

CS196-1: Algorithmic Foundations of Computational Biology

CIT 368 · Tuesday and Thursday, 9:00-10:20

Introduction

The course is an introduction to Computational Biology and Bioinformatics and is open to computer and mathematical sciences students as well as biological and medical students. Both advanced undergraduates and graduate students are welcome. The instructor taught evolutionary versions of this course in Departments of Biology, Computer Science, and Biochemistry and Cell Biology (Medical School).

Although “Computational Biology” and “Bioinformatics” are often used interchangeably, we will refer to Bioinformatics as primarily being about applying computational genomics tools, with well-established biological relevance, in large-scale applications in molecular biology and medical laboratories. Computational Biology, in turn, is about building these powerful genomics tools. Although there is significant overlap, both areas are essential: Computational Biology is more about the science, while Bioinformatics is more about the technology and engineering. To use a metaphor, Bioinformatics is about expeditions on the sea shores to find precious metals, or at the bottom of the sea to find oil, while Computational Biology is about building metal detectors or oil detection systems.

This Course is about Computational Biology and Bioinformatics in the Genome Era, which started in 2001 when the first assemblies of the Human Genome were published. Since then high-throughput biotechnologies are generating genomics and systems biology data sets of unprecedented magnitude and complexity. An exciting Informatics Renaissance is emerging where computer science is leading the way through new algorithmic, visualization, and computing paradigms. Building upon modeling insights from experimental molecular biology, physics, chemistry, and economics, Computational Biology is harnessing the complexity of biological systems by providing computational models and genomics tools that transform knowledge into understanding.

In this Course we will study computational and statistical methods in addition to genomics tools for bio-molecular (DNA, protein) sequence analysis. We will focus on understanding the design of the tools and methods, as well as on their applications.