Multivariate Estimation of Distribution Algorithms applied in Protein Structure Prediction

Daniel Rodrigo Ferraz Bonetti (Sao Carlos, Brazil)

Vortrag im Rahmen des Institutskolloquiums
Dienstag, 22 Juli 2014, 17:15 Uhr
Seminarraum, Insittut für Statistik

Proteins are essential in living organisms. They are responsible for maintaining life in living beings. Knowing the structure of a protein responsible for a disease, for example, it is possible to develop treatments. However, discovering the structure of a protein is a very difficult and expensive task, requiring highly skilled professionals, since it uses experimental methods to find the protein structures. Computational methods have been developed to find proteins structures. They require several calculations to predict even a small protein, since it's difficult to properly explore the large search space. We use metaheuristics to predict proteins. We did a comparison between the metaheuristics Random Walk, Monte Carlo, Genetic Algorithm, Differential Evolution and Estimation of Distribution Algorithm (EDA). We noticed that even a poor metaheuristic such as RW can find good solutions, but this would require much prior knowledge. Otherwise, EDAs are able to find good solutions with no prior knowledge at all, characterizing such a prediction as pure ab initio. We concluded that there is an increase of quality from RW going through Monte Carlo, Genetic Algorithm, Differential Evolution and EDA due to the increase of the number of the potential solutions used to compose new ones. Thus, we can show that EDAs are properly suitable for pure ab initio Protein Structure Prediction problem.