The Biochemical Abstract Machine BIOCHAM

Nathalie Chabrier-Rivier, François Fages, and Sylvain Soliman

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

Abstract. In this article we present the Biochemical Abstract Machine BIOCHAM and advocate its use as a formal modeling environment for networks biology. Biocham provides a precise semantics to biomolecular interaction maps. Based on this formal semantics, the Biocham system offers automated reasoning tools for querying the temporal properties of the system under all its possible behaviors. We present the main features of Biocham, provide details on a simple example of the MAPK signaling cascade and prove some results on the equivalence of models w.r.t. their temporal properties.

1 Introduction

In networks biology, the complexity of the systems at hand (metabolic networks, extracellular and intracellular networks, networks of gene regulation) clearly shows the necessity of software tools for reasoning globally about biological systems [1]. Several formalisms have been proposed in recent years for modeling biochemical processes either qualitatively [2, 3, 4] or quantitatively [5, 6, 7, 8, 9]. State-of-the-art tools integrate a graphical user interface and a simulator, yet few formal tools are available for reasoning about these processes and proving properties about them. Our focus in Biocham has been on the design of a biochemical rule language and a query language of the model in temporal logic, that are intended to be used by biologists.

Biocham has been designed in the framework of the ARC CPBIO on "Process Calculi and Biology of Molecular Networks" [10] which aims at pushing forward a declarative and compositional approach to modeling languages in Systems Biology. Biocham is a language and a programming environment for modeling biochemical systems, making simulations, and checking temporal properties. It is composed of :

1. a rule-based language for modeling biochemical systems, allowing patterns and constraints in the definition of rules;
2. a simple simulator;
3. a powerful query language based on Computation Tree Logic CTL;
4. an interface to the NuSMV [11] model checker for automatically evaluating CTL queries.

V. Danos and V. Schachter (Eds.): CMSB 2004, LNBI 3082, pp. 172–191, 2005.

The use of Computation Tree Logic (CTL) [12] for querying the temporal properties of the system provides an alternative technique to numerical models based on differential equations, in particular when numerical data are missing. The model-checking tools associated to CTL automate reasoning on all the possible behaviors of the system modeled in a purely qualitative way. The semantics of Biocham ensures that the set of possible behaviors of the model over-approximates the set of all behaviors of the system corresponding to different kinetic parameters.

Biocham shares several similarities with the Pathway Logic system [4] implemented in Maude. Both systems rely on an algebraic syntax and are rule-based languages. One difference is the use in Biocham of CTL logic which allows us to express a wide variety of biological queries, and the use of a state-of-the-art symbolic model checker for handling the complexity of highly non-deterministic models.

The first experimental results of this approach for querying models of biochemical networks in temporal logic have been reported in [13, 14], on a qualitative model of the mammalian cell cycle control [15, 16] and in [14] on a quantitative model of gene expression [9]. In this paper we describe the Biocham system which provides a modeling environment supporting this methodology.

The next section defines the syntax of Biocham objects, rules and patterns, and their semantics. The following section describes the CTL query language and the expression of biological queries. Then we detail Biocham functionalities on a simple model of the MAPK signaling cascade. In section 5 we discuss the comparison of different models of given biological systems and show two equivalence results w.r.t. CTL properties. Section 6 reports our on-going experience in applying inductive logic programming techniques to learning reaction rules from temporal properties, and learning rule patterns from a given set of reaction rules. Finally we compare our approach with related work and conclude on the perspectives of this work.

2 Syntax and Semantics

2.1 A simple Algebra of Biochemical Compounds

Biocham manipulates formal objects which represent chemical or biochemical compounds, ranging from ions, to small molecules, macromolecules and genes. Biocham objects can be used also to represent control variables and abstract biological processes.

Syntax:

$$\begin{aligned} \text{object} &= \text{molecule} \mid \text{abstract} \\ \text{molecule} &= \text{name} \mid \text{molecule-molecule} \mid \text{molecule}{\sim}\{\text{name},\ldots,\text{name}\} \\ &\quad \mid \text{gene} \mid (\ \text{molecule}\ ) \\ \text{gene} &= \#\text{name} \\ \text{abstract} &= @\text{name} \end{aligned}$$