

Computational Optimization Methods



Jianlin Cheng, PhD
Computer Science Department
University of Missouri, Columbia
Fall, 2014

Overview of Syllabus

Objective I

The course covers typical computational optimization methods widely used in many computing domains, such as *bioinformatics*, *data mining* and *machine learning*. The theoretical foundation of each optimization method is rigorously studied, followed by typical real-world applications in one or more domains.

Objective II

Development of transferable professional skills

- Communication
- Information collection & problem solving
- Project management
- Presentation
- Collaboration
- Team work
- Leadership

Instructor, Office Hours, Course Web



- **Instructor:**

Prof. Jianlin Cheng (

<http://www.cs.missouri.edu/~chengji>)

Deb Bhattacharya

Office Hours

Wed & Fri, 4 pm – 5 pm



Renzhi Cao

- **Course Website**

<http://www.cs.missouri.edu/~chengji/com2014/>

Topics

- 1. Markov chain Monte Carlo methods (MCMC) and their applications in sequence motif search**
- 2. Incremental optimization methods and their applications in travel salesman problem**
- 3. Dynamic programming and its applications in graph theory and sequence alignment**
- 4. Linear and integer programming and its applications in network flow**
- 5. Quadratic programming and its applications in kernel learning methods**
- 6. Contrastive divergence optimization with applications in deep learning networks**

Assignments

- **Reading Assignments:** There is possibly one reading assignment for each of some topics. Students are required to read one paper regarding a topic and write a half-page overview of the method and application described in the paper.
- **Group Projects:** There is one group project for each of the first five topics. Under the instructor's guidance, students work in a group to design and implement one optimization method for a topic and apply it to solve one computing problem. Each group has five students.
– start to form your group!

Assignment Submission

- **Email:** mumachinelearning@gmail.com
- **Data sharing server:** one account per group on a server for sharing data and document. (maybe should use JustCloud, GoogleDrive, BOX, SourceForge, GitHub, DropBox, or other cloud storage: create your group's account).

Problem Solving Based Active Teaching and Learning

Two Phases for Each Topic

- **Phase I – Theory:** a lecture (problem, algorithms, data structures) by instructor, reviewing a paper by students, $1/4 - 1/3$ class time
- **Phase II – Practice:** Under the direction of the faculty, students will apply the techniques learned in the first phase to develop and apply a computational optimization method of each topic by working on a group software development project.

Given a topic,

In Class

After Class

Brainstorm Discussion

- Introduce of a project problem (by faculty)
- Discuss solutions and tasks by students
- Electing a coordinator of project by rotation
- Present initial Solution and tasks by coordinator

- In-depth group discussion of project plan
- Make slides for tasks and solutions of the project (algorithms, implementation, testing, task assignment)

Plan Presentation

- Presentation of Solution and Plan
 - Feedback

- *Finalize the plan*
- Start Implementation and Evaluation
- Create a topic report

Results Report

- Informal Presentation of Results
- Discussions and Feedback

- Improve implementation
- *Finalize topic project report*



Final Presentation and Report

- Cover all the topics
- All members give a final presentation together
- A final report assembling all the results and methods of all the topics together

Introduction to Computational Optimization

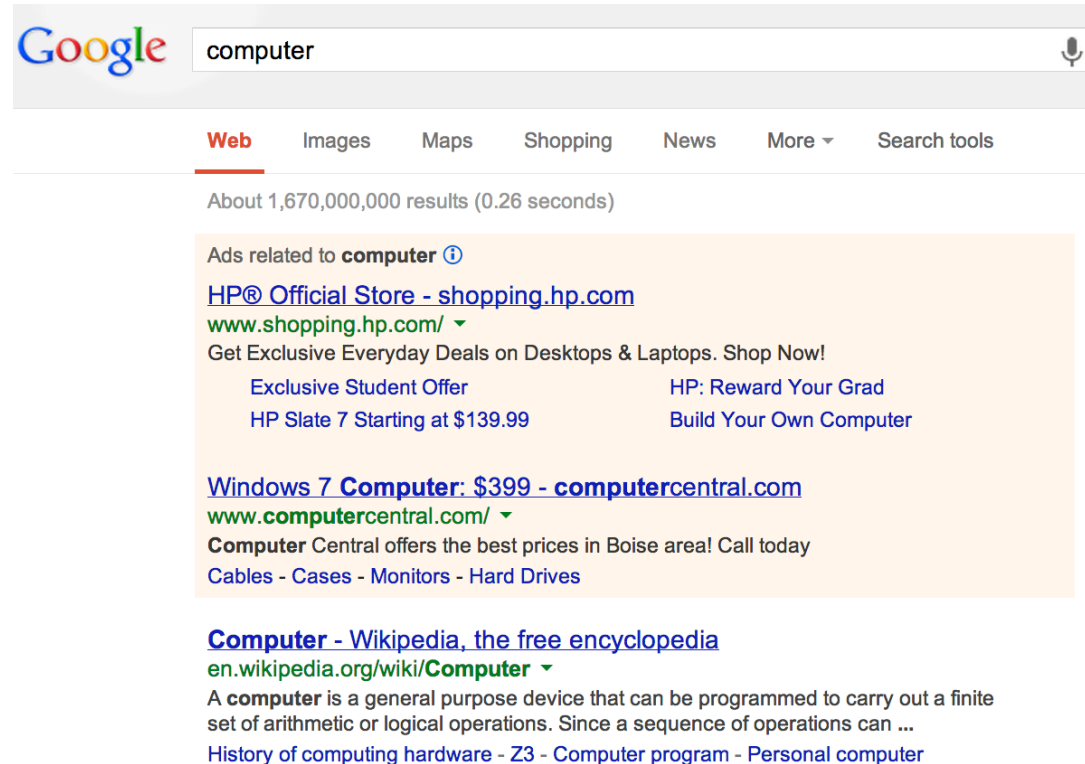
Examples of Real World Optimization Problems - 1

- Find directions
- Given a map of Europe, find the shortest driving path from Paris to Rome.



Examples of Real World Optimization Problems - 2

- A Google Interview Question
- Given a list of ads and their probability being viewed by a user and the fees offered to show them, decide which ad to show by Google?



The image shows a screenshot of a Google search for the word "computer". The search bar at the top contains the word "computer" and a microphone icon. Below the search bar, there are navigation tabs for "Web", "Images", "Maps", "Shopping", "News", "More", and "Search tools". The "Web" tab is selected. Below the tabs, it says "About 1,670,000,000 results (0.26 seconds)".

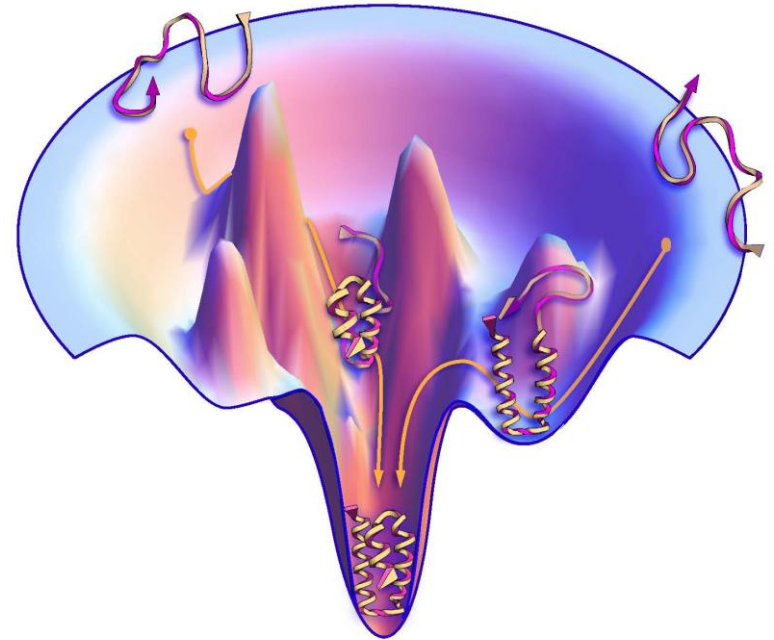
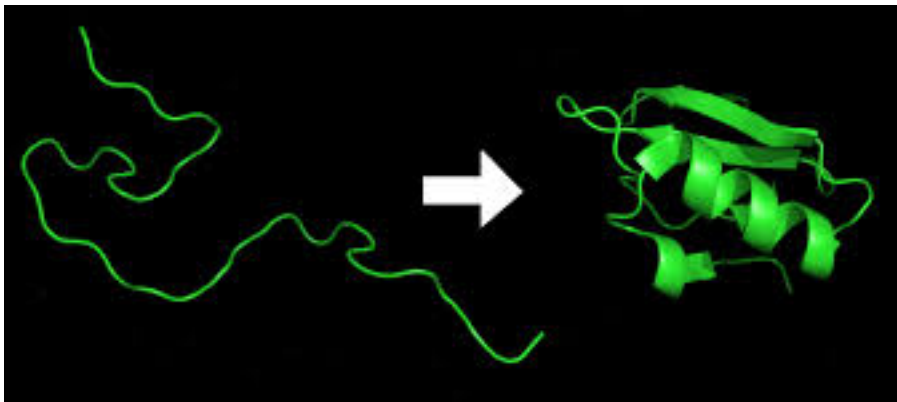
Underneath, there is a section for "Ads related to computer". The first ad is from HP's official store, with the URL www.shopping.hp.com/. The ad text says "Get Exclusive Everyday Deals on Desktops & Laptops. Shop Now!" and lists two offers: "Exclusive Student Offer" with "HP Slate 7 Starting at \$139.99" and "HP: Reward Your Grad" with "Build Your Own Computer".

The second ad is from Computer Central, with the URL www.computercentral.com/. The ad text says "Computer Central offers the best prices in Boise area! Call today" and lists "Cables - Cases - Monitors - Hard Drives".

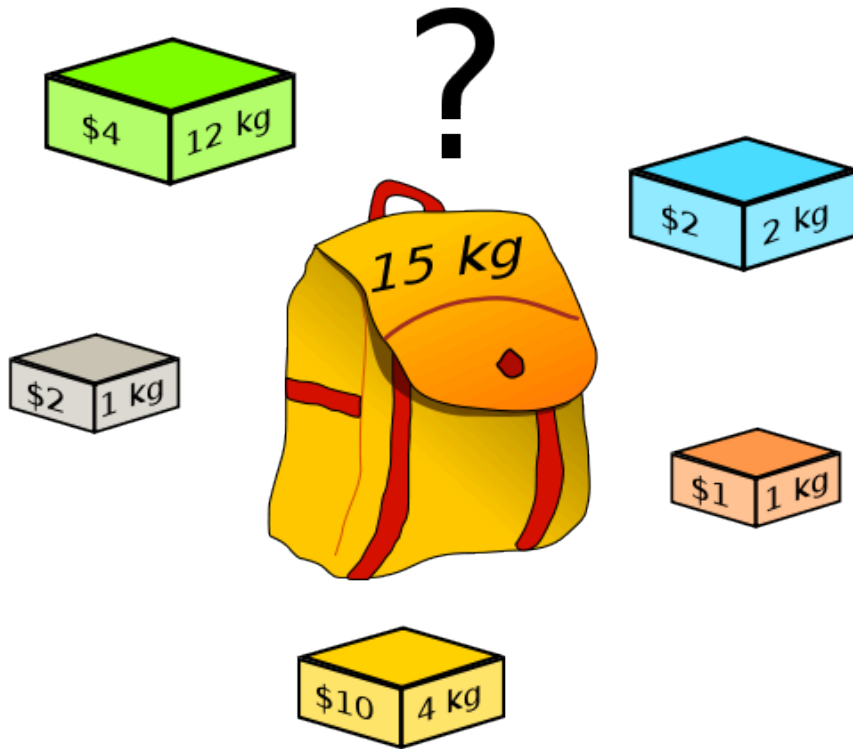
Below the ads, there is a link to the Wikipedia page for "Computer", with the URL en.wikipedia.org/wiki/Computer. The text below the link says "A computer is a general purpose device that can be programmed to carry out a finite set of arithmetic or logical operations. Since a sequence of operations can ..." and "History of computing hardware - Z3 - Computer program - Personal computer".

Examples of Real World Optimization Problems - 3

- Protein Folding
- Given an energy function, find shape / conformation of a protein with lowest energy

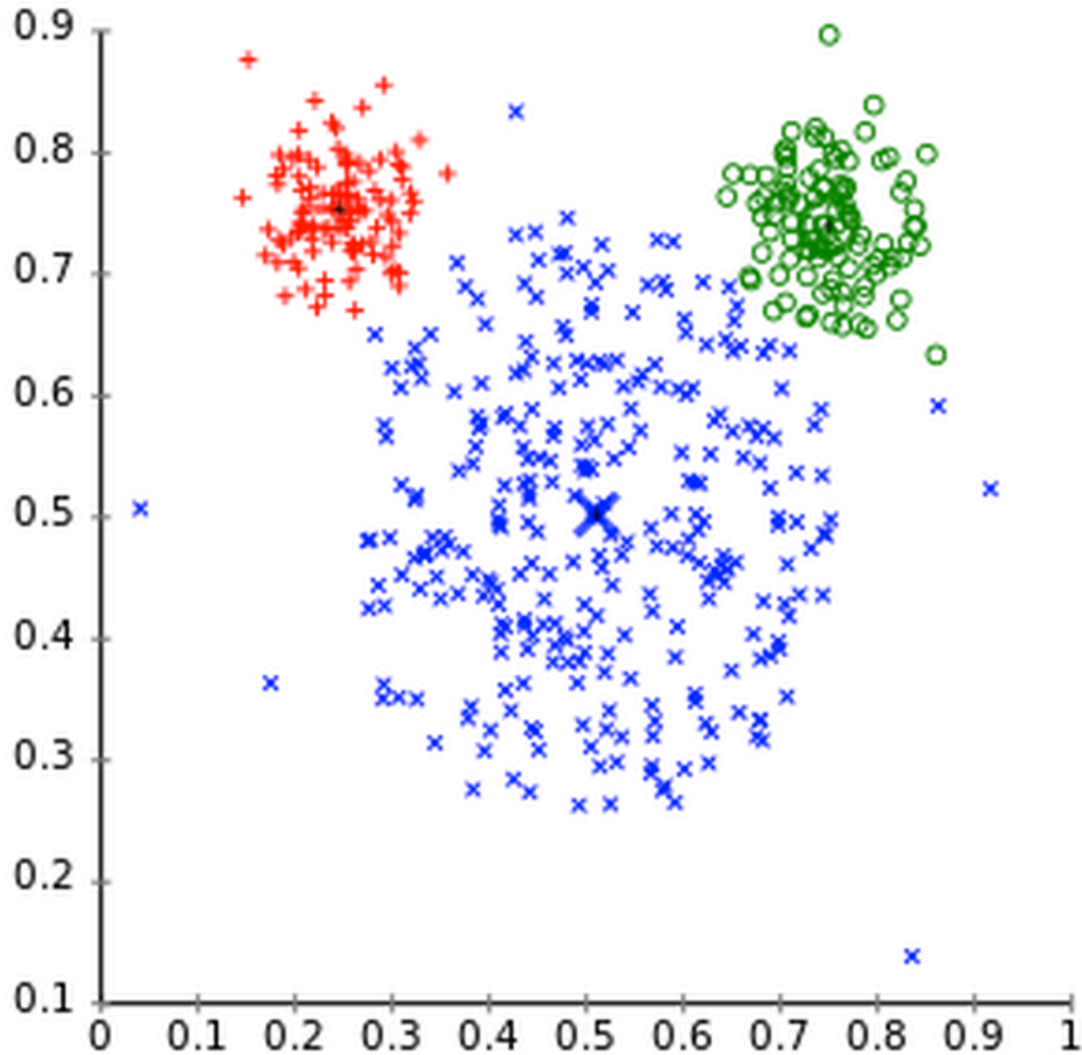


Knapsack Problem – Problem 4

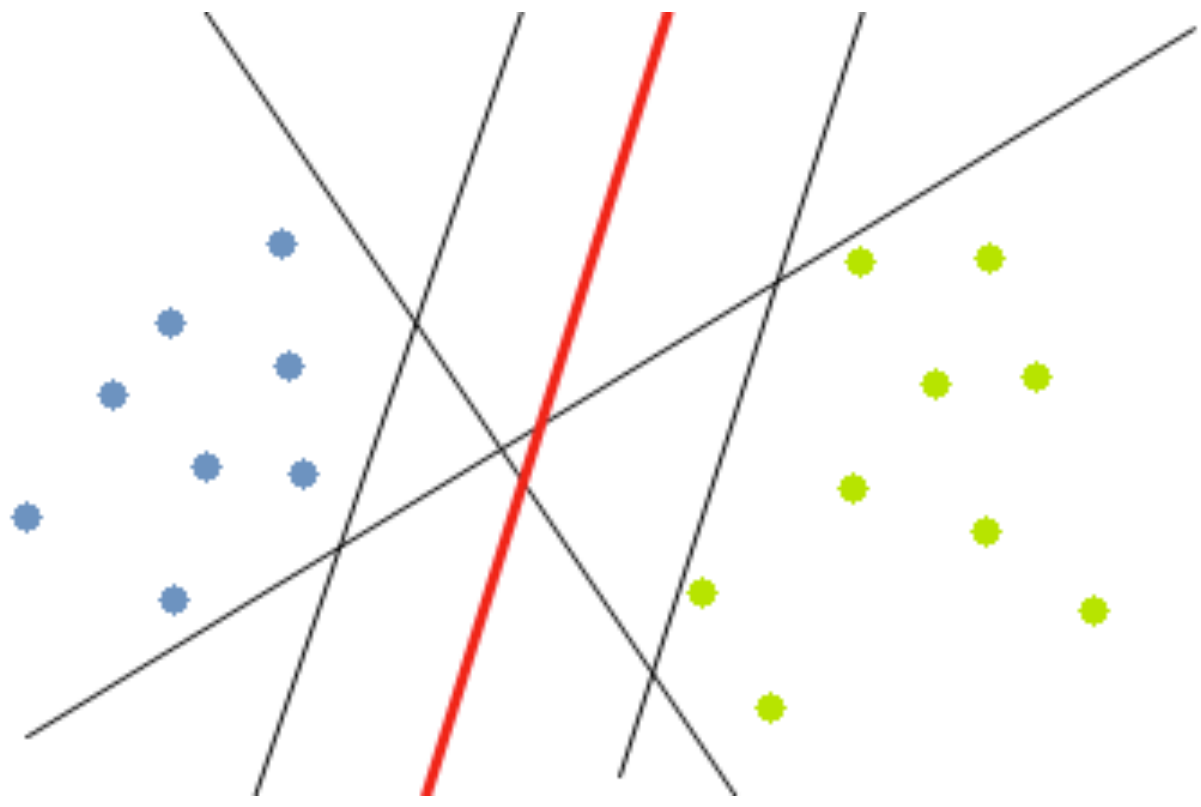


Related to investment portfolio
Maximize profit with limited resource

Problem 5 - Clustering



Problem 6 - Classification



Definition of Optimization Problems

- An objective function defining a quantity to be optimized (maximize or minimize)
- A set of variables to be changed in order to optimize the objective function
- (Optionally) a set of constraints on the variables

Formulation of An Integer Problem

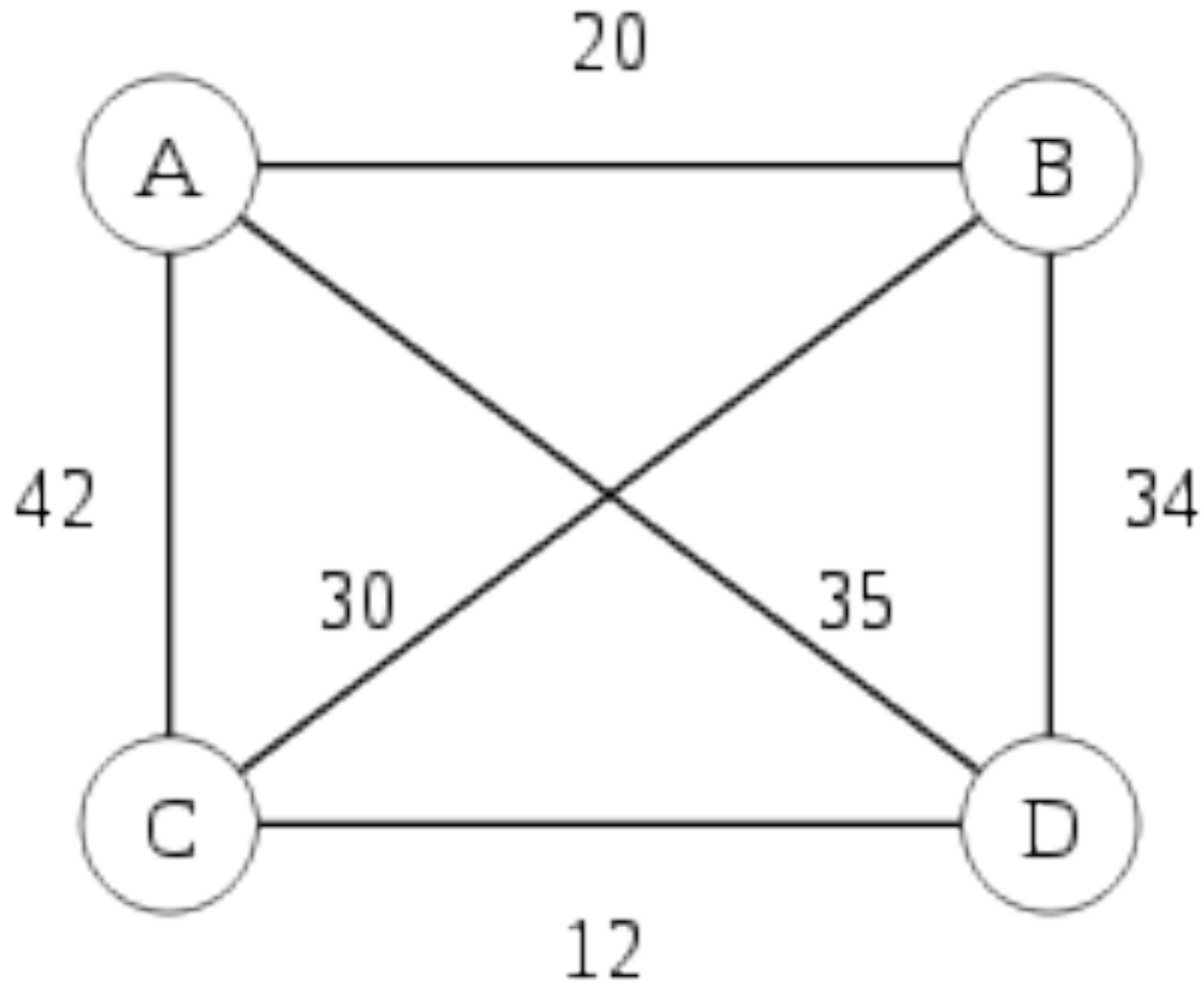
Problem

- An optimization problem in which some or all variables are restricted to be integers.
- Traveling salesman problem: Given a list of cities and the distances between each pair of cities, what is the shortest possible route that visits each city exactly once and returns to the origin city?

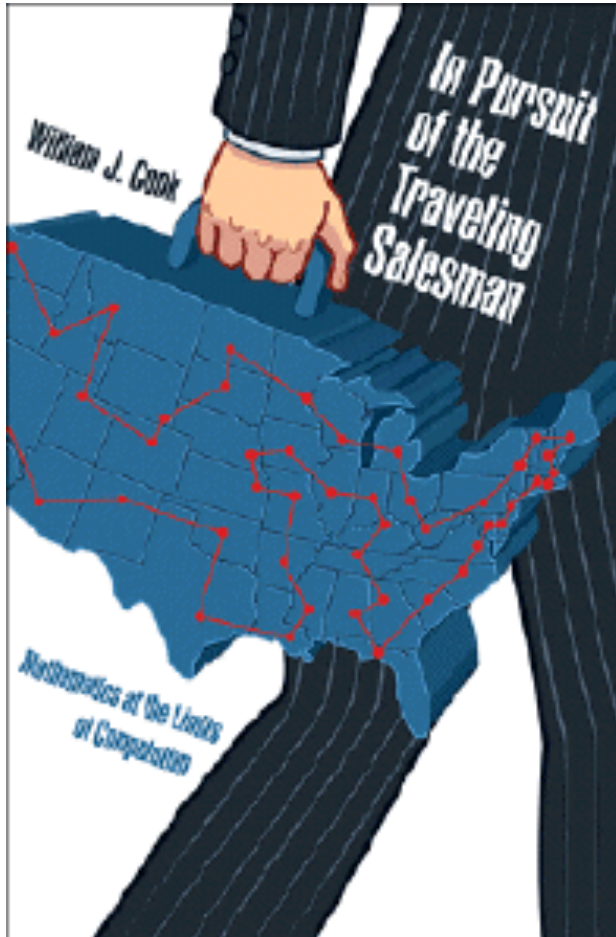


William Rowan Hamilton

Graph Visualization of a Toy Example



Informal Definition



Objective:

Minimize the total distance of the tour over all the cities

Variables:

take a path between any two cities or not (0/1)

Constraints:

one city can be visited only once and all the selected paths form a round tour

Demo of TSP Problem

- <https://www.youtube.com/watch?v=RtqoGeXDDbQ&feature=youtu.be>