SMTCoq: a modular integration of SAT/SMT solvers to Coq

Valentin Blot

IRIF - Université Paris-Diderot

LRI - Université Paris-Sud

Chantal Keller
(project leader)

LRI - Université Paris-Sud

Two approaches

Certified ATP:

- prove correctness of the ATP's code

\+ once and for all

\+ completeness possible

\- not flexible nor modular

\- hard

Certifying ATP:

- the ATP gives certificates that can be checked

\- certificates to check each time (but efficient)

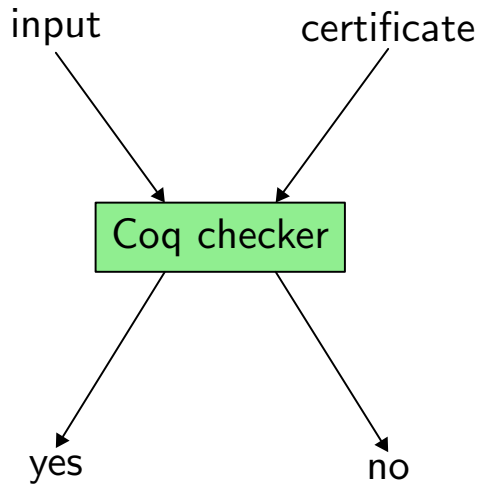
\- no completeness (or at the meta-level)

\+ very flexible and modular

\+ easier (certified checker)

Outline

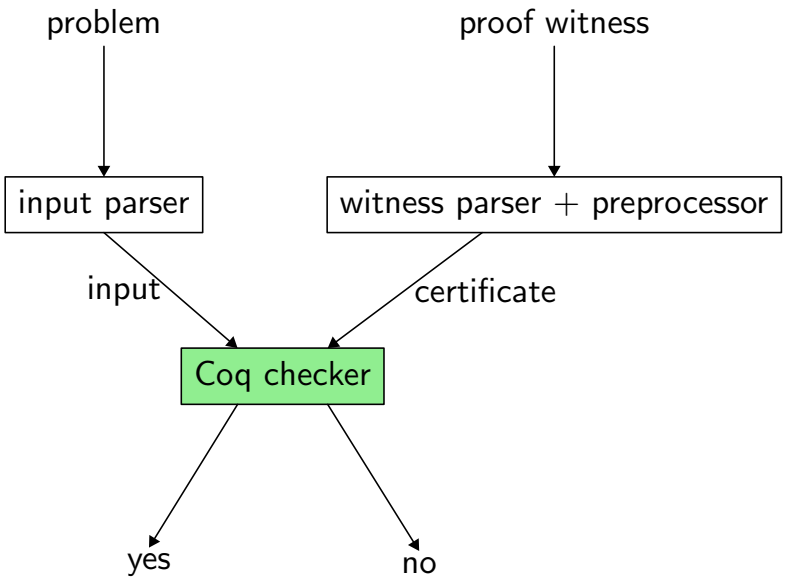
The heart: a certified checker for unsatisfiability

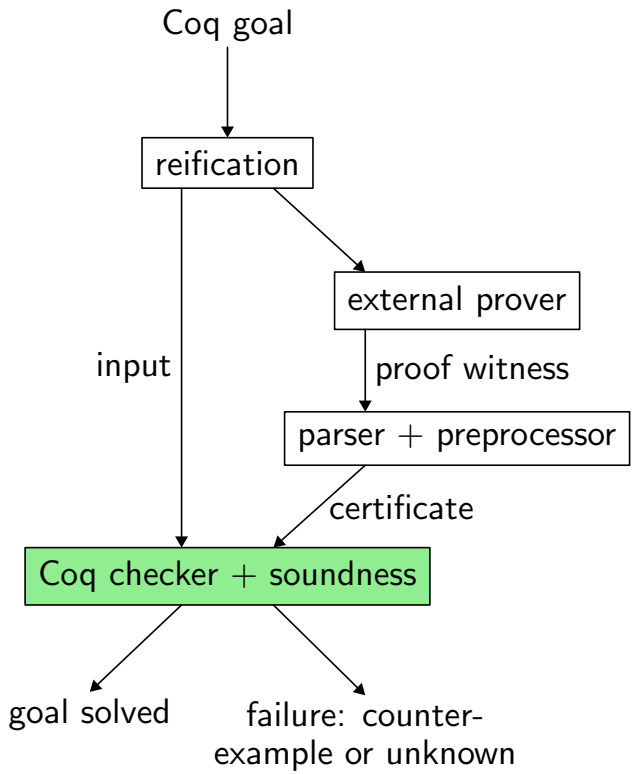


Certification:

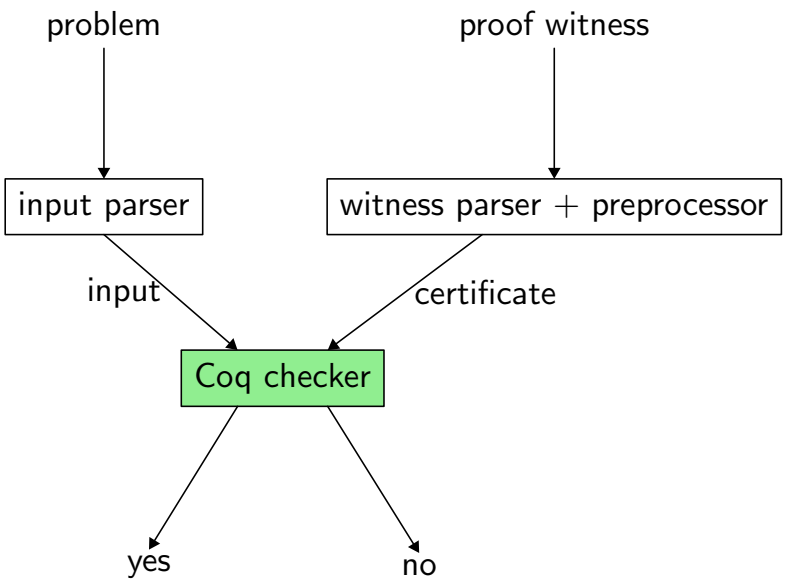
- `checker` : $\texttt{formula} \to \texttt{certif} \to \texttt{bool}$
- `correctness` :
  $$\forall \phi\, \texttt{c}, \texttt{checker}\ \phi\ \texttt{c} = \texttt{true} \to \forall \rho, |\phi|_\rho = \texttt{false}$$
- $|\bullet|_\rho : \texttt{formula} \to \texttt{bool}$ is an interpretation function
- can be extracted to ML

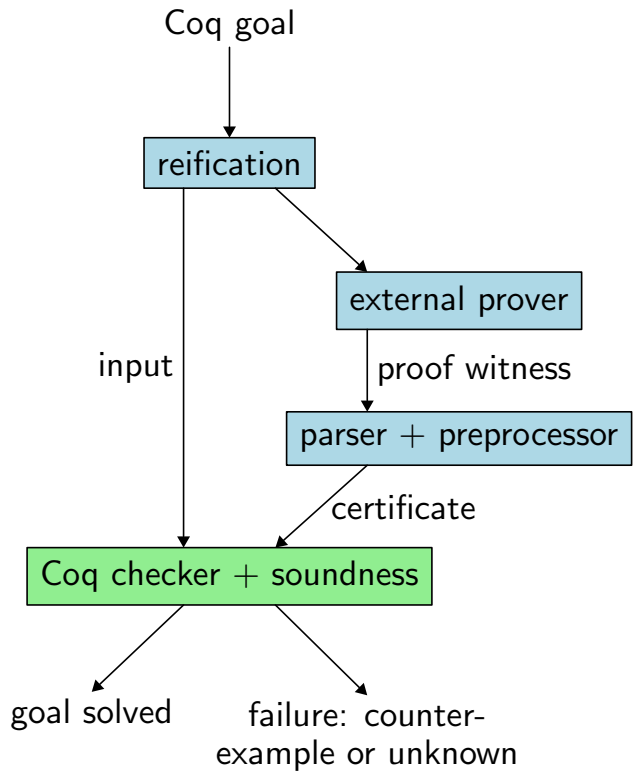
(At least) two usages



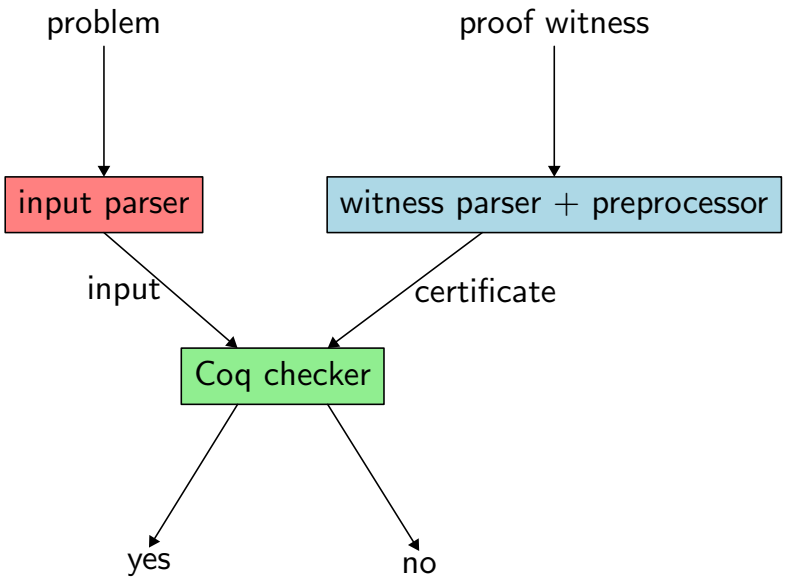


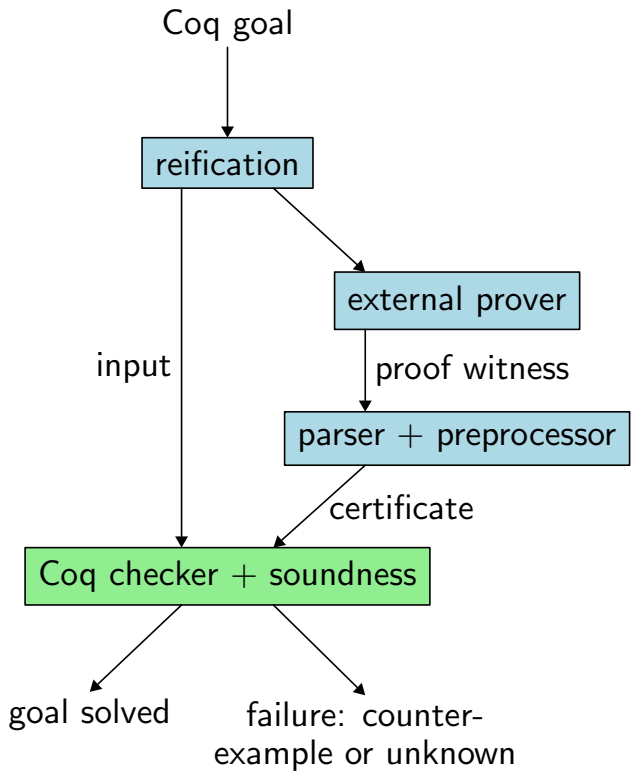
(At least) two usages





(At least) two usages





SMTCoq input and certificate formats

Input:

- a first-order formula ϕ in a combination of theories (SMT-LIB2)

Certificate:

- a resolution proof of the unsatisfiability of ϕ (seen as a set of clauses) with "theory lemmas"

Resolution and theory lemmas

The resolution rule:

$$\frac{x \vee C \qquad \neg x \vee D}{C \vee D}\ \textsc{Reso}$$

Theory lemmas:

$$\text{CNF}\ \frac{\chi \bigwedge \psi}{\chi} \qquad\qquad \frac{-}{\phi}\ \text{LIA}$$

if ϕ is a valid conjunction of atoms in LIA

Example

Unsatisfiability of (the conjunction of): $x \geq 7 \bigwedge y \leq -4 \quad \neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{\square}\,\textsc{Reso}$$

Example

Unsatisfiability of (the conjunction of): $x \geq 7 \bigwedge y \leq -4 \quad \neg x \geq 2$

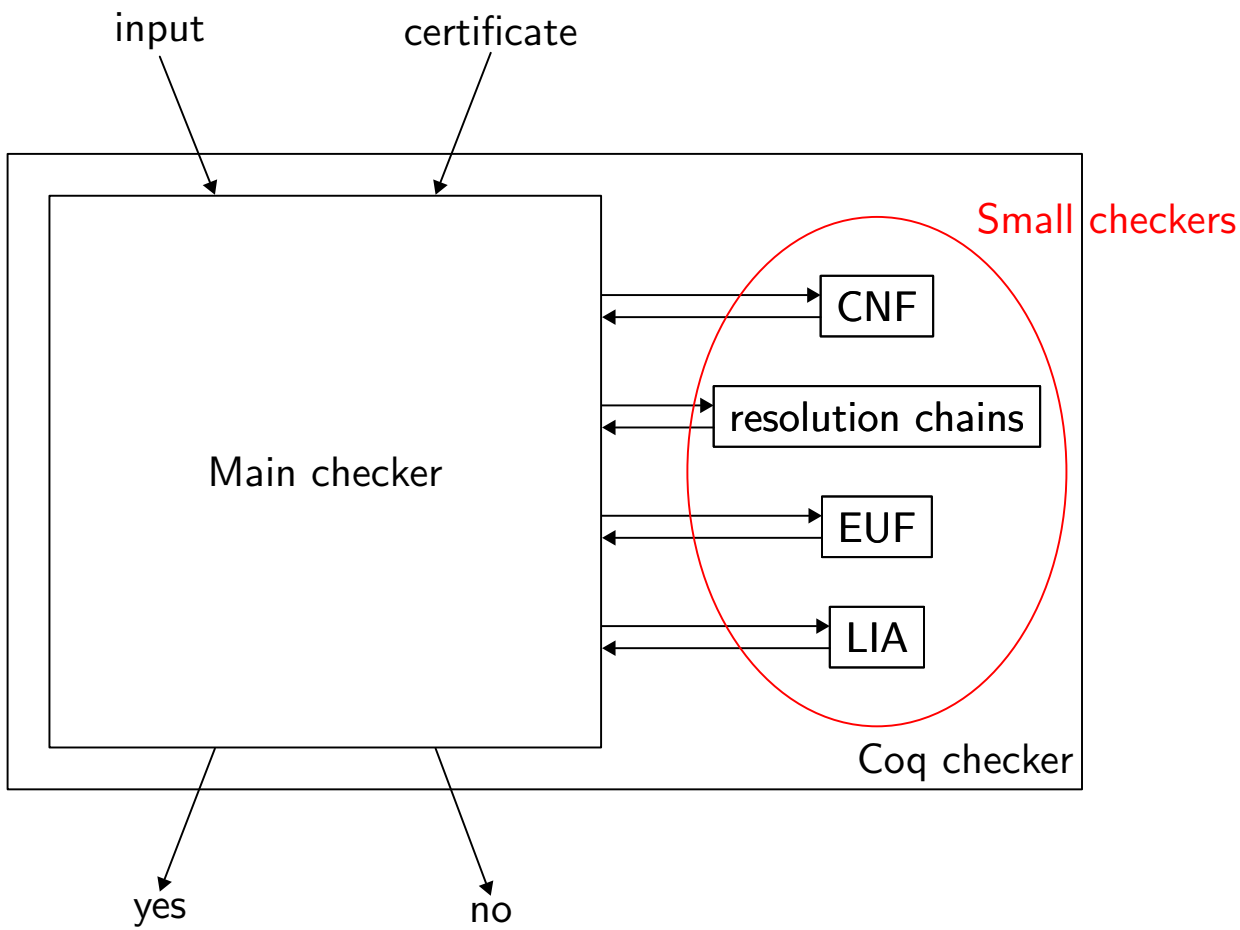
$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\square}\,\text{Reso}$$

Concrete syntax:

1:	INPUT	$x \geq 7 \bigwedge y \leq -4$
2:	INPUT	$\neg x \geq 2$
3:	LIA	$\neg x \geq 7 \vee x \geq 2$
4:	RESO	$[2; 3]$
5:	CNF_PROJ	1
6:	RESO	$[4; 5]$

Outline

A modular checker



The small checkers and the main checker

A small checker:

- takes some clauses and a piece of certificate as arguments
- returns a clause that is implied

The main checker:

- maintains a set of clauses, initialized with the input
- sequentially shares out the certificate steps to the corresponding small checkers
- checks that the last obtained clause is the empty clause

The correctness of each small checkers

implies the correctness of the whole checker

The main checker by example

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

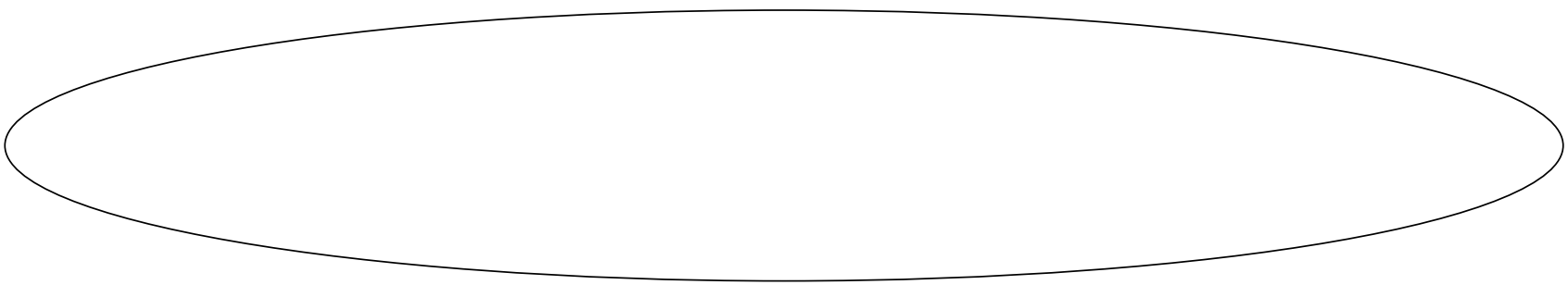
$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{\square}\,\textsc{Reso}$$

The main checker by example

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\square}\,\text{Reso}$$

A set of clauses:



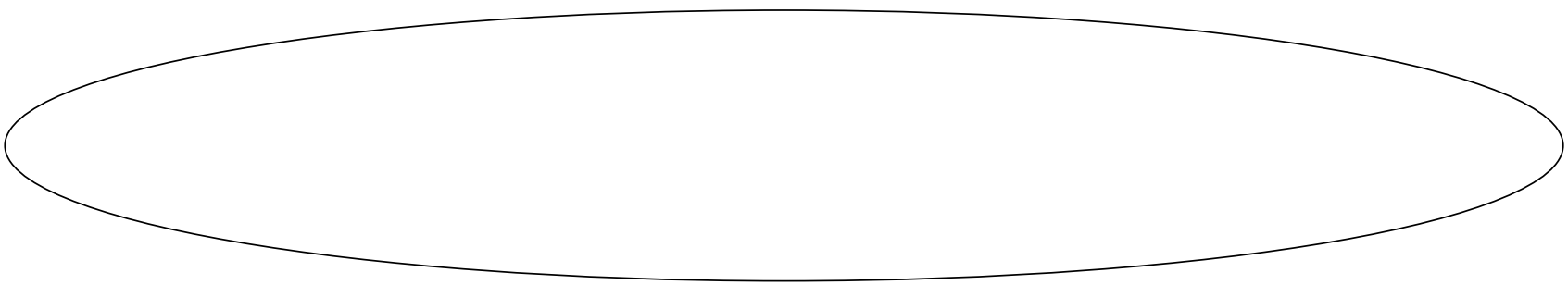
The main checker by example

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \qquad \dfrac{}{\neg x \geq 2}\,\text{LIA}}{}\,\text{Reso}$$

$$\dfrac{\qquad\qquad\qquad}{}\,\text{Reso}$$

A set of clauses:



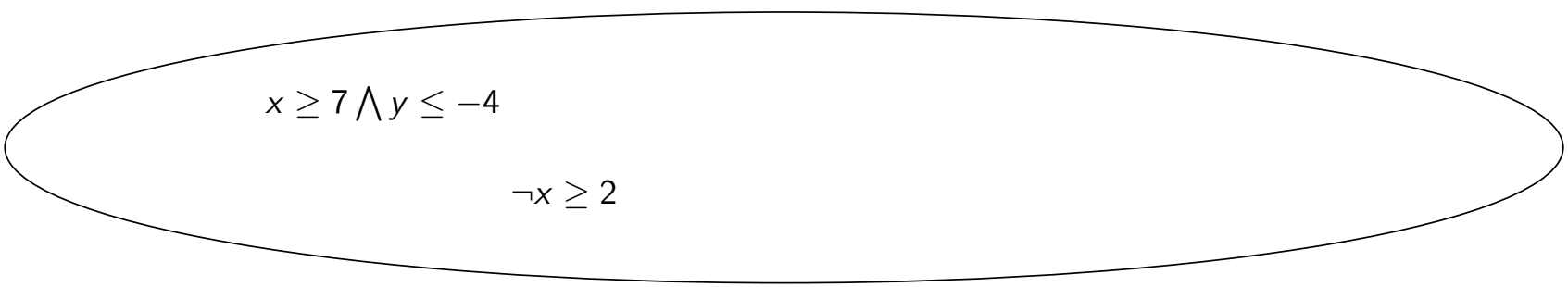
The main checker by example

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \qquad \dfrac{}{\neg x \geq 2}\,\text{LIA}}{}\,\text{Reso}$$

$$\dfrac{\qquad\qquad\qquad\qquad}{}\,\text{Reso}$$

A set of clauses:

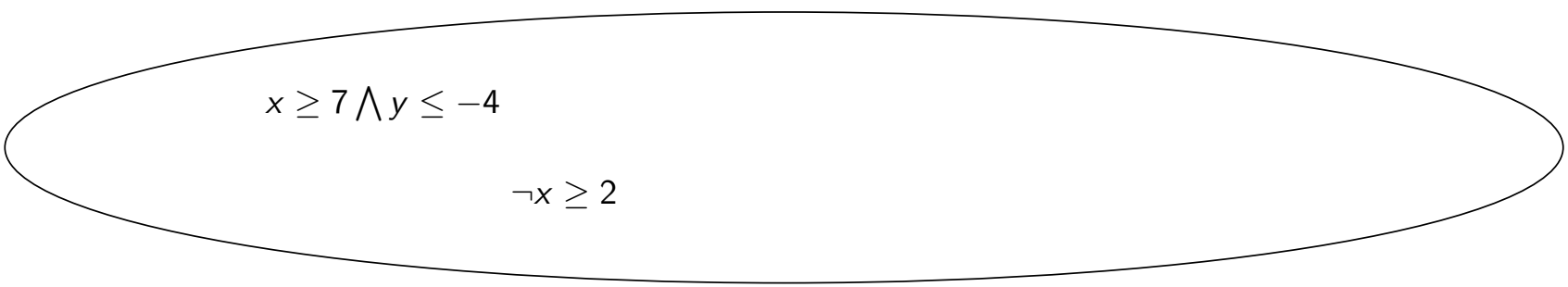


The main checker by example

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{}\,\text{Reso}}{}\,\text{Reso}$$

A set of clauses:

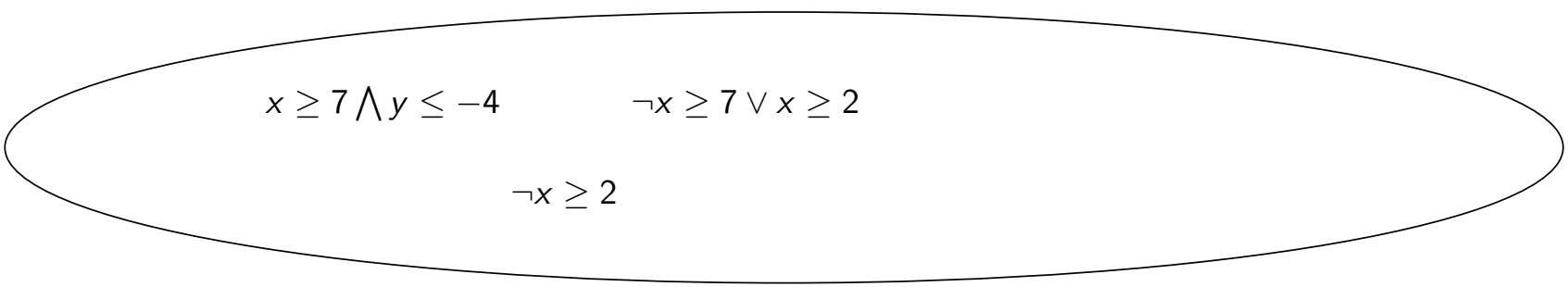


The main checker by example

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{}\,\textsc{Reso}}{}\,\textsc{Reso}$$

A set of clauses:

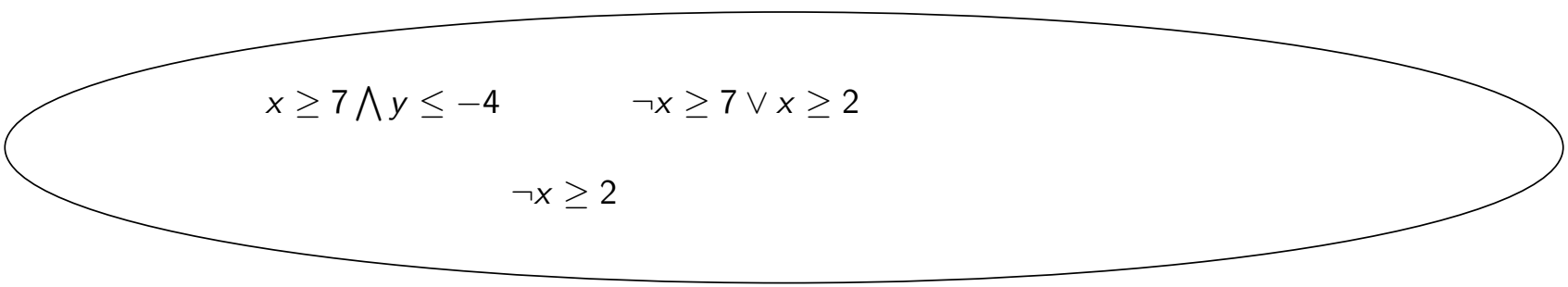


The main checker by example

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{}\,\textsc{Reso}$$

A set of clauses:

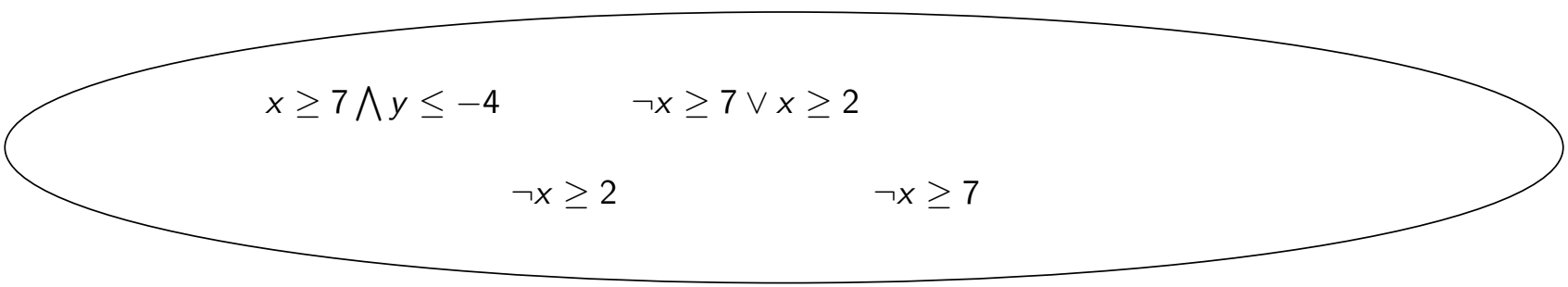


The main checker by example

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{}\,\textsc{Reso}$$

A set of clauses:

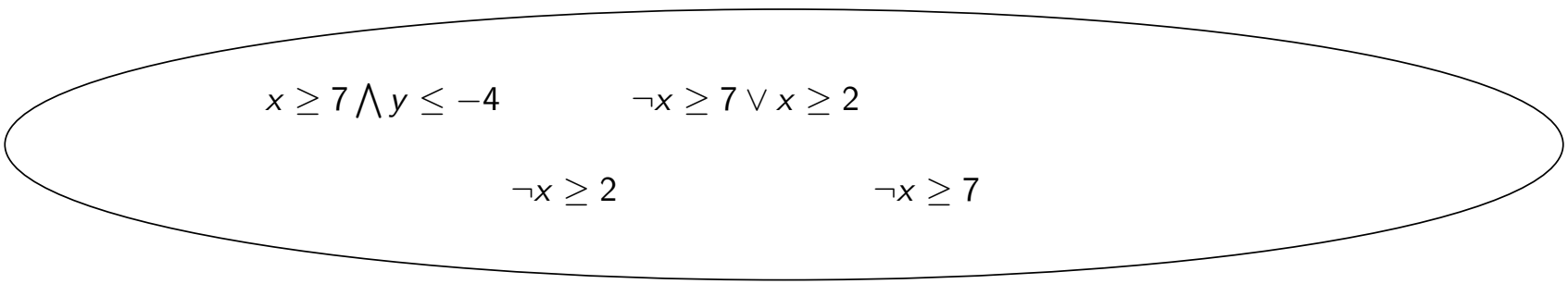


The main checker by example

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{}\,\text{Reso}$$

A set of clauses:

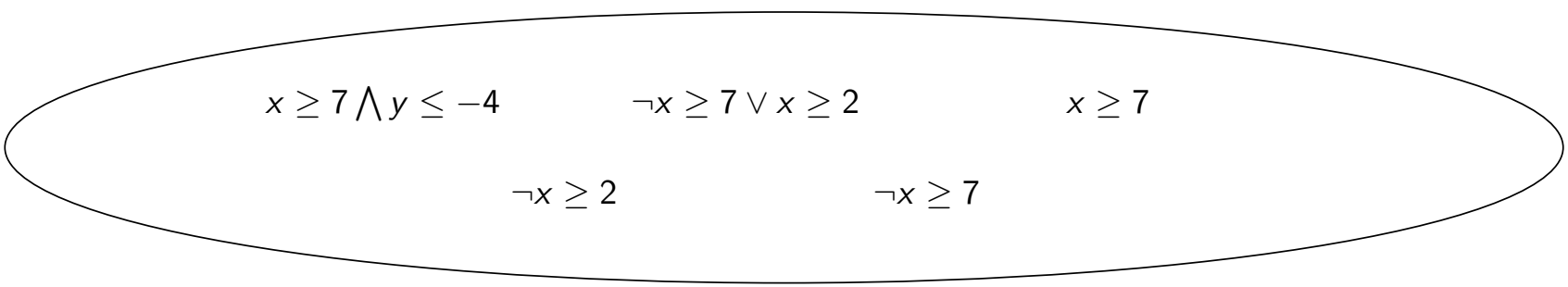


The main checker by example

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{}\,\text{Reso}$$

A set of clauses:

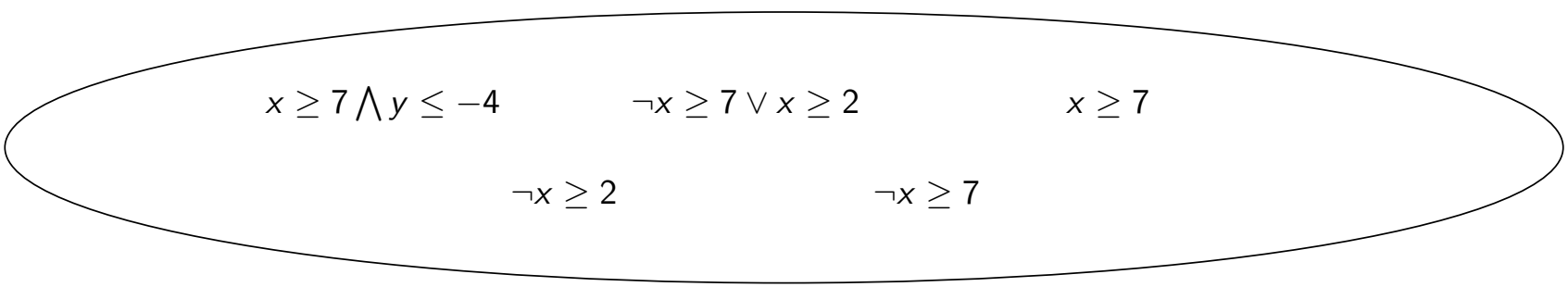


The main checker by example

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{\square}\,\textsc{Reso}$$

A set of clauses:

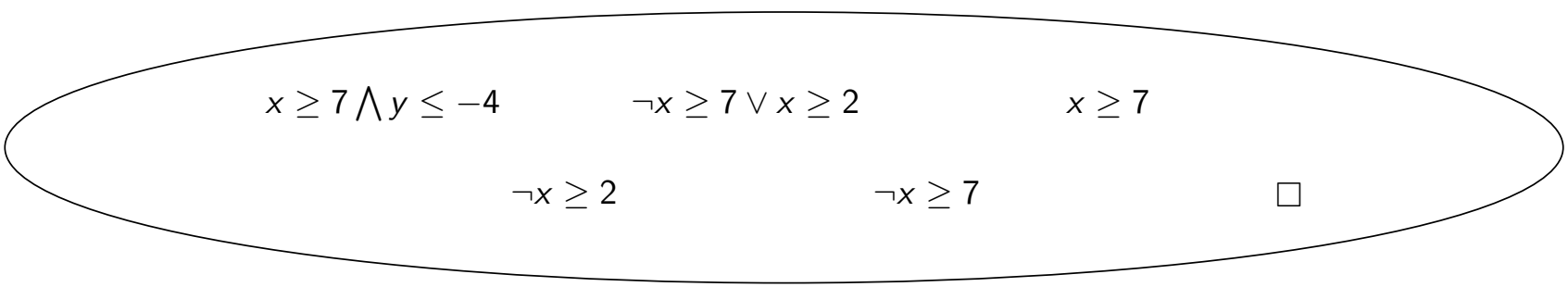


The main checker by example

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{\square}\,\textsc{Reso}$$

A set of clauses:

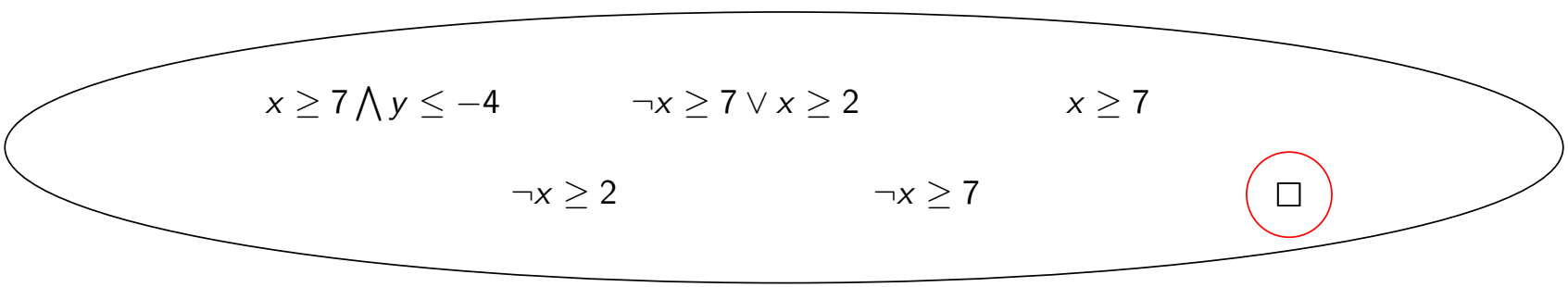


The main checker by example

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $\neg x \geq 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{\square}\,\textsc{Reso}$$

A set of clauses:



Improvement

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $x < 2$

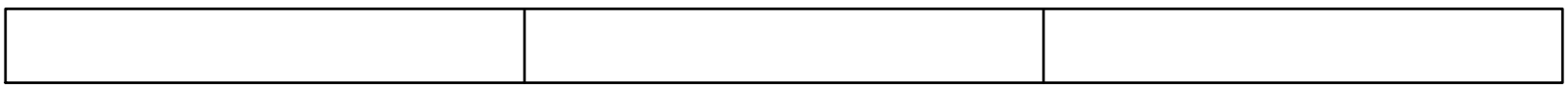
$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\square}\,\text{Reso}$$

Improvement

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{\square}\,\textsc{Reso}$$

3 clauses alive at the same time:



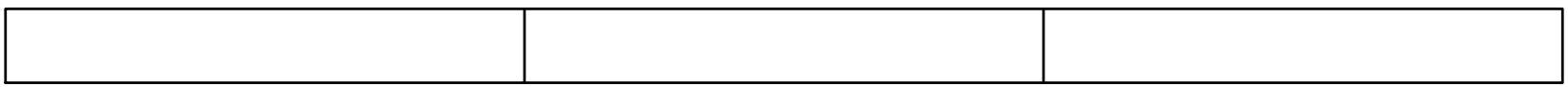
Improvement

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \qquad \dfrac{}{\neg x \geq 2}\,\text{LIA}}{}\,\text{Reso}$$

$$\dfrac{\qquad\qquad\qquad\qquad}{}\,\text{Reso}$$

3 clauses alive at the same time:

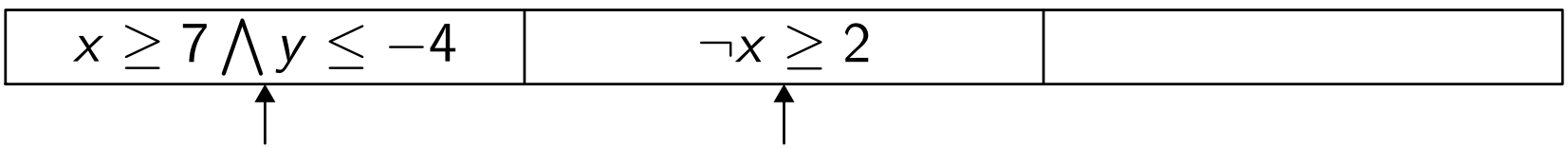


Improvement

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \qquad \dfrac{}{}\,\text{LIA}}{\dfrac{}{\neg x \geq 2}\,\text{Reso}}\,\text{Reso}$$

3 clauses alive at the same time:



Improvement

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{}\,\text{Reso}}{}\,\text{Reso}$$

3 clauses alive at the same time:

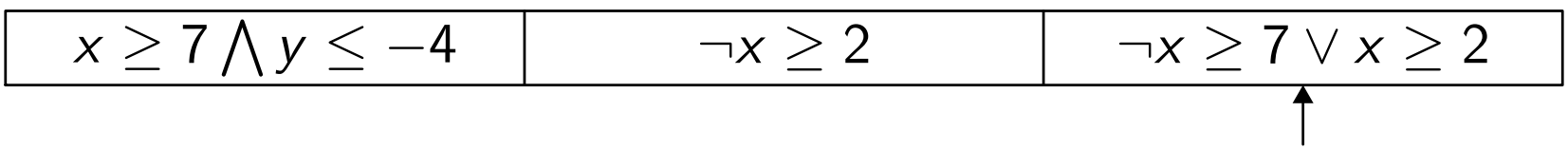
$x \geq 7 \bigwedge y \leq -4$	$\neg x \geq 2$	

Improvement

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{}\,\textsc{Reso}}{}\,\textsc{Reso}$$

3 clauses alive at the same time:



Improvement

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{}\,\textsc{Reso}$$

3 clauses alive at the same time:

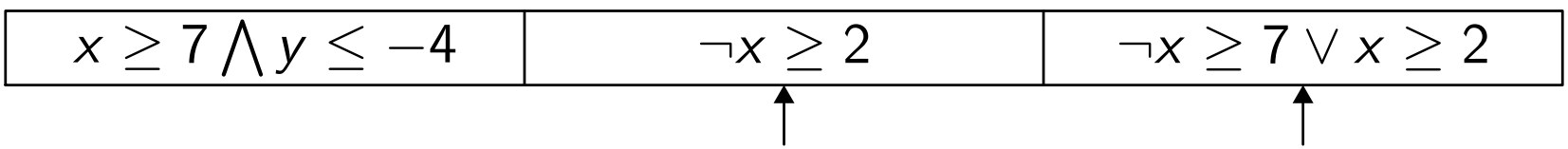
$x \geq 7 \bigwedge y \leq -4$	$\neg x \geq 2$	$\neg x \geq 7 \vee x \geq 2$

Improvement

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{}\,\textsc{Reso}$$

3 clauses alive at the same time:

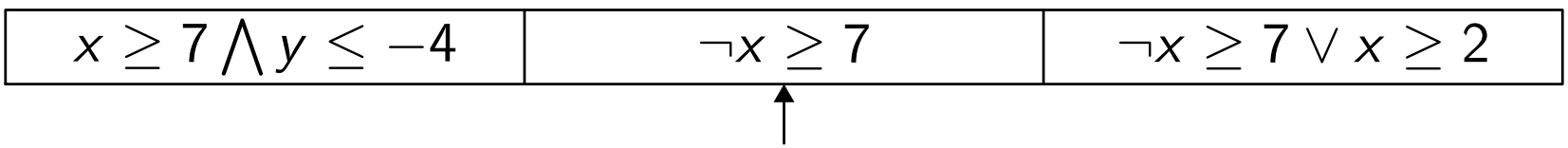


Improvement

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{}\,\text{Reso}$$

3 clauses alive at the same time:



Improvement

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{}\,\textsc{Reso}$$

3 clauses alive at the same time:

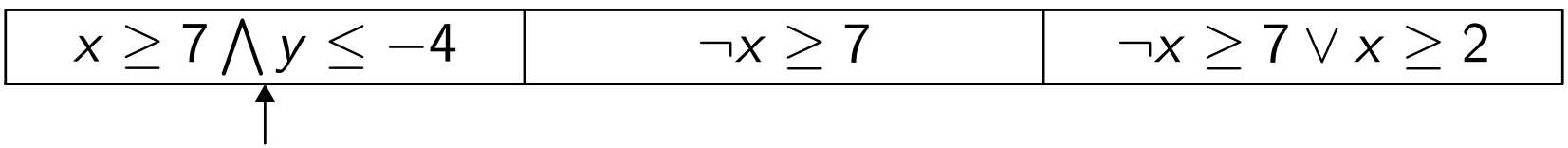
$x \geq 7 \bigwedge y \leq -4$	$\neg x \geq 7$	$\neg x \geq 7 \vee x \geq 2$

Improvement

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{}\,\textsc{Reso}$$

3 clauses alive at the same time:

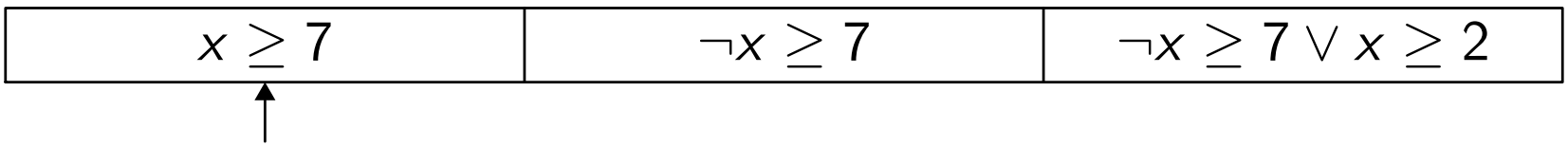


Improvement

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{}\,\textsc{Reso}$$

3 clauses alive at the same time:



Improvement

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{\square}\,\textsc{Reso}$$

3 clauses alive at the same time:

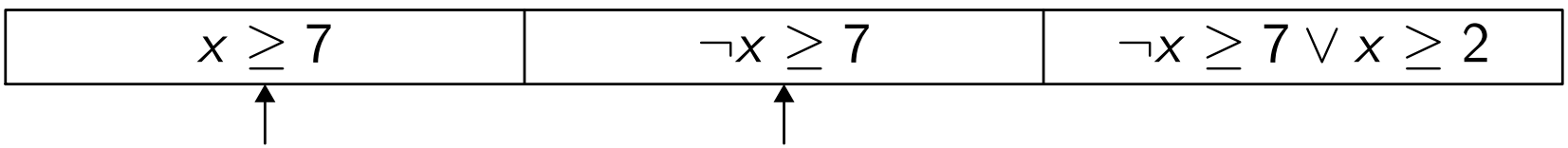
$x \geq 7$	$\neg x \geq 7$	$\neg x \geq 7 \vee x \geq 2$

Improvement

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{\square}\,\textsc{Reso}$$

3 clauses alive at the same time:

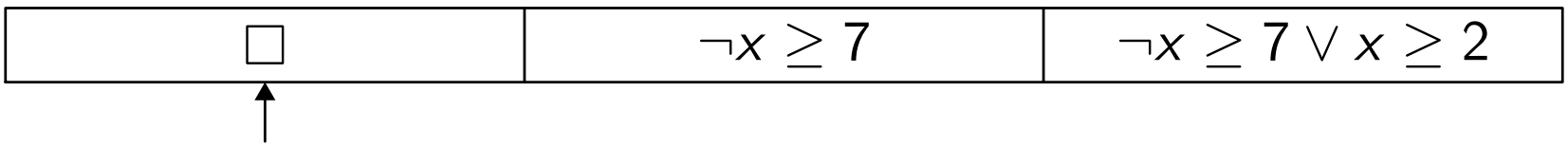


Improvement

Unsatisfiability of: $x \geq 7 \bigwedge y \leq -4$ $x < 2$

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \quad \neg x \geq 2}{\neg x \geq 7}\,\text{Reso}}{\square}\,\text{Reso}$$

3 clauses alive at the same time:



After preprocessing

1:	INPUT	$x \geq 7 \bigwedge y \leq -4$
2:	INPUT	$\neg x \geq 2$
3:	LIA	$\neg x \geq 7 \vee x \geq 2$
4:	RESO	$[2; 3]$
5:	CNF_PROJ	1
6:	RESO	$[4; 5]$

After preprocessing

1:	INPUT	$x \geq 7 \bigwedge y \leq -4$	1
2:	INPUT	$\neg x \geq 2$	2
3:	LIA	$\neg x \geq 7 \vee x \geq 2$	3
4:	RESO	$[2; 3]$	2
5:	CNF_PROJ	1	1
6:	RESO	$[4; 5]$	1

After preprocessing

1:	INPUT	$x \geq 7 \bigwedge y \leq -4$	1
2:	INPUT	$\neg x \geq 2$	2
3:	LIA	$\neg x \geq 7 \vee x \geq 2$	3
4:	RESO	$[2; 3]$	2
5:	CNF_PROJ	1	1
6:	RESO	$[1; 2]$	1

Outline

Current small checkers

Current small checkers:

- resolution chains
- CNF computation
- Linear Integer Arithmetic (using `Micromega`)
- Simplifications (eg. $x < 2 \rightsquigarrow \neg x \geq 2$)
- Equality of Uninterpreted Functions

Current small checkers

Current small checkers:

- resolution chains
- CNF computation
- Linear Integer Arithmetic (using `Micromega`)
- Simplifications (eg. $x < 2 \rightsquigarrow \neg x \geq 2$)
- Equality of Uninterpreted Functions

A need for a good representation of atoms and literals

Running example:

$$\dfrac{\dfrac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\,\text{CNF} \quad \dfrac{\dfrac{}{\neg x \geq 7 \vee x \geq 2}\,\text{LIA} \qquad \neg x \geq 2}{\neg x \geq 7}\,\textsc{Reso}}{\square}\,\textsc{Reso}$$

A need for a good representation of atoms and literals

Running example:

$$\dfrac{x \geq 7 \qquad \dfrac{\neg x \geq 7 \vee x \geq 2 \qquad \neg x \geq 2}{\neg x \geq 7}\ \textsc{Reso}}{\square}\ \textsc{Reso}$$

A need for a good representation of atoms and literals

Running example:

$$\dfrac{x \geq 7 \qquad \dfrac{\neg x \geq 7 \vee x \geq 2 \qquad \neg x \geq 2}{\neg x \geq 7}\ \textsc{Reso}}{\square}\ \textsc{Reso}$$

- the resolution checker does not need to know the content of atoms

Atoms and literals

Atoms:

- entirely hashed (maximum sharing)
- $x \geq 7 \bigwedge y \leq -4$ is $\mathcal{A}(5)$ in

0	1	2	3	4	5
x	y	$\mathcal{A}(0) \geq 7$	$\mathcal{A}(1) \leq -4$	$\mathcal{A}(0) \geq 2$	$\mathcal{A}(2) \bigwedge \mathcal{A}(3)$

Literals:

- $2i$ represents $\mathcal{A}(i)$
- $2i + 1$ represents $\neg\mathcal{A}(i)$
- l is the negation of m iff $l \oplus m = 1 \ldots$ etc.

Back to the example

0	1	2	3	4	5
x	y	$\mathcal{A}(0) \geq 7$	$\mathcal{A}(1) \leq -4$	$\mathcal{A}(0) \geq 2$	$\mathcal{A}(2) \bigwedge \mathcal{A}(3)$

The tree becomes:

$$\dfrac{x \geq 7 \qquad \dfrac{\neg x \geq 7 \vee x \geq 2 \qquad \neg x \geq 2}{\neg x \geq 7}\ \textsc{Reso}}{\square}\ \textsc{Reso}$$

$$\dfrac{4 \qquad \dfrac{5 \vee 8 \qquad 9}{5}\ \textsc{Reso}}{\square}\ \textsc{Reso}$$

Clauses and resolution

Clause:

- ordered list of integers (representing literals)

Resolution:

- does not need to know the hash-table of atoms
- is computed by running through the lists
- is efficient because comparison is based on bitwise operations

Current small checkers

Current small checkers:

- resolution chains
- CNF computation
- Linear Integer Arithmetic (using `Micromega`)
- Simplifications (eg. $x < 2 \rightsquigarrow \neg x \geq 2$)
- Equality of Uninterpreted Functions

Tseitin variables by example

Given the formula $(a \bigwedge b) \bigvee (c \Rightarrow d)$:

- introduce a new variable for each subformula:
  - $F_1 = a \bigwedge b$
  - $F_2 = c \Rightarrow d$
  - $F_3 = F_1 \bigvee F_2$
- introduce projection rules and tautologies associated to each of the variables:

$$\frac{F_1}{a} \quad \frac{F_1}{b} \quad \frac{\neg F_1}{\neg a \vee \neg b} \qquad \frac{}{\neg F_1 \vee a} \quad \frac{}{\neg F_1 \vee b} \quad \frac{}{F_1 \vee \neg a \vee \neg b}$$

$$\frac{F_2}{\neg c \vee d} \quad \frac{\neg F_2}{c} \quad \frac{\neg F_2}{\neg d} \qquad \frac{}{\neg F_2 \vee \neg c \vee d} \quad \frac{}{F_2 \vee c} \quad \frac{}{F_2 \vee \neg d}$$

$$\frac{F_3}{F_1 \vee F_2} \quad \frac{\neg F_3}{\neg F_1} \quad \frac{\neg F_3}{\neg F_2} \qquad \frac{}{\neg F_3 \vee F_1 \vee F_2} \quad \frac{}{F_3 \vee \neg F_1} \quad \frac{}{F_3 \vee \neg F_2}$$

In our representation of terms

- Tseitin variables are already introduced by the hash of atoms
- the CNF checker computes projections and tautologies by unfolding atoms at first level
- running example:

0	1	2	3	4	5
x	y	$\mathcal{A}(0) \geq 7$	$\mathcal{A}(1) \leq -4$	$\mathcal{A}(0) \geq 2$	$\mathcal{A}(2) \bigwedge \mathcal{A}(3)$

$$\frac{10}{4} \qquad \frac{10}{6} \qquad \frac{11}{5 \vee 7}$$

In our representation of terms

- Tseitin variables are already introduced by the hash of atoms
- the CNF checker computes projections and tautologies by unfolding atoms at first level
- running example:

0	1	2	3	4	5
x	y	$\mathcal{A}(0) \geq 7$	$\mathcal{A}(1) \leq -4$	$\mathcal{A}(0) \geq 2$	$\mathcal{A}(2) \bigwedge \mathcal{A}(3)$

$$\frac{10}{4} \qquad \frac{10}{6} \qquad \frac{11}{5 \vee 7} \qquad \frac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\ \mathrm{CNF}$$

In our representation of terms

- Tseitin variables are already introduced by the hash of atoms
- the CNF checker computes projections and tautologies by unfolding atoms at first level
- running example:

0	1	2	3	4	5
x	y	$\mathcal{A}(0) \geq 7$	$\mathcal{A}(1) \leq -4$	$\mathcal{A}(0) \geq 2$	$\mathcal{A}(2) \bigwedge \mathcal{A}(3)$

$$\frac{10}{4} \qquad \frac{10}{6} \qquad \frac{11}{5 \vee 7} \qquad \frac{x \geq 7 \bigwedge y \leq -4}{x \geq 7}\ \text{CNF}$$

Current small checkers

Current small checkers:

- resolution chains
- CNF computation
- Linear Integer Arithmetic (using `Micromega`)
- Simplifications (eg. $x < 2 \rightsquigarrow \neg x \geq 2$)
- Equality of Uninterpreted Functions

LIA: using a decision procedure as a blackbox

Micromega works like SMTCoq:

- call of an external solver that produces a certificate
- call a Coq checker on the formula and the certificate

Very easy to use:

- translate our representation of formulas into Micromega's representation
- call the external solver during preprocessing
- call the Coq checker during checking
- no need for a certificate from the SMT solver!

Current small checkers

Current small checkers:

- resolution chains
- CNF computation
- Linear Integer Arithmetic (using `Micromega`)
- Simplifications (eg. $x < 2 \rightsquigarrow \neg x \geq 2$)
- Equality of Uninterpreted Functions

The simplification problem

Simplifications:

- silent simplification at the beginning (the formula that is proved is not the input: flattening, canonical form)
- simplifications inside the proof
- usually no certificate

Problem:

- given two terms
- prove their interpretations are "equivalent"

Two approaches

Syntactical approach:

- simultaneous descent of the two terms
- usage: associativity of conjunction and disjunction, double negation, simple rewriting of linear equations ($a \geq b \equiv b \leq a$)
- possibility to mix theories

Semantical approach:

- send the equivalence between the terms to a decision procedure
- chose the decision procedure by looking at the head symbol of the terms
- usage: $3 + x + y + 7 + x \equiv 2x + 10 + y$
- no mixing of theories

Current small checkers

Current small checkers:

- resolution chains
- CNF computation
- Linear Integer Arithmetic (using `Micromega`)
- Simplifications (eg. $x < 2 \rightsquigarrow \neg x \geq 2$)
- Equality of Uninterpreted Functions

Congruence closure: a theory at the heart of SMT solvers

Equality of Uninterpreted functions:

- if $x = y$ then $f(x) = f(y)$ and $P(x) \Leftrightarrow P(y)$
- used in SMT solvers to share equalities between theories

A theory which is easy to check

With certificates:

- three rules considered modulo the symmetry of equality

$$\frac{}{x_1 \neq x_2 \vee \cdots \vee x_{n-1} \neq x_n \vee x_1 = x_n}\ \textsc{Trans}$$

$$\frac{}{x_1 \neq y_1 \vee \cdots \vee x_n \neq y_n \vee f\ x_1\ \ldots\ x_n = f\ y_1\ \ldots\ y_n}\ \textsc{CongF}$$

$$\frac{}{x_1 \neq y_1 \vee \cdots \vee x_n \neq y_n \vee \neg P\ x_1\ \ldots\ x_n \vee P\ y_1\ \ldots\ y_n}\ \textsc{CongP}$$

Without certificates:

- not difficult to write a decision procedure

Conclusion about small checkers

Small checkers:

- independent from each other
- large notion of “theory”
- different small checkers for overlapping theories might cohabit
- can choose the most appropriate one

Outline

Certifying SMT solvers

In addition to finding a proof:

- generation of the proof witness
- pre-processing
- checking

Benchmarks:

- finding + generation $>$ pre-processing + checking
- faster than the state-of-the-art certified SMT solvers

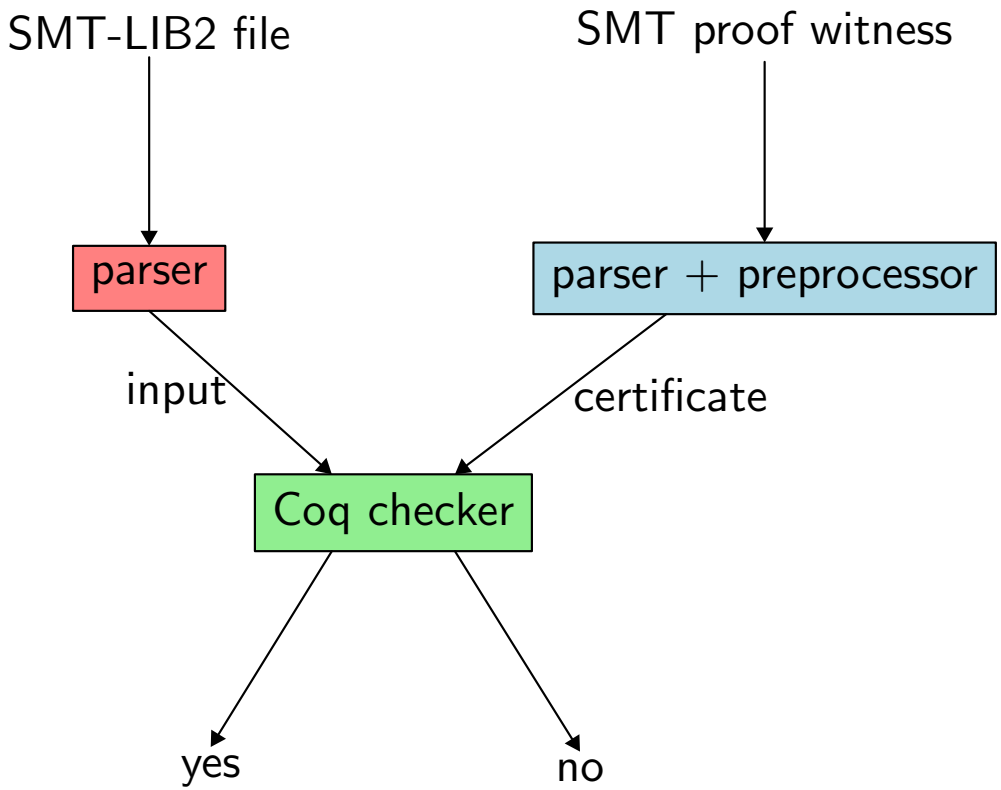
Outline

To sum-up

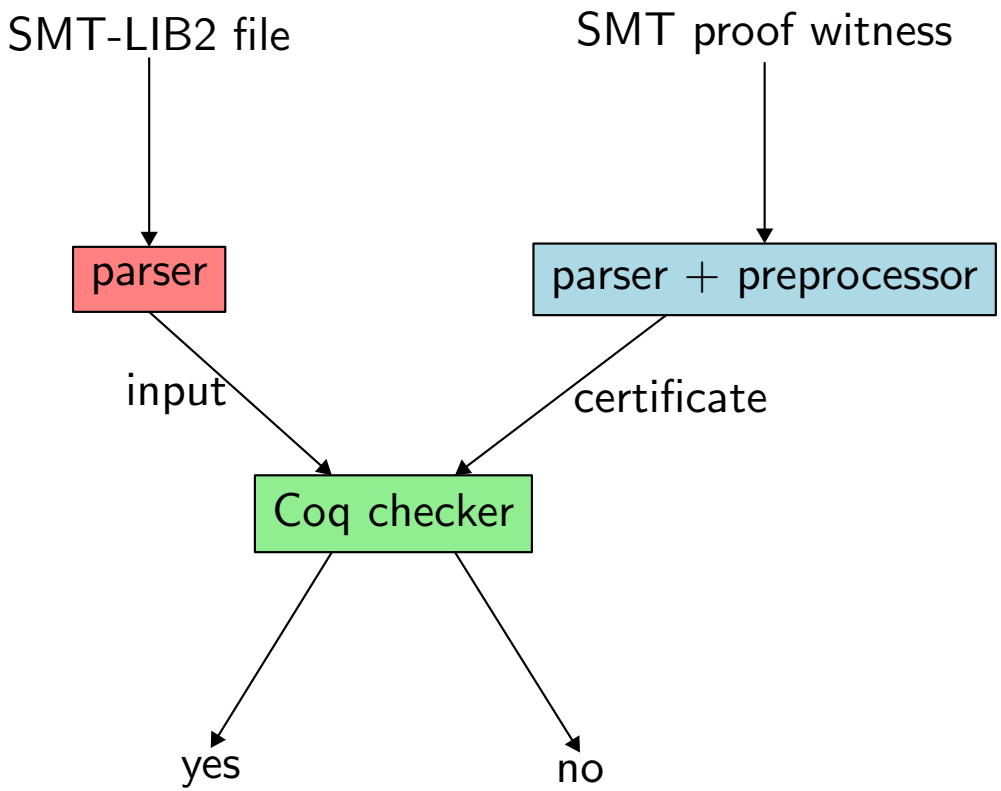
SMTCoq:

- efficient checker for a general notion of certificates
- on top of it, preprocessors for zChaff and veriT
- Coq tactics
- modular at two levels: in terms of provers and in terms of small checkers

Integration of new solvers



Integration of new solvers



other SMT solvers

first-order provers

tableaux, BDDs

An API for solvers

Some certificates can be very big (up to 200TB)

- Connect more tightly SMTCoq with the solvers
- Develop an API for solvers
- Replace writing a line in the certificate by calling the API

Using lemmas

SMT solvers can take advantage in using universally-quantified lemmas

- Include support for these in SMTCoq
- Pick lemmas from Coq standard library to send to the prover
- Recognize their instances in the certificate
- Apply the appropriate lemmas whn reconstructing the proof

Other work in progress

- Coq package
- integration of CVC4
- new theories: bit vectors, quantifiers, reals
  (arithmetic, differential equations)

Outline

Conclusion

`smtcoq.github.io`

Long-term perspectives:

- improve the tactics
- mix SMT solvers
- design a new certificate format
  - handle other provers (LF, HOL, ...)
  - standard?