Optimization revisited

Prof. Noah Snavely
CS1114
http://www.cs.cornell.edu/courses/cs1114/

Cornell University
Computer Science

Administrivia

- Assignments:
  - A6 due tomorrow (demo slots available)
- Prelim 3 next Thursday
  - Review in class on Tuesday

Administrivia

- Final projects
  - Demo session: Tuesday, May 15, 1:30-3:30
  - For the demo session, prepare a 7-minute demo for your project
  - Sign-up slots available soon
  - Best demo will receive the "Best Demo" prize

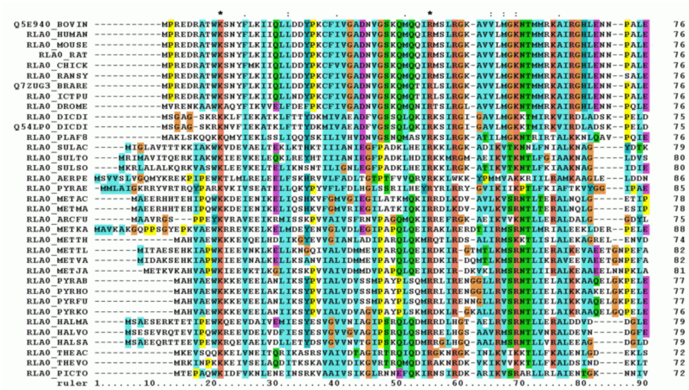
Optimization problems

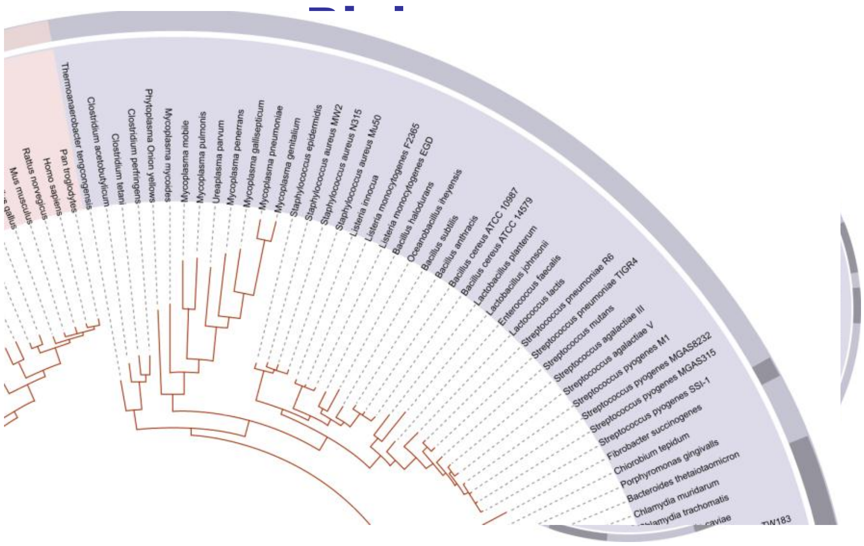
- Very important concept for many different areas of science and engineering

- City planning
  - How do I figure out what bus routes to put in Ithaca?
  - How do I decide where to put the next hospital?

Biology

- DNA sequence alignment

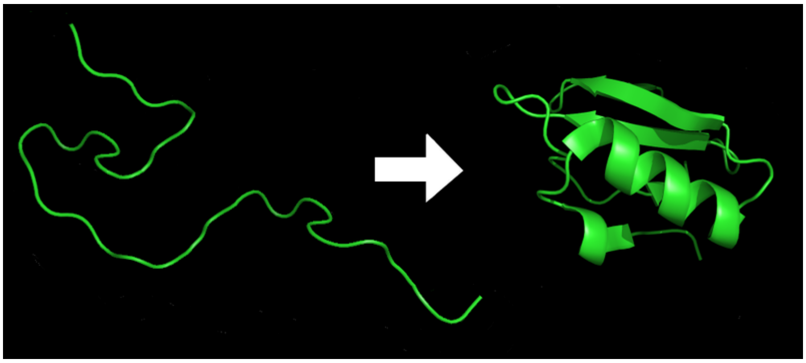




Optimal Phylogenic tree

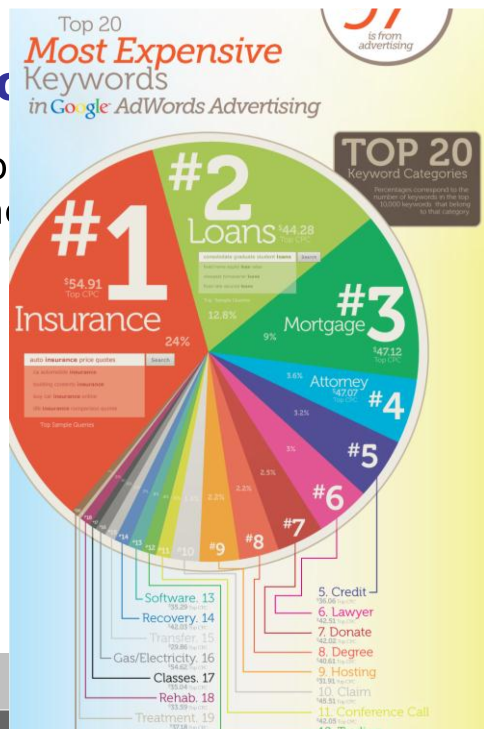
Biology

- Protein folding



Econ

- How much does Goo
keyword advertisem



Games

- What is the optimal strategy for a given game?
  - Rock / paper / scissors?
  - Chess?
    - Is there a winning strategy for white?

Computer Vision

Robotics

Optimization problems

- We've seen several optimization-type problems in this class:
  - *k*-means
  - *k*-centers
  - Fitting a line to a set of points
  - Coloring a graph with the minimum number of colors / fitting animals into the minimum set of cages
  - Finding the minimum spanning tree of a graph
  - Others?

Optimization problems

- Key elements:

1. A set of valid solutions
   - Line fitting: all possible lines
   - Minimum spanning tree: all possible spanning trees of a graph
   - *k*-means: all possible sets of *k* means

Optimization problems

- Key elements:

2. A way to measure how good/bad a solution is (an *objective function*):
   - Line fitting: sum of squared residuals
   - Minimum spanning tree: sum of edge weights
   - *k*-means: sum of squared distances from each input point to its assigned mean

   – Sometimes coming up with a good objective function is very difficult in itself (e.g., protein folding)

k-means objective function

- Find the centers that minimize the sum of squared distances to the points
- Objective function:

Given input points $x_1, x_2, x_3, \ldots, x_n$, find the clusters $C_1, C_2, \ldots C_k$ and the cluster centers $\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_k$ that minimize

$$\sum_{j=1}^{k} \sum_{x_i \in C_j} |x_i - \bar{x}_j|^2$$

Optimization algorithms

- For any given optimization *problem*, we would like to come up with a (hopefully efficient) algorithm
  - That ideally finds the global minimum of the objective function

Optimization algorithms

- Example: minimum spanning tree (MST)
  - First algorithm: Borůvka's algorithm (1926)
    - For constructing an efficient electrical network for Moravia
  - Prim's algorithm
  - Kruskal's algorithm

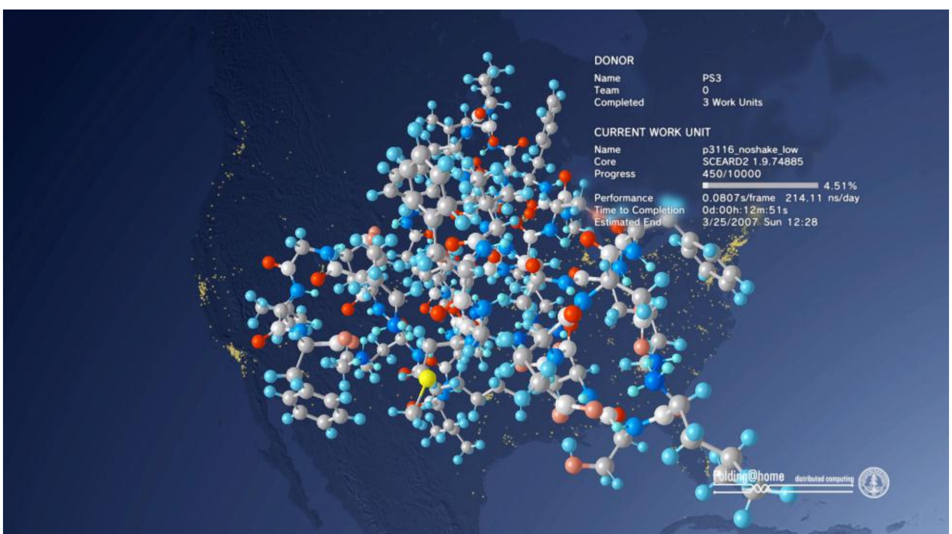
Optimization algorithms

- Possible algorithm types:
  - Greedy algorithms (e.g. Prim and Kruskal)
  - Iterative algorithms (e.g. bubble sort, gradient descent)
  - Guess-and-check (e.g. RANSAC)
- Some algorithms return the global optimum, others just return a "good" answer

Another optimization problem

- Dense box
  - One statement: Given a set of 2D points, compute a good bounding box, that is not too big and contains most of the points
  - What objective function should we use?

Parallel optimization algorithms



Human optimization algorithms

nature International weekly journal of science

Search This journa

nature.com ▸ journal home ▸ current issue ▸ letter ▸ full text

NATURE | LETTER

Predicting protein structures with a multiplayer online game

You did it!

Seth Cooper, Firas Khatib, Adrien Treuille, Janos Barbero, Jeehyung Lee, Michael Beenen, Andrew Leaver-Fay, David Baker, Zoran Popović & Foldit players

Affiliations | Contributions | Corresponding author

Nature **466,** 756–760 (05 August 2010) | doi:10.1038/nature09304

Received 22 January 2010 | Accepted 30 June 2010

Questions?