{dt} = \frac{k \cdot [F] \cdot \prod_{i \neq k} [C_i]}{\mathrm{SYM}(C_1, \ldots, C_n) \cdot \mathrm{SYM}(F)}.$$

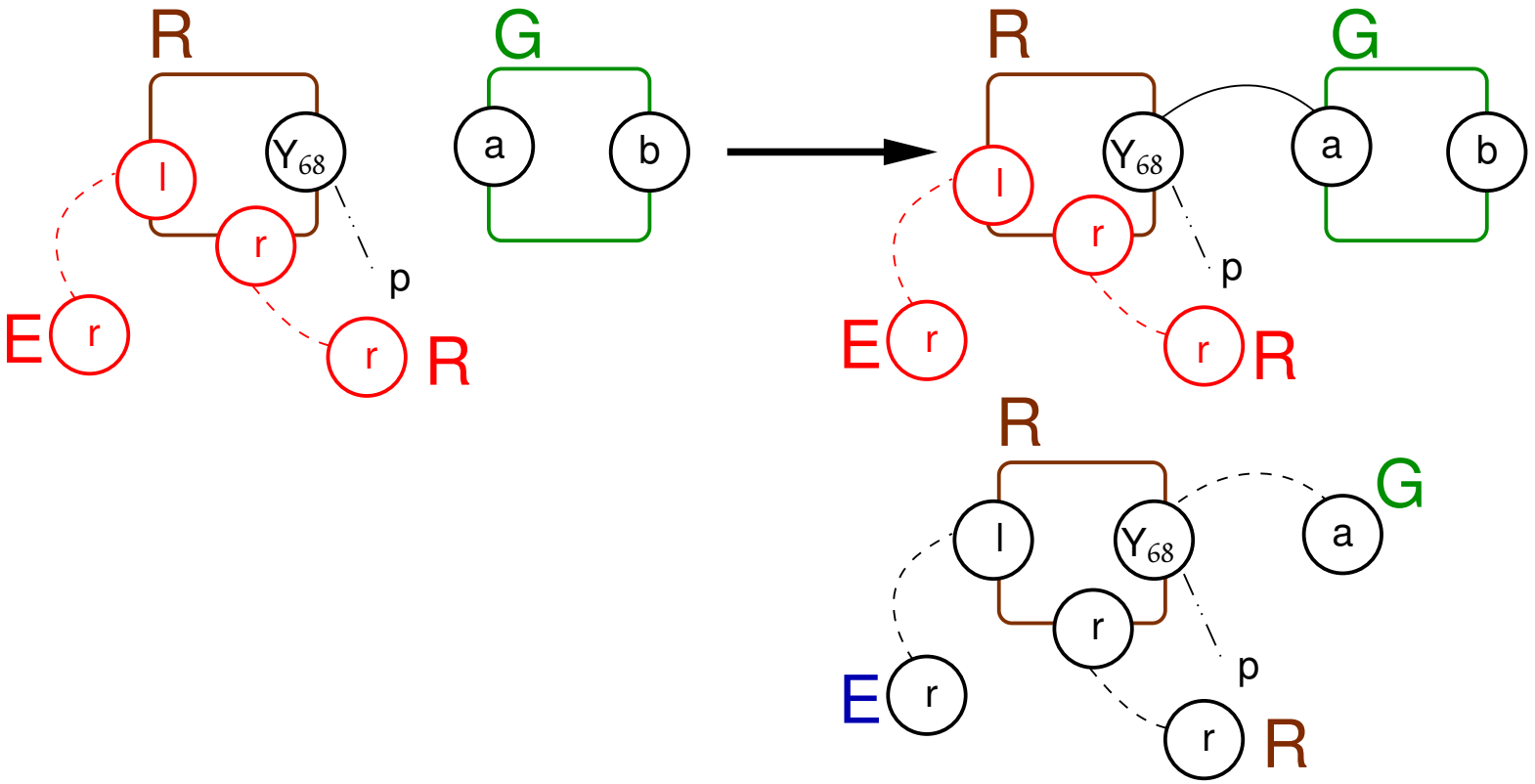
Fragment production





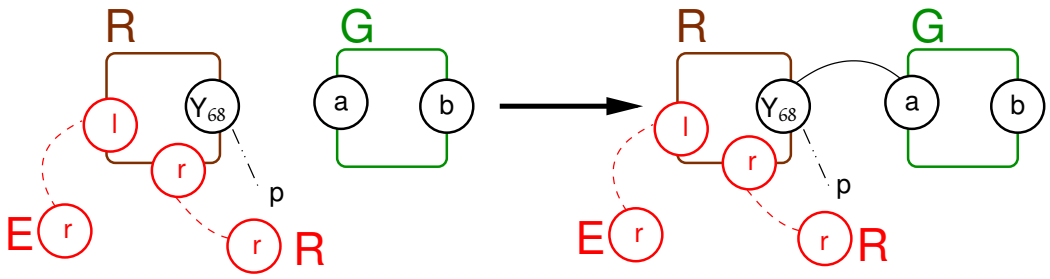
Can we express the amount (per time unit) of this fragment (bellow) concentration that is produced by the rule (above)?

Fragment production
Proper intersection (bis)



Yes, if the connected components of the lhs of the refinement are prefragments. This is already satisfied thanks to the previous syntactic criteria.

Fragment production



For any rule:

$$\mathit{rule}: \ C_1, \ldots, C_m \to \mathit{rhs} \quad k$$

and any overlap between a fragment F and *rhs* on a modified site,
we write $C'_1, \ldots, C'_n$ the lhs of the refined rule;
if $m = n$, then we get:

$$\frac{d[F]}{dt} \stackrel{+}{=} \frac{k \cdot \prod_i \left[C'_i\right]}{SYM(C_1, \ldots, C_m) \cdot SYM(F)};$$

otherwise, we get no contribution.

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:

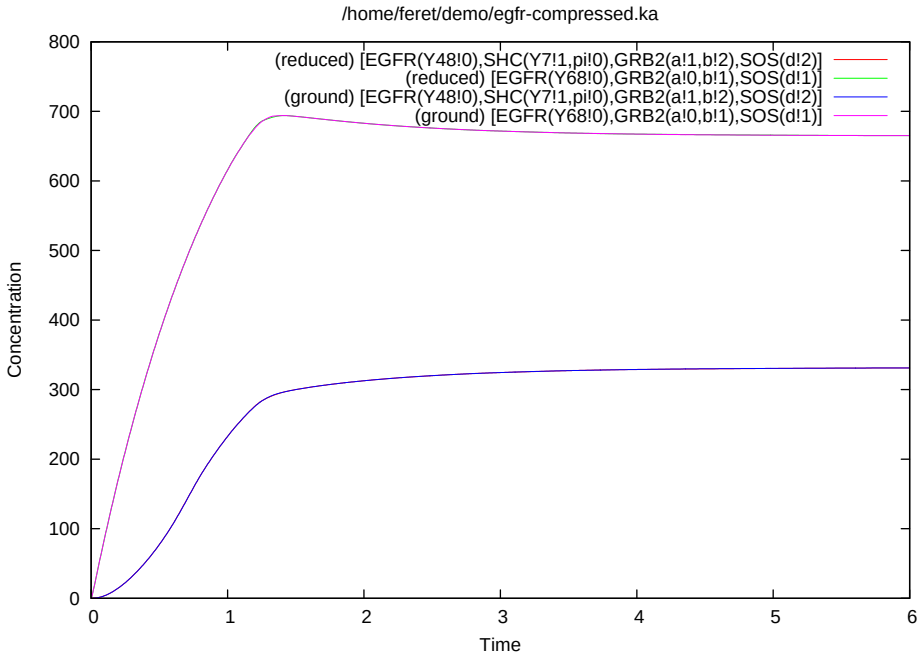
- we can express the concentration of any connected component occuring in lhss,
- we can express fragment proper consumption,
- we can express fragment proper production,
- WE HAVE A CONSTRUCTIVE DEFINITION FOR $\mathbb{F}^\sharp$.

Overview

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

Experimental results

Model	early EGF	EGF/Insulin	SFB
#species	356	2899	$\sim 2.10^{19}$
#fragments (ODEs)	38	208	$\sim 2.10^{5}$
#fragments (CTMC)	356	618	$\sim 2.10^{19}$



Both differential semantics
(4 curves with match pairwise)

Related issues

1. Model reduction of the ODE semantics:
   Joint work with Ferdinanda Camporesi
   - Less syntactic approximation of the flow of information
   - A hierarchy of abstractions tuned by the level of context-sensitivity

2. Model reduction of the stochastic semantics:
   Joint work with Thomas Henzinger, Heinz Koeppl, Tatjana Petrov
   - a framework that preserves the trace distribution
     (lumpability, backward bisimulation, equiprobability of equivalent concrete configurations)
   - Compositionality of the framework
   - Symmetry reduction

Cours MPRI

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

Friday, the 7th of February, 2014

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

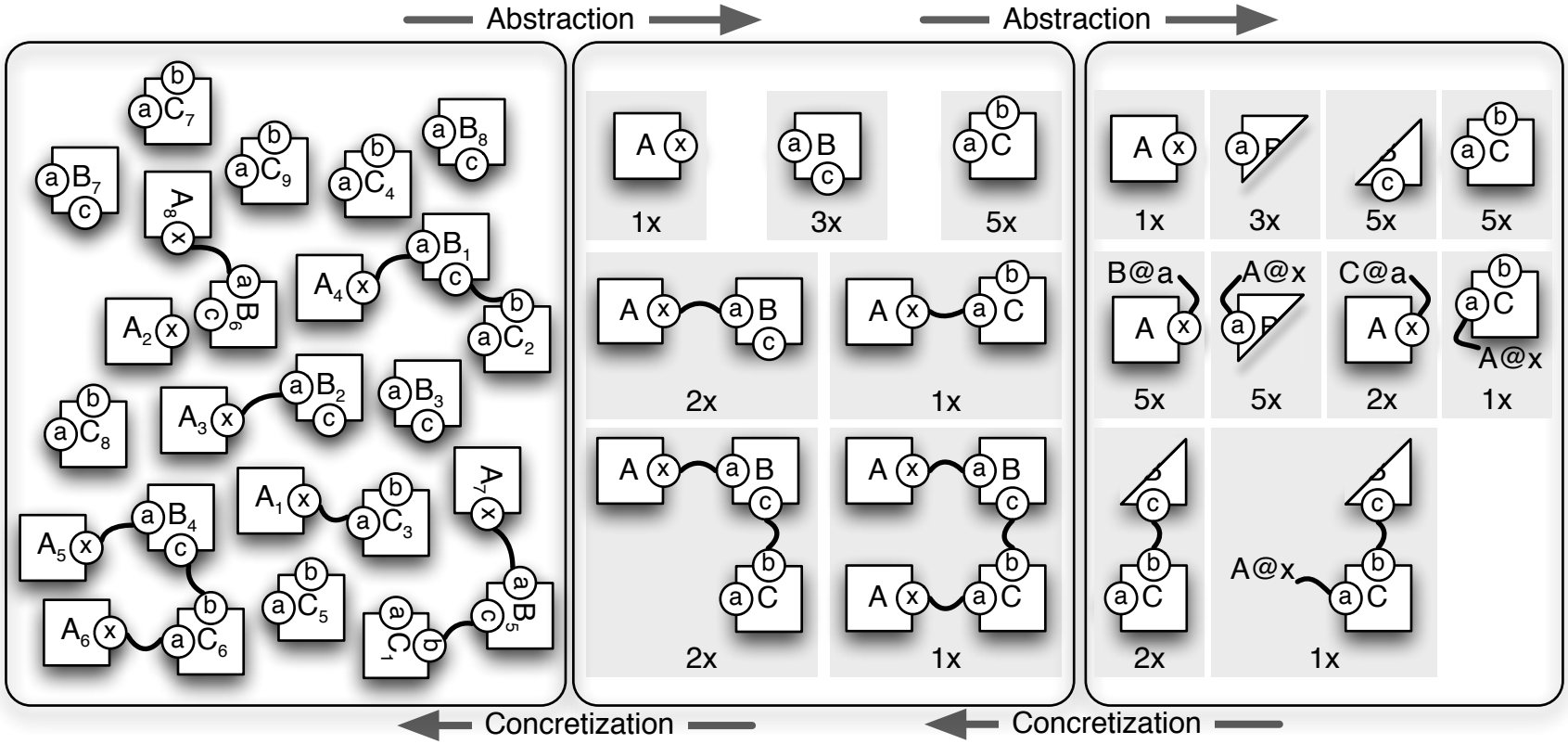
ODE fragments

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?

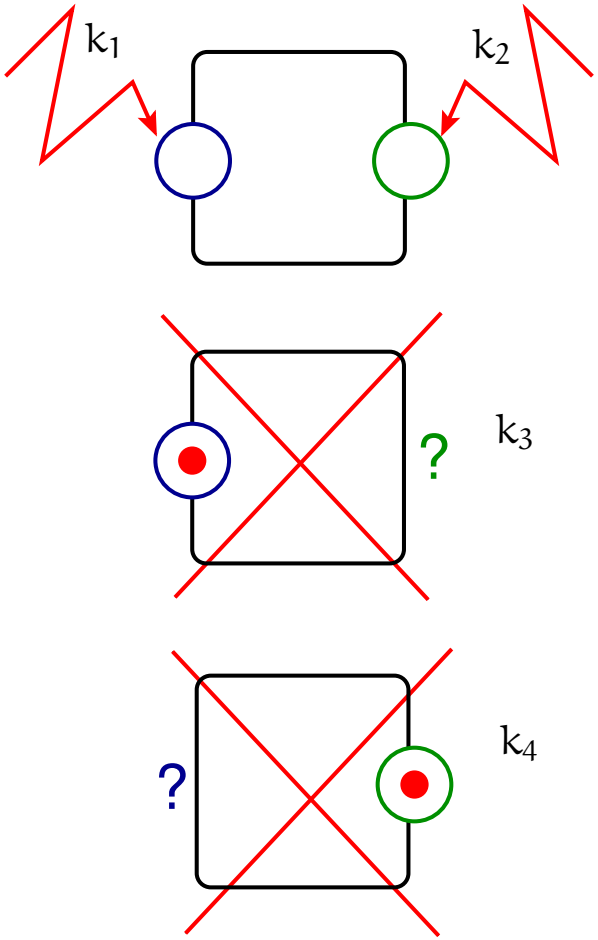
Stochastic fragments ?



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



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

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

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

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

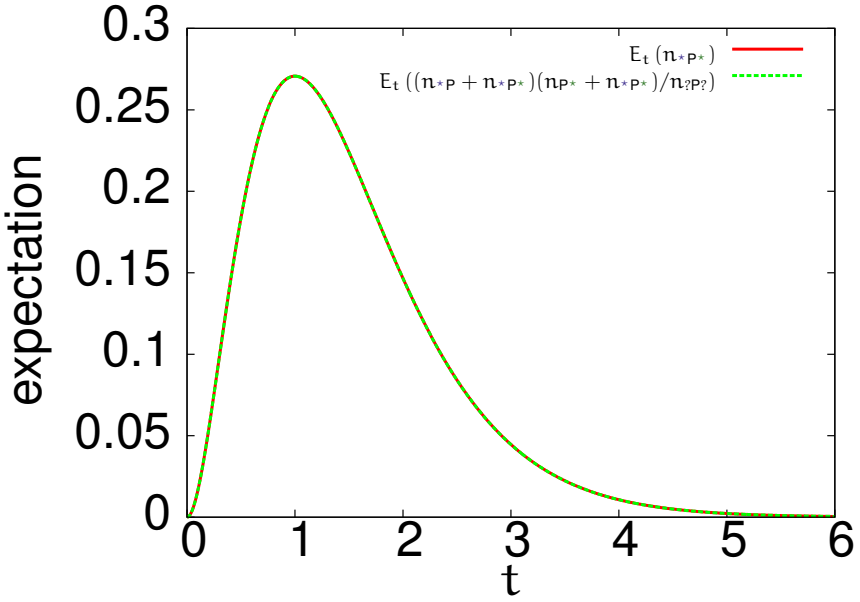
$$P^{\star} \xrightarrow{k_4} \emptyset$$

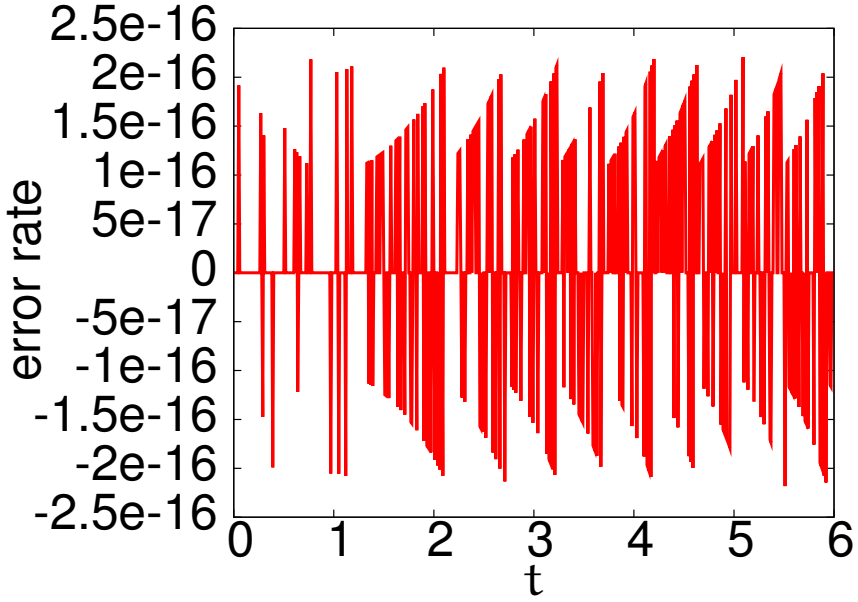
$${}^{\star}P^{\star} \xrightarrow{k_4} \emptyset$$

Statistical independence

We check numerically that:

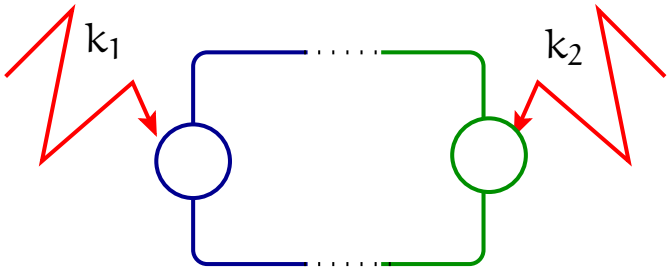
$$\mathrm{E}_t\left(n_{\star P\star}\right) = \mathrm{E}_t\left(\frac{(n_{\star P} + n_{\star P\star})(n_{P\star} + n_{\star P\star})}{n_P + n_{P\star} + n_{\star P} + n_{\star P\star}}\right).$$





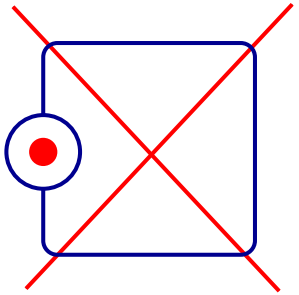
with $k_1 = k_2 = k_3 = k_4 = 1$
and two instances of P at time $t = 0$.

Reduced model



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

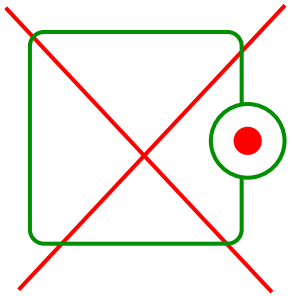
$$P \xrightarrow{k_2} P^{\star}$$



k_3

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

+ side effect: remove one P

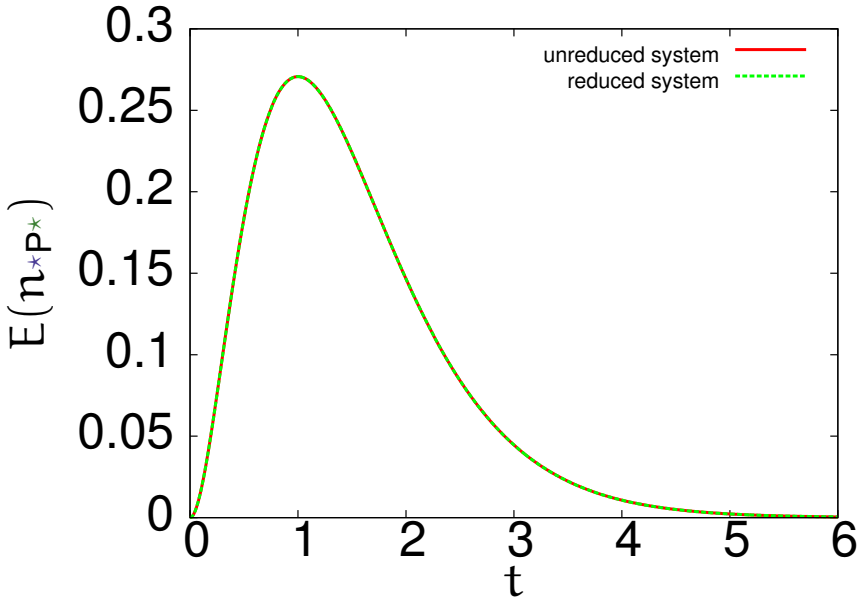


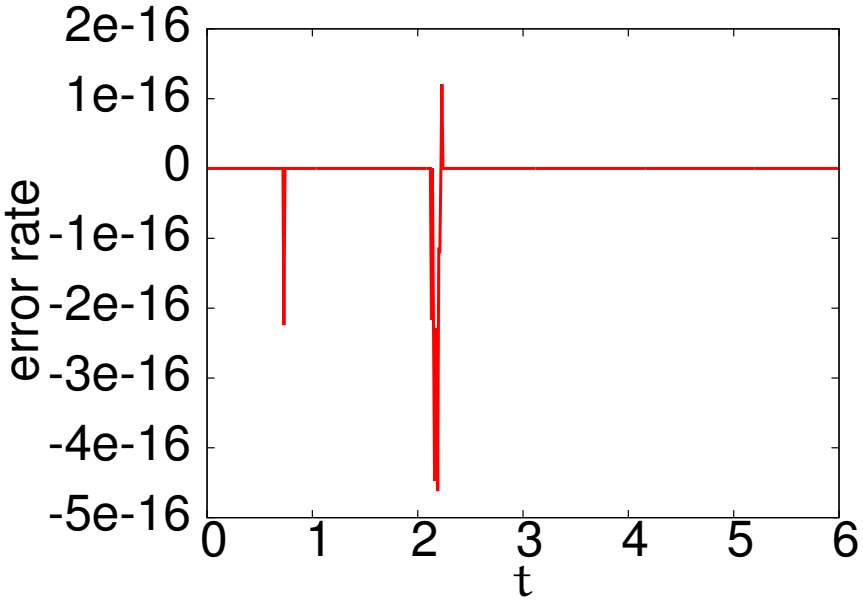
k_4

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

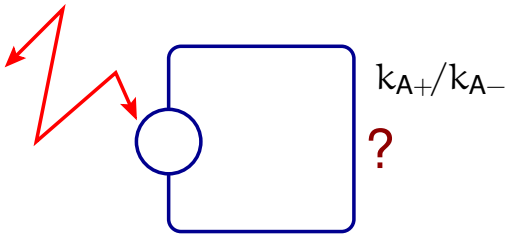
+ side effect: remove one P

Comparison between the two models

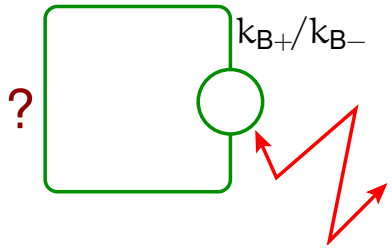




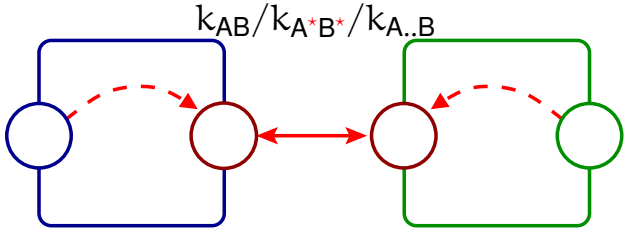
Coupled semi-reactions



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



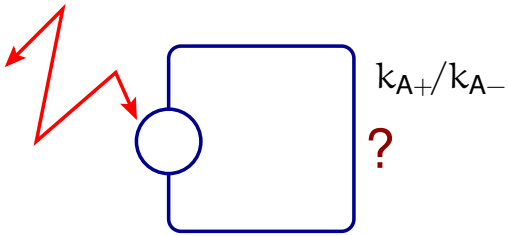
$$B \underset{k_{B-}}{\overset{k_{B+}}{\rightleftharpoons}} B^\star, \quad AB \underset{k_{B-}}{\overset{k_{B+}}{\rightleftharpoons}} AB^\star, \quad A^\star B \underset{k_{B-}}{\overset{k_{B+}}{\rightleftharpoons}} A^\star B^\star$$

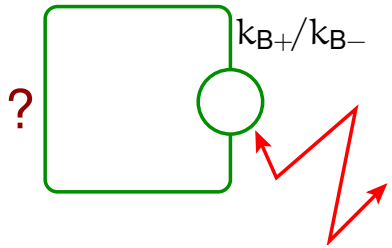


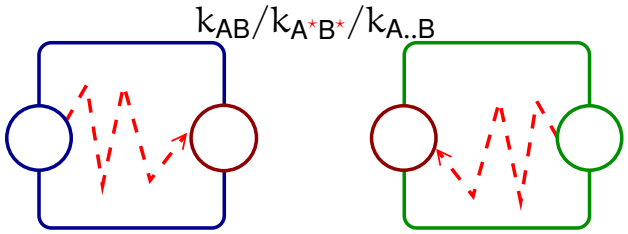
$$A + B \underset{k_{A..B}}{\overset{k_{AB}}{\rightleftharpoons}} AB, \quad A^\star + B \underset{k_{A..B}}{\overset{k_{AB}}{\rightleftharpoons}} A^\star B,$$

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

Reduced model







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

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

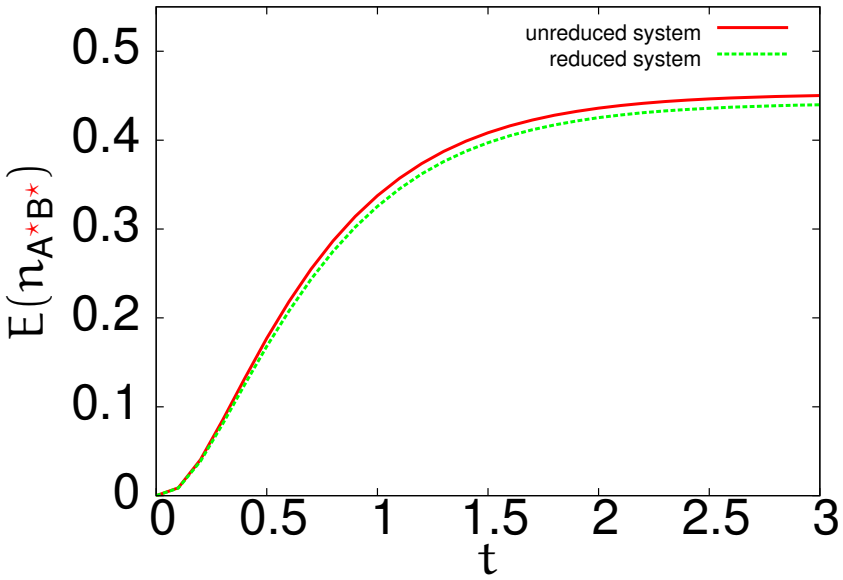
$$A + B \underset{k_{A..B}/(n_{A^\diamond B}+n_{A^\diamond B^\star})}{\overset{k_{AB}}{\rightleftharpoons}} AB^\diamond + A^\diamond B,$$

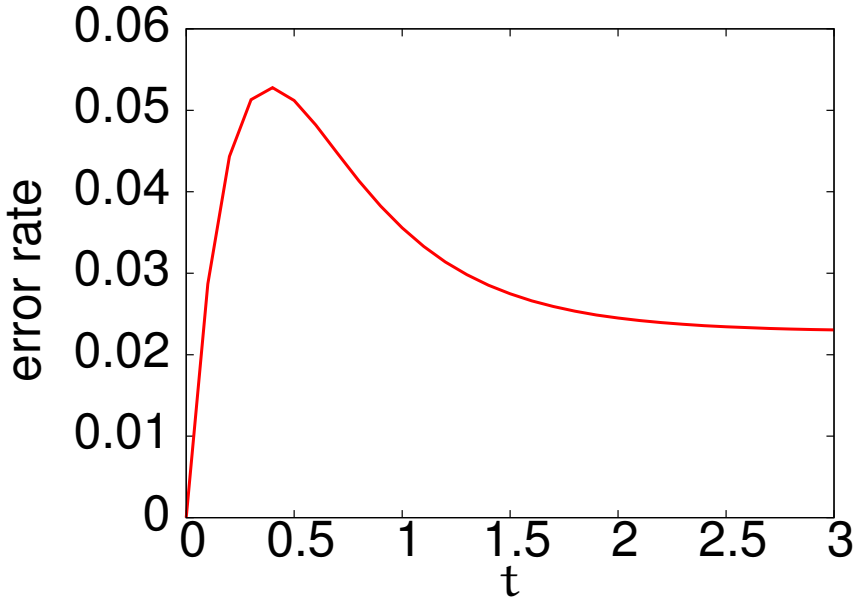
$$A^\star + B \underset{k_{A..B}/(n_{A^\diamond B}+n_{A^\diamond B^\star})}{\overset{k_{AB}}{\rightleftharpoons}} A^\star B^\diamond + A^\diamond B,$$

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

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

Comparison between the two models

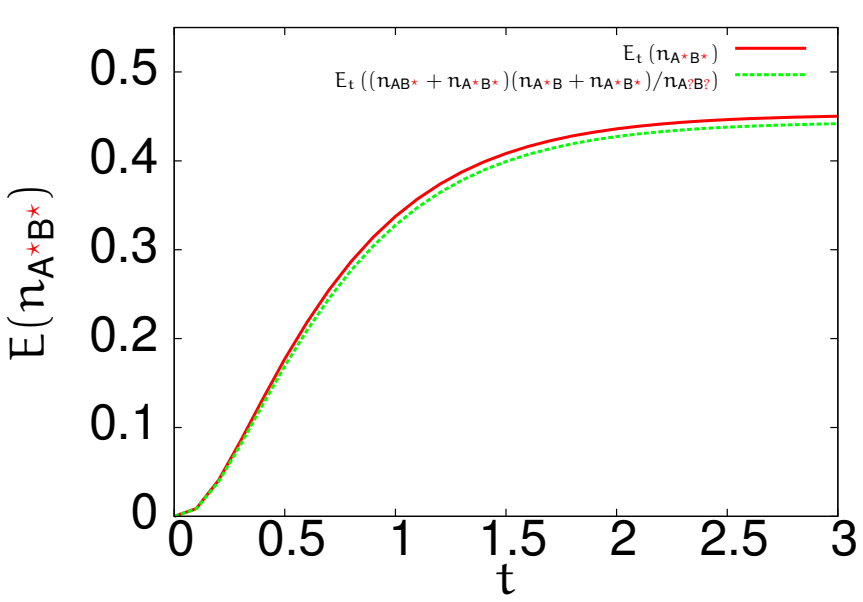


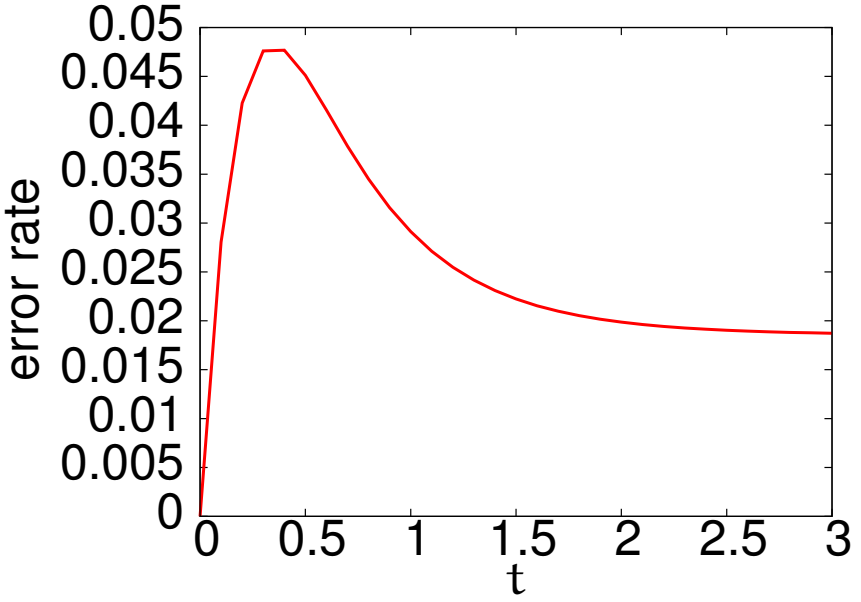


with $k_{A+} = k_{A-} = k_{B+} = k_{B-} = k_{AB} = k_{A..B} = 1$, $k_{A^\star B^\star} = 10$,
and two instances of A and B at time $t = 0$.

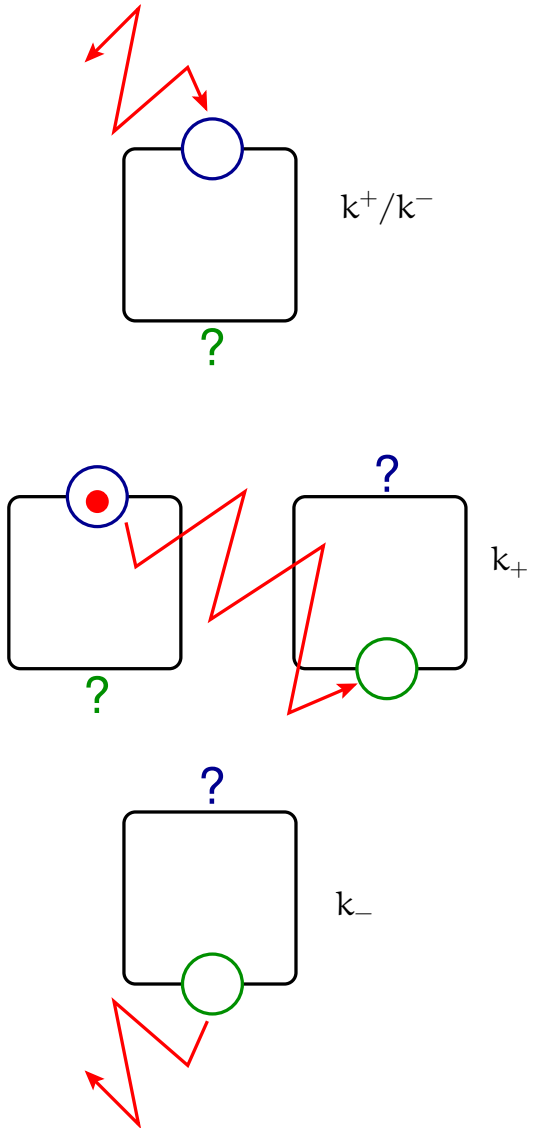
Although the reduction is correct in the ODE semantics.

Degree of correlation (in the unreduced model)





Distant control



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

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

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

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

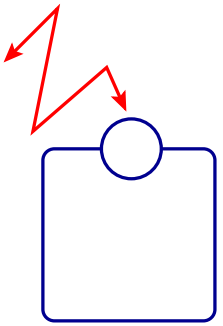
$$A + A^\star_\star \xrightarrow{k_+} A_\star + A^\star_\star$$

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

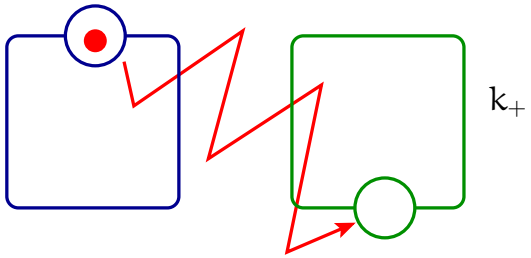
$$A^\star_\star \xrightarrow{k_-} A^\star$$

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

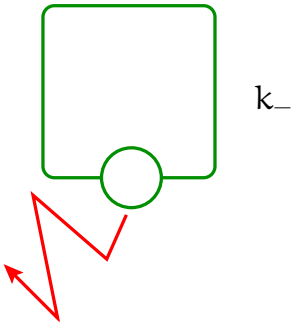
Reduced model



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

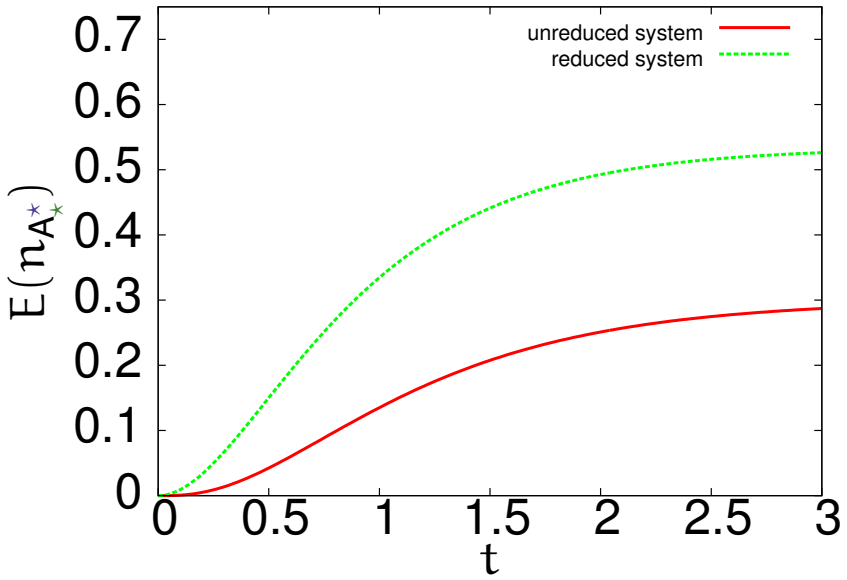


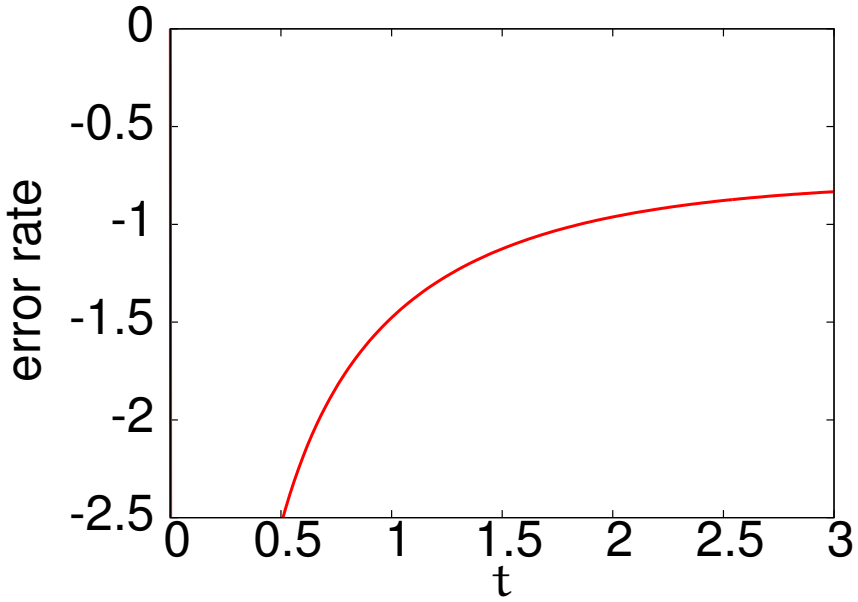
$$A + A^\star \xrightarrow{k_+} A_\star + A^\star$$



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

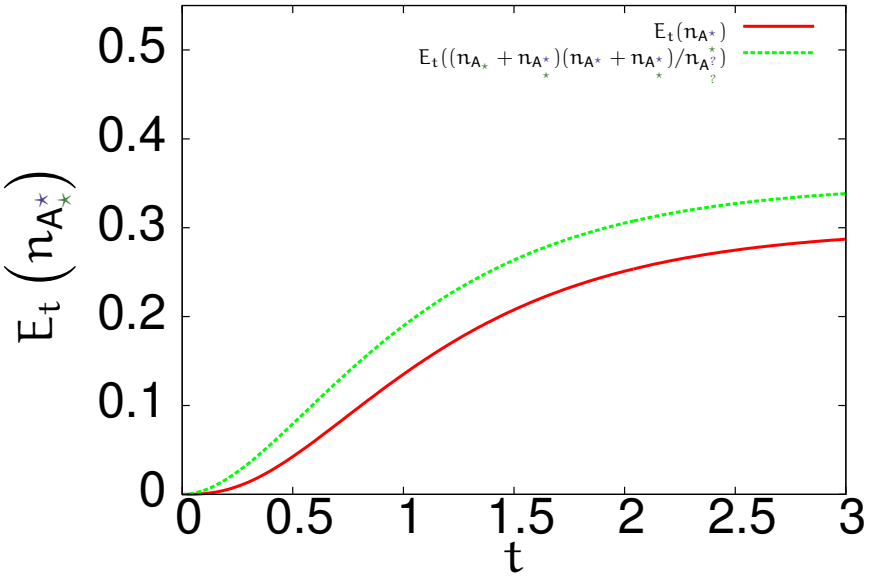
Comparison between the two models

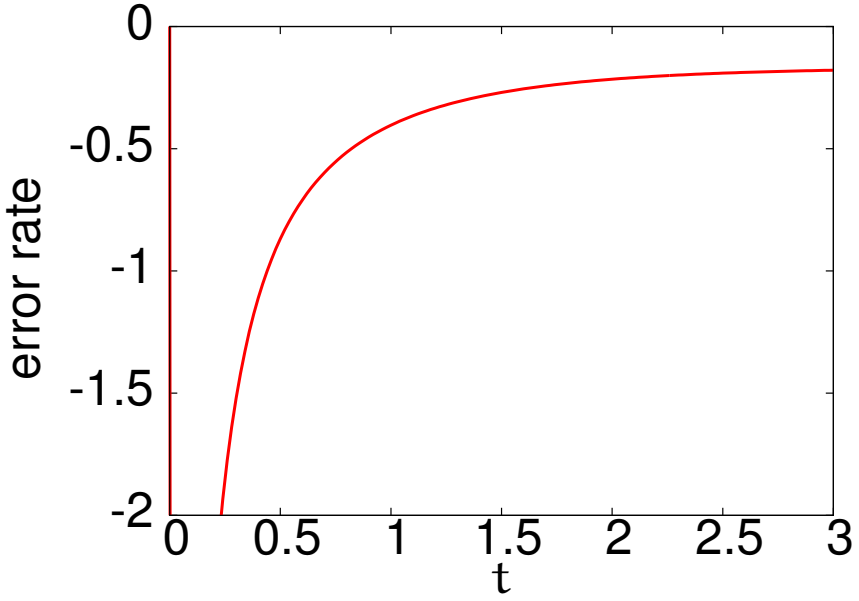




with $k^+ = k^- = k_+ = k_- = 1$,
and two instances of A at time $t = 0$.

Degree of correlation (in the unreduced model)

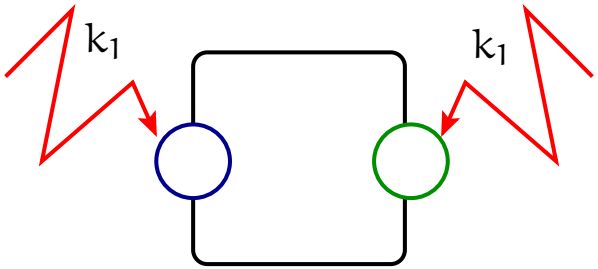




Overview

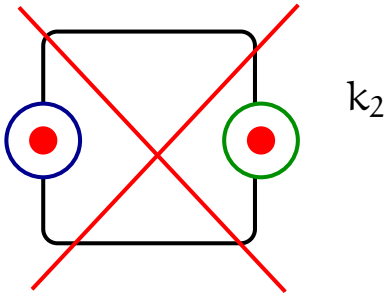
1. Introduction
2. Examples of information flow
3. Symmetric sites
4. Stochastic semantics
5. Lumpability
6. Bisimulations
7. Hierarchy of semantics
8. Conclusion

A model with symmetries



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

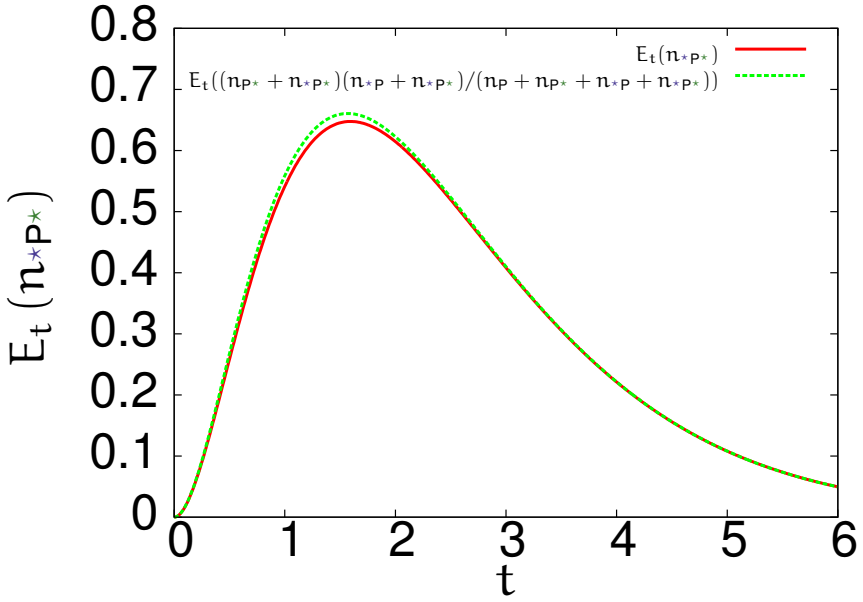
$$P \xrightarrow{k_1} P^{\star} \qquad {\star}P \xrightarrow{k_1} {\star}P^{\star}$$

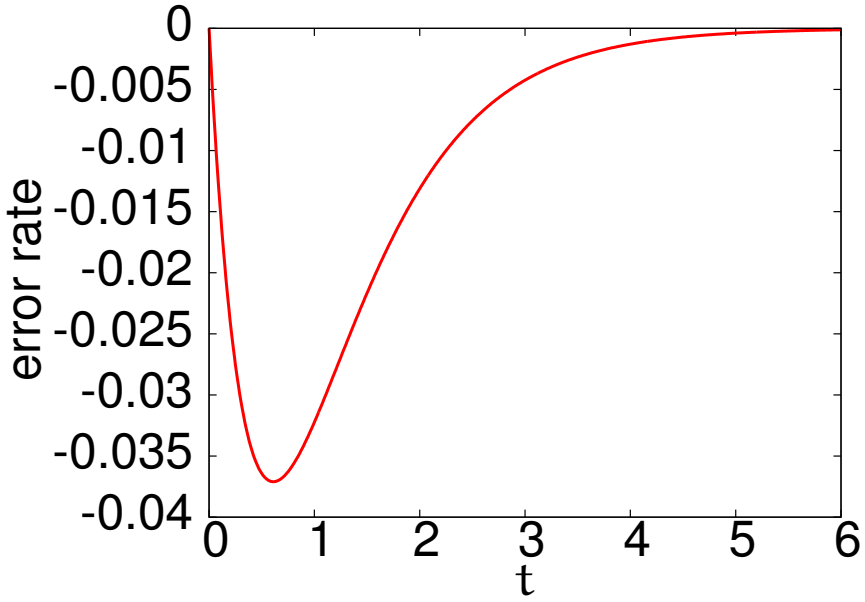


$${\star}P^{\star} \xrightarrow{k_2} \emptyset$$

Degree of correlation (in the unreduced model)

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



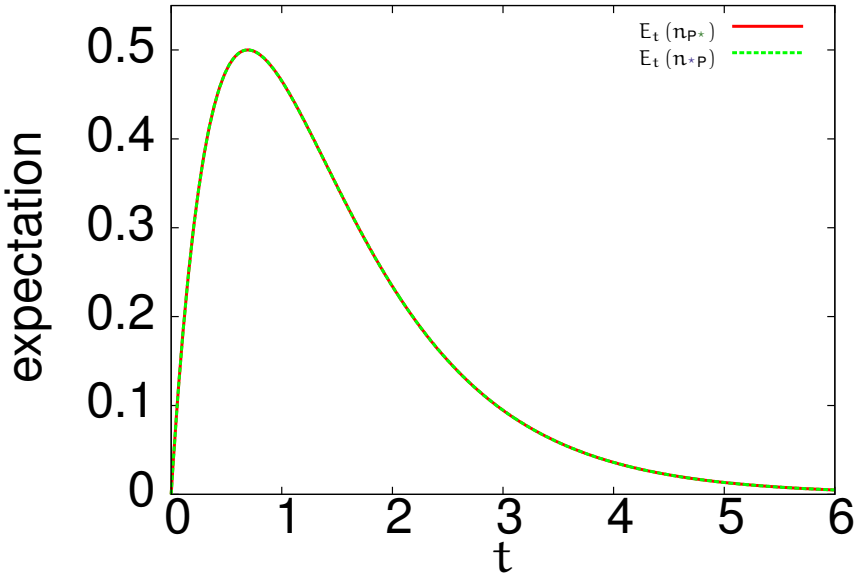


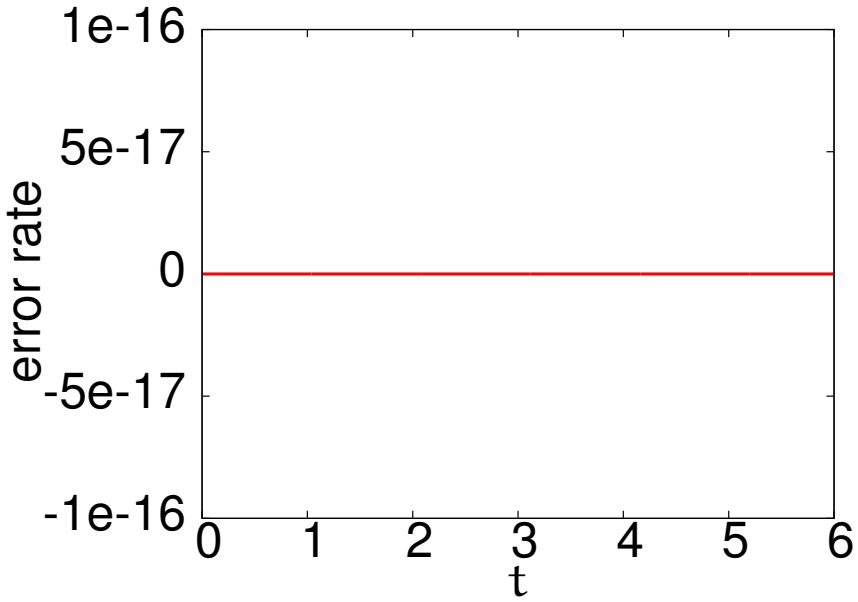
with $k_1 = k_2 = 1$
and two instances of P at time $t = 0$.

Equivalent chemical species

We check numerically that:

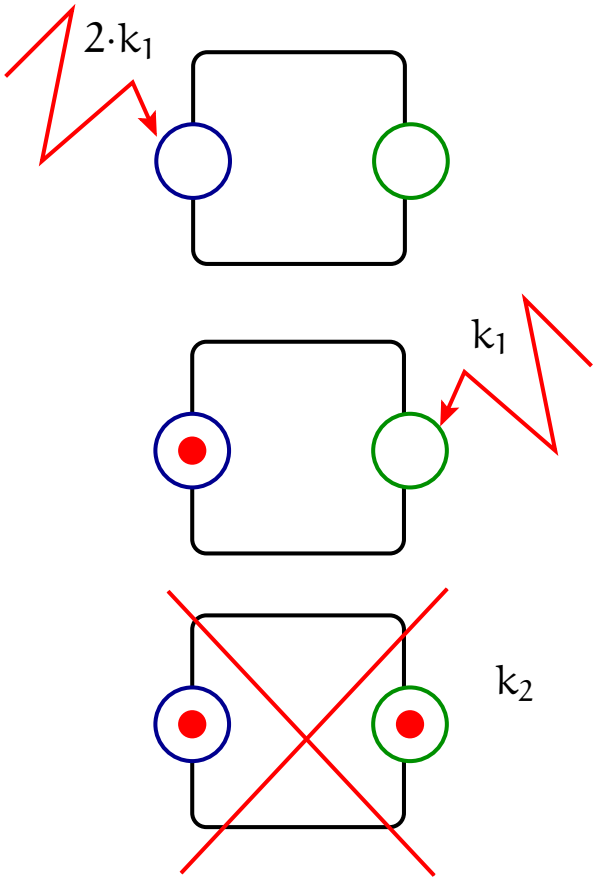
$$\mathrm{E}_t\left(n_{P\star}\right) = \mathrm{E}_t\left(n_{\star P}\right).$$





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

Reduced model



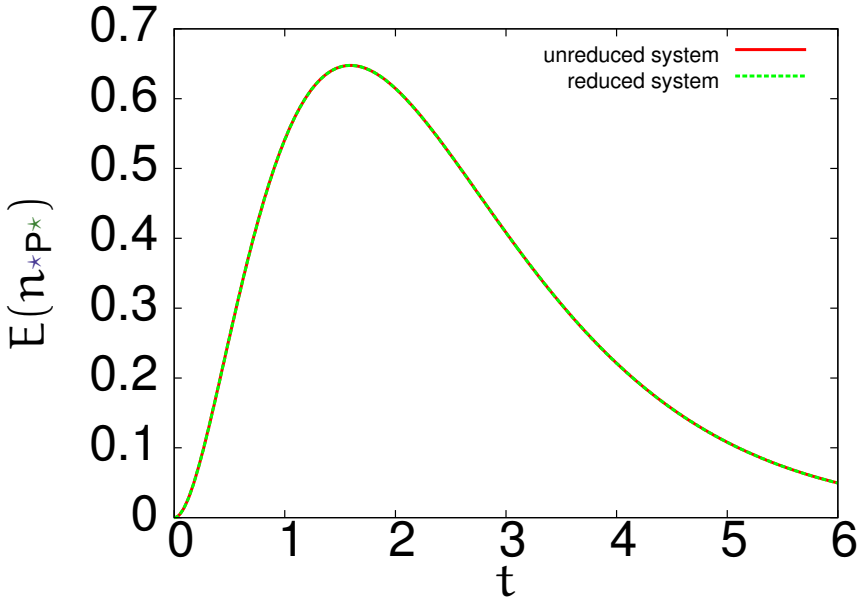
$$P \xrightarrow{2 \cdot k_1} {}^{\star}P$$

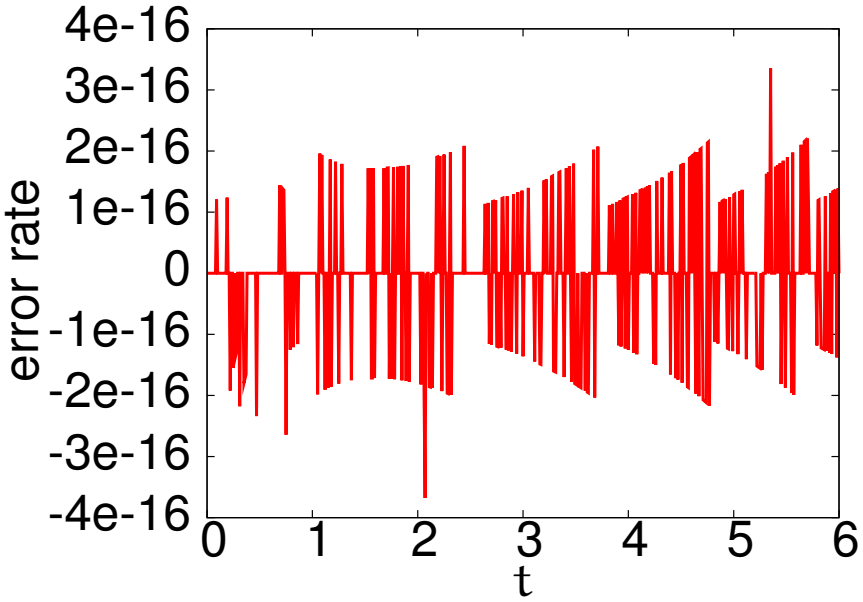
$${}^{\star}P \xrightarrow{k_1} {}^{\star}P^{\star}$$

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

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:

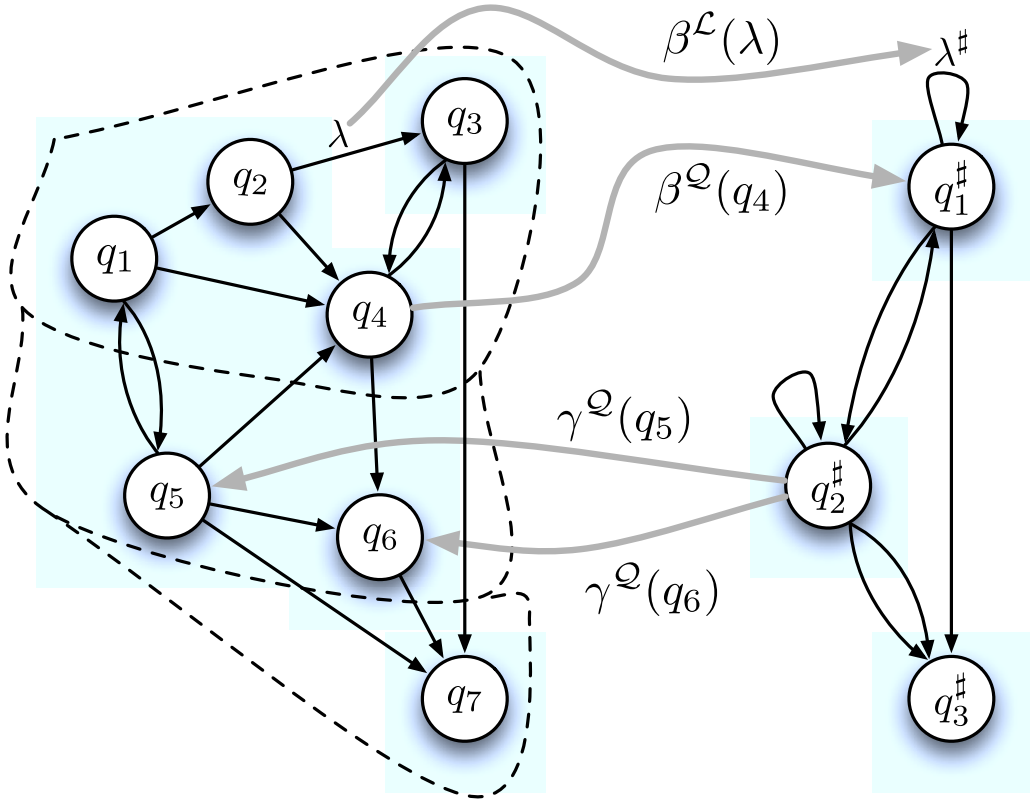
- $(q_i)_{0 \leq i \leq k} \in \mathcal{Q}^{k+1}$ and $(\lambda_i)_{1 \leq i \leq k} \in \mathcal{L}^k$,
- $(I_i)_{1 \leq i \leq k}$ is a family of open intervals in $\mathbb{R}_0^+$.

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

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

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:

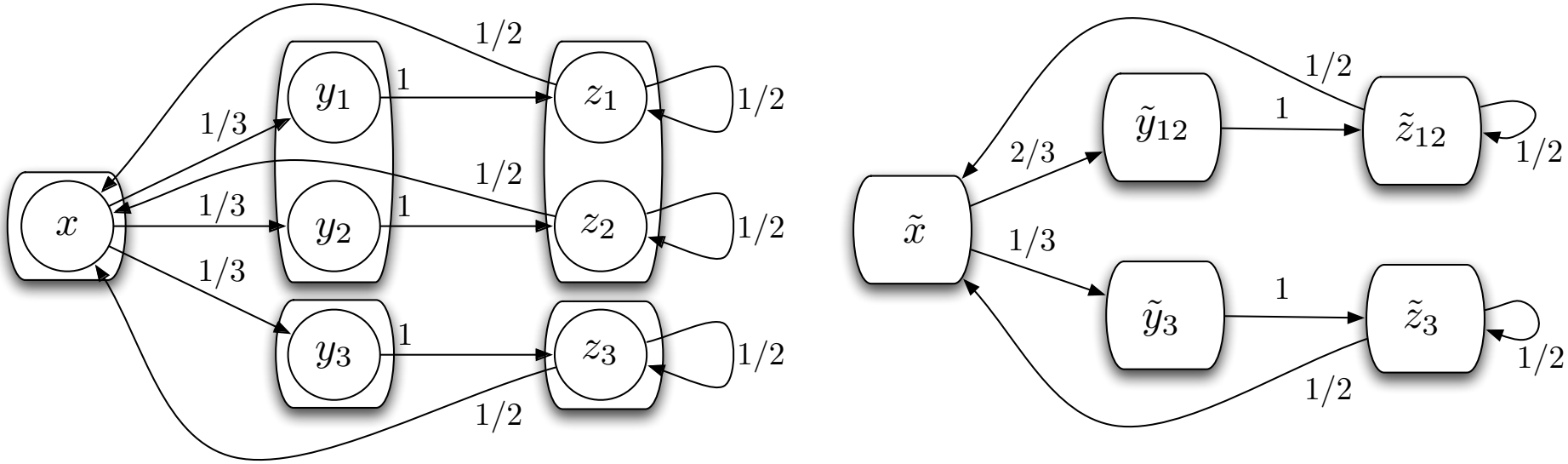
- a stochastic process $(\mathcal{X}_t)$ which satisfies the Markovian property,
- an initial distribution $\pi_0 : \mathcal{Q} \to [0, 1]$,
- an equivalence relation $\sim$ over $\mathcal{Q}$,

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.

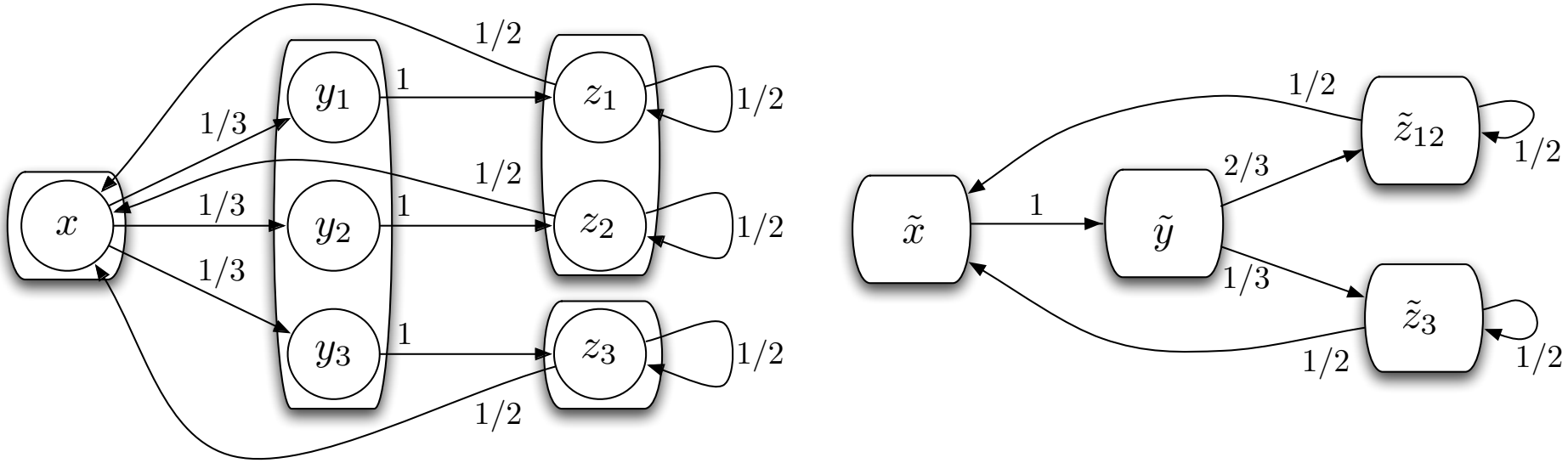
Strong lumpability



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

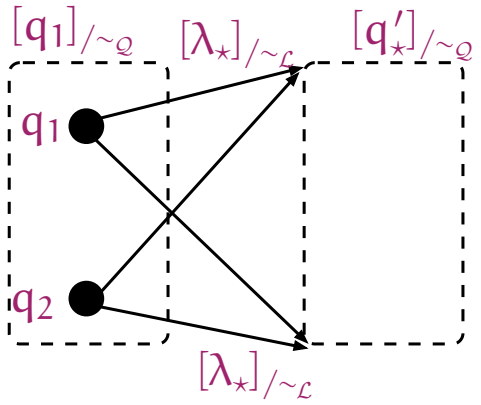
1. Introduction
2. Examples of information flow
3. Symmetric sites
4. Stochastic semantics
5. Lumpability
6. Bisimulations
7. Hierarchy of semantics
8. Conclusion

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

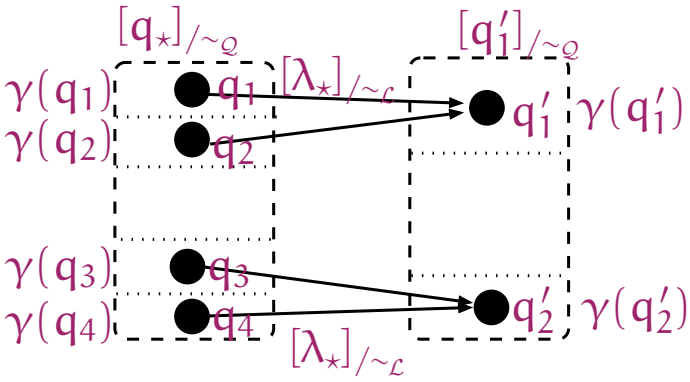


where: $\mathsf{fwd}(q, [\lambda_\star]_{/\sim_{\mathcal{L}}}, [q'_\star]_{/\sim_{\mathcal{Q}}}) = \sum_{\lambda', q'} (w(q, \lambda', q') \mid \lambda' \sim_{\mathcal{L}} \lambda_\star,\ q' \sim_{\mathcal{Q}} q'_\star)$.

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,

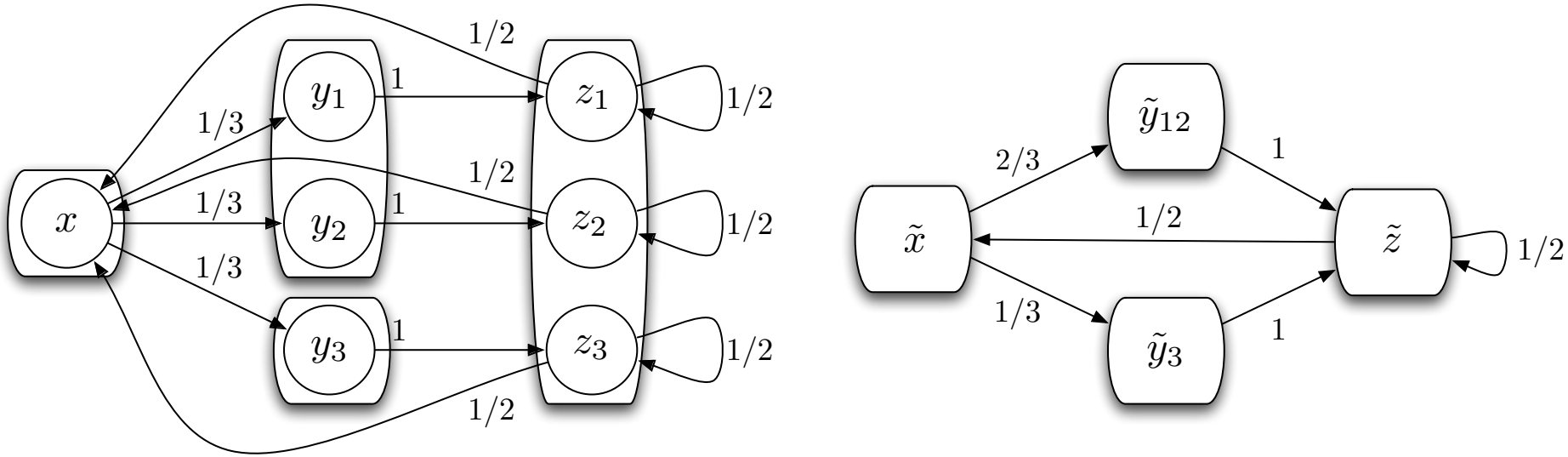
  moreover, it induces a sound abstraction;
- if $(\sim_{\mathcal{Q}}, \sim_{\mathcal{L}})$ is a backward bisimulation, then the process is $\sim_{\mathcal{Q}}$-weakly lumpable, for the initial distributions which satisfy:

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

  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.

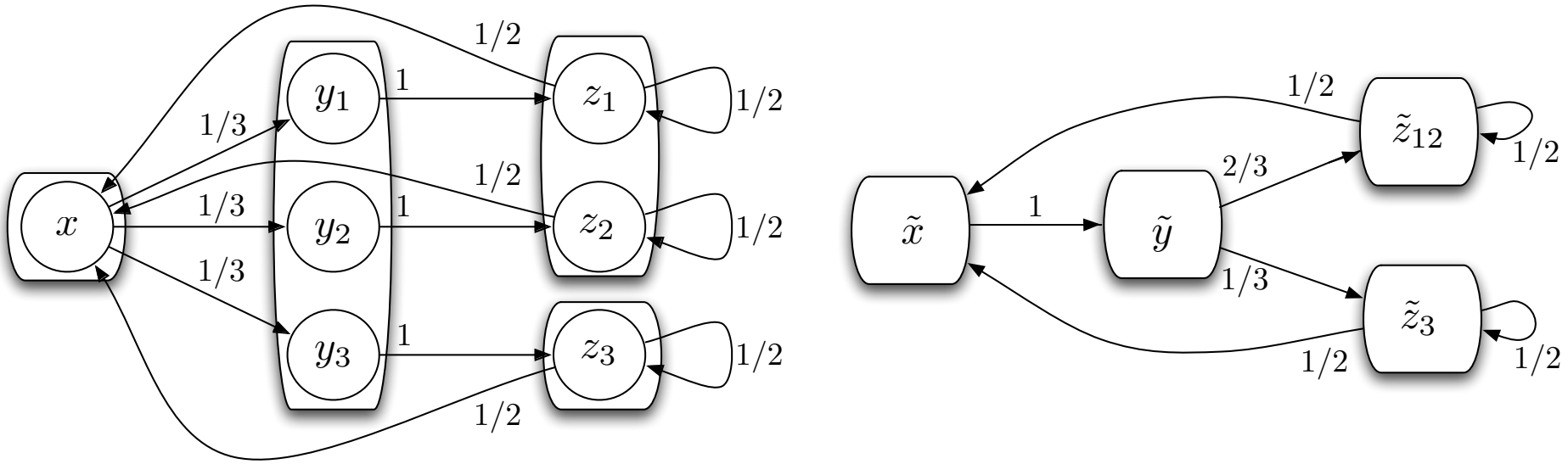
Counter-example I

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,

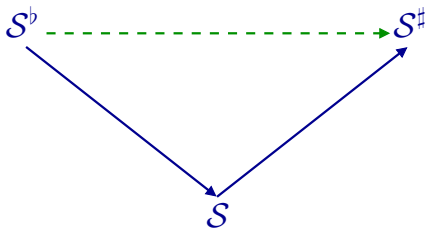
then:

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

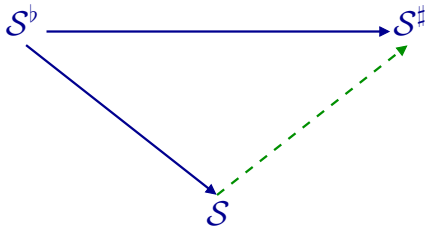
Abstraction algebra

(Sound) abstractions can be:

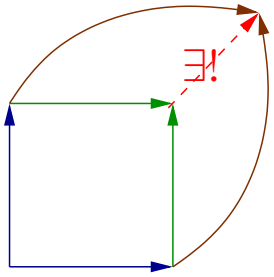
- composed:



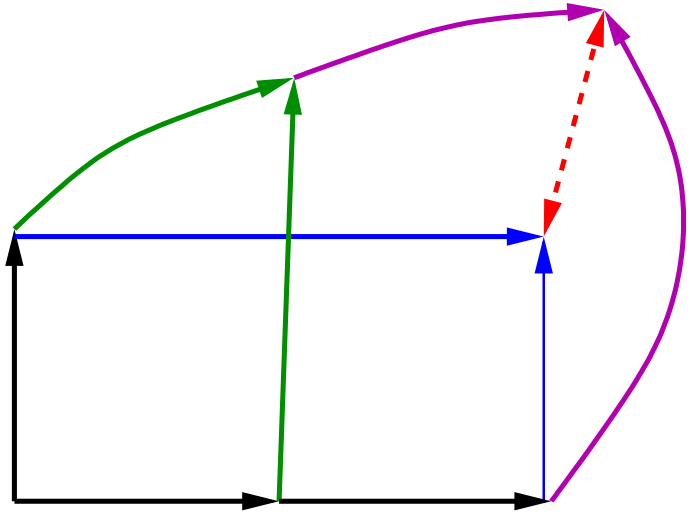
- factored:



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

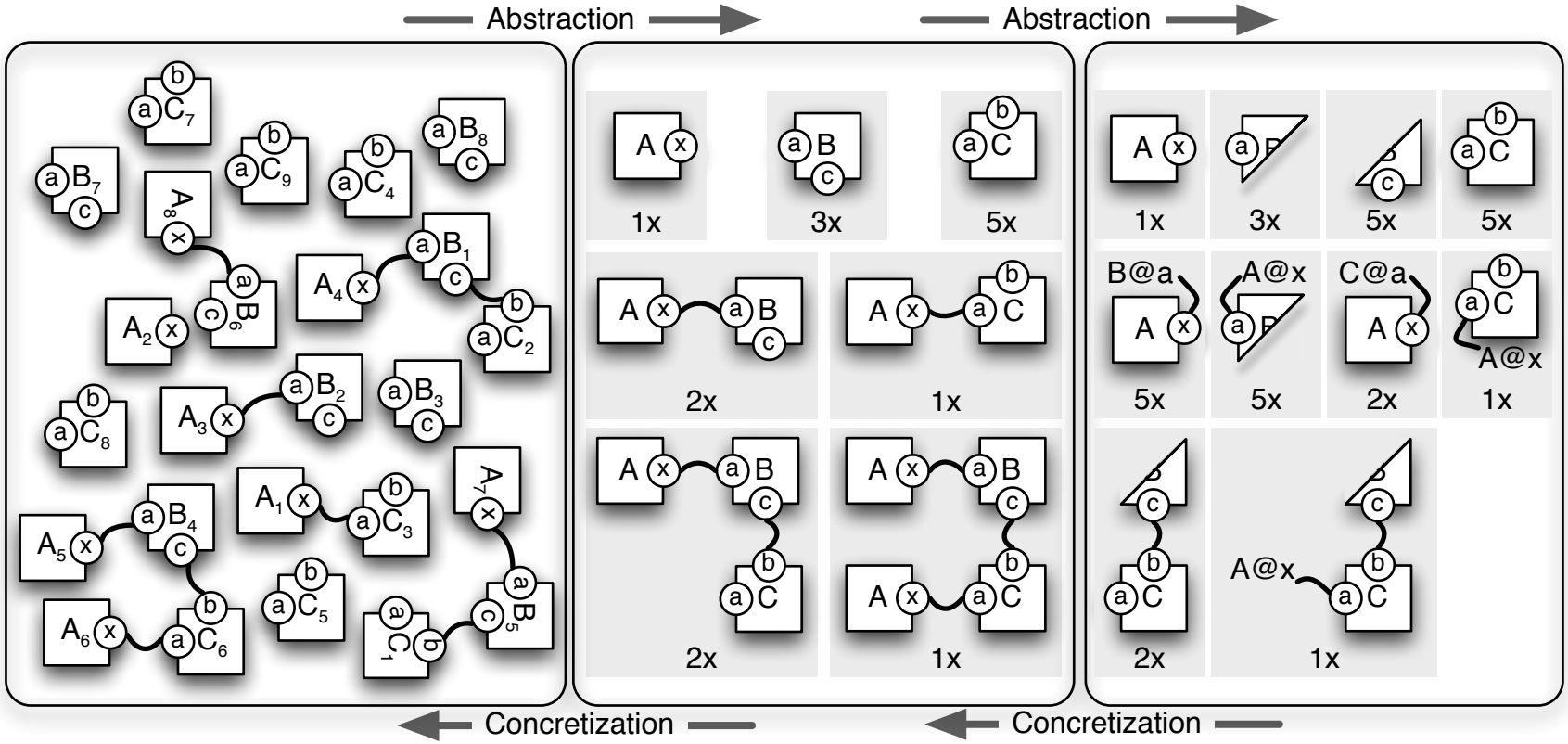


Compatibility between composition and pushout



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
5x
2x
1x
2x
1x
B@a
A@x
C@a
A@x
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:

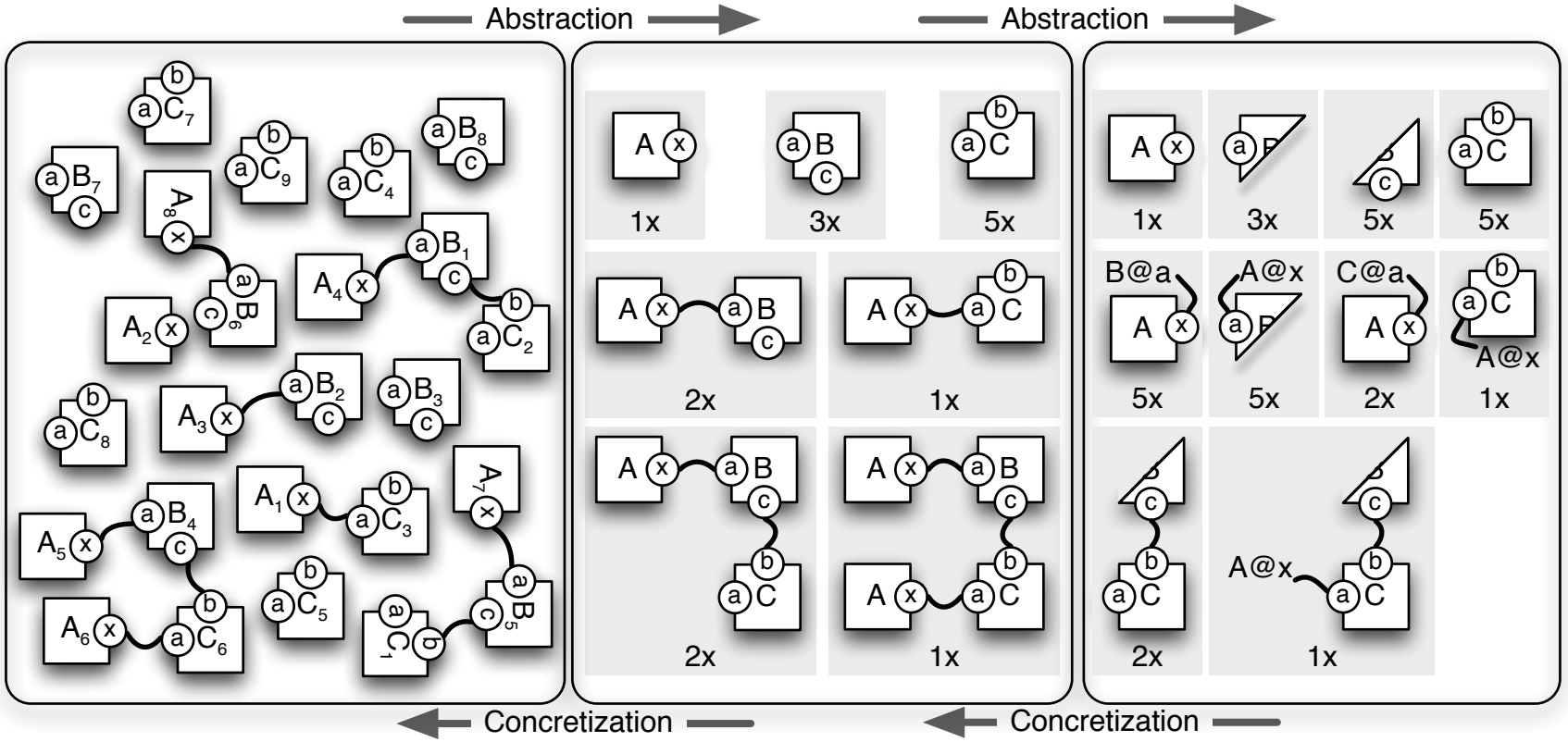
- Whenever the site x and y both occurs in the same or in distinct agent of type A in a rule, then, there should be a bidirectional edge between the site x and the y of A.
- Whenever there is a bond between two sites, each of which either carries an internal state of, is connected to some other sites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

From population to fragments

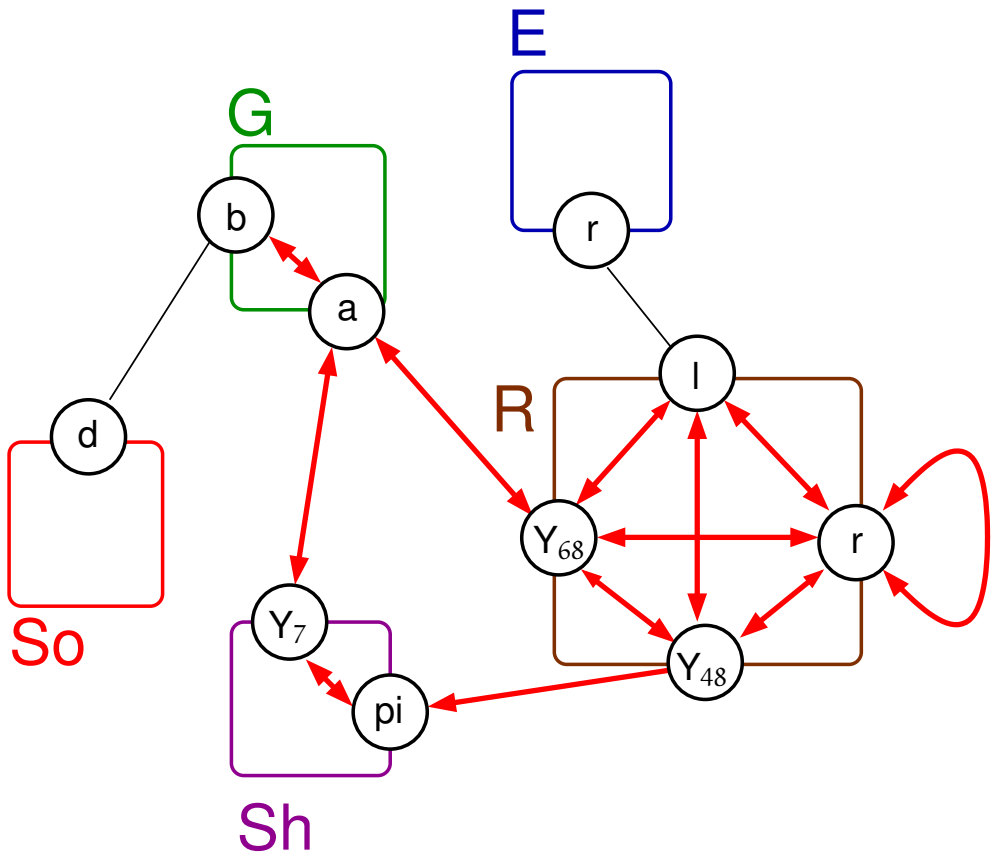
- Population of fragments:

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
5x
3x
5x
2x
1x
2x
1x
A@x
A@x
Concretization
Concretization

Example



Symmetries among sites

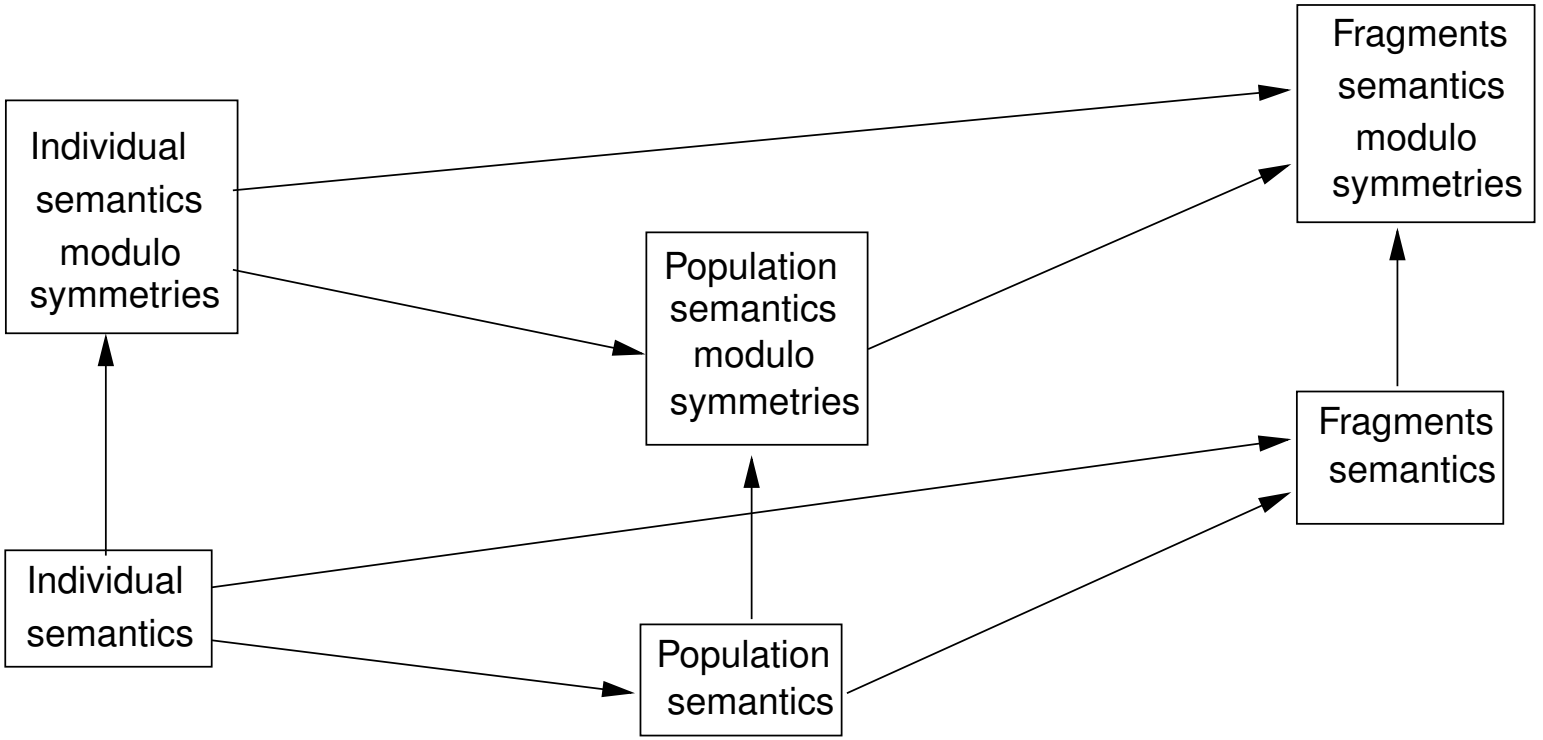
Let $\mathcal{R}$ be a set of rules and $\mathcal{M}_0$ be an initial mixture.

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,
  - ...