ROSAEC workshop

First session

Reachability Analysis of Kappa Models

Jérôme Feret

Laboratoire d'Informatique de l'École Normale Supérieure

INRIA, ÉNS, CNRS

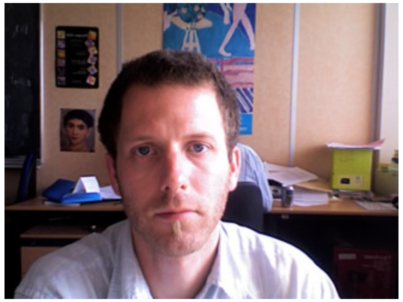
http://www.di.ens.fr/~feret

January, 2009

Joint-work with...

Walter Fontana
Harvard Medical School

Vincent Danos
Edinburgh



Russ Harmer
Paris VII

Jean Krivine
Harvard Medical School

Overview

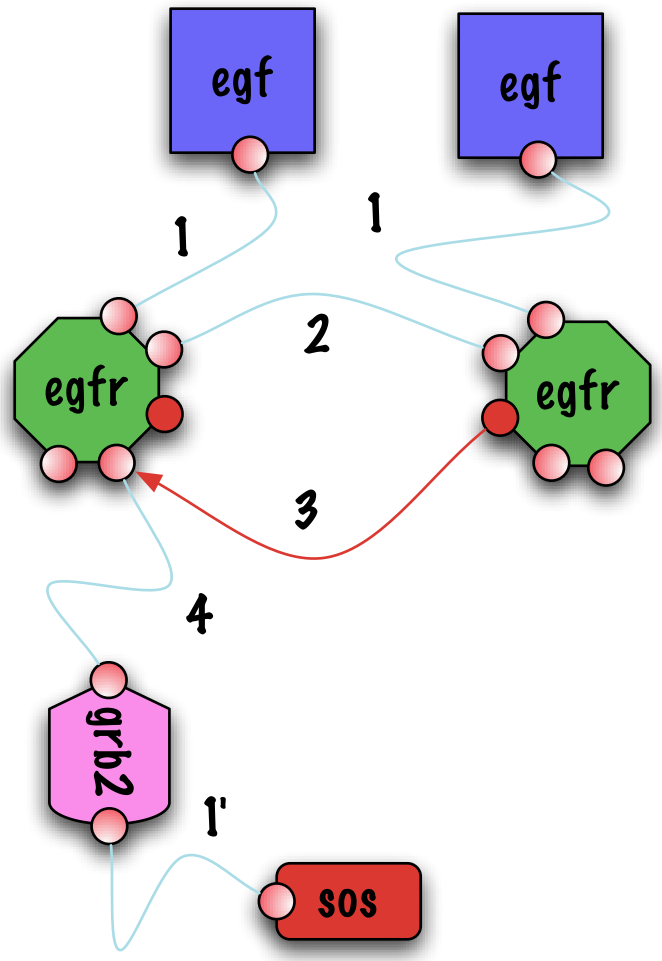
1. Introduction
2. Outline
3. Kappa language
4. Local views
5. Local set of complexes
6. Local rule systems
7. Decontextualization
8. Conclusion

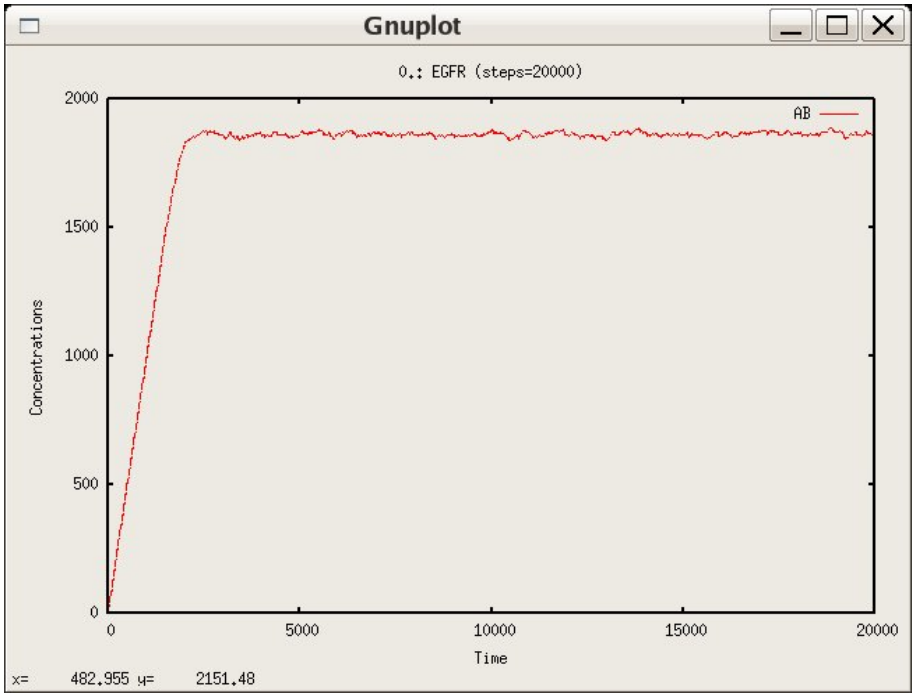
Modeling signaling pathway

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

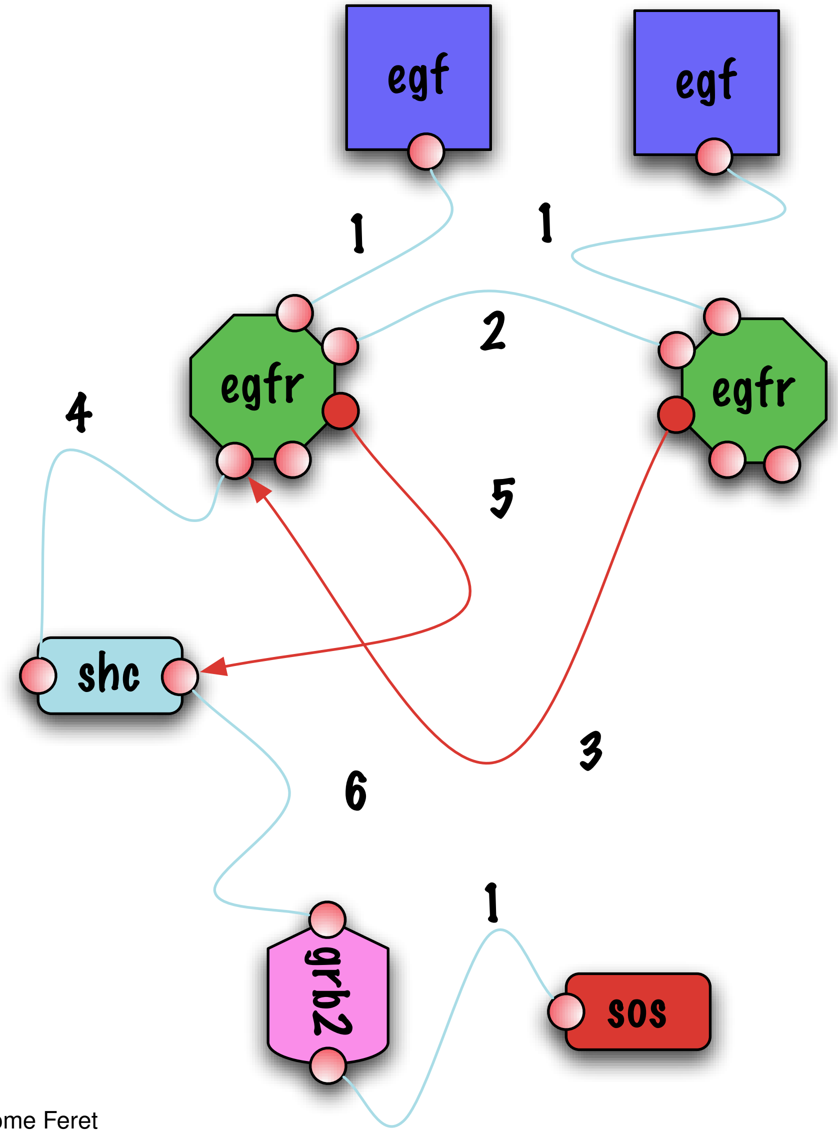
  We propose some static analysis tools in order to:
  - help the design of Kappa models;
  - compute (*abstract*) the properties of Kappa models.

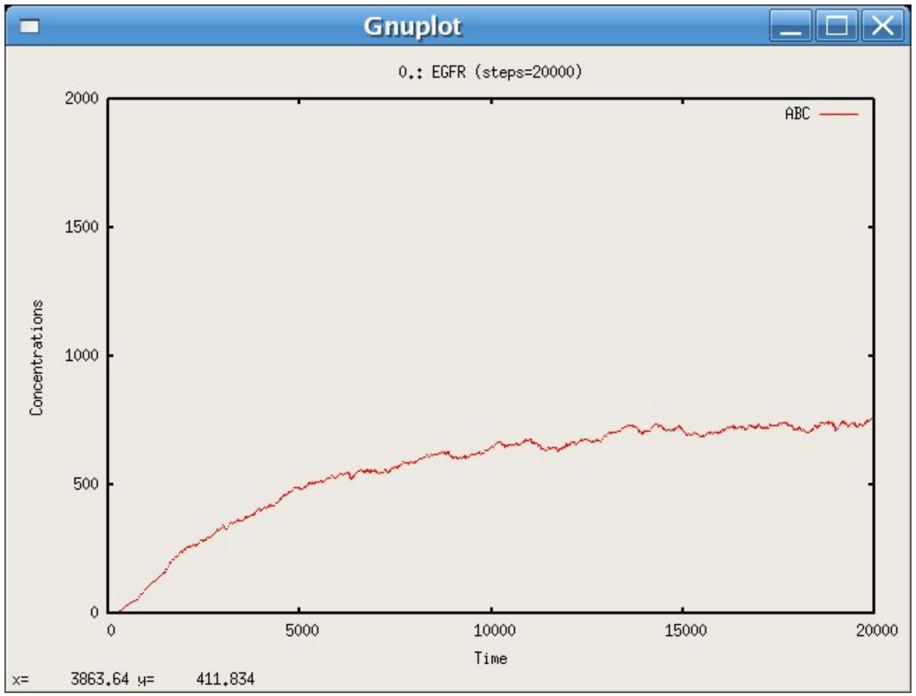
A single story





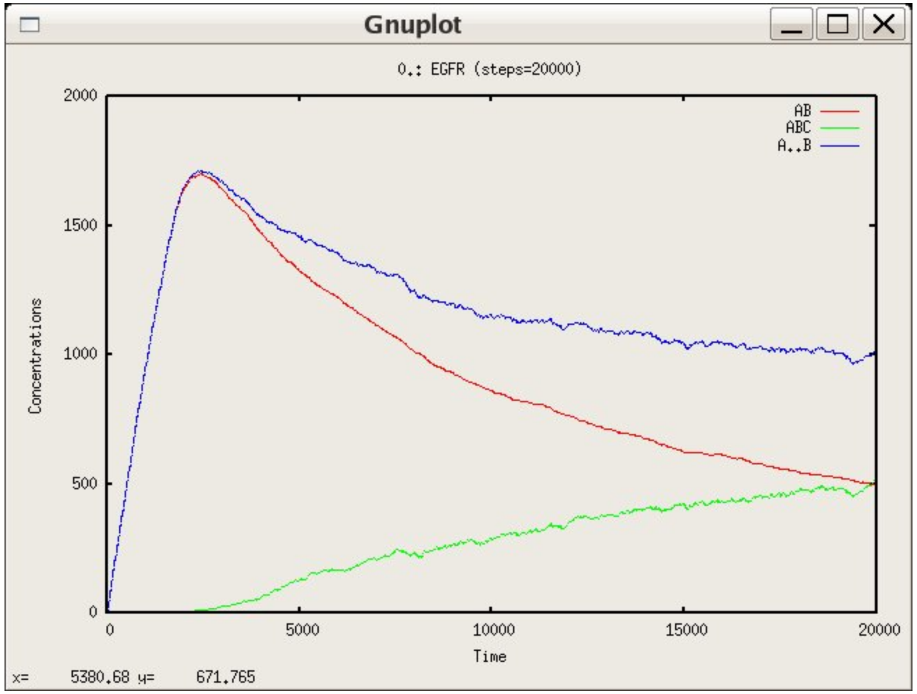
Another concurrent story





Overshoot

When we combine the two stories. . .



. . . we get an overshoot.

Overview

1. Introduction
2. Outline
3. Kappa language
4. Local views
5. Local set of complexes
6. Local rule systems
7. Decontextualization
8. Conclusion

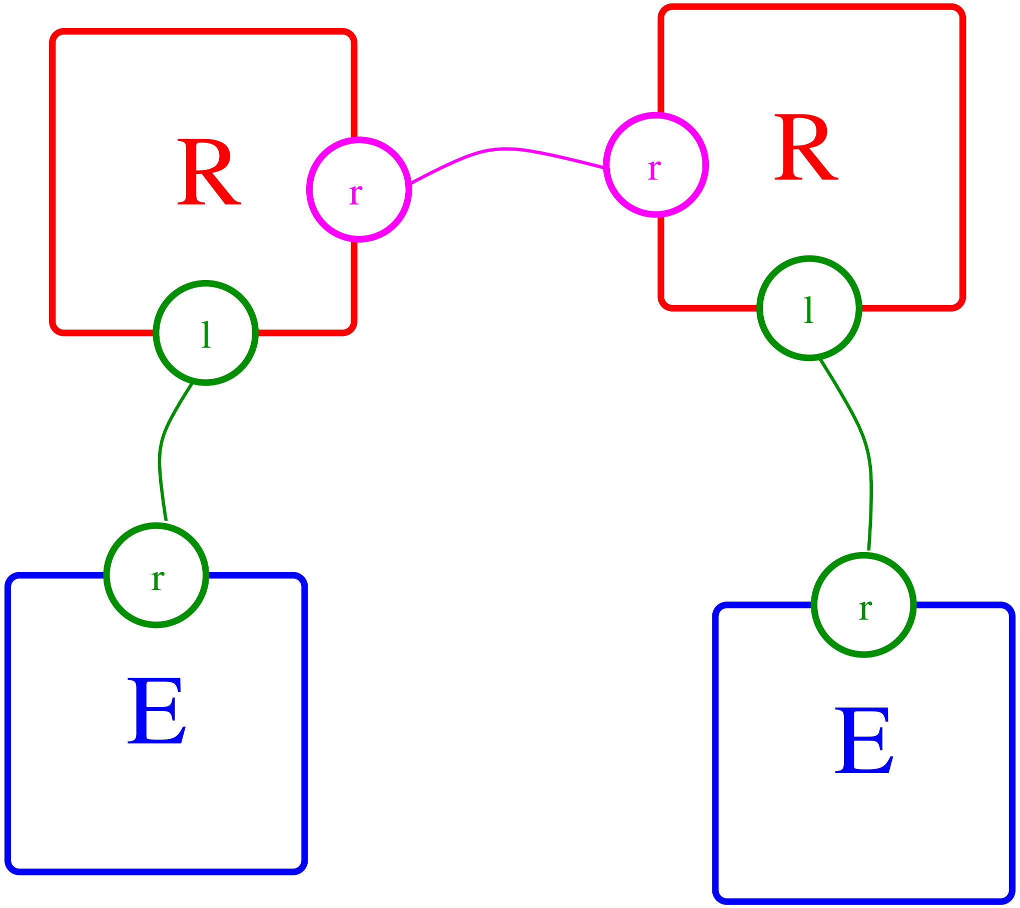
Outline

- Language: $\mathcal{L}$
- Semantics domains: $D \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} D^\sharp$
- Semantics functions: $[\![p]\!] \subseteq \gamma([\![p]\!]^\sharp)$
- (*Strong*) completeness ?
  for which $p \in \mathcal{L}$ do we have $[\![p]\!] = \gamma([\![p]\!]^\sharp)$?
  - Semantics characterization:
    $A \subseteq D$ such that: if $[\![p]\!] \in A$ then $[\![p]\!] = \gamma([\![p]\!]^\sharp)$;
  - Syntactic characterization:
    $\mathcal{K} \subseteq \mathcal{L}$ such that: if $p \in \mathcal{K}$ then $[\![p]\!] \in A$;
  - Driven program transformation:
    $T : D^\sharp \times \mathcal{L} \to \mathcal{L}$ such that: if $[\![p]\!] \subseteq \gamma(A^\sharp)$ then $\begin{cases} [\![p]\!] = [\![T(A^\sharp, p)]\!], \\ p \in \mathcal{K} \Rightarrow T(A^\sharp, p) \in \mathcal{K}. \end{cases}$

Overview

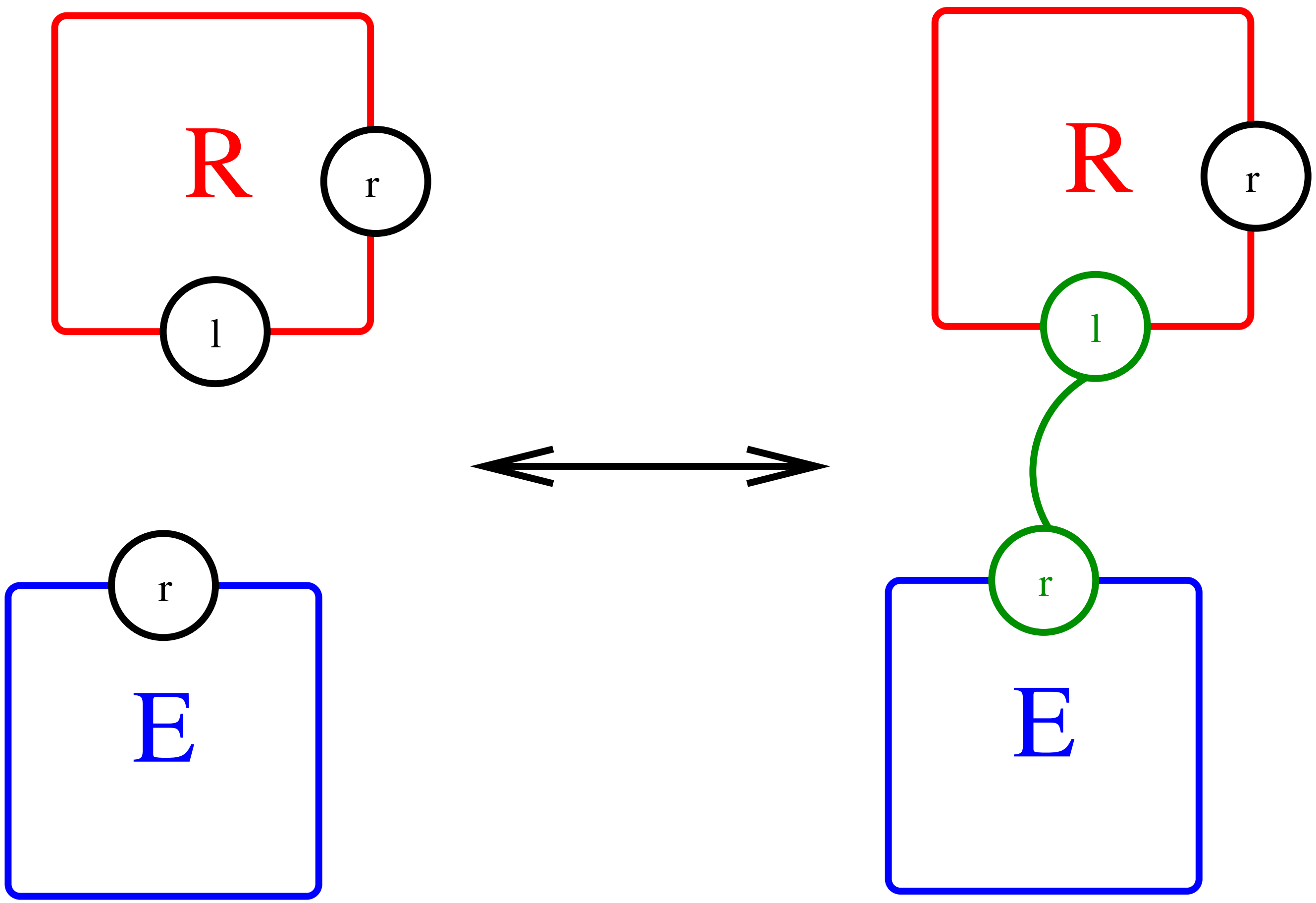
1. Introduction
2. Outline
3. Kappa language
4. Local views
5. Local set of complexes
6. Local rule systems
7. Decontextualization
8. Conclusion

A complex



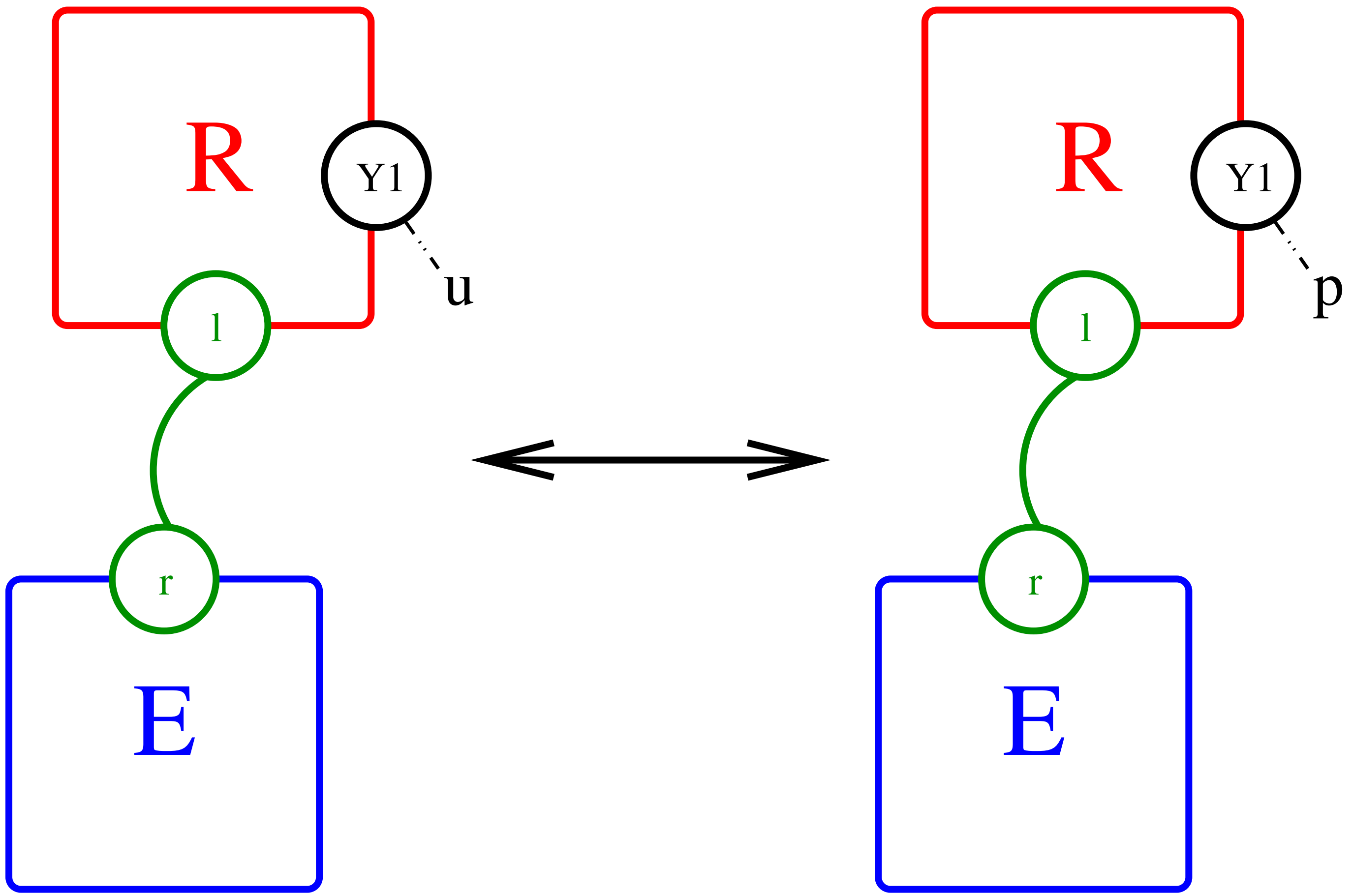
E(r!1), R(l!1,r!2), R(r!2,l!3), E(r!3)

A Unbinding/Binding Rule



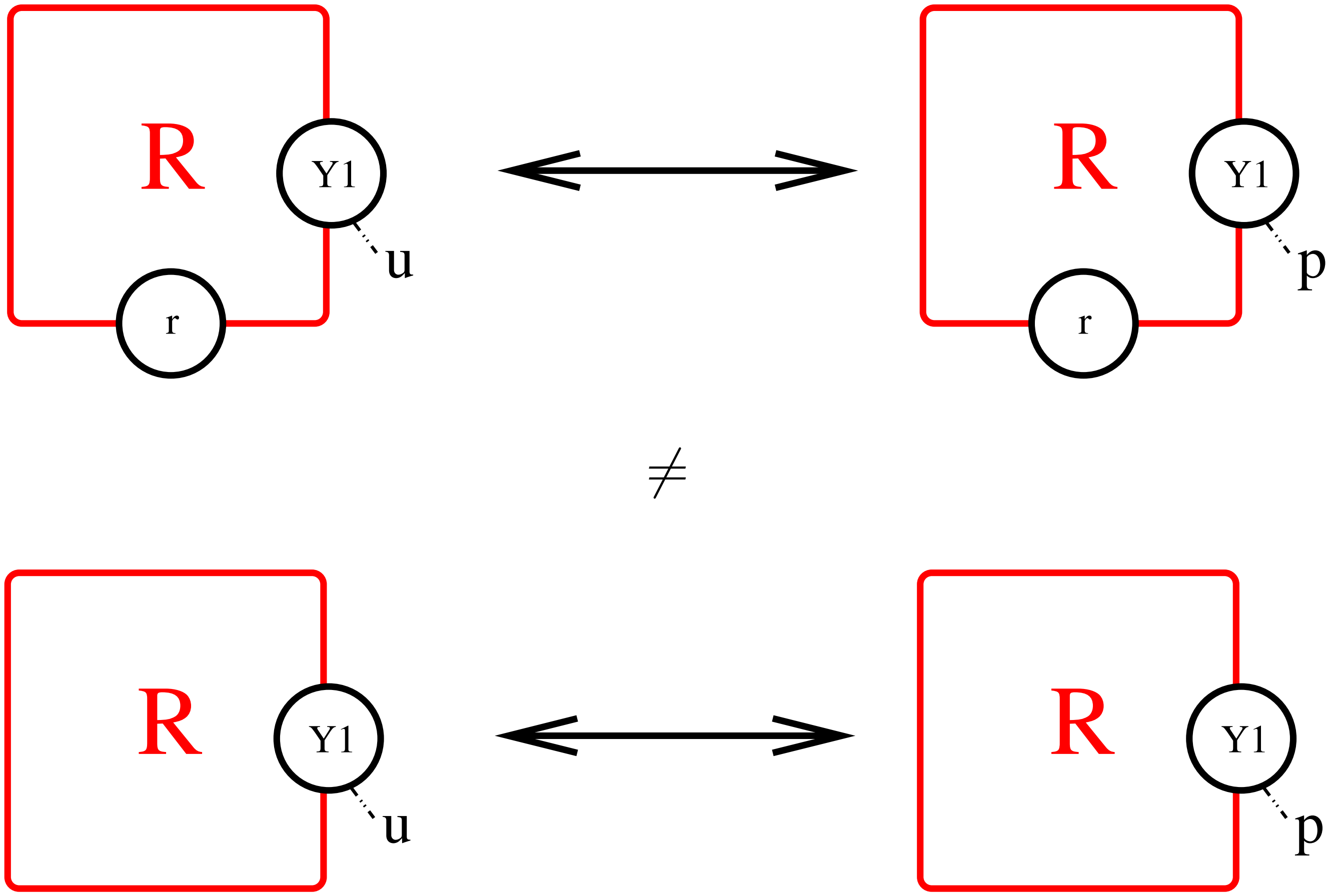
E(r), R(l,r) $\longleftrightarrow$ E(r!1), R(l!1,r)

Internal state

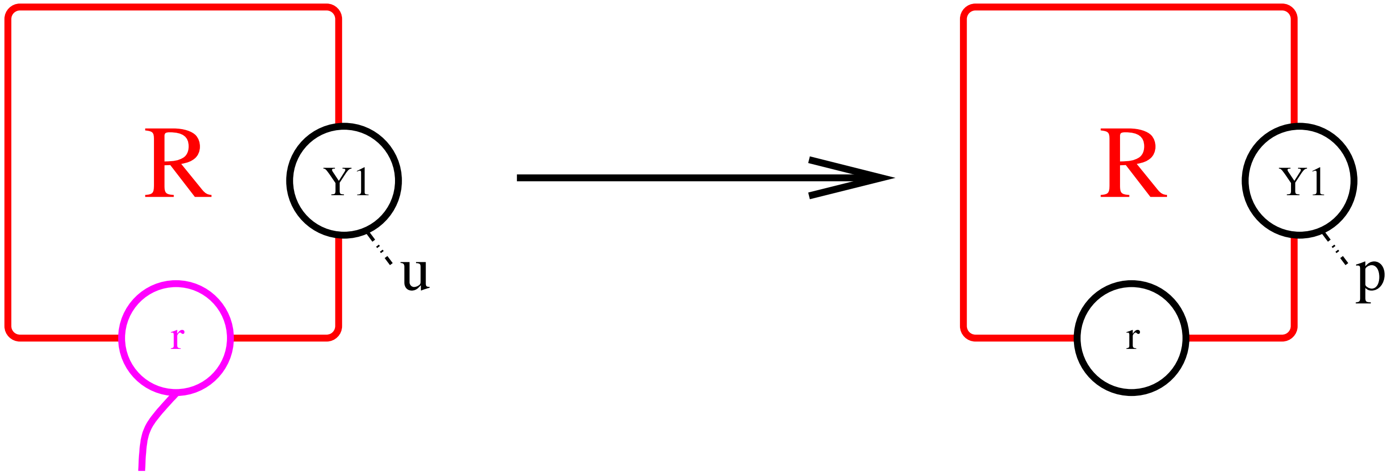


R(Y1~u,l!1), E(r!1) $\longleftrightarrow$ R(Y1~p,l!1), E(r!1)

Don’t care, Don’t write

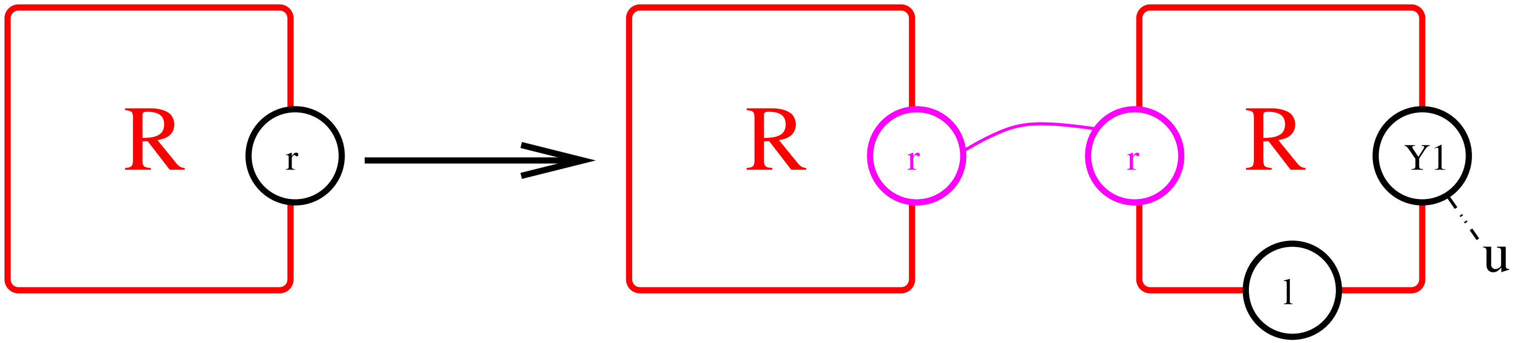


A contextual rule

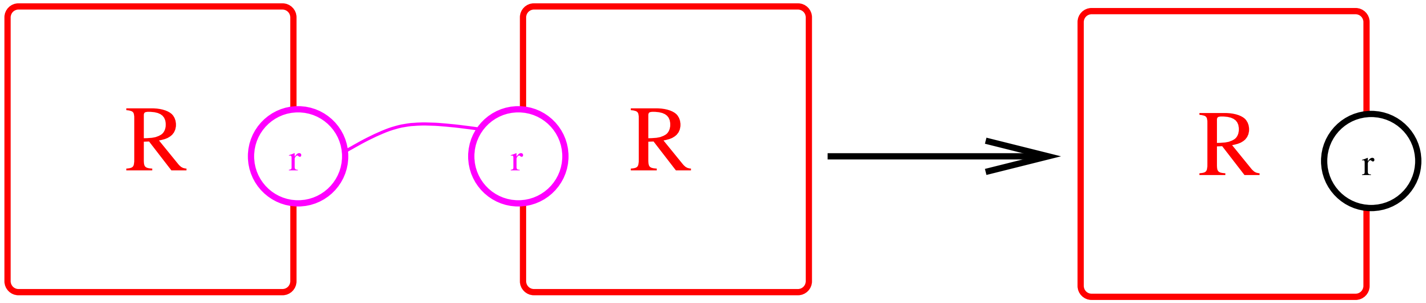


R(Y1~u,r!_) $\rightarrow$ R(Y1~p,r)

Creation/Suppression



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



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

Early EGF example

egf rules 1

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

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

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

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

refined from
$Sh(Y7^p) \rightarrow Sh(Y7)$

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

refined from
$R(Y68^p)+G(a) \leftrightarrow R(Y68^{p1})+G(a^1)$

Early EGF example

egf rules 2

refined from
So(d)+G(b)<->So(d¹)+G(b¹)

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

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

why not simply $G(b^3)$??

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

refined from
$Sh(\pi), G(a)$<->$Sh(\pi^1), G(a^1)$

Overview

1. Introduction
2. Outline
3. Kappa language
4. Local views
5. Local set of complexes
6. Local rule systems
7. Decontextualization
8. Conclusion

Set of reachable complexes

Let $\mathcal{R} = \{R_i\}$ be a set of rules.
Let *Complex* be the set of all complexes ($C, c_1, c_1', \ldots, c_k, c_k', \ldots \in$ *Complex*).
Let *Complex*$_0$ be the set of initial complexes.
We write:

$$c_1, \ldots, c_m \rightarrow_{R_k} c_1', \ldots, c_n'$$

whenever:

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

We are interested in *Complex*$_\omega$ the set of all complexes that can be constructed in one or several applications of rules in $\mathcal{R}$ starting from the set *Complex*$_0$ of initial complexes.

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

Inductive definition

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

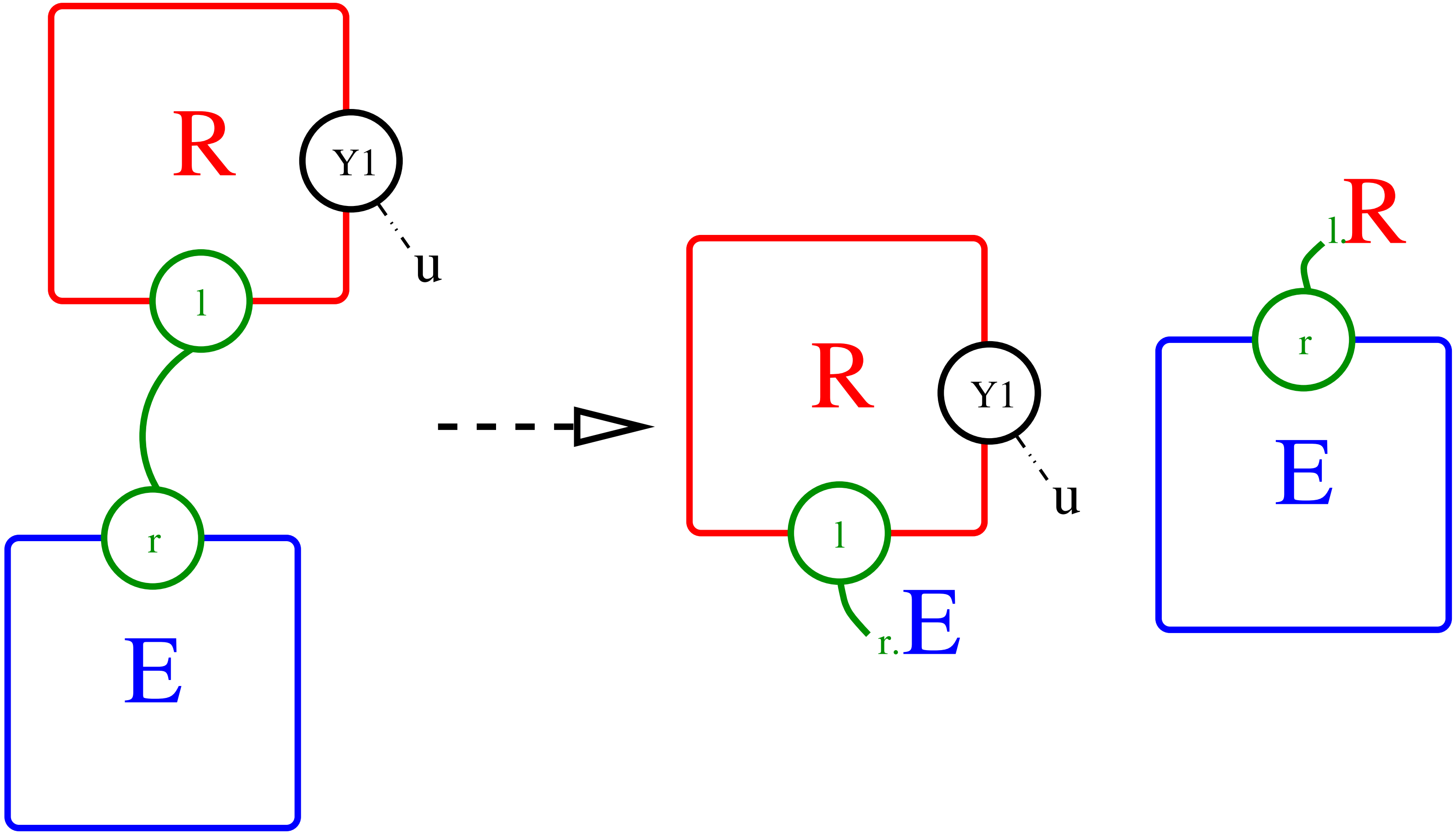
$$\mathbb{F} : \begin{cases} \wp(\mathit{Complex}) & \to \wp(\mathit{Complex}) \\ X & \mapsto X \cup \left\{ c'_j \;\middle|\; \begin{array}{l} \exists R_k \in \mathcal{R}, c_1, \ldots, c_m \in X, \\ c_1, \ldots, c_m \to_{R_k} c'_1, \ldots, c'_n \end{array} \right\} \end{cases}.$$

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

We have:

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

Local views



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

Galois connexion

Let *local_view* be the set of all local views.

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

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

Thus, it defines a Galois connexion:

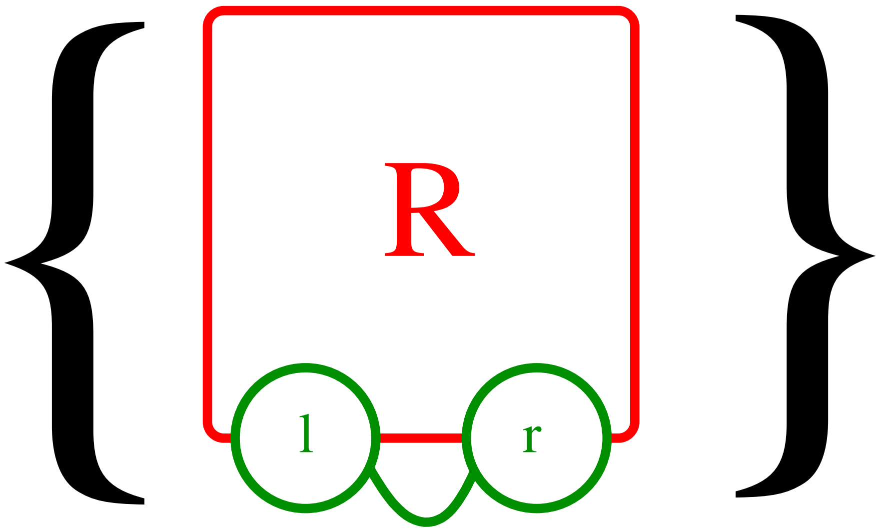
$$\wp(\mathit{Complex}) \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \wp(\mathit{local\_view}).$$

(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.

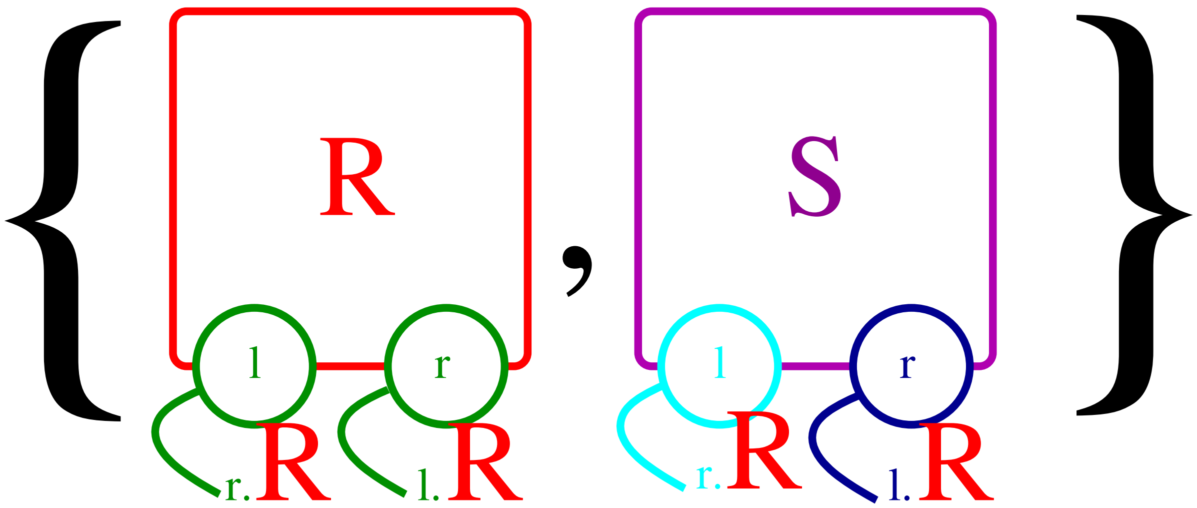
Guess the image of the following set of complexes ?



$\alpha \circ \gamma$

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

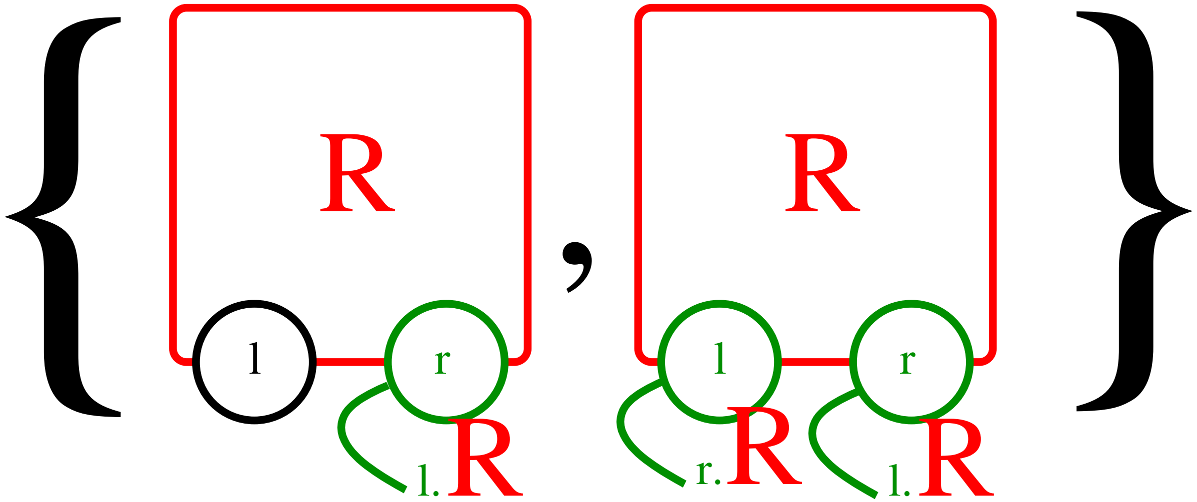
Guess the image of the following set of local views ?



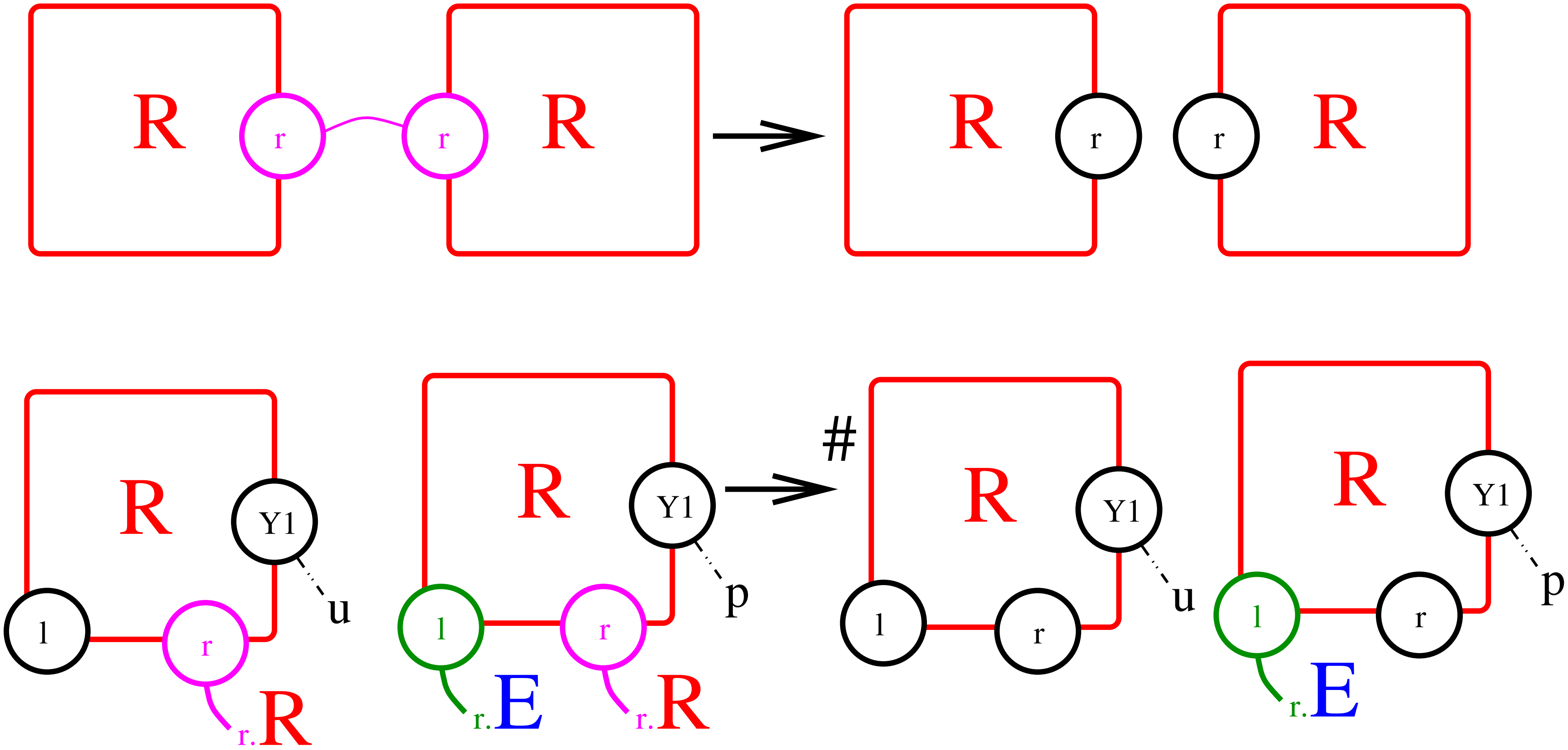
One more question

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

Guess the image of the following set of local views ?



Abstract rules



Abstract counterpart to $\mathbb{F}$

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

$$\mathbb{F}^\sharp : \begin{cases} \wp(\mathit{local\_view}) & \rightarrow \wp(\mathit{local\_view}) \\ X & \mapsto X \cup \left\{ lv'_j \;\middle|\; \begin{array}{l} \exists R_k \in \mathcal{R}, lv_1, \ldots, lv_m \in X, \\ lv_1, \ldots, lv_m \rightarrow^\sharp_{R_k} lv'_1, \ldots, lv'_n \end{array} \right\}. \end{cases}$$

We have:

- $\mathbb{F}^\sharp$ is extensive;
- $\mathbb{F}^\sharp$ is monotonic;
- $\mathbb{F} \circ \gamma \mathrel{\dot{\subseteq}} \gamma \circ \mathbb{F}^\sharp$;
- $\mathbb{F}^\sharp \circ \alpha = \alpha \circ \mathbb{F} \circ \gamma \circ \alpha$.

Soundness

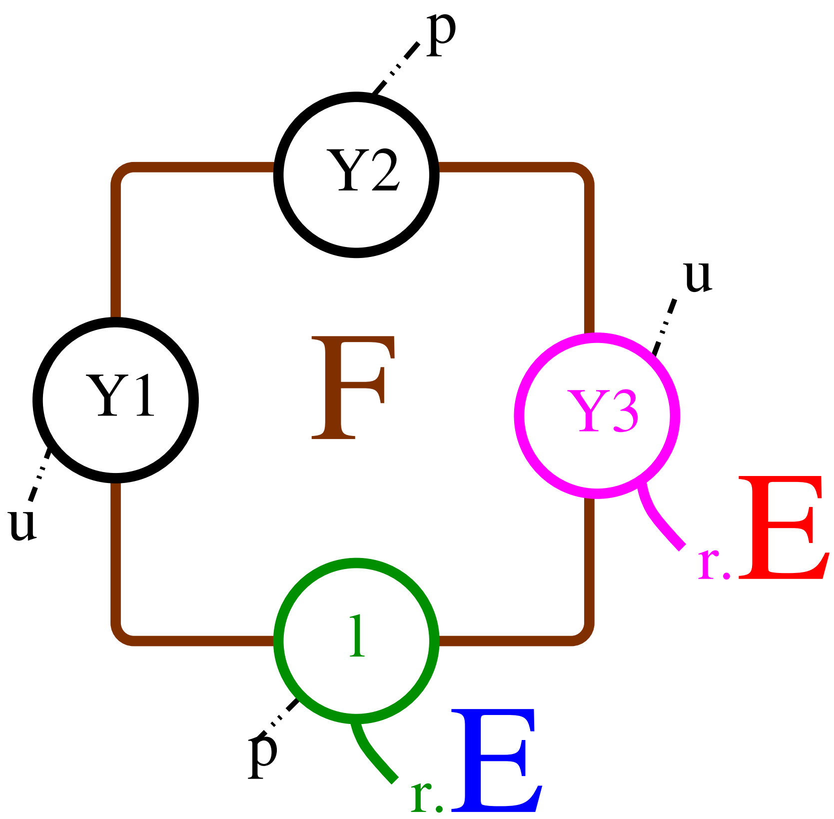
Theorem 1 Let:

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

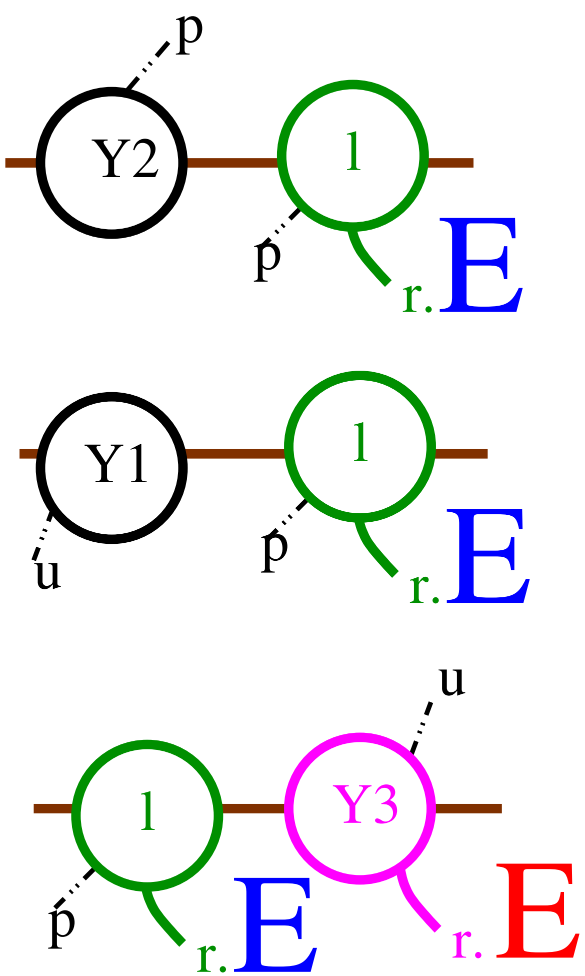
Then:

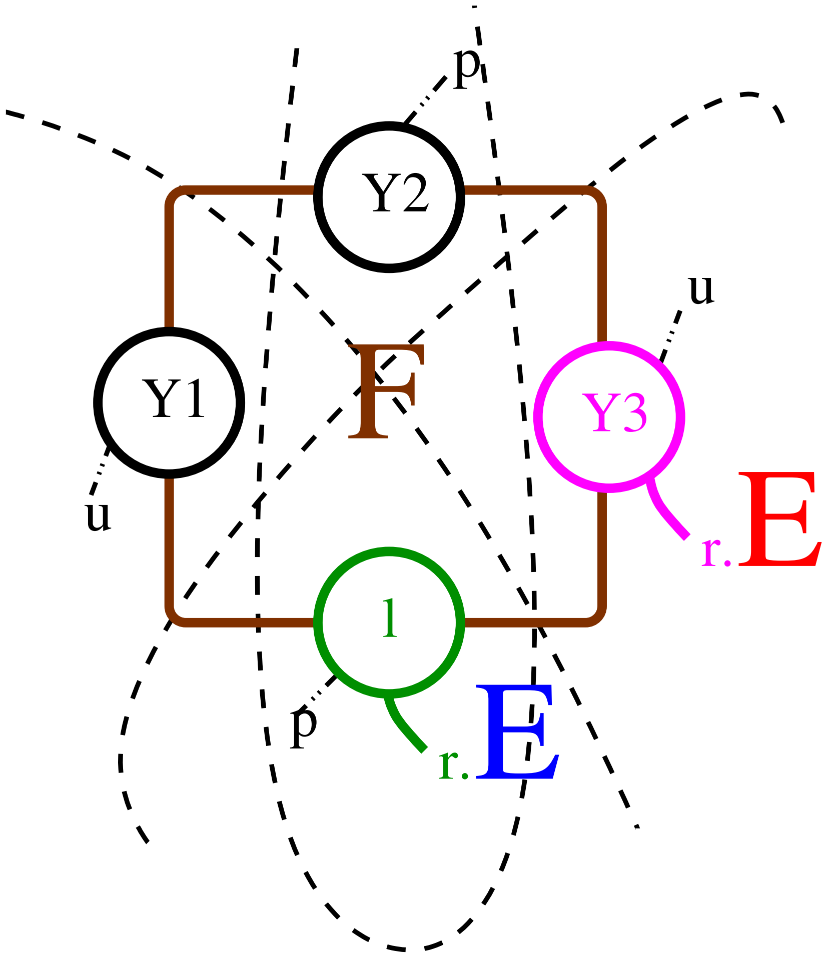
1. both $\mathit{lfp}_{x_0}\mathbb{F}$ and $\mathit{lfp}_{\alpha(x_0)}\mathbb{F}^\sharp$ exist,
2. $\mathit{lfp}_{x_0}\mathbb{F} \subseteq \gamma(\mathit{lfp}_{\alpha(x_0)}\mathbb{F}^\sharp)$.

Combinatorial blow up



Avoiding combinatorial blow up





Packing

- packing strategy: a covering $\mathcal{P}$ of the set of sites.
- template: a class $p \in \mathcal{P}$ in the covering.
- projection: for any $p \in \mathcal{P}$, the function Π_p restricts a local view to p.
- sub-local-view: $\textit{sub\_local\_view} = \{\Pi_p(\textit{lv}) \mid \textit{lv} \in \textit{local\_view},\ p \in \mathcal{P}\}$.
- abstraction: the following function

$$\alpha_{\text{PACKS}} : \begin{cases} \wp(\textit{local\_view}) & \rightarrow \wp(\textit{sub\_local\_view}) \\ X & \mapsto \{\Pi_p(\textit{lv}) \mid \textit{lv} \in X,\ p \in \mathcal{P}\}. \end{cases}$$

  is a $\cup$-complete morphism.
  It defines a Galois connexion $\wp(\textit{local\_view}) \underset{\alpha_{\text{PACK}}}{\overset{\gamma_{\text{PACK}}}{\leftrightarrows}} \wp(\textit{sub\_local\_view})$.
- counterpart: $\mathbb{F}^\sharp_{\text{PACK}} = \alpha_{\text{PACK}} \circ \mathbb{F}^\sharp \circ \gamma_{\text{PACK}}$.

Automatic packing

Let P be a covering such that:

1. each site is at least in one class of P;
2. whenever a rule modifies two sites a and b, then, for any $p \in P$, $a \in p \iff b \in p$;
3. whenever a rule modifies a site a and tests a site b, then for any $p \in P$, $a \in p \implies b \in p$;
4. $\mathit{Complex}_0 \in \gamma_{\mathsf{PACKS}}(\wp(\mathit{sub\_local\_view}))$.

Then, for any $n \in \mathbb{N}$,

$$\gamma_{\mathsf{PACKS}}(\mathbb{F}^{\sharp\, n}_{\mathsf{PACKS}}(\alpha_{\mathsf{PACKS}}(\mathit{Complex}_0))) = \mathbb{F}^{\sharp n}(\mathit{Complex}_0).$$

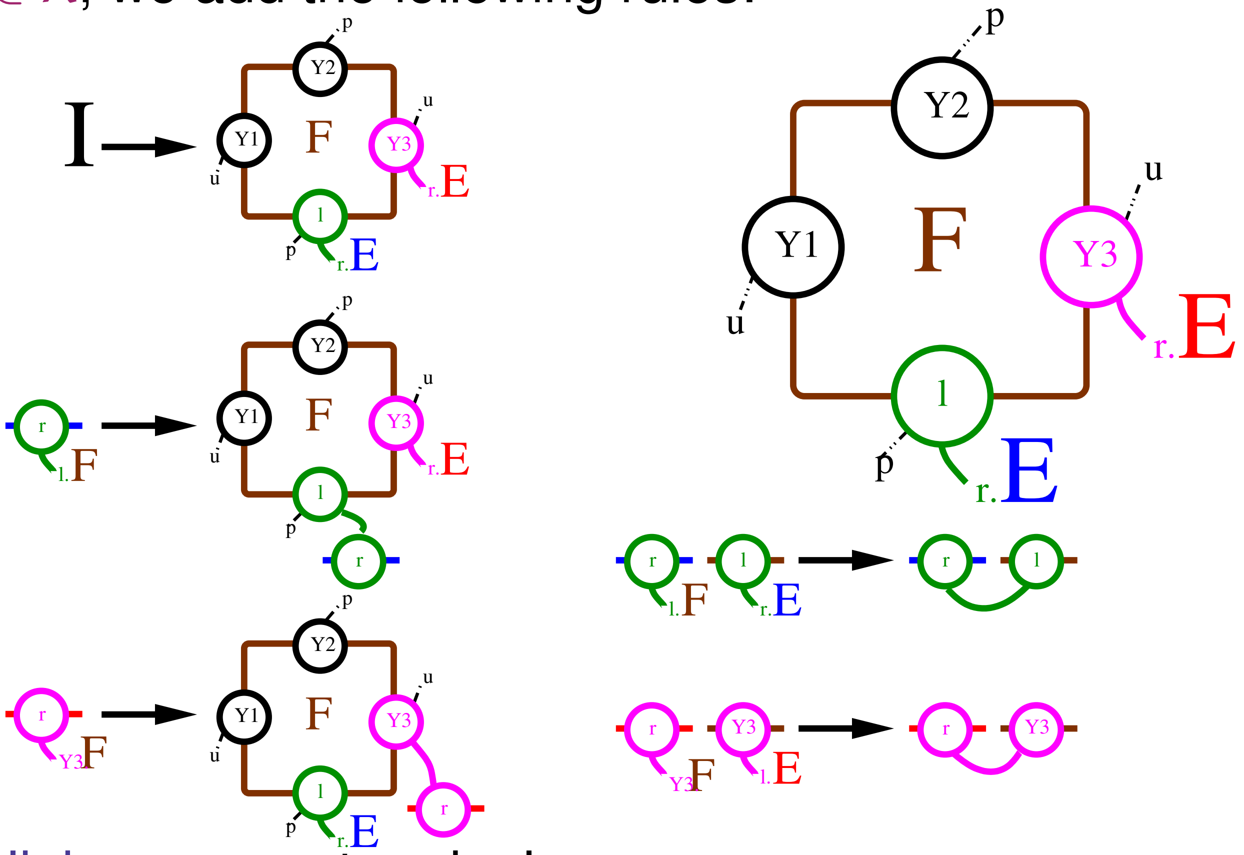
So, automatic packing never changes analysis results !!!

Overview

1. Introduction
2. Outline
3. Kappa language
4. Local views
5. Local set of complexes
6. Local rule systems
7. Decontextualization
8. Conclusion

Concretization

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



I and semi-links are non-terminal.

Pumping lemma

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

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

     can be built.
- We only enumerate complexes that are reached through an acyclic rewriting computation.
- We prove later that: if $X \in \alpha(\wp(\mathit{Complex}))$ then each partial rewrite sequence is the prefix of a terminating rewrite sequence.

Examples

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

Which information is abstracted away ?

Our analysis is exact (no false positive):

- for EGF cascade (356 complexes);
- for FGF cascade (79080 complexes);

We know how to build systems with false positives. . .

. . .but they seem to be biologically meaningless.

This raises the following issues:

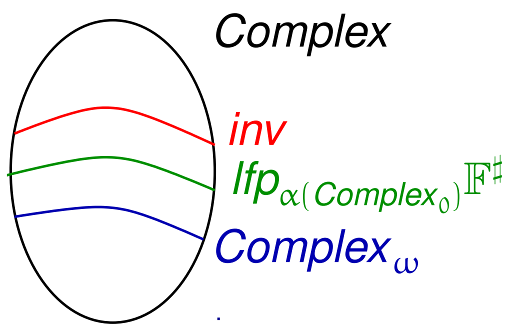
- Can we characterize which information is abstracted away ?
- Which is the form of the systems, for which we have no false positive ?
- Do we learn something about the biological systems that we describe ?

Which information is abstracted away ?

Theorem 2 We suppose that:

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

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



When is there no false positive ?

Theorem 3 We suppose that:

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

Then:

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

Local set of complexes

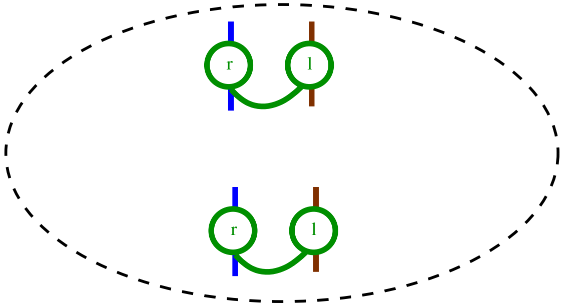
Definition 1 We say that a set $X \in \wp(\mathit{Complex})$ of complexes is local if and only if it satisfies $X = \gamma(\alpha(X))$.

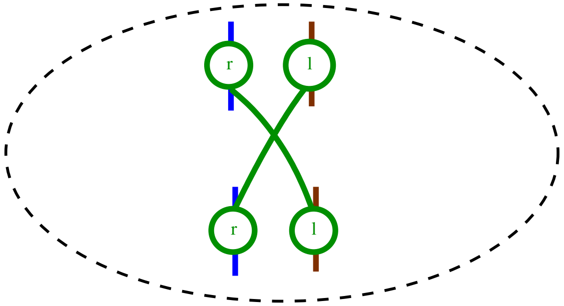
Remark 1 *A set* $X \in \wp(\mathit{Complex})$ *is local if and only if* $X \in \gamma(\wp(\mathit{local\_view}))$*.*

Swapping relation

We define the binary relation $\overset{\text{SWAP}}{\sim}$ among tuples *Complex** of complexes.

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 $X \in \wp(\textit{Complex})$ be a set of complexes.
The two following assertions are equivalent:

1. the set $X \in \wp(\textit{Complex})$ is local (i.e. $\gamma(\alpha(X)) = X$);
2. for any tuples $(C_i), (D_j) \in \textit{Complex}^*$ 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:

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

Then:

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

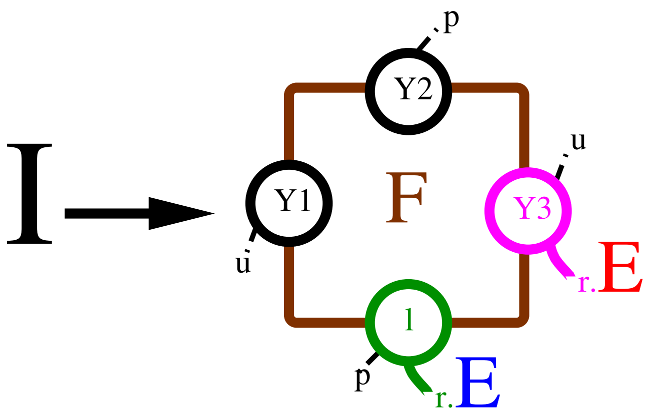
Proof (more difficult implication way)

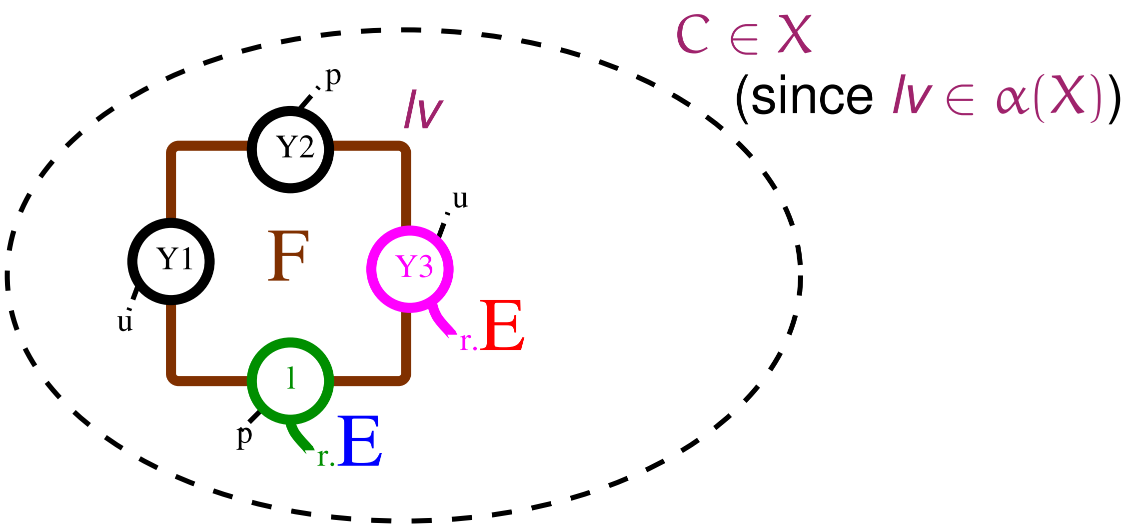
We suppose that X is close with respect to $\overset{\text{SWAP}}{\sim}$.
We want to prove that $\gamma(\alpha(X)) \subseteq X$.

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

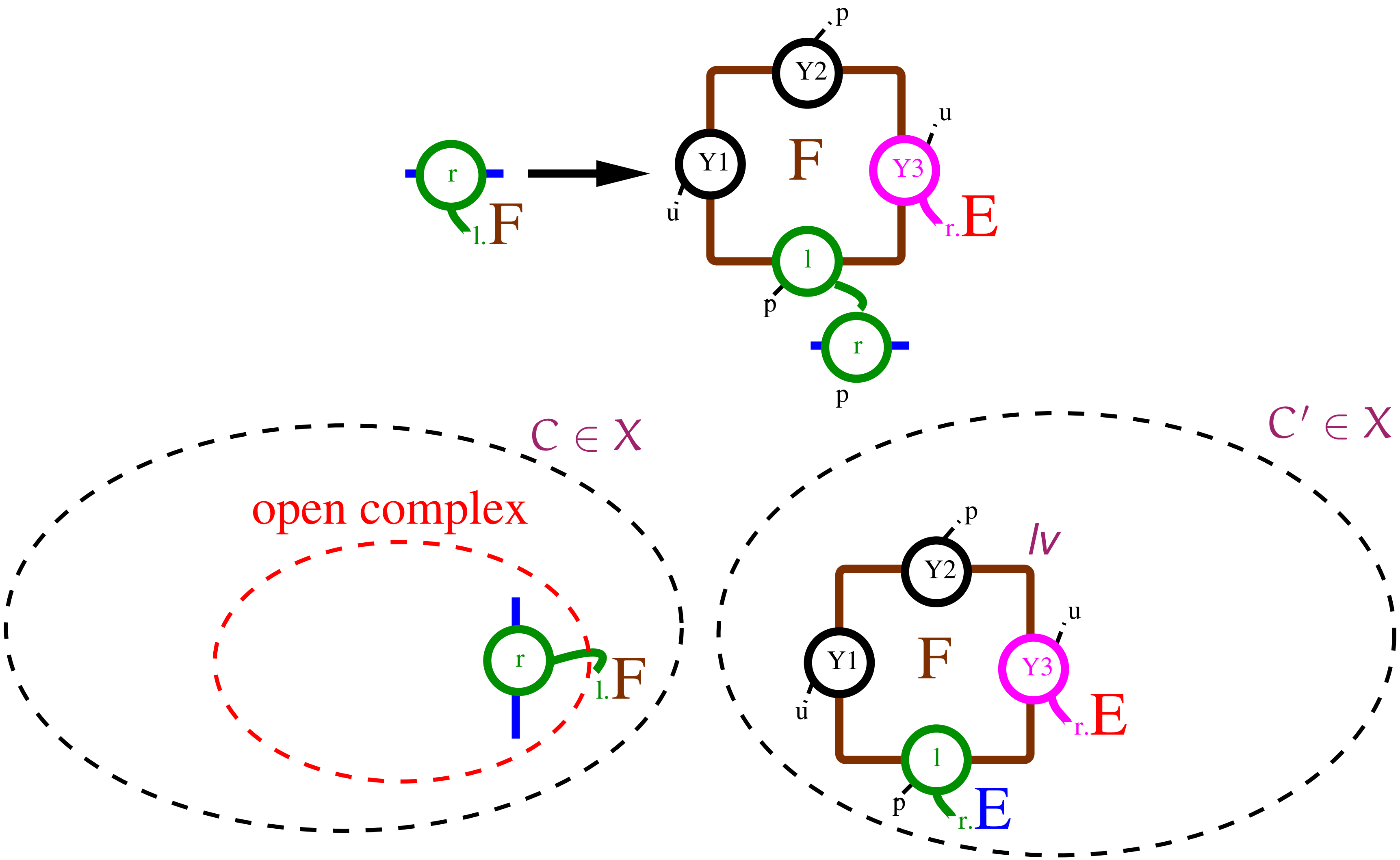
- By def. of α, this is satisfied for any local view in $\alpha(X)$;
- This remains satisfied after unfolding a semi-link with a local view;
- This remains satisfied after binding two semi-links.

Initialization

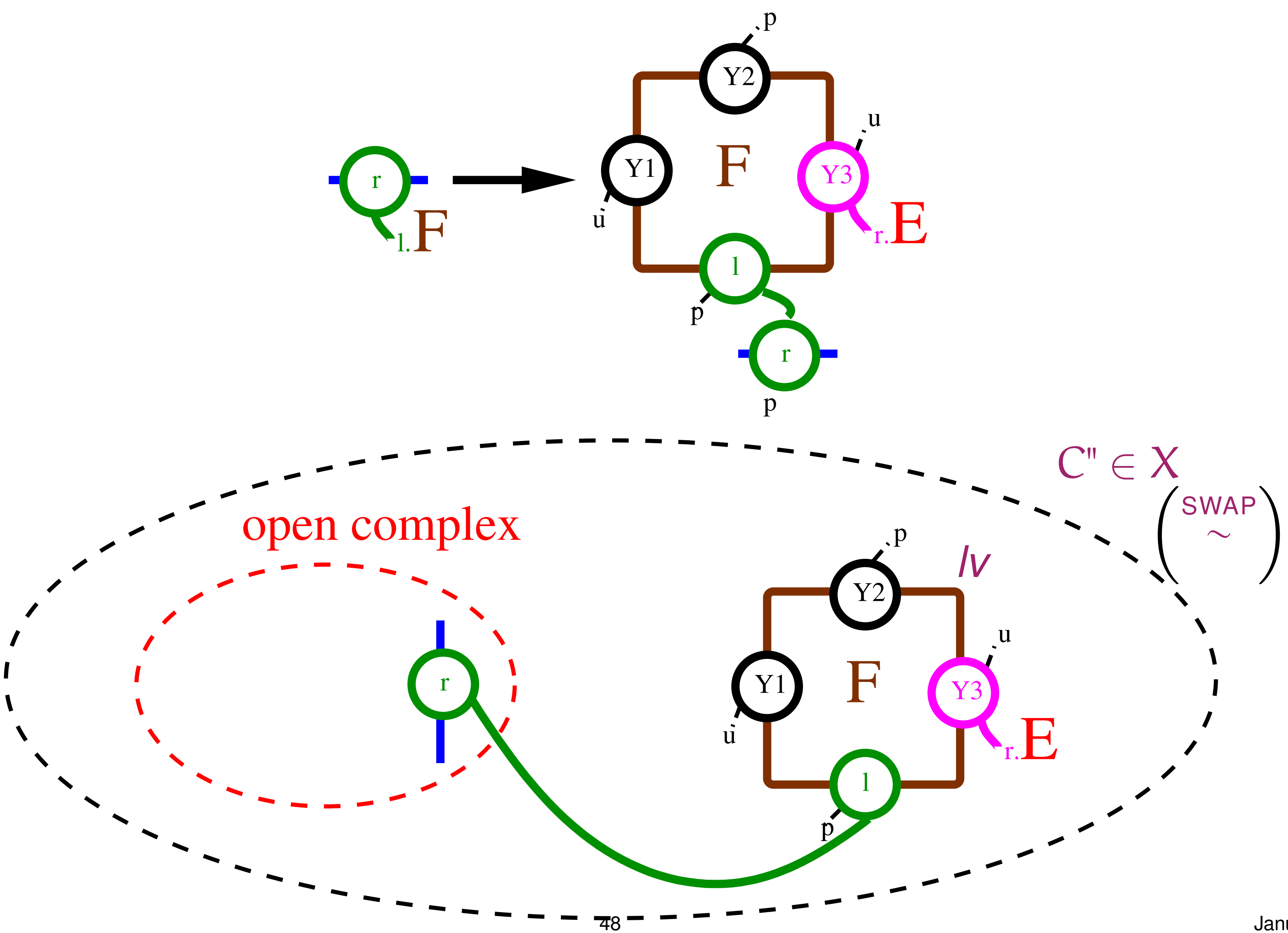




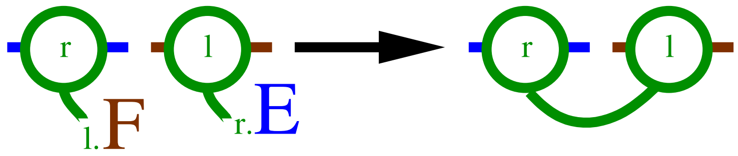
Unfolding a semi-link

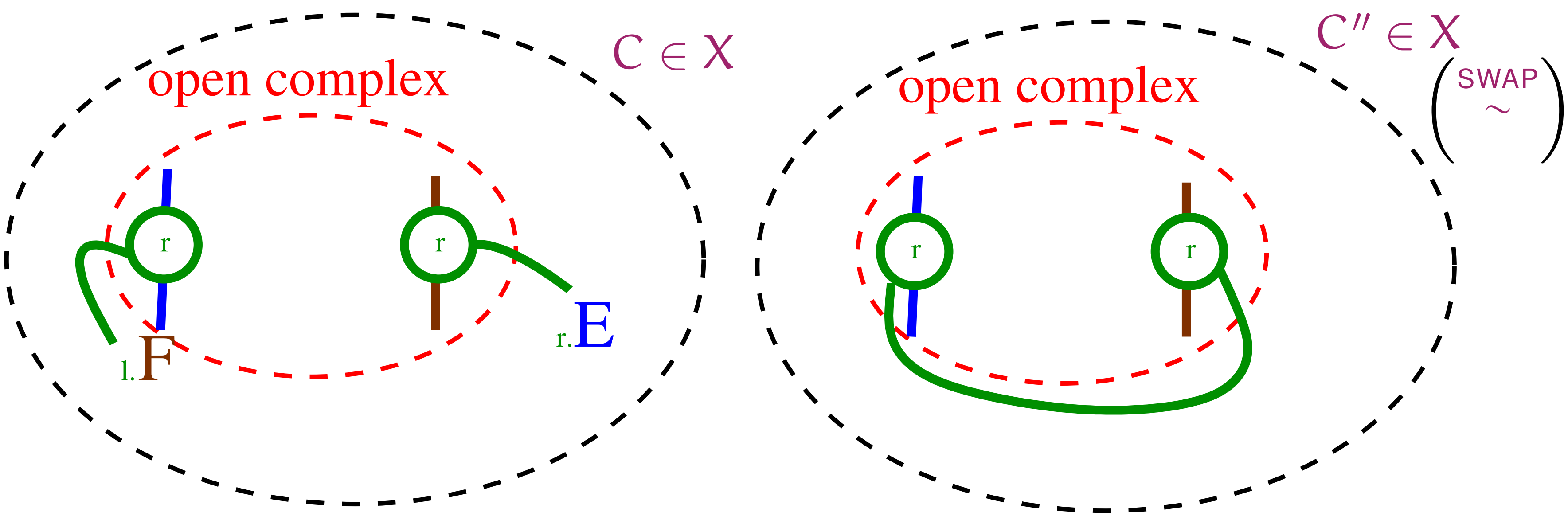


Unfolding a semi-link



Binding two semi-links





Consequences

Let $X \in \alpha(\wp(\mathit{Complex}))$ be a set of complexes.

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

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

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

4. We have $\mathit{lfp}_{\alpha(x_0)}\mathbb{F}^\sharp \in \alpha(\wp(\mathit{Complex}))$.

Overview

1. Introduction
2. Outline
3. Kappa language
4. Local views
5. Local set of complexes
6. Local rule systems
7. Decontextualization
8. Conclusion

Outline

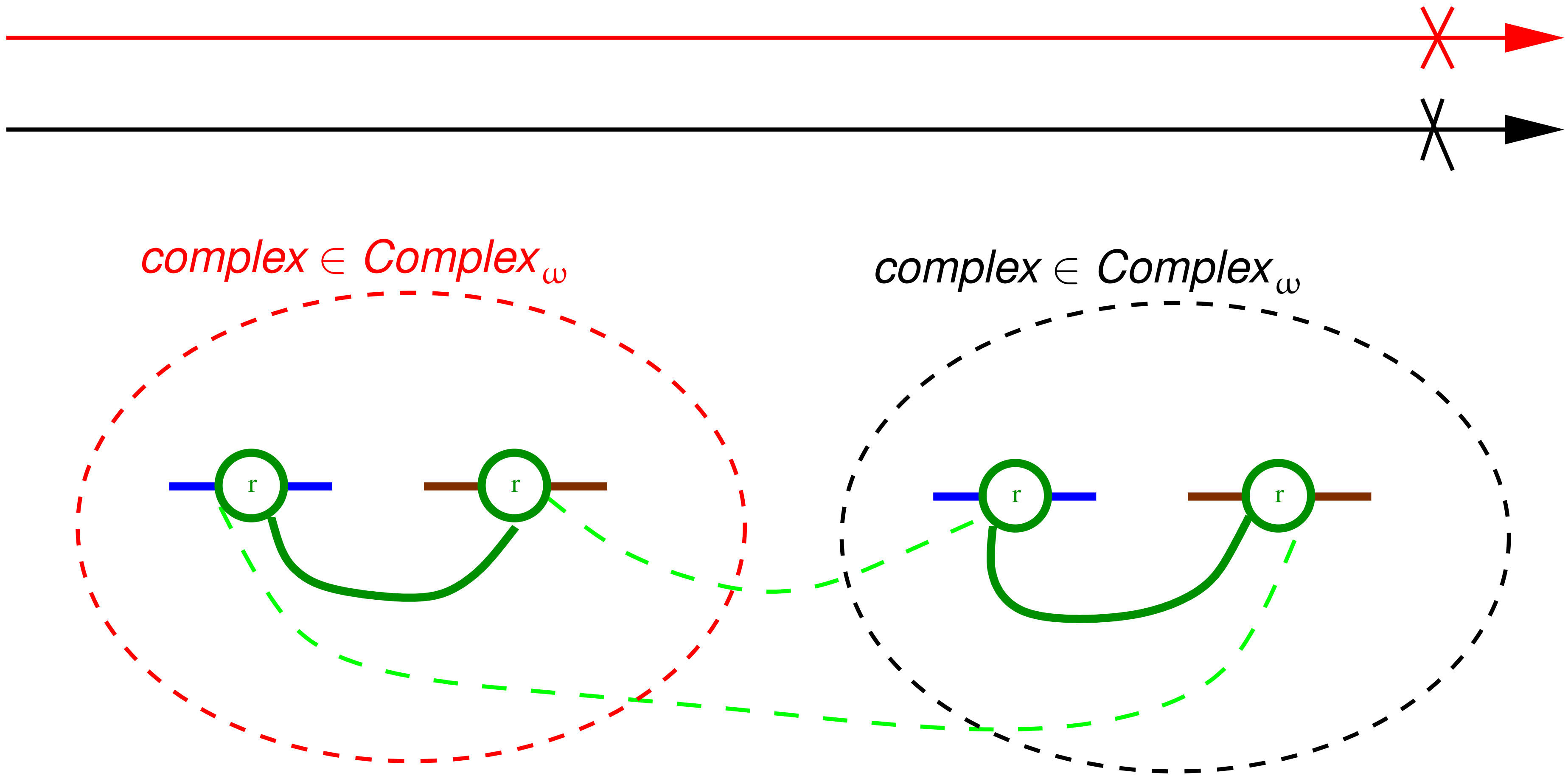
We have proved that:

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

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

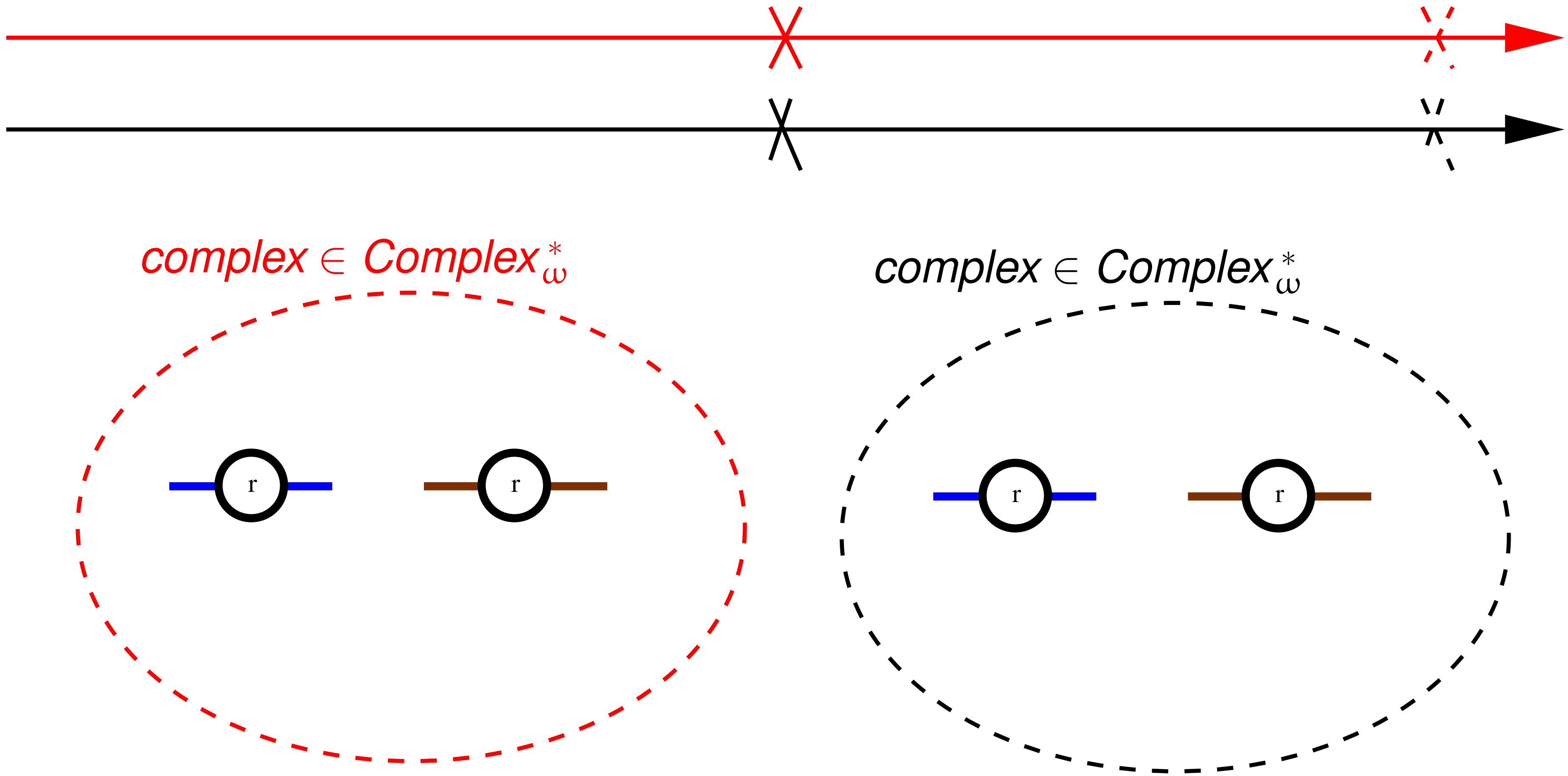
- We consider tuples of complexes in which the same kind of links occur twice.
- We want to swap these links.
- We introduce the history of their computation.
- There are several cases. . .

First case (I/V)



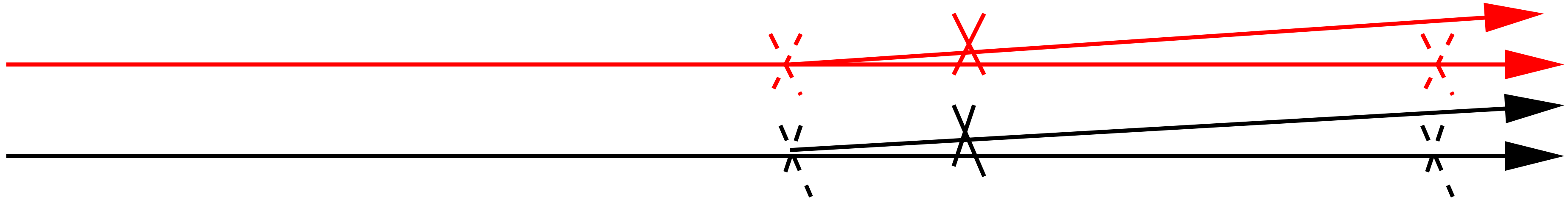
First case (II/V)

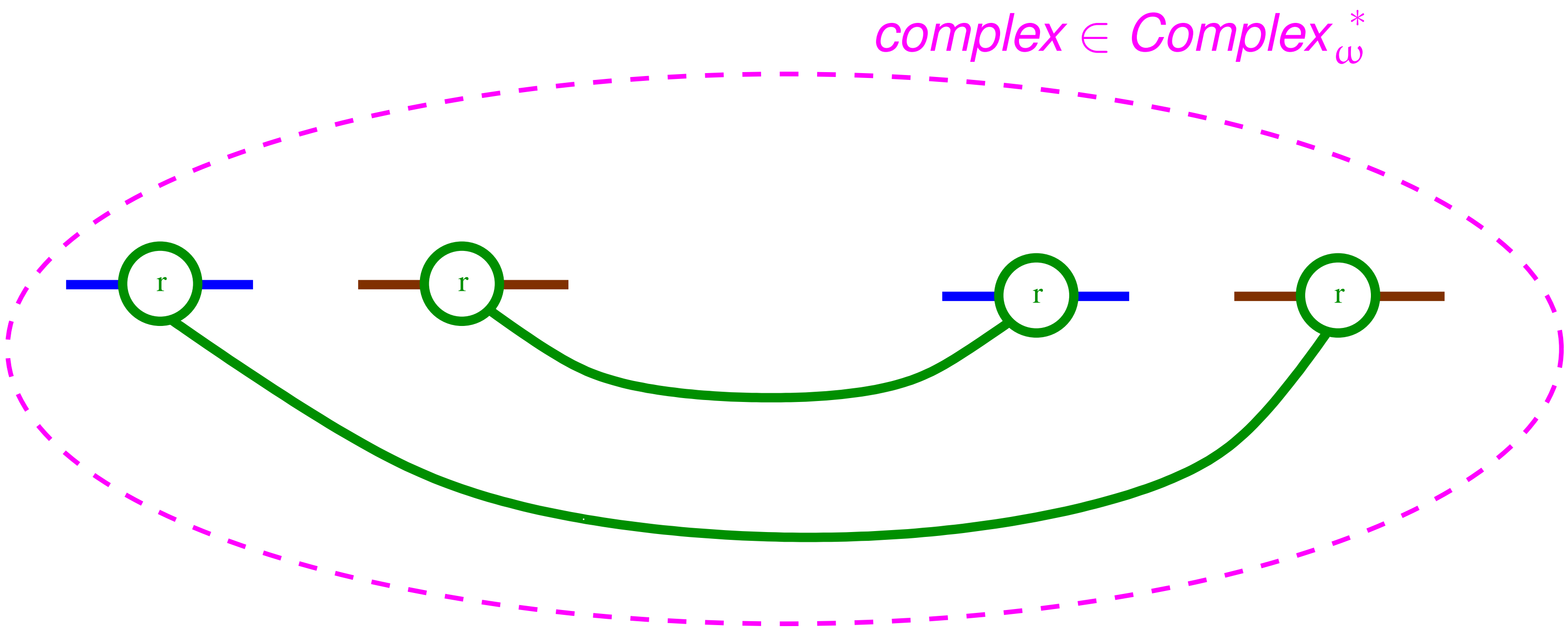
just before the links are made



First case (III/V)

we suppose we can swap the links





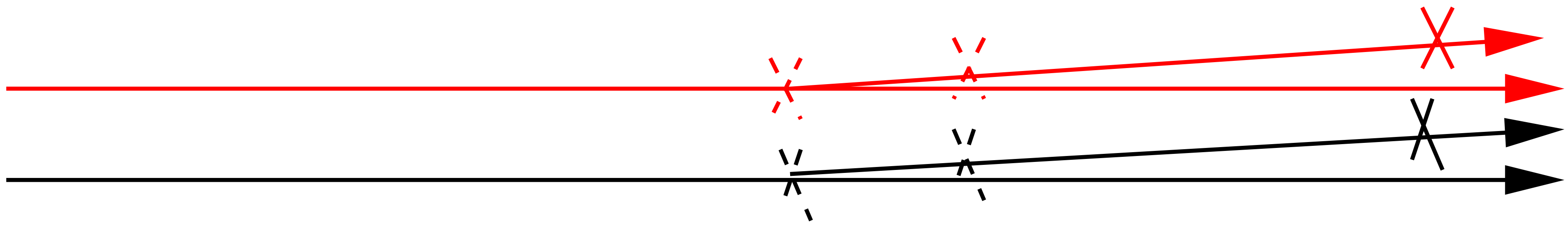
First case (IV/V)

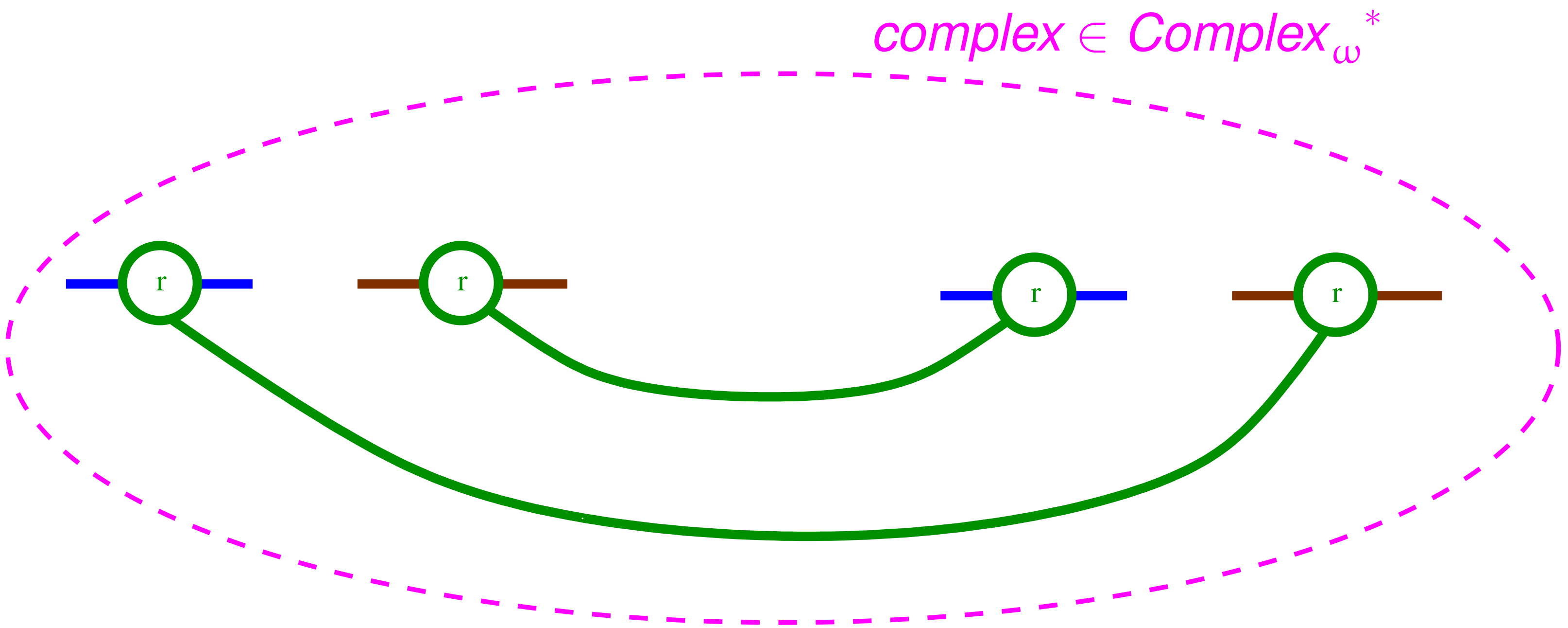
Then, we ensure that further computation steps:

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

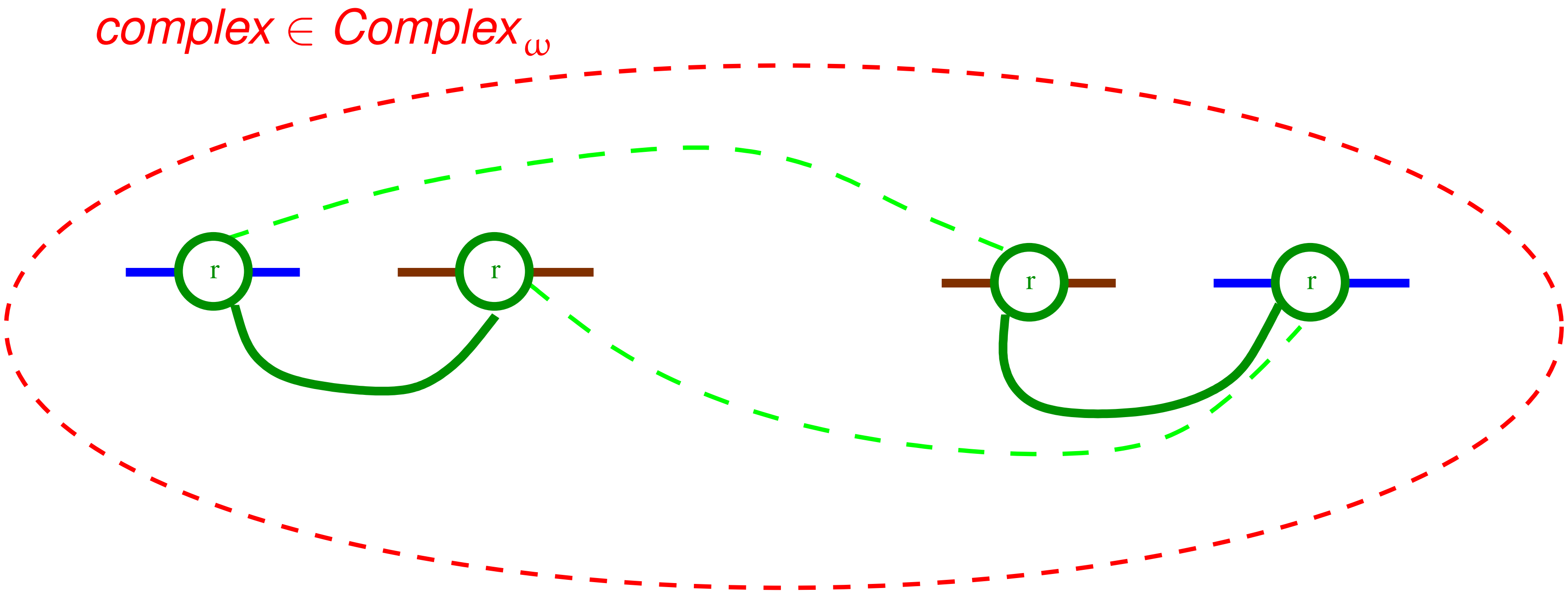
$$
\begin{tikzcd}
C_n \arrow[r, "\overset{\text{SWAP}}{\sim}{,}\sigma"] \arrow[d, "R{,}\phi"] & C'_n \arrow[d, "R{,}\phi"] \\
C_{n+1} \arrow[r, dotted, "\overset{\text{SWAP}}{\sim}{,}\sigma"] & C'_{n+1}
\end{tikzcd}
$$

First case (V/V)



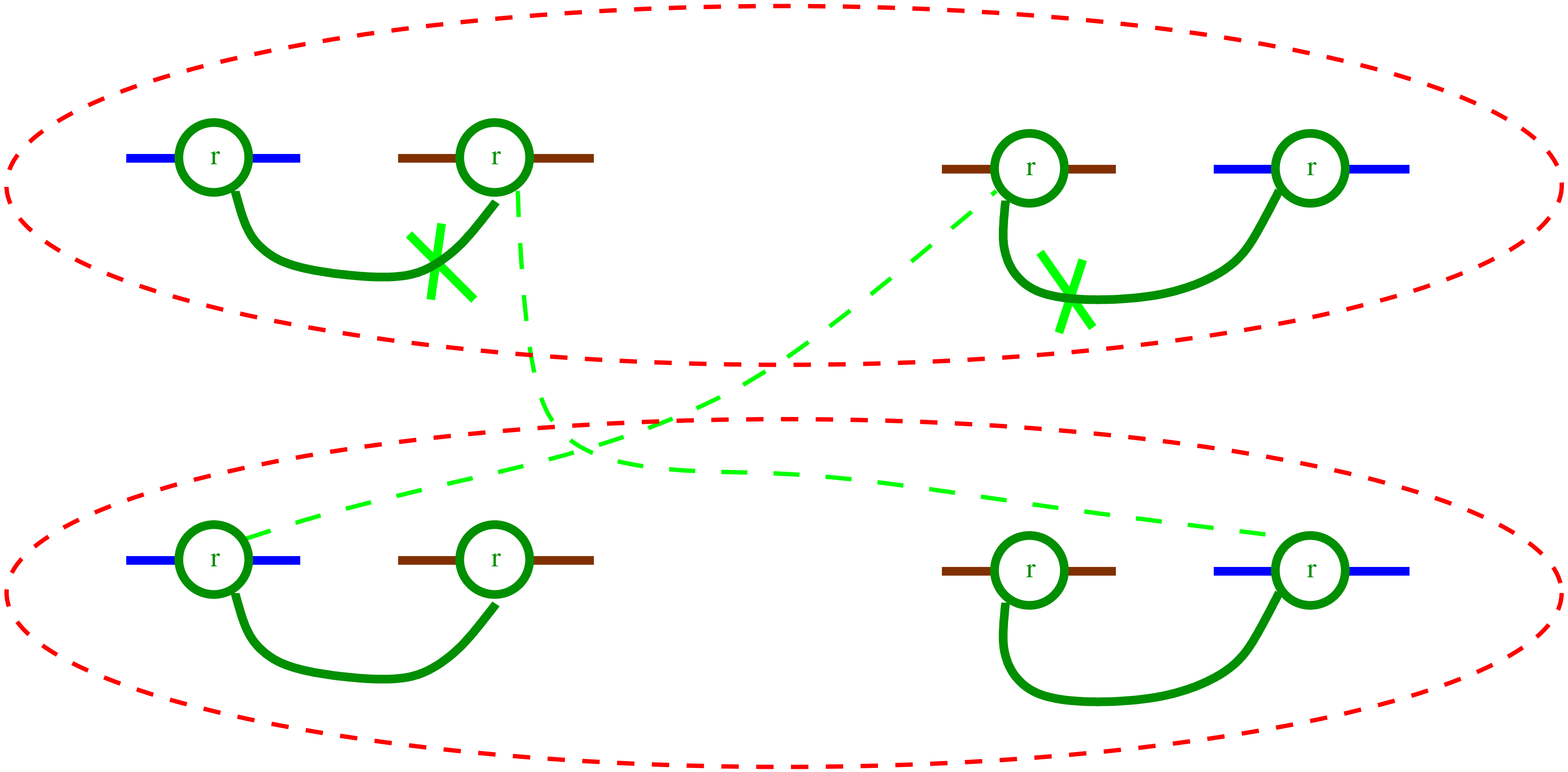


Second case (I/II)



we assume that the complex *complex* 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$\sim$m,S),B(b$\sim$n,T) $\rightarrow$ A(a$\sim$m!1,S),B(b$\sim$n!1,T)
   - and A(a$\sim$m',S'),B(b$\sim$n',T') $\rightarrow$ A(a$\sim$m'!1,S'),B(b$\sim$n'!1,T'),

   then:

   - A(a$\sim$m,S),B(b$\sim$n',T') $\rightarrow$ A(a$\sim$m!1,S),B(b$\sim$n'!1,T');
5. ...

are satisfied, the set of reachable complexes is local.

Non local systems

$$\mathit{Complex}_0 \triangleq \mathrm{R(a{\sim}u)}$$

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

R(a~u!1),R(a~u!1) $\in$ *Complex*$_\omega$
R(a~p!1),R(a~p!1) $\in$ *Complex*$_\omega$
But R(a~u!1),R(a~p!1) $\notin$ *Complex*$_\omega$.

Non local systems

$$Complex_0 \stackrel{\Delta}{=} \text{A(a$\sim$u),B(a$\sim$u)}$$

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

A(a~u!1),B(a~p!1) $\in$ $Complex_\omega$
A(a~p!1),B(a~u!1) $\in$ $Complex_\omega$
But A(a~p!1),B(a~p!1) $\notin$ $Complex_\omega$.

Non local systems

$$\mathit{Complex}_0 \triangleq \text{A(a}{\sim}\text{u)}$$

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

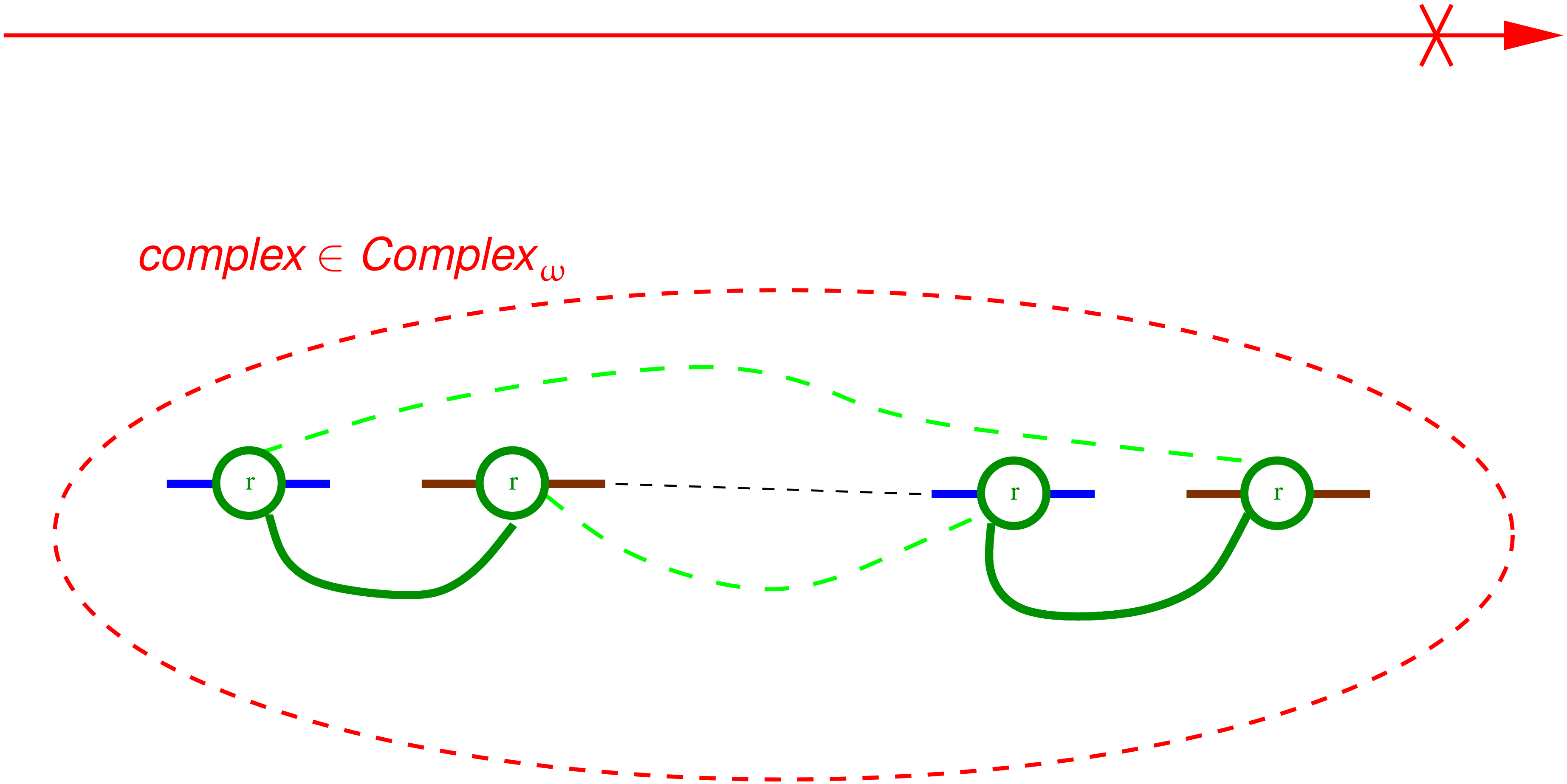
A(a~u!1),A(a~p!1) $\in$ *Complex*$_\omega$
But A(a~p!1),A(a~p!1) $\notin$ *Complex*$_\omega$.

Non local systems

$$
\begin{array}{lcl}
\textit{Complex}_0 & \triangleq & \mathrm{R(a,b)} \\
\textit{Rules} & \triangleq & \{\ \mathrm{R(a,b),R(a)} \rightarrow \mathrm{R(a,b!1),R(a!1)}\}
\end{array}
$$

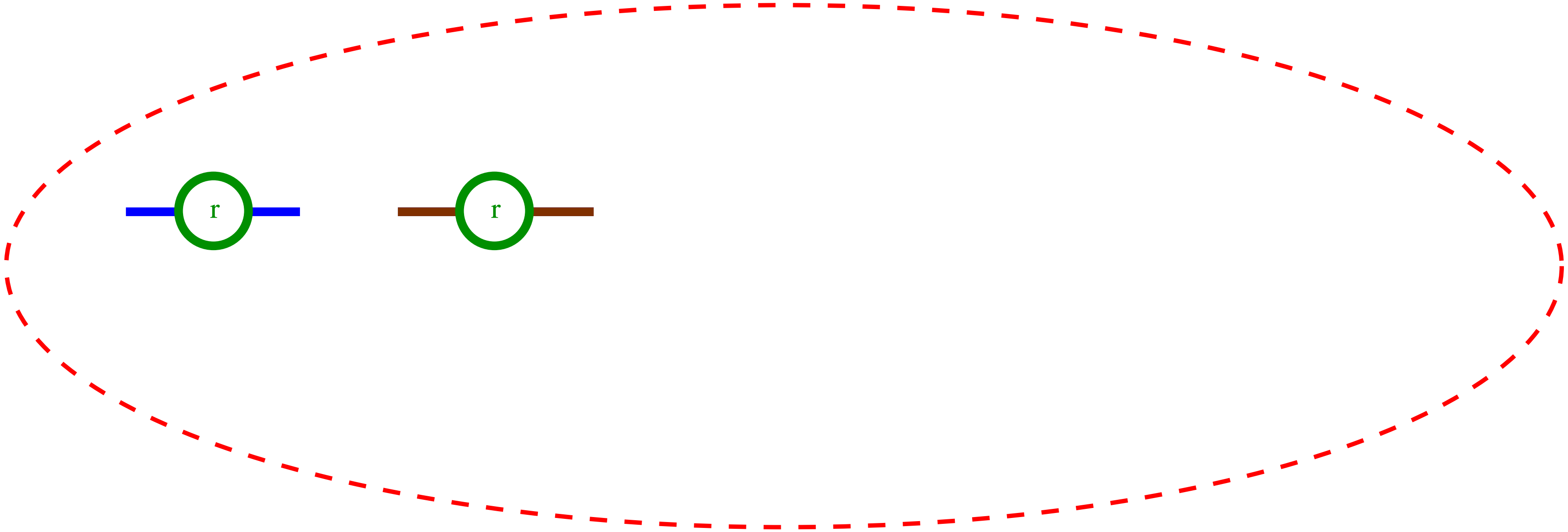
R(a,b!2),R(a!2,b!1),R(a!1,b)$\in$ *Complex*$_\omega$
But R(a!1,b!1) $\notin$ *Complex*$_\omega$.

Third case (I/III)

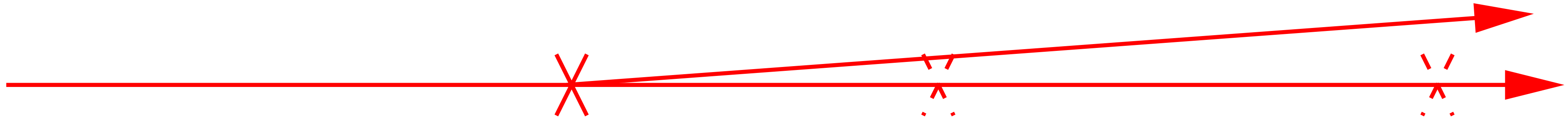


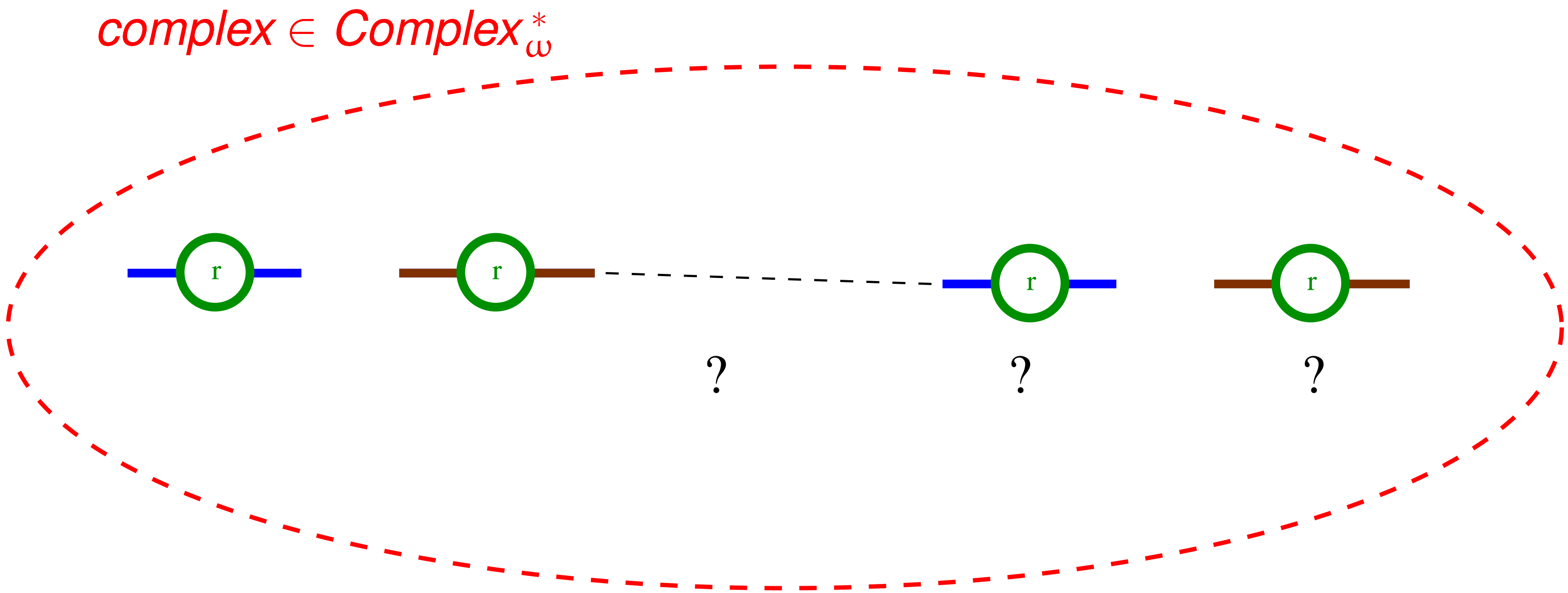
Third case (II/III)

complex $\in$ *Complex*$_{\omega}^{*}$



Third case (II/III)





Dangerous sites

A site is dangerous if it may occur in a cycle within a complex ($\in \gamma(\alpha(\mathit{Complex}_{\omega}))$).

To ensure that the set of reachable complexes is local, we also require that:

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

Then, we prove that:

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

Overview

Outline

- we have a syntactic criterion in order to ensure that the set of reachable complexes 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 complexes;
     - 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:

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

Decontextualized rule:

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

We can remove exhaustive enumerations.

How does it work ?

To remove a test, we prove that:

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

More formally

More formally:

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

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

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

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

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

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. Outline
3. Kappa language
4. Local views
5. Local set of complexes
6. Local rule systems
7. Decontextualization
8. Conclusion

Conclusion

- A scalable static analysis to abstract the reachable complexes.
- A class of models for which the abstraction is complete.
- Many applications:
  - idiomatic description of reachable complexes;
  - 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 complexes);
  - on the fly flat rule generation (for large/unbounded set)

Future works

- Reachability analysis.
  - Shape analysis: three valued logic, Mauborgne's graph abstract domains.
- Trace abstraction.
  - Can we lift the local-view abstraction to trace semantics ?
  - Which information do we obtain ?
- Quantitative/Semi-quantitative abstractions.
  - Can we design abstract domains to discover semi-quantitative properties (i.e. overshoot detection) ?