Cells need to be able to sustain themselves, divide, and adapt to new stimuli. Proteins are key agents in regulating these processes. Cell behavior is regulated by signaling pathways and proteins called transcription factors which regulate what and how much of a protein should be manufactured. Anytime a new stimulus arises, it can activate multiple signaling pathways by interacting with proteins on the cell surface (if it is an external stimulus) or proteins within the cell (if it is a virus for example).

Disruption in signaling pathways can lead to a myriad of diseases. Knowledge of which signaling pathways play a role in which condition, is thus key to comprehending how cells develop, react to environmental stimulus, and are able to carry out their normal functions. Another biological process that can play a role in such reaction is epigenetics — modification of the DNA structure that do not involve changing the sequence itself. Epigenetics has been shown to regulate transcription, however, how epigenetic changes are regulated and the exact impact these changes have on transcriptional regulation are still open questions.

In this thesis we present a suite of computational techniques that are focused on modeling the dynamic regulation of biological processes. These methods address the various aspects of the problem mentioned above focusing on the reconstruction of dynamic signaling and regulatory networks. In many cases, the amount of biological data available for a specific condition can be very small compared to the number of variables. We present an algorithm which uses multi-task learning to learn signaling networks from many related conditions. There are also very few tools that attempt to take temporal dynamics into account when inferring signaling networks. The thesis presents a new algorithm that utilizes and extends Integer Programming methods for inferring such dynamic regulation. Finally, we present a new strategy to integrate epigenetic data with other temporal datasets using deep neural networks. We use this new method to reconstruct dynamic disease progression networks in Idiopathic Pulmonary Fibrosis (IPF).

Thesis Committee:
Ziv Bar-Joseph, Chair
Eric Xing
Jaime Carbonell
Naftali Kaminski, Yale University