5th International Workshop on
Static Analysis and Systems Biology

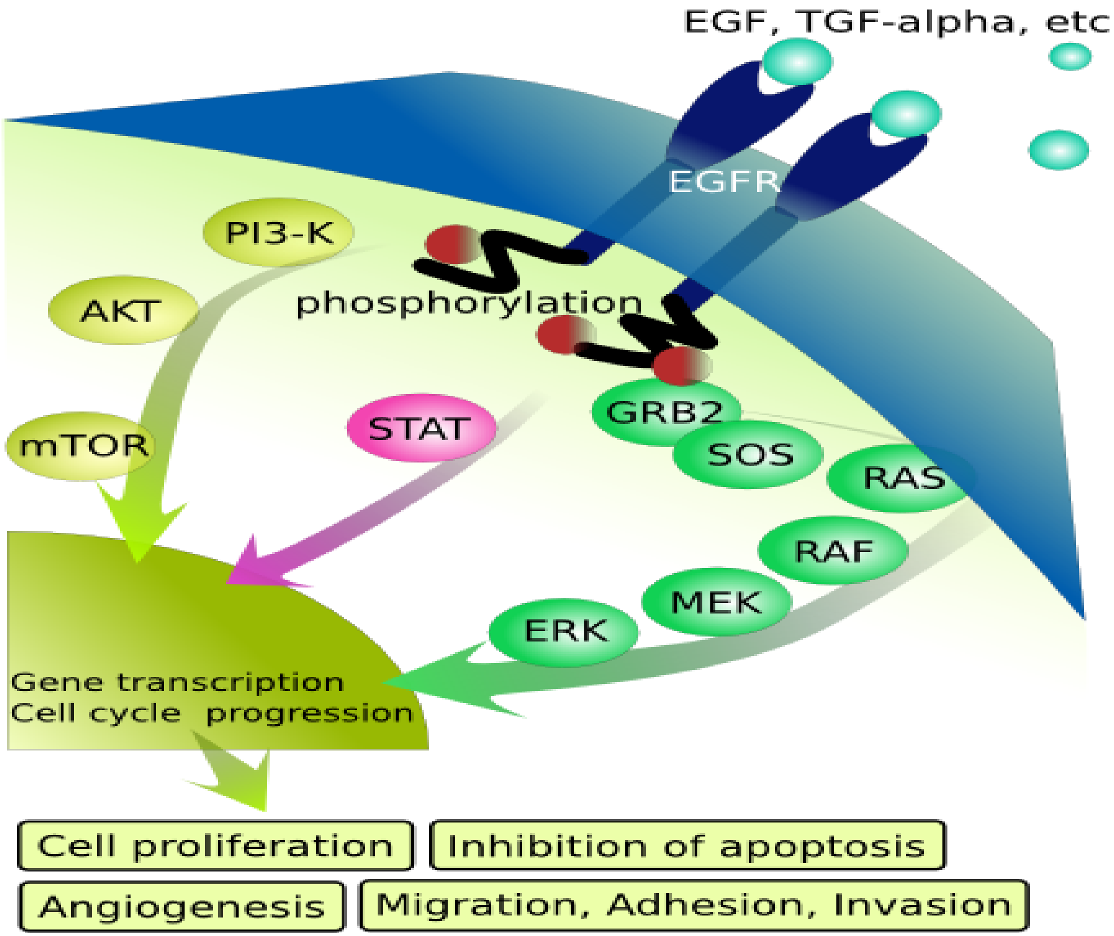
An algebraic approach for inferring and using symmetries in rule-based models

Jérôme Feret
Département d'informatique de l'école normale supérieure
INRIA, ÉNS, CNRS

2014, September 10

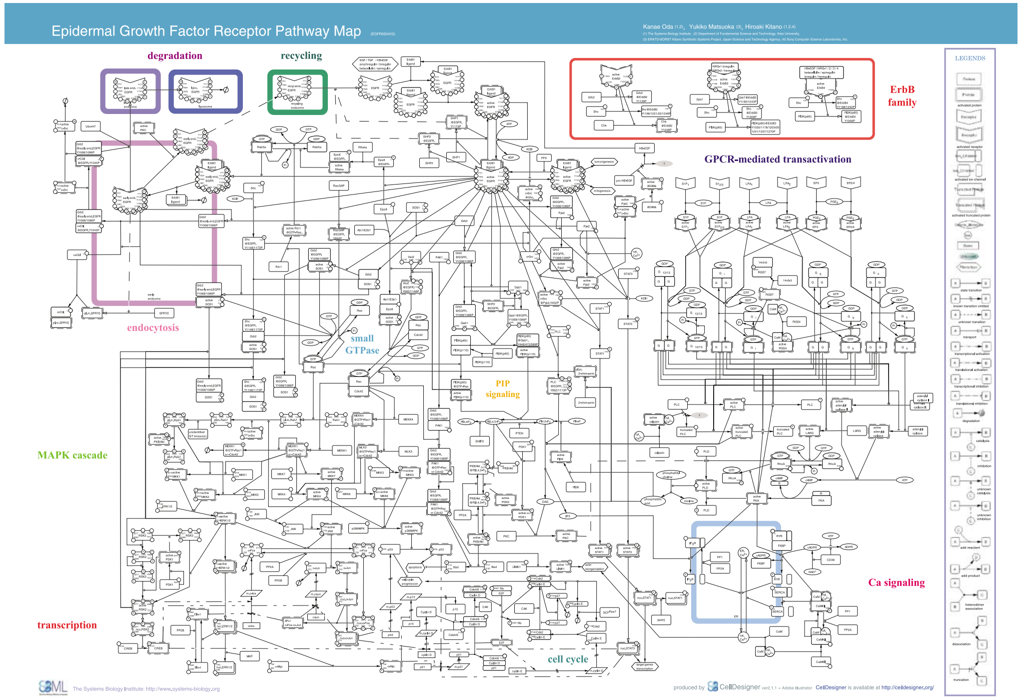
Overview

Signalling Pathways



Eikuch, 2007

A gap between two modeling methods

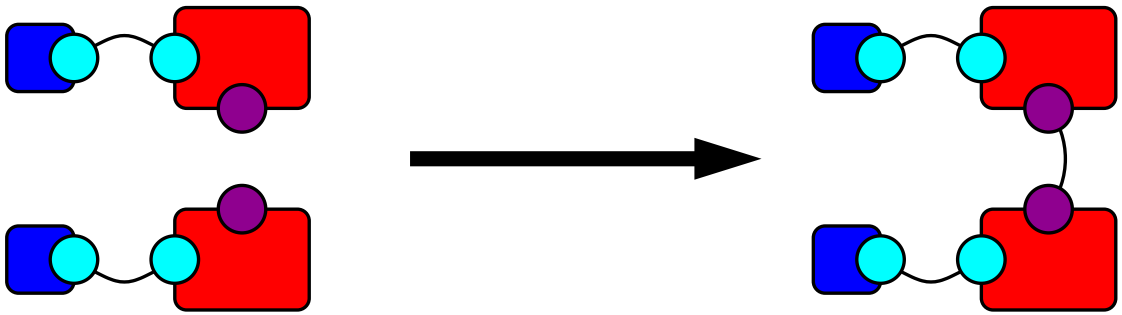


Oda, Matsuoka, Funahashi, Kitano, Molecular Systems Biology, 2005

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

Rule-based approach

We use site graph rewrite systems

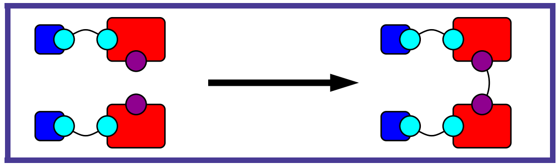


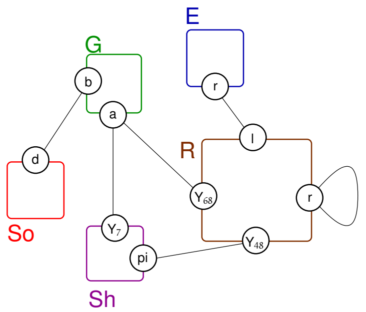
1. The description level matches with both
   - the observation level
   - and the intervention level

   of the biologist.

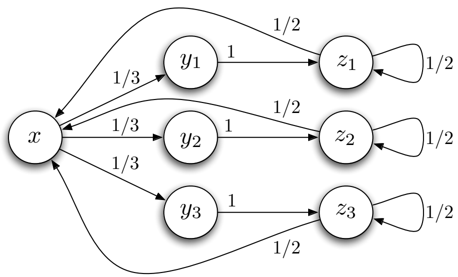
   We can tune the model easily.
2. Model description is very compact.

Rule-based models





Interaction map

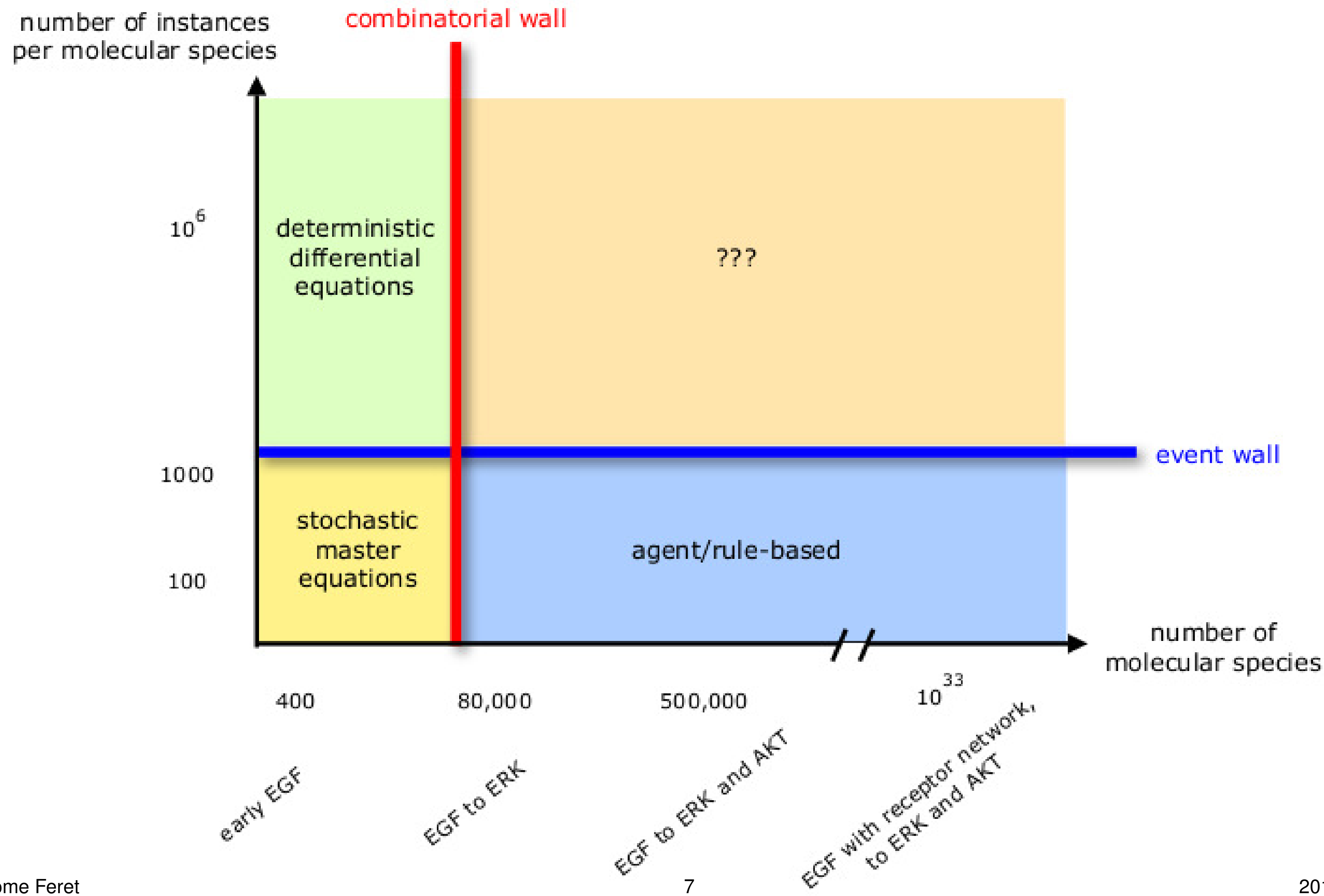


CTMC

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

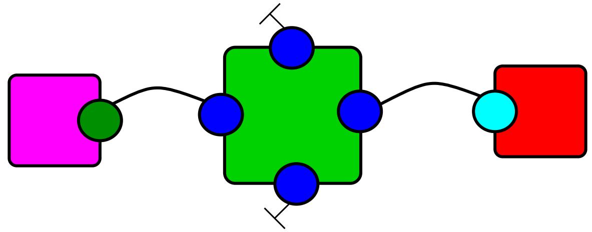
ODEs

Complexity walls

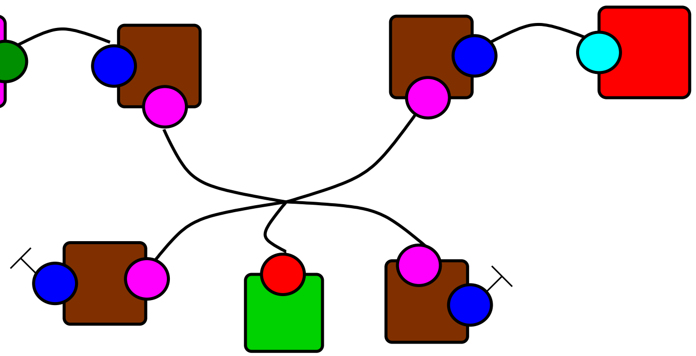


Symmetric sites

- in BNGL or MetaKappa (multiple-occurrences of sites):

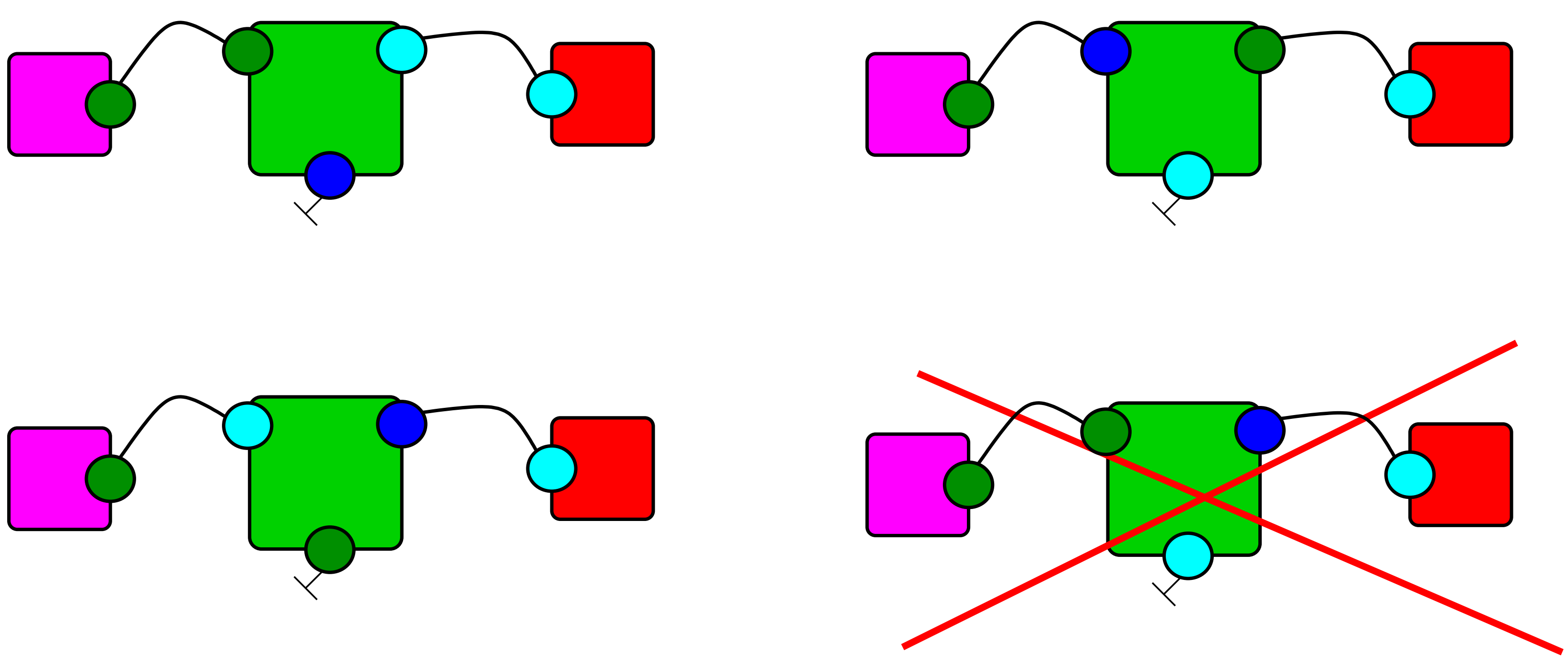


- in Formal Cellular Machinery or React(C) (hyper-edges):

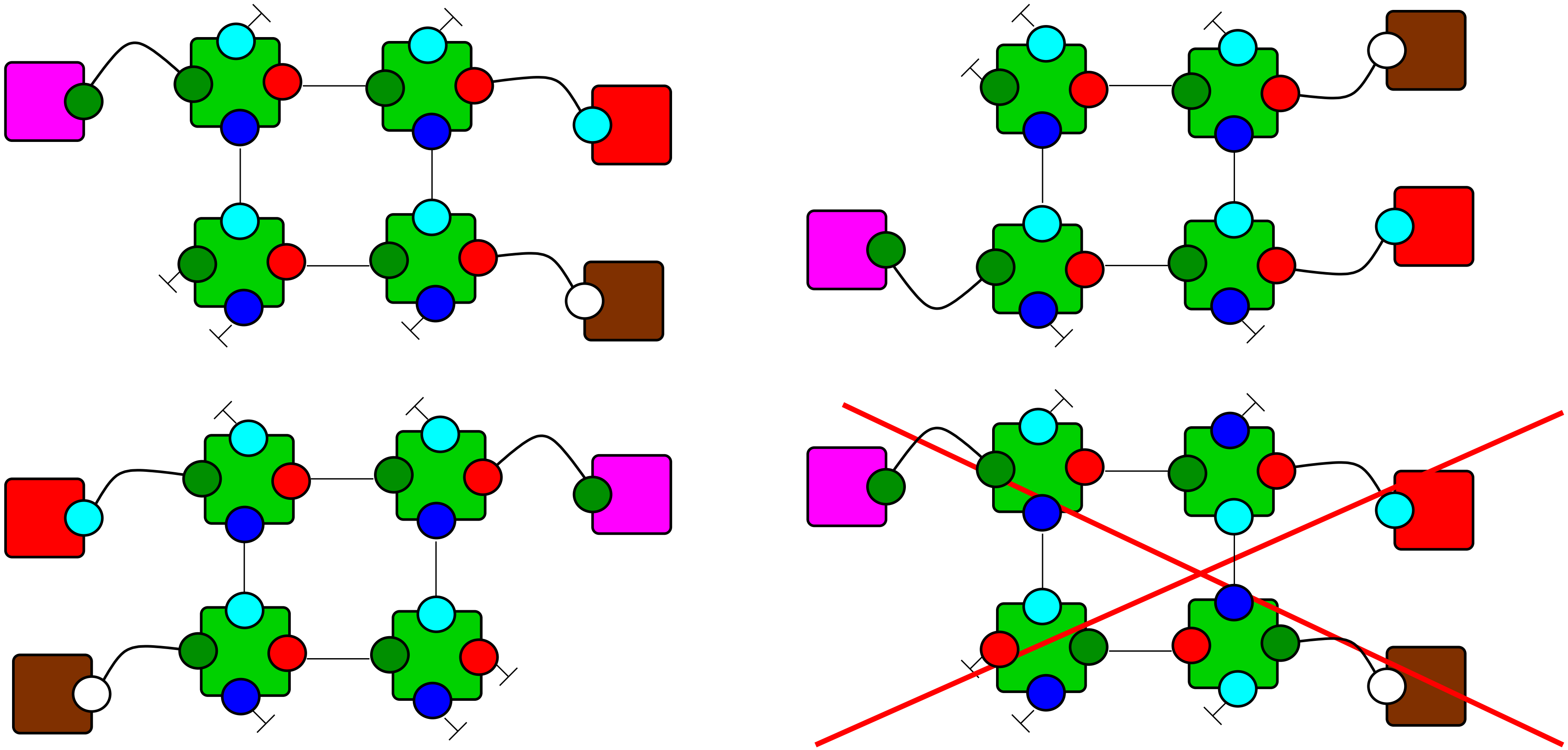


Blinov et al., BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains, Bioinformatics 2004
Danos et al., Rule-Based Modelling and Model Perturbation, TCSB 2009
Damgaard et al., Formal cellular machinery, Damgaard et al., SASB 2011
John et al., Biochemical Reaction Rules with Constraints, ESOP 2011

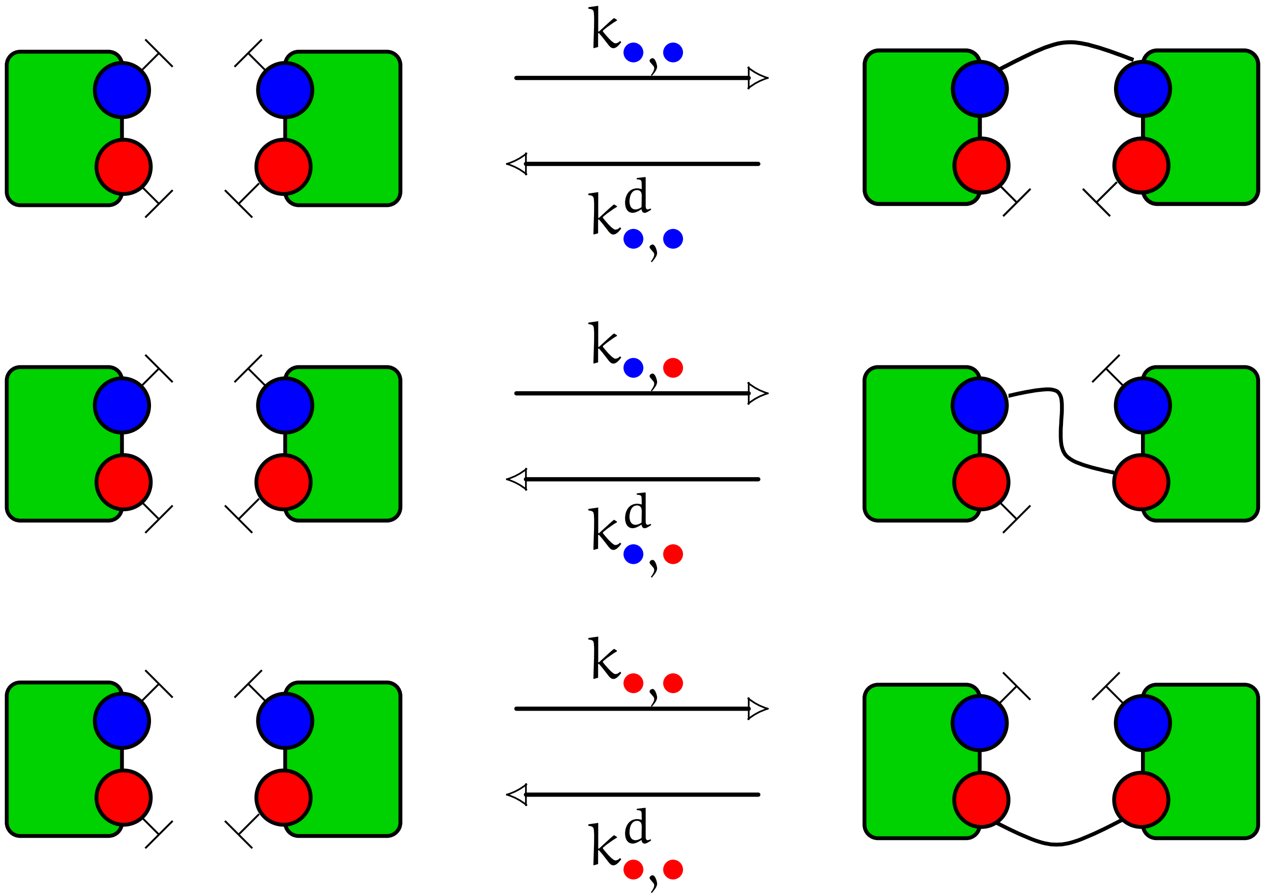
Other kinds of symmetries:
Circular permutations



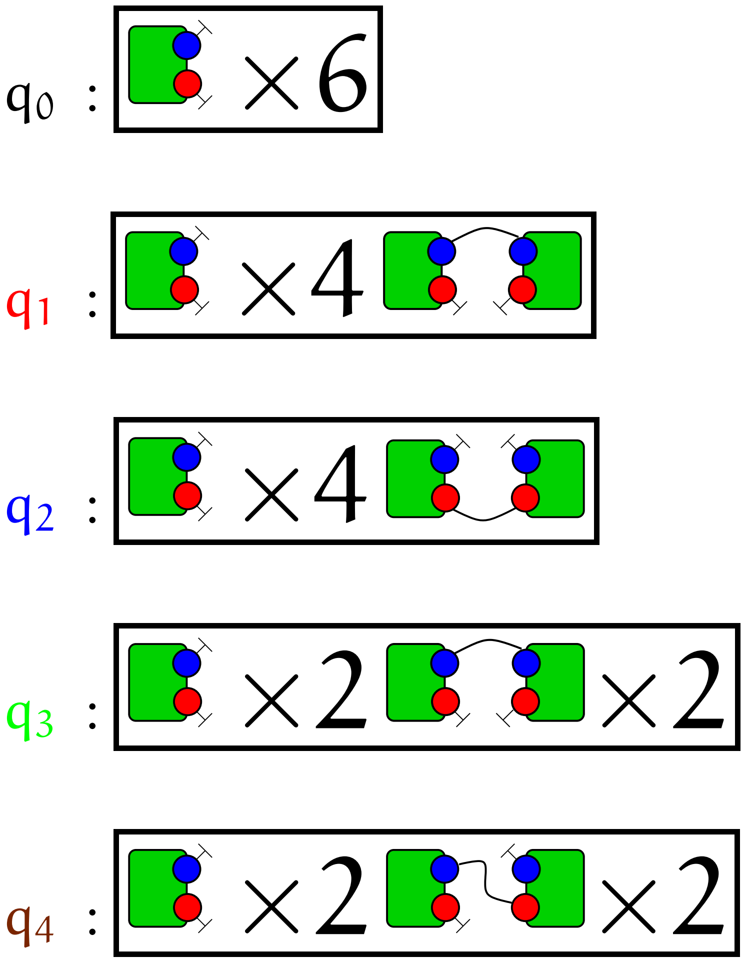
Other kinds of symmetries: Homogeneous symmetries

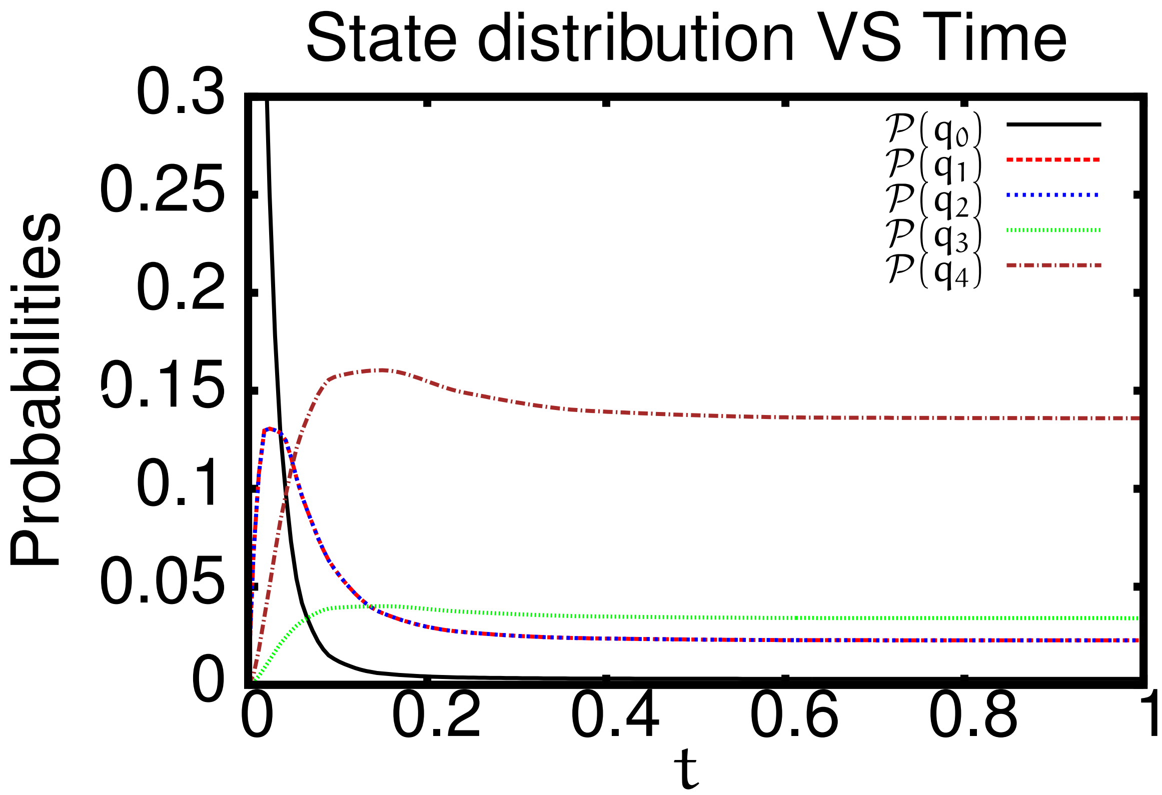


Case study



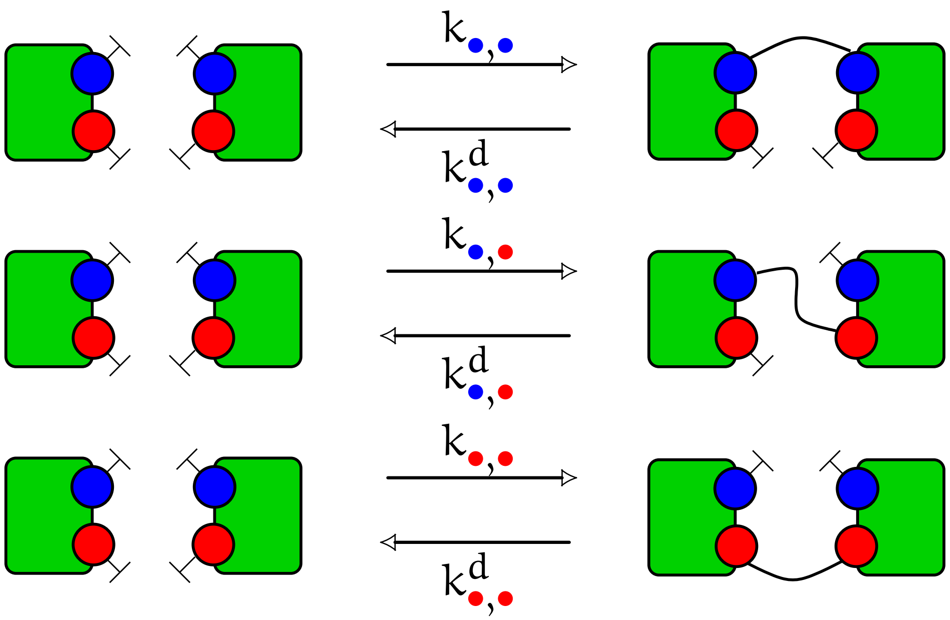
State distribution





with: $\begin{cases} k_{\color{blue}\bullet,\color{blue}\bullet} = k_{\color{red}\bullet,\color{red}\bullet} = 1 \\ k_{\color{blue}\bullet,\color{red}\bullet} = k^d_{\color{blue}\bullet,\color{blue}\bullet} = k^d_{\color{red}\bullet,\color{red}\bullet} = k^d_{\color{blue}\bullet,\color{red}\bullet} = 2 \\ P(q_0 \mid t = 0) = 1 \end{cases}$

Lumpability

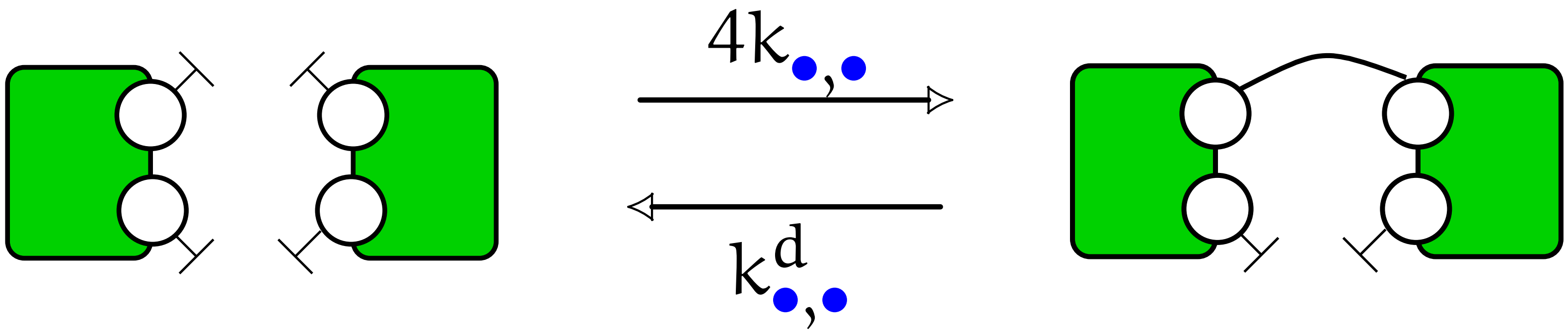


Whenever:

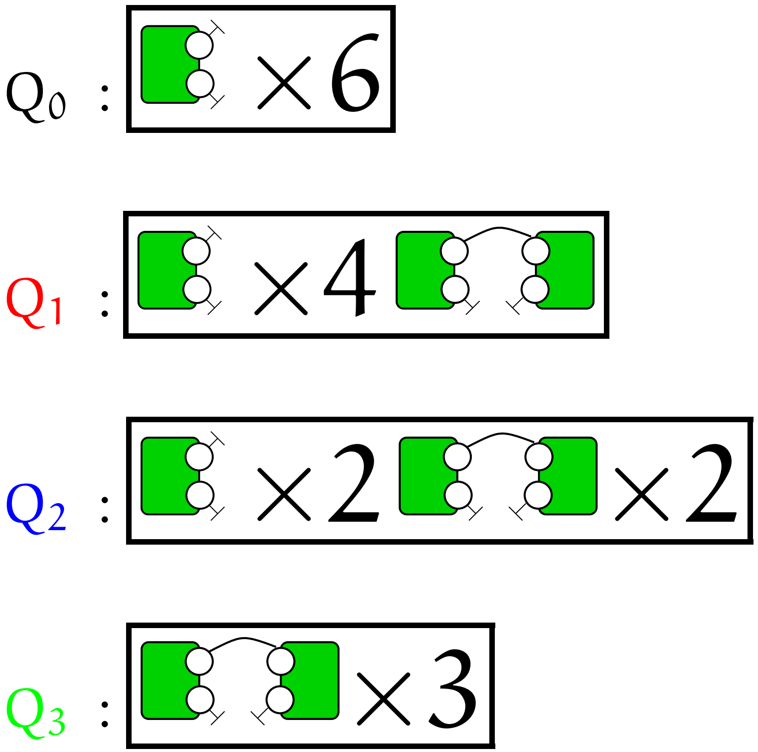
$$\begin{cases} 2k_{{\color{blue}\bullet},{\color{blue}\bullet}} = 2k_{{\color{red}\bullet},{\color{red}\bullet}} = k_{{\color{blue}\bullet},{\color{red}\bullet}} \\ k^{d}_{{\color{blue}\bullet},{\color{blue}\bullet}} = k^{d}_{{\color{blue}\bullet},{\color{red}\bullet}} = k^{d}_{{\color{red}\bullet},{\color{red}\bullet}} \end{cases}$$

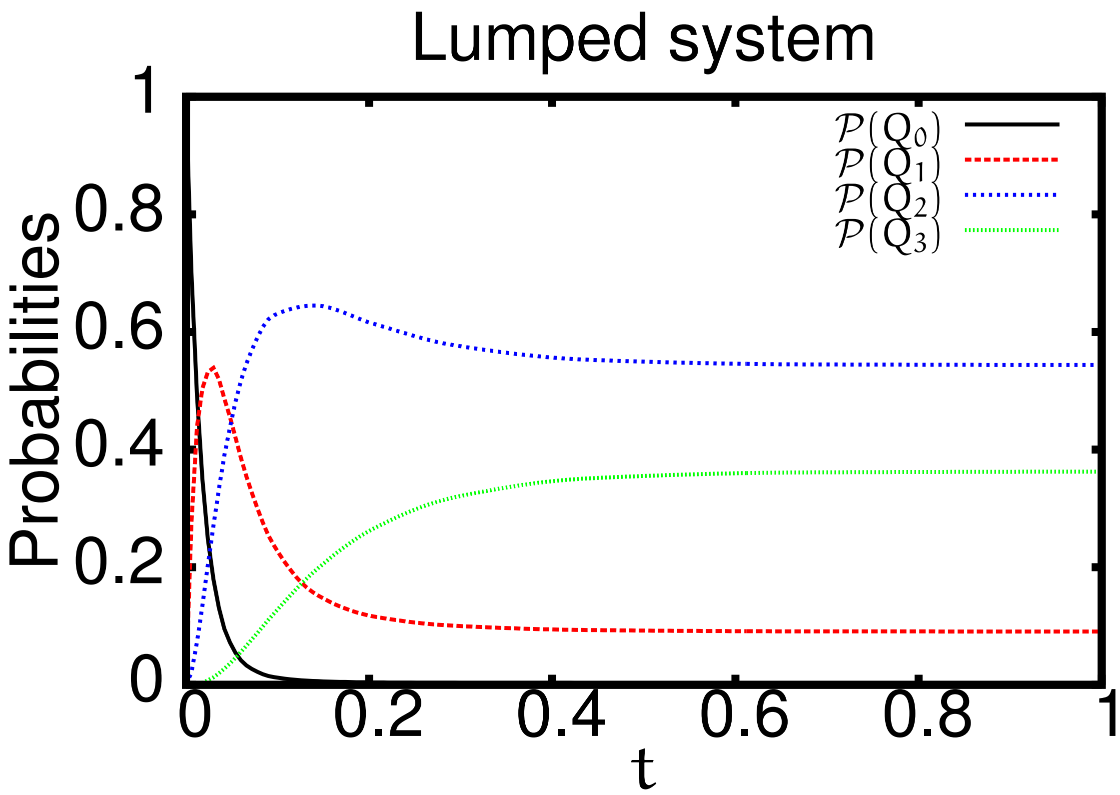
We can lump the system.

Lumped system



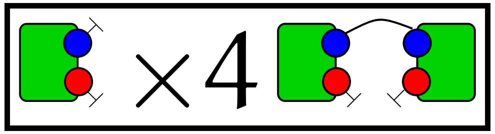
Macrostate distribution

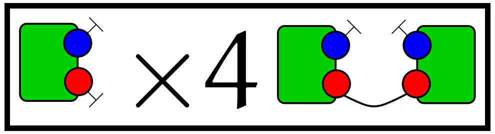


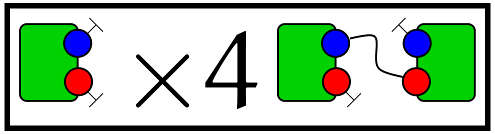


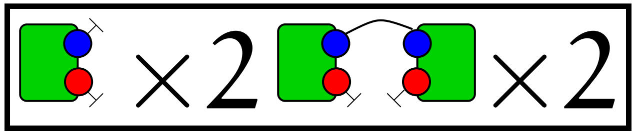
with: $\begin{cases} k_{\bullet,\bullet} = 1 \\ k^{d}_{\bullet,\bullet} = 2 \\ P(q_0 \mid t = 0) = 1 \end{cases}$

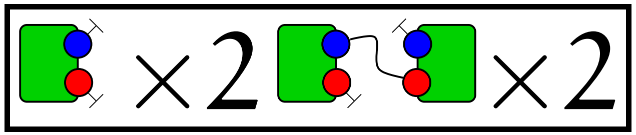
Probability ratios

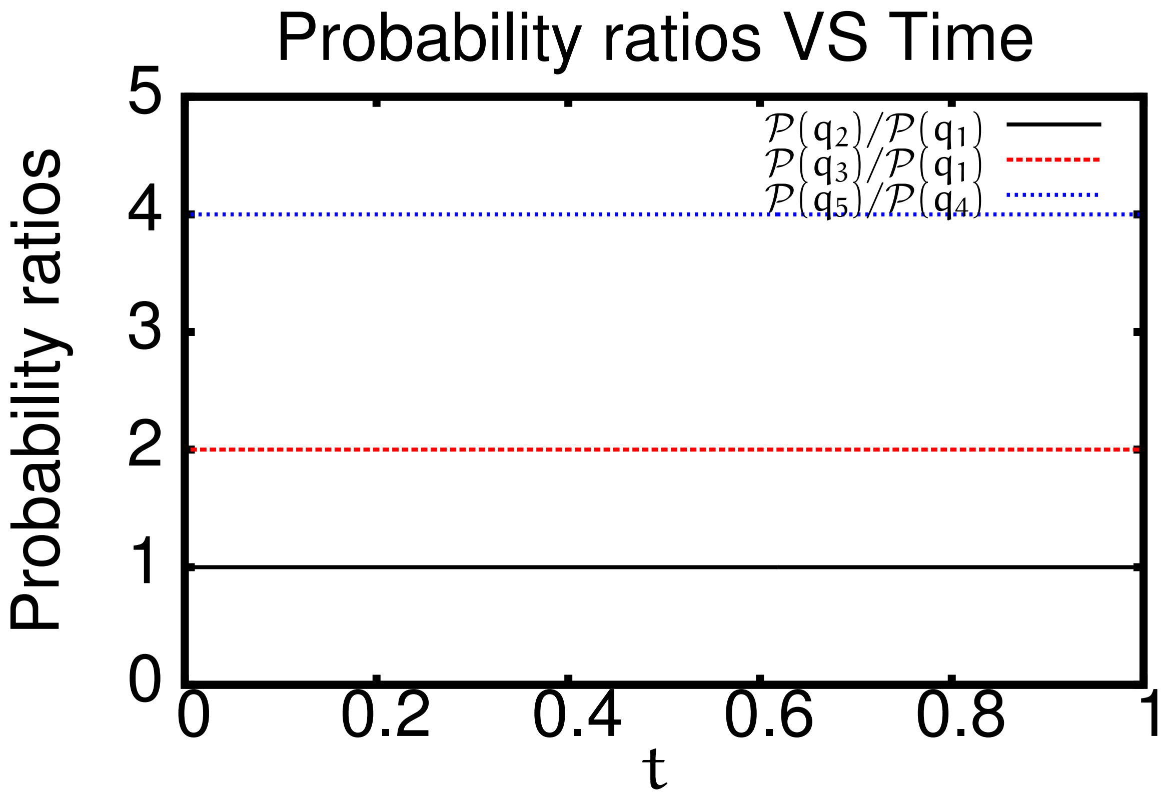
q_1 :


q_2 :


q_3 :


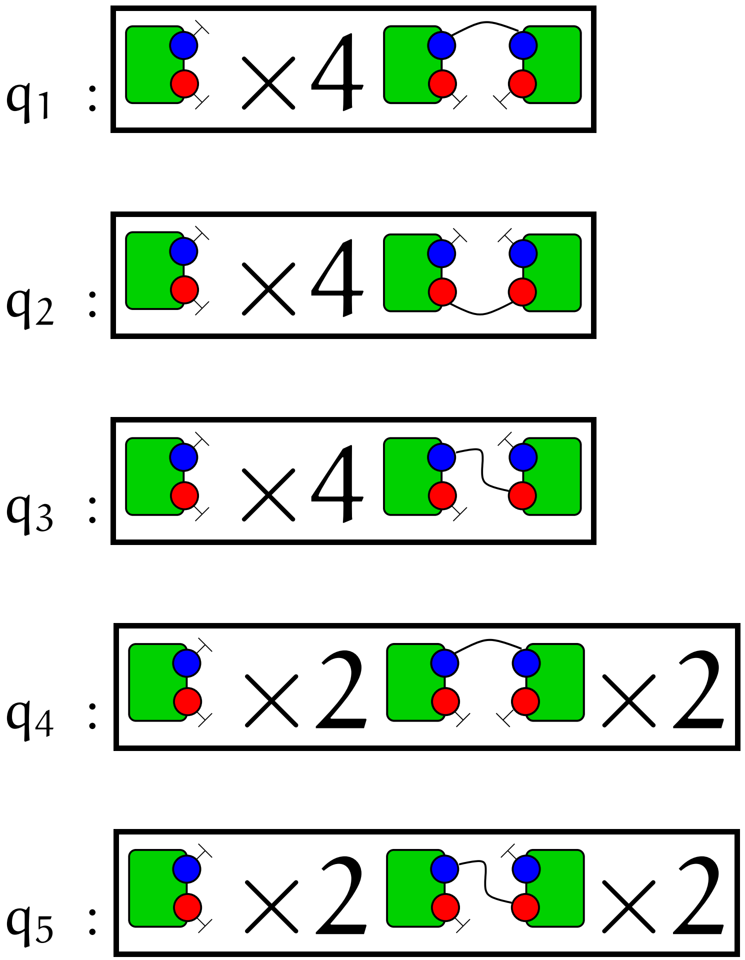
q_4 :


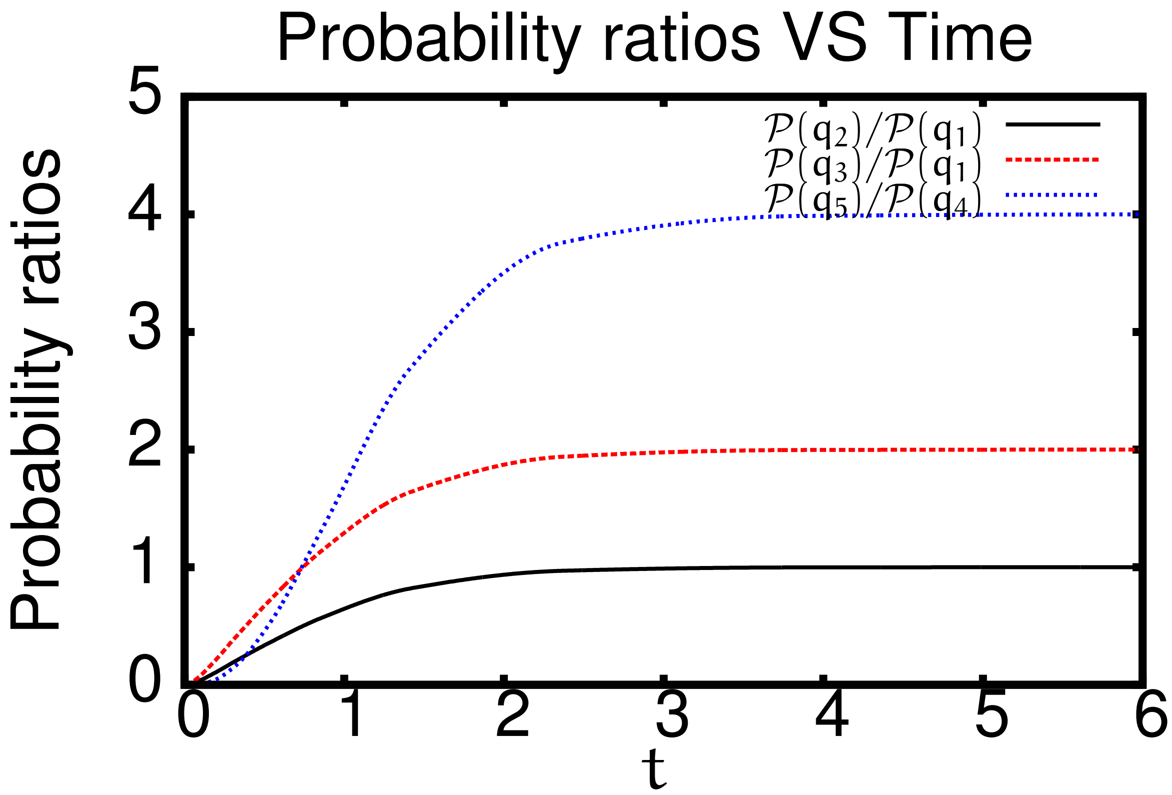
q_5 :




with:
$$\begin{cases} k_{{\color{blue}\bullet},{\color{blue}\bullet}} = k_{{\color{red}\bullet},{\color{red}\bullet}} = 1 \\ k_{{\color{blue}\bullet},{\color{red}\bullet}} = k^d_{{\color{blue}\bullet},{\color{blue}\bullet}} = k^d_{{\color{red}\bullet},{\color{red}\bullet}} = k^d_{{\color{blue}\bullet},{\color{red}\bullet}} = 2 \\ P(q_0 \mid t = 0) = 1 \end{cases}$$

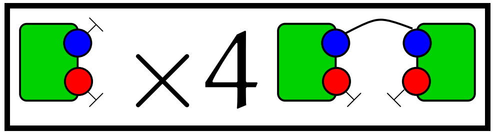
Probability ratios (wrong initial condition)

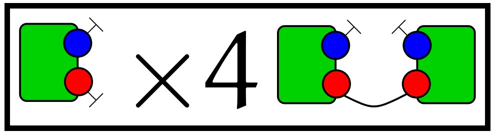


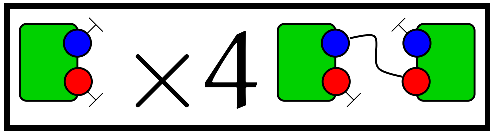


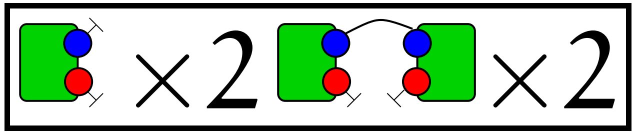
with: $\begin{cases} k_{{\color{blue}\bullet},{\color{blue}\bullet}} = k_{{\color{red}\bullet},{\color{red}\bullet}} = 1 \\ k_{{\color{blue}\bullet},{\color{red}\bullet}} = k^{d}_{{\color{blue}\bullet},{\color{blue}\bullet}} = k^{d}_{{\color{red}\bullet},{\color{red}\bullet}} = k^{d}_{{\color{blue}\bullet},{\color{red}\bullet}} = 2 \\ P(q_4 \mid t = 0) = 1 \end{cases}$

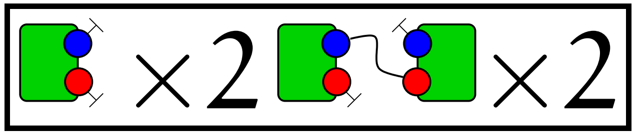
Probability ratios (wrong coefficients)

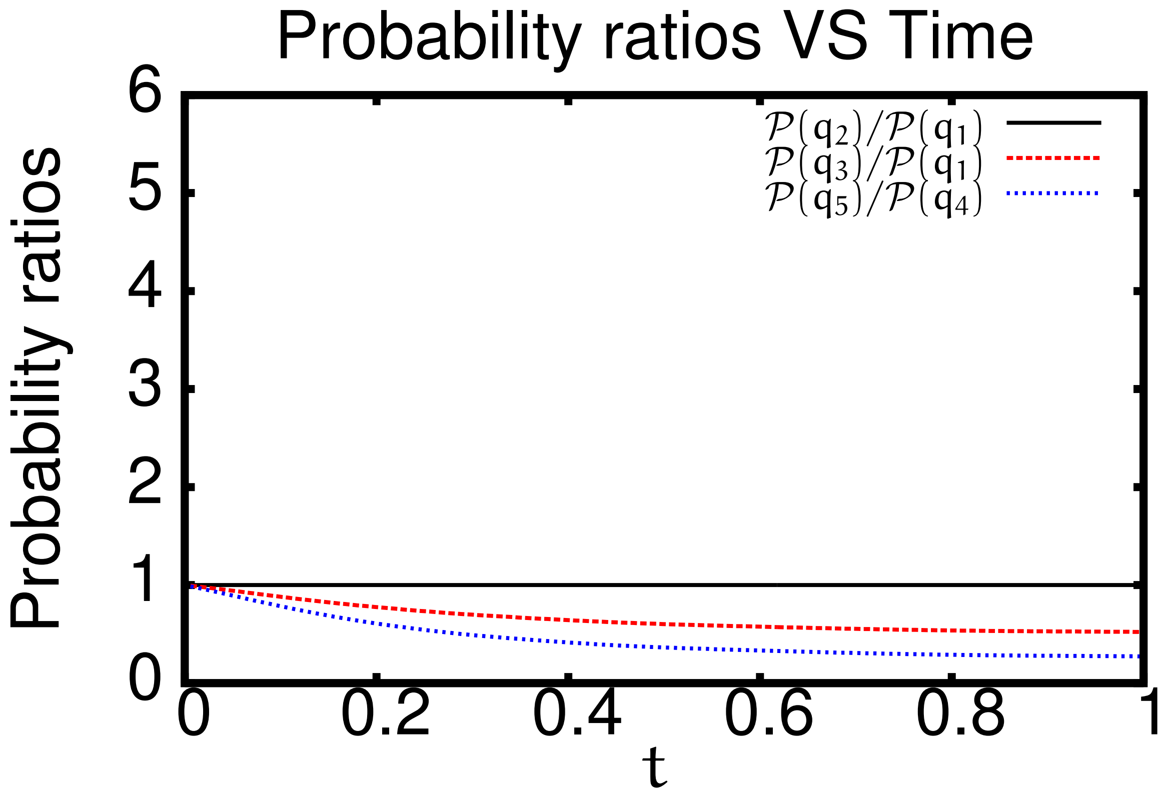
q_1 : 

q_2 : 

q_3 : 

q_4 : 

q_5 : 



with:
$$\begin{cases} k_{\color{blue}\bullet,\color{blue}\bullet} = k_{\color{red}\bullet,\color{red}\bullet} = k_{\color{blue}\bullet,\color{red}\bullet} = 1 \\ k^d_{\color{blue}\bullet,\color{blue}\bullet} = k^d_{\color{red}\bullet,\color{red}\bullet} = 2 \\ k^d_{\color{blue}\bullet,\color{red}\bullet} = 4 \\ P(q_0 \mid t = 0)) = 1 \end{cases}$$

In this talk

An algebraic notion of symmetries over site graphs:

- compatible with the SPO (Single Push-Out) semantics of Kappa;
- with a notion of subgroups of symmetries;
- with a notion of symmetric models.

Some conditions so that symmetries over a model induce

- a forward bisimulation;
- a backward bisimulation.

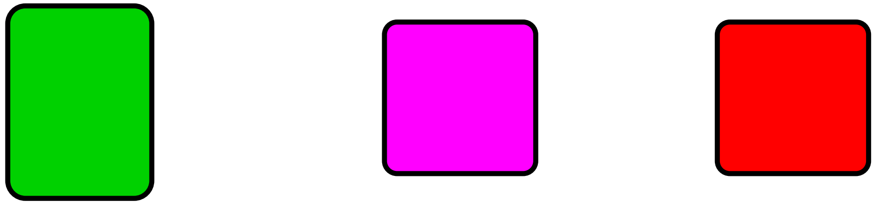
In this talk, we consider only a side-effect free fragment of Kappa.
The full language is handled with in the paper.

Overview

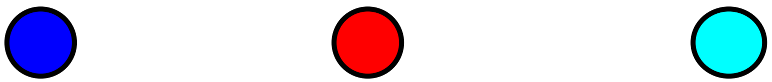
1. Context and motivations
2. Kappa semantics
3. Symmetries in site-graphs
4. Symmetric models
5. Conclusion

Signature

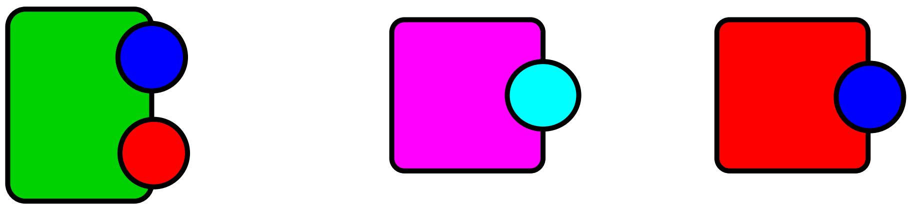
Agents:



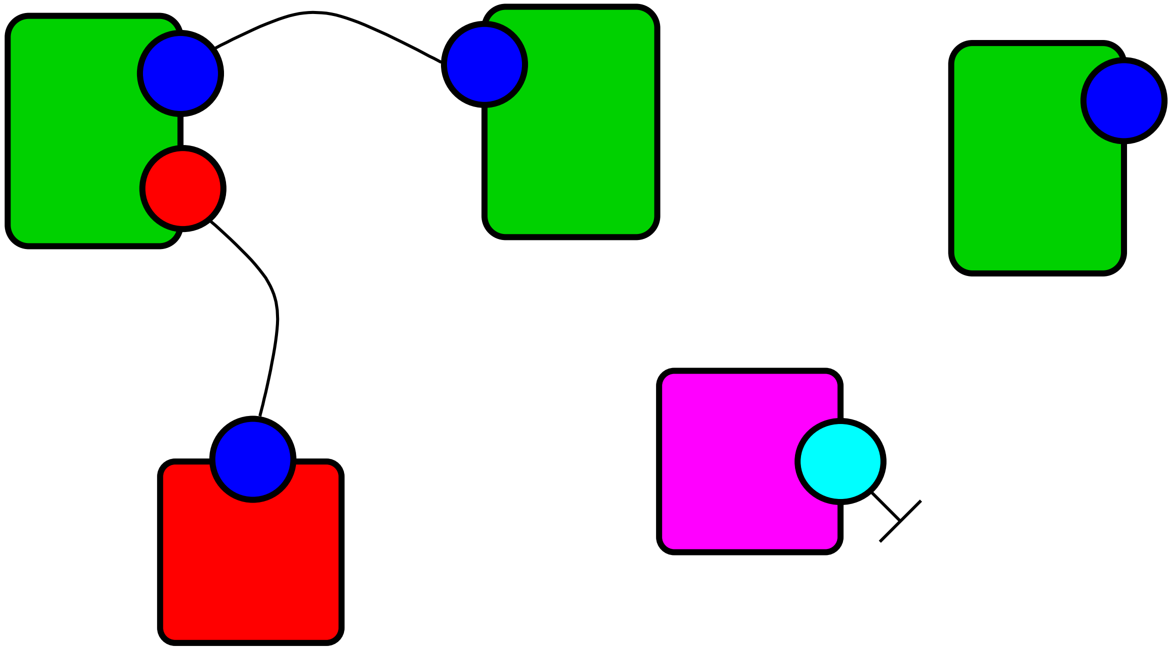
Sites:



Interface:

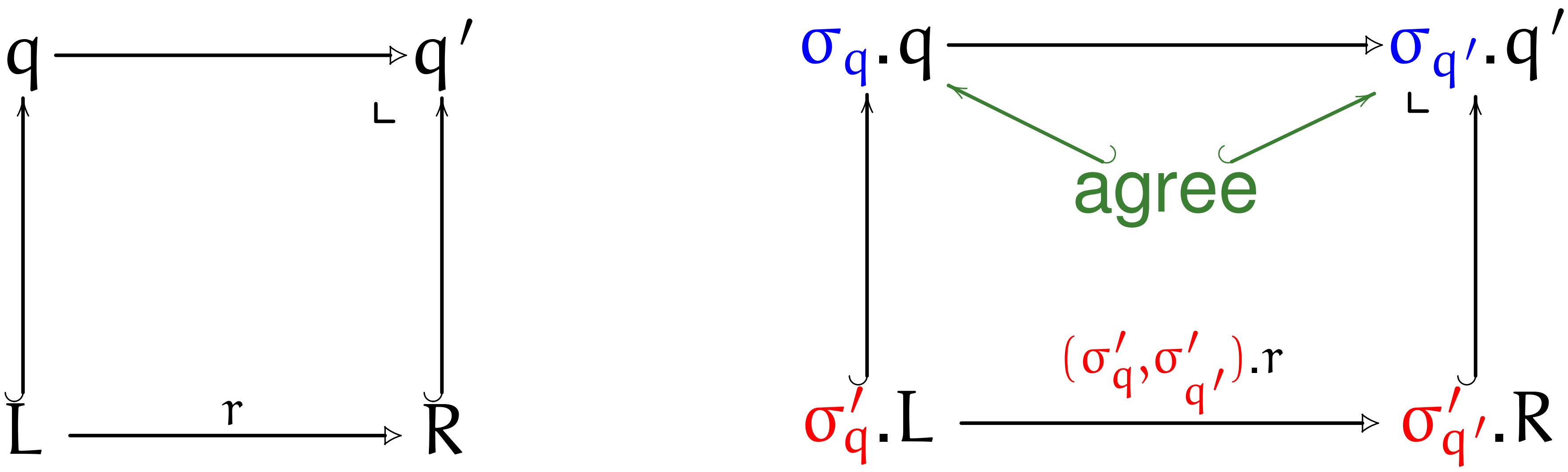


Site graph



Applying symmetries

We would like to make pairs of symmetries act over push-outs,

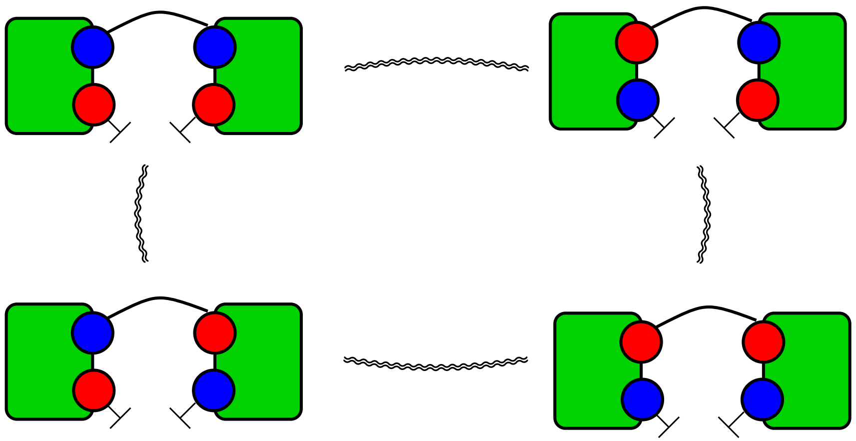


whenever they act the same way on preserved agents.

Overview

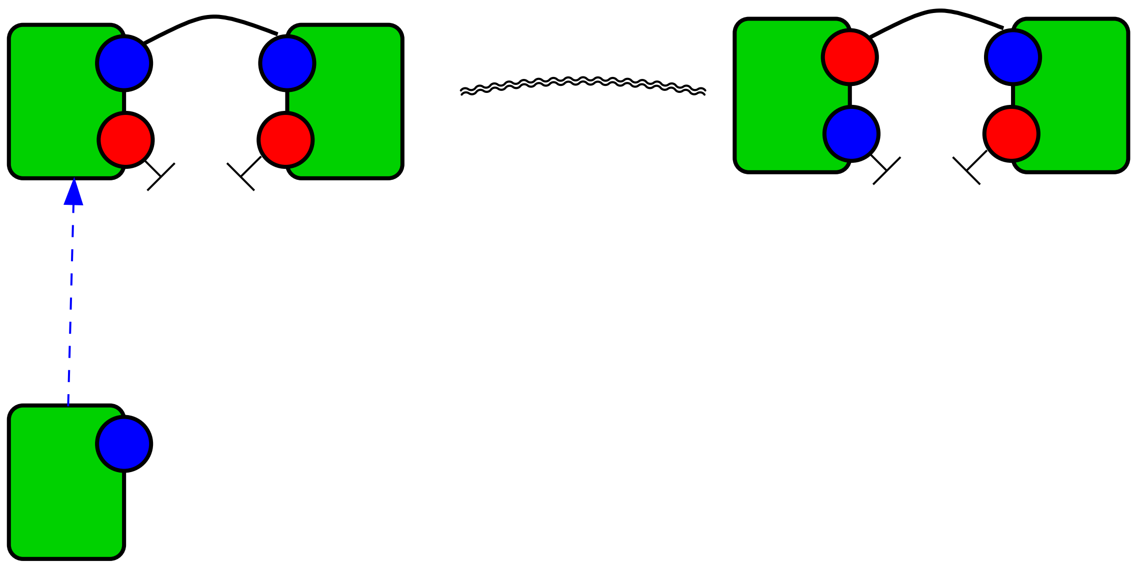
Symmetries over site graphs

- For any site graph G, we introduce a finite group of symmetries $\mathbb{G}_G$.

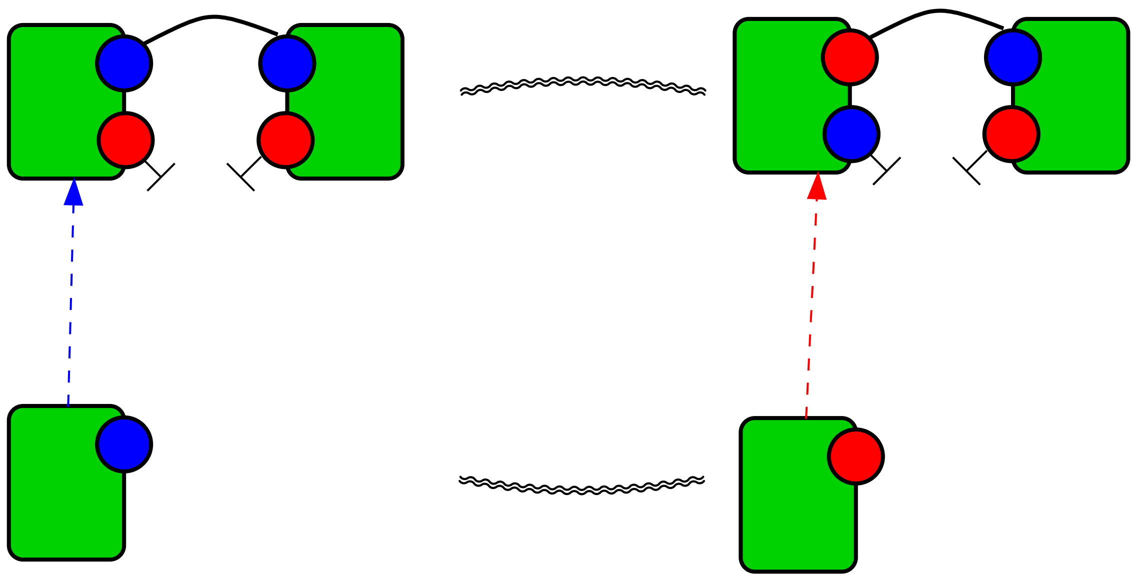


- For any site graph G and any symmetry $\sigma \in \mathbb{G}_G$, we introduce the site graph $\sigma.G$ and we call it the symmetric of G by σ.

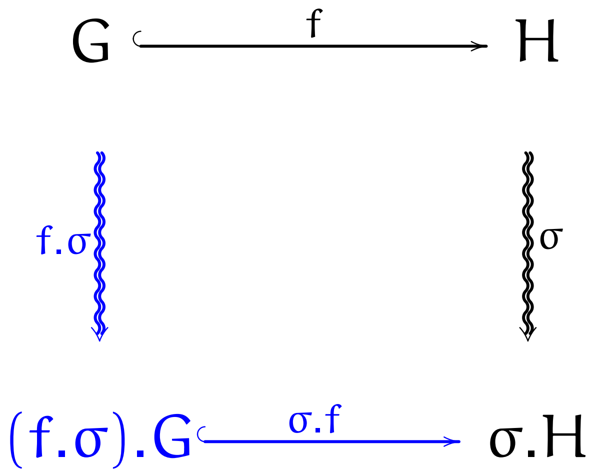
Restricting a symmetry to the domain of an embedding



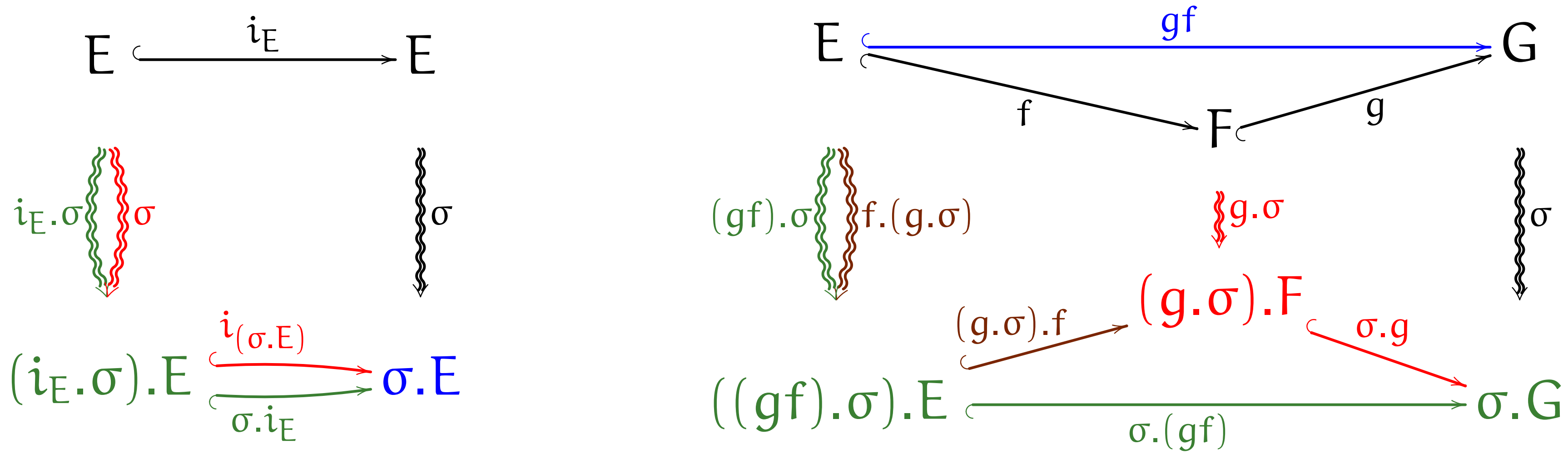
Restricting a symmetry to the domain of an embedding



Restriction of symmetry to the domain of an embedding



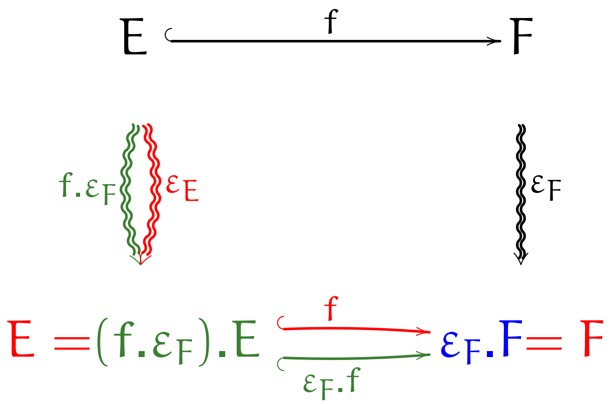
Symmetry VS embedding composition

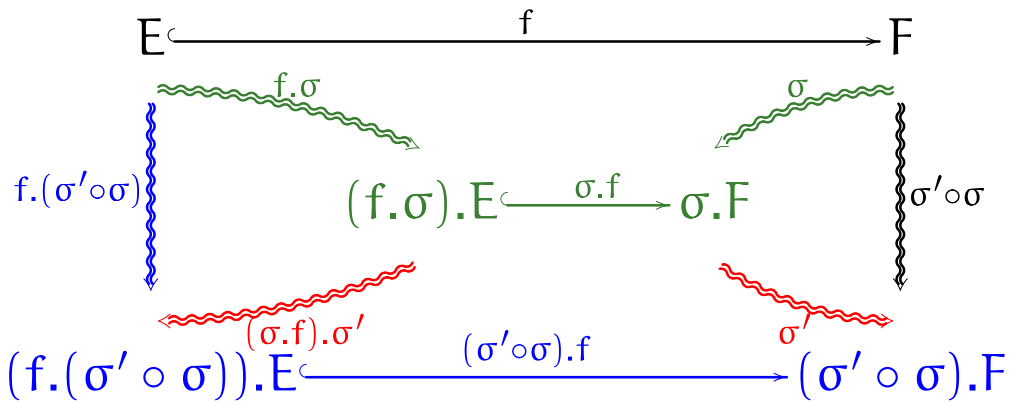


- $i_E.\sigma = \sigma$
- $\sigma.i_E = i_{(\sigma.E)}$

- $(gf).\sigma = f.(g.\sigma)$
- $\sigma.(gf) = (\sigma.g)((g.\sigma).f)$

Symmetry product
VS restriction to embedding domain



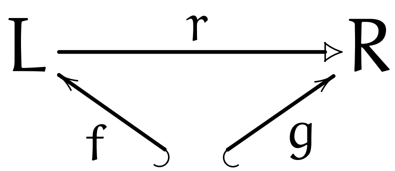


- $\varepsilon_F.F = F$
- $f.\varepsilon_F = \varepsilon_E$
- $\varepsilon_F.f = f$

- $(\sigma' \circ \sigma).F = \sigma'.(\sigma.F)$
- $f.(\sigma' \circ \sigma) = ((f.\sigma).\sigma') \circ (f.\sigma)$
- $(\sigma' \circ \sigma).f = \sigma'.(\sigma.f)$

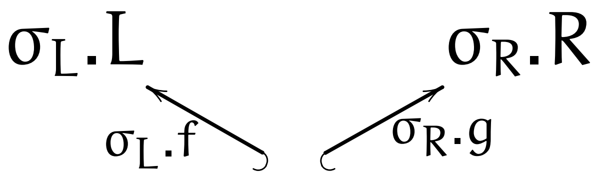
Symmetries VS rules

For any rule:



We introduce:

- $\mathbb{G}_r \stackrel{\Delta}{=} \{(\sigma_L, \sigma_R) \in \mathbb{G}_L \times \mathbb{G}_R \mid f.\sigma_L = g.\sigma_R\}$;
- $(\sigma_L, \sigma_R).r \stackrel{\Delta}{=}$



  (for any $(\sigma_L, \sigma_R) \in \mathbb{G}_r$).

We assume that:

1. $\mathbb{G}_r$ is stable upon pairwise product;
2. $\sigma.r$ is a rule, for any pair of symmetries $\sigma \in \mathbb{G}_r$ (and we write $r \approx_{\mathbb{G}} \sigma.r$).

Group actions over push-out

Theorem 1 Let r be a rule. The function which maps each pair of symmetries $(\sigma_L, \sigma_R) \in \mathbb{G}_r$ and each push-out of the form:

$$\begin{array}{ccc} L' & \xrightarrow{\quad r' \quad} & R' \\ {\scriptstyle h_L}\uparrow & & \uparrow{\scriptstyle h_R} \\ L & \xrightarrow[\quad r'' \quad]{} & R \end{array}$$

with $r' \approx_{\mathbb{G}} r$, to the push-out:

$$\begin{array}{ccc} \sigma_L.L' & \xrightarrow{\quad (\sigma_L,\sigma_R).r' \quad} & \sigma_R.R' \\ {\scriptstyle \sigma_L.h_L}\uparrow & & \uparrow{\scriptstyle \sigma_R.h_R} \\ (h_L.\sigma_L).L & \xrightarrow[(h_L.\sigma_L,h_R.\sigma_R).r'']{} & (h_R.\sigma_R).R \end{array}$$

is a group action.

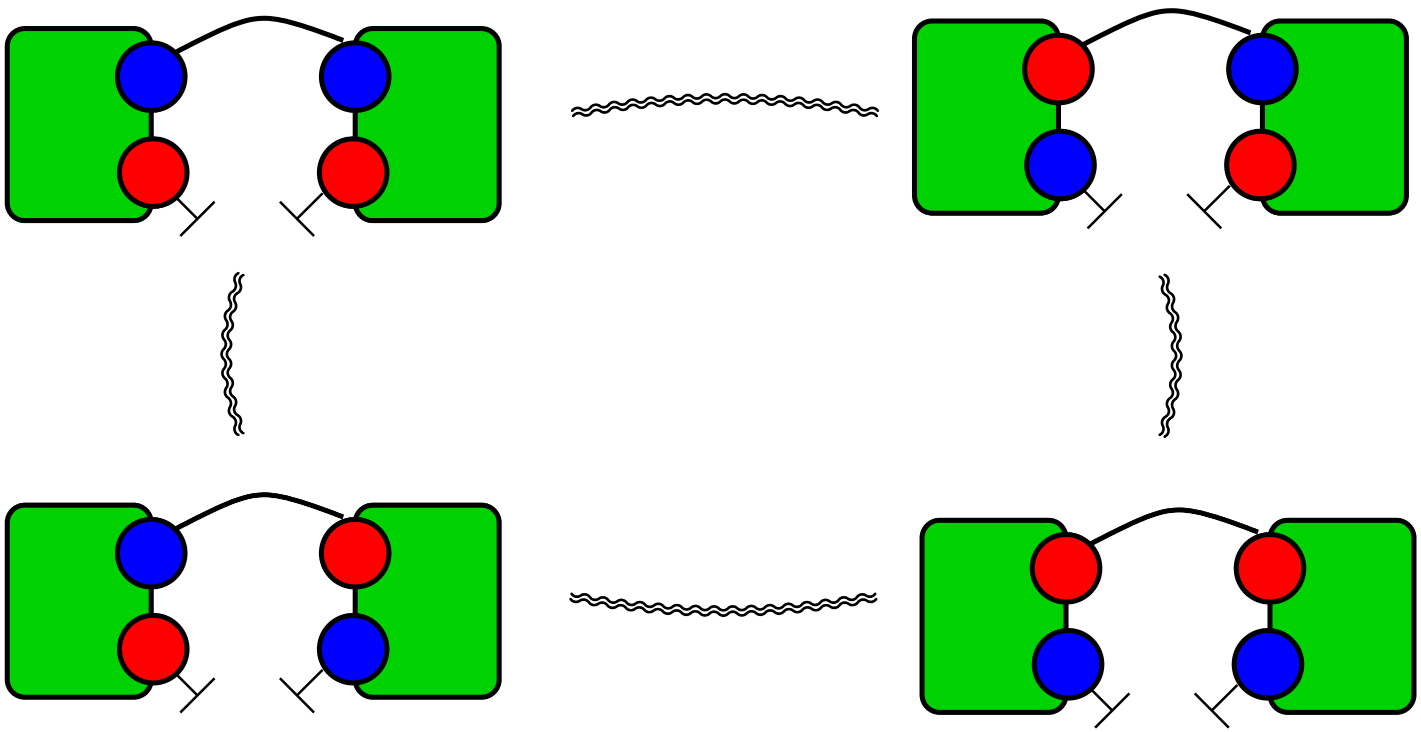
Subgroups of symmetries

Theorem 2
If, for any embedding h between two site graphs G and H:

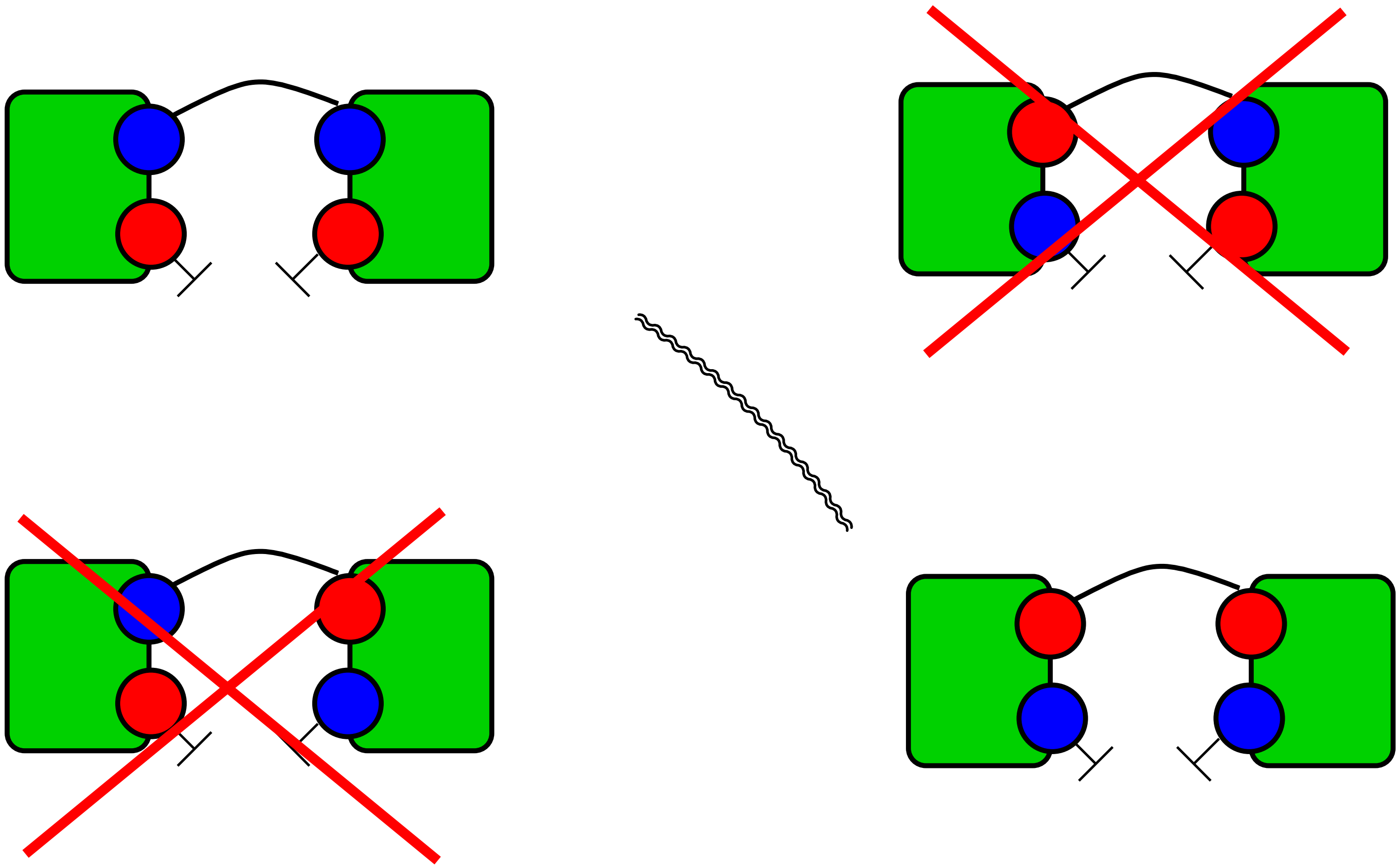
- we have a subset $\mathbb{G}^{'}_{G}$ of $\mathbb{G}_{G}$;
- for any symmetry $\sigma \in \mathbb{G}^{'}_{G}$, $\mathbb{G}^{'}_{G} = \mathbb{G}^{'}_{(\sigma.G)}$;
- for any two σ, σ' symmetries in $\mathbb{G}^{'}_{G}$, $\sigma \circ \sigma' \in \mathbb{G}^{'}_{G}$;
- for any symmetry $\sigma \in \mathbb{G}^{'}_{H}$, $h.\sigma \in \mathbb{G}^{'}_{G}$;

then the groups $(\mathbb{G}^{'}_{G})$ define a set of symmetries.

Example:
Heterogeneous site permutations



Example:
Homogeneous site permutations



Overview

Symmetric model

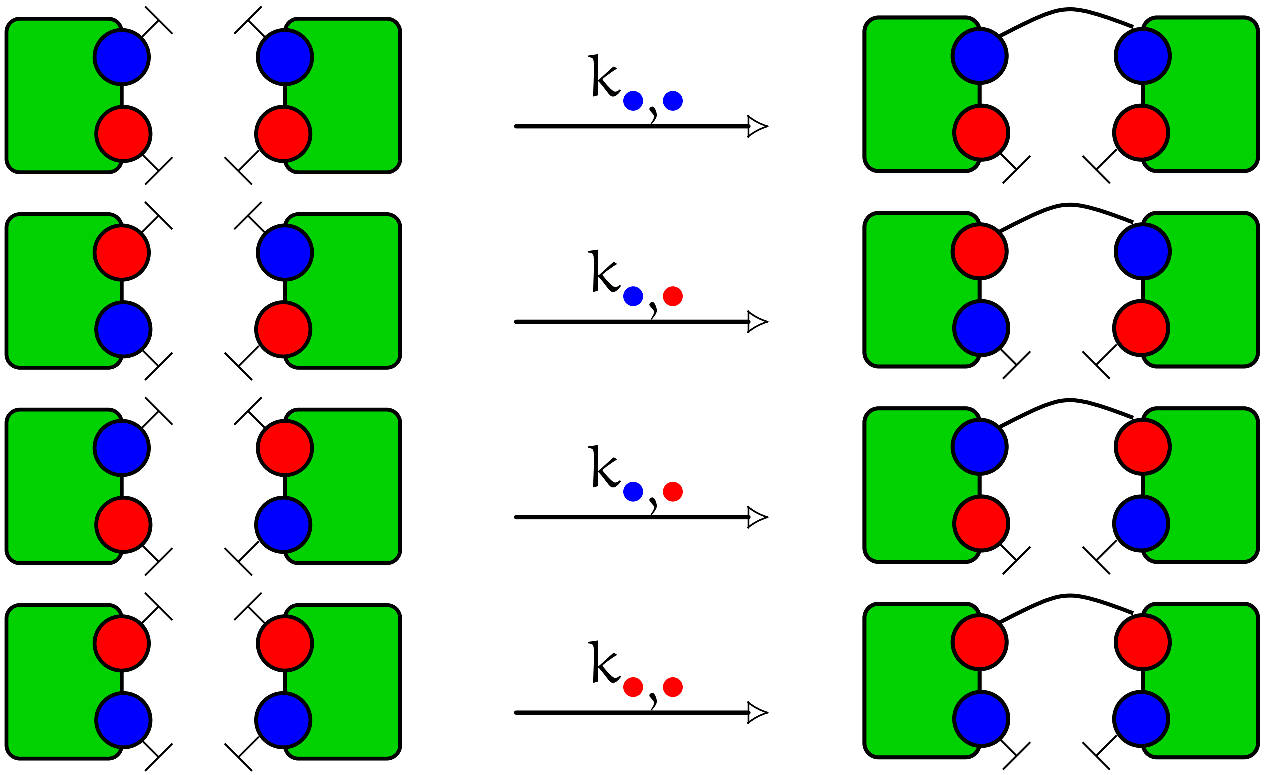
We assume that the model contains atmost one rule per isomorphism class.

A model is $\mathbb{G}$-symmetric if and only if:

- for any rule r in the model and any pair of symmetries $\sigma \in \mathbb{G}_r$, there is (unique) a rule r' in the model that is isomorphic to the rule $\sigma.r$.
- and, with the same notations, we have $g(r) = g(r')$ where:

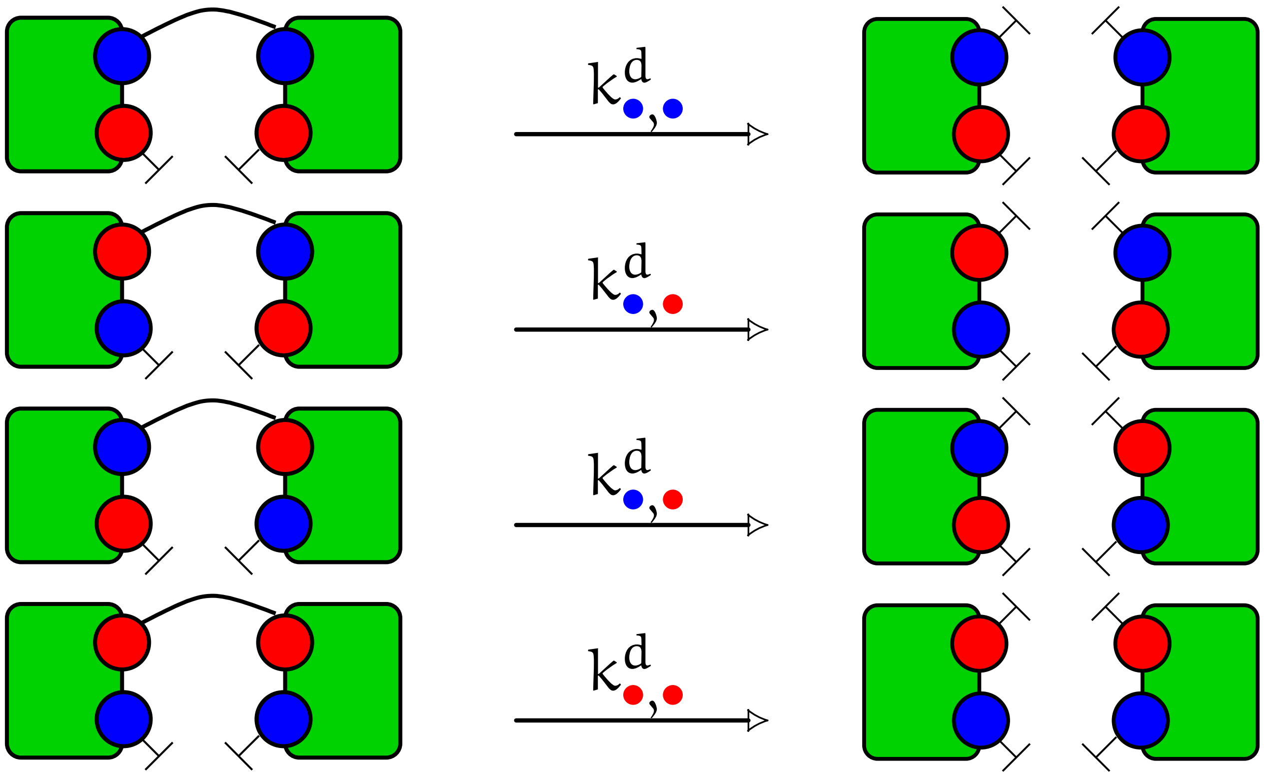
$$g(r) \stackrel{\Delta}{=} \frac{k(r)}{\textit{card}(\{\sigma \in \mathbb{G}_r \mid \sigma.r \text{ and } r \text{ are isomorphic}\})[\textit{lhs}(r), \textit{lhs}(r)]}.$$

Binding rules



$$\frac{k_{\color{blue}\bullet,\color{blue}\bullet}}{1\cdot 2} = \frac{k_{\color{red}\bullet,\color{red}\bullet}}{1\cdot 2} = \frac{k_{\color{blue}\bullet,\color{red}\bullet}}{2\cdot 2}$$

Unbinding rules



$$\frac{k^{d}_{{\color{blue}\bullet},{\color{blue}\bullet}}}{1 \cdot 2} = \frac{k^{d}_{{\color{red}\bullet},{\color{red}\bullet}}}{1 \cdot 2} = \frac{k^{d}_{{\color{blue}\bullet},{\color{red}\bullet}}}{2 \cdot 1}$$

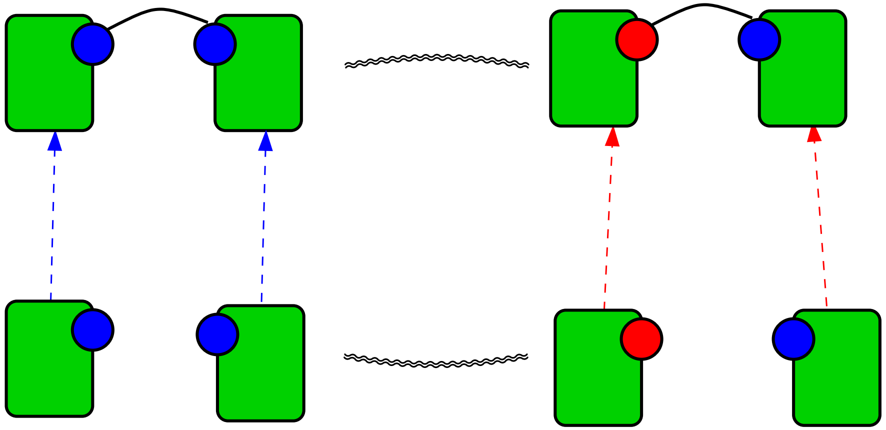
Compatible embeddings

An embedding f between two site graphs G and H is said compatible if and only if:

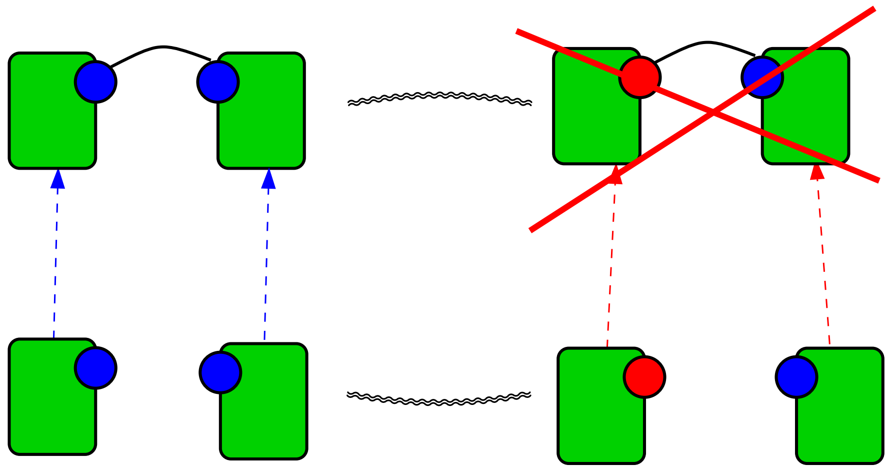
$$\mathbb{G}_G = \{f.\sigma \mid \sigma \in \mathbb{G}_H\}$$

(that is to say that any symmetry that can be applied to the domain of f can be extended to the image of f).

Compatible embeddings may not be preserved by subgroups of symmetries:



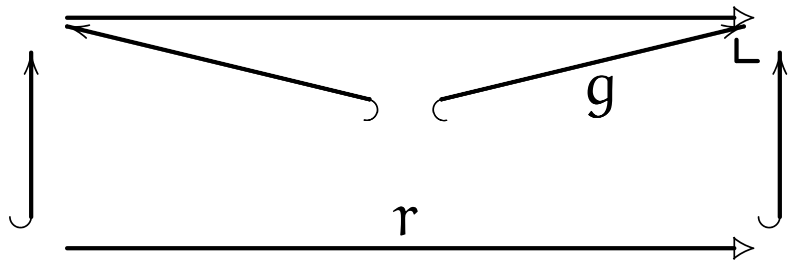
Heterogeneous permutations



Homogeneous permutations

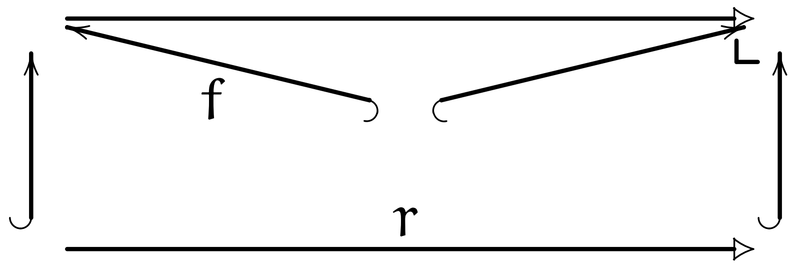
Compatible rules

We say that a rule r is forward-compatible if and only if, for any push-out of the following form:



the embedding g is compatible.

We say that a rule r is backward-compatible if and only if, for any push-out of the following form:



the embedding f is compatible.

Quantitative properties

Theorem 3 Let $\mathbb{G}$ be a set of symmetries and $\mathcal{M}$ be a $\mathbb{G}$-symmetric model. Then:

1. if each rule of r is forward compatible, then we can lump the system. (the proof relies on a forward bisimulation)

2. if each rule of r is both forward and backward compatible, then the following property:

   $$\left[\mathcal{P}(\mathrm{q})[\mathrm{q},\mathrm{q}] = \mathcal{P}(\sigma.\mathrm{q})[\sigma.\mathrm{q},\sigma.\mathrm{q}], \text{for any state } \mathrm{q} \text{ and any symmetry } \sigma \in \mathbb{G}_{\mathrm{q}}\right]$$

   is an invariant of the system. (the proof relies on a backward bisimulation)

Overview

1. Context and motivations
2. Kappa semantics
3. Symmetries in site-graphs
4. Symmetric models
5. Conclusion

Conclusion

A fully algebraic framework to infer and use symmetries in Kappa;

- Compatible with the SPO semantics (see [FSTTCS'2012]);
- Can handle side-effects (see the paper);
- Induces forward and/or back and forth bisimulations;
- Can be applied to discover model reductions for the qualitative semantics, the ODEs semantics, and the stochastic semantics [MFPSXXVII];
- Can be combined with other exact model reductions [MFPSXXVI].

This framework is cleaner and more general that the process algebra based one [MFPSXXVII].

Camporesi et al., Combining model reductions. MFPS XXVI (2010)
Camporesi et al., Formal reduction of rule-based models, MFPS XXVII (2011)
Danos et al., Rewriting and Pathway Reconstruction for Rule-Based Models, FSTTCS 2012

Future work

- Investigate which specific classes of symmetries and which specific classes of rules ensure that rules are forward and/or backward compatible with the symmetries;
- Check the compatibility with the DPO (Double Push-Out) semantics;
- Design approximate symmetries using bisimulation metrics
  (ask Norman Ferns, Post-doc at ÉNS).

Thank you !!!

We acknowledge the support of:

1. the AbstractCell ANR-Chair of Excellence 2009-2013
2. the ExeK project (Big Mechanism DARPA Program) 2014-2018.

We have open positions:

- at ÉNS-Ulm,
- at ÉNS-Lyon,
- at Paris-Diderot university,
- and at Harvard Medical School

for post-doc researchers and research engineers on the ExeK DARPA project.