DRAT proofs, propagation redundancy and extended resolution

Neil Thapen

Institute of Mathematics
Czech Academy of Sciences

Joint work with Sam Buss

SAT 2019, Lisbon

Background

Main results

Lower bound proof

Background

Main results

Lower bound proof

Some references

O. Kullmann
On a generalization of extended resolution, 1999

M. Järvisalo, M.J.H. Heule, A. Biere
Inprocessing rules, 2012

M.J.H. Heule, B. Kiesl, M. Seidl, A. Biere
PRuning through satisfaction, 2017

B. Kiesl, A. Rebola-Pardo, M.J.H. Heule
Extended resolution simulates DRAT, 2018

M.J.H. Heule, A. Biere
What a difference a variable makes, 2018

M.J.H. Heule, B. Kiesl, A. Biere
Strong extension-free proof systems, 2019

Proof complexity

We consider *refutations* of unsatisfiable sets of clauses Γ.

Proof complexity

We consider *refutations* of unsatisfiable sets of clauses Γ.

A *derivation* of a clause C from Γ is a sequence $\Gamma_0, \ldots, \Gamma_t$ where

- $\Gamma_0 = \Gamma$ and $C \in \Gamma_t$
- each Γ_{i+1} is derivable from Γ_i by a *rule* of the system being considered.

Rules will always preserve satisfiability.

Proof complexity

We consider *refutations* of unsatisfiable sets of clauses Γ.

A *derivation* of a clause C from Γ is a sequence $\Gamma_0, \ldots, \Gamma_t$ where

- $\Gamma_0 = \Gamma$ and $C \in \Gamma_t$
- each Γ_{i+1} is derivable from Γ_i by a *rule* of the system being considered.

Rules will always preserve satisfiability.

A *refutation* of Γ is a derivation of the empty clause $\bot$ from Γ.

Proof complexity

We consider *refutations* of unsatisfiable sets of clauses Γ.

A *derivation* of a clause C from Γ is a sequence $\Gamma_0, \ldots, \Gamma_t$ where

- $\Gamma_0 = \Gamma$ and $C \in \Gamma_t$
- each Γ_{i+1} is derivable from Γ_i by a *rule* of the system being considered.

Rules will always preserve satisfiability.

A *refutation* of Γ is a derivation of the empty clause $\bot$ from Γ.

A system $\mathcal{P}$ *simulates* a system $\mathcal{Q}$ if every $\mathcal{Q}$-refutation of Γ can be changed into a $\mathcal{P}$-refutation of Γ in polynomial time.

We write things like $\mathcal{Q} \leq \mathcal{P}, \quad \mathcal{Q} < \mathcal{P}, \quad \mathcal{Q} \equiv \mathcal{P}$.

Some notation

- p, q are literals, $\overline{p}, \overline{q}$ are their negations
- C, D are clauses
- α, β are partial assignments
- we can interpret partial assignments as sets of unit clauses
  e.g. if $\alpha : x \mapsto 1, y \mapsto 0$, undefined elsewhere
  then α corresponds to $\{\{x\}, \{\overline{y}\}\}$
- $\overline{C}$ is the partial assignment expressing the negation of C

Resolution, UP and RUP

Resolution rule

If Γ_i contains $C \vee p$ and $D \vee \overline{p}$, derive $\Gamma_{i+1} = \Gamma \cup \{ C \vee D \}$.

If C or D is empty, this is a *unit propagation* inference.

Resolution, UP and RUP

Resolution rule

If Γ_i contains $C \vee p$ and $D \vee \overline{p}$, derive $\Gamma_{i+1} = \Gamma \cup \{C \vee D\}$.

If C or D is empty, this is a *unit propagation* inference.

Definition

- $\Gamma \vdash_1 \bot$ means Γ is refutable using only unit propagations
- $\Gamma \vdash_1 C$ means $\Gamma \cup \overline{C} \vdash_1 \bot$

  We say C *is derivable from* Γ *by reverse unit propagation*.

Resolution, UP and RUP

Resolution rule

If Γ_i contains $C \vee p$ and $D \vee \overline{p}$, derive $\Gamma_{i+1} = \Gamma \cup \{C \vee D\}$.

If C or D is empty, this is a *unit propagation* inference.

Definition

- $\Gamma \vdash_1 \bot$ means Γ is refutable using only unit propagations
- $\Gamma \vdash_1 C$ means $\Gamma \cup \overline{C} \vdash_1 \bot$

  We say C *is derivable from* Γ *by reverse unit propagation*.

$\Gamma \vdash_1 C$ implies (but is not equivalent to) $\Gamma \vDash C$.
The relation $\vdash_1$ is decidable in polynomial time.

Initial rules

$\vdash_1$ rule (reverse unit propagation rule)

From Γ derive $\Gamma \cup \{C\}$, if $\Gamma \vdash_1 C$.

Initial rules

$\vdash_1$ rule (reverse unit propagation rule)

From Γ derive $\Gamma \cup \{C\}$, if $\Gamma \vdash_1 C$.

Deletion rule

From Γ derive any $\Delta \subseteq \Gamma$.

Initial rules

$\vdash_1$ rule (reverse unit propagation rule)

From Γ derive $\Gamma \cup \{C\}$, if $\Gamma \vdash_1 C$.

Deletion rule

From Γ derive any $\Delta \subseteq \Gamma$.

Resolution is equivalent to the system with just the $\vdash_1$ rule.

Neither system gets stronger if we also add the deletion rule.

Extended resolution (ER)

Extension rule

If r does not occur in Γ, for any p, q we can add to Γ the clauses

$$\overline{p} \vee \overline{q} \vee r \qquad \overline{r} \vee p \qquad \overline{r} \vee q$$

expressing $r \leftrightarrow (p \wedge q)$.

Extended resolution (ER)

Extension rule

If r does not occur in Γ, for any p, q we can add to Γ the clauses

$$\overline{p} \vee \overline{q} \vee r \qquad \overline{r} \vee p \qquad \overline{r} \vee q$$

expressing $r \leftrightarrow (p \wedge q)$.

This is a very strong system.

No non-trivial lower bounds are known.

The RAT rule

Definition

Let C contain a literal p. C is a *resolution asymmetric tautology* (RAT) w.r.t. Γ and p if

$$\Gamma \vdash_1 C \vee D$$

for every clause of the form $D \vee \overline{p}$ in Γ.

The RAT rule

Definition

Let C contain a literal p. C is a *resolution asymmetric tautology* (RAT) w.r.t. Γ and p if

$$\Gamma \vdash_1 C \vee D$$

for every clause of the form $D \vee \overline{p}$ in Γ.

RAT rule

Let C be any clause (even in new variables). If C is RAT w.r.t. Γ and p for some $p \in C$, we can derive $\Gamma \cup \{C\}$ from Γ.

The RAT rule

Definition

Let C contain a literal p. C is a *resolution asymmetric tautology* (RAT) w.r.t. Γ and p if

$$\Gamma \vdash_1 C \vee D$$

for every clause of the form $D \vee \overline{p}$ in Γ.

RAT rule

Let C be any clause (even in new variables). If C is RAT w.r.t. Γ and p for some $p \in C$, we can derive $\Gamma \cup \{C\}$ from Γ.

Deletion can make more clauses RAT.

Soundness of RAT

Lemma

If Γ is satisfiable, and C is RAT w.r.t. Γ and a literal p, then $\Gamma \cup \{C\}$ is satisfiable.

Soundness of RAT

Lemma

If Γ is satisfiable, and C is RAT w.r.t. Γ and a literal p, then $\Gamma \cup \{C\}$ is satisfiable.

Proof. Recall C contains p.

Let τ be a **total** assignment with $\tau \models \Gamma$. If $\tau \models C$ we are done.

Soundness of RAT

Lemma

If Γ is satisfiable, and C is RAT w.r.t. Γ and a literal p, then $\Gamma \cup \{C\}$ is satisfiable.

Proof. Recall C contains p.

Let τ be a **total** assignment with $\tau \models \Gamma$. If $\tau \models C$ we are done.

Otherwise, $\tau(p) = 0$. Let τ' be τ with p flipped to 1.

Then τ' satisfies C and also every clause in Γ not containing $\overline{p}$.

Soundness of RAT

Lemma

If Γ is satisfiable, and C is RAT w.r.t. Γ and a literal p, then $\Gamma \cup \{C\}$ is satisfiable.

Proof. Recall C contains p.

Let τ be a **total** assignment with $\tau \vDash \Gamma$. If $\tau \vDash C$ we are done.

Otherwise, $\tau(p) = 0$. Let τ' be τ with p flipped to 1.

Then τ' satisfies C and also every clause in Γ not containing $\overline{p}$.

It follows directly from the RAT condition that τ' also satisfies every clause in Γ which contains $\overline{p}$.

Hence $\tau' \vDash \Gamma \cup \{C\}$. □

Propagation redundancy

Definition

A clause C is *propagation redundant* w.r.t. Γ if there is a partial assignment τ such that, setting $\alpha = \overline{C}$,

$$\tau \vDash C \qquad \text{and} \qquad \Gamma_{|\alpha} \vdash_1 \Gamma_{|\tau}.$$

Propagation redundancy

Definition

A clause C is *propagation redundant* w.r.t. Γ if there is a partial assignment τ such that, setting $\alpha = \overline{C}$,

$$\tau \vDash C \qquad \text{and} \qquad \Gamma_{|\alpha} \vdash_1 \Gamma_{|\tau}.$$

PR rule

Let C be any clause (even in new variables). If C is PR w.r.t. Γ, we can derive $\Gamma \cup \{C\}$ from Γ.

Propagation redundancy

Definition

A clause C is *propagation redundant* w.r.t. Γ if there is a partial assignment τ such that, setting $\alpha = \overline{C}$,

$$\tau \vDash C \qquad \text{and} \qquad \Gamma_{|\alpha} \vdash_1 \Gamma_{|\tau}.$$

PR rule

Let C be any clause (even in new variables). If C is PR w.r.t. Γ, we can derive $\Gamma \cup \{C\}$ from Γ.

The PR rule generalizes the RAT rule.

It also preserves satisfiability.

Systems

- RAT has the RAT and $\vdash_1$ rules
- PR has the PR and $\vdash_1$ rules
- DRAT has the RAT, $\vdash_1$ and deletion rules
- DPR has the PR, $\vdash_1$ and deletion rules

Easy to show these are all equivalent to ER and thus very strong.

Systems

- RAT has the RAT and $\vdash_1$ rules
- PR has the PR and $\vdash_1$ rules
- DRAT has the RAT, $\vdash_1$ and deletion rules
- DPR has the PR, $\vdash_1$ and deletion rules

Easy to show these are all equivalent to ER and thus very strong.

"No new variables"

We study **weakened** systems where, in refutations of Γ, we may **only use variables from** Γ. We consider, amongst others:

- RAT^- has the RAT and $\vdash_1$ rules, with no new variables
- PR^- etc.
- DRAT^-
- DPR^-

Question

Basic picture

$$\mathrm{Res} < \mathrm{DRAT}^- \leq \mathrm{DPR}^- \leq \mathrm{ER}$$

$$\mathrm{Res} \; < \; \mathrm{RAT}^- \; \leq \; \mathrm{PR}^- \; \leq \; \mathrm{ER}$$

What more can be said?

Background

Main results

Lower bound proof

Results 1 : restrictions

Known: RAT $\equiv$ ER

Results 1 : restrictions

Known: RAT $\equiv$ ER

Proposition

Any proof system which simulates RAT^-, and which is closed under restrictions, also simulates ER.

Results 1 : restrictions

Known: RAT $\equiv$ ER

Proposition

Any proof system which simulates RAT^-, and which is closed under restrictions, also simulates ER.

Most commonly studied proof systems are *closed under restrictions*.

That is, short refutations of Γ imply short refutations of $\Gamma_{|\alpha}$.

Results 2 : deletion collapses sytems

Known: DRAT$^-$ **almost** simulates DPR$^-$.
The simulation works if we allow **one** new variable.

Results 2 : deletion collapses sytems

Known: DRAT^- **almost** simulates DPR^-.
The simulation works if we allow **one** new variable.

Proposition

$\mathsf{DRAT}^- \equiv \mathsf{DPR}^-$

Results 2 : deletion collapses sytems

Known: DRAT^{-} **almost** simulates DPR^{-}.
The simulation works if we allow **one** new variable.

Proposition

$\mathrm{DRAT}^{-} \equiv \mathrm{DPR}^{-}$

Idea: manipulate PR steps to free one variable.

Results 3 : various upper bounds

Many standard hard tautologies, used to prove size lower bounds, have polynomial size refutations in PR^-.
(Some of these were already known)

- The pigeonhole principle
- Tseitin contradictions

Results 3 : various upper bounds

Many standard hard tautologies, used to prove size lower bounds, have polynomial size refutations in PR^-.
(Some of these were already known)

- The pigeonhole principle
- Tseitin contradictions
- The bit pigeonhole principle
- The parity principle
- The clique-colouring principle
- OR-ifications and XOR-ifications.

Idea: these are all very symmetrical, so we can find many useful partial assignment pairs α and τ with $\Gamma_{|\alpha} = \Gamma_{|\tau}$.

Results 3 : various upper bounds

Many standard hard tautologies, used to prove size lower bounds, have polynomial size refutations in PR^-.
(Some of these were already known)

- The pigeonhole principle
- Tseitin contradictions
- The bit pigeonhole principle
- The parity principle
- The clique-colouring principle
- OR-ifications and XOR-ifications.

Idea: these are all very symmetrical, so we can find many useful partial assignment pairs α and τ with $\Gamma_{|\alpha} = \Gamma_{|\tau}$.

Question: what is a plausible hard principle for PR^-?

Results 4 : a lower bound

Theorem

RAT^{-} refutations of the bit pigeonhole principle BPHP_n require size exponential in n.

Results 4 : a lower bound

Theorem

RAT^- refutations of the bit pigeonhole principle BPHP_n require size exponential in n.

New picture

$$\mathsf{Res} < \mathsf{DRAT}^- \equiv \mathsf{DPR}^- \leq \mathsf{ER}$$

$$\mathsf{Res} \; < \; \mathsf{RAT}^- \; < \; \mathsf{PR}^- \; \leq \; \mathsf{ER}$$

Background

Main results

Lower bound proof

Bit pigeonhole principle

Definition

Let $n = 2^k$. The propositional contradiction BPHP_n asserts that each of $n + 1$ pigeons maps to a distinct k-bit binary string.

Bit pigeonhole principle

Definition

Let $n = 2^k$. The propositional contradiction BPHP_n asserts that each of $n+1$ pigeons maps to a distinct k-bit binary string.

Variables: $p_1^x, \ldots, p_k^x$ for the string assigned to pigeon x

BPHP_n consists of $O(n^3)$ "hole" clauses, each of width $2k$, asserting that pigeons x and x' do not both map to string y.

Goal

We define the *pigeon width*, or *p-width*, of a clause or assignment to be the number of pigeons it mentions.

Goal

We define the *pigeon width*, or *p-width*, of a clause or assignment to be the number of pigeons it mentions.

Lemma

There is no RAT^- refutation of BPHP_n in which every clause has p-width $\leq n/3$.

Goal

We define the *pigeon width*, or *p-width*, of a clause or assignment to be the number of pigeons it mentions.

Lemma

There is no RAT^- refutation of BPHP_n in which every clause has p-width $\leq n/3$.

Corollary

There is no RAT^- refutation of BPHP_n of size $2^{n/80}$.

Pigeon facts

A *partial matching* β is a partial assignment assigning some distinct pigeons to some distinct holes (by setting all their bits).

Pigeon facts

A *partial matching* β is a partial assignment assigning some distinct pigeons to some distinct holes (by setting all their bits).

Fact 1

If C has p-width m and $\overline{C}$ has no extension to a partial matching, then C is derivable from BPHP_n in resolution in p-width m.

Idea: it is easy to falsify BPHP_n, starting from $\overline{C}$.

Pigeon facts

A *partial matching* β is a partial assignment assigning some distinct pigeons to some distinct holes (by setting all their bits).

Fact 1

If C has p-width m and $\overline{C}$ has no extension to a partial matching, then C is derivable from BPHP_n in resolution in p-width m.

Idea: it is easy to falsify BPHP_n, starting from $\overline{C}$.

Fact 2

BPHP_n has no resolution refutation of p-width $\leq n$.

Further suppose β is a partial matching setting $\leq n/2$ pigeons. Then $\mathrm{BPHP}_n \cup \beta$ has no resolution refutation of p-width $\leq n/2$.

Idea: $\mathrm{BPHP}_n \cup \beta$ looks like $\mathrm{BPHP}_{n/2}$.

Width lower bound

Claim

Let $\Gamma_0, \ldots, \Gamma_t$ be a resolution derivation with $\Gamma_0 = \text{BPHP}_n$ s.t.

- all clauses have p-width $\leq n/3$
- no clause of BPHP_n is ever deleted.

Let C have p-width $\leq n/3$ and be RAT w.r.t. Γ_t and some p.

Then C is derivable from BPHP_n in **resolution** in p-width $n/3$.

Width lower bound

Claim

Let $\Gamma_0, \ldots, \Gamma_t$ be a resolution derivation with $\Gamma_0 = \mathrm{BPHP}_n$ s.t.

- all clauses have p-width $\leq n/3$
- no clause of BPHP_n is ever deleted.

Let C have p-width $\leq n/3$ and be RAT w.r.t. Γ_t and some p.

Then C is derivable from BPHP_n in **resolution** in p-width $n/3$.

It follows that a RAT^- refutation of BPHP_n of small p-width can be turned step-by-step into a resolution refutation of BPHP_n of small p-width.

By Fact 2, there can be no such refutation.

Width lower bound continued

By Fact 1, to prove the claim it is enough to show $\overline{C}$ **cannot** be extended to a partial matching.

So suppose $\overline{C}$ **can** be extended to a partial matching.

Width lower bound continued

By Fact 1, to prove the claim it is enough to show $\overline{C}$ **cannot** be extended to a partial matching.

So suppose $\overline{C}$ **can** be extended to a partial matching.

We are given that C is RAT w.r.t. Γ_t and some literal p.

Trick: find a hole axiom of the form $\overline{p} \vee H$ such that $\overline{C} \cup \overline{H}$ can be extended to a partial matching β setting $\leq n/2$ pigeons.

Width lower bound continued

By Fact 1, to prove the claim it is enough to show $\overline{C}$ **cannot** be extended to a partial matching.

So suppose $\overline{C}$ **can** be extended to a partial matching.

We are given that C is RAT w.r.t. Γ_t and some literal p.

Trick: find a hole axiom of the form $\overline{p} \vee H$ such that $\overline{C} \cup \overline{H}$ can be extended to a partial matching β setting $\leq n/2$ pigeons.

By the RAT assumption, $\Gamma_t \vdash_1 C \vee H$. That is, $\Gamma_t \cup \overline{C} \cup \overline{H} \vdash_1 \bot$.

Width lower bound continued

By Fact 1, to prove the claim it is enough to show $\overline{C}$ **cannot** be extended to a partial matching.

So suppose $\overline{C}$ **can** be extended to a partial matching.

We are given that C is RAT w.r.t. Γ_t and some literal p.

Trick: find a hole axiom of the form $\overline{p} \vee H$ such that $\overline{C} \cup \overline{H}$ can be extended to a partial matching β setting $\leq n/2$ pigeons.

By the RAT assumption, $\Gamma_t \vdash_1 C \vee H$. That is, $\Gamma_t \cup \overline{C} \cup \overline{H} \vdash_1 \bot$.

Hence $\Gamma_t \cup \beta \vdash_1 \bot$, since $\overline{C} \cup \overline{H} \subseteq \beta$.

Width lower bound continued

By Fact 1, to prove the claim it is enough to show $\overline{C}$ **cannot** be extended to a partial matching.

So suppose $\overline{C}$ **can** be extended to a partial matching.

We are given that C is RAT w.r.t. Γ_t and some literal p.

Trick: find a hole axiom of the form $\overline{p} \vee H$ such that $\overline{C} \cup \overline{H}$ can be extended to a partial matching β setting $\leq n/2$ pigeons.

By the RAT assumption, $\Gamma_t \vdash_1 C \vee H$. That is, $\Gamma_t \cup \overline{C} \cup \overline{H} \vdash_1 \bot$.

Hence $\Gamma_t \cup \beta \vdash_1 \bot$, since $\overline{C} \cup \overline{H} \subseteq \beta$.

But Γ_t is derivable from $BPHP_n$ in resolution in p-width $n/3$.

Width lower bound continued

By Fact 1, to prove the claim it is enough to show $\overline{C}$ **cannot** be extended to a partial matching.

So suppose $\overline{C}$ **can** be extended to a partial matching.

We are given that C is RAT w.r.t. Γ_t and some literal p.

Trick: find a hole axiom of the form $\overline{p} \vee H$ such that $\overline{C} \cup \overline{H}$ can be extended to a partial matching β setting $\leq n/2$ pigeons.

By the RAT assumption, $\Gamma_t \vdash_1 C \vee H$. That is, $\Gamma_t \cup \overline{C} \cup \overline{H} \vdash_1 \perp$.

Hence $\Gamma_t \cup \beta \vdash_1 \perp$, since $\overline{C} \cup \overline{H} \subseteq \beta$.

But Γ_t is derivable from BPHP_n in resolution in p-width $n/3$.

Therefore $\mathrm{BPHP}_n \cup \beta$ is refutable in resolution in p-width $n/3$, contradicting Fact 2. □

Summary of main results

$$\mathsf{Res} < \mathsf{DBC}^- \equiv \mathsf{DRAT}^- \equiv \mathsf{DSPR}^- \equiv \mathsf{DPR}^- \leq \mathsf{DSR}^- \leq \mathsf{ER}$$

$$\mathsf{Res} \; < \; \mathsf{BC}^- \; \leq \; \mathsf{RAT}^- \; < \; \mathsf{SPR}^- \; \leq^* \; \mathsf{PR}^- \; \leq \; \mathsf{SR}^- \; \leq \; \mathsf{ER}$$

The full paper is on Sam's webpage:

`www.math.ucsd.edu/~sbuss/ResearchWeb/DRAT_PR/`