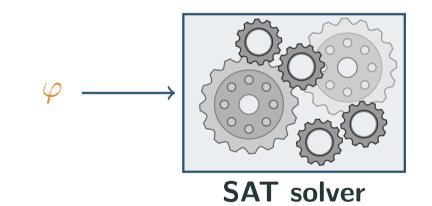
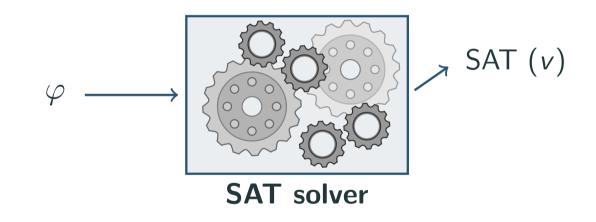
SAT Solving and Its Applications

Copyright: Sarah Winkler (Free University of Bozen-Bolzano)

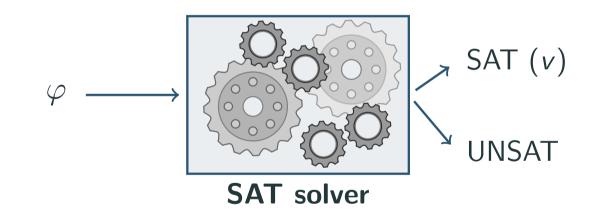
input: propositional formula φ



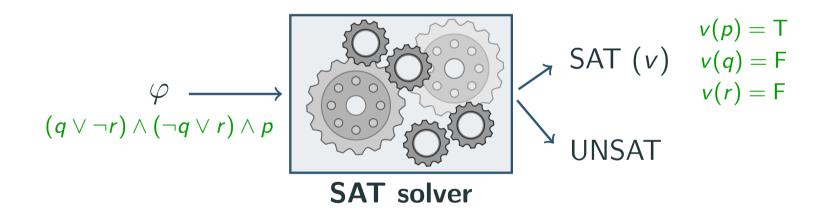
input:propositional formula φ output:SAT + valuation v such that $v(\varphi) = T$ if φ satisfiable



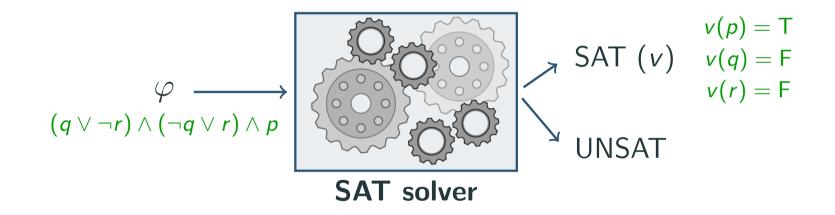
input: output: propositional formula φ SAT + valuation v such that $v(\varphi) = T$ if φ satisfiableUNSATotherwise



input: output: propositional formula φ SAT + valuation v such that $v(\varphi) = T$ if φ satisfiableUNSATotherwise



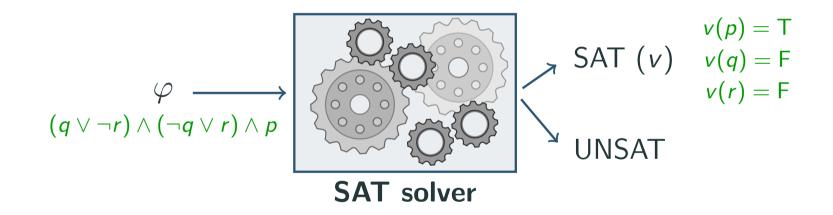
input:propositional formula φ output:SAT + valuation v such that $v(\varphi) = T$ if φ satisfiableUNSATotherwise



Terminology

decision problem P is problem with answer yes or no

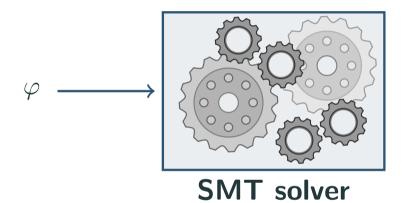
input:propositional formula φ output:SAT + valuation v such that $v(\varphi) = T$ if φ satisfiableUNSATotherwise



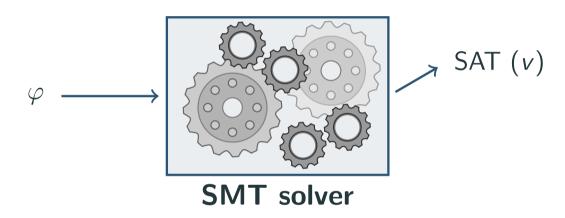
Terminology

- decision problem P is problem with answer yes or no
- ► SAT encoding of decision problem P is propositional formula φ_P such that answer to P is yes $\iff \varphi_P$ is satisfiable

input: formula φ involving theory T

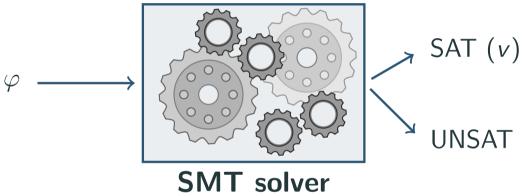


input: output: formula φ involving theory TSAT + valuation v such that $v(\varphi) = T$ if φ is T-satisfiable

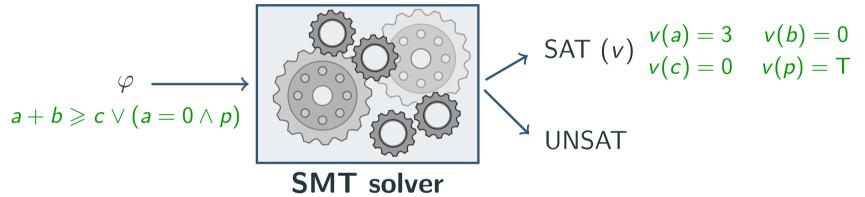


input: output: formula φ involving theory TSAT + valuation v such that $v(\varphi) = T$ UNSAT

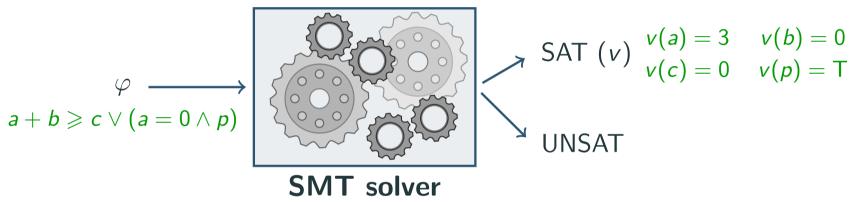
if φ is T-satisfiable otherwise



input: output: formula φ involving theory TSAT + valuation v such that $v(\varphi) = T$ if φ is T-satisfiable UNSAT otherwise



input: output: formula φ involving theory TSAT + valuation v such that $v(\varphi) = T$ if φ is T-satisfiable UNSAT otherwise

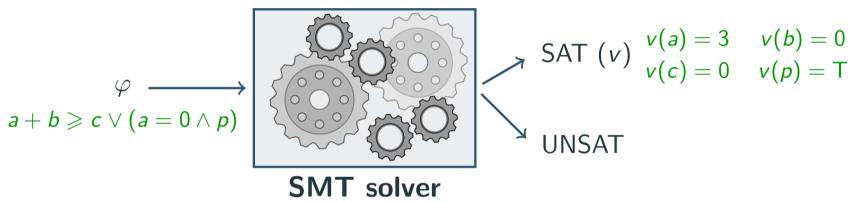


Example (Theories)

► arithmetic

 $2a + b \ge c \lor (a = 0 \land p)$

input: output: formula φ involving theory TSAT + valuation v such that $v(\varphi) = T$ if φ is T-satisfiable UNSAT otherwise

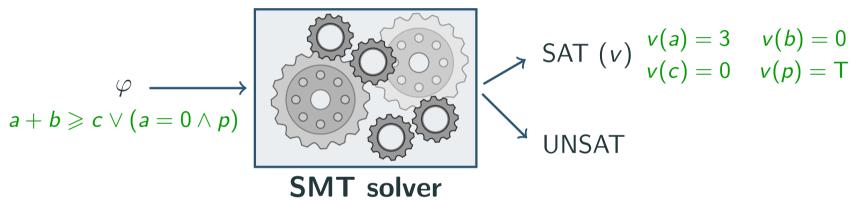


Example (Theories)

- ► arithmetic
- uninterpreted functions

 $2a + b \ge c \lor (a = 0 \land p)$ f(x,y) \ne f(y,x) \lapha g(f(x,x)) = g(y)

input: output: formula φ involving theory TSAT + valuation v such that $v(\varphi) = T$ if φ is T-satisfiable UNSAT otherwise

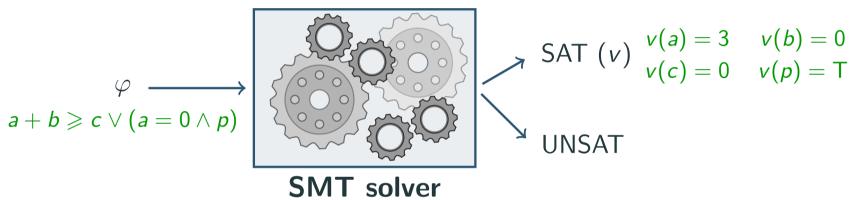


Example (Theories)

- ► arithmetic
- uninterpreted functions
- bit vectors

 $2a + b \ge c \lor (a = 0 \land p)$ $f(x, y) \ne f(y, x) \land g(f(x, x)) = g(y)$ $((\text{zext}_{32} a_8) + b_{32}) \times c_{32} \ge_u 0_{32}$

input: output: formula φ involving theory TSAT + valuation v such that $v(\varphi) = T$ if φ is T-satisfiable UNSAT otherwise



Example (Theories)

- ► arithmetic
- uninterpreted functions
- ► bit vectors

 $2a + b \ge c \lor (a = 0 \land p)$ $f(x, y) \ne f(y, x) \land g(f(x, x)) = g(y)$ $((zext_{32} a_8) + b_{32}) \times c_{32} \ge_u 0_{32}$

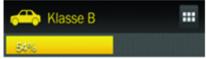
Terminology

▶ SMT encoding over theory T of decision problem P is formula φ_P such that

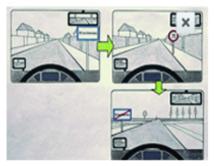
answer to P is yes $\iff \varphi_P$ is satisfiable

Austrian driving license test consists of 80 questions out of 1500 such that the following conditions are satisfied:

- ▶ 30 questions "main questions" with 3 sub-questions each
- ▶ at least 12 main questions must be about crossroads
- ► at least 12 questions must have pictures
- ▶ at least 5 "hard", "medium", and "easy" main questions

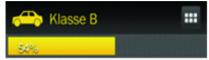


Frage 16.0 3 Punkte

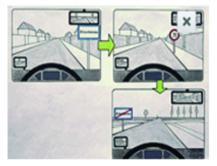


Austrian driving license test consists of 80 questions out of 1500 such that the following conditions are satisfied:

- ▶ 30 questions "main questions" with 3 sub-questions each
- ▶ at least 12 main questions must be about crossroads
- ► at least 12 questions must have pictures
- ▶ at least 5 "hard", "medium", and "easy" main questions
- ▶ how can software find valid question set?



Frage 16.0 3 Punkte

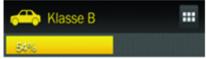


Austrian driving license test consists of 80 questions out of 1500 such that the following conditions are satisfied:

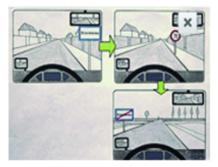
- ▶ 30 questions "main questions" with 3 sub-questions each
- ▶ at least 12 main questions must be about crossroads
- ► at least 12 questions must have pictures
- ▶ at least 5 "hard", "medium", and "easy" main questions
- ▶ how can software find valid question set?

SAT Encoding

• variables q_i for $1 \leq i \leq 1500$



Frage 16.0 3 Punkte



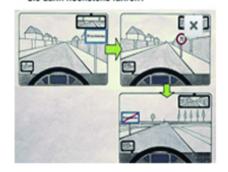
Austrian driving license test consists of 80 questions out of 1500 such that the following conditions are satisfied:

- ▶ 30 questions "main questions" with 3 sub-questions each
- ▶ at least 12 main questions must be about crossroads
- ► at least 12 questions must have pictures
- ▶ at least 5 "hard", "medium", and "easy" main questions
- how can software find valid question set?

SAT Encoding

- variables q_i for $1 \leq i \leq 1500$
- ▶ idea: valuation v sets $v(q_i) = T$ if question i is included, $v(q_i) = F$ otherwise





Austrian driving license test consists of 80 questions out of 1500 such that the following conditions are satisfied:

- ▶ 30 questions "main questions" with 3 sub-questions each
- ► at least 12 main questions must be about crossroads
- ► at least 12 questions must have pictures
- ▶ at least 5 "hard", "medium", and "easy" main questions
- ▶ how can software find valid question set?

SAT Encoding

► $\sum_{i \in Q_{\text{vroads}}} q_i \ge 12$

- variables q_i for $1 \leq i \leq 1500$
- ▶ idea: valuation v sets $v(q_i) = T$ if question i is included, $v(q_i) = F$ otherwise





Austrian driving license test consists of 80 questions out of 1500 such that the following conditions are satisfied:

- ▶ 30 questions "main questions" with 3 sub-questions each
- ▶ at least 12 main questions must be about crossroads
- ► at least 12 questions must have pictures
- ▶ at least 5 "hard", "medium", and "easy" main questions
- ▶ how can software find valid question set?

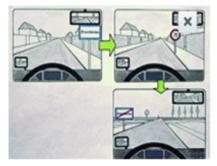
SAT Encoding

- variables q_i for $1 \leq i \leq 1500$
- ▶ idea: valuation v sets $v(q_i) = T$ if question i is included, $v(q_i) = F$ otherwise

$$\blacktriangleright \sum_{i \in Q_{\text{xroads}}} q_i \ge 12 \qquad \blacktriangleright \sum_{i \in Q_{\text{pictures}}} q_i \ge 12$$



biegen nach dem Ve



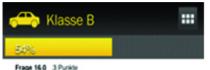
Austrian driving license test consists of 80 questions out of 1500 such that the following conditions are satisfied:

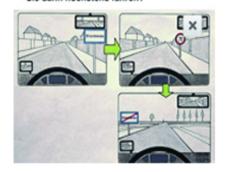
- ▶ 30 questions "main questions" with 3 sub-questions each
- ▶ at least 12 main questions must be about crossroads
- ► at least 12 questions must have pictures
- ▶ at least 5 "hard", "medium", and "easy" main questions
- how can software find valid question set?

SAT Encoding

- variables q_i for $1 \leq i \leq 1500$
- ▶ idea: valuation v sets $v(q_i) = T$ if question i is included, $v(q_i) = F$ otherwise

$$\blacktriangleright \sum_{i \in Q_{\text{xroads}}} q_i \geqslant 12 \qquad \blacktriangleright \sum_{i \in Q_{\text{pictures}}} q_i \geqslant 12 \qquad \blacktriangleright \sum_{i \in Q_{\text{hard}}} q_i \geqslant 5 \qquad \blacktriangleright \ldots$$





Austrian driving license test consists of 80 questions out of 1500 such that the following conditions are satisfied:

- ▶ 30 questions "main questions" with 3 sub-questions each
- ▶ at least 12 main questions must be about crossroads
- ► at least 12 questions must have pictures
- ▶ at least 5 "hard", "medium", and "easy" main questions
- ▶ how can software find valid question set?

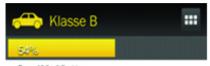
SAT Encoding

- variables q_i for $1 \leq i \leq 1500$
- ▶ idea: valuation v sets $v(q_i) = T$ if question i is included, $v(q_i) = F$ otherwise

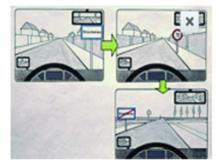
$$\blacktriangleright \sum_{i \in Q_{\text{xroads}}} q_i \geqslant 12 \qquad \blacktriangleright \sum_{i \in Q_{\text{pictures}}} q_i \geqslant 12 \qquad \blacktriangleright \sum_{i \in Q_{\text{hard}}} q_i \geqslant 5 \qquad \blacktriangleright \ldots$$

Result

easy generation of valid question sets (with some random preselection)



Frage 16.0 3 Punkte



Can one color all natural numbers with two colors such that whenever $x^2 + y^2 = z^2$ not all of x, y, and z have same color?

Can one color all natural numbers with two colors such that whenever $x^2 + y^2 = z^2$ not all of x, y, and z have same color?

Example

$$3^{2} + 4^{2} = 5^{2} \qquad 5^{2} + 12^{2} = 13^{2}$$
(a) 1 2 3 4 5 6 7 8 9 10 11 12 13 ... \checkmark
(b) 1 2 3 4 5 6 7 8 9 10 11 12 13 ... \checkmark

Can one color all natural numbers with two colors such that whenever $x^2 + y^2 = z^2$ not all of x, y, and z have same color?

Example

 $3^{2} + 4^{2} = 5^{2} \qquad 5^{2} + 12^{2} = 13^{2}$ (a) 1 2 3 4 5 6 7 8 9 10 11 12 13 ... \checkmark (b) 1 2 3 4 5 6 7 8 9 10 11 12 13 ... \checkmark

SAT Encoding

▶ variables x_i for $1 \leq i \leq n$ such that x_i becomes true iff it is colored red

Can one color all natural numbers with two colors such that whenever $x^2 + y^2 = z^2$ not all of x, y, and z have same color?

Example

 $3^{2} + 4^{2} = 5^{2} \qquad 5^{2} + 12^{2} = 13^{2}$ (a) 1 2 3 4 5 6 7 8 9 10 11 12 13 ... \checkmark (b) 1 2 3 4 5 6 7 8 9 10 11 12 13 ... \checkmark

SAT Encoding

- ▶ variables x_i for $1 \leq i \leq n$ such that x_i becomes true iff it is colored red
- SAT encoding: for all a² + b² = c² include (x_a ∨ x_b ∨ x_c) ∧ (x̄_a ∨ x̄_b ∨ x̄_c) (+ symmetry breaking, simplification, heuristics)

Can one color all natural numbers with two colors such that whenever $x^2 + y^2 = z^2$ not all of x, y, and z have same color?

Example

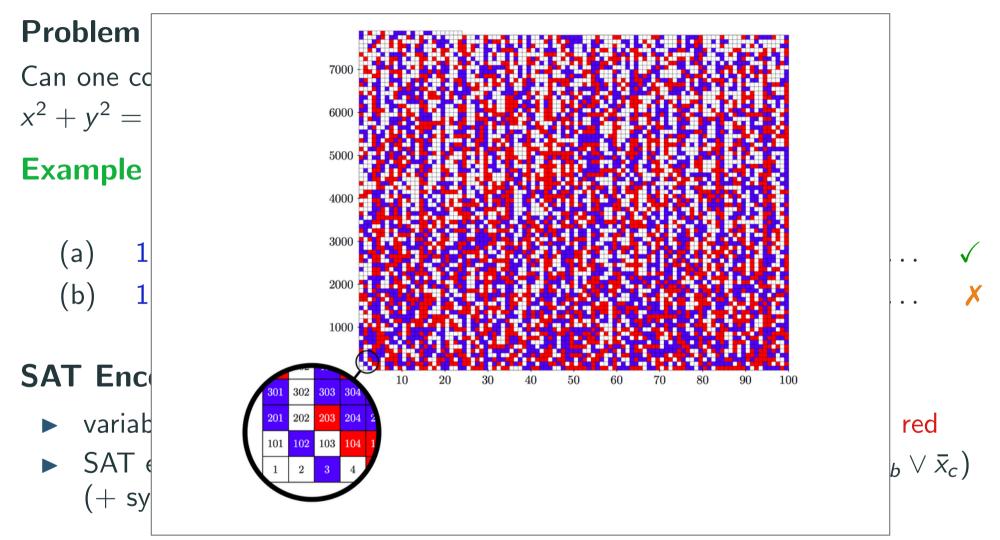
 $3^{2} + 4^{2} = 5^{2} \qquad 5^{2} + 12^{2} = 13^{2}$ (a) 1 2 3 4 5 6 7 8 9 10 11 12 13 ... \checkmark (b) 1 2 3 4 5 6 7 8 9 10 11 12 13 ... \checkmark

SAT Encoding

- ▶ variables x_i for $1 \leq i \leq n$ such that x_i becomes true iff it is colored red
- SAT encoding: for all a² + b² = c² include (x_a ∨ x_b ∨ x_c) ∧ (x̄_a ∨ x̄_b ∨ x̄_c) (+ symmetry breaking, simplification, heuristics)

Result: No. Coloring exists only up to 7,825.

Application: Pythagorean Triples



Result: No. Coloring exists only up to 7,825.

Can one color all natural numbers with two colors such that whenever $x^2 + y^2 = z^2$ not all of x, y, and z have same color?

Example

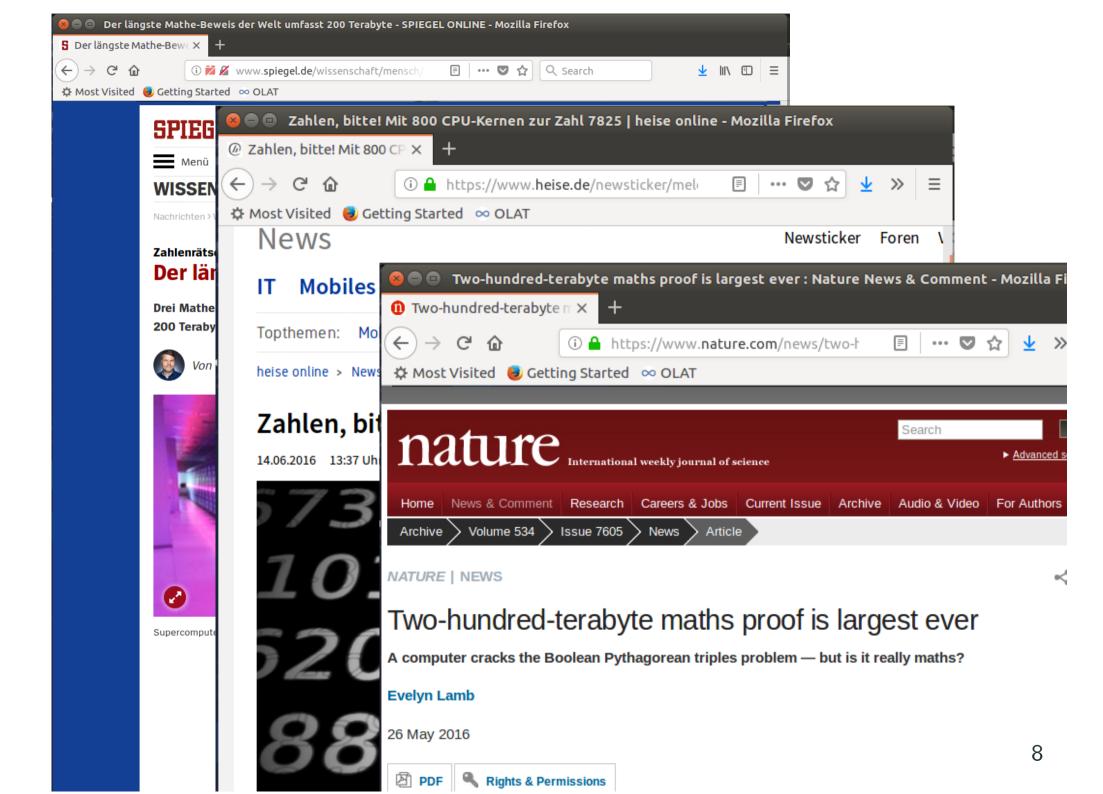
 $3^{2} + 4^{2} = 5^{2} \qquad 5^{2} + 12^{2} = 13^{2}$ (a) 1 2 3 4 5 6 7 8 9 10 11 12 13 ... \checkmark (b) 1 2 3 4 5 6 7 8 9 10 11 12 13 ... \checkmark

SAT Encoding

- ▶ variables x_i for $1 \leq i \leq n$ such that x_i becomes true iff it is colored red
- SAT encoding: for all a² + b² = c² include (x_a ∨ x_b ∨ x_c) ∧ (x̄_a ∨ x̄_b ∨ x̄_c) (+ symmetry breaking, simplification, heuristics)

Result: No. Coloring exists only up to 7,825.

1000s of variables, solving time 2 days with 800 processors, 200 TB of proof



Schedule sports league tournament for *n* teams, *p* periods of n - 1 rounds each (+ venue restrictions, break restrictions, ...)

Schedule sports league tournament for *n* teams, *p* periods of n - 1 rounds each (+ venue restrictions, break restrictions, ...)

Example (Österreichische Fußball-Bundesliga)

10 teams play in 4 periods (9 rounds each), periods 1 & 2 and 3 & 4 mirrored

Schedule sports league tournament for *n* teams, *p* periods of n - 1 rounds each (+ venue restrictions, break restrictions, ...)

Example (Österreichische Fußball-Bundesliga)

10 teams play in 4 periods (9 rounds each), periods 1 & 2 and 3 & 4 mirrored

(Part of) SAT Encoding

▶ variable x_{ijpr} is true if team *i* plays team *j* at home in period *p*, round *r*

Schedule sports league tournament for *n* teams, *p* periods of n - 1 rounds each (+ venue restrictions, break restrictions, ...)

Example (Österreichische Fußball-Bundesliga)

10 teams play in 4 periods (9 rounds each), periods 1 & 2 and 3 & 4 mirrored

(Part of) SAT Encoding

▶ variable x_{ijpr} is true if team *i* plays team *j* at home in period *p*, round *r*

 $\blacktriangleright \qquad \bigwedge_{i,p,r} \bigvee_{j \neq i} (x_{ijpr} \lor x_{jipr})$

each team plays in every round

Schedule sports league tournament for n teams, p periods of n - 1 rounds each (+ venue restrictions, break restrictions, ...)

Example (Österreichische Fußball-Bundesliga)

10 teams play in 4 periods (9 rounds each), periods 1 & 2 and 3 & 4 mirrored

(Part of) SAT Encoding

▶ variable x_{ijpr} is true if team *i* plays team *j* at home in period *p*, round *r*

$$\bigwedge_{i,p,r} \bigvee_{j \neq i} (x_{ijpr} \lor x_{jipr})$$
each team plays in every round

 $\bigwedge_{i,p,r} \bigwedge_{j \neq i} \bigwedge_{k \neq i \land k \neq j} (x_{ijpr} \to \neg (x_{ikpr} \lor x_{kipr})) \quad \text{each team plays at most once in every round}$

Problem: Round Robin Scheduling

Schedule sports league tournament for *n* teams, *p* periods of n - 1 rounds each (+ venue restrictions, break restrictions, ...)

Example (Österreichische Fußball-Bundesliga)

10 teams play in 4 periods (9 rounds each), periods 1 & 2 and 3 & 4 mirrored

(Part of) SAT Encoding

▶ variable x_{ijpr} is true if team *i* plays team *j* at home in period *p*, round *r*

$$\bigwedge_{i,p,r} \bigvee_{j \neq i} (x_{ijpr} \lor x_{jipr})$$
 each team plays in every round
$$\bigwedge_{i,p,r} \bigwedge_{j \neq i} \bigwedge_{k \neq i \land k \neq j} (x_{ijpr} \rightarrow \neg (x_{ikpr} \lor x_{kipr}))$$
 each team plays at most once in every round
$$\bigwedge_{i,j,r} (x_{ij1r} \rightarrow x_{ji2r}) \land (x_{ij3r} \rightarrow x_{ji4r})$$
 mirror rounds 1& 2 and 3& 4

Problem: Round Robin Scheduling

Schedule sports league tournament for n teams, p periods of n - 1 rounds each (+ venue restrictions, break restrictions, ...)

Example (Österreichische Fußball-Bundesliga)

10 teams play in 4 periods (9 rounds each), periods 1 & 2 and 3 & 4 mirrored

(Part of) SAT Encoding

▶ variable x_{ijpr} is true if team *i* plays team *j* at home in period *p*, round *r*

$$\bigwedge_{i,p,r} \bigvee_{j \neq i} (x_{ijpr} \lor x_{jipr})$$
 each team plays in every round
$$\bigwedge_{i,p,r} \bigwedge_{j \neq i} \bigwedge_{k \neq i \land k \neq j} (x_{ijpr} \rightarrow \neg (x_{ikpr} \lor x_{kipr}))$$
 each team plays at most once in every round
$$\bigwedge_{i,j,r} (x_{ij1r} \rightarrow x_{ji2r}) \land (x_{ij3r} \rightarrow x_{ji4r})$$
 mirror rounds 1& 2 and 3& 4

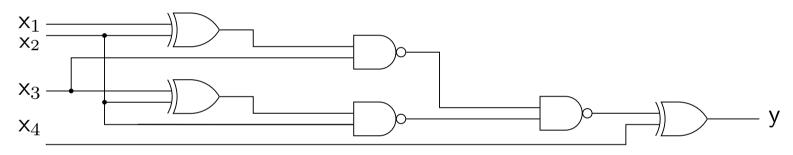
Result

SAT scheduling is 100x faster than previous industrial scheduling tools

Problem

errors in hardware chips are costly (Intel paid \$475 million for FDIV bug)

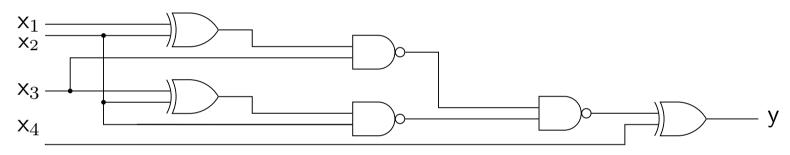
Example (Formal Circuit Model)



Problem

- errors in hardware chips are costly (Intel paid \$475 million for FDIV bug)
- testing is not enough to guarantee desired behavior

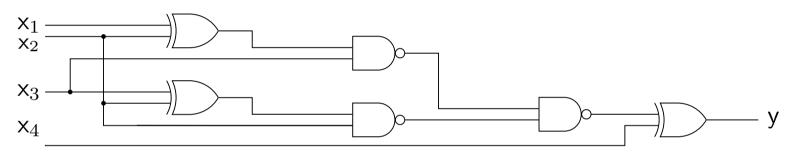
Example (Formal Circuit Model)



Problem

- ▶ errors in hardware chips are costly (Intel paid \$475 million for FDIV bug)
- testing is not enough to guarantee desired behavior

Example (Formal Circuit Model)



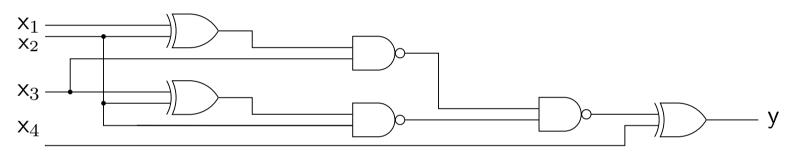
SAT Encoding

- variables for input and output
- SAT formulas for implemented behavior and expected behavior (specification)
- check for equivalence

Problem

- ▶ errors in hardware chips are costly (Intel paid \$475 million for FDIV bug)
- testing is not enough to guarantee desired behavior

Example (Formal Circuit Model)



SAT Encoding

- variables for input and output
- SAT formulas for implemented behavior and expected behavior (specification)
- check for equivalence

Impact

- ensured correctness, more reliable hardware components (formal verification)
- manufacturers rely on SAT-based verification since beginning of 2000s e.g., Intel Core i7 implements over 2700 distinct verified microinstructions ¹⁰

Concepts

- ► literal
- ► formula
- assignment
- satisfiability and validity
- negation normal form (NNF)
- conjunctive normal form (CNF)
- disjunctive normal form (DNF)

propositional formulas are built form

▶ atoms $p, q, r, p_1, p_2, ...$

propositional formulas are built form

▶ atoms $p, q, r, p_1, p_2, ...$

 \perp, \top

► constants

propositional formulas are built form

- ▶ atoms $p, q, r, p_1, p_2, ...$
- ▶ constants \bot, \top
- ▶ negation $\neg p$ "not p"

propositional formulas are built form

- ▶ atoms $p, q, r, p_1, p_2, ...$
- ▶ constants \bot, \top
- ▶ negation $\neg p$ "not p"
- ► conjunction $p \land q$ "p and q"

propositional formulas are built form

- atoms $p, q, r, p_1, p_2, \ldots$ \perp, \top constants "not *p*" negation $\neg p$
- conjunction $p \wedge q$ "*p* and *q*" "p or q"
- disjunction

 $p \lor q$

propositional formulas are built form

atoms	$p, q, r, p_1, p_2, \ldots$	
constants	\perp, \top	
negation	eg p	"not <i>p</i> "
conjunction	$oldsymbol{p}\wedgeoldsymbol{q}$	" <i>p</i> and <i>q</i> "
disjunction	$oldsymbol{p}ee oldsymbol{q}$	" <i>p</i> or <i>q</i> "
implication	p ightarrow q	"if <i>p</i> then <i>q</i> holds"

propositional formulas are built form

atoms	$p, q, r, p_1, p_2, \ldots$	
constants	\perp, \top	
negation	eg p	"not <i>p</i> "
conjunction	$oldsymbol{p}\wedgeoldsymbol{q}$	" <i>p</i> and <i>q</i> "
disjunction	$oldsymbol{p}ee oldsymbol{q}$	" <i>p</i> or <i>q</i> "
implication	ho ightarrow q	"if <i>p</i> then <i>q</i> holds"
equivalence	$p \leftrightarrow q$	" <i>p</i> if and only if <i>q</i>

propositional formulas are built form

atoms	$p, q, r, p_1, p_2, \ldots$	
constants	\perp, \top	
negation	eg p	"not <i>p</i> "
conjunction	$oldsymbol{p}\wedgeoldsymbol{q}$	" <i>p</i> and <i>q</i> "
disjunction	$p \lor q$	" <i>p</i> or <i>q</i> "
implication	ho ightarrow q	"if <i>p</i> then <i>q</i> holds"
equivalence	$p \leftrightarrow q$	" <i>p</i> if and only if <i>q</i> "

according to the BNF grammar

 $\varphi ::= p \mid \bot \mid \top \mid (\neg \varphi) \mid (\varphi \land \varphi) \mid (\varphi \lor \varphi) \mid (\varphi \to \varphi) \mid (\varphi \leftrightarrow \varphi)$

propositional formulas are built form

atoms	$p, q, r, p_1, p_2, \ldots$	
constants	\perp, \top	
negation	$\neg p$	"not <i>p</i> "
conjunction	$p \wedge q$	" <i>p</i> and <i>q</i> "
disjunction	$p \lor q$	" <i>p</i> or <i>q</i> "
implication	p ightarrow q	"if <i>p</i> then <i>q</i> holds"
equivalence	$p \leftrightarrow q$	" <i>p</i> if and only if <i>q</i> "

according to the BNF grammar

 $\varphi ::= p \mid \bot \mid \top \mid (\neg \varphi) \mid (\varphi \land \varphi) \mid (\varphi \lor \varphi) \mid (\varphi \to \varphi) \mid (\varphi \leftrightarrow \varphi)$

Conventions

▶ binding precedence \neg > \land > \lor →, ↔

propositional formulas are built form

atoms	$p, q, r, p_1, p_2, \ldots$	
constants	\perp, \top	
negation	$\neg p$	"not <i>p</i> "
conjunction	$p \wedge q$	" <i>p</i> and <i>q</i> "
disjunction	$p \lor q$	" <i>p</i> or <i>q</i> "
implication	p ightarrow q	"if <i>p</i> then <i>q</i> holds"
equivalence	$p \leftrightarrow q$	" <i>p</i> if and only if <i>q</i> "

according to the BNF grammar

 $\varphi ::= p \mid \bot \mid \top \mid (\neg \varphi) \mid (\varphi \land \varphi) \mid (\varphi \lor \varphi) \mid (\varphi \to \varphi) \mid (\varphi \leftrightarrow \varphi)$

Conventions

- ▶ binding precedence \neg > \land > \lor →, ↔
- omit outer parantheses

propositional formulas are built form

atoms	$p, q, r, p_1, p_2, \ldots$	
constants	\perp, \top	
negation	$\neg p$	"not <i>p</i> "
conjunction	$p \wedge q$	" <i>p</i> and <i>q</i> "
disjunction	$p \lor q$	" <i>p</i> or <i>q</i> "
implication	ho ightarrow q	"if <i>p</i> then <i>q</i> holds"
equivalence	$p \leftrightarrow q$	" <i>p</i> if and only if <i>q</i> "

according to the BNF grammar

 $\varphi ::= p \mid \bot \mid \top \mid (\neg \varphi) \mid (\varphi \land \varphi) \mid (\varphi \lor \varphi) \mid (\varphi \to \varphi) \mid (\varphi \leftrightarrow \varphi)$

Conventions

- ▶ binding precedence \neg > \land > \lor \rightarrow , \leftrightarrow
- omit outer parantheses
- ▶ \rightarrow , \land , \lor are right-associative: $p \rightarrow q \rightarrow r$ denotes $p \rightarrow (q \rightarrow r)$

▶ valuation (truth assignment) is mapping v : {p, q, r, ...} → {F, T}
 from atoms to truth values

- ▶ valuation (truth assignment) is mapping v : {p, q, r, ...} → {F, T}
 from atoms to truth values
- ► extension to formulas:

$$v(\perp) = F$$

- ▶ valuation (truth assignment) is mapping v : {p, q, r, ...} → {F, T}
 from atoms to truth values
- ► extension to formulas:

$$v(\perp) = \mathsf{F}$$
 $v(\top) = \mathsf{T}$

▶ valuation (truth assignment) is mapping v : {p, q, r, ...} → {F, T}
 from atoms to truth values

 $v(\top) = \mathsf{T}$

► extension to formulas:

$$m{v}(\perp) = \mathsf{F}$$

 $m{v}(arphi \wedge \psi) = egin{cases} \mathsf{T} & ext{if } m{v}(arphi) = m{v}(\psi) = \mathsf{T} \ \mathsf{F} & ext{otherwise} \end{cases}$

- valuation (truth assignment) is mapping v : {p, q, r, ...} → {F, T}
 from atoms to truth values
- ► extension to formulas:

$$egin{aligned} & v(ot) = \mathsf{F} \ & v(arphi \wedge \psi) = \left\{ egin{aligned} \mathsf{T} & ext{if } v(arphi) = v(\psi) = \mathsf{T} \ & \mathsf{F} & ext{otherwise} \end{aligned}
ight. \end{aligned}$$

 $v(\top) = \mathsf{T}$ $v(\neg \varphi) = \begin{cases} \mathsf{T} & \text{if } v(\varphi) = \mathsf{F} \\ \mathsf{F} & \text{if } v(\varphi) = \mathsf{T} \end{cases}$

- ▶ valuation (truth assignment) is mapping v : {p, q, r, ...} → {F, T}
 from atoms to truth values
- ► extension to formulas:

$$v(\perp) = \mathsf{F} \qquad v(\top) = \mathsf{T}$$

$$v(\varphi \land \psi) = \begin{cases} \mathsf{T} & \text{if } v(\varphi) = v(\psi) = \mathsf{T} \\ \mathsf{F} & \text{otherwise} \end{cases} \qquad v(\neg \varphi) = \begin{cases} \mathsf{T} & \text{if } v(\varphi) = \mathsf{F} \\ \mathsf{F} & \text{if } v(\varphi) = \mathsf{T} \end{cases}$$

$$v(\varphi \lor \psi) = \begin{cases} \mathsf{F} & \text{if } v(\varphi) = v(\psi) = \mathsf{F} \\ \mathsf{T} & \text{otherwise} \end{cases}$$

- ▶ valuation (truth assignment) is mapping v : {p, q, r, ...} → {F, T}
 from atoms to truth values
- ► extension to formulas:

$$v(\perp) = \mathsf{F} \qquad v(\top) = \mathsf{T}$$

$$v(\varphi \land \psi) = \begin{cases} \mathsf{T} & \text{if } v(\varphi) = v(\psi) = \mathsf{T} \\ \mathsf{F} & \text{otherwise} \end{cases} \qquad v(\neg \varphi) = \begin{cases} \mathsf{T} & \text{if } v(\varphi) = \mathsf{F} \\ \mathsf{F} & \text{if } v(\varphi) = \mathsf{T} \end{cases}$$

$$v(\varphi \lor \psi) = \begin{cases} \mathsf{F} & \text{if } v(\varphi) = v(\psi) = \mathsf{F} \\ \mathsf{T} & \text{otherwise} \end{cases} \qquad v(\varphi \leftrightarrow \psi) = \begin{cases} \mathsf{T} & \text{if } v(\varphi) = v(\psi) \\ \mathsf{F} & \text{otherwise} \end{cases}$$

- ▶ valuation (truth assignment) is mapping v : {p, q, r, ...} → {F, T}
 from atoms to truth values
- ► extension to formulas:

$$v(\perp) = \mathsf{F} \qquad v(\top) = \mathsf{T}$$

$$v(\varphi \land \psi) = \begin{cases} \mathsf{T} & \text{if } v(\varphi) = v(\psi) = \mathsf{T} \\ \mathsf{F} & \text{otherwise} \end{cases} \qquad v(\neg\varphi) = \begin{cases} \mathsf{T} & \text{if } v(\varphi) = \mathsf{F} \\ \mathsf{F} & \text{if } v(\varphi) = \mathsf{T} \end{cases}$$

$$v(\varphi \lor \psi) = \begin{cases} \mathsf{F} & \text{if } v(\varphi) = v(\psi) = \mathsf{F} \\ \mathsf{T} & \text{otherwise} \end{cases} \qquad v(\varphi \leftrightarrow \psi) = \begin{cases} \mathsf{T} & \text{if } v(\varphi) = v(\psi) \\ \mathsf{F} & \text{otherwise} \end{cases}$$

$$v(\varphi \rightarrow \psi) = \begin{cases} \mathsf{F} & \text{if } v(\varphi) = \mathsf{T}, \ v(\psi) = \mathsf{F} \\ \mathsf{T} & \text{otherwise} \end{cases}$$

▶ formula φ is satisfiable if $v(\varphi) = T$ for some valuation v

- formula φ is satisfiable if $v(\varphi) = T$ for some valuation v
- ▶ formula φ is valid if $v(\varphi) = T$ for every valuation v

- formula φ is satisfiable if $v(\varphi) = T$ for some valuation v
- ▶ formula φ is valid if $v(\varphi) = T$ for every valuation v
- semantic entailment $\varphi_1, \ldots, \varphi_n \vDash \psi$

if $v(\psi) = \mathsf{T}$ whenever $v(\varphi_1) = v(\varphi_2) = \cdots = v(\varphi_n) = \mathsf{T}$

- formula φ is satisfiable if $v(\varphi) = T$ for some valuation v
- ▶ formula φ is valid if $v(\varphi) = T$ for every valuation v
- ► semantic entailment $\varphi_1, \ldots, \varphi_n \vDash \psi$ if $v(\psi) = \mathsf{T}$ whenever $v(\varphi_1) = v(\varphi_2) = \cdots = v(\varphi_n) = \mathsf{T}$
- ▶ formulas φ and ψ are equivalent ($\varphi \equiv \psi$) if $v(\varphi) = v(\psi)$ for every valuation v

- formula φ is satisfiable if $v(\varphi) = T$ for some valuation v
- ▶ formula φ is valid if $v(\varphi) = T$ for every valuation v
- ► semantic entailment $\varphi_1, \ldots, \varphi_n \vDash \psi$ if $v(\psi) = \mathsf{T}$ whenever $v(\varphi_1) = v(\varphi_2) = \cdots = v(\varphi_n) = \mathsf{T}$
- ▶ formulas φ and ψ are equivalent ($\varphi \equiv \psi$) if $v(\varphi) = v(\psi)$ for every valuation v
- ▶ formulas φ and ψ are equisatisfiable ($\varphi \approx \psi$) if

 φ is satisfiable $\iff \psi$ is satisfiable

- formula φ is satisfiable if $v(\varphi) = T$ for some valuation v
- ▶ formula φ is valid if $v(\varphi) = T$ for every valuation v
- ► semantic entailment $\varphi_1, \ldots, \varphi_n \vDash \psi$ if $v(\psi) = \mathsf{T}$ whenever $v(\varphi_1) = v(\varphi_2) = \cdots = v(\varphi_n) = \mathsf{T}$
- ▶ formulas φ and ψ are equivalent ($\varphi \equiv \psi$) if $v(\varphi) = v(\psi)$ for every valuation v
- formulas φ and ψ are equisatisfiable ($\varphi \approx \psi$) if

$$\varphi$$
 is satisfiable $\iff \psi$ is satisfiable

Theorem

formula φ is unsatisfiable if and only if $\neg\varphi$ is valid

- formula φ is satisfiable if $v(\varphi) = T$ for some valuation v
- ▶ formula φ is valid if $v(\varphi) = T$ for every valuation v
- ► semantic entailment $\varphi_1, \ldots, \varphi_n \vDash \psi$ if $v(\psi) = \mathsf{T}$ whenever $v(\varphi_1) = v(\varphi_2) = \cdots = v(\varphi_n) = \mathsf{T}$
- ▶ formulas φ and ψ are equivalent ($\varphi \equiv \psi$) if $v(\varphi) = v(\psi)$ for every valuation v
- formulas φ and ψ are equisatisfiable ($\varphi \approx \psi$) if

$$\varphi$$
 is satisfiable $\iff \psi$ is satisfiable

Theorem

formula φ is unsatisfiable if and only if $\neg\varphi$ is valid

Theorem

satisfiability and validity are decidable

Definition (Literal)

▶ literal is atom p or negation of atom $\neg p$

Definition (Literal)

- ▶ literal is atom p or negation of atom $\neg p$
- ▶ literals l_1 and l_2 are complementary if $l_1 = \neg l_2$ or $l_2 = \neg l_1$

Definition (Literal)

- ▶ literal is atom p or negation of atom $\neg p$
- ▶ literals l_1 and l_2 are complementary if $l_1 = \neg l_2$ or $l_2 = \neg l_1$

Definitions

negation normal form (NNF) if formula with negation only applied to atoms

- ▶ literal is atom p or negation of atom $\neg p$
- ▶ literals l_1 and l_2 are complementary if $l_1 = \neg l_2$ or $l_2 = \neg l_1$

Definitions

- negation normal form (NNF) if formula with negation only applied to atoms
- conjunctive normal form (CNF) is conjunction of disjunctions

- ▶ literal is atom p or negation of atom $\neg p$
- ▶ literals l_1 and l_2 are complementary if $l_1 = \neg l_2$ or $l_2 = \neg l_1$

Definitions

- negation normal form (NNF) if formula with negation only applied to atoms
- conjunctive normal form (CNF) is conjunction of disjunctions
- ▶ 3-CNF is conjunction of disjunctions with 3 literals: $\bigwedge_i (a_i \lor b_i \lor c_i)$

- ▶ literal is atom p or negation of atom $\neg p$
- ▶ literals l_1 and l_2 are complementary if $l_1 = \neg l_2$ or $l_2 = \neg l_1$

Definitions

- negation normal form (NNF) if formula with negation only applied to atoms
- conjunctive normal form (CNF) is conjunction of disjunctions
- ▶ 3-CNF is conjunction of disjunctions with 3 literals: $\bigwedge_i (a_i \lor b_i \lor c_i)$
- disjunctive normal form (DNF) is disjunction of conjunctions

- ▶ literal is atom p or negation of atom $\neg p$
- ▶ literals l_1 and l_2 are complementary if $l_1 = \neg l_2$ or $l_2 = \neg l_1$

Definitions

- negation normal form (NNF) if formula with negation only applied to atoms
- conjunctive normal form (CNF) is conjunction of disjunctions
- ▶ 3-CNF is conjunction of disjunctions with 3 literals: $\bigwedge_i (a_i \lor b_i \lor c_i)$
- disjunctive normal form (DNF) is disjunction of conjunctions

Theorem

for every formula φ there is CNF ψ , 3-CNF χ and DNF η such that $\varphi \equiv \psi \equiv \chi \equiv \eta$

- ▶ literal is atom p or negation of atom $\neg p$
- ▶ literals l_1 and l_2 are complementary if $l_1 = \neg l_2$ or $l_2 = \neg l_1$

Definitions

- negation normal form (NNF) if formula with negation only applied to atoms
- conjunctive normal form (CNF) is conjunction of disjunctions
- ▶ 3-CNF is conjunction of disjunctions with 3 literals: $\bigwedge_i (a_i \lor b_i \lor c_i)$
- disjunctive normal form (DNF) is disjunction of conjunctions

Theorem

for every formula φ there is CNF ψ , 3-CNF χ and DNF η such that $\varphi \equiv \psi \equiv \chi \equiv \eta$

Remarks

translation from formula to CNF can result in exponential blowup

- ▶ literal is atom p or negation of atom $\neg p$
- ▶ literals l_1 and l_2 are complementary if $l_1 = \neg l_2$ or $l_2 = \neg l_1$

Definitions

- negation normal form (NNF) if formula with negation only applied to atoms
- conjunctive normal form (CNF) is conjunction of disjunctions
- ▶ 3-CNF is conjunction of disjunctions with 3 literals: $\bigwedge_i (a_i \lor b_i \lor c_i)$
- disjunctive normal form (DNF) is disjunction of conjunctions

Theorem

for every formula φ there is CNF ψ , 3-CNF χ and DNF η such that $\varphi \equiv \psi \equiv \chi \equiv \eta$

Remarks

- translation from formula to CNF can result in exponential blowup
- Tseitin's transformation is linear and produces equisatisfiable formula

instance: propositional formula φ question: is φ satisfiable?

instance: propositional formula φ question: is φ satisfiable?

3-Satisfiability (3-SAT)

instance: propositional formula φ question: is φ satisfiable?

3-Satisfiability (3-SAT)

instance: propositional formula φ in 3-CNF question: is φ satisfiable?

Theorem

SAT and 3-SAT are NP-complete problems

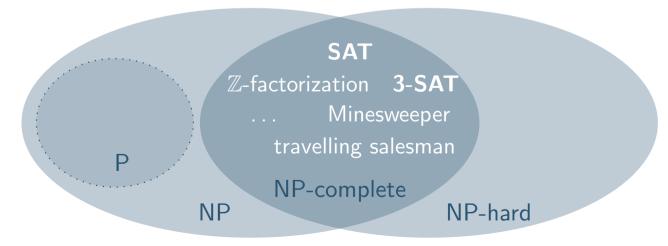
instance: propositional formula φ question: is φ satisfiable?

3-Satisfiability (3-SAT)

instance: propositional formula φ in 3-CNF question: is φ satisfiable?

Theorem

SAT and 3-SAT are NP-complete problems



▶ 1 million \$ prize money awarded for solution to $\mathbf{P} = {}^{?} \mathbf{NP}$

most SAT solvers require input to be in CNF

most SAT solvers require input to be in CNF

Remarks

- transforming formula to equivalent CNF can cause exponential blowup
- transforming formula into equisatisfiable CNF is possible in linear time

most SAT solvers require input to be in CNF

Remarks

- transforming formula to equivalent CNF can cause exponential blowup
- transforming formula into equisatisfiable CNF is possible in linear time

Definition

formulas φ and ψ are equisatisfiable ($\varphi\approx\psi)$ if

 φ is satisfiable $\iff \psi$ is satisfiable

most SAT solvers require input to be in CNF

Remarks

- transforming formula to equivalent CNF can cause exponential blowup
- transforming formula into equisatisfiable CNF is possible in linear time

Definition

formulas φ and ψ are equisatisfiable ($\varphi \approx \psi$) if

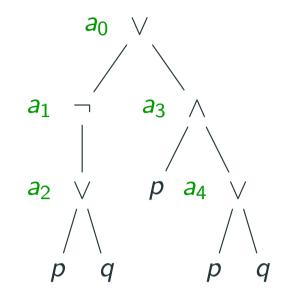
 φ is satisfiable $\iff \psi$ is satisfiable

Example

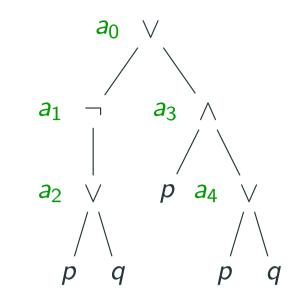
 $p \lor q \approx \top$ $p \land \neg p \approx q \land \neg q$ $p \land \neg p \not\approx p \land \neg q$

 $\blacktriangleright \quad \varphi = \neg (p \lor q) \lor (p \land (p \lor q))$

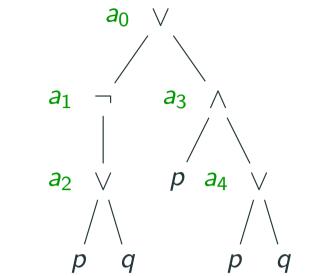
 $\blacktriangleright \quad \varphi = \neg (p \lor q) \lor (p \land (p \lor q))$



- $\blacktriangleright \quad \varphi = \neg (p \lor q) \lor (p \land (p \lor q))$
- ► use fresh propositional variable for every connective $a_0: \neg(p \lor q) \lor (p \land (p \lor q)) \quad a_1: \neg(p \lor q)$ $a_2: p \lor q \quad a_3: p \land (p \lor q)$ $a_4: p \lor q$



- $\blacktriangleright \quad \varphi = \neg (p \lor q) \lor (p \land (p \lor q))$
- use fresh propositional variable for every connective $a_0: \neg (p \lor q) \lor (p \land (p \lor q)) \quad a_1: \neg (p \lor q)$ $a_2: p \lor q \quad a_3: p \land (p \lor q)$ $a_4: p \lor q$



 $\begin{array}{ll} \bullet & \varphi \approx & a_0 \wedge (a_0 \leftrightarrow a_1 \vee a_3) \wedge (a_1 \leftrightarrow \neg a_2) \wedge (a_2 \leftrightarrow p \vee q) \wedge \\ & (a_3 \leftrightarrow p \wedge a_4) \wedge (a_4 \leftrightarrow p \vee q) \end{array}$

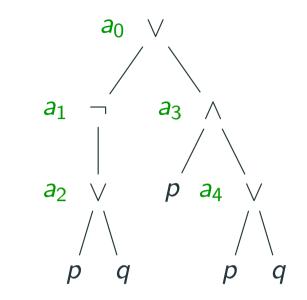
- $\blacktriangleright \quad \varphi = \neg (p \lor q) \lor (p \land (p \lor q))$
- use fresh propositional variable for every connective $a_0: \neg (p \lor q) \lor (p \land (p \lor q)) \quad a_1: \neg (p \lor q)$ $a_2: p \lor q \quad a_3: p \land (p \lor q)$ $a_4: p \lor q$

$$\begin{array}{c}
a_{0} \lor \\
& \swarrow \\
a_{1} \neg & a_{3} \land \\
& \downarrow & \swarrow \\
a_{2} \lor & p a_{4} \lor \\
& \swarrow \\
& p q & p q
\end{array}$$

• every \leftrightarrow subexpression can be replaced by at most three clauses:

$$\begin{array}{ll} a \leftrightarrow b \wedge c & \equiv & (\neg a \lor b) \wedge (\neg a \lor c) \wedge (a \lor \neg b \lor \neg c) \\ a \leftrightarrow b \lor c & \equiv & (\neg a \lor b \lor c) \wedge (a \lor \neg b) \wedge (a \lor \neg c) \\ a \leftrightarrow \neg b & \equiv & (\neg a \lor \neg b) \wedge (a \lor b) \end{array}$$

- $\blacktriangleright \quad \varphi = \neg (p \lor q) \lor (p \land (p \lor q))$
- use fresh propositional variable for every connective $a_0: \neg (p \lor q) \lor (p \land (p \lor q)) \quad a_1: \neg (p \lor q)$ $a_2: p \lor q \quad a_3: p \land (p \lor q)$ $a_4: p \lor q$



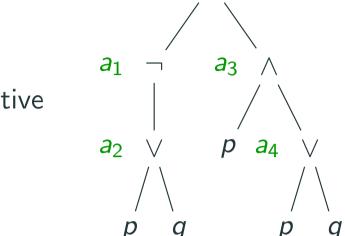
$$\varphi \approx a_0 \wedge (a_0 \leftrightarrow a_1 \vee a_3) \wedge (a_1 \leftrightarrow \neg a_2) \wedge (a_2 \leftrightarrow p \vee q) \wedge (a_3 \leftrightarrow p \wedge a_4) \wedge (a_4 \leftrightarrow p \vee q)$$

• every \leftrightarrow subexpression can be replaced by at most three clauses:

$$\begin{array}{ll} a \leftrightarrow b \wedge c & \equiv & (\neg a \lor b) \land (\neg a \lor c) \land (a \lor \neg b \lor \neg c) \\ a \leftrightarrow b \lor c & \equiv & (\neg a \lor b \lor c) \land (a \lor \neg b) \land (a \lor \neg c) \\ a \leftrightarrow \neg b & \equiv & (\neg a \lor \neg b) \land (a \lor b) \end{array}$$

common subexpressions can be shared

- $\blacktriangleright \quad \varphi = \neg (p \lor q) \lor (p \land (p \lor q))$
- use fresh propositional variable for every connective $a_0: \neg(p \lor q) \lor (p \land (p \lor q)) \quad a_1: \neg(p \lor q)$ $a_2: p \lor q \quad a_3: p \land (p \lor q)$ $a_4: p \lor q$



 $a_0 \vee$

$$\varphi \approx a_0 \wedge (a_0 \leftrightarrow a_1 \vee a_3) \wedge (a_1 \leftrightarrow \neg a_2) \wedge (a_2 \leftrightarrow p \vee q) \wedge (a_3 \leftrightarrow p \wedge a_2)$$

- every \leftrightarrow subexpression can be replaced by at most three clauses: $a \leftrightarrow b \wedge c \equiv (\neg a \lor b) \land (\neg a \lor c) \land (a \lor \neg b \lor \neg c)$ $a \leftrightarrow b \lor c \equiv (\neg a \lor b \lor c) \land (a \lor \neg b) \land (a \lor \neg c)$ $a \leftrightarrow \neg b \equiv (\neg a \lor \neg b) \land (a \lor \neg b)$
- common subexpressions can be shared

Minisat, Glucose, CaDiCaL, Glu_VC, Plingeling, MapleLRB LCM, MapleCOMPSPS, Riss, Lingeling, Treengeling, CryptoMiniSat, abcdSAT, Dimetheus, Kiel, MapleCOMSPS, Rsat, SWDiA5BY, BlackBox, SWDiA5BY, pprobSAT, glueSplit_clasp, BalancedZ, SApperloT, PeneLoPe, MXC, ROKKminisat, MiniSat_HACK_999ED, ZENN, CSHCrandMC, MiniGolf, march_rw, sattime2011, mphasesat64, sparrow2011, pmcSAT, CSHCpar8, gluebit_clasp, clasp, precosat, gNovelty, SATzilla, SatELite, Score2SAT, YalSAT, tch glucose3, ...

Minisat, Glucose, CaDiCaL, Glu_VC, Plingeling, MapleLRB LCM, MapleCOMPSPS, Riss, Lingeling, Treengeling, CryptoMiniSat, abcdSAT, Dimetheus, Kiel, MapleCOMSPS, Rsat, SWDiA5BY, BlackBox, SWDiA5BY, pprobSAT, glueSplit_clasp, BalancedZ, SApperloT, PeneLoPe, MXC, ROKKminisat, MiniSat_HACK_999ED, ZENN, CSHCrandMC, MiniGolf, march_rw, sattime2011, mphasesat64, sparrow2011, pmcSAT, CSHCpar8, gluebit_clasp, clasp, precosat, gNovelty, SATzilla, SatELite, Score2SAT, YalSAT, tch glucose3, ...

SAT Competition

- ▶ annual competition for different tracks (main, parallel, no-limit, ...)
- increasing set of benchmarks from industry, mathematics, cryptography, ...
- standardized input format DIMACS and proof format DRAT

Minisat, Glucose, CaDiCaL, Glu_VC, Plingeling, MapleLRB LCM, MapleCOMPSPS, Riss, Lingeling, Treengeling, CryptoMiniSat, abcdSAT, Dimetheus, Kiel, MapleCOMSPS, Rsat, SWDiA5BY, BlackBox, SWDiA5BY, pprobSAT, glueSplit_clasp, BalancedZ, SApperloT, PeneLoPe, MXC, ROKKminisat, MiniSat_HACK_999ED, ZENN, CSHCrandMC, MiniGolf, march_rw, sattime2011, mphasesat64, sparrow2011, pmcSAT, CSHCpar8, gluebit_clasp, clasp, precosat, gNovelty, SATzilla, SatELite, Score2SAT, YalSAT, tch glucose3, ...

SAT Competition

- ▶ annual competition for different tracks (main, parallel, no-limit, ...)
- increasing set of benchmarks from industry, mathematics, cryptography, ...
- standardized input format DIMACS and proof format DRAT

http://www.satcompetition.org/

Minisat, Glucose, CaDiCaL, Glu_VC, Plingeling, MapleLRB LCM, MapleCOMPSPS, Riss, Lingeling, Treengeling, CryptoMiniSat, abcdSAT, Dimetheus, Kiel, MapleCOMSPS, Rsat, SWDiA5BY, BlackBox, SWDiA5BY, pprobSAT, glueSplit_clasp, BalancedZ, SApperloT, PeneLoPe, MXC, ROKKminisat, MiniSat_HACK_999ED, ZENN, CSHCrandMC, MiniGolf, march_rw, sattime2011, mphasesat64, sparrow2011, pmcSAT, CSHCpar8, gluebit_clasp, clasp, precosat, gNovelty, SATzilla, SatELite, Score2SAT, YalSAT, tch glucose3, ...

SAT Competition

- ▶ annual competition for different tracks (main, parallel, no-limit, ...)
- increasing set of benchmarks from industry, mathematics, cryptography, ...
- standardized input format DIMACS and proof format DRAT

```
http://www.satcompetition.org/
```

Minisat

minimalistic open source solver (http://minisat.se/ or apt, yum,...)

\$ minisat test.sat result.txt

► web interface

formula $(x_1 \vee \neg x_3) \land (x_2 \vee x_3 \vee \neg x_1) \land (\neg x_1 \vee x_2 \vee x_4)$ can be expressed by

c a very simple example p cnf 4 3 1 -3 0 2 3 -1 0 -1 2 4 0

formula $(x_1 \vee \neg x_3) \land (x_2 \vee x_3 \vee \neg x_1) \land (\neg x_1 \vee x_2 \vee x_4)$ can be expressed by

c a very simple example p cnf 4 3 1 -3 0 2 3 -1 0 -1 2 4 0

The **DIMACS** Format

• header p cnf n m specifies number of variables n and number of clauses m

formula $(x_1 \lor \neg x_3) \land (x_2 \lor x_3 \lor \neg x_1) \land (\neg x_1 \lor x_2 \lor x_4)$ can be expressed by

c a very simple example p cnf 4 3 1 -3 0 2 3 -1 0 -1 2 4 0

- header p cnf n m specifies number of variables n and number of clauses m
- ▶ variables (atoms) are assumed to be x_1, \ldots, x_n

formula $(x_1 \lor \neg x_3) \land (x_2 \lor x_3 \lor \neg x_1) \land (\neg x_1 \lor x_2 \lor x_4)$ can be expressed by

c a very simple example p cnf 4 3 1 -3 0 2 3 -1 0 -1 2 4 0

- header p cnf n m specifies number of variables n and number of clauses m
- ▶ variables (atoms) are assumed to be x_1, \ldots, x_n
- ▶ literal x_i is denoted i and literal $\neg x_i$ is denoted -i

formula $(x_1 \lor \neg x_3) \land (x_2 \lor x_3 \lor \neg x_1) \land (\neg x_1 \lor x_2 \lor x_4)$ can be expressed by

c a very simple example p cnf 4 3 1 -3 0 2 3 -1 0 -1 2 4 0

- header p cnf n m specifies number of variables n and number of clauses m
- ▶ variables (atoms) are assumed to be x_1, \ldots, x_n
- ▶ literal x_i is denoted i and literal $\neg x_i$ is denoted -i
- ► a clause is a list of literals terminated by 0

formula $(x_1 \lor \neg x_3) \land (x_2 \lor x_3 \lor \neg x_1) \land (\neg x_1 \lor x_2 \lor x_4)$ can be expressed by

c a very simple example p cnf 4 3 1 -3 0 2 3 -1 0 -1 2 4 0

- header p cnf n m specifies number of variables n and number of clauses m
- ▶ variables (atoms) are assumed to be x_1, \ldots, x_n
- ▶ literal x_i is denoted i and literal $\neg x_i$ is denoted -i
- ► a clause is a list of literals terminated by 0
- lines starting with c are considered comments