Consistency Techniques for Hybrid Simulations

Marco Bottalico[1]

1 **Dipartimento di Scienze,**
Università "G. d'Annunzio" di Chieti-Pescara, Italy
bottalic@sci.unich.it

Abstract

The goal of this paper is to show consistency techniques methods and hybrid stochastic/deterministic models to describe biochemical systems and their behaviour through the ordinary differential equations.

1998 ACM Subject Classification I.6.1 Simulation Theory

Keywords and phrases Consistency techniques, deterministic approach, stochastic approach

Digital Object Identifier 10.4230/LIPIcs.ICLP.2011.255

1 Introduction and problem description

In this paper, we investigate hybrid methods based on simulation of stochastic and deterministic models for biochemical systems, with consistency techniques in ordinary differential equations to have a preliminary vision on dissimilar methods to simulate different biochemical systems in Biocham.

2 Background and overview of the existing literature

System biology is an interdisciplinary science, integrating experimental activity and mathematical modeling, which studies the dynamical behaviors of biological systems. An important problem in the modeling these systems is to characterize the dependence of certain properties on time and space. One frequently applied strategy is the description of the change of state variables by differential equations. If only temporal changes are considered, *ordinary differential equations* (ODEs) are used; for changes in time and space, partial differential equations are appropriate [3].

A variety of formalisms for modeling biological systems has been proposed in literature but in this paper we want to investigate only the consistency techniques in ordinary differential equations [2] and a new hybrid stochastic and deterministic model for biochemical systems [1]. There are two formalisms for mathematically describing the time behavior of a spatially homogeneous chemical system: the *deterministic approach* and the *stochastic approach*. The deterministic approach regards the time evolution as a continuous, wholly predictable process which is governed by a set of coupled, ordinary differential equations (the "reaction-rate equations"). The stochastic approach regards the time evolution as a kind of random-walk process which is governed by a single differential-difference equation (the "master equation"). Fairly simple kinetic theory arguments show that the stochastic formulation of chemical kinetics has a firmer physical basis than the deterministic formulation, but unfortunately the stochastic master equation is often mathematically intractable [7].

There is also a way to make exact numerical calculations within the framework of the stochastic formulation without having to deal with the master equation directly. We are


© Marco Bottalico;
licensed under Creative Commons License NC-ND
Technical Communications of the 27th International Conference on Logic Programming (ICLP'11).
Editors: John P. Gallagher, Michael Gelfond; pp. 255–260

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

talking about the Monte Carlo procedure to numerically simulate the time evolution of the given chemical system. Like the master equation, this "stochastic simulation algorithm" correctly accounts for the inherent fluctuations and correlations that are necessarily ignored in the deterministic formulation. Moreover this algorithm never approximates infinitesimal time increments dt by finite time steps Δt. The feasibility and utility of the simulation algorithm are demonstrated by applying it to several well-known model chemical systems, including the Lotka model, the Brusselator, and the Oregonator [7].

3 Goal of the research

How we have explained in the previous section, the ordinary differential equations (ODEs) play a crucial role in the deterministic model. A first order (ODE) system $\mathcal{O}$ is a system of the form

$$\begin{array}{rcl} u_1^{'}(t) & = & f_1(t, u_1(t), ..., u_n(t)) \\ u_2^{'}(t) & = & f_2(t, u_1(t), ..., u_n(t)) \\ & \vdots & \\ u_n^{'}(t) & = & f_n(t, u_1(t), ..., u_n(t)) \end{array}$$

In [2] the author uses the vector representation $u^{'}(t) = f(t, u(t))$ or more simply $u^{'} = f(t, u)$. At this point, there are two assumptions:

(1) the function f is sufficiently smooth;

(2) the existence and uniqueness of a solution.

Now, given an initial condition $u(t_0) = u_0$ and for the second assumption, the solution of $\mathcal{O}$ is a function $s^* : R \to R^s$ satisfying $\mathcal{O}$ and the initial condition $s^*(t_0) = u_0$.

Although for some classes of ODEs the solution can be represented in closed form, most ODE systems cannot be solved explicitly [2]. The *discrete variable method* aim at approximating the solution $s^*(t)$ of any ODE system, not over a continuous range of t, but only at some points $t_0, t_1, ..., t_m$. This method include *one-step methods* and *multi-step methods*; in general these methods do not guarantee the existence of a solution within a given bound.

The *interval analysis method* instead, was introduced by Moore [16] in 1966. These methods provide numerically reliable enclosures of the exact solution at points $t_0, t_1, ..., t_m$. To achieve the result, they typically apply a one-step Taylor interval method and make extensive use of automatic differentiation to obtain the Taylor coefficients[2].

The major problem of interval analysis methods on ODE systems is the explosion of the size of resulting boxes at point $t_0, t_1, ..., t_m$. For the author, there are two reasons for this explosion: at first this method has a tendency to accumulate errors from point to point, second the approximation of an arbitrary region by a box (wrapping effect) may introduce considerable loss of accuracy after a number of steps.

For all these reasons, in[17, 2] they show how to provide a unifying framework to extend traditional numerical techniques to intervals providing reliable enclosures. The first contribution is to extend explicit and implicit, one-step and multi-step methods to intervals. The second one is to generalize interval techniques into a two-step process: a forward process (to compute an enclosure) and a backward process (to reduce this enclosure).

4 Current status of the research

The stochastic effects play an important role in biological processes leading to an increase in stochastic modelling attempts. The main problem related to the stochastic simulations regards times and computations which are very expensive [1].
The stochastic models have gained considerable attention when experiments conducted at the level of single cells showed the existence of a non-negligible level of noise in intracellular processes, like transcriptions and translation [4]. The dynamics of a stochastic system is described by the *chemical master equation* and in the 1976 Gillespie devised two exact algorithms to numerically simulate the stochastic time evolution of coupled chemical reactions, which are equivalent to solving the chemical master equation [7]. Only recently, modifications to the original chemical master equation have been proposed to further speed up simulations. The most important methods involve the averaging over fast reactions [8], application of quasi-steady-state theory [9], grouping together reactions that occur in fast succession [10].

Another strategy is to model those processes that either involve large number of particles or have fast rates, in a deterministic way, keeping stochastic the remaining ones [1]. There are two recent algorithms to simulate biochemical systems in such hybrid framework that have been proposed [11, 12]. In both cases, the main idea is to first predict the time in which a stochastic event should occur and then evolve the system of ordinary differential equations. At specific instant in time, the system is updated, and it is checked whether the stochastic event has to be performed or not. Instead in [1] the authors propose a rigorous mathematical ground for hybrid stochastic and deterministic modelling in a natural way. There are three different algorithms: the direct hybrid method, the first reaction hybrid method and the next reaction hybrid method. The main difference between the first two approaches and the second one is essentially one: they are based on a prediction correction heuristic for the realization of the stochastic part that can be seen as an approximation to the simultaneous solution of the system of ODEs which in [1] are precisely calculated.

Consider N chemical species $S_1, ..., S_N$ involved in M reactions $R_1, ..., R_M$. Chemical species are modelled in terms of number of molecules $X(t) = (X_1(t), ...,$
$X_N(t))$. The reaction rate for each reaction R_j is specified by a so-called propensity function $a_j = a_j(X(t), t)$, which is equal to the product rate constant c_j and the number of possible combinations of reactant molecules involved in reaction R_j. Once a reaction R_j is performed, the number of molecules for each species is updated according to the state change vector v_j, i.e., $X(t) \leftarrow X(t) + v_j$ [1].

The **deterministic model** is based on the law of mass action, where a system of coupled ordinary differential equations (ODEs) is established for the time evolution of the number of molecules $X(t) \in R_+^N$

$$\frac{d}{dt}X(t) = \sum_{j=1}^{M} v_j a_j(X(t), t) \tag{1}$$

with some initial value $X(t_0) \in R_+^N$. While the system should be described as a vector of integers, this model needs real values for $X(t)$. This is however acceptable under the assumption of large number of molecules ($X_i(t) >> 1$) so that the relative error can be neglected [1].

The **stochastic model** is based on physical laws and the idea that chemical reactions are essentially random processes, the stochastic formulation of chemical reactions is given

in terms of a Markov jump process $X(t) \in N^N$ [13]. Its characterization is based on the probability $a_j(X(t),t)dt$ of a reaction R_j occurring in the next infinitesimal time interval $[t, t+dt]$. Denoting by $T_j(t)$ the time at which reaction R_j first occur after t, this amounts to write that

$$\mathbf{P}[T_j(t) \in [t, t+dt] | X(t)] = a_j(X(t),t)dt. \tag{2}$$

In [1] the authors consider a partition of the reactions $R_1, ..., R_M$ into those modelled stochastically (labeled with index $j \in \mathcal{S}$) and those modelled deterministically (labeled with index $j \in \mathcal{D}$). Now we can write the evolution equation for $X(t) \in R^N$ which is given by the following hybrid system

$$dX(t) = \sum_{j \in \mathcal{D}} v_j a_j(X(t),t)dt + \sum_{j \in \mathcal{S}} v_j dN_j(t) \tag{3}$$

To partition the reactions the authors suggest some methods:

- run a fully stochastic realization and analyze the frequencies/propensities of each reaction;
- use biological insight (i.e. in [1] the authors say that seems reasonable to model gene regulatory parts stochastically, while metabolic reactions deterministically);
- for each reaction choose adaptively between two approaches using a criterion based on the number of the molecules and its propensity function.

To check if the algorithms based on hybrid model (direct hybrid method, first and next reaction methods) obtained good results they tested them in a intracellular growth of bacteriophage T7 derived by [14]. From the experiment appears that the hybrid simulations are about 100 times as fast as the fully stochastic ones without compromising the results accuracy (fig. 1).

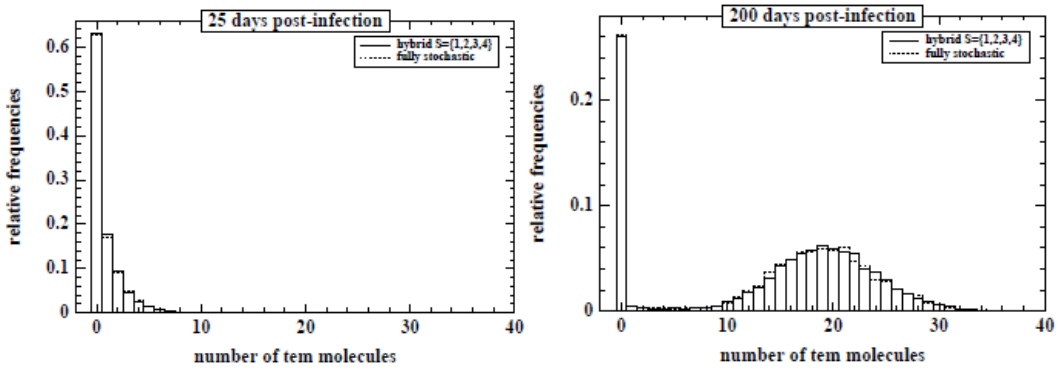

Figure 1 Hybrid kinetics for the bacteriophage T7 model (reaction R_1, R_2, R_3 and R_4 modelled stochastically, reactions R_5 and R_6 modelled deterministically) compared to the the reference fully stochastic model (based on 10^4 realizations) [1].

5 Preliminary results accomplished

The goal is to implement in Biocham some techniques to realize hybrid simulation, combining different kinds and different nature models, in a qualitative and quantitative optical, with discrete and continue dynamics. The solution is to provide the specific language with a

multi level description mechanism for the modelization; the second step is to distinguish in the formalism, the common characteristics from the details. At last we want to specify the criteria to change, during the simulation, the formalism.

6 Open issues and expected achievements

The importance of precise analysis to study and comprise biological phenomena involve different kind of models. On the one hand, it is necessary to describe some parts in a rigorous and accurate numerical method (for example methods based on ordinary differential equations or stochastic methods). On the other hand, the lack of evidence, drives the analysis on purely qualitative models (boolean or discrete models).

References

1. A. Alfonsi and E. Cancès and G. Turinici and B. Di Ventura and W. Huisinga, Exact simulation of hybrid stochastic and deterministic models for biochemical systems, English, MICMAC - INRIA Rocquencourt - INRIA - Ecole Nationale des Ponts et Chaussées , 20, Research Report, INRIA, RR-5435, 2004
2. Y. Deville and M. Janssen and P. Van Hentenrycy, Consistency Techniques in Ordinary Differential Equations, Constraints, 7, 3-4, 2002, 289-315
3. K. Edda and H. Ralf and K. Axel and W. Christoph and L. Hans, 1, Hardcover, 3527310789, Systems Biology in Practice: Concepts, Implementation and Application 2005
4. M.B. Elowitz and A.J. Levine and E.D. Siggia and P.S. Swain, Science, 5584, 1183-1186, Stochastic gene expression in a single cell, 297, 2002
5. V. Gupta and R. Jagadeesan and V.A. Saraswat and DG. Bobrow, Programming in Hybrid Constraint Languages, Hybrid Systems, 1994 ,226-251, P.J. Antsaklis and W. Kohn and A. Nerode and S. Sastry, Hybrid Systems II, Springer, Lecture Notes in Computer Science, 999, 1995, 3-540-60472-3
6. V.A. Saraswat, Concurrent Constraint Programming Languages, The MIT Press, Cambridge, MA, 1993
7. D.T. Gillespie, J. Phys. Chem., 25„ 2340-2361, Exact stochastic simulation of coupled chemical reactions, 81, 1977
8. E.L. Haseltine and J.B. Rawlings, The Journal of Chemical Physics, gillespie_algorithm, kinetics, simulation, stochastic, 6959–6969, 2007-12-11 22:31:52, 2, AIP, Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics, 117, 2002
9. C.V. Rao and A.P. Arkin, The Journal of Chemical Physics, 4999–5010, 2007-10-23, AIP, Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm, 118, 2003
10. K. Burrage and T. Tian and P. Burrage, Progress Biophys Mol Biol, 217-234, A multi-scale approach for simulation chemical reaction systems, 85, 2004
11. T.R. Kiehl and R.M. Mattheyses and M.K. Simmons, Hybrid simulation of cellular behavior, Bioinformatics, 20, 3, 2004, 316-322
12. K. Takahashi and K. Kaizu and B. Hu and M. Tomita, A Multi-Algorithm, Multi-Timescale Method for Cell Simulation, Bioinformatics, 2004, 20, 538–546
13. D.T. Gillespie, Markov Processes-An Introduction for Physical Scientists, Academic Press, 1992
14. Srivastava, R. and You, L. and Summers, J. and Yin, J., Journal of Theoretical Biology, deterministic, modeling, stochastic_simulation, 3, 309–321, 2008-10-20 , 2, Stochastic vs. Deterministic Modeling of Intracellular Viral Kinetics, 218, 2002

15 L. Calzone and F. Fages and S. Soliman, BIOCHAM: an environment for modeling biological systems and formalizing experimental knowledge, Bioinformatics, 22, 14, 2006, 1805-1807

16 R.E. Moore, Interval Analysis, Prentice-Hall, Englewood Cliffs, N.J., 1966

17 Y. Deville and M. Janssen and P. Van Hentenryck, Multistep Filtering Operators for Ordinary Differential Equations, CP, 1999, 246-260, J. Jaffar, Principles and Practice of Constraint Programming - CP'99, 5th International Conference, Alexandria, Virginia, USA, October 11-14, 1999, Proceedings, CP, Springer, Lecture Notes in Computer Science, 1713, 1999