"Fast Free Energy Calculations Using Markov Random Fields"

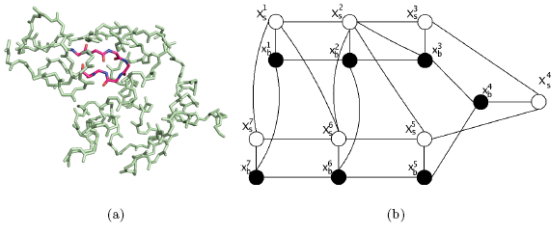
Christopher James Langmead
Department of Computer Science, Carnegie Mellon University

Summary

We have developed a new class of algorithms for molecular modeling based on Markov Random Fields (MRF). The primary advantages of the approach are its speed and flexibility. For example, MRF-based free-energy calculations are orders of magnitude faster than those based on molecular dynamics or Monte Carlo simulations. Moreover, a MRF-based approach to molecular modeling provides a unified framework for performing a wide range of activities including: structure prediction, protein-ligand interactions, and protein sequence design.

Free energy is a thermodynamic quantity that either directly or indirectly governs many of the physical properties of interest to scientists. Consequently, one of the most fundamental and important tasks in molecular modeling is the calculation of free energies. Unfortunately, traditional methods for performing free energy calculations, based on either molecular dynamics (MD) or Monte Carlo (MC) simulations, are expensive, requiring hours or days on multi-processor machines.

We have developed a novel, physics-based approach to molecular modeling. Our approach begins by compactly encoding the phase-space and Boltzmann distribution for a given system as a Markov Random Field (MRF). Free-energy calculations are then performed using algorithms for probabilistic inference over MRFs. These algorithms are physically sound. In particular, they are a formally exact approximation to the Helmholtz free-energy. Moreover, MRF-based free-energy calculations are orders of magnitude faster than MD or MC simulations.



MRF-based Molecular Modeling – (a) Structure of the backbone atoms of the protein Lysozyme with a few residues highlighted. (b) Part of the MRF induced by the highlighted residues. The structure of each residue is modeled using two multivariate random variables (black and white nodes). Edges encode physical interactions among the constituent atoms of each node.

Due to the speed of our approach, it is now possible to incorporate free-energy calculations in domains where simple potential-energy calculations are now used, such as protein structure prediction. In

*412-268-7571,cjl@cs.cmu.edu

particular, we have shown [1] that our free-energy calculations are sufficiently accurate to improve the accuracy of protein structure prediction.

In addition to their ability to perform fast free-energy calculations, MRFs provide remarkable flexibility. For example, we are presently using MRFs to perform protein-protein docking, protein-ligand docking, and automatic protein sequence design.

References

[1] "*Free Energy Estimates of All-atom Protein Structures Using Generalized Belief Propagation*", H. Kamisetty, E.P. Xing, C.J. Langmead; *Proceedings of The Eleventh Annual International Conference on Computational Molecular Biology (RECOMB),* pp 366-380

For further information on this subject contact:

Dr. Christopher James Langmead
Carnegie Mellon University
cjl@cs.cmu.edu
412-268-7571

Or

Dr. Anil Deane
Applied Mathematics Research Program
Office of Advanced Scientific Computing
Phone: 301-903-1465
deane@ascr.doe.gov