The Computational Modeling Assistant: semi-automating the generation and implementation of dynamic computational models

Scott Christley and Gary An, Dept. of Surgery, University of Chicago

Introduction: The translational challenge in biomedical research lies in developing effective and efficient means of transferring mechanistic knowledge across multiple scales from one biological context to another, namely from bench to bedside. A central deficiency of the current scientific process is the inability of traditional experimental procedures to sufficiently explore and evaluate the range of mechanistic hypotheses suggested by vast new data sets. Addressing this challenge requires technological augmentation of the ability to evaluate hypotheses to identify plausible subsets of hypotheses and suggest future experiments to further refine those hypotheses. Unfortunately there is a crucial barrier to the scalable application of such technology: the actual process of constructing a model. Overcoming this barrier cannot be simply addressed by an improved user interface, a software toolkit, or a new modeling language; rather it is a cognitive issue central to the modeling process itself, i.e. mapping a conceptual biological model to a specific mathematical modeling method. We propose that an artificial intelligence (AI) software agent, the Computational Modeling Assistant (CMA), can fill the critical link between biological knowledge and computational models, accelerating the modeling process by treating the mapping between biology and computation as an AI planning task to identify suitable modeling methods (including potential hybrid model specifications) and then creating the selected models. The CMA will also output to existing markup languages, such as SBML/CellML, when applicable. The CMA will provide these capabilities to the researcher without the requirement of mathematical expertise, as hypotheses are expressed using natural language. This will, facilitate their entry into computational modeling, allowing them to evaluate their hypotheses through the design of new experiments and dramatically increasing the scalability of computational biology in general. We demonstrate a prototype of the CMA using two examples: gut mucous dynamics and *Pseudomonas aeruginosa* virulence activation.

Methods: The CMA uses a logical rewrite system, Maude (http://maude.cs.uiuc.edu/), to construct an implementation plan consisting of: interpreting a biological hypothesis/conceptual model presented to it by a researcher in a constrained natural language syntax, utilizing knowledge concepts from bio-ontologies, selecting appropriate Modeling and Simulation (M&S) methods based on the researcher's requirements (including hybrid model structures), identifying input data requirements, capabilities and limitations of selected methods, and generating simulation code. The researcher interfaces with the CMA via their input of biological knowledge, selection among presented simulation methods based on desired level of resolution and available data, and interpretation of the simulation output.

Results: The first example examines the dynamics of gut mucous stratification concerning the generation of stratified gut mucous layers. The components of this hypothesis include goblet cells, putative genes regulating secreted compounds forming the mucus and the compounds themselves. The hypothesis is expressed in the constrained natural language syntax with attached bio-ontological meta-data, and an ordinary differential/partial differential equation reaction-diffusion method selected as the mathematical model specification. The second example examines virulence activation in *Pseudomonas aeruginosa* by host stress. Host tissue factors released in response to physiologic stresses are sensed by *P. aeruginosa* with the hypothesis that virulence pathways are activated at an earlier threshold when those host tissue factors are present. The CMA provides a Petri net model and an ODE model of this conceptual biological model.

Conclusion: We present an intelligent computational agent that can perform a composition planning operation utilizing biomedical and M&S ontologies to semi-automate computational model construction. Future development on the CMA's planning capability can improve the modularity, robustness and scale-ability of knowledge integration by maintaining interoperability with established and ongoing development in formal semantics/knowledge representation and M&S methods. We believe that the process automation advances offered by the CMA will lead towards the development of cyber-environments providing scale-able high-throughput hypothesis evaluation that will fundamentally advance the process of biomedical research.

Support: NSF 0830-370-V601