Modular Construction of Rule-Based Models

John A.P. Sekar[1] and James R. Faeder[2]

***Short Abstract* — Rule-based modeling provides a compressed and modular approach to modeling reaction networks. Here we extend the modular nature of reaction rules to rule-sets, enabling piecewise creation and flexible assembly of rule-based models. Creating and modifying modular rule-sets mimics evolutionary design and enables strict documentation of structural assumptions. Automated assembly of rule-sets allows synthesis of combinatorial design spaces such as perturbations and uncertain structural hypotheses. We demonstrate this approach using a model exploring multiple structural hypotheses for unregulated activation of mutant Ras in cancer.**

***Keywords* — Rule based model, model aggregation, model composition, logical modeling, reaction kinetics, piecewise modeling, model construction, evolutionary design**

I. Introduction

RULE-BASED MODELING is inherently a modular approach to building reaction networks[1,2], where reactions with identical kinetics on identical molecular substructures are grouped into reaction rules. Recently, there have been several innovative attempts to scale-up the rule-based framework to accommodate increasingly large biochemical models. Approaches have included wrappers for rule-building routines [3] and formal hierarchies of molecule types [4]. Our approach is complementary and involves creating generic rule-sets or modules and then syntactically modifying them to create variants that can be flexibly and automatically aggregated.

II. Modules (Rule-Sets)

Modules are *created* by arbitrarily grouping sets of rules isolated to a specific interaction between molecule types, for example, rules governing the interaction between a specific substrate and its kinase. These "generic" modules can be syntactically *diversified* by sequentially adding molecular context and modifying rate-expressions. This creates a hierarchy of modules akin to evolutionary diversification of protein structures and interactions. The modules can then be flexibly *aggregated* by a modeler-defined configuration that pairs modules to Boolean values (0-never load, 1-always load) and Boolean expressions (conditional loading).

III. Example

We examine the construction of a model of unregulated Ras activation in cancer [5]. Intrinsic Ras activity and Ras interactions with GAP, GEF and effectors are compiled in the form of 16 generic rules collated into 4 modules. These modules were used as a starting point for the module diversification process.

Two contextual variants are constructed for each module simply by adding wild-type or mutant context. Additional variation is encountered in the form of independent structural hypotheses for the mutant. These variants were constructed by sequential instructions to modify the rate expressions in the mutant modules. The choice to load any combination of these hypotheses was encoded as Boolean parameters in the aggregator, i.e. b_{RGA} for reduced GTPase activity, b_{GI} for GAP insensitivity and b_{IEA} for increased effector affinity.

In all, 61 rule variants were possible with unique structure, context and rate expression. The aggregator *automatically* compiles the rate-expressions of the mutant variants into unified expressions (with model choices reflected as Boolean weights). For example, the k_{cat} of the Ras-GAP interaction in the mutant depends on whether the GAP-insensitivity and reduced GTPase activity assumptions are deployed:

$$k_{cat,mut} = (1- b_{GI})\, k_{cat} + b_{GI}((1-b_{RGA})k_{hyd} + b_{RGA}f_{RGA}k_{hyd})$$

The final model is *automatically* generated and has 32 reaction rules (2*16) with rate expressions dependent on 3 Boolean parameters, enabling 8 (2^3) different embedded models. We are in the process of using this approach to build models of large systems with many related molecule types, e.g., the ErbB family of receptors and their ligands.

IV. Advantages

Modularity minimizes the number of rules that need to be manually written. It enables independent creation, debugging and curation of independent parts of the model. Module diversification is suited to rapidly create homologs, mutants and structural variants. Conditional aggregation enables embedding families or spaces of model structures within the same rule-based model. Experimental design on these spaces can be exploited for selection and comparative analysis of models sharing significant network structure.

Acknowledgements: This work was funded in part by NSF grant CCF-0829788.

[1,2]Department of Computational and Systems Biology, University of Pittsburgh. E-mail: {jas237,faeder}@pitt.edu

References

[1] Hlavacek W., Faeder J.R. and Blinov M. (2006), Rules for modeling signal-transduction systems, *Science's STKE* **344**, 6

[2] Danos V. and Laneve C. (2004), Formal molecular biology, *Theoretical Computer Science* **325**, 69-110.

[3] Muhlich JL, Lopez CF, Bachman JA, PySB, pysb.org

[4] Danos V., Feret J., Fontana W., Harmer R. and Krivine J. (2009) *Transactions on Computational Systems Biology XI* **5750**, 116-137

[5] Stites E.C., Trampont P.C., Zhong M., Ravichandran K.S. (2007) *Science* **318**, 463-467