

ECS 129

Assignment: Option 3 (Programming)

Protein Structure Geometry: Knots

Due: Friday, March 13th, 2020

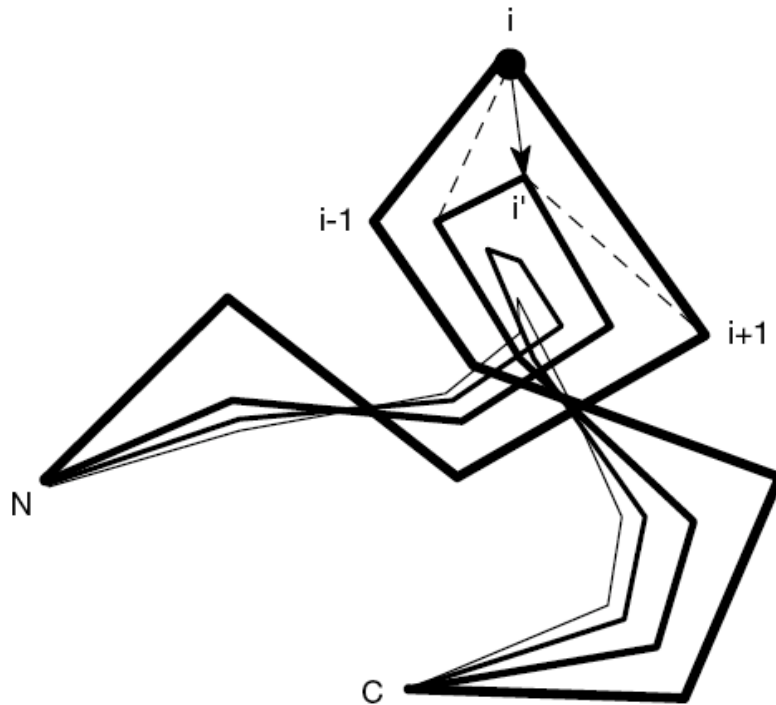
Detecting knots in proteins

Think of the main-chain of a protein as being a long entangled string. The question is: does that string form a knot? (i.e. if you were to hold the two extremities of the string and pull, would it result in the formation of a knot, or would the string become linear?) Some proteins do form knots, and it remains unclear at this time if these knots play a role in defining the functions of these proteins.

The purpose of this assignment is to write a program (your choice of language) that detects knots in a protein chain. You will follow the method proposed in the paper by William Taylor “A deeply knotted structure and how it might fold” (available as supplementary information on the class web site), and test it on the cases described in the paper.

Briefly, the method works as follows (see illustration below):

- The protein chain is defined as a polyline joining the central CA atoms of each residue running from the N-terminal to the C-terminal. Coordinates of the CA atoms are available in the PDB file for the protein of interest
- The algorithm attempts to smooth the line as long as it can: if it reaches a fully flat line, the protein does not contain a knot; otherwise, it is considered knotted.
- Beginning at the second residue, for each residue point (i) in the starting conformation, the average coordinate of i, i - 1 and i + 1 is taken as the new position (i') for the residue. This procedure is then repeated, and the results of this are progressively smoother chains
- Residue point (i) is updated to position (i') only if the triangles {i'-1,i, i'} and {i, i', i+1} are not crossed by any line segments defining the protein



This program relies heavily on a function that checks if a triangle (ABC) is crossed by a line segment [EF] or not. Note that there are several cases to consider:

- EF is parallel to the plane ABC: the line segment does not cross the triangle
- EF is not parallel to the plane:
 - Both E and F are on the same side of the plane ABC: the line segment does not cross the triangle
 - The segment EF crosses the plane containing ABC at a point I:
 - I is outside the triangle: EF does not cross the triangle
 - I is inside the triangle: EF crosses the triangle

Problem:

For this option, I am asking you to:

- Describe the method you use to implement the triangle-segment checking function
- Implement this method, and test it
- Write the full program for detecting knot
- Test it (on one of the proteins mentioned in Taylor's paper, provided on the web page)

This option is longer than the others, and I will definitely take this into account.

Please provide both the source code of the program you wrote, and a report describing

the results. There is no need to send a lengthy write-up, but it should definitely include an introduction, results and analysis, a conclusion, and references to published work, if needed.

Good Luck !