Geometry Seminar
Tuesday, October 18, 2011
Room 512 WWH at 6:00 P.M.

Dynamic Algorithms for Half-Space Depth

Michael Burr
Fordham University

Data depth is a statistical analysis method that assigns a depth value to each point x in R^d measuring how central the point x is with respect to a probability distribution or to a finite data set. I will focus on the case of a finite data set. Data depth measures have been designed to require no prior assumptions on the underlying probability distribution and to be robust even in the presence of outliers. These data depth measures for data sets are, by definition, geometric and have recently attracted attention from the computational geometry community.

Data depth contours are polyhedral regions which enclose regions of increasing depth and provide a tool to visualize, quantify, and compare data sets. There are two competing definitions for depth contours, termed rank-based and cover-based. The rank-based contours focus on the data points while the cover-based contours focus on the ambient space. These definitions provide distinct collections of contours (and each have advantages and disadvantages).

In this talk, I will begin with an overview of several standard geometric data depth measures. I will then focus on halfspace depth (also called Tukey depth or location depth). This data depth measure is particularly nice to work with because it possesses many desirable properties for data depth functions. I will present algorithms for dynamically maintaining (under insertions or deletions of points from the data set) the half-space depth of points and contours (for either definition). I will also discuss a structural result, of independent interest, describing how little the cover-based contours can change after an insertion or deletion.

For more information please visit the seminar website at:
`http://www.math.nyu.edu/seminars/geometry_seminar.html`.