Deducing Interactions in Partially Unspecified Biological Systems

P. Baldan[1] A. Bracciali[2] L. Brodo[3] R. Bruni[2]

[1]Università di Padova

[2]Università di Pisa

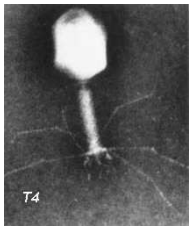
[3]Università di Sassari

Algebraic Biology - Hagenberg, July 2-4, 2007

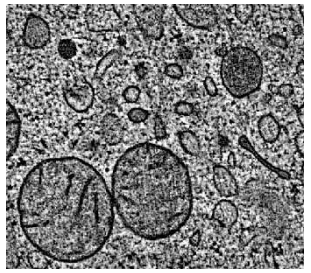
Outline

biological systems are *open* in different meanings

- **perturbable** - any living system expresses an external behaviour



- **incomplete** - our knowledge of biological systems is incomplete

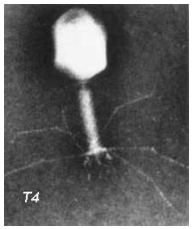


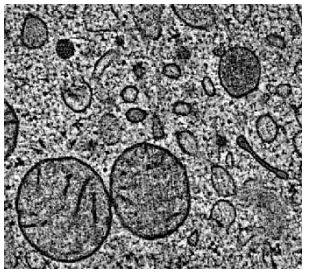
- **built by hand** - we may always assume to add some external components to a given complex

Partial information imposes that the approach is to predict the behaviours (of missing knowledge).

biological systems are *open* in different meanings

- **perturbable** - any living system expresses an external behaviour





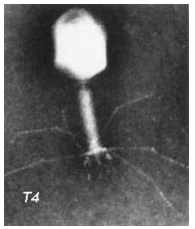
- **incomplete** - our knowledge of biological systems is incomplete

- **built by hand** - we may always assume to add some external components to a given complex

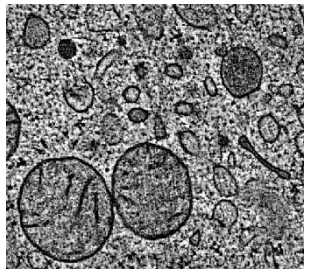
Partial information imposes that the approach is to predict the behaviours (of missing knowledge).

biological systems are *open* in different meanings

- **perturbable** - any living system expresses an external behaviour



- **incomplete** - our knowledge of biological systems is incomplete

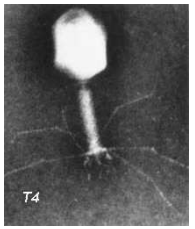


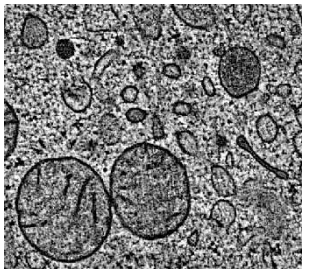
- **built by hand** - we may always assume to add some external components to a given complex

Partial information imposes that the approach is to predict the behaviours (of missing knowledge).

biological systems are *open* in different meanings

- **perturbable** - any living system expresses an external behaviour





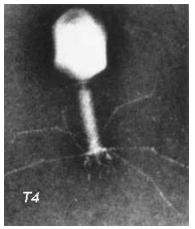
- **incomplete** - our knowledge of biological systems is incomplete

- **built by hand** - we may always assume to add some external components to a given complex

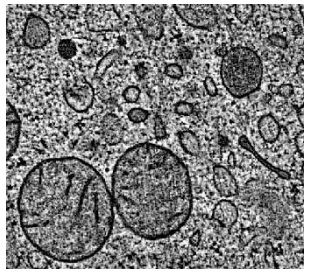
Partial information imposes that the approach is to predict the behaviours (of missing knowledge).

biological systems are *open* in different meanings

- **perturbable** - any living system expresses an external behaviour



- **incomplete** - our knowledge of biological systems is incomplete



- **built by hand** - we may always assume to add some external components to a given complex

Partial information imposes that the approach is to predict the behaviours (of missing knowledge).

partial biological knowledge, an analogy

molecule - as - computation

System of interacting molecular entities is described and modelled by a system of interacting computational entities.
Regev, Shapiro - *Nature, September 2002*.

- **relevant**: essential properties
- **computable**: computational knowledge
- **understandable**: conceptual framework
- **extensible**: capture other real properties

partial knwoledge - as - partial modelling

Modelling *incomplete systems* by applying a theory for *open* systems.

partial biological knowledge, an analogy

molecule - as - computation

System of interacting molecular entities is described and modelled by a system of interacting computational entities.
Regev, Shapiro - *Nature, September 2002*.

- **relevant**: essential properties
- **computable**: computational knowledge
- **understandable**: conceptual framework
- **extensible**: capture other real properties

partial knwoledge - as - partial modelling

Modelling *incomplete systems* by applying a theory for *open* systems.

partial biological knowledge, an analogy

molecule - as - computation

System of interacting molecular entities is described and modelled by a system of interacting computational entities.
Regev, Shapiro - *Nature, September 2002*.

- **relevant**: essential properties
- **computable**: computational knowledge
- **understandable**: conceptual framework
- **extensible**: capture other real properties

partial knwoledge - as - partial modelling

Modelling *incomplete systems* by applying a theory for *open* systems.

partial biological knowledge, an analogy

molecule - as - computation

System of interacting molecular entities is described and modelled by a system of interacting computational entities.
Regev, Shapiro - *Nature, September 2002*.

- **relevant**: essential properties
- **computable**: computational knowledge
- **understandable**: conceptual framework
- **extensible**: capture other real properties

partial knwoledge - as - partial modelling

Modelling *incomplete systems* by applying a theory for *open* systems.

Outline

Distributed and concurrent systems

Mobile and distributed programming (interaction, coordination, open)

- open systems can evolve (*no universal closure, higher level of dynamics*);
- constructive methodology (*unifi cation based - the most general unifi er is choosed to infer the transition*)
- complementary with contextualization techniques (*using contexts as labels to derive LTS, for which bisimilarity is a congruence, leads to use universal closure*);
- generality and friendly notation

The advantage is to import:

- theories
- automatic tools

Distributed and concurrent systems

Mobile and distributed programming (interaction, coordination, open)

- open systems can evolve (*no universal closure, higher level of dynamics*);
- constructive methodology (*unifi cation based - the most general unifi er is choosed to infer the transition*)
- complementary with contextualization techniques (*using contexts as labels to derive LTS, for which bisimilarity is a congruence, leads to use universal closure*);
- generality and friendly notation

The advantage is to import:

- theories
- automatic tools

Contextes

Main ingredients: contexts and process variables.

Processes with holes

$$P[X_1]|Q[X_2]$$

The evolution of the system depends on the particular components will be substituted with X_1 and X_2.

Computational steps, involving external components, require some hypothesis to fi x the particular behaviour needed to compute the desidered action.
A simulation trace is equipped with the required hypothesis.

Formula

Required behaviour

$$Q[X] \rightarrow Q'[Y]$$

Recording

- how many *holes* (process variables) there are;
- the hypothesis assumed at each computation step;
- how the computation is modifi ed by the new assumption taken

by the way of formula:

$$\varphi ::= X \mid \diamond a\, \varphi \mid f(\varphi_1, \ldots, \varphi_n)$$

Formulae as labels

General structrure of labels:

$$C[X_1, \ldots, X_n] \overset{(\varphi_1,\ldots,\varphi_n)}{\longrightarrow}_a D[Y_1, \ldots, Y_m]$$

The formula of a label transition, from one state to another, keeps trace of all the structure modifi cations required by the assumptions made hitherto.

formulae

Formula defi nition depends on the particular process algebra adopted:

- basic actions
- structural operators
- compositional operators

Outline

Simulating systems

Operational semantics

Inference rules

$$\frac{}{m[in\ n.Q \mid R] \mid n[P] \rightarrow_\tau n[m[Q \mid R] \mid P]} \ (in\ capability)$$

$$\frac{}{n[P] \mid open\ n.Q \rightarrow_\tau P \mid Q} \ (open\ capability)$$

$$\frac{P_1 \rightarrow_\alpha Q_1 \quad P_2 \rightarrow_{\alpha^\perp} Q_2}{P_1 \mid P_2 \rightarrow_\tau Q_1 \mid Q_2} \ (communication)$$

specifi c formula set

$$\varphi ::= X \mid \diamond in\,\varphi \mid \diamond open\,\varphi \mid (\varphi_1 \mid \varphi_2)$$

Simulating systems

Operational semantics

Inference rules

$$\frac{}{m[in\ n.Q \mid R] \mid n[P] \rightarrow_\tau n[m[Q \mid R] \mid P]} \ (in\ capability)$$

$$\frac{}{n[P] \mid open\ n.Q \rightarrow_\tau P \mid Q} \ (open\ capability)$$

$$\frac{P_1 \rightarrow_\alpha Q_1 \quad P_2 \rightarrow_{\alpha^\perp} Q_2}{P_1 \mid P_2 \rightarrow_\tau Q_1 \mid Q_2} \ (communication)$$

specifi c formula set

$$\varphi ::= X \mid \diamond in\,\varphi \mid \diamond open\,\varphi \mid (\varphi_1 \mid \varphi_2)$$

Symbolic Transition

$$v[\ X\ ] \mid c[open\ v.(prot \mid rna^{\perp})]$$

initial state

Assuming we want the *virus* $v[\ldots]$ to enter the *cell* $c[\ldots]$, the required hypothesis is that the virus were the capability to do that.

$$\xrightarrow{in\ c.Y|Z}_{\tau}$$

label

The formula abstracts over the structure of the virus in the most general way!

$$c[v[Y \mid Z] \mid open\ v.(prot \mid rna^{\perp})]$$

residual

In the next state (i.e. the continuation) there is a modifi cation of the confi guration of the process variables, due to the assumpions made.

Symbolic Transition

$$v[\ X\ ] \mid c[open\ v.(prot \mid rna^{\perp})]$$

initial state

Assuming we want the *virus* $v[\ldots]$ to enter the *cell* $c[\ldots]$, the required hypothesis is that the virus were the capability to do that.

$$\xrightarrow{in\ c.Y|Z}_{\tau}$$

label

The formula abstracts over the structure of the virus in the most general way!

$$c[v[Y \mid Z] \mid open\ v.(prot \mid rna^{\perp})]$$

residual

In the next state (i.e. the continuation) there is a modifi cation of the confi guration of the process variables, due to the assumpions made.

Symbolic Transition

$$v[\ X\ ] \mid c[open\ v.(prot \mid rna^{\perp})]$$

initial state

Assuming we want the *virus* $v[\ldots]$ to enter the *cell* $c[\ldots]$, the required hypothesis is that the virus were the capability to do that.

$$\xrightarrow{in\ c.Y|Z}_{\tau}$$

label

The formula abstracts over the structure of the virus in the most general way!

$$c[v[Y \mid Z] \mid open\ v.(prot \mid rna^{\perp})]$$

residual

In the next state (i.e. the continuation) there is a modifi cation of the confi guration of the process variables, due to the assumpions made.

Symbolic Transition

$$v[\ X\ ] \mid c[open\ v.(prot \mid rna^{\perp})]$$

initial state

Assuming we want the *virus* $v[\ldots]$ to enter the *cell* $c[\ldots]$, the required hypothesis is that the virus were the capability to do that.

$$\xrightarrow{in\ c.Y|Z}_{\tau}$$

label

The formula abstracts over the structure of the virus in the most general way!

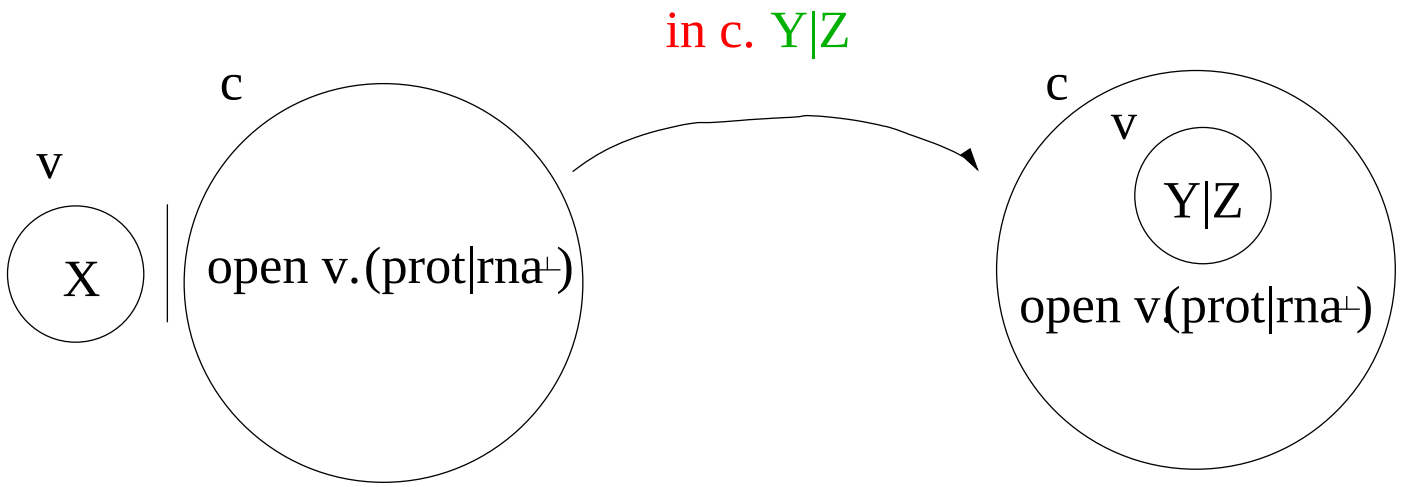
$$c[v[Y \mid Z] \mid open\ v.(prot \mid rna^{\perp})]$$

residual

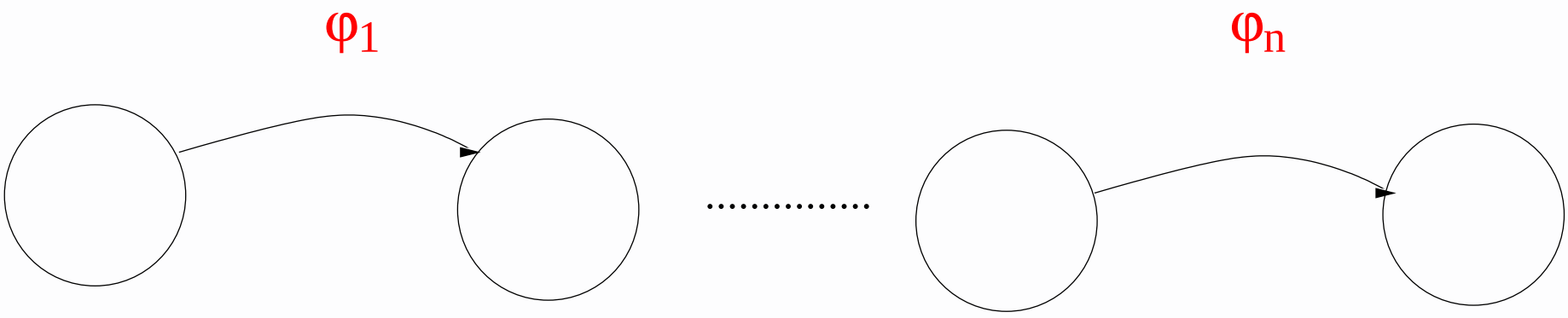
In the next state (i.e. the continuation) there is a modifi cation of the confi guration of the process variables, due to the assumpions made.

Symbolic Transition

At each step of the computation, new hypotheses are required

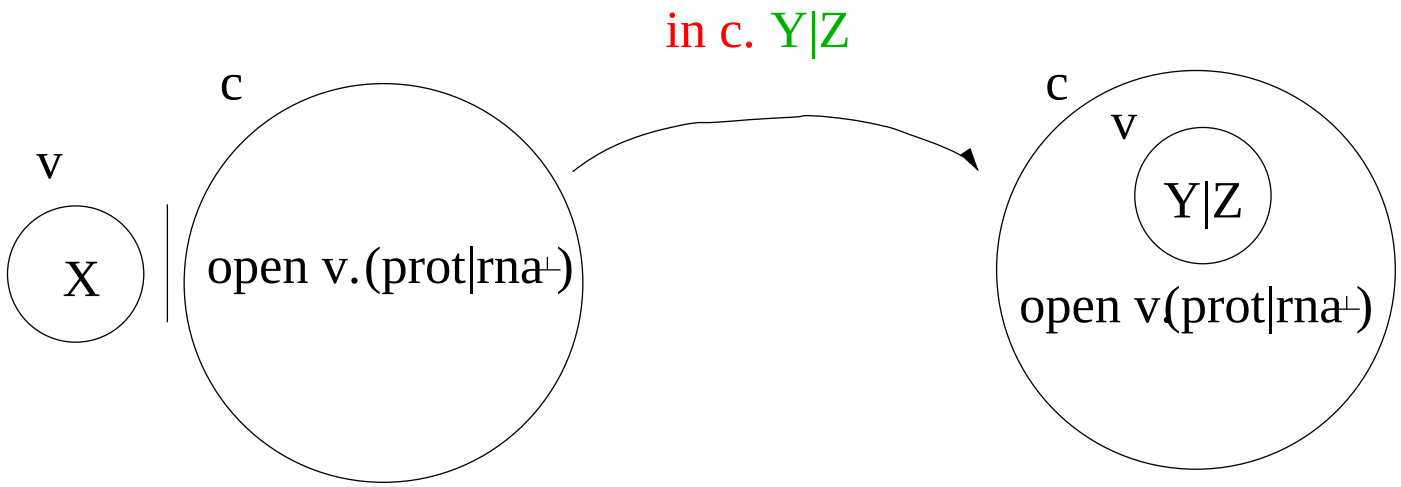


how to build a trace ?

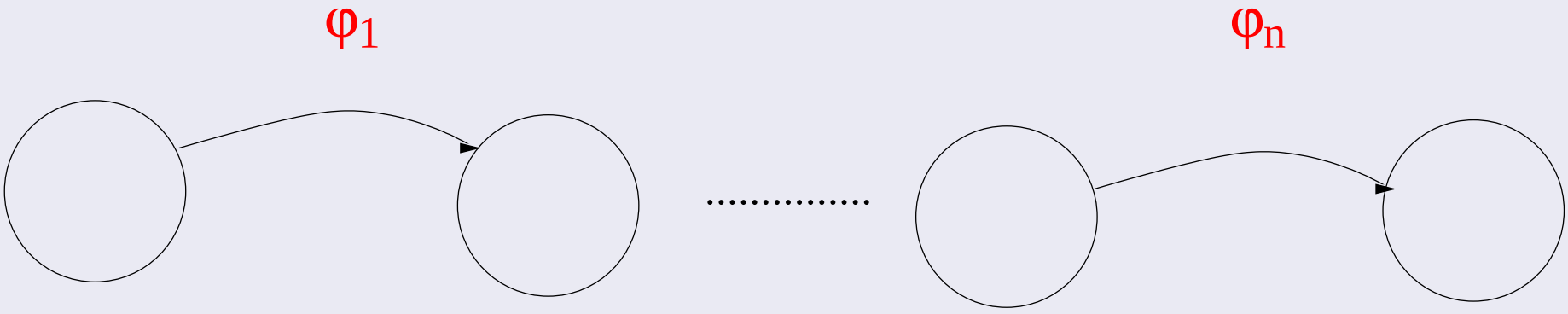


Symbolic Transition

At each step of the computation, new hypotheses are required



how to build a trace ?



Symbolic Trace

$$E[X] \;=\; v[\,X\,] \mid c[open\ v.(prot \mid rna^{\perp})]$$

fi rst transition

$$E[X] \xrightarrow{in\ c.Y|Z}_{\tau} c[v[Y \mid Z] \mid open\ v.(prot \mid rna^{\perp})]$$

second transition

$$c[v[Y \mid Z] \mid open\ v.(prot \mid rna^{\perp})] \xrightarrow{Y,Z}_{\tau} c[Y \mid Z \mid prot \mid rna^{\perp}]$$

third transition

$$c[Y \mid Z \mid prot \mid rna^{\perp}] \xrightarrow{\diamond rnaW,Z}_{c} [W \mid Z \mid prot]$$

label composition

$$in\ c.(\diamond\ rna\ W)|Y$$

Valid theory

- Correctness and completness of Symbolic Transition Systems
  - **correctness** every concrete behaviour (of a full specifi ed system) has an abstract representation (the corresponding in STS)
  - **completness** every instance of an abstract behaviour (STS) correspond to a concrete behaviour
- symbolic bisimulation for open systems (different from universal closure)
- Simbolic Transition System is built by unifi cation (Prolog)

bibliography

- Paolo Baldan, Andrea Bracciali, Roberto Bruni. **Bisimulation by unification**. AMAST 2002, LNCS vol.2422: 254-270.
- Paolo Baldan, Andrea Bracciali, Roberto Bruni. **Symbolic Equivalence for Open Systems**. Global Computing 2004, LNCS vol.3267: 1 -17.

Valid theory

- Correctness and completness of Symbolic Transition Systems
  - **correctness** every concrete behaviour (of a full specifi ed system) has an abstract representation (the corresponding in STS)
  - **completness** every instance of an abstract behaviour (STS) correspond to a concrete behaviour
- symbolic bisimulation for open systems (different from universal closure)
- Simbolic Transition System is built by unifi cation (Prolog)

bibliography

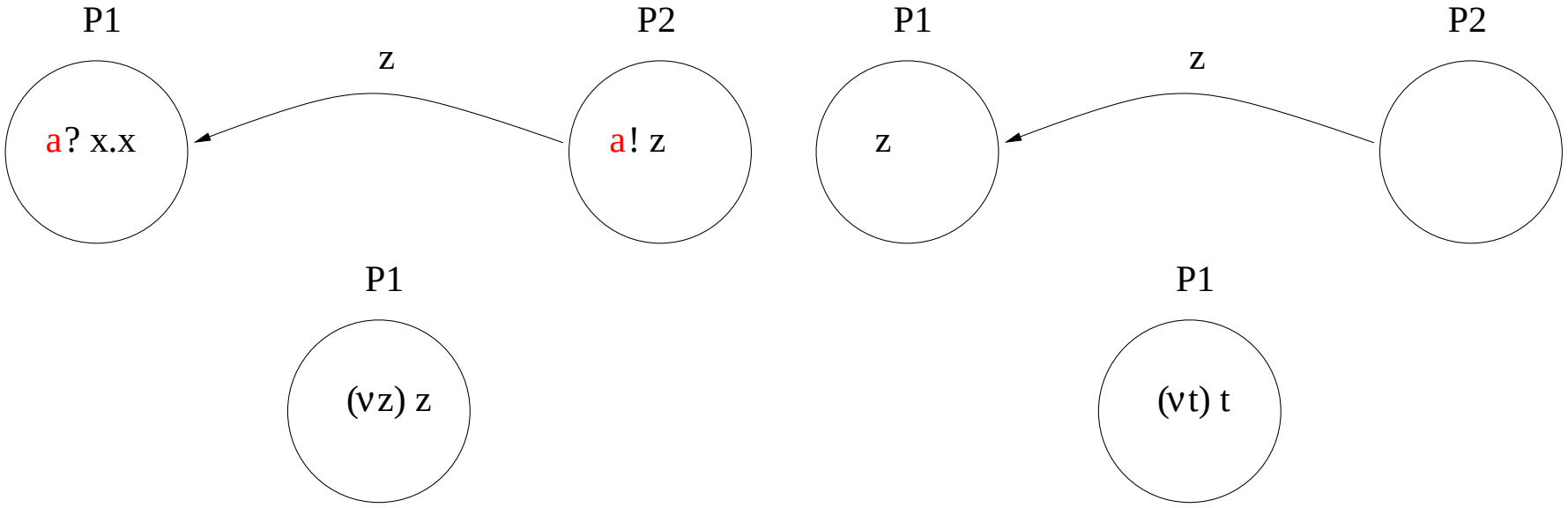
- Paolo Baldan, Andrea Bracciali, Roberto Bruni. **Bisimulation by unification**. AMAST 2002, LNCS vol.2422: 254-270.
- Paolo Baldan, Andrea Bracciali, Roberto Bruni. **Symbolic Equivalence for Open Systems**. Global Computing 2004, LNCS vol.3267: 1 -17.

On going work

STS works with a subset of the modeling languages: not able to handle names, **yet**!

π-calculus

$a\,!\,b$	output
$a\,?\,x$	input
$(\nu\, n)P$	private names



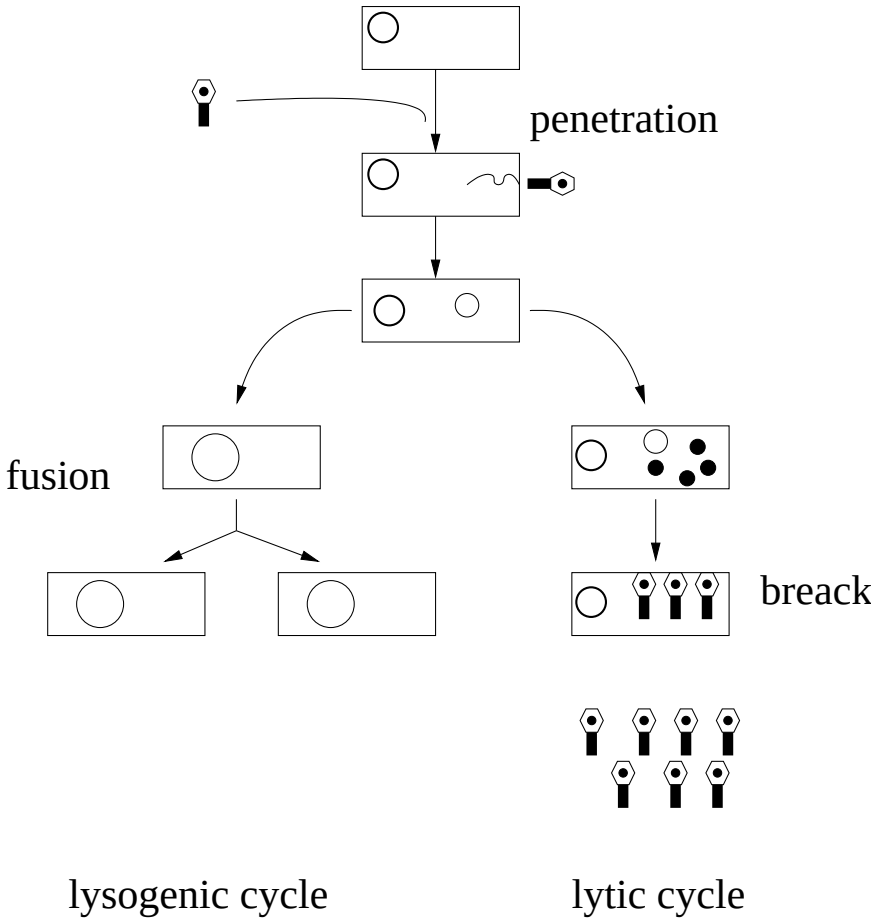
Outline

BioAmbients

Introduced for the modelling of biological membranes

$[\ldots]$	empty membrane
$[a?x \mid a!z]$	with internal behaviour
$[\cdots \mid a!z] \mid [a?x \mid \ldots]$	different membranes with internal behaviour
$[\cdots \mid \mathit{sibling}\ a!z] \mid [\mathit{sibling}\ a?x \mid \ldots]$	trapassing mebranes
$[\cdots \mid \mathit{merge}^{+}\ a] \mid [\mathit{merge}^{-}\ a \mid \ldots]$	merging membranes
$[\cdots \mid \mathit{enter}\ a] \mid [\mathit{accept}\ a \mid \ldots]$	nesting membranes
$[[\cdots \mid \mathit{exit}\ a] \mid \mathit{allow}\ a \mid \ldots]$	nesting membranes

Toy example



each protein is an ambient
activation as communication
inhibition as encapsulation

very simplified hyphoteses

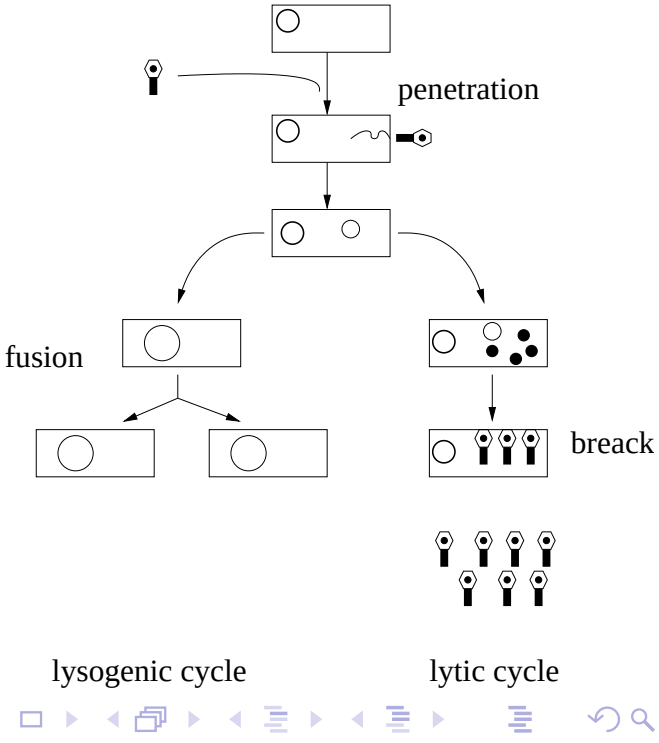
1 - high concentration of **CI** determines lysogeny

2 - absence of **CI** determines lysis
3 - **CII** promotes the production of **CI**
4 - **CIII** can inhibit **HFL**
5 - low concentration of **CRO** stimulates **CI** production

6 - high concentration of **CRO** inhibits **CI** production

BioAmbient code

$$
\begin{array}{lcl}
[\mathrm{VIRUS}] & = & [\mathrm{merge}^{+}\ \mathrm{virus}.([\mathrm{C3}] \mid [\mathrm{C2}] \mid [\mathrm{C1}] \mid [\mathrm{CRO}]) \mid [\mathrm{DNA}\lambda\ ]] \\
\mathrm{DNA}\lambda & = & (\mathrm{lyso?}.\mathrm{enter}\ \mathrm{dnae}.0) \\
 & & + \\
 & & \mathrm{lysi?}.(\lambda[\mathrm{exit}\ \mathrm{newph}.\mathrm{VIRUS}] \mid \mathrm{expel}\ \mathrm{newph}) \\
[\mathrm{ECOLI}] & = & [\mathrm{merge}^{-}\mathrm{virus} \mid {}_{Dna_e}[\mathrm{accept}\ \mathrm{dnae}] \mid [\mathrm{HFL}]] \\
\mathrm{HFL} & = & \mathrm{enter}\ \mathrm{h\_c3}.0 + X
\end{array}
$$

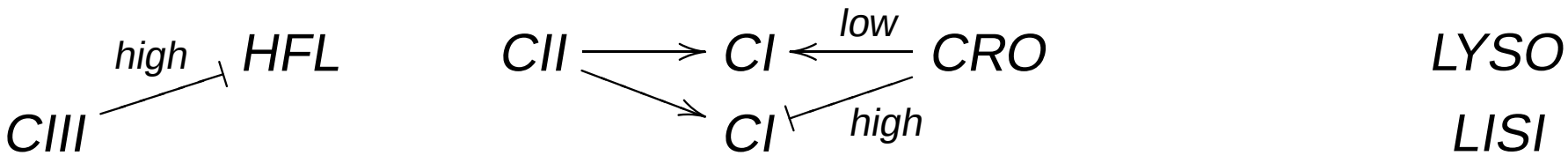


Bioambients - logical formulae

$$\varphi ::= X \mid \diamond a\,\varphi \mid \varphi_1 + \varphi_2 \mid \varphi_1 \mid \varphi_n \mid a.\varphi$$

$$a ::= n? \mid n! \mid \mathit{enter}\ n \mid \mathit{accept}\ n \mid \ldots$$

Assuming partial knowledge



C3 = l_c3!.0 + accept h_c3. pro_c2!.0
C2 = pro_c2?. pro_c1!.0 + enter c2.0

C1 = pro_c1?.(h_cro?. lysi!.0 + l_cro?. lyso!.0)
CRO = l_cro!.0 + h_cro.0

λ phago

$$Ecoli[([C3] \mid [C2] \mid [C1] \mid [CRO]) | [DNA\lambda] \mid_{Dna_e} [accept\ dnae] \mid [enter\ h\_c3.0 + X]]$$

$$\xrightarrow{(I\_c3?.Y_1 + Y_2) | Y_3}$$

$$Ecoli[_{CIII}[0] \mid [C2] \mid [C1] \mid [CRO]) | [DNA\lambda] \mid {}_{Dna_e}[accept\ dnae] \mid {}_{hfl}[Y_1 \mid Y_3]]$$

$$\xrightarrow{(Y_4 + accept\ c2.Y_5 | Y_6), Y_3}$$

$$Ecoli[(_{CIII}[0] \mid [C1] \mid [CRO]) | [DNA\lambda] \mid_{Dna_e} [accept\ dnae] \mid_{hfl} [_{CII}[0] | (Y_5 \mid Y_6) \mid Y_3]]$$

$$\xrightarrow{(Y_7 + lysi!.Y_8), Y_6, Y_3}$$

$$Ecoli[(_{CIII}[0] \mid [C1] \mid [CRO]) | [_\lambda[exit\ newph.VIRUS] \mid expel\ newph] \mid$$
$$_{Dna_e}[accept\ dnae] \mid_{hfl} [_{CII}[0] | (Y_8 \mid Y_6) \mid Y_3]]$$

$$\xrightarrow{Y_8, Y_6, Y_3}$$

$$Ecoli[(_{CIII}[0] \mid [C1] \mid [CRO]) |_\lambda [VIRUS] \mid_{Dna_e} [accept\ dnae] \mid_{hfl} [_{CII}[0] | (Y_8 \mid Y_6) \mid Y_3]]$$

The deduced symbolic trace

$$(l\_c3?.(Y_4 + accept\ c2(Y_7 + lysi!.Y_8) \mid Y_6) + Y_2) \mid Y_3$$

As theory is correct, the trace computed can be obtained in a fully specifi ed system, assuming the same biological hyphoteses.

Possible applications

How to use our approach:

Before in silico experiments

Preliminary study of the systems.

Partial data

When is not possible to design a complete model.

Extention of existing models

To build new complexes by studying the requirements the new components should have to interact in the desidered way

Future work

- **names**: extend approach to make the STS also works with process algebras that handle name restriction and substitution;
- **automatic tool** build automatic simulator working with any process algebra of interest (based on unifi cation);
- **quantitative analysis** adding values for calculating the probability of a transition;

The end

Thanks for your attention !

Future work II

- building a *signifi cative* trace requires to follow a criterium to discriminate between infi nite moves.
- ... quantitative values may help —