Mini-tutorial on proof complexity

Sam Buss

Theoretical Foundations of Applied Sat Solving
Banff International Research Station
January 20, 2014

This talk discusses:

Proof systems:
Frege proofs, extended Frege proofs, abstract proof systems, resolution, cutting planes, nullstellensatz, the polynomial calculus.

The extension rule:
Frege versus extended resolution (equivalent to extended Frege).
Resolution versus extended resolution

Interpolation and lower bounds:
Resolution.
Cutting planes.

Automatizability and conditional lower bounds.

The **Frege proof system** $\mathcal{F}$ is a “textbook-style” propositional proof system with Modus Ponens as its only rule of inference.

Variables: x, y, z, . . . range over *True*/*False*.
Connectives: $\neg$, $\wedge$, $\vee$, $\rightarrow$.

Modus Ponens: $$\frac{\varphi \qquad \varphi \rightarrow \psi}{\psi}.$$

Axiom Schemes:
$$\varphi \rightarrow \psi \rightarrow \varphi$$
$$(\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \psi \rightarrow \chi) \rightarrow (\varphi \rightarrow \chi)$$
and 8 more axiom schemes.

Defn: The **size** of a Frege proof is the number of symbols in the proof. $\mathcal{F} \stackrel{m}{\vdash} \varphi$ means φ has an $\mathcal{F}$ proof of size m.
The **size** of a formula is the number of symbols in the formula.

Thm: $\mathcal{F}$ is (implicationally) sound and (implicationally) complete.

More generally, a **Frege system** is specified by any finite complete set of Boolean connectives and finite set of axiom schemes and rule schemes, provided it is implicationally sound and implicationally complete.

By completeness, every tautology has an $\mathcal{F}$proof.

Open problem: Is there a polynomial $p(n)$ such that every tautology has an $\mathcal{F}$-proof of size $\leq p(n)$?
That is, is $\mathcal{F}$ polynomially bounded?

The answer is the same for all Frege systems, in that any two Frege systems "p-simulate" each other.
[Reckhow'76; Cook-Reckhow'79]

Defn: An **abstract proof system** is a polynomial time function f mapping $\{0,1\}^*$ *onto* the set of tautologies.
w is an f-proof of φ iff $f(w) = \varphi$. The **size** of w is $|w|$, i.e. the length of w.

Example: For the Frege system $\mathcal{F}$:

$$f_{\mathcal{F}}(w) = \begin{cases} \text{the last line of } w & \text{if } w \text{ is an } \mathcal{F}\text{-proof} \\ (x \vee \neg x) & \text{otherwise} \end{cases}$$

Similar constructions allow very strong systems, e.g. ZF set theory, to be abstract proof systems.

Thm. [CR'79] There ia polynomially bounded abstract proof system iff $\mathrm{NP} = \mathrm{coNP}$.

Defn: Let f, g be abstract proof systems.

f **simulates** g if there is a polynomial $q(n)$ s.t., whenever $g(w) = \varphi$, there is a v, $|v| \le q(|w|)$ such that $f(v) = \varphi$.

f **p-simulates** g if there is a polynomial-time computable $h(w)$, such that, whenever $g(w) = \varphi$, we have $f(h(w)) = \varphi$.

f is **polynomially bounded** if, for some polynomial $q(n)$, every tautology φ has an f-proof w of size $\le q(|\varphi|)$.

1. Resolution is not polynomially bounded.

2. Regular resolution does not simulate resolution.

2. Resolution does not simulate $\mathcal{F}$.

3. Any two Frege systems p-simulate each other.

4. It is open whether $\mathcal{F}$ is polynomially bounded.

5. It is open whether there is a maximum abstract proof system which p-simulates all abstract proof systems.

Defn: [ess. Tseitin '68] Extension allows introduction of new variables for formulas; namely the **extension rule**:

$$z \leftrightarrow \varphi$$

where z is a variable not appearing in earlier lines the proof, in φ, or in the last line of the proof.
The **extended Frege system ($e\mathcal{F}$)** is Frege ($\mathcal{F}$) plus the extension rule.

Thm. [Statman'77] If $\mathcal{F} \mathrel{\mid\!\!\frac{m \text{ steps}}{\quad}} \varphi$, then φ has a $e\mathcal{F}$-proof of size $O(m+|\varphi|^2)$, that is $e\mathcal{F} \mathrel{\mid\!\!\frac{O(m+|\varphi|^2)}{\quad}} \varphi$.

Thus the *size* of extended Frege proofs is essentially the same as the *number of lines* in Frege proofs.

Proof idea: Introduce extension variables for the formulas in the Frege proof; thereby reduce all lines to constant size. □

So:

- A Frege proof is a proof in which each line is a Boolean *formula*.
- An extended Frege proof is a proofs in which each line is a Boolean *circuit*.

It is conjectured that circuits cannot be converted into polynomial size equivalent formulas; the corresponding conjecture is that $\mathcal{F}$ does not (p-)simulate $e\mathcal{F}$.

Resolution

Resolution is a refutation system, refuting sets of clauses. Thus, resolution is a system for refuting CNF formulas, equivalently, a system for proving DNF formulas are tautologies.

Defn: Extension rule for resolution: For z a *new* variable, x and y literals, introduce $z \leftrightarrow (x \wedge y)$ by adding the clauses:

$$\{\overline{x}, \overline{y}, z\} \qquad \{\overline{z}, x\} \qquad \{\overline{z}, y\}.$$

Resolution as an abstract proof system: Given φ, introduce clauses Γ for the extension variables z_ψ for all subformulas ψ of φ. A resolution proof of φ is a resolution refutation of $\overline{z}_\varphi, \Gamma$.

Extended resolution is resolution augmented with unrestricted use of the extension rule.

Thm: Extended resolution and extended Frege p-simulate each other.

Can the extension rule help?

For resolution: Yes.

Thm: [Haken'86, BIKPPW'92, Raz'02, Razborov'03, many others]
The pigeonhole principle (PHP) requires resolution proofs of size 2^{n^ϵ} (even PHP_n^m for $m \gg n$).
Furthermore, depth d Frege proofs of PHP_n^{n+1} require size $2^{n^{\epsilon_d}}$.

Thm: [CR'79].
PHP_n^{n+1} has extended resolution proofs of polynomial size.

Can extension rule help?

For Frege: Open question.

Thm: [B'86]. PHP^{n+1}_n has Frege proofs of polynomial size.

Thm: [B'??]. The [CR'79] $e\mathcal{F}$ proofs of PHP^{n+1}_n can be translated into quasi-polynomial size $\mathcal{F}$-proofs of PHP^{n+1}_n.

[Bonet-B-Pitassi'95] Suggested several other tautologies for separating Frege and extended Frege systems. However, these all now have been shown to have quasi-polynomial size Frege proofs:

Thm: [Hrubes-Tzameret'12] The matrix identity over $\mathbb{Z}_2$, $AB = I \Rightarrow BA = I$ (c.f. BBP'91) has quasi-polynomial size Frege proofs.

Thm: [Aisenberg-Bonet-B'??] Frankl's Theorem (c.f. BBP'91) has quasipolynomial size Frege proofs.

Cook-Reckhow's $e\mathcal{F}$ proof of PHP_n^{n+1}

Code the graph of $f : [n+1] \to [n]$ with variables $x_{i,j}$ indicating that $f(i) = j$.

$PHP_n^{n+1}(\vec{x})$: "f is not both total and injective"

Use extension to introduce new variables

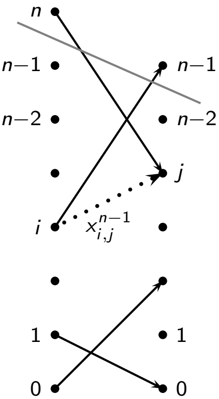
$$x_{i,j}^{\ell-1} \leftrightarrow x_{i,j}^{\ell} \vee (x_{i,\ell-1}^{\ell} \wedge x_{\ell,j}^{\ell}).$$

for $i \le \ell$, $j < \ell$; where $x_{i,j}^{n} \leftrightarrow x_{i,j}$.

Prove, for each ℓ that

$$\neg PHP_{\ell}^{\ell+1}(\vec{x}^{\ell}) \to \neg PHP_{\ell-1}^{\ell}(\vec{x}^{\ell-1}).$$

Finally derive $PHP_n^{n+1}(\vec{x})$ from $PHP_1^2(\vec{x}^1)$. □



Cook-Reckhow's proof of PHP_n^{n+1} as a Frege proof

Let G^ℓ be the directed graph with:
edges $(\langle i,0\rangle,\langle j,1\rangle)$ such that $x_{i,j}$ holds, and
edges $(\langle i,1\rangle,\langle i{+}1,0\rangle)$ such that $i{\geq}\ell$ (blue edges).

For $i \leq \ell$, $j < \ell$, let $\varphi^\ell_{i,j}$ express

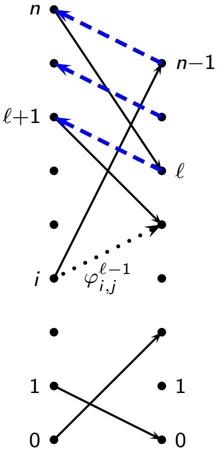
"Range node $\langle j,1\rangle$ is reachable
from domain node $\langle i,0\rangle$ in G^ℓ".

$\varphi^\ell_{i,j}$ is a quasi-polynomial size formula via an NC^2 definition of reachability.

For each ℓ, prove that

$$\neg PHP^{\ell+1}_\ell(\vec{\varphi}^\ell) \to \neg PHP^\ell_{\ell-1}(\vec{\varphi}^{\ell-1}).$$

Finally derive $PHP^{n+1}_n(\vec{x})$ from $PHP^2_1(\vec{\varphi}^1)$. □



Implications for using extension in CDCL????

- Concentrate on extension that represents polynomial size formulas? (Instead of polynomial size circuits.)

  Rationale: We have no conjectured examples of exponential improvement of $e\mathcal{F}$ over $\mathcal{F}$ except partial consistency statements.

- Extension should not be restricted to just conjunctions or disjunctions or original literals.

  Rationale: Constant depth Frege proofs require exponential size for combinatorial principles such as pigeonhole principle.

- However, it would reasonable to concentrate extension to representing low depth formulas.

Regular resolution versus resolution

Thm: [Goerdt'92,Alekhnovich-Johannsen-Pitassi-Urquhart'07, Urquhart'11]
Regular resolution does not simulate resolution.

[AJPU'07,U'11] proved the separation using modified ("guarded") graph tautologies and pebbling principles, and using a "Stone" principle. Both are based on well-foundedness conditions in directed acyclic graphs.

CDCL versus resolution

Thm: [Beame-Kautz-Sabharwal'04, Atserias-Fichte-Thurley'11, Pipatsrisawat-Darwiche'11]
CDCL with restarts simulates, and is simulated by, resolution.

Open: Does CDCL without restarts simulate resolution?

Thm: [Bonet-B'12, Bonet-Johannsen-B'??,B-Kołodziejczyk'??]
The guarded pebbling and graph tautologies, and Stone principles, of [AJPU'07,U'11] have polynomial size refuations in RegWRTI, a proof system that closely corresponds to CDCL without restarts.

Thus, we have no conjectured examples for separating resolution from CDCL without restarts, or from RegWRTI.
On the other hand, no simulation of resolution by CDCL without restarts has been found.

Cutting planes

Variables $x_1, x_2, \ldots$ are 0/1 valued (0="False", 1="True").

Lines are linear inequalities with integer coefficients:

$$a_1 x_1 + a_2 x_2 + \cdots a_n x_n \geq a_0.$$

Clauses become inequalities:
e.g., $x \vee \overline{y} \vee z$ becomes the inequality $x - y + z \geq 0$.
Note that $\overline{y}$ becomes $1 - y$.

Axioms: $x_i \geq 0$ and $-x_1 \geq -1$.

Addition rule: $\dfrac{\sum a_i x_i \geq a_0 \qquad \sum b_i x_i \geq b_0}{\sum (a_i + b_i) x_i \geq a_0 + b_0}$

Division rule: If $c | a_i$ for all $i > 0$, $\dfrac{\sum a_i x_i \geq a_0}{\sum (a_i / c) x_i \geq \lceil a_0 / c \rceil}$

Cutting planes is a *refutation system*: initial lines are axioms, or encode initial clauses. The final line of a refutation has the form $0 \geq 1$.

Thm: Cutting planes p-simulates resolution.

Thm: Resolution, and bounded depth Frege systems, do not simulate cutting planes.

Thm: [Goerdt'92] Cutting planes is p-simulated by Frege systems.

Thm: [Pudlák'90] Cutting planes is not polynomially bounded.

Nullstellensatz and polynomial calculus

Work over a finite field, characteristic p.

Variables $x_1, x_2, \ldots$ are $0/1$ valued.

A polynomial f is identified with the assertion $f = 0$.

A set of *initial* polynomials $\{f_j\}_j$ is **refuted** in the **Nullstellensatz system** by polynomials g_j, h_i such that

$$\sum f_j \cdot g_j + \sum (x_i^2 - x_i) \cdot h_i = 1,$$

where equality indicates equality as polynomials.

A **polynomial calculus** refutation uses the inferences of addition and multiplication:

$$\frac{f \quad g}{f+g} \qquad\qquad \frac{f}{f\cdot g}$$

A **polynomial calculus refutation** of a set of polynomials $\{f_j\}_j$ is a derivation of 1 from the f_j's and the polynomials $(x_i^2 - x_i)$.

It is more common to work with the *degree* of nullstellensatz or polynomial calculus proofs, rather than their size. These systems are known to not be simulated by resolution or bounded depth Frege; conversely, several lower bounds are known.

One sample result:

Thm: [Razborov'98] Any polynomial calculus proof of PHP_n^{n+1} must have degree $\Omega(n)$.

Craig Interpolation

[Bonet-Pitassi-Raz'95, Razborov'95, Krajíček'97, Pudlák'97]

Defn: Suppose $A(\vec{p},\vec{r}) \wedge B(\vec{q},\vec{r})$ is unsatisfiable, where A and B depend only on the variables indicated.
A **Craig interpolant** for this formula is a predicate $C(\vec{r})$ such that

- If $\neg C(\vec{r})$, then $A(\vec{p},\vec{r})$ is unsatisfiable.
- If $C(\vec{r})$, then $B(\vec{q},\vec{r})$ is unsatisfiable.

Remark: A Craig interpolant always exists when $A(\vec{p},\vec{r}) \wedge B(\vec{q},\vec{r})$ is unsatisfiable, but may not be a feasible predicate of $\vec{r}$.

Thm: [Krajíček'97] Suppose a set of clauses $A(\vec{p}, \vec{r}), B(\vec{q}, \vec{r})$ has a resolution refutation of size m, and that variables $\vec{r}$ appear only positively in the clauses in $A(\vec{p}, \vec{r})$ or negatively in the clauses in $B(\vec{q}, \vec{r})$. Then, there is Craig interpolant which is computed by a *monotone* Boolean circuit of size $m^{O(1)}$.

Corollary: (Using [Razborov'85, Alon-Boppana'87.) Resolution does not have polynomial size refutations of the Clique-Coloring clauses expressing that a graph both is k-colorable and has a $k+1$ clique.

Thm: [Pudlák'97] Suppose a set of clauses $A(\vec{p},\vec{r}), B(\vec{q},\vec{r})$ has a cutting planes refutation of m steps, and that the variables $\vec{r}$ appear only positively in the clauses in $A(\vec{p},\vec{r})$ or negatively in the clauses in $B(\vec{q},\vec{r})$. Then, there is Craig interpolant which is computed by a monotone real circuit of size $m^{O(1)}$.

Corollary: (Using [Razborov'85, Alon-Boppana'87.) Resolution does not have polynomial size refutations of the Clique-Coloring clauses expressing that a graph both is k-colorable and has a $k+1$ clique.

Automatizability

Defn: A proof system T is **automatizable** (in polynomial time) if there is a procedure, which given a formula φ, produces a T-proof of φ in time bounded by a polynomial of the size of the shortest T-proof of φ (if any).

Defn: A proof system T has **feasible interpolation** if there is polynomial time procedure $C(-,-)$ so that if P is a T-proof of $\neg(A(\vec{p},\vec{r}) \land B(\vec{q},\vec{r}))$, then $C(P,\vec{r})$ is a Craig interpolant for $A(\vec{p},\vec{r}) \land B(\vec{q},\vec{r})$.

Thm: [Bonet-Pitassi-Raz'00] If T is automatizable, then T has feasible interpolation.

Thm: [Krajíček-Pudlák'95, also B'97] The extended Frege system $e\mathcal{F}$ does not have feasible interpolation and thus is not automatizable, unless the RSA encryption function, the discrete logarithm encryption function, and the Rabin encryption function can be inverted in polynomial time.

Thm: [Bonet-Pitassi-Raz'00] The Frege system $\mathcal{F}$ does not have feasible interpolation and thus is not automatizable, unless Blum integers can be factored in polynomial time.

Defn: **Blum integers** are products of two primes, each congruent to 3 mod 4.

A related theorem holds for bounded depth Frege systems under a stronger hardness assumption about Blum integers. [BDGMP'03].

Thm: [Alekhnovich-Razborov'03] Resolution and tree-like resolution are not automatizable unless the parameterized hierarchy class $W[P]$ is fixed-parameter tractable via randomized algorithms with one-sided error.

On the other hand:

Thm: [Beame-Pitassi'96; building on CEI'96]
Tree-like resolution is automatizable in time $n^{\log S}$ where n is the number of variables, and S is the size of the shortest tree-like resolution refutation. (This is quasipolynomial time.)

Resolution is automatizable in time $n^{\sqrt{n \log S}}$.

Open: Is resolution automatizable in quasi-polynomial time?

Thank you!

Some survey articles:

- S. Buss, “Towards NP-P via Proof Complexity and Search”, Annals of Pure and Applied Logic 163, 7 (2012) 906-917.
- S. Buss, “Propositional Proof Complexity: An Introduction”, In Computational Logic, edited by U. Berger and H. Schwichtenberg, Springer-Verlag, Berlin, 1999, pp. 127-178.
- P. Beame and T. Pitassi, “Propositional Proof Complexity: Past, Present and Future”, In Current Trends in Theoretical Computer Science Entering the 21st Century, World Scientific, 2001, pp. 42-70.
- P. Beame with A. Sabharwal, “Propositional Proof Complexity”, In Computational Complexity Theory, IAS/Park City Clay Mathematics Series 10, 2000, pp. 199-246.
- P. Pudlák, “Twelve problems in proof complexity”, In Proc. 3rd International Computer Science Symposium in Russia, CSR 2008, pp.13-27
- N.Segerlind, “The Complexity of Propositional Proofs”, Bulletin of Symbolic Logic 13 (2007) 417-481.