Composition and Aggregation for Biological Pathway Modeling

Ranjit Randhawa*, Clifford A. Shaffer*, and John J. Tyson**
Departments of Computer Science* and Biological Sciences**
Virginia Tech
Blacksburg, VA 24061, USA
rrandhawa|shaffer|tyson@vt.edu

1 Abstract

Today's pathway models are small compared to the amount of information known about a particular cellular pathway, in part because current modeling languages and tools are unable to handle significantly larger models. Thus, most pathway modeling work today focuses on building small models of individual pathways since they are easy to construct and manage. The hope is someday to put these pieces together to create a more complete picture of the underlying molecular machinery. While efforts to make large models benefit from reusing existing components, unfortunately, there currently exists little tool or representational support for combining or composing models. We have identified four distinct modeling processes related to model composition.

Fusion is the process of combining two of more models into a single "flat" model. Fusion enables modelers to incorporate information from another model into an existing one, to create larger models. We are prototyping a tool that assists modelers to easily fuse multiple models together. However, the resulting monolithic model is inherently as complex as the sum of the complexities of the submodels, and there is a limit to how large such a model can become and still be comprehensible.

Composition allows modelers combine a collection of submodels. Submodels are themselves complete models, not elements within a model such as species, parameters, etc. We have defined a small set of features to be added to the SBML language [2] to support hierarchical composition.

Aggregation is a variation on composition that allows users to provide interfaces to components so as to restrict access to the information of a submodel by exposing only certain variables. An aggregated model contains a list of input and output ports (interfaces) that link to internal species and parameters. Submodels can then be connected via their interfaces to create larger models, also with restricted interfaces. Like composition, we implement aggregation through added SBML language constructs.

Flattening converts a composed or aggregated model with its hierarchy of connections to one without any hierarchy or connections. In SBML terms, flattening a composed or aggregated model will result in a valid SBML Level 2 [1] model. Flattening converts composed or aggregated models to fused models, so they can use existing simulators that lack support for composition or aggregation.

References

[1] A Finney, M Hucka, and H Bolouri. Systems Biology Markup Language (SBML) Level 2: Structures and Facilities for Model Definitions. Available at http://sbml.org/specifications/sbml-level-2/version-1/html/sbml-level-2.html, 2003.

[2] M Hucka, A Finney, H M Sauro, H Bolouri, J C Doyle, H Kitano, A P Arkin, B J Bornstein, D Bray, A Cornish-Bowden, A A Cuellar, S Dronov, E D Gilles, M Ginkel, V Gor, I I Goryanin, W J Hedley, T C Hodgman, J-H Hofmeyr, P J Hunter, N S Juty, J L Kasberger, A Kremling, U Kummer, N Le Novere, L M Loew, D Lucio, P Mendes, E Minch, E D Mjolsness, Y Nakayama, M R Nelson, P F Nielsen, T Sakurada, J C Schaff, B E Shapiro, T S Shimizu, H D Spence, J Stelling, K Takahashi, M Tomita, J Wagner, and J Wang. The Systems Biology Markup Language (SBML): A Medium for Representation and Exchange of Biochemical Network Models. *Bioinformatics*, 19(4):524–531, 2003.