Logic and Mechanized Reasoning

Introduction with a focus on mathematics

Marijn J.H. Heule

Carnegie Mellon University

Mechanized Reasoning Has Many Applications

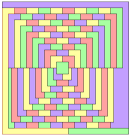
formal verification

security

bioinformatics

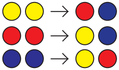
planning and scheduling

train safety



automated theorem proving

exploit generation



term rewriting termination

encode

mechanized reasoning

decode

Mechanized Reasoning Has Many Applications

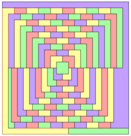
formal verification

security

bioinformatics

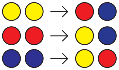
planning and scheduling

train safety



automated theorem proving

exploit generation



term rewriting termination

encode → **mechanized reasoning** → **decode**

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)

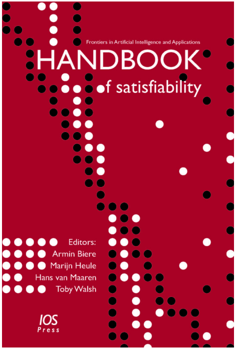
2018 Schur Number Five (using a SAT solver)

2019 Keller's Conjecture (using a SAT solver)

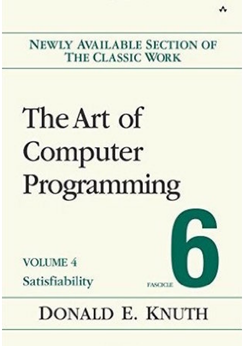
Breakthrough in SAT Solving in the Last 20 Years

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, and Walsh '09]



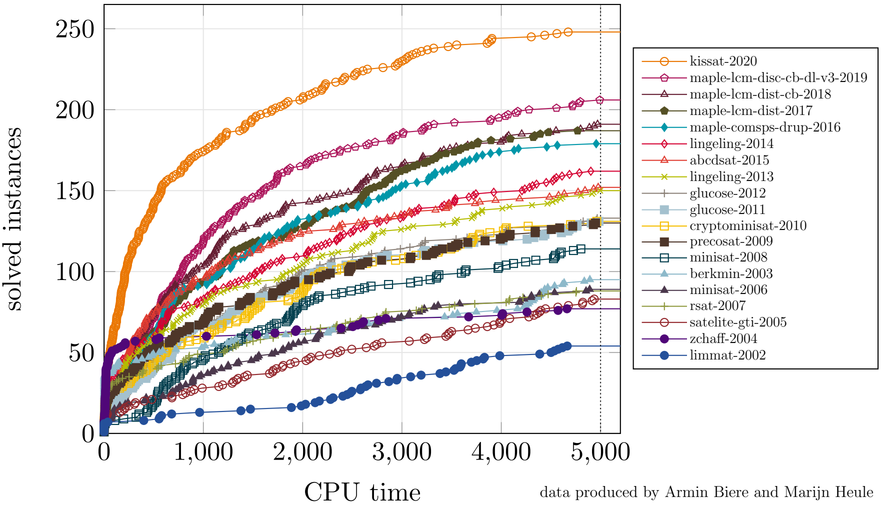
Donald Knuth: *"evidently a killer app, because it is key to the solution of so many other problems"* [Knuth '15]

Truth Table

$$F := (p \vee \overline{q}) \wedge (q \vee r) \wedge (\overline{r} \vee \overline{p})$$

p	q	r	falsifies	eval(F)
0	0	0	$(q \vee r)$	0
0	0	1	—	1
0	1	0	$(p \vee \overline{q})$	0
0	1	1	$(p \vee \overline{q})$	0
1	0	0	$(q \vee r)$	0
1	0	1	$(\overline{r} \vee \overline{p})$	0
1	1	0	—	1
1	1	1	$(\overline{r} \vee \overline{p})$	0

SAT Competition Winners on the SC2020 Benchmark Suite

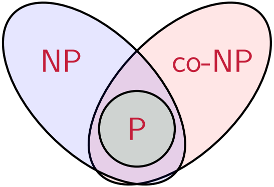


data produced by Armin Biere and Marijn Heule

Satisfiability and Complexity

Complexity classes of decision problems:

P : efficiently computable answers.
NP : efficiently checkable yes-answers.
co-NP : efficiently checkable no-answers.



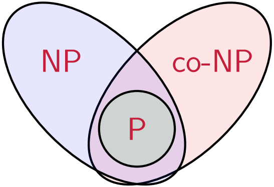
Cook-Levin Theorem [1971]: SAT is NP-complete.

Solving the $\mathsf{P} \stackrel{?}{=} \mathsf{NP}$ question is worth \$1,000,000 [Clay MI '00].

Satisfiability and Complexity

Complexity classes of decision problems:

P : efficiently computable answers.
NP : efficiently checkable yes-answers.
co-NP : efficiently checkable no-answers.



Cook-Levin Theorem [1971]: SAT is NP-complete.

Solving the $P \stackrel{?}{=} NP$ question is worth \$1,000,000 [Clay MI '00].

The effectiveness of SAT solving: fast solutions in practice.

The beauty of NP: guaranteed short solutions.

"NP is the new P!"

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

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

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 $(\overline{x}_a \vee \overline{x}_b \vee \overline{x}_c)$.

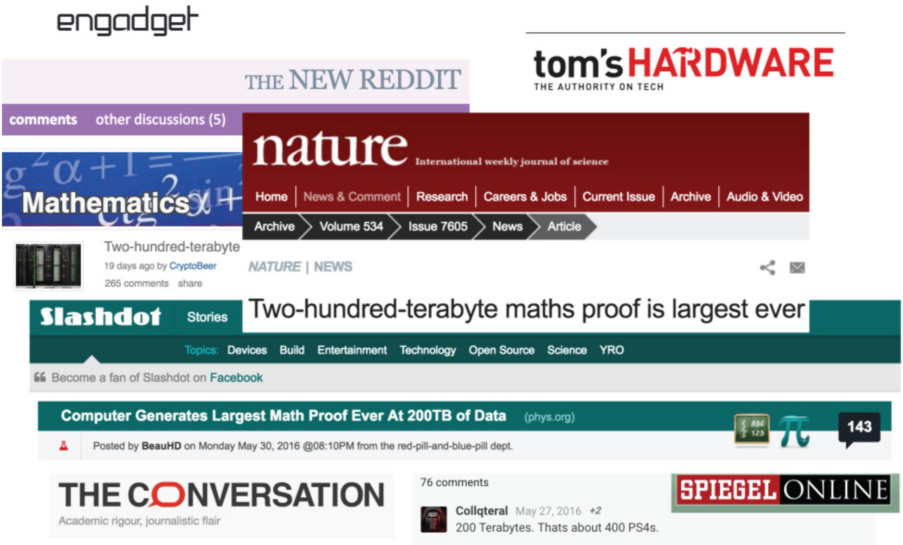
Theorem ([Heule, Kullmann, and Marek (2016)])

$[1, 7824]$ *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)

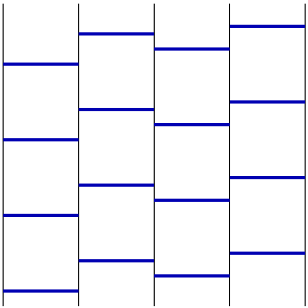
200 terabytes proof, but validated with verified checker

Media: “The Largest Math Proof Ever”



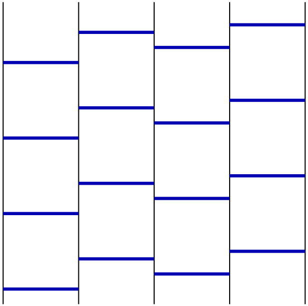
Keller's Conjecture: A Tiling Problem

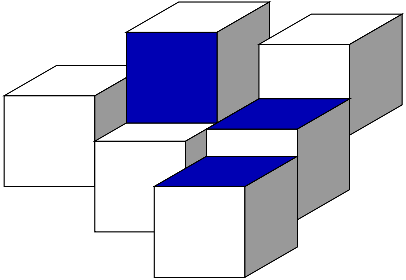
Consider tiling a floor with square tiles, all of the same size. Is it the case that any gap-free tiling results in at least two fully connected tiles, i.e., tiles that have an entire edge in common?



Keller's Conjecture: A Tiling Problem

Consider tiling a floor with square tiles, all of the same size. Is it the case that any gap-free tiling results in at least two fully connected tiles, i.e., tiles that have an entire edge in common?

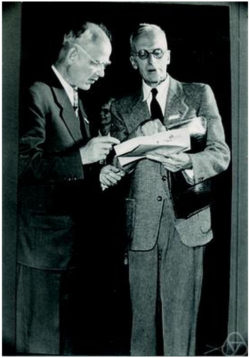




Keller's Conjecture: Resolved

In 1930, Ott-Heinrich Keller conjectured that this phenomenon holds in every dimension.

Keller's Conjecture.
For all $n \geq 1$, every tiling of the n-dimensional space with unit cubes has two which fully share a face.



[Wikipedia, CC BY-SA]

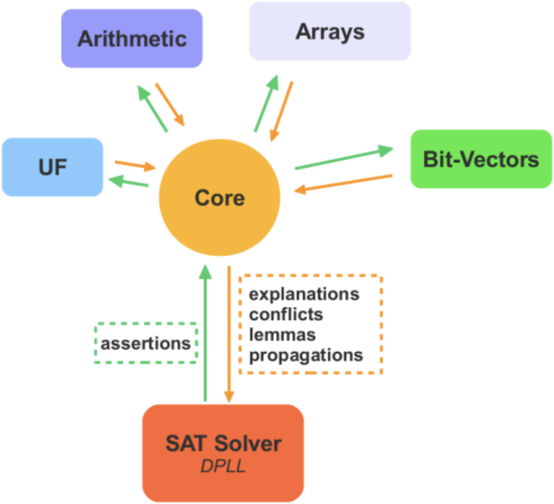
GEOMETRY

Computer Search Settles 90-Year-Old Math Problem

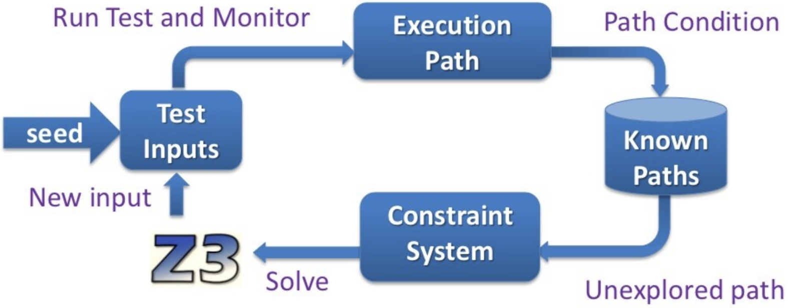
10

By translating Keller's conjecture into a computer-friendly search for a type of graph, researchers have finally resolved a problem about covering spaces with tiles.

Satisfiability Modulo Theories (SMT)



SMT at Microsoft: Test Input Generation




I Programmer

Microsoft Z3 Theorem Prover Wins Award

Microsoft Research's Z3 theorem prover has been awarded the 2015 ACM SIGPLAN Programming Languages Software Award. Z3banner.

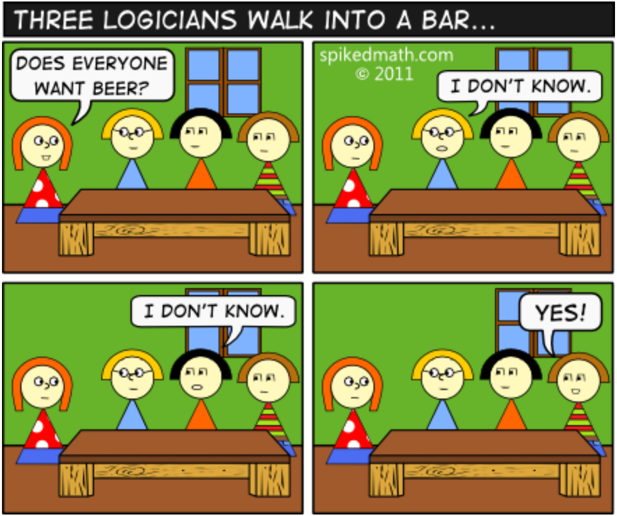
Jun 24, 2015

SMT at Amazon Web Services: Provable Security

Automated reasoning versus machine learning: How AWS provides secure access control without data

VIDEO EXCLUSIVE BY BETSY AMY-VOGT

First-Order and Higher-Order Logic



http://spikedmath.com/445.html

Automating Gödel's Ontological Proof of God's Existence

abcNEWS VIDEO LIVE SHOWS CORONAVIRUS

Computer Scientists 'Prove' God Exists

Can proof of God be proven in mathematical equations?

By **David Knight, SPIEGEL**
October 27, 2013, 3:30 AM • 5 min read

Getty Images

Two scientists believe they have formalized a theorem confirming the existence of God.

Lean Embraced by Mathematicians

nature

Explore content ⌄ Journal information ⌄ Publish with us ⌄ Subscribe

nature > news > article

NEWS | 18 June 2021

Mathematicians welcome computer-assisted proof in 'grand unification' theory

Proof-assistant software handles an abstract concept at the cutting edge of research, revealing a bigger role for software in mathematics.

Future of Computer-Aided Mathematics

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 Search Technology
  - Automatically determine whether a conjecture holds
  - Recent improvement: Linear speedups on thousands of cores
- Proof Checking Technology
  - Mechanized validation of all details
  - Recent improvement: Formally verified checking of huge proofs

Future of Computer-Aided Mathematics

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 Search Technology
  - Automatically determine whether a conjecture holds
  - Recent improvement: Linear speedups on thousands of cores
- Proof Checking Technology
  - Mechanized validation of all details
  - Recent improvement: Formally verified checking of huge proofs

Classic problems ready for mechanization:

- Chromatic number of the plane
- Ramsey number five
- Collatz Conjecture (maybe?)