Joint Prediction of RNA Secondary Structure

Gaurav Sharma  
Department of Electrical and Computer Engineering  
University of Rochester  
Rochester, New York

Abstract: With the discovery of noncoding RNAs that serve a direct functional role in cellular biology, the problem of determining the structure of these molecules has emerged as one of the fundamental problems in molecular biology. Due to the difficulty and cost associated with experimental procedures for determining structure, computational methods for predicting structure are of significant research interest. In this talk, we focus on computational techniques for RNA secondary structure prediction – the first step in the hierarchy of RNA structure estimation that predicts the folding configuration of a linear RNA chain. This folding of an RNA molecule is determined by the pairings of nucleotide bases through the formation of hydrogen bonds between them. Specifically, we consider methods for joint alignment and secondary structure prediction for two homologous RNA sequences that combine thermodynamic models of structure stability with common secondary structure constraints. The comparative analysis implicit in this process provides significant improvements in accuracy over single sequence prediction methods. We highlight our recent work in this area that: a) develops principled heuristics for computational simplification that are inspired by methods in error correction coding and offer significant computational speed-ups without a performance penalty, b) generalize the class of allowable correspondences between the structures for the two sequences to better represent observed homologous sequences, c) provide “local” estimates of confidence in predicted structure by assessing posterior base pairing probabilities, and d) utilize a joint stochastic sampling methodology to identify the distinct structural configurations among common secondary structures as alternate hypotheses. Finally, we outline our continuing research in this direction which utilizes a maximum a posteriori probability formulation for the estimation of RNA secondary structure with the goal of developing turbo-decoding style iterative estimation algorithms with high accuracy and low computational cost.

Wednesday, February 4, 2009  
10:00 a.m.  
Bryan Hall, room 305  
Host: Arye Nehorai

Short Bio: Gaurav Sharma received the PhD degree in electrical and computer engineering and an MS in Applied Mathematics from North Carolina State University, Raleigh in 1996 and 1995, respectively. Prior to that, he received the ME degree in Electrical Communication Engineering from Indian Institute of Science, Bangalore in 1992 and a BE in Electronics and Communication from Indian Institute of Technology, Roorkee in 1990. From Aug. 1996 to Aug. 2003 he was employed at Xerox Corporation's Research and Technology Division in Webster, NY, first as a member of the research and staff and then as a principal scientist and leader of a research project on color imaging. Since Aug. 2003, he is an Associate Professor in the Electrical and Computer Engineering Dept at the University of Rochester. His research interests include genomic signal processing, color science and imaging, and print/multimedia security.