Introduction to Automated Reasoning and Satisfiability

Marijn J.H. Heule

Carnegie Mellon University

http://www.cs.cmu.edu/~mheule/15816-f19/

Automated Reasoning and Satisfiability, September 3, 2019

Automated Reasoning Has Many Applications

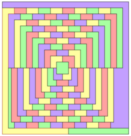
formal verification

security

bioinformatics

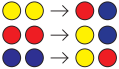
planning and scheduling

train safety



automated theorem proving

exploit generation



term rewriting termination

encode → **automated reasoning** → decode

Automated Reasoning Has Many Applications

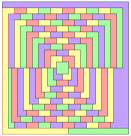
formal verification

security

bioinformatics

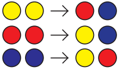
planning and scheduling

train safety



automated theorem proving

exploit generation



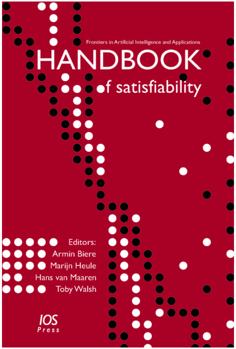
term rewriting termination

encode → **automated reasoning** → decode

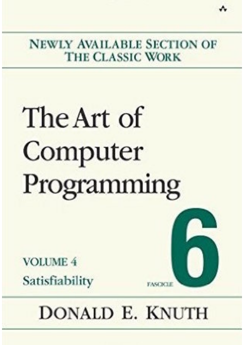
Breakthrough in SAT Solving in the Last 20 Years

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]

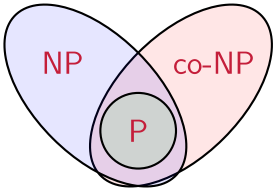


Donald Knuth: *"evidently a killer app, because it is key to the solution of so many other problems"* [Knuth '15]

Satisfiability and Complexity

Complexity classes of decision problems:

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



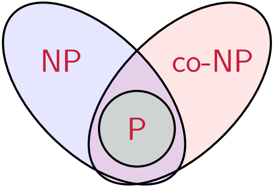
Cook-Levin Theorem [1971]: SAT is NP-complete.

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 beauty of NP: guaranteed short solutions.

The effectiveness of SAT solving: fast solutions in practice.

"NP is the new P!"

Course Overview

Lecture slides will be posted after each class meeting.

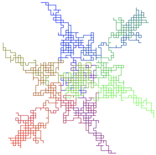
This schedule may change throughout the semester. Check back regularly for updates, including assignment deadlines and other important dates.

date	*topic*	*slides*	*notes*
9/3/19	Introduction to Automated Reasoning		
9/5/19	Applications for Automated Reasoning		
9/10/19	Representations for Automated Reasoning		
9/12/19	Proof Systems and Proof Complexity		*Homework 1 assigned*
9/17/19	Local Search Techniques		
9/19/19	Lookahead Techniques		*Homework 1 due*
9/24/19	Conflict-Driven Clause Learning		*Homework 2 assigned*
9/26/19	Preprocessing Techniuqes		
10/1/19	Maximum Satisfiability		*Homework 2 due*
10/3/19	Parallel Automated Reasoning		*Homework 3 assigned*
10/8/19	Quantified Boolean Formulas		
10/10/19	Verifying Automated Reasoning Results		*Homework 3 due*
10/15/19	**Select topic for final project and form groups**		

Course Reports

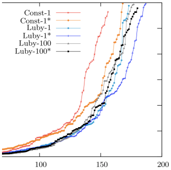
The second half of the course consists of a project

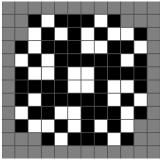
- A group of 2/3 students work on a research question
- The results will be presented in a scientific report
- Several have been published in journals and at conferences



Paul Herwig, Marijn Heule, Martijn van Lambalgen, and Hans van Maaren: A New Method to Construct Lower Bounds for Van der Waerden Numbers (2007). The Electronic Journal of Combinatorics 14 (R6).

Peter van der Tak, Antonio Ramos, and Marijn Heule: Reusing the Assignment Trail in CDCL Solvers (2011). Journal on Satisfiability, Boolean Modeling and Computation 7(4): 133-138.





Christiaan Hartman, Marijn Heule, Kees Kwekkeboom, and Alain Noels: Symmetry in Gardens of Eden (2013). The Electronic Journal of Combinatorics 20 (P16).

Introduction

Terminology

Basic Solving Techniques

Solvers and Benchmarks

Diplomacy Problem

"You are chief of protocol for the embassy ball. The crown prince instructs you either to invite *Peru* or to exclude *Qatar*. The queen asks you to invite either *Qatar* or *Romania* or both. The king, in a spiteful mood, wants to snub either *Romania* or *Peru* or both. Is there a guest list that will satisfy the whims of the entire royal family?"

Diplomacy Problem

"You are chief of protocol for the embassy ball. The crown prince instructs you either to invite *Peru* or to exclude *Qatar*. The queen asks you to invite either *Qatar* or *Romania* or both. The king, in a spiteful mood, wants to snub either *Romania* or *Peru* or both. Is there a guest list that will satisfy the whims of the entire royal family?"

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

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

Slightly Harder Example

Slightly Harder Example 1

What are the solutions for the following formula?

$$
\begin{aligned}
&(a \vee b \vee \overline{c}) \wedge \\
&(\overline{a} \vee \overline{b} \vee c) \wedge \\
&(b \vee c \vee \overline{d}) \wedge \\
&(\overline{b} \vee \overline{c} \vee d) \wedge \\
&(a \vee c \vee d) \wedge \\
&(\overline{a} \vee \overline{c} \vee \overline{d}) \wedge \\
&(\overline{a} \vee b \vee d)
\end{aligned}
$$

Slightly Harder Example

Slightly Harder Example 1

What are the solutions for the following formula?

$$(a \vee b \vee \overline{c}) \wedge (\overline{a} \vee \overline{b} \vee c) \wedge (b \vee c \vee \overline{d}) \wedge (\overline{b} \vee \overline{c} \vee d) \wedge (a \vee c \vee d) \wedge (\overline{a} \vee \overline{c} \vee \overline{d}) \wedge (\overline{a} \vee b \vee d)$$

a	b	c	d
0	0	0	0
0	0	0	1
0	0	1	0
0	0	1	1
0	1	0	0
0	1	0	1
0	1	1	0
0	1	1	1

a	b	c	d
1	0	0	0
1	0	0	1
1	0	1	0
1	0	1	1
1	1	0	0
1	1	0	1
1	1	1	0
1	1	1	1

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”

Introduction

Terminology

Basic Solving Techniques

Solvers and Benchmarks

Given a *CNF formula*,
does there exist an *assignment*
to the *Boolean variables*
that satisfies all *clauses*?

Terminology: Variables and literals

Boolean variable x_i

- can be assigned the Boolean values 0 or 1

Literal

- refers either to x_i or its complement $\overline{x}_i$
- literals x_i are satisfied if variable x_i is assigned to 1 (true)
- literals $\overline{x}_i$ are satisfied if variable x_i is assigned to 0 (false)

Terminology: Clauses

Clause

- Disjunction of literals: E.g. $C_j = (l_1 \vee l_2 \vee l_3)$
- Can be falsified with only *one* assignment to its literals: All literals assigned to false
- Can be satisfied with $2^k - 1$ assignment to its k literals
- One special clause - the empty clause (denoted by $\perp$) - which is always falsified

Terminology: Formulae

Formula

- Conjunction of clauses: E.g. $\mathcal{F} = C_1 \wedge C_2 \wedge C_3$
- Is satisfiable if there exists an assignment satisfying all clauses, otherwise unsatisfiable
- Formulae are defined in Conjunction Normal Form (CNF) and generally also stored as such - also learned information
- Any propositional formula can be efficiently transformed into CNF [Tseitin '70]

Terminology: Assignments

Assignment

- Mapping of the values 0 and 1 to the variables
- $\varphi \circ \mathcal{F}$ results in a reduced formula $\mathcal{F}_{\text{reduced}}$:
  - all satisfied clauses are removed
  - all falsified literals are removed
- satisfying assignment $\leftrightarrow$ $\mathcal{F}_{\text{reduced}}$ is empty
- falsifying assignment $\leftrightarrow$ $\mathcal{F}_{\text{reduced}}$ contains $\perp$
- partial assignment versus full assignment

Resolution

The most commonly used inference rule in propositional logic is the resolution rule (the operation is denoted by $\bowtie$)

$$\frac{C \vee x \qquad \bar{x} \vee D}{C \vee D}$$

Resolution

The most commonly used inference rule in propositional logic is the resolution rule (the operation is denoted by $\bowtie$)

$$\frac{C \vee x \qquad \bar{x} \vee D}{C \vee D}$$

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

- $(\overline{q} \vee p) \bowtie (\overline{p} \vee \overline{r}) = (\overline{q} \vee \overline{r})$
- $(p \vee \overline{q}) \bowtie (q \vee r) = (p \vee r)$
- $(q \vee r) \bowtie (\overline{r} \vee \overline{p}) = (q \vee \overline{p})$

Resolution

The most commonly used inference rule in propositional logic is the resolution rule (the operation is denoted by $\bowtie$)

$$\frac{C \vee x \qquad \bar{x} \vee D}{C \vee D}$$

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

- $(\overline{q} \vee p) \bowtie (\overline{p} \vee \overline{r}) = (\overline{q} \vee \overline{r})$
- $(p \vee \overline{q}) \bowtie (q \vee r) = (p \vee r)$
- $(q \vee r) \bowtie (\overline{r} \vee \overline{p}) = (q \vee \overline{p})$

Adding (non-redundant) resolvents until fixpoint, is a complete proof procedure. It produces the empty clause if and only if the formula is unsatisfiable

Tautology

A clause C is a tautology if it contains
for some variable x, both the literals x and $\overline{x}$.

Slightly Harder Example 2

Compute all non-tautological resolvents for:

$$
\begin{aligned}
&(a \vee b \vee \overline{c}) \wedge (\overline{a} \vee \overline{b} \vee c) \wedge \\
&(b \vee c \vee \overline{d}) \wedge (\overline{b} \vee \overline{c} \vee d) \wedge \\
&(a \vee c \vee d) \wedge (\overline{a} \vee \overline{c} \vee \overline{d}) \wedge \\
&(\overline{a} \vee b \vee d)
\end{aligned}
$$

Which resolvents remain after removing the supersets?

SAT solving: Unit propagation

A *unit clause* is a clause of size 1

UnitPropagation $(\varphi, \mathcal{F})$:

while $\bot \notin \mathcal{F}$ **and** unit clause y exists **do**
    expand φ by adding $y = 1$ and simplify $\mathcal{F}$
end while
return $\varphi, \mathcal{F}$

Unit Propagation: Example

$$\mathcal{F}_{\text{unit}} := (\overline{x}_1 \vee \overline{x}_3 \vee x_4) \wedge (\overline{x}_1 \vee \overline{x}_2 \vee x_3) \wedge (\overline{x}_1 \vee x_2) \wedge (x_1 \vee x_3 \vee x_6) \wedge (\overline{x}_1 \vee x_4 \vee \overline{x}_5) \wedge (x_1 \vee \overline{x}_6) \wedge (x_4 \vee x_5 \vee x_6) \wedge (x_5 \vee \overline{x}_6)$$

Unit Propagation: Example

$$\mathcal{F}_{\text{unit}} := (\overline{x}_1 \vee \overline{x}_3 \vee x_4) \wedge (\overline{x}_1 \vee \overline{x}_2 \vee x_3) \wedge (\overline{x}_1 \vee x_2) \wedge (x_1 \vee x_3 \vee x_6) \wedge (\overline{x}_1 \vee x_4 \vee \overline{x}_5) \wedge (x_1 \vee \overline{x}_6) \wedge (x_4 \vee x_5 \vee x_6) \wedge (x_5 \vee \overline{x}_6)$$

$$\varphi = \{x_1{=}1\}$$

Unit Propagation: Example

$$\mathcal{F}_{\text{unit}} := (\overline{x}_1 \vee \overline{x}_3 \vee x_4) \wedge (\overline{x}_1 \vee \overline{x}_2 \vee x_3) \wedge (\overline{x}_1 \vee x_2) \wedge (x_1 \vee x_3 \vee x_6) \wedge (\overline{x}_1 \vee x_4 \vee \overline{x}_5) \wedge (x_1 \vee \overline{x}_6) \wedge (x_4 \vee x_5 \vee x_6) \wedge (x_5 \vee \overline{x}_6)$$

$$\varphi = \{x_1{=}1, x_2{=}1\}$$

Unit Propagation: Example

$$\mathcal{F}_{\text{unit}} := (\overline{x}_1 \vee \overline{x}_3 \vee x_4) \wedge (\overline{x}_1 \vee \overline{x}_2 \vee x_3) \wedge (\overline{x}_1 \vee x_2) \wedge (x_1 \vee x_3 \vee x_6) \wedge (\overline{x}_1 \vee x_4 \vee \overline{x}_5) \wedge (x_1 \vee \overline{x}_6) \wedge (x_4 \vee x_5 \vee x_6) \wedge (x_5 \vee \overline{x}_6)$$

$$\varphi = \{x_1{=}1, x_2{=}1, x_3{=}1\}$$

Unit Propagation: Example

$$\mathcal{F}_{\text{unit}} := (\overline{x}_1 \vee \overline{x}_3 \vee x_4) \wedge (\overline{x}_1 \vee \overline{x}_2 \vee x_3) \wedge (\overline{x}_1 \vee x_2) \wedge (x_1 \vee x_3 \vee x_6) \wedge (\overline{x}_1 \vee x_4 \vee \overline{x}_5) \wedge (x_1 \vee \overline{x}_6) \wedge (x_4 \vee x_5 \vee x_6) \wedge (x_5 \vee \overline{x}_6)$$

$$\varphi = \{x_1{=}1, x_2{=}1, x_3{=}1, x_4{=}1\}$$

SAT Solving: DPLL

Davis Putnam Logemann Loveland [DP60,DLL62]

Recursive procedure that in each recursive call:

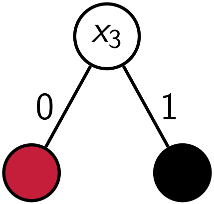
- Simplifies the formula (using unit propagation)
- Splits the formula into two subformulas
  - Variable selection heuristics (which variable to split on)
  - Direction heuristics (which subformula to explore first)

DPLL: Example

$$\mathcal{F}_{\mathrm{DPLL}} := (x_1 \vee x_2 \vee \overline{x}_3) \wedge (\overline{x}_1 \vee x_2 \vee x_3) \wedge (\overline{x}_1 \vee \overline{x}_2 \vee x_3) \wedge (x_1 \vee x_3) \wedge (\overline{x}_1 \vee \overline{x}_3)$$

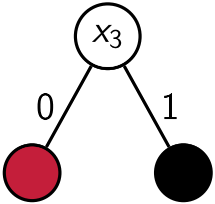
DPLL: Example

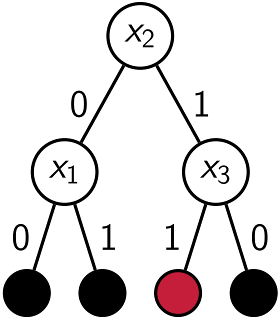
$$\mathcal{F}_{\text{DPLL}} := (x_1 \vee x_2 \vee \overline{x}_3) \wedge (\overline{x}_1 \vee x_2 \vee x_3) \wedge (\overline{x}_1 \vee \overline{x}_2 \vee x_3) \wedge (x_1 \vee x_3) \wedge (\overline{x}_1 \vee \overline{x}_3)$$



DPLL: Example

$$\mathcal{F}_{\mathrm{DPLL}} := (x_1 \vee x_2 \vee \overline{x}_3) \wedge (\overline{x}_1 \vee x_2 \vee x_3) \wedge (\overline{x}_1 \vee \overline{x}_2 \vee x_3) \wedge (x_1 \vee x_3) \wedge (\overline{x}_1 \vee \overline{x}_3)$$





DPLL: Slightly Harder Example

Slightly Harder Example 3

Construct a DPLL tree for:

$$(a \vee b \vee \overline{c}) \wedge (\overline{a} \vee \overline{b} \vee c) \wedge$$
$$(b \vee c \vee \overline{d}) \wedge (\overline{b} \vee \overline{c} \vee d) \wedge$$
$$(a \vee c \vee d) \wedge (\overline{a} \vee \overline{c} \vee \overline{d}) \wedge$$
$$(\overline{a} \vee b \vee d)$$

SAT Solving: Decision and Implications

Decision variables

- Variable selection heuristics and direction heuristics
- Play a crucial role in performance

Implied variables

- Assigned by reasoning (e.g. unit propagation)
- Maximizing the number of implied variables is an important aspect of look-ahead SAT solvers

SAT Solving: Clauses $\leftrightarrow$ assignments

- A clause C represents a set of falsified assignments, i.e. those assignments that falsify all literals in C
- A falsifying assignment φ for a given formula represents a set of clauses that follow from the formula
  - For instance with all decision variables
  - Important feature of conflict-driven SAT solvers

SAT Solving Paradigms

Conflict-driven

- search for short refutation, complete
- examples: lingeling, glucose, CaDiCaL

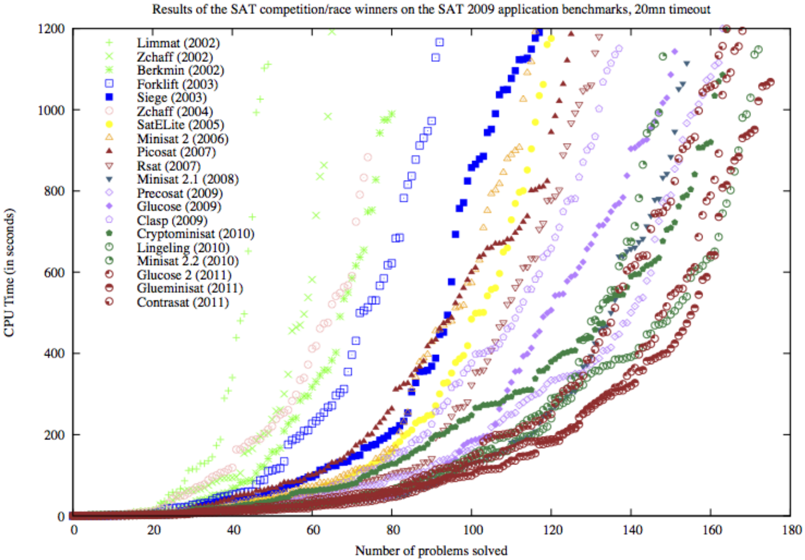
Look-ahead

- extensive inference, complete
- examples: march, OKsolver, kcnfs

Local search

- local optimizations, incomplete
- examples: probSAT, UnitWalk, Dimetheus

Progress of SAT Solvers



Applications: Industrial

- Model checking
  - Turing award '07 Clarke, Emerson, and Sifakis
- Software verification
- Hardware verification
- Equivalence checking
- Planning and scheduling
- Cryptography
- Car configuration
- Railway interlocking

Applications: Crafted

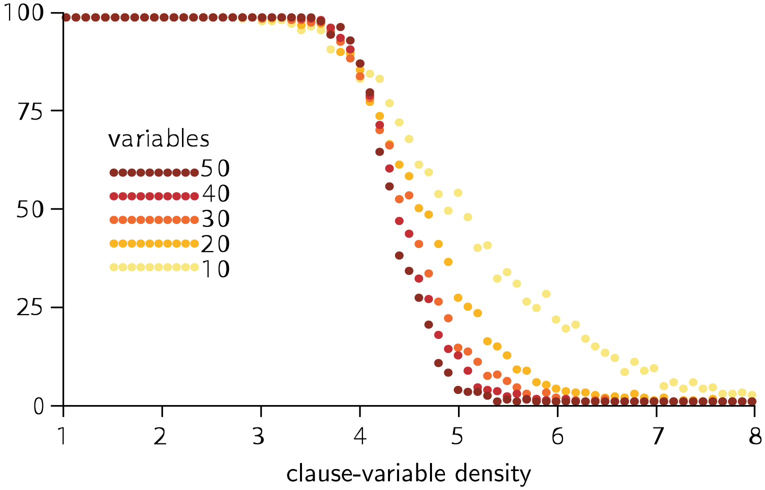
Combinatorial challenges and solver obstruction instances

- Pigeon-hole problems
- Tseitin problems
- Mutilated chessboard problems
- Sudoku
- Factorization problems
- Ramsey theory
- Rubik's cube puzzles

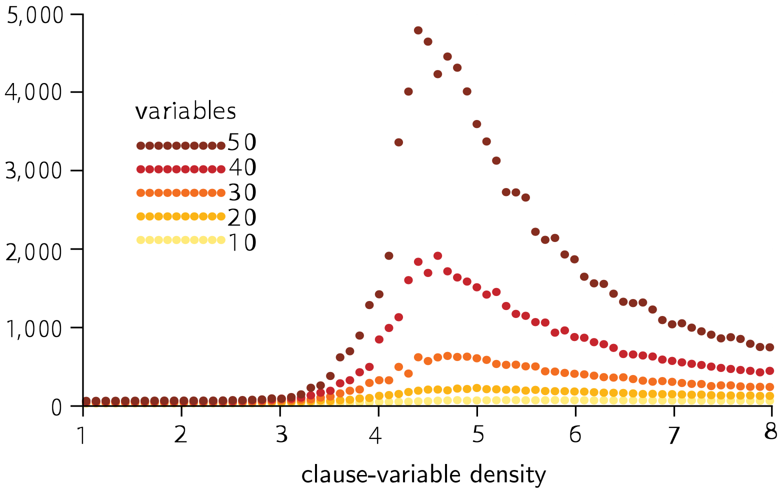
Random k-SAT: Introduction

- All clauses have length k
- Variables have the same probability to occur
- Each literal is negated with probability of 50%
- Density is ratio Clauses to Variables

Random 3-SAT: % satisfiable, the phase transition



Random 3-SAT: exponential runtime, the threshold



SAT Game

by Olivier Roussel

http://www.cs.utexas.edu/~marijn/game/

Introduction to Automated Reasoning and Satisfiability

Marijn J.H. Heule

Carnegie Mellon University

http://www.cs.cmu.edu/~mheule/15816-f19/