STANFORD UNIVERSITY

DEPARTMENT OF STATISTICS

BERKELEY-STANFORD JOINT COLLOQUIUM

In honor of Jerry Friedman's coming retirement

From 3:15 p.m., Tuesday, April 25, 2006
Packard 101 with a reception in Sequoia Hall

Schedule:

3:15-4:15	Lecture by Andreas Buja
4:15-4:30	Break
4:30-5:30	Lecture by Michael Jordan
5:45-7:30	Reception with food and speeches

Andreas Buja
Statistics Department; The Wharton School University of Pennsylvania

Five Ideas of Jerry Friedman

For the last thirty years, Jerry Friedman has been one of the most creative minds in our profession. He has contributed an astounding number of ideas with vast impact. A shared feature of many of these ideas is their algorithmic character combined with unusual conceptual insight.

In this talk I will discuss some of Jerry's older work: PRIM-9, CART, ACE, MARS, and EPP, and maybe one recent piece of work: the interpretation of boosting as additive modeling. Although much of Jerry's work is seen as the pursuit of predictive performance, many of his methods have another angle: they are also tools for interpreting data. This is an aspect that may have been lost somewhat under the impact of the recent encounter with machine learning. It may be time to restore a balance between predictive performance and the more qualitative aspects of data analysis, and Jerry's work is most suitable for making this point.

Michael Jordan
UC Berkeley

A Berkeley-Stanford Statistical Smorgasbord

In this talk I will present recent work on a number of statistical themes that have their origins at Berkeley or Stanford. The first line of work that I will discuss goes back to Blackwell, who uncovered links between the risk based on 0-1 loss and the class of f-divergences. We extend this programme by linking f-divergences to the broader class of so-called surrogate loss functions—computationally-inspired upper bounds on 0-1 loss that have become central in the machine learning literature on classification. I will present applications to distributed detection and sufficient dimension reduction. In a second line of work, I discuss some of the issues that arise when Stein's and Efron's shrinkage and empirical Bayes ideas are deployed in the nonparametric Bayesian setting of Dirichlet process priors, a setting whose foundations were laid by Freedman, Doksum, Ferguson and others. The Chinese restaurant process studied by Berkeley and Stanford probabilists (e.g., Aldous, Diaconis, Dubins and Pitman) turns out to be particularly helpful here. I discuss applications in bioinformatics to haplotype inference across subpopulations and to admixture problems. The overall theme is computationally-scalable statistical algorithms that target significant real-world data analysis problems—and here I acknowledge a general intellectual debt to another great Bay Area statistician: Jerry Friedman.

Joint work with Peter Bartlett, Jon McAuliffe, XuanLong Nguyen, Martin Wainwright, Yee Whye Teh, David Blei, and Eric Xing