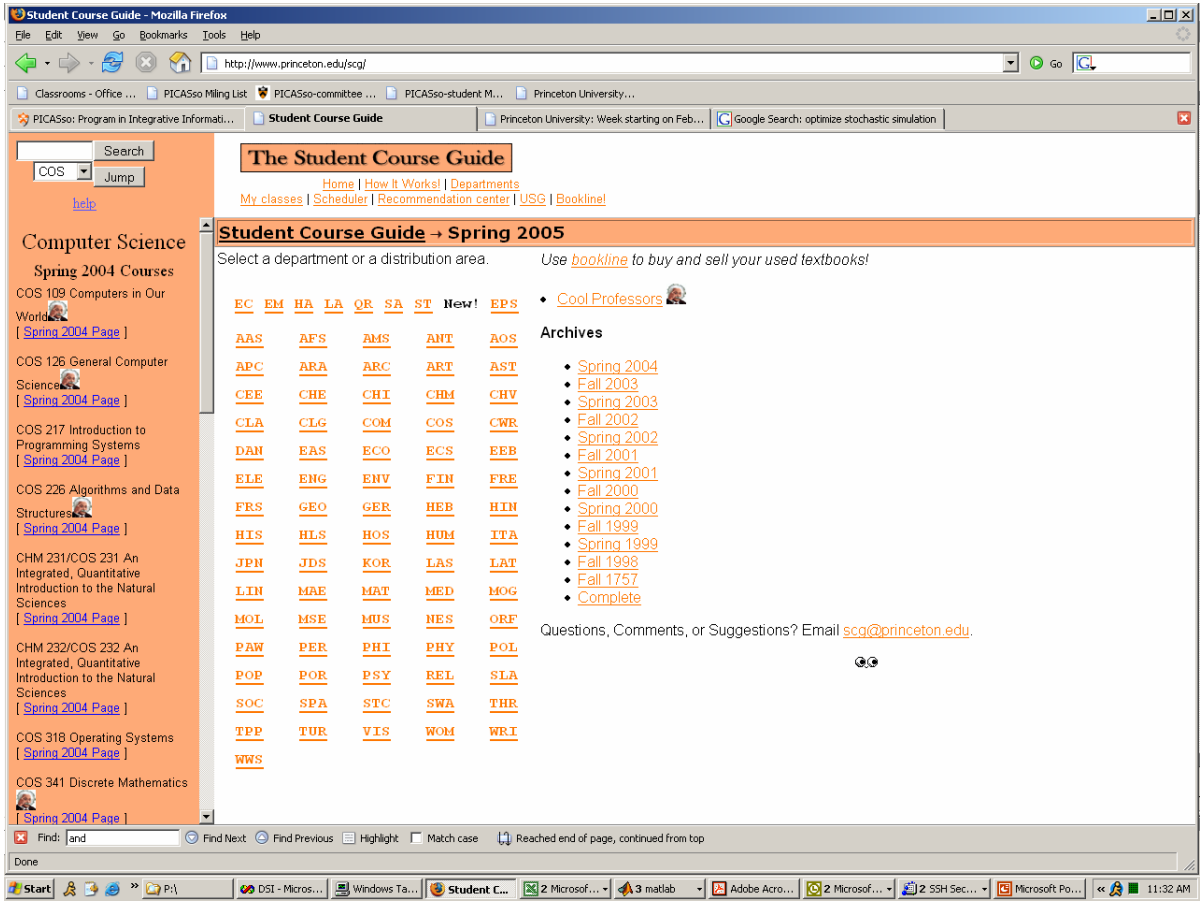
Project #1: Identifying and Visualizing Related Courses

Current interface for course searching is limited – department, keyword, distribution



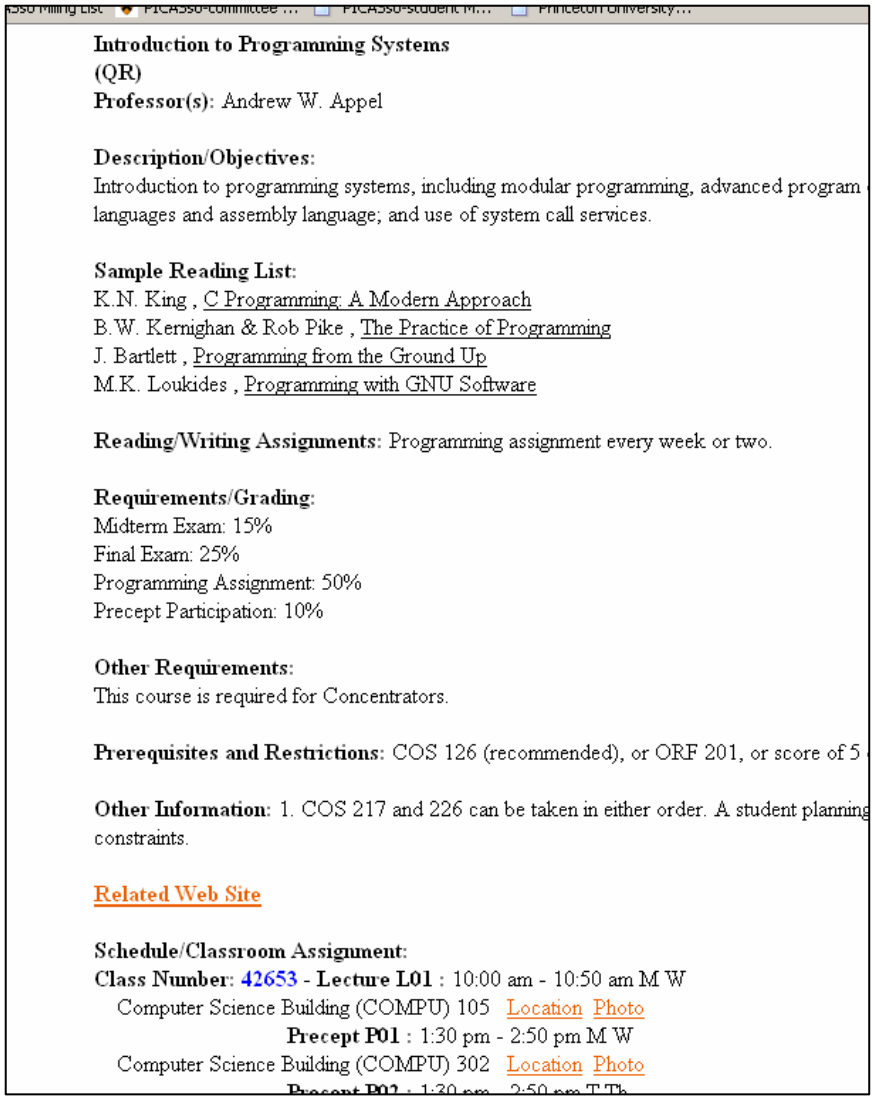
How can we identify/visualize related courses across departments?

Steve Kleinstein (stevenk@cs.princeton.edu)

Project #1: Identifying and Visualizing Related Courses

Part I: Defining a distance metric between courses, D = f(course 1, course 2)

- Information to use
  - Prerequisites
  - Keywords in name/description
  - Professor
- What is most useful measure?
- How can we combine information?



Introduction to Programming Systems
(QR)
Professor(s): Andrew W. Appel

Description/Objectives:
Introduction to programming systems, including modular programming, advanced program
languages and assembly language; and use of system call services.

Sample Reading List:
K.N. King , C Programming: A Modern Approach
B.W. Kernighan & Rob Pike , The Practice of Programming
J. Bartlett , Programming from the Ground Up
M.K. Loukides , Programming with GNU Software

Reading/Writing Assignments: Programming assignment every week or two.

Requirements/Grading:
Midterm Exam: 15%
Final Exam: 25%
Programming Assignment: 50%
Precept Participation: 10%

Other Requirements:
This course is required for Concentrators.

Prerequisites and Restrictions: COS 126 (recommended), or ORF 201, or score of 5

Other Information: 1. COS 217 and 226 can be taken in either order. A student planning
constraints.

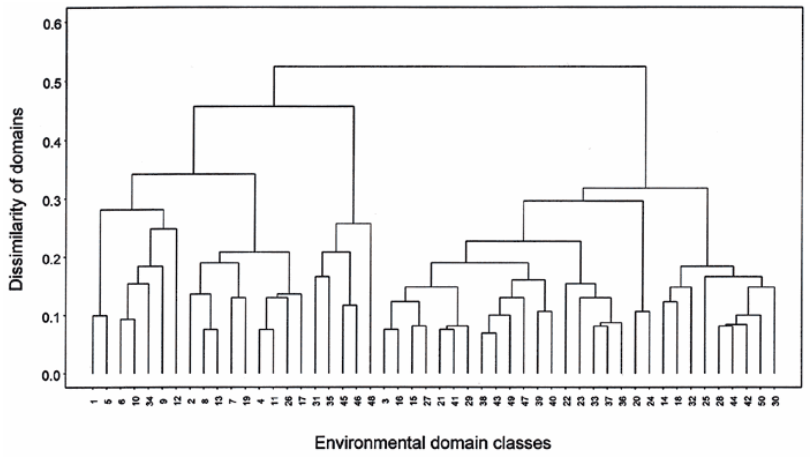
Related Web Site

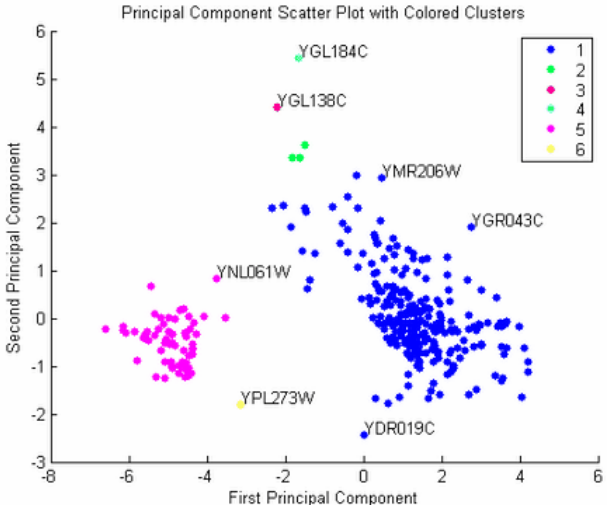
Schedule/Classroom Assignment:
Class Number: 42653 - **Lecture L01** : 10:00 am - 10:50 am M W
Computer Science Building (COMPU) 105 Location Photo
Precept P01 : 1:30 pm - 2:50 pm M W
Computer Science Building (COMPU) 302 Location Photo

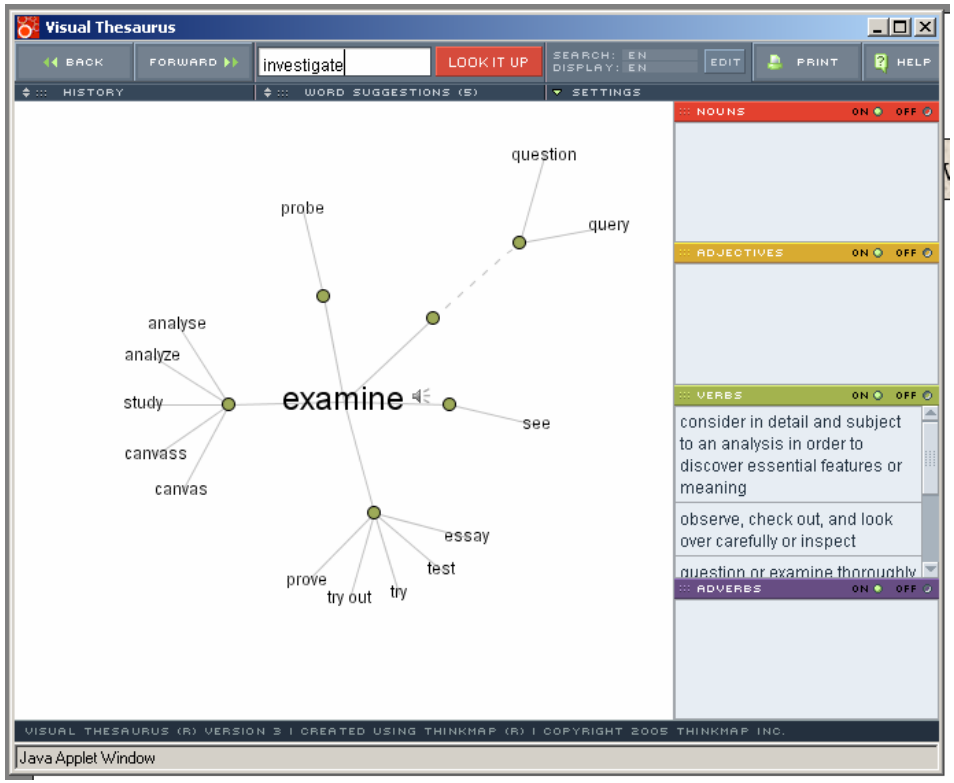
Steve Kleinstein (stevenk@cs.princeton.edu)

Project #1: Identifying and Visualizing Related Courses

Part II: Given D = f(course 1, course 2), how can we visualize courses







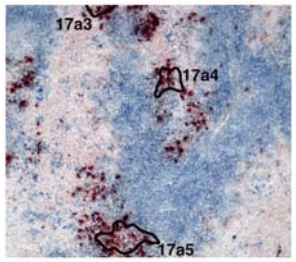
http://www.visualthesaurus.com

Potential to become useful resource for students

Steve Kleinstein (stevenk@cs.princeton.edu)

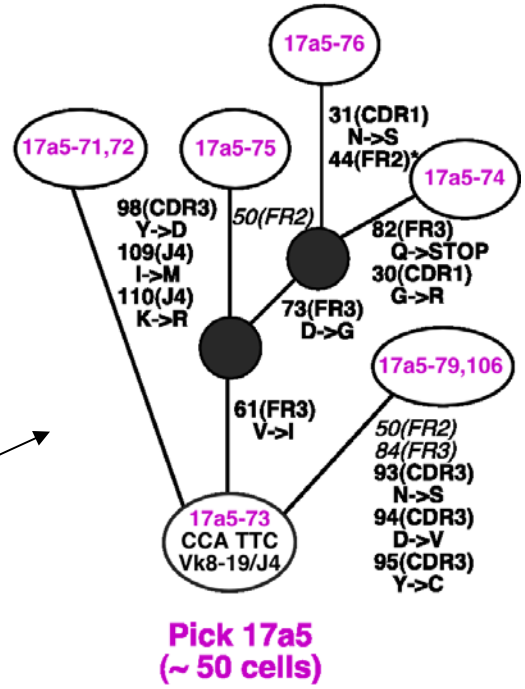
Project #2: Optimizing Stochastic Simulations

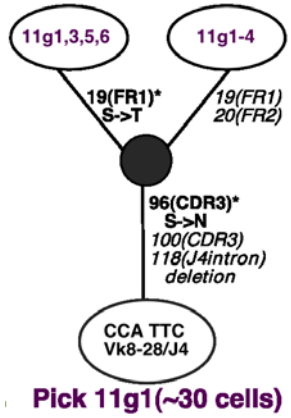
Motivation: Understand complex data from immune response dynamics



B cells T cells

Germline	GGGATTCTC
1	-C-----G-
2	-------G-
3	A------GA
4	A---C--GA





Fit computational model to experimental data
To estimate parameters, such as mutation rate (μ)

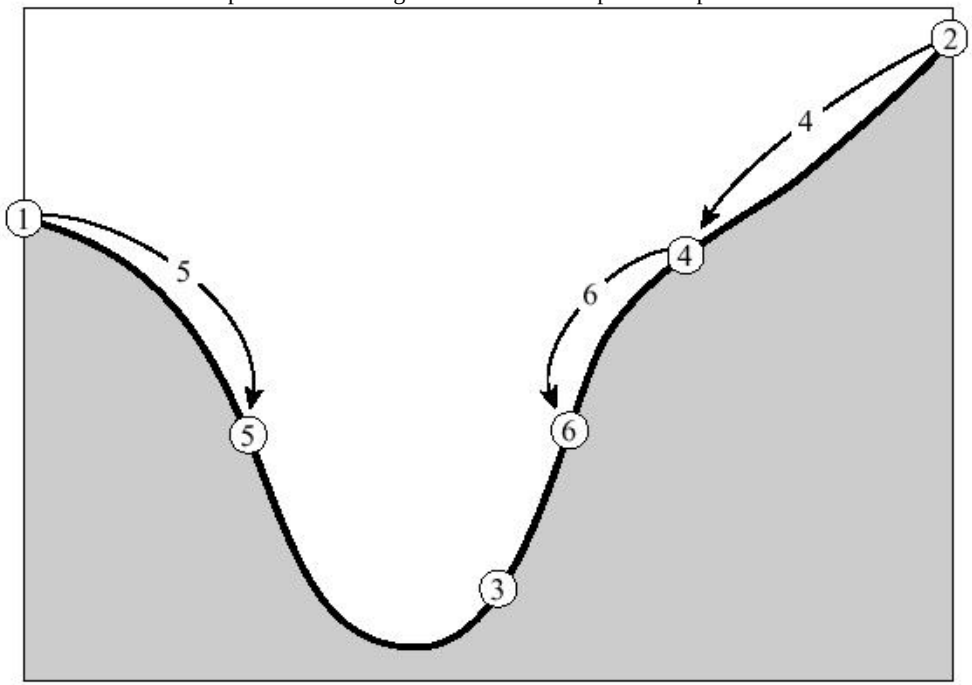
Steve Kleinstein (stevenk@cs.princeton.edu)

Project #2: Optimizing Stochastic Simulations

Current approach: Run simulation many times at each parameter value

http://lib-www.lanl.gov/numerical/bookcpdf/c10-1.pdf

Difference between simulation and experiment



Mutation Rate (μ)

Method requires >128,000 simulations per evaluation

Steve Kleinstein (stevenk@cs.princeton.edu)

Project #2: Optimizing Stochastic Simulations

Better approach would take noise into account during the optimization

- Implement and compare methods for optimization of noisy functions
  - Pairwise bisection, Anderson/Ferris Method, etc.
- Test on computational immunology model
- Investigate possible improvements

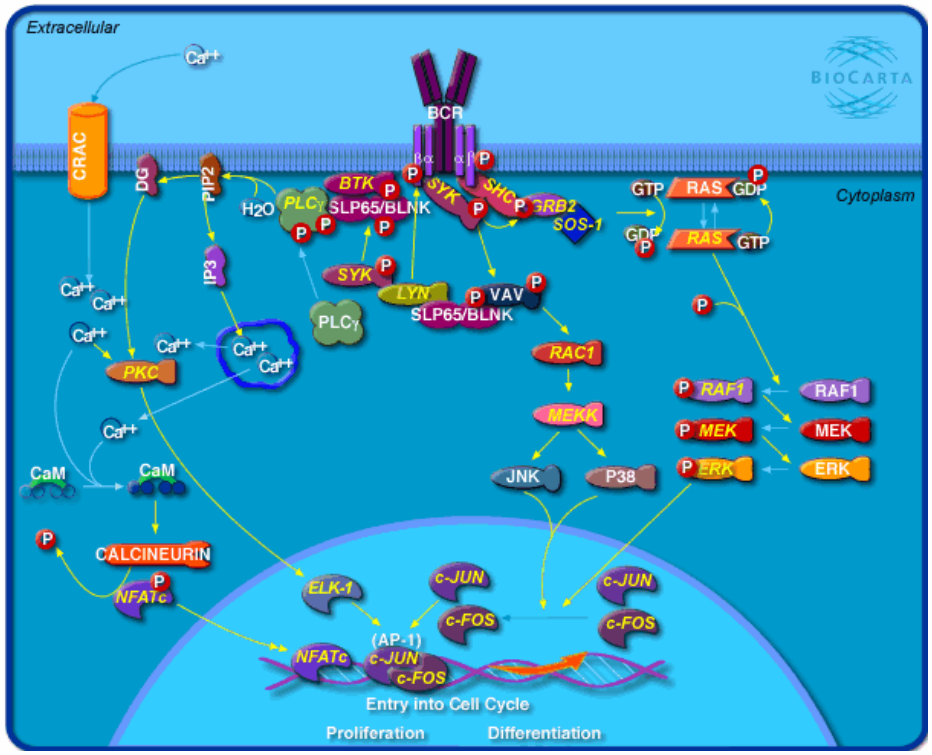
Simulations in many disciplines need such methods

Steve Kleinstein (stevenk@cs.princeton.edu)

Project #3: Parallel Global Optimization

Estimate parameters for large biological simulations

B Cell Signaling Pathway



- Estimate rate constants
- Evaluate effect of drugs
- Estimate drug dosing

Global optimization important since many local minima, but requires many simulations (expensive)

Steve Kleinstein (stevenk@cs.princeton.edu)

Project #3: Parallel Global Optimization

Implement and analyze parallel version of global optimization algorithm (for Matlab)

- Collaboration with BioAnalytics Group and Novartis
- Several Matlab Global Optimization Routines
  - Differential Evolution, SRES, Others…
- Implement different parallel versions
  - Coarse-grain: objective function level, population level
  - Fine-grain: elements level
- Compare performance on set of biological problems
  - Biopathway model, PK/PD simulation, etc.

Make use of CS Beowulf cluster (>64 processors)

Steve Kleinstein (stevenk@cs.princeton.edu)

Summary of Potential Projects

- **Project #1**: Identifying/Visualizing Related Courses
  - With Andrea LaPaugh
- **Project #2**: Optimizing Stochastic Simulations
  - With JP Singh (+ Michael Ferris, U. of Wisconsin ?)
- **Project #3**: Parallel Global Optimization
  - With JP Singh (+ Scott Lett, BioAnalytics Group)
- Also projects in Computational Immunology

For more information

Steve Kleinstein

stevenk@cs.princeton.edu