Exploring the parameter landscape in biological pathways

Namita Gupta*, Bobby Karp, Jeremy Gunawardena
Department of Systems Biology, Harvard Medical School

FAS Center for Systems Biology

Introduction

When dealing with complex biological pathways, it is useful to have a method for analyzing such systems quantitatively. One method for emulating the dynamics of a system is to use mass-action kinetics. This is most often used in biochemistry and is based on the fact that the rate at which a product is formed is proportional to the concentrations of the reactants, as seen below. Given a set of reaction with rate constants k_1, k_{-1}, and k_2:

$$A + B \rightleftarrows C \longrightarrow D$$

the corresponding rate equations are the following:

In general, given species $x_1,\ldots,x_n$ in a pathway, the rate equations are polynomials with nonlinear terms consisting of species concentrations. The rate equations are a system of differential equations; setting these polynomials to zero represents a steady-state, where no concentrations are changing anymore:

We are only interested in the biologically relevant steady-states that are *stable*, meaning if near the steady-state the system will move towards it as time progresses. The concentrations at a steady-state equilibrium depend on parameters such as rate constants and initial concentrations. For a given set of rate constants, it is possible that based on initial concentrations, the system could end up at different steady-states. This is referred to as *multi-stability.* If two cells seem like they should be identical, multi-stability can account for differences between them. Looking at how the pattern of steady-states changes as parameters vary is referred to as *exploring the parameter landscape*.

Moving through parameter space and describing steady-states is very relevant in biology because unlike in other sciences, such as physics, biological systems *do* have variable parameters. There is variation between individuals in a population, between different cell types, and even between cells of the same type. It is precisely the effect of these parameter variations that we are attempting to characterize, possibly leading to the answer of questions like how diseases (such as cancer) that stem from mutated cells arise.

Acknowledgments

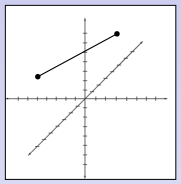
I would like to thank Dr. Bobby Karp for teaching me so much and giving me interesting and relevant projects to work on. I am equally grateful to Dr. Jeremy Gunawardena for welcoming me into his lab and giving me his full support. Finally, I thank the Harvard Systems Biology department for providing me with such a wonderful opportunity to learn about this area of research.

Methods

We developed a set of programs that take a textual description of a network, create a flow chart depicting the system, and generate the corresponding rate equations and conservation laws. Since these polynomials are nonlinear, it is hard to find all possible roots. To solve these equations, we use techniques from Algebraic geometry; in particular, *homotopy continuation*, by first solving the simpler system below.

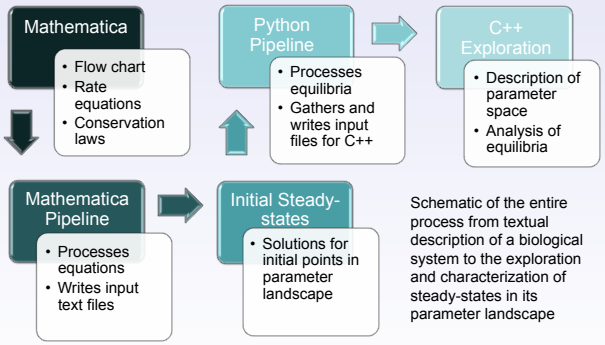
Then, we consider the equation

and, using multi-dimensional Newton's method, approximate its solution as t slowly moves from 0 to 1, ending at



Starting at $\mathbf{F}_0$, we increase t from 0 to 1, moving along the path to $\mathbf{F}_1$ in n-dimensional space (n is the number of species).

Once the solution for one system is found with a given set of parameters, we can keep using homotopy continuation from our initial steady-state as we slowly change parameters from the initial set $\mathbf{P}_0$ to the new set $\mathbf{P}_1$. This method of efficiently solving systems of polynomial equations is implemented using a highly optimized C++ program we developed, which requires specially formatted text files as input, describing the biological network in question. The Mathematica code pipelines by processing the output of the initial Mathematica program to create input for the initial use of homotopy continuation with a random set of parameter values. The Python code pipelines by processing the initial points on the parameter landscape to create input for the exploration by the C++ program.



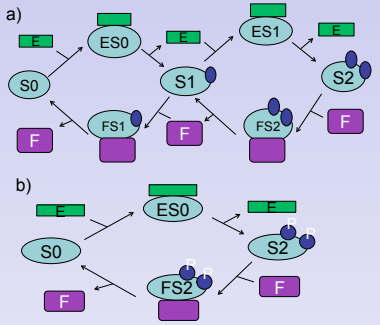
Schematic of the entire process from textual description of a biological system to the exploration and characterization of steady-states in its parameter landscape

Biological Applications

N-site phosphorylation:

A cell receives communication from an external stimulus through a process called *signal transduction*, or the relay of a signal into another type of signal. One of the main ways to transmit a signal is *phosphorylation* of proteins by a kinase or de-phosphorylation by a phosphatase. Oftentimes proteins have multiple sites on which they can be (de-)phosphorylated. In addition, a kinase/phosphatase could add/remove either one phosphate at a time (distributive) or many at once (processive). All of these factors contribute to the overall complexity of an *n-site phosphorylation system*.

It is known that in parts of the landscape there is multi-stability. Our exploration will reveal the size of the multi-stable portion of the landscape and from that, we will learn about how much information can be stored by the n-site phosphorylation system, a notion of central importance to the idea of histone code.

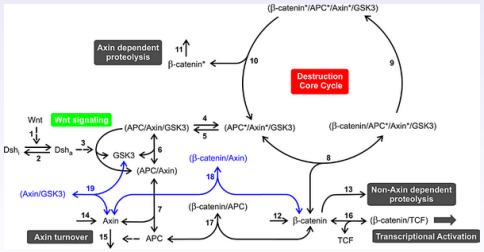


Kinases and phosphatases can be either a) distributive, where phosphate groups are 'distributed' one at a time or b) processive, where the substrate is 'processed' with multiple phosphate groups in one reaction.

Wnt Pathway:

A very important example of a signal transduction pathway is the *Wnt pathway*. Wnt, a small secreted molecule, acts as a signal that eventually leads to the up-regulation of the transcription factor β-catenin. Using the data that are available regarding this pathway, we have created a model using our Mathematica program. From this model, we will look into what negative effects mutations in the pathway can have and whether the effect of multiple mutations is synergistic.

Example of a model of the Wnt pathway



Lee E, et al. PLoS Biol. 2003 Oct;1(1):E10.