Zachary Nichols

Major: Math and Biology
University: University of South Carolina
Faculty Mentor: Dr. Dong Xu
Mentor Department: Computer Science
Funded by: NSF-REU Program in Biosystems Modeling and Analysis

Protein secondary structure prediction: Creating a meta-tool

Zachary Nichols, Rajkumar Bondugula and Dong Xu

Protein structure prediction is a growing field of interest for a many varied reasons, owing not only to its obvious utility, but also the success that applying newer mathematical tools has garnered in recent years. Despite the intractability of determining optimal protein structure directly by finding a lowest-energy conformation among a huge amount of candidates, many heuristic methods have emerged that sacrifice some degree of accuracy for reasonable speed of execution. Through the use of numerical techniques such as neural networks(1), neural networks bolstered by position-specific scoring matrices generated by psi-blast(2), and k-nearest neighbor algorithms(3), the success rate of protein structure prediction has been increasing over the past decade and a half. Each of these tools has particular strengths and weaknesses. To address this and to improve prediction accuracy, we are constructing a three-part meta-tool that combines k-nearest neighbor methods, neural network methods, and hidden markov models to predict the secondary structure of proteins based on their position-specific scoring matrices. The results from each of the individual tools will be integrated and filtered to form a final prediction. This tool will be available on the web through a simple interface for those wishing to evaluate or utilize it. References: 1: Rost and Sander. Predictions of protein secondary structure at better than 70% Accuracy; J. Mol. Biol. (1993) 232, 584-599 2: Jones. Protein secondary structure prediction based on position-specific scoring matrices; J. Mol. Boil. (1999) 292, 195-202 3: Bondugula, Duzlevski, Xu. Profiles and fuzzy k-nearest neighbor algorithm for protein secondary structure prediction; (unpublished).