

1077-92-835

Noah M. Daniels, Department of Computer Science, Tufts University, Medford, MA 02155, **Raghavendra Hosur**, CSAIL, Massachusetts Institute of Technology, Cambridge, MA 02139, **Lenore J. Cowen*** (lenore.cowen@tufts.edu), Department of Computer Science and, Department of Mathematics, Tufts University, Medford, MA 02155, and **Bonnie Berger** (bab@mit.edu), Department of Mathematics and, CSAIL, Massachusetts Institute of Technology, Cambridge, MA 02139. *Combining simplified Markov random fields with simulated evolution improves remote homology detection for beta-structural proteins into the twilight zone.* Preliminary report.

One of the most successful methods to date for recognizing protein sequences that are evolutionarily related has been profile Hidden Markov Models (HMMs). However, these models do not capture pairwise statistical preferences of residues that are hydrogen bonded in beta sheets. These dependencies can be partially captured in the HMM setting by simulated evolution in the training phase, and can be fully captured by Markov Random Fields (MRFs). However, the MRFs can be computationally prohibitive when beta strands are interleaved in complex topologies.

We introduce SMURFLite, a method that combines both simplified Markov Random Fields and simulated evolution to substantially improve remote homology detection for beta structures. Unlike previous MRF-based methods, SMURFLite is computationally feasible on any beta-structural motif. We show a median 24 percent improvement in AUC for beta-structural motif recognition as compared to HMMer (a well-known HMM method), a median 16 percent improvement in AUC as compared to Raptor (a well-known threading method) and even a median 9 percent improvement in AUC as compared to HHpred, despite its use of additional training data, in a stringent cross-validation experiment. (Received September 13, 2011)