

STANFORD UNIVERSITY
DEPARTMENT OF STATISTICS
DEPARTMENTAL SEMINAR

4:15 p.m., Tuesday, April 23, 2002
Sequoia Hall Room 200
(Cookies at 3:45 in 1st Floor Lounge)

Marc Coram
Department of Statistics
Stanford University

Nonparametric Bayesian Classification

A Bayesian approach to the classification problem is proposed in which random partitions play a central role. It is argued that the partitioning approach has the capacity to take advantage of a variety of large-scale spatial structures, if they are present in the unknown regression function f_0 . An idealized one-dimensional problem is considered in detail. The proposed nonparametric prior is found to provide a consistent estimate of the regression function in the L^p topology, for any $1 \leq p < \infty$, and for arbitrary measurable $f_0 : [0, 1] \rightarrow [0, 1]$. An MCMC implementation is outlined and simulation experiments are conducted to show that the proposed estimate compares favorably with CART and bagged CART estimates. A generalized prior is discussed which employs a random Voronoi partition of the covariate-space. The resulting estimate displays promise on a two-dimensional problem, and extends with a minimum of additional computational effort to arbitrary metric spaces.