May 31st, 6:00 PM - 6:30 PM

Inference and Control in regulatory genomics

Siddharth Sharma
University of Maryland College Park, ssharma6@umd.edu

Over the past two decades, the significance of data for biology in general and genetics in particular has seen a remarkable increase due to advances in computing. The primary goal of any study involving one of these big data applications is to unlock the inapparent complexity. The relevance of any data set is therefore closely tied to the ability of the analytical tools used to extract meaningful information from it. One such important piece of information is the set of interactions in a multivariate system. This is a central problem of network biology with large datasets representing network configurations(microarrays, neural spikes etc.) and almost no information of the governing dynamics. In the present work, we solve the statistical mechanics analog of the network inference problem i.e. the inverse Ising problem. Recent solution methods have focussed on solving the inverse problem in its thermal equilibrium version. This makes use of the Boltzmann distribution for parameter learning. Biological networks are open systems with active exchange of matter and energy. This means that the recorded configurations don't correspond to thermodynamic equilibrium but to a non-equilibrium steady state(NESS). We use a recently proposed likelihood based method to characterize the NESS. Our calculations successfully reconstruct the networks with asymmetric interactions in addition to the symmetric version corresponding to detailed balance. This paves the way towards designing a nonequilibrium control problem. Guided by recent works for controlling a symmetric network through geometry based methods, we discuss the necessary variations needed to extend the present methods to asymmetric networks, getting us further closer to in-vivo biological control.