Formal Methods and the Chromatic Number of the Plane

Marijn J.H. Heule

Carnegie Mellon University

Formal Methods in Mathematics January 8, 2020

Computer-Aided Mathematics

Chromatic Number of the Plane

Clausal Proof Optimization

Observed Patterns in $\mathbb{Q}[\sqrt{3}, \sqrt{11}] \times \mathbb{Q}[\sqrt{3}, \sqrt{11}]$

Small UD Graphs with Chromatic Number 5

Conclusions and Future Work

40 Years of Successes in Computer-Aided Mathematics

1976 Four-Color Theorem

1998 Kepler Conjecture

2010 "God's Number = 20": Optimal Rubik's cube strategy

2012 At least 17 clues for a solvable Sudoku puzzle

2014 Boolean Erdős discrepancy problem

2016 Boolean Pythagorean triples problem

2018 Schur Number Five

2019 Keller's Conjecture

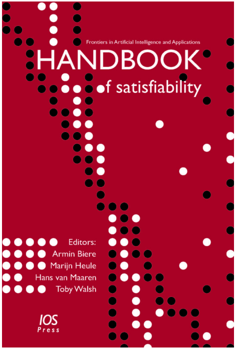
40 Years of Successes in Computer-Aided Mathematics

- 1976 Four-Color Theorem
- 1998 Kepler Conjecture
- 2010 "God's Number = 20": Optimal Rubik's cube strategy
- 2012 At least 17 clues for a solvable Sudoku puzzle
- 2014 Boolean Erdős discrepancy problem (using a SAT solver)
- 2016 Boolean Pythagorean triples problem (using a SAT solver)
- 2018 Schur Number Five (using a SAT solver)
- 2019 Keller's Conjecture (using a SAT solver)

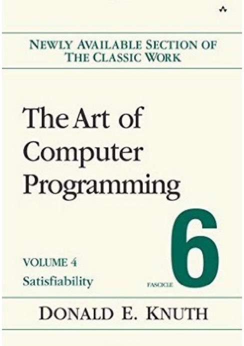
Breakthrough in SAT Solving in the Last 20 Years

Satisfiability (SAT) problem: Can a Boolean formula be satisfied?

mid '90s: formulas solvable with thousands of variables and clauses
now: formulas solvable with millions of variables and clauses



Edmund Clarke: *"a key technology of the 21st century"* [Biere, Heule, vanMaaren, Walsh '09]



Donald Knuth: *"evidently a killer app, because it is key to the solution of so many other problems"* [Knuth '15]

Pythagorean Triples Problem (I) [Ronald Graham, early 80's]

Will any coloring of the positive integers with red and blue result in a monochromatic Pythagorean Triple $a^2 + b^2 = c^2$?

$3^2 + 4^2 = 5^2$	$6^2 + 8^2 = 10^2$	$5^2 + 12^2 = 13^2$	$9^2 + 12^2 = 15^2$
$8^2 + 15^2 = 17^2$	$12^2 + 16^2 = 20^2$	$15^2 + 20^2 = 25^2$	$7^2 + 24^2 = 25^2$
$10^2 + 24^2 = 26^2$	$20^2 + 21^2 = 29^2$	$18^2 + 24^2 = 30^2$	$16^2 + 30^2 = 34^2$
$21^2 + 28^2 = 35^2$	$12^2 + 35^2 = 37^2$	$15^2 + 36^2 = 39^2$	$24^2 + 32^2 = 40^2$

Pythagorean Triples Problem (I) [Ronald Graham, early 80's]

Will any coloring of the positive integers with red and blue result in a monochromatic Pythagorean Triple $a^2 + b^2 = c^2$?

$3^2 + 4^2 = 5^2$	$6^2 + 8^2 = 10^2$	$5^2 + 12^2 = 13^2$	$9^2 + 12^2 = 15^2$
$8^2 + 15^2 = 17^2$	$12^2 + 16^2 = 20^2$	$15^2 + 20^2 = 25^2$	$7^2 + 24^2 = 25^2$
$10^2 + 24^2 = 26^2$	$20^2 + 21^2 = 29^2$	$18^2 + 24^2 = 30^2$	$16^2 + 30^2 = 34^2$
$21^2 + 28^2 = 35^2$	$12^2 + 35^2 = 37^2$	$15^2 + 36^2 = 39^2$	$24^2 + 32^2 = 40^2$

Best lower bound: a bi-coloring of $[1, 7664]$ s.t. there is no monochromatic Pythagorean Triple [Cooper & Overstreet 2015].

Myers conjectures that the answer is No [PhD thesis, 2015].

Pythagorean Triples Problem (II) [Ronald Graham, early 80's]

Will any coloring of the positive integers with red and blue result in a monochromatic Pythagorean Triple $a^2 + b^2 = c^2$?

A bi-coloring of $[1, n]$ is encoded using Boolean variables x_i with $i \in \{1, 2, \ldots, n\}$ such that $x_i = 1$ $(= 0)$ means that i is colored red (blue). For each Pythagorean Triple $a^2 + b^2 = c^2$, two clauses are added: $(x_a \vee x_b \vee x_c)$ and $(\overline{x}_a \vee \overline{x}_b \vee \overline{x}_c)$.

Pythagorean Triples Problem (II) [Ronald Graham, early 80's]

Will any coloring of the positive integers with red and blue result in a monochromatic Pythagorean Triple $a^2 + b^2 = c^2$?

A bi-coloring of $[1, n]$ is encoded using Boolean variables x_i with $i \in \{1, 2, \ldots, n\}$ such that $x_i = 1$ $(= 0)$ means that i is colored red (blue). For each Pythagorean Triple $a^2 + b^2 = c^2$, two clauses are added: $(x_a \vee x_b \vee x_c)$ and $(\overline{x}_a \vee \overline{x}_b \vee \overline{x}_c)$.

Theorem ([Heule, Kullmann, and Marek (2016)])

$[1, 7824]$ *can be bi-colored s.t. there is no monochromatic Pythagorean Triple. This is impossible for* $[1, 7825]$.

Pythagorean Triples Problem (II) [Ronald Graham, early 80's]

Will any coloring of the positive integers with red and blue result in a monochromatic Pythagorean Triple $a^2 + b^2 = c^2$?

A bi-coloring of $[1, n]$ is encoded using Boolean variables x_i with $i \in \{1, 2, \ldots, n\}$ such that $x_i = 1$ $(= 0)$ means that i is colored red (blue). For each Pythagorean Triple $a^2 + b^2 = c^2$, two clauses are added: $(x_a \vee x_b \vee x_c)$ and $(\overline{x}_a \vee \overline{x}_b \vee \overline{x}_c)$.

Theorem ([Heule, Kullmann, and Marek (2016)])

$[1, 7824]$ *can be bi-colored s.t. there is no monochromatic Pythagorean Triple. This is impossible for* $[1, 7825]$.

4 CPU years computation, but 2 days on cluster (800 cores)

Pythagorean Triples Problem (II) [Ronald Graham, early 80's]

Will any coloring of the positive integers with red and blue result in a monochromatic Pythagorean Triple $a^2 + b^2 = c^2$?

A bi-coloring of $[1, n]$ is encoded using Boolean variables x_i with $i \in \{1, 2, \ldots, n\}$ such that $x_i = 1$ $(= 0)$ means that i is colored red (blue). For each Pythagorean Triple $a^2 + b^2 = c^2$, two clauses are added: $(x_a \vee x_b \vee x_c)$ and $(\overline{x}_a \vee \overline{x}_b \vee \overline{x}_c)$.

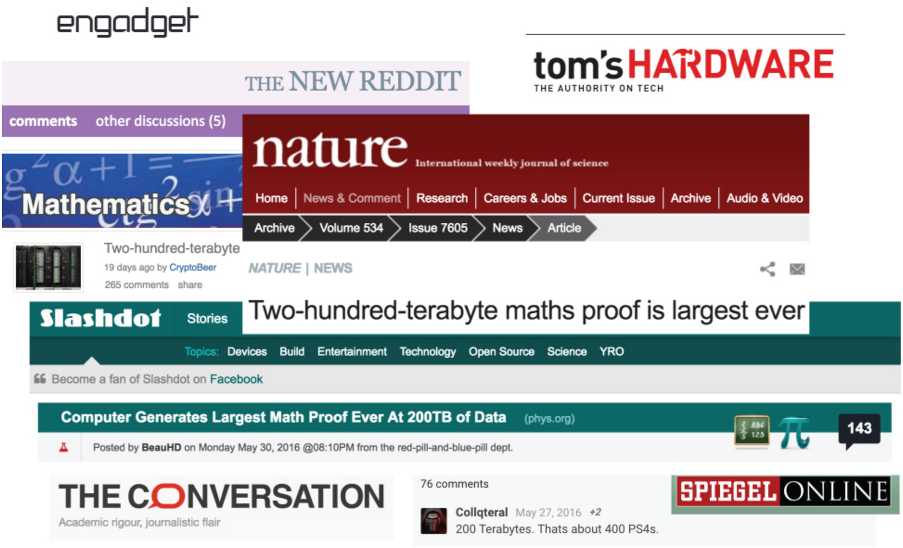
Theorem ([Heule, Kullmann, and Marek (2016)])

$[1, 7824]$ *can be bi-colored s.t. there is no monochromatic Pythagorean Triple. This is impossible for* $[1, 7825]$.

4 CPU years computation, but 2 days on cluster (800 cores)

200 terabytes proof, but validated with verified checker

Media: "The Largest Math Proof Ever"



Computer-Aided Mathematics Technologies

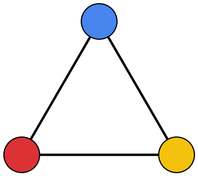
Fields Medalist Timothy Gowers stated that mathematicians would like to use three kinds of technology [Big Proof 2017]:

- Proof Assistant Technology
  - Prove any lemma that a graduate student can work out
- Proof Checking Technology
  - Mechanized validation of all details
- Proof Search Technology
  - Automatically determine whether a conjecture holds
  - This talk: Find small counter-examples

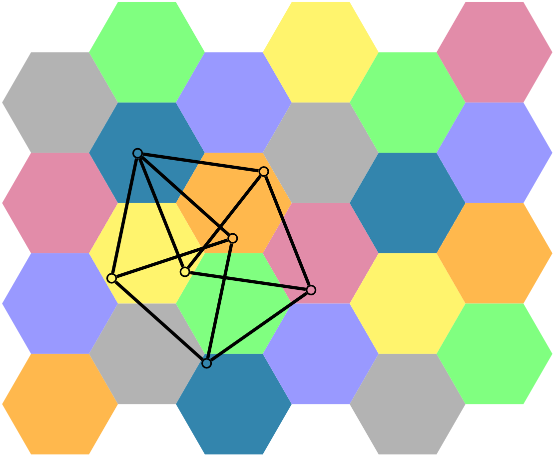
Chromatic Number of the Plane

The Hadwiger-Nelson problem:
How many colors are required to color the plane such that each pair of points that are exactly 1 apart are colored differently?

The answer must be three or more because three points can be mutually 1 apart—and thus must be colored differently.



Bounds since the 1950s



- The Moser Spindle graph shows the lower bound of 4
- A coloring of the plane showing the upper bound of 7

First progress in decades

Recently enormous progress:

- Lower bound of 5 [DeGrey '18] based on a 1581-vertex graph
- This breakthrough started a polymath project
- Improved bounds of the fractional chromatic number of the plane

First progress in decades

Recently enormous progress:

- Lower bound of 5 [DeGrey '18] based on a 1581-vertex graph
- This breakthrough started a polymath project
- Improved bounds of the fractional chromatic number of the plane

We found smaller graphs with SAT:

- 874 vertices on April 14, 2018
- 803 vertices on April 30, 2018
- 610 vertices on May 14, 2018

Validation

Check 1: Are two given points exactly 1 apart? For example:

- $\left(\frac{19+3\sqrt{5}}{16}, \frac{5\sqrt{15}-7\sqrt{3}}{16}\right)$
- $\left(\frac{135+21\sqrt{5}-7\sqrt{33}+3\sqrt{165}}{96}, \frac{33\sqrt{15}-49\sqrt{3}-21\sqrt{11}-3\sqrt{55}}{96}\right)$

Our method: An approach based on Groebner basis theory developed by Armin Biere, Manuel Kauers, Daniela Ritirc

Validation

Check 1: Are two given points exactly 1 apart? For example:

- $\left(\frac{19+3\sqrt{5}}{16}, \frac{5\sqrt{15}-7\sqrt{3}}{16}\right)$
- $\left(\frac{135+21\sqrt{5}-7\sqrt{33}+3\sqrt{165}}{96}, \frac{33\sqrt{15}-49\sqrt{3}-21\sqrt{11}-3\sqrt{55}}{96}\right)$

Our method: An approach based on Groebner basis theory developed by Armin Biere, Manuel Kauers, Daniela Ritirc

Check 2: Given a graph G, has it chromatic number k?

Our method: Construct two Boolean formulas: one asking whether G can be colored with $k-1$ colors (must be UNSAT) and one asking whether G can be colored with k colors (SAT).

Validation

Check 1: Are two given points exactly 1 apart? For example:

- $\left(\frac{19+3\sqrt{5}}{16}, \frac{5\sqrt{15}-7\sqrt{3}}{16}\right)$
- $\left(\frac{135+21\sqrt{5}-7\sqrt{33}+3\sqrt{165}}{96}, \frac{33\sqrt{15}-49\sqrt{3}-21\sqrt{11}-3\sqrt{55}}{96}\right)$

Our method: An approach based on Groebner basis theory developed by Armin Biere, Manuel Kauers, Daniela Ritirc

Check 2: Given a graph G, has it chromatic number k?

Our method: Construct two Boolean formulas: one asking whether G can be colored with $k-1$ colors (must be UNSAT) and one asking whether G can be colored with k colors (SAT).

Validation can provide more than correctness

Extracting Subgraphs from a Proof of Unsatisfiability

The validation method to check whether a graph has (at least) chromatic number k construct a SAT formula asking whether the graph G can be colored with $k - 1$ colors.

The resulting formula is unsatisfiable.

Most SAT solvers can emit a proof of unsatisfiability.

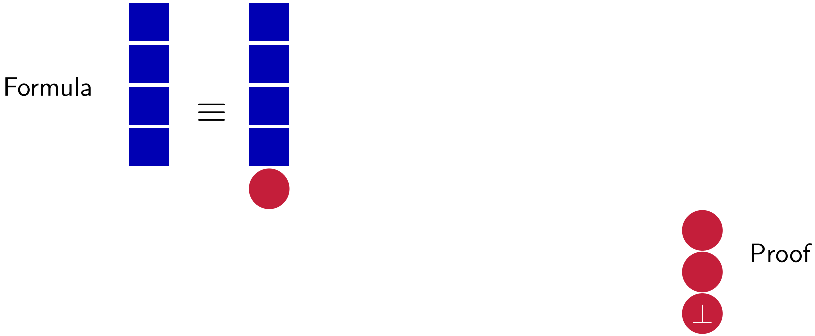
Proof checkers can extract an unsatisfiable core of the problem, which represents a subgraph of G.

Clausal Proofs of Unsatisfiability

Clause C is redundant w.r.t. formula F if F and $F \wedge C$ are equisatisfiable

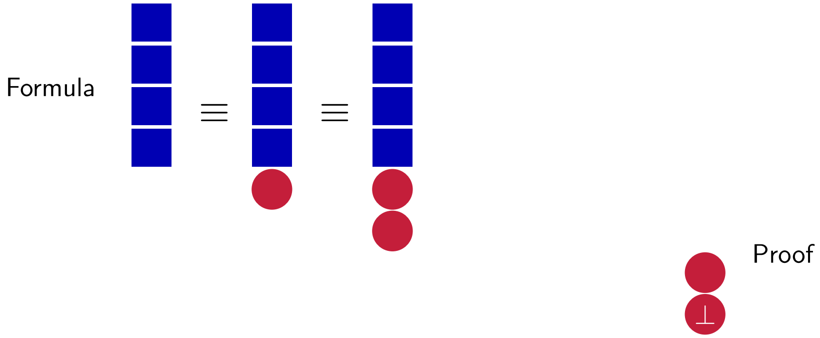
Clausal Proofs of Unsatisfiability

Clause C is redundant w.r.t. formula F if F and $F \land C$ are equisatisfiable



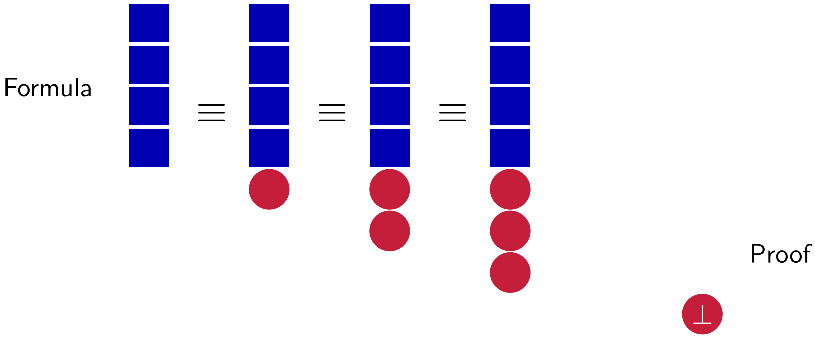
Clausal Proofs of Unsatisfiability

Clause C is redundant w.r.t. formula F if F and $F \wedge C$ are equisatisfiable



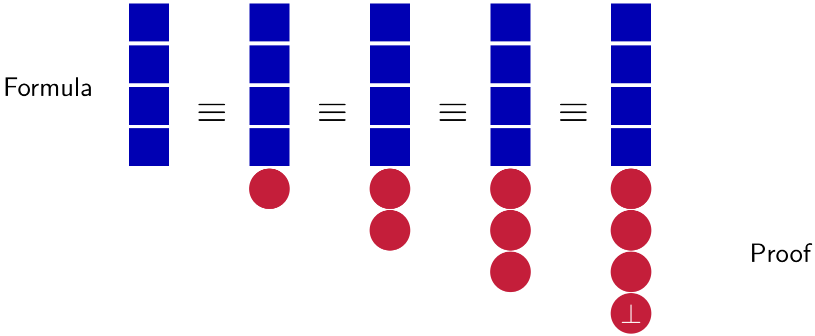
Clausal Proofs of Unsatisfiability

Clause C is redundant w.r.t. formula F if F and $F \wedge C$ are equisatisfiable



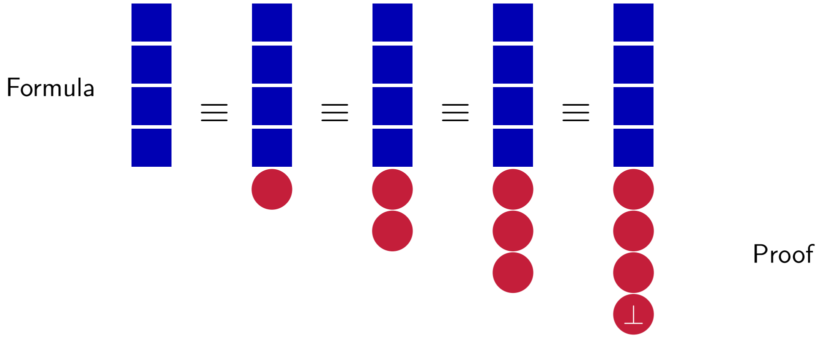
Clausal Proofs of Unsatisfiability

Clause C is redundant w.r.t. formula F if F and $F \land C$ are equisatisfiable



Clausal Proofs of Unsatisfiability

Clause C is redundant w.r.t. formula F if F and $F \wedge C$ are equisatisfiable



- Checking the redundancy of a clause in polynomial time
- Clausal proofs are easy to emit from modern SAT solvers
- A clausal proof usually covers many resolution proofs

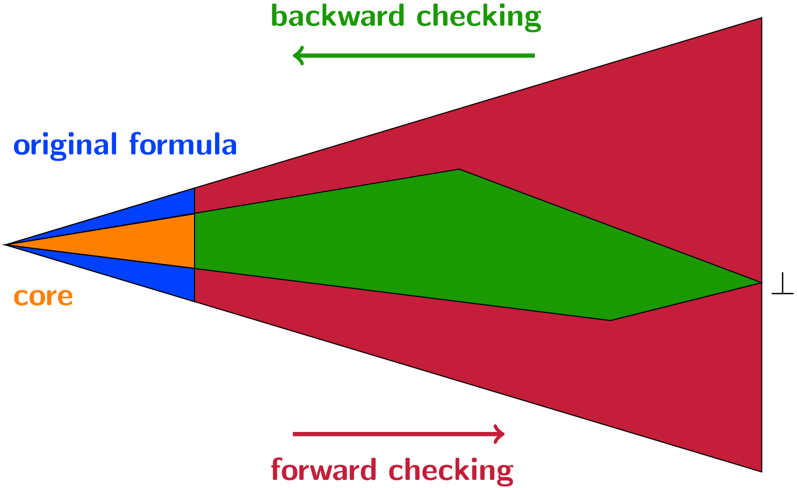
Proof Checking Techniques Advances

Proof checking techniques have improved significantly in recent years.

Clausal proofs of petabytes is size can now be validated reasonably efficiently, even with formally-verified checkers.

Long-standing open math problems —including the Erdős discrepancy problem, the Boolean Pythagorean triples problem, and Schur number five— have solved with SAT and their proofs have been constructed and validated.

Backward Proof Checking: Remove Redundancy



OptimizeProof

The order of the clauses in the proof and the order of the literals in clauses have a big impact on reduced proof.

- Optimize the proof by checking it multiple times;
- Each iteration uses the reduced proof; and
- Clauses are literals are shuffled.

OptimizeProof

The order of the clauses in the proof and the order of the literals in clauses have a big impact on reduced proof.

- Optimize the proof by checking it multiple times;
- Each iteration uses the reduced proof; and
- Clauses are literals are shuffled.

Shuffling of clauses is somewhat limited:

- A clause must occur after all clauses on which it depends;
- A clause must occur before all clauses that depend on it.

OptimizeProof

The order of the clauses in the proof and the order of the literals in clauses have a big impact on reduced proof.

- Optimize the proof by checking it multiple times;
- Each iteration uses the reduced proof; and
- Clauses are literals are shuffled.

Shuffling of clauses is somewhat limited:

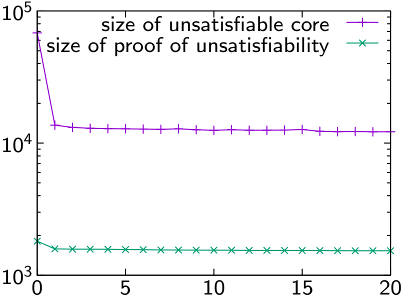
- A clause must occur after all clauses on which it depends;
- A clause must occur before all clauses that depend on it.

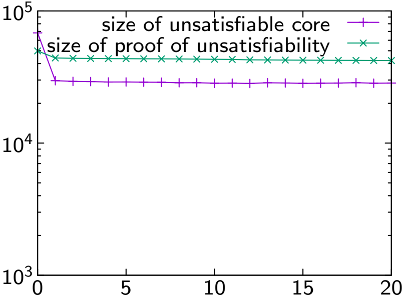
The OptimizeProof procedure repeats proof reduction until the size no longer decreases.

Quality of the Proof

The order of the clauses also influences the SAT solver

Left the smallest proof (100 random clause orders) and right the largest proof and 20 iterations of the OptimizeProof method



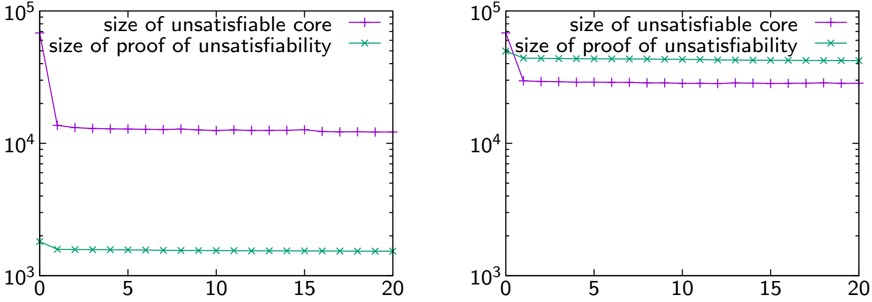


the size of the proof correlates with the size of the core

Quality of the Proof

The order of the clauses also influences the SAT solver

Left the smallest proof (100 random clause orders) and right the largest proof and 20 iterations of the OptimizeProof method



the size of the proof correlates with the size of the core

- Solve the problem multiple times with different clause orders
- Select the smallest proof for proof optimization

TrimFormulaPlain

Input: formula F
Output: an unsatisfiable core of F

```
F_core := F
do
    P := Solve (F_core)
    P := OptimizeProof (P, F_core)
    F_core := ExtractCore (P, F_core)
while (progress)
return F_core
```

problem: useful clauses may be removed from F_{core}

TrimFormulaInteract

Input: formula F
Output: an unsatisfiable core of F

$F_{\text{core}} := F$
do
  $P :=$ Solve (F_{core})
  $P :=$ OptimizeProof (P, F_{core})
  $P :=$ OptimizeProof (P, F)
  $F_{\text{core}} :=$ ExtractCore (P, F)
while (progress)
return F_{core}

solution: useful clauses can be pulled back in F_{core}

Graph Operations

Two operations are use to construct bigger and bigger graph:

- Minkowski sum of A and B ($A \oplus B$): $\{a + b \mid a \in A, b \in B\}$
- Two rotated copies of a graph with a common point

Example

Let $A = \{(0,0), (1,0)\}$ and $B = \{(0,0), (1/2, \sqrt{3}/2)\}$

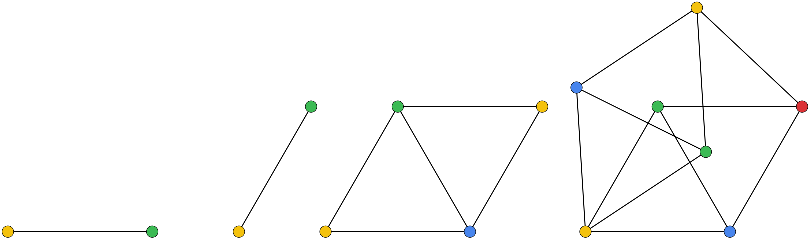
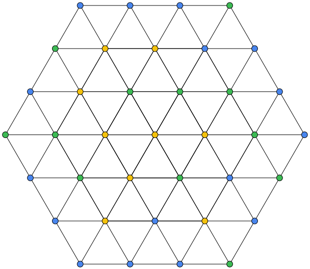


Figure: From left to right: UD-graphs A, B, $A \oplus B$, and the Moser Spindle.

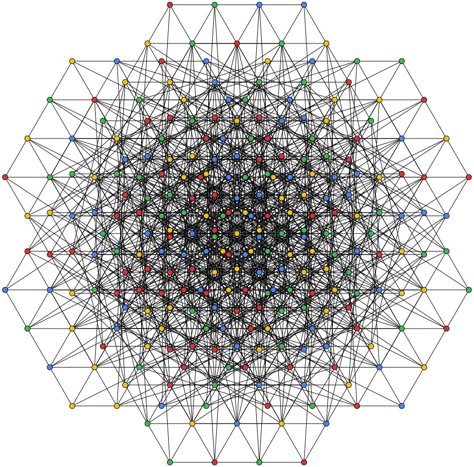
Small graphs in $\mathbb{Q}[\sqrt{3}, \sqrt{11}] \times \mathbb{Q}[\sqrt{3}, \sqrt{11}]$

Graph H_i is the 6-wheel with all edges of length i.

Graph H'_i is a copy of H_i rotated by 90 degrees.

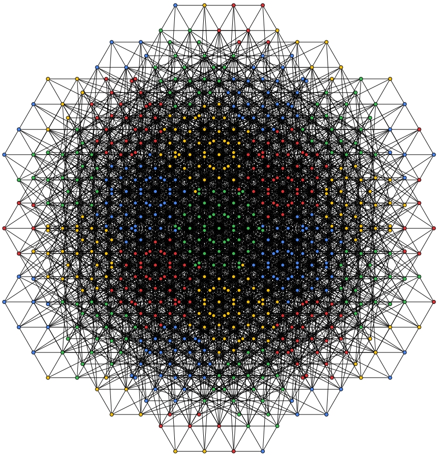


$H_{\frac{1}{3}} \oplus H_{\frac{1}{3}} \oplus H_{\frac{1}{3}}$

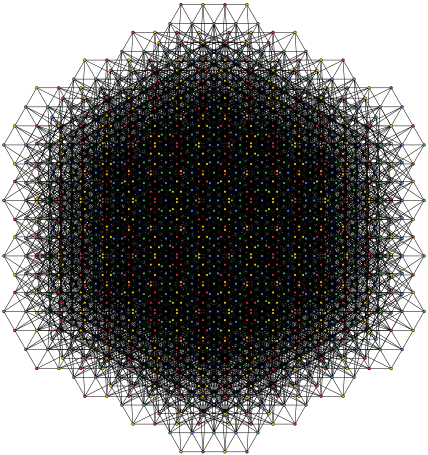


$H_{\frac{1}{3}} \oplus H_{\frac{1}{3}} \oplus H_{\frac{1}{3}} \oplus H'_{\frac{\sqrt{3}+\sqrt{11}}{6}}$

Larger graphs in $\mathbb{Q}[\sqrt{3}, \sqrt{11}] \times \mathbb{Q}[\sqrt{3}, \sqrt{11}]$

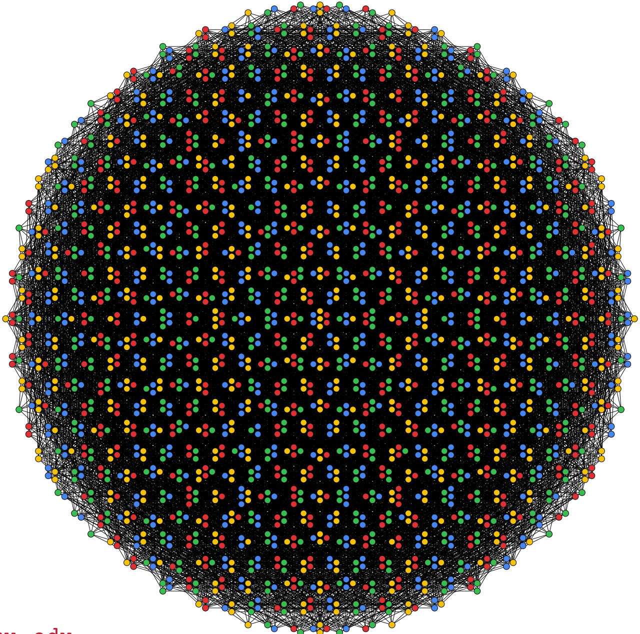


$H_{\frac{1}{3}} \oplus H_{\frac{1}{3}} \oplus H_{\frac{1}{3}} \oplus H'_{\frac{\sqrt{3}+\sqrt{11}}{6}} \oplus H'_{\frac{\sqrt{3}+\sqrt{11}}{6}}$



$H_{\frac{1}{3}} \oplus H_{\frac{1}{3}} \oplus H_{\frac{1}{3}} \oplus H'_{\frac{\sqrt{3}+\sqrt{11}}{6}} \oplus H'_{\frac{\sqrt{3}+\sqrt{11}}{6}} \oplus H'_{\frac{\sqrt{3}+\sqrt{11}}{6}}$

Graph G_{2167}



Computer-Aided Mathematics

Chromatic Number of the Plane

Clausal Proof Optimization

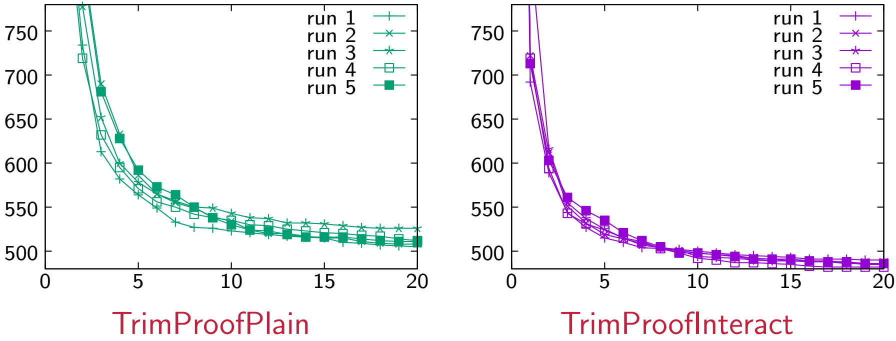
Observed Patterns in $\mathbb{Q}[\sqrt{3}, \sqrt{11}] \times \mathbb{Q}[\sqrt{3}, \sqrt{11}]$

Small UD Graphs with Chromatic Number 5

Conclusions and Future Work

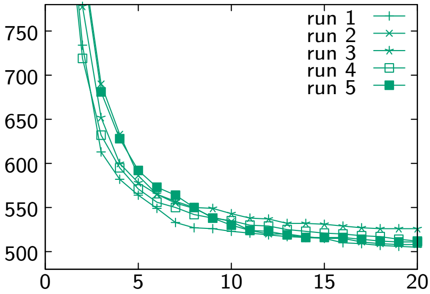
Impact of the Trimming Algorithms

We started with G_{2167} and reduced it using the proof trimming algorithms: TrimProofInteract outperforms TrimProofPlain.

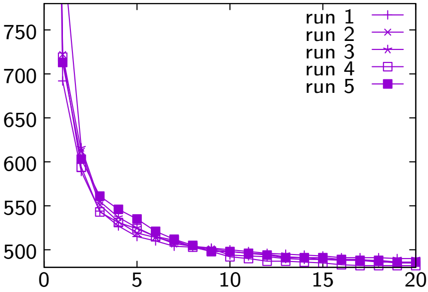


Impact of the Trimming Algorithms

We started with G_{2167} and reduced it using the proof trimming algorithms: TrimProofInteract outperforms TrimProofPlain.



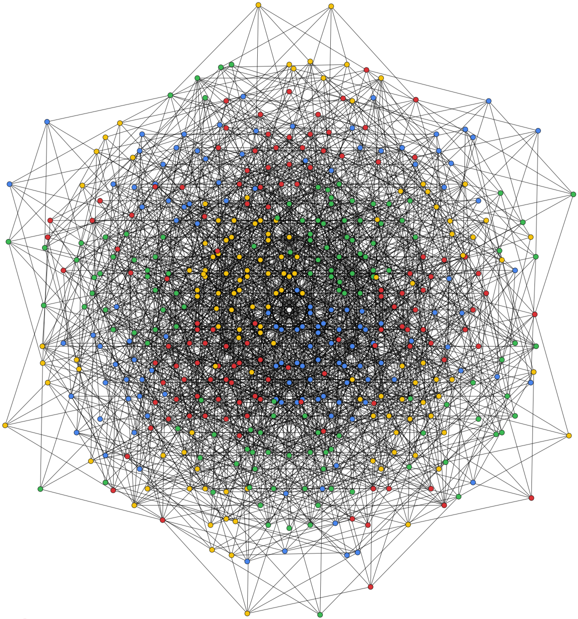
TrimProofPlain



TrimProofInteract

- The smallest subgraph with desired properties: 375 vertices
- We added 135 vertices to remove all 4-colorings

Graph G_{510}



Computer-Aided Mathematics

Chromatic Number of the Plane

Clausal Proof Optimization

Observed Patterns in $\mathbb{Q}[\sqrt{3}, \sqrt{11}] \times \mathbb{Q}[\sqrt{3}, \sqrt{11}]$

Small UD Graphs with Chromatic Number 5

Conclusions and Future Work

Conclusions and Future Work

Aubrey de Grey showed that the chromatic number of the plane is at least 5 using a 1581-vertex unit-distance graph.

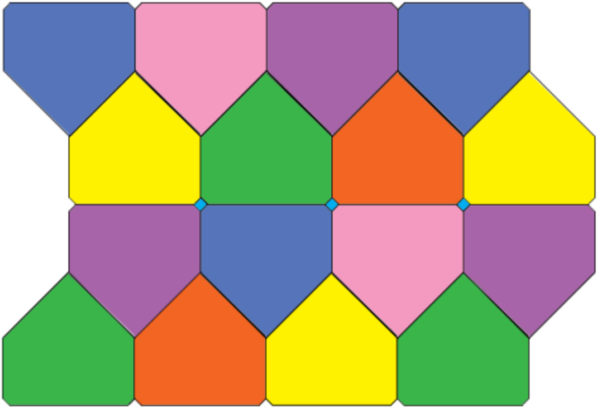
SAT technology can not only validate the result, but also reduce the size of the graph.

Our proof minimization techniques were able to construct a 510-vertex unit-distance graph with chromatic number 5.

Open questions regarding unit-distance graphs:

- What it is the smallest graph with chromatic number 5?
- Can we compute a graph that is human-understandable?
- Is there such a graph with chromatic number 6 (or even 7)?

Improve the Upper Bound?



A 7-coloring with one color covering 0.3% of the plane.

[Pritikin 1998]

Can SAT techniques be used to improve the upper bound?

A Page of God's Book on Theorems

"For many years now I am convinced that the chromatic number will be 7 or 6. One day, Paul Erdős said that God has an endless book that contains all the theorems and best of their evidence, and to some He shows it for a moment. If I had been awarded such an honor and I would have had a choice, I would have asked to look at the page with the problem of the chromatic number of the plane. And you?"

Alexander Soifer