# Multiscale Simulation of Rule-Based Models

Leonard A. Harris[1] and James R. Faeder[1]

***Short Abstract* — Rule-based modeling is a systematic approach for overcoming problems of combinatorial complexity in biological systems. Rules are most commonly used as generators of reaction networks. It has recently been demonstrated, however, that rules can also be used as generators of reaction firings within a kinetic Monte Carlo context, resulting in significant gains in computational efficiency. Here, we extend this approach to "leaping" techniques, a popular class of accelerated-stochastic simulation methods. Strategies for implementing this non-trivial extension are discussed and results demonstrating the utility of rule-based leaping are shown.**

***Keywords* — Rule-based modeling, stochastic simulation, tau-leaping.**

BIOLOGICAL molecules can undergo a variety of post-translational modifications and interact to form multitudes of heterogeneous complexes. As a result, a relatively small number of molecular components can generate a huge number of molecular species that are coupled through an even larger number of reactions [1]. This "combinatorial explosion" in the reactive state-space severely limits the ability of a modeler to compose a mechanistic model of a biological system as well as to analyze its dynamical behavior.

In reality, however, interactions involving complex biological molecules are often localized to specific regions on the molecules. As such, one often need not know the full specification of a molecule's state in order to enumerate its possible interactions. In other words, the problem of combinatorial complexity can often be overcome by recognizing that different molecular species can participate in the same interactions if they contain the same reactive sub-units. This concept can be formalized within the context of interaction "rules." Rules represent the minimal, or canonical, representation of a complex network of interacting entities and can be used to generate a network through iterative application to a set of initial species [2]. The main shortcoming of this approach, however, is that because all possible species and reactions are generated, the cost of simulating the network, which is often the computational bottleneck, is unaffected. Recently, an alternative use of rules has been demonstrated in which rules rather than reactions act as the event generators within a kinetic Monte Carlo (kMC) context [3, 4]. The approach significantly accelerates simulations relative to traditional kMC implementations because there are generally far fewer rules than reactions.

[1]Department of Computational Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15260. E-mail: lharris@pitt.edu, faeder@pitt.edu.

A bottleneck in the "rule-based" kMC approach, however, is its "one reaction at a time" nature. It is well known that kMC methods become prohibitively slow when even one reactive species in a system exists in large numbers [5, 6]. A number of algorithms have been proposed with the aim of overcoming this problem, of which "leaping" methods are a particularly interesting type. Leaping methods, which achieve accelerations by firing multiple reactions in parallel, operate by determining a period of time over which all reaction rates in a system remain constant within a specified tolerance and then choosing the number of times each reaction fires within that interval from an appropriate probability distribution [6]. For systems with wide disparities in species populations, this approach has been shown to lead to orders-of-magnitude accelerations relative to traditional kMC [7, 8].

Here, we extend this approach to rule-based models. The extension is non-trivial because full knowledge of the interconnectivity of a network is necessary for calculating an appropriate time step in a leaping simulation. At the level of rules, this interconnectivity is hidden to some extent because rules may apply to non-overlapping sets of reactants and products. Thus, the effect of firing a rule on the quantities used to compute the rates of other rules cannot always be determined directly from the variables explicitly tracked in a simulation. Additional "auxiliary" quantities may need to be calculated as the simulation proceeds in order to perform the time step calculation accurately. Moreover, it may be advantageous to approximate certain quantities and then apply so-called "post-leap checking" in order to correct for any errors introduced by the approximation. We discuss these strategies and automated schemes for implementing them and present results demonstrating the utility of rule-based leaping.

## REFERENCES

[1] Faeder JR, Blinov ML, Goldstein B, Hlavacek WS (2005) Rule-based modeling of biochemical networks. *Complexity* **10**, 22-41.

[2] Hlavacek WS, et al. (2006) Rules for modeling signal-transduction systems. *Science STKE* **2006**, re6.

[3] Yang J, Monine MI, Faeder JR, Hlavacek WS (2007) Kinetic Monte Carlo method for rule-based modeling of biochemical networks. arXiv:0712.3773.

[4] Danos V, Feret J, Fontana W, Krivine J (2007) Scalable Simulation of Cellular Signaling Networks. *Lect. Notes Comput. Sci* **4807**, 139-157.

[5] Gillespie DT, (1976) A general method for numerically simulating the stochastic evolution of coupled chemical reactions. *J. Comput. Phys.* **22**, 403-434.

[6] Gillespie DT (2001) Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.* **115**, 1716-1733.

[7] Harris LA, Clancy P (2006) A "partitioned leaping" approach for multiscale modeling of chemical reaction dynamics. *J. Chem. Phys.* **125**, 144107.

[8] Harris LA, Piccirilli AM, Majusiak ER, Clancy P (2007) Probing the Effects of Stochasticity in Biochemical Reaction Networks using Multiscale Simulation. arXiv:0709.0561.