On Division Versus Saturation in Pseudo-Boolean Solving

Stephan Gocht, Jakob Nordström, Amir Yehudayoff

21.05.2019

The SAT Problem

- find assignment to Boolean variables that satisfies constraints
- SAT expressive formalism
  $\Rightarrow$ captures many real-world problems
- theoretically hard, in practice often feasible
- state-of-the-art:
  conflict driven clause learning (CDCL) SAT solvers
  [MS96, BS97, MMZ+01, ...]

The SAT Problem

- find assignment to Boolean variables that satisfies constraints
- SAT expressive formalism
  $\Rightarrow$ captures many real-world problems
- theoretically hard, in practice often feasible
- state-of-the-art:
  conflict driven clause learning (CDCL) SAT solvers
  [MS96, BS97, MMZ+01, ...]

Potential for improvement?

- CDCL SAT solvers essentially based on resolution
- resolution very simple
  + simple data structures allow efficient implementation
  − weak method of reasoning

Pseudo-Boolean SAT Solving

- exponential stronger reasoning via cutting planes proof system
- constraints: pseudo-Boolean (PB) (0-1 linear inequalities)

Pseudo-Boolean SAT Solving

- exponential stronger reasoning via cutting planes proof system
- constraints: pseudo-Boolean (PB) (0-1 linear inequalities)

In practice:

- PB solvers often worse than resolution based solvers
- only implementation details?

Pseudo-Boolean SAT Solving

- exponential stronger reasoning via cutting planes proof system
- constraints: pseudo-Boolean (PB) (0-1 linear inequalities)

In practice:

- PB solvers often worse than resolution based solvers
- only implementation details? no
- different solver use different variants of cutting planes
- none as strong as full cutting planes

$\Rightarrow$ study cutting-planes subsystems used in PB solvers

Method of Reasoning Underlying CDCL: Resolution

Literal a: a variable x or its negation $\bar{x}$
Clause $C = a_1 \vee \cdots \vee a_k$: disjunction ($\vee$) of variables
CNF $F = C_1 \wedge \ldots \wedge C_m$: conjunction ($\wedge$) of clauses

Method of Reasoning Underlying CDCL: Resolution

Literal a: a variable x or its negation $\bar{x}$
Clause $C = a_1 \vee \cdots \vee a_k$: disjunction ($\vee$) of variables
CNF $F = C_1 \wedge \ldots \wedge C_m$: conjunction ($\wedge$) of clauses

Goal: Show F unsatisfiable using:

Resolution rule $\dfrac{C \vee x \qquad \bar{x} \vee D}{C \vee D}$

Example:

$$\dfrac{\dfrac{y \vee x \qquad \bar{x} \vee y}{y} \qquad \bar{y}}{\bot}$$

Method of Reasoning Underlying CDCL: Resolution

Literal a: a variable x or its negation $\bar{x}$
Clause $C = a_1 \vee \cdots \vee a_k$: disjunction ($\vee$) of variables
CNF $F = C_1 \wedge \ldots \wedge C_m$: conjunction ($\wedge$) of clauses

Goal: Show F unsatisfiable using:

Resolution rule $\dfrac{C \vee x \qquad \bar{x} \vee D}{C \vee D}$

Example:

$$\dfrac{\dfrac{y \vee x \qquad \bar{x} \vee y}{y} \qquad \bar{y}}{\bot}$$

- *implicationally complete*, i.e. can drive all consequences
- in particular, can refute all unsatisfiable CNF formulas

Pseudo-Boolean Constraints

- integer linear inequalities over 0-1 variables
- $1 = \text{true}, 0 = \text{false}$

Example: $3\bar{x} + 2\bar{y} + 2z \geq 5$

Normalized Form

- use literals, i.e. $x, \bar{x}$ with $\bar{x} = (1 - x) \Rightarrow x + \bar{x} = 1$
- only positive coefficients and "$\geq$"
- degree (of falsity): right hand side of normalized constraint

Pseudo-Boolean Constraints

- integer linear inequalities over 0-1 variables
- $1 = \text{true}, 0 = \text{false}$

Example: $3\bar{x} + 2\bar{y} + 2z \geq 5$

Normalized Form

- use literals, i.e. $x, \bar{x}$ with $\bar{x} = (1 - x) \Rightarrow x + \bar{x} = 1$
- only positive coefficients and "$\geq$"
- degree (of falsity): right hand side of normalized constraint

Representing Clauses:

$$\bar{x} \vee y \vee z \quad \rightsquigarrow \quad \bar{x} + y + z \geq 1$$

Cardinality constraints, e.g at-least 2:

$$(x \vee y) \wedge (x \vee z) \wedge (y \vee z) \quad \rightsquigarrow \quad x + y + z \geq 2$$

Cutting Planes: Linear Combination

Literal Axioms

$$\frac{}{x \geq 0} \quad \frac{}{\bar{x} \geq 0}$$

Cutting Planes: Linear Combination

Literal Axioms

$$\frac{}{x \geq 0} \qquad \frac{}{\bar{x} \geq 0}$$

Positive Linear Combination — Remember: $x + \bar{x} = 1$

$$\frac{3 \cdot (y + z + \bar{x} \geq 1) \qquad 1 \cdot (2x + z \geq 3)}{3y + 4z + \underbrace{2x + 3\bar{x}}_{=2+\bar{x}} \geq 6}$$

Cutting Planes: Linear Combination

Literal Axioms

$$\frac{}{x \geq 0} \qquad \frac{}{\bar{x} \geq 0}$$

Positive Linear Combination — Remember: $x + \bar{x} = 1$

$$\frac{3 \cdot (y + z + \bar{x} \geq 1) \qquad 1 \cdot (2x + z \geq 3)}{3y + 4z + \qquad \bar{x} \geq 4}$$

Cutting Planes: Linear Combination

Literal Axioms

$$\frac{}{x \geq 0} \qquad \frac{}{\bar{x} \geq 0}$$

Positive Linear Combination — Remember: $x + \bar{x} = 1$

$$\frac{3 \cdot (y + z + \bar{x} \geq 1) \qquad 1 \cdot (2x + z \geq 3)}{3y + 4z + \qquad \bar{x} \geq 4}$$

Generalized Resolution

$$\frac{2\bar{x} + v + \overline{w} \geq 2 \qquad 2x + y + z \geq 2}{v + \overline{w} + y + z \geq 2}$$

Cutting Planes: Boolean Rule

Division (divide and round up)

$$\frac{x + 2y + 2z \geq 3}{x + y + z \geq 2} \text{ Divide by 2}$$

Cutting Planes: Boolean Rule

Division (divide and round up)

$$\frac{x + 2y + 2z \geq 3}{x + y + z \geq 2} \text{ Divide by 2}$$

Saturation (set coefficient to value $\leq$ degree of falsity)

$$\frac{4x + 4\bar{y} + z \geq 2}{2x + 2\bar{y} + z \geq 2}$$

Example

$$\bar{x}_1 + \bar{x}_2 \geq 1 \qquad \bar{x}_1 + \bar{x}_3 \geq 1$$

$$\bar{x}_2 + \bar{x}_3 \geq 1$$

$$x_1 + x_2 + x_3 \geq 2$$

Example

$$\dfrac{\dfrac{\dfrac{\bar{x}_1 + \bar{x}_2 \geq 1 \qquad \bar{x}_1 + \bar{x}_3 \geq 1}{2\bar{x}_1 + \bar{x}_2 + \bar{x}_3 \geq 2} \qquad \bar{x}_2 + \bar{x}_3 \geq 1}{}}{} \qquad x_1 + x_2 + x_3 \geq 2$$

Example

$$\cfrac{\cfrac{\cfrac{\cfrac{\bar{x}_1 + \bar{x}_2 \geq 1 \qquad \bar{x}_1 + \bar{x}_3 \geq 1}{2\bar{x}_1 + \bar{x}_2 + \bar{x}_3 \geq 2} \qquad \bar{x}_2 + \bar{x}_3 \geq 1}{2\bar{x}_1 + 2\bar{x}_2 + 2\bar{x}_3 \geq 3}}{} \qquad x_1 + x_2 + x_3 \geq 2}{}$$

Example

$$\cfrac{\cfrac{\cfrac{\cfrac{\bar{x}_1 + \bar{x}_2 \geq 1 \qquad \bar{x}_1 + \bar{x}_3 \geq 1}{2\bar{x}_1 + \bar{x}_2 + \bar{x}_3 \geq 2} \qquad \bar{x}_2 + \bar{x}_3 \geq 1}{2\bar{x}_1 + 2\bar{x}_2 + 2\bar{x}_3 \geq 3}}{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 \geq 2} \qquad x_1 + x_2 + x_3 \geq 2}{}$$

Example

$$\dfrac{\dfrac{\dfrac{\dfrac{\bar{x}_1 + \bar{x}_2 \geq 1 \qquad \bar{x}_1 + \bar{x}_3 \geq 1}{2\bar{x}_1 + \bar{x}_2 + \bar{x}_3 \geq 2} \qquad \bar{x}_2 + \bar{x}_3 \geq 1}{2\bar{x}_1 + 2\bar{x}_2 + 2\bar{x}_3 \geq 3}}{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 \geq 2} \qquad x_1 + x_2 + x_3 \geq 2}{0 \geq 1}$$

Pseudo-Boolean Solvers and the Subsystem Used

not all rules used by implementations

generalized resolution and saturation:

- PRS [DG02]
- Galena [CK05]
- Pueblo [SS06]
- Sat4j [LP10]

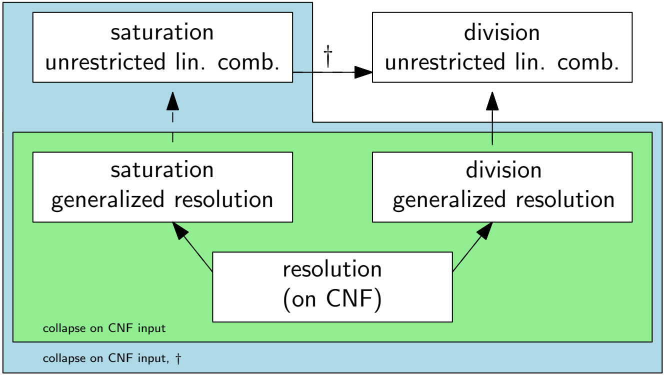
generalized resolution and division:

- RoundingSat [EN18]

linear combination and division (full cutting planes):

- $\emptyset$

Systematic Overview [VEG+18]



$\mathcal{A}$ - - ► $\mathcal{B}$ $\mathcal{B}$ can do everything $\mathcal{A}$ can

$\mathcal{A}$ ——► $\mathcal{B}$ $\mathcal{B}$ can do everything $\mathcal{A}$ can and $\mathcal{B}$ can do things $\mathcal{A}$ can't

† polynomial-sized coefficients

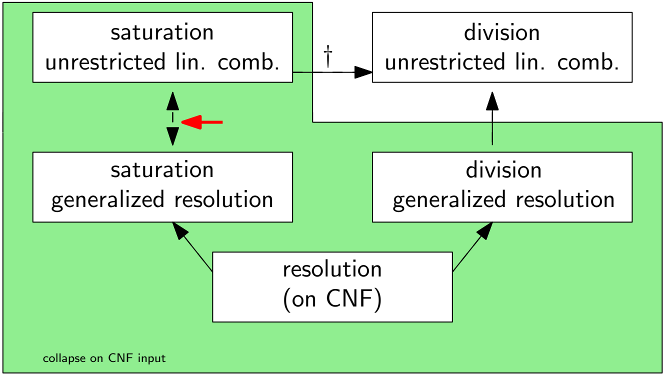
Our Results

We want to study reasoning power depending on choice of:

- boolean rule: (a) division, (b) saturation
- linear combination: (a) generalized resolution, (b) unrestricted

- saturation as boolean rule $\Rightarrow$ generalized resolution as powerful as unrestricted linear combinations
- division + generalized resolution can be exponentially stronger than saturation + unrestricted linear combinations
- replacing single saturation, requires large # divisions

Our Result: Strength of Generalized Resolution



$\mathcal{A}$ - - ► $\mathcal{B}$ $\mathcal{B}$ can do everything $\mathcal{A}$ can

$\mathcal{A}$ ——► $\mathcal{B}$ $\mathcal{B}$ can do everything $\mathcal{A}$ can and $\mathcal{B}$ can do things $\mathcal{A}$ can't

† polynomial-sized coefficients

Proof Sketch: “Rewriting” to Generalized Resolution

$$\frac{\bar{x} + 2y \geq 2 \qquad 2x + 2y + z \geq 2}{x + 4y + z \geq 3}$$

- rewrite as generalized resolution:

Proof Sketch: "Rewriting" to Generalized Resolution

$$\frac{\bar{x} + 2y \geq 2 \qquad 2x + 2y + z \geq 2}{x + 4y + z \geq 3}$$

- rewrite as generalized resolution:

$$\text{generalized res.}\ \frac{2 \cdot (\bar{x} + 2y \geq 2) \qquad 2x + 2y + z \geq 2}{6y + z \geq 4}$$

Proof Sketch: “Rewriting” to Generalized Resolution

$$\frac{\bar{x} + 2y \geq 2 \qquad 2x + 2y + z \geq 2}{x + 4y + z \geq 3}$$

- rewrite as generalized resolution:

$$\text{postponed step}\ \frac{\text{generalized res.}\ \dfrac{2 \cdot (\bar{x} + 2y \geq 2) \qquad 2x + 2y + z \geq 2}{6y + z \geq 4} \qquad 2x + 2y + z \geq 2}{2x + 8y + 2z \geq 6}$$

- postpone addition of constraints that can’t be rewritten

Proof Sketch: "Rewriting" to Generalized Resolution

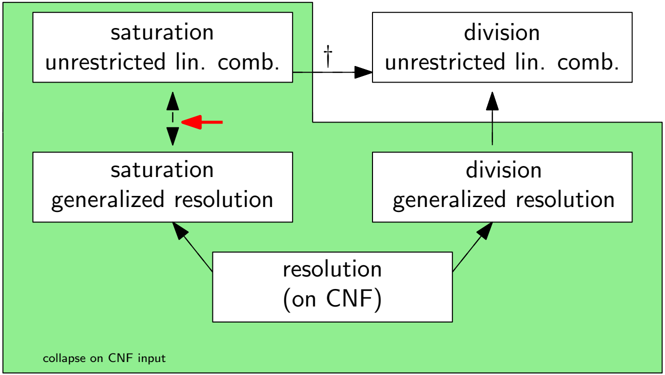
$$\frac{\bar{x} + 2y \geq 2 \qquad 2x + 2y + z \geq 2}{x + 3y + z \geq 3}$$

- rewrite as generalized resolution:

$$\text{postponed step}\ \frac{\text{generalized res.}\ \dfrac{2 \cdot (\bar{x} + 2y \geq 2) \qquad 2x + 2y + z \geq 2}{4y + z \geq 4} \qquad 2x + 2y + z \geq 2}{2x + 6y + 2z \geq 6}$$

- postpone addition of constraints that can't be rewritten
- saturation not affected by postponing

Our Result: Strength of Generalized Resolution

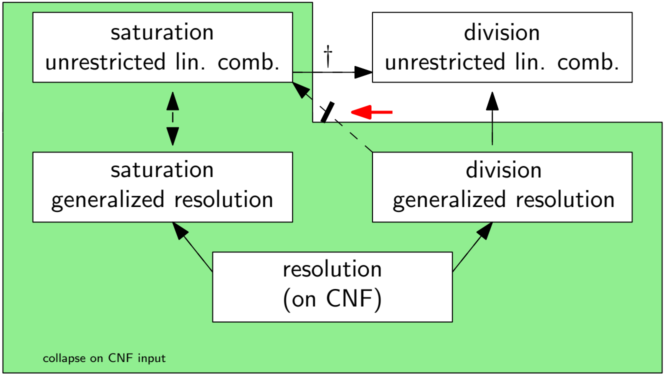


$\mathcal{A}$ - - ► $\mathcal{B}$ $\mathcal{B}$ can do everything $\mathcal{A}$ can

$\mathcal{A}$ ——► $\mathcal{B}$ $\mathcal{B}$ can do everything $\mathcal{A}$ can and $\mathcal{B}$ can do things $\mathcal{A}$ can't

†   polynomial-sized coefficients

Our Result: Strength of Division



$\mathcal{A}$ - - ► $\mathcal{B}$ $\mathcal{B}$ can do everything $\mathcal{A}$ can

$\mathcal{A}$ ——► $\mathcal{B}$ $\mathcal{B}$ can do everything $\mathcal{A}$ can and $\mathcal{B}$ can do things $\mathcal{A}$ can't

† polynomial-sized coefficients -/► negation

Strength of Division

- cutting planes stronger than resolution because it can “count”
- requires division + unrestricted linear combination
- for example, recovering cardinality constraints:

$$\begin{array}{r} \bar{x} + \bar{y} \phantom{+ \bar{z}} \geq 1 \\ \bar{y} + \bar{z} \geq 1 \\ \bar{x} + \phantom{\bar{y} +} \bar{z} \geq 1 \\ \hline 2\bar{x} + 2\bar{y} + 2\bar{z} \geq 3 \\ \hline \bar{x} + \bar{y} + \bar{z} \geq 2 \end{array}$$

Strength of Division

- cutting planes stronger than resolution because it can "count"
- requires division + unrestricted linear combination
- for example, recovering cardinality constraints:

$$\begin{array}{rrrrl}
h_1 + h_2 + & \bar{x} + & \bar{y} & & \geq 1 \\
\bar{h}_1 + & & \bar{y} + & \bar{z} & \geq 2 \\
\bar{h}_2 + & \bar{x} + & & \bar{z} & \geq 2 \\
\hline
& 2\bar{x} + & 2\bar{y} + & 2\bar{z} & \geq 3 \\
\hline
& \bar{x} + & \bar{y} + & \bar{z} & \geq 2
\end{array}$$

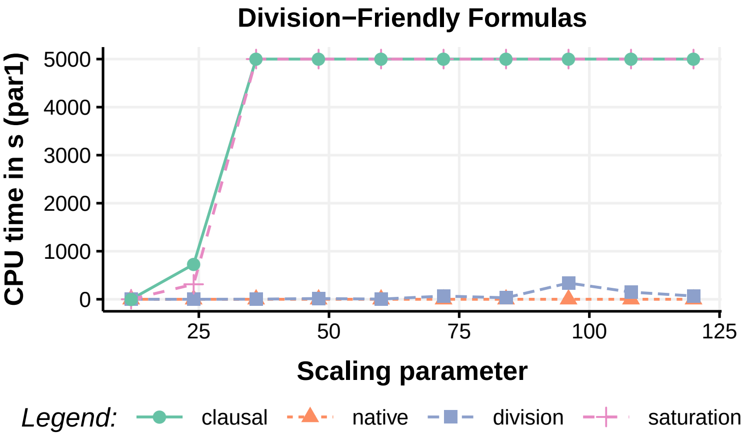
Achieve separation by modifying benchmark:

- introduce helper variables to allow generalized resolution
- still hard for saturation, easy for division + generalized res.

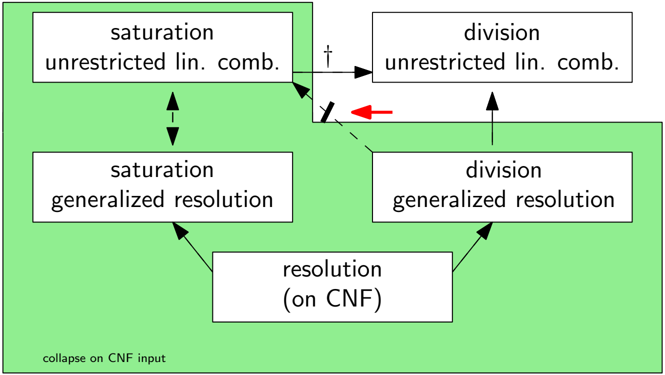
Practical Result for Division-Friendly Formulas

- apply "trick" to subset cardinality formulas

  [Spe10, VS10, MN14]



Our Result: Strength of Division

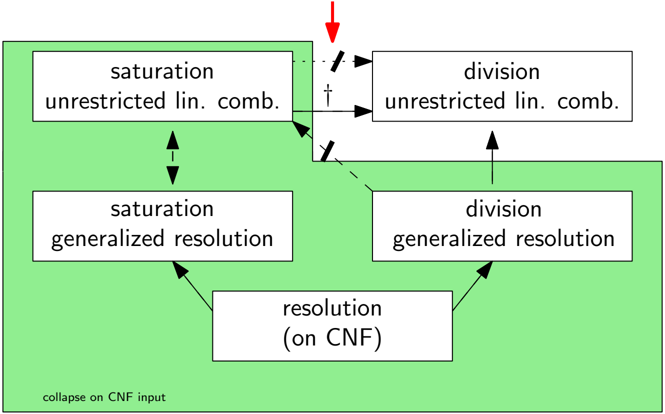


$\mathcal{A}$ - - ► $\mathcal{B}$ $\mathcal{B}$ can do everything $\mathcal{A}$ can

$\mathcal{A}$ ——► $\mathcal{B}$ $\mathcal{B}$ can do everything $\mathcal{A}$ can and $\mathcal{B}$ can do things $\mathcal{A}$ can't

† polynomial-sized coefficients -/► negation

Our Result: Strength of Saturation



$\mathcal{A}$ - - ► $\mathcal{B}$ $\mathcal{B}$ can do everything $\mathcal{A}$ can

$\mathcal{A}$ ——► $\mathcal{B}$ $\mathcal{B}$ can do everything $\mathcal{A}$ can and $\mathcal{B}$ can do things $\mathcal{A}$ can't

† polynomial-sized coefficients -/-► negation

Simulating Saturation with Division, Lower Bound

To replace one saturation step

$$\frac{2Rx + \sum_{i=1}^{2R} z_i \geq R}{Rx + \sum_{i=1}^{2R} z_i \geq R}$$

by division. . .

Simulating Saturation with Division, Lower Bound

To replace one saturation step

$$\frac{2Rx + \sum_{i=1}^{2R} z_i \geq R}{Rx + \sum_{i=1}^{2R} z_i \geq R}$$

by division...

- it takes $\Omega(\sqrt{R})$ division steps
- still true if we add generalized resolution step to obtain unsaturated constraint, i.e., start from
  $Rx + Ry + \sum_{i=1}^{R} z_i \geq R$ $\quad$ $Rx + R\bar{y} + \sum_{i=R+1}^{2R} z_i \geq R$
- does not show that cutting planes with saturation can be exponentially stronger than cutting planes with division

Proof Sketch

define potential function $\mathcal{P}(C)$ such that:

- needs to change:
  $\mathcal{P}(C_{\text{start}}) - \mathcal{P}(C_{\text{end}}) \geq 1/6$
- doesn't change with linear combination:
  $\mathcal{P}(C_1 + C_2) \geq \min\{\mathcal{P}(C_1), \mathcal{P}(C_2)\}$
- changes by a small amount by division:
  $\mathcal{P}(C/k) \geq \mathcal{P}(C) - 1/\sqrt{R}$

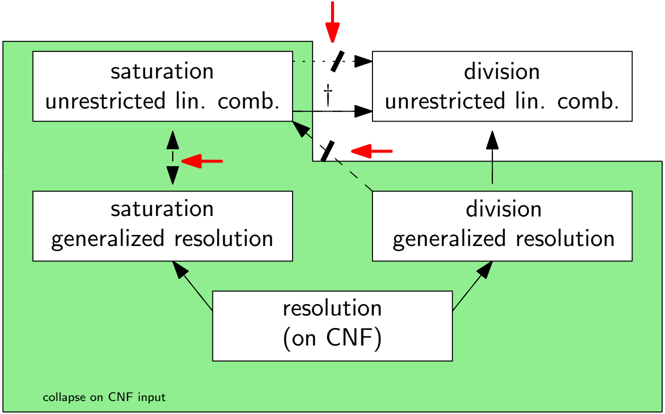
Proof Sketch

define potential function $\mathcal{P}(C)$ such that:

- needs to change:
  $\mathcal{P}(C_{\text{start}}) - \mathcal{P}(C_{\text{end}}) \geq 1/6$
- doesn't change with linear combination:
  $\mathcal{P}(C_1 + C_2) \geq \min\{\mathcal{P}(C_1), \mathcal{P}(C_2)\}$
- changes by a small amount by division:
  $\mathcal{P}(C/k) \geq \mathcal{P}(C) - 1/\sqrt{R}$

$$\mathcal{P}(ax + \sum b_i z_i \geq A) := \ln(a/A)$$

Conclusion

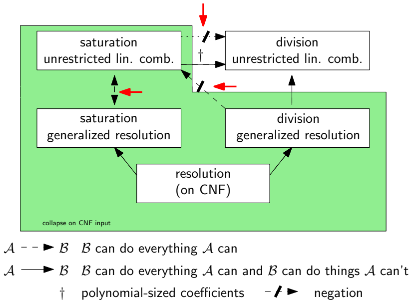


$\mathcal{A}$ - - ► $\mathcal{B}$ — $\mathcal{B}$ can do everything $\mathcal{A}$ can

$\mathcal{A}$ ──► $\mathcal{B}$ — $\mathcal{B}$ can do everything $\mathcal{A}$ can and $\mathcal{B}$ can do things $\mathcal{A}$ can't

† polynomial-sized coefficients — -/-► negation

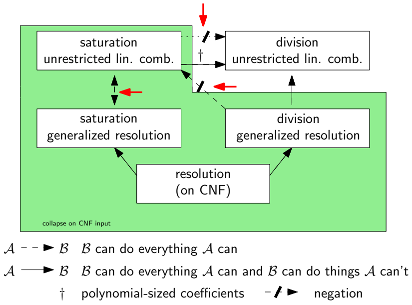
Conclusion



Future Research Directions

- Can saturation be stronger than division in proving UNSAT?
- implement adaptive choice between division and saturation
- to supersede resolution on CNF we need "natural way" / heuristic for unrestricted linear combination + divisions

Conclusion



Future Research Directions

- Can saturation be stronger than division in proving UNSAT?
- implement adaptive choice between division and saturation
- to supersede resolution on CNF we need "natural way" / heuristic for unrestricted linear combination + divisions

Thank you for your attention!

Simulating Saturation with Division

- saturation can be simulated by repeated division

$$\cfrac{\cfrac{\cfrac{\cfrac{49x + 50y + 51z + 200y \geq 100}{49x + 50y + 51z + 199y \geq 100} \times 100 \div 101}{49x + 50y + 51z + 198y \geq 100} \times 100 \div 101}{49x + 50y + 51z + 197y \geq 100} \times 100 \div 101}{\cdots} \times 100 \div 101$$

- only guaranteed to reduce coefficient by 1 per iteration
  $\Rightarrow$ only efficient if coefficients small

References I

[BS97] Roberto J. Bayardo Jr. and Robert Schrag.
Using CSP look-back techniques to solve real-world SAT instances.
In *Proceedings of the 14th National Conference on Artificial Intelligence (AAAI '97)*, pages 203–208, July 1997.

[CK05] Donald Chai and Andreas Kuehlmann.
A fast pseudo-Boolean constraint solver.
IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 24(3):305–317, March 2005.
Preliminary version in *DAC '03*.

[DG02] Heidi E. Dixon and Matthew L. Ginsberg.
Inference methods for a pseudo-Boolean satisfiability solver.
In *Proceedings of the 18th National Conference on Artificial Intelligence (AAAI '02)*, pages 635–640, July 2002.

[EN18] Jan Elffers and Jakob Nordström.
Divide and conquer: Towards faster pseudo-Boolean solving.
In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI '18)*, pages 1291–1299, July 2018.

References II

[LP10] Daniel Le Berre and Anne Parrain.
The Sat4j library, release 2.2.
Journal on Satisfiability, Boolean Modeling and Computation, 7:59–64, 2010.

[MMZ+01] Matthew W. Moskewicz, Conor F. Madigan, Ying Zhao, Lintao Zhang, and Sharad Malik.
Chaff: Engineering an efficient SAT solver.
In *Proceedings of the 38th Design Automation Conference (DAC '01)*, pages 530–535, June 2001.

[MN14] Mladen Mikša and Jakob Nordström.
Long proofs of (seemingly) simple formulas.
In *Proceedings of the 17th International Conference on Theory and Applications of Satisfiability Testing (SAT '14)*, volume 8561 of *Lecture Notes in Computer Science*, pages 121–137. Springer, July 2014.

[MS96] João P. Marques-Silva and Karem A. Sakallah.
GRASP—a new search algorithm for satisfiability.
In *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design (ICCAD '96)*, pages 220–227, November 1996.

[Spe10] Ivor Spence.
sgen1: A generator of small but difficult satisfiability benchmarks.
Journal of Experimental Algorithmics, 15:1.2:1–1.2:15, March 2010.

References III

[SS06] Hossein M. Sheini and Karem A. Sakallah.
Pueblo: A hybrid pseudo-Boolean SAT solver.
Journal on Satisfiability, Boolean Modeling and Computation, 2(1-4):165–189, March 2006.
Preliminary version in *DATE '05*.

[VEG+18] Marc Vinyals, Jan Elffers, Jesús Giráldez-Cru, Stephan Gocht, and Jakob Nordström.
In between resolution and cutting planes: A study of proof systems for pseudo-Boolean SAT solving.
In *Proceedings of the 21st International Conference on Theory and Applications of Satisfiability Testing (SAT '18)*, volume 10929 of *Lecture Notes in Computer Science*, pages 292–310. Springer, July 2018.

[VS10] Allen Van Gelder and Ivor Spence.
Zero-one designs produce small hard SAT instances.
In *Proceedings of the 13th International Conference on Theory and Applications of Satisfiability Testing (SAT '10)*, volume 6175 of *Lecture Notes in Computer Science*, pages 388–397. Springer, July 2010.