Mechanistic analysis of reaction network models

José-Juan Tapia[1], John A.P. Sekar[2] and James R. Faeder[3]

***Abstract*— Chemical reaction networks have been used for a long time to model biochemical systems. Rule-based modeling is a newer graph-based approach that allows explicit consideration of molecular structures and reaction configurations. We have previously developed a tool called Atomizer that infers molecular structures and interactions from reaction network models to generate rule-based representations. Here, we have atomized reaction network models in the BioModels database and performed a statistical analysis of their composition in terms of the five basic graph operations that can occur in reactions.**

I. Background

RULE-BASED MODELING is an alternative approach to building kinetic models, that explicitly encodes information about structures and reaction mechanisms [1]. In the rule-based framework, molecules and complexes are explicitly represented as graphs. Reactions, modeled as *reaction rules*, are explicit graph rewritings using five basic operations: add and delete molecules, add and remove bonds and change internal states [2]. However, in the reaction network framework, each reaction naïvely models the creation and deletion of structureless chemical species. Any structure present in the species has to be manually encoded, either using annotations or using an *ad hoc* labeling convention (e.g. `A_B` to represent a complex of `A` and `B`). The Atomizer algorithm [4] attempts to learn these conventions and extract this hidden structure from each RNM. Using clues from reaction stoichiometry and common naming conventions, the algorithm relearns the explicit mechanistic interactions that were implicitly encoded in the reactions.

BioModels is a repository of user-submitted reaction network models (RNMs) focusing on cellular biochemistry and other biological processes [3]. It includes 540 models curated by the BioModels group and 650 uncurated models. Here, we use Atomizer to analyze the composition of RNMs in both the curated and uncurated sets of the BioModels database and compare this to a control set of rule-based models.

II. Results

The balance of composition between bond and state changes versus creation and deletion is indicative of the degree to which mechanistic information can be extracted from a particular reaction network. The fractional occurrence of a particular graph operation within a model can be compared across a whole suite of models. Figure 1 shows a histogram of the fractional occurrence for each graph operation in three sets of models: atomized curated BioModels, atomized non-curated BioModels, and a control set of BioNetGen rule-based models. The control set is predominantly composed of bond and state changes, reflecting the explicit encoding involved in the rule-based framework. The non-curated set is predominantly composed of molecule creation and deletion. These operations do not provide information about the underlying structure of the species involved and their extensive use in a model limits the amount of structural information that can be recovered. The curated set lies between the two, showing that manual curation can, to some extent, be used to resolve molecular interactions.

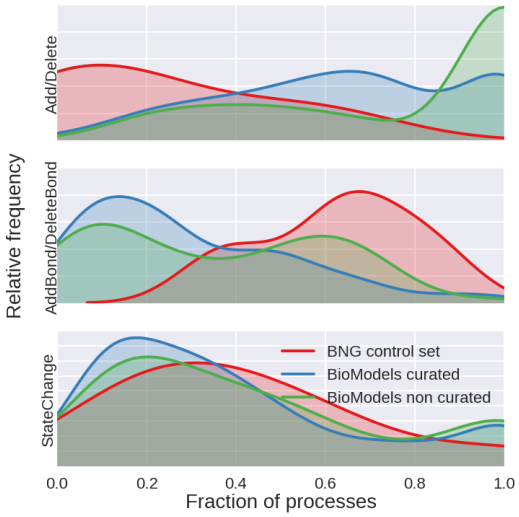

Fig. 1. Distribution in the fractional occurrence of three basic sets of graph operations over the curated (blue) and non-curated (green) sections of the BioModels database and a control set of rule-based models (red).

III. Conclusion

The results of our analysis suggest that the network approach does not limit the use of explicit mechanisms, but makes it harder to recapitulate or infer them after construction. This is because the network abstraction is not optimal for structured species, requiring manual encoding for each species and reaction. On the other hand, the rule-based framework was designed to be a more appropriate abstraction for hierarchically structured biochemical entities such as proteins and signaling complexes.

References

[1] J. R. Faeder et al. Rule-based modeling of biochemical systems with bionetgen. In *Systems Biology*, pages 113–167. Springer, 2009.

[2] J. S. Hogg et al. Exact hybrid particle/population simulation of rule-based models of biochemical systems. *PLoS computational biology*, 10(4):e1003544, 2014.

[3] C. Li et al. BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Systems Biology*, 4:92, Jun 2010.

[4] J.-J. Tapia and J. R. Faeder. The atomizer: Extracting implicit molecular structure from reaction network models. *Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine*, 2013.

[1,2,3]Department of Computational and Systems Biology, University of Pittsburgh E-mail: {jjtapia[1],jas237[2],faeder[3]}@pitt.edu

Acknowledgements: We gratefully acknowledge funding from NIH grant P41 GM103712 and NSF Expeditions in Computing Grant (award 0926181).