Smoothed Analysis of Algorithms Part II: Binary and Multiobjective Optimization

Heiko Röglin Department of Computer Science



16 January 2013

Outline

Binary Optimization Problems When does a binary optimization problem have polynomial smoothed complexity?

Multiobjective Optimization

How many Pareto-optimal solutions do usually exist?

Conclusions

Outline

Binary Optimization Problems When does a binary optimization problem have polynomial smoothed complexity?

Multiobjective Optimization How many Pareto-optimal solutions do usually ex

Conclusions

Model

Linear Binary Optimization Problem

- set of feasible solutions S ⊆ {0, 1}ⁿ
 solution x = (x₁,..., x_n) ∈ S consists of *n* binary variables
- linear objective function max $c^T x = c_1 x_1 + \dots + c_n x_n$

Model

Linear Binary Optimization Problem

- set of feasible solutions S ⊆ {0, 1}ⁿ
 solution x = (x₁,..., x_n) ∈ S consists of *n* binary variables
- linear objective function max $c^T x = c_1 x_1 + \dots + c_n x_n$

S can encode arbitrary combinatorial structure, e.g., for a given graph, all paths from *s* to *t*, all Hamiltonian cycles, all spanning trees, ...



• Knapsack Problem: variable $x_i \in \{0, 1\}$ for each item i $S = \{x \mid w_1x_1 + \cdots + w_nx_n \le t\}$



• TSP: variable $x_e \in \{0, 1\}$ for each $e \in E$ $S = \{x \mid x \text{ encodes Hamiltonian cycle}\}$

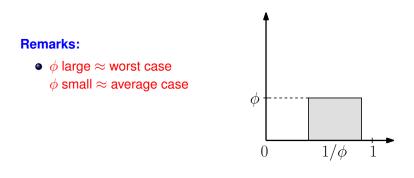
Worst-case Analysis: Adversary chooses S and $c_i \in [-1, 1]$.

Worst-case Analysis: Adversary chooses S and $c_i \in [-1, 1]$.

Smoothed Analysis:

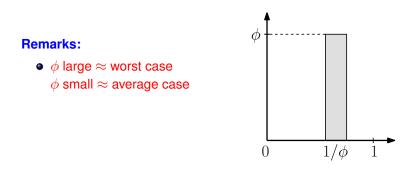
Worst-case Analysis: Adversary chooses S and $c_i \in [-1, 1]$.

Smoothed Analysis:



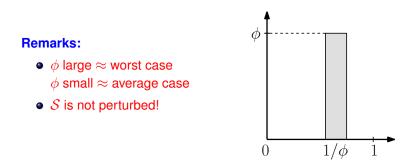
Worst-case Analysis: Adversary chooses S and $c_i \in [-1, 1]$.

Smoothed Analysis:



Worst-case Analysis: Adversary chooses S and $c_i \in [-1, 1]$.

Smoothed Analysis:



Theorem [Beier, Vöcking (STOC 2004)] linear binary opt. problem has polynomial smoothed complexity \iff pseudo-polynomial time $poly(n, max\{|c_i|\})$ in the worst case Theorem [Beier, Vöcking (STOC 2004)] linear binary opt. problem has polynomial smoothed complexity \iff pseudo-polynomial time $poly(n, max\{|c_i|\})$ in the worst case



- Knapsack Problem: Can be solved in time $O(n^2 P)$, where P is the largest profit.
 - \Rightarrow polynomial smoothed complexity

Theorem [Beier, Vöcking (STOC 2004)] linear binary opt. problem has polynomial smoothed complexity \iff pseudo-polynomial time $poly(n, max\{|c_i|\})$ in the worst case



- Knapsack Problem: Can be solved in time $O(n^2 P)$, where *P* is the largest profit.
 - \Rightarrow polynomial smoothed complexity



- TSP: strongly NP-hard (even if all edge lengths are 1 or 2)
 - \Rightarrow no polynomial smoothed complexity

- A =algorithm
- \mathcal{I}_n = set of inputs of length *n*
- $per_{\phi}(I) = perturbation of instance I$
- $T_A(I)$ = running time of A on instance I

- *A* = algorithm
- \mathcal{I}_n = set of inputs of length *n*
- $per_{\phi}(I) = perturbation of instance I$
- $T_A(I) =$ running time of A on instance I

Definition (first attempt):

Polyn. smoothed compl. $\iff \max_{I \in \mathcal{I}_n} \mathsf{E} \big[\mathcal{T}_A(\operatorname{per}_{\phi}(I)) \big] = \operatorname{poly}(n, \phi)$

- *A* = algorithm
- \mathcal{I}_n = set of inputs of length *n*
- $per_{\phi}(I) = perturbation of instance I$
- $T_A(I) =$ running time of A on instance I

Definition (first attempt):

Polyn. smoothed compl. $\iff \max_{l \in \mathcal{I}_n} \mathsf{E}[T_A(\operatorname{per}_{\phi}(l))] = \operatorname{poly}(n, \phi)$ Problem: Not robust against change of machine model.

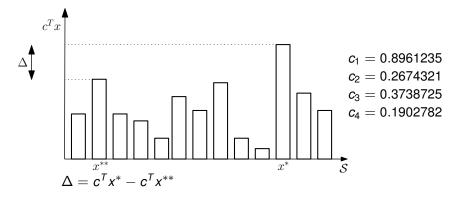
- A =algorithm
- \mathcal{I}_n = set of inputs of length *n*
- $per_{\phi}(I) = perturbation of instance I$
- $T_A(I) =$ running time of A on instance I

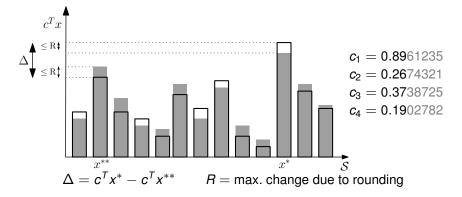
Definition (first attempt):

Polyn. smoothed compl. $\iff \max_{l \in \mathcal{I}_n} \mathsf{E} [T_A(\operatorname{per}_{\phi}(l))] = \operatorname{poly}(n, \phi)$ Problem: Not robust against change of machine model.

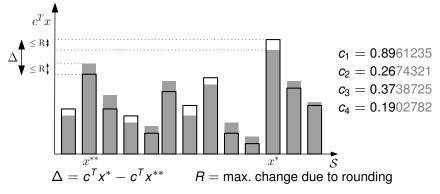
Definition

Algorithm *A* has polynomial smoothed complexity if there exist $\alpha > 0$ and $\beta > 0$ with $\max_{l \in \mathcal{I}_n} \mathbf{E} \left[T_A(\operatorname{per}_{\phi}(l))^{\alpha} \right] \leq \beta n \phi.$

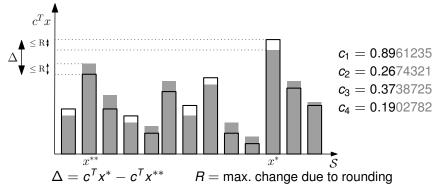




Idea: Round coefficients c_i and apply pseudo-polyn. algo



• Rounding after the *b*-th bit $\Rightarrow |c_i - [c_i]| \le 2^{-b}$ $\Rightarrow \forall x \in S : |c^T x - [c]^T x| \le n2^{-b} = R$



- Rounding after the *b*-th bit $\Rightarrow |c_i [c_i]| \le 2^{-b}$ $\Rightarrow \forall x \in S : |c^T x - [c]^T x| \le n2^{-b} = R$
- $\Delta > 2R \Rightarrow$ rounding does not change optimal solution

For all S and all densities $f_i : [-1, 1] \rightarrow [0, \phi]$

 $\Pr[\Delta < \varepsilon] \le 2n\phi\varepsilon.$

For all S and all densities $f_i : [-1, 1] \rightarrow [0, \phi]$

$$\Pr[\Delta < \varepsilon] \le 2n\phi\varepsilon.$$

Corollary

For every
$$p \in (0, 1]$$
 and $b \ge \log \left(\frac{n^2 \phi}{p}\right) + 2$,

Pr[rounding changes optimal solution] $\leq p$.

For all S and all densities $f_i : [-1, 1] \rightarrow [0, \phi]$

$$\Pr[\Delta < \varepsilon] \le 2n\phi\varepsilon.$$

Corollary

For every
$$p \in (0, 1]$$
 and $b \ge \log \left(\frac{n^2 \phi}{p}\right) + 2$,

Pr[rounding changes optimal solution] $\leq p$.

pseudo-polynomial algorithm \Rightarrow polynomial smoothed complexity

For all S and all densities $f_i : [-1, 1] \rightarrow [0, \phi]$

$$\Pr[\Delta < \varepsilon] \le 2n\phi\varepsilon.$$

Corollary

For every
$$p \in (0, 1]$$
 and $b \ge \log \left(\frac{n^2 \phi}{p}\right) + 2$,

Pr[rounding changes optimal solution] $\leq p$.

pseudo-polynomial algorithm \Rightarrow polynomial smoothed complexity

 Round coefficients after a logarithmic number of bits and call pseudo-polynomial algorithm.

For all S and all densities $f_i : [-1, 1] \rightarrow [0, \phi]$

$$\Pr[\Delta < \varepsilon] \le 2n\phi\varepsilon.$$

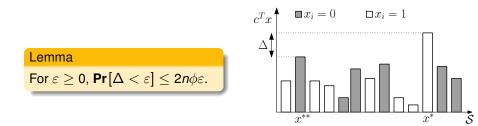
Corollary

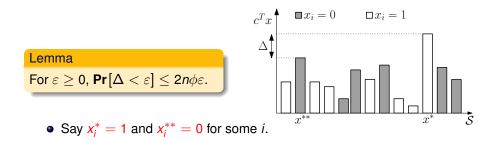
For every
$$p \in (0, 1]$$
 and $b \ge \log \left(\frac{n^2 \phi}{p}\right) + 2$,

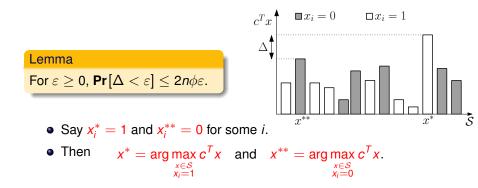
Pr[rounding changes optimal solution] $\leq p$.

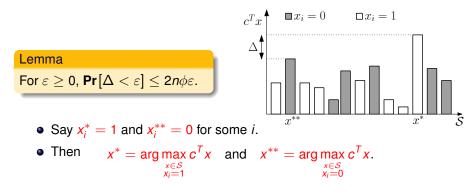
pseudo-polynomial algorithm \Rightarrow polynomial smoothed complexity

- Round coefficients after a logarithmic number of bits and call pseudo-polynomial algorithm.
- If necessary, increase precision and repeat.

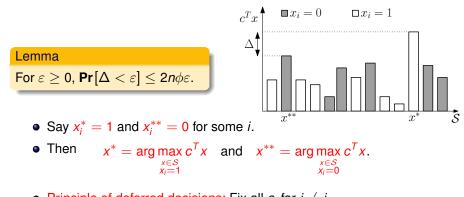




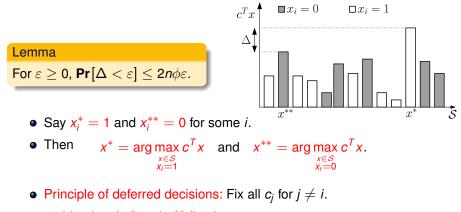




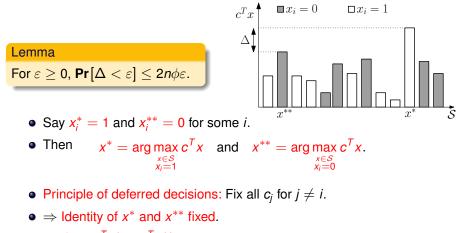
• Principle of deferred decisions: Fix all c_j for $j \neq i$.



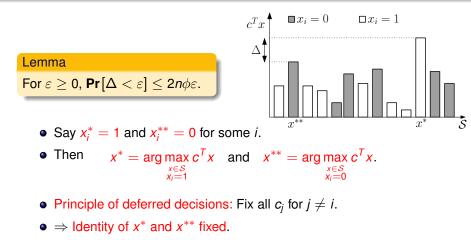
- Principle of deferred decisions: Fix all c_j for $j \neq i$.
- \Rightarrow Identity of x^* and x^{**} fixed.



- \Rightarrow Identity of x^* and x^{**} fixed.
- $\Rightarrow \Delta = c^T x^* c^T x^{**} = \kappa + c_i$ for constant κ



- $\Rightarrow \Delta = c^T x^* c^T x^{**} = \kappa + c_i$ for constant κ
- $\Pr[\Delta \in [0, \varepsilon])] = \Pr[c_i \in [-\kappa, -\kappa + \varepsilon])] \le \varepsilon \phi$



- $\Rightarrow \Delta = c^T x^* c^T x^{**} = \kappa + c_i$ for constant κ
- $\Pr[\Delta \in [0, \varepsilon])] = \Pr[c_i \in [-\kappa, -\kappa + \varepsilon])] \le \varepsilon \phi$
- Union Bound over all *n* choices for *i*.

Theorem [Beier, Vöcking (STOC 2004)] linear binary opt. problem has polynomial smoothed complexity \iff pseudo-polynomial time poly $(n, \max\{|c_i|\})$ in the worst case

[Beier, Vöcking (STOC 2004)]

Theorem remains true if linear constraints are perturbed.

[R., Vöcking (IPCO 2005)]

Theorem remains true for integer optimization problems.

Outline

Binary Optimization Problems When does a binary optimization problem have polynomial smoothed complexity?

Multiobjective Optimization

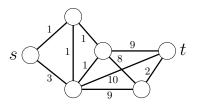
How many Pareto-optimal solutions do usually exist?

Conclusions

Optimization Problems

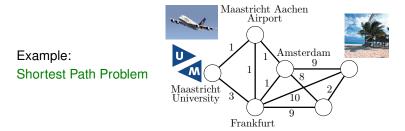
Single-criterion Optimization Problem: min f(x) subject to $x \in S$.

Example: Shortest Path Problem



Optimization Problems

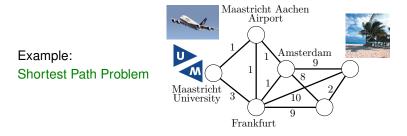
Single-criterion Optimization Problem: min f(x) subject to $x \in S$.



Real-life logistical problems often involve multiple objectives. (travel time, fare, departure time, etc.)

Optimization Problems

Single-criterion Optimization Problem: min f(x) subject to $x \in S$.



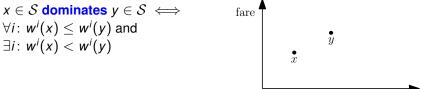
Real-life logistical problems often involve multiple objectives. (travel time, fare, departure time, etc.)

Multiobjective Opt. Problem: min $f_1(x), \ldots, \min f_d(x)$ s.t. $x \in S$. Usually, there is no solution that is simultaneously optimal for all f_i .

Question

What can we do algorithmically to support the decision maker?

Multiobjective Opt. Problem: min $w^1(x), \ldots$, min $w^d(x)$ s.t. $x \in S$

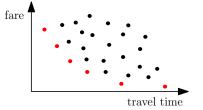


travel time

Multiobjective Opt. Problem: min $w^1(x), \ldots, \min w^d(x)$ s.t. $x \in S$

 $x \in S$ dominates $y \in S \iff$ $\forall i: w^i(x) \le w^i(y)$ and $\exists i: w^i(x) < w^i(y)$

 $x \in S$ Pareto-optimal \iff $\exists y \in S: y$ dominates x



Multiobjective Opt. Problem: min $w^1(x), \ldots, \min w^d(x)$ s.t. $x \in S$

 $x \in S$ dominates $y \in S \iff$ $\forall i: w^i(x) \le w^i(y)$ and $\exists i: w^i(x) < w^i(y)$

 $x \in S$ Pareto-optimal \iff $\exists y \in S: y$ dominates x fare travel time

Often the Pareto curve is generated:

- Pareto curve limits options for decision maker.
- Monotone functions are optimized by Pareto-optimal solutions, e.g., $\lambda_1 w^1(x) + \ldots + \lambda_d w^d(x)$ or $w^1(x) \cdots w^d(x)$.
- Tool for solving single-criterion problems

Multiobjective Opt. Problem: min $w^1(x), \ldots, \min w^d(x)$ s.t. $x \in S$

 $x \in S$ dominates $y \in S \iff$ $\forall i: w^i(x) \le w^i(y)$ and $\exists i: w^i(x) < w^i(y)$

 $x \in S$ **Pareto-optimal** \iff $\exists y \in S: y$ dominates x fare travel time

Often the Pareto curve is generated:

- Pareto curve limits options for decision maker.
- Monotone functions are optimized by Pareto-optimal solutions, e.g., $\lambda_1 w^1(x) + \ldots + \lambda_d w^d(x)$ or $w^1(x) \cdots w^d(x)$.
- Tool for solving single-criterion problems

Central Question

How large is the Pareto curve?

Model

Linear Binary Optimization Problem

- set of feasible solutions S ⊆ {0, 1}ⁿ
 solution x = (x₁,..., x_n) ∈ S consists of *n* binary variables
- *d* linear objective functions: $\forall i \in \{1, \dots, d\}$: min $w^i(x) = w_1^i x_1 + \dots + w_n^i x_n$

Model

Linear Binary Optimization Problem

- set of feasible solutions S ⊆ {0, 1}ⁿ
 solution x = (x₁,..., x_n) ∈ S consists of *n* binary variables
- *d* linear objective functions: $\forall i \in \{1, \dots, d\}$: min $w^i(x) = w_1^i x_1 + \dots + w_n^i x_n$

How large is the Pareto curve?

- Exponential in the worst case for almost all problems.
- In practice, often few Pareto optimal solutions.



Example: Train Connections

w.r.t. travel time, fare, number of train changes [Müller-Hannemann, Weihe 2001]

Results (Bicriteria Optimization)

Adversary chooses S and a probability density $f_j^i : [-1, 1] \rightarrow [0, \phi]$ for every w_j^i and some $\phi \ge 1$. Every w_j^i is **drawn independently** according to f_j^i .

 $P_d(n, \phi) = \max_{\mathcal{S}, f_j^i} \mathbf{E} \left[\text{number of Pareto-optimal sol. for } \mathcal{S} \text{ and } f_j^i \right]$

Results (Bicriteria Optimization)

Adversary chooses S and a probability density $f_j^i : [-1, 1] \rightarrow [0, \phi]$ for every w_j^i and some $\phi \ge 1$. Every w_i^i is **drawn independently** according to f_i^i .

 $P_d(n, \phi) = \max_{\mathcal{S}, f_j^i} \mathbf{E} \left[\text{number of Pareto-optimal sol. for } \mathcal{S} \text{ and } f_j^i \right]$

Bicriteria Optimization (d = 2):

Theorem [Beier, Vöcking (STOC 2003)]

 $P_2(n,\phi) = O(n^4\phi)$ $P_2(n,\phi) = \Omega(n^2)$

Results (Bicriteria Optimization)

Adversary chooses S and a probability density $f_j^i : [-1, 1] \rightarrow [0, \phi]$ for every w_j^i and some $\phi \ge 1$. Every w_i^i is **drawn independently** according to f_i^i .

 $P_d(n, \phi) = \max_{\mathcal{S}, f_j^i} \mathbf{E} \left[\text{number of Pareto-optimal sol. for } \mathcal{S} \text{ and } f_j^i \right]$

Bicriteria Optimization (d = 2):

Theorem [Beier, Vöcking (STOC 2003)]

 $P_2(n,\phi) = O(n^4\phi)$ $P_2(n,\phi) = \Omega(n^2)$

Theorem [Beier, R., Vöcking (IPCO 2007)]

 $P_2(n,\phi) = O(n^2\phi)$

extension to integer optimization problems

Multiobjective Optimization (d arbitrary constant):

Theorem [R., Teng (FOCS 2009)]

 $P_d(n, \phi) = O((n\phi)^{h(d)})$ for some function h

Multiobjective Optimization (d arbitrary constant):

Theorem [R., Teng (FOCS 2009)]

 $P_d(n, \phi) = O((n\phi)^{h(d)})$ for some function h

Theorem [Moitra, O'Donnell (STOC 2011)]

 $P_d(n,\phi) = O(n^{2d}\phi^{\Theta(d^2)})$

Multiobjective Optimization (d arbitrary constant):

Theorem [R., Teng (FOCS 2009)]

 $P_d(n, \phi) = O((n\phi)^{h(d)})$ for some function h

Theorem [Moitra, O'Donnell (STOC 2011)]

 $P_d(n,\phi) = O(n^{2d}\phi^{\Theta(d^2)})$

Theorem [Brunsch, R. (TAMC 2011, STOC 2012)]

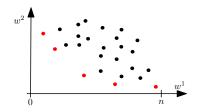
 $P_d(n,\phi) = O(n^{2d}\phi^d)$ $P_d(n,\phi) = \Omega(n^{d-1.5}\phi^d)$

extension to non-linear objective functions

Beier, R., Vöcking (IPCO 2007)

- min $w^1(x) = w_1 x_1 + \cdots + w_n x_n$ and min $w^2(x)$
- subject to $x \in S \subseteq \{0, 1\}^n$, *S* arbitrary
- w_j drawn according to $f_j : [0, 1] \rightarrow [0, \phi]$ for $\phi \ge 1$

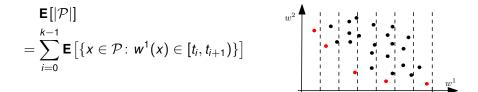
 $\mathsf{P}_2(\mathsf{n},\phi)=\mathsf{O}(\mathsf{n}^2\phi)$



Beier, R., Vöcking (IPCO 2007)

- min $w^1(x) = w_1 x_1 + \dots + w_n x_n$ and min $w^2(x)$
- subject to $x \in S \subseteq \{0, 1\}^n$, *S* arbitrary
- w_j drawn according to $f_j \colon [0, 1] \to [0, \phi]$ for $\phi \ge 1$

 $\mathsf{P}_2(\mathsf{n},\phi)=\mathsf{O}(\mathsf{n}^2\phi)$



 $t_0 = \dot{0}$

 t_1

 t_5

Beier, R., Vöcking (IPCO 2007)

- min $w^1(x) = w_1 x_1 + \cdots + w_n x_n$ and min $w^2(x)$
- subject to $x \in S \subseteq \{0, 1\}^n$, *S* arbitrary
- w_j drawn according to $f_j \colon [0, 1] \to [0, \phi]$ for $\phi \ge 1$

 $\mathsf{P}_2(\mathsf{n},\phi)=\mathsf{O}(\mathsf{n}^2\phi)$

$$\mathbf{E}[|\mathcal{P}|] = \sum_{i=0}^{k-1} \mathbf{E}\left[\{x \in \mathcal{P} : w^{1}(x) \in [t_{i}, t_{i+1})\}\right]$$

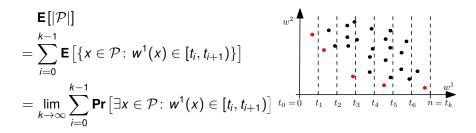
$$\approx \sum_{i=0}^{k-1} \mathbf{Pr}\left[\exists x \in \mathcal{P} : w^{1}(x) \in [t_{i}, t_{i+1})\right]$$

$$w^{2}$$

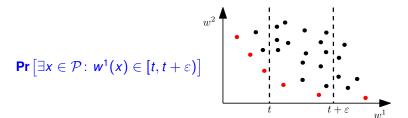
Beier, R., Vöcking (IPCO 2007)

- min $w^1(x) = w_1 x_1 + \dots + w_n x_n$ and min $w^2(x)$
- subject to $x \in S \subseteq \{0, 1\}^n$, *S* arbitrary
- w_j drawn according to $f_j : [0, 1] \rightarrow [0, \phi]$ for $\phi \ge 1$

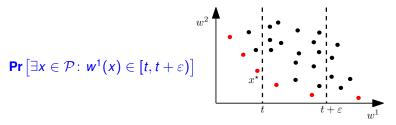
 $\mathsf{P}_2(\mathsf{n},\phi)=\mathsf{O}(\mathsf{n}^2\phi)$



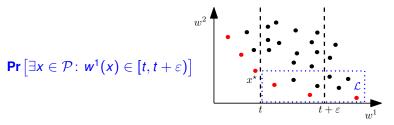
Loser Gap



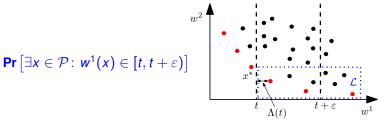
Loser Gap



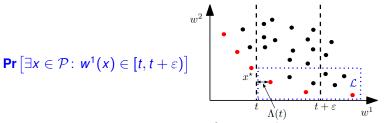
- single-criterion problem: min $w^2(x)$ s.t. $w^1(x) \le t$ and $x \in S$
- winner: $x^* = optimal solution$



- single-criterion problem: min $w^2(x)$ s.t. $w^1(x) \le t$ and $x \in S$
- winner: $x^* = optimal solution$
- loser set: \mathcal{L} = all solutions $x \in S$ with $w^2(x) < w^2(x^*)$

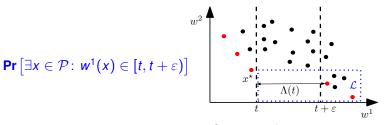


- single-criterion problem: min $w^2(x)$ s.t. $w^1(x) \le t$ and $x \in S$
- winner: $x^* = optimal solution$
- loser set: \mathcal{L} = all solutions $x \in S$ with $w^2(x) < w^2(x^*)$
- loser gap: $\Lambda(t)$ = distance of loser set \mathcal{L} from t



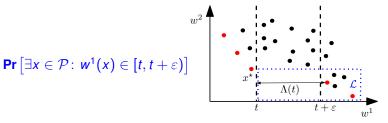
- single-criterion problem: min $w^2(x)$ s.t. $w^1(x) \le t$ and $x \in S$
- winner: $x^* = optimal solution$
- loser set: \mathcal{L} = all solutions $x \in S$ with $w^2(x) < w^2(x^*)$
- loser gap: $\Lambda(t)$ = distance of loser set \mathcal{L} from t

$$\exists x \in \mathcal{P} \colon w^{1}(x) \in [t, t + \varepsilon) \iff \Lambda(t) \leq \varepsilon$$



- single-criterion problem: min $w^2(x)$ s.t. $w^1(x) \le t$ and $x \in S$
- winner: $x^* = optimal solution$
- loser set: \mathcal{L} = all solutions $x \in S$ with $w^2(x) < w^2(x^*)$
- loser gap: $\Lambda(t)$ = distance of loser set \mathcal{L} from t

$$\exists x \in \mathcal{P} \colon w^1(x) \in [t, t + \varepsilon) \iff \Lambda(t) \le \varepsilon$$



- single-criterion problem: min $w^2(x)$ s.t. $w^1(x) \le t$ and $x \in S$
- winner: $x^* = optimal solution$
- loser set: \mathcal{L} = all solutions $x \in \mathcal{S}$ with $w^2(x) < w^2(x^*)$
- loser gap: $\Lambda(t)$ = distance of loser set \mathcal{L} from t

$$\exists x \in \mathcal{P} \colon w^1(x) \in [t, t + \varepsilon) \iff \Lambda(t) \le \varepsilon$$

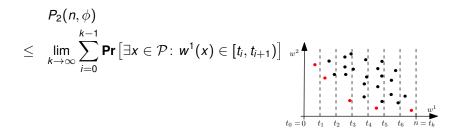
Lemma [Beier, Vöcking (STOC 2004)]

For every $\varepsilon \geq 0$ and $t \in \mathbb{R}$, $\Pr[\Lambda(t) \leq \varepsilon] \leq n\phi\varepsilon$.

Bicriteria Optimization – Proof

Lemma [Beier, Vöcking (STOC 2004)]

For every $\varepsilon \geq 0$ and $t \in \mathbb{R}$, $\Pr[\Lambda(t) \leq \varepsilon] \leq n\phi\varepsilon$.



Bicriteria Optimization – Proof

Lemma [Beier, Vöcking (STOC 2004)]

For every $\varepsilon \geq 0$ and $t \in \mathbb{R}$, $\Pr[\Lambda(t) \leq \varepsilon] \leq n\phi\varepsilon$.

Bicriteria Optimization – Proof

Lemma [Beier, Vöcking (STOC 2004)]

For every $\varepsilon \geq 0$ and $t \in \mathbb{R}$, $\Pr[\Lambda(t) \leq \varepsilon] \leq n\phi\varepsilon$.

Beier, R., Vöcking (IPCO 2007) $P_2(n, \phi) = O(n^2 \phi)$

Outline

Binary Optimization Problems When does a binary optimization problem have polynomial smoothed complexity?

Multiobjective Optimization How many Pareto-optimal solutions do usually exist?

Conclusions

Summary

Smoothed analysis is a promising framework for a more realistic theory of algorithms. It explains success of simplex algorithm, 2-Opt, and many other algorithms.

Summary

Smoothed analysis is a promising framework for a more realistic theory of algorithms. It explains success of simplex algorithm, 2-Opt, and many other algorithms.

Open Questions

- analyze other pivot rules for simplex method
- improve exponents of smoothed running time for 2-Opt etc.
- analyze your favorite problem/algo that is hard in the worst case
- use insights to develop better algorithms
- explore other frameworks for realistic theory