Hybrid particle-population simulations for rule-based models

Justin S. Hogg[1] and James R. Faeder[2]

Network-based simulation is limited by the number of unique species in the network, whereas particle-based simulation is limited by the total number of molecules in the system. Hybrid particle-population methods can address these limits by treating a subset of common species as population variables. Bonding between populations and particles is an obstacle to efficient hybrid simulation in rule-based models. In a naïve formulation, each dissociation reaction entails an expensive graph labeling step to determine if the products are population species. We present an approach that avoids the labeling step by representing population-particle bonds with component state labels.

***Keywords* — hybrid simulation, agent compression, rule-based modeling, combinatorial complexity.**

I. Combinatorial complexity & simulation limits

Eukaryotic signal transduction involves complex assemblies of macromolecules whose molecular constituency and post-transcriptional modification state changes dynamically. Such systems can be compactly represented in a rule-based modeling framework, such as BioNetGen [1], which treats complexes as hierarchical graphs and reactions as transformations on subgraphs. The combinatorial possibilities in signal transduction often exceed the capacity of network-based simulation approaches, such as the ODE and SSA methods. An alternative is particle-based simulation, in which each molecule is represented as an individual object that can bind other objects to form complexes [2]. Combinatorial complexity is efficiently simulated in a particle-based approach, but the size of the simulation is limited by the number of objects that can be represented in physical memory.

II. Hybrid simulation & Rule-based models

Species with high populations, *e.g.* soluble ligand, lead to inefficiency in particle simulations. Hybrid particle-population simulation, *i.e.* agent compression [3], attempts to improve efficiency by treating common species as population variables. Naïve implementation of a rule-based hybrid is inefficient due to bonding between populations and particles. The dissociation of a complex particle into disjoint products necessitates an expensive graph labeling algorithm to determine if the products should be lumped into a population variable. In reverse, the association of a population to a particle requires instantiation of the population species as a particle prior to bond formation.

III. An efficient Rule-based Hybrid method

We propose a method of re-writing a rule-based model in a form that represents population species bound to particles by component state labels. The revised model is composed of a network component and a particle component. Interactions between the components are limited to state changes on particles coupled to spawning/elimination of population species. Explicit bonds between populations and particles are not permitted. This eliminates the need for graph labeling and particle instantiation.

Initially, a subset of molecule types are selected to be treated as population variables. Network generation is applied to the population molecules to establish the network component of the model. Next, for each component site on the particle molecules, we enumerate the population species that bind at the site. A unique state label is generated for each possible binding partner. Finally, the reaction rules are rewritten to operate on state labels. This last step is the most complicated and requires careful calculation of the overlaps between population species graphs and rule patterns.

IV. Application to EGF signaling model

We apply the hybrid approach to an EGF signaling model [4] with 5 molecule types and 23 reaction rules, which induces a network with 356 species and 3749. All molecule types except the EGF receptor are treated as populations. The original model with 2.1×10^6 molecules is reduced to a system with 1.8×10^5 particles and 8 population variables.

References

[1] Faeder JR, Blinov ML, Hlavacek WS (2009) Rule-based modeling of biochemical systems with BioNetGen. *Meth. Mol. Biol* **500**, 113-167.

[2] Yang J, Monine MI, Faeder JR, Hlavacek WS (2009) Kinetic Monte Carlo method for rule-based modeling of biochemical networks. *Phys Rev E Stat Nonlin Soft Matter Phys* **78**, 031910.

[3] Wendel S, Dibble C (2007) Dynamic Agent Compression. *J. Artificial Societies and Social Simulation* **10**.

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

Acknowledgements: JSH was funded as a trainee under NIH T32 EB009403 as part of the HHMI-NIBIB Interfaces Initiative. JRF acknowledges support from NSF CCF-0829788.

[1]Carnegie Mellon – University of Pittsburgh Ph. D. Program in Computational Biology, Pittsburgh, PA 15260. E-mail: jsh32@pitt.edu

[2]Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, PA 15260. E-mail: faeder@pitt.edu