Computer-Aided Mathematics

Successes, Advances, and Trust

Marijn J.H. Heule

Carnegie Mellon University

34^{th} Midwestern Conference on Combinatorics
and Combinatorial Computing
October 21, 2022

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)

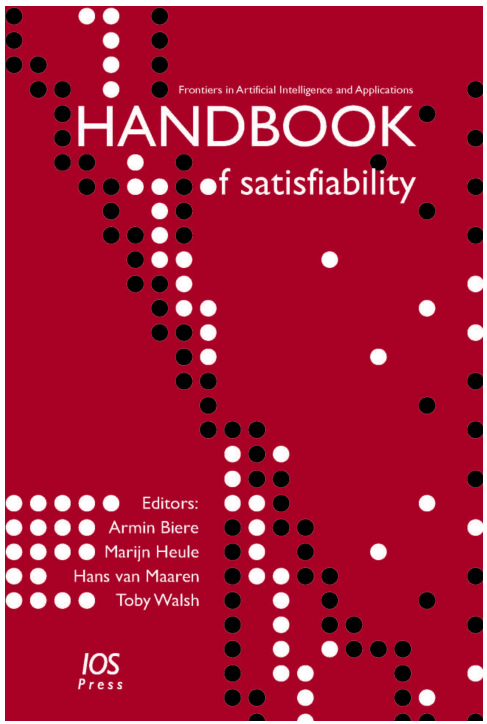
2018 Schur Number Five (using a SAT solver)

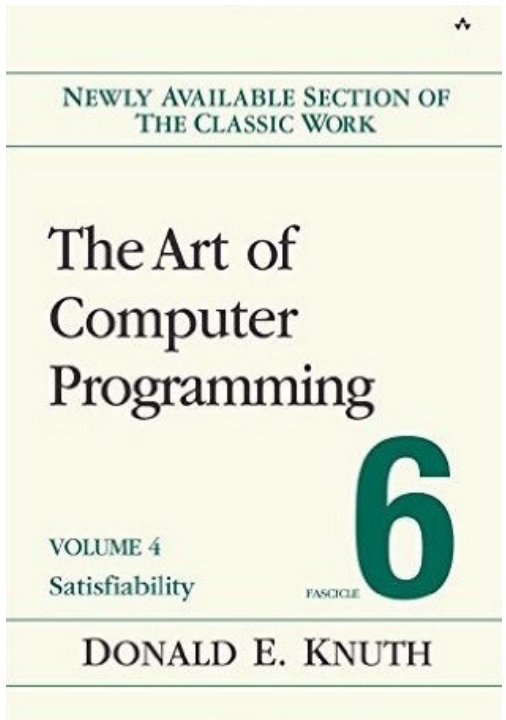
2019 Keller's Conjecture (using a SAT solver)

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

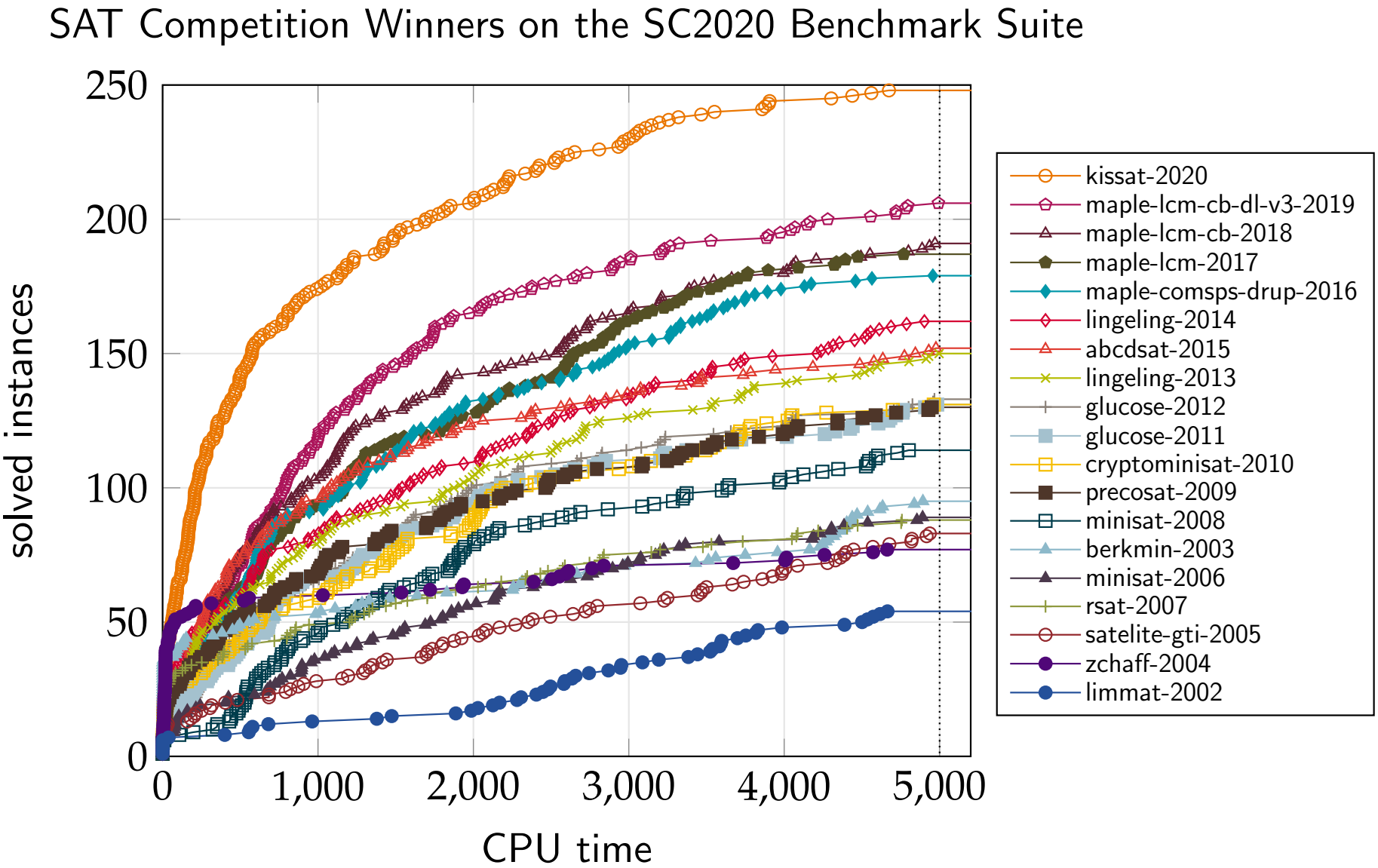




Edmund Clarke: *"a key technology of the 21st century"* [Biere, Heule, vanMaaren, and Walsh '09]

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

Progress of SAT Solvers



Satisfiability and Mathematics

Proofs of Unsatisfiability

Future and Challenges

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 $a + b = c$?

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

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

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

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].

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

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

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)])

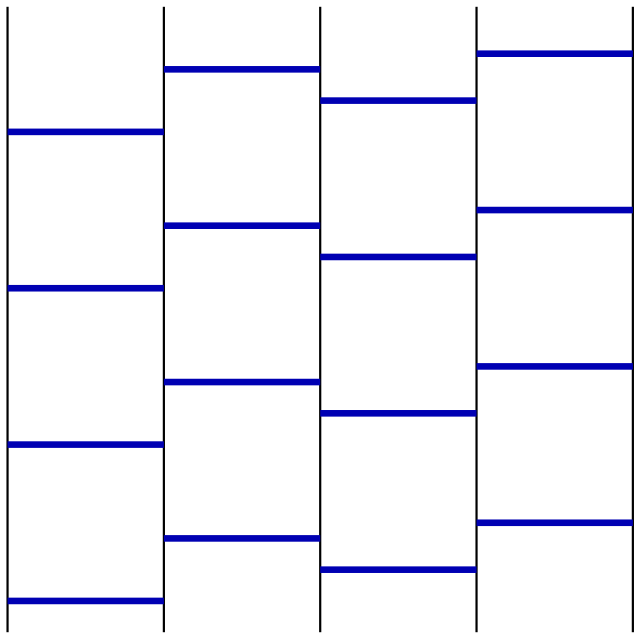
$[1, 7824]$ *can be bi-colored s.t. there is no monochromatic Pythagorean Triple. This is impossible for* $[1, 7825]$.

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

200 terabytes proof, but validated with verified checker

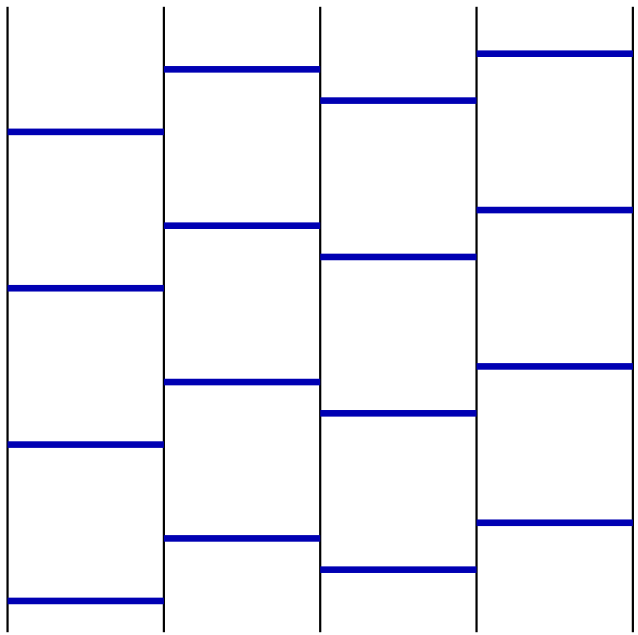
Keller's Conjecture: A Tiling Problem

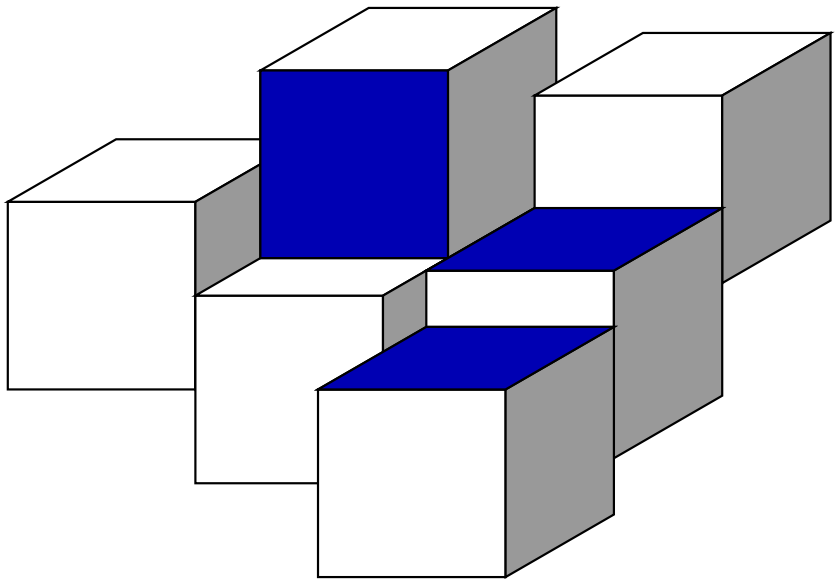
Consider tiling a floor with square tiles, all of the same size. Is it the case that any gap-free tiling results in at least two fully connected tiles, i.e., tiles that have an entire edge in common?



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?



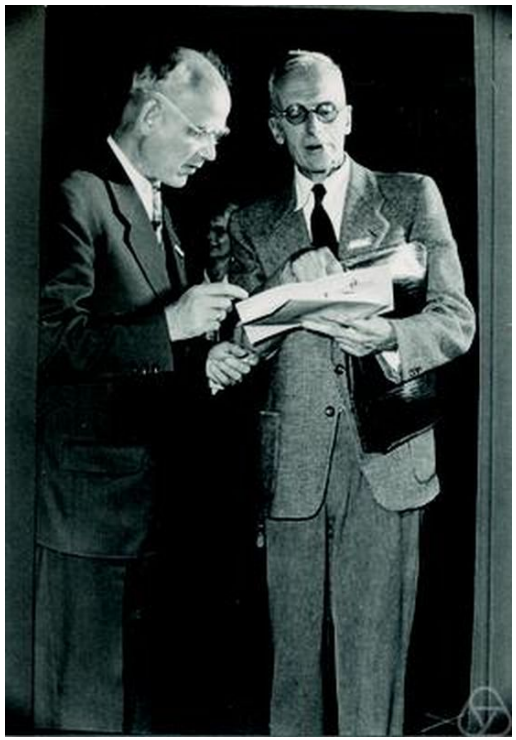


Keller's Conjecture: Resolved

[Brakensiek, Heule, Mackey, & Narvaez 2019]

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 and Mathematics

Proofs of Unsatisfiability

Future and Challenges

Media: "The Largest Math Proof Ever"

engadget

THE NEW REDDIT

comments other discussions (5)

tom's HARDWARE
THE AUTHORITY ON TECH

Mathematics

nature
International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video

Archive > Volume 534 > Issue 7605 > News > Article

Two-hundred-terabyte
19 days ago by CryptoBeer
265 comments share

NATURE | NEWS

Slashdot Stories

Two-hundred-terabyte maths proof is largest ever

Topics: Devices Build Entertainment Technology Open Source Science YRO

Become a fan of Slashdot on Facebook

Computer Generates Largest Math Proof Ever At 200TB of Data (phys.org)

Posted by BeauHD on Monday May 30, 2016 @08:10PM from the red-pill-and-blue-pill dept.

143

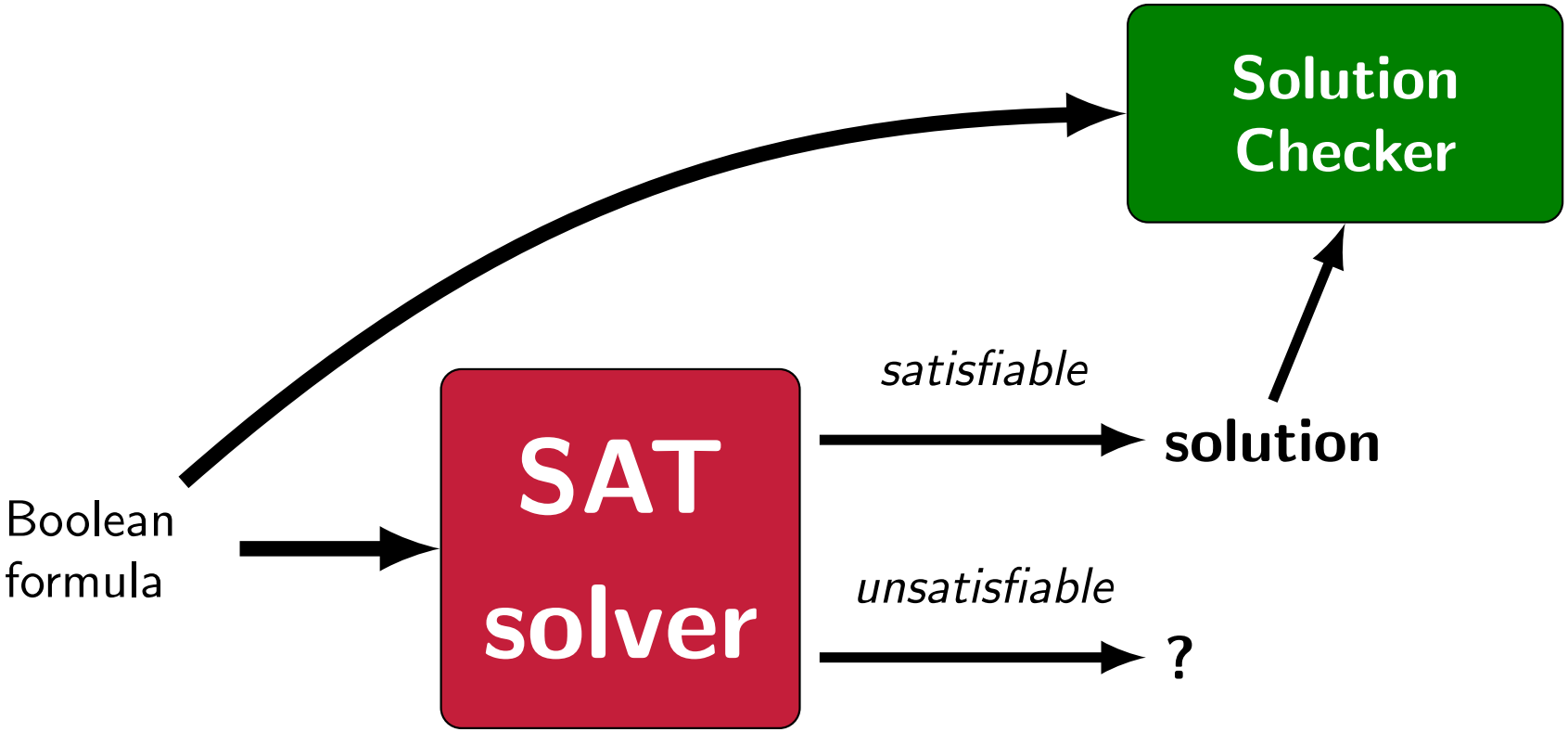
THE CONVERSATION
Academic rigour, journalistic flair

76 comments

Collqteral May 27, 2016 +2
200 Terabytes. Thats about 400 PS4s.

SPIEGEL ONLINE

Proofs of Unsatisfiability: Checking Satisfiability is Easy



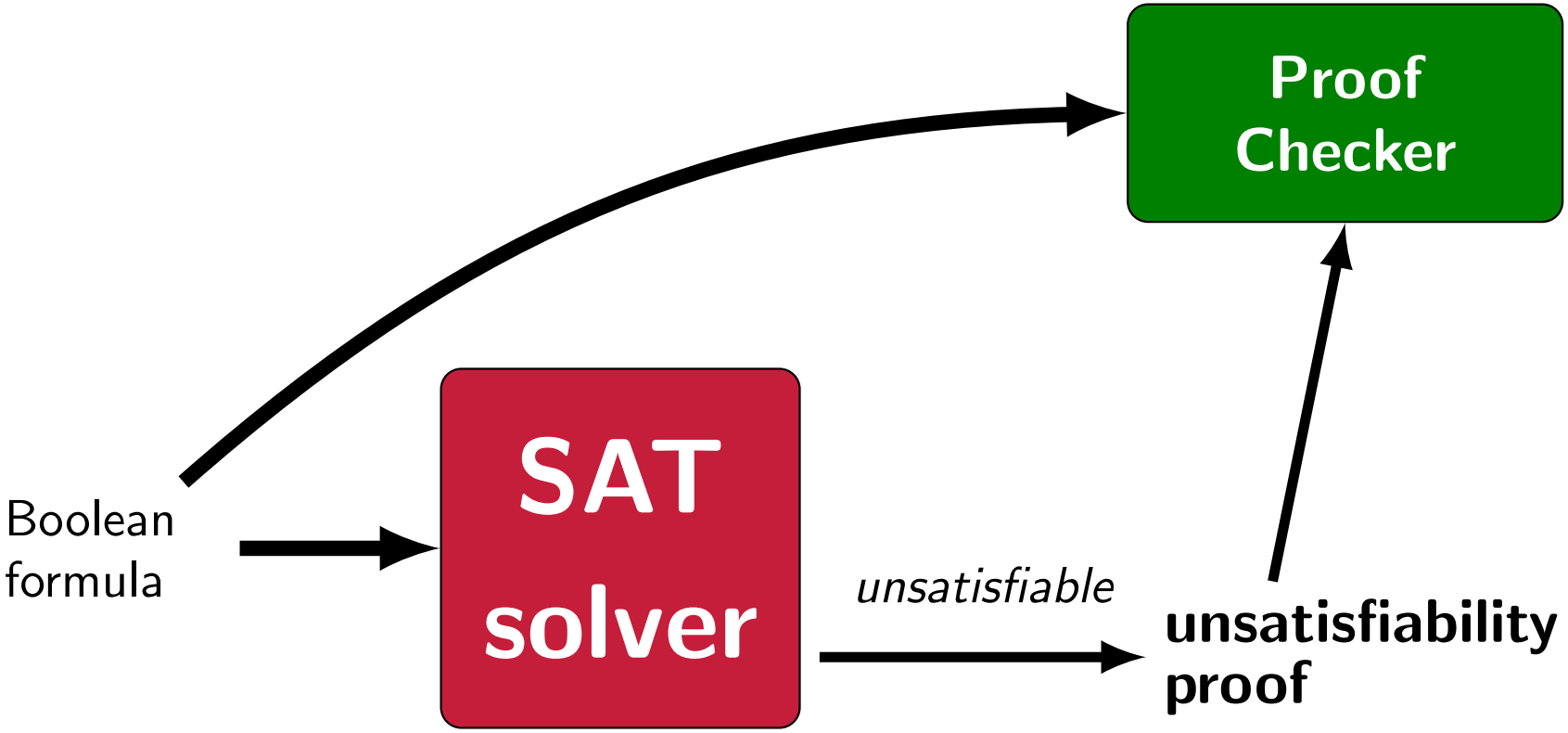
SAT Solvers Useful & Powerful

- Formal verification
- Security verification
- Mathematics

Can We Trust Them?

- No!
- Complex software with lots of optimizations

Proofs of Unsatisfiability: Proof Generating Solvers



Unsatisfiability Proof

- Step-by-step proof in some logical framework

Proof Checker

- Simple program
- May be formally verified

Proofs of Unsatisfiability: Motivation

Automated reasoning tools may give incorrect answers.

- Documented bugs in SAT, SMT, and QSAT solvers;
  [Brummayer and Biere, 2009; Brummayer et al., 2010]
- Claims of correctness could be due to bugs;
- Misconception that only weak tools are buggy;
- Implementation errors often imply conceptual errors;
- Proofs now mandatory in some competitive events;
- Mathematical results require a stronger justification than a simple yes/no by a tool. Answers must be verifiable.

Proofs of Unsatisfiability: Verified Solvers?

Verifying efficient automated reasoning tools is a daunting task:

- Tools are constantly modified and improved; and
- Even top-tier and "experimentally correct" solvers turned out to be buggy. [Järvisalo, Heule, Biere '12]

Proofs of Unsatisfiability: Verified Solvers?

Verifying efficient automated reasoning tools is a daunting task:

- Tools are constantly modified and improved; and
- Even top-tier and "experimentally correct" solvers turned out to be buggy. [Järvisalo, Heule, Biere '12]

Various simple solvers can verified, but they lack performance

- DPLL [Shankar and Vaucher '11]
- CDCL [Fleury, Blanchette, Lammich '18]

Proofs of Unsatisfiability: Verified Solvers?

Verifying efficient automated reasoning tools is a daunting task:

- Tools are constantly modified and improved; and
- Even top-tier and "experimentally correct" solvers turned out to be buggy. [Järvisalo, Heule, Biere '12]

Various simple solvers can verified, but they lack performance

- DPLL [Shankar and Vaucher '11]
- CDCL [Fleury, Blanchette, Lammich '18]

Validating proof is the more effective approach

- Solving + proof logging + proof verification is much faster compared to running a verified solver
- One verified tool can validate the results of many solvers

Proofs of Unsatisfiability: Initial Challenges

Theoretical challenges:

- Some "simple" problems have exponentially large proofs in the resolution proof system [Urquhart '87, Buss and Pitassi '98];
- While some dedicated techniques can quickly solve them.

Solution: A proof system to compactly express all techniques.

Proofs of Unsatisfiability: Initial Challenges

Theoretical challenges:

- Some "simple" problems have exponentially large proofs in the resolution proof system [Urquhart '87, Buss and Pitassi '98];
- While some dedicated techniques can quickly solve them.

Solution: A proof system to compactly express all techniques.

Practical challenges:

- Earlier efforts failed due to complexity and overhead
- Convince developers to support proof logging

Solution:

- The computational burden and complexity is in the checker
- A reference implementation of proof logging

Proofs of Unsatisfiability: Arbitrarily Complex Solvers

Verified checkers of certificates in strong proof systems:

- 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

Satisfiability and Mathematics

Proofs of Unsatisfiability

Future and Challenges

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

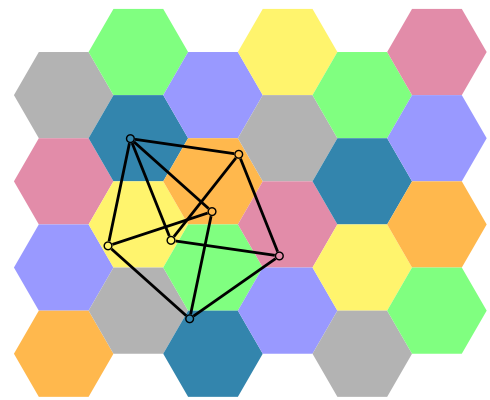
Future of Computer-Aided Mathematics

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?

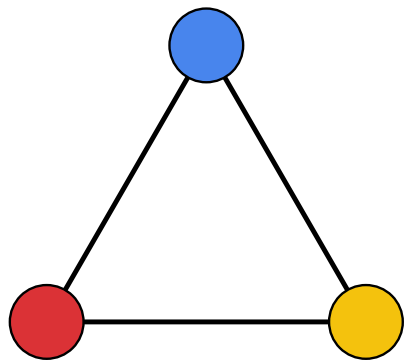
- Chromatic number of the plane
- Optimal matrix multiplication
- Collatz Conjecture



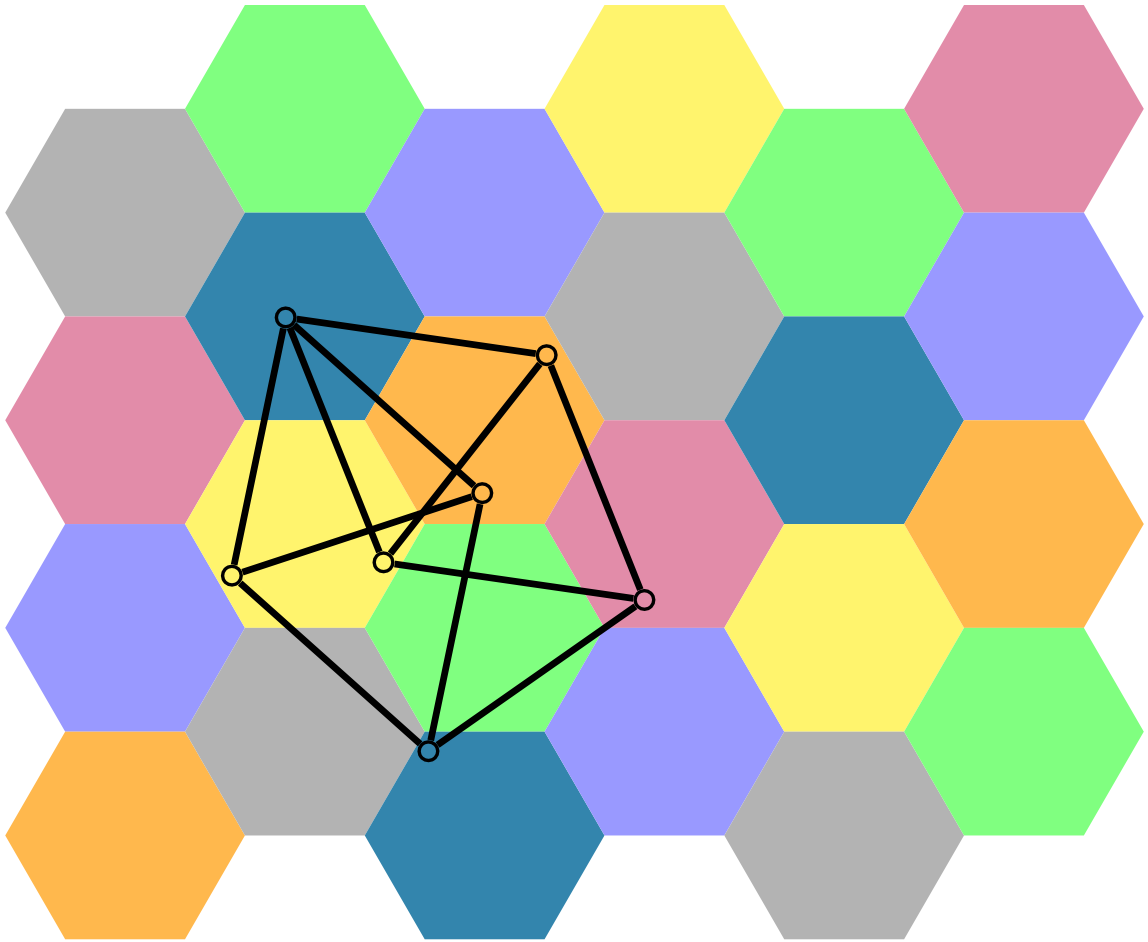
Chromatic Number of the Plane (CNP)

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.



CNP: 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

CNP: 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

CNP: 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

Quantamagazine Physics Mathematics

業餘數學家為一道填色難題帶來突破！

2018/4/26 • TNL • 四色定理、填色難題、數學

Раскраска для математиков

Как покрасить плоскость?

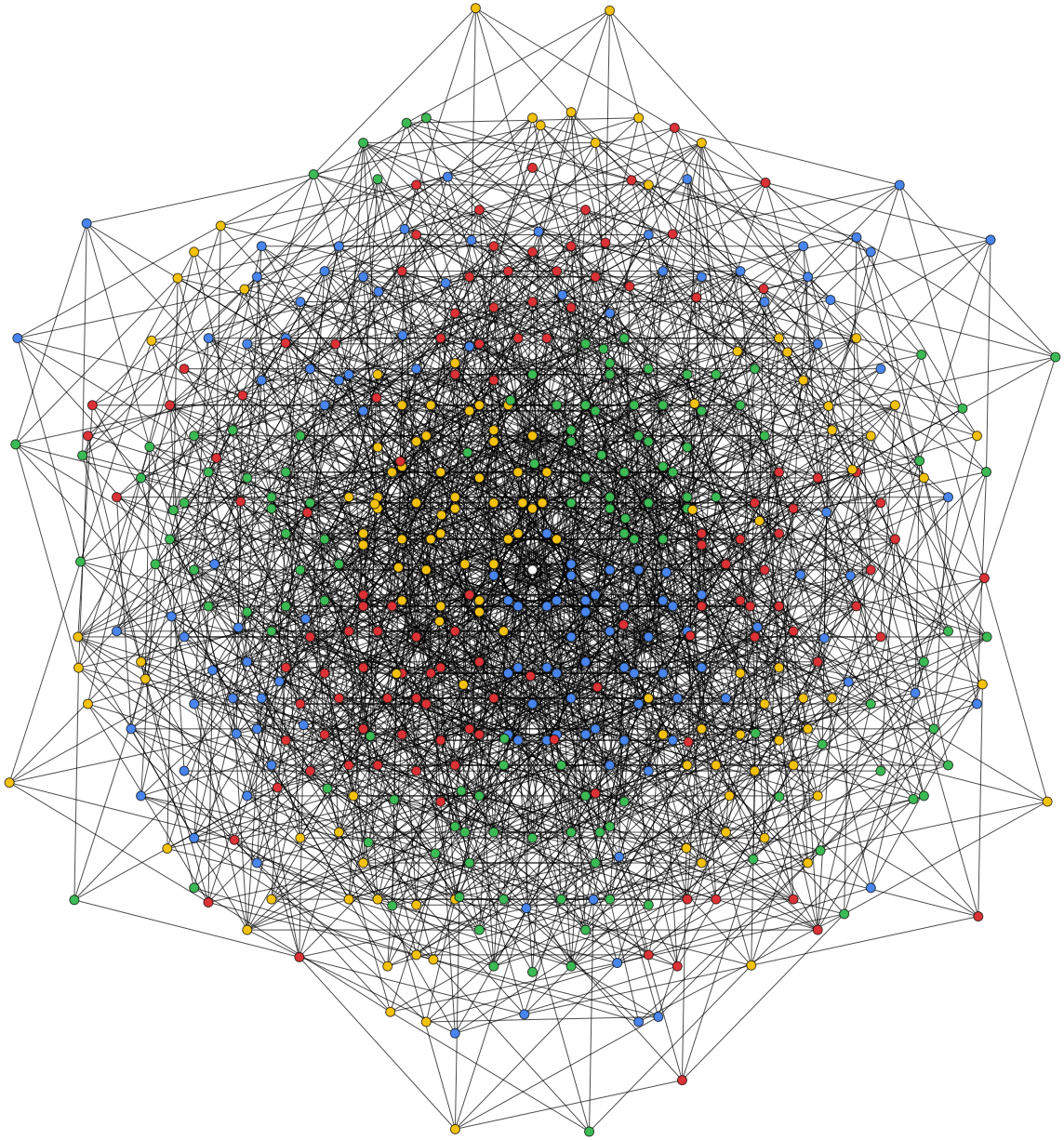
WIRED

Marijn Heule, a computer scientist at the University of Texas, Austin, found one with just 874 vertices. Yesterday he lowered this number to 826 vertices.

We found smaller graphs with SAT:

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

Proof Minimization: 510 Vertices [Heule 2021]



Matrix Multiplication: Introduction

$$\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix} = \begin{pmatrix} c_{1,1} & c_{1,2} \\ c_{2,1} & c_{2,2} \end{pmatrix}$$

$$\begin{aligned} c_{1,1} &= a_{1,1} \cdot b_{1,1} + a_{1,2} \cdot b_{2,1} \\ c_{1,2} &= a_{1,1} \cdot b_{1,2} + a_{1,2} \cdot b_{2,2} \\ c_{2,1} &= a_{2,1} \cdot b_{1,1} + a_{2,2} \cdot b_{2,1} \\ c_{2,2} &= a_{2,1} \cdot b_{1,2} + a_{2,2} \cdot b_{2,2} \end{aligned}$$

Matrix Multiplication: Introduction

$$\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix} = \begin{pmatrix} c_{1,1} & c_{1,2} \\ c_{2,1} & c_{2,2} \end{pmatrix}$$

$$\begin{aligned} c_{1,1} &= M_1 + M_4 - M_5 + M_7 \\ c_{1,2} &= M_3 + M_5 \\ c_{2,1} &= M_2 + M_4 \\ c_{2,2} &= M_1 - M_2 + M_3 + M_6 \end{aligned}$$

Matrix Multiplication: Introduction

$$\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix} = \begin{pmatrix} c_{1,1} & c_{1,2} \\ c_{2,1} & c_{2,2} \end{pmatrix}$$

. . . where

$$\begin{aligned}
M_1 &= (a_{1,1} + a_{2,2}) \cdot (b_{1,1} + b_{2,2}) \\
M_2 &= (a_{2,1} + a_{2,2}) \cdot b_{1,1} \\
M_3 &= a_{1,1} \cdot (b_{1,2} - b_{2,2}) \\
M_4 &= a_{2,2} \cdot (b_{2,1} - b_{1,1}) \\
M_5 &= (a_{1,1} + a_{1,2}) \cdot b_{2,2} \\
M_6 &= (a_{2,1} - a_{1,1}) \cdot (b_{1,1} + b_{1,2}) \\
M_7 &= (a_{1,2} - a_{2,2}) \cdot (b_{2,1} + b_{2,2})
\end{aligned}$$

The 3×3 Case is Still Open

7 multiplications for 2×2 matrices is optimal and unique

Question: What's the minimal number of multiplications needed to multiply two 3×3 matrices?

- naive algorithm: 27
- padd with zeros, use Strassen twice, cleanup: 25
- best known upper bound: 23 (Laderman 1976)
- best known lower bound: 19 (Bläser 2003)
- maximal number of multiplications allowed if we want to beat Strassen: 21 (because $\log_3 21 < \log_2 7 < \log_3 22$).

Other schemes [Heule, Kauers, & Seidl 2019]

- Using integer coefficients, there have so far been only three other schemes for 3×3 matrices and 23 multiplications.
- Using altogether about 35 years of computation time, we found more than 13000 new schemes for 3×3 and 23 using SAT, and we expect that there are many others.
- Unfortunately we found no scheme with only 22 multiplications yet.

Other schemes [Heule, Kauers, & Seidl 2019]

- Using integer coefficients, there have so far been only three other schemes for 3×3 matrices and 23 multiplications.
- Using altogether about 35 years of computation time, we found more than 13000 new schemes for 3×3 and 23 using SAT, and we expect that there are many others.
- Unfortunately we found no scheme with only 22 multiplications yet.

nature

Discovering faster matrix multiplication algorithms with reinforcement learning

- Kauers already improved the 5x5 bound days later to 95
- Next step: Use SAT to further improve these bounds

Beyond NP: The Collatz Conjecture

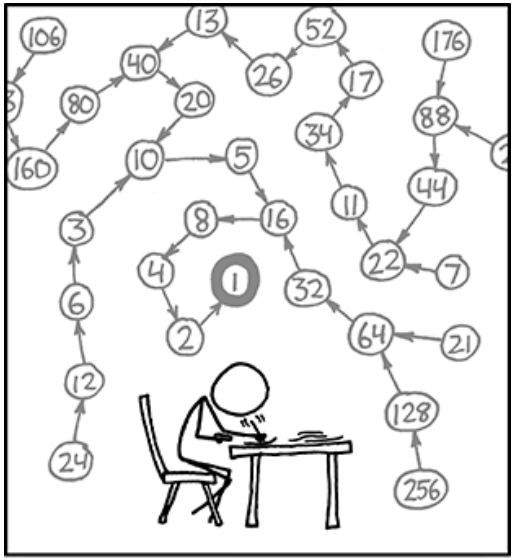
Resolving foundational algorithm questions

$$Col(n) = \begin{cases} n/2 & \text{if } n \text{ is even} \\ (3n+1)/2 & \text{if } n \text{ is odd} \end{cases}$$

Does `while`$(n > 1)\ n = Col(n);$ terminate?

Find a non-negative function $fun(n)$ s.t.

$$\forall n > 1 : fun(n) > fun(Col(n))$$



THE COLLATZ CONJECTURE STATES THAT IF YOU PICK A NUMBER, AND IF IT'S EVEN DIVIDE IT BY TWO AND IF IT'S ODD MULTIPLY IT BY THREE AND ADD ONE, AND YOU REPEAT THIS PROCEDURE LONG ENOUGH, EVENTUALLY YOUR FRIENDS WILL STOP CALLING TO SEE IF YOU WANT TO HANG OUT.

source: xkcd.com/710

Beyond NP: The Collatz Conjecture

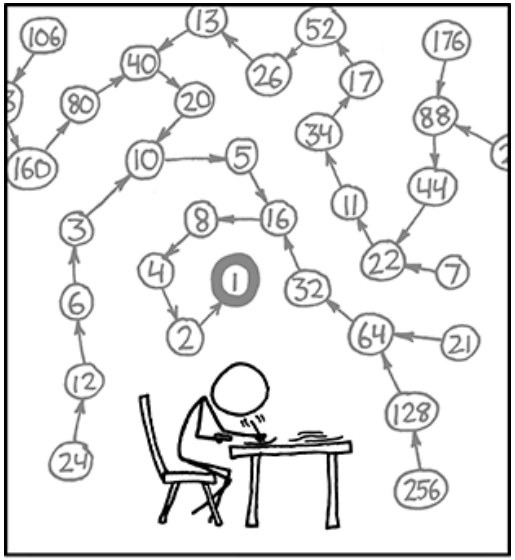
Resolving foundational algorithm questions

$$Col(n) = \begin{cases} n/2 & \text{if } n \text{ is even} \\ (3n+1)/2 & \text{if } n \text{ is odd} \end{cases}$$

Does `while`$(n > 1)\ n = Col(n);$ terminate?

Find a non-negative function $fun(n)$ s.t.

$$\forall n > 1 : fun(n) > fun(Col(n))$$



THE COLLATZ CONJECTURE STATES THAT IF YOU PICK A NUMBER, AND IF IT'S EVEN DIVIDE IT BY TWO AND IF IT'S ODD MULTIPLY IT BY THREE AND ADD ONE, AND YOU REPEAT THIS PROCEDURE LONG ENOUGH, EVENTUALLY YOUR FRIENDS WILL STOP CALLING TO SEE IF YOU WANT TO HANG OUT.

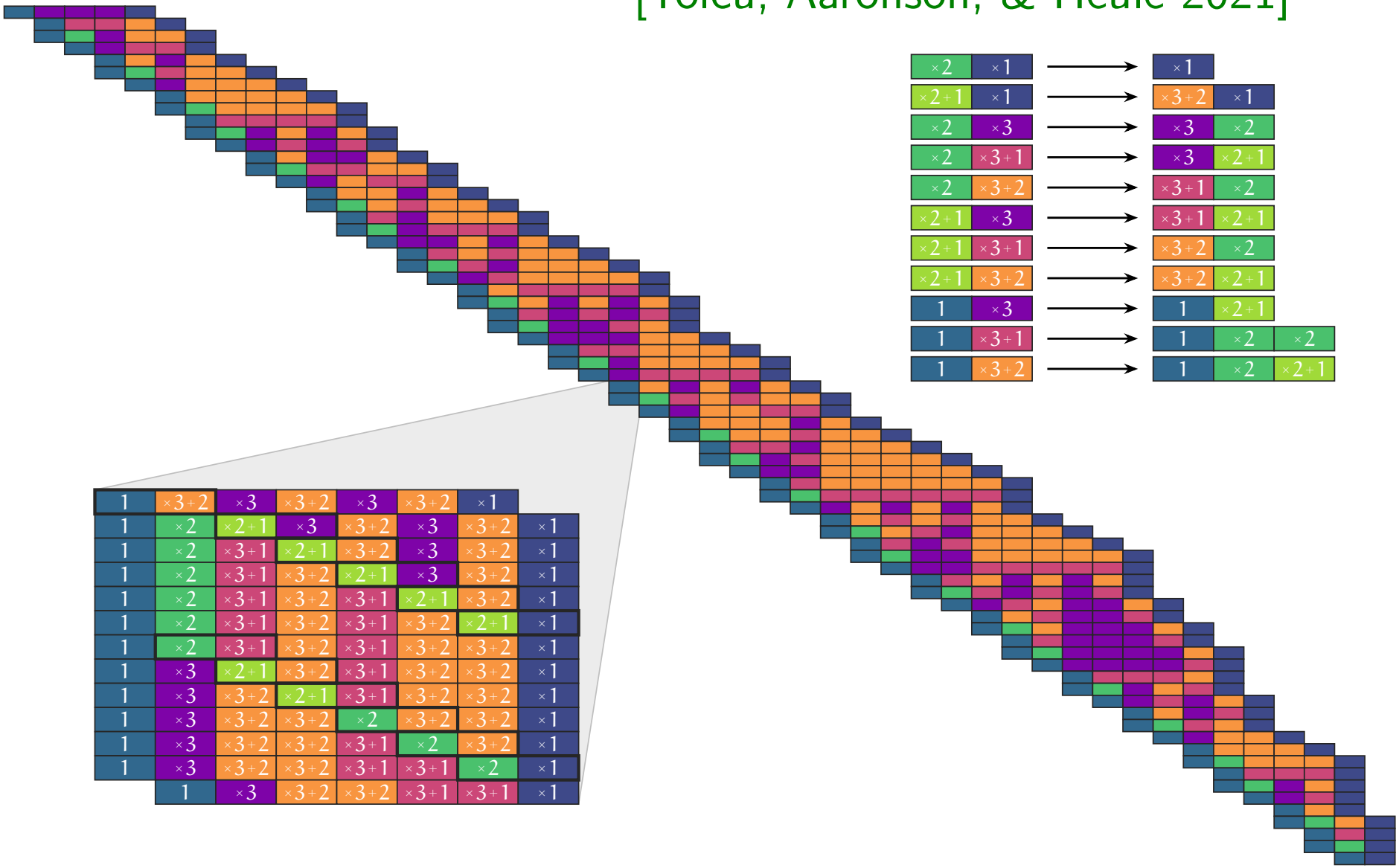
source: xkcd.com/710

Can we construct a function s.t. $fun(n) > fun(Col(n))$ holds?

$fun(3)$	$fun(5)$	$fun(8)$	$fun(4)$	$fun(2)$	$fun(1)$
5	4	3	2	1	0

Collatz Conjecture: Studying a Rewrite System

[Yolcu, Aaronson, & Heule 2021]



Collatz Conjecture: Successes and Challenge

Success. Rewrite system with 11 rules: Their termination solves Collatz. Our tool proves termination of any subset of 10 rules.

Collatz Conjecture: Successes and Challenge

Success. Rewrite system with 11 rules: Their termination solves Collatz. Our tool proves termination of any subset of 10 rules.

Success. Our tool proves termination of Farkas' variant:

$$F(n) = \begin{cases} \frac{n-1}{3} & \text{if } n \equiv 1 \pmod 3 \\ \frac{n}{2} & \text{if } n \equiv 0 \text{ or } n \equiv 2 \pmod 6 \\ \frac{3n+1}{2} & \text{if } n \equiv 3 \text{ or } n \equiv 5 \pmod 6 \end{cases}$$

Collatz Conjecture: Successes and Challenge

Success. Rewrite system with 11 rules: Their termination solves Collatz. Our tool proves termination of any subset of 10 rules.

Success. Our tool proves termination of Farkas' variant:

$$F(n) = \begin{cases} \frac{n-1}{3} & \text{if } n \equiv 1 \pmod 3 \\ \frac{n}{2} & \text{if } n \equiv 0 \text{ or } n \equiv 2 \pmod 6 \\ \frac{3n+1}{2} & \text{if } n \equiv 3 \text{ or } n \equiv 5 \pmod 6 \end{cases}$$

Challenge ($500). An easier generalized Collatz problem is open:

$$H(n) = \begin{cases} \frac{3n}{4} & \text{if } n \equiv 0 \pmod 4 \\ \frac{9n+1}{8} & \text{if } n \equiv 7 \pmod 8 \\ \bot & \text{otherwise} \end{cases}$$

Conclusions

Successes, Advances, and Trust:

- A performance boost of SAT technology allows solving new problems in mathematics
- Problems beyond NP are ready for an automated approach
- Some proofs may be gigantic, but can be validated using formally-verified checkers

Classic problems ready for mechanization?

- Chromatic number of the plane
- Optimal matrix multiplication
- Collatz Conjecture