A Novel Strategy to Accelerate the Modeling and Analysis of Complex Biological Systems

Michael A. Savageau

Department of Microbiology & Molecular Genetics and Department of Biomedical Engineering, University of California, Davis USA

Although we have a well-defined concept of an organism's genotype, its phenotypes – the biological functions implemented by its underlying biochemistry – are difficult to define and predict. We ultimately want to represent the phenotype by a mechanistic model that accurately describes the changes in the concentrations of the compounds under various conditions. The 'architecture' of a system can be inferred from high-throughput data, but numerous unknown kinetic parameters influence exactly how change in one concentration affects others. Often the phenotype becomes clear only when those parameters are known; even a simple model can exhibit many phenotypes given different parameter choices. Current approaches to determining phenotype thus focus *first* on finding parameter values for the underlying biochemistry, typically through a mixture of ad-hoc experimentation and computationally inefficient high-dimensional numerical search. While these strategies have been used to fully characterize small systems in the pre-genomic era, a mechanistic understanding of systems, even of moderate size, derived from genotype data remains elusive. We propose a fundamental shift towards a post-genomic computational paradigm in which we first *analytically* determine the space of possible phenotypes for a given network architecture and then predict parameter values for their realization, predictions that can guide experimentation and further numerical analysis. This *'phenotype-centric'* paradigm combines four innovations with the potential to accelerate our understanding of complex biological systems: (1) a rigorous mathematical definition of biochemical phenotypes, (2) a method for enumerating the phenotypic repertoire based on the biomolecular network architecture, (3) an integrated suite of computational algorithms for the efficient prediction of parameter values and analysis of the phenotypic repertoire, and (4) a user-focused environment for navigating the resulting space of phenotypes and identifying biologically relevant features. These innovations will facilitate deterministic and stochastic simulations that require parameter values, will accelerate both hypothesis discrimination in systems biology and the design cycle in synthetic biology, and will enable investigators to achieve predictive understanding of biomolecular phenotypes from genotype.

In the lecture, I will describe the underlying theory, review recent progress toward the realization of its potential, and outline some of the remaining challenges.

In the breakout session, I will work through a few specific applications, highlight areas of theory in need computer implementation, and explore connections to other modeling approaches and issues of scalability.