On the Interpretation of Delays in Delay Stochastic Simulation of Biological Systems

Giulio Caravagna

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

CompMod 2009
Eindhoven, November 2009
Systems Biology

An interdisciplinary field of research regarding:

- the study of the **interactions** between the **components** of **biological systems**, and how these interactions give **rise** to the function and behavior of that system.

  i.e.: metabolic pathways, cellular evolution, ....

- The main aim consists of building **real** models:
  - via observations / measurements / interactions with experts;
  - by using specific languages.

- models have to be analyzed:
  - via **simulation** (deterministic, stochastic);
  - model checking;
Systems Biology Workflow

- Biological systems
  - Deterministic model
    - ODEs
  - Stochastic model
    - CME
    - SSA
**The Stochastic Simulation Algorithm [Gillespie, 1977]**

Any reaction $R_j$ is associated with a propensity function $a_j(x)$ providing the probability of $R_j$ to fire in state $x$.

1. Initialize the time $t = t_0$ and the system state $x = x_0$.
2. With the system in state $x$ at time $t$, evaluate all the $a_j(x)$ and their sum $a_0(x) = \sum_{j=1}^{M} a_j(x)$.
3. Given two random numbers $r_1, r_2 \in U[0;1]$, generate values for $\tau$ and $j$ accordingly to

   $$\tau = \frac{1}{a_0(x)} \ln\left(\frac{1}{r_1}\right)$$

   $$\sum_{i=1}^{j-1} a_i(x) < r_2 \cdot a_0(x) \leq \sum_{i=1}^{j} a_i(x)$$

   then update $x = x + \nu_j$ and $t = t + \tau$, go to step 2.
The use of delays: a dual view

Given a complex dynamics composed by many sub-events (i.e. the cell-cycle). Suppose you know, at least, the average time $\sigma$ for completion of the whole complex dynamics.
The use of delays: a dual view

Given a complex dynamics composed by many sub-events (i.e. the cell-cycle). Suppose you know, at least, the average time $\sigma$ for completion of the whole complex dynamics.

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

G. Caravagna (Pisa)
The use of delays: a dual view

Given a complex dynamics composed by many sub-events (i.e. the cell-cycle). Suppose you know, at least, the average time $\sigma$ for completion of the whole complex dynamics.

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.
Systems Biology Workflow (extended with delays)

- **biological systems**
  - deterministic model
    - ODEs
    - DDEs
  - stochastic model
    - DCME
    - CME
    - DSSA
    - SSA
Delay Stochastic Simulation Algorithms (DSSAs)

- extensions of the Gillespie’s Algorithm;
- related to the Delayed Chemical Master Equation;
- combine delayed and non–delayed reactions;

Barrio et al., 2006: DSSA with an interpretation of delays as duration. Cai and Anderson, 2007: improve the performance of Barrio’s DSSA.

All based on an interpretation of delays as duration.
Delays as duration

Given the reaction: $A + B \xrightleftharpoons[k,\sigma]{\rightarrow} C$

(1) current state: $X(t) = x$;
(2) next state: $X(t + \tau) = x - \{A, B\}$, (i.e. removed reactants);
(3) scheduled event: $X(t + \tau + \sigma) = x' + \{C\}$, (i.e. inserted products).
Delays as duration

Given the reaction: $A + B \overset{k,\sigma}{\rightarrow} C$

(1) current state: $X(t) = x$;

(2) next state: $X(t + \tau) = x - \{A, B\}$, (i.e. removed reactants);

(3) scheduled event: $X(t + \tau + \sigma) = x' + \{C\}$, (i.e. inserted products).

- any scheduled reaction in $[t, t + \tau]$ has priority.
- reactants excluded from any event in $[t, t + \tau + \sigma]$;
Delays as duration: 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 $x = x + \nu^p_k$ and $t = t + \tau_k$;
   
   (B) else:
      
      (B1) If $R_j \in \mathcal{R}_{nd}$ then update $x = x + \nu_j$ and $t = t + \tau$;
      
      (B2) If $R_j \in \mathcal{R}_{c}$, schedule $R_j$ at time $t + \sigma_j + \tau$, update $x = x + \nu^r_k$ and set time to $t + \tau$;

5. go to step 2.
An example: 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:

- in any phase a cell may die (i.e. via apoptosis);
- a cell divides (one \( T_M \) becomes two \( T_I \)).

Delayed events:

- 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 \( \sigma \)).
The DDEs model by Villasana et al. is:

\[
\begin{align*}
\frac{dT_I}{dt} &= 2a_4 T_M - d_2 T_I - a_1 T_I(t - \sigma) & T_I(t) &= \phi_0(t) \text{ for } t \in [-\sigma, 0] \\
\frac{dT_M}{dt} &= a_1 T_I(t - \sigma) - d_3 T_M - a_4 T_M & T_M(t) &= \phi_1(t) \text{ for } t \in [-\sigma, 0]
\end{align*}
\]

with extended initial conditions in \([-\sigma, 0]\) and explicit number of cells.

Where are delays?

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

Delays correspond to dependencies to past–states of the system.
Deterministic Simulation

\[
\sigma = 1.0, \quad (0,0), \quad (0,0), \quad (0,0)
\]
\[
\sigma = 10.0, \quad \text{osc.}, (0,0), \quad \text{osc.}, (0,0)
\]
Deterministic Simulation

\[
\begin{array}{c|cccc}
\sigma = 1.0 & R-I & R-II & R-III / V & R-IV \\
\infty & 50 & 15 & 238 \\
\end{array}
\]

First observable day with $T_I(t)$ and $T_M(t) < 1.0$. 

G.Caravagna (Pisa)
Deterministic Simulation

$$\sigma = 10.0 \quad \infty \quad 59 \quad 12 \quad 440$$

First observable day with $T_I(t)$ and $T_M(t) < 1.0$
The equivalent stochastic model is given by three non–delayed reaction:

\[ T_M \xrightarrow{a_4} 2T_I \]
\[ T_I \xrightarrow{d_2} \]
\[ T_M \xrightarrow{d_3} \]

and the delayed reaction:

\[ T_I \xrightarrow{a_1} T_M \text{ with delay } \sigma \]
Stochastic Simulation (delays as duration)

<table>
<thead>
<tr>
<th></th>
<th>R-I</th>
<th>R-II</th>
<th>R-III / V</th>
<th>R-IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma = 1.0$</td>
<td>$\infty$</td>
<td>64</td>
<td>29</td>
<td>302</td>
</tr>
</tbody>
</table>
Stochastic Simulation (delays as duration)

\[ \sigma = 10.0 \]

<table>
<thead>
<tr>
<th></th>
<th>R-I</th>
<th>R-II</th>
<th>R-III / V</th>
<th>R-IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma = 10.0 )</td>
<td>( \infty )</td>
<td>224</td>
<td>126</td>
<td>1072</td>
</tr>
</tbody>
</table>
DDEs v.s. Delay as Duration

- in general, comparison with DDEs is not necessary meaningful;
- for this simulations, with $10^5$ starting cells, we obtain the same **qualitative** behaviour (i.e. growth and eradication);

Is also **quantitatively** equivalent?

<table>
<thead>
<tr>
<th></th>
<th>DDEs</th>
<th>DDA Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-II with $\sigma = 1.0$</td>
<td>50</td>
<td>64</td>
</tr>
<tr>
<td>R-II with $\sigma = 10.0$</td>
<td>59</td>
<td>224</td>
</tr>
<tr>
<td>R-III with $\sigma = 1.0$</td>
<td>15</td>
<td>29</td>
</tr>
<tr>
<td>R-III with $\sigma = 10.0$</td>
<td>12</td>
<td>126</td>
</tr>
<tr>
<td>R-IV with $\sigma = 1.0$</td>
<td>238</td>
<td>302</td>
</tr>
<tr>
<td>R-IV with $\sigma = 10.0$</td>
<td>440</td>
<td>1072</td>
</tr>
</tbody>
</table>
The purely delayed approach

Given the reaction: $A + B \xrightleftharpoons[k,\sigma]{} C$

1. current state: $X(t) = x$;
2. next state: $X(t + \tau) = x$, (i.e. no action here);
3. scheduled event: $X(t + \tau + \sigma) = x' - \{A, B\} + \{C\}$, (i.e. application if still possible, $\{A, B\} \in x'$).
The purely delayed approach

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

(1) current state: \( X(t) = x \);
(2) next state: \( X(t+\tau) = x \), (i.e. no action here);
(3) scheduled event: \( X(t+\tau+\sigma) = x' - \{A, B\} + \{C\} \), (i.e. application if still possible, \( \{A, B\} \in x' \)).

- any scheduled reaction in \([t, t+\tau]\) has priority;
- reactants can participate in reactions firing in \([t, t+\tau+\sigma]\) (i.e. cell death);
The purely delayed approach: 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 applicable delayed reaction $R_k$ is scheduled at time $t + \tau_k$ and $\tau_k < \tau$ then update $x = x + \nu_k$ and $t = t + \tau_k$;
   
   (B) else:
      
      (B1) If $R_j \in \mathcal{R}_{nd}$ then update $x = x + \nu_j$ and $t = t + \tau$;
      
      (B2) If $R_j \in \mathcal{R}_c$, schedule $R_j$ at time $t + \sigma_j + \tau$ and set time to $t + \tau$;

5. go to step 2.
Stochastic Simulation (purely delayed)

\[
\begin{array}{|c|c|c|c|}
\hline
\sigma &= 1.0 & R-I & \infty \\
R-II & 51 & R-III / V & 17 \\
R-IV & 214 & \\
\hline
\end{array}
\]
Stochastic Simulation (purely delayed)

\[ \sigma = 10.0 \]

<table>
<thead>
<tr>
<th></th>
<th>R-I</th>
<th>R-II</th>
<th>R-III / V</th>
<th>R-IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma = 10.0 )</td>
<td>( \infty )</td>
<td>67</td>
<td>20</td>
<td>248</td>
</tr>
<tr>
<td></td>
<td>DDEs</td>
<td>DDA Simulation</td>
<td>PDA Simulation</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>------</td>
<td>----------------</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>R-II with $\sigma = 1.0$</td>
<td>50</td>
<td>64</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>R-II with $\sigma = 10.0$</td>
<td>59</td>
<td>224</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>R-III with $\sigma = 1.0$</td>
<td>15</td>
<td>29</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>R-III with $\sigma = 10.0$</td>
<td>12</td>
<td>126</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>R-IV with $\sigma = 1.0$</td>
<td>238</td>
<td>302</td>
<td>214</td>
<td></td>
</tr>
<tr>
<td>R-IV with $\sigma = 10.0$</td>
<td>440</td>
<td>1072</td>
<td>248</td>
<td></td>
</tr>
</tbody>
</table>

**Table:** Average eradication times given in *days* for DDE model, DDA and PDA stochastic models. For the stochastic models the entries represent the sample of 100 simulations.
Conclusions and Future Work

Our approach has to be improved:

- it may schedule multiple times the same reactants;
- hard checking applicability (fresh reactants);

Possible solution is a marking procedure (≠ removal) of the reactants.

It is also of interest the combination of both the approaches.

All these simulation algorithms must be embedded in the semantics of languages for the specification of models.

We could inherit advices from already existing languages (i.e. Petri Nets have got similar notions of reaction firing with delays).