

An Alternative to Gillespie's Algorithm for Simulating Chemical Reactions

Roberto Barbuti Andrea Maggiolo-Schettini
Paolo Milazzo Angelo Troina

Dipartimento di Informatica, Università di Pisa, Italy

Edinburgh – April 4, 2005

Outline

Introduction

Simulating chemical reactions

- Gillespie's algorithm
- Our algorithm
- A comparison

Examples and applications

- Probabilistic MultiSet
- Rewriting
- Simulation results

Conclusions

Introduction

Simulating chemical reactions

Gillespie's algorithm

Our algorithm

A comparison

Examples and applications

Probabilistic MultiSet Rewriting

Simulation results

Conclusions

Outline

Introduction

Simulating chemical reactions

Gillespie's algorithm

Our algorithm

A comparison

Examples and applications

Probabilistic MultiSet
Rewriting

Simulation results

Conclusions

Introduction

Chemical reactions are described by the *law of mass action*

- ▶ the speed of a reaction is proportional to the concentrations of the individual reactants involved
- ▶ differential equations

Gillespie's simulation algorithm

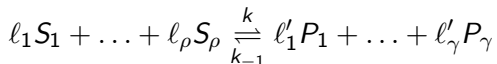
- ▶ stochastic method based on the theory of collisions
- ▶ each reaction takes a (continuous) random time which is exponentially distributed

The simulation algorithm we propose

- ▶ performs discrete time steps of (fixed) length Δt ;
- ▶ assume that at each step at most one reaction may occur (randomly chosen).

Background

Usual notation for chemical reactions:



where:

- ▶ S_i, P_i are molecules
- ▶ ℓ_i, ℓ'_i are stoichiometric coefficients
- ▶ k, k_{-1} are the kinetic constants

For the *law of mass action*, the forward rate of a reaction is:

$$\frac{dP}{dt} = k[S_1]^{\ell_1} \dots [S_\rho]^{\ell_\rho}$$

and the backward rate is:

$$\frac{dS}{dt} = k_{-1}[P_1]^{\ell'_1} \dots [P_\gamma]^{\ell'_\gamma}$$

Gillespie's stochastic approach (1)

- ▶ assumes a stochastic reaction constant c_μ for each chemical reaction R_μ
- ▶ $c_\mu dt$ is the probability that a particular combination of reactant molecules of R_μ react in an infinitesimal time interval dt

The total probability (denoted $a_\mu dt$) of R_μ to occur in the infinitesimal time interval dt is

$$a_\mu dt = h_\mu c_\mu dt$$

where h_μ is the number of distinct molecular reactant combinations.

Gillespie's stochastic approach (2)

Example:

solution with X_1 molecules S_1 and X_2 molecules S_2

reaction $R_1 : S_1 + S_2 \rightarrow 2S_1$

- ▶ $h_1 = X_1 X_2$
- ▶ $a_1 = X_1 X_2 c_1$
- ▶ $a_1 dt = X_1 X_2 c_1 dt$

reaction $R_2 : 2S_1 \rightarrow S_1 + S_2$

- ▶ $h_2 = \frac{X_1(X_1-1)}{2}$
- ▶ $a_2 = \frac{X_1(X_1-1)}{2} c_2$
- ▶ $a_2 dt = \frac{X_1(X_1-1)}{2} c_2 dt$

Gillespie's algorithm

Given a set of reactions $\{R_1, \dots, R_M\}$ and a current time t

1. The time $t + \tau$ at which the next reaction will occur is randomly chosen with τ exponentially distributed with parameter $\sum_{\nu=1}^M a_{\nu}$;
2. The reaction R_{μ} that has to occur at time $t + \tau$ is randomly chosen with probability $a_{\mu} dt$.

At each step t is incremented by τ and the chemical solution is updated.

At each step the *probability density function*

$$P_g(\tau, \mu) = \exp\left(-\sum_{\nu=1}^M a_{\nu} \tau\right) \cdot a_{\mu} dt$$

gives the probability that the next reaction will occur in the time interval $(t + \tau, t + \tau + dt)$ and will be R_{μ} .

Our algorithm (1)

- ▶ assumes that in a very small (fixed) time interval Δt at most one reaction may occur
- ▶ Δt depends on the number and on the rates of the chemical reactions

Basic idea:

- ▶ divide the rate of each reaction (given by the law of mass action) by an arbitrarily great integer value N
- ▶ use the result as the probability of each reaction to occur in Δt

Our algorithm (2)

Given a set of reactions $\{R_1, \dots, R_M\}$, and assuming a volume of 1 litre

- ▶ Δt has to be fixed to $\frac{1}{MN}$
- ▶ N is such that

$$0 < \frac{k_\mu [S_{\mu 1}]^{\ell_{\mu 1}} \dots [S_{\mu \rho}]^{\ell_{\mu \rho}}}{N} \leq 1$$

for $1 \leq \mu \leq M$ and for all the possible concentrations (assumed to be finite) of $S_{\mu 1}, \dots, S_{\mu \rho}$

The probability of R_μ is

$$P(R_\mu) = \begin{cases} \frac{k_\mu [S_{\mu 1}]^{\ell_{\mu 1}} \dots [S_{\mu \rho}]^{\ell_{\mu \rho}}}{N} & \text{if } R_\mu \text{ can occur} \\ 0 & \text{otherwise} \end{cases}$$

Our algorithm (3)

The algorithm iterates the following steps:

1. A reaction R_μ is randomly chosen (all the reactions are equiprobable);
2. The chosen R_μ is performed with probability $P(R_\mu)$.

The probability of choosing and performing R_μ in Δt is

$$P(\mu) = \frac{1}{M} P(R_\mu)$$

and the probability of performing no reactions in Δt is

$$P_0 = 1 - \sum_{\nu=1}^M P(\nu)$$

Comparing the two algorithms (1)

Assume Δt infinitesimal. Results:

- ▶ The probability of performing R_μ in Gillespie's algorithm is equivalent (with an approximation) to the probability of choosing and performing R_μ in our algorithm, that is

$$a_\mu dt \approx P(\mu)$$

- ▶ A step in Gillespie's algorithm can be simulated by a sequence of steps in our algorithm having (approximatively) the same probability

Approximations are introduced by deriving c_μ from k_μ

Comparing the two algorithms (2)

Gillespie's algorithm	Our Algorithm
<ul style="list-style-type: none">- Assumes c_μ in general unknown derived by k_μ (approx)	<ul style="list-style-type: none">+ Uses k_μ usually well-known
<ul style="list-style-type: none">- Considers reactions individually	<ul style="list-style-type: none">+ Is based on the law of mass action does not depend on the measure units (scalability)
<ul style="list-style-type: none">+ Precise reaction times	<ul style="list-style-type: none">- Assumes at most one reaction in Δt

Probabilistic MultiSet Rewriting (1)

The Probabilistic MultiSet Rewriting formalism (PMSR)

- ▶ is a simple example of formalism using our algorithm
- ▶ describes the behaviour of a chemical solution as a probabilistic transition system

Definition (PMSR Rule). A *Probabilistic MultiSet Rewriting rule* is a triple (M_1, p, M_2) where:

- ▶ M_1, M_2 are two different multisets;
- ▶ $p \in]0, 1]$ is the *probabilistic constant* of the rule.

A probabilistic rewriting rule (M_1, p, M_2) can be denoted also with the more usual notation $M_1 \rightarrow_p M_2$.

Probabilistic MultiSet Rewriting (2)

Definition (PMSR System). A *Probabilistic MultiSet Rewriting system* is a pair (M, \mathcal{R}) , where M is a multiset and \mathcal{R} is a finite set of rewriting rules.

Definition (Semantics). The semantics of PMSR is the probabilistic transition system in which states are PMSR systems and transitions are described by the following inference rules:

$$\frac{R_\mu \in \mathcal{R} \quad R_\mu = M_\mu \rightarrow_{p_\mu} M'_\mu \quad M_\mu \subseteq M}{(M, \mathcal{R}) \xrightarrow{P(\mu)} ((M \setminus M_\mu) \cup M'_\mu, \mathcal{R})}$$

$$\frac{}{(M, \mathcal{R}) \xrightarrow{P_0} (M, \mathcal{R})}$$

Lotka and Brusselator reactions

An Alternative to
Gillespie's
Algorithm for
Simulating
Chemical
Reactions

Barbuti
Maggiolo-Schettini
Milazzo
Troina

Outline

Introduction

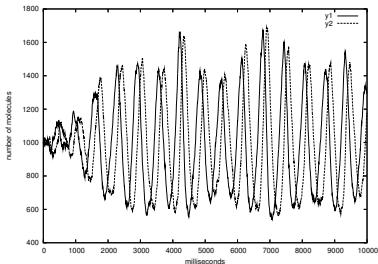
Simulating
chemical reactions

Gillespie's algorithm
Our algorithm
A comparison

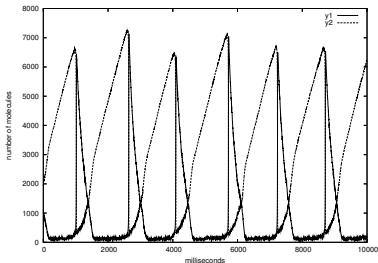
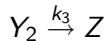
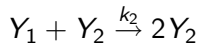
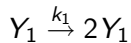
Examples and
applications

Probabilistic MultiSet
Rewriting
Simulation results

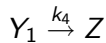
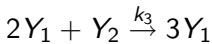
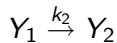
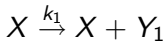
Conclusions



Lotka reactions

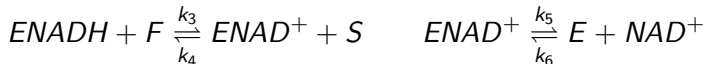


Brusselator reactions



Sorbitol Dehydrogenase

Real enzymatic activity involving Fructose and Sorbitol.



$$k_1 = 6.2 \times 10^{-6} \text{ s}^{-1} \text{ pM}^{-1}$$

$$k_2 = 33 \text{ s}^{-1}$$

$$k_3 = 2.2 \times 10^{-9} \text{ s}^{-1} \text{ pM}^{-1}$$

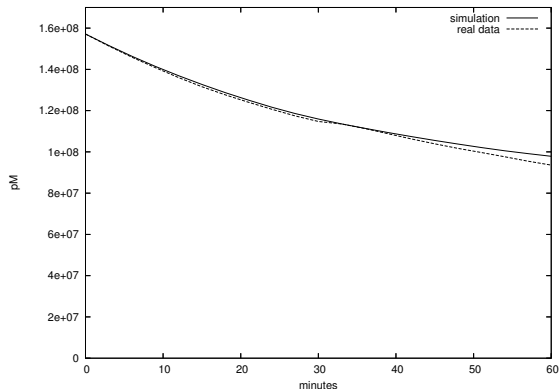
$$k_4 = 7.9 \times 10^{-9} \text{ s}^{-1} \text{ pM}^{-1}$$

$$k_5 = 227 \text{ s}^{-1}$$

$$k_6 = 6.1 \times 10^{-7} \text{ s}^{-1} \text{ pM}^{-1}$$

$$k_7 = 1.9 \times 10^{-3} \text{ s}^{-1}$$

Sorbitol Dehydrogenase – simulation 1



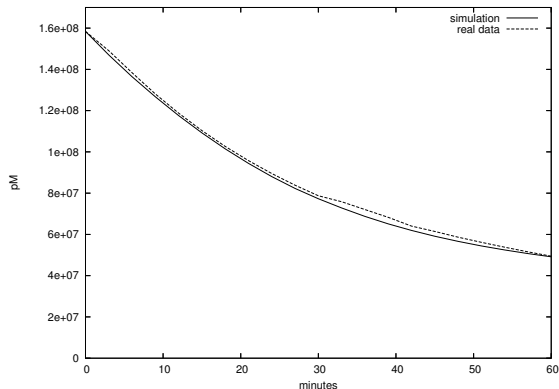
concentration of NADH over time
simulation (solid line) vs experiments (dashed line)

Initial (picomolar) concentrations

$$[E] = 210 \quad [F] = 4 \times 10^{11} \quad [NADH] = 1.6 \times 10^8$$

$$[S] = [NAD^+] = [ENADH] = [ENAD^+] = 0$$

Sorbitol Dehydrogenase – simulation 2



concentration of NADH over time
simulation (solid line) vs experiments (dashed line)

Initial (picomolar) concentrations

$$[E] = 430 \quad [F] = 4 \times 10^{11} \quad [NADH] = 1.6 \times 10^8$$

$$[S] = [NAD^+] = [ENADH] = [ENAD^+] = 0$$

An Alternative to
Gillespie's
Algorithm for
Simulating
Chemical
Reactions

Barbuti
Maggiolo-Schettini
Milazzo
Troina

Outline

Introduction

Simulating
chemical reactions

Gillespie's algorithm
Our algorithm
A comparison

Examples and
applications

Probabilistic MultiSet
Rewriting
Simulation results

Conclusions

Conclusions

We have:

- ▶ introduced a probabilistic algorithm for simulating chemical reactions
- ▶ compared our algorithm with Gillespie's one
- ▶ shown simulation results

Advantages of our algorithm:

- ▶ based on the law of mass action
- ▶ scalable (on the measure unit of concentrations)

Prototype implementation:

<http://www.di.unipi.it/~milazzo/biosims/>