Using Hybrid Concurrent Constraint Programming to Model Dynamic Biological Systems

Alexander Bockmayr and Arnaud Courtois

Université Henri Poincaré, LORIA
B.P. 239, F-54506 Vandœuvre-lès-Nancy, France
{bockmayr,acourtoi}@loria.fr

Abstract. Systems biology is a new area in biology that aims at achieving a systems-level understanding of biological systems. While current genome projects provide a huge amount of data on genes or proteins, lots of research is still necessary to understand how the different parts of a biological system interact in order to perform complex biological functions. Computational models that help to analyze, explain or predict the behavior of biological systems play a crucial role in systems biology. The goal of this paper is to show that hybrid concurrent constraint programming [11] may be a promising alternative to existing modeling approaches in systems biology. Hybrid cc is a declarative compositional programming language with a well-defined semantics. It allows one to model and simulate the dynamics of hybrid systems, which exhibit both discrete and continuous change. We show that Hybrid cc can be used naturally to model a variety of biological phenomena, such as reaching thresholds, kinetics, gene interaction or biological pathways.

1 Introduction

The last decades have seen a tremendous progress in molecular biology, the most spectacular result being the announcement of a first draft of the entire human genome sequence in June 2000, with analyses published in February 2001. Current genome, transcriptome or proteome projects, whose goal is to determine completely all the genes, RNA or proteins in a given organism, produce an exponentially growing amount of data. Storing, maintaining, and accessing these data represents already a challenge to computer science. But the real work - with an enormous impact on medicine and pharmacy - consists in exploiting all these data and in understanding how the various components of a biological system (i.e. genes, RNA, proteins etc.) interact in order to perform complex biological functions.

Systems biology is a new area in biology, which aims at a system-level understanding of biological systems [16]. While traditional biology examines single genes or proteins in isolation, system biology simultaneously studies the complex interaction of many levels of biological information - genomic DNA, mRNA,

P. J. Stuckey (Ed.): ICLP 2002, LNCS 2401, pp. 85–99, 2002.

proteins, informational pathways and networks - to understand how they work together, see [14] for a recent example.

The development of computational models of biological systems plays a crucial role in systems biology [3]. A number of projects like BioSpice, Cellerator, DBSolve, E-Cell, Gepasi, Jarnac, ProMot/DIVA, StochSim or Virtual Cell aim at modeling and simulating biological processes. The Systems Biology Workbench [13] is a software platform currently being developed in order to enable the different tools to interact with each other. From a programming language perspective, a *fundamental question* arises: what is the semantics underlying these different approaches and the possible combinations between them?

The goal of this paper is to present hybrid concurrent constraint programming (`Hybrid cc`) [11] as a promising alternative to existing modeling and simulation approaches in systems biology. `Hybrid cc` is a very powerful framework for modeling, analyzing and simulating *hybrid systems*, i.e., systems that exhibit both discrete and continuous change. It is a declarative compositional programming language based on the `cc` paradigm. From a computer science perspective, a major advantage of `Hybrid cc` compared to other approaches is that it is a full programming language with a well-defined semantics, based on a small number of primitives. From the viewpoint of systems biology, these basic constructs may help to identify key computational concepts needed to represent and to understand biological systems at the molecular and cellular level.

The organization of this paper is as follows. We start in Sect. 2 by giving a short overview of modeling approaches for molecular and cell biology. We emphasize the role of hybrid systems, which can cover both discrete and continuous phenomena. Sect. 3 recalls some of the basic ideas underlying hybrid concurrent constraint programming and gives a short introduction into the system `Hcc` that we used in our experiments. Sect. 4 is the core of the paper, explaining how `Hybrid cc` can be used to model biological systems in a high-level and declarative way. We show that various phenomena in biology, like thresholds, kinetics, gene interactions, or biological pathways have their natural counterpart in `Hybrid cc`. Sect. 5 summarizes the discussion and points out directions for further research. The results presented in this paper were first announced in [2], see also [5].

2 Existing Modeling Approaches

A variety of formalisms for modeling biological systems has been proposed in the literature. A detailed discussion goes far beyond the scope of this paper, we refer to [6,3] for an overview. Following [9], we may distinguish three basic approaches

- discrete,
- continuous,
- stochastic,

and various combinations between them.