Bioinformatics Seminar

Speaker: Mona Singh, Dept. of Computer Science and Lewis-Sigler Institute of Integrative Genomics Princeton University
Title: Computational Methods Towards Predicting Aspects of Protein Structure and Interactions
Date: Monday, 8 November 2004
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:

A central challenge in bioinformatics is the development of genomic scale methods for predicting protein structure, function and interactions. The difficulty of the general protein structure prediction problem suggests that one promising approach is to identify important subproblems that lend themselves to effective solutions. I will present two simplifications that my group has been pursuing in predicting protein interactions and protein structure.

First, I will discuss methods my group has developed for predicting protein-protein interactions mediated by the coiled-coil motif, an important motif that is found in proteins that participate in transcription, oncogenesis, cell structure, and cell-cell and viral-cell fusion events. We have introduced an optimization framework for predicting these types of protein interactions that uses both genomic sequence data and experimental data. In testing on coiled-coil interactions among nearly all human and yeast bZIP transcription factors, we show that our method is able to make large-scale, high-confidence predictions.

Second, I will discuss recent hardness results and computational methods for the side-chain positioning problem, a central component of both protein structure prediction and protein design. I will present an integer linear programming formulation of the problem, and then show empirically that, surprisingly, in many interesting cases the linear program relaxation finds optimal solutions to the integer program. Our analysis demonstrates that LP-based approaches are highly effective in finding optimal (and near-optimal) solutions for the side-chain positioning problem.