Computational Challenges for Discovering Homologies between Genome Sequences

Sakti Pramanik[1, 2], James Cole[3], Min Zhang[2], Eric Torng[2], Patrick Gallagher[2]

1. Principle Investigator   2. Computer Science Department   3. Center for Microbial Ecology

Homologies between genome sequences provide a reasonable basis for discovering important biological features. Computation complexity for discovering them is generally very high and requires effective techniques to model them. One of the goals of this research is to investigate several fundamental properties of data distribution in high dimensional Euclidean space, and apply them to develop effective methodologies for finding sequence homologies in large databases. We are also developing abstraction-based visual computer models to aid in discovering related biological features between large genome sequences. Relatedness between genome sequences using restriction maps and restriction patterns is also being investigated.