

Roni Rosenfeld

Professor

Learning Technology Institute, Machine Learning Department, Computer Science Department

2:00, August 28, 2007

Research Topic

Computational Genetics

Research Problem

How can we rapidly analyze the human genome?

Problem Statement

Given a large database of RNA viruses and the ability to track their mutations, construct a model that can predict genetic sequences from limited portions of the genome.

Operational Definitions

RNA Viruses: Viruses that have RNA, not DNA, as their genetic sequence and which are known for extremely rapid evolution.

Genome: The complete genetic code of a given organism.

Problem Description

Since the Human Genome Project first successfully sequenced human DNA, the time and money needed to analyze the genome has begun to decrease, and the amount of genetic information available for analysis has started to grow. By working with RNA viruses, which mutate rapidly and thus provide a rich data source, it is possible to study the analysis of large volumes of genetic data, prediction of gene sequences given partial information, and ways to visualize genetic sequences.

Computer Science Perspective

Artificial Intelligence, theoretical computer science, and visualization are connected to this research. To identify viruses quickly, fast algorithms for searching large amounts of data are required, as well as new algorithms for inference from incomplete data and for intuitively displaying the data, all of which fall into the above three domains.

Disciplines Actively Involved

Genetics

Computer Science

Operational Definition

Actively Involved Disciplines: A Discipline from which a member would be acknowledged in the research paper or any discipline involved for which there is a professional association that fosters more knowledge of it through the scientific method.

Description of Disciplines Involved

Genetics: This research actively involves Genetics, as it relies on current research into Genetics and is intended to improve our ability to analyze an individual's Genetic code and thus make predictions about our health.

Computer Science: This research involves developing new, faster algorithms for improved predictive modeling, a topic which is highly relevant to Computer Science.

References

Presenter web page:

<http://www.cs.cmu.edu/~roni/>

Project GATTACA website:

<http://www.cs.cmu.edu/~gattaca/>

By plandweh

Updated plandweh