The Problem: Biological Development

How does an organism's genetic makeup interact with its environment to shape the intricate developmental processes that lead to functional tissues, organs and organisms from undifferentiated cells?

The Problem: Biological Development

This question has long challenged biologists. Researchers have traditionally used microscopy, mutants and other methods to understand the molecular and cellular bases of development. The amount of data acquired has rapidly increased with recent advances in instrumentation and genomics.

We have therefore embarked upon a multidisciplinary computational based approach to integrate these data.

Software Architecture

The architecture, aimed at producing-scale model inference, is illustrated in the figure to the right. We generate simulation code from high-level models specified in biological and/or mathematical language. Other computational tools are used to analyze expression imagery and other data sources, and the simulator combined with nonlinear optimization is used to fit the models to the data. Key elements include:

• a mathematical framework combining transcriptional regulation, signal transduction, and dynamical mechanical models,
• a model generation package (Cellerator) based on a computer algebra representation, including subcellular and tissue-level representations,
• extensions to SBML (Systems Biology Markup Language), an exchangeable model representation format, to include dynamic objects and relationships, and
• a C++ code generator to translate SBML into highly efficient simulation modules; a simulation engine including standard numerical solvers and plot capabilities; and
• a nonlinear optimizer, and
• ad hoc image processing and data mining tools.

Images are acquired with a Zeiss LSM 510 meta upright laser scanning confocal microscope.

Mathematical Framework

Let $b$ be a vector of concentrations in a particular cell, and let $v_i$ be the corresponding vector in a neighboring cell. Then the following genetic regulatory network (GRN) is fit to the data:

$$v_i = f(v_i, b)$$

where

$$f(v, b) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} v_j + \sum_{i=1}^{n} d_i b_i$$

Here $T$, $P$, and $Q$ are connection matrices, describing respectively the effect of $T$ species on species $i$, $P$ species in neighboring cell, and $Q$ species, $P$ the effect of receptor $j$ activation on species $i$, and $Q$ the effect of ligand $k$, produced in neighboring cell, on receptor $j$ activation. $g_i, v_i, b_i$ and $d_i$ are differential equations produced by cellerator from known or postulated specific biochemical reaction networks. Spatial dynamics are modeled by a breakable nonlinear “spring” force between neighboring cells that is set to zero between non-interacting cells.

As an example, consider the naive four protein GRN: $v_i$ is illustrated below to the left, in which each region and the stem are identified by differential expression of region-specific proteins. Despite its simplicity, this model produces remarkably realistic-looking simulations. The resulting differentiation pattern at three different time points is shown below on the right.

Code Autogeneration

Cellerator is used to design signal transduction networks (STNs) based on traditional kinetic interactions (e.g., Mass Action, Michaelis-Menten, etc) as well as the GRN framework. Input to Cellerator is via a set of arrow-based reactions representing the network; cellerator then uses R programming language to translate the STN into a system of differential equations. Cellerator then uses MathSBML to write the model as extended SBML. A C++ program parses and generates derivative functions for each chemical species in the model. The def functions are then linked with the simulation engine (numerical solver) and run to generate time course predictions.

Educational Outreach

We are currently developing a new set of techniques for high school, pre-service science teachers, and undergraduate students through our partnership with the Huntington Botanical Gardens in San Marino, CA. Outreach activities will culminate in a summer institute in which 30 high school students will develop a public kiosk to display a computable plant model for exhibit at the Huntington, which hosts 500,000 visitors per year. This program holds remarkable promise for linking cutting-edge knowledge and technologies with K-12 teachers’ and students’ understanding of plant development and integrative biology.

References


