CSC 2429 Winter 2018

Proof Complexity, Mathematical Programming, and Algorithms

Toni Pitassi
www.cs.toronto.edu/~toni

Course webpage: click on teaching

This course is about large classes of (sound) algorithms for solving NP-hard optimization problems (+ decision problems)

Marriage between complexity theory (proof complexity) and algorithms

which has emerged over the last 10-15 years

Very Roughly: Proof system gives rise to a family of sound and feasible algorithms

can be used to rule out major approaches ← Extended Formulations

Can also be used to give new algorithms! ← automat. of SOS plus proof upper bounds

Example 1 MaxSAT / MaxCSP

given $f = C_1 \wedge C_2 \wedge \dots \wedge C_m$ over $x_1 \dots x_n$

find 0/1 assignment to $x_1 \dots x_n$ that maximizes the number of satisfied clauses

Easy to formulate as an integer program:

Say $f = (x_1 \vee x_2 \vee \bar{x}_3)(x_1 \vee x_3)(x_1 \vee \bar{x}_2)(\bar{x}_1)$

$$\text{Max} \quad c_1 + c_2 + c_3 + c_4$$

$$\begin{aligned}
\text{s.t.} \quad & x_1 + x_2 + (1 - x_3) \geq c_1 \\
& x_1 + x_3 \geq c_2 \\
& x_1 + (1 - x_2) \geq c_3 \\
& (1 - x_1) \geq c_4 \qquad , \quad x_i, c_j \in \{0, 1\}
\end{aligned}$$

what happens if we relax constraints to

$0 \le x_i \le 1 \quad 0 \le c_j \le 1$?

Then we have an LP so can solve in polytime
(ellipsoid alg Khachiyàn '79)

How do solutions compare?

For unsat 3CNF f, integer OPT is $\frac{7}{8}m$

But LP has a fractional OPT of m

Our example: set $x_1 = x_2 = x_3 = \frac{1}{2}$, $c_1 = c_2 = c_3 = c_4 = 1$
satisfies all constraints & has fractional OPT = 4

Note we could alternatively formulate MAXSAT as a decision problem

Constraints are as above **plus**

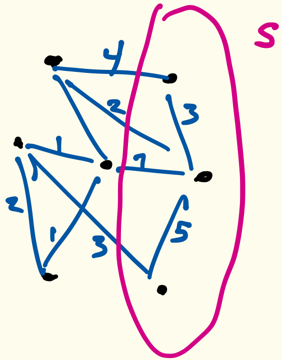
$c_1 + c_2 + c_3 + c_4 \geq 4$

over integers this system of inequalities is infeasible, but there are feasible solutions whenever $0 \leq x_i \leq 1$

Example 2 MAX-CUT

Let $G = (V, E)$ $V = \{v_1 \ldots v_n\}$ with non-neg edge weights w_{ij}

Find $S \subset V$ that maximizes $\mathrm{cut}(S)$



Cut: edges from $\bar{S}$ to S

value$(S) = 4+2+7+3 = 16$

Easy to formulate as a quadratic program:

$$\max \ \frac{1}{2} \sum_{i<j} w_{ij} (1 - y_i y_j)$$

$$y_i \in \{-1, 1\}$$

Relaxation of Quadratic Program to SDP:

$$\max \ \tfrac{1}{2} \sum_{i<j} w_{ij}(1 - u_i^T u_j)$$

$$u_i \in \mathbb{R}^n, \quad \|u_i\| = 1$$

Let X = gram matrix of $u_1 \dots u_n$

So $X = U^T U$ where $U = [u_1 \ u_2 \dots u_n]$

Fact $X = U^T U$ is equivalent to X being positive-semidef PSD ($z^T X z \geq 0 \ \forall z$)

Thus relaxation becomes

$$\max \ \tfrac{1}{2} \sum_{i<j} w_{ij}(1 - X_{ij}),$$

$$X \succeq 0$$

PSD

The relaxation is a SDP and can be solved in polynomial time (Ellypsoid Alg, Khachiyan '79)

Goemans-Williamson gave an amazing approx alg for MaxCUT:

1. Solve above SDP
2. apply randomized rounding to convert solution to 0/1 solution

They prove that the resulting (rounded) solution is always $\geq .872 \cdot OPT$

This course:

We study systematic techniques to improve the relaxations in order to get better algorithms

Improvements involve:

adding new variables + new constraints to original set to cut away feasible fractional (non-integer) solutions

Need to keep all feasible integer solutions (this is soundness)

Proof Systems are systematic, sound ways to do this

Ones we will study:

- Resolution
- Poly Calculus/Nullsatz
- Sherali Adams, Cutting Planes
- SOS

Viewing them as refutation systems corresponds to analyzing decision problems/feasibility

Viewing them as derivation systems corresponds to analyzing optimization problems

New Algorithms via Proof Complexity UPPER Bounds

1. We will study <u>automatizability</u> — how hard it is to find proofs in a particular pf system.

   Nearly all proof systems we will study will be (somewhat) automatizable

   So if a proof of small size/degree exists there is an efficient algorithm to find one

2. automatizability + small proofs of definability
   $\Rightarrow$ efficient algorithm

Limitations of Algorithmic Methods via Proof Complexity Lower Bounds

1. DPLL / Res LBs rule out large family of SAT algs

2. SA lower bounds $\Rightarrow$ LBs for a large class of LPs

   $\Rightarrow$ LBs for large class of extended formulations

3. SOS lower bounds $\Rightarrow$ LBs for large class of SDPs + SDP extended formulations

Proof System Basics

Input: a set of constraints over $x_1 \ldots x_n$

[usually each $x_i \in \{0,1\}$, but we will also be interested in other cases: Finite field, $\mathbb{R}$]

A proof system $P(y) \to F$ polytime

y: proof in P F: what y proves

P as a refutation system

soundness: Range $\subseteq$ UNSAT

completeness: UNSAT $\subseteq$ Range

$P(y) \to (h \to F)$

as a derivation system

h: hypotheses

y: P-proof from h

$y \to F$ what is being proven

Soundness: Range $\subseteq$ TAUT, COMPLETENESS: TAUT $\subseteq$ RANGE

Proof Length, Automatizability

$S_P(F)$ = size of shortest P-refutation of F

P is polynomially automatizable if $\exists A$ for all n suff large and all F over n vars

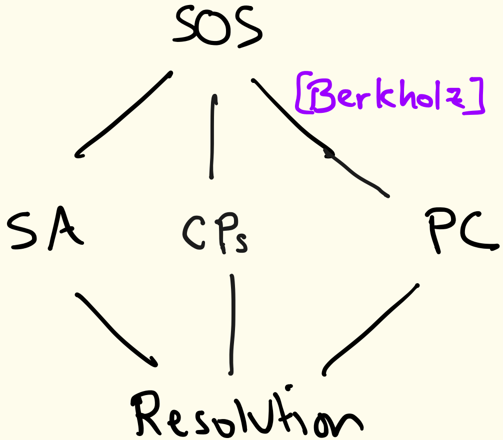
$A(F) \longrightarrow Y$, y is a P-proof of F

Runtime of A is $\mathrm{poly}(S_P(F))$

Can define $g(n)$-automatizability for other runtime functions $g(n)$

★ automatizable means short proofs $\Rightarrow$ efficient algorithms

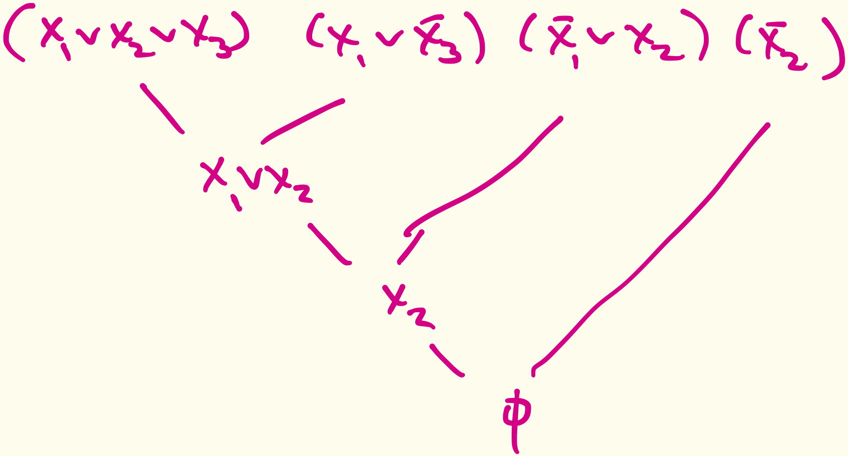
Proof Systems



Resolution (as refutation system) over $\{0,1\}$

$$f = (x_1 \vee x_2 \vee x_3)(x_1 \vee \bar{x}_3)(\bar{x}_1 \vee x_2)(\bar{x}_2)$$

Rule: $(x \vee C)\ (\bar{x} \vee D) \implies C \vee D$



Lets view Resolution a bit differently over $[0,1]$

$$f = (x_1 \vee x_2 \vee x_3)(x_1 \vee \bar{x}_3)(\bar{x}_1 \vee x_2)(\bar{x}_2)$$

$$x_1 + x_2 + x_3 \geq 1 \;,\quad x_1 + (1 - x_3) \geq 1 \;,\quad (1 - x_1) + x_2 \geq 1 \;,\quad 1 - x_2 \geq 1 \;,\quad 0 \leq x_i \leq 1$$

$$x_1 + x_2 \geq 1$$

$$x_2 \geq 1$$

$$0 \geq 1$$

rule preserves all 0/1 feasible solutions

Lets view Resolution a bit differently

$$f = (x_1 \vee x_2 \vee x_3)(x_1 \vee \bar{x}_3)(\bar{x}_1 \vee x_2)(\bar{x}_2)$$

$$x_1 + x_2 + x_3 \geq 1, \quad x_1 + (1 - x_3) \geq 1, \quad (1 - x_1) + x_2 \geq 1, \quad 1 - x_2 \geq 1, \quad 0 \leq x_i \leq 1$$

$$x_1 + x_2 \geq 1$$

$$x_2 \geq 1$$

$$0 \geq 1$$

Original constraints : have a fractional solution

Extended constraints : No fractional solutions

- Resolution in the worst-case requires $2^{\Omega(n)}$ length refutations.

- Width-size relationship (for kCNFs F)

  width w proof $\Rightarrow$ size $2^{O(w)}$ Tree-Res

  width w proof $\Rightarrow$ size $2^{\sqrt{n \log S(F)}}$ Res

  size of shortest Res ref of F

- automatizability:

  width w proofs $\Rightarrow$ can be found in $n^{O(w)}$ time

  Tree Res is quasi-poly automatizable

  Res is automatizable in time $\exp(\sqrt{n \log S(F)})$

Example where Resolution tightening helps

Paturi-Pudlak-Saks-Zane

$(1.364)^n$ algorithm for 3SAT (Hertli - $(1.308)^n$)

$(1.308)^n$ alg for unique 3SAT

Idea Start with f.

Run bounded-width Resolution to "tighten" the constraints

Apply simple randomized search repeatedly

• set unit clauses, ow pick random var + assignment

Intuition: unique sat ass ⇒ lots of unit clauses

after tightening by bded width Res ⇒ more unit clauses

LP (as a proof system)

Sound, complete, poly automatizable proof system for linear inequalities over $\mathbb{R}^{\geq 0}$

LP: $\max \quad c^T x$ ← linear objective function

$$\text{s.t.} \quad \left.\begin{aligned} Ax &\leq b \\ x &\geq 0 \end{aligned}\right\} \text{linear constraints} \quad (*)$$

Decision version:

Is there a value of x satisfying $(*)$?

Farkas' Lemma (Completeness of LP, decision version)

A set $\{Ax \le b, \; x \ge 0\}$ of linear inequalities is unsat (over $\mathbb{R}$) iff $\exists y \ge 0$ s.t. $y^T A = 0$ and $y^T b = -1$

Duality (Implicational completeness of LP)

Consider any non-neg y^T s.t. $y^T A \ge c^T$

Then for any feasible solution x to $\{Ax \le b, x \ge 0\}$ we have

$$c^T x \le y^T A x \le y^T b$$

So $y \ge 0$ witnesses the upper bound $b^T y$

How tight is such an upper bound?

Duality

(P) primal:

$\max \ c^T x$

s.t. $Ax \le b, \ x \ge 0$

(D) dual:

$\min \ b^T y$

s.t. $A^T y \ge c, \ y \ge 0$

Duality theorem (implied by Farkas' Lemma)

Exactly one of the following holds

(i) Neither (P) nor (D) have a feas. soln

(ii) (P) has solns with arbitrarily large values, & (D) unsat

(iii) (D) has " " " small " , (P) unsat

(iv) Both (P) and (D) have optimal solns, x^* and y^*

Then $c^T x^* = b^T y^*$

so there is a soln to dual that witnesses tight bound

So an LP "refutation" of $\{Ax \le b,\ x \ge 0\}$ is
a non-neg linear comb. of these inequalities that equals -1.

Soundness → easy
Completeness → Farkas' Lemma

An LP "derivation" of $\{Ax \le b,\ x \ge 0\} \to c^T x \le c_0$
is nonneg y^* s.t. $(y^*)^T b = c_0$

$$\left(\text{since } c^T x \le (y^*)^T A x \le (y^*)^T b = c_0\right)$$

Soundness → easy
Completeness → Duality Thm

Polynomial automatizability of LP

Satisfiability of linear inequalities : in NP

Farkas' Lemma / Duality Thm : in coNP

so in NP ∩ coNP

Ellipsoid Alg [Khachiyan '79]

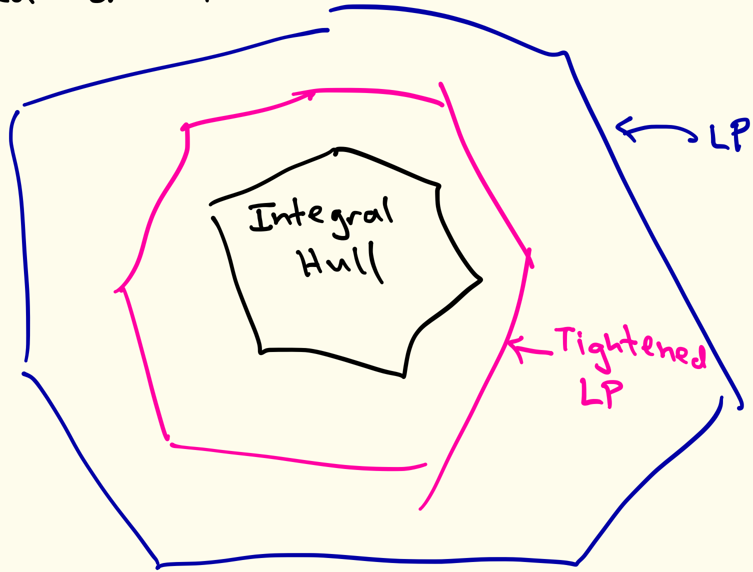
LP in P

LP as relaxation of Integer Program

As mentioned earlier, we can view LP as a relaxation of an LP.

Then we have Fractional OPT $\geq$ Integral OPT

Interested in Ratio



Cutting Planes [LP tightening]

Let $\{Ax \geq 0\}$ be a set of linear inequalities

A refutation of $\{Ax \geq 0\}$ (over $x_i \in \{0,1\}$) is a sequence of inequalities s.t. each is either from $\{Ax \geq 0\}$ or follows from previous lines by a rule, and final line is $-1 \geq 0$

Rules: ① can take positive linear comb's of previously derived ineq's

② Division with Rounding: $\sum c_i x_i \geq b$, each c_i / k

$$\Rightarrow \sum \frac{c_i}{k} x_i \geq \left\lceil \frac{b}{k} \right\rceil$$

on <u>Automatizability of CPs</u>

Not known to be automatizable
in any sense

(So we won't say much more about it)

Sherali Adams (LP tightening)

add new variables to represent all low degree ($\leq d$) juntas (Junta: $x_1 \bar{x}_2 x_3 \approx x_1(1-x_2)x_3$)

This "lifts" LP from n dimensions to $\leq n^d$ dimensions
projection back to $x_1 \ldots x_n$ presences all 0/1 solns
and will hopefully remove a lot of fractional solns

Sherali Adams degree d tightening

Original LP: $\max c^T x$

s.t. $Ax \geq b$

$0 \leq x \leq 1$

add new variables $J_{S,T}$ $\forall S,T$ $S \cap T = \phi$

$|S|+|T| \leq d$

$J_{S,T}$ represents $\prod_{i \in S} x_i \prod_{j \in T} (1-x_j)$

New constraints: $J_{S,T} \geq 0$, $1 - J_{S,T} \geq 0$

$J_{S \cup v, T} + J_{S, T \cup v} \geq J_{S,T}$

$J_{S,T} \geq J_{S \cup v, T} + J_{S, T \cup v}$

$J_{S,T} \left(\Sigma a_{ij} x_j \right) \geq J_{S,T} b_j$

Sherali-Adams (static)

Let $\mathcal{F}$ be a set of poly equalities (includes $x_i^2 - x_i = 0$)
$\mathcal{H}$ a set of poly inequalities

a SA derivation of f from $(\mathcal{F}, \mathcal{H})$ is
$(g_1 \ldots g_m, P_1 \ldots P_s)$ s.t.

$$\sum_{i=1}^{m} g_i f_i + \sum_{\ell=1}^{s} P_\ell h_\ell = f$$

arbitrary poly's

nonnegative linear combination of juntas

$$\prod_{j \in A} x_j \cdot \prod_{j \in B} (1 - x_j)$$

Degree = max degree of $g_i f_i$, $P_\ell h_\ell$

SA automatizability

By Farkas' Lemma, Duality degree d

SA refutations/derivations are automatizable

in time $n^{O(d)}$.

Sum-of-Squares SOS (Static)

Let $\mathcal{F}$ be a set of poly equalities (includes $x_i^2 - x_i = 0$)
$\mathcal{H}$ a set of poly inequalities

an SOS derivation of f from $(\mathcal{F}, \mathcal{H})$ is

$(g_1 \ldots g_m, p_1 \ldots p_s)$ s.t.

$$\sum_{i=1}^{m} g_i f_i + \sum_{\ell=1}^{s} p_\ell h_\ell = f$$

arbitrary poly's

sum of squares
ie. $\sum_i (p_{\ell,i})^2$

Degree = max degree of $g_i f_i$, $p_\ell h_\ell$

Sum-of-Squares SOS (static)

Nullstellensatz (Static)

Start with polynomials $\{f_1 = 0, \dots\}$ including $x_i^2 - x_i = 0$

a Nullstellensatz derivation of f is $(g_1, \dots)$ s.t.

$$\sum g_i f_i = f$$

$f = -1$: refutation of $\mathcal{F} = \{f_1 = 0, \dots\}$

Poly calculus (Dynamic)

Axioms : $f = 0 \;\forall f \in \mathcal{F}$ (including $x_i^2 - x_i = 0$)

Rules : $f = 0 \Rightarrow x_j \cdot f = 0$

$f = 0,\; g = 0 \Rightarrow ag + bf = 0$

Automatizability of Nullsatz + PC

Proofs of degree d can be found in time $n^{O(d)}$

Nullsatz : solve system of linear eqns

PC : bded-degree version of gröbner basis alg