

Hybrid approximation of stochastic concurrent constraint programs

Luca Bortolussi* Alberto Policriti†

Modeling in computational systems biology is dominated by two formalisms, the first one being (mainly ordinary) differential equations and the second one being continuous-time stochastic processes [1, 2]. Both methodologies have their roots reaching back to physical and chemical arguments that, at least for modeling biochemical reactions, give strong foundations to the approach. Recently, stochastic modeling has received a lot of attention, thanks also to the use of process algebras (in their stochastic variants) as modeling languages that allow to construct models in a compositional and reusable way [3, 4].

It is well known that the choice of “mathematics” to use is not just a matter of personal taste, but rather it is connected with the specific features of the system under analysis. Differential models are appealing, as they count on a (mature) bulk of mathematical instruments for their analysis, numerical solvers above all. However, biological processes often have inherently discrete components (they involve molecules interacting) and hence a continuous approximation may lead to incorrect results, especially when small-sized populations are taken into account. Stochastic models can be more precise, as they are intrinsically discrete and also able to capture the effects of noise. However, their exact simulation with the (celebrated) Gillespie algorithm [5] is computationally more expensive and, in general, the analysis of stochastic models is more difficult.

Relating stochastic process algebras (SPA) and differential equation models is a difficult and interesting problem. Specifically, we are interested in finding methods that automatically translate SPA models into ODEs. This translation must somehow preserve both the structure and the behavior of the models: if we start from a SPA program, the associated set of ODEs should be behaviorally equivalent to it, in the sense of showing the same (qualitative) dynamics. Some techniques have been proposed in the literature to cope with this problem. The first one that appeared has been developed by Hillston [6] for PEPA [7]. Subsequently, this technique has been extended to stochastic π -calculus [8, 9] and to stochastic concurrent constraint programming (sCCP [10]) [11, 12]. Unfortunately, all these methods, though working well for some systems, fail to preserve behavior in general.

A paradigmatic example is the Repressilator [13], a synthetic genetic circuit involving three genes repressing cyclically each other, which exhibits oscillatory behavior. A simple stochastic model for Repressilator can be written using the formalism of gene gates [14], where genes are considered as logical gates pro-

*Dept. of Mathematics and Computer Science, University of Trieste, Italy.
Email: luca@dimi.units.it

†Dept. of Mathematics and Computer Science, University of Udine, Italy; Istituto di Genomica Applicata, Udine, Italy. Email: alberto.policriti@dimi.uniud.it

ducing a protein as output with a speed regulated positively or negatively by several input proteins. The model of Repressilator, involving three gene gates negatively regulating each other in a cyclic fashion, manifests neat oscillations. If we convert this model into ODEs, using the method of [8] or [9], we obtain a set of equations that does not oscillate at all. In this procedure, we are approximating boolean states of our processes with a continuous variable, thus making a questionable assumption. Indeed, also the average trace of the stochastic model does not show any oscillatory pattern: the quantity of produced proteins converges to a stationary value. Actually, negative gates can be in two states, active or inhibited, thus averaging over all traces destroys the information about the discrete switching dynamics within a gene, which is the inducer of the oscillatory pattern. Avoiding the approximation of inner states with continuous variables leads straight on to a decoupling of a finite number of possible scenarios and thereafter to hybrid automata.

Hybrid automata (HA) are a formalism mixing discrete and continuous ingredients: essentially, they are finite automata extended with a set of variables evolving continuously in each state, according to state-specific differential equations. The discrete dynamics is given by transitions between states, which are triggered when activation conditions on variables are satisfied and, when taking place, also reset the value of some variables.

Our proposal consists in a translation machinery from sCCP to hybrid automata. The basic idea is to identify a finite set of “states” of the stochastic system, each characterized by a specific dynamics. These “states” will constitute the discrete skeleton of the hybrid automaton, while continuous laws and discrete transitions will be defined according to the dynamics within each “state”. In particular, we will show that the hybrid version of the Repressilator, obtained from the model based on gene gates, preserves the oscillatory behavior (its main behavioral property). These preliminary results and the flexibility offered by the interplay between discrete and continuous dynamics suggest HAs an interesting target formalism for the approximation of stochastic process algebras and, perhaps, for modeling techniques with a stochastic ingredient.

The translation defined here certainly does not solve all the problems. Most notably, there is still a qualitative divergence between stochastic models and the corresponding hybrid automata when the concentration of species is low. However, the mixed discrete/continuous nature of dynamics of hybrid automata offers the possibility of defining hybrid models that should be able to capture the behavior (of stochastic systems) in a more refined way. Moreover, we are investigating the use of non-determinism in order to encode some stochastic variability of the original system.

References

- [1] H. Kitano, *Foundations of Systems Biology*. MIT Press, 2001.
- [2] H. Kitano, “Computational systems biology,” *Nature*, vol. 420, pp. 206–210, 2002.
- [3] C. Priami and P. Quaglia, “Modelling the dynamics of biosystems,” *Briefings in Bioinformatics*, vol. 5, no. 3, pp. 259–269, 2004.

- [4] A. Regev and E. Shapiro, “Cellular abstractions: Cells as computation,” *Nature*, vol. 419, 2002.
- [5] D. Gillespie, “Exact stochastic simulation of coupled chemical reactions,” *J. of Physical Chemistry*, vol. 81, no. 25, 1977.
- [6] J. Hillston, “Fluid flow approximation of pepa models,” in *Proceedings of the Second International Conference on the Quantitative Evaluation of Systems (QEST05)*, 2005.
- [7] J. Hillston, *A Compositional Approach to Performance Modelling*. Cambridge University Press, 1996.
- [8] L. Bortolussi and A. Policriti, “Relating stochastic process algebras and differential equations for biological modeling,” *Proceedings of PASTA 2006*, 2006.
- [9] L. Cardelli, “From processes to odes by chemistry,” *Draft*, 2006.
- [10] L. Bortolussi, “Stochastic concurrent constraint programming,” in *Proceedings of 4th International Workshop on Quantitative Aspects of Programming Languages, QAPL 2006, ENTCS*, vol. 164, pp. 65–80, 2006.
- [11] L. Bortolussi, *Constraint-based approaches to stochastic dynamics of biological systems*. PhD thesis, PhD in Computer Science, University of Udine, 2007. Available at <http://www.dmi.units.it/~bortolu/files/reps/Bortolussi-PhDThesis.pdf>.
- [12] L. Bortolussi and A. Policriti, “Stochastic concurrent constraint programming and differential equations,” in *Proceedings of QAPL 2007*, 2007.
- [13] M. Elowitz and S. Leibler, “A synthetic oscillatory network of transcriptional regulators,” *Nature*, vol. 403, pp. 335–338, 2000.
- [14] R. Blossey, L. Cardelli, and A. Phillips, “A compositional approach to the stochastic dynamics of gene networks,” *T. Comp. Sys. Biology*, pp. 99–122, 2006.