ICARM Tutorial

Efficient Search for Combinatorial Objects I

Marijn J.H. Heule and Bernardo Subercaseaux

Institute for Computer-Aided Reasoning in Mathematics

`https://icarm.io/` JMM booth 409

Joint Mathematics Meetings January 5, 2026

`sat4math.com/tutorials/`

AI for Mathematics

A.I. Is Coming for Mathematics, Too

For thousands of years, mathematicians have adapted to the latest advances in logic and reasoning. Are they ready for artificial intelligence?

Move Over, Mathematicians, Here Comes AlphaProof

A.I. is getting good at math — and might soon make a worthy collaborator for humans.

AI for Mathematics

A.I. Is Coming for Mathematics, Too

For thousands of years, mathematicians have adapted to the latest advances in logic and reasoning. Are they ready for artificial intelligence?

Move Over, Mathematicians, Here Comes AlphaProof

A.I. is getting good at math — and might soon make a worthy collaborator for humans.

Mathematics is the perfect playground to get AI right

- Formal methods offers essential logic-based reasoning
- Highly trustworthy results thanks to (formal) proofs

50 Years of Successes in Computer-Aided Mathematics

1976 Four-Color Theorem

1998 Kepler Conjecture

2014 Boolean Erdős discrepancy problem

2016 Boolean Pythagorean triples problem

2018 Schur Number Five

2019 Keller's Conjecture

2021 Kaplansky's Unit Conjecture

2022 Packing Number of Square Grid

2023 Empty Hexagon in Every 30 Points

50 Years of Successes in Computer-Aided Mathematics

1976 Four-Color Theorem

1998 Kepler Conjecture

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)

2021 Kaplansky's Unit Conjecture (using a SAT solver)

2022 Packing Number of Square Grid (using a SAT solver)

2023 Empty Hexagon in Every 30 Points (using a SAT solver)

50 Years of Successes in Computer-Aided Mathematics

1976 Four-Color Theorem

1998 Kepler Conjecture

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)

2021 Kaplansky's Unit Conjecture (using a SAT solver)

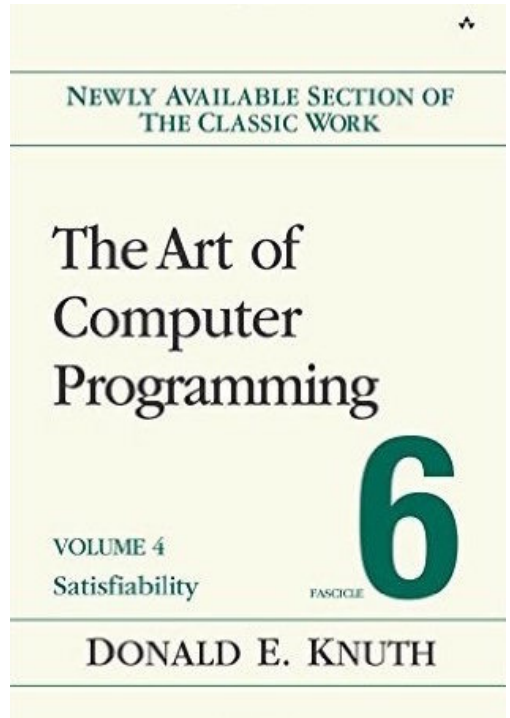
2022 Packing Number of Square Grid (using a SAT solver)

2023 Empty Hexagon in Every 30 Points (using a SAT solver)

Breakthrough in SAT Solving in the Last 30 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/'21]

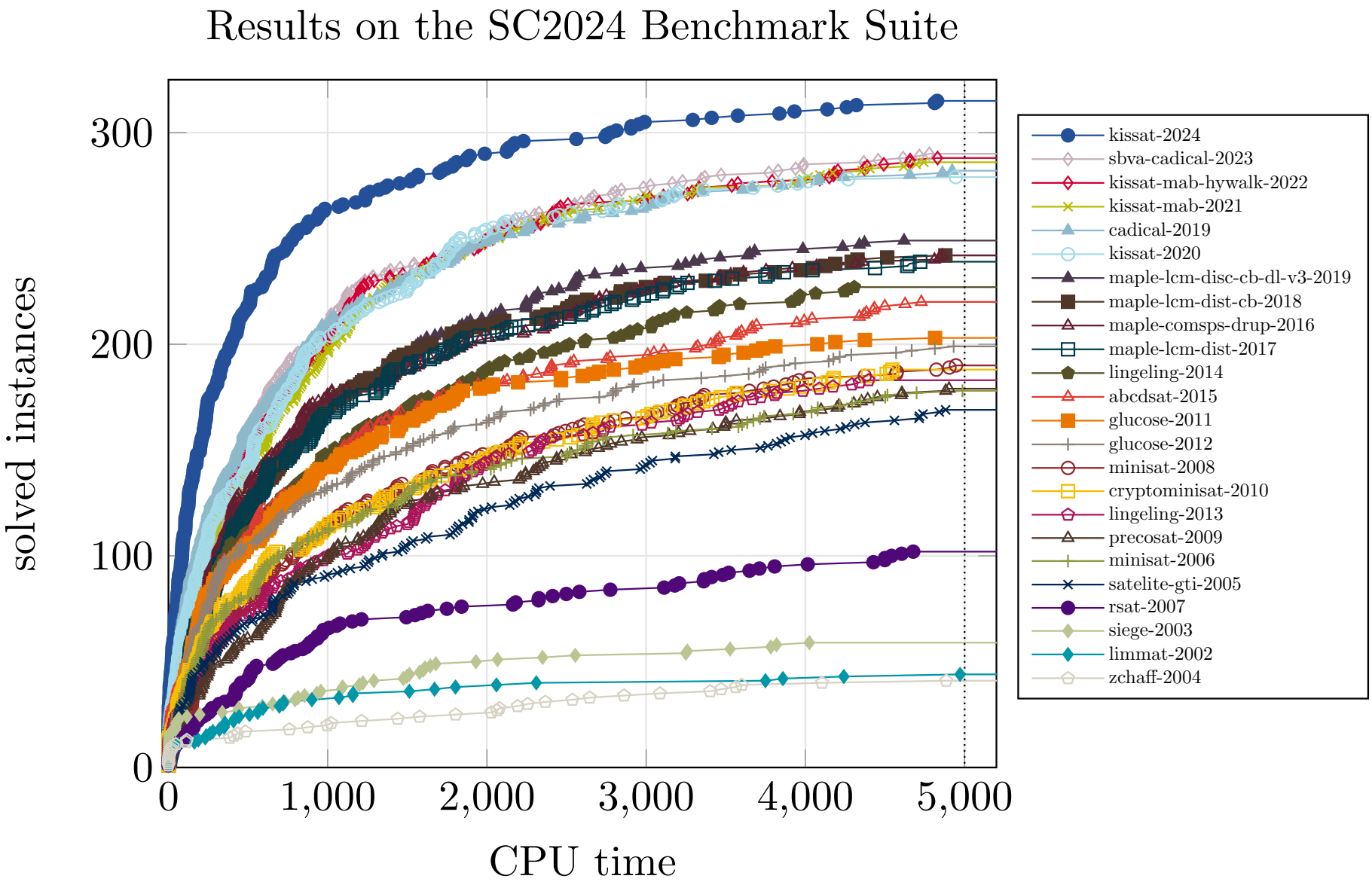
Donald Knuth: *"evidently a killer app, because it is key to the solution of so many other problems"* [Knuth '15]

Naive SAT Solving: Truth Table

$$\Gamma := (p \vee \neg q) \wedge (q \vee r) \wedge (\neg r \vee \neg p)$$

p	q	r	falsifies	eval(Γ)
$\bot$	$\bot$	$\bot$	$q \vee r$	$\bot$
$\bot$	$\bot$	$\top$	—	$\top$
$\bot$	$\top$	$\bot$	$p \vee \neg q$	$\bot$
$\bot$	$\top$	$\top$	$p \vee \neg q$	$\bot$
$\top$	$\bot$	$\bot$	$q \vee r$	$\bot$
$\top$	$\bot$	$\top$	$\neg r \vee \neg p$	$\bot$
$\top$	$\top$	$\bot$	—	$\top$
$\top$	$\top$	$\top$	$\neg r \vee \neg p$	$\bot$

Progress of SAT Solvers



Schur's Theorem [Schur 1916]

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

$$1+1=2 \qquad 1+2=3 \qquad 1+3=4$$
$$1+4=5 \qquad 2+2=4 \qquad 2+3=5$$

Schur's Theorem [Schur 1916]

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

$$1 + 1 = 2 \qquad 1 + 2 = 3 \qquad 1 + 3 = 4$$
$$1 + 4 = 5 \qquad 2 + 2 = 4 \qquad 2 + 3 = 5$$

Schur's Theorem [Schur 1916]

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

$$1 + 1 = 2 \qquad 1 + 2 = 3 \qquad 1 + 3 = 4$$
$$1 + 4 = 5 \qquad 2 + 2 = 4 \qquad 2 + 3 = 5$$

Schur's Theorem [Schur 1916]

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

$1 + 1 = 2$ $\quad$ $1 + 2 = 3$ $\quad$ $1 + 3 = 4$

$1 + 4 = 5$ $\quad$ $2 + 2 = 4$ $\quad$ $2 + 3 = 5$

Schur's Theorem [Schur 1916]

Will any coloring of the positive integers with red and blue result in a monochromatic solution of $a + b = c$? Yes

$$1+1=2 \qquad 1+2=3 \qquad 1+3=4$$
$$1+4=5 \qquad 2+2=4 \qquad 2+3=5$$

Schur's Theorem [Schur 1916]

Will any coloring of the positive integers with red and blue result in a monochromatic solution of $a+b=c$? Yes

$$1+1=2 \qquad 1+2=3 \qquad 1+3=4$$
$$1+4=5 \qquad 2+2=4 \qquad 2+3=5$$

Theorem (Schur's Theorem)

For every positive integer k, there exists a number $S(k)$, such that $[1, S(k)]$ can be colored with k colors while avoiding a monochromatic solution of $a+b=c$ with $a,b,c \leq S(k)$, while this is impossible for $[1, S(k)+1]$.

$S(1)=1, S(2)=4, S(3)=13,$

Schur's Theorem [Schur 1916]

Will any coloring of the positive integers with red and blue result in a monochromatic solution of $a+b=c$? Yes

$$1+1=2 \qquad 1+2=3 \qquad 1+3=4$$
$$1+4=5 \qquad 2+2=4 \qquad 2+3=5$$

Theorem (Schur's Theorem)

For every positive integer k, there exists a number $S(k)$, such that $[1, S(k)]$ can be colored with k colors while avoiding a monochromatic solution of $a+b=c$ with $a,b,c \leq S(k)$, while this is impossible for $[1, S(k)+1]$.

$S(1)=1, S(2)=4, S(3)=13, S(4)=44$ [Baumert 1965].

- We will prove this during the tutorial.

Schur's Theorem [Schur 1916]

Will any coloring of the positive integers with red and blue result in a monochromatic solution of $a+b=c$? Yes

$$1+1=2 \qquad 1+2=3 \qquad 1+3=4$$
$$1+4=5 \qquad 2+2=4 \qquad 2+3=5$$

Theorem (Schur's Theorem)

For every positive integer k, there exists a number $S(k)$, such that $[1, S(k)]$ can be colored with k colors while avoiding a monochromatic solution of $a+b=c$ with $a,b,c \leq S(k)$, while this is impossible for $[1, S(k)+1]$.

$S(1)=1, S(2)=4, S(3)=13, S(4)=44$ [Baumert 1965].

- We will prove this during the tutorial.

We show that $S(5)=160$ [Heule 2018].

Schur's Theorem [Schur 1916]

Will any coloring of the positive integers with red and blue result in a monochromatic solution of $a+b=c$? Yes

$$1+1=2 \qquad 1+2=3 \qquad 1+3=4$$
$$1+4=5 \qquad 2+2=4 \qquad 2+3=5$$

Theorem (Schur's Theorem)

For every positive integer k, there exists a number $S(k)$, such that $[1, S(k)]$ can be colored with k colors while avoiding a monochromatic solution of $a+b=c$ with $a,b,c \leq S(k)$, while this is impossible for $[1, S(k)+1]$.

$S(1)=1, S(2)=4, S(3)=13, S(4)=44$ [Baumert 1965].

- We will prove this during the tutorial.

We show that $S(5)=160$ [Heule 2018]. Proof: 2 petabytes

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 $a^2 + b^2 = c^2$?

$3^2 + 4^2 = 5^2$ $\quad 6^2 + 8^2 = 10^2$ $\quad 5^2 + 12^2 = 13^2$ $\quad 9^2 + 12^2 = 15^2$

$8^2 + 15^2 = 17^2$ $\quad 12^2 + 16^2 = 20^2$ $\quad 15^2 + 20^2 = 25^2$ $\quad 7^2 + 24^2 = 25^2$

$10^2 + 24^2 = 26^2$ $\quad 20^2 + 21^2 = 29^2$ $\quad 18^2 + 24^2 = 30^2$ $\quad 16^2 + 30^2 = 34^2$

$21^2 + 28^2 = 35^2$ $\quad 12^2 + 35^2 = 37^2$ $\quad 15^2 + 36^2 = 39^2$ $\quad 24^2 + 32^2 = 40^2$

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 $a^2 + b^2 = c^2$?

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

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 p_i with $i \in \{1, 2, \ldots, n\}$ such that $p_i = 1$ $(= 0)$ means that i is colored red (blue). For each Pythagorean Triple $a^2 + b^2 = c^2$, two clauses are added: $(p_a \vee p_b \vee p_c)$ and $(\neg p_a \vee \neg p_b \vee \neg p_c)$.

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 p_i with $i \in \{1, 2, \ldots, n\}$ such that $p_i = 1$ $(= 0)$ means that i is colored red (blue). For each Pythagorean Triple $a^2 + b^2 = c^2$, two clauses are added: $(p_a \vee p_b \vee p_c)$ and $(\neg p_a \vee \neg p_b \vee \neg p_c)$.

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 p_i with $i \in \{1, 2, \ldots, n\}$ such that $p_i = 1$ $(= 0)$ means that i is colored red (blue). For each Pythagorean Triple $a^2 + b^2 = c^2$, two clauses are added: $(p_a \vee p_b \vee p_c)$ and $(\neg p_a \vee \neg p_b \vee \neg p_c)$.

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 p_i with $i \in \{1, 2, \ldots, n\}$ such that $p_i = 1$ $(= 0)$ means that i is colored red (blue). For each Pythagorean Triple $a^2 + b^2 = c^2$, two clauses are added: $(p_a \vee p_b \vee p_c)$ and $(\neg p_a \vee \neg p_b \vee \neg p_c)$.

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

Tutorials

Today:

- Small Ramsey numbers
- Small Schur numbers (if there is time)

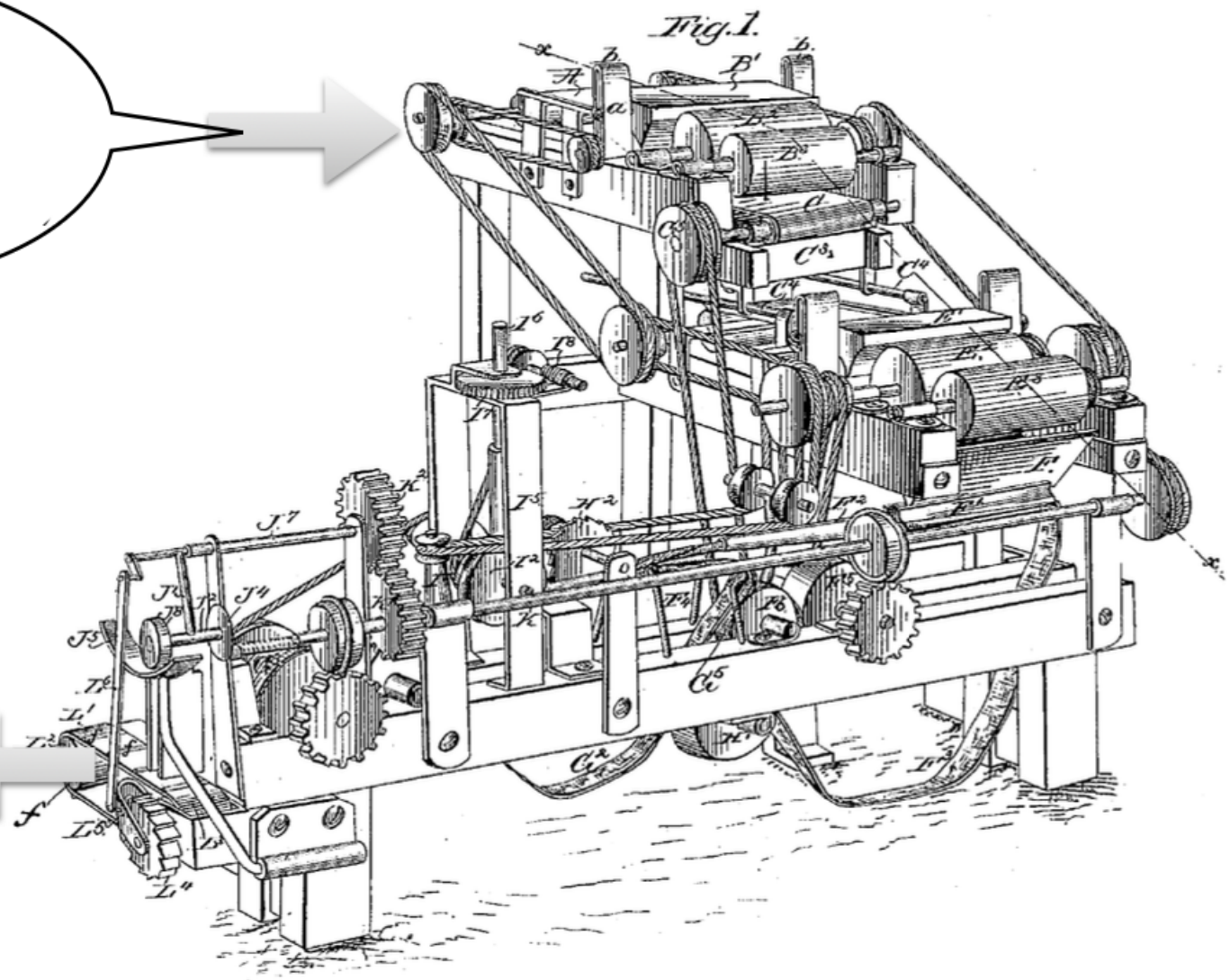
Tomorrow (same time, same room):

- Kaplansky's unit problem
- Erdős discrepancy problem

`sat4math.com/tutorials/`

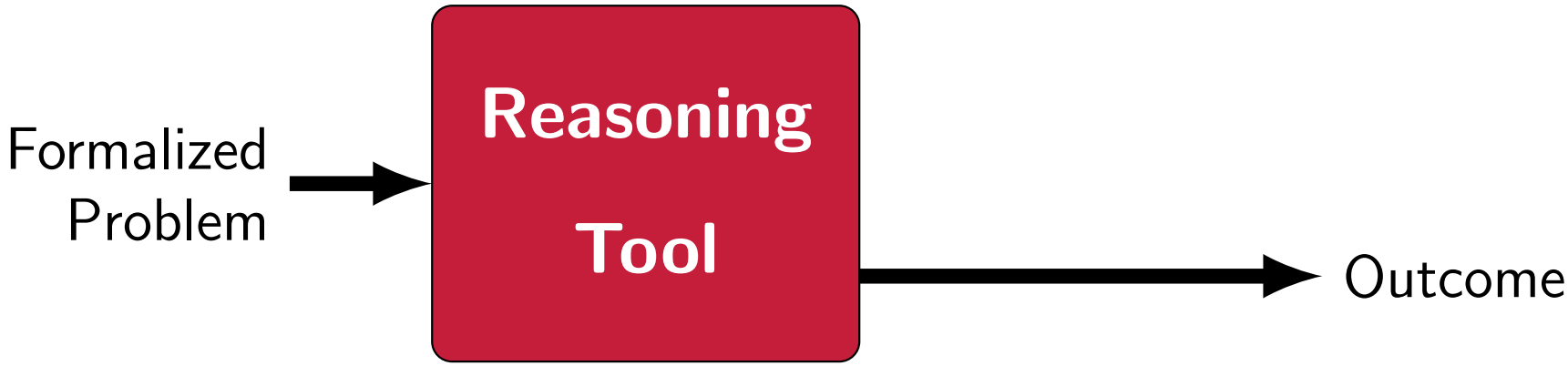
SAT Solvers are Complex Tools

Γ



Proofs

Automated Reasoning Programs



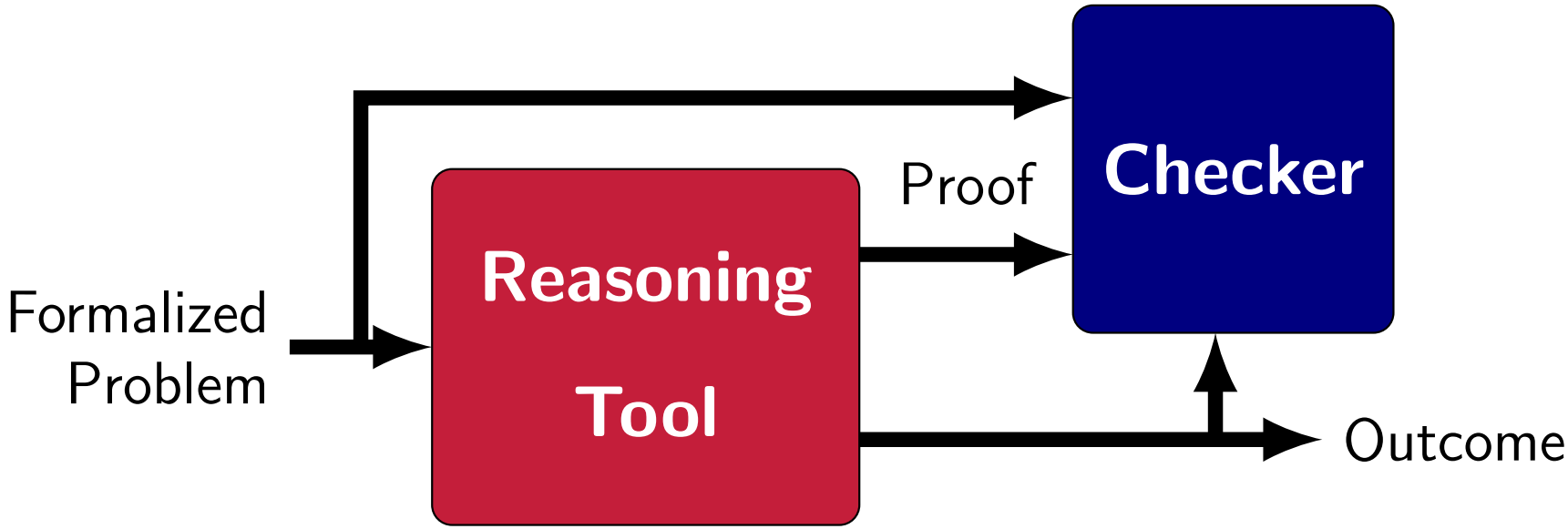
Standard Implementations

- Lingering doubt about whether the result can be trusted
- If a bug is found in a tool, rerun all prior verifications

Formally Verified Tools

- Hard to develop
- Hard to make scalable

Proof-Generating Automated Reasoning Programs



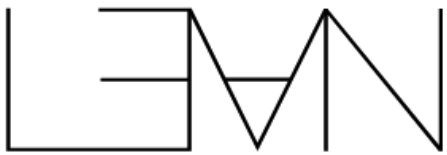
Proof-Generating Tools

- Only need to prove individual executions, not entire program
- Can have bugs in the tool, but still trust the result
- Can we trust the checker?
  - Simple algorithms and implementation
  - Ideally formally verified

Proof-Generating Tools: Arbitrarily Complex Solvers

Proof-generating tools with verified checkers is a powerful idea:

- Don't worry about correctness or completeness of tools;
- Facilitates making tools more complex and efficient; while
- Full confidence in results. [Heule, Hunt, Kaufmann, Wetzler '17]

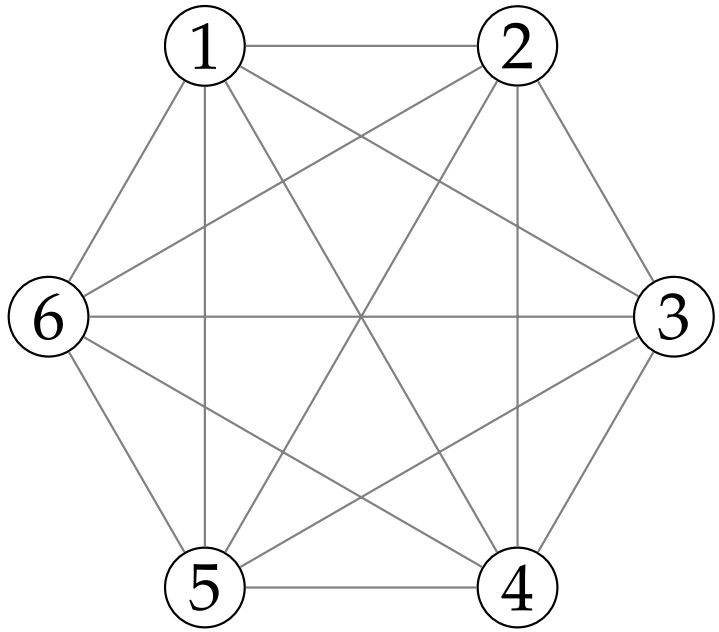



Formally-verified checkers now also used in industry

First Tutorial: Small Ramsey Numbers

Ramsey Number $R(k)$: What is the smallest n such that any graph with n vertices has either a clique or a co-clique of size k?

$$\begin{aligned} & R(3) = 6 \\ & R(4) = 18 \\ 43 \leq \quad & R(5) \leq 46 \end{aligned}$$



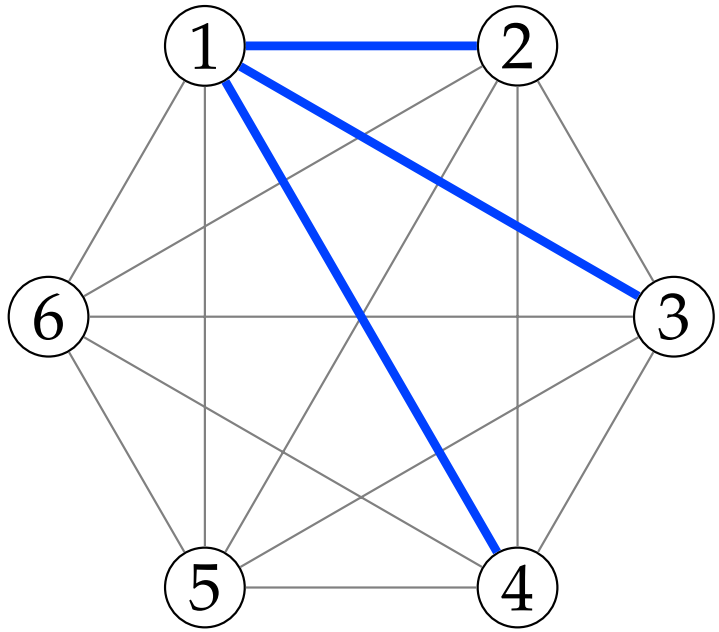
SAT solvers can determine that $R(4) = 18$ in 1 second using symmetry breaking; w/o symmetry breaking, it requires weeks.

Symmetry breaking validated by proof checker [CADE'15]

First Tutorial: Small Ramsey Numbers

Ramsey Number $R(k)$: What is the smallest n such that any graph with n vertices has either a clique or a co-clique of size k?

$$\begin{aligned} & R(3) = 6 \\ & R(4) = 18 \\ 43 \leq \quad & R(5) \leq 46 \end{aligned}$$



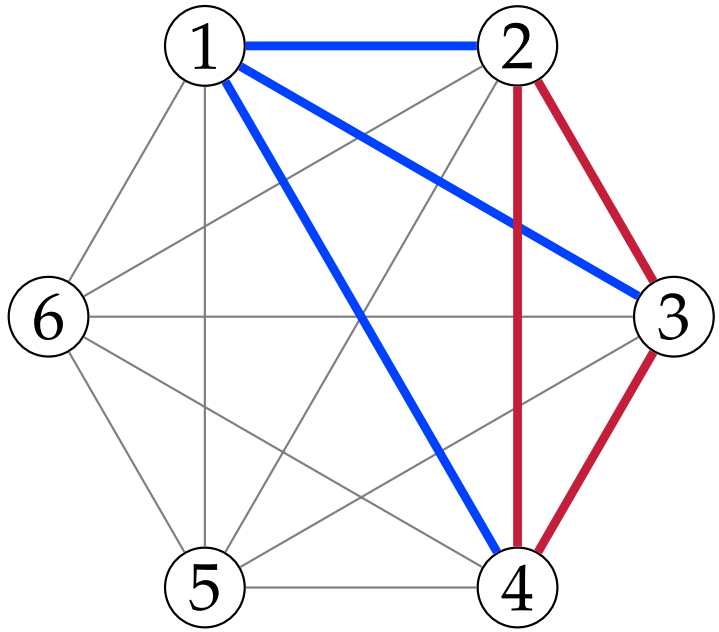
SAT solvers can determine that $R(4) = 18$ in 1 second using symmetry breaking; w/o symmetry breaking, it requires weeks.

Symmetry breaking validated by proof checker [CADE'15]

First Tutorial: Small Ramsey Numbers

Ramsey Number $R(k)$: What is the smallest n such that any graph with n vertices has either a clique or a co-clique of size k?

$$\begin{aligned} & R(3) = 6 \\ & R(4) = 18 \\ 43 \leq \quad & R(5) \leq 46 \end{aligned}$$



SAT solvers can determine that $R(4) = 18$ in 1 second using symmetry breaking; w/o symmetry breaking, it requires weeks.

Symmetry breaking validated by proof checker [CADE'15]

sat4math.com/tutorials/

Ramsey Numbers Tutorial

sat4math.com/tutorials/

Schur Numbers Tutorial