



# All Science is Computer Science!

Vasant Honavar  
Artificial Intelligence Research Laboratory  
Department of Computer Science  
Bioinformatics and Computational Biology Graduate Program  
Computational Intelligence, Learning, & Discovery Program  
Iowa State University

[honavar@cs.iastate.edu](mailto:honavar@cs.iastate.edu)  
[www.cs.iastate.edu](http://www.cs.iastate.edu)  
[www.cild.iastate.edu](http://www.cild.iastate.edu)  
[www.bcb.iastate.edu](http://www.bcb.iastate.edu)



# Outline

- Current state of Computer Science Education
- Conceptual impact of Computer Science
- All Science is Computer Science!
- ISU Bioinformatics Experience
- Implications for Computer Science Education



## Computer Science

Computer science is often equated with the information technologies enabled by it:

Hardware and software for personal computing

Internet and the world-wide web

Electronic commerce

Data mining

Digital libraries

Precision farming

Computer assisted surgery

Workflow management

Databases and information systems

Virtual reality

Smart homes

.....



# Computer Science Education

Computer Science encompasses a body of knowledge concerning algorithms, communication, languages, software and information systems.

Computer science education at present tends to focus on

- Imparting the necessary skills for creating information technology solutions  
and / or
- Teaching the mathematical foundations of computer science - theory of computation, logic, algorithms, and complexity



# Computer Science Education

Computer science education at present largely ignores the **conceptual impact** of computer science on science, technology, humanities, society

The result is a computer science curriculum that aims to train either technicians or mathematicians but not a new generation of scientists and scholars

**This needs to change!**



# Conceptual impact of Computer Science

Computer science, broadly defined, is the theory and practice of representation, processing, and use of information.

Computer Science offers a powerful paradigm for modeling complex phenomena such as cognition and life, and representing, processing, acquiring, and communicating knowledge that is new in the history of humanity.



# The road to Computer Science

History of computer science is really a history of humankind's attempts to understand **nous** (the rational mind) -- intelligence – processes of acquiring, processing, and using information and knowledge.

Aristotle (384-322 BC) distinguishes **matter** from **form** thereby laying the foundations of **representation**

Panini (350 BC) develops a formal grammar for Sanskrit

Al Khowarizmi (825) introduces algorithms in his text on mathematics

Descartes (1556-1650) – **Cogito ergo sum!**



# The road to Computer Science

Hobbs (1650) suggests that thinking is a rule-based process analogous to arithmetic

Leibnitz (1646-1716) seeks **a general method for reducing all truths to a kind of calculation**

Boole (1815-1864) **proposes logic and probability as the basis of laws of thought**

Frege (1848-1925) further develops first order logic

Tarski (1902-1983) introduces a theory of reference for relating objects in a logic to objects in the world



# The road to Computer Science

Hilbert (1862-1943) presents the **decision problem** – Is there an effective procedure for determining whether or not a given theorem logically follows from a given set of axioms?

Godel (1906-1978) shows the existence of an effective procedure to prove any theorem in Frege's logic and proves the **incompleteness theorem**

Turing (1912-1954) invents the Turing Machine to formalize the notion of an **effective procedure**



# The road to Computer Science

Church, Kleene, Post, Markov (1930-1950) develop other models of computation based on alternative formalizations of **effective procedures**.

Turing and Church put forth the **Church-Turing thesis** that Turing machines are **universal** computers.

Several special purpose analog and digital computers are built (including the Atanasoff-Berry Computer)

Von Neumann (1956) works out a detailed design for a stored program digital computer

Chomsky (1956) develops the Chomsky hierarchy of languages



# The road to Computer Science

Several digital computers are constructed and universal languages for programming them are developed – Lisp, Snobol, Fortran...

Von Neumann, McCulloch, Rashevsky (1940-1956), investigate the relationship between the brain and the computer

Von Neumann and Morgenstern develop a formal framework for rational decision making under uncertainty

Von Neumann (1956) develops a theory of self-reproducing automata



# The road to Computer Science

McCarthy, Minsky, Selfridge, Simon, Newell, Uhr et al (1956) begin to investigate the possibility of **artificial intelligence**

Dantzig and Edmunds (1960-62) introduce **reduction** – a general transformation from one class of problems to another

Cobham and Edmunds (1964-65) introduce **polynomial and exponential complexity**

Cook and Karp (1971-72) develop the theory of NP-completeness which helps recognize problems that are **intractable**

The rest is... recent history 😊



# Conceptual impact of computer science

The **language** of computation is the best language we have so far for describing how **information** is encoded, stored, manipulated and used by natural as well as synthetic systems

**Algorithmic or information processing models provide for biological, cognitive, and social sciences what calculus provided for classical physics**



# Conceptual impact of Computer Science

Computation: Cognition : : Calculus : Physics (Artificial Intelligence, Cognitive Science)

What are the information requirements of learning?

What is the algorithmic basis of learning?

What is the algorithmic basis of rational decision making?

Can we automate scientific discovery?

Can we automate creativity?



# Conceptual impact of Computer Science

Computation: Cognitive Science :: Calculus : Physics  
Computer science offers fundamentally new ways to  
understand cognitive processes –

Perception

Memory and learning

Reasoning and planning

Rational decisions and problem solving

Communication and language

Behavior



# Conceptual impact of Computer Science

Computation: Life :: Calculus : Physics

(Computational Biology, Computational Ecology ..)

How is information acquired, stored, processed, and used in living systems – in gene expression, protein folding, protein-protein interaction, reproduction?

How do brains process information?

How do genes and environment determine behavior?

What does the genetic program for fetal development look like?



# Conceptual impact of Computer Science

Computation: Biology :: Calculus : Physics

Computer science offers fundamentally new ways to understand biological processes

Reproduction

Development

Molecular function

Gene regulation and expression

Cellular function

Signal transduction

Brain function

Adaptation and Evolution



# Conceptual impact of Computer Science

Computation: Society :: Calculus : Physics

(Computational Economics, Computational Organization Theory)

What is the informational and algorithmic basis of inter-agent interaction, communication, and coordination?

Under what conditions can self-interested rational agents cooperate to achieve a common good?

How do groups and coalitions form?

How do different social organizations (democracies, economies, etc.) differ in terms of how they process information?



# Conceptual impact of Computer Science

Computation: Social Sciences :: Calculus : Physics

Computer science offers fundamentally new ways to understand economic, and organizational, and social phenomena

Cooperation and competition

Bounded rationality and economic behavior

Community, coalition, and organization formation

Social contract

Organization, interaction, and communication

Rise and fall of cultures e.g., the Anasazi Indians



# Conceptual impact of Computer Science

Computer science offers fundamentally new ways to model and understand

cognitive, biological, and social processes

through computational or information processing or algorithmic models



# Algorithms as Theories

We will have a theory of

**learning** when we have precise information processing models of learning (computer programs that learn from experience)

**protein folding** when we have an algorithm that accepts a linear sequence of amino acids as input and produces a description of the 3-dimensional structure of a protein as output

**bounded rationality** when we have an algorithm for rational decision making under limited information, memory or computation



# Conceptual impact of Computer Science

## Pre-Turing

Focus on **physical basis** of the universe with the objective of **explaining all natural phenomena in terms of physical processes**

## Post-Turing

Focus on **informational** and **algorithmic basis** of the universe with the objective of **explaining natural phenomena in terms of processes that manipulate information**

**We understand a phenomenon when we can write a computer program that models it at the desired level of detail**

When theories and explanations take the form of algorithms, all science becomes computer science!



# All Science is Computer Science!

Is it any surprise then that Computer Science has given birth to:

Computational molecular biology and bioinformatics

Computational neuroscience and neuro-informatics

Computational developmental biology

Cognitive science

Computational economics

Computational chemistry and chemo-informatics

Computational organization theory

Medical informatics

Agricultural informatics

Geo-informatics



# All Science is Computer Science!

Is it any surprise then that computer scientists are being hired by departments of

Biological Sciences

Chemistry

Economics

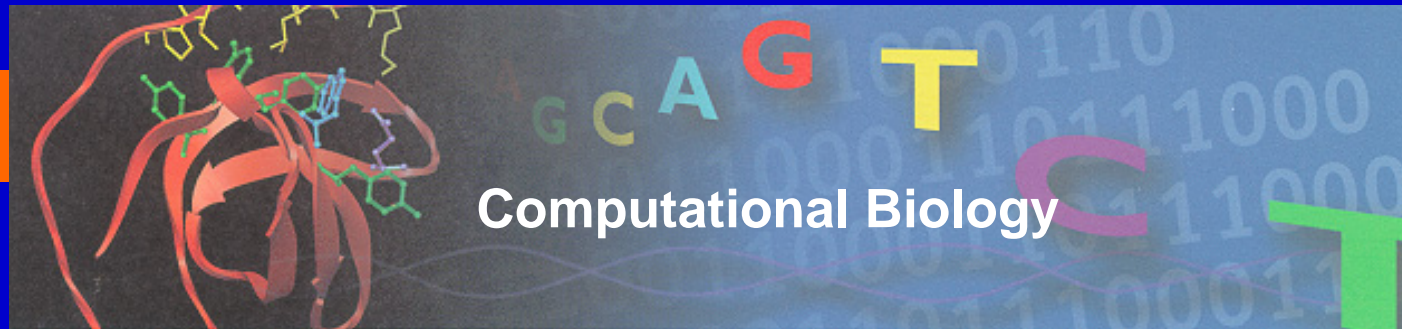
Engineering

Philosophy

Physics

Psychology

Sociology .....



## Case Study -- The ISU Bioinformatics and Computational Biology Program

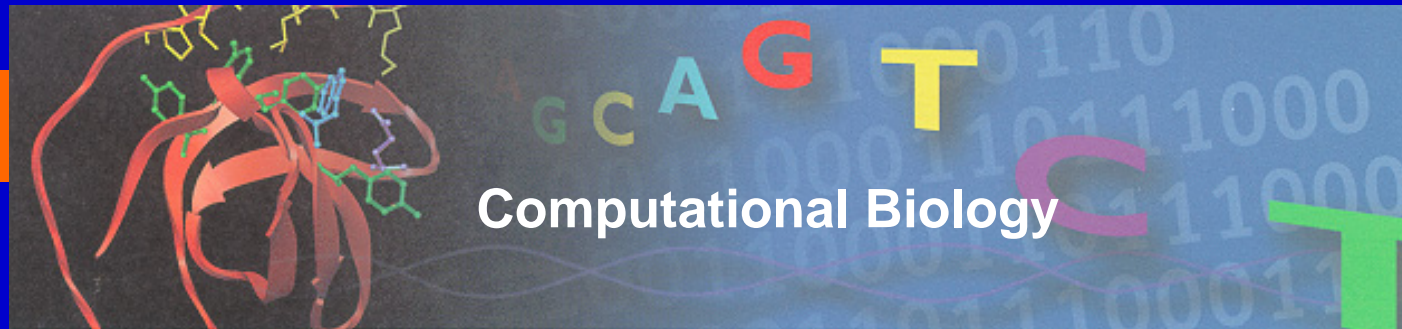
What is Bioinformatics?

What is Computational Biology?

ISU Bioinformatics and Biology Program

Computer Science BCB Curriculum

Computer Science BCB Research



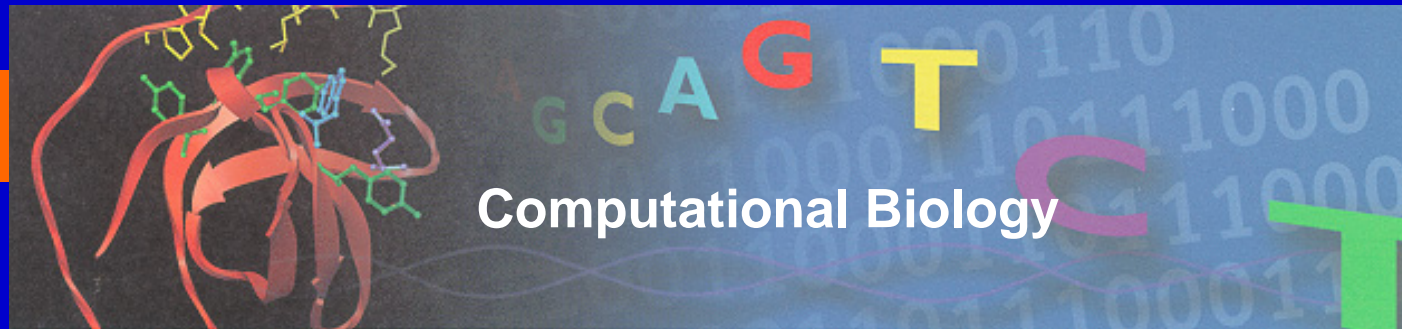
## Computational Biology

Algorithmic models provide for biological sciences what calculus provided for classical physics

The *language* of computation is the best language we have so far for describing how *information* is encoded, stored, manipulated and used by biological systems

Central problem: Given genomic sequences – text in a language with known alphabet but unknown syntax and semantics, and some additional clues, discover the syntax and semantics!

Goal is to develop *information processing* or *computational* models of biological processes (protein folding, gene regulation, protein-protein interaction)



## Representative Problems

Inference of tree of life from DNA sequence data

Characterization of protein sequence – structure – function relationships e.g., discovery of sequence and structural correlates of protein-protein interactions

Genetic network inference from gene expression data

Inference of metabolic pathways

Modeling and prediction of cellular processes

Modeling host-pathogen interactions



Transformation of Biology from a *data poor* science to a *data rich* science

- High throughput Data Acquisition
- Processors, Storage, and Communication Technologies
- Algorithms for information processing

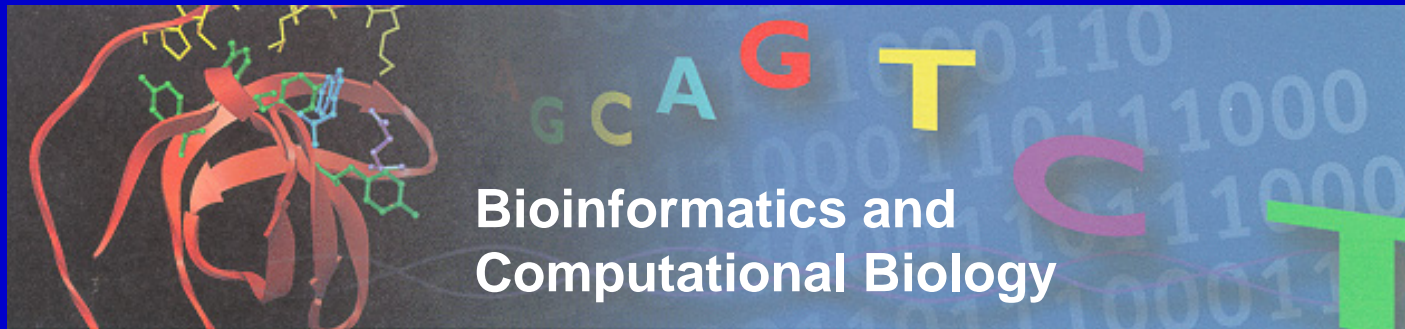
*In principle, it is possible to gather, store, access, and analyze large volumes of data (e.g., sequence data, structure data, expression data)*

The focus of bioinformatics is on the design and implementation of software tools for data driven knowledge discovery in data rich biological sciences



Leveraging the ability to gather, store, and process large volumes of data at increasing rates into scientific advances requires new algorithms and software for

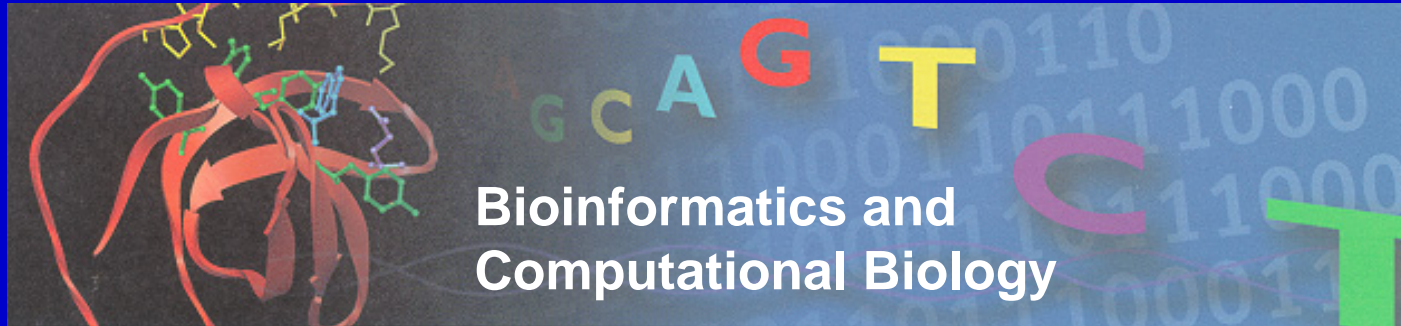
- Data description, organization, visualization
- Pattern matching, retrieval
- Information extraction and integration
- Knowledge representation
- Data mining and hypothesis generation
- Computer assisted collaborative discovery



The focus of bioinformatics is on the design and implementation of algorithmic and systems solutions to support data driven knowledge discovery in data rich biological sciences

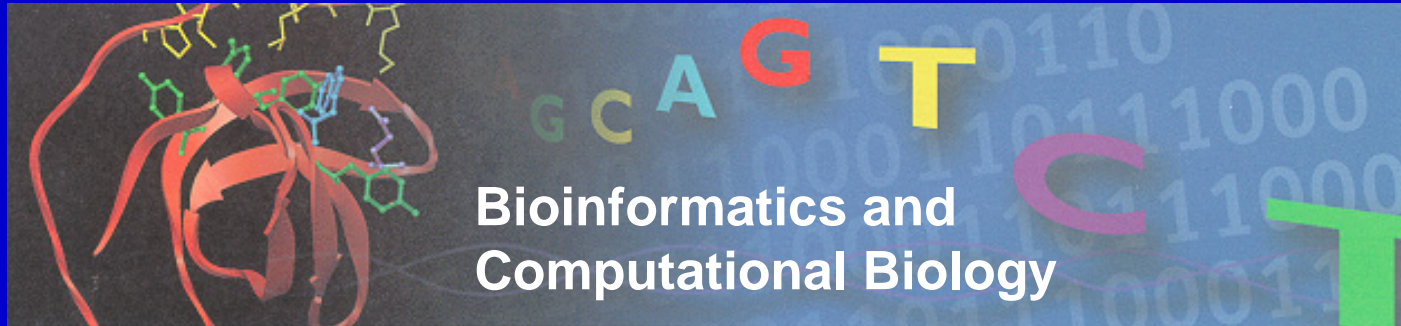
The emphasis of computational biology is on the development of *information processing* or *computational* models of biological processes (protein folding, gene regulation, protein-protein interaction)

Since hypothesis or model construction is generally data driven, bioinformatics tools are essential for computational biology



## Mission of the BCB Graduate Program

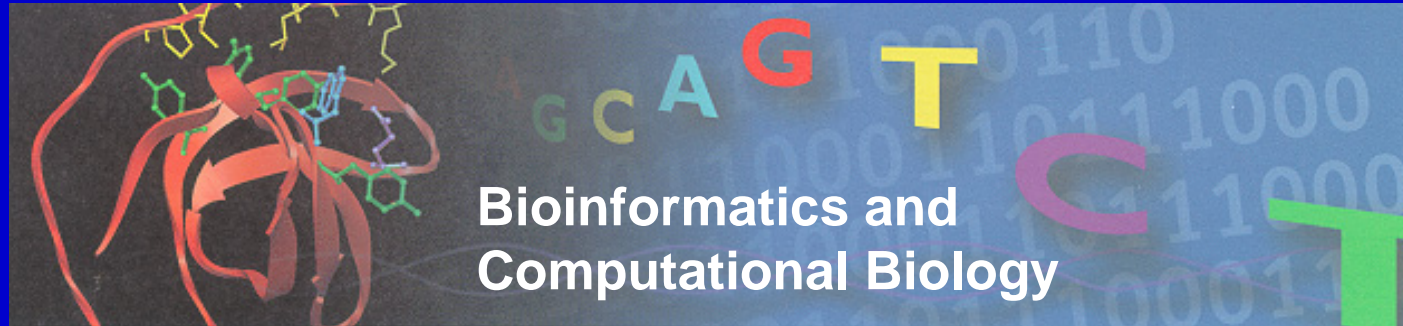
To enable students to explore fundamental research questions by providing training in biological sciences and computer science, mathematics, and statistics



## Bioinformatics and Computational Biology

### Program History

- 1997 Iowa Computational Biology Laboratory
- 1998 Formal Coursework Begins
- 1998 Graduate programs offer Areas of Specialization in Computational Molecular Biology
- 1998 Began hiring computational biologists (12 to date)
- 1999 NSF-IGERT training grant
- 1999 Bioinformatics and Computational Biology Graduate Program
- 2000 Laurence H. Baker Center for Bioinformatics and Biological Statistics
- 2002 BCB Program Review
- 2004 Board of Regents Review



## Research areas

Bioinformatics

Functional and structural genomics

Genome evolution

Macromolecular structure and function

Metabolic and regulatory networks

Mathematical biology

Biological statistics



## Program Overview

One of the first Bioinformatics Ph.D. programs in the US

Over 60 Ph.D. students, one of the largest and strongest Bioinformatics Ph.D. programs in the US

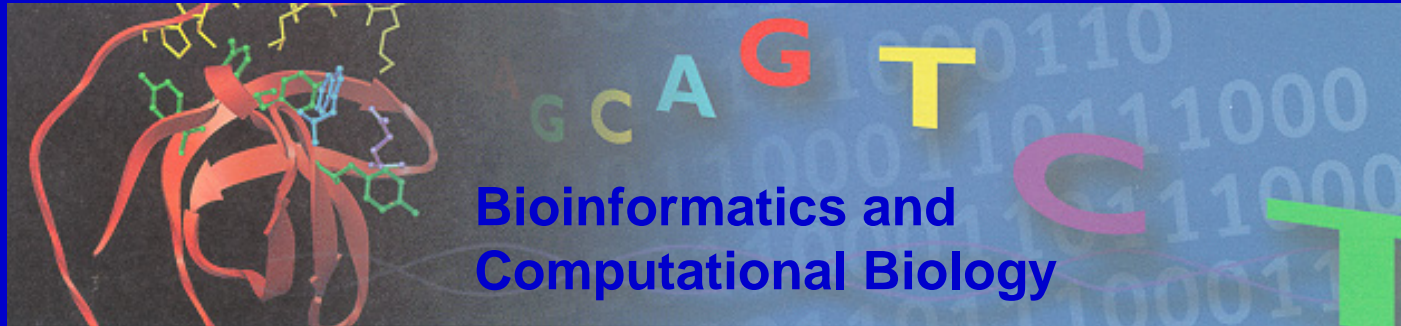
Over 72 BCB faculty in 14 different departments

\$30 million in grants between July 1999-July 2004

2 major training grants

(NSF-IGERT, \$2.7M; USDA-MGET, \$1.7M)

Bioinformatics Summer Institute (NSF-NIH)



## Bioinformatics and Computational Biology

### Interdisciplinary Training

Background & ramp-up courses

Core courses in computational molecular biology and bioinformatics

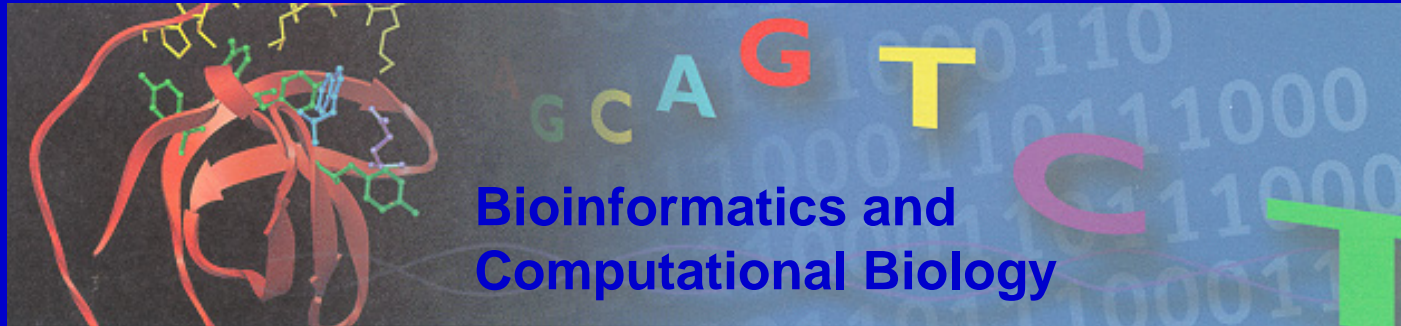
Bioethics courses & training sessions

Electives in Statistics, Computer Science, Biology

Faculty Research Seminar

Student Seminar

Computational molecular biology seminars and symposia



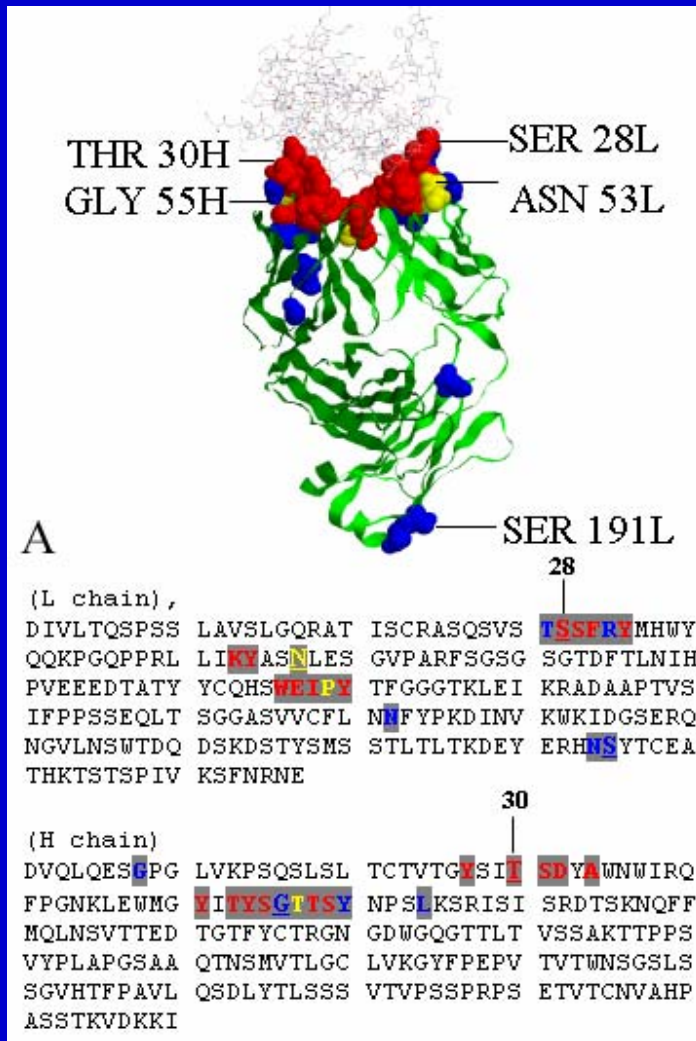
## Interdisciplinary Research

Research exploration rotations

Joint mentoring -- 'Wet' and 'dry' lab research experiences and mentoring by a major and co-major professor (from computational and biological sciences)

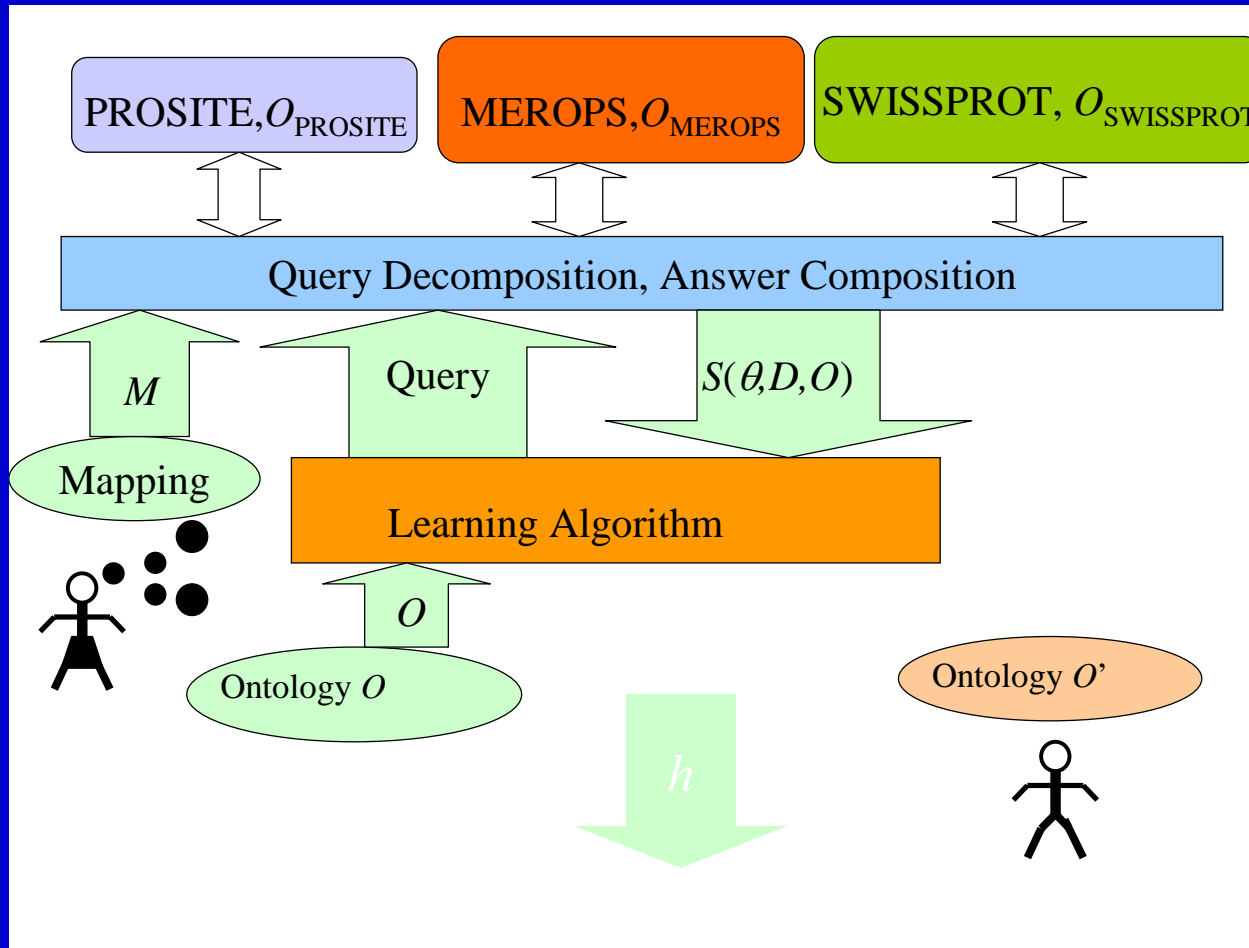
International and industrial internships

# Representative Computational Biology research



Computational approach to prediction of protein-protein interface residues from protein sequence

# Representative bioinformatics research



Infrastructure for collaborative discovery from distributed, semantically heterogeneous autonomous information sources



# The ISU Bioinformatics and Computational Biology Experience

It is possible at an institution like ISU

- to train a new generation of biologists -- computational biologists – who are proficient in both computer science and biology to pursue a fundamentally new approach to answering basic research questions in biology
- to train a new generation of computer scientists – bioinformaticists – to develop new information technologies for storage, retrieval, and analysis of diverse types of biological data and knowledge to facilitate collaborative scientific discovery in biology



# Conceptual impact of Computer Science

## Pre-Turing

Focus on **physical basis** of the universe with the objective of **explaining all natural phenomena in terms of physical processes**

## Post-Turing

Focus on **informational** and **algorithmic basis** of the universe with the objective of **explaining natural phenomena in terms of processes that manipulate information**

**We understand a phenomenon when we can write a computer program that models it at the desired level of detail**

When theories and explanations take the form of algorithms, all science becomes computer science!



## What Next?

Can we replicate the Bioinformatics and Computational Biology experience at ISU in

- **Agriculture** – Agricultural Informatics
- **Basic Sciences** – Computational Chemistry, Computational Physics, Chemoinformatics, Computational Neuroscience, Neuroinformatics
- **Engineering** – Engineering informatics
- **Social Sciences** – Social Informatics, Computational Economics

How?



# What should a Computer Scientist know?

Computer science and a subset of:

**Biology, Chemistry, and Physics** – molecular biology and neuroscience

**Mathematics and Statistics** – Logic, information theory, probability theory, statistics, game theory, decision theory

**Cognitive science** – perception, cognition, language, action

**Social sciences** – economics, organizational theory, communication theory ..

**Philosophy** – epistemology, philosophy of mind, philosophy of science



# What should every literate person know?

Elements of the theory of computation

Algorithms and information processing

Elements of Programming

Information system design and use

Use of computational models in his or her discipline



## Implications for Computer Science Education

**Computer Science undergraduate and graduate students** need significant exposure to physical, biological, cognitive, and social sciences!

**All undergraduates and graduate students** need significant exposure to **computer science** – not just information technology – across the curriculum!