MIT
Department of Mathematics
& The Theory of
Computation Group
At CSAIL

Bioinformatics Seminar

Speaker: Charles (Chip) E. Lawrence, Director- Center for
Computational Molecular Biology, and Professor of Applied
Mathematics, Brown University
Title: Abuse of the Mode and an Ensemble Alternative
Date: Monday, 7 November 2005
Time & Location:
Refreshments: 11 am in the Theory of Computation Lab at MIT's
Building 32, Stata Center Room G-575
Talk: 11:30 am the Theory of Computation Lab at MIT's Building 32,
Stata Center, Room G-575
URL: http://www-math.mit.edu/compbiosem/

Abstract:

Advances in data collection technologies have rendered increasingly large data sets available for
analysis. While the emergence of such large data sets would seem to lead to increasingly more
precise estimates of parameters, paradoxically just the opposite seems to becoming increasingly
common. This paradoxical circumstance has emerged because these technologies have
simultaneously opened opportunities to draw inferences on previously unanswerable high
dimensional questions. For decades optimization has been employed as the major tool of most
inference procedures. It has been clearly recognized for some time now that the favorable
properties of optimization based inferences rest on an asymptotic foundation that requires the
data to grow in comparison with the number of unknowns. Nevertheless, optimization very often
continues as the method of choice even when these supporting conditions are not present.
Genomics and computational molecular biology are among the more predominate fields
experiencing the duality of the growth in data resources and inference expectations. In fact,
prediction and inference of high dimensional objects are now arguably the most important
activities in these allied new biological fields, and the inspiration for this paper. RNA secondary
structure prediction offers a very special lens to examine the untoward consequences of the
reliance on the mode in high dimensional inferences because polynomial time algorithms are
available to comprehensively characterize the space of solutions, and a references set of
structures is available for the comparison of alternative prediction methods. Through this lens we
will examine these untoward effects, consider their boarder implications, and present an
alternative "ensemble based" inference procedure.

The seminar is co-hosted by Professor Peter Clote of Boston College's Biology and Computer Science
Departments and MIT Professor of Applied Math Bonnie Berger. Professor Berger is also affiliated with CSAIL &
HST.

Massachusetts Institute
of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

For General Questions, please contact kvdickey@mit.edu