RuleHub:An environment for developing and sharing rule-based models

José Juan Tapia[1] and James R. Faeder[1]

***Abstract*— Rule-based modeling (RBM) is a molecule centric representation of biological systems that enables a direct translation of knowledge about protein-protein interactions, post-translational modifications, and enzyme catalysis into executable models. Many models have been developed using this paradigm, but presently there is no centralized database or standard annotation system that allows modelers to exchange model information. In this paper we present RuleHub, a repository that allows for annotation, visualization, simulation, and sharing of RBMs. Encoding of models using rules facilitates model querying and comparison.**

I. RuleHub

RuleHub provides access to a collection of both user-contributed and automatically translated quantitative models of biochemical systems encoded in the rule-based modeling paradigm (RBM [1], [2]). Although there is an existing need for models to exchange information through tools like a centralized knowledge base and a common annotation system, previous attempts to develop a rule-based modeling database (GetBonnie,Rulebase) are now defunct. RuleHub attempts to fill this void by offering similar capabilities (hosting, editing, analysis, and executing) along with other planned functionality (model alignment, comparison, merging, entity based queries). RuleHub is implemented using the Google App Engine and available at http://dragon-001.appspot.com/.

A. Features

We implemented an annotation system based on the MIRIAM [4] format that describes the structure of the model and its constituent components. RuleHub contains an interface that allows users to upload pre-annotated models, to add this information during the submission process, or to automatically retrieve it server-side if the modeler used identifiable naming conventions. An uploaded model is also associated with a simulation time series and a model contact map, which RuleHub generates automatically and is a concise graphical representation of an RBM displaying the molecules, molecular components, component states, and component binding interactions (Fig. 1). Model annotation information allows users of RuleHub to retrieve models based on any biological keyword.

The database will be populated using existing models in the BioNetGen library, the RuleBase and the GetBonNie database. Additionally we will import reaction network models in SBML format using a structural translator we have developed called the "Atomizer," [5] which recovers implicit molecular structure

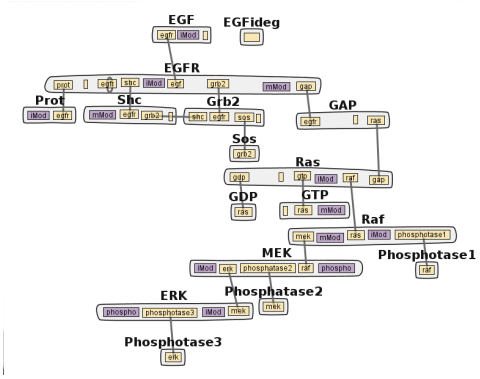

Fig. 1. Contact map for the atomized Biomodels 19

information, as shown in Fig. 1 for a classic model of EGFR signal transduction. Atomizer has been used to generate rule-based versions of about 60% of the models in the BioModels [7] database and work is in progress to generalize Atomizer to work on any model in SBML format.

B. Future Work

Beyond storing models, RuleHub is envisioned as a way for modelers to exchange information and enhance the creation of new models using existing information. As such, additional to the model-level querying capabilities currently available we will add an entity-level query, where the user is able to search for relevant rules and species based on a BioNetGen pattern. This feature can be used to extend previous models through model fusion. (merging two models through their common entities [8]). Additionally, this can be used to compare the assumptions and differences of models that study the same system through model alignment.

References

[1] Faeder, JR, et al. (2009). In Systems biology(pp. 113-167). Humana Press.

[2] Hogg, JS et al. (2013). arXiv preprint arXiv:1301.6854.

[3] Danos, V. et al. (2008). In Formal Methods in Systems Biology(pp. 103-122). Springer Berlin Heidelberg.

[4] Novere, NL et al. (2005). Nature biotechnology,23(12), 1509-1515.

[5] Tapia, JJ and Faeder, JR (2013). In Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics(p. 726). ACM.

[6] Hucka, M. et al. (2003). Bioinformatics,19(4), 524-531.

[7] Novere, NL et al. (2006). Nucleic acids research, 34(suppl 1), D689-D691.

[8] Randhawa, R, et al. (2010): Computational Biology and Bioinformatics

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

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