Modeling biological systems with delays in Bio-PEPA

Giulio Caravagna

Department of Computer Science, University of Pisa
caravagn@di.unipi.it

PEPA Club meeting,
Index

Introduction

Bio-PEPAd
   The Syntax
   A Structural Operational Semantics

Analysis of Bio-PEPAd systems
   Translation to Delay Differential Equations
   Delay Stochastic Simulation Algorithms

Examples
   A toy example
   A model of the cell cycle

Conclusions
   Future work
The use of delays: a dual view

Delays can be used to have abstractions of complex dynamics composed by many sub-events.

1. If the knowledge of the sub-events is incomplete (unobservable, unmeasurable) then the whole dynamics cannot be modeled:
   ▶ Use $\sigma$ as an abstraction.

2. If the whole dynamics is computationally too expensive to be simulated;
   ▶ Use $\sigma$ as a simplification.
Delays-as-duration approach

Given the reaction: \( A + B \overset{k,\sigma}{\longrightarrow} C \)

\[ \begin{align*}
\text{(1) current state: } & \quad X(t) = x; \\
\text{(2) next state: } & \quad X(t + \tau) = x - \{A, B\}, \text{ (i.e. removed reactants)}; \\
\text{(3) scheduled event: } & \quad X(t + \tau + \sigma) = x' + \{C\}, \text{ (i.e. inserted products).}
\end{align*} \]

- any scheduled reaction in \([t, t + \tau]\) has priority.
- reactants excluded from any event in \([t, t + \tau + \sigma]\);
Index

Introduction

Bio-PEPAd

The Syntax

A Structural Operational Semantics

Analysis of Bio-PEPAd systems

Translation to Delay Differential Equations

Delay Stochastic Simulation Algorithms

Examples

A toy example

A model of the cell cycle

Conclusions

Future work
Processes as in Bio-PEPA

\[ S ::= (\alpha, \kappa)_{\text{op}} S \mid S + S \mid C \]
\[ P ::= P \bigotimes_{\mathcal{L}} P \mid S(l) \]

- where \( \text{op} = \downarrow \mid \uparrow \mid \odot \mid \oplus \mid \ominus \);
- \( \kappa, l \in \mathbb{N} \);
- \( \mathcal{L} \subseteq \text{Act}^* \).

Delays are properties of the \textbf{actions}.

Hence, we can use the same Bio-PEPA processes syntax.
Delays require that, for a species $S$, we must know in which scheduled actions it is currently involved:

$$
C_S ::= (\alpha, \kappa) \circ \; C_S \quad | \quad C_S + C_S \quad | \quad C
$$

$$
C_P ::= C_P \prod L C_P \quad | \quad C_S(l, L)
$$

where $L$ is a list of 4-tuples $(l, \kappa, \alpha, \circ)$. 

For instance $S(3, [(2, 1, \alpha, \uparrow)])$ is:

- a species with current concentration level 3;
- involved in a scheduled action $\alpha$:
  - producing 1 concentration level;
  - started when its concentration level was 2.
Bio-PEPAd systems

As in Bio-PEPA, we have a notion of system

\[ \langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, \text{Comp}, \sigma, P_C \rangle \]

where:

- \( \langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, \text{Comp}, P \rangle \) is a Bio-PEPA system;
  - \( \mathcal{V} \): set of compartments;
  - \( \mathcal{N} \): species info(levels, max. conc...);
  - \( \mathcal{K} \): constants;
  - \( \mathcal{F} \) functional rate definitions;
  - \( \text{Comp} \): sequential components.
- \( \sigma \) is a function

\[ \sigma : A \mapsto \{ r \in \mathbb{R} \mid r \geq 0 \} \]

such that \( \sigma(\alpha) \) denotes the delay of action \( \alpha \).
Reusing the Bio-PEPA syntax and notion of systems was crucial.

Moving from Bio-PEPA to Bio-PEPAd models is straightforward.

Given any Bio-PEPA system \( \langle V, N, K, F, Comp, P \rangle \) we have to

1. specify the delays by means of a function \( \sigma \);
2. encode the initial Bio-PEPA process \( P \) in the corresponding initial process configuration \( \mu(P) \);

\[
S_1(l_1) \stackrel{\mathcal{L}}{\longrightarrow} S_2(l_2) \quad \text{corresponds to} \quad S_1(l_1, []) \stackrel{\mathcal{L}}{\longrightarrow} S_2(l_2, [])
\]

hence \( \langle V, N, K, F, Comp, \sigma, \mu(P) \rangle \) is a Bio-PEPAd model.
Index

Introduction

Bio-PEPAd
- The Syntax
  - A Structural Operational Semantics

Analysis of Bio-PEPAd systems
- Translation to Delay Differential Equations
- Delay Stochastic Simulation Algorithms

Examples
- A toy example
- A model of the cell cycle

Conclusions
- Future work
Starting-Terminating (ST) Semantics

In Bio-PEPAd start/completion of an action are detached events.

ST semantics is such that:

- when an action $\alpha$ starts the system exhibits an $\alpha^+$ label;
- when an action $\alpha$ completes the system exhibits an $\alpha^-$ label;

$$S_1 \xrightarrow{\alpha^+} S_2 \xrightarrow{\beta^+} S_3 \xrightarrow{\beta^-} S_n \xrightarrow{\alpha^-} S_{n+1} \ldots$$

- $\alpha$ starts;
- $\beta$ starts;
- $\beta$ completes;
- $\alpha$ completes;
In Bio-PEPA the SOS is given by means of two relations:
  ▶ capability : models both start and completion of an action;
  ▶ stochastic : associates rates to transitions.

Bio-PEPAd SOS is defined by means of three relations:
  ▶ start ($\rightarrow_{st}$):
    ▶ models the start of an action (delay-as-duration approach);
    ▶ labels for computing the rates.
  ▶ completion ($\rightarrow_{co}$):
    ▶ models the completion of an action;
    ▶ labels for re-computing the rates at the start.
  ▶ stochastic ($\rightarrow_{s}$):
    ▶ models both start and completion of an action;
    ▶ associates rates to transitions of the system.
Excerpts from the start relation $\rightarrow_{st} \subseteq \mathcal{C} \times \Theta^+ \times \mathcal{C}$

For $S$ acting as a reactant in "$\kappa S + \ldots \rightarrow \ldots$"

$((\alpha, \kappa)\downarrow S)(l, L) \xrightarrow{(\alpha^+, [S:\downarrow(l,\kappa)])} \rightarrow_{st} (S)(l-\kappa, L@[((l, \kappa, \alpha, \downarrow)]) \quad \kappa \leq l \leq N$

- action $\alpha$ starts exhibiting label $\alpha^+$;
- the concentration level changes from $l$ to $l - \kappa$;
- $S : \downarrow(l, \kappa)$ is used to compute the rate like in Bio-PEPA;
- $(l, \kappa, \alpha, \downarrow)$ is enqueued in $L$;
- the constraint for starting $\alpha$ are like in Bio-PEPA.

Information are take from $((\alpha, \kappa)\downarrow S)(l, L)$ since we model the start!
For $S$ acting as a product in $\ldots \rightarrow \kappa S + \ldots$

$\left( (\alpha, \kappa) \uparrow S \right)(l, L) \xrightarrow{\left( \alpha^+, \left[ S: \uparrow(l, \kappa) \right] \right)}_{st} (S)(l, L@\left[ (l, \kappa, \alpha, \uparrow) \right]) \quad 0 \leq l + \rho \pi L \leq N$

- action $\alpha$ starts exhibiting label $\alpha^+$;
- the concentration level is not changing (delay-as-duration);
- $S : \uparrow(l, \kappa)$ is used to compute the rate like in Bio-PEPA;
- $(l, \kappa, \alpha, \uparrow)$ is enqueued in $L$;
- $\rho \pi L$ is the number of levels scheduled for production.
Excerpts from the completion relation $\rightarrow_{co} \subseteq C \times \Theta^- \times C$

For $S$ acting as a reactant in "$\kappa S + \ldots \rightarrow \ldots$"

$$\phi \alpha L = (l, \kappa, \alpha, \downarrow)$$

$$S(l', L) \xrightarrow{(\alpha^-, [S:\downarrow(l, \kappa)])} S(l', \zeta \alpha L)$$

- action $\alpha$ completes exhibiting label $\alpha^-$;
- the concentration level is not changing;
- $\phi \alpha L$ denotes the first entry in $L$ (to complete) regarding $\alpha$;
- $S : \downarrow(l, \kappa)$ is the entry in $L$ (conc. level is not $l'$);
- $\zeta \alpha L$ is $L$ where $\phi \alpha L$ has been removed;

Information are take from $L$ since we model completion!
For $S$ acting as a product in "... $\rightarrow \kappa S + ...$"

\[
\phi \alpha L = (l, \kappa, \alpha, \uparrow)
\]

\[
S(l', L) \xrightarrow{(\alpha^-, [S:\uparrow(l,\kappa)])} \text{co } S(l' + k, \zeta \alpha L)
\]

- action $\alpha$ completes exhibiting label $\alpha^-$;
- the concentration level is **now** changing from $l'$ to $l' + \kappa$;
- $\phi \alpha L$ denotes the first entry in $L$ (to complete) regarding $\alpha$;
- $S : \uparrow(l, \kappa)$ is the entry in $L$ (conc. level is not $l'$);
- $\zeta \alpha L$ is $L$ where $\phi \alpha L$ has been removed;
The stochastic relation $\rightarrow_s \subseteq \overline{P} \times \Gamma \times \overline{P}$

\[
P \xrightarrow{\alpha^+, w}_{st} P'
\]

\[
\langle V, N, K, F, Comp, \sigma, P \rangle \xrightarrow{\alpha^+, r_\alpha[w, N, K], \sigma(\alpha)}_s \langle V, N, K, F, Comp, \sigma, P' \rangle
\]

\[
P \xrightarrow{\alpha^-, w}_{co} P'
\]

\[
\langle V, N, K, F, Comp, \sigma, P \rangle \xrightarrow{\alpha^-, r_\alpha[w, N, K], \sigma(\alpha)}_s \langle V, N, K, F, Comp, \sigma, P' \rangle
\]

- $\alpha^+ / \alpha^-$ for start/completion of an $\alpha$;
- $r_\alpha[w, N, K]$ computed as in Bio-PEPA;
- $\sigma(\alpha)$ exhibited as a label;
Analysis available in Bio-PEPA
Analysis currently available in Bio-PEPAd
Index

Introduction

Bio-PEPAd
The Syntax
A Structural Operational Semantics

Analysis of Bio-PEPAd systems
Translation to Delay Differential Equations
Delay Stochastic Simulation Algorithms

Examples
A toy example
A model of the cell cycle

Conclusions
Future work
Delay Differential Equations (DDEs)

\[ \frac{dX}{dt} = f_X(t, X(t), X(t - \sigma_1), \ldots, X(t - \sigma_n)) \]

- \( \sigma_1 > \ldots > \sigma_n \geq 0, \sigma_i \in \mathbb{R} \) are delays;
- \( X(t - \sigma_i) \) is the state of the system at the past time \( t - \sigma_i \);

We now how to encode a Bio-PEPA system \( \langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, \text{Comp}, P \rangle \) in a set of ODEs:

1. definition of the stoichiometry matrix;
2. derivation of the kinetic law vector \( \nu_{KL} \);
3. association of the deterministic variables with the components.

Steps (1) and (3) are unaffected by the use of delays.
Index

Introduction

Bio-PEPAd
   The Syntax
   A Structural Operational Semantics

Analysis of Bio-PEPAd systems
   Translation to Delay Differential Equations
   Delay Stochastic Simulation Algorithms

Examples
   A toy example
   A model of the cell cycle

Conclusions
   Future work
Our target for simulation is:

- an extension of the Gillespie’s SSA;
- DSSA with a delay-as-duration approach (Barrio et al.).

We now how to encode a Bio-PEPA system $\langle V, N, K, F, Comp, P \rangle$ in an input of the SSA.

This conservative extension permits to use the same encoding.

Of course, the system is now simulated by the DSSA.
1. Initialize the time $t = t_0$ and the system state $X(t_0) = x_0$.
2. Evaluate all the propensity functions;
3. Compute time for next reaction, $\tau$;
4. Select next reaction, $R_j$;
   (A) If delayed reaction $R_k$ is scheduled at time $t + \tau_k$ and $\tau_k < \tau$ then update $x = x + \nu_k^p$ and $t = t + \tau_k$;
   (B) else:
     (B1) If $R_j \in R_{nd}$ then update $x = x + \nu_j$ and $t = t + \tau$;
     (B2) If $R_j \in R_{c}$, schedule $R_j$ at time $t + \sigma_j + \tau$, update $x = x + \nu_k^r$ and set time to $t + \tau$;
5. go to step 2.
Index

Introduction

Bio-PEPAd
  The Syntax
  A Structural Operational Semantics

Analysis of Bio-PEPAd systems
  Translation to Delay Differential Equations
  Delay Stochastic Simulation Algorithms

Examples
  A toy example
  A model of the cell cycle

Conclusions
  Future work
A toy example

\[ A \xrightarrow{k,\sigma'} B \]

- kinetic law is \( f_{MA}(k) \);
- transformation has delay \( \sigma' \);
- the initial state is described by the vector \( \mathbf{X}(t_0) = (3, 0)^T \).

The Bio-PEPA processes are

\[ A \overset{\text{def}}{=} (\alpha, 1) \downarrow A \quad B \overset{\text{def}}{=} (\alpha, 1) \uparrow B \]

and the initial Bio-PEPAd process configuration is

\[ A(3, [\!\!\!\!\!\!] \{\alpha\}) \overset{\text{merge}}{\longrightarrow} B(0, [\!\!\!\!\!\!]) \].
Figure: A graphical representation of the stochastic process obtained by applying the semantics to the process configuration for the toy example.

Notation \((n_1, n_2)\) : \(m\) is such that:

- \(n_1\) is the concentration of the species \(A\);
- \(n_2\) is the concentration of the species \(B\);
- \(m\) is the number of instances of \(\alpha\) currently scheduled.
Index

Introduction

Bio-PEPAd
  The Syntax
  A Structural Operational Semantics

Analysis of Bio-PEPAd systems
  Translation to Delay Differential Equations
  Delay Stochastic Simulation Algorithms

Examples
  A toy example
    A model of the cell cycle

Conclusions
  Future work
A model of the cell cycle

Four phases: resting ($G_1$), DNA replication ($S$), gap ($G_2$), mitosis ($M$).

We classify cells in two populations (Vilassana et al.):

- $T_I$: cells in the interphase ($G_1$, $S$ and $G_2$);
- $T_M$: cells in the mitotic phase ($M$).
Non–delayed events (mass-action kinetics):

- in any phase a cell may die (i.e. via apoptosis):
  - (action $\gamma$) a $T_I$ die with rate $d_2$;
  - (action $\delta$) a $T_M$ die with rate $d_3$.

- a cell divides:
  - (action $\beta$) one $T_M$ becomes two $T_I$ with rate $a_4$. 
Delayed events (mass-action kinetics):

- a cell lasts $\sigma'$ time units in the interphase, then starts mitosis (the passage from $G_1$ to $S$ and from $S$ to $G_2$ are abstracted by using $\sigma'$).

  - (action $\alpha$) one $T_I$ becomes one $T_M$ with rate $a_1$. 
Model specification in Bio-PEPA/Ad

The Bio-PEPA/Bio-PEPAAd processes are:

\[ T_I \overset{\text{def}}{=} (\alpha, 1)\downarrow + (\beta, 2)\uparrow + (\gamma, 1)\downarrow \]
\[ T_M \overset{\text{def}}{=} (\alpha, 1)\uparrow + (\beta, 1)\downarrow + (\delta, 1)\downarrow . \]

With the following kinetic laws

\[ f_\alpha = f_{MA}(a_1) \quad f_\beta = f_{MA}(a_4) \quad f_\gamma = f_{MA}(d_2) \quad f_\delta = f_{MA}(d_3). \]

The Bio-PEPA/Bio-PEPAAd process modeling the interactions is

\[ T_I(n_I^0) \overset{\{\alpha, \beta\}}{\otimes} T_M(n_M^0) \]

where:

- \( n_I^0 \) is the initial concentration levels for the \( T_I \) species;
- \( n_M^0 \) is the initial concentration levels for the \( T_M \) species.
The delay are specified by a function \( \sigma \):

\[
\sigma(\alpha) = \sigma' \quad \sigma(\beta) = \sigma(\gamma) = \sigma(\delta) = 0.
\]

The Bio-PEPAd initial process configuration is

\[
T_I(n_{0I}, [ ]) \{\alpha, \beta\} T_M(n_{0M}, [ ])
\]

How can we analyze this model?

1. translation into DDEs;
2. stochastic simulation.
Translation into DDEs

Given the stoichiometry matrix $D$ computed as in Bio-PEPAd, we compute the vector of the kinetic laws:

$$
\nu_{KL} = \begin{pmatrix}
    a_1 T_I(t - \sigma(\alpha)) \\
    a_4 T_M(t - \sigma(\beta)) \\
    d_2 T_I(t - \sigma(\gamma)) \\
    d_3 T_M(t - \sigma(\delta))
\end{pmatrix}
= \begin{pmatrix}
    a_1 T_I(t - \sigma') \\
    a_4 T_M(t) \\
    d_2 T_I(t) \\
    d_3 T_M(t)
\end{pmatrix}.
$$

The following set of DDEs can be computed ($d\bar{x}/dt = D \times \nu_{KL}$):

$$
\frac{dT_I}{dt} = 2a_4 T_M - d_2 T_I - a_1 T_I(t - \sigma')
$$

$$
\frac{dT_M}{dt} = a_1 T_I(t - \sigma') - d_3 T_M - a_4 T_M.
$$

The number of cells that enter mitosis at time $t$ depends on the number of cells that entered the interphase $\sigma$ time units before, $T(t - \sigma)$. 
Stochastic simulation

We can encode the system as if it were a Bio-PEPA one:

\[ T_I \xrightarrow{a_1} T_M \text{ with delay } \sigma' \quad T_M \xrightarrow{a_4} 2T_I \quad T_I \xrightarrow{d_2} T_M \xrightarrow{d_3}. \]

With initial conditions as follows:

\[ X(t_0) = \begin{pmatrix} n_0^I \times h \times N_A \times v \\ n_0^M \times h \times N_A \times v \end{pmatrix} \]

And we can then apply the DSSA with the delay-as-duration approach.
Conclusions

What is Bio-PEPAd:

- a conservative extension of Bio-PEPA;
- a non-Markovian PA;
- a language whose models can be analyzed in different frameworks (DDEs+DSSAs).

This will be part of my Ph.D. thesis plus:

- DSSAs (the PDA, the mPDA, the DDA+mPDA);
- A PA for multiscale modelling of biological systems.

Ongoing/Future work

Ongoing work (for an extended version):
- which non–Markov processes are we describing?
- are equivalence notions for Bio-PEPA usable in Bio-PEPAd?
- relation between Bio-PEPA/Bio-PEPAd semantics?

Thanks to Jane for giving me the opportunity to visit you.