Applications of Formal Methods in Biology

Amir Pnueli

Department of Applied Mathematics and Computer Science
The Weizmann Institute of Science, Rehovot, Israel 76100
`amir@wisdom.weizmann.ac.il`

Abstract. From the first introduction of the notion of "Reactive Systems" and development of specification languages (such as Temporal Logic and Statecharts) and verification methods for this class of systems, it has been stated that this notion encompasses a wider class of systems than just programs or hardware designs, and should be applicable to other complex systems unrelated to computers. In a similar vein, the acronym UML talks about "modeling language" rather than "programming language", implying that the approach should be applicable to a more general class of systems than just computer-related.

While this claim of wider applicability has been always implied, it was never before seriously substantiated. In this talk, I will describe some recent attempts to apply the discipline of formal methods to the modeling, analysis, and prediction of biological systems. This corresponds to an emerging trend in Biology, according to which Biology in the 21st century will have to direct its attention towards understanding how component parts collaborate to create a whole system or organism. The transition from identifying building blocks (analysis) to integrating the parts into a whole (synthesis) will have to use mathematics and algorithmics. We need a language that is legible both to biologists and computers, and that is faithful to the logic of the biological system of interest.

In search for an appropriate rigorous approach to modeling biological systems, we examined formal modeling methods in computer science that were originally developed for specification, design, and analysis of reactive systems. We found that the visual formalism of statecharts can address this challenge, within the general framework of object-oriented modeling. This conclusion followed an initial study we carried out, in which we constructed a detailed executable model for T cell activation, and were able, using verification techniques to find and correct a flaw in the original model.

Following this preliminary study, we have now undertaken the more challenging project of applying and extending this methodology for constructing a detailed model of the developmental processes that lead to the formation of the egg-laying system in the nematode C. elegans. The model is built to capture in a natural yet rigorous and analyzable way the aspects of concurrency, multi scalar data, and hierarchical organization. This project involves a close collaboration with Naaman Kam, David Harel, and Irun Cohen from the department of Immunology at the Weizmann Institute.