

CMBSlib: A Library for Comparing Formalisms and Models of Biological Systems

Sylvain Soliman and François Fages

Projet Contraintes, INRIA Rocquencourt,
BP105, 78153 Le Chesnay Cedex, France
{Sylvain.Soliman, Francois.Fages}@inria.fr
<http://contraintes.inria.fr>

Abstract. We present CMBSlib, a library of Computational Models of Biological Systems. It is aimed at providing a list of test problems for formalisms, modeling issues and implementation issues in systems biology.

The main motivation for CMBSlib is to stimulate research on the formal modeling of biological systems, by facilitating the exchange of formal models between researchers, and by providing a forum of comparison and validation of not only models, but also modeling formalisms and implementations.

Unlike a standardization effort, CMBSlib welcomes the most exotic formalisms and models provided they attack the modeling of well documented biological systems. Models of biological systems written in any referenced formalism can be submitted to CMBSlib. No special format or standard is required.

We discuss the advantages of and problems encountered in building such a library, give an example of typical entry in the library, and most of all we invite the community to become active contributors to CMBSlib.

1 Introduction

As the first CMSB workshop proved, modeling for Systems Biology is an important new task for computer scientists, mathematicians and biologists; and to quote the Call For Papers of this second edition:

As the field matures, it is becoming increasingly obvious that there is probably no 'one-size fits all' formal language for molecular biology, but rather several meddling paradigms, each with its strengths and weaknesses relatively to specific analytical goals.

It therefore seems that the need to compare all these formal languages, and their corresponding analytical capabilities, is also increasing.

On the other hand, most of the modeling work done around systems biology is quite difficult to transfer since lots of articles mentioning new models only describe the resulting analyses (usually a simulation plot) and the biological lessons learned from it. The models themselves often lack proper publication.

We thus advocate the need for a general framework allowing to compare:

- on the one hand, different formalisms, associated tools, resulting analyses;
- on the other hand, different models of the same or related biological systems.

CMBSlib aims at becoming such a framework, and provides at the same time a global facility to store and publish models. This is in contrast to the already existing model repositories, like that of the SBML-capable tool, Cellerator [1] for instance. All model repositories are indeed oriented towards one single formalism, whether ODEs, petri-nets or process calculi [2], while CMBSlib aims at confronting various formalisms useful for Systems Biology. Furthermore, existing repositories maintain one single model of a given biological system in a given organism, there is thus no facility for comparing different models of the same biological system.

There are also attempts at unifying the current mass of languages, such as that of the BioPAX group¹ for sharing pathway information, and if such unification succeeds we would be very happy; but until then, it seems necessary to allow the use of different formalisms, description languages and tools. CMBSlib was designed for that purpose in the framework of the ARC CPBIO [3] whose more general aim is to study new languages suited to Systems Biology.

One should also remark that the creation of analogous libraries in other domains of computer science has usually resulted in a big progress in comprehension of the issues involved. These repositories are mostly benchmarking libraries, but our constraint programming origin made us aware of CSPLIB [4], a library of constraint satisfaction problems, where “representation” is also a determining factor. That existing library was an important source of inspiration for building CMBSlib.

2 Comparing Formalisms

As made clear above, there are currently many formalisms used for modeling biological systems, and it is much too early to throw away all but one.

There is however currently no existing repository trying to encompass many formalisms in order to allow the user to choose the one the most appropriate to what he wants to do with the corresponding model.

Moreover, many models were often built about one given biological system, or even similar systems, but they usually are kept separate because they are expressed using different formalisms or description languages. This separation makes it impossible to benefit in one of those models from the enrichments brought to another one; it also impeaches any meaningful comparison.

All these barriers not only impact the users of the models, since they do not know about the real reasons to choose one or another formalism, and do not benefit from cross-improvement of the existing models, they also have a negative effect on the designers of formalism who need to build test-cases and showcases

¹ <http://www.biopax.org>