

Name: Peiying Ruan
Title: Prediction of protein complex structures using neural networks
Institute: Bioinformatics Center (Akutsu Laboratory), Kyoto University Institute for Chemical Research; Kyoto, Uji, Gokasho 611-0011, Japan
Partner institute: Macromolecular Modelling Group (Knapp Laboratory), Free University of Berlin; Fabeckstrassen 36a, D-14195, Germany, Berlin-Dahlem
Duration : April 8, 2011~July 5, 2011
<p>Plan :</p> <p><i>Background</i></p> <p>In previous research, I along with other collaborators, have proposed a quadsection type algorithm for grammar-based image compression. We apply this methodology to compression of rectangular image data. In many existing methods, compressed data are difficult to extract some patterns, which exist in the original image, from compressed data. Therefore, in this abstract, we propose approximation algorithm QUADSECTION (QSN) for finding a small CFRIG which is obtained by extending BISECTION algorithm for text data compression.[1][2] Furthermore, we show that QSN computes in polynomial time a grammar of size $O(g^*n^{4/3})$ for a given image of size $O(n) \times O(n)$, where the g^* is the size of a minimum grammar generating the given image. As we know, it is important to understand protein structures and functions. Therefore, we present the application of this new method for comparing of distance maps of protein structures. Besides, I also began doing research about predicting complex structures which has some common place with the research I'm going to do in Knapp laboratory.</p> <p><i>Research objectives</i></p> <p>At the Free University of Berlin, I will collaborate with Knapp's group on the following projects:</p> <ol style="list-style-type: none"> (1) There is some work to be done with the neural network described in predicting protein complex geometries with a neural network by Chae et al. (2) They have a new neural network program written in C++ written by Florian Krull (one of the coauthors). This program has to be tested, used and probably integrated in a framework that makes it easy to use for protein-protein docking. (3) Optimize parameters of the neural network to reach quality the neural network already published. (4) Find suitable stop function to stop training of neural network with optimal recall but without over-fitting.