

Prolog-based XML Processing for Compact Representation and Deductive Evaluation of Biological Reaction Networks

René Haberland

Max-Planck Institute for Informatics
Saarbrücken, Germany

Thomas Hinze

Department of Bioinformatics
Friedrich Schiller University
Jena, Germany

The Systems Biology Markup Language (SBML) [1] has been established as quasi-standard for representation of biological reaction networks including formalisation of dynamical behaviour. A variety of software tools enables further processing or specific analysis of SBML data. Particularly in case of complex reaction networks composed from primarily separate local models, a certain degree of redundancy or even ambiguity can easily occur. Subsequent studies can become complicated or impossible by that. Having at hand a toolbox able to transform SBML files into compact and consistent counterparts turns out to be a reasonable first step of processing. Moreover, a deductive evaluation of SBML data in terms of a knowledge base along with user-defined queries can gain new insights into the functionality of reaction networks from a systems biological point of view. A simple illustrating example is answering the question whether a given reaction network is able to exhibit stable oscillating behaviour [2]. Here, the network description along with a dedicated mathematical toolbox acts as knowledge base. A sequence of reasoning stages then checks required properties with increasing complexity (e.g. topological presence of at least one feedback loop, steady-state conditions, eigenvalues of the corresponding Jacobian matrix).

Motivated by the idea to employ Prolog-based reasoning, we propose a general method for evaluation and transformation of SBML data that benefits from the advantages of a logical approach in comparison to pure imperative or recent functional techniques. These advantages include lazy evaluation, parallelisation and maintenance of soundness and completeness in concert with a redundancy-free and consistent representation. Additionally, the semi-structural nature of SBML features heterogeneous architectures by its capability of syntax-independent interpretation and integration. Since SBML is defined upon a XML schema by means of parameterised graphs, we focus on extension of XML-based transformation techniques towards a general framework that does not lack of abstraction from the target model [3].

Our Prolog-based XML processing framework incorporates model transformations within SBML, and also for MathML applications like symbolic integration. SBML models are currently validated against several XML validation schemas. Therefore it can also be applied for intermediation and bridging of different schemas. A current multiparadigmatic Java-Prolog implementation allows the integration of Java code in Prolog and vice versa which may improve the staging of SBML annotations. Besides the openness, Prolog allows term unification and complex data types such as lists or tuples for free due to its language definition. Functional XML transformation languages such as XSL-T or XQuery in its present form have to overcome serious feature restrictions. So, XSL-T ties user-defined functions preventing an efficient control flow. Specification of premise and conclusion within rules is restricted and new operator definitions in current working sets are forbidden.

REFERENCES

- [1] M. Hucka, A. Finney, H.M. Sauro et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19(4):524-531, 2003
- [2] T. Hinze, R. Fassler, T. Lenser, P. Dittrich. Register Machine Computations on Binary Numbers by Oscillating and Catalytic Chemical Reactions Modelled using Mass-Action Kinetics. *International Journal of Foundations of Computer Science* 20(3):411-426, 2009
- [3] R. Haberland, I.L. Bratchikov. Transformation of XML-documents with Prolog. *Advances in Methods of Information and Communication Technology*, 10:99-109, Helsinki/Petrozavodsk, 2008