

STANFORD UNIVERSITY
DEPARTMENT OF STATISTICS
DEPARTMENTAL SEMINAR

4:15 p.m., Tuesday, May 23, 2006
Sequoia Hall Room 200
(Cookies at 3:45 in 1st Floor Lounge)

Elizabeth Purdom
Stanford University

Data Analysis using Graphical Structures

In biological experiments, researchers often have information in the form of a graph that supplements observed numerical data. Common types of graphs encountered might be a graph depicting genomic regulatory networks or a phylogenetic tree depicting the interspecies relationships among the species of interest. Incorporating the information contained in these graphs into an analysis of the numerical data is an important and non trivial task. A promising method recently introduced for these kinds of data-graph problems rely on kernel methods common in the data mining literature to correlate graphical and numerical data.

In this talk, we will show how to explicitly create a regression framework for these data-graph problems using some kernel methods ideas. In particular, we will discuss the example of using these techniques for data connected to phylogenetic trees. We will discuss how in this case the kernel methods relate to other methods found in phylogenetics and ecology. We demonstrate these results on a genomic analysis of microbacterial communities found within the human intestinal track (Eckburg et. al, Science, 2005).