Complexity and Advanced Algorithms
Monsoon 2011

A Last Overview

Overview

- Important items we could not cover
- Activities and opportunities@CSTAR

Approximation Algorithms

- Of late, most problems have become tough enough that finding an exact solution is a near-impossibility.
  - Think of also applied areas such as speech recognition, handwriting recognition, computer vision, ...
- Many interesting and practical problems are quickly turning out to be either in NP or beyond.
  - Games, IBM Watson, ...

Approximation Algorithms

- But practice demands some solution at least!
- Enter approximation algorithms.

- Let P be a maximization optimization problem.
- Let OPT(P) be the value of the best possible solution. Let A(P) be the value of the solution produced by an algorithm A.
- Algorithm A is called a c-approximation to P if the ratio OPT(P)/A(P) is at most c for all instances of P.
  - Consider the ration A(P)/OPT(P) if P is a minimization problem.

Quick Example

- Given a graph G = (V, E), a subset of vertices U such that every edge in E has at least one endpoint in U.
- The vertex cover problem is to find a cover of the smallest size.
- Finds importance in several problems with wireless networks.
- Also, one of the first problems shown to be NP-complete.

An Easy Approximation Algorithm

- Known before even the term approximation algorithms was coined.
- Compute a maximal matching of G.
- A matching is a subset M of edges so that no two edges in M share any common end point.
  - A matching M is maximal if no proper superset of M is also a matching.

An Easy Approximation Algorithm

- Take all the endpoints of the edges in M as a vertex cover.
- Two claims to be shown.

An Easy Approximation Algorithm

- Claim 1. The proposed set is a cover.
- Proof: Otherwise, M is not a maximal matching.

- Claim 2. The proposed cover is at most twice the size of the best possible vertex cover.
- Proof: Let OPT be the best possible vertex cover.
- Then, OPT must cover all end points of edges in M.
- OPT has to include at least one endpoint for every edge in M.
- So, |OPT| is at least |M|.

An Easy Approximation Algorithm

- On the other hand, the size of our cover is exactly 2|M|.

- So, the ratio of |A(P)|/|OPT(P)| is at most 2.

Approximation

- More is known about approximation algorithms in general.
- Good approximation algorithms exist for several NP-complete problems such as TSP, Steiner trees, facility location, ...
- A rich complexity flavor too.

Yet Another Category

- Rent or Buy?
- You want to use a facility by renting it out. Costs R rupees each time you rent. You may instead own the facility at an initial cost of S rupees.
- You do not know apriori how many times you would use the facility.
- But, you are always looking to minimize the overall cost incurred.

Yet Another Category

- Can algorithmically study when to rent and when to buy.
- Can also compare with a clairvoyant solution.
- Again, the goal is to perform as close to possible as the clairvoyant solution.
- In this case, rent for S/R times and then buy.
- The overall cost is 2S.
- Can show that this is best possible.

Rent or Buy?

- Let there be k usages of the facility.
- Case k at least S/R : A clairvoyant strategy would immediately buy the facility at cost S.
- The cost incurred by our strategy is R. S/R + S = 2S.

- Case k < S/R : The clairvoyant strategy will just rent the facility at a cost of kR. Equals our cost.

Rent or Buy?

- This simple technique has found application also in power management in computer systems.
- The underlying question is when should your computer go to a sleep state?

A CS View of Life

- Suppose you park your vehicle in a large parking lot and forget where your parked.
- The lot has some geometry, say concentric grids from the building entrance.
- How do you locate your vehicle while spending the least possible energy in the worst case?
- Idea: Try an ever increasing grid in each round.
- Can show that the strategy is very competitive compared to a strategy which knows where the vehicle is.
  - So, can afford to forget things :-)

Online Algorithms

- This field of algorithms is called online algorithms.
- Can be used in several other interesting settings
  - Cache page replacement
  - K-server problems
  - ...

Work and Opportunities@CSTAR

- A listing by research subfields within algorithms is what is shown here.
- Other areas such as information security not included.

Work@CSTAR

- Resource budgeted distributed algorithms
  - Think of any problem P to be solved in the distributed setting.
  - The algorithm requires some r(n) rounds, using M(n) messages, and requiring A(n) auxiliary resources.
  - Suppose we have a budget on any of these, say only r'(n) rounds are allowed.
  - Can we still find a solution? If so, what should we trade off?

Work @CSTAR

- A more concrete example
  - Let P be graph coloring. A(n) could be the number of colors used.
  - We have shown recently that A(n) can be expressed as a function of the number of rounds r(n) used by a distributed coloring algorithm.
- Other problems that can be attacked in this framework
  - Message broadcast
  - Symmetry breaking
  - Spanning trees and other structures.

Work@CSTAR

- Fault tolerance in distributed computing
  - Given a fault model, can a particular problem be solved in the distributed setting?
  - Fault models typically indicate the nature and extent of the fault.
  - Model practical issues such as routers in the Internet failing, nodes in a sensor network not alive, functional units in a chip that have errors, ...
  - The question therefore is, what kind of faults can an algorithm tolerate.

Work@CSTAR

- Parallel Computing
  - Efficient algorithmics on current architectures such as GPUs and combinations of GPUs and CPUs.
  - Have several recent successes on a variety of problems.
    - Fastest list ranking in practice
    - Fastest routine to multiply a sparse matrix with a vector
    - Fastest way to generate (pseudo) random numbers
    - Fastest graph traversal and spanning tree
    - Several others....

Work@CSTAR

- Several interesting works in the pipeline.
  - Multiplying sparse matrices
  - Sorting, histogram, ...
  - Image processing and graphics
  - ...

Work@CSTAR

- Data Structures for querying and reporting.
- Recall the range minima problem.
- The idea is to preprocess a given data set to quickly handle queries on the data set.
  - To answer a query, time (poly)-logarithmic plus the size of the output is ideal.
  - In symbols, $O(\log^c n + k)$ for a constant c.
- Several applications
  - List all facilities in a given area.
  - List all facilties in a given area by their category.
  - List the top-k facilities in a given area, by their category.
  - ...

Work@CSTAR

- Offering a follow-up course titled "Topics in Distributed Algorithms", jointly by Dr Srinathan and self.
- Can help you build background on several research topics.
- We have open student research positions in all these areas, plus more.

- Reminder: Friday is a review lecture. So, please send in any questions you have.