SBFM'12

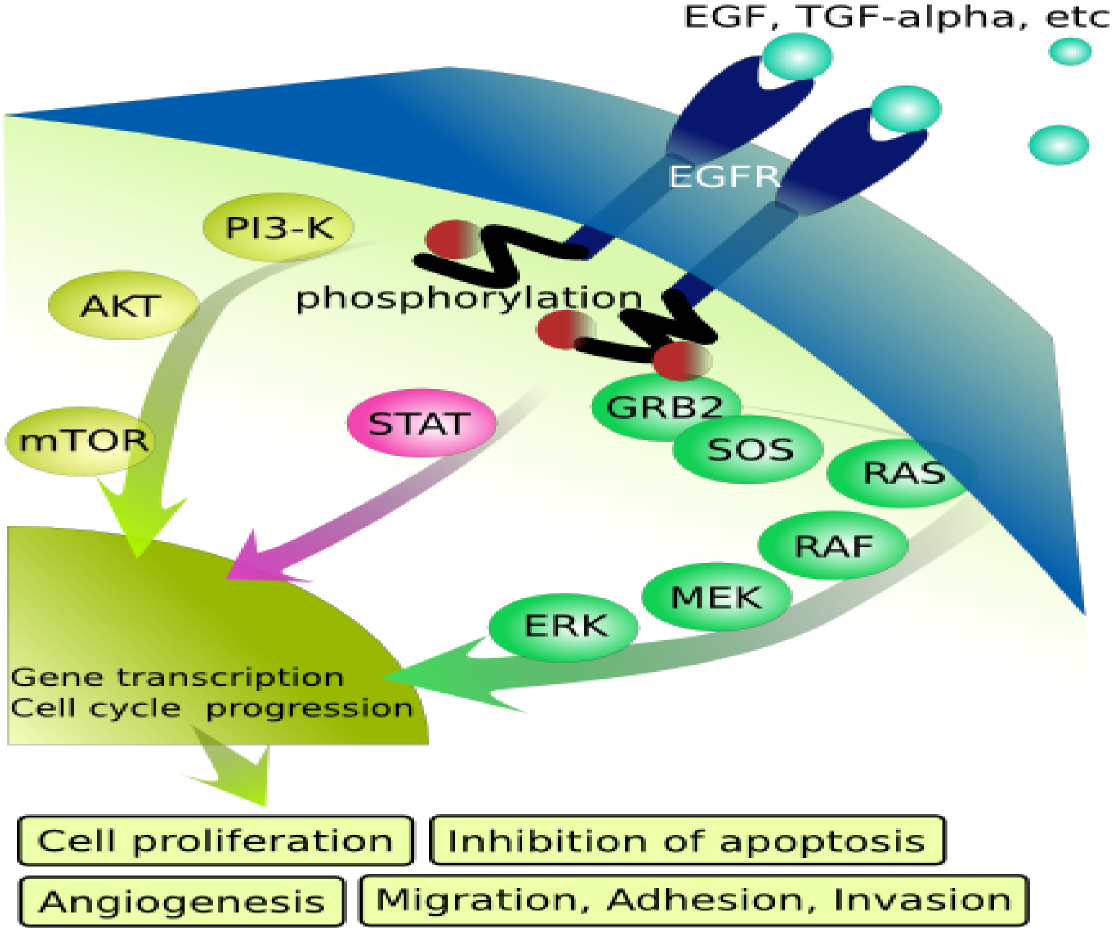
Formal model reduction

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

29 March 2012

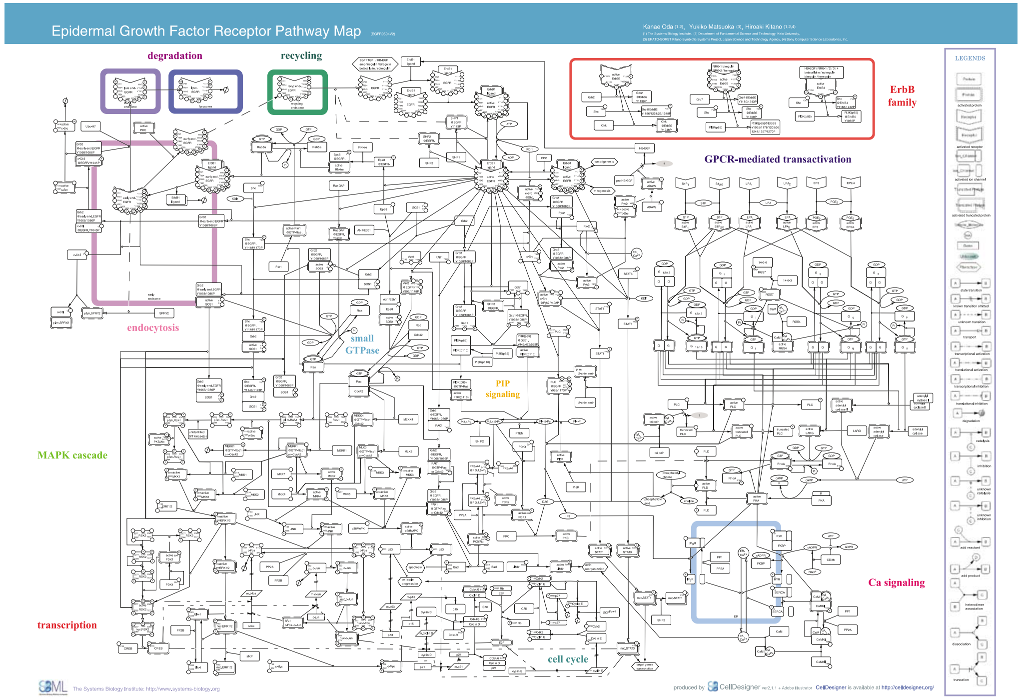
Overview

Signalling Pathways



Eikuch, 2007

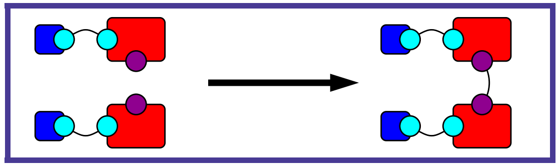
Bridge the gap between...

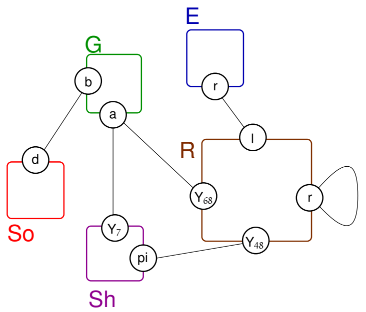


Oda, Matsuoka, Funahashi, Kitano, Molecular Systems Biology, 2005

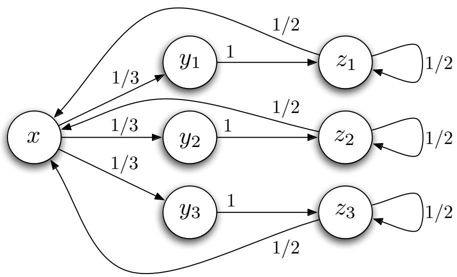
$$
\begin{cases}
\frac{dx_1}{dt} = -k_1 \cdot x_1 \cdot x_2 + k_{-1} \cdot x_3 \\
\frac{dx_2}{dt} = -k_1 \cdot x_1 \cdot x_2 + k_{-1} \cdot x_3 \\
\frac{dx_3}{dt} = k_1 \cdot x_1 \cdot x_2 - k_{-1} \cdot x_3 + 2 \cdot k_2 \cdot x_3 \cdot x_3 - k_{-2} \cdot x_4 \\
\frac{dx_4}{dt} = k_2 \cdot x_3^2 - k_2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

Rule-based models





Interaction map

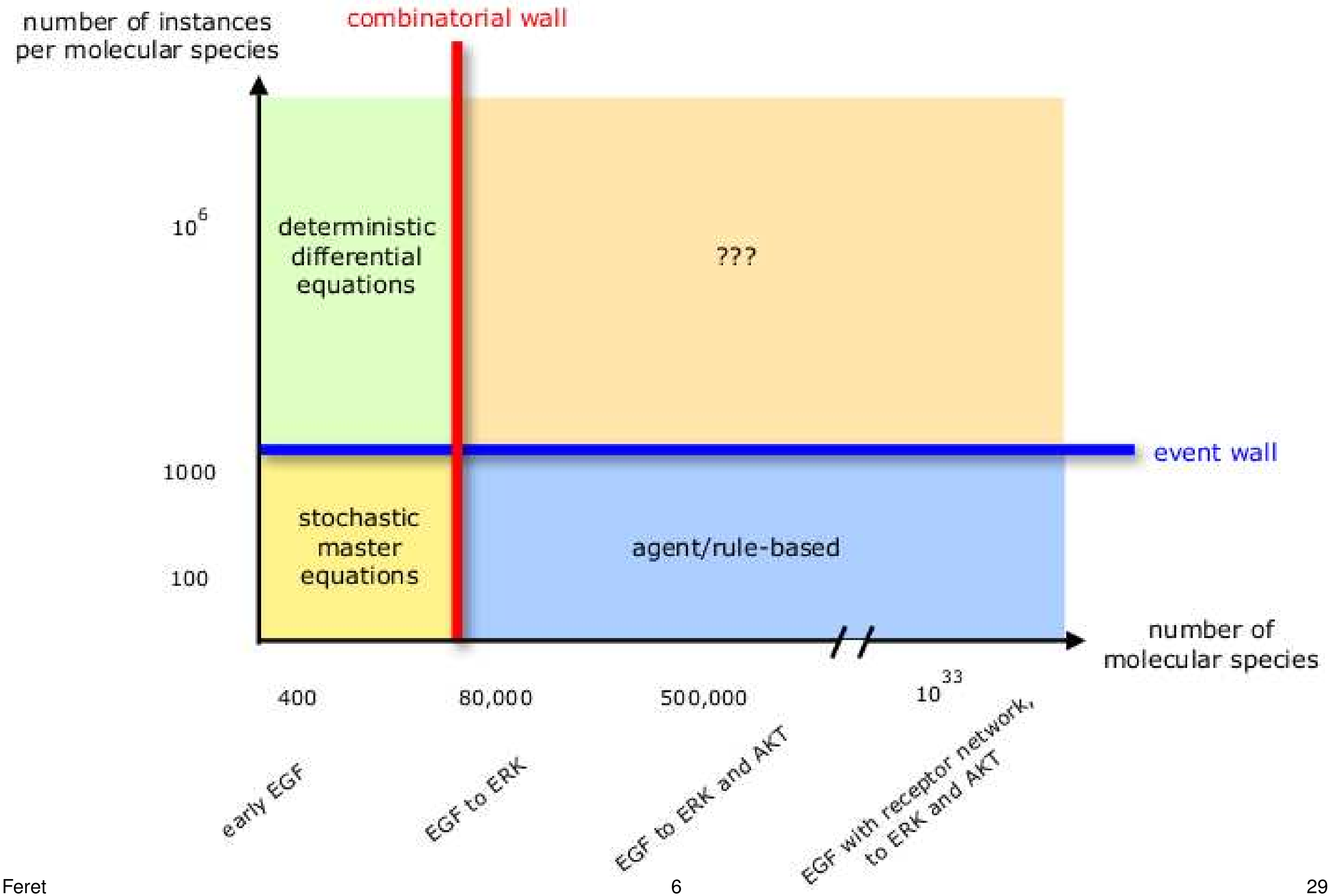


CTMC

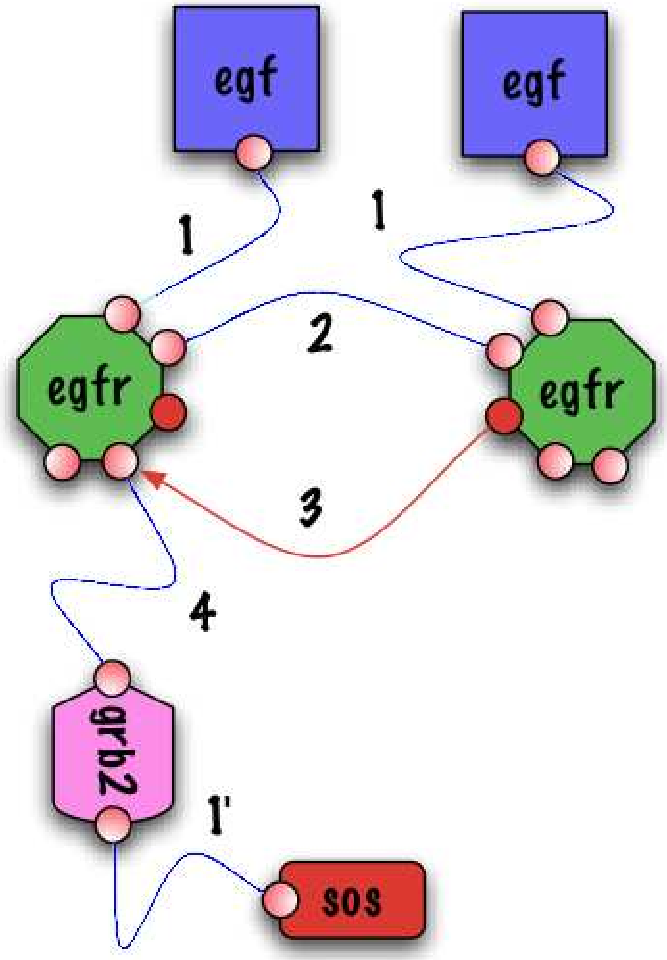
$$
\begin{cases}
\frac{dx_1}{dt} = -k_1 \cdot x_1 \cdot x_2 + k_{-1} \cdot x_3 \\
\frac{dx_2}{dt} = -k_1 \cdot x_1 \cdot x_2 + k_{-1} \cdot x_3 \\
\frac{dx_3}{dt} = k_1 \cdot x_1 \cdot x_2 - k_{-1} \cdot x_3 + 2 \cdot k_2 \cdot x_3 \cdot x_3 - k_{-2} \cdot x_4 \\
\frac{dx_4}{dt} = k_2 \cdot x_3^2 - k_2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

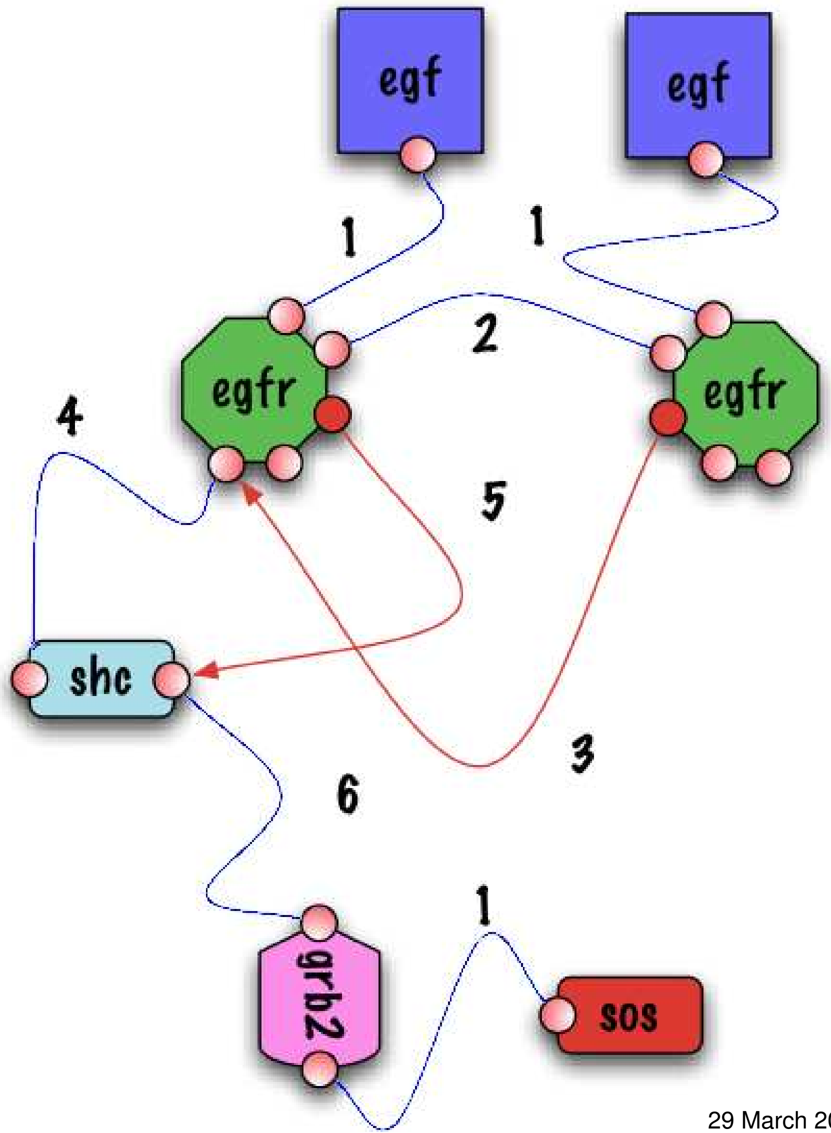
ODEs

Complexity walls



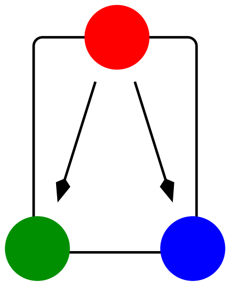
A breach in the wall(s) ?



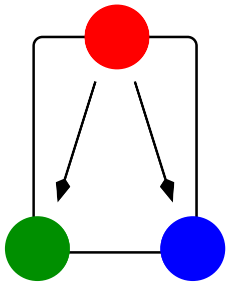


Overview

A system with a switch

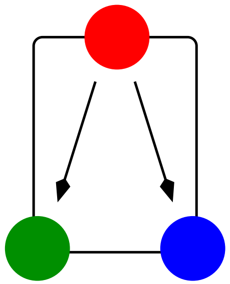


A system with a switch



(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

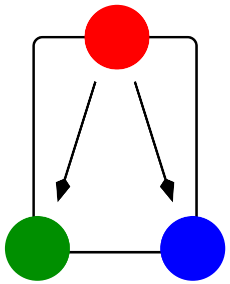
A system with a switch



$$(u,u,u) \longrightarrow (u,p,u) \quad k^c$$
$$(u,p,u) \longrightarrow (p,p,u) \quad k^l$$
$$(u,p,p) \longrightarrow (p,p,p) \quad k^l$$
$$(u,p,u) \longrightarrow (u,p,p) \quad k^r$$
$$(p,p,u) \longrightarrow (p,p,p) \quad k^r$$

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

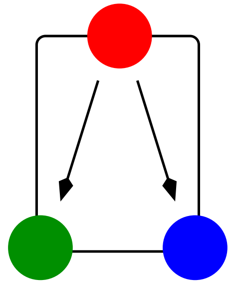
A system with a switch



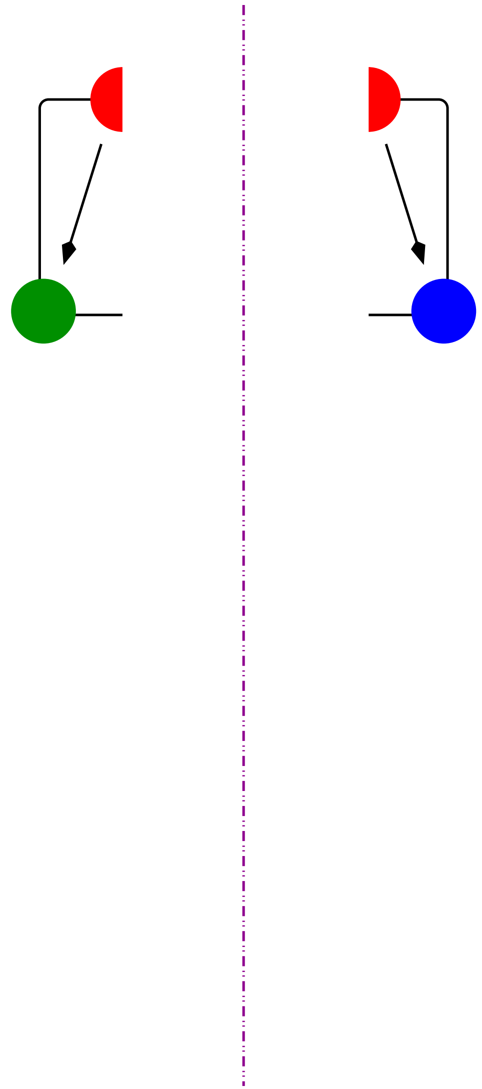
(u,u,u)	$\longrightarrow$	(u,p,u)	k^c
(u,p,u)	$\longrightarrow$	(p,p,u)	k^l
(u,p,p)	$\longrightarrow$	(p,p,p)	k^l
(u,p,u)	$\longrightarrow$	(u,p,p)	k^r
(p,p,u)	$\longrightarrow$	(p,p,p)	k^r

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

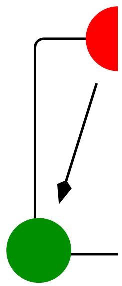
Two subsystems

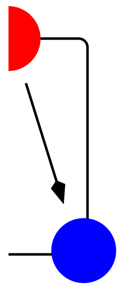


Two subsystems



Two subsystems





$$[(u,u,u)] = [(u,u,u)]$$
$$[(u,p,?)] \triangleq [(u,p,u)] + [(u,p,p)]$$
$$[(p,p,?)] \triangleq [(p,p,u)] + [(p,p,p)]$$

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

$$[(u,u,u)] = [(u,u,u)]$$
$$[(?,p,u)] \triangleq [(u,p,u)] + [(p,p,u)]$$
$$[(?,p,p)] \triangleq [(u,p,p)] + [(p,p,p)]$$

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

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:

$$X \triangleq [(p,p,p)] \cdot ([(?,p,u)] + [(?,p,p)]) - [(?,p,p)] \cdot [(p,p,?)].$$

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.

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
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 $\mathcal{V} \rightarrow \mathbb{R}^+$ into $\mathcal{V} \rightarrow \mathbb{R}$,
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 $X_0 \in \mathcal{V} \rightarrow \mathbb{R}^+$ to the maximal solution $X_{X_0} \in [0, T^{max}_{X_0}[\rightarrow (\mathcal{V} \rightarrow \mathbb{R}^+)$ which satisfies:

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

Overview

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 $\mathcal{V} \rightarrow \mathbb{R}$ into $\mathcal{V}^\sharp \rightarrow \mathbb{R}$,
- $\mathbb{F}^\sharp$: a $\mathcal{C}^\infty$ mapping from $\mathcal{V}^\sharp \rightarrow \mathbb{R}^+$ into $\mathcal{V}^\sharp \rightarrow \mathbb{R}$;

such that:

- ψ is linear with positive coefficients,
- the following diagram commutes:

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

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

Abstraction example

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

(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

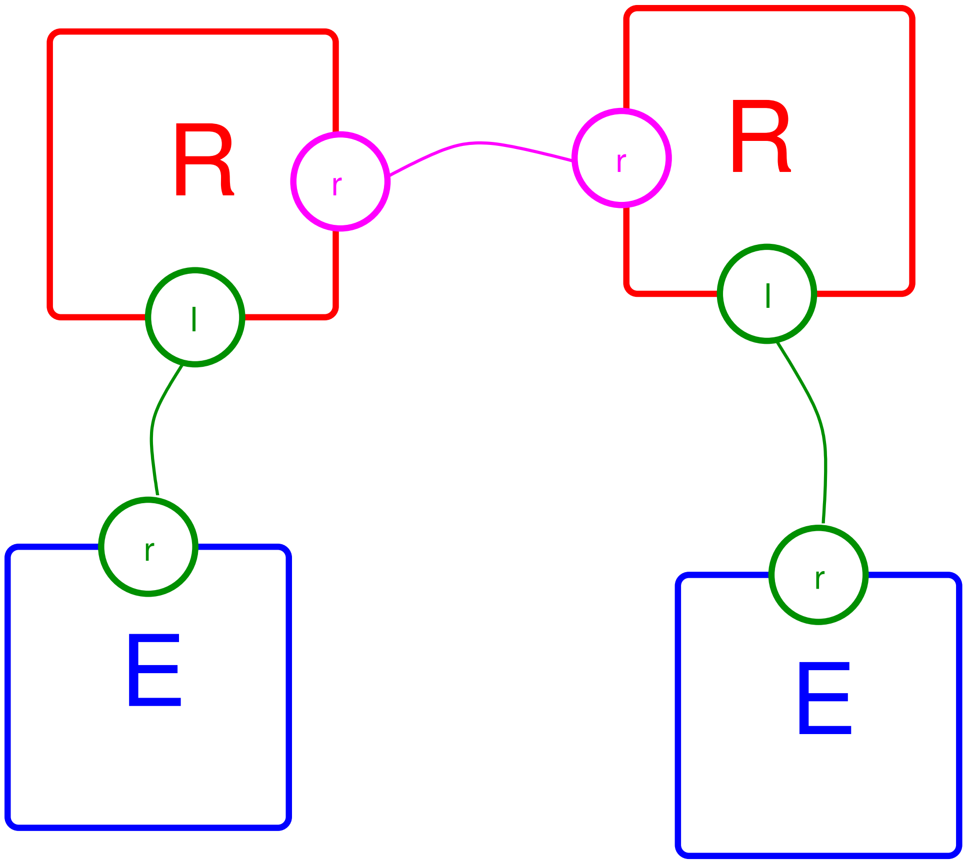
$$X_{X_0}(T) = X_0 + \int_{t=0}^{T} \mathbb{F} \left(X_{X_0}(t) \right) \cdot dt.$$

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

Overview

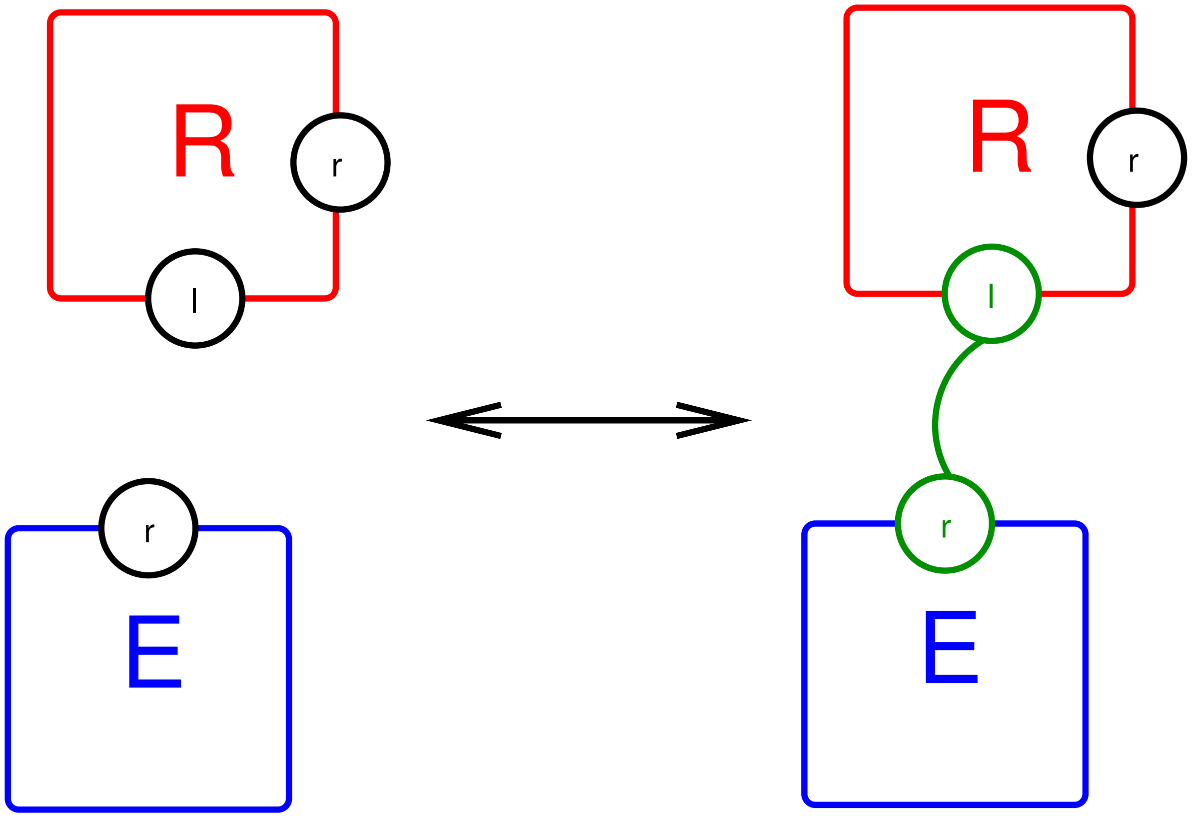
1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
7. Conclusion

A species



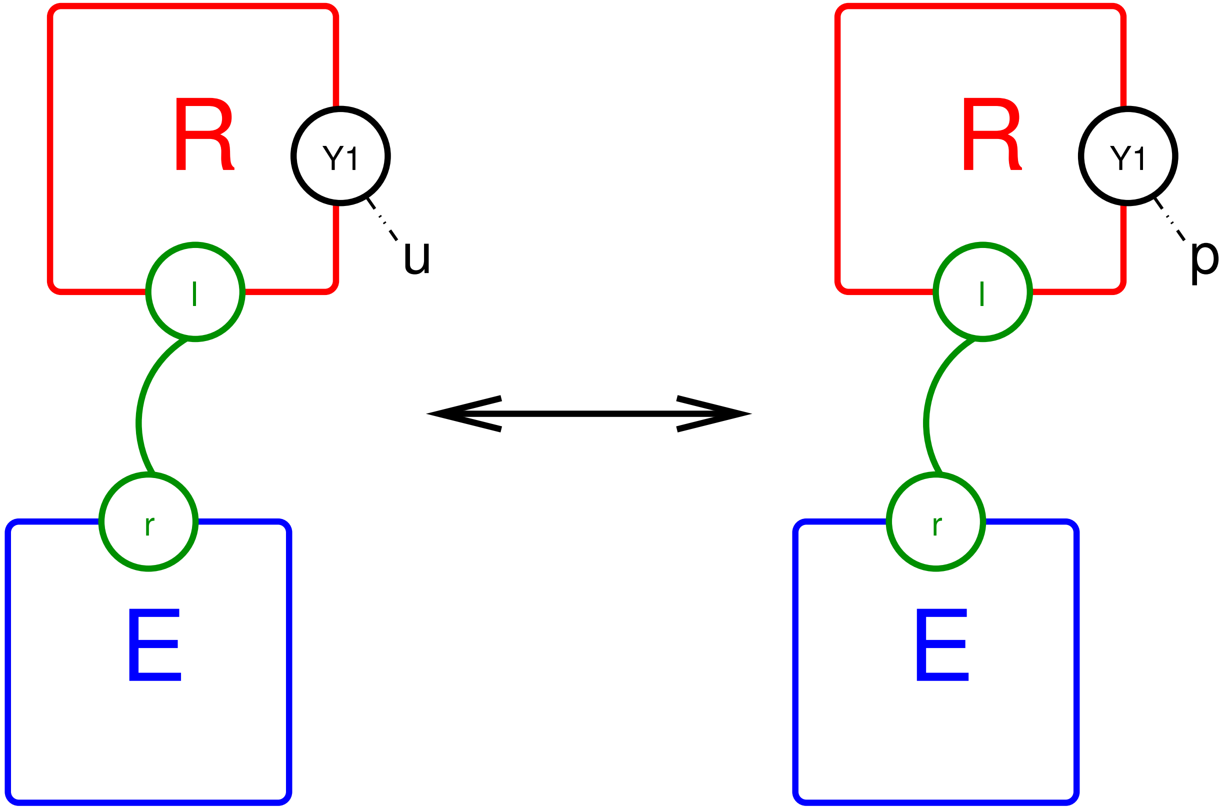
E(r!1), R(l!1,r!2), R(r!2,l!3), E(r!3)

A Unbinding/Binding Rule



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

Internal state



R(Y1~u,l!1), E(r!1) $\longleftrightarrow$ R(Y1~p,l!1), E(r!1)

Overview

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:

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

where $\mathrm{SYM}(E)$ 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) Flow of information
   - (c) Abstract counterpart
7. Conclusion

Abstract domain

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

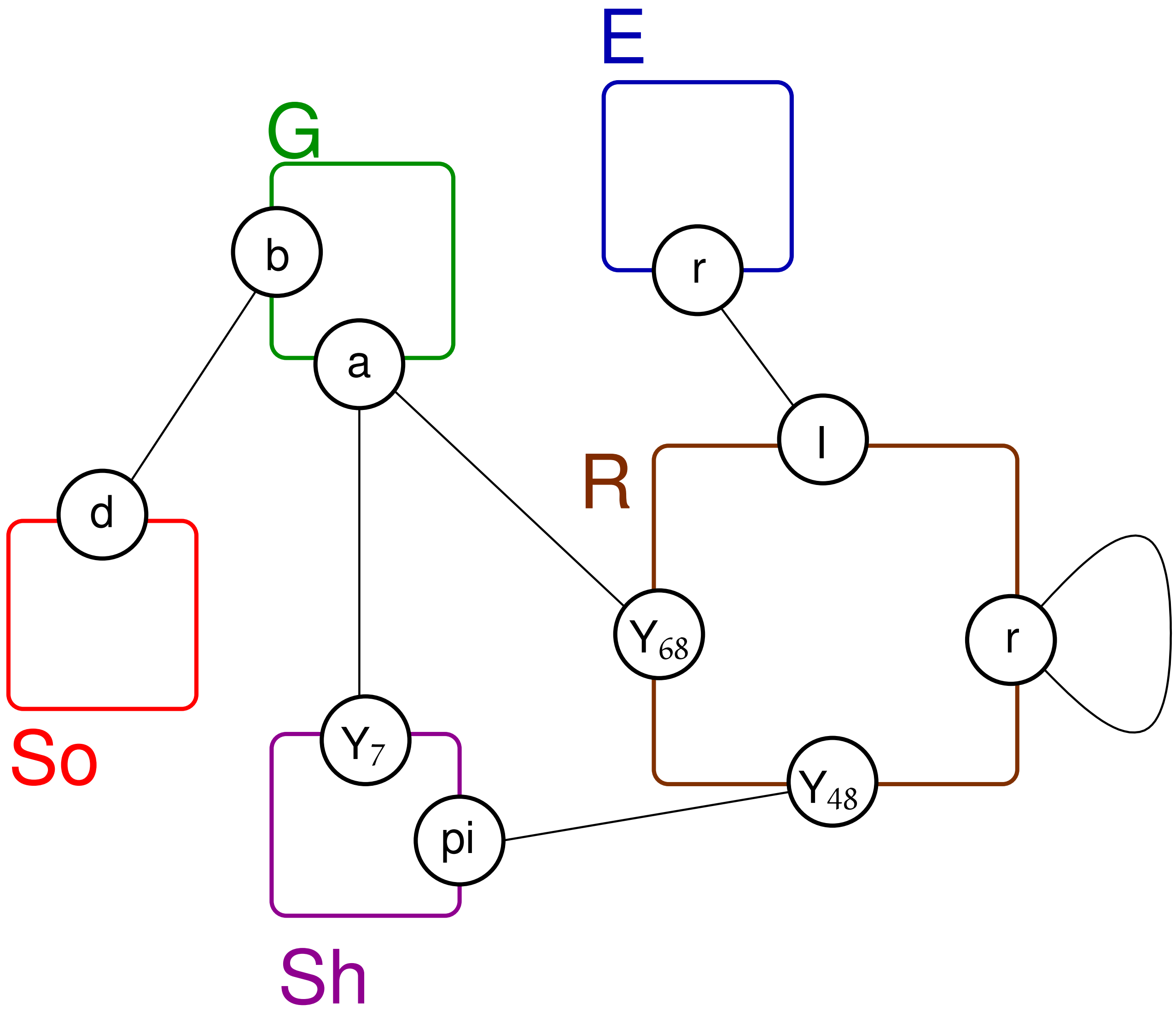
The set of linear variable replacements is too big to be explored.

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

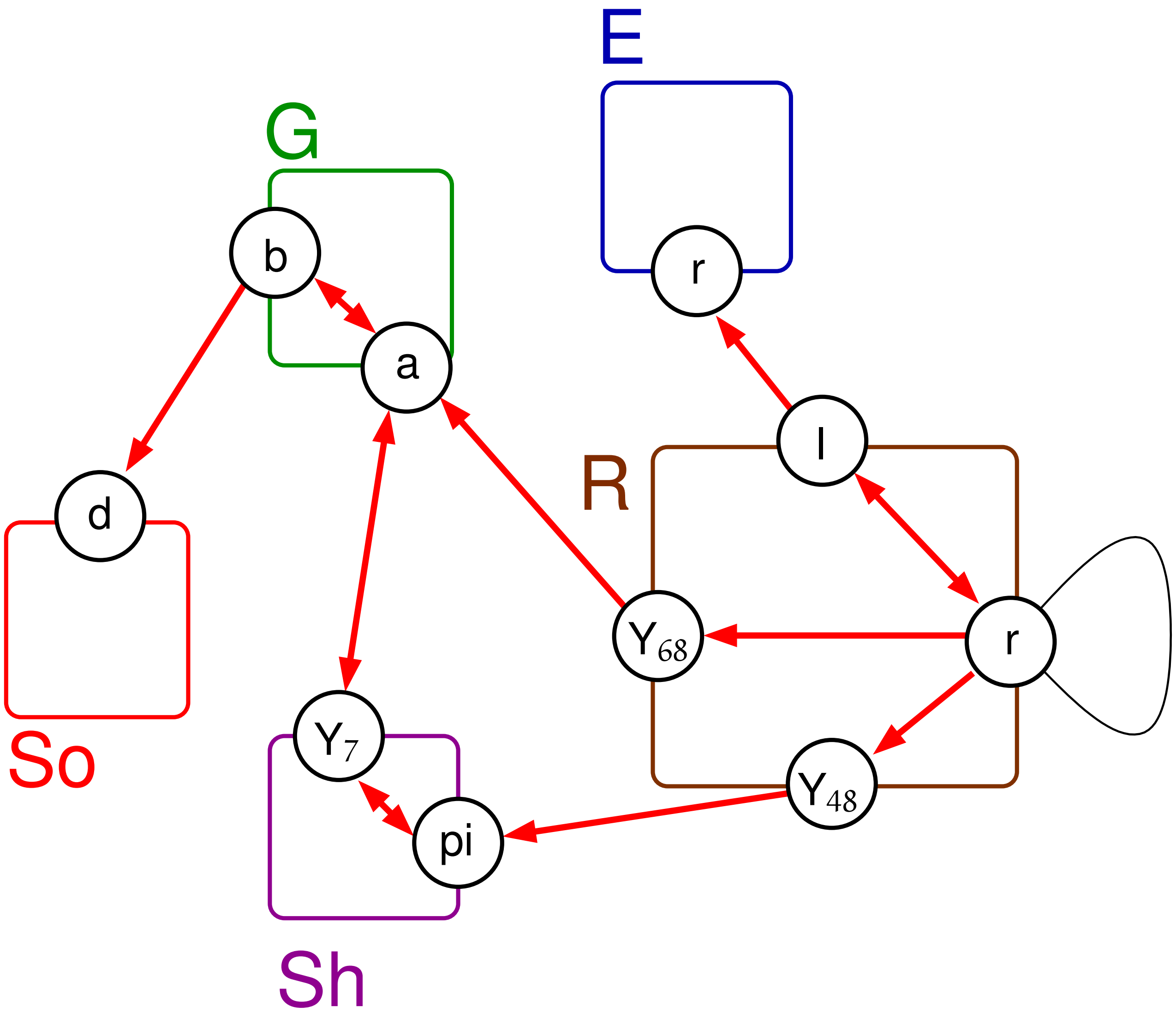
- restrict the exploration;
- drive the intuition (by using the flow of information);
- having efficient way to find suitable abstractions $(\mathcal{V}^\sharp, \psi)$ and to compute $\mathbb{F}^\sharp$.

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

Contact map



Annotated contact map



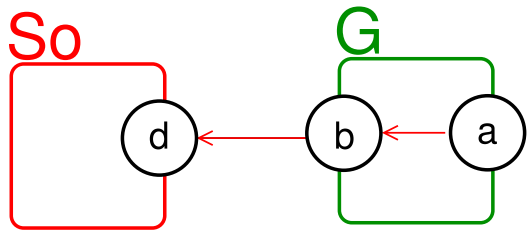
Fragments and prefragments

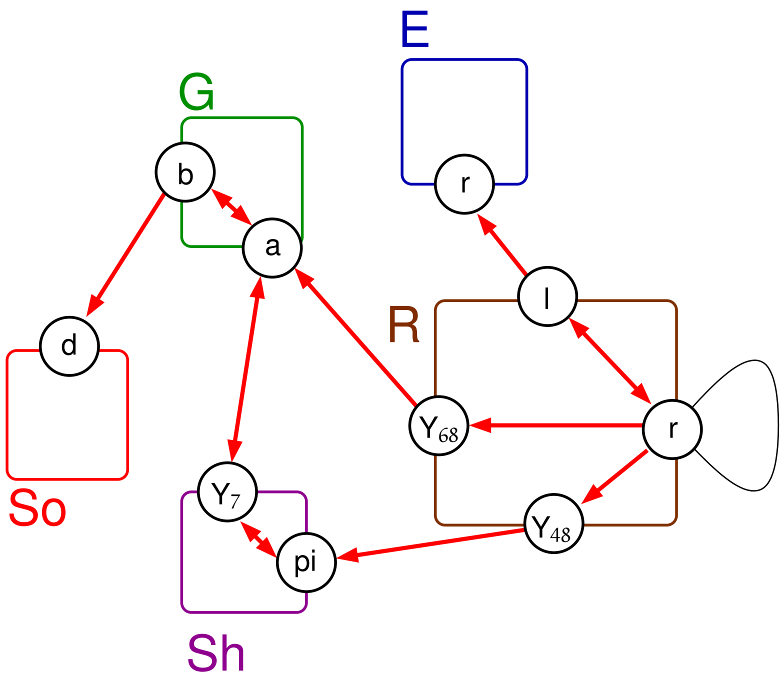
A prefragment is a connected site graph for which there exists a binary relations $\to$ between sites such that:

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

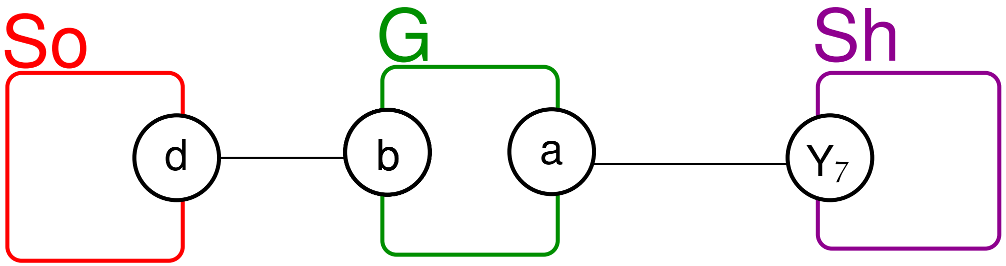
A fragment is a prefragment F such that:

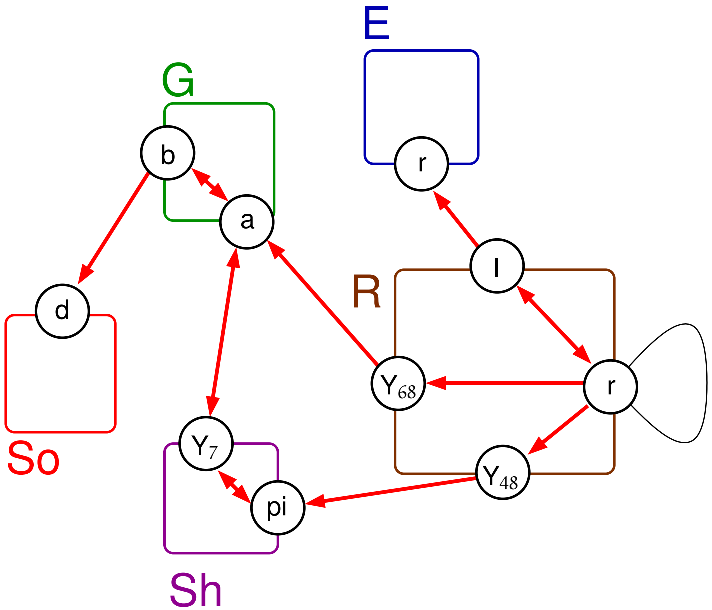
- Parsimoniousness: for any prefragment F' such that F embeds in F', F' also embeds into F.



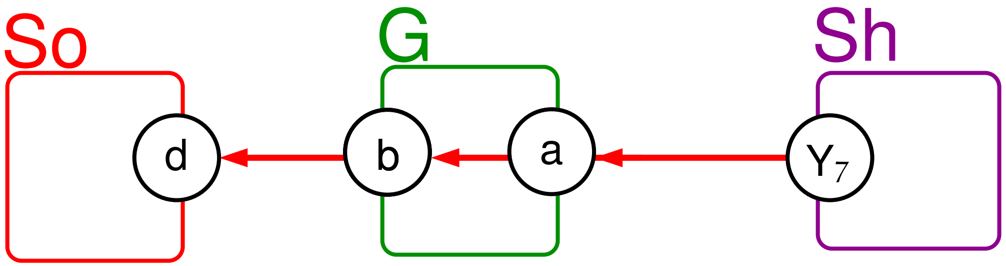


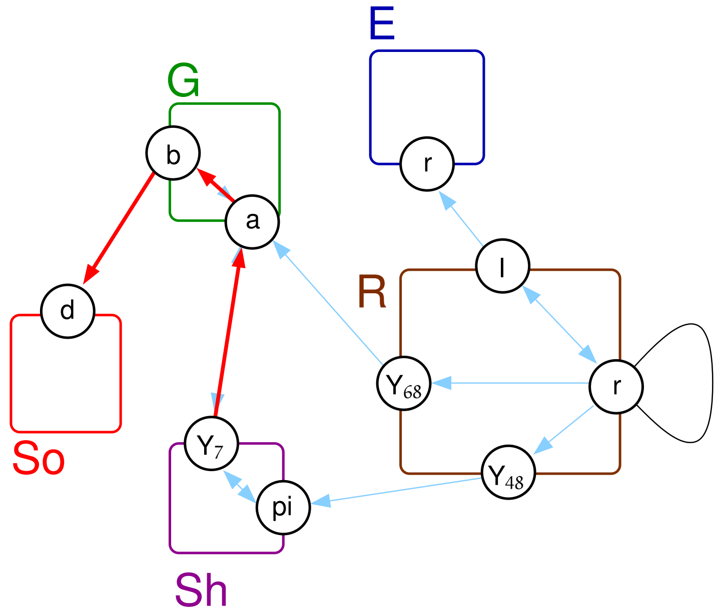
Are they fragments ?





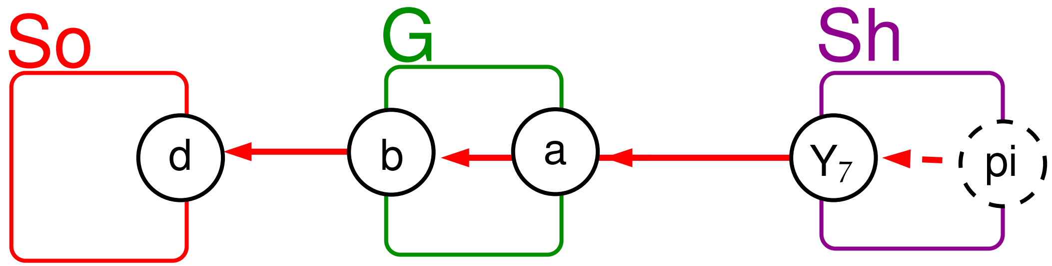
Are they fragments ?



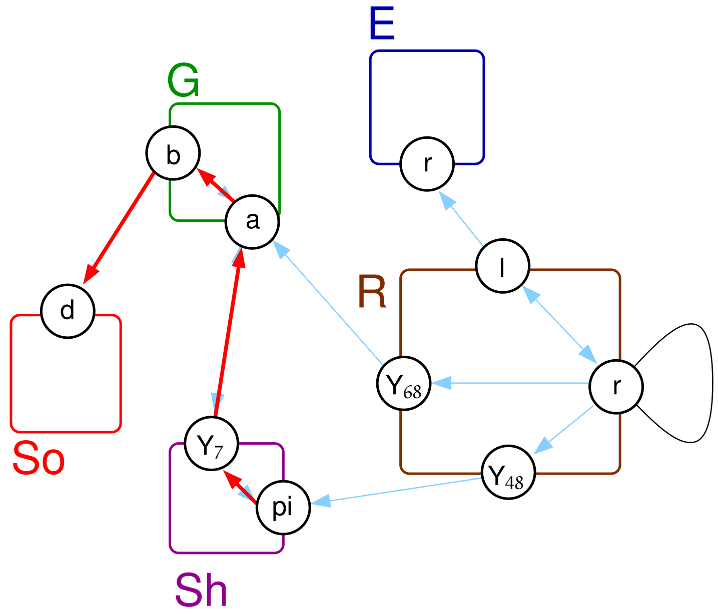


Thus, it is a prefragment.

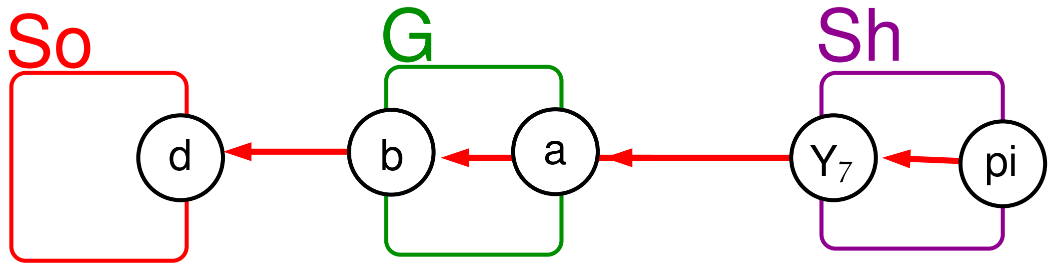
Are they fragments ?

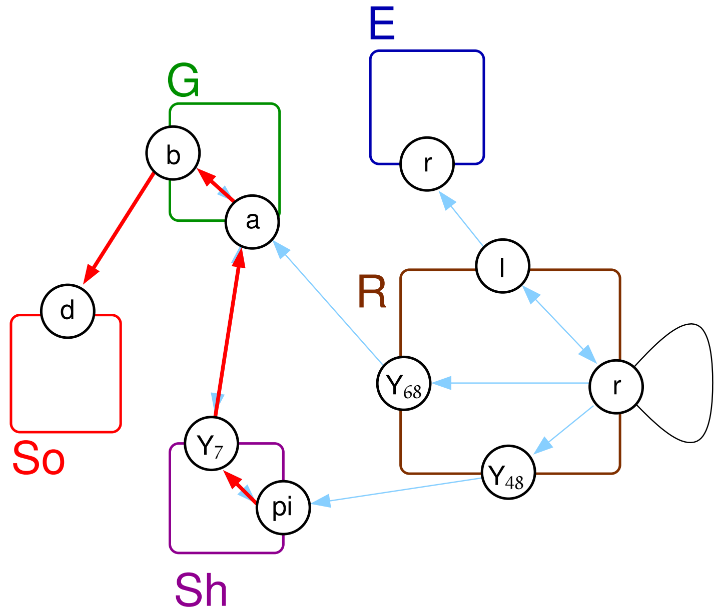


It can be refined into another prefragment.
Thus, it is not a fragment.



Are they fragments ?





It is maximally specified.
Thus it is a fragment.

Orthogonal refinement

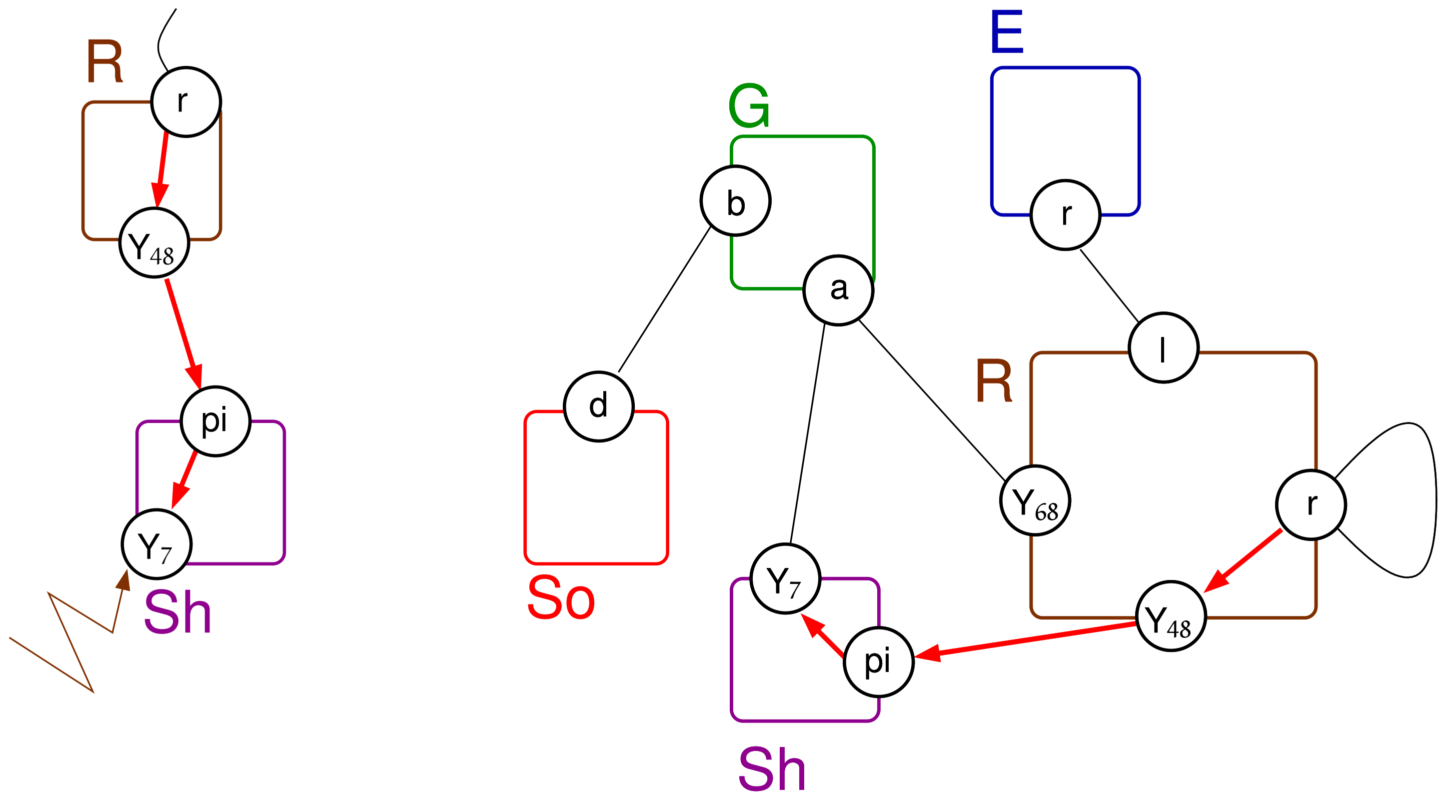
Property 1 (prefragment) The concentration of any prefragment can be expressed as a linear combination of the concentration of some fragments.

Which constraints shall we impose 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) Flow of information
   (c) Abstract counterpart
7. Conclusion

Flow of information

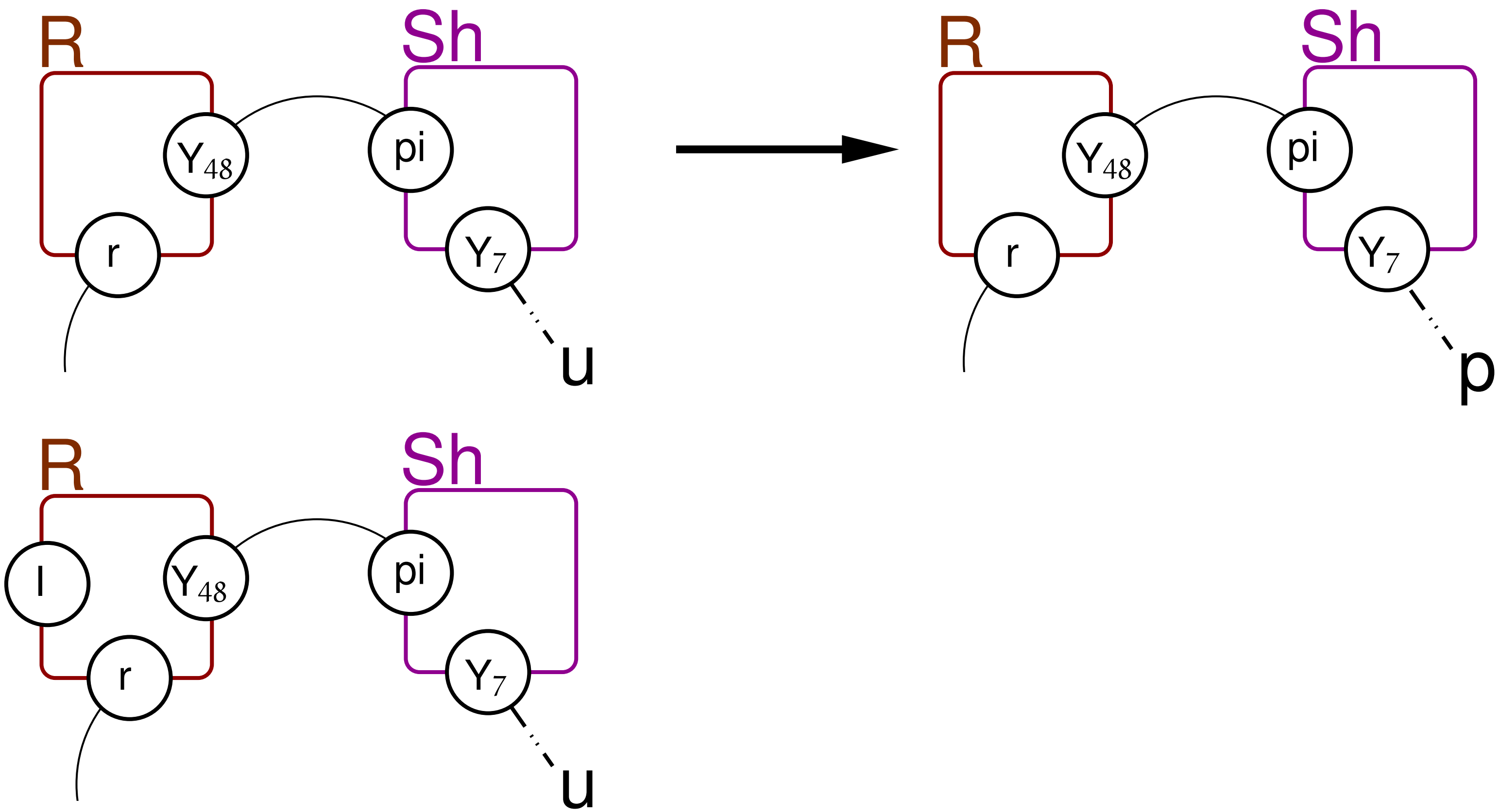


We reflect, in the annotated contact map, each path that stems from a site that is tested to a site that is modified.

Overview

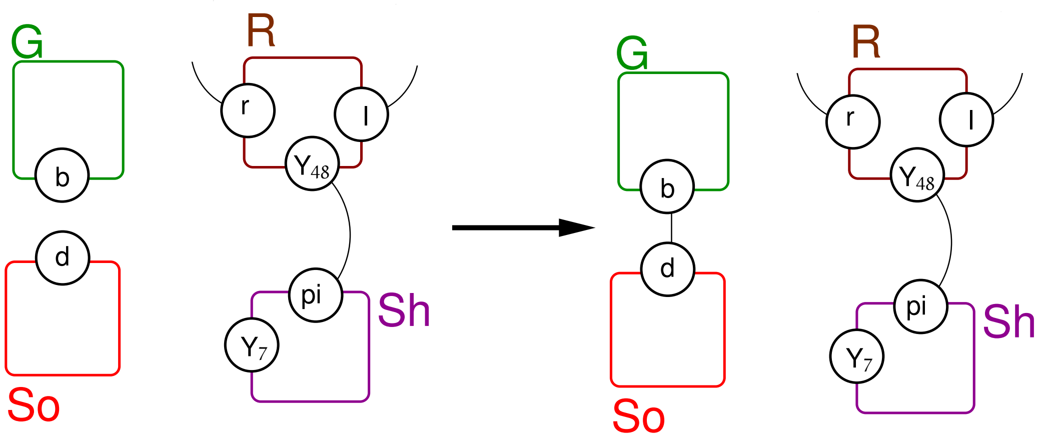
1. Context and motivations
2. Handmade ODEs
3. Abstract interpretation framework
4. Kappa
5. Concrete semantics
6. Abstract semantics
    (a) Fragments
    (b) Flow of information
    (c) Abstract counterpart
7. Conclusion

Fragments consumption
Proper intersection



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

Fragment consumption



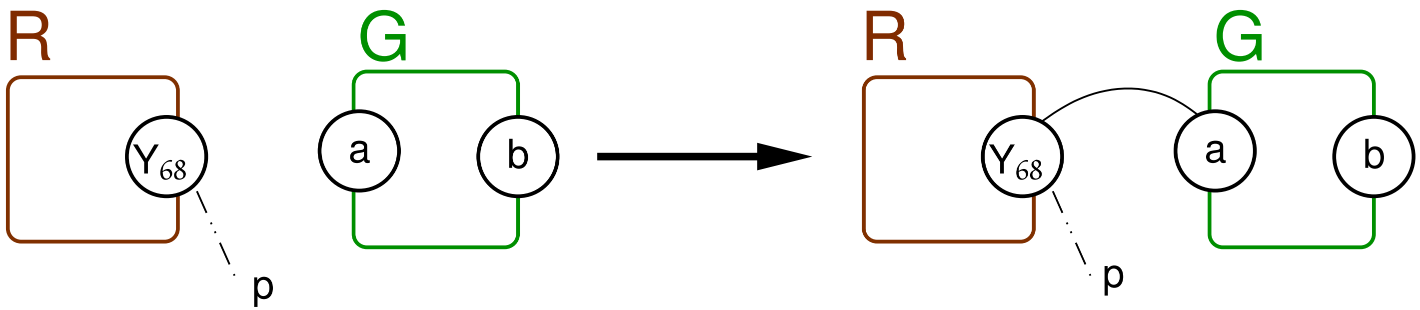
For any rule:

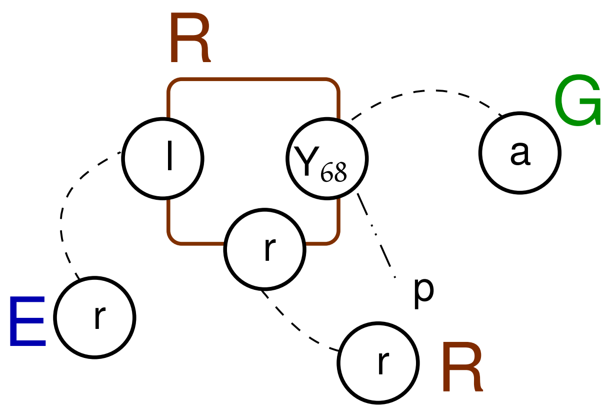
$$\textit{rule}: \; C_1, \ldots, C_n \to \textit{rhs} \quad k$$

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

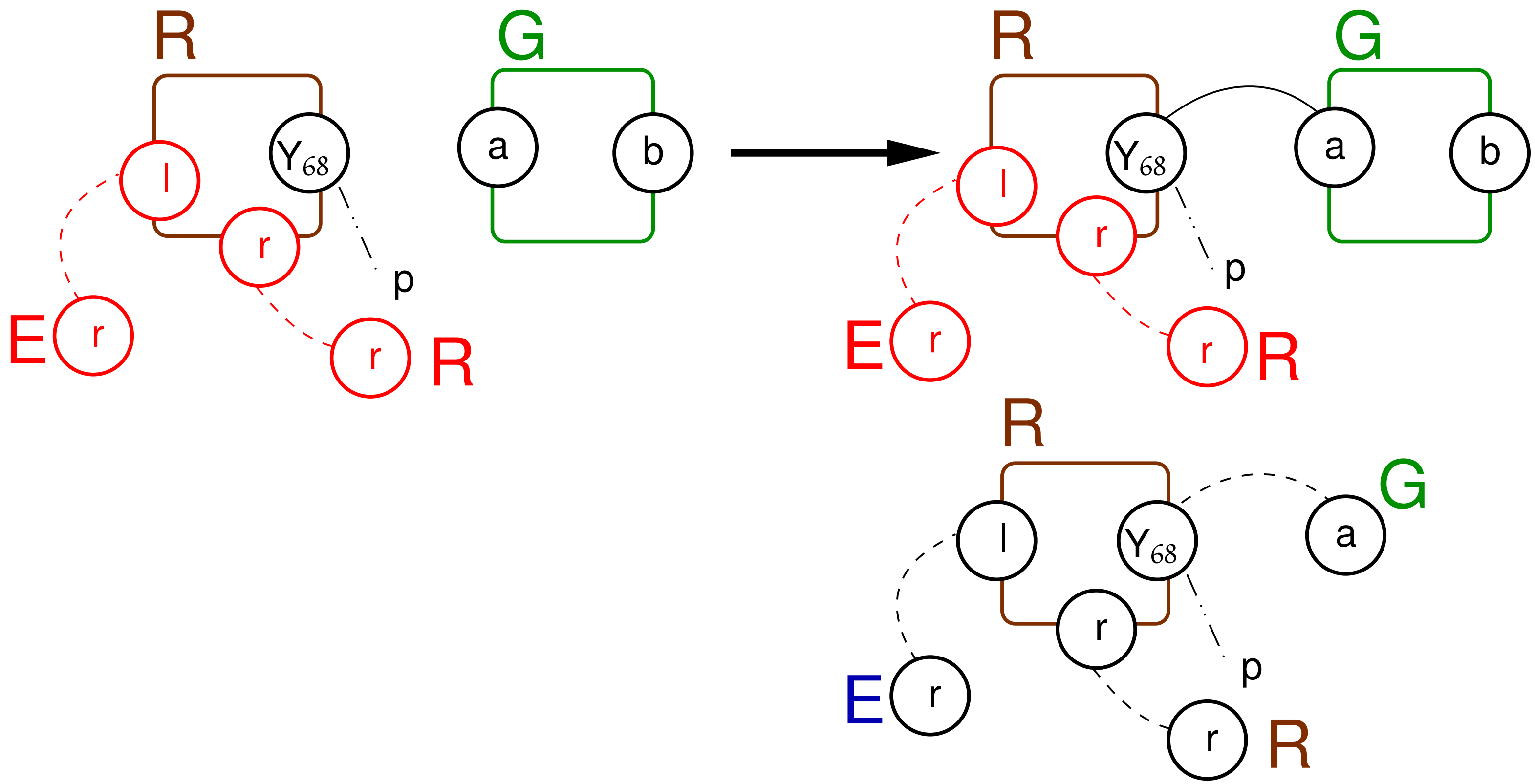
$$\frac{d[F]}{dt} \stackrel{-}{=} \frac{k \cdot [F] \cdot \prod_{i \neq k} [C_i]}{\mathsf{SYM}(C_1, \ldots, C_n) \cdot \mathsf{SYM}(F)}.$$

Fragment production



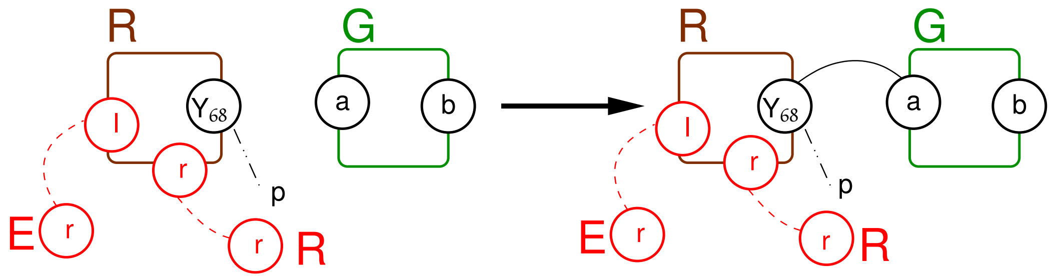


Fragment production
Proper intersection (bis)



Any connected component of the lhs of the refinement is prefragments.

Fragment production



For any rule:

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

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.

We get:

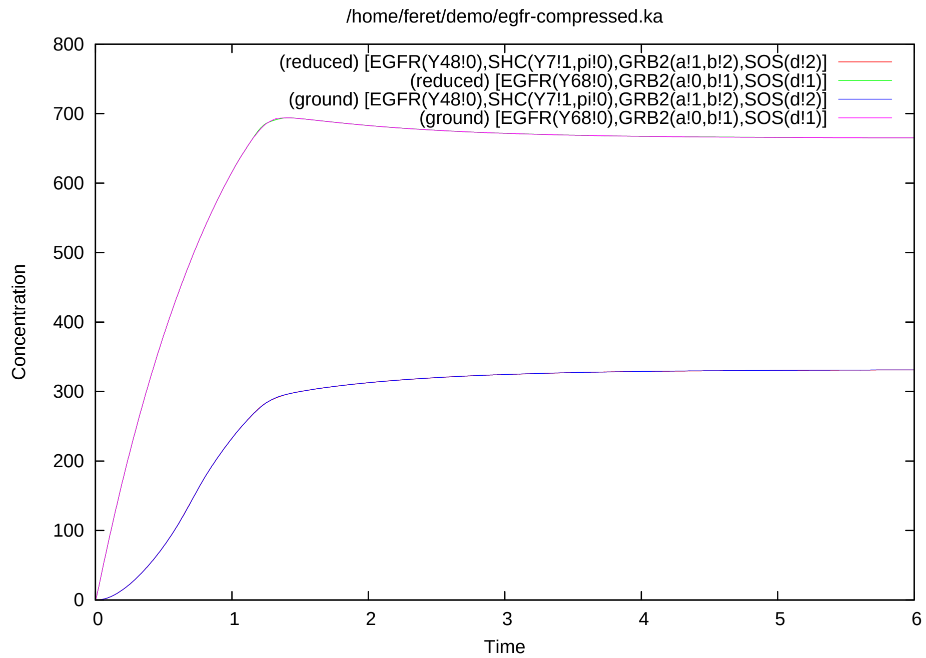
$$\frac{d[F]}{dt} \stackrel{+}{=} \frac{k \cdot \prod_i \left[C'_i\right]}{\mathsf{SYM}(C_1, \ldots, C_m) \cdot \mathsf{SYM}(F)}.$$

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. ODE approximations:
   - Less syntactic approximation of the flow of information.
   - A hierarchy of abstractions tuned by the level of context-sensitivity.

   Joint work with Ferdinanda Camporesi (Bologna/ÉNS)

2. Model reduction of the stochastic semantics:
   - See the poster of Tatjana Petrov.

SASB 2012

Third International Workshop on Static Analysis and Systems Biology
http://www.di.ens.fr/sasb2012/

September the 10th,
Deauville,
France,

Abstract Submission: 25th of May
Paper Submission: 1st of June

Co-chaired by:

- Jérôme Feret
- Andre Levchenko.

Keynote speakers:

- Russ Harmer
- Andre Levchenko.