Master internship proposal

Laboratory: Inria Sophia Antipolis - Méditerranée

Supervisors: Frederic.Cazals@inria.fr (Inria, ABS, `http://team.inria.fr/abs`)
Dorian.Mazauric@inria.fr (Inria, ABS, `http://team.inria.fr/abs`)
Nicolas.Nisse@inria.fr (Inria, COATI, `https://team.inria.fr/coati`)

Exploration Algorithms of Combinatorial Landscapes

Context in Computational Structural Biology (CSB). A potential energy landscape (PEL) of a molecular system is a function that represents the potential energy for every conformation of the system. The variables are the 3n coordinates of the n atoms composing the molecular system. Exploring such a landscape is a central problem in CSB because identifying the low local minima of this function will permit to know the corresponding molecular conformations of high interest (that can be very helpful for folding and docking problems). State-of-the art methods for exploring PEL have been developed in ABS [RDRC16]. See `https://sbl.inria.fr/doc/Landscape_explorer-user-manual.html` for the Landscape Explorer Application in the Structural Bioinformatics Library (SBL) `sbl.inria.fr` [CD17].

Context in Graph Algorithms Techniques. Graphs (and hypergraphs) can model many structures: roads between cities, links in a telecommunication network, contacts between sub-units of a protein, friendship between users in social networks... Important problems in such networks can be translated into combinatorial problems on the corresponding graphs (see concrete examples below). Unfortunately, many problems are known to be NP-complete. Intuitively, there is no polynomial time algorithm for solving them, unless P=NP. Many approximation algorithms, algorithms for particular instances... have been developed. See for instance [Vaz01] for approximation algorithms.

Exploration Algorithms of Combinatorial Landscapes. Given an important combinatorial problem Π (see concrete examples below), the idea is first to model/translate instances of Π into landscapes (called combinatorial lanscapes) for which optimal solutions for Π represent local minima. Then, exploration algorithms will be used to identify some low local minima and so optimal (or near) optimal solutions of Π.

Research programme. The aim of this internship is first to translate a combinatorial problem into a problem of exploration of an auxiliary combinatorial landscape such that low local minima are optimal solutions (or near optimal solutions). During the internship, we will focus on the following combinatorial problems: problems related to graph decompositions (e.g. pathwidth, treewidth), problem of computing maximum independent set of graphs, and internal distance matching problem. The second part of the internship consists in integrating (in the SBL) the methods/algorithms for designing the combinatorial landscapes and test them for simple instances, by using the explorations methods of the SBL. We think that proteins are able to solve, very quickly, very hard problems!

Background. Theoretical computer science and/or bioinformatics and/or applied maths. Abilities for Programming will be very appreciated.

Misc. The internship may be followed-up by a PhD thesis.

References

[CD17] F. Cazals and T. Dreyfus. The Structural Bioinformatics Library: modeling in biomolecular science and beyond. *Bioinformatics*, 7(33):1–8, 2017.

[RDRC16] A. Roth, T. Dreyfus, C.H. Robert, and F. Cazals. Hybridizing rapidly growing random trees and basin hopping yields an improved exploration of energy landscapes. *J. Comp. Chem.*, 37(8):739–752, 2016.

[Vaz01] V. V. Vazirani. *Approximation Algorithms*. Springer-Verlag New York, Inc., New York, NY, USA, 2001.