Reflections on Approximation
or
How I (almost) finished my PhD thesis

Mihalis Yannakakis

Stanford University

Some History …

NP-completeness

- Cook, The complexity of theorem-proving procedures, STOC 1971
- Karp, Reducibility among combinatorial problems, 1972

NP-completeness

- Satisfiability
- Many optimization problems
  - Max Clique / Independent Set
  - Min Node Cover; Set Cover
  - Min Graph Coloring
  - Traveling Salesman Problem
  - ….
- In the following years, tons of other NP-complete problems

Approximation Algorithms

- Ullman, The performance of memory allocation algorithms, Princeton U. TR 100, 1971.
- Garey, Graham, Ullman: Worst-case analysis of memory allocation algorithms, STOC 1972

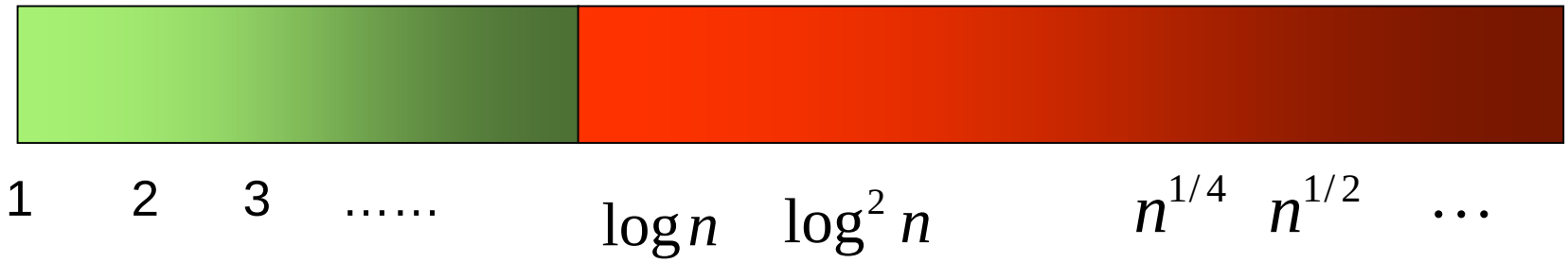
$$\text{approximation ratio} = \frac{\text{approximate cost}}{\text{optimal cost}}$$

Bin Packing

First Fit algorithm: ratio 17/10

Approximability spectrum

approximation ratio



All kinds of different behavior: For some problems:

- can get any ratio arbitrarily close to 1 (PTAS)
- others, can get some constant ratio, but no better
- others, can get no constant ratio unless P=NP

Picture ca 1976: Many unknowns

- Node Cover: can get 2; better? 1+ε?
- Metric TSP: can get 3/2 ; better? 1+ε?
- Set Cover: can get log n ; constant ?
- Clique: either can get any constant 1+ε arbitrarily close to 1 or no constant at all
- Graph Coloring: can't get ratio 2 ; some constant?

Theory of approximation?

- NP-completeness theory: ties together different hard computational problems
- Analogue for approximation?

Here's the material you gave me on reductivities. I have very little in the way of technical errors. The main problems are with organization and exposition. You need to define terms that are apparently left undefined. You need examples. Most importantly, you need a uniform framework in which to couch your results. Let's talk about what that might be (after generals if you prefer).

uniform framework?

Maximum Subgraph & Minimum Node Deletion Problems

- Find maximum number of nodes which induces a subgraph that satisfies a desired property π
- Find minimum number of nodes whose deletion leaves a subgraph that satisfies desired property π

- Examples:

  π = no edges • Max Independent Set, Min Node Cover

  π = clique • Clique problem

  acyclic • Feedback Node Set problem

  planar

  bipartite, …

Hereditary Properties

- Inherited by subgraphs: If a graph satisfies the property, then deleting nodes does not destroy the property

- For every nontrivial hereditary property, the maximum subgraph and minimum node deletion problems are NP-hard

Approximation?

- Complementary problems (max subgraph – min node deletion problem for same property π) are equivalent in terms of exact optimization, but not in terms of approximation.

- Node Cover has ratio 2 approximation does not mean that Max Independent Set (Clique) does too.

Connected version

- If we require the remaining subgraph to be a *connected* graph that satisfies property π, then the node deletion problem is not approximable within ratio $n^{1-\varepsilon}$ unless P=NP, for any hereditary property π that is "determined by the biconnected components"

  (eg. tree, planar, bipartite – but not clique).

Problem traces back to Satisfiability

- Max Independent set = max “consistent” set, with pairwise consistency constraints (edges=conflicts)
- Other properties – more complex consistency constraints

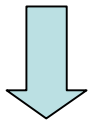
- Satisfiability: pick for each clause a literal (a way of satisfying the clause), so that the local choices are mutually consistent

- Need a gap in Satisfiability:

  Can we tell apart formulas in which we can satisfy all (or most) clauses from those where we can satisfy much fewer?

and further back to computations, proofs

- Computation of nondeterministic machine on given input, written down step-by-step



Satisfiability formula which checks that the steps of the computation satisfy all the rules of the Turing machine and that they end in acceptance

- Fragility of computations
- Robust version of NP ?

MAX SNP [PY'88]

- Class of optimization problems
- Defined logically
- Max-3SAT complete,

  various others problems complete or hard

  → if Max-3SAT has an approximation barrier, then so do they

  (Node Cover, Metric TSP, Clique , …)

Probabilistically Checkable Proofs (PCP) [ALMSS'92, AS, FGLS]

- Robust version of Cook's theorem
  - NP-hard to approximate Max-3SAT within some constant
    - yes instances • satisfiable formulas
    - no instances • 'very' unsatisfiable formulas
- Robust proofs (of yes instances)
  - Can check few random places in proof and gain high confidence on its correctness

- Opened door to many inapproximability results (Clique, Coloring, Set Cover, ….)

Maximum Subgraph Problems

- For every nontrivial hereditary property π we cannot approximate the maximum π subgraph problem within any constant factor (and much worse) (LY'93)

Node Deletion Problems

(the last chapter)

- All at least as hard to approximate as Node Cover
  → all have a barrier
- Picture is still quite blurry

  π = acyclic (Feedback Node Set)
  - Undirected graphs : ratio 2 [BBF'95]
  - Directed graphs: logn loglogn [S'95] ; better? constant?

  π = bipartite
  - logn (GVY'95); better? constant?
    (related to multicommodity cut / flow problems)

- What's missing?