MAR05-2005-020169

Abstract for an Invited Paper
for the MAR05 Meeting of
the American Physical Society

A Predictive Approach to Network Reverse-Engineering

CHRIS WIGGINS, Columbia University

A central challenge of systems biology is the "reverse engineering" of transcriptional networks: inferring which genes exert regulatory control over which other genes. Attempting such inference at the genomic scale has only recently become feasible, via data-intensive biological innovations such as DNA microrrays ("DNA chips") and the sequencing of whole genomes. In this talk we present a predictive approach to network reverse-engineering, in which we integrate DNA chip data and sequence data to build a model of the transcriptional network of the yeast S. cerevisiae capable of predicting the response of genes in unseen experiments. The technique can also be used to extract "motifs," sequence elements which act as binding sites for regulatory proteins. We validate by a number of approaches and present comparison of theoretical prediction vs. experimental data, along with biological interpretations of the resulting model. En route, we will illustrate some basic notions in statistical learning theory (fitting vs. over-fitting; cross- validation; assessing statistical significance), highlighting ways in which physicists can make a unique contribution in data- driven approaches to reverse engineering.