Aspect of multiscale modelling in a process algebra for biological systems

Roberto Barbuti[1], Giulio Caravagna[1], Paolo Milazzo[1],
Andrea Maggiolo-Schettini[1] and Simone Tini[2]

1. Dipartimento di Informatica, Università di Pisa, Italy
2. Dipartimento di Informatica e Comunicazione, Università dell'Insubria, Como, Italy

MeCBIC 2010 – Jena, Germany – August 23rd, 2010

Introduction (1)

Formal modelling notations of computer science are nowadays often applied to the description of biological systems.

- Process calculi
- Term rewriting systems
- Petri nets
- Automata
- ...

The dynamics of the systems is described as a sequence of events (usually biochemical reactions) described either as communications between processes or as applications of rewrite rules.

Stochastic extensions allow rates of events to be taken into account.

Introduction (2)

The description of the dynamics of biological systems as a sequence of events assumes that the occurrence of one such events can be described as an instantaneous change in the system state.

Even in the stochastic extension, the only time information associated with events is frequency (rather than duration)

This way of describing the dynamics of biological systems is suitable if the duration of events can be hidden in the time interval between two subsequent events.

The multiscale approach to modelling

Phenomena of interest in the study of biological systems often include processes at different levels of abstraction.

A typical example: cell signalling phenomena

- **Intra-cellular level:** Gene regulation and protein interaction processes;
- **Inter-cellular level:** Signal diffusion among cells
- **Tissue level:** Macroscopic changes (e.g. Tumour growth, formation of new blood vessels, etc.)

The processes at the different levels influence each other.

The multiscale approach to modelling consists in:

- constructing different models for the different abstraction levels;
- connecting the different models by defining suitable simulation workflows.

Aims (1)

We face the problem of applying formal modelling notations to the description of multiscale biological systems.

- We focus on process algebras.

The description of the dynamics of biological systems with a sequence of instantaneous events is not suitable for multiscale models:

- events at a scale level may take much more time than events at a lower level;
- while a component is involved in a long-lasting high-level event, it may be involved also in several faster lower level events;
- the completion of some events may (or may not) interfere with other events in which the same component is concurrently involved.

Key features for a new process algebra: non-instantaneous actions and possibility for a process to perform several concurrent actions.

Aims (2)

Events at different levels may (or may not) interfere with each other:

- Example: cell death may interrupt all intra-cellular processes

Key feature for a new process algebra: the completion of some actions may interrupt other running actions of the same processes

In this work we propose a simple process algebra that tries to incorporate all of these features.

Our approach is foundational and minimalist:

- We aim at understanding the semantic implications of the new features rather than at proposing a rich modelling language

Let's start from CCS

Let Act be an infinite set of actions $a, b, c, \ldots$ and co-actions $\bar{a}, \bar{b}, \bar{c}, \ldots$.

We start from the following fragment of CCS:

$$P \; := \; 0 \;\;|\;\; \alpha.\,P \;\;|\;\; P + P \;\;|\;\; P \,|\, P \;\;|\;\; A$$

where $\alpha \in Act$ and A is a process constant.

This fragment is rich enogh to describe systems of chemical reactions: for example $E + S \rightleftharpoons ES \rightarrow E + P$ becomes:

$$P_E = a.P_{ES} \qquad P_S = \bar{a}.0 \qquad P_{ES} = b.(P_E \mid P_S) + c.(P_E \mid P_P)$$
$$Aux = \bar{b}.Aux + \bar{c}.Aux$$

with a process such as $P_E \mid P_E \mid P_E \mid P_S \mid P_S \mid P_S \mid P_S \mid P_S \mid P_S \mid Aux$ modelling a possible initial solution

Modelling event durations

The semantics of a syncronization in CCS usually corresponds to a single transition between processes:

$$a.P_{ES} \mid \bar{a}.0 \xrightarrow{\tau} P_{ES}$$

In order to model non-instantaneous events we consider separate transitions for action start and action completion (as in ST semantics)

$$a.P_{ES} \mid \bar{a}.0 \xrightarrow{\tau^+} \dots \qquad \dots \xrightarrow{\tau^-} P_{ES}$$

In order to keep the continuations of the involved processes frozen for the duration of the action we extend the syntax with a new operation $[\_]^l$ (where $l \in \mathbb{N}$ is a fresh index) as follows:

$$a.P_{ES} \mid \bar{a}.0 \xrightarrow{1,\tau^+} [a]^1 .P_{ES} \mid [\bar{a}]^1 .0 \rightarrow \dots \quad \dots \rightarrow [a]^1 .P_{ES} \mid [\bar{a}]^1 .0 \xrightarrow{1,\tau^-} P_{ES}$$

Process configurations

We obtain a new syntax consisting of processes P and process configurations C_P defined by the following grammar:

$$
\begin{array}{rcl}
P & ::= & 0 \;\big|\; \alpha.P \;\big|\; P+P \;\big|\; P \mid P \;\big|\; A \\
C_P & ::= & P \;\big|\; [\alpha]^l.P \;\big|\; C_P + C_P \;\big|\; C_P \mid C_P
\end{array}
$$

where $\alpha \in Act$ and $l \in \mathbb{N}$.

A minimal multiscale model

Let C be a cell, and P be some growth factor produced by the cell.

Minimal multiscale model:

- Low-level event: production of the growth factor

$$C \rightarrow C + P$$

- High-level events: cell duplication (mitosis) and cell death

$$C \rightarrow C + C \qquad\qquad C \rightarrow$$

This example would be modelled by the following CCS process:

$$P_C = a.(P_C \mid P_P) + b.(P_C \mid P_C) + c.0 \qquad Aux = \bar{a}.Aux + \bar{b}.Aux + \bar{c}.Aux$$

where P_P is a process modelling the growth factor P.

Note: With event durations there must be as many Auxs as P_Cs, hence Aux should be defined as $\bar{a}.Aux + \bar{b}.(Aux \mid Aux) + \bar{c}$

Desired behaviour

The minimal multiscale model

$$C \rightarrow C + P \qquad C \rightarrow C + C \qquad C \rightarrow$$

$$P_C = a.(P_C \mid P_P) + b.(P_C \mid P_C) + c.0$$

We would like to be able to describe the following possible behaviour:

1. A cell repeatedly duplicates
2. in the meanwhile it produces growth factors
3. and unexpectedly it dies (and interrupts its active processes)

The problem with choice

The minimal multiscale model

$$C \rightarrow C + P \qquad C \rightarrow C + C \qquad C \rightarrow$$

$$P_C = a.(P_C \mid P_P) + b.(P_C \mid P_C) + c.0$$

With non-instantaneous actions we have two possible times for resolving choices:

- at the action start (usual approach):

$$a.(P_C \mid P_P) + b.(P_C \mid P_C) + c.0 \xrightarrow{1,a^+} [a]^1 .(P_C \mid P_P) \xrightarrow{1,a^-} P_C \mid P_P$$

- at the action termination (the approach we shall follow):

$$a.(P_C \mid P_P) + b.(P_C \mid P_C) + c.0 \xrightarrow{1,a^+} [a]^1 .(P_C \mid P_P) + b.(P_C \mid P_C) + c.0 \xrightarrow{1,a^-} P_C \mid P_P$$

The usual approach is not suitable in our case:

- in the considered minimal example it would prevent P from being produced during a duplication

A problem with the late handling of choices

The minimal multiscale model

$$C \to C + P \qquad C \to C + C \qquad C \to$$

$$P_C = a.(P_C \mid P_P) + b.(P_C \mid P_C) + c.0$$

Resolving choices at the time of action termination raises another problem:

$$[a]^1 .(P_C \mid P_P) + [b]^2 .(P_C \mid P_C) + [c]^3 .0 \xrightarrow{3,c^-} 0$$

Actions $[a]^1$ and $[b]^2$ get interrupted

- their partners $[\bar{a}]^1$ and $\left[\bar{b}\right]^2$ (both in the context of the process) have to be notified.

Conservative actions

The minimal multiscale model

$$C \to C + P \qquad C \to C + C \qquad C \to$$

$$P_C = a.(P_C \mid P_P) + b.(P_C \mid P_C) + c.0$$

In some cases we would like to avoid interruptions:

$$[a]^1.(P_C \mid P_P) + [b]^2.(P_C \mid P_C) + c.0 \xrightarrow{1,a^-} \ldots$$

We introduce a new action prefixing operation : for conservative actions

- $a.(P_C \mid P_P)$ becomes $a\colon P_P$ (meaning C produces P without being consumed)
- the semantics should be: $a\colon P_P \xrightarrow{1,a^+} [a]^1\colon P_P \xrightarrow{1,a^-} a\colon P_P \mid P_P$
- the completion of a conservative action does not resolve choices

$$[a]^1\colon P_P + [b]^2.(P_C \mid P_C) + c.0 \xrightarrow{1,a^-} a\colon P_P + [b]^2.(P_C \mid P_C) + c.0 \mid P_P$$

Process Algebra with Preemptive and Conservative actions (PAPC)

The syntax of the Process Algebra with Preemptive and Conservative actions (PAPC) consists of processes P and process configurations C_P defined by the following grammar:

$$
\begin{array}{rcl}
P & := & 0 \quad | \quad \alpha.P \quad | \quad \alpha : P \quad | \quad P + P \quad | \quad P \mid P \quad | \quad A \\
C_P & := & P \quad | \quad [\alpha]^l.P \quad | \quad [\alpha]^l : P \quad | \quad C_P + C_P \quad | \quad C_P \mid C_P
\end{array}
$$

where $\alpha \in Act$ and $l \in \mathbb{N}$.

Note: the word preemptive refers to action with the standard prefixing operator

The minimal multiscale model

$$C \rightarrow C + P \qquad C \rightarrow C + C \qquad C \rightarrow$$

$$P_C = a\!:\!P_P + b\!:\!P_C + c.0$$

By taking the given considerations into account, the semantics of PAPC should allow us to obtain the desired behaviour:

$$
\begin{aligned}
P_C \mid Aux &\xrightarrow{1,\tau^+} a\!:\!P_P + [b]^1\!:\!P_C + c.0 \mid Aux^1 \\
&\xrightarrow{2,\tau^+} [a]^2\!:\!P_P + [b]^1\!:\!P_C + c.0 \mid Aux^{12} \\
&\xrightarrow{2,\tau^-} a\!:\!P_P + [b]^1\!:\!P_C + c.0 \mid Aux^1 \mid P_P \\
&\quad \ldots \\
&\xrightarrow{1,\tau^-} [a]^2\!:\!P_P + b\!:\!P_C + c.0 \mid Aux^2 \mid Aux \mid P_P \mid P_P \mid P_P \mid P_C \\
&\xrightarrow{1,\tau^+} [a]^2\!:\!P_P + [b]^1\!:\!P_C + c.0 \mid Aux^{12} \mid Aux \mid P_P \mid P_P \mid P_P \mid P_C \\
&\xrightarrow{3,\tau^+} [a]^2\!:\!P_P + [b]^1\!:\!P_C + [c]^3.0 \mid Aux^{123} \mid Aux \mid P_P \mid P_P \mid P_P \mid P_C \\
&\xrightarrow{3,\tau^-} Aux \mid P_P \mid P_P \mid P_P \mid P_C
\end{aligned}
$$

Semantics of PAPC: formal definition

We have defined a compositional semantics of PAPC in terms of a labelled transition system (LTS).

The LTS results from the combination of four transition relations:

- the handshaking relation $\longrightarrow_H$, describing the occurrence of an action start,
- the completion relation for preemtive actions $\longrightarrow_{CP}$
- the interruption relation $\longrightarrow_I$, describing the interruptions caused by the completion of a preemptive action
- the completion relation for conservative actions $\longrightarrow_{CC}$

In addition, we have a standard rule for recursion.

The handshaking relation

The handshaking relation describes the starting of an action and the coupling of the processes starting complementary actions.

The handshaking relation is $\rightarrow_H \subseteq \mathcal{C} \times \Theta^+ \times \mathcal{C}$, where Θ^+ contains labels of the form (l, α^+) with $l \in \mathbb{N}$ and $\alpha \in Act$.

$$\alpha.P \xrightarrow{1,\alpha^+}_H [\alpha]^1.P \qquad \alpha : P \xrightarrow{1,\alpha^+}_H [\alpha]^1 : P$$

$$\frac{P \xrightarrow{l,\alpha^+}_H P' \qquad l' = min\{\mathbb{N} - Id(P+Q)\} \quad \alpha \in Act}{P + Q \xrightarrow{l',\alpha^+}_H P'[l'/l] + Q}$$

$$\frac{P \xrightarrow{l,\alpha^+}_H P' \quad l' = min\{\mathbb{N} - Id(P \mid Q)\} \quad \alpha \in Act_\tau}{P \mid Q \xrightarrow{l',\alpha^+}_H P'[l'/l] \mid Q}$$

$$\frac{P \xrightarrow{l,\alpha^+}_H P' \qquad Q \xrightarrow{l',\overline{\alpha}^+}_H Q' \quad l'' = min\{\mathbb{N} - Id(P \mid Q)\}}{P \mid Q \xrightarrow{l'',\tau^+}_H P'[l''/l] \mid Q'[l''/l']}$$

Completion of preemptive actions and interruptions

The completion relation for preemptive actions is as follows:

$$[\alpha]^l.P \xrightarrow{l,\alpha^-,\emptyset}_{CP} P \qquad \frac{P \xrightarrow{l,\alpha^-,L}_{CP} P'}{P + Q \xrightarrow{l,\alpha^-,L \cup Id(Q)}_{CP} P'}$$

$$\frac{P \xrightarrow{l,\alpha^-,L}_{CP} P' \quad Q \xrightarrow{M}_{I} Q' \quad M \supseteq (Id(Q) \cap L) \quad \alpha \in Act_\tau}{P \mid Q \xrightarrow{l,\alpha^-,(L \cup M)\setminus(L \cap Id(Q))}_{CP} P' \mid Q'}$$

$$\frac{P \xrightarrow{l,\alpha^-,L}_{CP} P' \quad Q \xrightarrow{l,\overline{\alpha}^-,M}_{CP} Q' \quad N = L \cap M}{P \mid Q \xrightarrow{l,\tau^-,(L \cup M)\setminus N}_{CP} P' \mid Q'}$$

The interruption relation is as follows:

$$[\alpha]^l.P \xrightarrow{\{l\}}_{I} \alpha.P \qquad [\alpha]^l.P \xrightarrow{\emptyset}_{I} [\alpha]^l.P \qquad \alpha.P \xrightarrow{\emptyset}_{I} \alpha.P$$

$$[\alpha]^l : P \xrightarrow{\{l\}}_{I} \alpha : P \qquad [\alpha]^l : P \xrightarrow{\emptyset}_{I} [\alpha]^l : P \qquad \alpha : P \xrightarrow{\emptyset}_{I} \alpha : P$$

$$\frac{P \xrightarrow{L}_{I} P' \qquad Q \xrightarrow{M}_{I} Q'}{P + Q \xrightarrow{L \cup M}_{I} P' + Q'}$$

The completion relation for conservative actions

The completion relation for conservative actions is as follows:

$$[\alpha]^l : P \xrightarrow{l,\alpha^-,\emptyset,P}_{CC} \alpha : P \qquad \frac{P \xrightarrow{l,\alpha^-,L,P''}_{CC} P' \quad Q \xrightarrow{M}_I Q'}{P + Q \xrightarrow{l,\alpha^-,L\cup M,P''}_{CC} P' + Q'}$$

$$\frac{P \xrightarrow{l,\alpha^-,L,P''}_{CC} P'}{P \mid Q \xrightarrow{l,\alpha^-,L,P''}_{CC} P' \mid Q} \qquad \frac{P \xrightarrow{l,\alpha^-,\emptyset,P''}_{CC} P' \qquad Q \xrightarrow{l,\overline{\alpha}^-,\emptyset,Q''}_{CC} Q'}{P \mid Q \xrightarrow{l,\tau^-,\emptyset}_{CP} P' \mid Q' \mid P'' \mid Q''}$$

Completion of a hybrid (preemptive/conservative) synchronization is handled by the following rule:

$$\frac{P \xrightarrow{l,\alpha^-,L,P''}_{CC} P' \qquad Q \xrightarrow{l,\overline{\alpha}^-,M}_{CP} Q' \quad L \subseteq M}{P \mid Q \xrightarrow{l,\tau^-,M\setminus L}_{CP} P' \mid Q' \mid P''}$$

Finally, recursion is handled by the following standard rule:

$$\frac{P \xrightarrow{\ell}_r P'}{A \xrightarrow{\ell}_r P'} \qquad \text{if } A \stackrel{\text{def}}{=} P$$

Bisimulations (1)

A semantics given in terms of a LTS allows us to study behavioural equivalences.

We consider bisimulation in a higher order style.

A symmetric relation $\mathcal{R} \subseteq \mathcal{C} \times \mathcal{C}$ is a *bisimulation* iff whenever $(P, Q) \in \mathcal{R}$, then it holds that:

- if $P \xrightarrow{\ell}_r P'$ for any $r \in \{H, I, CP\}$, then $Q \xrightarrow{\ell}_r Q'$ with $(P', Q') \in \mathcal{R}$.
- if $P \xrightarrow{I,\alpha^-,L,P''}_{CC} P'$, then $Q \xrightarrow{I,\alpha^-,L,Q''}_{CC} Q'$ with $(P', Q') \in \mathcal{R}$ and $(P'', Q'') \in \mathcal{R}$.

The union of all bisimulations is denoted with $\approx$

Theorem. $\approx$ is a congruence

Bisimulations (2)

The bisimulation relation for PAPC turns out to be a very fine behavioural equivalence.

This implies that all ingredients in the process algebra play an important role

Example: $C = a.(C \mid C) \qquad C' = a : C'$

- the behaviours of C and C' seem to be similar
- $C \not\approx C'$
- in the context of a summation C and C' behave differently

Conclusions

We have faced the problem of taking peculiarities of multiscale models into account in the semantics of a process algebra.

We have proposed PAPC, an extension of CCS with

- non-instantaneous actions
- preemptive and conservative actions

We have followed a fundational and minimalist approach

- as future work we shall study richer modelling languages