1020-92-89   **Michael Erdmann*** (`me@cs.cmu.edu`), Computer Science Department, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213. *Protein Similarity from Line Weavings.*

Proteins provide a rich domain in which to test theories of shape similarity. Sometimes the detection of common local structure is sufficient to infer global alignment of two proteins; at other times it provides false information. Proteins with very low sequence identity may share large substructures, or perhaps just a central core. There are even examples of proteins with nearly identical primary sequence in which alpha-helices have become beta-sheets.

In this talk I will discuss a method for comparing protein shapes based on line weavings. The algorithm approximates protein secondary structures using straight lines, then compares the resulting line weavings. A key step is to filter incomparable proteins by observing inconsistencies in the crossing matrices. We have used this method to extract a basis for the four primary classes of known protein structures, and then have compared the elements of this basis pairwise. The resulting clusters capture many well-known protein folds.

A paper motivating this talk may be found at:

http://reports-archive.adm.cs.cmu.edu/anon/2004/abstracts/04-138.html

http://www.liebertonline.com/doi/abs/10.1089/cmb.2005.12.609

For comparison results see:

http://www.cs.cmu.edu/ me/A40/explorations/

(Received August 17, 2006)